



HAL
open science

Quantifying the effect of overland flow on *Escherichia coli* pulses during floods: use of a tracer-based approach in an erosion-prone tropical catchment

Laurie Boithias, Olivier Ribolzi, Guillaume Lacombe, Chanthamousone Thammahacksa, Norbert Silvera, Keooudone Latsachack, Bounsamay Soulileuth, Marion Viguier, Yves Auda, Elodie Robert, et al.

► To cite this version:

Laurie Boithias, Olivier Ribolzi, Guillaume Lacombe, Chanthamousone Thammahacksa, Norbert Silvera, et al.. Quantifying the effect of overland flow on *Escherichia coli* pulses during floods: use of a tracer-based approach in an erosion-prone tropical catchment. *Journal of Hydrology*, 2021, 594, pp.125935. 10.1016/j.jhydrol.2020.125935 . hal-03093896

HAL Id: hal-03093896

<https://hal.science/hal-03093896v1>

Submitted on 17 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

1 Quantifying the effect of overland flow on
2 *Escherichia coli* pulses during floods: use of a
3 tracer-based approach in an erosion-prone
4 tropical catchment

5
6 Laurie BOITHIAS^{1,*}, Olivier RIBOLZI¹, Guillaume LACOMBE², Chanthamousone THAMMAHACKSA³, Norbert
7 SILVERA⁴, Keoudone LATSACHACK³, Bounsamay SOULILEUTH³, Marion VIGUIER³, Yves AUDA¹, Elodie
8 ROBERT¹, Olivier EVRARD⁵, Sylvain HUON⁴, Thomas POMMIER⁶, Cyril ZOUITEN¹, Oloth
9 SENGTAHEUANGHOUNG⁷, Emma ROCHELLE-NEWALL⁴

10

11 ¹ Géosciences Environnement Toulouse, Université de Toulouse, CNES, CNRS, IRD, UPS, France

12 ² CIRAD, UMR G-EAU, F-34398 Montpellier, France

13 G-EAU, Univ Montpellier, AgroParisTech, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

14 ³ IRD, Department of Agricultural Land Management (DALaM), P.O. Box 4199, Ban Nogviengkham,
15 Xaythany District, Vientiane, Lao PDR

16 ⁴ Institute of Ecology and Environmental Sciences of Paris (iEES-Paris), Sorbonne Université, Univ Paris Est
17 Creteil, IRD, CNRS, INRA, 4 place Jussieu, 75005 Paris, France

18 ⁵ Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL), Unité Mixte de Recherche 8212
19 (CEA/CNRS/UVSQ), Université Paris-Saclay, Gif-sur-Yvette, France

20 ⁶ Laboratoire d'Ecologie Microbienne, UMR1418 INRA, UMR 5557 CNRS, UCBL, VetAgroSup, Université
21 de Lyon, Villeurbanne, France

22 ⁷ Department of Agricultural Land Management (DALaM), P.O. Box 4195, Ban Nogviengkham, Xaythany
23 District, Vientiane, Lao PDR

24 * Corresponding author: laurie.boithias@get.omp.eu

25

26

27 **Abstract**

28 Bacterial pathogens in surface waters threaten human health. The health risk is especially high in
29 developing countries where sanitation systems are often lacking or deficient. Considering twelve flash-
30 flood events sampled from 2011 to 2015 at the outlet of a 60-ha tropical montane headwater catchment
31 in Northern Lao PDR, and using *Escherichia coli* as a fecal indicator bacteria, our objective was to quantify
32 the contributions of both surface runoff and sub-surface flow to the in-stream concentration of *E. coli*
33 during flood events, by (1) investigating *E. coli* dynamics during flood events and among flood events and
34 (2) designing and comparing simple statistical and mixing models to predict *E. coli* concentration in
35 stream flow during flood events. We found that in-stream *E. coli* concentration is high regardless of the
36 contributions of both surface runoff and sub-surface flow to the flood event. However, we measured the
37 highest concentration of *E. coli* during the flood events that are predominantly driven by surface runoff.
38 This indicates that surface runoff, and causatively soil surface erosion, are the primary drivers of in-
39 stream *E. coli* contamination. This was further confirmed by the step-wise regression applied to
40 instantaneous *E. coli* concentration measured in individual water samples collected during the flood
41 events, and by the three models applied to each flood event (linear model, partial least square model,
42 and mixing model). The three models showed that the percentage of surface runoff in stream flow was
43 the best predictor of the flood event mean *E. coli* concentration. The mixing model yielded a Nash-
44 Sutcliffe efficiency of 0.65 and showed that on average, 89% of the in-stream concentration of *E. coli*
45 resulted from surface runoff, while the overall contribution of surface runoff to the stream flow was
46 41%. We also showed that stream flow turbidity and *E. coli* concentration were positively correlated, but
47 that turbidity was not a strong predictor of *E. coli* concentration during flood events. These findings will
48 help building adequate catchment-scale models to predict *E. coli* fate and transport, and mapping the
49 related risk of fecal contamination in a global changing context.

50

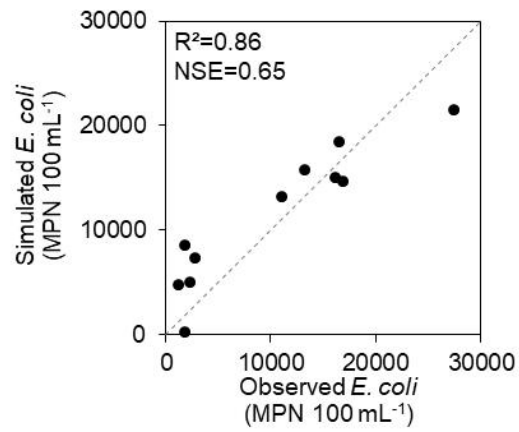
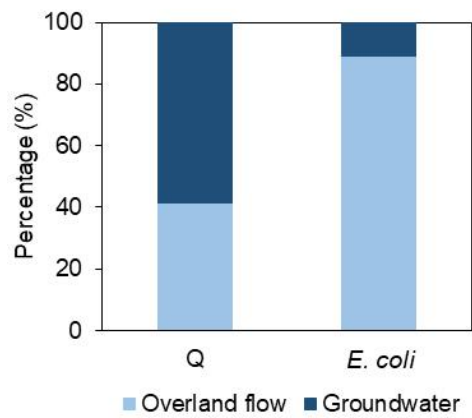
51 **Keywords**

52 Surface runoff; Fecal Indicator Bacteria; Storm flow; Land-use change; Surface-sub surface flow

53 separation; northern uplands of Lao PDR

54

55 **Graphical abstract**



56

57

58 1. Introduction

59 The presence of bacterial pathogens in surface waters threatens human health. Pathogenic bacteria are
60 etiological agents of several waterborne diseases such as diarrhea, which is globally a leading cause of
61 death among all ages (1.31 million deaths in 2015) (Troeger et al., 2017). These bacterial pathogens can
62 be of fecal origin, e.g. from cattle or from human feces where open defecation is practiced or where
63 sanitation systems are lacking or deficient (Exley et al., 2015; Tong et al., 2016). In developing countries,
64 untreated surface water is often used for drinking, cooking, bathing and washing, thus exposing the
65 population to a high health risk (Boithias et al., 2016).

66 Catchment microbial response depends on complex interactions between spatial patterns of land use,
67 soil types, antecedent conditions of soil, and rainfall characteristics (Kim et al., 2018; McKergow and
68 Davies-Colley, 2010; Pachepsky et al., 2018; Strauch et al., 2014). In the tropics, the presence in stream
69 water of Fecal Indicator Bacteria (FIB), such as *Escherichia coli* (*E. coli*), is known to be partly driven by
70 surface runoff (Causse et al., 2015; Ribolzi et al., 2016a) and groundwater flow (Chuah and Ziegler, 2018).
71 However, the underlying mechanisms of FIB dynamics during storms remain to be fully documented. For
72 example, there is a need to better quantify the respective contributions of surface and sub-surface flows
73 to stream flow (e.g. using geochemical tracers) to understand and model the transfers of *E. coli* at the
74 catchment scale during a flood event (Cho et al., 2016; Kim et al., 2017) and to predict the relative
75 contributions of surface and sub-surface flows to the overall in-stream *E. coli* contamination (Chin, 2011).

76 *E. coli* is known to be transported in water as free cells or cells attached to particles of soil, manure or
77 sediment (Garcia-Armisen and Servais, 2009; Krometis et al., 2007; Soupir et al., 2010). While stream
78 flow turbidity is often higher in tropical areas, as compared to temperate areas, suspended sediments
79 are known to carry *E. coli* (Nguyen et al., 2016). However, the relationships between stream flow
80 turbidity, suspended sediments concentration, and *E. coli* concentration, and their possible hysteretic
81 patterns during flood events, have not been yet fully explored.

82 Several studies in the tropics have investigated the relationships between suspended sediments
83 concentration and turbidity in rivers and lakes (Martinez et al., 2009; Robert et al., 2017, 2016; Ziegler et
84 al., 2014), and few studies investigated the relationships between suspended sediments concentration
85 and *E. coli* concentration (Boithias et al., 2016; Nguyen et al., 2016). However, to our knowledge, no
86 study has yet investigated the relationships between turbidity and *E. coli* concentration. Given that in-
87 stream *E. coli* concentration can vary rapidly during a flood event, a high-frequency sampling is required
88 to characterize this variability. The latter is both time consuming and expensive. Therefore, a proxy-
89 based approach is required and there is a need to verify if stream flow turbidity can be used as a proxy of
90 *E. coli* concentration during flood events.

91 In rural Southeast Asia, such as the uplands of Northern Lao PDR, livestock grazing and the lack of
92 adequate sanitation systems result in high levels of fecal contamination in surface waters during both
93 high and low flow seasons (Boithias et al., 2016; Nguyen et al., 2016; Rochelle-Newall et al., 2016). In
94 addition, Northern Lao PDR, like other mountainous regions in South-East Asia, is experiencing rapid
95 changes in land use (Riboldi et al., 2017; Turkelboom et al., 2008). Annual crops are replaced by tree
96 plantations (e.g. teak) with limited understorey and litter coverage. This land-use change decreases soil
97 water infiltration and increases surface runoff and soil erosion (Lacombe et al., 2018; Riboldi et al., 2017;
98 Song et al., 2020; Ziegler et al., 2004), but its impact on bacteria export has not yet been investigated.

99 Considering twelve flash-flood events sampled from 2011 to 2015 at the outlet of a 60-ha montane
100 tropical headwater catchment in Northern Lao PDR, our objective was thus to quantify the contributions
101 of both surface runoff and sub-surface flow to the in-stream concentration of *E. coli* during flood events,
102 by (1) investigating *E. coli* dynamics during flood events and among flood events and (2) designing and
103 comparing simple statistical and mixing models to predict *E. coli* concentration in stream flow during
104 flood events.

105 2. Materials and methods

106 2.1. Study design

107 In this paper, we quantify the relative roles of both surface runoff and sub-surface flow in the
108 contamination of stream water by *E. coli* in a tropical montane catchment. To do so, we measured fine
109 temporal-scale discharge and stream flow electrical conductivity to calculate the relative contributions of
110 surface runoff and sub-surface flow to total stream flow. For twelve flash-flood events from 2011 to
111 2015 (named with the letters A to L), we then investigated statistical relationships between surface
112 runoff, sub-surface flow, *E. coli* concentration, together with turbidity, total suspended sediments
113 concentration, land-use features (areal percentage of land cover and *E. coli* input), and flood event-
114 integrated variables, i.e. total rainfall, maximum rainfall intensity, rainfall duration, soil antecedent
115 conditions, peak discharge, total volume of stream water exported during the flood event, and flow
116 coefficient. For each flood event, we applied a mixing model to calculate the relative contributions of
117 surface runoff and of sub-surface flow to the overall in-stream contamination by *E. coli*, by assigning to
118 each one of these two flow components a characteristic bacteria concentration based on *E. coli*
119 measurements in overland flow (overland flow *E. coli* end-member) and in groundwater (groundwater *E.*
120 *coli* end-member).

121 2.2. Study area: the Houay Pano catchment

122 The 0.6 km² Houay Pano headwater catchment is located 10 km south of Luang Prabang city in Northern
123 Lao PDR (Fig. 1a), within the 800 000 km² Mekong basin. This experimental site (Boithias et al., 2020) is
124 part of the critical zone observatories' network named Multiscale TROPICAL Catchments (M-TROPICS),
125 which belongs to the French Research Infrastructure OZCAR (Gaillardet et al., 2018). This catchment can
126 be considered as being representative of the montane agro-ecosystems of South-East Asia. The climate is
127 sub-tropical humid and is characterized by a monsoon regime with a dry season from November to May,
128 and a wet season from June to October. The mean annual (2001-2019) temperature is 23.4 °C while the

129 mean annual rainfall is 1 366 mm (CV=0.23), about 71 % (CV=0.09) of which falls during the wet season.
130 Altitude within the catchment is 435-716 m (Fig. 1b) and the slope gradient is 1-135 % (mean=52 %). The
131 Laksip village, located downstream the S4 station (Fig. 1b), includes 484 inhabitants (Census of 2015).

