RELATIONSHIPS BETWEEN STREAMWATER *E. COLI* CONCENTRATIONS AND ENVIRONMENTAL FACTORS IN NEW ZEALAND

R. Collins

National Institute of Water and Atmospheric Research Limited (NIWA)
PO Box 11 115, Hamilton, New Zealand (R.Collins@niwa.co.nz)

ABSTRACT

Stream water *E. coli* concentrations, collected from 73 sites across the Waikato region, North Island of New Zealand, have been analysed in conjunction with a range of environmental variables. Faecal contamination of streams and rivers occurs throughout the region and median *E. coli* concentrations range between 1 and 1,300 cfu/100 mL. At 75% of the sites, the median concentration exceeds the national guideline of 126 cfu/100 mL for contact recreation. The pattern of faecal contamination across the region is strongly influenced by grazing livestock. No strong correlation is apparent between median *E. coli* and point sources; this is attributed to a relatively low number of consented discharges that can cause faecal contamination. The amount of poorly drained soil within a watershed explains a relatively large proportion of the variation in streamwater faecal contamination. Median turbidity is a relatively strong correlate of median *E. coli* across the region and may be a useful surrogate variable for *E. coli* at a given site. A statistical model to determine median *E. coli* concentrations as a function of environmental factors explains c. 70% of the observed variance.

Keywords faecal contamination; grazing livestock; soil drainage; turbidity

INTRODUCTION

The sources of faecal contamination of freshwaters are diverse and include grazing livestock (Baxter-Potter and Gilliland 1988), wild mammals and birds (Niemi and Niemi 1991), and point source discharges of wastewater (Smith et al. 2001). The impact of each source of faceal contamination upon freshwater is potentially influenced by a range of interacting environmental factors that include the physical characteristics of a catchment and the land management practices within it. However, understanding of the interactions between sources of contamination and the catchment processes acting upon them remains incomplete. This, in turn, limits the ability to identify and apply land management strategies to minimise faecal contamination. In an attempt to advance understanding of faecal contamination processes, stream water *E. coli* concentrations, collected in a regional water quality monitoring program, were analysed in conjunction with a range of environmental variables. The microbial data was collected from 73 stream sites throughout the Waikato region, North Island, New Zealand, that encompass a diverse range of contaminant sources, catchment characteristics and land management practices. As part of the analysis a statistical model has been developed to relate median *E. coli* concentrations across the region to basic environmental data, and to other water quality variables (turbidity).

STUDY REGION

The Waikato region covers approximately 25,000 km^2 of the central North Island of New Zealand. Pastoral farming dominates the region although large areas of plantation forest and indigenous vegetation are also evident. Sheep and beef cattle are farmed on the steeper hillslopes, but dairying dominates and is most intensive in the centre of the region near the city of Hamilton. The Waikato has a mean annual rainfall of 1,250 mm, but this ranges between 1,000 and 3,000 mm. Soils with contrasting drainage characteristics are found within the Waikato, ranging from free draining volcanic soils to poorly drained peats.

WATER QUALITY DATA

Streamwater *E. coli* and turbidity concentrations were provided by Environment Waikato, the regional environmental management authority. Quarterly spot samples have been collected since 1988 at each of the 73 sites across the region. *E. coli* analysis is done by membrane filtration with a count on MFC agar at 44.5°C after 24 hours, following the American Public Health Association method 9222G. Mean *E. coli* concentrations are often markedly higher than median values reflecting a positive skewness of the data distributions. Analysis is therefore focused upon the median values as they better describe the central tendency of the data. Median concentrations range from 1 to 1300 cfu/100mL with the highest values generally associated with the most intensive dairy farming in the centre of the region. The lowest median values are generally found in forested catchments bordering Lake Taupo in the south. Of the 73 sites, median values at 53 of them exceed the guideline for freshwater recreation (126 cfu/100mL). The range in values found at a site can be extremely large, reflecting marked temporal variation in factors influencing faecal contamination, including point sources, livestock grazing patterns, stream-flow, and die-off. Median turbidity varies markedly across the region, ranging from <1 to 55 Nephelometric Turbidity Units (NTU). This range reflects the diversity of soil character (especially clay content) in the region, and the diversity of processes determining hillslope and channel sediment dynamics.

