Water Research 194 (2021) 116921

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Investigation of relationships between fecal contamination, cattle grazing, human recreation, and microbial source tracking markers in a mixed-land-use rangeland watershed

Naveen Joseph[a], Jane Lucas[b], Nikhil Viswanath[a], Reed Findlay[c], Jim Sprinkle[d], Michael S. Strickland[b], Eric Winford[e], Alan S. Kolok[a,*]

[a] Idaho Water Resources Research Institute, University of Idaho, Moscow, ID, USA
[b] Department of Soil and Water Systems, University of Idaho, Moscow, ID, USA
[c] University of Idaho Extension – Eastern District, University of Idaho, Pocatello, ID, USA
[d] Nancy M. Cummings Research, Extension and Education Center, University of Idaho, Carmen, ID, USA
[e] Rangeland Center, University of Idaho, Boise, ID, USA

ARTICLE INFO

Article history:
Received 8 October 2020
Revised 8 February 2021
Accepted 9 February 2021
Available online 11 February 2021

Keywords:
Escherichia coli
Microbial source tracking markers
Quantitative polymerase chain reaction
Cattle DNA
Human DNA

ABSTRACT

The United States National Forests are mixed-use lands that support human recreation and cattle grazing. Overuse by humans or cattle, however, can lead to the fecal contamination of local waterways. Until recently, the source of these contaminants was a subject of conjecture; however, microbial source tracking tools have become widely used and are proving to be a valid methodology to identify the contamination source. This study aims to analyze and model the quantity and sources of fecal contamination in the Mink Creek watershed in southeastern Idaho. The U.S. Forest Service Caribou-Targhee National Forest (USFS) manages this watershed. Previous research has indicated that some localities within the watershed exceed US EPA standards for coliform bacteria. In 2019, water samples were collected before livestock began grazing and throughout the spring, summer, and fall after livestock grazing had ended. Fourteen sites were sampled seven times during the field season, allowing the water to be analyzed for total coliforms and E. coli bacteria. Microbial source tracking techniques using Bacteroides bacteria, which are known to live in specific digestive tracks, were used to identify the source of E. coli at each sampling location. The analysis indicated that E. coli counts exceeded state regulatory limits 35% of the time. These exceedances were associated with DNA source tracking markers for humans (58.8%), cattle (5.9%), or both cattle and humans (5.9%). Unknown sources were responsible for the Bacteroides bacteria 29.4% of the time. A statistical model was developed to estimate E. coli using the datasets of microbial source tracking measures, the presence or absence of humans, cattle, the proximity of the sampling date to a holiday, and other seasonal factors. The resulting model showed good performance indices at all the 14 sites based on a K-fold cross-validation scheme (R^2 = 0.83 and NSE = 0.69). The results demonstrated that E. coli exceedances have a close association with human recreation and unknown sources and negatively influenced by dissolved oxygen.

1. Introduction

United States Forest Service (USFS) supports human recreation and livestock grazing on National Forest System lands (Williams, 2000). Approximately 1.8 million livestock graze across the national forests of the western United States alone (Roche et al., 2013). This substantial grazing activity is part of the USFS's multiple-use mandate and frequently co-occurs with recreational use and wildlife activity (USFS, 2011). Recent studies identified that recreation in the U.S. has increased by 7% from 2000 to 2009, and the number of days public lands utilized for recreation has increased by 30% (Wolf et al., 2017). However, fecal contamination of water resources often occurs when there is heavy use by cattle, humans, and wildlife in rangeland watersheds (Roche et al., 2013). Fecal contamination, especially by E. coli (E. coli), can occur when fecal matter is deposited on or near streams and waterways and can lead to deterioration of water quality and water supply impairment (Petersen and Hubbart, 2020).

* Corresponding author.
E-mail address: akolok@uidaho.edu (A.S. Kolok).

https://doi.org/10.1016/j.watres.2021.116921
0043-1354/

Several studies have focused on the impact of human recreation and livestock grazing on chemical and bacteriological water quality in rangelands [e.g., Roche et al. (2013)]. Most of these assessments used fecal indicator bacteria (FIB) to monitor fecal contamination in drinking water (WHO, 2011). World Health Organization (WHO) and Joint Monitoring Program for Water Supply and Sanitation (JMP) lists E. coli as the commonly used FIB for monitoring fecal contamination (WHO/UNICEF, 2015). However, it has been historically difficult, if not impossible, to identify the primary source(s) of fecal pollution in watersheds with multiple potential inputs, as is often the case in rangeland watersheds (Leclerc et al., 2001). Microbial Source Tracking (MST) has been developed to differentiate various animal sources of fecal pollution and has been used to determine sources of fecal pollution in streams (Harwood et al., 2014). Several studies have employed MST to track the target members of contamination from humans [e.g., Reischer et al. (2007)], ruminant [e.g., Mieszkin et al. (2010)], and other sources [e.g., Green et al. (2012)], from the collected water samples. Library-independent source tracking methods using quantitative polymerase chain reaction (qPCR) has increased in popularity since they are cost-effective, rapid, and do not require organisms' culturing (Scott et al., 2002). Supporting this, several model validation studies have identified that qPCR is a better tool in identifying markers associated with various sources [e.g., Layton et al. (2013)].

Various studies have employed qPCR methods to track human marker genes and identified that human recreation and its associated fecal contamination pose significant public health issues (Harwood et al., 2014; Ponce-Terashima et al., 2014; Devane et al., 2019; Senkbeil et al., 2019). Recent studies such as Devane et al. (2019) identified that human-associated qPCR markers hold a significant positive correlation with E. coli in water samples. Similar to human recreation, cattle grazing cycles have a critical role in microbial contamination, measured by E. coli concentrations (Allocca et al., 2018). For example, Derlet and Carlson (2006) analyzed water samples collected downstream of grazed and non-grazed sites in Sierra Nevada, California, and found that detectable E. coli concentrations are highly probable in grazed, relative to non-grazed areas.

Besides human and cattle sources, seasonal changes and other abiotic factors influence the microbial indicators in surface water (Valeo et al., 2016). E. coli counts are found to hold a strong relationship with meteorological conditions, such as solar radiation, temperature (Cho et al., 2010), and precipitation (Buckerfield et al., 2019). Similarly, E. coli growth and metabolism are strongly related to dissolved oxygen in streams (Baez and Shiloach 2014). An increase in organic matter increases the microbes and their metabolism in the streams, thereby lowering dissolved oxygen (Mulholland et al., 2005). On the other hand, high dissolved oxygen in streams corresponds with reduced E. coli growth and results in a higher removal rate of coliform bacteria (Baez and Shiloach 2017). Moreover, dissolved oxygen in streams drops during hot summer days, coinciding with low stream flows, creating an ideal environment for E. coli to grow (Khaengraeng and Reed 2005).

Public land managers, including those with the USFS, seek ways to reduce the potential impairment of waterways and watersheds in their jurisdiction. In southeastern Idaho, the Mink Creek watershed had a history of E. coli contamination. In 2017, the Idaho Department of Environmental Quality (IDEQ) monitored seven locations in the watershed bi-monthly from June-October to determine E. coli counts (Harris 2017). They found out that E. coli counts exceeded the regulatory limit of 126 MPN/100 mL (Most Probable Number of E. coli organisms per 100 mL) at all locations except one (IDAPA 2014). Although the high E. coli counts coincided with cattle movement, there were spikes in coliform counts at several sites, weeks to months after the cattle had left the adjacent grazing units. Given that wildlife and humans periodically inhabit areas within the watershed, the dominant source of coliforms remained unresolved. This study's objective was to determine how the timing of cattle grazing and human recreation influenced the source and quantity of E. coli in the Mink Creek watershed.

Table 1
Temperature and precipitation anomaly comparison.

Anomaly	Month	2017	2019
Temperature (°C)	June	−0.6	0.9
	July	0.3	2
	August	0.1	0.6
Precipitation (mm)	June-August	−36.57	−14.73

[Source: NOAA (2020)].