132 2.3. Land use and land-use change

133 Detailed land-cover surveys and mapping were conducted each year from 2011 to 2015 within the
134 catchment (Boithias et al., 2020). The annual areal percentages of fallow (Fallow), teak trees (Teak),
135 annual crops (Annual crop), and annual crops grown under young teak trees (Teak+Crop) were calculated
136 using QGIS 2.6. We assessed the monthly *E. coli* input from 2011 to 2015 within the catchment with the
137 monthly counting of human and domestic animal (swine and poultry) traffic within the catchment, with
138 the daily feces production of humans and domestic animals, and with the *E. coli* content in their feces
139 (Causse et al., 2015).

140 2.4. Rainfall measurements

141 Rainfall was measured by an automatic weather station (Campbell BWS200 equipped with ARG100, 0.2
142 mm capacity tipping-bucket) located within the catchment (Fig. 1b). Data was recorded at 6-min time
143 interval from 2011 to 2012, and at 1-min time interval from 2013 to 2015. Data from 2013 to 2015 was
144 then cumulated into a 6-min time series for consistency with the 2011-2012 sub-period.

145 2.5. Flow rate measurements and water quality monitoring

146 We measured stream water level at the S4 gauging station of the catchment outlet (Fig. 1b) within a V-
147 notch weir equipped with a water level recorder (OTT Thalimedes) connected to a data logger, with 1-
148 mm vertical precision at a minimum of 3-min time interval. To relate water level to discharge, a control
149 rating curve was determined using both the velocity area method and the salt dilution method by slug
150 injection. Samples of stream water were collected at S4 gauging station in clean, plastic bottles using an

151 automatic sampler (Automatic Pumping Type Sediment Sampler, ICRISAT) for the measurement of Total
152 Suspended Solids concentration ([TSS]), turbidity (Turbidity), Electrical Conductivity at 25°C (EC), and
153 *E. coli* concentration ([*E. coli*]). The automatic sampler was triggered by the water level recorder to
154 collect water after every 2-cm water level change during flood rising and every 5-cm water level change
155 during flood recession.

156 [TSS] was measured in each sample after filtration on 0.2 µm porosity cellulose acetate filters (Sartorius)
157 and evaporation at 105 °C for 48 h. Turbidity was determined with a turbidity meter (EUTECH
158 Instruments TN-100) and EC with a conductivity meter (WTW340).

159 For nine flood events (events A, B, F-L), we measured [*E. coli*] with the standardized microplate method
160 (ISO 9308-3). A water sub-sample was incubated at four dilution rates (i.e. 1:2, 1:20, 1:200 and 1:2000) in
161 a 96-well microplate (MUG/EC, BIOKAR DIAGNOSTICS) for 48 h at 44 °C. Ringers' Lactate solution was
162 used for the dilutions and one plate was used per sample. The number of positive wells for each
163 microplate was noted and the Most Probable Number (MPN) per 100 mL was determined using the
164 Poisson distribution. For three flood events (events C-E), we measured [*E. coli*] following the membrane-
165 filter method (EPA Method 1603). A small quantity of each water sample was filtered through the filter
166 membrane, which retained the bacterial cells. After filtration, this membrane was placed on a selective
167 medium (Sartorius NKS Endo nutrient pads), and incubated at 44 °C for 24 h. Shiny green *E. coli* colonies
168 were directly counted and expressed as Colony Forming Units (CFU) per 100 mL. Although "colony
169 forming unit" techniques and "most probable number" techniques are known to give highly correlated
170 results (e.g. Cho et al., 2010; Lušić et al., 2016; Wohlsen et al., 2006), we previously ensured that the two
171 methods were consistent in the Houay Pano catchment conditions (unpublished work).

172 In addition, we monitored surface runoff in the different land uses of the catchment from 2011 to 2015
173 using 1-m² microplots (Patin et al., 2018). We annually installed between 6 and 29 microplots to collect

174 surface runoff water samples during each rainfall event and to measure EC with WTW340. In 2012 and
175 2014, we used sub-samples of surface runoff water to measure *E. coli* concentration in surface runoff as
176 described previously. We calculated the mean concentration of *E. coli* in surface runoff by weighing the
177 *E. coli* concentration exported from each land use by the annual areal percentage of each land use ($[E.$
178 $coli]_{OF-TOT}$). We also measured the *E. coli* concentration in water samples taken from three piezometers
179 (Ribolzi et al., 2018) and calculated the mean concentration of *E. coli* in groundwater ($[E. coli]_{GW-TOT}$).

180 2.6. Electrical conductivity-based hydrograph separation

181 We used a tracer-based approach to separate storm hydrographs into ‘event water’ and ‘pre-event
182 water’. This approach relies on a simple mixing model with two end-members and EC as a tracer. It is of
183 relatively low cost compared to e.g. isotopic tracers and was successfully tested in the study catchment
184 (Ribolzi et al., 2018). Based on previous field observations and measurements performed in the same
185 study catchment (Patin et al., 2012; Ribolzi et al., 2011; Vigiak et al., 2008), the two end-members of the
186 model (i.e. overland flow EC end-member in event water, and groundwater EC end-member in pre-event
187 water) can be interpreted in terms of hydrological processes. Event water mainly includes infiltration
188 excess that produces overland flow along hillslopes. Pre-event water relates to groundwater that feeds
189 the stream during the storm event, plus the water in the stream channel prior to the storm event, which
190 is also related to groundwater outflows. As suggested by Collins and Neal (1998), we verified the linearity
191 between EC and the concentration of a conservative tracer to control the relevance of the EC-based
192 approach in our context (Ribolzi et al., 2018). The mixing model applied to individual samples is
193 described by the following equations:

$$Q = Q_{OF} + Q_{GW} \quad \text{Eq. 1}$$

$$Q \times EC = Q_{OF} \times EC_{OF} + Q_{GW} \times EC_{GW} \quad \text{Eq. 2}$$

194 where Q is the instantaneous stream water discharge at the catchment outlet ($L s^{-1}$), Q_{OF} is the
195 instantaneous discharge of overland flow, i.e., event water or surface runoff ($L s^{-1}$), Q_{GW} is the

196 instantaneous discharge of groundwater, i.e., pre-event water or sub-surface flow ($L s^{-1}$), EC is the
197 instantaneous electrical conductivity measured in the stream ($\mu S cm^{-1}$), EC_{OF} is the electrical conductivity
198 in overland flow (overland flow EC end-member; $\mu S cm^{-1}$), approximated from electrical conductivity
199 measurements in samples of overland flow collected at the soil surface on hillslopes draining to the
200 stream, and EC_{GW} is the electrical conductivity in groundwater (groundwater EC end-member; $\mu S cm^{-1}$),
201 approximated from the stream electrical conductivity at the beginning of the flood event, since
202 groundwater is the only supply of water to the stream during inter-storm flow periods (Ribolzi et al.,
203 2005). For each individual sample, we calculated the relative contributions of Q_{OF} and of Q_{GW} to Q based
204 on Eq. 2, namely $Q_{OF}\%$ and of $Q_{GW}\%$ (in %).

205 We assessed the uncertainty in estimating the contributions of overland flow and groundwater flow to
206 the total stream flow using the formulation of Genereux (1998), which assigns an uncertainty specific to
207 each term in Eqs. 1 and 2. The accuracy of Q is considered to be within $\pm 10\%$ of the measured value,
208 while the uncertainty of EC is approximately $\pm 5\%$ (Ribolzi et al., 2018). The uncertainties of both EC_{OF} and
209 EC_{GW} are estimated from the coefficients of variation (CV) of overland flow and groundwater samples,
210 respectively.

211 2.7. Flood-event variables

212 For each flood event, the mean concentration of *E. coli* ($[E. coli]_{mean}$) is equivalent to the average of the
213 instantaneous *E. coli* concentrations ($[E. coli]$) measured during the event, weighted by the
214 corresponding measurements of instantaneous discharge. We considered 13 candidate explanatory
215 variables to predict $[E. coli]_{mean}$:

- 216 - Rainfall duration (R_D , in min), total rainfall (R_{TOT} , in mm), maximum rainfall intensity (RI_{MAX} , in mm
217 h^{-1});

- 218 - Antecedent Precipitation Index (API, in mm), as a proxy of soil moisture conditions, calculated as
 219 $API_n = (API_{n-1} + P_{n-1}) \times e^{-\alpha \times t}$ where API_{n-1} is the antecedent precipitation index prior to
 220 rain event n-1 (mm), P_{n-1} is the rainfall precipitated during rain event n-1 (mm), α is the
 221 calibration coefficient usually set at 0.5 (day^{-1}), and t is the duration between P_n and P_{n-1} (day)
 222 (Descroix et al., 2002);
- 223 - Peak discharge during flood event (Q_{MAX} , in L s^{-1}), total volume of stream water during flood
 224 event (Q_{TOT} , in m^3), total volume of surface runoff, i.e. overland flow, during flood event (Q_{OF-TOT} ,
 225 in m^3) and its contribution to Q_{TOT} ($Q_{OF-TOT}\%$, in %), and volume of sub-surface flow during flood
 226 event (Q_{GW-TOT} , in m^3) and its contribution to Q_{TOT} ($Q_{GW-TOT}\%$, in %);
- 227 - Flow coefficient of the flood event (K_E , no dimension), calculated as the ratio between Q_{TOT} and
 228 R_{TOT} ;
- 229 - Flood-event mean concentration of total suspended sediments concentration ($[TSS]_{mean}$, in g L^{-1}),
 230 calculated by weighting [TSS] by the discharge, and flood-event mean turbidity ($Turbidity_{mean}$, in
 231 NTU), calculated by weighting Turbidity by time.

232 2.8. Hysteresis

233 Combined with EC-based hydrograph separation, the analysis of rotational direction, curvature, and
 234 trend of hysteresis relationships can be used to interpret the relative contributions from surface water
 235 and groundwater, and analyte sources (Holz, 2010; Long et al., 2017). Here, we analyzed hysteresis
 236 patterns of EC, Turbidity, [TSS] and [*E. coli*] against discharge; [TSS], Turbidity and [*E. coli*] against Q_{OF} ;
 237 Turbidity and [*E. coli*] against [TSS]; and [*E. coli*] against Turbidity. For each flood, we considered straight
 238 relationships when the p-value of the Pearson correlation was below 0.001. The threshold was
 239 voluntarily chosen very demanding to distinguish straight relationships from other hysteretic patterns.

240 2.9. Statistical analysis

241 We calculated correlations and regressions to identify explanatory variables predicting [*E. coli*] and [*E.*
242 *coli*]_{mean}. Statistics were calculated with R statistical package version 3.4.3 (correlations), Minitab 18.1
243 (stepwise regression), and XLSTAT 20.1.1 (partial least square regression). An explanatory variable was
244 considered to be statistically significantly different from zero when its p-value, derived from Student's t-
245 test, was lower than 0.05.

246 2.9.1. Modelling instantaneous *E. coli* concentration [*E. coli*]

247 We first calculated Pearson correlations between [*E. coli*] and the eight hydrological variables measured
248 or calculated for each sample (EC, Q, Q_{OF}, Q_{GW}, Q_{OF}%, Q_{GW}%, [TSS], and Turbidity). In addition, to select
249 the best set of explanatory variables predicting [*E. coli*], we derived linear regressions from the ‘step-
250 wise regression’ selection algorithm. This selection intended to maximize the prediction R² (R²_{pred})
251 calculated by leave-one-out cross-validations. This performance criterion reflects the ability of the model
252 to predict observations that were not used in the model calibration. Its maximization leads to greater
253 parsimony in the number of explanatory variables (Helsel and Hirsch, 2002). We verified the required
254 homoscedasticity of the model’s residuals by visual inspection. Multi-collinearity among the selected
255 explanatory variables was avoided by ensuring that the Variance Inflation Factor (VIF) never exceeded
256 the value of 8 (Helsel and Hirsch, 2002). We didn’t mix instantaneous measurements with flood-event
257 variables or with variables assessed annually (Annual crop, Teak+Crop, Teak, and Fallow) or monthly (*E.*
258 *coli* input).

259 2.9.2. Modelling mean concentration of *E. coli* per flood event [*E. coli*]_{mean}

260 We first calculated Pearson correlations between [*E. coli*]_{mean} and the thirteen hydro-meteorological
261 variables measured or calculated for each flood event (Table 1) and the five land-use related variables
262 (Annual crop, Teak+Crop, Teak, Fallow, and *E. coli* input). In addition, we used Partial Least Square (PLS)

263 regression to detect dependencies between variables. PLS is able to handle datasets with a number of
264 variables higher than the number of observations. It is also poorly sensitive to multi-collinearity, and
265 handles missing data by imputation. The importance of each projected variable is estimated by the
266 Variable Importance in the Projection (VIP). We discarded the variables for which VIP values were below
267 1 (Ribolzi et al., 2016b; Wold, 1995).