METHODOLOGY

From a 30 m resolution digital elevation model the catchment boundaries upstream of each of the 73 sampling points were delineated using a GIS. These boundaries were then used to determine summary statistics of land use, livestock numbers, soil drainage properties, slope angle, and rainfall, for each catchment. Land management practices within each catchment were determined from datasets describing the location and volume of licensed (and not necessarily actual) effluent discharges to land, and point source discharges direct to water. Both types of discharge were divided into 3 categories based on their source; dairy farms, non-dairy agriculture, and utilities, including treated sewage wastewater. The catchments range from 2 to 2270 km^2 in size and cover approximately 50% of the Waikato region. They encompass a range of land uses, soil types, and potential sources of faecal contamination.

Bivariate relationships were used to examine the impact of each spatially varying environmental factor upon median *E. coli* across the region. Pearson (r) and Spearman Rank (r_s) correlation coefficients are given in Table 1 for each relationship and correlation (r) between the environmental factors is given in Table 2. The site with the lowest median *E. coli* concentration (1 cfu/100 mL) was excluded from the statistical analysis since geothermal water contributes to stream-flow, enhancing die-off.

RESULTS AND DISCUSSION

Livestock

Grazing animals are an important causal factor in the faecal contamination of Waikato streams. This is illustrated (Figure 1) through a comparison of median *E. coli* concentrations in streams draining pastoral catchments with those draining non-pastoral vegetation. Furthermore, a correlation exists between median *E. coli* concentration and the percentage of pastoral land within a catchment, across the region (r = 0.50, r_s = 0.62). This relationship strengthens (r = 0.64, r_s = 0.65) when cattle density is used as the independent variable, possibly because cattle are attracted to water, depositing faecal material directly to streams.

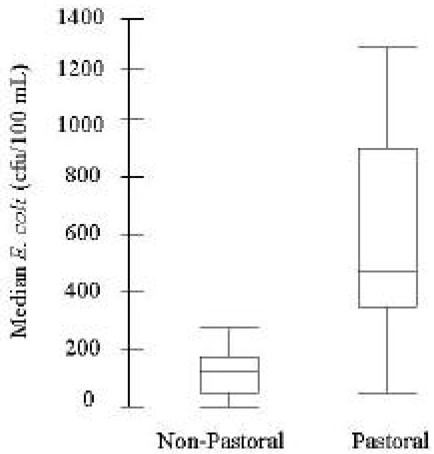

***Figure 1**. Boxplots illustrating median* E. coli *statistics from sites characterised by > 90% non-pastoral vegetation, and those characterised by > 90% pastoral vegetation. The outlined central box depicts the middle half of data between the 25th and 75th percentiles, the horizontal line marks the median, whiskers extend from the top and bottom of the box to depict the extent of the main body of data.*

Non-Pastoral Vegetation

Catchments characterised by non-pastoral vegetation exhibit relatively low median *E. coli* concentrations (Figure 1), and an inverse relationship exists (r = -0.61, r_s = -0.69) between the percentage of non-pastoral land and median *E. coli*, across the region. The absence of livestock from such land is likely to be the principal reason for this correlation. Non-zero *E. coli* concentrations are observed, however, even in fully forested catchments, and this can probably be attributed to birds and wild mammals.

