Scottish drinking water catchments: A proactive risk approach to protect raw water quality

Carolin Vorstius*, John Rowan*, Iain Brown*, Zoë Frogbrook, Javier Palarea*****

*University of Dundee, School of Social Sciences, Dundee, UK
**Scottish Water, Fairmilehead Office, Edinburgh, UK
***Biomathematics and Statistics Scotland, Edinburgh, UK

a.c.vorstius@dundee.ac.uk
www.hydronationscholars.scot

Climate change and drinking water quality

Freshwater ecosystems have a multitude of uses, one of the most important as drinking water supplies. Climate change, and related land use changes, are expected to increase pressures on water resources, with negative implications for water quality (ASC, 2016). It is important to understand how water quality may develop in future, and to anticipate emerging risks to raw water quality in the drinking water catchment. This will not only support a stable supply and efficient treatment of drinking water, but also a catchment management that creates wider benefits for the environment and society.

Rapid risk screening approach

A rapid, large-scale screening for drinking water supply catchments that are at risk of deteriorating raw water quality can be a first step in a more strategic inclusion of emerging risks into raw water management. It allows identifying catchments that need to be prioritized in terms of monitoring, research, and mitigation measures. An analysis of current raw water quality and pressures within the catchment forms the basis of such a screening that can then be combined with an assessment of future developments of these pressures. In a first application of this approach, raw water quality (described by concentrations of aluminium, iron, manganese, colour, turbidity, pH, coliform bacteria and *E. coli* for 2011-2016) of 154 surface water catchments used for drinking water by the public drinking water provider in Scotland, Scottish Water, was investigated. Relationships between catchment characteristics and raw water quality as well as spatial patterns of water quality were investigated using a mix of statistical multivariate methods, including Principal Component Analysis, Cluster Analysis, and Multi-Target Predictive Clustering Trees.

Principal Component Analysis

Principal component analysis (PCA) allows to identify overall conditions and pressures that explain most variability in the multivariate dataset, with Principal Components (PCs) reflecting key processes that are not directly observable in the original data (Selle et al., 2016). This points towards the most important current risk factors for water quality.

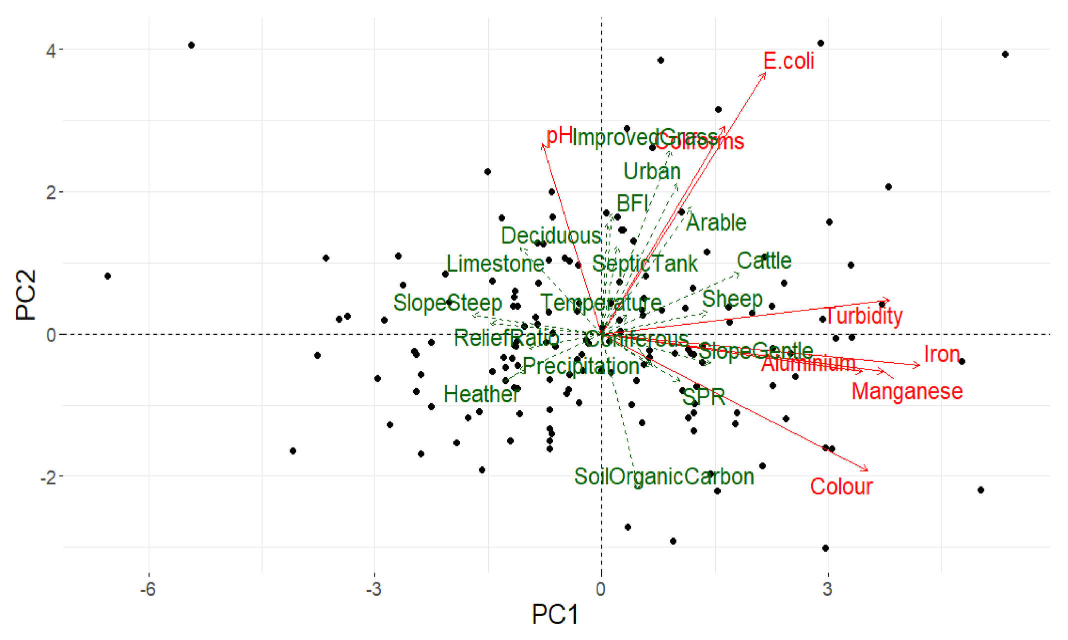

Figure 1: PCA biplot based on the first two PCs. These PCs explain 46% and 20% of the observed variability respectively. Red rays represent correlations between water quality variables and their associations with PC1 and PC2, dashed green rays reflect the correlation of these with catchment characteristics, and black dots display associations with catchments based on PC scores.