268 2.10. Mixing model of flood-event mean concentration of *E. coli*

269 We applied a simple mixing model with two end-members to predict $[E. coli]_{mean}$ and to separate the
270 contributions of both 'event water' and 'pre-event water' to the in-stream *E. coli* concentrations. Here
271 the two end-members of the model are *E. coli* concentration in event water (overland flow *E. coli* end-
272 member or $[E. coli]_{OF-TOT}$) and *E. coli* concentration in pre-event water (groundwater *E. coli* end-member
273 or $[E. coli]_{GW-TOT}$). The mixing model applied to flood-event mean variables is described by the following
274 equations:

$$Q_{TOT} = Q_{OF-TOT} + Q_{GW-TOT} \quad \text{Eq. 3}$$

$$Q_{TOT} \times [E. coli]_{mean} = Q_{OF-TOT} \times [E. coli]_{OF-TOT} + Q_{GW-TOT} \times [E. coli]_{GW-TOT} \quad \text{Eq. 4}$$

275 where Q_{TOT} is the total volume of stream water during flood event (m^3), Q_{OF-TOT} is the total volume of
276 surface runoff, or overland flow, during flood event (m^3), Q_{GW-TOT} is the total volume of sub-surface flow,
277 or groundwater flow, during flood event (m^3), $[E. coli]_{mean}$ is the flood-event mean concentration of *E. coli*
278 (MPN 100 mL⁻¹), $[E. coli]_{OF-TOT}$ is the mean concentration of *E. coli* exported in overland flow (MPN 100
279 mL⁻¹), calculated from *E. coli* concentration measured in samples of overland flow collected at the soil
280 surface on hillslopes draining to the stream, and $[E. coli]_{GW-TOT}$ is the mean concentration of *E. coli* in
281 groundwater (MPN 100 mL⁻¹), calculated from *E. coli* concentration measured in groundwater. This
282 approach was shown to yield a comparable or even higher performance when predicting in-stream
283 pathogen concentrations compared with more complex fate and transport models (Chin, 2011). We
284 applied the mixing model to the twelve flood events and compared its performance with the

285 performance of the PLS model. For each flood event, we calculated the relative contributions of [*E.*
286 *coli*]_{OF-TOT} and of [*E. coli*]_{GW-TOT} to [*E. coli*]_{mean} based on Eq. 4, namely [*E. coli*]_{OF-TOT}% and of [*E. coli*]_{GW-TOT}%
287 (in %).

288 3. Results

289 3.1. *E. coli* dynamics during flood events and among flood events

290 The areal percentage of land use evolved from 2011 to 2015 (Fig. SI1): annual crop decreased from 28%
291 to 4% of the catchment area, and teak and annual crop decreased from 12% to 0%, while teak increased
292 from 18% to 36%, and fallow increased from 29% to 46%. Areal percentage of forest was constant at 8%.
293 The *E. coli* input into the catchment decreased from a monthly average of 2×10^{15} in 2011 to a monthly
294 average of 3×10^{14} in 2015 (Fig. SI2). Over the 2011-2015 period, the cumulated *E. coli* input for swine and
295 poultry accounted for 0.01-0.38 % of the total *E. coli* input.

296 In the meantime, a total of 294 discharge peaks were recorded between January 1, 2011, and December
297 31, 2015 (Fig. 2), in response to a range of rainfall events (Fig. SI3, see interquartile ranges): 50 % of R_{TOT}
298 ranged between 7 and 22 mm, 50 % of R_{MAX} ranged between 24 and 60 mm h^{-1} , and 50 % of R_D ranged
299 between 49 and 235 min. Q ranged from 0 to 1 807.6 L s^{-1} ($30 \text{ L s}^{-1} \text{ ha}^{-1}$), with a mean Q of 33.3 L s^{-1} .

300 We monitored twelve flood events (A to L) for stream water quality. Datasets for each of the twelve
301 flood events are complete except EC measurements lacking for event A and Turbidity measurements
302 lacking for event G, due to measurement devices' breakdown. A total of 99 stream water samples was
303 collected: [TSS] ranged from 0.02 to 25.7 g L^{-1} while Turbidity ranged from 191.8 to 13 480 NTU (Fig. 3).
304 *E. coli* were detected in all samples and [*E. coli*] ranged between 160 and 74 000 MPN 100 mL^{-1} (Fig. 3).

305 For the twelve flood events, rainfall was followed by an increase of Q , Turbidity, [TSS] and [*E. coli*] (Fig.
306 3). However, the peaks of Q , Turbidity, [TSS] and [*E. coli*] were often asynchronous (full description is
307 given in the supplementary information). Hysteresis loops showed complex patterns mixing straight
308 lines, simple clockwise or anti-clockwise loops, figures-of-eight or multiple hysteresis loops (summary is
309 presented in Table SI1, full description is given in the supplementary information, including Figs. SI4-SI7).

310 The uncertainty bands enclosing Q_{OF} and Q_{GW} in Fig. 3 show that the hydrograph could be clearly
311 separated for each flood event. Flood event C was the only event where the stream flow was virtually
312 100% groundwater flow. Excluding flood event C, $Q_{OF-TOT}\%$ ranged between 17 and 80% (Fig. 4a). The
313 average $Q_{OF-TOT}\%$ was 41% (CV=0.59).

314 The twelve flood events corresponded to a range of storm characteristics in terms of R_{TOT} , $R_{I_{MAX}}$ and R_D
315 (Fig. SI3, Fig. 2). Q_{MAX} ranged from 31.7 to 967.9 $L s^{-1}$ (Table 1, Fig. 3). Three flood events occurred after
316 dry periods (events B, G, and I: API<10), whereas others occurred shortly after a previous flood event
317 (events D, E, and J: API>30) or later (events A, C, F, H, K, and L: 10<API<20) (Table 1). $[E. coli]_{mean}$ ranged
318 from 1 125 and 27 375 MPN 100 mL^{-1} .

319 Mean EC_{OF} was 55 $\mu S cm^{-1}$ (CV=0.7) over the 2011-2015 period (65 surface runoff water samples). Mean
320 EC_{GW} was 256 $\mu S cm^{-1}$ (CV=0.42) among the twelve flood events. $[E. coli]_{OF-TOT}$ was 24 880, 27 292, 27 441,
321 26 987, and 26 935 MPN 100 mL^{-1} in 2011, 2012, 2013, 2014, and 2015, respectively (26 surface runoff
322 water samples). $[E. coli]_{GW-TOT}$ was 277 MPN 100 mL^{-1} (6 groundwater samples).

323 3.2. Statistical analysis

324 3.2.1. Modelling instantaneous *E. coli* concentration [*E. coli*]

325 Considering all water samples collected during the flood events (Fig. SI8), [*E. coli*] was positively
326 correlated to Q ($r=0.54$, $p<0.001$), Q_{OF} ($r=0.56$, $p<0.001$), $Q_{OF}\%$ ($r=0.51$, $p<0.001$), [TSS] ($r=0.43$, $p<0.001$),
327 and Turbidity ($r=0.43$, $p<0.001$), and was negatively correlated to EC ($r=-0.31$, $p=0.003$), and to $Q_{GW}\%$ ($r=-$
328 0.51 , $p<0.001$). Turbidity was positively correlated to [TSS] ($r=0.82$, $p<0.001$).

329 From the stepwise regression, [*E. coli*] was best predicted by [TSS] and Q_{OF} (Eq. 5):

$$[E. coli] = 4995 + 689 \cdot [TSS] + 41.49 \cdot Q \quad \text{Eq. 5}$$

330 T-values and p-values of the coefficients are given in Table SI2. The R^2 of the model is 35.3 % and the
331 R^2_{pred} is 24.42 %.

332 3.2.2. Modelling mean concentration of *E. coli* per flood event $[E. coli]_{\text{mean}}$

333 $[E. coli]_{\text{mean}}$ was positively correlated to Q_{MAX} ($r=0.71$, $p=0.009$), Q_{TOT} ($r=0.59$, $p=0.042$), $Q_{\text{OF-TOT}}$ ($r=0.72$,
334 $p=0.012$), and $Q_{\text{OF-TOT}}\%$ ($r=0.93$, $p<0.001$), and was negatively correlated to $Q_{\text{GW-TOT}}\%$ ($r=-0.93$, $p<0.001$)
335 (Fig. SI9).

336 From the above, we found that $Q_{\text{OF-TOT}}\%$ was the variable the most positively correlated to $[E. coli]_{\text{mean}}$.

337 The linear model is:

$$[E. coli]_{\text{mean}} = 332 \cdot Q_{\text{OF-TOT}}\% - 3749 \quad \text{Eq. 6}$$

338 The R^2 between predicted and observed $[E. coli]_{\text{mean}}$ was 0.87 while the Nash-Sutcliffe efficiency (NSE)
339 was 0.84 (Fig. 5a). Excluding event C that was driven by groundwater flow only (and for which the linear
340 model predicted a negative value), $[E. coli]_{\text{mean}}$ predicted by the linear model ranged from 1981 to 22
341 825 MPN 100 mL⁻¹.

342 These latter trends are confirmed by the PLS regression (Fig. 6a), which shows that Axis 1 mostly explains
343 the variables $Q_{\text{OF-TOT}}\%$ ($r=0.863$), $Q_{\text{GW-TOT}}\%$ ($r=-0.863$), Q_{MAX} ($r=0.859$), and $Q_{\text{OF-TOT}}$ ($r=0.815$) whereas Axis 2
344 mostly explains the variable RI_{MAX} ($r=0.563$). Hence, Axis 1 corresponds to variables that are strongly
345 related to stream water whereas Axis 2 corresponds to variables describing rainfall and soil moisture
346 conditions. Accordingly, the twelve flood events were scattered along the two axes, with the flood
347 events of highest $[E. coli]_{\text{mean}}$ in the right panel (Fig. 6b). The variables with VIP values above 1 were $Q_{\text{OF-}}$
348 $\text{TOT}\%$, $Q_{\text{GW-TOT}}\%$, Q_{MAX} , $Q_{\text{OF-TOT}}$, and Q_{TOT} when considering one component and $Q_{\text{OF-TOT}}\%$, $Q_{\text{GW-TOT}}\%$,
349 $[\text{TSS}]_{\text{mean}}$, Q_{MAX} , and $Q_{\text{OF-TOT}}$ when considering two components (Fig. 6c; Table SI3). The statistical model
350 given by the PLS is given in Table SI3. The R^2 between predicted and observed $[E. coli]_{\text{mean}}$ was 0.83 while
351 the NSE was 0.78 (Fig. 5b). Excluding event C that was driven by groundwater flow only (and for which

352 the PLS model predicted a negative value), $[E. coli]_{\text{mean}}$ predicted by the PLS model ranged from 1 067 to
353 23 237 MPN 100 mL⁻¹.

354 $[E. coli]_{\text{mean}}$ predicted by the mixing model ranged from 277 to 21 572 MPN 100 mL⁻¹. The R² between
355 predicted and observed $[E. coli]_{\text{mean}}$ was 0.86 while the NSE was 0.65 (Fig. 5c). Flood event C was the only
356 event where the percentage of *E. coli* was 100% from groundwater flow. Excluding flood event C, $[E.$
357 $coli]_{\text{OF-TOT}}\%$ ranged from 95 to virtually 100% (Fig. 4b). The average $[E. coli]_{\text{OF-TOT}}\%$ was 89% (CV=0.33).

358

359

360 4. Discussion

361 4.1. High *E. coli* concentration pulses match high overland flow pulses

362 The analysis of the twelve monitored flood events (Table 1) shows that $[E. coli]_{\text{mean}}$ was high regardless of
363 the flood event characteristics: $[E. coli]_{\text{mean}}$ ranged between 1 125 and 27 375 MPN 100 mL⁻¹. The order of
364 magnitude of the maximum $[E. coli]$ is about 10⁵ MPN 100 mL⁻¹. All these values exceed the 1 000
365 MPN 100 mL⁻¹ threshold provided by e.g. the European Directive 2006/7/EC for bathing water quality.

366 Although rainfall is known to reactivate hydrological connectivity (Bracken et al., 2013), rainfall
367 characteristics such as R_{TOT} , RI_{MAX} , and R_{D} , did not appear as strong explanatory variables of $[E. coli]_{\text{mean}}$
368 (Fig. 6 and Fig. SI9). Variations in $[E. coli]$ and $[E. coli]_{\text{mean}}$ between flood events may be explained by
369 rainfall spatial distribution and the timing between rainfall events: the rainfall distribution will determine
370 which areas draining animal and human manure are activated, whereas the timing between events may
371 indicate which catchment microbial stocks have accumulated to high levels (McKergow and Davies-
372 Colley, 2010). In Houay Pano catchment, variations in concentration between flood events may also be
373 explained by *E. coli* input distribution at the soil surface, as the land use is mixed (Fig. SI1). Indeed, based
374 on this twelve-event dataset and hysteresis analysis, we could not find any clear, consistent trend at the
375 outlet of the Houay Pano catchment (Table SI1) and we thus could not identify any primary bacteria
376 source.

377 The land use change in Houay Pano catchment was rapid and led to the simultaneous increase of teak
378 tree plantations and of fallow areas (Riboldi et al., 2017). Previous plot- and catchment-scale results
379 showed higher surface runoff and suspended matter export when teak tree plantations increased
380 (Lacombe et al., 2018; Mügler et al., 2019; Riboldi et al., 2017; Song et al., 2020), suggesting that higher
381 numbers of *E. coli* could be transferred to the river. However, land use characteristics such as land use
382 variables (Annual crop, Teak+Crop, Teak, Fallow, and *E. coli* input) did not appear as strong explanatory

383 variables of $[E. coli]_{\text{mean}}$ (Fig. 6 and Fig. SI10). Growing teak trees and fallow requires less people in the
384 field compared to annual crops. Since field workers practice open defecation, less workers imply lower
385 input of *E. coli* into the catchment. The contribution of domestic animals to the overall *E. coli* input being
386 only 0.01-0.38 % of the total *E. coli* input, the consequence of the extension of teak tree plantation and
387 of fallow in Houay Pano from 2011 to 2015 (Fig. SI1) is the overall decrease of the *E. coli* input into the
388 catchment over the same period (Fig. SI2).

