Environmental Science Processes & Impacts

PAPER

View Article Online
View Journal

CrossMark click for updates

Cite this: DOI: 10.1039/c4em00327f

Response of discharge, TSS, and *E. coli* to rainfall events in urban, suburban, and rural watersheds

H. J. Chen[a] and H. Chang*[b]

Understanding dominant processes influencing microorganism responses to storm events aids in the development of effective management controls on pathogen contamination in surface water so that they are suitable for water supply, recreation, and aquatic habitat. Despite the urgent needs at present, numerous facets of microbial transport and fate are still poorly understood. Using correlation and multiple regression combined with spatial analyses, this paper evaluates the relationship between antecedent precipitation and discharge, TSS, and *E. coli* concentrations, and examines correlations between *E. coli* and TSS, as well as whether and how those relationships change along an urban and rural gradient. The urban watershed exhibited a faster and stronger response of streamflow, TSS, and *E. coli* to precipitation mainly due to its higher degree of imperviousness. In general, TSS was significantly correlated with *E. coli* concentrations, which linearly decreased as % developed area increased, with large variation in regions with a high percentage of development, implying the more complex stormwater infrastructure and more variable pollutant sources of *E. coli* in the urban watershed. Seasonal differences for *E. coli* were noted. Specifically, summer showed a higher level of *E. coli*, which might be attributed to the higher temperature since *E. coli* is more likely to persist and grow in a warmer environment. Further multiple linear regression analyses showed the best *E. coli* prediction result for the largest, suburban watershed, using antecedent precipitation, TSS, and temperature as independent variables. The models are capable of explaining 60% and 50% of the variability in the *E. coli* concentration for the dry and wet season, respectively. The study not only provides more detailed and accurate characterization of the storm-period response of *E. coli* across an urban and rural gradient, but also lays a foundation for predicting the concentration of *E. coli* in practice, potentially suggesting effective watershed management decisions.

Received 12th June 2014
Accepted 14th July 2014

DOI: 10.1039/c4em00327f

rsc.li/process-impacts

Environmental impact

E. Coli is an increasing concern in many watersheds. Since elevated levels of *E. coli* can negatively affect water supply, recreation, and aquatic habitat, identifying possible sources and transport of *E. coli* has been a primary concern in environmental sciences and management. Here we investigated how the sources and transport mechanisms might differ across different levels of urban development during storm events. Identifying potential factors that affect *E. coli* concentrations can not only help control sources, but also provide information to better predict changing levels of *E. coli* as they relate to other pollutants or meteorological factors. We attempt to unravel the dynamics of *E. coli* concentration using a combination of meteorological and landscape factors during storm events.

Introduction

Pathogens are the number one cause of impairment for Clean Water Act (CWA) Section 303(d) listed waters in the USA and pose a significant threat on human health and quality of life.[1] Fecal coliform bacteria, particularly *E. coli*, are often selected as critical biological indicators of the presence of pathogens. The higher the level of *E. coli* density in water bodies, the greater possibility that the water has been polluted by feces associated with pathogens. Simulation models can play an important role in the assessment and management of microbial contamination. However, the development of such a model requires an accurate understanding of the transport, build-up and persistence of microorganisms in the catchment system.[2]

Recent studies have shown that levels of *E. coli* dramatically increased in response to storm events,[2–4] indicating that wash-off models, in which stormwater runoff serves as a contributor to pollutants in surface waters, will partially explain the considerable inter-event variability in *E. coli* concentrations. Moreover, a significant correlation was observed between *E. coli* and preceding rainfall events.[5–7] These findings suggest that the

[a]*Department of Land, Air and Water Resources, University of California, Davis. One Shields Avenue, Davis, CA 95616-5270, USA*

[b]*Department of Geography, Portland State University, PO Box 751 – Geog, Portland, Oregon 97207-0751, USA. E-mail: changh@pdx.edu*

antecedent rainfall conditions are likely to impact both the amount of water and energy available for *E. coli* transport and the amount of moisture present in a watershed that is critical for *E. coli* survival.[6] In a study of Stock Creek, Tennessee, however, there was no statistically significant correlation between *E. coli* and precipitation.[8] Therefore, a more comprehensive understanding of the relationship between rainfall and the presence of *E. coli* across different types of watersheds is needed, which is critical for managing water systems so that water managers are able to provide potable water and water suitable for recreation and aquatic habitat.

Apart from inter-event variation, *E. coli* concentrations also significantly vary by season.[6,9] Temperature, which is known to influence the survival of *E. coli*, has been used to help explain the variation of *E. coli* levels between dry and wet periods. Many studies have shown that die off rates for *E. coli* are higher as temperature increases, which could be attributed to stronger sunlight radiation in warmer seasons.[10–12] However, reverse trends were observed in several studies where *E. coli* levels were higher during summer seasons. These contradictory results are likely consequences of the summer seasons having the warmest temperatures and less streamflow, which corresponds to higher growth and survival rates of *E. coli* bacteria[9,13] and less dilution effect.[5] In addition, higher *E. coli* concentrations in urban areas may also be attributed to the increase in animal and human activities during the warm season.[14]

The majority of *E. coli* in aquatic systems is also associated with sediments, and these associations increase the survival rate of fecal bacteria relative to those in the water column.[15,16] Positive correlations between total suspended solids (TSS) and *E. coli* concentrations have been observed by Anderson and Rounds (2003)[13] and Hamilton & Luffman (2009)[7] for highly urbanized watersheds (Fanno Creek watershed, Oregon; Little River watershed, Tennessee), and Muirhead *et al.* (2004)[17] for artificial flood events in pasture land. These findings indicate that *E. coli* were either transported to stream bound to particulate matter, adsorbed onto resuspended stream bed particles, or they had an affinity for sediments in water.[18] Therefore, sediments have the potential to serve as a surrogate for *E. coli* concentrations. In two other studies, however, only a weak relationship between *E. coli* and TSS was observed in artificial flood events in the northern England[19] and on marsh lands at Texas coast.[20] Given the large spatial variation of the correlation between *E. coli* and TSS, site-specific landscape patterns within a given watershed may have important impacts on the *E. coli*-sediment relationship, which has not been fully characterized.

Regression analysis has been applied to predict *E. coli* concentrations and determine the nature and causes of its variability. Several studies have reported that *E. coli* concentrations are strongly related to antecedent precipitation, sediments, streamflow characteristics, temperature, and season.[7,18,21] Linear regression is one of the most commonly used statistical methods in water quality research. Rasmussen and others (2008)[22] have developed simple linear regression models to perform real time prediction for 19 constituents, including fecal coliforms in streams of Johnson County, northeast Kansas, using turbidity as the only explanatory variable. The R^2 values of their models ranged from 0.67 to 0.84.[22] Hamilton and Luffman (2009)[7] have achieved relative success in using multiple linear regression analysis ($R^2 = 0.565$) to predict the concentration of *E. coli* using precipitation, discharge, and turbidity as predictors. In a study conducted along tributaries of Tualatin River, Oregon, discharge and turbidity were selected as predictors in the regression equation.[18] The study showed moderately successful prediction for *E. coli* bacteria with reasonably high R^2 values (0.586–0.713). Although several studies have reported that antecedent precipitation, stream discharge, sediment density, and water temperature could serve as potential predictors for *E. coli* concentrations, it is still unclear how that pattern would change across an urban and rural gradient.

