Towards an Adaptive Method for Spatial Interpolation of Global Rain Gauge Data

Reading, March 11-13, 2003

Jürgen Grieser

Global Precipitation Climatology Centre
Deutscher Wetterdienst, P.O.Box 10 04 65, 63004 Offenbach, Germany
Juergen.grieser@dwd.de

Introduction

Gridded precipitation data are important information for many applications. The analysis of droughts and floods are only two examples.

The quality of estimated precipitation fields depends on many influences. From a pure climatological view the precipitation pattern consists of a deterministic and a stochastic part. The first should be subject to deterministic methods, whereas the latter should be subject to geostatistical methods. Both kinds of methods should be flexible to deal with the regional climatic and geographic conditions as well as with the station density provided.

The Global Precipitation Climatology Center (GPCC) has a large database of monthly precipitation data (over 50.000 stations worldwide). These data are thoroughly tested for outliers as well as errors in the station meta data like location and elevation. For the main portion of these stations precipitation data are available since 1986. However, for some 1000 stations also longer time series are acquired.

Aims of the new Project

It is one aim of a current project to use this large and quality controlled data base to achieve more information about the optimal way to interpolate monthly global precipitation fields over land, with special regard to the different climatological and geographic conditions as well as station density. Furthermore, this knowledge finally may result in long time series of best estimates of gridded world wide precipitation data as well as estimates of their uncertainty.

In a first step local altitude dependencies are estimated and averaged for the Koeppen climate zones each by each and for each calendar month see (Fig. 1). Furthermore, averaged local horizontal climate gradients are estimated for different regions and seasons. As a result it can be seen that the deterministic portion of spatial variability depends on the season and region under investigation. This reveals that it should not be neglected as part of the interpolation procedure. However, the actual altitude dependency depends strongly on the local. Therefore, the altitude function has to be fitted for each grid point and month.

In a second step the empirical semivariograms are calculated again for the different climate zones and calendar months. This gives the opportunity to learn about the regional and seasonal scales of precipitation. A comparison of these scales with the regional effective station distance should reveal information about the sampling error. Eventually this may be used to learn about the station density necessary to obtain reliable estimates under certain conditions.

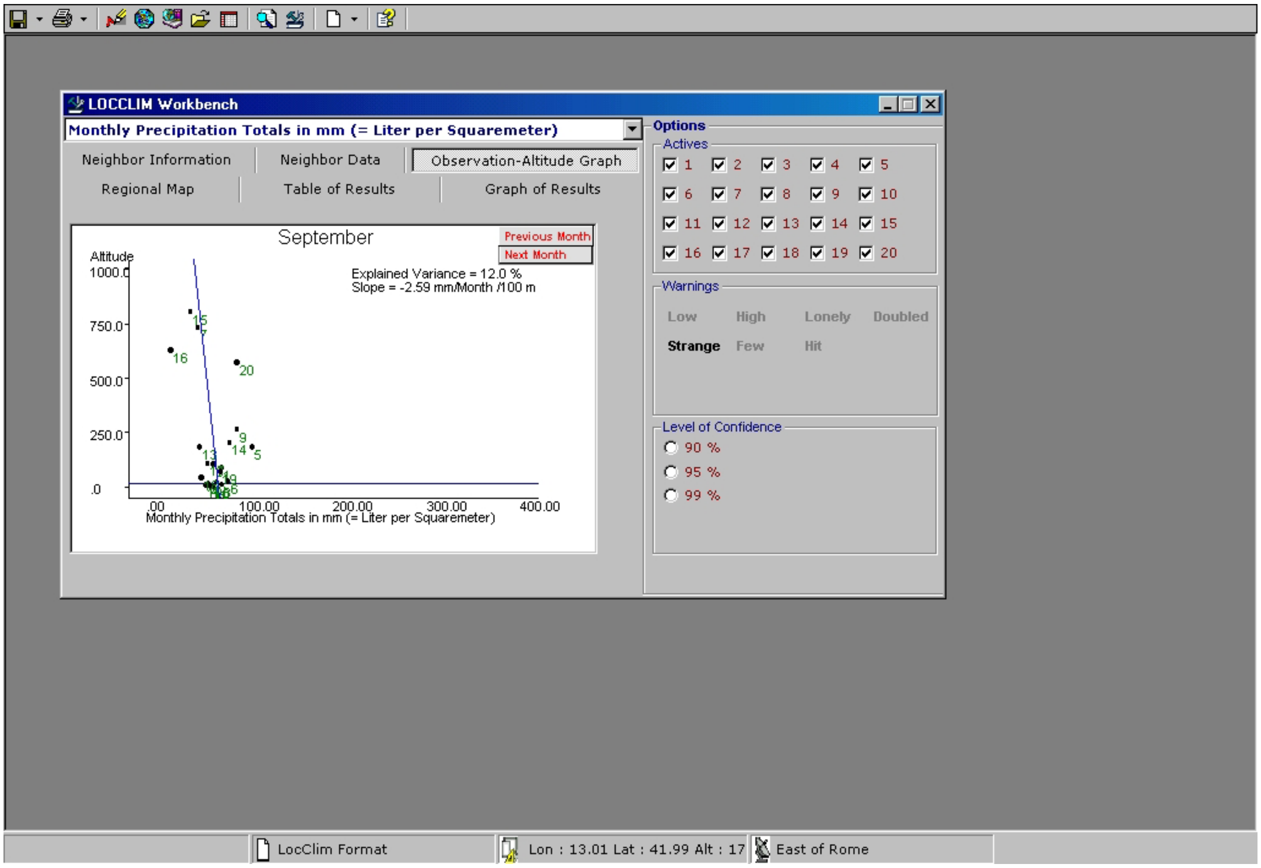

Figure 1: Investigation of local altitude dependence of precipitation. In this example the altitude dependence explains about 12% of variance. Outliers are detected and marked as strange data.