389 The highest $[E. coli]_{\text{mean}}$ was often associated to flood events with highest Q_{MAX} (events J and L) and
390 dominant surface runoff ($Q_{\text{OF-TOT}}\%>50$: events D, and I-L), confirming the role of surface runoff, and
391 subsequent soil surface erosion, in *E. coli* transfers to the stream (Causse et al., 2015) (Table 1, Fig. 6).
392 For flood events where sub-surface flow was dominating ($Q_{\text{OF-TOT}}\%<50$: events B, C, and E-H), $[E. coli]_{\text{mean}}$
393 was in general lower, although exceeding 1 000 MPN 100 mL⁻¹. High *E. coli* concentrations even when
394 sub-surface flow was dominating suggest that streambanks and the streambed may release stored *E. coli*
395 (Chu et al., 2011; Park et al., 2017; Stocker et al., 2018) since sediment deposited at the bottom of the
396 stream may act as an *E. coli* reservoir (Pachepsky et al., 2017; Rochelle-Newall et al., 2015; Smith et al.,
397 2008). In fact, the concentration of *E. coli* in Houay Pano streambed sediment is about 40 000 MPN g⁻¹
398 (Ribolzi et al., 2016a), but we cannot exclude that *E. coli* is simultaneously transferred from the hillslope
399 with surface runoff, if hillslope soil surface is highly contaminated. We also found both clockwise and
400 anti-clockwise $[E. coli]$ -[TSS] and $[E. coli]$ -Turbidity hysteresis loops, and in some cases figure-of-eight
401 patterns (Table SI1, Fig. SI6 and Fig. SI7). This further suggests that the erosion of bacteria stores is not
402 strictly driven by soil erosion, but also by in-stream sediment resuspension (Evrard et al., 2016; Gourdin
403 et al., 2015; Huon et al., 2017). In parallel, possibly less *E. coli*-contaminated sub-surface flow, or return
404 flow, may dilute in the stream the *E. coli* concentration originating from surface runoff and thus mitigate
405 the in-stream microbial contamination.

406 4.2. Predictive models of *E. coli* concentration: performances and usefulness

407 To our knowledge, few studies have reported predictive models using independent variables to explain
408 *E. coli* concentration and, to the best of our knowledge, they were all developed in temperate areas
409 (Chen and Chang, 2014; Hathaway et al., 2010). The models suggested by Chen and Chang (2014)
410 included antecedent precipitation, stream temperature, and TSS concentration, and gave an R^2 of 0.27-
411 0.61 depending on the season and the catchment. The model suggested by Hathaway et al. (2010)
412 included temperature, rainfall, and humidity, and had an R^2 of 0.7462. The variety of explanatory
413 variables among models reflects the variety of driving processes driving *E. coli* fate and transport among
414 catchments.

415 4.2.1. Modelling instantaneous *E. coli* concentration [*E. coli*]

416 [*E. coli*] was equally correlated to both [TSS] and Turbidity ($r=0.43$, $p<0.001$; Fig. SI8) while Turbidity was
417 correlated with [TSS] ($r=0.82$, $p<0.001$, Fig. SI8). The strong correlation between Turbidity and [TSS] is in
418 line with the results obtained in other studies in tropical areas, such as Ziegler et al. (2014) for rivers in
419 Thailand and Robert et al. (2017) in West African lakes and ponds. The relationships between [TSS] and
420 [*E. coli*] and between Turbidity and [*E. coli*] are qualified by hysteretic trends (Figs. SI6 and SI7,
421 respectively) during flood events. Notably, little hysteresis is observed between [TSS] and Turbidity (Fig.
422 SI6). Hysteretic trends may illustrate the uncertainty when predicting *E. coli* by a proxy such as turbidity:
423 sources of sediments may not exactly coincide with sources of *E. coli*. For example, these relationships
424 may not apply in large mixed-land use catchments, because of the multiple, geographically separated, *E.*
425 *coli* and turbidity sources. This may limit the usefulness of using turbidity as a proxy of *E. coli* near the
426 catchment outlet (McKergow and Davies-Colley, 2010). Another explanation of the hysteresis is the
427 nature and the properties of the suspended sediments, that may change during the flood event: bacteria
428 may be less prone to attach to sand-rich suspended particles (Oliver et al., 2007). Similarly, hysteresis in
429 Turbidity-[TSS] relationships may be explained by changing reflectance properties of the suspended

430 sediments depending on particle size, shape, mineralogy, aggregation/flocculation, dissolved light-
431 absorbing matter and bubbles, because of the spatial heterogeneity of rainfall and suspended sediment
432 sources at the catchment scale (Navratil et al., 2011; Ziegler et al., 2014). A last uncertainty source in the
433 relationships between [*E. coli*], [TSS], and Turbidity, is the possible exhaustion of the *E. coli* stock within
434 the catchment after a succession of flood events.

435 From the step-wise regression analysis, the predictive model for [*E. coli*] includes [TSS] and Q_{OF} as
436 explanatory variables (Eq. 5). The two explanatory variables reflect the two processes driving *E. coli*
437 transport within the catchment: Q_{OF} reflects the mobilization of *E. coli* with relative contributions of
438 surface and sub-surface flows along the flood event, whereas [TSS] reflects the attachment of *E. coli* to
439 soil particles and/or streambed resuspended sediments (Garcia-Armisen and Servais, 2009; Nguyen et
440 al., 2016). Overall, higher *E. coli* concentration is related to higher surface runoff and higher suspended
441 matter in the stream flow. The value of the intercept term in Eq. 5 corresponds to *E. coli* concentration
442 during base flow (Boithias et al., 2016; Kim et al., 2018), i.e. when [TSS] and Q_{OF} tend to zero because
443 there is no surface runoff during inter-storm periods, in other words when groundwater is the only
444 supply to stream flow and EC tends to EC_{GW} .

445 The two variables best predicting [*E. coli*], namely [TSS] and Q_{OF} , can be assessed with proxies. [TSS]
446 could be interchanged with Turbidity, since [TSS] and Turbidity are correlated, as discussed above.
447 Similarly, EC appears strongly related to Q_{OF} (and Q) by a hyperbolic function (Fig. SI8). Using EC as a
448 proxy of Q has already been proposed in Alpine headwaters (Cano-Paoli et al., 2019). Both of these
449 proxies can be monitored *in situ* at high frequencies.

450 4.2.2. Modelling mean concentration of *E. coli* per flood event [*E. coli*]_{mean}

451 The PLS regression analysis (Table SI3) confirmed the analysis of the correlation matrix (Fig. SI9): $Q_{OF-TOT}\%$
452 and $Q_{GW-TOT}\%$ were the best predictors of [*E. coli*]_{mean} ($r= 0.93$ and -0.93 , respectively). Land uses that

453 favor higher $Q_{OF-TOT}\%$ or lower $Q_{GW-TOT}\%$, such as erosion-prone teak trees plantations (Lacombe et al.,
454 2016; Ribolzi et al., 2017), or more intense rainfall events, will lead to higher in-stream concentrations of
455 *E. coli* (Fig. 6a and Fig. SI10) and thus higher fecal contamination risk. Hence, a model based on $Q_{OF-TOT}\%$
456 or $Q_{GW-TOT}\%$ appears relevant to predict $[E. coli]_{mean}$. Such a basic tool may help local and national
457 stakeholders to assess the fecal contamination risk by testing global change scenarios in target
458 catchments.

459 In this study, the mixing model of $[E. coli]_{mean}$ based on $Q_{OF-TOT}\%$ and $Q_{GW-TOT}\%$ had a R^2 of 0.86 and a NSE
460 of 0.65. The NSE value is lower than the NSE values' range reported by Chin (2011) in six nested
461 catchments in Georgia, United-States. However, in our study the values of both overland flow and
462 groundwater end-members were measured in the field, conversely to Chin (2011) who numerically
463 optimized the values of the two end-members. The performance of the mixing model is lower in terms of
464 NSE than those of the PLS and the linear models (NSE = 0.78 and 0.84, respectively). However, the PLS
465 model is difficult to implement because of the large number of variables it implies, and both the linear
466 and the PLS models are constrained by their validity domains.

467 Within a catchment, the sources of *E. coli* are the *E. coli* transferred from hillslopes with surface runoff,
468 as discussed in section 4.1, and the resuspension from the streambed, whereas the sinks of *E. coli* are the
469 deposition of the bacterium on the streambed and its decay in the water column. Since rainfall events in
470 the Houay Pano catchment last 0.5-5 hours, considering a decay rate of 0.3-0.6 d^{-1} (Nguyen et al., 2016),
471 the removal of *E. coli* in the water column from bacteria population decay is negligible during a flood
472 event. Since the mixing model is able to accurately predict the flood-event mean concentration of *E. coli*
473 at the catchment outlet, the deposition of *E. coli* on the streambed is either negligible or compensated
474 by the resuspension. The percentage of *E. coli* resuspended from streambed during a flood is about 11%
475 (Ribolzi et al., 2016a), which implies that the compensated *E. coli* deposition, e.g. in small wetlands along
476 the main stream, is about 11% as well. Finally, since the possible sources and sinks of *E. coli* are negligible

477 compared to the bacterium transport processes during flood events, the approximation of the simple
478 model that only considers two end-members, namely the *E. coli* concentration in both surface runoff and
479 groundwater, is acceptable.

480 Hence, we could calculate the relative contributions of surface runoff and of sub-surface flow to the in-
481 stream *E. coli* contamination. Excluding flood event C that was driven by groundwater flow, more than
482 95% of $[E. coli]_{\text{mean}}$ resulted from overland flow, even for flood events where the percentage of overland
483 flow was small (<20%, events E and G). The average contribution of overland flow to the flood-event
484 mean concentration of *E. coli* was 89%, while the average contribution of overland flow to the flood-
485 event mean stream flow was 41% (Fig. 4). In other words, the contribution of groundwater to the in-
486 stream *E. coli* contamination was low during flood events, even though the contribution of groundwater
487 to stream water flow was greater than that of overland flow, i.e. when base flow index (BFI) was over
488 0.5. In comparison, with similar BFI values, Chin (2011) reported bacteria loads from surface runoff in the
489 range of 80–90%, with the remainder mostly originating from base flow, and negligible background loads
490 from the catchment.

491

492 5. Conclusion

493 To our knowledge, few papers have reported the statistical relationships between *E. coli* and its
494 environmental drivers (meteorological, hydrological, and land use variables) and the present study is the
495 first one to investigate *E. coli* dynamics during flood events in a tropical humid catchment. We conclude
496 that:

- 497 - Regardless of the contribution of sub-surface flow to the flood, in-stream *E. coli* concentration is
498 high, suggesting that the streambed *E. coli* store is high. However, highest concentrations of *E.*
499 *coli* are measured for the flood events that are driven by surface runoff, suggesting that surface
500 runoff, and consequently soil erosion, are the primary drivers of in-stream *E. coli* contamination;
- 501 - The three predictive models (linear model, partial least square model, and mixing model) show
502 that the percentage of surface runoff in stream flow is the best predictor of the flood event
503 mean concentration of *E. coli*;
- 504 - A simple mixing model based on the relative contributions of both overland flow and
505 groundwater, and on the *E. coli* concentration in both overland flow and groundwater, is reliable
506 to predict in-stream flood-event mean *E. coli* concentration (NSE=0.65). On average, 89% of the
507 in-stream concentration of *E. coli* is supplied by surface runoff, while the overall contribution of
508 surface runoff to the stream flow is 41%.
- 509 - Stream flow turbidity and *E. coli* concentration are positively correlated, but turbidity is not a
510 strong predictor of *E. coli* concentration during flood events;

511 Predictive models based on turbidity and electrical conductivity may be used to provide real-time
512 estimates of in-stream *E. coli* concentration, whereas a mixing model may provide flood-event mean
513 information. Such simple models may help to assess the impact of global change on in-stream *E. coli*
514 contamination. They could thus be used to assess the risk of water borne diseases such as diarrhea, in
515 rural areas where mammals, including humans, practice open defecation, and to design early warning

516 systems. Applied over long periods of time, the models might be used to calculate *E. coli* input-output
517 load balances. Such basic modelling studies would thus help to assess the long-term impact of land-use
518 change on the microbial quality of surface water.

519 Future work is required to better understand the pathways of FIB at catchment scale. For example, there
520 is a need to characterize the interactions between ground and surface waters, and the role of the
521 hyporheic zone, as sources of contaminant. Furthermore, the partition between free and particle-
522 attached *E. coli* should be quantified to better understand and predict the probability that bacteria is
523 deposited on the streambed during the flood recession stage and re-suspended in the water column
524 during the next flood rising stage. These findings are of primary importance to build adequate
525 catchment-scale models to accurately simulate *E. coli* fate and transport, and thus better assess fecal
526 contamination risk in a global changing context.

527

528 Acknowledgments

529 The authors sincerely thank the Lao Department of Agricultural Land Management (DALaM) for its
530 support, including granting the permission for field access, and the M-TROPICS Critical Zone Observatory
531 (<https://mtropics.obs-mip.fr/>), which belongs to the French Research Infrastructure OZCAR
532 (<http://www.ozcar-ri.org/>), for data access. This study was funded by the GIS-Climat
533 (<http://www.gisclimat.fr/projet/pastek.html>) and by the French National Research Agency (TecltEasy
534 project; ANR-13-AGRO-0007). LB thanks the Research Institute for Development (IRD) for her
535 postdoctoral grant 2016-2017 (*Accueil de post-doctorants / Campagne 2015*; www.ird.fr).

