Escherichia coli concentration, multiscale monitoring over the decade 2011-2021 in the Mekong basin, Lao PDR

Laurie Boithias[1], Olivier Ribolzi[1], Emma Rochelle-Newall[2], Chanthanousone Thammahacksa[3], Paty Nakhle[1], Bounsamay Soulileuth[3], Anne Pando-Bahuon[3], Keooudone Latsachack[3], Norbert Silvera[3], Phabvilay Sounyafong[3], Khampaseuth Xayyathip[3], Rosalie Zimmermann[4,5,6], Sayaphet Rattanavong[4], Priscia Oliva[1], Thomas Pommier[7], Olivier Evrard[8], Sylvain Huon[2], Jean Causse[9], Thierry Henry-des-Tureaux[3], Oloth Sengtaheuanghoung[10], Nivong Sipaseuth[10], Alain Pierret[3]

[1]Géosciences Environnement Toulouse (GET), Université de Toulouse, CNRS, IRD, UPS, Toulouse, France
[2]Sorbonne Université, Univ. Paris Est Creteil, IRD, CNRS, INRAE, Institute of Ecology and Environmental Sciences of Paris (iEES-Paris), Paris, France
[3]IRD, Department of Agricultural Land Management (DALaM), Vientiane, Lao PDR
[4]Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Microbiology Laboratory, Mahosot Hospital, Vientiane, Lao PDR
[5]Department of Environmental Sciences, University of Basel, Basel, Switzerland
[6]Department of Medical Microbiology and Infection Prevention, Amsterdam University Medical Centers (UMC), Amsterdam, The Netherlands
[7]Laboratoire d'Ecologie Microbienne (LEM), CNRS, UCBL, VetAgroSup, Université de Lyon, Villeurbanne, France
[8]Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL), CEA, CNRS, UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[9]Société Transcender, Rennes, France / Ecole des Hautes Etudes en Santé Publique (EHESP), Laboratoire d'Etude et de Recherche en Environnement et Santé, IRSET-INSERM, Rennes, France
[10]Ministry of Agriculture and Forestry (MAF), Department of Agricultural Land Management (DALaM), Vientiane, Lao PDR

Correspondence to: Laurie Boithias (laurie.boithias@get.omp.eu) and Olivier Ribolzi (olivier.ribolzi@ird.fr)

Abstract. Bacterial pathogens in surface waters may threaten human health, especially in developing countries, where untreated surface water is often used for domestic needs. The objective of the long-term multiscale monitoring of *Escherichia coli* concentration in stream water, and that of associated variables (temperature, electrical conductance, dissolved oxygen concentration and saturation, pH, oxidation-reduction potential, turbidity, and total suspended sediment concentration), was to identify the drivers of bacterial dissemination across tropical catchments. This data description paper presents three datasets (see section Data availability) collected at 31 sampling stations located within the Mekong river and its tributaries in Lao PDR (0.6-25,946 km²) from 2011 to 2021. The 1,602 records have been used to describe the hydrological processes driving in-stream *Escherichia coli* concentration during flood events, to understand land-use impact on bacterial dissemination on small and large catchment scales, to relate stream water quality and diarrhea outbreaks, and to build numerical models. The database may be further used e.g. to interpret new variables measured in the monitored catchments, or to map the health risk posed by fecal pathogens.

1 Introduction

Bacterial pathogens, including fecal bacteria, are etiological agents of several waterborne diseases such as diarrhea. Primary sources of fecal bacteria in the environment are cattle droppings or human feces where open defecation is practiced or where sanitation systems are lacking or deficient (Exley et al., 2015; Tong et al., 2016; Rochelle-Newall et al., 2015). Fecal bacteria may threaten human health if present in surface water, especially in developing countries where population often uses untreated surface water for domestic needs (Boithias et al., 2016; Vos et al., 2020).

Prior to 2010, to our knowledge, no data had been published to document the occurrence of Fecal Indicator Bacteria (FIB), such as *Escherichia coli* (*E. coli*), in stream water of rural, tropical catchments, such as the area of the 800,000 km^2 Mekong river basin. Besides, very little information existed on microbial contamination and dissemination mechanisms in tropical environments. Accordingly, our research group initiated the systematic monitoring of surface water quality in 2011, focusing on Lao PDR, a landlocked country that contributes 41 and 54% of the total Mekong flow during the dry and the rainy seasons, respectively (MRC, 2009), and where about 67% of the population lives in rural areas (Lao Statistics Bureau, 2015).

The objective of the spatial and temporal monitoring of *E. coli* concentrations in stream water, and that of associated physico-chemical measurements (i.e., temperature, electrical conductance at 25 °C, dissolved oxygen concentration and saturation, pH, oxidation-reduction potential, turbidity, and total suspended sediment concentration), was to understand bacterial fate and transport, and underlying drivers, during storm and inter-storm flow periods within tropical catchments. The study design included three monitoring scales:

(1) A spatial survey conducted during both the dry and the rainy seasons in the Mekong river (6 sampling stations) and its tributaries (23 sampling stations) across Lao PDR. The survey aimed at assessing the spatial variability of in-stream *E. coli* concentration at large spatial scales (239-25,946 km^2 for the tributaries, up to 549,055 km^2 for the Mekong itself), during both low flow and high flow periods. Grab samples were collected twice: in March and in July 2016;

(2) A temporal monitoring at the outlet of six catchments (6 sampling stations) in the mountainous area of northern Lao PDR, initiated in 2011, to assess the temporal variability of in-stream *E. coli* concentration at large spatial scales (239-25,946 km^2 for the tributaries, up to 272,155 km^2 for the Mekong itself), during both low flow and high flow periods. Grab samples were first collected with a biweekly time-interval, thereafter modified to a 10-day time-interval in 2017;

(3) A temporal monitoring at the outlet of a 0.6 km^2 headwater catchment (1 sampling station) in the mountainous area of northern Lao PDR, initiated in 2011, which aimed at understanding the dynamics and the drivers of *E. coli* dissemination during low flow and flood events. Grab samples were collected during low flow with a biweekly time-interval, thereafter modified to a 10-day time-interval in 2017. During flood events, water sample collection was made with an automatic sampler triggered by water level change.

2 Methods

2.1 Study site

The Mekong river basin has experienced a dramatic economic and population growth over the past half-century, increasing pressure on land, water, and other natural resources (Pokhrel et al., 2018; Arias et al., 2014; Global Water Forum, 2015). Rapid land-use change, such as deforestation (Lyon et al., 2017) and conversion from traditional slash-and-burn agricultural systems to tree plantations (Ribolzi et al., 2017), have put soil fertility, agricultural productivity, and biodiversity conservation at stake (MA, 2005). Since the 1970s, dams have been constructed on the Mekong river and its tributaries (WLE, 2017), leading to severe impacts on hydrology (Le Meur et al., 2021; Hecht et al., 2019), on sediment transport (Kondolf et al., 2014; Shrestha et al., 2018), and on aquatic biodiversity (Sabo et al., 2017), in a climate change context.

The 31 sampling stations (Fig. 1, Table 1) of this dataset are located in Lao PDR, within the Mekong river basin. The choice of sampling stations in Mekong tributaries was made to encompass a broad range of catchment sizes (0.6-25,946 km²), and a large range of geological, topographical, and land-use features. In Lao PDR, the tropical climate (i.e., wet and dry Aw climate) is under the influence of the monsoon regime, dividing the year into two seasons: a dry season lasting from October to April, and a rainy season lasting from May to October. The average annual rainfall in Lao PDR varies from 1,300 to 2,500 mm and can exceed 3,500 mm in central and southwestern Lao PDR (Nakhle et al., 2021b).

