Understanding and modelling long-term spatio-temporal dynamics of faecal contamination in a mixed land-use stream network

Aaron J Neill[1], Doerthe Tetzlaff[1], Norval Strachan[2], Rupert Hough[3], Chris Soulsby[1]
[1]Northern Rivers Institute, University of Aberdeen, AB24 3UF; [2] School of Biological Sciences, University of Aberdeen, AB24 2TZ; [3]The James Hutton Institute, Aberdeen, AB15 8QH
Email: aaron.neill@abdn.ac.uk
www.crew.ac.uk/hydro-nationscholars

Scottish Government
gov.scot
Hydro Nation Scholars Programme

Introduction

- Maintaining microbiological quality of water sources used to supply drinking water is critical[1]. This is a key issue in rural areas where drinking water is often sourced from poorly-treated private supplies, yet risk of faecal contamination of water courses is elevated (e.g. due to high grazing densities, open farm yards and rudimentary sewage treatment systems[2]).

- Policy development and effective mitigation can be aided by increasing understanding of spatio-temporal dynamics of faecal contamination, and models which can make continuous predictions along river networks[1,3].

Aims (Work Package 1):

For a mixed land-use stream network, 1) Investigate spatio-temporal dynamics of faecal contamination using a long-term dataset; 2) Assess whether stream network spatial linear models (SLMs)[4,5] accounting for autocorrelation between flow-connected sites can improve prediction of long-term faecal contamination compared to simple multiple linear regression (MLR) models.

Results: Table and Figures

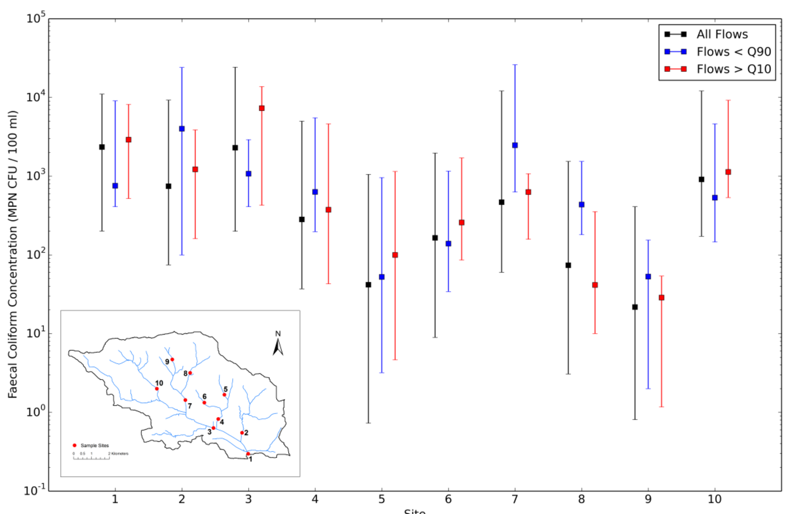

Fig 2: The 5^{th}, 50^{th} and 95th percentile most probable number (MPN) concentrations of faecal coliforms for each sample site for all, high and low flows. 50^{th} percentile is shown by square markers, 5th and 95th percentiles are given by line caps

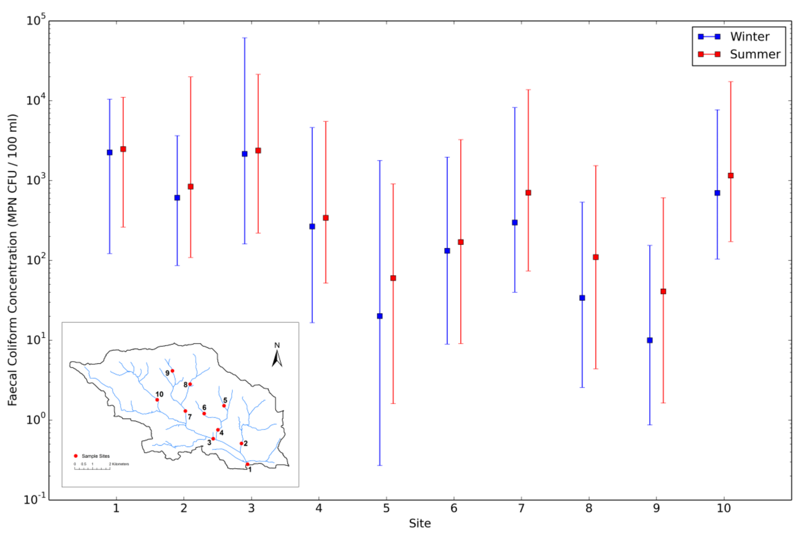

Fig 3: The 5^{th}, 50^{th} and 95th percentile most probable number (MPN) concentrations of faecal coliforms for each sample site for summer and winter. 50^{th} percentile is shown by square markers, 5th and 95th percentiles are given by line caps

