Revista Internacional de Ciencia y Tecnología de la Información Geográfica
International Review of Geographical Information Science and Technology

Fernández Valencia, M. L., Soto Hidalgo, K. (2025). Geospatial ComSAS model to evaluate the presence of faecal *Enterococcus* as an environmental risk factor from non-point sources in creeks of the Rio Grande de Loiza, in Trujillo Alto Municipality, Puerto Rico. *GeoFocus*, *Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (Articles), 35, 47-65. https://dx.doi.org/10.21138/GF.846

GEOSPATIAL COMSAS MODEL TO EVALUATE THE PRESENCE OF FAECAL *ENTEROCOCCUS* AS AN ENVIRONMENTAL RISK FACTOR FROM NON-POINT SOURCES IN CREEKS OF THE RIO GRANDE DE LOIZA, IN TRUJILLO ALTO MUNICIPALITY, PUERTO RICO.

[1a]María de Lourdes Fernández Valencia, **[2b]Keyla Soto Hidalgo**
[1]Universidad Politécnica de Puerto Rico. 377 Av. Juan Ponce de León, San Juan, 00918, Puerto Rico
[2]Departamento de Ciencias Biológicas, 12 Ave. Universidad STE. 1201, Universidad de Puerto Rico, Recinto de Rio Piedras, San Juan, 00925-2532
[a]lufeva@gmail.com, [b]keyla.soto@upr.edu

ABSTRACT

Water resource management and sanitation are global priorities, as established in the Sustainable Development Goals, as well as for the United States and its territories, including Puerto Rico, located in the Caribbean. A study was conducted using the Communities Without Sanitary Sewer Model, which was developed in a Geographic Information System, to identify creeks of higher risk areas in watershed in Puerto Rico Island. Surface water samples were taken to analyse the presence of faecal *Enterococcus*. Out of the results obtained, 8 out of 10 were positive and exceeded the regulatory values, that confirm the contamination of surface waters with pathogens associated with lack of sanitary infrastructure to collect wastewater, the presence of septic tanks, either through filtration, overflow, or direct discharge of wastewater from non-point sources from communities surrounding rivers. Continuous contamination downstream was found in areas where sanitary services are already provided, confirming that contaminants are transported through water flow. This study concludes that the quality of surface waters in these areas is poor and poses a risk to public health, as well as the aquatic life of river ecosystems.

Keywords: communities without sanitary sewer; ComSAS model; *Enterococcus*; watershed; septic water; domestic discharges; Puerto Rico.

MODELO GEOESPACIAL COMSAS PARA EVALUAR LA PRESENCIA DE ENTEROCOCOS FECALES, COMO FACTOR DE RIESGO AMBIENTAL, DE FUENTES NO PUNTUALES EN ARROYOS DEL RÍO GRANDE DE LOÍZA, EN EL MUNICIPIO TRUJILLO ALTO, PUERTO RICO

RESUMEN

El manejo y saneamiento del recurso agua es una prioridad mundial conforme se establece en los objetivos de desarrollo sostenible, así como, para los Estados Unidos y sus territorios, incluyendo a

Received: 15/02/2024 Accepted: 26/02/2025
Academic Editor: Dr. Miquel Ninyerola www.geofocus.org

Puerto Rico, que se ubica en el Caribe. Se realizó un estudio utilizando el Modelo de Comunidades Sin Alcantarillado Sanitario, desarrollado en un Sistema de Información Geográfica, para identificar las quebradas de mayor riesgo en cuencas hidrográficas en la Isla de Puerto Rico. Se tomaron muestras de aguas superficiales para analizar la presencia de Enterococos fecales. Los resultados mostraron que 8 de 10 fueron positivos y excedieron los valores reglamentarios, lo que confirma la contaminación de las aguas superficiales con patógenos asociados a la falta de infraestructura sanitaria para recolectar aguas residuales, la presencia de fosas sépticas ya sea por filtración, desbordamiento o descarga directa de aguas residuales de fuentes no puntuales en de las comunidades aledañas a los ríos. Se encontró contaminación continua aguas abajo, en áreas donde se brinda servicio sanitario, lo que confirma que los contaminantes son transportados a través del flujo de agua. Este estudio concluye que la calidad de las aguas superficiales en estas áreas es deficiente y representa un riesgo para la salud pública, así como para la vida acuática de los ecosistemas fluviales.

Palabras Clave*:* comunidades sin alcantarillado sanitario; modelo ComSAS; Enterococos; cuenca hidrográfica; agua séptica; descargas domésticas; Puerto Rico.

1. Introduction

The environmental issue of water sanitation and water resource management has been a global priority for various international organizations. The United Nations (UN, 2024a) has outlined action plans since the first Earth Summit in 1972 in Johannesburg, in which it declared sustainable development as the guiding principle for long-term global development and established guidelines for the comprehensive management of water resources. It is reaffirmed in its master plan with the objectives and goals for sustainable development of the UN 2030 agenda (2024b), which sets out objective number six to ensure the availability of water, its sustainable management, and sanitation for all, to achieve equitable access to these services and end open defecation. The United Nations Development Programme (UNEP, 2025) has stated that non-existent or insufficient infrastructure for clean water and sanitation poses a serious risk to public health, increasing infection rates of diseases like cholera, malaria, and diarrhoea.

According to Ferro (2017), Latin America and the Caribbean have made progress in incorporating sustainable development goals, to access water resources, sanitation, and wastewater treatment; however, rural and marginalized areas have less than 30 % access to these services, and it continues to be a challenge in the 2030 agenda.

On the other hand, the United States of America (USA) set a global precedent in environmental protection with its Environmental Policy of 1970, affecting its territories, including Puerto Rico and others. The USA declares that water quality is essential to guarantee the environmental health of ecosystems and human life. The Federal Environmental Protection Agency (EPA) regulates this natural resource through the Federal Clean Water Act (CWA). This act establishes federal standards, requirements, and programs to restore and maintain the chemical, physical, and biological integrity of U.S. waters (EPA, 2019).

Watersheds' quality of water depends on the influx of rivers and creeks. Septic tank malfunction, with overflow, infiltration, or direct discharges of non-point sources to water bodies, increases microbial contamination that affects human health, the environment, and decontamination processes. If the resource is contaminated, more exhaustive water treatment must be provided to comply with the federal Safe Drinking Water Act (SDWA) and distribute it to the population as safe and potable water.

In a historical context, Vasquez *et al.* (2020) indicated that the father of epidemiology, John Snow (1813-1858), correlated cholera deaths with a sewage drain as a vector agent of the outbreak, contributing to medicine and developing the first map that represented a public health phenomenon in a spatial context, observing the incidence of deaths with the place of residence and the source of contamination. Advances in Geographic Information Systems (GIS) technology began in the 1960s with research in the academic community, and in 1969 Jack Dangermond and his wife Laura founded the Environmental System Research Institute (ESRI) to create computer maps and spatial analysis, to assist planners and land resource managers in making informed decisions. In 1981 it was commercialized because of the improvements in the system, and over time, it has evolved to provide the opportunity to create maps, collect data, and solve real-world problems, while also allowing data to be shared and the creation of global databases (ESRI, 2022).

Improvements in GIS have given rise to geospatial models, as discussed by Miller (2019), which are processing tools that generate data derived from one or more existing data sets. One of the most powerful features of a geospatial processing model is that the output of one function can be the input of another function. Geospatial processing models assess the context of a study area concerning the suitability or vulnerability to a particular set of land uses and management strategies.

It is a fact that GIS is a useful tool for managing environmental risk and the geographic location of the aspect under analysis. Its data analysis potential, superposition shapefiles, and other program analyses allow for a preliminary diagnosis of the study area. The scope and accuracy depend on the quantity and quality of the data collected in the generation of the GIS (Fernandez-Valencia *et al* 2022).

The Communities Without Sanitary Sewers Model ComSAS (*Comunidades Sin Alacantarillado Sanitario*, ComSAS, for its acronym in Spanish) was developed in GIS using the available data. ComSAS was designed as a geospatial tool to prioritize environmental risk factors related to areas without Sanitary infrastructure, and to facilitate decision-making in the process of compliance with the CWA in the Rio Grande de Loiza (RGL) watershed in Puerto Rico (PR). The model risk map correlated the criteria of density of structures/km^2, proximity to rivers, land use classification and the presence of type D water soils, resulting in the geographic information layer, which identified that 27 % of the area of the basin lacks sanitary service, and was identified as high and very high risk areas. Keeping that in mind, this methodology has the potential to be applied to other watersheds and research, if GIS data of their local governments is available (Fernández-Valencia *et al*. 2023).

According to data from the two-year integrated report of the Department of Natural and Environmental Resources of PR (DNER, 2023) monitoring network, in the RGL basin there was water contamination by pathogens, from dispersed sources of domestic origin, which cause non-compliance with the parameters for the bacteriological indicator: " Surface waters intended for use as a raw source of public water supply, propagation and maintenance of desirable species, including threatened or endangered species, as well as primary and secondary contact recreation". The compliance standard for "enterococcus density, in terms of geometric mean shall not exceed 35 colonies/100 mL in any 90-day interval: neither the 90th percentile of the samples taken shall exceed 130 colonies/100 mL in the same 90-day interval". EPA's National Drinking Water Regulation (2022) has set maximum levels of contaminants and risk to human health effects, from exposure above allowable limits. It is clarified that the criteria for pathogenic *Enterococci* may also be referred to as faecal *Enterococcus*, and the units can be expressed as colony forming units (CFU) or in Most Probable Number (MPN), and one CFU/100 mL is equal to 1MNP/100 mL.