Median concentrations per water quality parameter and catchment were used as input for the PCA, and catchment characteristics were projected on the biplot of PC1 and PC2 (Figure 1) as supplementary variables. The biplot shows the majority of catchments relatively evenly scattered around the origin of the biplot (representing the overall mean across variables) in all directions. Catchments located towards the upper part are generally associated with increasing values of coliform and *E. coli* medians, as well as improved grassland, arable and urban cover. Catchments placed more toward the right-hand side present relatively higher median concentration values of metals, colour and turbidity. Toward the lower right corner, catchments were mostly characterised by increasing values in colour, going together with lower pH values and high organic carbon contents in the soils.

Cluster Analysis

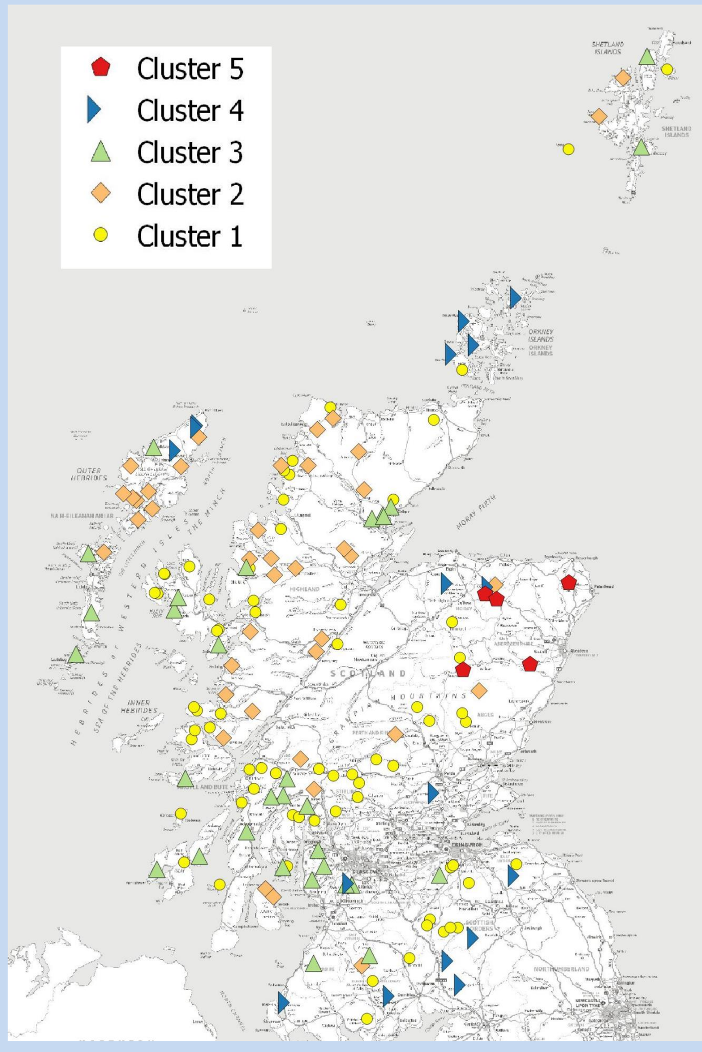

Figure 2: Distribution of catchments per cluster. Average silhouette width of clustering structure: 0.21.

Cluster analysis of the catchments was conducted in order to distinguish typical water quality profiles, to examine potential geographic influences, and to make general inferences with regard to risk factors on a national scale. For the cluster analysis, the partitioning around medoids (Kaufman and Rousseeuw, 1990) clustering algorithm was applied on the standardized (z-transformed) catchment medians, and their dissimilarity was measured using Euclidean distances. From the distribution of clusters (Figure 2), a spatial pattern is observable, e.g. a distribution of catchments in cluster 3 mainly along the West coast, in cluster 4 along the East, in the South and on Orkney, and cluster 5 being restricted to the Northeast.

