Medicon Agriculture & Environmental Sciences
Volume 8 Issue 6 June 2025
Article Type: Research Article
ISSN: 2972-2691

Metagenomics Study of Kamvari River and Varal Devi Lake in Bhiwandi, Maharashtra to Apply Nature Based Strategies to Human-induced Challenges of Untreated Sewage Discharge and Enhance the Climate Change Resilience

Snehal Donde[1]* and Upendra Muni Raval[2]

[1]*Fulbright SIR, US & Dean Administrative Affairs, Bhaktivedant Research Centre, Palghar, Mumbai, India*

[2]*Ecologist, India*

***Corresponding Author:** Snehal Donde, Fulbright SIR, US & Dean Administrative Affairs, Bhaktivedant Research Centre, Palghar, Mumbai, India.

Received: May 21, 2025; **Published:** June 28, 2025

DOI: 10.55162/MCAES.08.248

Abstract

India is suffering from the worst water crisis in its history with some 6000 million facing acute water shortage. This crisis will worsen as the demand is projected to be twice the available supply in 2030. The extent to which surface water and ground water is becoming polluted by anthropogenic activities including untreated sewage discharge, is adding to the complexities. Thus building understanding about the basic fundamentals of water, waterbodies as a bioreactor and role of microbial community are pertinent for the revival of lakes or riverine system. Prioritizing the nature based strategies this paper attempts to identify the characteristics, diversity and distribution of microbes in waterbodies to work on it, for an appropriate treatment mechanism to ensure its use as an alternative water resource. Metagenomics tool was first time used to study sewage dynamics in contaminated Varal devi lake and Kamvari river in Bhiwandi city, Thane dist. MS, as Varal devi lake supply 5 MLD potable water and may have adverse consequences. Findings of soil and water samples revealed potential environmental and health risk as the lake and river water showed presence of novel or less-characterized Leptospira and Turneriella species dominant in the system localized in Varal devi lake pumping and inlet discharge site. Hydrogenotrophic Methanogens were detected in soil samples, where as it was absent in water. A detailed microbiome taxonomic breakdown, abundance across samples, and insights into its ecological and health significance is comprehended with the recommendations to take appropriate measures of nature based strategies for ecological restoration of the lake and riverine ecosystem. Decentralized sewage treatment by septic tanks interception and constructed wetlands is suggested, simultaneously with Kamvari river rehabilitation from zero order stream onwards. Restoration of flood plains and buffer zones are also necessary by action on the encroachments, for self-regulation mechanism of the nature to prevail. This paper is an attempt to create awareness among the stakeholders about the waterbodies with river basin approach and strengthen scientific communication for good water governance. Engineering infrastructures like conventional STPs lack understanding and design of biological concepts and operationalization, this has led to the deterioration of almost all waterbodies. Hence it is imperative to work with integrated approach for future demand and water resources management by handholding between the community and government agencies. Disseminating the idea of decentralized community driven water conservation, river rejuvenation initiatives and to encourage use of advanced genomic technology to understand sewage degradation science for wastewater treatment, was the main goal of this research work.

Introduction

The conditions of the waterbodies across the country is worrisome and need immediate attention keeping in view the increasing demand and looming water scarcity. The reason being the poorly managed water distribution, wastewater management, concretization of roads & river beds impacting percolation and silt deposition in dam's and waterbodies reducing the carrying capacity. Additionally, untreated industry effluents discharged into the rivers, rampant mining and climate change impacts are catastrophic. The adverse effects of anthropogenic activities and climate change on water resources could be only neutralized with good water treatment methods and governance policies.

Demand for water is rising exponentially while supply is become more erratic and uncertain. Unless action is taken soon water will become scarce even in the region where it is currently abundant. Food prices spikes caused by droughts can influence latent conflicts and drive migration. Economic growth impacted by erratic rainfalls and episodes of droughts and floods have generated waves of migration and rise in violence within the countries. Climate -resilient economies require better planning of water, resources allocation, adoption of incentives for increasing water efficiency and investments in right infrastructures for more secure water, supplies and availability.

Implementing nature-based strategies can enhance the resilience of various climatic zones ecosystems against climate change and human-induced challenges of untreated sewage discharge. Therefore, it is essentially required that standards of water quality be maintained along with waste water treatment standards, with site specific ecosystems. Although there is formation of high level committees for monitoring the water quality, planning for controlling, charging penalties to industries or domestic tenants, collection of penalties, and use of these revenues for the rejuvenation of water bodies & file case against the violators etc. Also there are various agencies having role in Municipal corporations, Pollution Control Boards, etc. to maintain the storage, water utilisation & water charge collection, pollution monitoring etc. But there is a huge gap in the governance system and academia about the Science of water, water bodies as a bioreactor and understanding of sewage degradation for its apt treatment, and maintaining the quality standards of the rivers. Kumbh Mela 2025 was celebrated from 13th Jan to 26th Feb at Prayagraj, UP in India. Millions of people took dip in Sangam area where Ganga, Yamuna and Saraswati confluences. Reports by the Central Pollution Control Board (CPCB) had revealed alarming levels of bacterial contamination, which raised concerns over public health and safety. However, religious sentiments were very high as this holy confluence of the Ganga, Yamuna, and the mythical Saraswati has been a spiritual epicentre for centuries. CPCB's latest submission to the National Green Tribunal (NGT) confirms that high levels of faecal coliform (microbes from human and animal excreta) were found in the river where devotees took their holy dip. The presence of these bacteria is a clear indicator of sewage contamination, making the water unfit for bathing and exposing devotees to serious health risks, including gastrointestinal infections and skin diseases. The reports confirmed that faecal coliform levels in the Ganga reached 11,000 MPN/100ml near Shastri Bridge and 7,900 MPN/100ml at Sangam, far exceeding the recommended maximum of 2,500 MPN/100ml. And Yamuna showed alarming levels of 4,900 MPN/100ml near the Old Naini Bridge before it merges with the Ganga. This had put high burden on rivers survival too.