The presence of faecal bacteria is indicative of water contamination and can cause diarrhoea, cramps, nausea, headaches, or other symptoms. On the other hand, the Centers for Disease Control and Prevention (CDC, 2022) warns that "workers who manage human waste or wastewater may be at increased risk for waterborne diseases." According to Said *et al*. (2022), faecal *Enterococci* are found in the soil, water, food, sewage, plants, human skin, the oral cavity, and the large intestine, constituting less than 1 % of the total microbiota. It is the third most common cause of community-acquired endocarditis in North America and causes 15 to 20 % of urinary tract infections in the hospital setting. Intra-abdominal collections and peritonitis are associated with those pathogens.

Given the importance of water resource management at a global level, and the usefulness of GIS to conduct geospatial analysis, the ComSAS Model allowed for to identification of the greatest risk areas with no sanitary sewerage system, and apply it as a tool to select sampling water points or other investigations. This article discusses the methodology, findings, and conclusions of the research carried out in Puerto Rico, in the Río Grande de Loíza watershed, at the Municipality of Trujillo Alto. Water samples were collected and analysed to determine water quality and the presence of pathogen contaminants associated with waste water sources without sanitary infrastructure. Under this premise, the hypothesis and objectives of the study were established. Despite being local research, it can be replicated in other places with ComSAS sectors.

1.1. Hypothesis

If water samples are collected in a creek of high-risk ComSAS areas, then, there will be presence concentrations of Enterococcus faecalis above regulatory standards.

1.2. General Objective

The general objective of the research was to evaluate the presence and concentration of in water Enterococcus faecalis bodies in water adjacent to communities without sanitary sewerage and compare it with areas that have sanitary sewerage.

1.2.1. Objectives

A) Use ComSAS Model to identify risk areas at creeks adjacent to the communities without sanitary service to collect samples.

The ComSAS Model was used to select risk area sampling points. For decades, pathogens have been evaluated by regulatory agencies in Puerto Rico, including the main sub-basins of the RGL, however, they have not yet been evaluated in the tributary creeks adjacent to the communities whose water flows into the river.

B) Determinate presence and concentration of faecal Enterococcus.

Elaborate a sampling plan map in GIS, and take samples to determine the presence and concentration of faecal *Enterococcus* on selected areas and compare with regulatory standards. Three samples were taken on ComSAS high risk area, and two samples down river on areas with sanitary sewers during December 2023 and January 2024.

C) Correlate faecal Enterococcus *with presence of structures in ComSAS areas.*

GIS allows multi-criteria analysis to be conducted, to determine risk and vulnerability, and make decisions in an effective, visual, and sustained way.

D) Correlate faecal Enterococcus *with epidemiological human health effects.*

Identify the results of food- and/or water transmissible diseases of the Department of Health of PR and correlate with the priority system for clean water state revolving fund projects criteria of critical health problems.

2. Materials, data, and methods

2.1 Case Study: faecal *Enterococcus* sampling in Puerto Rico, on RGL watershed at Trujillo Alto municipality.

Puerto Rico (PR) is an island located in the Caribbean (see Figure 1), and it is a US territory. The island has 54 major watersheds, the largest one is Rio Grande de Loiza (RGL) with an area of 751 km^2 and a population of 494,366 inhabitants (SDC-PR 2020). RGL is a river that receives agricultural runoff, Waste Water Treatment Plant (WWTP) effluent (secondary treatment only), and septic system effluent (Laureano-Rosario *et. all* 2017). It has the Carraizo dam, from which 357,721 m^3 per day is extracted to be purified at the treatment plant located in the municipality of Trujillo Alto. The Caribbean EPA (2011) estimated that 65 % of the population in this basin lacks sanitary sewerage, and has septic tanks or direct discharges to water bodies, which represents a public health and environmental problem. These septic tanks discharge approximately 165 million gallons of water per day directly into streams entering coastal waters (Laureano-Rosario *et. all* 2017).

This situation affects the quality of water reservoirs in PR. Otherwise, the transportation and disposal of sanitary water, through trucks of private or municipal companies, to the WWTP represents another complication, in addition to the costs, each truck can extract the contents of one or two septic tank residences at the time, and transport them to the treatment plant of the PR Aqueduct and Sewer Authority (PRASA) located in San Juan, and Caguas Municipalities.

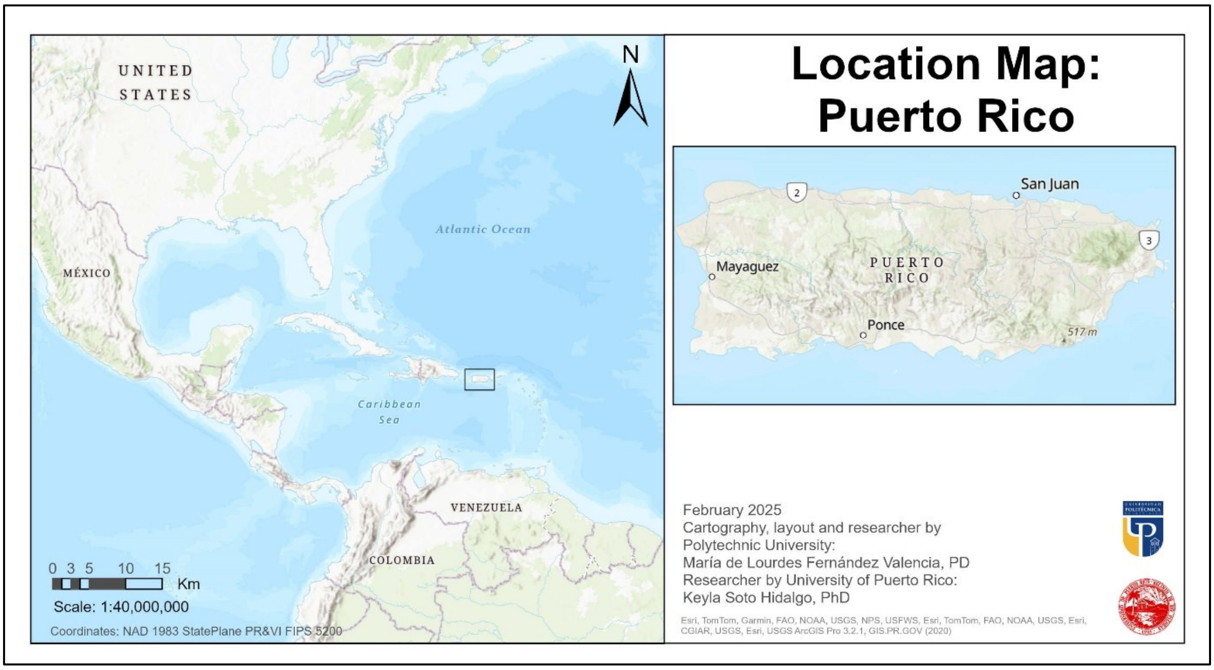

Figure 1. Location map: Puerto Rico

Study area selection was based on the findings of Fernández-Valencia (2023) research, which concluded that 79 % of the territorial area of the RGL basin does not have the infrastructure to collect used water, and generated a high-risk areas map using ComSAS Model in GIS. In the RGL watershed, those areas are located in the municipality of Trujillo Alto, PR. They are a hydric reserve with the presence of Loiza Lake, from which 350,000 m^3 of water is extracted for treatment and supply to the inhabitants of the metropolitan area.

The study area is in the Watershed of RGL, which is composed of 15 sub-basins, three of them, Loiza Lake, Quebrada Grande, and RGL-North segment, are part of the Trujillo Alto municipality, where creeks with high-risk ComSAS were selected to take samples. Loiza Lake is also located in this municipality, which is a source of water supply for part of the metropolitan area of the island. A location map can be appreciated in Figure 2, which includes the Watershed of RGL limit on the green, with principal rivers, and the sub-basins (in yellow colour) that are part of the territory of the municipality of Trujillo Alto (in pink colour) within the basin.

The study area is represented on Figure 3 with a map that includes Trujillo Alto Municipality with a pink perimeter, and sub-basins of Lake Loiza, Quebrada Grande, and RGL North. The complex hydrology can be appreciated more closely by focusing on the sub-basins of Loíza Lake and Carraizo Creek, Quebrada Grande sub-basin with Infierno Creek, and a segment of the RGL North sub-basins, where samples of faecal *Enterococcus* were collected.

