Geophysical Research Abstracts
Vol. 17, EGU2015-10462, 2015
EGU General Assembly 2015

Data-infilling in daily mean river flow records: first results using a visual analytics tool (gapIT)

Laura Giustarini, Olivier Parisot, Mohammad Ghoniem, Ivonne Trebs, Nicolas Médoc, Olivier Faber, Renaud Hostache, Patrick Matgen, and Benoît Otjacques

Luxembourg Institute of Science and Technology (LIST), Environmental Research and Innovation Department (ERIN), 41 rue du Brill, L-4422 Belvaux, Luxembourg (laura.giustarini@list.lu)

Missing data in river flow records represent a loss of information and a serious drawback in water management. An incomplete time series prevents the computation of hydrological statistics and indicators. Also, records with data gaps are not suitable as input or validation data for hydrological or hydrodynamic modelling.

In this work we present a visual analytics tool (gapIT), which supports experts to find the most adequate data-infilling technique for daily mean river flow records. The tool performs an automated calculation of river flow estimates using different data-infilling techniques. Donor station(s) are automatically selected based on Dynamic Time Warping, geographical proximity and upstream/downstream relationships. For each gap the tool computes several flow estimates through various data-infilling techniques, including interpolation, multiple regression, regression trees and neural networks. The visual application provides the possibility for the user to select different donor station(s) w.r.t. those automatically selected.

The gapIT software was applied to 24 daily time series of river discharge recorded in Luxembourg over the period 01/01/2007 – 31/12/2013. The method was validated by randomly creating artificial gaps of different lengths and positions along the entire records. Using the RMSE and the Nash-Sutcliffe (NS) coefficient as performance measures, the method is evaluated based on a comparison with the actual measured discharge values. The application of the gapIT software to artificial gaps led to satisfactory results in terms of performance indicators (NS>0.8 for more than half of the artificial gaps). A case-by-case analysis revealed that the limited number of reconstructed record gaps characterized by a high RMSE values (NS>0.8) were caused by the temporary unavailability of the most appropriate donor station. On the other hand, some of the gaps characterized by a high accuracy of the reconstructed record were filled by using the data from an unexpected station, i.e. a donor station not belonging to the same basin as the target station. A subsequent analysis of the catchment characteristics of those stations revealed similarities in terms of catchment area, land cover and geology that could have led to a similar basin response, particularly when considering the relatively small size of the study area. A key finding was the overall better performance of neural networks and regression trees compared to other methods. As expected, best performances were obtained when middle and low flow conditions prevailed or when the gaps were comparatively short.

The gapIT tool represents a consistent approach for infilling data gaps in river flow records. The automated approach, coupled with a visual inspection system for user-defined refinement, allows for a standardized objective infilling, where subjective decisions are allowed but are at the same time traceable.