The constitution of any nation is drafted for maintaining a peaceful society. Does that consider constitution of the planet, in other words, laws of the nature (the planet)? It is the latter we are focussing on. If a national constitution does not cover it, any decisions taken on nature's safety will be misleading. Because a society's (nation's) safety depends on the nature's balance. We are at war with nature we can never win. Keeping in view the complexities of climate change and environmental issues it is pertinent that the national constitution should throw light on restoration of nature's organ/s and not replacement. And the policy framework process must include a microbiologist and a molecular scientist to be on the team for laying guidelines.

All small drains, nallahs, rivulets, water reservoirs, primary, secondary, tertiary streams, are hierarchical components of a river. The zero order streams have been destroyed by urban sprawling. Hence it is necessary to explain at length what zero order streams are. Taking away any one component destroys the entire system of a river (vessels of the planet). And we have no substitute for lost streams/rivers. Such a distortion of natural structure has unforeseen consequences, at times, with unimaginable economical catastro-

phes. At the most Govt. can allot funding, but nature does not read nor care for a budget. This is happening world over. And is it a river if its water all gone? As rivers are drying and dying. All rivers, soil with microbes, once gone do we have a design to recreate or recover them?

Rivers and lakes have many other planetary obligations to fulfil which mankind is not aware of. Chief of which, they hold the keys to climate stability. Untreated discharge of sewage has impacted the river climate stability. Understanding sewage will ease treat aquatic systems. World over sewage management is standing on an unexamined foundation. Conventional sophisticated approach is unaffordable to a pocket and to the nature, hence not a solution. Sewage is a climate linked microbial science governing public health and pathogen elimination, breaking worms life-cycles. Inadequate sanitation for 40% world population now, where are we for 10 billion in 2050? Conventional STPs use excess Water and aeration, whereas they have no role in organic degradation and infrastructure built on such parameters are a liability with no solution in sight. Water crisis is a fallout of pushing the poop with pristine water. Gene transcription dynamics forms the very basis of sewage degradation, which only colon microbes are capable of performing, this is yet to be understood. Nutrient and water recovery from a wastewater treatment plant is a proof of sewage illiteracy rendering a society poorer. Advanced technology like Metagenomics and shotgun have sequenced genome to unveil the microbial dark matter (Zhang et. al 2023, Pavlopoulos et. al, 2023).

Hence this paper is focusing on sewage dynamics using metagenomics tool for analysis of water and soil samples of Varal Devi lake in Bhiwandi Taluka Thane Dist, Maharashtra. This lake is spread in almost 62 acres and flows into Kamvari river and its water is supplied for drinking purpose. Kamvari river which was once a riverine port is diminished into a drainage and polluted water from this is used for agriculture and farming in the region. Both waterbodies are posing immense health risk threat to human being, soil, ground water and surface water. The main reason of using metagenomics tool is to identify the diversity and distribution of bacteria and pathogens, to provide nature based solution to sewage treatment as an alternative to the STPs.

Methodology

Environmental samples were collected from soil and water sources at Varhaldevi Lake and Kamvari Nadi, for microbial DNA extraction for sequencing. The sequencing depth ranged from 0.1 to 0.2 million reads per sample to ensure sufficient taxonomic coverage. 16S rRNA amplicon sequencing was conducted to analyze bacterial and archaeal communities, specifically targeting the V3-V4 variable region. High-throughput sequencing using Illumina technology ensured accurate and in-depth analysis.

Raw sequencing reads underwent quality control using FastQC, evaluating base quality scores, GC content, and sequence length distribution. Low-quality sequences were filtered out based on a Phred score threshold of >19. QIIME was used for OTU/ASV clustering, referencing the GreenGenes 13_8_97 database, and generating an OTU table to quantify microbial abundance.

Taxonomic classification was performed using the GreenGenes 16S rRNA reference database, assigning microbial communities at multiple taxonomic levels. Alpha diversity analysis (Shannon Index, Simpson Index, Richness) measured species diversity within samples, while Beta diversity analysis (Unweighted & Weighted UniFrac distances) compared microbial compositions between samples. G-Test, Kruskal-Wallis test, T-test, and Mann-Whitney U test evaluated significant differences between microbial communities.

To visualize microbial composition, relative abundance plots, heatmaps, Krona plots, and dimensionality reduction techniques (t-SNE, PCA, PCoA) were used. Rarefaction curves assessed sequencing depth adequacy, and STAMP analysis identified statistically significant microbial differences. PICRUSt was utilized for functional predictions, estimating microbial metabolic pathways. Environmental risk assessment categorized sites into high, moderate, and low risk based on the presence of pathogenic bacteria.

A comprehensive report was generated, incorporating sequencing statistics, taxonomic assignments, diversity indices, and visual representations to provide an in-depth analysis of microbial diversity, environmental health, and potential contamination risks.

Findings and discussion

World over current practice is to flush toilet waste with grey water (other use namely bath, laundry, and kitchen). This is called MFSL, mix first separate latter, for convenience called WW- wastewater (meaning domestic WW). This is unscientific and once mixed there is no successful management for the treatment of domestic WW. Because all BW & GW (kitchen liquid waste) have different physical, chemical, biochemical and microbial characteristics. Black and Grey waters cannot be mixed, must be kept separate, and be treated separately. There is no method (STP) that can offer a solution. One single economic decision of using a common pipe (Cloaca Maxima) to carry all liquids has created this mess. More dangerously BW as a source of pathogen is not addressed adequately. Repeated pandemics world over (4 in last 100 and others over 1000 years) all urban in origin and sewage as a common source of pathogens is the evidence.

This challenge is further complicated by mixing industrial wastewater and storm water to meet the convenience is misleading with unforeseen consequences. To cover it up, new terms like sustainable development, asset management, etc., and science is neglected, economy suffers as well. Slums and eutrophication are urban gift to this planet. Such an approach with only short-term economy does not guarantee civility. Sewage is a climate linked microbial science. Handling it with mere hydraulic equations, tools, leads us nowhere. Flushing human excreta with pristine water (leading to water scarcity) is no brainer.

Therefore, to get clarity metagenomics of Varal devi lake & Kamvari iver water and soil sample was done as this lake and river in Bhiwandi Taluka in Thane District Maharastra, India is completely polluted mainly due to untreated sewage discharge and other anthropogenic activities. Analyzing the 16S V3-V4 region in metagenomics is a common approach for studying microbial communities. The 16S rRNA gene is used as a marker gene to identify and classify bacteria and archaea. The V3-V4 region of the 16S rRNA gene is often targeted for sequencing because it provides a good balance between taxonomic resolution and sequencing cost.