2. Methods

2.1. Sampling locations and timing

The Mink Creek watershed predominately contains lands managed by the Caribou-Targhee National Forest, with smaller sections of private and state-managed land, in which grazing and recreation co-occur. Samples were collected from 14 locations across the watershed in 2019, of which IDEQ had previously sampled seven sites in 2017 (Fig. 1). The sites sampled by IDEQ in 2017 include site 2, site 4, site 5, site 6, site 7, site 9, and site 11. The sites added for this study include site 1, site 3, site 8, site 10, site 12, site 13, and site 14. The new seven sites were selected based on cattle grazing and recreational locations.

The temperature and precipitation anomaly for the years 2017 and 2019 compared to the period 1981–2010 are shown in Table 1 (NOAA 2020). Both the years have negative precipitation anomalies indicating low rainfall compared to the mean rainfall of 1981 to 2010. The year 2019 has a positive temperature anomaly for June, July, and August. The year 2017 has a negative temperature anomaly for June and a positive temperature anomaly for July and August. In comparison between 2017 and 2019, the year 2019 was warmer and received relatively high rainfall.

Recreation activities (mountain biking, hiking, dog walking, camping, day-use) occurred throughout the watershed, while target shooting was mostly confined to sites 9 and 10. Some sites were in closer proximity to camping (site 11 and site 13), while other sites were close to day-use activities (site 7). Site 8 was near summer homes. Grazing occurred adjacent to site 1, site 2, site 3, site 5, site 9, site 10, site 11, site 12, site 13, and site 14.

Each location was sampled seven times over six months (May-October), resulting in 98 total samples. The first sampling period was at the end of May, before the start of the grazing and recreational seasons. The remaining sampling periods were set to mirror the movement of cattle and recreationists throughout the watershed. The primary goal was to sample at each site before, during, and after the cattle were present. The secondary goal was to time the sampling before and after holiday weekends.

Fig. 2 shows the spatial maps of cattle and humans' movement with E. coli exceedances marked for various measurement days. There were two herds of cattle in the watershed. Table 2 shows the movement of cattle in the watershed. The West herd had 362 cow/calf pairs; they were in the Catch unit (758 ha) from June 1 to June 30 (14 Animal Unit Days per Hectare, AUD/ha) and the Highway unit (1298 ha) July 1 to August 15 (13 AUD/ha). The Middle herd had 588 cow/calf pairs; they were in the Lead Draw unit (1350 ha) from June 3 to June 23 (9 AUD/ha), the Lower Cow Camp unit (670 ha) from June 24 to July 9 (13 AUD/ha), the Upper Cow Camp unit (1800 ha) from July 10 to August 12 (11 AUD/ha), and

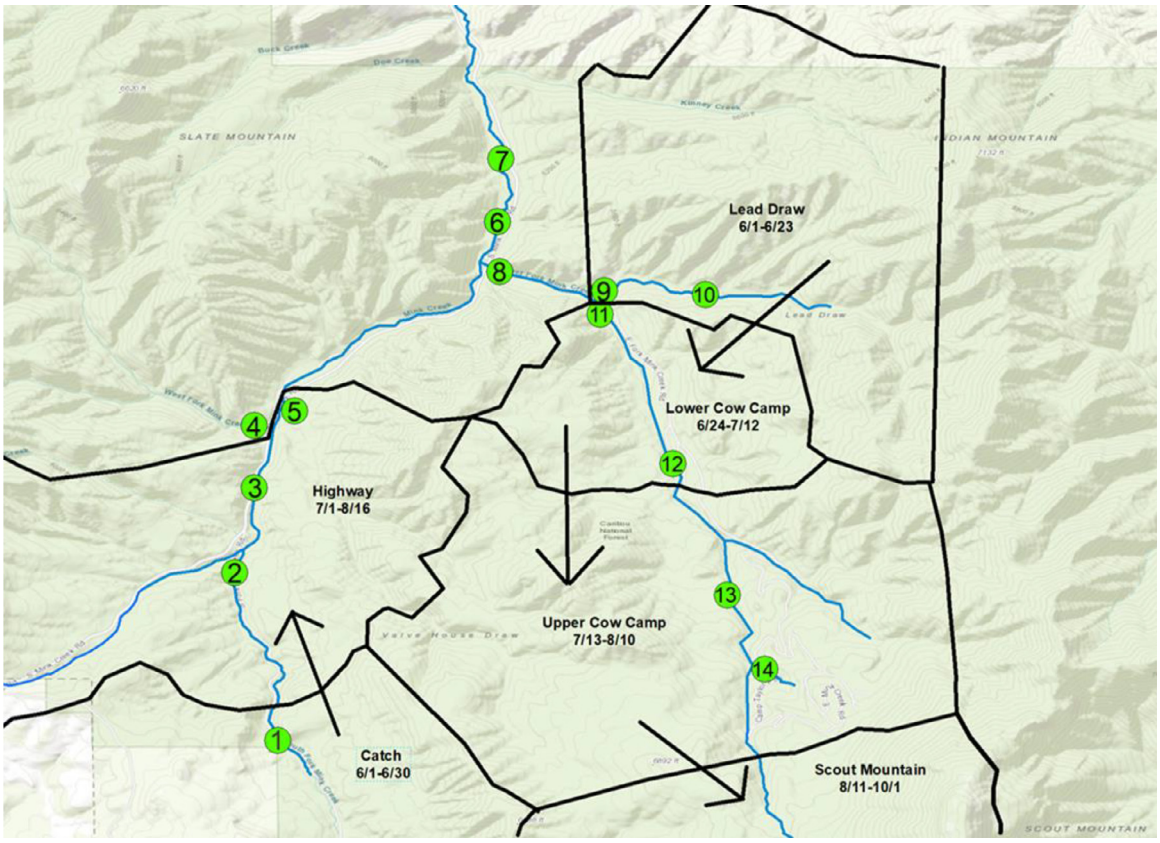

Fig. 1. Spatial map of sample locations and the movement of the two cattle herds throughout the grazing season.