Soil Drainage

A relatively strong correlation exists between median *E. coli* and the percentage of a catchment characterised by poorly drained soil (r = 0.78, r_s = 0.64). There are two possible explanations for this. Firstly, the hydrological characteristics of these soils promotes the generation of a relatively large volume of surface runoff during rainfall, which can rapidly transport entrained faecal material to surface waters. Secondly, the correlation may arise because of artificial drainage. The analysis of soil drainage properties in this study did not account for the presence of subsurface drains, often installed under otherwise poorly drained dairying land. Although these drains reduce ponding and overland flow of surface water, they may also act to provide a flowpath by which faecal contamination can be rapidly transported to the channel network. An inverse relationship is apparent between median *E. coli* and the percentage of well-drained soils (r = -0.51, r_s = -0.61) in a catchment. Well-drained soils will minimise the generation of surface runoff, promoting a vertical movement of water and microbes through the soil. This reduces contamination since the soil matrix is generally effective at filtering particulates in soil water.

Rainfall and Slope Angle
Both mean annual rainfall and the percentage of steep land (>15°) show an inverse relationship with median *E. coli*. This is due, in part, to the relative lack of pastoral land (as reflected by the negative correlation coefficients in Table 2) upon steep slopes and areas of high rainfall.

Urban Land
The presence of urban areas appears to correlate relatively weakly (r = 0.42, r_s = 0.44) with median *E. coli* concentrations. Urban runoff is known to have relatively high levels of faecal contamination, attributed, for example, to bird droppings and dog faeces. Furthermore, point source discharges from sewage treatment plants (and some industries) within urban areas are likely to provide an input of faecal contamination direct to a stream or river. Few watersheds within the Waikato monitoring program, however, are characterised by much urban land.

Point Source Discharges
No strong correlation is apparent between median *E. coli* and the 3 categories of point sources across the region. The consented volume of non-dairy point sources apparently provides the strongest linear relationship (Table 1) primarily due to a very large discharge at one site only. Excluding this site from the bivariate relationship extinguishes the correlation between median *E. coli* and non-dairy point sources. The direct discharge of dairy effluent to surface water is relatively low across the region. This reflects a policy in recent years to reduce the direct discharge of effluent to surface water, and is the likely reason for the lack of a strong correlation (Table 1) with median *E. coli* concentrations. The lack of a strong relationship between point source discharges from utilities and median *E. coli* probably reflects recent improvements in the treatment of sewage and industrial wastewater (Vant 2001).

Effluent Discharge to Land
No strong correlation is apparent between the 3 sources of effluent discharge to land, and median *E. coli* across the region. Due to incomplete records this variable was necessarily formulated in terms of the number of occurrences of discharge to land within each watershed, rather than the volume of discharge. This may have masked stronger correlation with the faecal contamination of streams.

Turbidity
Median turbidity is a relatively strong predictor (r = 0.73, r_s = 0.77) of median *E. coli*, across the region. Correlation is to be expected given that both stream bed-sediments and the microbes settled within them are subject to entrainment (Nagels et al. 2002) as flow velocity increases. In addition, the processes by which overland flow detaches and transports soil particles on the hillslope broadly apply to faecal material.

STATISTICAL MODELLING
Multiple regression was used to examine relationships between median *E. coli* and environmental factors, and to attempt to derive a statistical tool with which to predict median concentrations across the region. Median *E. coli* concentration was used as the dependent variable, and the environmental factors as the independent variables. The strength of relationships was assessed using the coefficient of determination (r^2), expressed as a percentage and adjusted for the number of predictors in the regression equation. All environmental factors were examined in an interactive stepwise selection procedure, using DataDesk software (http://www.datadesk.com). During this process independent variables were retained even if they were correlated with other independent variables in the model. However, DataDesk checks for severe co-linearity and aborts the regression computation, if necessary. The analysis derived a predictive model whereby four factors together explained 69 % of the variance in median *E. coli* across the region. Each factor, and the partial r^2 associated with its addition to the regression model, is given in Table 3.

