01 LD 0032
1.10.2001 – 30.4.2006

DEVELOPMENT OF A DATA BASIS FOR DEKLIM AND STATISTICAL ANALYSIS OF CLIMATE VARIABILITY

C.-D. SCHÖNWIESE[1], B. RUDOLF[2], S. TRÖMEL[1], C. BECK[2], J. GRIESER[2]

[1]*Institut für Atmosphäre und Umwelt, Universität Frankfurt, Postf. 111932,D-60054 Frankfurt*
[2]*Weltzentrum für Niederschlagsklimatologie (WZN), Deutscher Wetterdienst,Postf. 100465, D-63004 Offenbach*

Schönwiese@meteor.uni-frankfurt.de; http://www.geo.uni-frankfurt.de/iau/klima/Deklim/index.html

Key words: observed climate variability, time series analysis, spatial interpolation, gridded data, quality control, extremes

1 Summary

Existing global data sets of observed climate data are not sufficient in order to provide reliable gridded products. This project provides gridded data sets on the basis of the largest data collections world-wide merged into one relational database. These data are thoroughly quality controlled. In addition to the best estimates also error estimates are provided.

Up to now part of the data collections are merged, meta data are quality checked, station records are homogenized, outliers are removed and a global land surface precipitation climatology on a grid resolution of .5° for the period 1951-2000 is produced and published.

The analysis of climate variability realized in times series of observational data needs adequate statistical methods. Consequently, a generalized time series decomposition theory is developed, allowing a free choice of the underlying probability function. Depending on the characteristics of the climate variable, time series are interpreted as realization of a Gaussian-distributed, Gumbel-distributed or Weibull-distributed random variable with two time-dependent parameters. In particular, an analytical description of monthly precipitation time series can be achieved, which allows probability assessments of extreme values for every time step during the observation period. Furthermore, trend estimators with an reduced mean squared error compared to the common least-squared estimator can be deduced.

The project meets the milestone planning and can successfully reach its goals within the time frame given.

2 Aim of the research in the framework of DEKLIM

The project comprises two parallel parts which aim first (A) at the creation of a global climate data set based on observations with special regard to data quality and homogeneity, and second (B) at a diagnostic study on the global and European climate variability.

Within the first project part (A), a comprehensive data basis concerning precipitation, snow depth, surface air temperature, and mean sea level air pressure time series is under development. Large global data collections are achieved. Gaps are to be filled with national data sets. The data are to be quality controlled with respect to both, data and meta data (homogeneity, outliers, station locations incl. elevation, etc.). Finally gridded global data sets will be provided to all scientific users inside and outside DEKLIM.

Within the second project part (B) a generalized time series decomposition theory is introduced making the detection of significant changes in climate time series possible, even if Gaussian assumptions are not fulfilled. Only a time series decomposition based on a suitable statistical model provides the complete description of the time series under consideration and leads to reliable trend estimators for example or probability assessments.

3 Principal results

The project now has the most comprehensive collection of historical observed monthly precipitation data world-wide. A large portion of it is already quality controlled with respect to meta data and merged into a relational data base.

The station records are tested for homogeneity as well as for outliers. Both, inhomogeneities and outliers are removed. A subset of 9343 stations (out of about 50.000) passes all quality checks. This subset is used in order to produce a gridded global land surface precipitation climatology covering the period 1951-2000 with a grid resolution of .5°. It is freely available at http://www.dwd.de/vasclimo

The global data basis including surface air temperature, precipitation, snow depth and mean seal level pressure in process in project part (A) represents an adequate basis for analyzing climate variability but impose a high degree of flexibility on time series analysis concerning the statistical models and assumptions the statistical tools are based on. Within a generalized time series decomposition technique introduced the signal (structured components like trends, annual cycle, episodic component etc.) is detected in two instead of one parameter of a probability density function (PDF), which can be chosen without any further restriction. On the basis of the least-squares estimator, a successful statistical modelling can only be achieved for temperature time series.

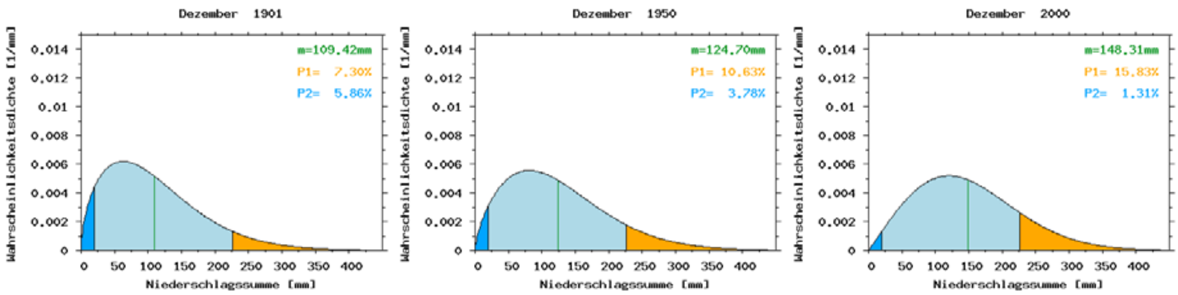

Figure 1. Three snapshots of the analytical description (PDF) of the monthly precipitation time series observed in Memphis (35.05N, 90.00E), USA. Integration gives the probability P1 for exceeding the 95th percentile (marked in orange) and the probability P2 for falling under the 5th percentile (marked in dark blue). The green line shows the expected value m at a given time.