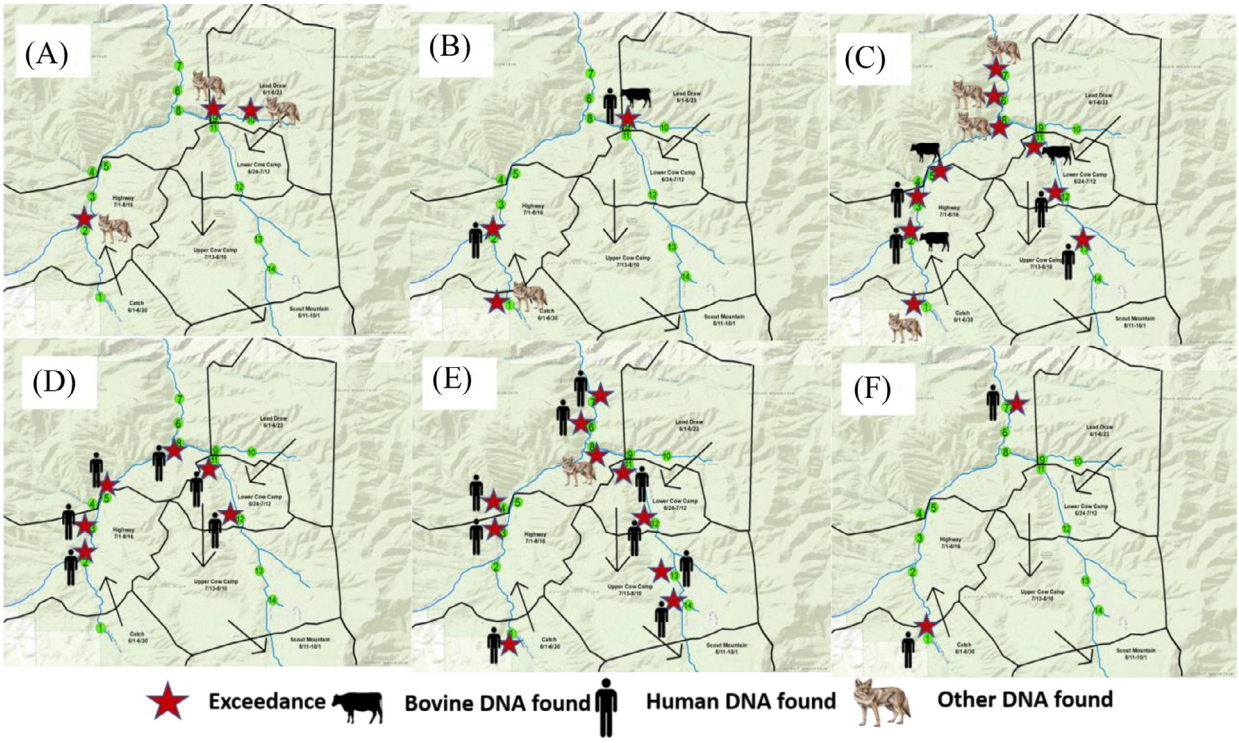

Fig. 2. Sites of E. coli exceedances marked with the presence of bovine DNA, human DNA, and other DNA for (A) June 5, (B) June 20, (C) July 8, (D) August 8, (E) September 5, (F) October 16.

Table 2
Cattle movement in the watershed.

Region	Site	Area (ha)	Cow/Calf pairs	Stocking rate (AUD/ha)	Grazing period
West Herd	Catch	758	362	14	June 1 - June 30
	Highway	1298	362	13	July 1 - August 15
Middle Herd	Lead Draw	1350	588	9	June 3 - June 23
	Lower Cow Camp	670	588	13	June 24 - July 9
	Upper Cow Camp	1800	588	11	July 10 - August 12
	Scout Mountain	2030	588	14	August 13 - October 2

the Scout Mountain unit (2030 ha) from August 13 to October 2 (14 AUD/ha).

Sites 4, 6, 7, and 8 were areas fenced off and free from grazing cattle. Site 4 was located in West Fork Mink Creek, where a substantial recreational activity occurred. Site 8 was located in the East Fork Mink Creek, directly downstream from units with grazing activity, and adjacent to several homes with septic systems. Site 6 was located on the Mink Creek's main stem, downstream of the East Fork and Mink Creek intersection. Finally, site 7 was located at Cherry Springs, which hosts considerable recreational activity.

As E. coli tends not to be carried far downstream (Clark et al., 2012), there is less possibility of cumulative loading of E. coli on the streams. While there is evidence that E. coli can survive for months in soil (Ishii et al., 2007), its survival in streams is typically measured in weeks unless they become part of the sediments at the bottom of the stream (Garzio-Hadzick et al., 2010). Hence the E. coli from upstream grazing activity on the downstream area was minimal.

2.2. Sample collection

Two replicate water samples of 350 mL each were collected at each sampling location for E. coli enumeration and MST marker quantification. Standard protocols per ISO 19458:2006 was followed for sample collection, and samples were immediately placed on ice until analysis (Standardization 2006). Efforts were taken to prevent cross-contamination among the sites. The samples collected for E. coli quantification were processed the same day. The collected samples were put in a −80 °c freezer and analyzed with qPCR (Quantitative Polymerase Chain Reaction) after collection.

The physical water quality parameters, including dissolved oxygen (DO; mg/L), water temperature (°C), specific conductivity (S/cm), and pH (YSI Pro20 and YSI Pro1030 instruments, YSI Inc., Yellow Springs, OH, USA) were also measured at each site following calibration of the equipment. In addition to the water quality parameters, daily maximum air temperature (°C), average air temperature (°C), and daily rain accumulation (converted to cm) were obtained from the Pocatello airport weather station, a distance of approximately 25 km from the sampled areas.

2.3. E coli enumeration and DNA extraction

E. coli counts (MPN/100 mL) were determined according to ISO 9308-2:2012 with an IDEXX, Colilert 18 quantification system (IDEXX Laboratories, Maine, USA) (Pitkänen, 2012). In one case at site 7, the difference between the two replicate estimates of E. coli exceeded 1000 MPN/100 mL. Hence, this data was excluded from additional analysis.

An additional 250 mL of each sample was filtered through 0.2 μm cellulose nitrate membranes (47 mm diameter). The resulting filter for each sample was aseptically transferred to 2 mL cryovials and stored at −80°C until DNA extraction. Negative control was generated during each sampling period by running sterile water through the same procedures conducted on the samples.

DNA from the filters was extracted using MoBio Powerwater® DNA Isolation Kit (MoBio Laboratories, California, USA) following the manufacturer's instructions. The total amount of DNA present in each sample was quantified using a Qubit 2.0 fluorometer (Thermo Fischer Scientific, New York, USA).

2.4. Quantitative PCR

Quantitative Polymerase Chain Reaction (qPCR) was used to assess the human and bovine contribution to the bacterial load in the water. The amplified genes are from Bacteroides bacteria and are strain-specific to human or cattle digestive tracks (Bernhard and Field 2000). Each sample was screened for the bovine bacterial gene (COWM2) following the protocol of Shanks et al. (2008). Similarly, the human bacterial gene (HF183) was identified following the protocol of Green et al. (2014).

Table 3 lists the primers and probes used for each assay. For each sample, we created 25 µl qPCR reactions that comprised of 10 µl of TaqMan Environmental Master Mix (version 2.0, Applied Biosystems, Thermo Fischer, New York, USA), 1 µl of 1 µM of each primer, 0.1 µl of 1 µM of the probe, 4.9 µl of nuclease-free H_2O and 3 µl of 3 ng μl^{-1} DNA template. PCR conditions were 10 mins at 95°C, followed by 40 cycles of 95°C for 30 s and 30 s at 60°C for annealing. Each sample was run in triplicate with appropriate negative controls to confirm that no contamination, primer-dimers, or other artifacts amplified. If DNA amplification occurred in only one of the three triplicates in the rare case, those data were excluded from further analysis.

A double-stranded gBlock standard (Integrated DNA Technologies, Iowa, USA) fragment of DNA was designed for each target region to generate our standard curves (Table 3). Standard curves were generated using 10-fold dilutions that ranged from 10×10^{-8} to 10×10^{-2} ng of DNA reaction^{-1}. The target copy number for each gene target was then estimated from the respective standard curve. In this study, we use the term cattle DNA and human DNA throughout the manuscript, which corresponds to the Bacteroide markers associated with cattle and human sources, respectively.

2.5. Statistical analysis

Data collected were analyzed as either continuous data (pH, DO, conductivity, maximum air temperature, and Pocatello precipitation data) or as binary (presence/absence) data (cattle DNA, human DNA, grazing, holiday), or as count-based (E. coli). Since three out of eleven variables under consideration were temperature variables (stream temperature, maximum air temperature, and average air temperature), which are highly correlated with each other (average correlation coefficient, $R = 0.7$), we selected the maximum air temperature. We excluded the other two temperature variables for the remaining analysis. The maximum air temperature was selected as it showed the highest correlation with the E. coli. Table 4 shows the correlation matrix of the datasets used, and it is evident that most of these datasets are cross-correlated. Among the 45 comparisons, 20 were significant at the P-value < 0.01 suggesting considerable cross-correlation among the variables. E. coli was significantly correlated with cattle DNA, holiday, grazing, DO, and conductivity.

Table 3
The oligotide sequences that were used as primers for the HF183 (Human) gene and the CowM2 (cattle) gene.