536

537

538 **References**

- 539 Boithias, L., Auda, Y., Audry, S., et al., 2020. Dataset from the Multiscale TROPICAL CatchmentS critical
540 zone observatory M-TROPICS II: land use, hydrology and sediment production monitoring in
541 Houay Pano, northern Lao PDR. *Hydrol. Process.* Submitted.
- 542 Boithias, L., Choisy, M., Souliyaseng, N., Jourden, M., Quet, F., Buisson, Y., Thammahacksa, C., Silvera,
543 N., Latsachack, K., Sengtaheuanghoung, O., Pierret, A., Rochelle-Newall, E., Becerra, S., Ribolzi,
544 O., 2016. Hydrological regime and water shortage as drivers of the seasonal incidence of
545 diarrheal diseases in a tropical montane environment. *PLoS Negl. Trop. Dis.* 10, e0005195.
546 <https://doi.org/10.1371/journal.pntd.0005195>
- 547 Bracken, L.J., Wainwright, J., Ali, G.A., Tetzlaff, D., Smith, M.W., Reaney, S.M., Roy, A.G., 2013. Concepts
548 of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Sci. Rev.*
549 119, 17–34. <https://doi.org/10.1016/j.earscirev.2013.02.001>
- 550 Cano-Paoli, K., Chiogna, G., Bellin, A., 2019. Convenient use of electrical conductivity measurements to
551 investigate hydrological processes in Alpine headwaters. *Sci. Total Environ.* 685, 37–49.
552 <https://doi.org/10.1016/j.scitotenv.2019.05.166>
- 553 Causse, J., Billen, G., Garnier, J., Henri-des-Tureaux, T., Olasa, X., Thammahacksa, C., Latsachak, K.O.,
554 Soulileuth, B., Sengtaheuanghoung, O., Rochelle-Newall, E., Ribolzi, O., 2015. Field and modelling
555 studies of *Escherichia coli* loads in tropical streams of montane agro-ecosystems. *J. Hydro-*
556 *Environ. Res.* 9, 496–507. <https://doi.org/10.1016/j.jher.2015.03.003>
- 557 Chen, H.J., Chang, H., 2014. Response of discharge, TSS, and *E. coli* to rainfall events in urban, suburban,
558 and rural watersheds. *Env. Sci Process. Impacts* 16, 2313–2324.
559 <https://doi.org/10.1039/C4EM00327F>
- 560 Chin, D.A., 2011. Quantifying pathogen sources in streams by hydrograph separation. *J. Environ. Eng.*
561 137, 770–781.
- 562 Cho, K.H., Han, D., Park, Y., Lee, S.W., Cha, S.M., Kang, J.-H., Kim, J.H., 2010. Evaluation of the
563 relationship between two different methods for enumeration fecal indicator bacteria: Colony-
564 forming unit and most probable number. *J. Environ. Sci.* 22, 846–850.
565 [https://doi.org/10.1016/S1001-0742\(09\)60187-X](https://doi.org/10.1016/S1001-0742(09)60187-X)
- 566 Cho, K.H., Pachepsky, Y.A., Oliver, D.M., Muirhead, R.W., Park, Y., Quilliam, R.S., Shelton, D.R., 2016.
567 Modeling fate and transport of fecally-derived microorganisms at the watershed scale: State of
568 the science and future opportunities. *Water Res.* 100, 38–56.
569 <https://doi.org/10.1016/j.watres.2016.04.064>
- 570 Chu, Y., Salles, C., Tournoud, M.-G., Got, P., Troussellier, M., Rodier, C., Caro, A., 2011. Faecal bacterial
571 loads during flood events in Northwestern Mediterranean coastal rivers. *J. Hydrol.* 405, 501–511.
572 <https://doi.org/10.1016/j.jhydrol.2011.05.047>
- 573 Chuah, C.J., Ziegler, A.D., 2018. Temporal Variability of Faecal Contamination from On-Site Sanitation
574 Systems in the Groundwater of Northern Thailand. *Environ. Manage.*
575 <https://doi.org/10.1007/s00267-018-1016-7>
- 576 Collins, R.U., Neal, C., 1998. The hydrochemical impacts of terraced agriculture, Nepal. *Sci. Total Environ.*
577 212. [https://doi.org/10.1016/S0048-9697\(97\)00342-2](https://doi.org/10.1016/S0048-9697(97)00342-2)
- 578 Descroix, L., Nouvelot, J.-F., Vauclin, M., 2002. Evaluation of an antecedent precipitation index to model
579 runoff yield in the western Sierra Madre (North-west Mexico). *J. Hydrol.* 263, 114–130.
580 [https://doi.org/10.1016/S0022-1694\(02\)00047-1](https://doi.org/10.1016/S0022-1694(02)00047-1)
- 581 Evrard, O., Laceby, J.P., Huon, S., Lefèvre, I., Sengtaheuanghoung, O., Ribolzi, O., 2016. Combining
582 multiple fallout radionuclides (¹³⁷Cs, ⁷Be, ²¹⁰Pb_{xs}) to investigate temporal sediment source
583 dynamics in tropical, ephemeral riverine systems. *J. Soils Sediments* 16, 1130–1144.
584 <https://doi.org/10.1007/s11368-015-1316-y>

585 Exley, J.L.R., Liseka, B., Cumming, O., Ensink, J.H.J., 2015. The Sanitation Ladder, What Constitutes an
586 Improved Form of Sanitation? *Environ. Sci. Technol.* 49, 1086–1094.
587 <https://doi.org/10.1021/es503945x>

588 Gaillardet, J., Braud, I., Hankard, F., Anquetin, S., Bour, O., Dorfliger, N., de Dreuzy, J.R., Galle, S., Galy, C.,
589 Gogo, S., Gourcy, L., Habets, F., Laggoun, F., Longuevergne, L., Le Borgne, T., Naaïm-Bouvet, F.,
590 Nord, G., Simonneaux, V., Six, D., Tallec, T., Valentin, C., Abril, G., Allemand, P., Arènes, A., Arfib,
591 B., Arnaud, L., Arnaud, N., Arnaud, P., Audry, S., Comte, V.B., Batiot, C., Battais, A., Bellot, H.,
592 Bernard, E., Bertrand, C., Bessière, H., Binet, S., Bodin, J., Bodin, X., Boithias, L., Bouchez, J.,
593 Boudevillain, B., Moussa, I.B., Branger, F., Braun, J.J., Brunet, P., Caceres, B., Calmels, D.,
594 Cappelaere, B., Celle-Jeanton, H., Chabaux, F., Chalikakis, K., Champollion, C., Copard, Y., Cotel,
595 C., Davy, P., Deline, P., Delrieu, G., Demarty, J., Dessert, C., Dumont, M., Emblanch, C., Ezzahar, J.,
596 Estèves, M., Favier, V., Faucheux, M., Filizola, N., Flammarion, P., Floury, P., Fovet, O., Fournier,
597 M., Francez, A.J., Gandois, L., Gascuel, C., Gayer, E., Genthon, C., Gérard, M.F., Gilbert, D.,
598 Gouttevin, I., Grippa, M., Gruau, G., Jardani, A., Jeanneau, L., Join, J.L., Jourde, H., Karbou, F.,
599 Labat, D., Lagadeuc, Y., Lajeunesse, E., Lastennet, R., Lavado, W., Lawin, E., Lebel, T., Le
600 Bouteiller, C., Legout, C., Lejeune, Y., Le Meur, E., Le Moigne, N., Lions, J., Lucas, A., Malet, J.P.,
601 Marais-Sicre, C., Maréchal, J.C., Marlin, C., Martin, P., Martins, J., Martinez, J.M., Massei, N.,
602 Mauclerc, A., Mazzilli, N., Molénat, J., Moreira-Turcq, P., Mougin, E., Morin, S., Ngoupayou, J.N.,
603 Panthou, G., Peugeot, C., Picard, G., Pierret, M.C., Porel, G., Probst, A., Probst, J.L., Rabatel, A.,
604 Raclot, D., Ravanel, L., Rejiba, F., René, P., Ribolzi, O., Riotte, J., Rivière, A., Robain, H., Ruiz, L.,
605 Sanchez-Perez, J.M., Santini, W., Sauvage, S., Schoeneich, P., Seidel, J.L., Sekhar, M.,
606 Sengtaeuanghoung, O., Silvera, N., Steinmann, M., Soruco, A., Tallec, G., Thibert, E., Lao, D.V.,
607 Vincent, C., Viville, D., Wagnon, P., Zitouna, R., 2018. OZCAR: The French Network of Critical Zone
608 Observatories. *Vadose Zone J.* 17, 180067. <https://doi.org/10.2136/vzj2018.04.0067>

609 Garcia-Armisen, T., Servais, P., 2009. Partitioning and Fate of Particle-Associated *E. coli* in River Waters.
610 *Water Environ. Res.* 81, 21–28. <https://doi.org/10.2175/106143008X304613>

611 Genereux, D., 1998. Quantifying uncertainty in tracer-based hydrograph separations. *Water Resour. Res.*
612 34, 915–919. <https://doi.org/10.1029/98WR00010>

613 Gourdin, E., Huon, S., Evrard, O., Ribolzi, O., Bariac, T., Sengtaeuanghoung, O., Ayrault, S., 2015. Sources
614 and export of particle-borne organic matter during a monsoon flood in a catchment of northern
615 Laos. *Biogeosciences* 12, 1073–1089. <https://doi.org/10.5194/bg-12-1073-2015>

616 Hathaway, J.M., Hunt, W.F., Simmons, O.D., 2010. Statistical Evaluation of Factors Affecting Indicator
617 Bacteria in Urban Storm-Water Runoff. *J. Environ. Eng.* 136, 1360–1368.
618 [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000278](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000278)

619 Helsel, D.R., Hirsch, R.M., 2002. Chapter A3: Statistical Methods in Water Resources, in: *Techniques of*
620 *Water-Resources Investigations of the United States Geological Survey*. Book 4, Hydrologic
621 *Analysis and Interpretation*.

622 Holz, G.K., 2010. Sources and processes of contaminant loss from an intensively grazed catchment
623 inferred from patterns in discharge and concentration of thirteen analytes using high intensity
624 sampling. *J. Hydrol.* 383, 194–208. <https://doi.org/10.1016/j.jhydrol.2009.12.036>

625 Huon, S., Evrard, O., Gourdin, E., Lefèvre, I., Bariac, T., Reyss, J.-L., Henry des Tureaux, T.,
626 Sengtaeuanghoung, O., Ayrault, S., Ribolzi, O., 2017. Suspended sediment source and
627 propagation during monsoon events across nested sub-catchments with contrasted land uses in
628 Laos. *J. Hydrol. Reg. Stud.* 9, 69–84. <https://doi.org/10.1016/j.ejrh.2016.11.018>

629 Kim, M., Boithias, L., Cho, K.H., Sengtaeuanghoung, O., Ribolzi, O., 2018. Modeling the Impact of Land
630 Use Change on Basin-scale Transfer of Fecal Indicator Bacteria: SWAT Model Performance. *J.*
631 *Environ. Qual.* 47, 1115–1122. <https://doi.org/10.2134/jeq2017.11.0456>

632 Kim, M., Boithias, L., Cho, K.H., Silvera, N., Thammahacksa, C., Latsachack, K., Rochelle-Newall, E.,
633 Sengtaeuanghoung, O., Pierret, A., Pachepsky, Y.A., Ribolzi, O., 2017. Hydrological modeling of
634 Fecal Indicator Bacteria in a tropical mountain catchment. *Water Res.* 119, 102–113.
635 <https://doi.org/10.1016/j.watres.2017.04.038>

636 Krometis, L.-A.H., Characklis, G.W., Simmons, O.D., Dilts, M.J., Likirdopulos, C.A., Sobsey, M.D., 2007.
637 Intra-storm variability in microbial partitioning and microbial loading rates. *Water Res.* 41, 506–
638 516. <https://doi.org/10.1016/j.watres.2006.09.029>

639 Lacombe, G., Ribolzi, O., de Rouw, A., Pierret, A., Latsachak, K., Silvera, N., Pham Dinh, R., Orange, D.,
640 Janeau, J.-L., Soullieuth, B., Robain, H., Taccoen, A., Sengphaathith, P., Mouche, E.,
641 Sengtaeuanghoung, O., Tran Duc, T., Valentin, C., 2016. Contradictory hydrological impacts of
642 afforestation in the humid tropics evidenced by long-term field monitoring and simulation
643 modelling. *Hydrol. Earth Syst. Sci.* 20, 2691–2704. <https://doi.org/10.5194/hess-20-2691-2016>

644 Lacombe, G., Valentin, C., Sounyafong, P., de Rouw, A., Soullieuth, B., Silvera, N., Pierret, A.,
645 Sengtaeuanghoung, O., Ribolzi, O., 2018. Linking crop structure, throughfall, soil surface
646 conditions, runoff and soil detachment: 10 land uses analyzed in Northern Laos. *Sci. Total
647 Environ.* 616–617, 1330–1338. <https://doi.org/10.1016/j.scitotenv.2017.10.185>

648 Long, D.T., Voice, T.C., Xagaroraki, I., Chen, A., Wu, H., Lee, E., Oun, A., Xing, F., 2017. Patterns of c-q
649 hysteresis loops and within an integrative pollutograph for selected inorganic and organic solutes
650 and *E. coli* in an urban salted watershed during winter-early spring periods. *Appl. Geochem.* 83,
651 93–107. <https://doi.org/10.1016/j.apgeochem.2017.03.002>