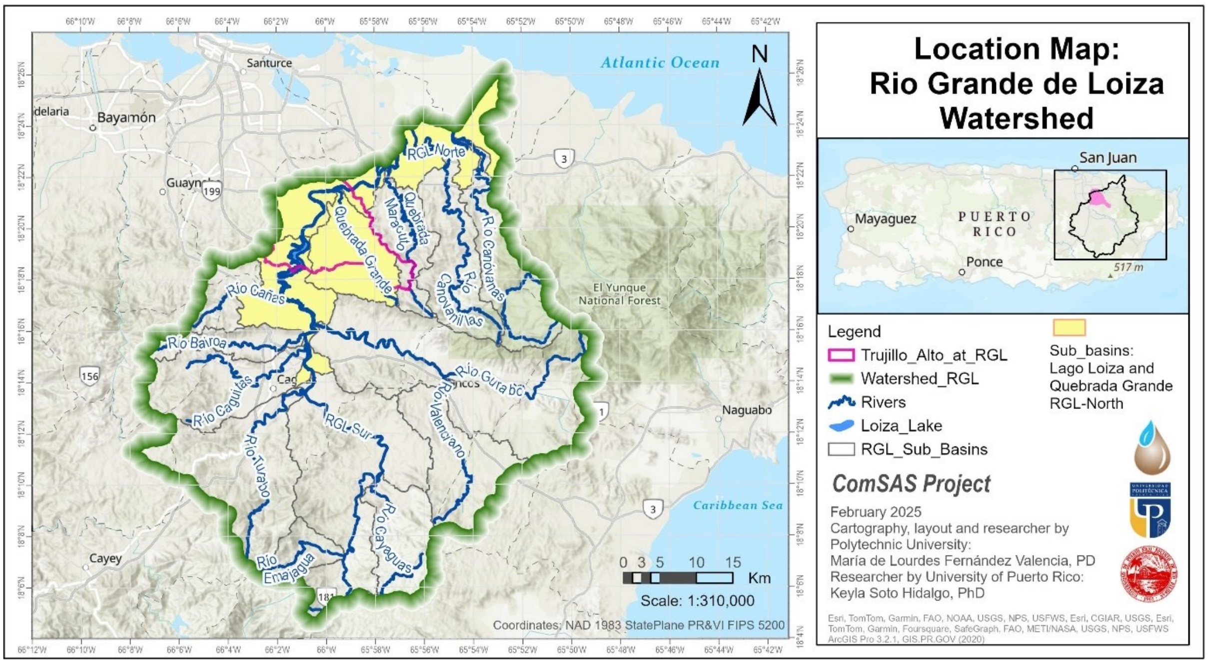

Figure 2. Location map: **Rio Grande de Loiza Watershed**

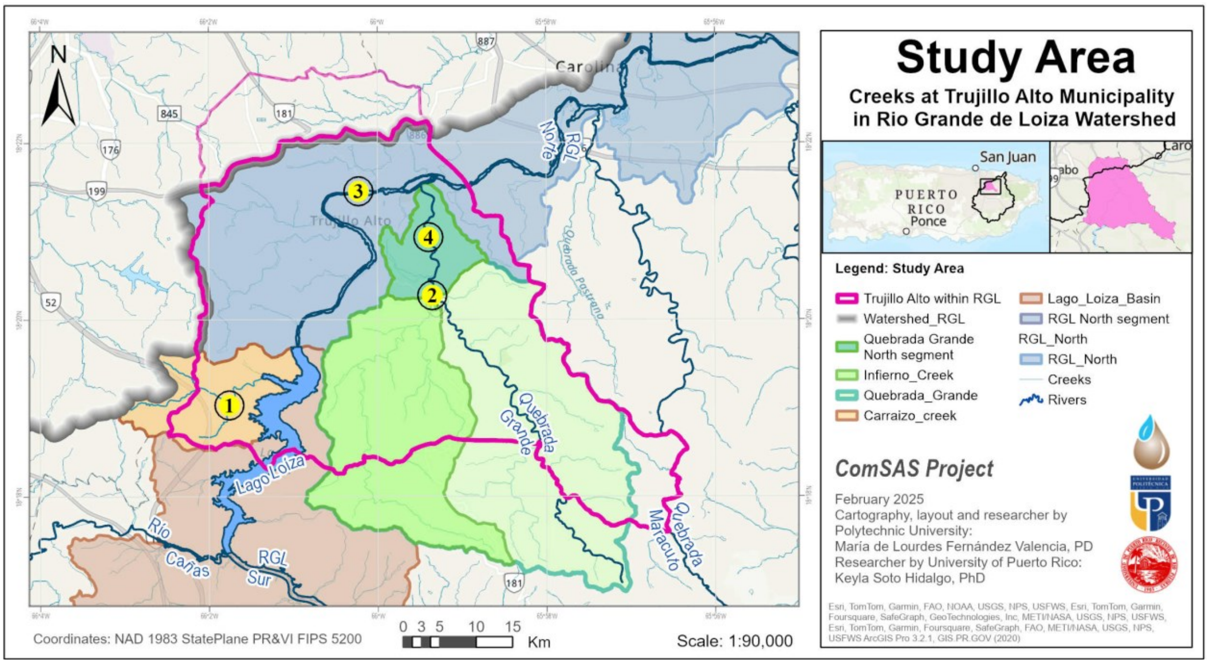

Figure 3. Sampling map.

2.2.1. Study area characteristics

Trujillo Alto municipality has 67,740 populations (SDC-PR, 2020) and a territorial extension of 54 Km^2, and 91 % of its territory is within the RGL watershed. The study area represents 76 % of the territorial extension of Trujillo Alto within the watershed. Starting from the CIG of topographic contours and rivers, the streams and segments of the sub-basins corresponding to the study areas were delimited.

The territorial extension areas of selected creeks are described in Table 1. It is important to consider the total area of the stream or that flows into the sampling point as a water unit. According to the information available in the GIS, the topography ranges at the study area are 10 m to 310 m above mean sea level (mamsl), with the highest points (280 mamsl-320 mamsl) in the Quebrada Grande, and

Infierno Creek. Sampling points were at 80 mamsl the highest one at Carraizo Creek and the lowest one was at RGL TA North area at 10 mamsl (see table 1 and figure 4).

Table 1. Territorial Extension Areas.

Point	Creek area	Area km^2	% of Study Area	Elevation(mamsl)
1	Carraizo creek	5.22	13 %	80
2	Infierno creek	13.46	33 %	20
3	RGL TA North Area	19.65	48 %	10
4	Quebrada Grande North	2.93	7 %	15
	Total Study Area	**41.26**	100 %	

Source: Own source based on available data from GIS.PR.GOV (2020).

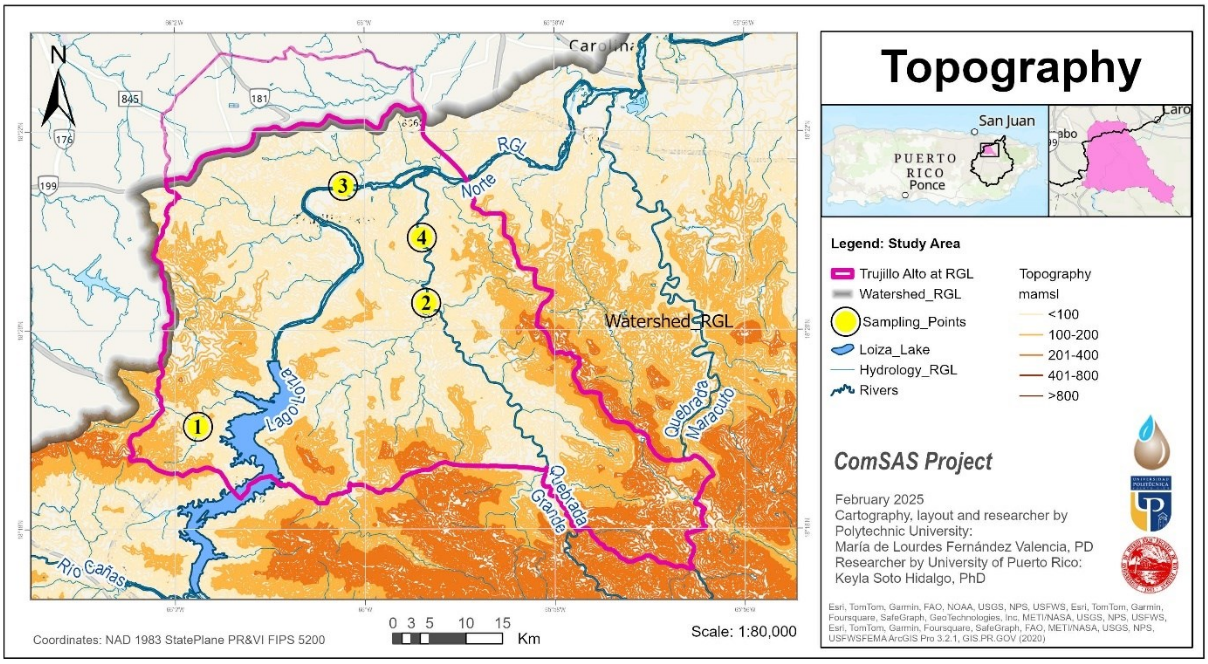

Figure 4. Topography map.