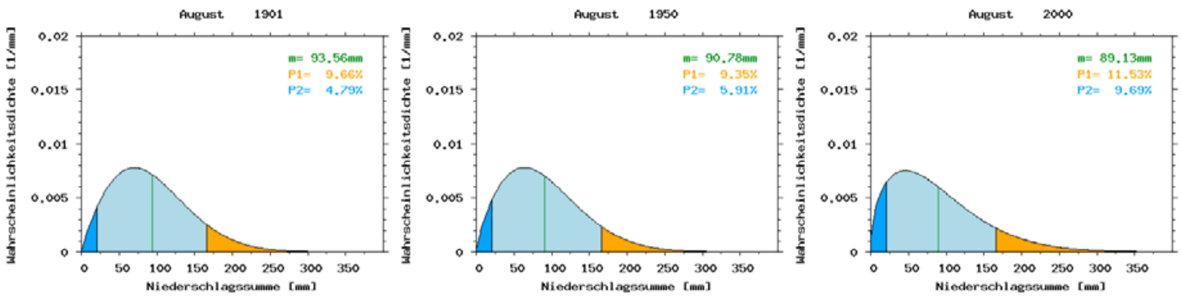

Figure 2. Similar to Figure 1, three snapshots of the analytical description of the monthly precipitation time series, 1901-2000, observed in Martinsburg (39.4N, 77.98W), USA, are shown.

Only a shift to higher or smaller values of the Gaussian distribution with constant variance in the course of time is necessary to describe the series.

But Figure 1 illustrates the description of the monthly precipitation times series observed at Memphis (USA) with the help of three snapshots in December of 1901, 1950 and 2000. Obviously, a positive trend in the scale parameter and a positive trend in the shape parameter of the Weibull distribution takes influence on the probabilities P1 and P2 of relatively high or relatively low precipitation totals and the expected value m either. Figure 2 shows analogous snapshots of the analytical description of the precipitation time series observed in Memphis. An increase in the probabilities for relatively high and relatively low precipitation totals P1 and P2 goes along with an decrease in the expected value m. Only the generalized time series decomposition technique provides the high degree of flexibility for the description of the observed changes in precipitation time series.

4 Main conclusions

Gridded climatological data sets published so far suffer from both, gaps in spatial data distribution and inhomogeneities in time. The product of this project can be seen as a significant improvement especially with respect to inhomogeneities. Since these are removed the gridded data set is suitable for the investigation of changes in global and regional precipitation. It reveals that the global land surface mean precipitation of about 2.1 mm/day does not show a significant change within the period 1951-2000 (Fig. 3). However, significant regional and seasonal changes become visible.

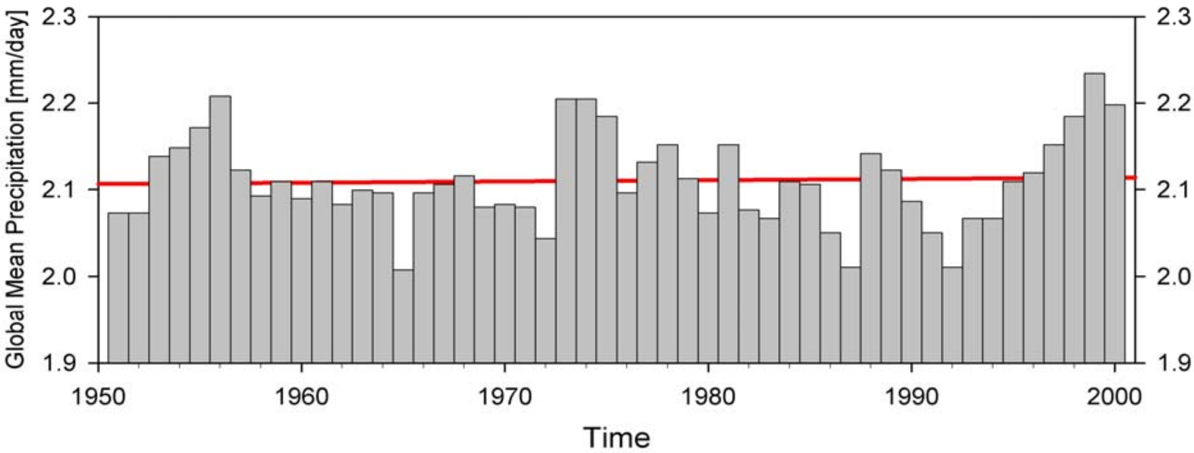

Figure 3. Time series of the annual mean of monthly precipitation totals averaged over the global land surface together with the linear trend from 1951 to 2000 (bold line)**.**