652 Lušić, D.V., Jozić, S., Cenov, A., Glad, M., Bulić, M., Lušić, D., 2016. *Escherichia coli* in marine water:
653 Comparison of methods for the assessment of recreational bathing water samples. *Mar. Pollut.
654 Bull.* 113, 438–443. <https://doi.org/10.1016/j.marpolbul.2016.10.044>

655 Martinez, J.M., Guyot, J.L., Filizola, N., Sondag, F., 2009. Increase in suspended sediment discharge of the
656 Amazon River assessed by monitoring network and satellite data. *Catena* 79, 257–264.
657 <https://doi.org/10.1016/j.catena.2009.05.011>

658 McKergow, L.A., Davies-Colley, R.J., 2010. Stormflow dynamics and loads of *Escherichia coli* in a large
659 mixed land use catchment. *Hydrol. Process.* 24, 276–289. <https://doi.org/10.1002/hyp.7480>

660 Mügler, C., Ribolzi, O., Janeau, J.-L., Rochelle-Newall, E., Latsachack, K., Thammahacksa, C., Viguier, M.,
661 Jardé, E., Henri-Des-Tureaux, T., Sengtaeuanghoung, O., Valentin, C., 2019. Experimental and
662 modelling evidence of short-term effect of raindrop impact on hydraulic conductivity and
663 overland flow intensity. *J. Hydrol.* 570, 401–410. <https://doi.org/10.1016/j.jhydrol.2018.12.046>

664 Navratil, O., Esteves, M., Legout, C., Gratiot, N., Nemery, J., Willmore, S., Grangeon, T., 2011. Global
665 uncertainty analysis of suspended sediment monitoring using turbidimeter in a small
666 mountainous river catchment. *J. Hydrol.* 398, 246–259.
667 <https://doi.org/10.1016/j.jhydrol.2010.12.025>

668 Nguyen, H.T.M., Le, Q.T.P., Garnier, J., Janeau, J.-L., Rochelle-Newall, E., 2016. Seasonal variability of
669 faecal indicator bacteria numbers and die-off rates in the Red River basin, North Viet Nam. *Sci.
670 Rep.* 6, 21644. <https://doi.org/10.1038/srep21644>

671 Oliver, D.M., Clegg, C.D., Heathwaite, A.L., Haygarth, P.M., 2007. Preferential Attachment of *Escherichia*
672 *coli* to Different Particle Size Fractions of an Agricultural Grassland Soil. *Water, Air, Soil Pollut.*
673 185, 369–375. <https://doi.org/10.1007/s11270-007-9451-8>

674 Pachepsky, Y., Stocker, M., Saldaña, M.O., Shelton, D., 2017. Enrichment of stream water with fecal
675 indicator organisms during baseflow periods. *Environ. Monit. Assess.* 189, 51.
676 <https://doi.org/10.1007/s10661-016-5763-8>

677 Pachepsky, Y.A., Allende, A., Boithias, L., Cho, K., Jamieson, R., Hofstra, N., Molina, M., 2018. Microbial
678 Water Quality: Monitoring and Modeling. *J. Environ. Qual.* 47, 931–938.
679 <https://doi.org/10.2134/jeq2018.07.0277>

680 Park, Y., Pachepsky, Y., Hong, E.-M., Shelton, D., Coppock, C., 2017. Release from Streambed to Water
681 Column during Baseflow Periods: A Modeling Study. *J. Environ. Qual.* 46, 219–226.
682 <https://doi.org/10.2134/jeq2016.03.0114>

683 Patin, J., Mouche, E., Ribolzi, O., Chaplot, V., Sengtahevanghoun, O., Latsachak, K.O., Soullileuth, B.,
684 Valentin, C., 2012. Analysis of runoff production at the plot scale during a long-term survey of a
685 small agricultural catchment in Lao PDR. *J. Hydrol.* 426–427, 79–92.
686 <https://doi.org/10.1016/j.jhydrol.2012.01.015>

687 Patin, J., Mouche, E., Ribolzi, O., Sengtahevanghoun, O., Latsachak, K.O., Soullileuth, B., Chaplot, V.,
688 Valentin, C., 2018. Effect of land use on interrill erosion in a montane catchment of Northern
689 Laos: An analysis based on a pluri-annual runoff and soil loss database. *J. Hydrol.* 563, 480–494.
690 <https://doi.org/10.1016/j.jhydrol.2018.05.044>

691 Ribolzi, O., Evrard, O., Huon, S., de Rouw, A., Silvera, N., Latsachack, K.O., Soullileuth, B., Lefèvre, I.,
692 Pierret, A., Lacombe, G., Sengtahevanghoun, O., Valentin, C., 2017. From shifting cultivation to
693 teak plantation: effect on overland flow and sediment yield in a montane tropical catchment. *Sci.*
694 *Rep.* 7, 3987. <https://doi.org/10.1038/s41598-017-04385-2>

695 Ribolzi, O., Evrard, O., Huon, S., Rochelle-Newall, E., Henri-des-Tureaux, T., Silvera, N., Thammahacksac,
696 C., Sengtahevanghoun, O., 2016a. Use of fallout radionuclides (⁷Be, ²¹⁰Pb) to estimate
697 resuspension of *Escherichia coli* from streambed sediments during floods in a tropical montane
698 catchment. *Environ. Sci. Pollut. Res.* 23, 3427–3435. <https://doi.org/10.1007/s11356-015-5595-z>

699 Ribolzi, O., Lacombe, G., Pierret, A., Robain, H., Sounyafong, P., de Rouw, A., Soullileuth, B., Mouche, E.,
700 Huon, S., Silvera, N., Latsachak, K.O., Sengtahevanghoun, O., Valentin, C., 2018. Interacting land
701 use and soil surface dynamics control groundwater outflow in a montane catchment of the lower
702 Mekong basin. *Agric. Ecosyst. Environ.* 268, 90–102. <https://doi.org/10.1016/j.agee.2018.09.005>

703 Ribolzi, O., Patin, J., Bresson, L.M., Latsachack, K.O., Mouche, E., Sengtahevanghoun, O., Silvera, N.,
704 Thiébaux, J.P., Valentin, C., 2011. Impact of slope gradient on soil surface features and
705 infiltration on steep slopes in northern Laos. *Geomorphology* 127, 53–63.
706 <https://doi.org/10.1016/j.geomorph.2010.12.004>

707 Ribolzi, O., Rochelle-Newall, E., Dittrich, S., Auda, Y., Newton, P.N., Rattanavong, S., Knappik, M.,
708 Soullileuth, B., Sengtahevanghoun, O., Dance, D.A.B., Pierret, A., 2016b. Land use and soil type
709 determine the presence of the pathogen *Burkholderia pseudomallei* in tropical rivers. *Environ.*
710 *Sci. Pollut. Res.* 23, 7828–7839. <https://doi.org/10.1007/s11356-015-5943-z>

711 Ribolzi, O., Silvera, N., Xayyakummanh, K., Latsachak, K., Tasaketh, S., Vanethongkham, K., 2005. The
712 use of pH to spot groundwater inflows along the stream of a cultivated catchment in the
713 northern Lao PDR. *Lao J. Agric. For.* 10, 72–84.

714 Robert, E., Grippa, M., Kergoat, L., Pinet, S., Gal, L., Cochonneau, G., Martinez, J.-M., 2016. Monitoring
715 water turbidity and surface suspended sediment concentration of the Bagre Reservoir (Burkina
716 Faso) using MODIS and field reflectance data. *Int. J. Appl. Earth Obs. Geoinformation* 52, 243–
717 251. <https://doi.org/10.1016/j.jag.2016.06.016>

718 Robert, E., Kergoat, L., Soumaguel, N., Merlet, S., Martinez, J.-M., Diawara, M., Grippa, M., 2017. Analysis
719 of Suspended Particulate Matter and Its Drivers in Sahelian Ponds and Lakes by Remote Sensing
720 (Landsat and MODIS): Gourma Region, Mali. *Remote Sens.* 9, 1272.
721 <https://doi.org/10.3390/rs9121272>

722 Rochelle-Newall, E., Nguyen, T.M.H., Le, T.P.Q., Sengtahevanghoun, O., Ribolzi, O., 2015. A short review
723 of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions.
724 *Front. Microbiol.* 6, 308. <https://doi.org/10.3389/fmicb.2015.00308>

725 Rochelle-Newall, E.J., Ribolzi, O., Viguier, M., Thammahacksa, C., Silvera, N., Latsachack, K., Dinh, R.P.,
726 Naporn, P., Sy, H.T., Soullileuth, B., Hmaimum, N., Sisouvanh, P., Robain, H., Janeau, J.-L.,
727 Valentin, C., Boithias, L., Pierret, A., 2016. Effect of land use and hydrological processes on

728 Escherichia coli concentrations in streams of tropical, humid headwater catchments. *Sci. Rep.* 6,
729 32974. <https://doi.org/10.1038/srep32974>

730 Smith, J., Edwards, J., Hilger, H., Steck, T.R., 2008. Sediment can be a reservoir for coliform bacteria
731 released into streams. *J. Gen. Appl. Microbiol.* 54, 173–179.
732 <https://doi.org/10.2323/jgam.54.173>

733 Song, L., Boithias, L., Sengtaheuanghoung, O., Oeurng, C., Valentin, C., Souksavath, B., Sounyafong, P., de
734 Rouw, A., Soullileuth, B., Silvera, N., Lattanavongkot, B., Pierret, A., Ribolzi, O., 2020. Understory
735 Limits Surface Runoff and Soil Loss in Teak Tree Plantations of Northern Lao PDR. *Water* 12,
736 2327. <https://doi.org/10.3390/w12092327>

737 Soupir, M.L., Mostaghimi, S., Dillaha, T., 2010. Attachment of *Escherichia coli* and Enterococci to Particles
738 in Runoff. *J. Environ. Qual.* 39, 1019–1027. <https://doi.org/10.2134/jeq2009.0296>

739 Stocker, M.D., Penrose, M., Pachepsky, Y.A., 2018. Spatial Patterns of Concentrations in Sediment before
740 and after High-Flow Events in a First-Order Creek. *J. Environ. Qual.* 47, 958–966.
741 <https://doi.org/10.2134/jeq2017.11.0451>

742 Strauch, A.M., Mackenzie, R.A., Bruland, G.L., Tingley, R., Giardina, C.P., 2014. Climate Change and Land
743 Use Drivers of Fecal Bacteria in Tropical Hawaiian Rivers. *J. Environ. Qual.* 43, 1475–1483.
744 <https://doi.org/10.2134/jeq2014.01.0025>

745 Tong, Y., Yao, R., He, W., Zhou, F., Chen, C., Liu, X., Lu, Y., Zhang, W., Wang, X., Lin, Y., Zhou, M., 2016.
746 Impacts of sanitation upgrading to the decrease of fecal coliforms entering into the environment
747 in China. *Environ. Res.* 149, 57–65. <https://doi.org/10.1016/j.envres.2016.05.009>

748 Troeger, C., Forouzanfar, M., Rao, P.C., 65 other authors, 2017. Estimates of global, regional, and
749 national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the
750 Global Burden of Disease Study 2015. *Lancet Infect. Dis.* 17, 909–948.
751 [https://doi.org/10.1016/S1473-3099\(17\)30276-1](https://doi.org/10.1016/S1473-3099(17)30276-1)

752 Turkelboom, F., Poesen, J., Trébuil, G., 2008. The multiple land degradation effects caused by land-use
753 intensification in tropical steepplands: A catchment study from northern Thailand. *Catena* 75,
754 102–116. <https://doi.org/10.1016/j.catena.2008.04.012>

755 Vigiak, O., Ribolzi, O., Pierret, A., Sengtaheuanghoung, O., Valentin, C., 2008. Trapping Efficiencies of
756 Cultivated and Natural Riparian Vegetation of Northern Laos. *J. Environ. Qual.* 37, 889–897.
757 <https://doi.org/10.2134/jeq2007.0251>

758 Wohlsen, T., Bates, J., Vesey, G., Robinson, W.A., Katouli, M., 2006. Evaluation of the methods for
759 enumerating coliform bacteria from water samples using precise reference standards. *Lett. Appl.*
760 *Microbiol.* 42, 350–356. <https://doi.org/10.1111/j.1472-765X.2006.01854.x>

761 Wold, S., 1995. PLS for Multivariate Linear Modelling. In: van de Waterbeemd H. (ed.), *QSAR:*
762 *Chemometric Methods in Molecular Design.* Vol 2. Wiley-VCH, Weinheim, Germany. pp. 195–
763 218.

