Hydrology and Earth System Sciences Discussions

Interactive comment on "Regional-scale identification of groundwater-surface water interaction using hydrochemistry and multivariate statistical methods, Wairarapa Valley, New Zealand" *by* M. R. Guggenmos et al.

Anonymous Referee #2

Received and published: 6 September 2011

In this paper, the authors use multivariate statistical methods to infer spatial patterns of surface water-groundwater interaction using a regional, historical water quality data set. The authors first refine the large chemical data set (of both surface and ground waters) by integrating and reducing reported field and laboratory results, summarizing multiple samples from individual sites using medians, and assessing data quality using a charge balance approach. They then apply hierarchical cluster analysis (HCA) to group sets of sites that are statistically similar and then use principal component analysis to better

understand what aspects of water quality differentiate the clusters defined by HCA. Finally, they look at the spatial distribution of sites that fall into the 7 clusters to infer patterns of surface water-groundwater interaction in the region.

Overall, this paper is of high quality. The presentation, both in the text and figures/tables, is strong and clear. I found the approach and results interesting and I think this manuscript will be of interest to HESS readers. Particular strengths of the research include the refinement of the large, regional data set, the novel use of multivariate statistical methods, the logical distribution of clusters in space (which supports the strength of the multivariate statistical methods used), and the validation of results using independent estimates of water age.

My few, specific comments for improvement are: - In the fourth paragraph of the introduction, the authors could include reference to the following article, which features the use of PCA to infer patterns of SW-GW interaction, albeit at a much, much smaller scale. Lautz LK, Fanelli RM. 2008. Seasonal biogeochemical hotspots in the streambed around restoration structures. Biogeochemistry, 91(1): 85-104. DOI: 10.1007/s10533-008-9235-2

- The first line of the last paragraph on page 6450 requires editing for clarity. "pastoral agriculture covers approximately 76% of the valley floor" – does this include small viticulture and market gardening projects OR are these projects also present?

- I would suggest re-naming section 3.2 as "Multivariate statistical methods." The authors could then add a paragraph describing their PCA approach, using text that is currently provided in the results and discussion section (see below).

- The first paragraph of section 4.1 discusses "outliers." The text references Table 1, where the outliers are listed, with bold values for chemical parameters that likely caused these sites to be classified as outliers. From the table, it is unclear why these values are bold (or why these values cause the sites to be outliers). For example, why would a conductivity value of 183 (first outlier site) cause a site to be identified as an

outlier, particularly when sites with an even lower conductivity (105) are not tagged as outlier and the mean for non-outlier sites is 313? Also, for Ca, why are values of 8 and 14 outliers, when values of 5 and 18 are not outliers? If these values are going to appear in bold in the table, some better explanation for why they are in bold is required. Also note that site 26/0001 has no values in bold – so there is no explanation for that outlier. Perhaps don't include bold values on the table.

- The majority of the last paragraph on page 6459 (which continues on the following page) is a discussion of the PCA methods. As such, it should be moved to the methods, per my earlier comment.

- On page 6462, line 22, a cluster is referred to as 1A, rather than A1.

- Line 3 of page 6463 – a "to" is missing between compared and monitoring sites.

- Page 6465, line 9 – I couldn't find a "Lake Ferry" on the map.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 6443, 2011.

Discussion Paper