In a third step different geostatistical methods are intercompared for well defined artificial spatial patterns (see Fig. 2). This allows to investigate how well different patterns are reproduced by different methods. Also from this artificial examples information can be drawn about the impact of different station densities and distributions, given different methods of interpolation. As a result this step should offer enough information to decide which method provides the most reliable gridded data. Furthermore it may provide the information necessary to investigate the average error of the method under certain conditions.

Another point under investigation is the station distribution as well as the amount of stations used to interpolate a grid point value. Though more data means a better statistical basis, it also means that more distant stations have to be taken into consideration. With respect to this, more data does not necessarily mean better results. Experiments are planned to get the optimal number of stations to be used or the optimal distance up to which stations should be taken into account.

By means of a shadow correction stations hidden behind other stations should get less weight in the averaging procedures compared to stations that represent a large portion of a grid points neighbourhood. First experiments reveal that this leads to a further improvement of the gridded data.

Dubious data, which may not fit to the group of neighbouring stations should be indicated and treated separately. We saw that dubious data can be found by a simple cross check.

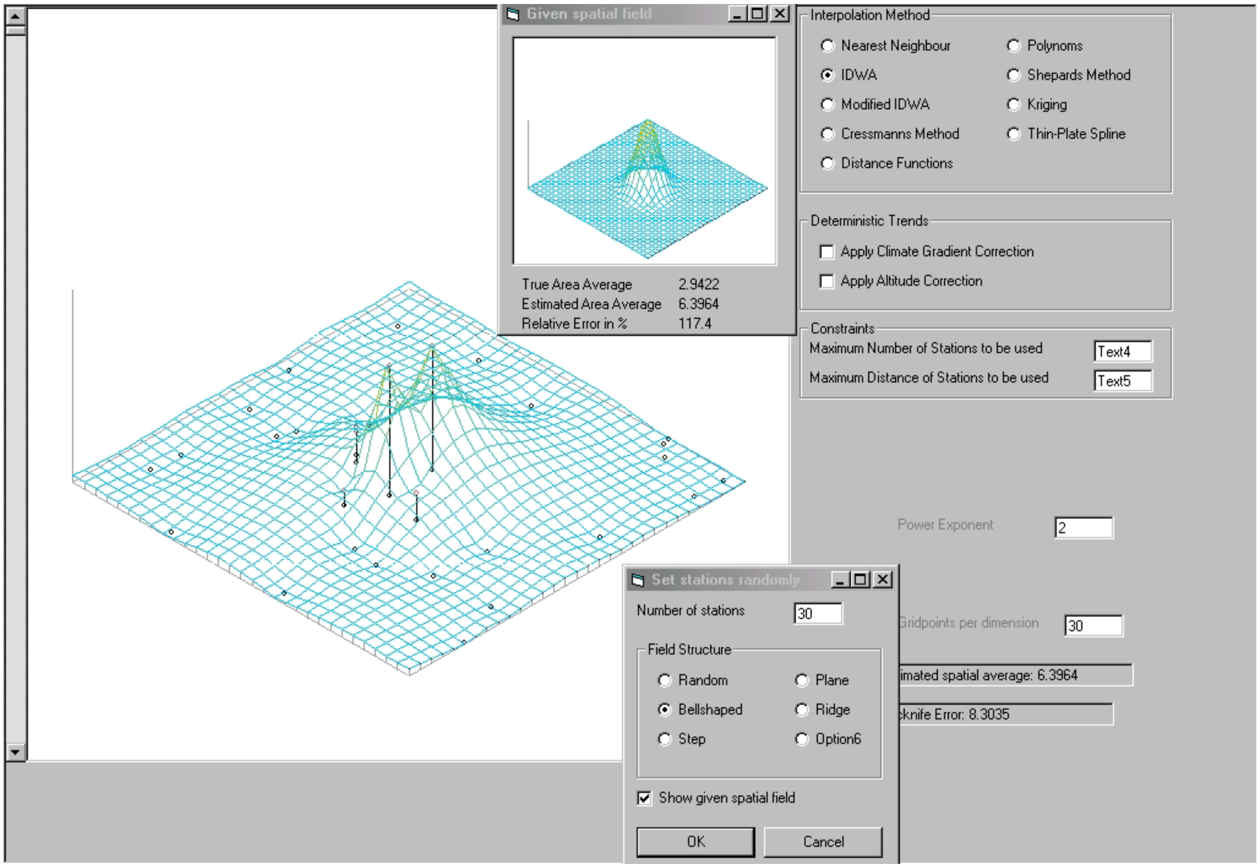

Figure 2: Investigation of the influence of station density and interpolation method. For the example of a bell shaped precipitation pattern and effective station distances equal to the scale of the bell the regional precipitation is overestimated by a factor of 2. Moreover, the shape is not correctly reproduced.

Furthermore, the jackknife error should be given for all the estimates in order to provide a measure for the accuracy of the actually interpolated value.

Finally all the information gathered from the experiments will be used to obtain an adaptive interpolation procedure to optimally fit the local and seasonal climatic conditions, the given station densities and the geographic margins best.