The current Land Use Plan from the Planning Board of Puerto Rico (PB, 2015) is illustrated in Figure 5, which highlights the Trujillo Alto municipality boundary in pink. The land use plan designates the areas around Loiza Lake as hydric reserves, classified as Special Rustic Protected Land (SRPL), including categories such as Hydric (H), Ecologic/Hydric (EH), and Water (which encompasses rivers, creeks, and the lake). Natural reserves categorized under SRPL as Ecologic (E) and Agricultural (a) land. Others are Common Rustic Land (CRL), which can be utilized for various purposes, including residential, commercial (such as schools, hospitals, shops, and offices), and industrial uses. Urban Land is represented in grey, alongside the area's vital infrastructure, with sanitary infrastructure, while the other land use zones in the selected municipality lack sanitary sewerage services. The study areas, marked by yellow points, are within the hydric reserve, CRL, and urban land zones. According to the PB (2015), this municipality does not have Planned Urban Land (PUL) or Not Planned Urban Land (NPUL).

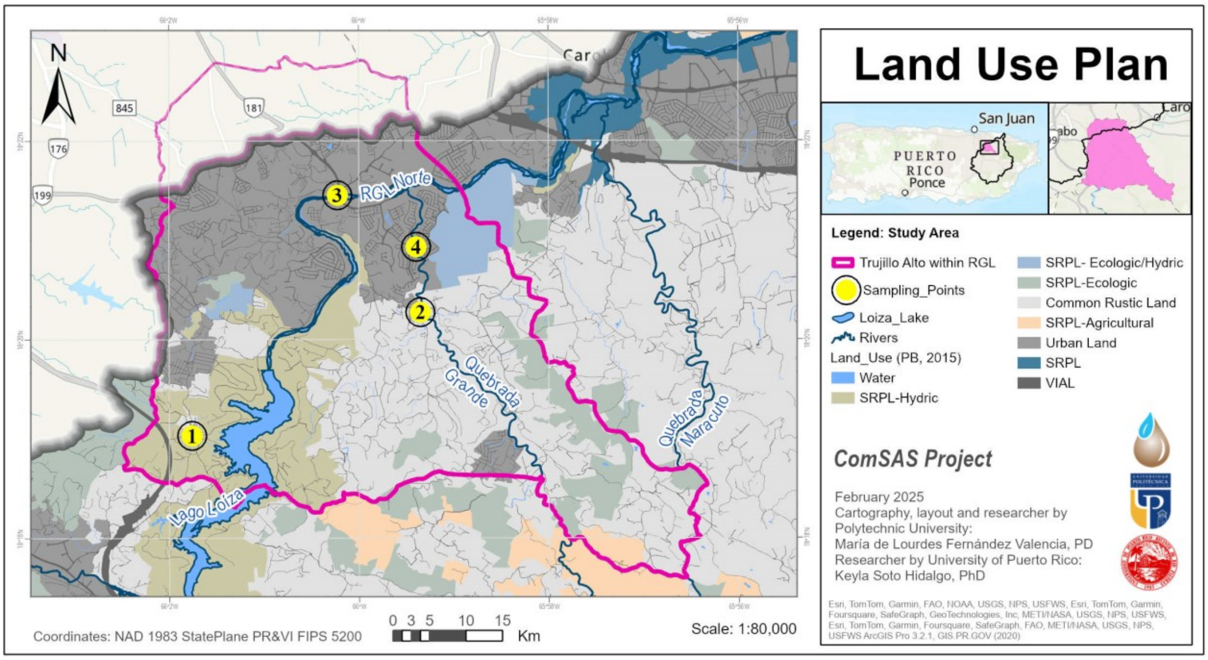

Figure 5. Land Use Plan map.

The study area is located on a hydric soil type D. According to National Resources Conservation Services (NRCS, 2009), those soil types have a very slow infiltration rate, which means they have a high runoff potential when completely wet. These consist primarily of clays that have a high shrinkage potential, soils that have a high-water table, soils that have a layer of clay or mud at or near the surface, and shallow soils over impermeable material. They have a very slow rate of water transmission. Soil type C has a slow infiltration rate when thoroughly wet. The flood zones, 0.2 % surround the lake and main rivers. Figure 6 shows hydric soil types: D in green colour, C in purple colour, and the flooding zones in orange colour, which coincide with the course of the rivers.

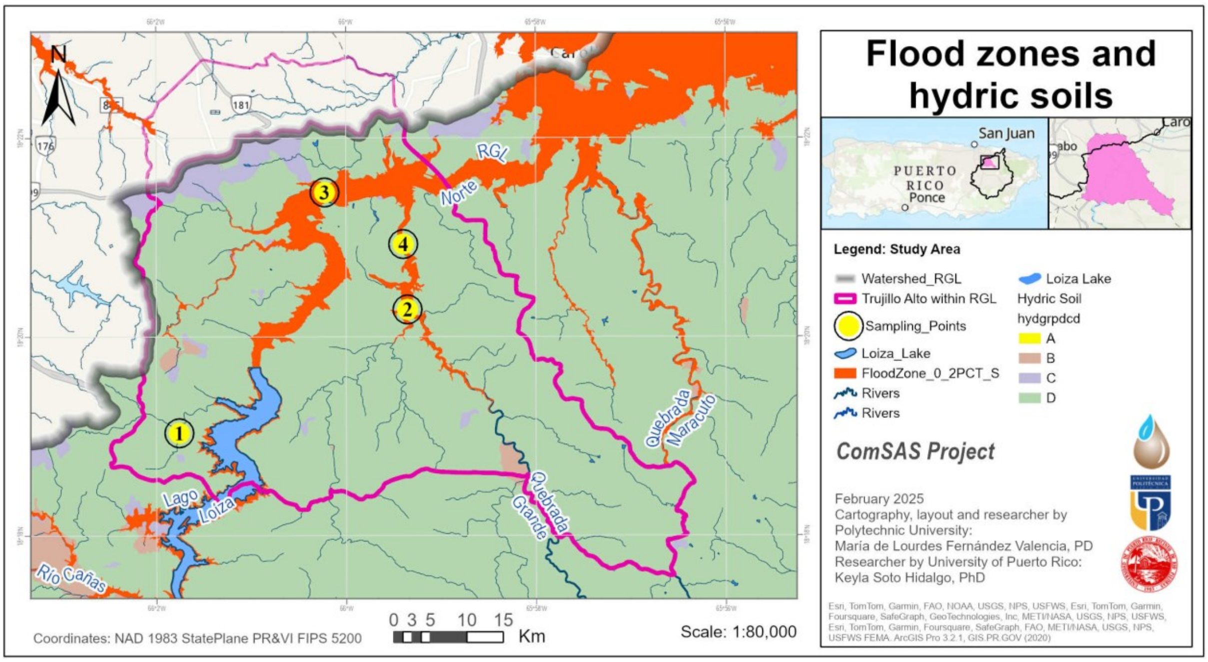

Figure 6. Flood zones and hydric soils map.

2.2. Data analysis and techniques

Data analysis included field samples collection, analysis in a certified laboratory, location points in GIS with ComSAS Model, and new shapefiles generated. Geospatial analysis was performed in GIS with ArcGIS Pro 3.2.1. using available data from the GIS.PR.GOV portal and data generated by the ComSAS Model (Fernández-Valencia, M.L., 2023).

2.2.1. Sampling Plan and ComSAS Model

Fernandez-Valencia (2023) developed a correlational matrix of risk factors and prioritization parameters as part of the ComSAS Model, which was the result of the federal and state regulatory framework analysis that allowed the identification of environmental factors based on the characteristics of the basin. It provides a GIS map with the risk area; therefore, it was used for the selection of criteria in the prioritization and correlation of risk factors of this research. This spatial model made it possible to select sampling points in very high-risk areas. Road access to creeks was crucial in selecting areas to collect samples safely. The sampling collection criteria included areas in ComSAS and areas with sanitary infrastructure service on the same rivers to compare results:

A) Samples collection at ComSAS high-risk areas corresponds to the Carraizo Creek located in the sub-basin of Loiza Lake, and the second to the Infierno Creek located in the sub-basin of the Quebrada Grande, both in the municipality of Trujillo Alto. In both streams, samples were collected biweekly for a total of 3 samples in each one.

B) Samples collection at areas with sanitary infrastructure service: One sample was taken downstream, where there is sanitary sewerage infrastructure, in the main river RGL, and another in the Quebrada Grande.

Table 2 includes the sampling plan and location:

Table 2. Sampling Plan.

Point	Basin Name	Location	Total Samples
1	Carraizo Creek	PR-175 highway intersection with PR-843, in the Carraizo suburb.	3
2	Infierno Creek	PR-181 intersection with PR-852 at La Gloria suburb.	3
3	Rio Grande de Loiza (RGL)	Downstream of the confluence of the Carraizo Creek on PR-181 at the historic bridge.	2
4	RGL TA Area	downstream of the confluence of the Quebrada Infierno at Via del Este road, at Encantada suburb.	2

In addition, one sample was taken at Punta Picuas WWTP from the affluent of the raw wastewater to compare the concentration of faecal *Enterococcus* present in used water vs. streams in the ComSAS areas. This affluent collects wastewater from 422 apartments. This sample reflected a concentration of 2,800,000 MPN/100 mL of faecal *Enterococcus* that was expected because of the inflow of raw wastewater.