Host Species	Gene	Primer	Oligonucleotide Sequence (5′−>3′)	Annealing Temperature, °C
Human	HF183	HF183F	ATCATGAGTTCACATGTCCG	60
		BacR287	CTTCCTCTCAGAACCCCTATCC	60
		Probe	FAM-CTAATGGAACGCATCCC-MGB	
		Standard	ATCATGAGTTCACATGTCCGCATGATTAAAGGTATTTT CCGGTAGACGATGGGGATGCGTTCCATTAGCTCGAGATAG TAGGCGGGGTAACGGCCCACCTAGTCAACGATGGATAG GGGTTCTGAGAGGAAGGTCCCCCACATTGGAACTGAGACACGGT CCAAACTCCTACG	
Cow	CowM2	CowM2F	CGGCCAAATACTCCTGATCGT	60
		CowM2R	GCTTGTTGCGTTCCTTGAGATAAT	60
		Probe	[6FAM]AGGCACCTATGTCCTTTACCTCATCAACTACAGACA[TAM]	
		Standard	ATCGCGGCCAAATACTCCTGATCGTACTCGAGATAG GCACCTATGTCCTTTACCTCATCAACTACAGACAAAATTAT CTCAAGGAACGCAACAAGCATCGCCTCTAATGGAA AATGGATGGTATCTTTGGAGCCTTTGAAAGCACTCGA GCCTTATGCATTGAGCATCGAGGCCGGAAAGCAGGAACTTATATAT AATAAGGTATTAGCAGGCGAAGTATGGATCG	

Table 4
Correlation matrix for the ten variables measured during the field season. An asterisk denotes a correlation that is statistically significant at the P-value < 0.01.

Variables	E.Coli	Cattle DNA	Human DNA	Holiday	Grazed	pH	DO (mg/L)	Conductivity (mS/cm)	Pocatello Air temp (Max)	Pocatello Precipitation
E. coli*	1.00	0.45*	0.04	0.28*	0.46*	0.18	−0.52*	0.30*	0.24	0.17
Cattle DNA		1.00	−0.02	0.13	0.46*	0.07	−0.21	0.11	−0.13	−0.16
Human DNA			1.00	−0.23	0.01	0.39*	−0.21	0.20	0.19	0.37*
Holiday				1.00	−0.10	−0.10	−0.42*	0.05	0.32*	0.03
Grazed					1.00	0.11	−0.17	0.04	−0.07	0.06
pH						1.00	−0.39*	0.67*	0.28*	0.40*
DO (mg/L)							1.00	−0.48*	−0.63*	−0.43*
Conductivity (mS/cm)								1.00	0.28*	0.37*
Pocatello Air temp (Max)									1.00	0.67*
Pocatello Precipitation										1.00

Table 5
Model specifics for mixed-effects regression approach.

Number of observations	97
Fixed effects coefficients	10
Random effects coefficients	21
Covariance parameters	2
Distribution	Poisson
Link	Log
Fit method	Laplace

Since the variables include continuous, binary, and count-based, a mixed-effects regression approach was adopted to model this dataset. MATLAB R2019a tool was used in this study for model simulation (MATLAB, 2019). Since the outcomes are discrete counts, either Poisson or negative binomial distribution could be adopted. Pearson Chi-squared dispersion statistic was then used to choose between Poisson and negative binomial distribution (Loukas and Kemp, 1986). Since the statistic was equal to one, Poisson distribution was selected for this study.

Table 5 shows the model details adopted for the mixed regression approach. Since the distribution selected was Poisson, the link function selected to map the relationship between the mean response and the linear combination of the predictors was logarithmic. The fit method selected was Laplace, with site and sample date as intercepts. The number of random intercepts was 21, which includes 14 sites and 7 dates. The binary variables cattle DNA, human DNA, holiday, and grazing were considered as categorical variables. A generalized mixed-effects model was developed, and the model coefficients were evaluated using t-statistics, P-value, and confidence intervals. The model output (E. coli) was calibrated against 97 observations from the 14 sites in Mink Creek. The model output was further evaluated using R-squared and Nash-Sutcliffe Efficiency (NSE). The random effects were analyzed using covariance parameters for site and sample date.

The model performance was also evaluated using K-fold cross-validation (Fushiki, 2011). This validation scheme was applied to each site by randomly choosing 70% training data and 30% testing data. This was repeated for the maximum possible combinations of training and testing datasets (21 simulations per each site) to avoid selection bias (Rodriguez et al., 2009). The model results were further evaluated based on the statistical performance indices. The statistical performance indices used in this study as a measure of model validation include the coefficient of determination, R^2 (Glantz and Slinker 1990) and Nash-Sutcliffe Efficiency, NSE (Nash and Sutcliffe 1970).

3. Results

3.1. E. coli exceedances

The variation of E. coli by 14 sites and seven measurement days are shown in Fig. 3A and Fig. 3B. Of the 98 water samples collected, E. coli counts exceeded the IDEQ primary contact regulatory limit of 126 MPN/100 mL 34 times (35%). The number of exceedances by date and site is also listed in Table 6. All 14 sites

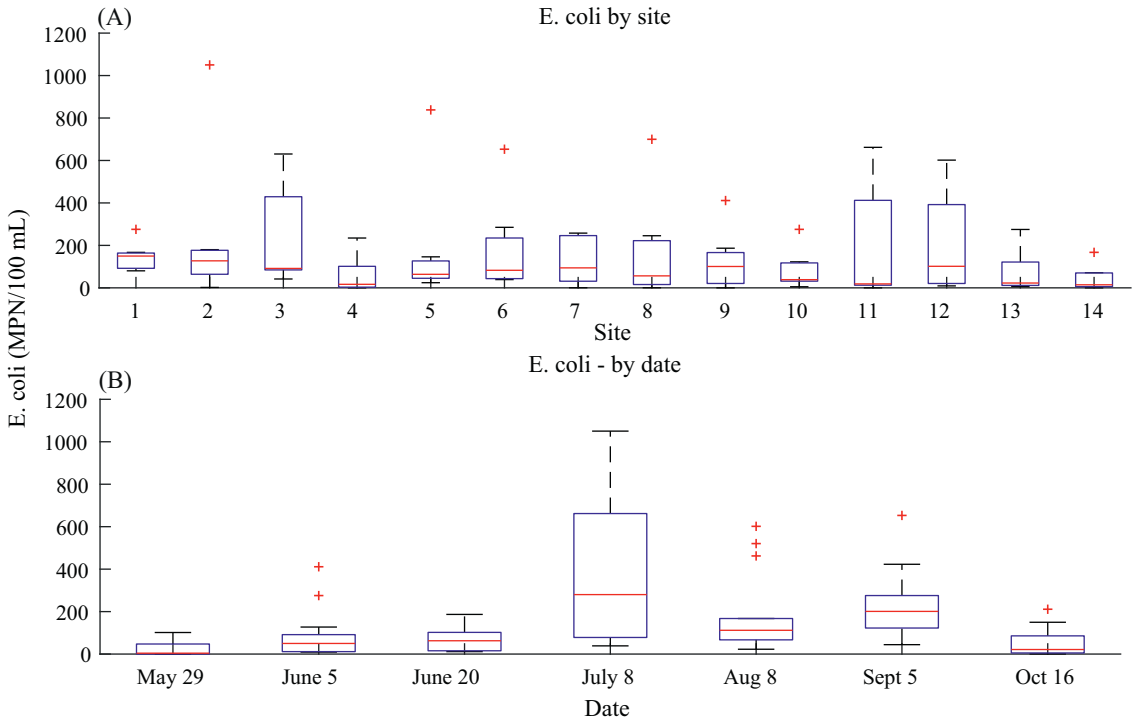

Fig. 3. (A) Boxplot of E. coli by 14 sites, (B) Box plot of E. coli by date [E. coli counts are in MPN/100 mL].

Table 6
E. coli exceedances at 14 sites for each sampling date [When cattle DNA was found in a sample, it was marked with a "#', while human DNA found at the site was marked with a '@'.].