Response Variable	Spatial Covariance Structure	Fixed Effects Parameters: L_{All}	Fixed Effects Parameters: Intercept	Spatial Covariance Parameters: Nugget (θ_0)	Partial Sill (θ_1)	Range (θ_2)	Performance: [a] LOOCV RMSPE (log_{10} MPN CFU / 100 ml)
log_{10} 50^{th}	None	* 0.7059	* 1.5012	0.126	-	-	0.371
	Exponential	* 0.691	* 1.5745	0.007	0.118	33745.58	0.317
	Spherical	* 0.6948	* 1.5817	0.004	0.113	33745.58	0.310
	Linear-with-sill	* 0.6997	* 1.5873	0.000	0.129	33745.58	0.304
log_{10} 5^{th}	None	* 0.889	0.1089	0.330	-	-	0.600
	Exponential	* 0.93094	0.09394	0.000	0.361	10723.51	0.626
	Spherical	* 0.96508	0.04557	0.000	0.369	8482.14	0.621
	Linear-with-sill	* 0.9909	0.01558	0.000	0.384	7361.53	0.616
log_{10} 95^{th}	None	* 0.5337	* 2.9012	0.125	-	-	0.390
	Exponential	* 0.5579	* 2.8854	0.000	0.134	8110.68	0.394
	Spherical	* 0.582	* 2.8525	0.000	0.140	7712.88	0.390
	Linear-with-sill	* 0.5884	* 2.843	0.000	0.142	5926.96	0.386

*significant with at least $p < 0.05$; [a] Leave one out cross validation, root mean squared predictive error

Table 1: Summary of predictive statistics for models predicting 5^{th}, 50^{th} and 95th percentile concentrations of faecal coliforms (log_{10} MPN CFU / 100 ml)

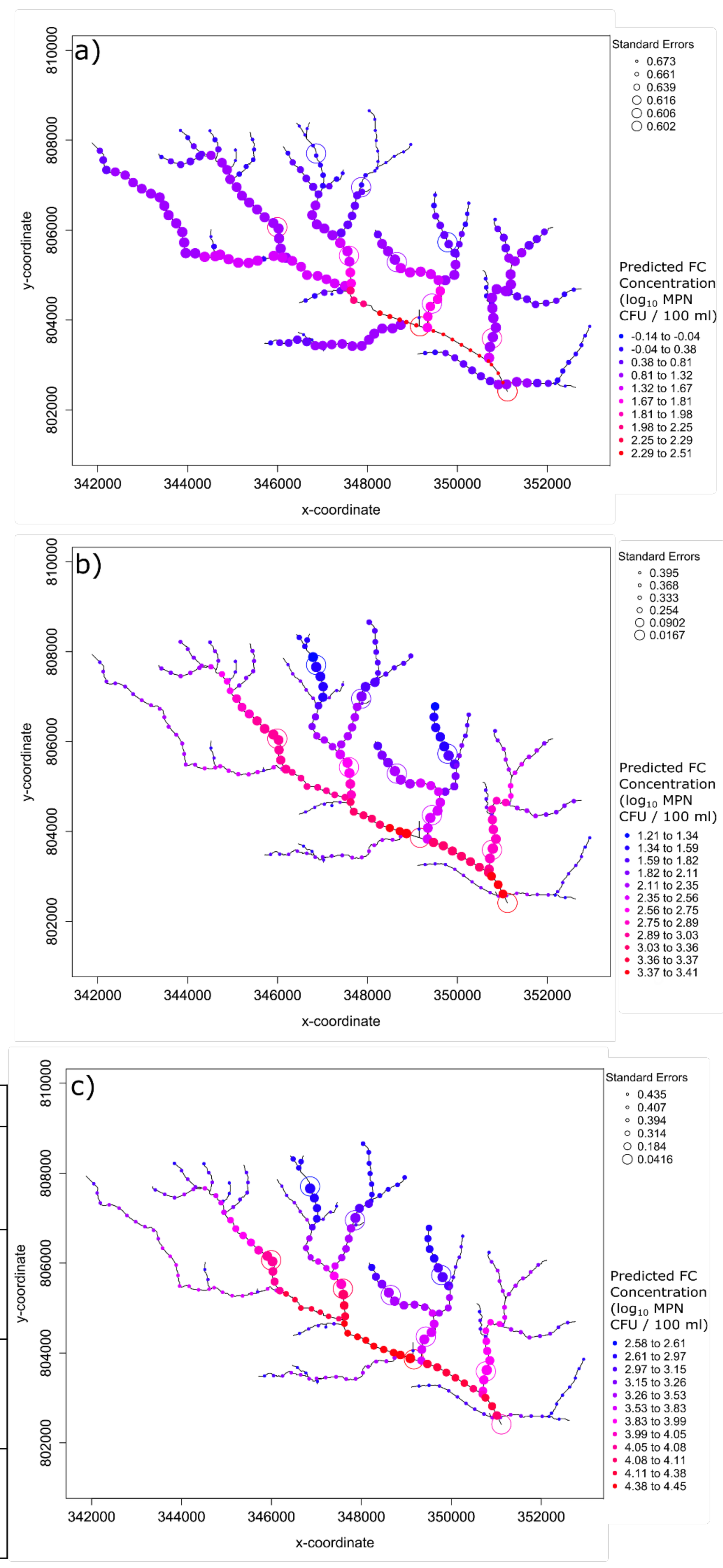

Fig 4: Predictions of a) 5th percentile; b) 50^{th} percentile; c) 95th percentile concentrations of faecal coliforms. Filled circles are prediction sites, with circle size representing the standard error of prediction. Large open circles are observed sites.

