Systematic visual analysis of groundwater hydrographs: potential benefits and challenges

Roland Barthel [a], Ezra Haaf [b], Michelle Nygren[a], Markus Giese [a]

[a] Department of Earth Sciences, University of Gothenburg, Gothenburg, SE-40530, Sweden

[b] Department of Architecture and Civil Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

* Roland Barthel (roland.barthel@gu.se)

This document provides supplementary results that are not immediately needed to support the conclusions provided in the main article, but give nevertheless interesting insights into how the methods shown in the article could be extended and improved. In the main article, the present document could form a subsection in section "*Visual comparison and classification of time series*".

"Fuzzy" classification

The results presented in the main article represent an "exclusive" hierarchical classification – each time series belongs to one Type only. However, results remain unsatisfactory – there are many cases where one time series fits in several Types, depending on which of its visual features is weighted highest. This leads to an ambiguity of classifications, which can be explained by the mixture of patterns as demonstrated in Figure 4 of the main article. In-between-Type similarities across Groups and Subgroups exist, making the hierarchical approach questionable. This is a significant observation in relation to the application of the similarity concept. To explore this further, the between-similarities and within-similarities between/in Types, Subgroups and Groups were determined visually using a similarity matrix.

In this "fuzzy" approach a "degree of membership" (to a type, group) was assigned and each time series could be a member of several Types, Subgroups and Groups. To achieve this, a pairwise comparison of each Type with each other Type was carried out, and a between-Type similarity value assigned. The between-Type similarity values ranged from 0 (not similar at all) to 100 (identical). A between-Type similarity value of 25 could be interpreted as describing a situation where 25% of plots could realistically belong in either of the compared Types. These similarity values were visualized in a matrix showing both between-Type and within-Type similarity. The procedure is best explained with reference to the actual results (see section '*Visual comparison and classification of time series*' of the main article, and Figures S1, S2 and S3 in the present document).

Figure S1 shows a section from the 82 by 82 similarity matrix at Type level (the entire matrix is available in the figshare online supplementary material (Barthel, et al., 2020)). The matrix reveals that, as expected, Types within a Subgroup are somewhat similar but also that there can be some similarity between Types that are in different Subgroups or even Groups. Some Types show little similarity with other Types (e.g. Type 1.5.1) while others have similarities with many other Types (e.g. Type 2.2.4).

	Type01.1.1	Type01.1.2	Type01.1.3	Type01.1.4	Type01.2.1	Type01.3.1	Type01.3.2	Type01.4.1	Type01.4.2	Type01.5.1	Type01.5.2	Type02.2.1	Type02.2.2	Type02.2.3	Type02.2.4	Type02.3.1	Type02.3.2	Type02.3.3	Type02.3.4	Type03.1.1	Type03.1.2	Type03.1.3	Type03.1.4	Type03.2.1	Type03.2.2	Type03.2.3	Type03.2.4	Type04.1.1	Type04.1.2	Type04.1.3	Type04.2.1	Type04.2.2	Type05.1.2	Type05.2.1	Type06.1.1	Type06.1.3	Type06.2.1	Type06.2.2
Type01.1.1	100	20	0	20	0	20	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.1.2	20	100	0	10	0	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.1.3	0	0	100	20	10	20	10	10	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.1.4	20	10	20	100	5	15	10	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.2.1	0	0	10	5	100	10	5	5	0	20	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.3.1	20	5	20	15	10	100	50	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.3.2	10	5	10	10	5	50	100	5	0	0	0	10	10	0	0	0	0	0	0	0	0	0	0	20	0	10	0	0	0	10	0	5	0	0	0	0	0	0
Type01.4.1	0	0	10	10	5	10	5	100	30	0	0	15	5	5	0	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type01.4.2	0	0	5	0	0	0	0	30	100	0	10	10	20	10	0	10	10	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0
Type01.5.1	0	0	0	0	20	0	0	0	0	100	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
Type01.5.2	0	0	0	0	10	0	0	0	10	10	100	0	0	0	0	0	0	0	0	10	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type02.2.1	0	0	0	0	0	0	10	15	10	0	0	100	30	15	0	0	0	0	10	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0
Type02.2.2	0	0	0	0	0	0	10	5	20	0	0	30	100	15	20	5	10	0	20	0	0	0	0	10	15	0	0	0	0	0	0	0	0	0	0	0	0	0
Type02.2.3	0	0	0	0	0	0	5	10	0	0	0	15	15	100	25	15	20	10	20	0	0	0	0	0	25	10	10	0	0	0	0	0	0	0	0	0	0	0
Type02.2.4	0	0	0	0	0	0	0	0	0	0	0	0	20	25	100	20	20	15	30	5	10	10	20	10	25	10	5	0	0	0	0	0	0	0	0	0	0	0
Type02.3.1	0	0	0	0	0	0	5	10	0	10	0	0	5	15	20	100	40	15	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Type02.3.2	0	0	0	0	0	0	5	10	0	10	0	0	10	20	20	40	100	15	25	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0
Type02.3.3	0	0	0	0	0	0	0	0	0	0	0	0	0	10	15	15	15	100	20	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
Type02.3.4	0	0	0	0	0	0	0	10	0	0	0	10	20	20	30	25	25	20	100	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0