The objective of this study was to develop more detailed and accurate characterization of the storm-period response of *E. coli* in urban, suburban, and rural watersheds, potentially allowing effective management controls on pathogen contamination in these watersheds. Specifically, the goals are to:

(1) Evaluate the relationship between antecedent precipitation and the three other parameters: discharge, TSS, and *E. coli* concentrations.

(2) Determine the correlations between *E. coli* and TSS, as well as whether and how those relationships change along an urban and rural gradient.

(3) Investigate the variation of discharge, TSS, temperature and *E. coli* between dry and wet seasons.

(4) Determine the correlation between discharge, TSS, temperature, *E. coli* and land use type.

(5) Construct regression models to predict the concentrations of *E. coli* using antecedent precipitation, stream discharge, TSS and temperature as easy-to-measure proxies, and compare model significance and parameters among the three watersheds.

Study area

The sample sites were located in the Portland Metropolitan area (Fig. 1), which comprises Clackamas, Columbia, Multnomah, Washington, and Yamhill Counties. Climates of these counties are heavily influenced by the Pacific Ocean, and are characterized by mild, humid winter and hot, dry summer. Future climate is projected to be drier and hotter in summer and wetter in winter,[23] increasing the probability of droughts[24] and floods.[25] Being one of the fastest growing metropolitan areas of the USA,[26] ongoing urban development is posing water quality concerns, particularly in the urban–rural fringe area.[27] The land use varies from heavily developed urban areas in the middle part to rural and agricultural in the eastern and far western extents. We selected three watersheds that represent a gradient of urban development. They are Fanno Creek, Johnson Creek, and Balch Creek. We refer to Fanno Creek as urban, Johnson Creek as mixed, and Balch Creek as forested, according to their land use patterns.

Fanno Creek is a 24 km tributary of the Tualatin River and flows west from its headwaters in Hillsdale to its confluence with the Tualatin River near Durham, draining an area of

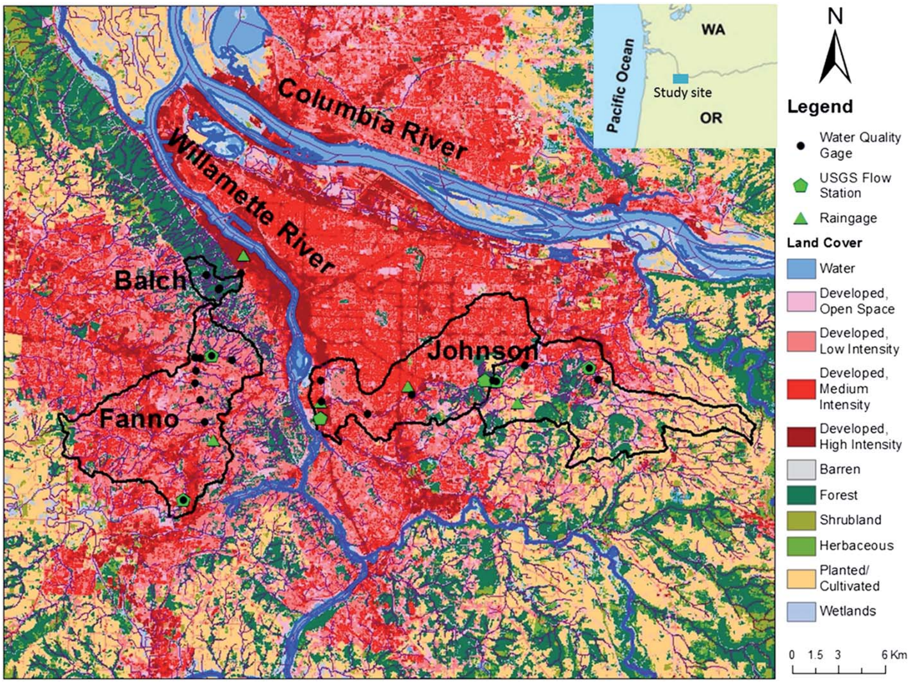

Fig. 1 Study area and location of sampling sites.

83 km^2. Fanno Creek was chosen as the study site because it serves as a representative for highly urbanized watersheds. Covering several major cities in Washington County including Portland, Beaverton, Tigard and Durham, the Fanno Creek watershed is highly developed with 84% urban land use.[28] Fanno Creek is also a typical stream system that is impaired by stormwater runoff from existing point sources and development[29] and is listed on the Oregon's 303(d) list (Table 1) due to high levels of ammonia, nutrients, total solids, and *E. coli* which exceeded the water quality standard in 50% of samples in summer and 25% during winter.[29] Although heavily polluted, the creek supports aquatic life in upper reaches and passes through or close to 14 parks in several jurisdictions, serving a recreational function for nearby residents. Despite tremendous efforts in stormwater management and increase in public awareness, no remarkable improvement in water quality in Fanno Creek was observed,[30] primarily owing to the high proportion of impervious surfaces as well as the loss of riparian areas in the watershed. The locations and land cover of the sample sites are shown in Fig. 1 and summarized in Tables 2 and 3.

Located on the eastern side of the Portland Metropolitan region, the Johnson Creek is a 40.2 km tributary of the Willamette River. Beginning at Clackamas County, east of Boring, it flows west and drains an area of approximately 137.6 km^2 with 180 000 residents. The Johnson Creek watershed is moderately developed with approximately 50% of the watershed urbanized, mostly within the urban growth boundary in the lower and middle reaches of the catchment. There are quite a few creek-side parks along the Johnson Creek, including sports fields, picnic areas, and trails. For example, the 20 mile spring water corridor, which runs in parallel with the Johnson Creek between the mouth of the Johnson Creek and the mid-point of the mainstem Johnson Creek, has been a popular place for running, cycling, walking, and other human activities. Kids and pets play with the creek water in these recreational areas adjacent to the creek. The upper part is primarily used for rural and agricultural lands. The Johnson Creek fails to meet the state health standards for contact recreation because of high levels of *E. coli* bacteria.[31] *E. coli* contamination is a major concern for both stream health and property sales price. A recent study shows that higher *E. coli* concentrations have negative influence on home sales price.[32]

The Balch Creek is a 5.6 km tributary of the Willamette River. It drains a small basin of 9.1 km^2 at the central part of the Portland Metropolitan area. Originating from the crest of the Tualatin Mountains, the stream flows east through the Macleay Park section of Forest Park, a large municipal park in Portland.

Table 1 Oregon's 303(d) list – Fanno Creek

Parameter related	Season	Beneficial uses	Status
E. coli	Fall winter spring	Water contact recreation	Cat 4A: Water quality limited, TMDL approved
E. coli	Summer	Water contact recreation	Cat 4A: Water quality limited, TMDL approved
Temperature	Summer	Salmonid fish rearing; anadromous fish passage	

Table 2 Summary of rainfall and discharge gages

	Rainfall			Discharge		
Watershed	Station name	Long.	Lat.	USGS site number	Long.	Lat.
Fanno	Sylvania PCC	−122.73	45.44	14206900	−122.73	45.49
				14206950	−122.75	45.40
Johnson	Kelly School	−122.57	45.47	14211550	−122.64	45.45
	Hayney	−122.64	45.46	14211500	−122.51	45.48
	Pleasant Valley School	−122.48	45.46	14211400	−122.42	45.49
				14211499	−122.5	45.48
Balch	Yeon	−122.71	45.55	NA		

After entering a pipe at the lower end of the park, the creek remains underground until reaching the river. Most parts of the Balch Creek watershed remain as a rural and open area as well as forestry (Fig. 1). Only a tiny part of the watershed is used for residential and commercial land uses.