The 0.6 km² Houay Pano headwater catchment (station S4) is located in northern Lao PDR, 10 km south of Luang Prabang city (Fig. 1a). This experimental site (Boithias et al., 2021b) is part of the Multiscale TROPIcal CatchmentS Critical Zone Observatory (M-TROPICS CZO; https://mtropics.obs-mip.fr/), a network of observatories under the French Research Infrastructure OZCAR (Gaillardet et al., 2018). This catchment is representative of the montane agro-ecosystems of Southeast Asia. Altitude within the catchment is 435-716 m (Fig. 1b) and the slope gradient is 1-135 % (mean=52 %). Over the last decade, land-use change in the catchment mostly consisted of an increase of teak tree plantations at the expense of shifting cultivation (with slash-and-burn method). Overall, the areal percentage of annual crops decreased from 28% to virtually zero, while the areal percentage of teak tree plantations increased from 17% to 33% from 2011 to 2021.

2.2 Data collection

The dataset comprises *E. coli* concentrations ([*E. coli*]) with lower and upper limits of the confidence interval ($[E. coli]_{LL}$ and $[E. coli]_{UL}$, respectively), and physico-chemical measurements in stream water, recorded from May 25, 2011, to May 25, 2021. Physico-chemical measurements include temperature (T), electrical conductance (EC) at 25 °C, dissolved oxygen concentration ([DO]) and saturation (DO%), pH (pH), oxidation-reduction potential (ORP), turbidity (Turb), and total suspended sediment concentration ([TSS]). The units of the eleven variables are given in Table 2.

For the spatial survey during both the dry and the rainy seasons, we chose sampling sites so as to ensure a broad geographical coverage of Lao PDR, and to represent a large range of geological, topographical, and land-use features. We also chose the sampling sites for being accessible from the road, in order to achieve the sampling campaign within a relatively short time.

For the temporal monitoring, we initiated the sampling at stations S4, NK20, NK26, and A6, in May 2011 with a biweekly time-interval, changed to a 10-day time-interval in 2017. We initiated the monitoring at stations Nou_1, Nse_1, and MK_17 in July 2017 according to the 10-day time-interval.

At the station S4, we collected water samples during flood events using an automatic sampler (Automatic Pumping Type Sediment Sampler, ICRISAT). The automatic sampler was triggered by a water level recorder to collect water after every 2 cm water level change during flood rising and every 4 cm water level change during flood recession. We measured EC, DO%, [DO], pH, and ORP once back to the laboratory, within 6 hours after automatic sampling, with a portable multi-probe system (YSI 556), and Turb with a portable turbidity meter (EUTECH Instruments TN-100). Maximal storage duration of 6 hours ensured that [*E. coli*] variation was not significant before laboratory analysis (Nakhle et al., 2021a).

At the other stations, and at the station S4 during low flow, we manually sampled water with a bailer sampler and measured *in situ* T, EC, DO%, [DO], pH, and ORP, with the multi-probe system and Turb with the turbidity meter. Bailer-sampled water was typically collected 5-10 m from the stream bank, except for station S4 where water was collected in the middle of the stream, about 0.5 m from the stream bank, and for stations MK_17, MK_3, and MK_5, where we used a boat to collect the water sample in the middle of the stream. We stored water samples in clean plastic bottles in an opaque icebox until laboratory analysis within 6 hours.

We measured [*E. coli*] in the laboratory with the standardized microplate method (ISO 9308-3). For each water sample, we incubated a water sub-sample at four dilution rates (i.e., 1:2, 1:20, 1:200, and 1:2000) in a 96-well microplate (MUG/EC, BIOKAR DIAGNOSTICS) for 48 h at 44 °C. Ringers' Lactate solution was used for the dilutions and one plate was used per sample. We tested the sterility of Ringers' Lactate solution and found no *E. coli*. The number of positive wells for each microplate was noted and the Most Probable Number (MPN) per 100 mL was determined using the Poisson distribution. Given the four dilution ratios, the detection limit of [*E. coli*] was 38 MPN 100 mL^{-1} with lower confidence limit at 5.4 MPN 100 mL^{-1} and upper confidence limit at 270 MPN 100 mL^{-1}. Among the three datasets, the number of water samples with [*E. coli*] below the limit of detection was 5.9 %.

We measured [TSS] in the laboratory after the filtration of 100 mL of sample water on 0.2 µm porosity cellulose acetate filters (Sartorius) and evaporation at 105 °C for 48 h. We used 0.2 µm filters to ensure that filtration trapped the clay fraction of the water samples, knowing that clay fraction is up to 72.9 % in adjacent soils (Chaplot and Poesen, 2012).

3 Results

The spatial survey (Ribolzi et al., 2021c) conducted in both the dry and the rainy seasons in 2016 resulted in 58 records (Fig. 2). The dataset shows contrasted values of [*E. coli*] depending on the season (Nakhle et al., 2021b). Median [*E. coli*] are higher during the rainy season, similarly to [TSS], Turb, ORP, and T, while EC, DO%, and pH show smaller median values during the rainy season compared to the dry season.

The temporal monitoring at the outlet of the six catchments (Ribolzi et al., 2021a), initiated in May 2011 at stations NK20, NK26, and A6 (Fig. 3), and in July 2017 at stations Nou_1, Nse_1, and MK_17 (Fig. 4), resulted in 1,131 records until May 2021. The dataset shows seasonal variations of [*E. coli*] and of the other variables (Boithias et al., 2016). In general, T and EC were increasing throughout the dry season and decreasing throughout the rainy season. The highest and lowest values of Turb, [TSS], and [*E. coli*], were measured during the rainy and the dry seasons, respectively. The temporal dynamics of DO%, [DO], pH, and ORP, are less clear. At the station NK26 the number of water samples with [*E. coli*] below the detection limit increased after 2016 (Fig. 3). The data gap between March 12 and June 20, 2020, is due to COVID-19 lockdown and traffic restrictions. The temporal monitoring at the station S4, outlet of the Houay Pano headwater catchment (Ribolzi et al., 2021b), initiated in May 2011, resulted in 413 records until May 2021 (Fig. 5). The dataset shows seasonal variations of [*E. coli*] and of the other variables (Boithias et al., 2016). Similar to the large catchments, T and EC were generally increasing throughout the dry season and decreasing throughout the rainy season. The highest and lowest values of Turb, [TSS], and [*E. coli*], were measured during the rainy and the dry seasons, respectively. The temporal dynamics of DO%, [DO], pH, and ORP, are less clear. We monitored [*E. coli*] dynamics during 14 flood events (Boithias et al., 2021a). Values of [*E. coli*] measured during flood events with the automatic sampler are in general higher than those measured during low flow periods in grab samples.

4 Technical validation

We calibrated the multi-probe system each day before measurement, i.e. every day during the spatial survey (Ribolzi et al., 2021c) and at a biweekly or 10-day frequency for the two other datasets (Ribolzi et al., 2021b, a). EC probe was calibrated with a 1,413 µS cm^{-1} solution. [DO] probe was calibrated following the air-calibration chamber in air method (USGS, 2006). DO% was then automatically calibrated based on [DO] and the barometric pressure. pH probe was calibrated using a 3-point calibration (pH = 4.01, 7.01, and 9.18). ORP probe was calibrated with a 240 mV solution. For the turbidity meter, we verified each day before measurement that the turbidity measured using the 100 NTU calibration solution was correct. If a discrepancy was observed, we calibrated the turbidity meter with 4 calibration solutions at 0.02, 20, 100, and 800 NTU. The accuracy of the suspended sediment mass on the 0.2 µm filters was ensured by the use of a 10^{-4} g precision balance. We assessed the uncertainty of [*E. coli*] by using the MPN statistical method, which supplies the upper and the lower limits of the confidence interval.

We collected water samples as far as possible from the stream bank, to avoid any influence of the latter. Similarly, we chose sampling stations so as to not be affected by an upstream confluence. Physico-chemistry and suspended sediment concentration can be heterogeneous along a stream cross-section (Santini et al., 2019), but for reasons of logistical capacity of sampling, the measurements could only be made from a single sample at the different stations. However, to ensure that measurements at sampling point were within the range of variation of the values measured along the stream cross-section, we performed a 10-point survey across the river section at station NK20, which is a referenced station of the Lao PDR national hydrological monitoring network (Ribolzi et al., 2022).