Data quality analysis

The raw data quality was assessed using the FastQC program, a tool widely employed for this purpose in bioinformatics. The results of this assessment, which provided a comprehensive overview of the quality of individual reads. This study provides valuable information about various quality metrics, such as sequence length distribution, per-base sequence quality, overrepresented sequences, and more, all of which are crucial for evaluating the integrity of the raw data before further analysis or downstream processing. The FASTQC analysis revealed suboptimal sequencing quality with the presence of poor-quality sequences. Subsequently, we eliminated these low- quality sequences, retaining only those bases with a Phred score greater than 19 for OTU picking. Quality filter results are shared.

Sample ID	*Total Input Sequences*	*Reads Too Short After Quality Truncation*	*Median Sequence Length (bp)*	*Total Sequences Retained After Filtering*	*Percentage of Reads Retained*
S-5	17,894	364	460.00	17,530	97.97%
S-6	18,309	74	459.00	18,235	99.60%
S-7	18,825	325	460.00	18,500	98.27%
S-10	20,417	1,540	460.00	18,877	92.46%
S-11	18,418	644	454.00	17,774	96.50%

Table 1: Quality Filtering Summary: Quality filtering results across samples. Most samples retained over 95% of sequences, except S-10, which had the highest read loss (7.54%). S-6 showed the best retention (99.60%). Median sequence length remained consistent (460 bp), except for S-11, which had a slightly shorter length (454 bp).

The quality filtering results show that most samples retained over 95% of their sequences, indicating high sequencing quality. S-6 had the best retention rate (99.60%), suggesting minimal read loss, while S-10 had the highest sequence loss (7.54%), indicating lower initial sequencing quality. S-11 had the shortest median sequence length (454 bp) compared to others (~460 bp), possibly due to minor truncation. Despite some variation in read loss, all samples maintained sufficient sequence quality for downstream microbial analysis.

Sample ID	*Total Reads*	*OTUs Observed*	*Interpretation*
S-10 (Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil)	3,572	468	Lowest read count but moderate OTU count, indicating stable microbial diversity.
S-11 (Varhaldevi lake soil, pumping site)	4,658	569	Highest OTU count, suggesting the most diverse microbial community.
S-6 (Varhaldevi lake water, pumping site)	6,352	383	Moderate read count but lower OTU count, indicating a less diverse but stable community.
S-5 (Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water)	6,952	480	High read count with a relatively high OTU count, suggesting good microbial richness.
S-7 (Kamvari Nadi, Eidgah site)	7,050	287	Highest read count but lowest OTU count, indicating dominance of a few microbial species.

Table 2: Reads and OTU Counts Per Sample with Interpretation: Summary of sequence reads and OTU counts across samples. S-11 (Varhaldevi lake soil, pumping site) had the highest microbial diversity (569 OTUs), while S-7 (Kamvari Nadi) had the lowest (287 OTUs) despite having the highest read count (7,050), suggesting microbial dominance by a few species. S-5 (Inlet Discharge water) and S-10 (Inlet Discharge soil) showed relatively high OTU counts, indicating good microbial richness.

Statistic	*Reads*	*OTUs*	*Interpretation*
Total Across Samples	28,584	1,694	Indicates a substantial sequencing effort capturing diverse microbial communities.
Minimum	3,572 (S-10)	287 (S-7)	S-10 had the lowest reads, while S-7 had the lowest OTU count, suggesting limited microbial diversity in S-7.
Maximum	7,050 (S-7)	569 (S-11)	S-7 had the most reads but the least diversity, while S-11 had the most OTUs, indicating a highly diverse microbial population.
Median	6,352	468	The middle values suggest a balanced distribution of reads and OTUs across samples.
Mean	5,716.8	437.4	The average OTU count shows a relatively diverse microbial community overall.
Standard Deviation	1,373.16	95.55	The variation in reads is greater than in OTUs, showing some inconsistency in sequencing depth but relatively stable microbial diversity.

Table 3: Statistical Summary of Reads and OTUs with Interpretation: Statistical summary of sequence reads and OTU observations. The dataset contains a total of 28,584 reads and 1,694 OTUs. S-11 exhibited the highest microbial diversity (569 OTUs), while S-7 had the lowest (287 OTUs), despite having the most reads. The variation in reads is greater than in OTUs, indicating that while sequencing depth varies, microbial diversity remains relatively stable across samples.

The table shows S-11 has the highest microbial diversity, making it the richest microbial sample. S-7 has the most reads but the lowest OTU count, suggesting a microbial community dominated by a few species. S-10 has the lowest read count but a moderate OTU count, meaning sequencing depth may have limited species detection. Overall, the dataset captures a diverse microbial community, with variation in sequencing depth across the samples.

Statistical analysis

A series of conducted for statistical analyses to explore and draw meaningful insights from the abundance data of Operational Taxonomic Units (OTUs). These analyses aimed to assess the significance of differences and associations within the OTU abundance data, using a variety of statistical tests tailored to different research questions and hypotheses:

G-Test: also known as the G-statistic or likelihood ratio test, was used to compare the observed distribution of OTUs to an expected distribution. It assesses whether there are significant differences between groups in terms of the proportions of OTUs they contain. This test is particularly useful when dealing with categorical data and can highlight associations between categorical variables and OTU abundance.

Kruskal-Wallis Test: is a non-parametric test employed to evaluate whether there are statistically significant differences in the distribution of OTU abundances among multiple groups or conditions.

Test	***p-value***
G-Test (OTU Distributions)	0.000 (Highly significant, OTU distributions vary significantly across samples)
Kruskal-Wallis (Shannon Index)	0.406 (Not significant)
Kruskal-Wallis (Simpson Index)	0.406 (Not significant)
Kruskal-Wallis (Richness)	0.406 (Not significant)
T-Test (Shannon Index)	0.624 (Not significant)
T-Test (Simpson Index)	0.982 (Not significant)
T-Test (Richness)	0.292 (Not significant)
Mann-Whitney (Shannon Index)	0.800 (Not significant)
Mann-Whitney (Simpson Index)	0.800 (Not significant)
Mann-Whitney (Richness)	0.200 (Not significant)

Table 4: Overall Statistical Tests for Diversity & OTU Distribution.