For the purposes of this analysis, the Fanno Creek is the most developed watershed in the Portland Metropolitan area, which has more than 84% of urban land use and little agricultural or forestry usage (Table 3). The Johnson Creek has approximately 50% of urban area, and yet retains a certain amount of forested and agricultural land use (Table 3). The Balch Creek is the only watershed of the three with significant rural and forest coverage, and very little residential and business usage (Table 3).

Method

Data sources

We obtained data from three primary sources (Table 4). Daily precipitation data came from the City of Portland HYDRA (Hydrological Data Retrieval and Alarm) Rainfall Network operated and maintained by the City of Portland's Bureau of Environmental Services (BES). Daily stream discharge data were obtained from USGS (the United States Geological Survey). Water quality data such as TSS, stream temperature, and *E. coli* concentrations were obtained from the datasets maintained by BES and Clean Water Services, the managing agency responsible for water and sewer management in Washington County, Oregon. The summaries of monitoring stations are listed in Tables 2 and 3.

Data were extracted from July 2002 to June 2010. Precipitation data were recorded in millimeter (mm) as a 24 hour total. The three-day, five-day, and seven-day antecedent precipitation data were calculated by adding up the prior precipitation correspondingly. Data for streamflow were recorded through an automatic system at one-minute intervals, transmitted from each station at intervals of 3–6 hours, and then loaded onto the USGS computer system. Daily discharge measurements were then calculated as the average of discharge measurements over a 24 hour period and are reported in cubic meter per second.[33]

Table 3 Summary of water quality monitor stations

Watershed	Station name	Long.	Lat.	Slope (degree)	%IMP	%Dev	%Forest
Fanno	N Ash	−122.74	45.46	16.7	39.4	100.0	0.0
	S Ash	−122.74	45.45	13.5	39.0	100.0	0.0
	Main 3975	−122.72	45.49	14.6	55.0	100.0	0.0
	Main 4916	−122.73	45.49	5.2	56.1	100.0	0.0
	Main 6900	−122.75	45.49	5.8	91.0	100.0	0.0
	Pendleton	−122.74	45.49	7.2	41.5	100.0	0.0
	Vermont	−122.75	45.48	20.0	28.7	100.0	0.0
	Woods	−122.75	45.47	18.2	14.0	100.0	0.0
	Durham	−122.75	45.40	11.2	21.8	92.2	0.0
Johnson	Crystal	−122.64	45.47	1.0	61.0	100.0	0.0
	SE Regner	−122.42	45.49	11.2	9.4	33.1	20.4
	SE Umatilla	−122.64	45.46	11.6	51.8	97.2	1.8
	East of Johnson	−122.60	45.46	12.4	44.4	88.1	9.5
	SE 92	−122.57	45.47	17.0	86.0	100.0	0.0
	SE 158	−122.50	45.48	0.3	30.7	69.1	24.3
	SW Pleasant	−122.48	45.49	24.7	43.0	66.1	33.9
	SE Hogan	−122.41	45.48	13.2	14.1	100.0	0.0
	SE 159	−122.50	45.48	10.5	10.7	37.3	28.2
Balch	Thompson	−122.74	45.53	16.9	4.0	24.1	75.1
	East of Bones	−122.73	45.53	21.8	44.4	88.1	9.5
	Cornell	−122.73	45.53	12.4	14.0	63.6	36.4
	L Macleay	−122.71	45.54	9.6	36.7	100.0	0.0

Table 4 Summary of data sources

Data	Source
Precipitation	City of Portland HYDRA (Hydrological Data Retrieval and Alarm) Rainfall Network
Stream discharge	USGS (the United States Geological Survey)
Water quality	City of Portland's Bureau of Environmental Services (BES) and Clean Water Services, Oregon
Land cover, imperviousness estimate, and digital elevation model	USGS (the United States Geological Survey)
Stream network	Metro's RLIS (Regional Land Information System)

All discharge data used in the analyses have been quality-assured and approved for publication by USGS.

Water quality data were collected monthly, on site for water temperature, total suspended solids (TSS), and *E. coli* concentrations. Laboratory analyses were conducted by the Clean Water Services water-quality laboratory in Hillsboro, Oregon, using protocols described in Standard Methods for the Examination of Water and Wastewater, 2005.[34] TSS was measured according to SM 2540 D. Well-mixed samples were filtered through a weighed standard glass-fiber filter. The retained residue was dried to a constant weight at 103 to 105 °C, and TSS was calculated as the increase in weight of the filter. Temperature measurements were made with a mercury-filled Celsius thermometer (SM 2550 B). *E. coli* concentrations were measured using a 9223B Enzyme Substrate Test. Samples were mixed with enzyme substrates and incubated at 35 ± 0.5 °C. Beta-glucuronidase, an enzyme produced by *E. coli*, was detected by hydrolysis of the fluorescent substrate MUG (4-methylumbelliferyl-beta-D-glucuronide). Hydrolyzed MUG was seen as blue fluorescence when viewed under long-wavelength (366 nm) ultraviolet light, indicating a positive test for *E. coli*. At least 10% of all samples were analyzed independently in duplicate, which agreed within 5% of their average values.[34]

Land cover and imperviousness estimate layers were provided by the US Geological Survey (http://www.mrlc.gov/nlcd06_data.php). A thirty meter digital elevation model was also obtained from USGS. A stream network layer was obtained from Metro's RLIS (Regional Land Information System).

Spatial analysis

Each study catchment area was derived using ArcMap 10.1 geographic information system software. The subwatershed drained by each location was determined using the ArcHydro extension and digital elevation models produced by the US Geologic Survey at 30 m resolution (http://ned.usgs.gov/). Land parcel polygons within each delineated drainage basin were selected and summarized by the land use type. Those types utilized in this analysis are: developed land (DEV), forest, and impervious surfaces (IMP).

Statistical analysis

According to the Shapiro–Wilk test, our dataset is not normally distributed. Hence, the Spearman's rank correlation method—which makes no assumptions about the distribution of the data[35]—was applied to examine the correlations between precipitation and the three other parameters respectively: discharge, TSS, and *E. coli*. The one-day, three-day, five-day, and seven-day antecedent precipitation were all included in order to fully consider the rainfall events that last longer than one day.[7] Second, the cross-correlations between *E. coli* and discharge, TSS, and water temperature were determined to further characterize the possible association between *E. coli* and other environmental variables.

The dry–wet seasonal differences in water quality were then investigated. Again, our non-normally distributed data only allow us to perform the Mann–Whitney U test, which is used to compare differences between two independent groups without assumption about normality. The dry and wet seasons were assigned as follows: dry season = May–October; wet season = November (of the preceding year)–April, following a previous study in the study area.[36,37]

The Spearman ranking correlation between water quality and landscape variables was performed in both dry and wet seasons, across the Fanno, Johnson, and Balch Creek watersheds. Specifically, we examined the relationship between *E. coli* and TSS, with %DEV, %IMP and %Forest across all delineated subwatersheds. Correlation results are considered significant at the 0.05 level.

