REGIONAL-SCALE IDENTIFICATION OF SURFACE WATER AND GROUNDWATER INTERACTION USING HYDROCHEMISTRY AND MULTIVARIATE STATISTICAL METHODS, WAIRARAPA VALLEY, NEW ZEALAND.

MICHAEL R. GUGGENMOS, CHRISTOPHER J. DAUGHNEY, BETHANNA M. JACKSON AND UWE MORGENSTERN

Response to comments from Reviewer 5

We thank the reviewer for the positive comments on the manuscript. The reviewer makes eight specific comments, to which we reply below (our responses to the reviewer's last two comments are combined in the list below).

1. The reviewer suggests that we try more than two components in PCA.

   We performed PCA using variables that were standardised according to their z-scores. Hence it is reasonable for us to apply the Kaiser criterion, whereby we retain only those components with eigenvalues greater than one. The first component has an eigenvalue of 7.1 and explains 57.5% of the variation in the dataset. The second component has an eigenvalue of 1.9 and explains 15.8% of the variation in the dataset. The third component, which we opted not to retain in our final model, has an eigenvalue of 0.8 and explains only 6.9% of the variation in the dataset. The third component therefore explains less variance than any single variable from the original dataset, and hence it is not worthy of interpretation.

2. The reviewer suggests that we try K-means clustering analysis.

   Our study aimed to identify locations and styles of groundwater-surface water interaction. We selected multivariate statistical techniques that allowed us to achieve that aim. It would be interesting to compare the different statistical techniques that could be used instead of HCA and PCA, but this would be beyond the scope of our study.

   Those interested in comparison of different graphical and statistical techniques that can be applied to water quality data are referred to Güler and Thyne (2002) *Evaluation of graphical and statistical methods for classification of water chemistry data*, Hydrogeol. J. 10: 455-474. These authors conclude that HCA and K-means clustering (KMC) produce very similar results but that HCA offers certain important advantages. Notably, KMC does not produce easily interpretable visual output whereas HCA produces a dendrogram. In addition, KMC requires the user to specify the number of clusters that will be generated, which introduces a subjective bias, whereas HCA has no such requirement.

3. The reviewer points out that Table 1 doesn't explain which values have led to sites being classified as outliers.

   We acknowledge that the use of bold text in Table 1 was hard to follow. We have taken the advice of Reviewer 2 and removed the bold highlighting from Table 1.

4. The reviewer states that there is no need to compile TDS in Table 2 if EC is already provided.

   We acknowledge that there are generally good correlations between EC and TDS. However, we prefer to retain both values in Table 2 because the correlations between them are not perfect. It is also useful to include both values for readers that may be more familiar with one of the two types of measurement.

5. The reviewer points out that Table 3 should use bold text to show the weight for HCO_3 in component 1 and that weightings for PO_4 do not show any pattern amongst the components.

   This is correct and will be modified in our revised manuscript.

6. The reviewer states that "using multivariate statistical methods…is not widespread, but it is not neglected, many authors used this technique in studying groundwater and surface water interaction (see Woocay and Walton, 2006, 2008)."

   This seems to be more of a comment from the reviewer than a specific suggestion to change some part of our manuscript. Nevertheless, we agree with the reviewer. Our introduction states that techniques such as PCA and HCA have been used previously to interpret water chemistry, but use of these techniques is definitely not widespread for evaluation of groundwater-surface water interaction.

7. The reviewer notes two minor typographical errors in the manuscript.

   These minor errors will be corrected in our revised manuscript.