764 Ziegler, A.D., Benner, S.G., Tantasirin, C., Wood, S.H., Sutherland, R.A., Sidle, R.C., Jachowski, N., Nullet,
765 M.A., Xi, L.X., Snidvongs, A., Giambelluca, T.W., Fox, J.M., 2014. Turbidity-based sediment
766 monitoring in northern Thailand: Hysteresis, variability, and uncertainty. *J. Hydrol.* 519, 2020–
767 2039. <https://doi.org/10.1016/j.jhydrol.2014.09.010>

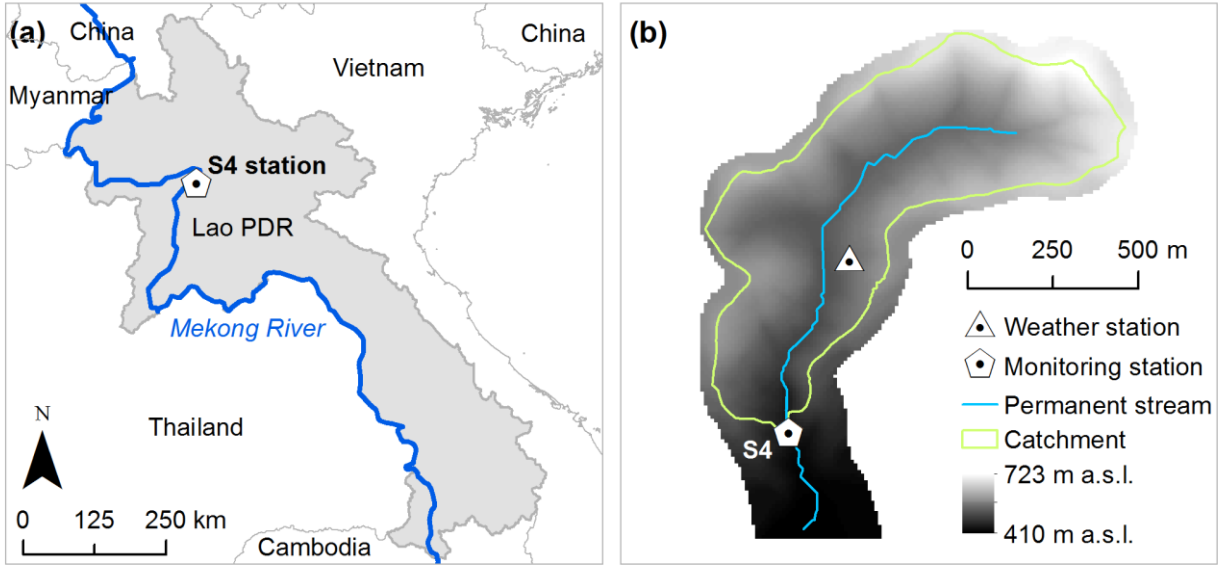
768 Ziegler, A.D., Giambelluca, T.W., Tran, L.T., Vana, T.T., Nullet, M.A., Fox, J., Vien, T.D., Pinthong, J.,
769 Maxwell, J., Evett, S., 2004. Hydrological consequences of landscape fragmentation in
770 mountainous northern Vietnam: evidence of accelerated overland flow generation. *J. Hydrol.*
771 287, 124–146. <https://doi.org/10.1016/j.jhydrol.2003.09.027>

772

773

774 List of figures

775



776

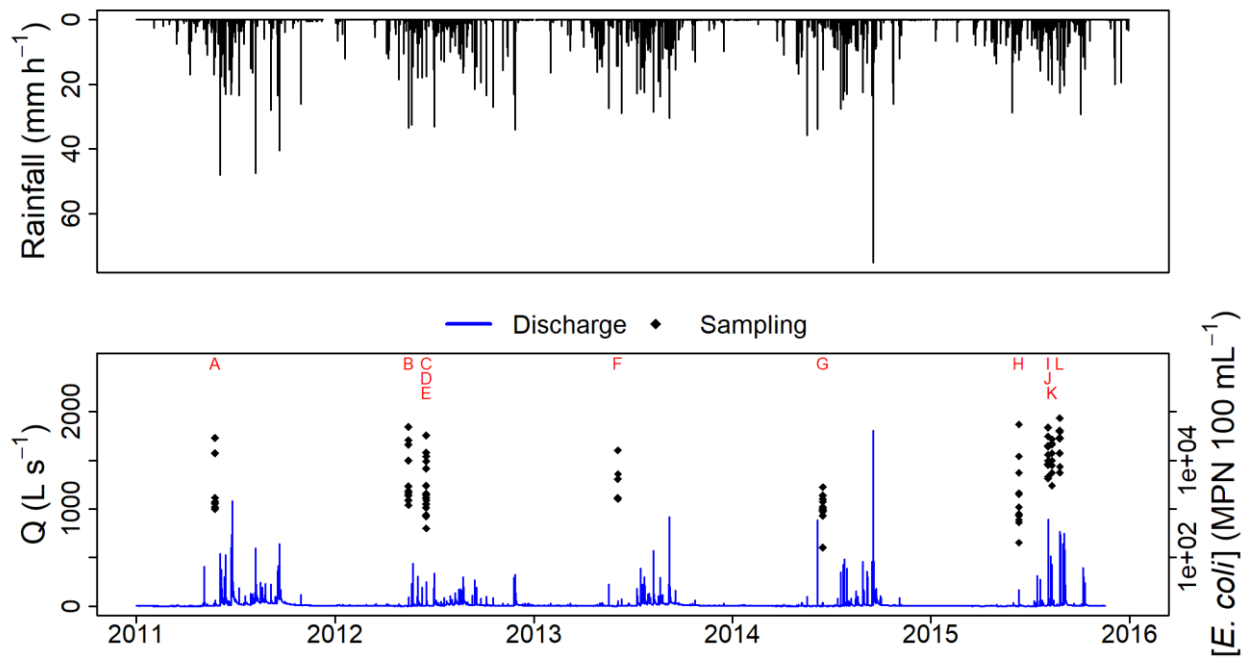
777 Fig. 1. (a) Location of the S4 outlet of the Houay Pano catchment in northern Lao PDR; (b) River gauging
778 and sampling station S4, weather station, and altitudes. Altitudes are given in meters above sea level (m
779 a.s.l.).

780

781

782

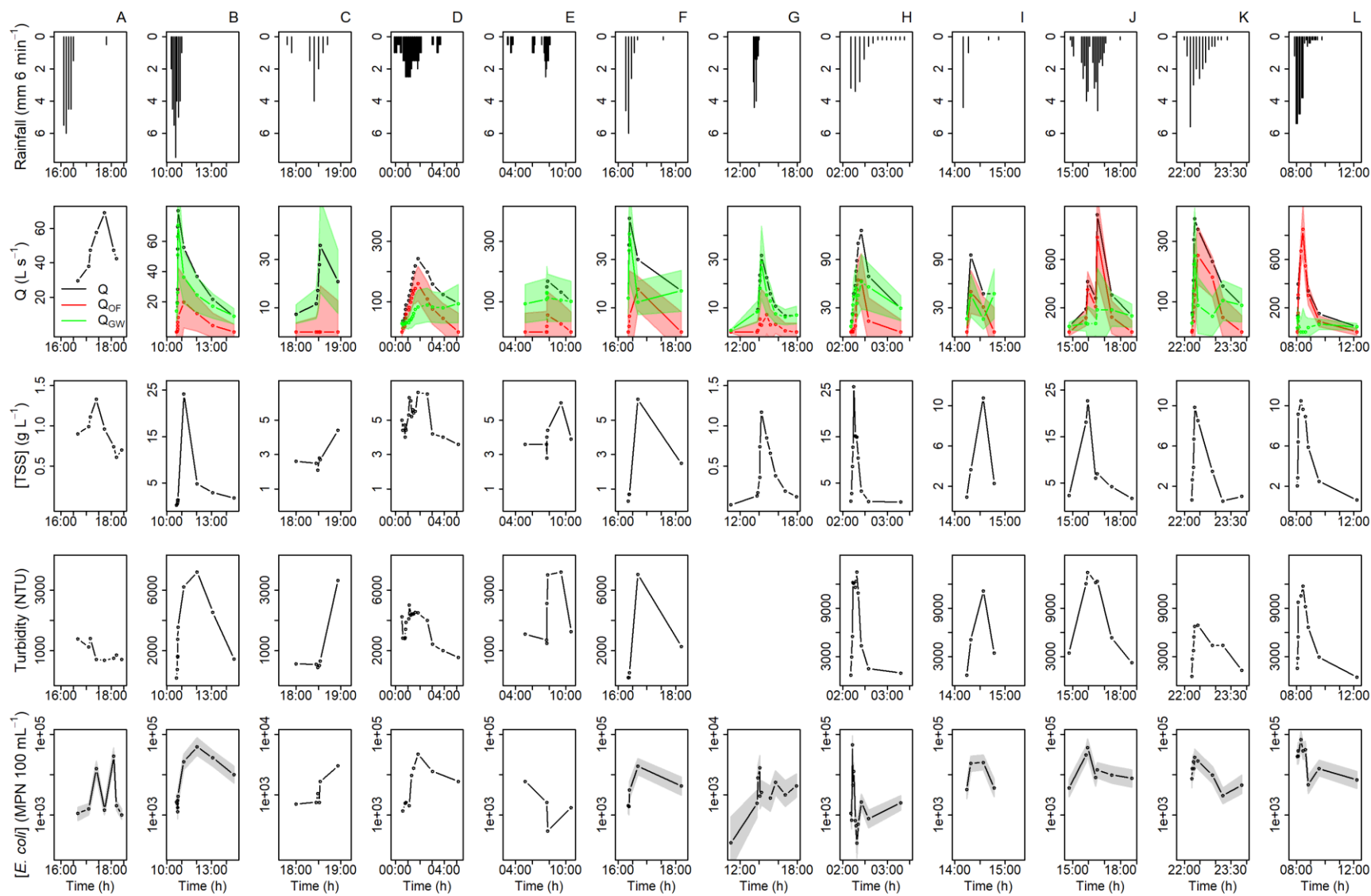
783



784
 785 Fig. 2. Rainfall (mm h^{-1}) and discharge (Q , in L s^{-1}) for the 2011-2015 period, and *E. coli* concentration ([*E.*
 786 *coli*], in $\text{MPN } 100 \text{ mL}^{-1}$) for twelve flood events from 2011 to 2015 at the outlet of the Houay Pano
 787 catchment, northern Lao PDR. The flood events of 16-17 June 2012 (C, D, and E) and of 4 August 2015 (I
 788 and J) have been decomposed into 3 and 2 separate flood events. Details of these two flood events is
 789 shown in Fig. 3. Although [*E. coli*] measures for events C-E were expressed in $\text{CFU } 100 \text{ mL}^{-1}$, they are
 790 reported as $\text{MPN } 100 \text{ mL}^{-1}$ for the sake of simplicity.

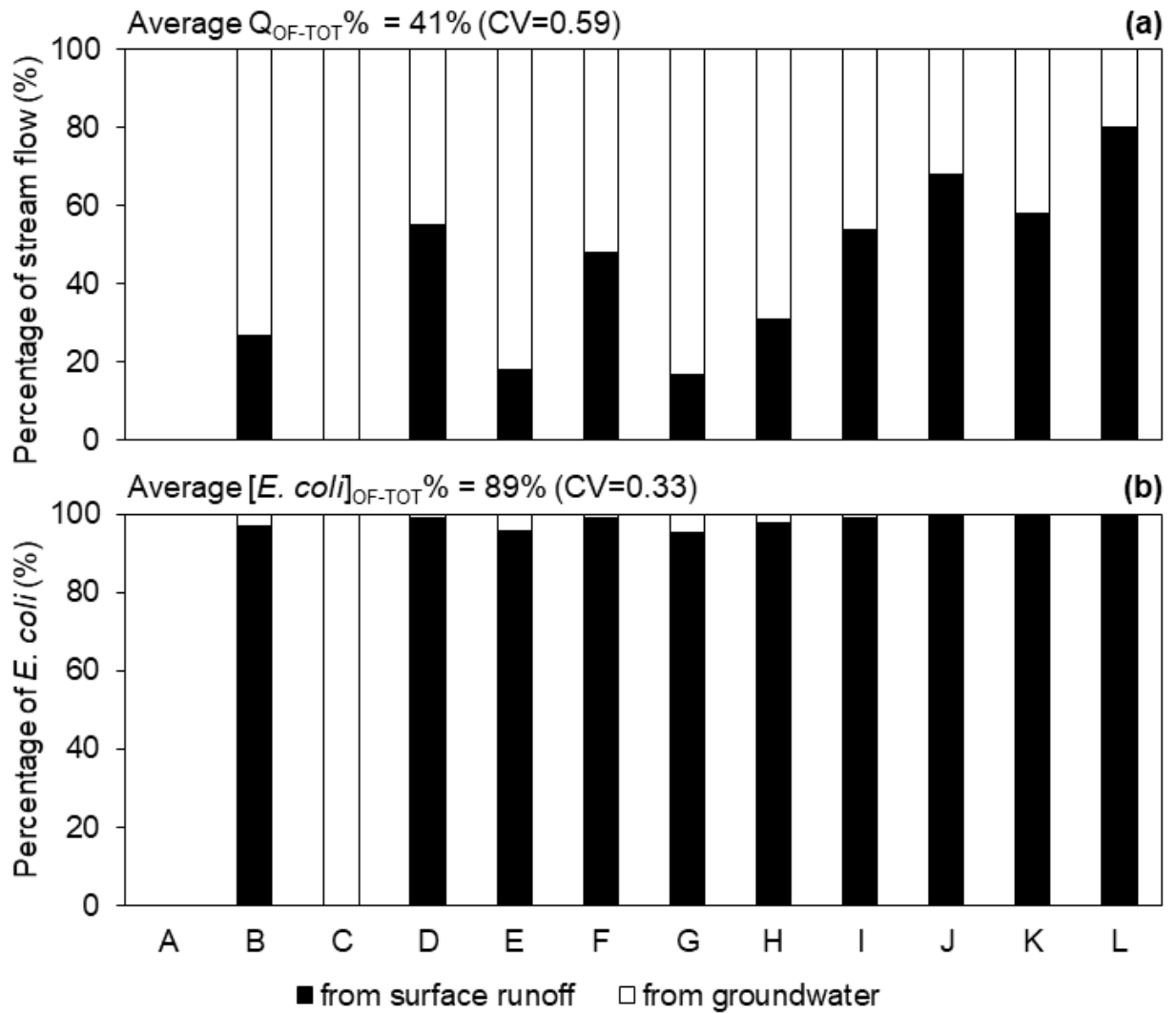
791

33



793 Fig. 3. Monitored flood events from 2011 to 2015 at the outlet of the Houay Pano catchment, northern Lao PDR. Rainfall: rainfall (mm); Q:
794 discharge ($L s^{-1}$); Q_{OF} : surface runoff ($L s^{-1}$); Q_{GW} : sub-surface flow ($L s^{-1}$); [TSS]: total suspended sediments concentration ($g L^{-1}$); Turbidity: turbidity
795 (NTU); [*E. coli*]: *E. coli* concentration (MPN 100 mL⁻¹). Measurements of turbidity were lacking for event G. Although [*E. coli*] measures for events
796 C-E were expressed in CFU 100 mL⁻¹, they are reported as MPN 100 mL⁻¹ for the sake of simplicity. Red and green bands for Q_{OF} and Q_{GW} ,
797 respectively, are uncertainty bands calculated with the Genereux (1998) method. Grey bands for [*E. coli*] are uncertainty intervals given by
798 Poisson distribution when using the standardized microplate method.