Regression analysis

Two types of multiple regression analysis were used. First, multiple linear regression was applied to determine the response of a correlation between TSS and *E. coli* concentrations to land cover types. The percentage of imperviousness (%IMP), percentage of Development (%DEV), and %Forest were selected as the independent variables, and seasons (dry and wet) as well as watersheds (Fanno and Johnson) were used as grouping variables. The Balch Creek watershed was excluded from the analysis due to the lack of significance data points. The full model was constructed as:

$$\text{Cor} = \text{season} + \text{watershed} + X + \text{season } X + \text{watershed} \times X + \text{season} \times \text{watershed},$$

where Cor stands for the correlation between TSS and *E. coli*, and X stands for %IMP or %DEV or %Forest. Variables selection was performed using stepwise AIC in software R. We would expect decrease in correlation strength as %IMP and %DEV go up, and %Forest goes down. The R^2 value was reported that represents the proportion of variation in the response variable explained by the fitted regression line.[38]

Second, we employed multiple linear regression to model the determinants of the concentrations of *E. coli* based on previous correlation analyses. Precipitation, discharge, TSS, and temperature are the likely candidates for explaining the concentration of *E. coli* as these variables could strongly influence the transport, build-up, and survival of *E. coli*. Log transformation, which can provide better homoscedasticity and result in a more symmetric dataset with normal residuals,[39] was performed on *E. coli* concentrations. This approach has been used successfully for *E. coli* concentrations and other selected variables in streams in Oregon.[13] After natural log transformation, our *E. coli* data are able to fit into a normal distribution, proving the applicability of regression analysis. The goodness-of-fit of predictive models was assessed with diagnostics statistics, including residual plots and the R^2. ANOVA tables were also employed to identify the significance of our models at the 95 percent confidence interval ($p = 0.05$).

Results and discussion

Correlation between discharge and antecedent precipitation amount

A strongest significant positive correlation was observed between stream discharge and the three-day antecedent precipitation in the Fanno Creek watershed for both dry and wet seasons ($\rho = 0.547$, 0.759 respectively, Table 5). This relatively fast response of streamflow to precipitation is not surprising in highly urbanized watersheds with a steep slope like the Fanno Creek watershed (Table 3), where higher impervious surfaces associated with high-density development are likely to promote hydraulic efficiency, resulting in a faster response of streamflow to precipitation events.[40]

For discharge–precipitation correlation, the Johnson Creek was more highly correlated with the seven-day antecedent precipitation than with the 3 day or 5 day antecedent precipitation for both seasons (Table 5). This slow response could be attributed to the large size, flat topography, and elongated shape of the Johnson Creek watershed. Moreover, unlike the Fanno Creek that is mostly covered by urban land use (Table 3), land cover upstream of monitoring stations in the Johnson Creek watershed is dominated by agricultural and rural lands where more time is needed for soil to be saturated, and then runoff can start. A similar result was found in small Pennsylvania watersheds.[41] This may also suggests a significant base flow component to Johnson Creek discharge from groundwater or septic systems in rural areas or old residential areas.

Correlation between TSS and antecedent precipitation

TSS was more strongly correlated with the one-day and three-day than the seven-day antecedent precipitation for dry and wet seasons in the Fanno Creek (Table 5). This fast response relative to discharge may indicate the existence of a first flush effect, which was also observed in urban stormwater in Raleigh, North Carolina, U.S.[3] In other words, most sediments were delivered to the stream at the initial stage (in terms of runoff volume) of the rainfall events, meaning that the sources were either near streams or from top soils that had been accumulated between rainfall events. The assumption seems to be reasonable since all monitoring stations are located in either high- or medium-developed areas in the Fanno Creek watershed, and most are close to parks and other green spaces where people typically accompany dogs or other pets.

The Johnson Creek showed a slow response in TSS with the strongest correlation with the five-day antecedent precipitation in the wet season, and no significant correlation in the dry season (Table 5). The results would accord with the previous discussion regarding streamflow; since most precipitation made its way down through the soil into groundwater, and the relatively small amount of surface runoff was not be able to carry sediments into the stream systems, resulting in little TSS response to rainfall events. This could also be related to the dominant contribution of baseflow during the dry period, which might attenuate or overshadow the pollutant trends in stormwater. Moreover, some sections of the Johnson Creek are armored from the WPA (Works Progress Administration) era. The consolidated and armored river bank tends to have less bank erosion and consequently fewer storm-associated sediments.

No significant rainfall correlation with TSS was observed in the Balch Creek during both seasons. Since the Balch Creek is the least urbanized stream system, possible explanations could be that sediment sources are scarce or soils can effectively absorb sediments. Also, similar to the Johnson Creek, the

Table 5 Spearman's correlation coefficient in the Fanno Creek, Johnson Creek and Balch Creek watersheds ($n_{\text{Fanno,dry}} = 66$, $n_{\text{Fanno,wet}} = 68$; $n_{\text{Johnson,dry}} = 52$, $n_{\text{Johnson,wet}} = 46$; $n_{\text{Balch,dry}} = 48$, $n_{\text{Balch,wet}} = 35$)[a]

	Discharge				TSS						*E. coli*					
	Fanno		Johnson		Fanno		Johnson		Balch		Fanno		Johnson		Balch	
	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet
Precip-1	0.525*	0.720*	0.480*	ns	ns	0.490*	ns	ns	ns	ns	**0.498***	**0.539***	ns	**0.658***	ns	ns
Precip-3	**0.547***	**0.759***	0.585*	0.585*	ns	**0.497***	ns	0.753*	ns	ns	0.393*	0.426*	ns	0.615*	ns	0.574*
Precip-5	0.527*	0.709*	0.560*	0.700*	ns	0.419*	ns	**0.796***	ns	ns	0.391*	0.350*	ns	0.399*	ns	**0.658***
Precip-7	0.547*	0.681*	**0.596***	**0.723***	ns	0.416*	ns	0.692*	ns	ns	0.322*	ns	ns	ns	ns	0.551*

[a] ns: non-significant; * significant at the 0.05 level; the highest coefficient values are in bold type.

higher surface permeability might generate too scarce runoff to wash off pollutants into streams.

Correlation between *E. coli* and antecedent precipitation

In all three watersheds, significant correlation was observed (Table 5) between *E. coli* and preceding rainfall events, at least during the wet season, which accords with previous studies.[5–7] This is probably because the antecedent rainfall conditions affect both the amount of water and energy available for *E. coli* transport and the amount of moisture present in a watershed that is critical for *E. coli* survival.[6] Similar to TSS, *E. coli* in the Fanno Creek was most strongly correlated with the one-day antecedent precipitation, which might also indicate a first flush effect on *E. coli*. In the Johnson Creek, however, *E. coli* showed a different response, compared to TSS response to precipitation, with the strongest correlation with the one-day antecedent precipitation during the wet season ($r = 0.658$).

