Electronic supplementary material – Hydrogeology Journal

Similarity-based approaches in hydrogeology: proposal of a new concept for data-scarce groundwater resource characterization and prediction

Roland Barthel[1*], Ezra Haaf[2], Markus Giese[1], Michelle Nygren[1], Benedikt Heudorfer[3], Kerstin Stahl[3]

[1] Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden; * roland.barthel@gu.se

[2] Department of Architecture and Civil Engineering, Chalmers University of Technology, Sweden

[3] Faculty of Environment and Natural Resources, University of Freiburg, Germany

* corresponding author

Explanations on how to use and read the supplementary material

1 Files and folders provided

The zip-archive contains two pdf-files and one folder:

1. This present file "Explanations Supplementary Materials.pdf"
2. A folder called "Visual classification results"

2 Purpose of the supplementary material

Using the similarity between groundwater hydrographs is one of the core principles of the proposed approach. In the paper as such, only a few examples of hydrographs can be shown. Therefore, this supplementary material presents a much larger selection of time series. To demonstrate the idea of time series similarity, we chose to present the hydrographs sorted in groups which were defined by a visual comparison (visual classification). Please see sections '*Time series similarity'* and '*Visual classification'* in the paper for further explanations of this approach. Visual classification (grouping) of time series is not an objective scientific approach but a great tool to get to know and understand a data set – and in particular to get an idea of the wide variety of different "time series behaviours" (that is shapes and features).

In particular we would like to highlight the folders "Irregular" which contains plots of hydrographs which show irregular behaviour of some kind. Figure S1 shows a selection of time series from this folder.

The plots are provided as a supplementary material to give the interested reader a flavour of the diversity of time series data used in the authors' various studies and the concept of time series similarity, which is essential to the approaches proposed in the manuscript.

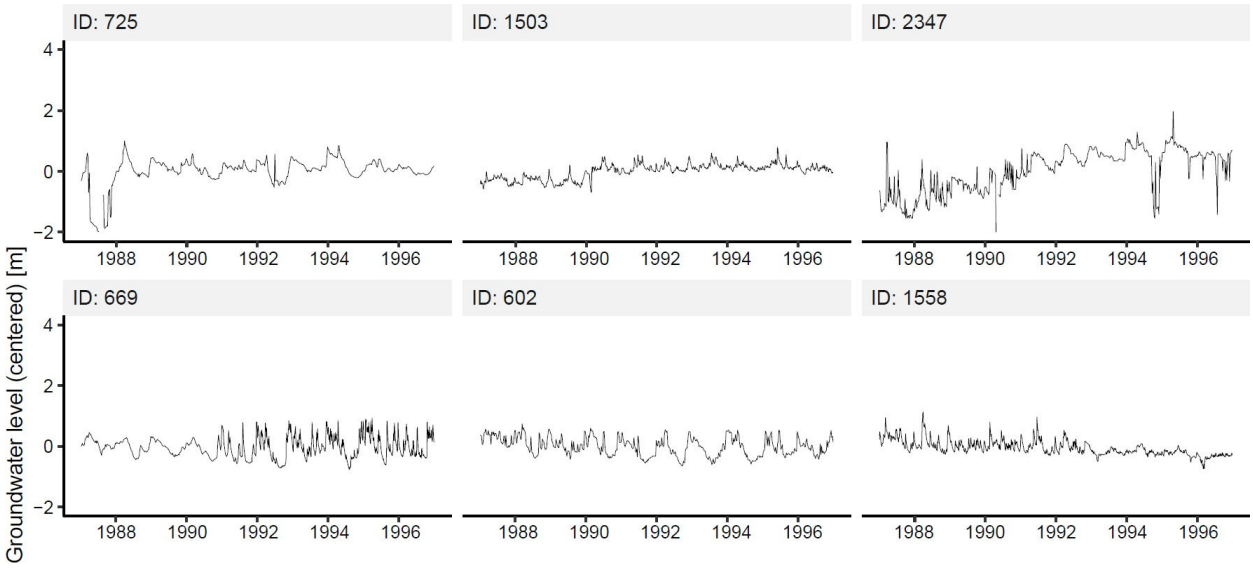

Figure S1: Examples of plots of time series that were classified as being "irregular" upon visual inspection. Irregularities can be sudden or continuous changes of magnitude, frequency or appearance in general.

3 Contents of the folder "Visual classification results"

The folder contains two subfolders:

1. z-scores
2. mean subtracted

Each of the two folders contains 936 plots of groundwater hydrographs from the study area (see section '*Locations and general description of the data'* of the paper) that were classified (grouped) based on visual classification (see section '*Visual classification'* of the paper), i.e. based on the perceived similarity upon visual comparison. The visual classification process resulted in a hierarchical classification scheme consisting of Groups, Subgroups and Types. Types were numbered with three digits X.Y.Z, where X stands for the Group, X.Y indicates the Subgroup, and X.Y.Z a Type.

- The plots in the folder "z-scores" show z-scores of the time series data, i.e. the mean was subtracted and divided by the standard deviation.
- The plots in the folder "mean subtracted" show the time series data, where the mean of each time series was subtracted from each value of the time series.

Please note: All plots provided contain a four digit ID with the suffix "GLOWA". These IDs are internal numbers used by the research groups and cannot be traced back to locations or the original IDs used by the agencies who provided the data. The raw data can unfortunately not be made available due to restrictions set by the data providers (please see acknowledgements section of the paper).