Site	May 29	June 5th	June 20th	July 8th	August 8th	September 5th	October 16th	Total
Site 1	0	0	1	1	0	1@	1@	4
Site 2	0	1	1@	1#@	1@	0	0	4
Site 3	0	0	0	1@	1	1@	0	3
Site 4	0	0	0	0	0	1@	0	1
Site 5	0	0	0	1#	1@	0	0	2
Site 6	0	0	0	1	0	1@	0	2
Site 7	0	0	0	1	0	1@	1@	3
Site 8	0	0	0	1	1@	1	0	3
Site 9	0	1	1#@	0	0	0	0	2
Site 10	0	1	0	0	0	0	0	1
Site 11	0	0	0	1#	1@	1@	0	3
Site 12	0	0	0	1@	1@	1@	0	3
Site 13	0	0	0	1@	0	1@	0	2
Site 14	0	0	0	0	0	1@	0	1
Total	0	3	3	10	6	10	2	34

had at least one exceedance, while sites 1 and 2 experiencing exceedances four out of the seven sampling dates (57%). Of the seven sample dates, only May 29 had no exceedances. The sampling dates of July 8 and September 5 experienced 10 exceedances out of the 14 sites sampled on that date (71%).

Relative to the secondary contact regulatory limit of 576 MPN/100 mL, eight sites exceeded the limit. The highest number of sites exceeding this limit occurred on July 8, with five out of 14 sites (36%) showing values over 576 MPN/100 mL. The highest E. coli count was recorded at Site 7 on September 5, with 1732 MPN/100 mL.

Fig. 4 shows the E. coli values for the 14 sites in the forward axes, with cattle DNA and human DNA presented on the reverse axes. Of the 34 sites that exceeded the IDEQ primary contact standard, human DNA was detected at 20 of those sites (58.8%), and cattle DNA was detected at two sites (5.9%). Both human and cattle DNA were detected at two sites (5.9%). Neither human nor cattle-specific DNA was found at ten sites (29.41%). The non-specific DNA may have been contributed by wildlife or domestic canines. Human DNA was found most frequently on the September 5th sampling date, with nine sites.

3.2. E. coli model results

A mixed-effects regression model using Poisson distribution was applied for all 14 sample sites at Mink Creek. Table 7 shows the fixed effects coefficients and confidence intervals for the variables – cattle DNA, human DNA, holiday, grazed, pH, DO, conductivity, maximum air temperature, precipitation, and intercept. The degree of freedom was 87, obtained by subtracting the number of variables (9 variables and intercept) from the total number of observations (97). The t-statistic for these variables was also estimated, and it was identified that cattle DNA, grazing, and DO have higher t-statistic (tStat$>$10). The P-value of these variable coefficients was also small (P-value$<<$0.01), suggesting high significance for these variables. Additionally, variables such as human DNA, pH, conductivity, and intercept shows high significance with P-value

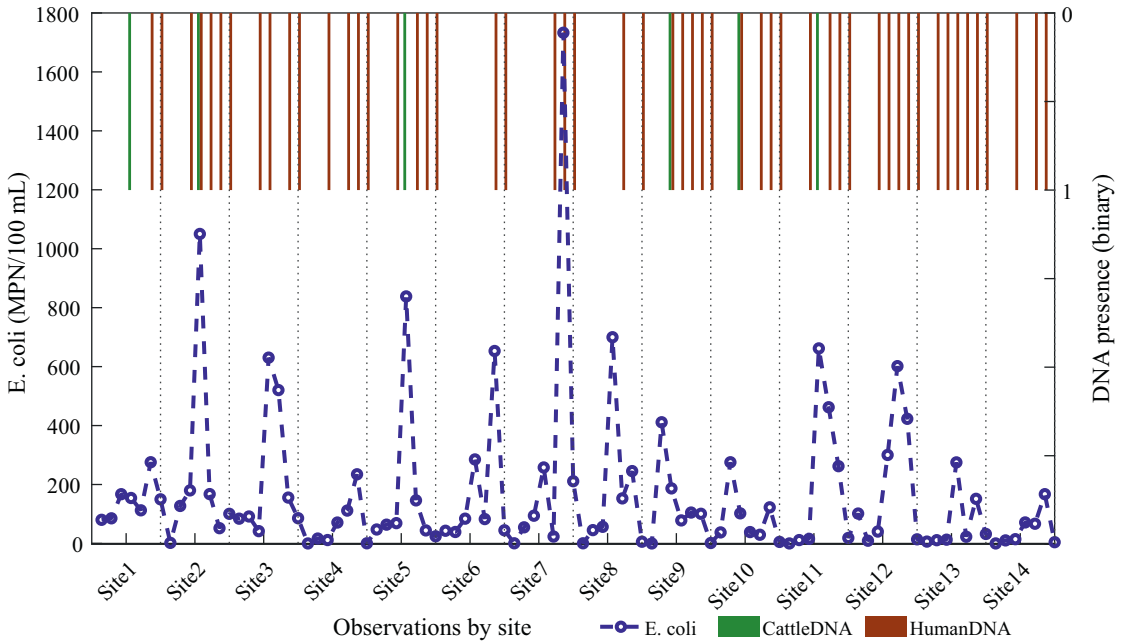

Fig. 4. Comparison of E. coli with Cattle DNA and Human DNA presence for all the 14 sites. The dots within each site represents the progression of the sampling season at that site. The sites are arranged following Fig. 1 and do not reflect upstream to downstream positioning or a stream profile.

Table 7
Fixed effects coefficients of the model with confidence intervals.

Name	Estimate	SE	tStat	DF	P-value	Lower C.I.	Upper C.I.
CattleDNA	−0.25	0.02	−13.88	87	8.83E-24	−0.29	−0.21
HumanDNA	0.06	0.01	4.39	87	3.19E-05	0.03	0.09
Holiday	0.08	0.13	0.58	87	0.56494	−0.18	0.34
Grazed	−0.36	0.01	−27.56	87	9.45E-45	−0.39	−0.34
pH	−0.72	0.14	−5.11	87	1.92E-06	−1.01	−0.44
DO	−0.51	0.03	−17.02	87	1.99E-29	−0.57	−0.45
Conductivity	0.68	0.23	3.02	87	0.00336	0.23	1.13
MaxAirTemp	0.02	0.02	0.79	87	0.43053	−0.03	0.06
Precipitation	6.81	5.20	1.31	87	0.19377	−3.53	17.16
Intercept	14.28	2.00	7.13	87	2.78E-10	10.30	18.27

SE – Standard error, tStat – t-statistic, DF – degree of freedom, C.I. – Confidence interval.

<0.01. The coefficients of the variables such as holiday, maximum air temperature, and precipitation were found to hold a less significant impact, as evidenced by low t-statistic and high P-value. The random covariance estimated was positive, with the covariance of 0.43 for site and 0.31 and sample date. The random effect was high for site 12 and site 3 and dates July 8 and June 20.

Fig. 5 shows the modeled and observed E. coli at each of the 14 sites. The dots within each site represents the progression of the sampling season at that site. The model results agree with the observed data for most sites by estimating the low E. coli counts and peak values reasonably good. Nevertheless, the E. coli peak for sites 5, 6, 8, 9, and 13 were slightly underestimated. The model against observed E. coli counts is also shown in Fig. 6. The model results agree with the observed data for most of the observations. Further, a K-fold cross-validation scheme was adopted to verify the model results using the available datasets. Based on the cross-validation scheme, R-squared and Nash-Sutcliffe Efficiency (NSE) measures were also estimated as 0.83 and 0.69, respectively, supporting the model selection.

4. Discussion

The findings of this study suggest that E. coli exceedances have a close association with the presence of human recreation, followed by unknown sources and cattle grazing. Out of the 34 E. coli exceedances, 20 exceedances corresponded to human DNA alone, two exceedances corresponded to cattle DNA alone, two exceedances corresponded to both cattle and human DNA, and ten exceedances corresponded unknown sources. The contribution of humans, cattle, unknown sources, and dissolved oxygen are further discussed in Sections 4.1, 4.2, 4.3., and 4.4, respectively.