Correlation between *E. coli* and TSS

As shown in Fig. 2, *E. coli* is generally positively associated with TSS, with higher *E. coli* and lower TSS concentrations in the dry season than that in the wet season. The correlations are all statistically significant in three watersheds in the dry season, while the correlation for the Balch Creek is not statistically significant in the wet season (Table 6). While the Balch Creek has the highest correlation in the dry season, the Johnson Creek has the highest correlation in the wet season.

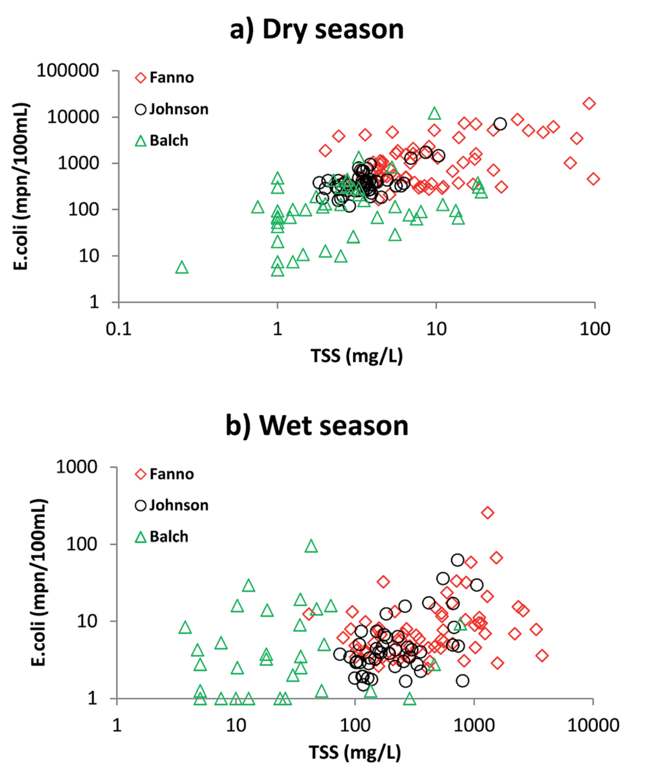

Fig. 2 Relationship between *E. coli* and TSS in three watersheds for (a) the dry season and (b) the wet season; all correlation results are significant at the 0.05 level.

Our results are largely consistent with previous findings from Anderson and Rounds (2003)[13] and Hamilton & Luffman (2009)[7] for highly urbanized watersheds in Oregon and eastern Tennessee, U.S., respectively; both determined that *E. coli* bacteria was positively correlated with turbidity, indicating that *E. coli* were either transported to stream bound to particulate matter, adsorbed onto resuspended stream bed particles, or they had an affinity for sediments in water.[18] It is reasonable to suggest that the majority of *E. coli* sources lie close to the Johnson Creek, especially considering spring corridor trails parallel to the creek in the middle and lower section of the creek and the agricultural activities in riparian areas of the upper section of the creek that apply a large amount of manures. It should also be noticed that *E. coli* is an indicator of recent pollution which could attenuate the correlations between *E. coli* concentrations and the seven-day antecedent precipitation. Samples and analyses for indicators of old pollution (*e.g.*, *Enterococcus faecalis*) are needed for further assessment.

However, one cannot simply infer that *E. coli* is transported with suspended sediments, especially given the distinct responses of TSS and *E. coli* to storm events in the Johnson Creek. The bifurcation in the *E. coli versus* TSS plot of the Balch Creek watershed in the wet season (Fig. 2) could suggest two different TSS sources with different amounts of *E. coli* associated with them. One possible source of sediment is from near streams such as the stream bank or stream bed, and the other potential source is from distant areas such as upstream areas or water delivered from storm pipes.

Seasonal differences in water quality

According to the Mann–Whitney *U* test, discharge, temperature, and *E. coli* showed significant dry–wet seasonal differences across Fanno, Johnson, and Balch Creek watersheds ($p = 0.01$, 2-tailed). Not surprisingly, higher discharge and lower temperature were observed during the wet season. *E. coli* concentrations are significantly higher in the dry season than in the wet season (Fig. 3). This is a likely consequence of higher flows and more frequent washout of stored bacteria in the wet season.[42] The increase in animal and human activities during the warm season could also lead to higher *E. coli* concentrations in streams. Most occurrences of elevated *E. coli* levels in urban watersheds originate from sources such as domestic pet waste.[14] People typically walk their pets more often during the warmer season, and therefore increasing the probability of feces contamination.

Table 6 Spearman's rank correlation coefficient between *E. coli* and TSS[a]

	Dry	Wet	All season
Fanno	0.246* ($n = 66$)	0.321* ($n = 68$)	**0.323*** ($n = 134$)
Johnson	0.359* ($n = 52$)	**0.423*** ($n = 46$)	0.299* ($n = 98$)
Balch	**0.380*** ($n = 48$)	0.222 ($n = 35$)	0.223* ($n = 83$)

[a] *: significant at the 0.05 level; the highest coefficient values are in bold type.

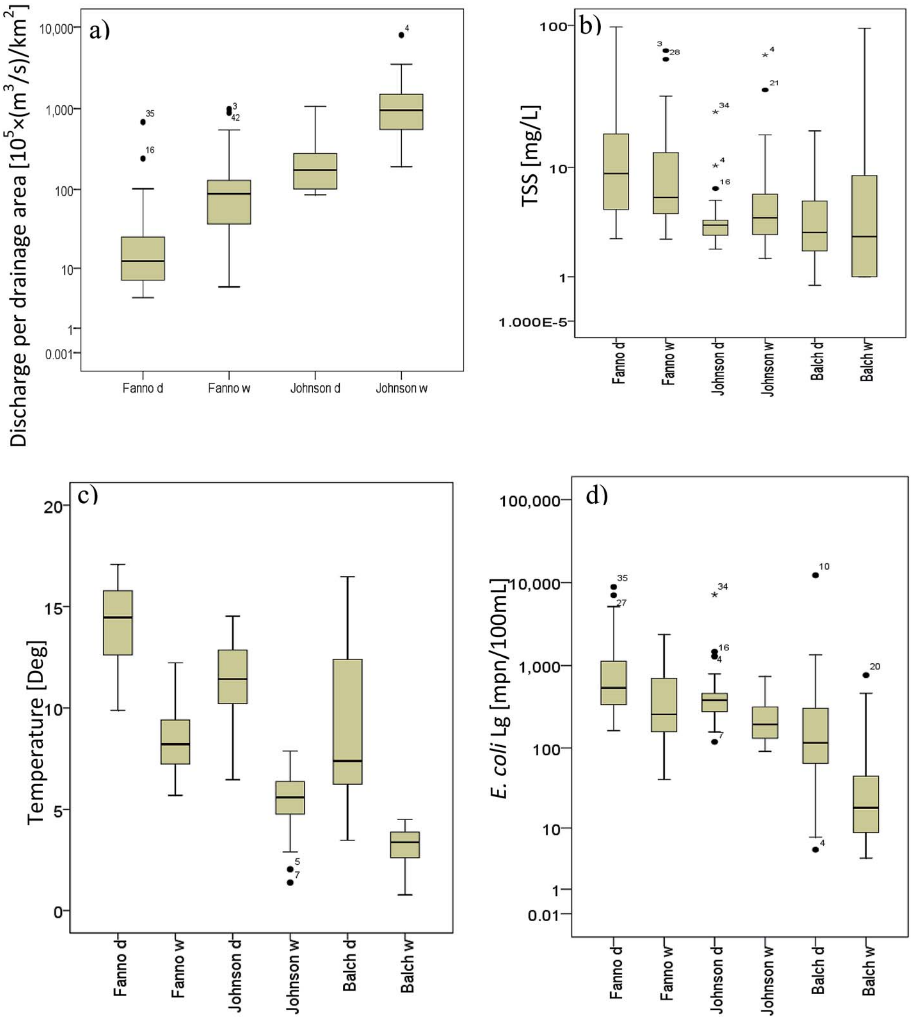

Fig. 3 Dry–wet seasonal differences ("d" for dry, "w" for wet) of (a) discharge $\times 10^5$, (b) TSS, (c) temperature, and (d) *E. coli* in the Fanno, Johnson, and Balch Creek.

