Geophysical Research Abstracts,
Vol. 10, EGU2008-A-02701, 2008
SRef-ID: 1607-7962/gra/EGU2008-A-02701
EGU General Assembly 2008

Filling missing values in time series using Singular System Analysis and application to hydrological time series

L. Gudmundsson (1), Dr. G. Lischeid (2), Dr. H. Lange (3)

(1) Department of Geo Sciences, University of Oslo, POBox 1047, Blindern, 316 Oslo, Norway, (lukas.gudmundsson@gmail.com), (2) Ökologische Modellbildung, Bayreuther Zentrum für Ökologie und Umweltforschung (BayCEER), Universität Bayreuth, 95440 Bayreuth, Germany, (gunnar.lischeid@bayceer.uni-bayreuth.de), (3) Norsk Institutt for Skog og Landskap, Postboks 115, N - 1431 Ås, Norway, (Holger.Lange@skogoglandskap.no)

Many time series analysis methods depend on equally spaced observations with no data point missing. If this condition is met, powerful techniques are available that identify temporal structures such as trends or periodic phenomena or nonlinear dynamics. Unfortunately, most of observations of natural systems, in particular over longer periods of time such as decades, are prone to sampling errors leading to missing points in the observations.

Singular System Analysis (SSA) is a powerful tool to extract the dynamics contained in time series at arbitrary temporal scales. In its original formulation, however, SSA relies as well on gapfree data. Recently several extensions to SSA have been proposed which are designed to fill the gaps, exploiting the dynamics contained in gapfree parts of the series to estimate the structure of the signal at the position of missing values. In order to evaluate the power of these methods we apply them to artificial data sets as well as to observed runoff time series. The performance of the methods under investigation is assessed for various distributions of missing points in the data sets (in particular, one large gap as opposed to a number of smaller ones adding up to the same gap length) as well as for different percentages of missing values.

We demonstrate that SSA successfully reproduces time series with up to 30% miss-

ing values. A number of distributed smaller gaps, a situation most likely to occur in observations, spoil the analysis to a much lesser degree than a single large gap. Thus, these new variants of SSA substantially enlarge the set of observational time series amenable to the analysis, and allow obtaining precise estimates of the signal at the position of missing data points. Based on our results, recommendations w.r.t. to the optimal approach will be given depending on length and distribution of gaps.