4.1. Human contribution to E. coli counts

Human DNA was found at 65% of the instances that had E. coli counts over 126 MPN/100 mL (22 out of 34 sites). Among this, 58.8% (20 instances) corresponded to human DNA alone, and 5.9% (2 instances) corresponded to both human and cattle DNA. The dates with the highest number of E. coli exceedances were July 8 and September 5 (10 exceedances each); human DNA was found at 4 and 9 sites, respectively. Those dates were also immediately after the summer holidays. The correlation matrix showed that the human DNA holds a statistically significant relationship with E. coli in Mink Creek. Moreover, the mixed-effects regression model found that a holiday's presence positively influenced E. coli, and the regression coefficient is statistically significant at P-value < 0.01. This study's results reinforce the findings of previous studies (e.g., Cao et al. (2018)) who inferred that exposure to human waste poses a higher health risk than other sources.

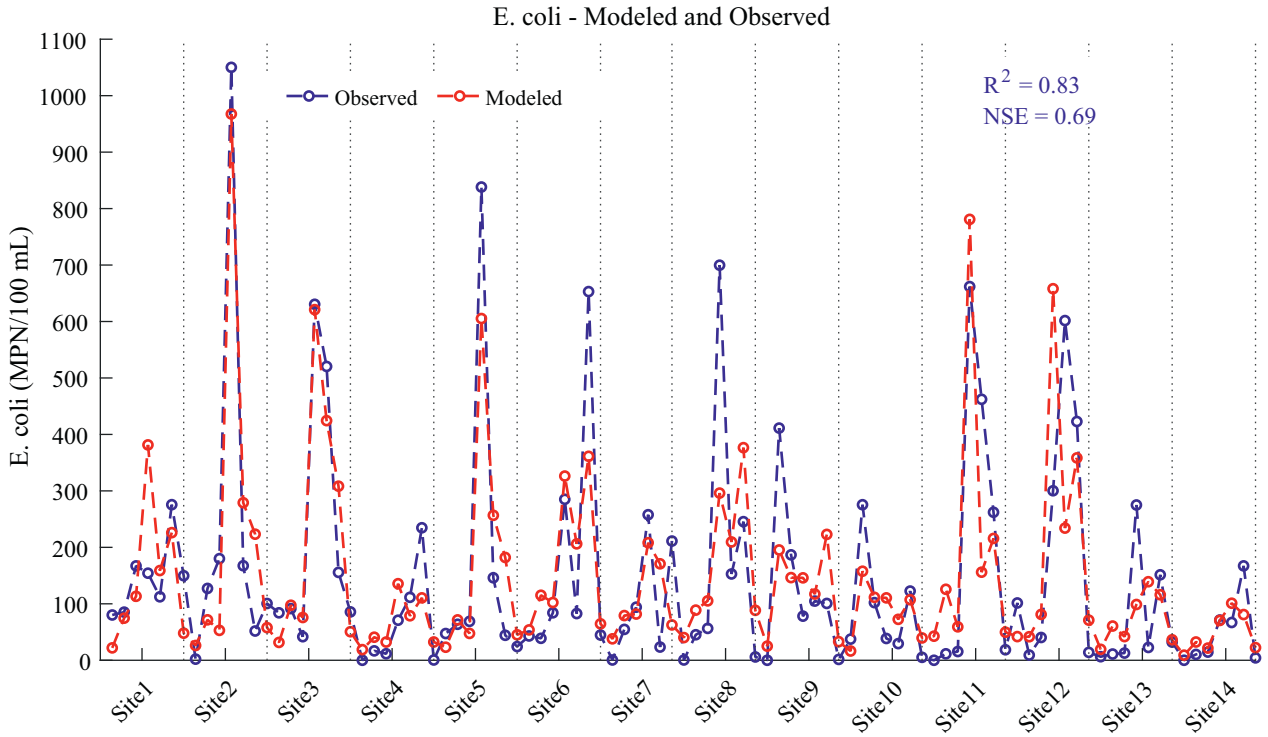

Fig. 5. E. coli modeled and observed for 14 sites in Mink Creek.

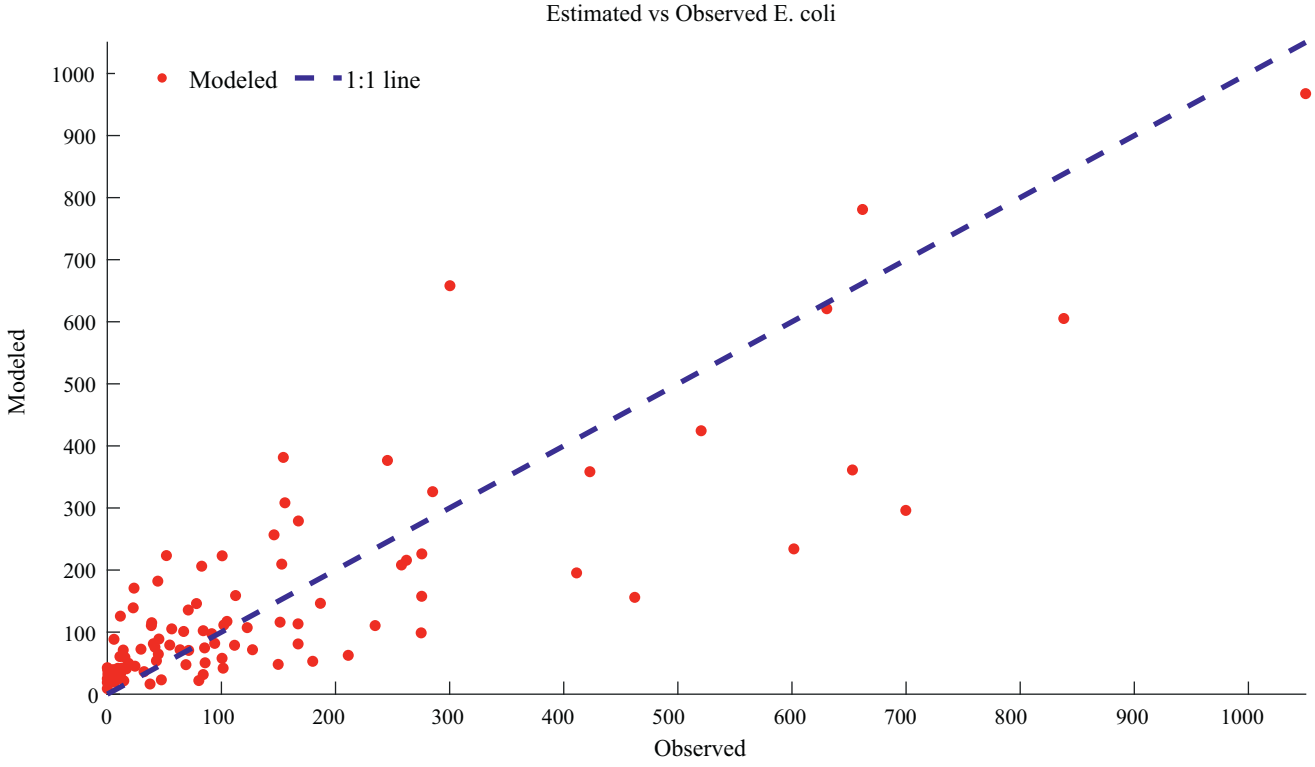

Fig. 6. Modeled against observed E. coli with the 1:1 line added.

4.2. *Cattle contribution to* E. coli *counts*

Cattle DNA was found at 12% of the instances that had E. coli counts that exceeded 126 MPN/100 mL (4 out of 34 instances), out of which two instances corresponded to cattle DNA alone, and the remaining two instances corresponded to both cattle and human DNA. Issues with stream loading of E. coli from cattle are strongly associated with periods in which ambient air temperatures are elevated. Similar to humans, cattle favor shady spots closer to water under these high-temperature conditions (Sprinkle et al., 2019). Cattle are considered to move into mild heat load when the temperature-humidity index exceeds 72 (Du Preez et al. 1990). For

Idaho, this occurs at around 79 to 80°F. The date with the most cattle-driven E. coli exceedances was July 8 (3), all of which exceeded the secondary regulatory limit (576 MPN/100 mL). It should be noted that climate data on July 7 indicated that cattle on Mink Creek likely experienced mild heat load for at least 6 h from midday to late afternoon.