In addition, higher *E. coli* concentrations may be attributed to the warmer temperature in the summer season, which corresponds to higher growth and survival rates of *E. coli* bacteria.[9,13] As *E. coli* bacteria are thermotolerant (tolerant of relatively high temperatures), the lower river temperatures during the cool season inhibit the growth and survival of *E. coli* bacteria in the river. The reproduction of *E. coli* outside the intestines of warm-blooded animals seems unlikely. However, there is a growing body of evidence suggesting that there exists a specialized subset of *E. coli* strains that can reproduce in secondary environments in both tropical[43,44] and temperate climates.[45–49] Therefore, the use of *E. coli* as an indicator of fecal pollution should be reevaluated.[50]

We were unable to identify a significant seasonal pattern for TSS, indicating that either no seasonal difference exists for the study period, or there are insufficient data points to identify a significant difference. It is possible that although the wet season generates more runoff to carry sediment into streams, it also has more dilution effect which might cancel out the additional input.

Landscape impacts

According to the Mann–Whitney *U* test, all variables surveyed (discharge, temperature, TSS, and *E. coli*) demonstrated significant differences across the Fanno, Johnson, and Balch Creek watersheds at the 0.01 level (2-tailed), except for TSS difference between the Johnson and Balch Creek watershed. The Johnson Creek has higher flow per drainage area than the Fanno Creek despite having a lower degree of imperviousness. This is probably because storm pipes reroute water further downstream to the mouth of the Fanno Creek where a wastewater discharge plant is located. The Fanno Creek has the highest levels of TSS, *E. coli*, and temperature, followed by the Johnson and Balch Creek in order (Fig. 3). The results are expected since most pollutants could be attributed largely to anthropogenic sources.

A few significant correlations were observed between water quality and landscape variables (Fig. 4). Correlations between %IMP and *E. coli* in the wet season, as well as %DEV with TSS indicate that high impervious surface coverage and degree of urban development carry increased concentrations of TSS and *E. coli* bacteria to surface waters. The results generally agree with a growing body of the scientific literature that predicts water quality degradation resulting from urbanization.[5] However, most of the correlations are not significant, suggesting that more study is needed to identify the underlying controls of transport and build-up of *E. coli* in water systems under various landscape regimes.

We went further to analyze the response of the correlation between TSS and *E. coli* to landscape characteristics. Consistent with our hypothesis, a correlation between TSS and *E. coli* significantly decreased as %DEV went up, albeit to a small extent (Table 7). The best model was listed as follows: Cor = season + %DEV. The whole model is significant at the 0.05 level with a R^2 of 39%, and both independent variables, season and %DEV, are significant at the 0.05 level (Table 7). Large variation was observed in regions with high percentage of development, shown as the deviation at the upper right on the normal Q–Q plot (Fig. 5). Most of those regions lie within the Fanno Creek watershed, implying the more complex stormwater infrastructure and more variable pollutant sources of *E. coli* in the urban watershed.[4] No significant impact, however, was identified in %IMP, suggesting complex transport and build-up processes of *E. coli* and TSS, which could hardly be represented as a linear function of degree of imperviousness. Those processes include changes in the stream route caused by urban storm drains in the Fanno Creek. %Forest also did not show any significant impact on *E. coli* and TSS correlation. Apart from the trend distortion caused by oversimplification, a possible explanation could be that our sample size is too small to capture the trend since we only have eight data points.

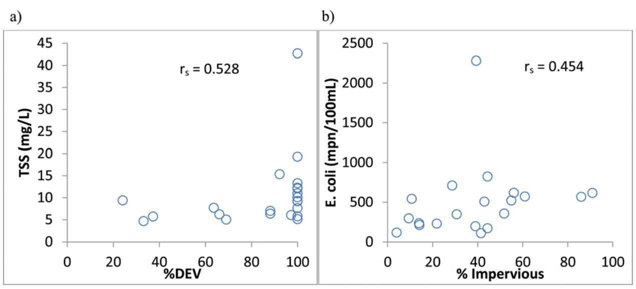

Fig. 4 Relationship between *E. coli*, TSS and landscape and the corresponding Spearman's rank correlation coefficients (two-tailed) for (a) %DEV *vs.* TSS in the dry season, and (b) %IMP *vs. E. coli* in the wet season; both correlation results are significant at the 0.05 level.

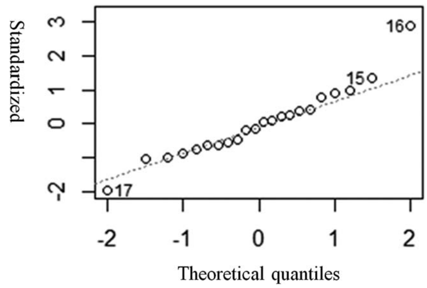

Fig. 5 Probability plot (Q–Q plot) of % developed land.

Multiple regression analysis

We ran several multiple linear regression models to predict the *E. coli* concentration based on the stepwise method and the knowledge gained from our correlation analyses. The Balch Creek was excluded from the analysis since no significant predictors were found within our targeted variables (antecedent precipitation, discharge, TSS, and water temperature). Table 8 summarizes the major information regarding our models. Models for the Johnson Creek watershed (eqn (2) and (4)) have much higher adjusted R^2 than those for the Fanno Creek watershed, which are capable of explaining 60% and 50% of the variability in the *E. coli* concentration for the dry and wet season, respectively. Again, the lower predictability of the Fanno Creek models are likely the result of a more complex stormwater infrastructure and more variable pollutant sources of *E. coli* in the urban watershed.

Best models for the Fanno, and Johnson Creek watersheds are listed as follows:

$$\text{Fanno dry: } \ln(E.\ coli) = 0.522 \times \mathrm{P}_1 \tag{1}$$

$$\text{Johnson dry: } \ln(E.\ coli) = 0.735 \times \mathrm{TSS} + 0.212 \times T \tag{2}$$

$$\text{Fanno wet: } \ln(E.\ coli) = 0.356 \times \mathrm{P}_3 + 0.309 \times \mathrm{P}_1 \tag{3}$$

$$\text{Johnson wet: } \ln(E.\ coli) = 0.448 \times \mathrm{P}_1 + 0.528 \times \mathrm{TSS} - 0.305 \times \mathrm{P}_7 \tag{4}$$

where P1, P3, P7, T, and TSS stand for the one-day, three-day, and seven-day antecedent precipitation, temperature, and total suspended solids after standardization. Variables are ordered by coefficient significance, and all of these selected variables are significant at the 0.05 level. The models for the Fanno Creek watershed show positive relationships between the *E. coli* concentration and the one-day, three-day antecedent

Table 7 Landscape model summary (n = 22)[a]

	Estimate	Std. error	t value	P-value		Durbin–Watson
(Intercept)	0.60749	0.08968	6.774	1.80×10^{-6}	*	2.16
Season	−0.0517	0.01981	−2.609	0.0173	*	
%DEV	−0.0023	0.00099	−2.298	0.0331	*	

[a] *: significant at the 0.05 level.

