



Spatial-stream-network models as novel tools for predicting catchment-scale spatial patterns of microbial water quality

Aaron Neill (1,2), Doerthe Tetzlaff (1,3,4), Norval Strachan (5), Rupert Hough (2), Lisa Avery (2), Helen Watson (2), and Chris Soulsby (1)

(1) University of Aberdeen, Northern Rivers Institute, School of Geosciences, Aberdeen, United Kingdom (aaron.neill@abdn.ac.uk), (2) The James Hutton Institute, Craigiebuckler, Aberdeen, United Kingdom, (3) IGB Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany, (4) Humboldt University Berlin, Berlin, Germany, (5) University of Aberdeen, School of Biological Sciences, Aberdeen, United Kingdom

Decisions regarding the management and mitigation of faecal contamination often start at the catchment scale. However, historical paucity of long-term data relating to faecal indicator organisms has caused the development of catchment-scale models of microbial pollution to lag behind other water quality models. Spatial-stream-network models (SSNMs) represent an advancement in geostatistical methods that allow spatial autocorrelation between observations along stream networks to be accounted for within a linear mixed-modelling approach. Here, we present the novel application of SSNMs to an 11-year dataset of concentrations of *E. coli* at 10 spatially-distributed sites in a 52 km² mixed land-use catchment in the NE of Scotland. This demonstrates the potential value such models have in understanding catchment-scale microbial water pollution.

SSNMs were developed for the log₁₀-transformed 5th, 50th and 95th percentile concentrations of *E. coli* at each site to represent low, average and high levels of faecal contamination. For comparison, more traditionally-used multiple linear regression models (MLRMs) based on land-cover proxies for contamination sources were also developed. The best SSNM for each percentile was then used to predict concentrations with associated uncertainty at 200 m intervals along the stream network. For each percentile, the best SSNM included an anthropogenic impact index, an indicator of contamination potential due to humans, as the sole fixed-effect, and spatially-autocorrelated random effects modelled by a “tail-up” linear-with-sill autocovariance structure. Inclusion of the latter allowed almost all of the variance in the observed data to be explained by the SSNMs, compared with 66% to 78% explained by the best MLRMs. By explaining more of the observed variance, SSNMs had value in providing more robust predictions of catchment-wide concentrations of *E. coli*. This aided the identification of potential “hot spots” of contamination and areas where contamination was reduced that predictions made by the MLRMs may not have captured. The inclusion of spatial autocorrelation in the SSNMs also allowed stream network interactions to be inferred, such as the downstream attenuation of *E. coli* inputs from a waste water treatment works during periods of high contamination. The SSNMs in this study may have been limited by the relatively sparse spatial nature of the *E. coli* dataset used in their development. However, the potential for such models to better characterise “hot spots” of contamination and offer insights into how these affect water quality, together with providing estimates of uncertainty for predictions, is likely to be of value to both stakeholders and researchers. In particular, one could use predictions made by SSNMs to more reliably target mitigation strategies, whilst estimates of uncertainty may be used to inform future monitoring efforts to better assess faecal contamination in the most uncertain areas of the stream network. Furthermore, both predictions and their associated uncertainty can be presented in a visually effective manner that will likely aid in communicating model results to a range of audiences. As such, it is likely that SSNMs have potential for wider application in catchment-scale microbial water quality studies.