NSTI score analysis: The NSTI score essentially quantifies the phylogenetic relatedness between the sequences in your dataset and the closest fully sequenced organisms in a reference database, typically a 16S rRNA gene database for bacterial and archaeal community analyses.

Table 5: NSTI score analysis showing taxonomic similarity between identified microbial species and known reference genomes. S-11 has a higher NSTI, indicating the presence of novel or less-studied species, while S-7 and S-10 have lower NSTI scores, suggesting their microbial species are well-characterized.

The above table show S-11 and S-6 have higher NSTI scores, meaning they contain some novel or poorly studied bacteria. S-7 and S-10 have the lowest NSTI scores, indicating their microbial species are well-represented in reference databases. And S-5 has a moderately low NSTI, meaning Leptospira and other identified microbes are fairly well-documented.

Sample ID	*NSTI Score*	*Interpretation*
S-11 (Varhaldevi lake soil, pumping site)	0.244	Moderately high, suggesting some well-characterized species but also potential novel taxa (e.g., Turneriella and Leptospira).
S-6 (Varhaldevi lake water, pumping site)	0.220	Indicates a diverse microbial community with some novel taxa.
S-7 (Kamvari Nadi, Eidgah site)	0.121	Lowest NSTI score, meaning most identified species match known reference genomes well.
S-5 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge water)	0.196	Moderately low, suggesting the presence of well-characterized species like Leptospira.
S-10 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge soil)	0.142	Low NSTI score, meaning most species are well-matched to reference databases.

Table 5: NSTI Score Per Sample.

Taxonomy classification

Bar plots have helped visualizing the relative abundance of different microbial taxa across samples. They allow for Comparing microbial composition at various taxonomic levels (Phylum, Class, Order, Family), Identifying dominant microbial groups in different environments and understanding differences in microbial diversity across samples.

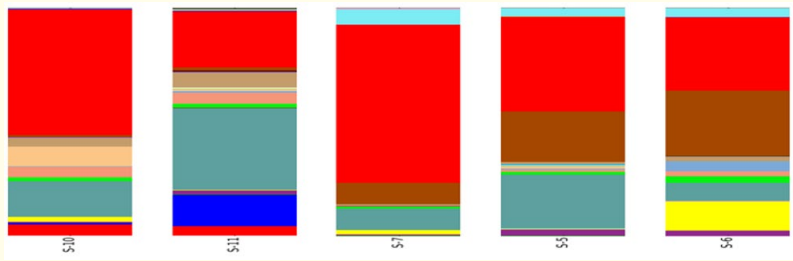


Figure 1: Bar plot: Phylum-level microbial composition across samples. Proteobacteria dominate most samples, especially in water environments (S-5: 69.4%, S-6: 41.6%), indicating a strong water-associated microbial presence. Bacteroidetes is highest in S-7 (35.8%), suggesting a role in organic matter degradation, while Actinobacteria is most abundant in S-6, reflecting soil microbial activity.

Phylum	*S-10 VBS*	*S-11 VPS*	*S-7 KI*	*S-5 VBW*	*S-6 VPW*	*Key Observations*
Proteobacteria	44.6%	55.3%	24.8%	69.4%	41.6%	Dominant across all samples, highest in S-5, indicating a strong water-associated microbial community.
Bacteroidetes	18.6%	15.7%	35.8%	9.5%	23.7%	High in S-7, suggesting its role in organic matter degradation.
Planctomycetes	12.5%	1.0%	1.1%	9.4%	22.2%	Highly abundant in S-6, indicating its role in nitrogen cycling.

Actinobacteria	3.4%	1.9%	0.3%	1.8%	0.2%	Low abundance across samples, but highest in S-6 (12.9%), suggesting presence of soil bacteria.
Firmicutes	2.1%	8.8%	0.8%	0.0%	1.0%	Highest in S-11, possibly indicating spore-forming bacteria.

Table 6: Phylum-level microbial composition across samples. Proteobacteria dominate, especially in water samples (S-5, S-6), while Bacteroidetes is high in S-7, suggesting its role in organic matter degradation and pollution.

The principle of a Sewage Treatment Plant (STP) involves the treatment of wastewater through a series of physical, biological, and chemical processes to remove pollutants and contaminants. The primary principle is to mimic and enhance natural processes that occur in the environment to purify water. Hence the metagenomics study is very important to understand the role of microbes and to maintain the hydraulics retention time (HRT) or solid retention time (SRT) for the impactful action for organic degradation, by detailed analysis of the waste water sample.

Proteobacteria are pathogens causing diseases in animals and human beings, but few are beneficial and essential. They play environmental role like Desulfovibrio (sulphate reduction), Geobacter (metal reduction), Nitrosomonas (nitrogen cycle), Rhizobium (nitrogen fixation), E.Coli some strains, Salmonella, Shigella, Helicobacter, Canpylobacter, Pseudomonas, Vibrio Cholera, Rickettsia etc are few examples. Bacteroidetes supress pathogens. They play role in soli health improvement, by degrading organic matter and mobilizing nutrients. Participating in nutrient cycling, degrading complex carbohydrates, polysaccharides like cellulose, starch, pectin and production of short-chain fatty acids, acetate, butrate, propionate etc. Whereas Planctomycetes play vital role in global carbon and nitrogen cycles, especially in aquatic environments, and are known for applications in wastewater treatment. Planctomycetes are key players in both the carbon and nitrogen cycles, contributing to the breakdown of complex carbon sources and performing important reactions like anaerobic ammonium oxidation crucial for nitrogen cycling. Some Planctomycetes have an endocytosis-like mechanism for macromolecule uptake. Actinobacteria, are a diverse phylum of Gram-positive bacteria which play crucial roles in ecosystem by decomposing organic matter, cycling nutrients like carbon, nitrogen, phosphorus, and potassium, etc. They are the key decomposers, breaking down complex organic materials like cellulose, chitin, and lignin, and are essential for nutrient cycling and soil health and degrade pollutants like hydrocarbons. Firmicutes are the Gram-positive bacteria, many of which are spore-forming and ferment sugars. Examples include genera like Bacillus, Clostridium, Enterococcus, Lactobacillus, and Ruminococcus. They play a role in degrading organic matter and potentially contribute to nitrogen removal through denitrification, especially under anaerobic conditions, indicating active hydrolysis and methanogenesis steps. They form biofilms, which are communities of microorganisms attached to surfaces, and these biofilms can enhance the biological nitrogen removal ability of the treatment system. In municipal wastewater treatment plants, Proteobacteria, Bacteroidetes, and Firmicutes are among the dominant bacterial phyla, all of which have the function of denitrification and dephosphorization to degrade organic matter.