Before laboratory analysis, water samples were stored in clean plastic bottles. These bottles were new empty bottles, supplied by a plastic bottle plant producing bottles for mineral water packaging. We regularly verified that the bottles were free from *E. coli*, and we triple rinsed the bottle with stream water before each water sampling.
Records were collated, curated, and cross-checked. If the information recorded in the databases was ambiguous or did not match between records, we traced the sample back using the original paper records as far as possible. Samples which could not be verified in this way were excluded from publication in the current dataset.

5 Usage note

We published the three datasets as .CSV files so that they can be accessed and processed with any data processing software. The three datasets are open access, licensed under a Creative Commons Attribution 4.0 International License. For the proper functioning and the sustainability of the CZO, M-TROPICS asks data users to (1) acknowledge the M-TROPICS CZO and its financing institutions in their publications, (2) cite both the appropriate DOI of the data they used and the related publications, and (3) send to the corresponding authors of the present data description paper (LB and/or OR) a copy of any produced material based on the data. The corresponding authors, who are intimately familiar with the background of these datasets, are at the disposal of the authors wishing to reuse the datasets.

6 Data availability

The database includes a total of 1,602 records and is publicly available online as a collection of three files (Ribolzi et al., 2021b, a, c) hosted within the DataSuds platform (https://dataverse.ird.fr/). The three repositories are:

(1) *Escherichia coli* concentrations and physico-chemical measurements at the outlet of 29 catchments of the Mekong river basin, Lao PDR, during dry and rainy seasons (2016) (Ribolzi et al., 2021c): https://doi.org/10.23708/ZRSBM4
(2) *Escherichia coli* concentrations and physico-chemical measurements (2011-2021) at the outlet of six catchments of the Mekong river basin, northern Lao PDR (Ribolzi et al., 2021a): https://doi.org/10.23708/1YZQHH
(3) *Escherichia coli* concentrations and physico-chemical measurements (2011-2021) at the outlet of the Houay Pano catchment, northern Lao PDR (Ribolzi et al., 2021b): https://doi.org/10.23708/EWOYNK

We published the three datasets as .CSV files and a data dictionary along with the data files. The column headings of the data dictionary files are listed in Table 2, together with the unit of each variable.

7 Conclusions

These three datasets together present a unique long-term spatiotemporal and multiscale surface water quality monitoring within the Mekong river basin. So far, the datasets have been used: (1) to describe the hydrological processes driving in-stream *E.*

coli concentration during flood events (Boithias et al., 2021a; Ribolzi et al., 2016b), (2) to understand the role of land use in bacterial dissemination on small and large catchment scales, e.g. *E. coli* (Causse et al., 2015; Rochelle-Newall et al., 2016; Nakhle et al., 2021b; Ribolzi et al., 2011) and *Burkholderia pseudomallei* (Ribolzi et al., 2016a; Zimmermann et al., 2018; Liechti et al., 2021), (3) to relate stream water quality and diarrhea outbreaks (Boithias et al., 2016), and (4) to build catchment-scale numerical models focused on water quality (Kim et al., 2017, 2018; Abbas et al., 2021, 2022).

The dataset may be further used (1) to assess the role of headwater catchments as *E. coli* source in large tropical river basins, (2) to interpret new variables measured in the monitored catchments (e.g. contaminants other than *E. coli*), (3) to assess the impact of dams on downstream *E. coli* concentration, (4) to map the health risk posed by fecal pathogens, and (5) to assess the relative contributions of both climate and land-use change on changes in in-stream *E. coli* concentration.

Author contribution

O.R. designed the study. L.B., O.R., and A.P. coordinated the project. O.R., R.Z., S.R., P.O., and A.P. collected water samples and performed field measurements for dataset https://doi.org/10.23708/ZRSBM4. L.B., O.R., C.T., B.S., K.L., N.Sil., P.S., K.X., J.C., and T.H.d.T collected water samples and performed field measurements for dataset https://doi.org/10.23708/1YZQHH. L.B., O.R., E.R.N., C.T., B.S., K.L., N.Sil., P.S., K.X., T.P., O.E., and S.H. collected water samples and performed field measurements for dataset https://doi.org/10.23708/EWOYNK. C.T. and A.P.B. performed microbial laboratory analysis. L.B. and O.R. validated the data and curated the database. L.B., O.R., E.R.N., and P.N. analyzed the data. O.S. and N.Sip. provided institutional support. L.B. wrote the original draft of the manuscript. All other authors reviewed and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

These three datasets have been collected thanks to the long-term partnership with the Department of Agricultural Land Management (DALaM), Lao PDR, which granted the permission for field access, and to the financial, scientific, technical, and logistical support of the Multiscale TROPIcal CatchmentS Critical Zone Observatory (M-TROPICS CZO, previously MSEC; https://mtropics.obs-mip.fr/).

The authors acknowledge the financial support of the Consultative Group on International Agricultural Research (CGIAR; Humidtropics program; http://humidtropics.cgiar.org/), the GIS-Climat (Pastek program, http://www.gisclimat.fr/projet/pastek.html), the French National Research Agency (TecItEasy project ANR-13-AGRO-0007;

https://anr.fr/), the Institut de Recherche pour le Développement (IRD), including through the International Joint Laboratory on the impact of rapid Land-use change on Soil Ecosystem Services (LMI LUSES; ECOFILTER program; https://luses.ird.fr/) and the regional pilot program Soils, Waters, Coastal Zones and Societies in Southern and Southeast Asia (SELTAR-RPP), the International Center for Tropical Agriculture (CIAT) through the Payment for Environmental Services (PES) program in Lao PDR, the French Centre National de la Recherche Scientifique (CNRS) through the national grants EC2CO-Biohefect/Ecodyn//Dril/MicrobiEn (Belcrue, Belkong, and NARIBACT projects), the US Defence Threat Reduction Agency Cooperative Biological Engagement Programme (contract HDTRA-16-C-0017), the Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit funded by the Wellcome Trust of Great Britain (Grant number 089275/H/09/Z), the Li Ka Shing Foundation of the University of Oxford (grant SM40), and the Nam Theun 2 Power Company (NTPC).

References

Abbas, A., Baek, S., Silvera, N., Soulileuth, B., Pachepsky, Y., Ribolzi, O., Boithias, L., and Cho, K. H.: In-stream *Escherichia coli* modeling using high-temporal-resolution data with deep learning and process-based models, Hydrol. Earth Syst. Sci., 25, 6185–6202, https://doi.org/10.5194/hess-25-6185-2021, 2021.

Abbas, A., Boithias, L., Pachepsky, Y., Kim, K., Chun, J. A., and Cho, K. H.: AI4Water v1.0: an open-source python package for modeling hydrological time series using data-driven methods, Geosci. Model Dev., 15, 3021–3039, https://doi.org/10.5194/gmd-15-3021-2022, 2022.

Arias, M. E., Cochrane, T. A., Kummu, M., Lauri, H., Holtgrieve, G. W., Koponen, J., and Piman, T.: Impacts of hydropower and climate change on drivers of ecological productivity of Southeast Asia's most important wetland, Ecol. Model., 272, 252–263, https://doi.org/10.1016/j.ecolmodel.2013.10.015, 2014.

Boithias, L., Choisy, M., Souliyaseng, N., Jourdren, M., Quet, F., Buisson, Y., Thammahacksa, C., Silvera, N., Latsachack, K., Sengtaheuanghoung, O., Pierret, A., Rochelle-Newall, E., Becerra, S., and Ribolzi, O.: Hydrological regime and water shortage as drivers of the seasonal incidence of diarrheal diseases in a tropical montane environment, PLoS Negl. Trop. Dis., 10, e0005195, https://doi.org/10.1371/journal.pntd.0005195, 2016.

Boithias, L., Ribolzi, O., Lacombe, G., Thammahacksa, C., Silvera, N., Latsachack, K., Soulileuth, B., Viguier, M., Auda, Y., Robert, E., Evrard, O., Huon, S., Pommier, T., Zouiten, C., Sengtaheuanghoung, O., and Rochelle-Newall, E.: Quantifying the effect of overland flow on *Escherichia coli* pulses during floods: Use of a tracer-based approach in an erosion-prone tropical catchment, J. Hydrol., 594, 125935, https://doi.org/10.1016/j.jhydrol.2020.125935, 2021a.