The factors were: the percentage of land with poorly drained soil (*PoD*), median turbidity (*Turb*), cattle density (*Catt*), and the percentage of well drained soil *(WeD)*, providing the following relationship, where *Ec* is median *E. coli*.

$$Ec = 188 + (3.8\,PoD) + (8.6\,Turb) + (0.19\,Catt) - (1.27\,WeD) \tag{1}$$

Co-linearity is apparent between the independent variables within the models (Table 2), notably between the percentage of poorly drained soil and median turbidity (r = 0.79). The strength of this co-linearity means that the model should not be used to draw inferences about the relative contributions to median *E. coli* concentrations made by each of the independent variables. The equations will be most reliable for predicting *E. coli* concentrations in unmonitored catchments where the relationships between independent variables are similar to those in the original dataset. This is likely to hold in most places throughout the Waikato Region, but may not apply elsewhere in New Zealand.

Table 1. Summary of the Pearson (*r*), and Spearman Rank correlation coefficients (r_s) between median *E. coli* and each independent variable.

Abbrev.	Independent variables	Units	*r*	r_s
Past	% Pasture	dimensionless	0.50	0.62
NPV	% Non-Pastoral Vegetation	dimensionless	-0.61	-0.69
Catt	Cattle density	stockunits km^{-2}	0.64	0.65
PoD	%Poorly drained soil	dimensionless	0.78	0.64
WeD	%Well drained soil	dimensionless	-0.51	-0.61
Rain	Rainfall	mm yr^{-1}	-0.37	-0.32
Steep	% Steep slope	dimensionless	-0.54	-0.34
Urb	%Urban	dimensionless	0.42	0.44
DP	Density of Dairy Point Source Volumes /km^2	m^3 d^{-1} km^{-2}	0.38	0.57
NDP	Density of Non-Dairy Point Source Volumes /km^2	m^3 d^{-1} km^{-2}	0.39	0.34
	Density of Utility Point Source Volumes /km^2	m^3 d^{-1} km^{-2}	-0.07	0.36
	Density of Dairy effluent discharges to land	km^{-2}	0.22	0.35
NDE	Density of Non-Dairy effluent discharges to land	km^{-2}	0.47	0.48
	Density of Utility effluent discharges to land	km^{-2}	-0.05	0.15
Turb	Median turbidity	NTU	0.73	0.77

Table 2. Correlation (r) between the key independent variables.

	Past	NPV	Catt	PoD	WeD	Rain	Steep	Urb	DP	NDP	NDE
Past	1.00										
NPV	-0.99	1.00									
Catt	0.78	-0.78	1.00								
PoD	0.53	-0.58	0.65	1.00							
WeD	-0.31	0.33	-0.23	-0.52	1.00						
Rain	-0.40	0.43	-0.48	-0.45	-0.06	1.00					
Steep	-0.51	0.56	-0.65	-0.70	0.01	0.60	1.00				
Urb	0.10	-0.25	0.17	0.46	-0.23	-0.21	-0.38	1.00			
DP	0.50	-0.48	0.61	0.35	0.01	-0.14	-0.47	0.00	1.00		
NDP	-0.02	0.02	0.06	-0.05	0.13	-0.03	0.00	0.00	-0.01	1.00	
NDE	0.37	-0.35	0.60	0.42	-0.10	-0.31	-0.45	-0.04	0.32	0.09	1.00
Turb	0.41	-0.45	0.44	0.79	-0.49	-0.28	-0.46	0.36	0.26	-0.06	0.28

Table 3. The statistical model for median *E. coli*

Variable	Coefficient	Partial r^2	Comments
Constant	188		Intercept
PoD	3.8	60.4	% of poorly drained soil
Turb	8.6	63.5	Median turbidity
Catt	0.19	67.3	Cattle stock units
WeD	-1.27	68.8	% of well drained soil

CONCLUSIONS

Analysis of the *E. coli* dataset for the Waikato region in New Zealand has demonstrated that faecal contamination of streams and rivers occurs widely, even in forested catchments. Median concentrations range from 1 to 1,300 cfu/100 mL and at 53 of the 73 sites sampled the median concentration exceeded the national guideline of 126 cfu/100 mL for contact recreation.