The above table shows that the pathogens percentage is overall high in all the samples, except S-11.

Streptomyces 3.5%, 7.4%, 1.0%, 0.5%, 2.2% High in S-11, important for soil health.

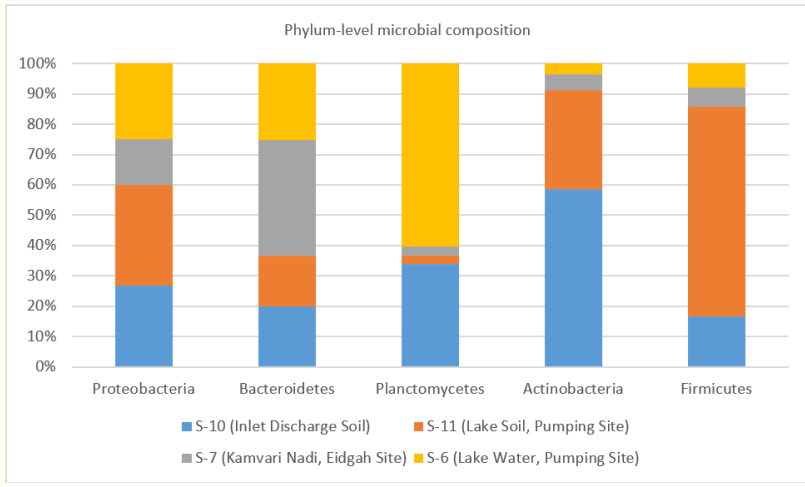


Figure 2: Phylum level microbial composition in the various sites of Varal devi lake and Kamvari river.

Class	***S-10***	***S-11***	***S-7***	***S-5***	***S-6***	***Key Observations***
Gammaproteobacteria	25.6%	30.5%	12.9%	48.1%	31.4%	Common in water samples, involved in organic decomposition.
Bacteroidia	13.5%	10.4%	32.1%	6.7%	19.8%	Dominates S-7, linked to organic matter breakdown.
Actinobacteria	3.2%	1.5%	0.2%	1.6%	12.9%	High in S-6, indicating soil bacteria presence.

Table 7: Class-level microbial composition showing Gammaproteobacteria dominance in water samples (S-5, S-6), while Bacteroidia is highest in S-7, suggesting an organic degradation role.

Order	***S-10***	***S-11***	***S-7***	***S-5***	***S-6***	***Key Observations***
Pseudomonadales	21.2%	25.3%	10.1%	36.9%	25.7%	High in S-5, common in water environments.
Flavobacteriales	9.8%	7.1%	25.6%	5.2%	12.4%	Dominates S-7, linked to organic decomposition.
Streptomycetales	6.3%	9.5%	1.2%	0.7%	3.4%	High in S-11, supports soil health and antibiotic production.

Table 8: Order-level microbial composition. Pseudomonadales dominates in water samples, while Flavobacteriales is high in S-7, suggesting an organic degradation role.

Family	***S-10***	***S-11***	***S-7***	***S-5***	***S-6***	***Key Observations***
Pseudomonadaceae	18.5%	22.7%	8.6%	32.4%	21.8%	High in S-5, aquatic decomposers.
Flavobacteriaceae	7.1%	5.8%	23.4%	4.6%	10.3%	Dominant in S-7, suggests organic degradation.

Table 9: Family-level microbial composition. Pseudomonadaceae dominates in water samples, while Flavobacteriaceae is high in S-7, suggesting an organic degradation role.

Genus	*S-10*	*S-11*	*S-7*	*S-5*	*S-6*	*Key Observations*
Pseudomonas	12.8%	14.9%	6.3%	21.4%	15.6%	Common in water, important for organic decomposition.
Flavobacterium	4.9%	3.7%	18.6%	2.9%	6.7%	High in S-7, suggests organic degradation role.
Streptomyces	3.5%	7.4%	1.0%	0.5%	2.2%	High in S-11, important for soil health.

Table 10: Genus-level microbial composition. Pseudomonas dominates water samples, while Flavobacterium is highest in S-7, suggesting an organic degradation role.

Krona Plot Analysis, Summary, and Interpretation

A Krona plot was created individually for each sample to visually represent and explore the microbial diversity present in the dataset. Krona plots are effective tools for conveying complex taxonomic information in an easily interpretable and interactive manner.

Krona plots provide a hierarchical and interactive visualization of microbial diversity, allowing users to explore taxonomic composition at different levels (phylum, class, order, family, genus, species). They are useful in: comparing microbial diversity across different samples, identifying dominant bacterial groups in each environment, understanding taxonomic distribution and ecological functions.