Boithias, L., Auda, Y., Audry, S., Bricquet, J., Chanhphengxay, A., Chaplot, V., de Rouw, A., Henry des Tureaux, T., Huon, S., Janeau, J., Latsachack, K., Le Troquer, Y., Lestrelin, G., Maeght, J., Marchand, P., Moreau, P., Noble, A., Pando-Bahuon, A., Phachomphon, K., Phanthavong, K., Pierret, A., Ribolzi, O., Riotte, J., Robain, H., Rochelle-Newall, E., Sayavong, S., Sengtaheuanghoung, O., Silvera, N., Sipaseuth, N., Soulileuth, B., Souliyavongsa, X., Sounyaphong, P., Tasaketh, S., Thammahacksa, C., Thiebaux, J., Valentin, C., Vigiak, O., Viguier, M., and Xayyathip, K.: The Multiscale TROPIcal CatchmentS critical zone observatory M-TROPICS dataset II: land use, hydrology and sediment production monitoring in Houay Pano, northern Lao PDR, Hydrol. Process., 35, e14126, https://doi.org/10.1002/hyp.14126, 2021b.

Causse, J., Billen, G., Garnier, J., Henri-des-Tureaux, T., Olasa, X., Thammahacksa, C., Latsachak, K. O., Soulileuth, B., Sengtaheuanghoung, O., Rochelle-Newall, E., and Ribolzi, O.: Field and modelling studies of *Escherichia coli* loads in tropical streams of montane agro-ecosystems, J. Hydro-environ. Res., 9, 496–507, https://doi.org/10.1016/j.jher.2015.03.003, 2015.

Chaplot, V. and Poesen, J.: Sediment, soil organic carbon and runoff delivery at various spatial scales, 88, 46–56, https://doi.org/10.1016/j.catena.2011.09.004, 2012.

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

Gaillardet, J., Braud, I., Hankard, F., Anquetin, S., Bour, O., Dorfliger, N., de Dreuzy, J. R., Galle, S., Galy, C., Gogo, S., Gourcy, L., Habets, F., Laggoun, F., Longuevergne, L., Le Borgne, T., Naaim-Bouvet, F., Nord, G., Simonneaux, V., Six, D., Tallec, T., Valentin, C., Abril, G., Allemand, P., Arènes, A., Arfib, B., Arnaud, L., Arnaud, N., Arnaud, P., Audry, S., Comte, V. B., Batiot, C., Battais, A., Bellot, H., Bernard, E., Bertrand, C., Bessière, H., Binet, S., Bodin, J., Bodin, X., Boithias, L., Bouchez, J., Boudevillain, B., Moussa, I. B., Branger, F., Braun, J. J., Brunet, P., Caceres, B., Calmels, D., Cappelaere, B., Celle-Jeanton, H., Chabaux, F., Chalikakis, K., Champollion, C., Copard, Y., Cotel, C., Davy, P., Deline, P., Delrieu, G., Demarty, J., Dessert, C., Dumont, M., Emblanch, C., Ezzahar, J., Estèves, M., Favier, V., Faucheux, M., Filizola, N., Flammarion, P., Floury, P., Fovet, O., Fournier, M., Francez, A. J., Gandois, L., Gascuel, C., Gayer, E., Genthon, C., Gérard, M. F., Gilbert, D., Gouttevin, I., Grippa, M., Gruau, G., Jardani, A., Jeanneau, L., Join, J. L., Jourde, H., Karbou, F., Labat, D., Lagadeuc, Y., Lajeunesse, E., Lastennet, R., Lavado, W., Lawin, E., Lebel, T., Le Bouteiller, C., Legout, C., Lejeune, Y., Le Meur, E., Le Moigne, N., Lions, J., et al.: OZCAR: The French Network of Critical Zone Observatories, Vadose Zone J., 17, 180067, https://doi.org/10.2136/vzj2018.04.0067, 2018.

Global Water Forum: Basins under pressure: the Mekong basin, 2015.

Hecht, J. S., Lacombe, G., Arias, M. E., Dang, T. D., and Piman, T.: Hydropower dams of the Mekong River basin: A review of their hydrological impacts, J. Hydrol., 568, 285–300, https://doi.org/10.1016/j.jhydrol.2018.10.045, 2019.

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

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

Kondolf, G. M., Rubin, Z. K., and Minear, J. T.: Dams on the Mekong: Cumulative sediment starvation, Water Resour. Res., 50, 5158–5169, https://doi.org/10.1002/2013WR014651, 2014.

Lao Statistics Bureau: Results of Population and Housing Census 2015 - The 4th Population and Housing Census (PHC), 2015.

Le Meur, M., Le Phu, V., and Nicolas, G.: What Is the Future of the Lower Mekong Basin Struggling against Human Activities? A Review, in: River Deltas - Recent Advances, IntechOpen, https://doi.org/10.5772/intechopen.95010, 2021.

Liechti, N., Zimmermann, R. E., Zopfi, J., Robinson, M. T., Pierret, A., Ribolzi, O., Rattanavong, S., Davong, V., Newton, P. N., Wittwer, M., and Dance, D. A. B.: Whole-Genome Assemblies of 16 *Burkholderia pseudomallei* Isolates from Rivers in Laos, Microbiol. Resour. Announc., 10, https://doi.org/10.1128/MRA.01226-20, 2021.

Lyon, S. W., King, K., Polpanich, O., and Lacombe, G.: Assessing hydrologic changes across the Lower Mekong Basin, J. Hydrol. Reg. Stud., 12, 303–314, https://doi.org/10.1016/j.ejrh.2017.06.007, 2017.

MA: Millennium Ecosystem Assessment. Ecosystems and Human Well-being: Synthesis. Island Press/World Resources Institute, Washington DC, 2005.

MRC: The flow of the Mekong. Mekong River Commission Management Information Booklet Series No. 2, 2009.

Nakhle, P., Boithias, L., Pando-Bahuon, A., Thammahacksa, C., Gallion, N., Sounyafong, P., Silvera, N., Latsachack, K., Soulileuth, B., Rochelle-Newall, E. J., Marcangeli, Y., Pierret, A., and Ribolzi, O.: Decay Rate of *Escherichia coli* in a Mountainous Tropical Headwater Wetland, Water, 13, 2068, https://doi.org/10.3390/w13152068, 2021a.

Nakhle, P., Ribolzi, O., Boithias, L., Rattanavong, S., Auda, Y., Sayavong, S., Zimmermann, R., Soulileuth, B., Pando, A., Thammahacksa, C., Rochelle-Newall, E., Santini, W., Martinez, J. M., Gratiot, N., and Pierret, A.: Effects of hydrological regime and land use on in-stream *Escherichia coli* concentration in the Mekong basin, Lao PDR, Sci. Rep., 11, 3460, https://doi.org/10.1038/s41598-021-82891-0, 2021b.

Pokhrel, Y., Burbano, M., Roush, J., Kang, H., Sridhar, V., and Hyndman, D.: A Review of the Integrated Effects of Changing Climate, Land Use, and Dams on Mekong River Hydrology, Water, 10, 266, https://doi.org/10.3390/w10030266, 2018.

Ribolzi, O., Cuny, J., Sengsoulichanh, P., Mousquès, C., Soulileuth, B., Pierret, A., Huon, S., and Sengtaheuanghoung, O.: Land Use and Water Quality Along a Mekong Tributary in Northern Lao P.D.R., Environ. Manage., 47, 291–302, https://doi.org/10.1007/s00267-010-9593-0, 2011.

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

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

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

Ribolzi, O., Boithias, L., Thammahacksa, C., Silvera, N., Pando-Bahuon, A., Sengtaheuanghoung, O., Sipaseuth, N., Latsachack, K., Soulileuth, B., Sounyafong, P., Khampaseuth, X., and Pierret, A.: *Escherichia coli* concentrations and physico-chemical measurements (2011-2021) at the outlet of six catchments of the Mekong river basin, northern Lao PDR [Data set], https://doi.org/10.23708/1YZQHH, 2021a.

