A monthly, three-dimensional data set of the global atmosphere, 1870-2005

S. BRONNIMANN, T. GRIESSER, T. EWEN, A. GRANT
Institute of Atmospheric and Climate Science, ETH Zurich

An important limitation of our understanding of 20th century climate variability is the availability of three-dimensional data sets of the atmosphere. This is necessary for analysing climate variability in the context of atmospheric circulation and for comparison with corresponding GCM simulations. Current observation-based data sets are limited to the reanalysis period, which reaches back to 1948 (NCEP/NCAR) or 1958 (ERA40). This period does not cover the first half of the 20th century which is thought to be essential for our understanding of the climate system because of strong trends (e.g., warming of the Arctic) and extremes (e.g., severe droughts). Here we present a new monthly, 3-D data set of the global atmosphere that reaches back to 1870. The data set comprises fields of temperature and geopotential height up to 200 hPa (for the northern extratropics up to 100 hPa). The data set is based on historical upper-air data from pilot balloons, kites, air craft, and radiosonde. In total, around three million profiles from around two thousand stations are now available prior to 1948. Some of the data were compiled from existing sources; some were digitised by us during the past few years. These data are supplemented with data from the Earth's surface (surface air temperature over land, SLP) in order to obtain better reconstructions. The data are used in a principal component regression approach to reconstruct the fields. Calibration is performed within the NCEP/NCAR or ERA40 reanalysis. The statistical models are validated within the calibration period using split-sample validations. They are also validated within the reconstruction period by leaving out predictors. Moreover, they are validated in a climate model simulation. Results show that reasonably good reconstructions can be obtained for analysing climate variability and for comparison with climate models. The reconstruction skill is better for geopotential height than for temperature, better for lower levels of the atmosphere than for the tropopause level, better for the northern extratropics and tropics than for the southern extratropics, better for the boreal winter season than for summer, and better for later periods (1930s and 1940s) than for the early decades. We present the reconstruction method as well as some analyses of the data set.