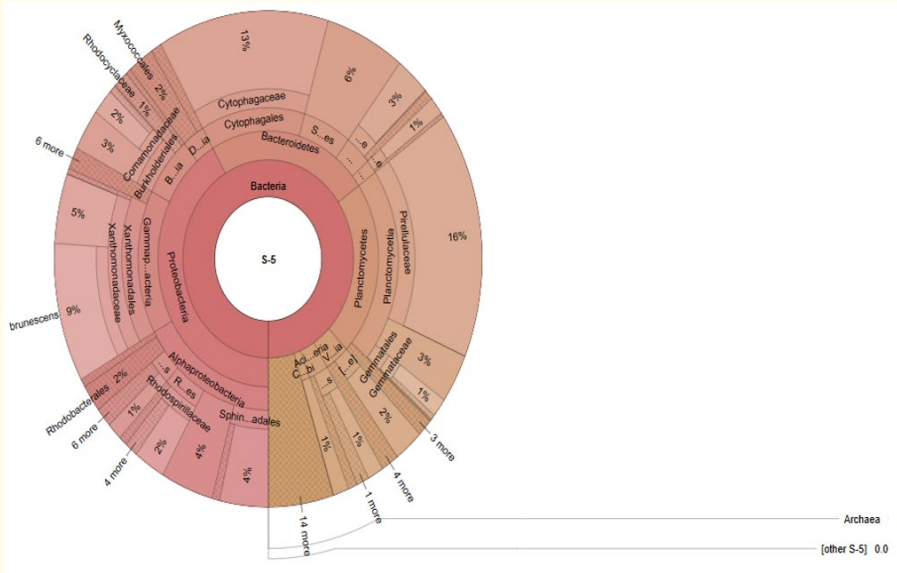


Figure 3: Krona plot of Sample S5 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge Water), showing Leptospira dominance, suggesting potential contamination. The microbial community lacks diversity, indicating a disturbed water ecosystem.

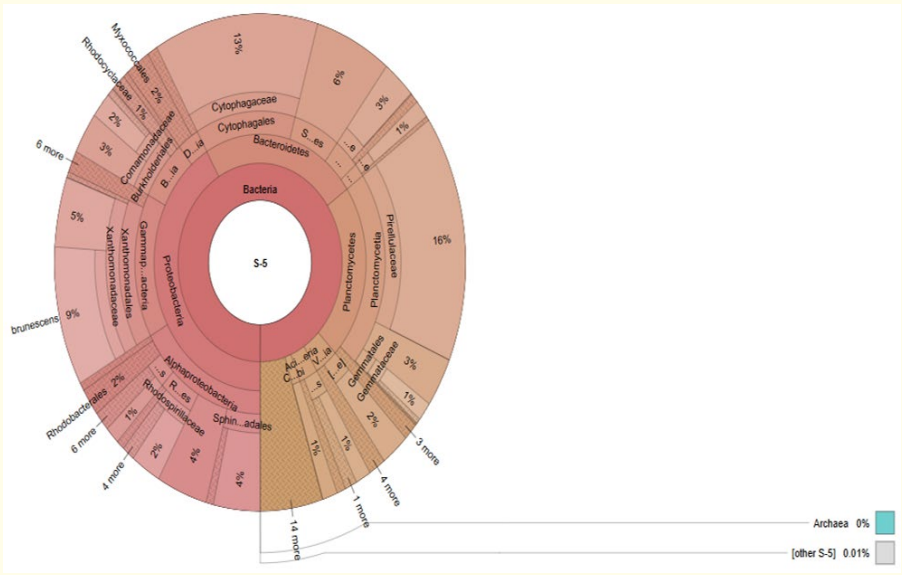


Figure 4: Krona plot of Sample S6 (Varhaldevi Lake Water, Pumping Site), revealing a balanced microbial community with Pseudomonas and Acinetobacter as dominant genera, indicating a stable aquatic microbiome.

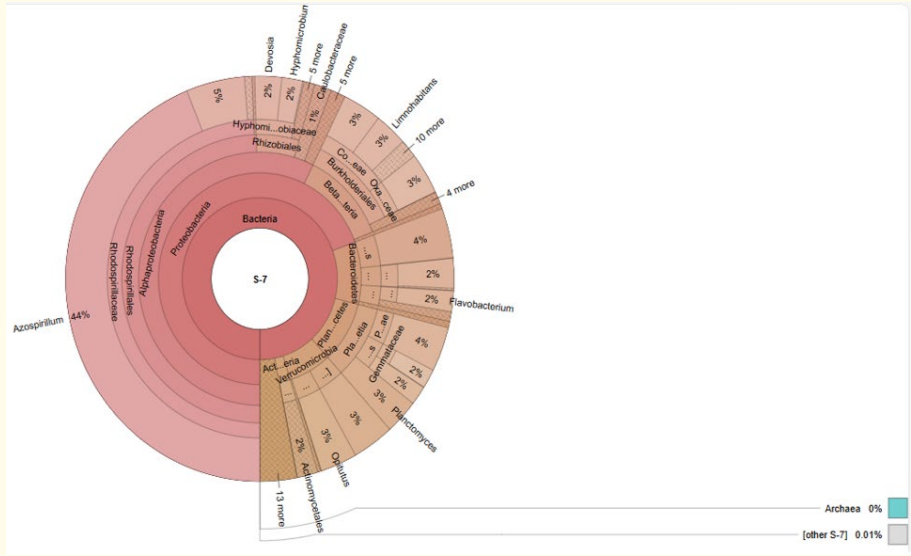


Figure 5: Krona plot of Sample S7 (Kamvari Nadi, Eidgah Site), showing low microbial diversity, with Bacillus and Aeromonas dominating. The limited species richness suggests environmental stress or pollution.

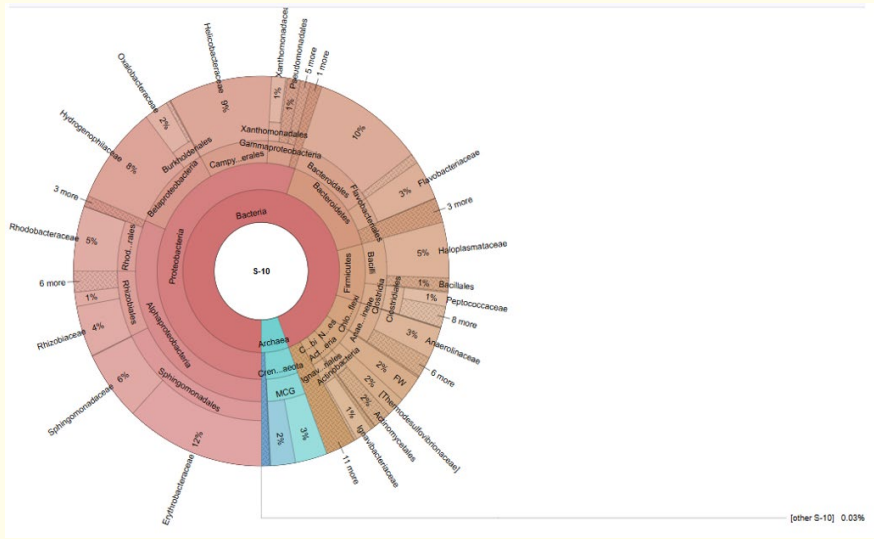


Figure 6: Krona plot of Sample S10 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge Soil), visualizing a diverse soil microbiome, dominated by Streptomyces and Pseudomonas, suggesting active nutrient cycling and potential for antibiotic production.