Ribolzi, O., Boithias, L., Thammahacksa, C., Rochelle-Newall, E., Pando-Bahuon, A., Silvera, N., Sengtaheuanghoung, O., Sipaseuth, N., and Pierret, A.: *Escherichia coli* concentrations and physico-chemical measurements (2011-2021) at the outlet of the Houay Pano catchment, northern Lao PDR [Data set], https://doi.org/10.23708/EWOYNK, 2021b.

Ribolzi, O., Zimmermann, R., Thammahacksa, C., Rattanavong, S., Oliva, P., Sengtaheuanghoung, O., and Pierret, A.: *Escherichia coli* concentrations and physico-chemical measurements at the outlet of 29 catchments of the Mekong river basin, Lao PDR, during dry and rainy seasons (2016) [Data set], https://doi.org/10.23708/ZRSBM4, 2021c.

Ribolzi, O., Causse, J., Thammahacksa, C., Latsachack, K., Huon, S., Henry-Des-Tureaux, T., Sengtaheuanghoung, O., Sipaseuth, N., and Pierret, A.: *Escherichia coli* concentrations and physico-chemical measurements (2011) along a crosssectional profile of the Nam Khan river, Mekong river basin, northern Lao PDR [Data set], https://doi.org/10.23708/RNY0LD, 2022.

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

Rochelle-Newall, E. J., Ribolzi, O., Viguier, M., Thammahacksa, C., Silvera, N., Latsachack, K., Dinh, R. P., Naporn, P., Sy, H. T., Soulileuth, B., Hmaimum, N., Sisouvanh, P., Robain, H., Janeau, J.-L., Valentin, C., Boithias, L., and Pierret, A.: Effect of land use and hydrological processes on *Escherichia coli* concentrations in streams of tropical, humid headwater catchments, Sci. Rep., 6, 32974, https://doi.org/10.1038/srep32974, 2016.

Sabo, J. L., Ruhi, A., Holtgrieve, G. W., Elliott, V., Arias, M. E., Ngor, P. B., Räsänen, T. A., and Nam, S.: Designing river flows to improve food security futures in the Lower Mekong Basin, Science, 358, eaao1053, https://doi.org/10.1126/science.aao1053, 2017.

Santini, W., Camenen, B., Le Coz, J., Vauchel, P., Guyot, J.-L., Lavado, W., Carranza, J., Paredes, M. A., Pérez Arévalo, J. J., Arévalo, N., Espinoza Villar, R., Julien, F., and Martinez, J.-M.: An index concentration method for suspended load monitoring in large rivers of the Amazonian foreland, Earth Surf. Dynam., 7, 515–536, https://doi.org/10.5194/esurf-7-515-2019, 2019.

Shrestha, B., Maskey, S., Babel, M. S., van Griensven, A., and Uhlenbrook, S.: Sediment related impacts of climate change and reservoir development in the Lower Mekong River Basin: a case study of the Nam Ou Basin, Lao PDR, Climatic Change, 149, 13–27, https://doi.org/10.1007/s10584-016-1874-z, 2018.

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

USGS: Dissolved oxygen: U.S. Geological Survey Techniques and Methods, book 9, chap. A6.2, version 2.0., 48 p., https://doi.org/10.3133/tm9A6.2, 2006.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., Abdollahi, M., Abdollahpour, I., Abolhassani, H., Aboyans, V., Abrams, E. M., Abreu, L. G., Abrigo, M. R. M., Abu-Raddad, L. J., Abushouk, A. I., Acebedo, A., Ackerman, I. N., Adabi, M., Adamu, A. A., Adebayo, O. M., Adekanmbi, V., Adelson, J. D., Adetokunboh, O. O., Adham, D., Afshari, M., Afshin, A., Agardh, E. E., Agarwal, G., Agesa, K. M., Aghaali, M., Aghamir, S. M. K., Agrawal, A., Ahmad, T., Ahmadi, A., Ahmadi, M., Ahmadieh, H., Ahmadpour, E., Akalu, T. Y., Akinyemi, R. O., Akinyemiju, T., Akombi, B., Al-Aly, Z., Alam, K., Alam, N., Alam, S., Alam, T., Alanzi, T. M., Albertson, S. B., Alcalde-Rabanal, J. E., Alema, N. M., Ali, M., Ali, S., Alicandro, G., Alijanzadeh, M., Alinia, C., Alipour, V., Aljunid, S. M., Alla, F., Allebeck, P., Almasi-Hashiani, A., Alonso, J., Al-Raddadi, R. M., Altirkawi, K. A., Alvis-Guzman, N., Alvis-Zakzuk, N. J., Amini, S., Amini-Rarani, M., Aminorroaya, A., Amiri, F., Amit, A. M. L., Amugsi, D. A., Amul, G. G. H., Anderlini, D., Andrei, C. L., Andrei, T., Anjomshoa, M., Ansari, F., Ansari, I., Ansari-Moghaddam, A., Antonio, C. A. T., Antony, C. M., Antriyandarti, E., Anvari, D., Anwer, R., Arabloo, J., Arab-Zozani, M., Aravkin, A. Y., Ariani, F., Ärnlöv, J., Aryal, K. K., Arzani, A., Asadi-Aliabadi, M., Asadi-Pooya, A. A., Asghari, B., Ashbaugh, C., et al.: Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, The Lancet, 396, 1204–1222, https://doi.org/10.1016/S0140-6736(20)30925-9, 2020.

WLE: Dataset on the Dams of the Irrawaddy, Mekong, Red and Salween River Basins. Vientiane, Lao PDR: CGIAR Research Program on Water, Land and Ecosystems - Greater Mekong, 2017.

Zimmermann, R. E., Ribolzi, O., Pierret, A., Rattanavong, S., Robinson, M. T., Newton, P. N., Davong, V., Auda, Y., Zopfi, J., and Dance, D. A. B.: Rivers as carriers and potential sentinels for *Burkholderia pseudomallei* in Laos, Sci. Rep., 8, https://doi.org/10.1038/s41598-018-26684-y, 2018.

