

Digitizing climate data in Germany: the progress of 6 years

H. Mächel (1) and A. Kapala (2)

(1) Deutscher Wetterdienst, Klima und Umwelt, Offenbach, Germany (hermann.maechel@dwd.de), (2) Meteorological Institute, University Bonn, Bonn, Germany

In this contribution the progress in digitization of German daily historical climate data are presented. Until 2005 the German Meteorological Service (Deutscher Wetterdienst, DWD) digitized historical data on request. In September 2005 the DWD project KLIDADIGI was established on the initiative of the Meteorological Institute of the University of Bonn. The goal of this project is a systematic digitization of German daily historical weather records.

In the first project year, an inventory of all archives belonging to the German Meteorological Service was carried out and the available inventory lists were digitized. The volume of the undigitized data for precipitation amounts to 165000 station-years (total potential) and for climate stations 23000 station-years. The labor input to digitize these documents amounts in total about 790 person-years. Because of high costs, only a portion of the documents can be digitized. Priority will be given to the long, complete series and the uniform spatial distribution after World War II.

The digitization is done in two different ways. First, photos are taken with a digital camera from the paper documents, and then the observed values are keyed. The digitization of the data with OCR software is still unsatisfactory because the records are handwritten or corrected, and thus difficult to recognize by the software. Therefore, Excel spreadsheets were prepared in which staff from the German Meteorological Service, and students are typing the values from the documents. These Excel spreadsheets calculate the basic statistics (maximum and minimum, monthly means and the number of days), that are also included in the paper records and that should be checked by the staff. The second quality control is a comparison of the monthly data calculated from these daily data and those monthly data that were digitized in the past. The third check of the digitized daily data is performed with neighboring stations.

Up to now, about 16500 station-years of rainfall were digitized (10.6% of the total potential, which is a prolongation of about 430 precipitation series) and 1800 years of climate station (7.8% of the total potential or 68 prolonged climate series). In the next few years, mainly climate records will be digitized, as the documents (most roll films from the World War II) are in a poor condition. At the end of 2013, the number of precipitation series should increase to 12% of the total potential and that of the climate series to 11% of the total potential. If the project is extended beyond 2013 the plan until 2025 is to increase the number of precipitation series of up to 30% of the total potential and the climate series of up to 24% of the total potential. This would increase the number of 100-year long precipitation series from 88 (in 2005) to 1000 and the number of the 100-year long climate series from 12 (in 2005) to 90 for Germany.