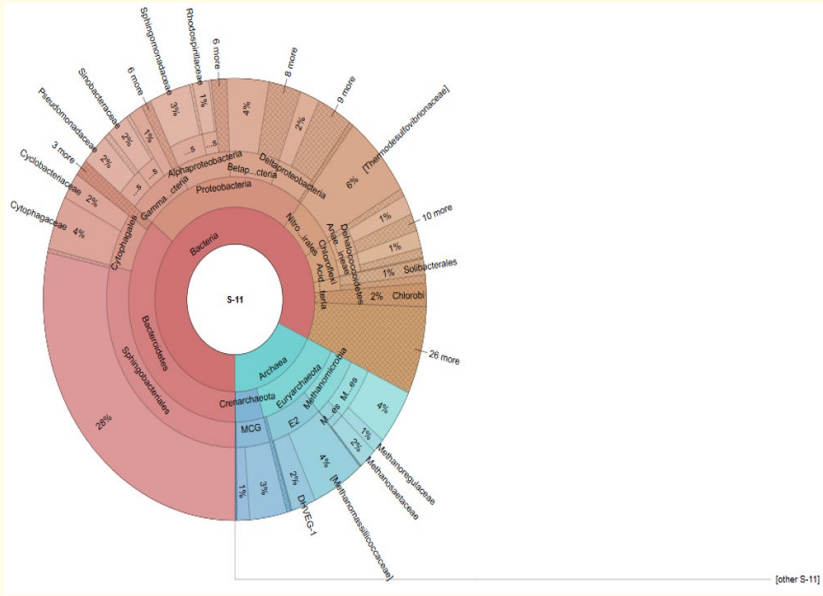


Figure 7: Krona plot of Sample S11 (Varhaldevi Lake Soil, Pumping Site), displaying the highest microbial diversity, with dominant taxa including Streptomyces, Bacillus, and Rhizobium, indicating a rich and ecologically stable soil community.

Sample ID	*Dominant Phylum*	*Dominant Genera*	*Key Observations*
S-5 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge Water)	Spirochaetes	Leptospira	Low diversity, Leptospira-dominated, indicating potential contamination.
S-6 (Varhaldevi Lake Water, Pumping Site)	Proteobacteria	Pseudomonas, Acinetobacter	Aquatic microbes, moderate diversity, stable water microbiome.
S-7 (Kamvari Nadi, Eidgah Site)	Proteobacteria, Firmicutes	Bacillus, Aeromonas	Lowest diversity, limited species richness, possibly due to high level of pollution.
S-10 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge Soil)	Actinobacteria, Proteobacteria	Streptomyces, Pseudomonas	High soil microbial diversity, likely involved in nutrient cycling.
S-11 (Varhaldevi Lake Soil, Pumping Site)	Actinobacteria, Firmicutes	Streptomyces, Bacillus, Rhizobium	Highest microbial diversity, rich soil community structure.

Table 11: Summary of Krona Plot Findings.

Interpretation of Krona Plot Data:

1. S-11 (Varhaldevi Lake Soil, Pumping Site) has the highest microbial diversity.
   - Dominated by *Actinobacteria, Firmicutes*, and *Proteobacteria*, which contribute to soil health, organic matter decomposition, and plant-microbe interactions.
   - *Streptomyces* and *Bacillus* are involved in antibiotic production and bioremediation.
2. S-5 (Inlet Discharge Water) is dominated by *Leptospira*, indicating contamination.
   - *Leptospira* is a pathogenic bacterium associated with waterborne diseases.
   - The low diversity suggests a disturbed microbial ecosystem, possibly due to pollution and rodents found in the holes in the periphery surrounding the lake, due to negligence of maintenance and repair of the reservior.
3. S-6 (Lake Water, Pumping Site) has a stable microbial community.
   - *Pseudomonas* and *Acinetobacter* are common waterborne bacteria, playing roles in organic matter decomposition.
   - Moderate diversity indicates a balanced aquatic microbiome.
4. S-7 (Kamvari Nadi, Eidgah Site) has the lowest diversity.
   - *Aeromonas* and *Bacillus* suggest limited diversity and possible pollution effects.
   - A low number of microbial species may indicate environmental stressors.
5. S-10 (Inlet Discharge Soil) has high microbial diversity, supporting soil health.
   - Presence of *Actinobacteria*, *Proteobacteria* suggests nutrient cycling and organic matter decomposition.
   - *Streptomyces* is known for antibiotic production and biodegradation.

Top most Abundant Genera in Each Sample

S-10 (Varaldevi Inlet Discharge Soil)

1. Unclassified (2,066 counts).
2. *Erythrobacter* (375) – Phototrophic bacterium.
3. *Thiobacillus* (292) – Sulfur-oxidizing bacteria.
4. *Sulfuricurvum* (101) – Sulfur-metabolizing bacteria.
5. *Anaerolinea* (83) – Found in anaerobic environments.
6. *Novosphingobium* (72) – Involved in bioremediation.

7. *Gramella* (68) – Marine bacterium, sometimes found in sediments.
8. *Salinimicrobium* (26) – Halotolerant bacteria.

S-11 (Varhaldevi Lake Soil, Pumping Site)

1. Unclassified (3,831 counts).
2. *Pseudomonas* (108) – Organic matter decomposer.
3. *Methanosaeta* (70) – Methanogenic archaea.
4. *Novosphingobium* (49) – Bioremediation bacterium.
5. *Candidatus Methanoregula* (40) – Methanogenic archaea.
6. *Thiobacillus* (38) – Sulfur-oxidizing bacteria.
7. *Nitrospira* (25) – Nitrifying bacteria.
8. *Perlucidibaca* (21) – Aquatic bacterium.