Figures

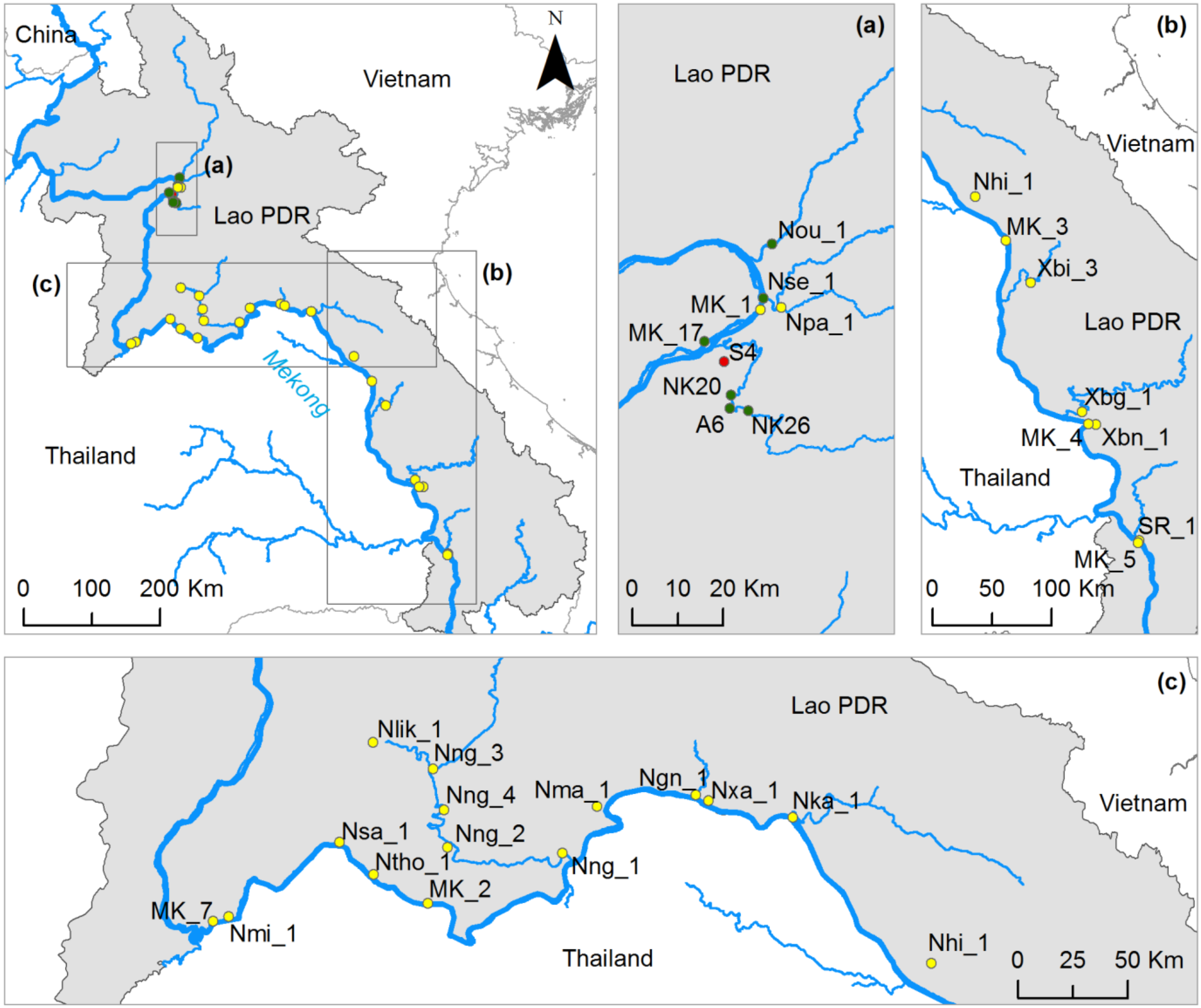

Figure 1: Location of the 31 sampling stations within the Mekong river basin in Lao PDR. Tributary names and catchment areas are given in Table 1. Red dot represents sampling station S4 and green dots represent sampling stations NK20, NK26, A6, Nou_1, Nse_1, and MK_17.

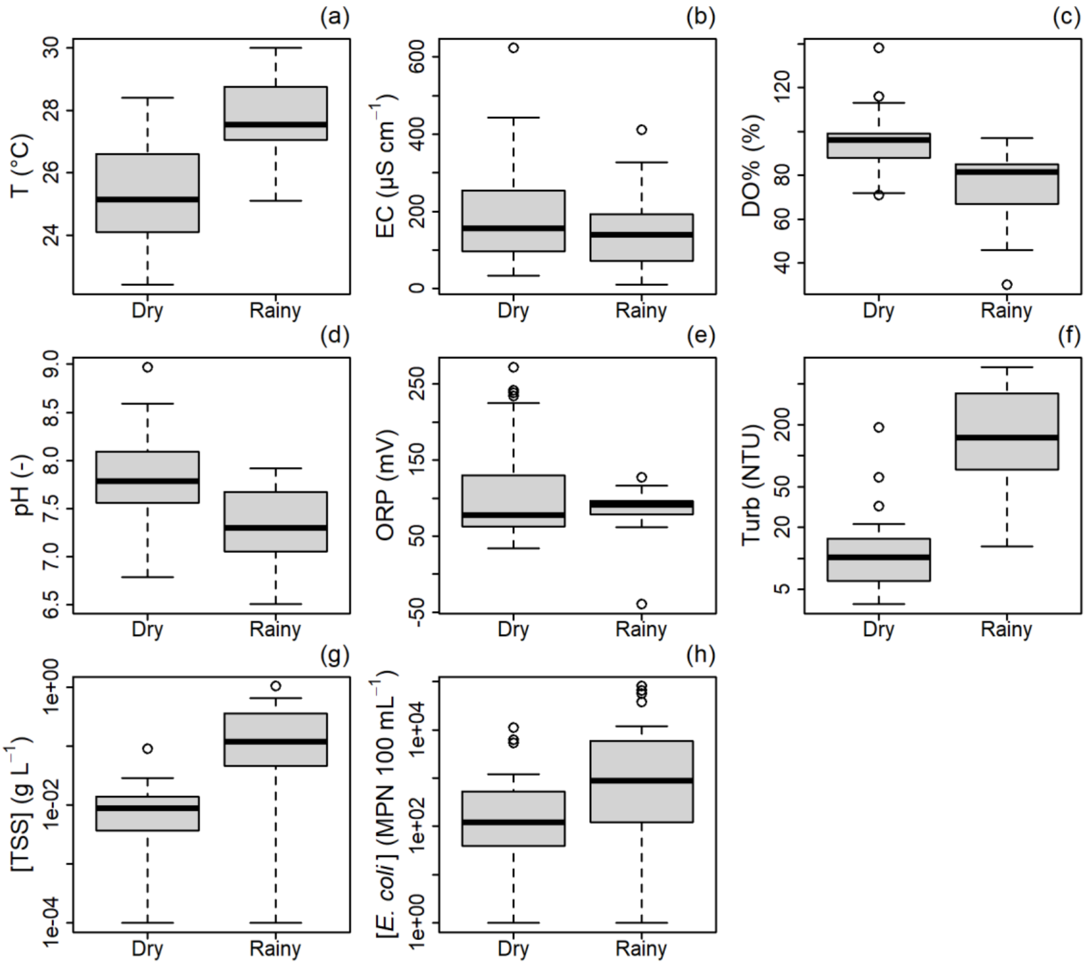

Figure 2: Stream water quality during dry and rainy seasons in 2016 along the Mekong river (6 sampling stations) and in 19 of its tributaries (23 sampling stations) in Lao PDR. T: temperature (°C); EC: electrical conductance at 25°C (µS cm^{-1}); DO%: oxygen saturation (%); pH: pH (-); ORP: oxidation-reduction potential (mV); Turb: turbidity (NTU); [TSS]: total suspended sediment concentration (g L^{-1}); [*E. coli*]: *Escherichia coli* concentration (MPN 100 mL^{-1}). Turb, [TSS], and [*E. coli*] are shown with Y axis as log scale. We added 0.0001 and 1 to all [TSS] and [*E. coli*] values, respectively, to present 0-values in a log scale, by convention.