Figure S1: Section from the Type-similarity matrix showing the perceived similarity between each of the Types. A value of zero means not similar at all, 100 means identical. The meaning of the red box is explained in the text below.

Figure S2 shows the between-Subgroup and the within-Subgroup similarity as perceived by visual comparison. As an example, the between-similarity between Subgroups 2.2 and 3.2 is defined as the sum of the similarity values in the red box in Figure S1, normalized with the number of values within the box. Values on the matrix's diagonal show the within-Subgroup similarity. A value of 100 (identity) is only possible when there is only one single Type in a Subgroup.

	1.1	1.2	1.3	1.4	1.5	2.2	2.3	3.1	3.2	4.1	4.2	5.1	5.2	6.1	6.2	7.1	7.2	7.3	7.4	7.5	8.2	8.3	8.4	9.1	9.2	9.3	9.4	9.5
1.1	33.8	3.8	11.9	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.2	3.8	100.0	7.5	2.5	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.3	11.9	7.5	75.0	3.8	0.0	2.5	0.0	0.0	3.8	1.7	1.3	0.0	0.0	0.0	0.0	6.7	1.0	0.0	6.3	5.0	0.0	0.0	0.0	1.7	0.0	1.3	1.4	10.0
1.4	3.1	2.5	3.8	65.0	2.5	8.1	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.5	0.0	15.0	0.0	2.5	55.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2.2	0.0	0.0	2.5	8.1	0.0	38.1	12.2	2.8	8.1	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2.3	0.0	0.0	0.0	5.0	2.5	12.2	42.5	0.9	0.6	0.0	0.0	0.0	0.0	2.5	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3.1	0.0	0.0	0.0	0.0	0.0	2.8	0.9	41.9	11.6	0.0	6.9	6.3	10.0	5.6	0.6	10.0	8.8	14.6	1.6	0.0	0.0	0.0	0.8	0.8	0.6	0.3	1.4	0.0
3.2	0.0	0.0	3.8	0.0	0.0	8.1	0.6	11.6	45.6	2.1	5.0	1.3	7.5	0.0	0.0	21.7	11.0	16.3	8.4	0.0	0.0	0.0	0.0	0.8	0.3	0.0	0.7	2.5
4.1	0.0	0.0	1.7	0.0	0.0	0.0	0.0	0.0	2.1	60.0	14.2	0.0	0.0	0.0	0.0	0.6	0.7	1.1	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4.2	0.0	0.0	1.3	0.0	0.0	0.0	0.0	6.9	5.0	14.2	65.0	40.0	30.0	12.5	0.0	3.3	10.5	21.7	12.5	0.0	0.6	0.0	2.0	0.0	4.4	0.6	6.1	10.0
5.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.3	1.3	0.0	40.0	100.0	60.0	5.0	0.0	1.7	14.0	33.3	18.8	0.0	1.3	0.0	10.0	0.0	13.8	1.3	10.7	15.0
5.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	7.5	0.0	30.0	60.0	100.0	0.0	0.0	5.0	30.0	36.7	10.0	0.0	1.3	0.0	6.0	0.0	7.5	6.3	18.6	10.0
6.1	0.0	0.0	0.0	0.0	0.0	0.0	2.5	5.6	0.0	0.0	12.5	5.0	0.0	75.0	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0
6.2	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	25.0	62.5	0.0	0.0	0.0	0.0	0.0	0.0	3.1	2.0	0.0	0.0	0.0	0.0	0.0
7.1	0.8	0.0	6.7	0.0	0.0	3.8	0.4	10.0	21.7	0.6	3.3	1.7	5.0	0.0	0.0	55.6	14.0	11.1	13.3	3.3	0.0	0.0	0.0	5.0	0.0	0.4	3.8	0.0
7.2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	8.8	11.0	0.7	10.5	14.0	30.0	0.0	0.0	14.0	44.8	38.0	18.3	7.0	0.0	0.0	0.0	2.3	4.0	2.3	17.1	2.0
7.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.6	16.3	1.1	21.7	33.3	36.7	0.0	0.0	11.1	38.0	55.6	17.5	0.0	0.0	0.0	0.0	1.1	5.8	1.3	17.6	8.3
7.4	0.0	0.0	6.3	0.0	0.0	0.0	0.0	1.6	8.4	1.7	12.5	18.8	10.0	0.0	0.0	13.3	18.3	17.5	40.6	7.5	0.0	0.0	0.0	9.2	5.6	5.3	13.6	17.5
7.5	2.5	0.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	7.0	0.0	7.5	100.0	0.0	0.0	0.0	23.3	0.0	7.5	0.7	0.0
8.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	1.3	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	37.5	1.6	7.3	5.0	0.0	7.2	2.7	6.3
8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	1.6	33.1	4.8	0.0	0.0	0.0	0.0	0.0
8.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	2.0	10.0	6.0	5.0	2.0	0.0	0.0	0.0	0.0	0.0	7.3	4.8	35.2	0.7	8.5	1.5	6.9	2.0
9.1	0.0	0.0	1.7	0.0	0.0	0.0	0.0	0.8	0.8	0.0	0.0	0.0	0.0	0.0	0.0	5.0	2.3	1.1	9.2	23.3	5.0	0.0	0.7	40.0	3.8	22.1	1.7	0.0
9.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.3	0.0	4.4	13.8	7.5	0.0	0.0	0.0	4.0	5.8	5.6	0.0	0.0	0.0	8.5	0.0	48.8	21.6	24.6	11.3
9.3	0.0	0.0	1.3	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.6	1.3	6.3	0.0	0.0	0.4	2.3	1.3	5.3	7.5	7.2	0.0	1.5	0.0	21.6	48.1	16.1	6.3
9.4	0.0	0.0	1.4	0.0	0.0	0.0	0.0	1.4	0.7	0.0	6.1	10.7	18.6	0.0	0.0	3.8	17.1	17.6	13.6	0.7	2.7	0.0	6.9	0.0	24.6	16.1	39.2	13.6
9.5	0.0	0.0	10.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	10.0	15.0	10.0	0.0	0.0	0.0	2.0	8.3	17.5	0.0	6.3	0.0	2.0	0.0	11.3	6.3	13.6	100.0