2.2.3. Sample analysis at a certified laboratory.

The samples were taken directly from the rivers using a rope and a clean bucket for the direct extraction of water, and from there, the sample was extracted using gloves and transferred to a clean and sterile container provided by EqLab Environmental Quality Laboratories, Inc. It was kept preserved in a refrigerator at 4°C, and taken to a certified laboratory to perform analysis of surface water, with the analytical Standard Method 9230b faecal *Enterococcus* Multiple Tube Technique, approved by EPA as specified in the Code of Federal Regulation 40 CFR 141.402(c)(2).

3. Results analysis and discussion.

3.1. Objective (A). Use the ComSAS Model to identify risk areas at creeks adjacent to the communities without sanitary service to collect samples.

As described above, the selected area is located in the Trujillo Alto municipality on the RGL watershed. According to Fernández-Valencia (2023), the prioritization of risk levels in the ComSAS Model was the result of the correlation of the risk criteria of structures, land use, proximity to bodies of water, and hydrological analysis. The parameters used are described in Table 3, and the reference for the risk level types is shown in Figure 7, sampling plan map. For acronyms, refer to the Land Use Plan section and Figure 5.

Table 3. ComSAS Model: Risk factors and prioritization parameters.

Criteria of Risk Factors	Parameters to prioritize the level and type of risk				
	Very high (4)	**High (3)**	**Medium (2)**	**Low (1)**	**Very Low (0)**
1. Structures density/km^2	> 400	301-400	201-300	100-200	<100
2. Land Use Plan	Water SRPL (H, AH & EH)	CRL, UL	SRPL (A, AE), PUL, NPUL	SRPL(E, EA, EP)	Vial
4. Proximity to water body (meters)	≤ 400	401- 800	801- 1200	1201-1600	1600-2000
4. Hydrologic Analysis					
A) Runoff, Curve Number (CN)	> 80	61-80	41-60	21-40	<20
B) Hydrologic group	D	C, CD	B, BC	A	A

Fernández-Valencia, M.L. (2023).

The Sampling Plan Map is represented in Figure 7. The Trujillo Alto Municipality perimeter and extension within the RGL watershed are highlighted in pink. Sampling points selected in high-risk areas using the ComSAS Model are in yellow points with numbers 1 to 4 as described in Table 2. Point 1 is located at Carraizo Creek, point 2 at Rio Grande de Loiza at the historic bridge, point 3 at Infierno Creek, and point 4 at Quebrada Grande. The map allows us to appreciate the areas where the selected sampling points (one and two) are located in red areas that correspond to very high risk, while points three and four are in urban areas that do have sanitary infrastructure (Gray colour).

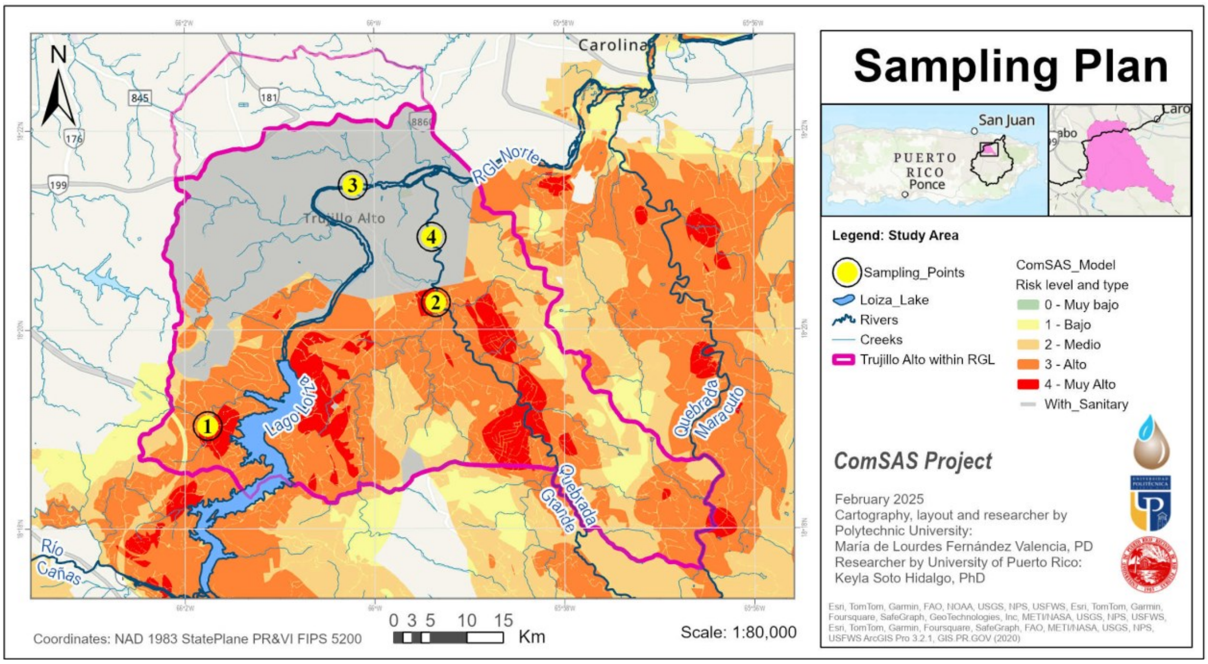

Figure 7. Sampling plan map.

Figures 8 to 11 were taken at the river to illustrate sampling points where surface water samples were taken. In the figures it is possible to appreciate the proximity of structures to the river. The photos were taken by the researchers. The samples were taken in places where a road bridge crossed the river or creek.

Figure 8. Sampling point 1: Carraizo Creek, suburb Carraizo, in Trujillo Alto, PR.

Figure 9. Sampling point 2: Infierno Creek, suburb La Gloria, in Trujillo Alto, PR.

Figure 10. Sampling point 3: Rio Grande de Loiza, at Historic Bridge, in Trujillo Alto, PR

Figure 11. Sampling point 4: Quebrada Grande, in Trujillo in Alto Municipality.

.

4.2. Objective (B). Determinate presence and concentration of faecal *Enterococcus*.

The presence of *Enterococcus* is evaluated by DNER to comply with CWA criteria. The U.S. Geological Survey (USGS) collects and analyses the data of the RGL and DNER takes them from Loiza Lake. This official data serve as a baseline to compare the samples collected data results, it is available at government websites.

The USGS (2024) includes the National Water Information System: Mapper with public access to data. In the study area, there are four monitoring stations, however, only station number 50059100, located in the RGL on the historic bridge of the municipality of Trujillo Alto, is the one that analyses samples for *Enterococci* (see figure 12). In addition, the DNER conducts special monitoring at Loiza Lake that includes three monitoring points (see table 4).

In 5 years, period from 2019 to 2023, a total of 38 samples of surface water were collected by authorities at Trujillo Alto, 27 of them complied and 11 were noncompliance samples with the CWA criteria. In the river station number 50059100, three to four samples were taken every year to totalize 19 samples, whose levels range between 1 to 2 800 CFU/100 mL of *Enterococci*. Nine of the 19 samples were above the regulatory standard. Otherwise in the lake, 19 samples were collected, with results levels ranging between 1 to 45.7 CFU/100 mL of this bacterium. Only two of the samples were above regulatory criteria. Lake sampling plan collected one to three samples every year except in 2020 because of pandemic event Covid-19, and 2022 because of the overgrowth of hyacinths on station 58800D. Results from federal and state monitoring station are in table 4, note that Lake Loiza in RGL station 5005910 were the highest ones.