Considering only those time series with 0,1 or 2 gaps, project part (A) provided 446 monthly precipitation time series covering 1901-2000 and 2607 for the observation period 1951-2000 of a global station network to project part (B). The generalized time series decomposition technique provides movies of the analytical description for nearly every time series (extreme aride areas are excluded). Changes in structured components in the different parameters of the underlying PDF, in the probability of extreme values or in the expected value within the observation periods can be represented in maps. As an example Figure 4 shows those changes in the expected value of the monthly precipitation total in August in France for the period 1951-2000. In the left map, results are based on the interpretation of the time series as a realization of a Weibull-distributed random variable. Consequently, changes in the scale and the shape parameter are considered and the data of the whole observation period is used (50x12 values of each time series). A smaller mean squared error is the consequence. In the left map the ordinary least-squares estimator (based on Gaussian assumptions) is used, considering only the precipitation totals of August (50 values of each time series). Intercomparison illustrates, that unfulfilled Gaussian assumptions, often lead to a positive bias.

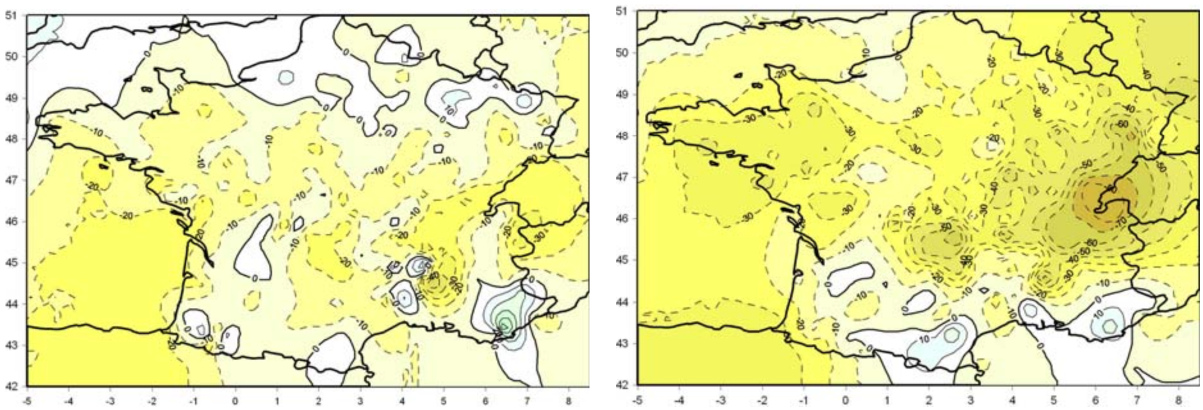

Figure 4. Trends in the monthly precipitation total in August observed in France for the observation period 1951-2000. Trend estimators are based on the generalized time series decomposition (left map) and the least-squares estimator (right map).

5 Next steps and completing activities

Though the first version of the globally gridded monthly precipitation is based on the largest collection of observations world wide there are still regions with poor station density. As a result of bilateral connections to the affected countries with the support of WMO these spatial gaps should be minimized within the frame of the project. Furthermore some statistical features of the data set are to be calculated and provided on the web page. Before investigating other variables, the investigations of precipitation include the analyses of pattern with respect to circulation indices like North Atlantic Oscillaten (NAO).

6 Policy relevance and application

The knowledge about the global and regional spatial patterns of climate variables as well as the knowledge about their observed changes within the last century is crucial for many applications. Aside the scientific aspects (evaluation and calibration of climate and ecosystem models as well as satellite data) gridded precipitation data are of great value for the investigation of global, regional and local water balances. In many regions of the Earth precipitation is the public source of palatable water. Furthermore, precipitation is one of the limiting factors with respect to agriculture. Thus, the knowledge of changes in regional and seasonal precipitation is of crucial importance for human life and prosperity. It is an important basis for climate impact studies and mitigation/adaptation strategies.

References

Beck, C., Grieser, J. and Rudolf, B., 2005, A new Monthly Precipitation Climatology for the Global Land Areas for the Period 1951 to 2000. In DWD: Klimastatusbericht 2004. In press.

Trömel, S., Schönwiese C.-D., 2004, A generalized method of time series decomposition into significant components including probability assessments of extreme events and application to observational German precipitation data. Submitted to Meteorol. Z.

Trömel, S., 2004, Statistische Modellierung von Klimazeitreihen. Dissertation, Institut für Atmosphäre und Umwelt, Universität Frankfurt a. M.