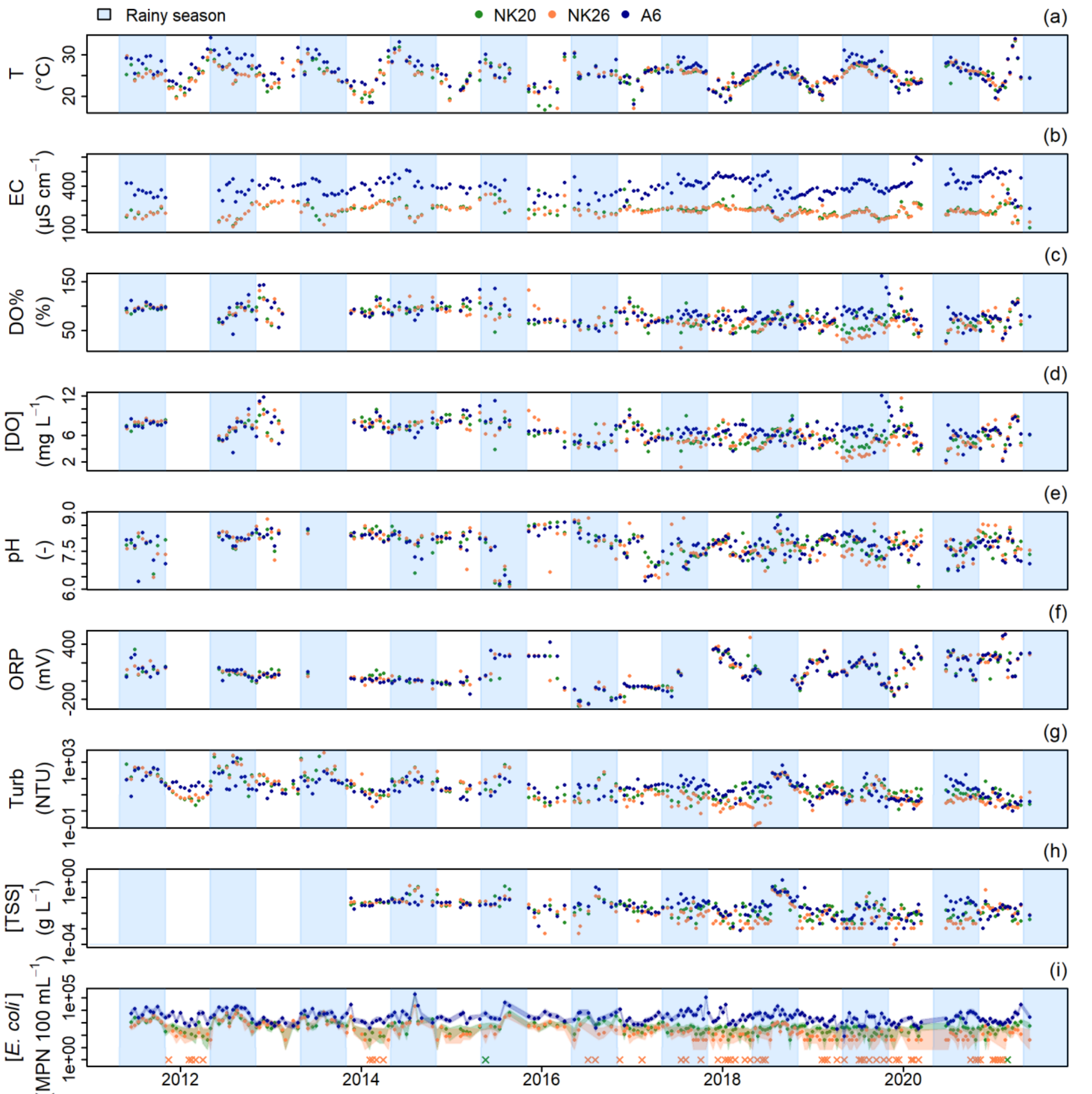

Figure 3: Stream water quality from 2011 to 2021 along the Nam Khan river (sampling stations NK20 and NK26) and its tributary Houay Khan (sampling station A6), northern Lao PDR. T: temperature (°C); EC: electrical conductance at 25°C (µS cm^{-1}); DO%: oxygen saturation (%); [DO]: oxygen concentration (mg L^{-1}); pH: pH (-); ORP: oxidation-reduction potential (mV); Turb: turbidity (NTU); [TSS]: total suspended sediment concentration (g L^{-1}); [*E. coli*]: *Escherichia coli* concentration (MPN 100 mL^{-1}) with lower and upper limits of the confidence interval given by Poisson distribution using the standardized microplate method. Turb, [TSS], and [*E. coli*] are shown with Y-axis as log scale. We added 0.0001 and 1 to all [TSS] and [*E. coli*] values, respectively, to present 0-

values in a log scale, by convention. Crosses for [*E. coli*] represent [*E. coli*] below the detection limit. Blue polygons represent the rainy season from May to October.

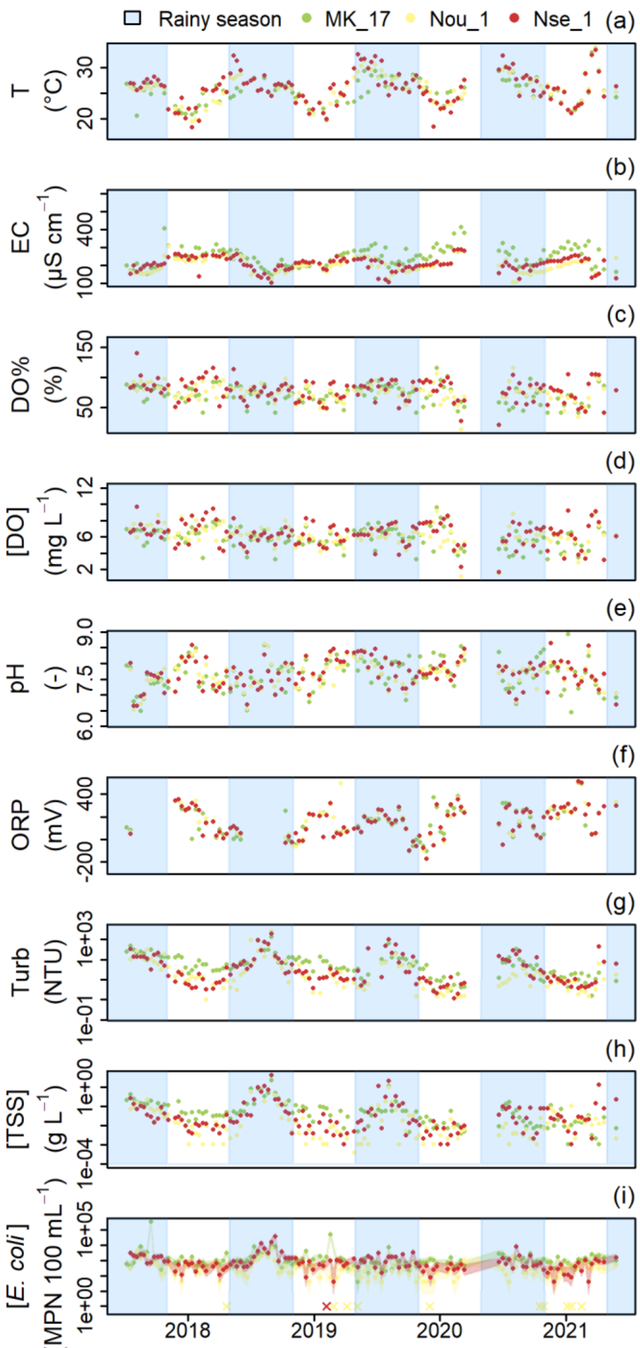

Figure 4: Stream water quality from 2017 to 2021 in the Mekong at Luang Prabang (sampling station MK_17) and its tributaries Nam Ou and Nam Seuang (sampling stations Nou_1 and Nse_1, respectively), northern Lao PDR. T: temperature (°C); EC: electrical conductance at 25°C (µS cm^{-1}); DO%: oxygen saturation (%); [DO]: oxygen concentration (mg L^{-1}); pH: pH (-); ORP: oxidation-reduction potential (mV); Turb: turbidity (NTU); [TSS]: total suspended sediment concentration (g L^{-1}); [*E. coli*]: *Escherichia coli* concentration (MPN 100 mL^{-1}) with lower and upper limits of the confidence interval given by Poisson distribution using the standardized microplate method. Turb, [TSS], and [*E. coli*] are shown with Y axis as log scale. We added 0.0001 and 1 to all [TSS]

and [*E. coli*] values, respectively, to present 0-values in a log scale, by convention. Crosses for [*E. coli*] represent [*E. coli*] below the detection limit. Blue polygons represent the rainy season from May to October.