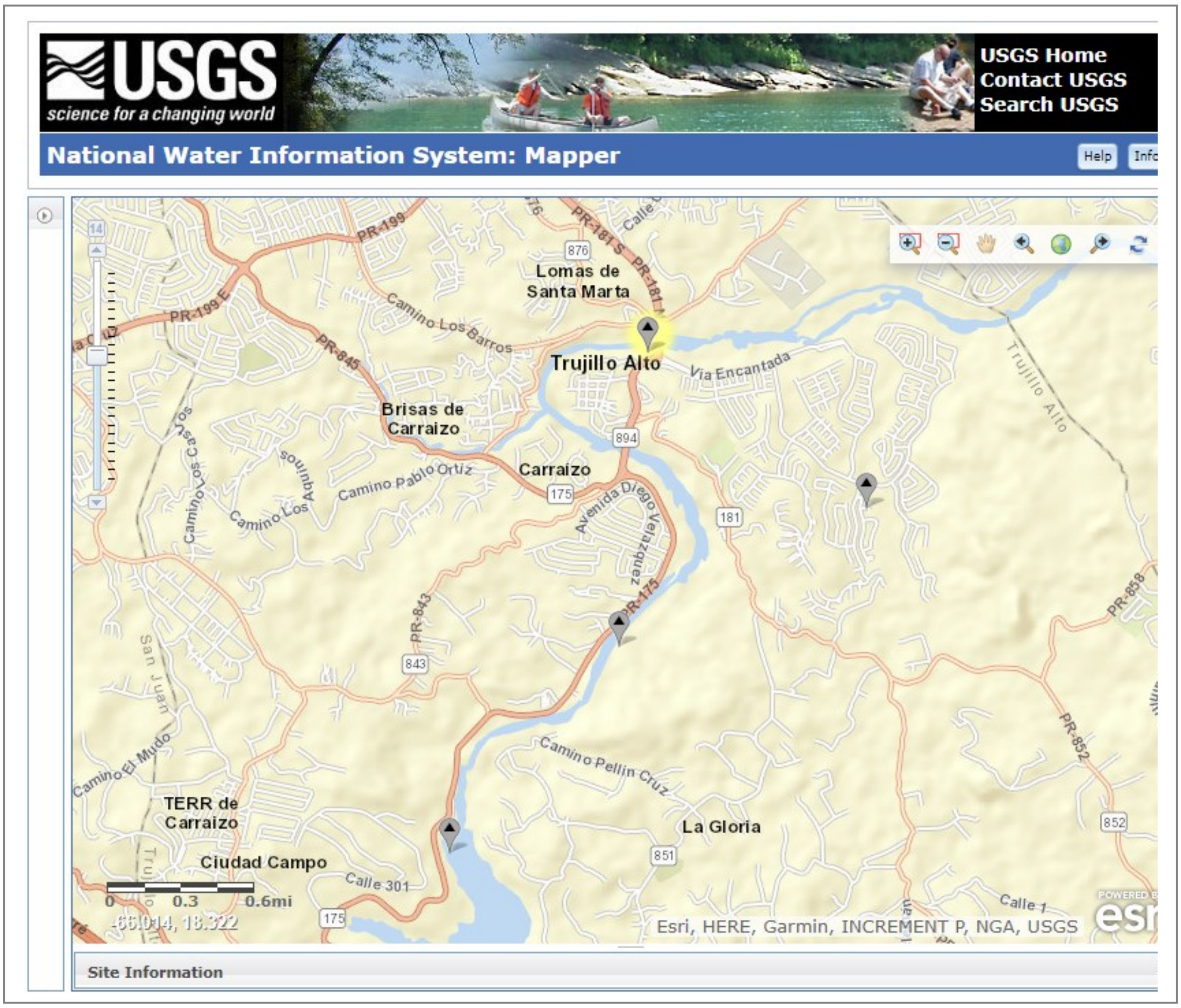

Figure 12. Monitoring Stations at Trujillo Alto municipality.
Source data USGS, 2020.

Table 4. Enterococci results of PR monitoring stations results from 2019 to 2023.

Station Id and Sampling years	Total Samples per year	Samples above 35 CFU/100 mL	Samples range CFU/100 mL
RGL 5005910			
2019	3	1	9 to 53
2020	4	1	10 to 40
2021	4	2	13 to 67
2022	4	3	1 to 2 800
2023	4	2	13 to 520
LAKE-57500			
2019	1	0	35
2020	0		none
2021	3	1	1 to 45.7
2022	2	0	15 to 17.1
2023	1	1	42.8
LAKE-58800			
2019	1	0	21
2020	0		
2021	3	none	2 to 6.3

2022	2	none	3 to 18
2023	1	none	9.7
LAKE-58800D			
2019	1	0	4
2020	0		
2021	3	none	1 to 3.1
2022*	0		Not Sampled
2023	1	none	2

Source: developed with public data retrieved from USGS (2020), DNER (2024), and 40 CFR131.40

As a part of this research, a total of 10 surface water samples were collected. The results only represent the location, date, and time they were collected. Figure 13 shows the bar chart with the presence of faecal *Enterococcus* in 8 of the 10 samples, whose levels ranged between 1 to 170, 000 MNP/100 mL. It was extremely high and exceeded the regulatory limits significantly. Results in ComSAS areas were as expected, above regulatory parameters except for the third one, which was below 1.8 MNP/100 mL. In the image, it is possible to observe a trend in exceeding the regulatory limit of 35 cfu/100 mL.

It was expected to obtain similar or lower values than upstream because those sampled points (3 and 4) are with sanitary infrastructure. The highest results were obtained at RGL point 3 with 170,000 cfu/100 mL, more than 4,000 times above the regulatory parameter. The results of the RGL samples are consistent with those reported by the PR government at station 5005910.

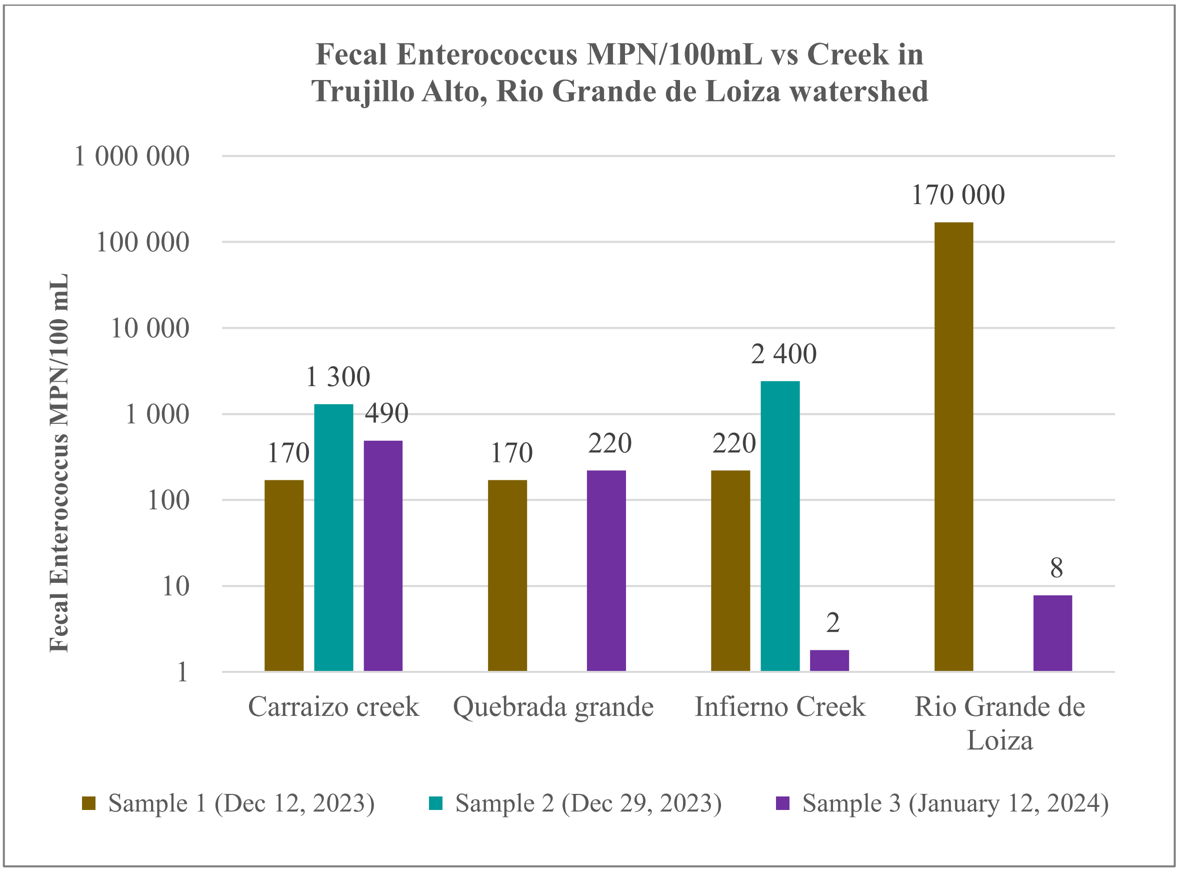

Figure 13. Fecal *Enterococcus* samples results graphic.

USGS (2024) station number 50059210 is in Quebrada Grande, in Encantada, Trujillo Alto. It corresponds to sampling point number 4 of the study. When observing the gage height and daily discharge data, it is seen that there is an increase in the month of December 2023 and decrease in

January. This confirms that the runoff carries pathogenic contaminants that are reflected in the results of the samples analysed (see figure 14).

