Geophysical Research Abstracts
Vol. 17, EGU2015-10256, 2015
EGU General Assembly 2015

Spatio-temporal statistical model for the optimal combination of precipitation measured at different time scales for estimating unobserved point values and disaggregating to finer timescales

Andràs Bàrdossy (1,2) and Geoffrey Pegram (2,3)

(1) University of Stuttgart, Institute for Hydraulic Engineering, Stuttgart, Germany (bardossy@iws.uni-stuttgart.de), (2) Civil Engineering, University of KZN, Durban, South Africa, (3) Pegram and Associates, Durban, South Africa

Precipitation observations are unique in space and time, so if not observed, the values can only be estimated. Many applications, such as the calculation of water balances, calibration of hydrological models or the provision of unbiased ground truth for remote sensing require full datasets. Thus a reliable estimation of the missing observations is of great importance. The problem is exacerbated by the ubiquitous decimation of gauge networks. We consider 2 problems as examples of the methodology: (i) infilling monthly data where some days are missing in the monthly records and (ii) infilling missing hourly values in daily records with the assistance of some nearby pluviometers. The key is that we need estimates of the distributions of the infilled values, not just their expectations, as we have found that the traditional 'best' values bias the spatial estimates.

We first performed monthly precipitation interpolation using 311 full records, 31 stations of which were randomly decimated to artificially create incomplete records as inequality constraints. Interpolation was carried out (i) without using these 31 in any way and (ii) using them as inequality constraints, in the sense that we determine a lower limit by aggregating the surviving data in a decimated record. We compare the errors if (i) the 31 stations with incomplete records are not considered against (ii) the errors if the incomplete records are considered as inequalities, and found that the partially decimated data add considerable value, as compared to neglecting them.

In a second application we performed a disaggregation in time. We take a set of complete hourly pluviometer data, then aggregate some stations to days. These then have their hourly missing data reconstructed and we evaluate the success of the procedure by cross-validation. In this application the daily sums for a location are considered as a constraint and the disaggregated daily data are compared to their observed hourly precipitation.

The techniques we have developed have the following advantages: they provide a flexible incorporation of information with reasonable uncertainty measures; no assumptions about marginal distributions are required and if appropriate, non-Gaussian spatial dependence could be used. The disadvantages are that: the procedure has a high computational demand, there is no public domain/commercial software available and there are a limited number of useful parametric spatial copulas.