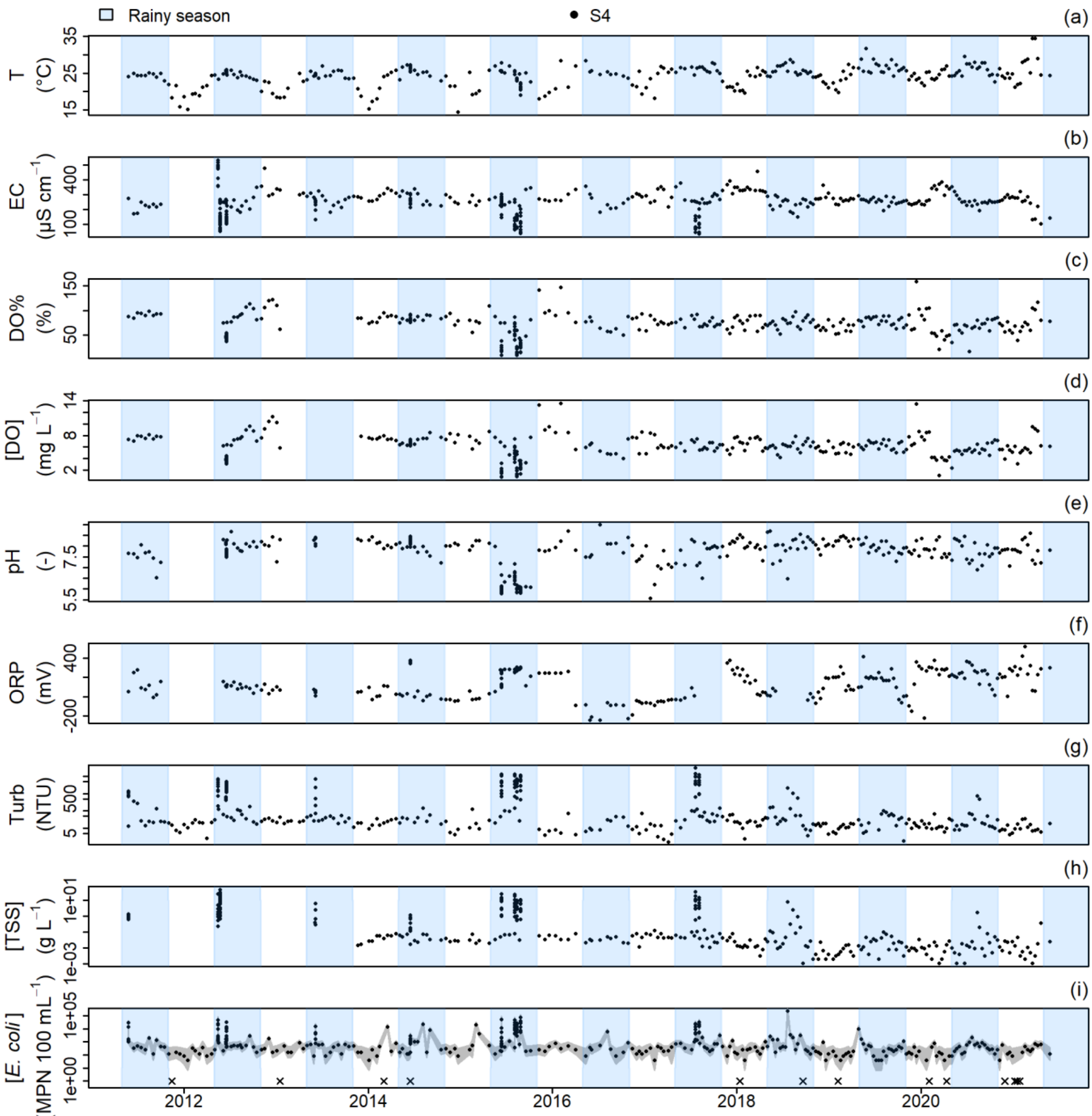

Figure 5: Stream water quality from 2011 to 2021 at the outlet of the Houay Pano catchment (sampling station S4), northern Lao PDR. T: temperature (°C); EC: electrical conductance at 25°C (µS cm^{-1}); DO%: oxygen saturation (%); [DO]: oxygen concentration (mg L^{-1}); pH: pH (-); ORP: oxidation-reduction potential (mV); Turb: turbidity (NTU); [TSS]: total suspended sediment concentration (g L^{-1}); [*E. coli*]: *Escherichia coli* concentration (MPN 100 mL^{-1}) with lower and upper limits of the confidence interval given by Poisson distribution using the standardized microplate method. Turb, [TSS], and [*E. coli*] are shown with Y axis as log scale. We added 0.0001 and 1 to all [TSS] and [*E. coli*] values, respectively, to present 0-values in a log scale, by convention. Crosses for [*E. coli*] represent [*E. coli*] below the detection limit. Blue polygons represent the rainy season from May to October.

Tables

Table 1: Description of the 31 sampling stations within the Mekong river basin in Lao PDR: sampling station name, river name, geographical coordinates of sampling station (i.e., latitude and longitude in degrees, WGS 1984), sampling period, and catchment drainage area in km².

Sampling station	River	Latitude (°)	Longitude (°)	Sampling period	Catchment area (km²)
S4	Houay Pano	19.85262	102.16912	2016, 2011-2021	0.6
NK26	Nam Khan	19.75536	102.21755	2011-2021	6,885
NK20	Nam Khan	19.78602	102.18336	2016, 2011-2021	7,236
A6	Houay Khan	19.76010	102.18112	2016, 2011-2021	239
Nou_1	Nam Ou	20.08642	102.26406	2016, 2017-2021	25,946
Nse_1	Nam Seuang	19.97931	102.24728	2016, 2017-2021	6,577
MK_17	Mekong	19.89267	102.13074	2017-2021	272,155
Npa_1	Nam Pa	19.96049	102.28289	2016	758
Nmi_1	Nam Mi	17.91917	101.68856	2016	1,021
Nsa_1	Nam Sang	18.22284	102.14222	2016	1,210
Ntho_1	Nam Thôn	18.09152	102.28159	2016	582
Nlik_1	Nam Lik	18.63280	102.28104	2016	3,022
Nng_3	Nam Ngum	18.52502	102.52631	2016	8,366
Nng_4	Nam Ngum	18.35581	102.57204	2016	14,318
Nng_2	Nam Ngum	18.20269	102.58588	2016	14,985
Nng_1	Nam Ngum	18.17879	103.05593	2016	16,841
Nma_1	Nam Mang	18.37019	103.19846	2016	1,793
Ngn_1	Nam Gniep	18.41756	103.60217	2016	4,564
Nxa_1	Nam Xan	18.39523	103.65408	2016	2,223
Nka_1	Nam Kadin	18.32517	103.99924	2016	14,820
Nhi_1	Nam Hin Boun	17.72699	104.56798	2016	2,152
Xbi_3	Xe Bang Fai	17.07782	104.98496	2016	9,433
Xbg_1	Xe Bang Hieng	16.09804	105.37625	2016	19,817
Xbn_1	Xe Bang Nouan	16.00290	105.47937	2016	1,351
SR_1	Nam Sedon	15.12390	105.80748	2016	7,225
MK_1	Mekong	19.95601	102.24113	2016	263,880
MK_7	Mekong	17.89870	101.62397	2016	295,246
MK_2	Mekong	17.97276	102.50410	2016	301,826
MK_3	Mekong	17.39714	104.79999	2016	373,368
MK_4	Mekong	16.00503	105.42449	2016	417,094
MK_5	Mekong	15.10721	105.79878	2016	549,055

Table 2: Description of the 18 column headings of the database files along with variables units.

Variable	Column heading	Description	Unit
	Outlet	Name of the sampling station	-
	LAT	Latitude, geographical coordinates of the sampling station	°
	LONG	Longitude, geographical coordinates of the sampling station	°
	River	Name of the river	-
	Date	Day of sampling	-
	Time	Time of sampling	-
	Date_Time	Day and time of sampling	-
T	T	Stream water temperature	°C
EC	EC	Stream water electrical conductance à 25°C	µS cm^{-1}
DO%	DOpercent	Stream water dissolved oxygen saturation	%
[DO]	DO	Stream water dissolved oxygen concentration	mg L^{-1}
pH	pH	Stream water pH	-
ORP	ORP	Stream water oxidation-reduction potential	mV
Turb	Turbidity	Stream water turbidity	NTU
[TSS]	TSS	Stream water total suspended sediment concentration	g L^{-1}
[*E. coli*]$_{LL}$	E-coli_4dilutions_95%-CI-LL	Lower limit of the confidence interval of the *Escherichia coli* concentration in water	MPN 100 mL^{-1}
[*E. coli*]	E-coli_4dilutions	Stream water *Escherichia coli* concentration	MPN 100 mL^{-1}
[*E. coli*]$_{UL}$	E-coli_4dilutions_95%-CI-UL	Upper limit of the confidence interval of the *Escherichia coli* concentration in water	MPN 100 mL^{-1}