4.3. Instances when the source of E. coli was not found

The DNA analysis did not find either human or cattle-specific DNA in 29% of the instances (10 out of 34). The qPCR design looked for Bacteroides specific to and ubiquitous in either cattle or humans; thus, we are confident that those two groups were not the source of this contamination. Other possible coliform sources were not directly tested using source tracking methods specific for other animals (e.g., dogs). Emanating from previous studies, the likely sources of this contamination could be dogs, wildlife, or soil-borne coliforms (Ishii et al., 2006; Ervin et al., 2014).

4.4. The role of dissolved oxygen in E. coli counts

When cattle graze or humans recreate in the vicinity of a stream, there is potential for an increase in sediment, nutrient, and fecal bacteria levels (Grudzinski et al., 2020). This will increase the biological oxygen demand that will be needed to metabolize the organic material ultimately. In such a degraded environment, microbes' metabolic activities will increase, as will their oxygen consumption (Mulholland et al., 2005). Consequently, when organic material is found in abundance within a stream reach, DO levels will be reduced, explaining the negative relationship between DO and E. coli counts found within this study.

In this study, DO was also inversely related to air temperature ($R = -0.63$). This decline in DO happens in parallel with decreasing water flows since most of this watershed is fed by snowmelt and spring rain. With the decrease in water flow and an increase in water temperature during the summer, the environment is optimal for the E. coli growth.

4.5. Management implications

Since most DEQ exceedances of E. coli in this study were attributed to humans, novel approaches to public education and collaborative land management are necessary (Wolf et al., 2017). US Forest Service employees should consider a recreation plan that helps minimize fecal coliform impacts during the holidays and when temperatures are elevated. This is necessary as the non-metropolitan population growth in the western U.S. is three times higher than the rest of the country and occurs disproportionately on forests and rangelands (Hansen et al., 2002). A collaborative effort from land managers and grazers for developing approaches for grazing and recreation without adversely affecting natural resources is required (Walker et al., 2002; Sayre 2005). Moreover, education and management methods such as public surveys, educational workshops, improved on-site signage, more user-friendly and up-to-date websites, and educational materials could be employed (Wolf et al., 2017). Further, keeping clean toilets available, locating temporary toilets in high use areas further up the stream reach, using occupancy targets at higher use areas, and educational kiosks could all be possible considerations.

Managing the distribution and timing of grazing can lead to improved soil parameters [e.g., Byrnes et al. (2018)] and stream water quality [e.g., Agouridis et al. (2005)]. Recent studies have identified that grazing management practices significantly impact water quality improvement at the watershed scale (Park et al., 2017). The management options include grazing lower elevation areas of Mink Creek when temperatures are cooler and grazing high elevation pastures when ambient temperatures are maximum. Over a longer period, the development of off-stream water sources (such as solar pumps, ram pumps) could be evaluated for feasibility. Another thing that could be considered by permittees over a long period is selecting replacement heifers with the capability and inclination to access higher elevation areas of the pasture when temperatures increase during the summer (Sprinkle et al., 2019).

5. Conclusions

The important findings of this study can be listed as:

- Microbial source tracking markers coupled with the fecal indicator bacteria E. coli was found to be an effective method of quantifying source-wise fecal contamination in the Mink Creek watershed. Both bovine and human-associated marker genes of E. coli were identified using quantitative polymerase chain reaction (qPCR).
- Most number of E. coli exceedances in the Mink Creek watershed corresponded to the presence of human DNA (58.8%), followed by unknown sources (29.4%), and cattle DNA (5.9%). Similarly, most E. coli exceedances occurred on the sampling dates of July 8 and September 5, each date with 10 exceedances in 14 sites (71%). It should be noted that the five exceedances on July 8 exceeded the secondary contact regulatory limit (576 MPN/100 mL).
- This study developed a mixed-effects regression model for estimating E. coli using microbial source tracking markers of cattle and humans. The model results show good agreement with the observed data based on a k-fold cross-validation scheme (R-squared = 0.83 and NSE = 0.69).
- Human DNA, dissolved oxygen, cattle DNA, and grazing were identified as the most significant predictors of E. coli (P-value < 0.01) in the model developed. While human and cattle presence drove the E. coli count increase, the DO reduction was caused by an increase in E. coli bacteria and reduced streamflow. E. coli was insensitive to the variations in maximum air temperature and precipitation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Rob Mickelsen, Hans Bastian, and Heidi Heyrend of the US Forest Service, Caribou-Targhee National Forest, and Curlew National Grassland for their financial and logistical support. Also, thanks to Idaho State University, who provided laboratory space to perform the coliform analysis. Partial support for this project was provided by the University of Idaho and the Idaho Water Rources Research Institute.

References

Agouridis, C.T., Workman, S.R., Warner, R.C., Jennings, G.D., 2005. Livestock grazing management impacts on stream water quality: a review 1. JAWRA J. Am. Water Resour. Assoc. 41 (3), 591–606.

Allocca, V., Marzano, E., Tramontano, M., Celico, F., 2018. Environmental impact of cattle grazing on a karst aquifer in the southern Apennines (Italy): quantification through the grey water footprint. Ecol. Indic. 93, 830–837.

Baez, A., Shiloach, J., 2014. Effect of elevated oxygen concentration on bacteria, yeasts, and cells propagated for production of biological compounds. Microb. Cell Fact. 13 (1), 181.

Baez, A., Shiloach, J., 2017. Increasing dissolved-oxygen disrupts iron homeostasis in production cultures of Escherichia coli. Antonie Van Leeuwenhoek 110 (1), 115–124.