Examination of the microbial dataset in conjunction with a range of environmental factors, including land use and soil type, has highlighted the key controls upon faecal contamination. The pattern of contamination across the Waikato is strongly influenced by the presence of grazing livestock, and the highest median *E. coli* concentrations are generally associated with the most intensive dairy farming in the centre of the region. Strategies to reduce faecal contamination of streams and rivers in the region need therefore to focus upon grazing livestock. Cattle access to streams and near-channel areas is likely to be important in determining the level of faecal contamination and mitigation measures may therefore be best directed at riparian zones.

Generally, there is no strong correlation between point source discharges and faecal contamination across the region. This is attributed to an environmental policy in recent years to reduce the number of consented discharges, and to a relatively high standard of wastewater treatment. Strong correlation was not apparent between median *E. coli* and the number of consented discharges to land, from which it is tentatively concluded that land disposal is not a major source of microbial contamination.

Soil drainage properties explain much of the variation in streamwater faecal contamination and the percentage of poorly drained soil within a watershed is a key factor within a statistical model of median *E. coli* across the region. Two possible mechanisms giving rise to this correlation are enhanced surface flow, and artificial drainage. It is not possible to determine which mechanism predominates in an analysis of this kind and experimental studies are desirable. Nevertheless, the strength of poorly drained land as an explanatory variable has important implications for land management, suggesting that bacterial water quality on poorly drained land would benefit from (1) targetted riparian protection to maximise filtering of faecal material within overland flow and, (2) interception of subsurface drainage flows, for example, in constructed wetlands.

Median turbidity is a relatively strong predictor of median *E. coli* across the region. This indicates that the processes mobilising fine sediment, both on the hillside and in-stream, apply also to bacteria. Turbidity may be a useful surrogate variable for *E. coli* at a given site.

A statistical model for median *E. coli* concentrations in the region, explains almost 70% of the observed variance. The percentage of poorly drained soil, the density of cattle within a catchment, and median turbidity, are the 3 key factors incorporated into the model. Strong co-linearity is apparent between the independent variables within the model, but the equations are likely to have utility in predicting *E. coli* concentrations in unmonitored catchments in the Waikato region.

ACKNOWLEDGEMENTS

This study was funded by the New Zealand Ministry of Agriculture and Forestry (MAF) and Phil Journeaux (MAF) is thanked for his support of the study. Todd Williston (NIWA) delineated the watersheds and summary statistics for them were provided by Environment Waikato Regional Council. Some of these statistics were derived from datasets purchased by Environment Waikato under formal licensing agreements with external agencies (Terralink New Zealand Limited, AgriQuality New Zealand Limited and Landcare Research New Zealand Limited). Len Brown and Bill Vant at Environment Waikato are thanked for providing data and contributing to the development of the methodology, respectively. Rob Davies-Colley, Kit Rutherford and Graham McBride (NIWA) provided constructive review comments.

REFERENCES

Baxter-Potter, W, R. and Gilliland, M.W. (1988). Bacterial pollution from agricultural lands. *J. Environ. Qual.* **17**, 27-34

Nagels, J.W., Davies-Colley, R.J., Donnison, A.M. and Muirhead, R.W. (2002). Faecal contamination over flood events in a pastoral agricultural stream in New Zealand. *Wat. Sci. Tech.* **45**, 12, 45-52

Niemi, R.M. and Niemi, J.S. (1991). Bacterial pollution in pristine and agricultural lands. *J. Environ. Qual.* **20**, 620-627

Smith, J.H., Wickham, J.D., Norton, D., Wade, T.G. and Jones, K.B. (2001). Utilization of landscape indicators to model potential pathogen impaired waters. *J. American Wat. Res. Assoc.*, **37**, 805-814.

Vant, W. N. (2001). New challenges for the management of plant nutrients and pathogens in the Waikato River, New Zealand. *Wat. Sci. Tech.* **43**, 137-142