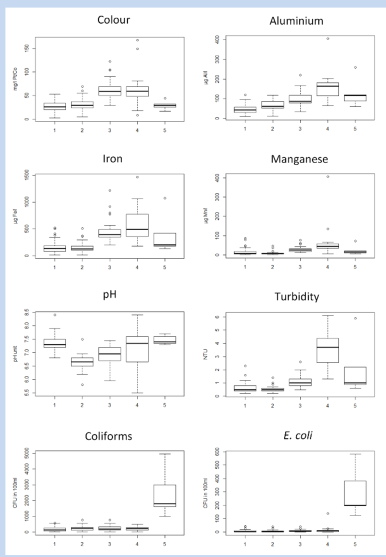

Figure 3: Distribution of catchment water quality concentrations per cluster.

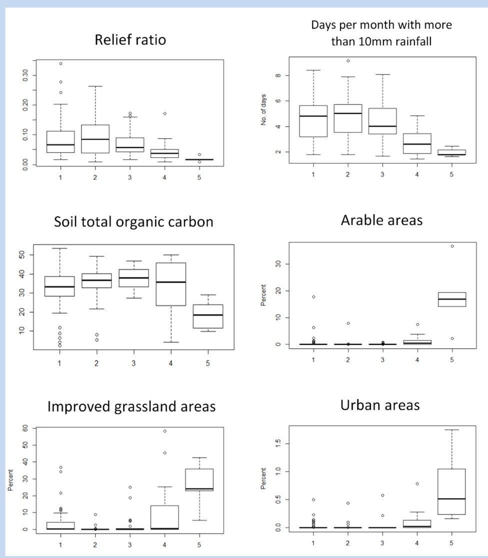

Figure 4: Distribution of catchment characteristic values per cluster.

Concentrations (Figure 3) were overall statistically significantly different between clusters (p<.001). There were also significant differences between catchment characteristics per cluster (Figure 4).

Multi-target Predictive Clustering Trees

In order to formally model the relationships between catchment characteristics and water quality parameters, regression-type trees were fitted as predictive models of water quality variables from catchment characteristics. The software package Clus (Struyf, Ženko, Blockeel, Vens & Džeroski, 2001) was used to produce multi-target predictive clustering trees (Figure 5). These models produced a predicted value for medians as well as 95th percentiles using catchment characteristics as predictors.