Bernhard, A.E., Field, K.G., 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA. Appl. Environ. Microbiol. 66 (10), 4571–4574.
Buckerfield, S.J., Quilliam, R.S., Waldron, S., Naylor, L.A., Li, S., Oliver, D.M., 2019. Rainfall-driven E. coli transfer to the stream-conduit network observed through increasing spatial scales in mixed land-use paddy farming karst terrain. Water Res. X 5, 100038.
Byrnes, R.C., Eastburn, D.J., Tate, K.W., Roche, L.M., 2018. A global meta-analysis of grazing impacts on soil health indicators. J. Environ. Qual. 47 (4), 758–765.
Cao, Y., Sivaganesan, M., Kelty, C.A., Wang, D., Boehm, A.B., Griffith, J.F., Weisberg, S.B., Shanks, O.C., 2018. A human fecal contamination score for ranking recreational sites using the HF183/BacR287 quantitative real-time PCR method. Water Res. 128, 148–156.
Cho, K.H., Cha, S.M., Kang, J.-.H., Lee, S.W., Park, Y., Kim, J.-.W., Kim, J.H., 2010. Meteorological effects on the levels of fecal indicator bacteria in an urban stream: a modeling approach. Water Res. 44 (7), 2189–2202.
Clark, P.E., Moffet, C.A., Lewis, G.S., Seyfried, M.S., Hardegree, S.P., Pierson, F.B., 2012. Water quality effects of herded stream crossings by domestic sheep bands. J. Environ. Qual. 41 (5), 1580–1590.
Derlet, R.W., Carlson, J.R., 2006. Coliform bacteria in Sierra Nevada wilderness lakes and streams: what is the impact of backpackers, pack animals, and cattle? Wilderness Environ. Med. 17 (1), 15–20.
Devane, M.L., Moriarty, E.M., Robson, B., Lin, S., Wood, D., Webster-Brown, J., Gilpin, B.J., 2019. Relationships between chemical and microbial faecal source tracking markers in urban river water and sediments during and post-discharge of human sewage. Sci. Total Environ. 651, 1588–1604.
Du Preez, J., Giesecke, W., Hattingh, P., 1990. Heat Stress in Dairy Cattle and Other Livestock Under Southern African conditions. I. Temperature-humidity index mean values during the four main seasons.
Ervin, J.S., Van De Werfhorst, L.C., Murray, J.L., Holden, P.A., 2014. Microbial source tracking in a coastal California watershed reveals canines as controllable sources of fecal contamination. Environ. Sci. Technol. 48 (16), 9043–9052.
Fushiki, T., 2011. Estimation of prediction error by using K-fold cross-validation. Stat. Comput. 21 (2), 137–146.
Garzio-Hadzick, A., Shelton, D., Hill, R., Pachepsky, Y., Guber, A., Rowland, R., 2010. Survival of manure-borne E. coli in streambed sediment: effects of temperature and sediment properties. Water Res. 44 (9), 2753–2762.
Glantz, S., Slinker, B., 1990. Primer of Applied Regression and Analysis of Variance. Mc Grawhill, New York 136–41.
Green, H.C., Dick, L.K., Gilpin, B., Samadpour, M., Field, K.G., 2012. Genetic markers for rapid PCR-based identification of gull, Canada goose, duck, and chicken fecal contamination in water. Appl. Environ. Microbiol. 78 (2), 503–510.
Green, H.C., Haugland, R.A., Varma, M., Millen, H.T., Borchardt, M.A., Field, K.G., Walters, W.A., Knight, R., Sivaganesan, M., Kelty, C.A., 2014. Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Appl. Environ. Microbiol. 80 (10), 3086–3094.
Grudzinski, B., Fritz, K., Dodds, W., 2020. Does Riparian fencing protect stream water quality in cattle-grazed lands? Environ. Manage. 1–15.
Hansen, A.J., Rasker, R., Maxwell, B., Rotella, J.J., Johnson, J.D., Parmenter, A.W., Langner, U., Cohen, W.B., Lawrence, R.L., Kraska, M.P., 2002. Ecological Causes and consequences of demographic change in the new west: as natural amenities attract people and commerce to the rural west, the resulting land-use changes threaten biodiversity, even in protected areas, and challenge efforts to sustain local communities and ecosystems. BioScience 52 (2), 151–162.
Harris, H., 2017. Escherichia Coli Monitoring in the Mink Creek Watershed Bannock Co., Idaho, 2017. Quality, S.o.I.D.o.E..
Harwood, V.J., Staley, C., Badgley, B.D., Borges, K., Korajkic, A., 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. FEMS Microbiol. Rev. 38 (1), 1–40.
IDAPA, 2014. Idaho water quality standards. Idaho administrative code. IDAPA 58 (01), 02.
Ishii, S., Hansen, D.L., Hicks, R.E., Sadowsky, M.J., 2007. Beach sand and sediments are temporal sinks and sources of Escherichia coli in Lake Superior. Environ. Sci. Technol. 41 (7), 2203–2209.
Ishii, S., Ksoll, W.B., Hicks, R.E., Sadowsky, M.J., 2006. Presence and growth of naturalized Escherichia coli in temperate soils from Lake Superior watersheds. Appl. Environ. Microbiol. 72 (1), 612–621.
Khaengraeng, R., Reed, R., 2005. Oxygen and photoinactivation of Escherichia coli in UVA and sunlight. J. Appl. Microbiol. 99 (1), 39–50.
Layton, B.A., Cao, Y., Ebentier, D.L., Hanley, K., Ballesté, E., Brandão, J., Byappanahalli, M., Converse, R., Farnleitner, A.H., Gentry-Shields, J., 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. Water Res. 47 (18), 6897–6908.
Leclerc, H., Mossel, D., Edberg, S., Struijk, C., 2001. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. Annu. Rev. Microbiol. 55 (1), 201–234.
Loukas, S., Kemp, C., 1986. The index of dispersion test for the bivariate Poisson distribution. Biometrics 941–948.
MATLAB (2019) 9.6. 0 (R2019a). The MathWorks Inc., Natick, Massachusetts 431.
Mieszkin, S., Yala, J.F., Joubrel, R., Gourmelon, M., 2010. Phylogenetic analysis of Bacteroidales 16S rRNA gene sequences from human and animal effluents and assessment of ruminant faecal pollution by real-time PCR. J. Appl. Microbiol. 108 (3), 974–984.
Mulholland, P.J., Houser, J.N., Maloney, K.O., 2005. Stream diurnal dissolved oxygen profiles as indicators of in-stream metabolism and disturbance effects: fort Benning as a case study. Ecol. Indic. 5 (3), 243–252.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—a discussion of principles. J. Hydrol. 10 (3), 282–290.
NOAA (2020) National Centers for Environmental information, Climate at a Glance: Global Mapping, https://www.ncdc.noaa.gov/cag/.
Park, J.-.Y., Ale, S., Teague, W.R., 2017. Simulated water quality effects of alternate grazing management practices at the ranch and watershed scales. Ecol. Modell. 360, 1–13.
Petersen, F., Hubbart, J.A., 2020. Advancing understanding of land use and physicochemical impacts on fecal contamination in mixed-land-use watersheds. Water 12 (4), 1094.
Pitkänen, T., 2012. Water Quality-Enumeration of Escheria coli and Coliform Bacteria-Part 2: Most probable Number Method.
Ponce-Terashima, R., Koskey, A.M., Reis, M.G., McLellan, S.L., Blanton, R.E., 2014. Sources and distribution of surface water fecal contamination and prevalence of schistosomiasis in a Brazilian village. PLoS Negl. Trop. Dis. 8 (10), e3186.
Reischer, G.H., Kasper, D.C., Steinborn, R., Farnleitner, A.H., Mach, R.L., 2007. A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. Lett. Appl. Microbiol. 44 (4), 351–356.
Roche, L.M., Kromschroeder, L., Atwill, E.R., Dahlgren, R.A., Tate, K.W., 2013. Water quality conditions associated with cattle grazing and recreation on national forest lands. PLoS One 8 (6), e68127.
Rodriguez, J.D., Perez, A., Lozano, J.A., 2009. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans. Pattern Anal. Mach. Intell. 32 (3), 569–575.
Sayre, N.F., 2005. Interacting effects of landownership, land use, and endangered species on conservation of southwestern US rangelands. Conserv. Biol. 19 (3), 783–792.
Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial source tracking: current methodology and future directions. Appl. Environ. Microbiol. 68 (12), 5796–5803.
Senkbeil, J.K., Ahmed, W., Conrad, J., Harwood, V.J., 2019. Use of Escherichia coli genes associated with human sewage to track fecal contamination source in subtropical waters. Sci. Total Environ. 686, 1069–1075.
Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Santo Domingo, J., Seifring, S., Sivaganesan, M., Haugland, R.A., 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. Appl. Environ. Microbiol. 74 (3), 745–752.
Sprinkle, J.E., Ellison, M.J., Hall, J.B., Yelich, J.V., Willmore, C.M., Brennan, J.R., 2019. Are low-residual feed intake cows adapted to rangelands? Transla. Anim. Sci. 3 (Supplement_1), 1797–1801.
Standardization, I.O.f., 2006. Water quality—sampling for microbiological analysis. ISO 19458 2006.
, 2011. Grazing Statistical Summary Fiscal Year 2009. Service, U.F., Washington, D.C, p. 108 ed.
Valeo, C., Checkley, S., He, J., Neumann, N., 2016. Rainfall and microbial contamination in Alberta well water. J. Environ. Eng. Sci. 11 (1), 18–28.
Walker, B., Carpenter, S., Anderies, J., Abel, N., Cumming, G., Janssen, M., Lebel, L., Norberg, J., Peterson, G.D., Pritchard, R., 2002. Resilience management in social-ecological systems: a working hypothesis for a participatory approach. Conserv. Ecol. 6 (1).
WHO, 2011. Guidelines for drinking-water quality. World Health Organ. Chronicle 38, 104–108 Fourth Edition(Fourth Edition).
WHO/UNICEF, 2015. Progress on Sanitation and Drinking water: 2015 Update and MDG Assessment. World Health Organization.
Williams, G.W., 2000. The USDA Forest Service: the First Century. USDA Forest Service.
Wolf, K.M., Baldwin, R.A., Barry, S., 2017. Compatibility of livestock grazing and recreational use on coastal California public lands: importance, interactions, and management solutions. Rangeland Ecol. Manage. 70 (2), 192–201.