Table 8 *E. coli* model summary

Season	Watershed	Independent variable	Standardized coefficients Beta	T	Sig.	Collinearity statistics Tolerance	VIF	R squared	Durbin–Watson
Dry	Fanno	(Constant)		54.41	0.00			0.27	1.15
		Precip-1	0.52	4.89	0.00	1.00	1.00		
	Johnson	(Constant)		15.63	0.00			0.61	1.85
		TSS	0.74	8.22	0.00	0.99	1.01		
		T	0.21	2.37	0.02	0.99	1.01		
Wet	Fanno	(Constant)		43.31	0.00			0.35	1.69
		Precip-3	0.36	2.93	0.01	0.68	1.47		
		Precip-1	0.31	2.55	0.01	0.68	1.47		
	Johnson	(Constant)		4.82	0.00			0.49	1.65
		Precip-1	0.45	3.48	0.00	0.73	1.37		
		TSS	0.53	3.30	0.00	0.47	2.13		
		Precip-7	−0.31	−2.05	0.05	0.54	1.84		

precipitation, which indicates the dominance of *E. coli* wash-off at the initial stage of rainfall. The seven-day antecedent discharge, on the other hand, is negatively associated with the concentration of *E. coli* in the Johnson Creek watershed. It may be inferred that after several days of rain, most of the accumulated *E. coli* have been flushed through the watershed.[7] Therefore, the previous concentration response is replaced by dilution effects. TSS plays an important role in predicting the *E. coli* concentration in the Johnson Creek watershed for both seasons (Fig. 2). The antecedent precipitation appears to be the most important predictor for both dry- and wet-season models in the Fanno Creek watershed and the wet-season model in the Johnson Creek watershed. This outcome is consistent with previous multiple linear regression for *E. coli* prediction that also yielded explanatory variables related to antecedent precipitation.[2,6,7] Moreover, the results emphasize the importance of antecedent weather, and therefore *E. coli* build-up, persistence and die-off processes, in microbial modeling.[4]

Usefulness of regression models

Our regression models, though simple, can be used in several ways. First, these models can be used to regulate the monitoring of Fanno and Johnson Creeks for the concentration of *E. coli* because they may provide a preliminary estimate of the concentration of *E. coli* using readily available data and a common, easily interpretable statistical method familiar to many. Since both creeks pass through several riparian parks and serve a recreational function, the timely *E. coli* data may be compared with the water quality criteria for body contact to determine whether or not the stream will pose a threat to human health. Second, the knowledge of the possible changes in water quality ahead of time is useful to ensure adequate sampling, and proactive watershed management practices are performed to maintain the standards required for surface water, and thereby facilitating proactive watershed management practices to prevent negative effects on human and aquatic life. With a monitoring station and a streamflow gaging station located together, constituent loads can be calculated as well which are useful for calculating total maximum daily loads (TMDLs), a mandatory criterion established for the concentration of *E. coli* in the Fanno Creek (Table 1). Moreover, the model could also be used for the evaluation of possible land-use management changes in the target watershed. If model coefficients of certain variables exhibit constantly increasing or declining trends over a period of time, then such trends can suggest a new or reduced source, or process-based changes in the basin.[18] Further investigation and model calibration are in need for a better understanding of the complex variations in the concentration of *E. coli* and its interactions with other weather and hydrologic variables before these models could be fully applied for practical use.

Conclusions

Correlations of precipitation, seasonal differences, landscape impacts, and regression analyses of *E. coli* concentrations were performed in this study across an urban, mixed, and forested watershed. From this, the following conclusions can be drawn:

(1) We investigated the relationships between precipitation and the three other parameters: stream discharge, TSS, and *E. coli* concentrations using the Spearman's ranking correlation coefficient. The Fanno Creek watershed exhibited a fast response of streamflow to precipitation, which could be attributed to its steep slope and high degree of imperviousness. The discharge of the Johnson Creek showed a slow or weak response to storm events, suggesting that there is a strong base flow component and that the basin size, shape and topography prolong the time of concentration. Moreover, there were likely first flush effects for TSS and *E. coli* in the Fanno Creek watershed, indicating nearby pollution sources. The weak response of TSS in the Johnson Creek could be attributed to high permeability, baseflow dilution, and river bank consolidation. The Balch Creek in general showed weak responses to storm events, which were likely to be the result of scarce sources.

(2) The Mann–Whitney U test of seasonal trends identified significant variations of discharge, temperature, and *E. coli* concentrations between dry and wet seasons. The higher *E. coli*

concentrations in the dry season could be attributed to the warmer temperature that provides better persistence or growing environment for *E. coli*, and therefore creating the opportunity for higher concentrations during subsequent runoff events. There were no significant seasonal differences in TSS. It is possible that although the wet season generates more runoff to carry sediment into streams, it also has more dilution effect which might cancel out the additional input.

(3) The urban watershed has the highest levels of TSS, *E. coli*, and temperature, followed by mixed and forested watersheds in order. The results are expected because most pollutants could be attributed largely to anthropogenic sources. In general, TSS was significantly correlated with *E. coli* concentrations, particularly during the dry season. Such correlations linearly decreased as %DEV went up, with a large variation in regions with high percentage of development, implying the more complex stormwater infrastructure and more variable pollutant sources of *E. coli* in urban watersheds. %IMP and %Forest did not show any significant impact on *E. coli* and TSS correlations, which could be attributed to limited sample sizes.

(4) Multiple linear regression models were developed using antecedent precipitation TSS, and temperature to predict the concentration of *E. coli*. Models for the Johnson Creek watershed have much higher adjusted R^2 than those for the Fanno Creek watershed, which are capable of explaining 60% and 50% of the variability in the *E. coli* concentration for the dry and wet season, respectively. These models can provide a preliminary estimate of the concentration and loads of *E. coli*, and therefore are able to facilitate the establishment of water quality criteria, enable proactive watershed management practices to prevent negative effects on human and aquatic life, and identify long-term water quality changes in the target watershed. The complex variations in the concentration of *E. coli* and its interactions with other climatic and hydrologic variables still need to be further investigated before the development of regression models with a higher predictive accuracy and confidence level.

Acknowledgements

This research was supported by the US National Science Foundation (grant # 0948983). Additional support was provided by Portland State University and Tongji University (China). We appreciate City of Portland's Bureau of Environmental Services for providing stream water quality data and Portland Metro for sharing GIS data. We appreciate Zbigniew Grabowski for proofreading the manuscript. Views expressed are our own and are not those of the sponsoring agencies.

