AUTHORS: We thank Referee 2 for her/his time, work and the constructive comments. This helped to improve and clarify our manuscript. For the review process we divided the comments in subitems.

Anonymous Referee 2

Received and published: 18th September 2019

REFEREE #1: The paper addresses time series analyses for piezometric heads measured on different wells belonging to a dense network. The topic of the paper is suitable for HESS. The analyses are performed with a known statistical method (Principal Component). The novelty of the paper is the application of this method to piezometric heads chronicle to detect peculiarities in hydrographs of groundwater head.

AUTHORS: We thank the referee for the generally positive valuation of our study.

REFEREE #2: 141 groundwater head timeseries were selected from 583 wells. The selection criteria are for me unclear. Were they selected because they reach different aquifers? Using a first screening using statistical criteria?

AUTHORS: The suggested approach requires that the analysed series exhibit, at least after preprocessing, the same dates (lines 83+84, 434-436). The 583 wells were not monitored all at the same time (line 151). Therefore, we chose a period during which the number of continuously monitored wells was preferably large, and during which the gaps in the series were preferably small.

To clarify this, we rephrased the sentence in line 156 to:

"For this study, we selected those 141 wells which were continuously monitored during the 20 years period from 1993-11-01 to 2013-10-22 (Figure 1)."

In the following lines, we provide now also a more detailed description on the consistency of the sampling interval during the analysed period (please see our reply to comment SRC #8 of referee 1).

REFEREE #3: Furthermore, piezometric fluctuations with known anthropogenic influences are excluded from the PCA. Does it mean that the method detects only "minor" (not obvious from the visual inspection) peculiarities?

AUTHORS: No, the method is not restricted to "minor" peculiarities which are not obvious from the visual inspection. Please see also our reply to your comment #6.

REFEREE #4: Groundwater heads are usually depending on groundwater recharge, the thickness of the unsaturated zone, exchange with rivers that can have different time characteristics. Under such very different conditions, the computation of the reference hydrograph is not obvious and need some more details (see §3.3). Could you provide some more details for two very different time series and how the PCs included in the calculation are chosen?

AUTHORS: For each single observed hydrograph the respective reference hydrograph is calculated individually by multiple linear regression with the same stable PCs (in this study the first four PCs) (section 3.3). Thus, the regression coefficients are site specific, but not the selection of the PCs used

in the regression. The stable PCs were shown to be broadly insensitive to the selection of single wells or of single sampling dates. Thus, it can be concluded that the stable PCs depict general features of groundwater head dynamics in the region. Local features which appear only at a small subset of the wells will be not assigned to the stable PCs. Consequently, local effects will be assigned to the residuals of the affected wells rather than to their reference hydrographs. For example, if there would be an influence of a river on the observed groundwater head dynamics only at a small subset of the wells than this effect would be assigned to the residuals of the affected wells.

REFEREE #5: Mean depth to the ground surface are analyzed. Therefore, systematic errors due to the vertical coordinate of the well reference cannot be detected.

AUTHORS: We are not quite sure what you mean. Data about mean depth to groundwater is provided as background information for the data set. But these data were neither required nor used for subsequent analysis. However, the skewed distribution of mean depth to groundwater affects the estimation of the "normal behaviour" with the reference hydrograph. In this study it is distorted towards lower depths, as well as it is distorted towards areas with higher density of observation wells (line 350). We changed the wording from "biased" to "distorted" and extended the statement in line 350 to clarify this. It reads now:

"Because all the series were equally weighted by z-scaling (section 3.1), the derived PCs and consequently the determined normal behaviour were distorted towards areas with higher density of observation wells (Karl et al., 1982) as well as towards lower mean depths to groundwater (Figure 2)."

REFEREE #6: Moreover, the provided examples show time limited peculiarities. Is the method suitable to detect long term peculiarities like drifts?

AUTHORS: Yes it is, as long as the drift is limited to no more than a small subset of the monitoring network. In that case the drift will be not included in the stable PCs which are used to calculate the reference hydrographs. Instead the drift will be assigned to the residuals of the respective well(s).

This is also why for this application it is not required that the residuals fulfil specific conditions like exhibiting white noise distribution, etc. Please see lines 233-236 in the manuscript. We rephrased it according to your comment and the comment GRC #2 of Referee 1. It reads now:

"Systematic structures in the residuals like drifts, trends, cyclic patterns, sudden shifts or distinct periods of deviations indicate that the respective pattern is not representative for the whole data set, but is a local peculiarity instead."

To clarify this we extended the sentence in lines 433 + 434 of the conclusion to:

"The assignment of local anomalies to the residuals is not restricted to specific types of temporal patterns. The residuals merely comprise what cannot be ascribed to the reference hydrographs by means of the stable PCs. This can be short term structures like sudden shifts or distinct periods of deviation as well as long term structures like drifts, trends or cyclic patterns. The presented approach also does not require an interpretation of single PCs as distinct physical processes or functional relationships."

REFEREE #7: L319 - Reference to Kaiser criterion is 3.2 and not 3.1.

 AUTHORS: Thank you. We corrected this.