Figure S2: Matrix showing between-Subgroup and within-Subgroup similarity (matrix diagonal).

Following the same principle as for between-Subgroup and within-Subgroup similarity, Figure S3 shows the between-Group and the within-Group similarity as perceived using visual comparison.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9
Group 1	15.2	1.6	0.3	0.3	0.0	0.2	0.7	0.0	0.3
Group 2	1.6	26.3	3.1	0.0	0.0	0.6	0.4	0.0	0.0
Group 3	0.3	3.1	27.7	3.0	6.3	1.6	10.2	0.1	0.7
Group 4	0.3	0.0	3.0	38.8	14.0	2.5	5.0	0.4	1.5
Group 5	0.0	0.0	6.3	14.0	80.0	1.3	17.7	3.5	9.1
Group 6	0.2	0.6	1.6	2.5	1.3	46.9	0.0	1.8	0.0
Group 7	0.7	0.4	10.2	5.0	17.7	0.0	24.4	0.0	7.3
Group 8	0.0	0.0	0.1	0.4	3.5	1.8	0.0	15.2	3.9
Group 9	0.3	0.0	0.7	1.5	9.1	0.0	7.3	3.9	21.9

Figure S3: Between-Group and within-Group similarity calculated as explained in Figure S2. Values on the matrix diagonal express the within-Group similarity.

The similarity matrices reveal a good deal about both visual similarity detection and similarity in general. For example, in some Groups, within-Group similarity (value on the diagonal in Figure S3) is not much higher than the between-Group similarity with another Group (e.g. Groups 7 and 5). That is, some of the Subgroups, Types or even individual time series, show strong similarities to other Subgroups, Types or individual time series which were classified differently. Essentially what this shows is that similarity of time series is complex and multifaceted – a strictly hierarchical and exclusive classification will not be able to capture this. Different aspects and patterns contribute in different ways to the perception of similarity and dissimilarity. Depending on how one perceives the "weight" of those aspects and patterns, similarity will be perceived to be weaker or stronger.

ESM Reference

Barthel R, Haaf E, Giese M, Nygren M (2020) Visual Classification Results. Figshare: https://doi.org/10.6084/m9.figshare.13281395.v1 (posted 24 November 2020)