Notes and references

1 United States Environmental Protection Agency, National Summary of Impaired Waters and TMDL Information, http://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T#causes_303dfiles/250/National Summary of Impaired Waters and TMDL Infor.html.

2 D. T. McCarthy, V. G. Mitchell, A. Deletic and C. Diaper, *Water Sci. Technol.*, 2007, **56**, 27–34.

3 J. M. Hathaway and W. F. Hunt, *Water, Air, Soil Pollut.*, 2011, **217**, 135–147.

4 D. T. McCarthy, J. M. Hathaway, W. F. Hunt and A. Deletic, *Water Res.*, 2012, **46**, 6661–6670.

5 M. A. Mallin, V. L. Johnson and S. H. Ensign, *Environ. Monit. Assess.*, 2009, **159**, 475–491.

6 J. M. Hathaway, W. Hunt and O. Simmons, *J. Environ. Eng.*, 2010, **136**, 1360–1368.

7 J. L. Hamilton and I. Luffman, *Phys. Geogr.*, 2009, **30**, 236–248.

8 R. W. Gentry, M. John, A. Layton, L. D. McKay, D. Williams, S. R. Koirala and G. S. Sayler, *J. Environ. Qual.*, 2006, **35**, 2244–2249.

9 D. Chin, *J. Environ. Eng.*, 2010, **136**, 249–253.

10 K. H. Cho, Y. A. Pachepsky, J. H. Kim, A. K. Guber, D. R. Shelton and R. Rowland, *J. Hydrol.*, 2010, **391**, 322–332.

11 S. J. Ki, S. Ensari and J. H. Kim, *Environ. Manage.*, 2007, **39**, 867–875.

12 J. L. Mancini, *J. - Water Pollut. Control Fed.*, 1978, **50**, 2477–2484.

13 C. W. Anderson and S. A. Rounds, *U.S. Geological Survey Water-Resources Investigations Report 02-4232*, 2003.

14 K. A. McCarthy, *U.S. Geological Survey Water-Resources Investigations Report 00-4062*, 2000.

15 T. Garcia-Armisen and P. Servais, *Water Environ. Res.*, 2009, **81**, 21–28.

16 J. Smith, J. Edwards, H. Hilger and T. R. Steck, *J. Gen. Appl. Microbiol.*, 2008, **54**, 173–179.

17 R. W. Muirhead, R. J. Davies-Colley, A. M. Donnison and J. W. Nagels, *Water Res.*, 2004, **38**, 1215–1224.

18 C. W. Anderson and S. A. Rounds, *U.S. Geological Survey Scientific Investigations Report 2010-5008*, 2010.

19 A. McDonald, D. Kay and A. Jenkins, *Appl. Environ. Microbiol.*, 1982, **44**, 292–300.

20 S. M. Goyal, C. P. Gerba and J. L. Melnick, *Appl. Environ. Microbiol.*, 1977, **34**, 139–149.

21 A. M. Desai and H. S. Rifai, *J. Environ. Eng.*, 2010, **136**, 1347–1359.

22 Teresa J. Rasmussen, C. J. Lee and C. Z. Andrew, *U.S. Geological Survey Scientific Investigations Report 2008–5014*, 2008.

23 H. Chang and I. W. Jung, *J. Hydrol.*, 2010, **388**, 186–207.

24 I. W. Jung and H. Chang, *Theor. Appl. Climatol.*, 2012, **108**, 355–371.

25 H. Chang, M. Lafrenz, I.-W. Jung, M. Figliozzi, D. Platman and C. Pederson, *Ann. Assoc. Am. Geogr.*, 2010, **100**, 938–952.

26 H. Chang, P. Thiers, N. R. Netusil, J. A. Yeakley, G. Rollwagen-Bollens, S. M. Bollens and S. Singh, *Hydrol. Earth Syst. Sci.*, 2014, **18**, 1383–1395.

27 R. Hoyer and H. Chang, *Land*, 2014, **3**, 322–341.

28 I. W. Jung, H. Chang and H. Moradkhani, *Hydrol. Earth Syst. Sci.*, 2011, **15**, 617–633.

29 Environmental Service, the City of Portland, Water Quality | About the Watershed|, The City of Portland, Oregon, http://

www.portlandoregon.gov/bes/article/315387files/212/315387.html.

30 S. Singh and H. Chang, *Int. J. Geospat. Environ. Res.*, 2014, **1**, 8.

31 Johnson Creek Watershed Council, 2012 State of the Watershed Report, http://jcwc.org/2012-state-of-the-watershed-report/files/342/2012-state-of-the-watershed-report.html.

32 N. R. Netusil, M. Kincaid and H. Chang, *Water Resour. Res.*, 2014, **50**, 4254–4268.

33 U.S. Geological Survey, Tualatin River Basin Monitoring Sites, http://or.water.usgs.gov/tualatin/monitors/files/200/monitors.html.

34 American Public Health Association (APHA), *Standard Methods for the Examination of Water and Waste Water*, American Public Health Association, Washington, D.C., 21st edn, 2005.

35 J. F. Pallant, *SPSS Survival Manual: a Step by Step Guide to Data Analysis Using SPSS for Windows*, Allen & Unwin, Crows Nest, N.S.W., 2013.

36 M. Boeder and H. Chang, *J. Environ. Manage.*, 2008, **87**, 567–581.

37 B. Pratt and H. Chang, *J. Hazard. Mater.*, 2012, **209–210**, 209–210.

38 N. J. Gotelli and A. M. Ellison, *A Primer of Ecological Statistics*, Sinauer Associates Publishers, Sunderland, Mass, 2004.

39 J. R. Gray, G. D. Glysson, L. M. Turcios and G. E. Schwarz, *Water-resources Investigations Report 00-4191*, 2000.

40 H. Chang, *Hydrol. Processes*, 2007, **21**, 211–222.

41 H. Chang and T. Carlson, *Hydrobiologia*, 2005, **544**, 321–332.

42 P. Rodgers, C. Soulsby, C. Hunter and J. Petry, *Sci. Total Environ.*, 2003, **314–316**, 289–302.

43 M. Bermúdez and T. C. Hazen, *Appl. Environ. Microbiol.*, 1988, **54**, 979–983.

44 H. M. Solo-Gabriele, M. A. Wolfert, T. R. Desmarais and C. J. Palmer, *Appl. Environ. Microbiol.*, 2000, **66**, 230–237.

45 M. N. Byappanahalli and R. S. Fujioka, *Water Sci. Technol.*, 1998, **38**, 171–174.

46 D. M. Gordon, S. Bauer and J. R. Johnson, *Microbiology*, 2002, **148**, 1513–1522.

47 S. Ishii, W. B. Ksoll, R. E. Hicks and M. J. Sadowsky, *Appl. Environ. Microbiol.*, 2006, **72**, 612–621.

48 P. K. Jjemba, L. A. Weinrich, W. Cheng, E. Giraldo and M. W. LeChevallier, *Appl. Environ. Microbiol.*, 2010, **76**, 4169–4178.

49 S. L. McLellan, *Appl. Environ. Microbiol.*, 2004, **70**, 4658–4665.

50 E. M. Anastasi, B. Matthews, H. M. Stratton and M. Katouli, *Appl. Environ. Microbiol.*, 2012, **78**, 5536–5541.