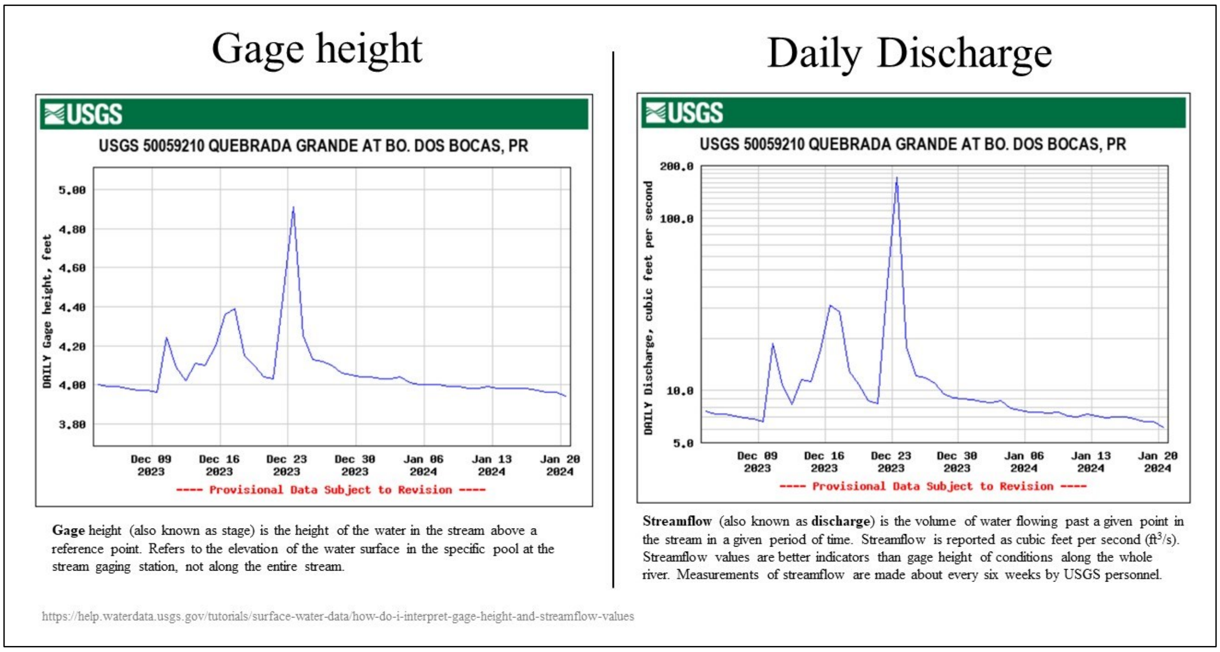

Figure 14. Gage height and daily discharge at Quebrada Grande Station (USGS, 2024).

4.3. Objective (C). Correlate faecal *Enterococcus* with presence of structures in ComSAS areas.

Structures within the creek areas were taken from the ComSAS Model. Each point represents a structure that may have a septic tank, mostly domestic; however, there are commercial zones with restaurants and other businesses that lack sanitary service from PRASA. In terms of planning, it is necessary to consider the sub-basins as hydrographic units, therefore, the results are expressed in terms of sub-basin and addition, which corresponds to the study area of Trujillo Alto Municipality. There is a total of 12 460 ComSAS structures in the selected watersheds, of which 69 % are in the study area. Carraizo creek has 8 % and Infierno creek has 25 % of total structures of the municipality. Table 5 includes total structures in selected creek areas with the representative percentile in main watersheds and within the Trujillo Alto municipality.

Table 5. Structures ComSAS at study area and main watersheds.

Sample point	Watershed area name	ComSAS Structures in study area	% ComSAS Structures main watersheds	% ComSAS Structures at Trujillo Alto
1	Carraizo creek	956	8 %	11 %
2	Infierno creek	3093	25 %	36 %
3	RGL North segment	2020	16 %	23 %
4	Quebrada Grande North	166	1 %	2 %
	Total structures	**6235**	50 %	72 %
Total structures main watersheds downstream		12 460	100 %	
ComSAS Structures at Trujillo Alto municipality		8611		
% ComSAS Structures at Trujillo Alto municipality		69 %		

Figure 15 illustrates the presence of ComSAS structures represented with wine-colored dots. In addition to 26 pumping stations, represented with orange dots, 13 of them are upstream of sampling point 3, in areas with sanitary sewers located in the segment of the RGL Norte subbasin.

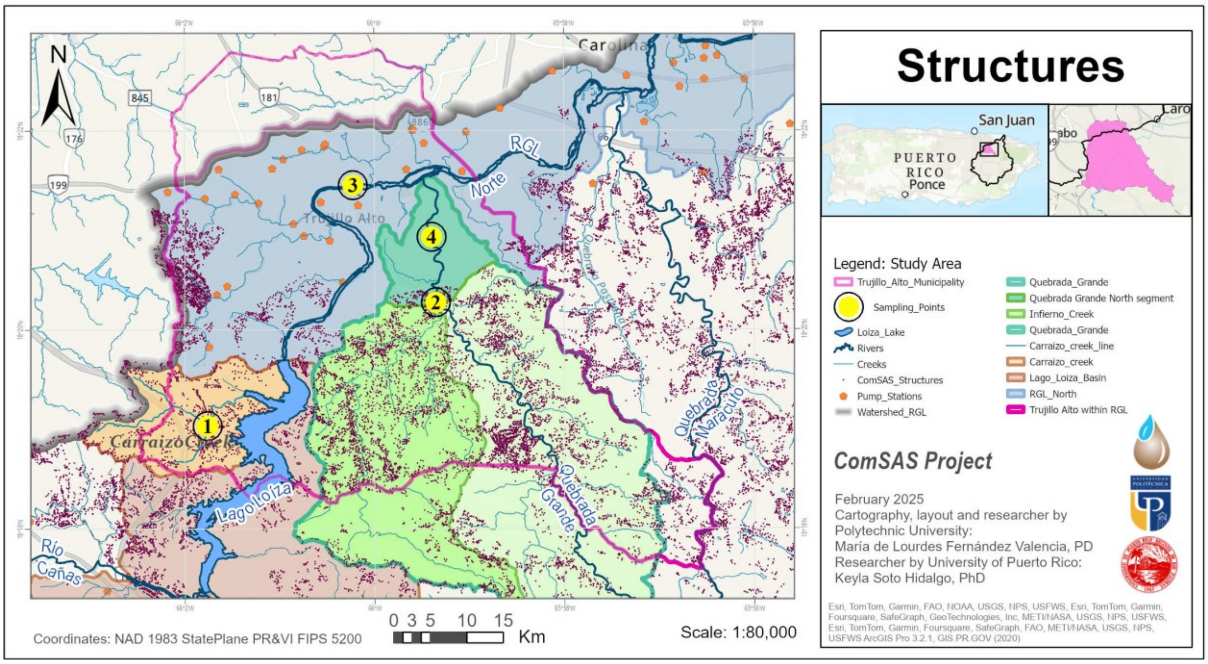

Figure 15. Structures in the study area map.

Once the structures were obtained, it was feasible to correlate with the averages of the analytical results of faecal *Enterococcus*. It was expected that the concentrations would decrease downstream; however, it was observed that the three samples from the Carraizo creek obtained a range of 170 to 1 300 MNP/100 mL, that means that, the creek is above the regulatory value and water quality is poor on it, and represents a severe health and environmental concern.

It is possible to observe on map of structures shown in figure 15, the following aspects: (1) the Carraizo creek flows into Lake Loiza near the Carraizo dam, and is 100 % ComSAS area; (2) Lake Loiza watershed receives the inflow by all the sub-basins of the RGL watershed from the south and southwest that flow into the lake; (3) The segment of the North RGL in Trujillo Alto, flows into the RGL and is composed of ComSAS areas at the south, and areas with sanitary sewers with pumping stations in north area. A significant increase in the concentration of the pollutant was observed between sampling points 1 and 2.

On the other hand, the three samples from the Quebrada Infierno were taken at ComSAS area, it obtained a range of 2 to 2 400 MNP/100 mL that means that creek was also above the regulatory value and the creek water quality is poor and represents severe health and environmental risk. Additionally, Quebrada Grande point 4 was above regulatory criteria. This river also has the inflow of the southeast area of the main watershed, which is 100 % ComSAS areas; it can be appreciated in figure 15. The results of Quebrada Grande reported 170 and 200 MNP/100 mL, in areas that have sanitary infrastructure, which would mean that the dilution of the contaminant has not occurred in that path of the river, and that runoff carries pollutants to the water on the river's flow from the ComSAS areas to the downstream sampling point 4.

4.4. Objective (D). Correlate faecal *Enterococcus* with epidemiological human health effects.

According to the Department of Health of PR (DH, 2022) it is estimated that half of people over the age of 65 who contract a foodborne illness, such as Salmonella, Campylobacter, Listeria or E. Coli, are hospitalized. 1 of every 7 children younger than 5 years of age who are diagnosed with E. Coli O157 infection have kidney failure. From the distribution of cumulative incidence transmissible diseases food and/or water by municipality of residence, epidemiological, that in the metropolitan region, corresponding to the municipalities Canovanas, Carolina, Guaynabo, Loiza, San Juan and Trujillo Alto, all of them except Guaynabo, are part of RGL Watershed. It was reported that there was a frequency of 195 cases (17.38 % from PR) from those, 7 (1.28 %) were in Trujillo Alto.

DNER (2019) considers a critical health problem in those areas with confirmed 10 to 20 % if incidence of water transmissible diseases. Under this premise, the island is in critical incidence, but not the municipality of Trujillo Alto.

5. Conclusions.

This study allowed for to validation of the usefulness of GIS. Specifically, the ComSAS Model was a valuable tool to select high-risk areas and design the sampling plan, to collect samples, and verify the presence of faecal Coliforms in the creeks adjacent to communities lacking sanitary sewers that are contaminating the water. Although the study was conducted in a sector of the municipality of Trujillo Alto, Puerto Rico, it has the potential to be replicated in other places in the world, using site-specific GIS data. It is important to include local and validated data, to relate the criteria of hydric reserve presence, hydric soil types, structure density, land use plan, and proximity to rivers on the geospatial correlation model.