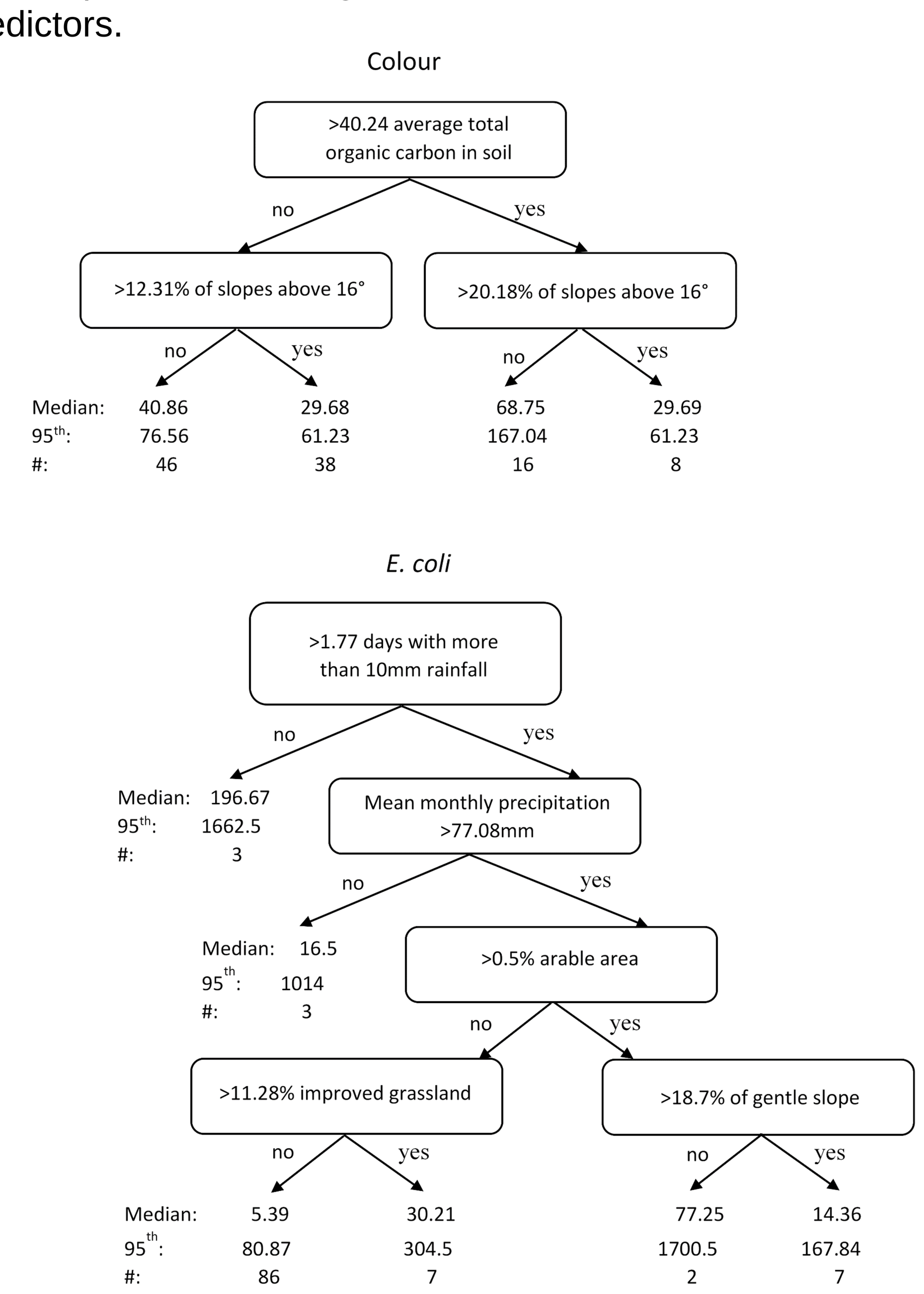

Figure 5: Multi-target predictive clustering trees for colour and *E. coli*. 'Median' shows the average median concentration for this group of catchments, '95th' is the average value for the 95th percentile, and '#' is the number of catchments per group. Root mean square errors: Colour 17.07, 39.06; *E. coli* 29.51, 341.15

Discussion

Catchments with comparatively high median bacteria concentrations, as well as catchments with elevated median colour concentrations could be identified in the PCA and associated with land use practices and soil organic carbon content, respectively. The cluster analysis reflected a general West vs. East split and a clear distinction of a small number of catchments with high bacteria concentrations in the Northeast, a region with comparatively high agricultural use. The MTPCTs for colour and *E. coli* reinforced the association with soil organic content and agricultural land uses, respectively. It is estimated that climate change will lead to intensified agricultural uses in parts of Scotland (Brown *et al.*, 2010), and that colour release from Scottish peatlands will rise further (Evans, Monteith & Cooper, 2005), exacerbating these issues for raw water quality.

Conclusion

The water quality issues identified for Scottish drinking water catchments highlight the importance of good land management practices, in terms of both reducing existing pressures but also to build additional resilience to cope with the expected increased exposure to extreme events. At a strategic level, those catchments can in a next step be identified that require additional attention, e.g. through increased monitoring or through targeted interventions, to facilitate ecosystem restoration for degraded habitats, or to avert detrimental changes in the first place.

Acknowledgements:
This research is funded by the Scottish Government through the Hydro Nation Scholars Programme.

References:
ASC (2016). UK Climate Change Risk Assessment 2017 Evidence Report – Summary for Scotland. Adaptation Sub-Committee of the Committee on Climate Change, London [pdf]. https://www.theccc.org.uk/wp-content/uploads/2016/07/UK-CCRA-2017-Scotland-National-Summary.pdf [06/07/2018].
Brown, I., Poggio, L., Gimona, A., & Castellazzi, M. (2010). Climate change, drought risk and land capability for agriculture: implications for land use in Scotland. Regional Environmental Change, 11, 503-518
Evans, C. D., Monteith, D. T., & Cooper, D. M. (2005). Long-term increases in surface water dissolved organic carbon: Oberservations, possible causes and environmental impacts. Environmental Pollution 137 (2005) 55-71
Kaufman, L. & Rousseeuw, P.J. (1990). Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York
Selle, B., Schwientek, M., & Lischeid, G. (2013). Understanding processes governing water quality in catchments using principal component scores. Journal of Hydrology 486 (2013) 31-38
Struyf, J., Ženko, B., Blockeel, H., Vens, C., & Džeroski, D. (2011). CLUS: User's Manual. https://www.researchgate.net/publication/265243854_Clus_User%27s_Manual [03/06/2018].