Methods

Study Site: The Tarland Burn catchment ($52km^2$). Mixed land-use, brown earth and humus-iron podzols, main settlement is Tarland Village (pop. 650).

Data: Faecal coliform (FC) concentrations at 10 sample sites (Fig 1) collected at monthly to 3 monthly intervals over 11 water years.

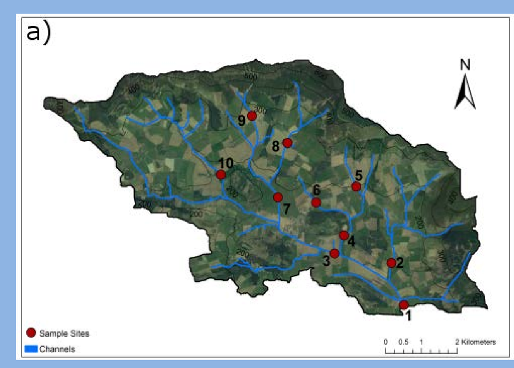

Fig 1: Tarland Burn catchment showing topography and locations of the 10 long-term sample sites

Spatio-temporal dynamics: Concentrations at each site summarised in relation to flow (all flows , high flows >Q10 and low flows <Q90) and season (summer: April – September).

Multiple linear regression modelling: Backwards stepwise procedure used to fit models predicting log_{10} 5^{th}, log_{10} 50^{th} and log_{10} 95^{th} all-flow FC concentrations, with % pasture, % arable and log_{10} Anthropogenic Impact Index (L_{All}) as initial predictors.

Stream network spatial linear modelling: Significant predictors from MLR models used as fixed effects. Flow-connected spatial autocorrelation modelled by fitting spatial covariance structure[4,5].

Results

FC – Flow Dynamics (Fig 2): No clear relationship between flow and concentrations of FC. However, low and high flow samples composed only 7.5% and 10% of all samples, respectively. Spatially, sites 1 and 3 have highest median concentrations during all and high flows, whilst sites 2, 3 and 7 have the highest during low flows.

FC Seasonal Dynamics (Fig 3): Lack of strong relationship between concentrations of FC and season, however concentrations tend to be elevated in summer. Spatially, sites 1 and 3 have highest concentrations in summer and winter.

Modelling long-term faecal contamination: L_{All} the only significant predictor in MLR models. SLMs only improved prediction of log_{10} 50^{th} and log_{10} 95^{th} concentrations (Table 1). Spatial pattern of MLR predictions of log_{10} 5^{th} percentile concentrations simply reflects distribution of L_{All}, with step-changes occurring with large increases in L_{All} (Fig 4a). SLM predictions of log_{10} 50^{th} and log_{10} 95^{th} concentrations (Fig 4b and 4c) change more smoothly, and spatial patterns are governed by distributions of L_{All} and residuals from fixed-effects part of the models. SLM predictions give greater accuracy and therefore confidence in identifying hot spot areas of contamination.

Future

Results presented here from Work Package (WP) 1 give a coarse-scale understanding of dynamics of faecal contamination in the Tarland. WP2 will adopt a smaller-scale, process-based approach, with isotope tracers, DNA sequencing and modelling being integrated to identify specific sources of faecal contaminants and the hydrological flow paths connecting them to sources for private water supplies. WP3 will use a tracer-aided model to investigate dynamics of faecal contamination in a montane catchment.

References
1: Kay *et al*. (2008) Quantification and control of microbial pollution from agriculture: a new policy challenge. *Environmental Science and Policy.* 11: 171-184
2: Tetzlaff *et al*. (2012) Land use and hydroclimatic influences on Faecal Indicator Organisms in two large Scottish catchments: Towards land use-based models as screening tools. *Sci Total Environ.* 434: 110-122
3: Oliver *et al*. (2016) Predicting microbial water quality with models: Over-arching questions for managing risk in agricultural catchments. *Sci Total Environ.* 544: 39-47
4: Ver Hoef *et al*. (2006) Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics.* 13: 449-464
5: Ver Hoef *et al*. (2010). SSN : An R Package for Spatial Statistical Modelling on Stream Networks.

Northern Rivers Institute | UNIVERSITY OF ABERDEEN | The James Hutton Institute | crew Scotland's centre of expertise for waters