799



800

801 Fig. 4. Contributions (in %) of overland flow and of groundwater flow in (a) flood-event total volume of

802 stream flow, and in (b) flood-event total *E. coli* number. The twelve flood events were sampled from

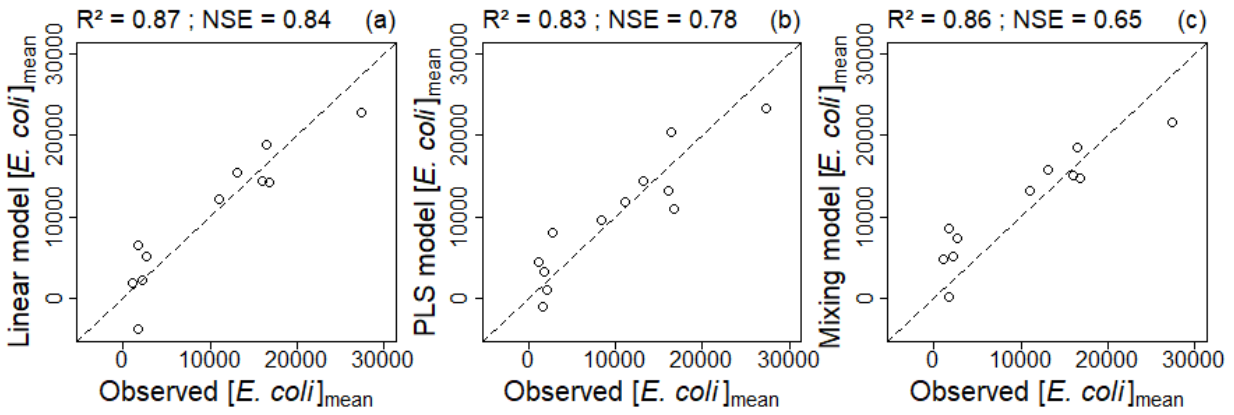
803 2011 to 2015 at the outlet of the Houay Pano catchment, northern Lao PDR. Measurements of electrical

804 conductivity were missing for event A. Average $Q_{OF-TOT}\%$ and average $[E. coli]_{OF-TOT}\%$ are the average

805 contribution of overland flow to the in-stream flow and the average percentage of *E. coli* resulting from

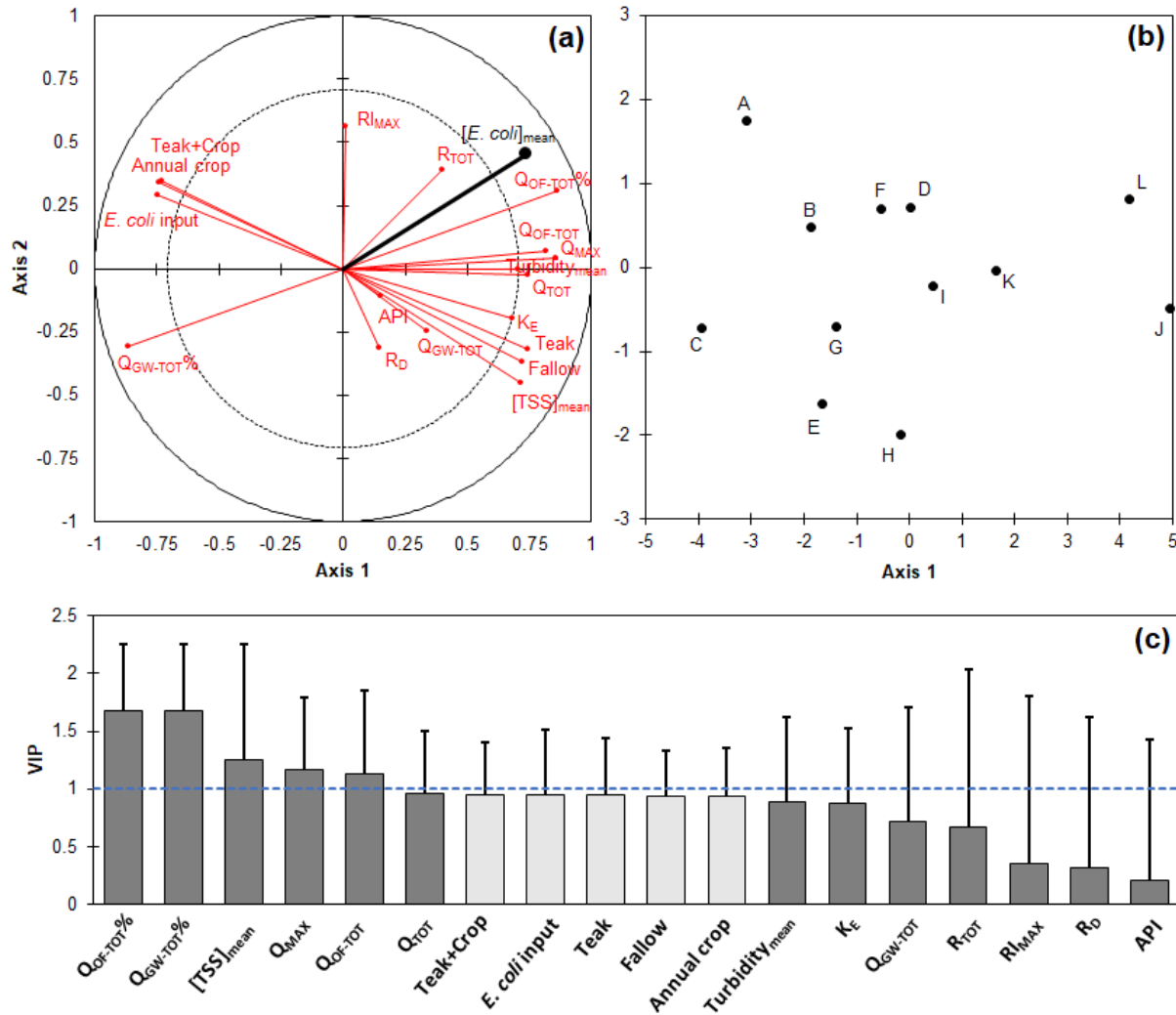
806 overland flow, respectively. CV: coefficient of variation.

807



808

809 Fig. 5. Mean *E. coli* concentration per flood event ($[E. coli]_{\text{mean}}$): comparison of models' predictions with
 810 observed values (MPN 100 mL⁻¹). (a) $Q_{\text{OF-TOT}}\%$ -based linear model against observed $[E. coli]_{\text{mean}}$; (b) PLS
 811 model against observed $[E. coli]_{\text{mean}}$; (c) Mixing model against observed $[E. coli]_{\text{mean}}$. R^2 : coefficient of
 812 determination; NSE: Nash-Sutcliffe Efficiency. The twelve flood events were sampled from 2011 to 2015
 813 at the outlet of the Houay Pano catchment, northern Lao PDR.



814

815 Fig. 6. Partial Least Square regression analysis where the flood-event mean concentration of *E. coli* ($[E.$

816 $coli]_{mean}$) is explained by hydro-meteorological and land-use variables: (a) *E. coli* located in the

817 correlations circle with 13 meteorological and hydrological variables and 5 land-use variables; (b) Map of

818 observations (twelve flood events from 2011 to 2015 at the outlet of the Houay Pano catchment,

819 northern Lao PDR); (c) Variable Importance in Projection (VIP) of the 18 variables for two components.

820 R_D : event rainfall duration; R_{TOT} : cumulated rainfall over the event; R_{MAX} : maximum rainfall intensity; API :

821 antecedent precipitation index; Q_{MAX} : peak discharge during the event; Q_{TOT} : total volume of stream

822 water exported during the event; Q_{OF-TOT} : total volume of surface runoff exported during the event; Q_{OF-

823 $TOT\%$: percentage of surface runoff during the event calculated as $Q_{OF-TOT}/Q_{TOT} * 100$; Q_{GW-TOT} : total volume

824 of sub-surface flow exported during the event; $Q_{GW-TOT}\%$: percentage of sub-surface flow during the
825 event calculated as $Q_{GW-TOT}/Q_{TOT} * 100$; K_E : flow coefficient calculated as Q_{TOT}/R_{TOT} ; $[TSS]_{mean}$: flood-event
826 mean total suspended sediments concentration; $Turbidity_{mean}$: flood-event mean turbidity; Fallow:
827 annual percentage of the catchment area covered by fallow; Teak: annual percentage of the catchment
828 area covered by teak trees; Annual crop: annual percentage of the catchment area covered by annual
829 crops; Teak+Crop: annual percentage of the catchment area covered by annual crops grown under young
830 teak trees; *E. coli* input: monthly *E. coli* input within the catchment.

831 List of tables

832 Table 1. Hydro-meteorological characteristics of the selected flood events from 2011 to 2015 at the outlet of the Houay Pano catchment,
833 northern Lao PDR.

Event	Rainfall			Discharge						Water quality				
	Start date Time	R _D (min)	R _{TOT} (mm)	RI _{MAX} (mm h ⁻¹)	API (mm)	Q _{MAX} (L s ⁻¹)	Q _{TOT} (m ³)	Q _{OF-TOT} (m ³)	Q _{OF-TOT} % (%)	K _E (%)	n (-)	[TSS] _{mean} (g L ⁻¹)	Turbidity _{mean} (NTU)	[<i>E. coli</i>] _{mean} (MPN 100 mL ⁻¹)
A	05/25/2011 16:06	24	22.5	60	15.5	69.1	305.5			2.3	8	1.0	993.6	8388
B	05/15/2012 10:18	42	33.5	75	2.8	80.3	553.4	146.7	27	2.7	10	4.9	5466.0	2672
C	06/16/2012 17:48	54	10.5	40	15.0	35.8	69.7	0.0	0	1.1	6	2.9	594.2	1682
D	06/17/2012 00:00	228	39	25	48.6	242.0	2620.8	1432.4	55	11.2	18	5.3	3252.9	16053
E	06/17/2012 03:00	306	19	25	40.6	168.4	2471.3	440.9	18	21.6	7	4.1	2578.0	2181
F	06/04/2013 16:11	82	14.6	60	12.3	47.0	215.6	103.3	48	2.5	5	2.0	5860.3	11020
G	06/16/2014 13:19	36	15.4	44	9.3	31.7	222.6	38.4	17	2.4	11	0.6		1125
H	06/12/2015 02:06	74	13.2	34	11.1	125.8	308.9	96.4	31	4.0	12	10.1	3696.0	1748
I	08/04/2015 13:22	50	6	44	4.3	95.2	127.4	69.1	54	3.8	4	4.8	6809.8	16776
J	08/04/2015 14:49	131	34.6	46	30.9	967.9	5554.2	3796.0	68	26.7	7	9.8	7790.9	16403
K	08/11/2015 21:59	81	20.8	56	17.4	374.7	1378.9	804.9	58	11.0	9	6.0	5205.8	13112
L	08/26/2015 07:48	116	32.4	54	16.2	849.6	3903.6	3118.3	80	20.0	10	8.2	4822.8	27375

834 R_D: duration of rainfall event; R_{TOT}: cumulated rainfall of the event; RI_{MAX}: maximum rainfall intensity; API: antecedent precipitation index; Q_{MAX}: peak discharge
835 during the flood event; Q_{TOT}: total volume of stream water exported during the event; Q_{OF-TOT}: total volume of surface runoff exported during the event; Q_{OF-}
836 _{TOT}%: percentage of surface runoff during the event calculated as Q_{OF-TOT}/Q_{TOT}*100; K_E: flow coefficient calculated as Q_{TOT}/R_{TOT}; n: sample size; [TSS]_{mean}: flood-
837 event mean total suspended sediments concentration; Turbidity_{mean}: flood-event mean turbidity; [*E. coli*]_{mean}: flood-event mean *E. coli* concentration.
838 Measurements of EC and of Turbidity were lacking for events A and G, respectively. Although *E. coli* measurements for events C-E were expressed in CFU 100
839 mL⁻¹, they are reported as MPN 100 mL⁻¹ for the sake of simplicity.