In sampled areas, it is confirmed the presence of faecal *Enterococcus*, in very high concentrations, which indicates poor water quality on those creeks and the main river. The water quality is not safe for human use, public water supply, recreational activities, fish consumption, and represents a risk to water life species. As expected, the ComSAS areas resulted in very high concentrations above the regulatory level, which may be associated with septic wells that are infiltrating or overflowing sanitary wastewater into the body of water.

The samples taken from points downstream of the ComSAS areas were found to exceed regulatory levels, even though there are regions along their path that have sanitary infrastructure. The main watershed of RGL flows through both ComSAS areas and zones with sanitary facilities. Therefore, it is essential to investigate the sources of these high concentrations of *Enterococci*.

It is recommended that future studies include sampling in the RGL streams upstream of the historical bridge, as well as in all the streams within the ComSAS areas. This will allow us to correlate and compare results and help to identify the origins of the microbial contamination. Due to budget constraints, this research had a limited number of samples. A future study could implement a comprehensive one-year sampling plan, which would include collecting one sample from a creek that flows into the main river inlet before reaching the RGL.

The analytical results for faecal *Enterococcus* did not include genetic testing to differentiate between human and animal sources. The predominant sources may be human waste from septic tanks and runoff from domestic animals. Further studies should be conducted to ascertain that these point sources are indeed domestic septic tanks.

The high number of structures in 100 % ComSAS study areas could be linked to elevated concentrations of pathogens, indicating a significant environmental and public health concern, particularly since this is a water reserve designated for supply. Consequently, water extraction must undergo thorough treatment by PRASA before being supplied to the population.

It was not possible to correlate the sampling findings of faecal *Enterococcus* with DH data on the cumulative incidence of transmissible food and waterborne diseases by the municipality, as the data were provided for entire municipalities rather than specific areas, dates, or times. More research is necessary to conclude in this context. Future epidemiological studies could also be conducted to confirm the presence of *Enterococcus* or E. coli-related diseases in the population of Trujillo Alto municipality, as well as other research zones. The samples taken from points downstream of the ComSAS areas were found to exceed regulatory levels, even though there are regions along their path that have sanitary infrastructure. The main watershed of RGL flows through both ComSAS areas and with sanitary facilities. Therefore, it is essential to investigate the sources of these high concentrations of *Enterococci*.

6. Acknowledgments

We appreciate the support of the Polytechnic University of Puerto Rico (PUPR) with the Promoting Postbaccalaureate Opportunities for Hispanic Americans Program (PPOHA) Peer-to-peer collaboration fellowship; and the University of Puerto Rico (UPR). This project was financially supported by NSF-CREST Center for Innovation, Research, and Education in Environmental Nanotechnology (CIRE2N) Grant Number NSF-HRD-1736093 of the UPR, Rio Piedras Campus. In addition to Sani Plant Co. Inc., and Environmental Technical Advisory Group (ETAG) for support and assistance in this research.

References

CDC. (2022). *Handling Human Waste or Sewage.* Center for Disease Control and Prevention. Available at https://www.cdc.gov/healthywater/emergency/sanitation-wastewater/workers_handlingwaste.html

CFR, (2024). 40 CFR 131.40 *Code of Federal Regulat*ion. Available at https://www.ecfr.gov/current/title-40/chapter-I/subchapter-D/part-131

DNER. (2019). *Priority system for clean water state revolving fund projects*. Unpublished public document. Environmental Quality Board. Department of Natural and Environmental Resources. Government of Puerto Rico.

DNER. (2023). *Puerto Rico 2022 - 305(b) and 303(d) Integrated Report*. Department of Natural and Environmental Resources, Government of Puerto Rico. Available at https://www.epa.gov/system/files/documents/2023-10/pr-2022-305b-303d-ir.pdf

DNER. (2024). *Lakes network data IR 2024.* Unpublished public documents of the Department of Natural and Environmental Resources.

DH. (2022). *Enfermedades transmisibles por alimentos y/o agua* [Transmissible Diseases by food and/or water]. Department of Health. Government of Puerto Rico. Available at https://www.salud.pr.gov/CMS/DOWNLOAD/7123

EPA. (2011). Río la Plata and Río Grande de Loíza Watershed. (Final Report Phase III). *Watershed Stewardship Program*. US Environmental Protection Agency, Caribbean Division.

EPA. (2019). *Summary of the Clean Water Act*. US Environmental Protection Agency. Available at https://www.epa.gov/laws-regulations/summary-clean-water-act

EPA. (2022). *National Primary Drinking Water Regulation*. Environmental Protection Agency. Available at https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations#Microorganisms

ESRI. (2022). *What is GIS*. Environmental Systems Research Institute: https://www.esri.com/en-us/what-is-gis/history-of-gis

Fernandez-Valencia, M.L. (2023). *Modelo Geoespacial para Priorizar los Factores de Riesgo Ambiental de las Comunidades Sin Alcantarillado Sanitario en la Cuenca del Río Grande de Loíza en Puerto Rico* [Geospatial Model to Prioritize Environmental Risk Factors of Communities Without Sanitary Sewers in the Río Grande de Loíza basin in Puerto Rico]. [Doctorate thesis unpublished]. San Juan Puerto Rico: Universidad Internacional Iberoamericana Puerto Rico (Unib), and Mexico: Universidad Internacional Iberoamericana, México (UNINI).

Fernandez-Valencia, M. L., & Rivera-Rivas, M. C. (2022). *Modelo Geoespacial para priorizar los factores de riesgo ambiental de las las Comunidades Sin Alcantarillado Sanitario en la cuenca del Río Grande de Loíza en Puerto Rico* [Geospatial model to prioritize environmental risk factors of

Communities Without Sanitary Sewers in the Río Grande de Loiza basin in Puerto Rico]. *Umbral*, 183-209. Available at https://umbral.uprrp.edu/modelo-geoespacial-para-priorizar-los-factores-de-riesgo-ambiental-de-las-comunidades-sin-alcantarillado-sanitario-en-la-cuenca-del-rio-grande-de-loiza-en-puerto-rico/

Ferro, G. (2017). *Latin America and the Caribbean towards sustainable development goals in water and sanitation. Recent reforms of sectoral policies.* Natural Resources and Infrastructure Series. Introduction. United Nations. Economic Comision for Latin America and Caribbean. https://repositorio.cepal.org/bitstream/handle/11362/41136/1/S1700195_es.pdf

Laureano-Rosario, A.E., Symonds, E.M., Rueda-Roa, D., Otis, D., and Muller-Karger F.E. (2017). Environmental Factors Correlated with Culturable *Enterococci* Concentrations in Tropical Recreational Waters: A Case Study in Escambron Beach, San Juan, Puerto Rico. *International Journal of Environmental Research and Public Health*. 2017 Dec 19;14(12):1602. PMID: 29257092; PMCID: PMC5751019. https://doi.org/10.3390/ijerph14121602.

Miller, W. (2019). Introducing Geodesign: The Concept [Introducción al Geodiseño: El Concepto]. Environmental System Researchs Institute. http://www.esri.com/products/technology-topics/geodesign/~/media/files/pdfs/technology-topics/introducing-geodesign.pdf

NRCS. (2009). Hydrologic Soil Groups. In *Hydrology National Engineering Handbook*. US Department of Agriculture. Natural Resources Conservation Service. Chapter 7. Available at https://directives.nrcs.usda.gov/sites/default/files2/1712930597/11905.pdf

PB (2015). *Land Use Plan.* Planning Board of Puerto Rico. https://jp.pr.gov/plan-de-usos-de-terrenos/

PR.GOV. (2020). *Portal de datos geográficos gubernamentales* [Government Geographic Data Portal] Government of Puerto Rico. Available at https://gis.pr.gov/Pages/default.aspx

Said, M. S., Tirthani, E., & Lesho, E. (2 de May de 2022). Enterococcus Infection. In *StatPearls Publishing LLC.* National Library of Medicine, National Center for Biotechnology Information. Available at https://www.ncbi.nlm.nih.gov/books/NBK567759/

SDC-PR. (2020). *Censo decenal de población y vivienda* [Ten-year population and housing census]. State Data Center Puerto Rico. Available at https://censo.estadisticas.pr/EstimadosPoblacionales

UN. (2024a). Sustainable Development Goals. *17 Goals to transform our world.* United Nations. https://www.un.org/sustainabledevelopment/

UN. (2024b). Sustainable Development Goals. *History.* United Nations. https://sdgs.un.org/goals

UNEP (2025). *Water and sanitation.* United Nations Environment Programme. https://www.unep.org/explore-topics/resource-efficiency/what-we-do/cities/water-and-sanitation

USGS. (2024). *National Water Information System: Mapper*. US Department of the Interior, US Geological Survey. Available at https://maps.waterdata.usgs.gov/mapper/index.html

Vasquez, D., Murillo, L. and Iglesias, A. (2020). *Cholera*. Journal of Medicine. 42(2), 226-239 https://www.revistamedicina.net/index.php/Medicina/article/view/1517/1919