S-7 (Kamvari Nadi, Eidgah Site - Water)

1. *Azospirillum* (3,093) – Nitrogen-fixing bacteria.
2. Unclassified (2,470 counts).
3. *Opitutus* (215) – Found in wetland environments.
4. *Planctomyces* (196) – Organic matter decomposer.
5. *Limnohabitans* (192) – Common freshwater bacterium.
6. *Devosia* (158) – Associated with nitrogen metabolism.
7. *Hyphomicrobium* (122) – Involved in organic carbon cycling.
8. *Flavobacterium* (117) – Plays a role in organic degradation.
9. *Caulobacter* (56) – Important in aquatic environments.
10. *Mycobacterium* (51) – Some species are pathogenic.

S-5 (Varaldevi Inlet Discharge Water)

1. Unclassified (5,303 counts).
2. *Lysobacter* (630) – Produces antimicrobial compounds.
3. *Planctomyces* (223) – Involved in organic matter degradation.
4. *Magnetospirillum* (165) – Magnetotactic bacteria.
5. *Bryobacter* (63) – Found in peat and soil environments.
6. *Rhodobacter* (61) – Photosynthetic bacterium.
7. *Nitrospira* (60) – Nitrifying bacteria.
8. *Prosthecobacter* (54) – Found in freshwater habitats.
9. *Azospira* (37) – Nitrogen-fixing bacteria.
10. *Clostridium* (33) – Some species are pathogenic.

S-6 (Varhaldevi Lake Water, Pumping Site)

1. Unclassified (5,119 counts).
2. *Planctomyces* (385) – Organic matter decomposer.
3. *Gemmata* (205) – Planctomycete.
4. *Rhodobacter* (159) – Photosynthetic bacterium.
5. *Prosthecobacter* (129) – Found in freshwater.

6. *Nitrospira* (110) – Nitrifying bacterium.
7. *Plesiocystis* (39) – Myxobacterium involved in predatory behaviour.
8. *Microcystis* (27) – Cyanobacteria, associated with blooms.
9. *Phenylobacterium* (22) – Degrades aromatic compounds.
10. *Hyphomicrobium* (19) – Organic carbon cycling.

Comparison: Soil vs. Water Samples

Soil samples (S-10, S-11) contain

- More sulphur-metabolizing bacteria (*Thiobacillus, Sulfuricurvum*).
- Methanogenic archaea (*Methanosaeta, Candidatus Methanoregula*) in S-11.
- Bioremediation bacteria (*Novosphingobium, Pseudomonas*).

Water samples (S-5, S-6, S-7) contain

- Nitrogen-fixing bacteria (*Azospirillum, Azospira, Nitrospira*).
- Photosynthetic organisms (*Rhodobacter, Microcystis*).
- Planctomycetes (*Planctomyces, Gemmata*), known for complex lifestyles.

Soil microbiomes are more diverse in sulphur and methane metabolism, while water microbiomes are richer in nitrogen cyclers and photosynthetic bacteria.

t-SNE Plot (Clustering of Microbial Communities)

A t-SNE plot for 16S metagenomics is a powerful visualization tool used to represent complex microbial community data in a reduced-dimensional space. t-SNE stands for t- Distributed Stochastic Neighbour Embedding, and it is a dimensionality reduction technique commonly employed in bioinformatics and metagenomics to visualize high-dimensional data such as microbiome composition.

Sample-Wise Interpretation:

- ***S-11 (Varhaldevi lake soil, pumping site)*** → Likely forms a distinct cluster, suggesting a unique microbial composition with high diversity (e.g., Leptospira and Turneriella present).
- ***S-6 (Varhaldevi lake water, pumping site)*** → Expected to cluster closer to S-10 and S-7, showing a moderate overlap in microbial composition.
- ***S-7 (Kamvari Nadi, Eidgah site)*** → Closer to S-6 and S-10, indicating shared microbial species.
- ***S-5 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge water)*** → Likely an outlier in the plot, as it has a low-diversity microbiome dominated by Leptospira.
- ***S-10 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge soil)*** → Expected to cluster near S-7 and S-6, showing a moderate similarity.

PCA Plot (Principal Component Analysis – Microbial Variability Across Samples)

PCA can help identify trends, groupings, or outliers within complex datasets, making it a valuable tool for understanding the structure of microbial communities.

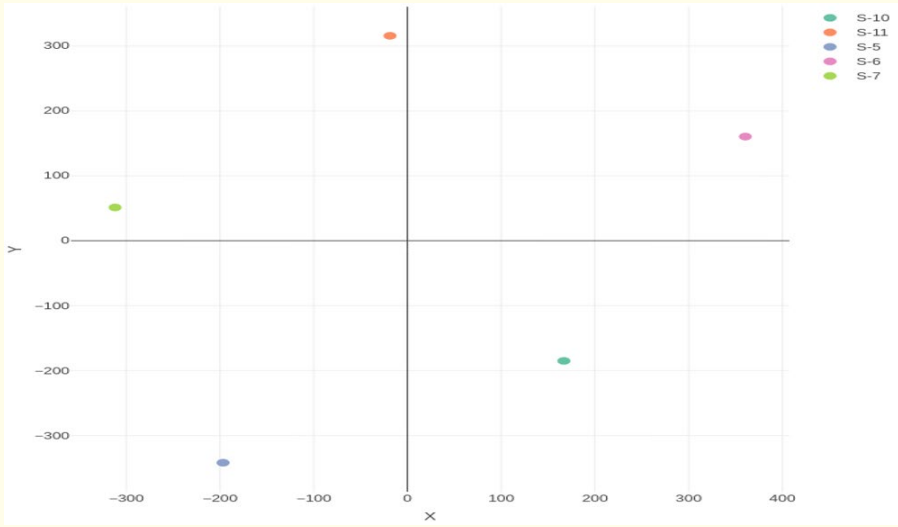


Figure 8: t-SNE plot visualizing microbial community clustering across samples. S-11 appears in a separate cluster, indicating a unique microbial composition, while S-5 is an outlier due to its low-diversity microbiome dominated by Leptospira. Other samples (S-6, S-7, S-10) cluster together, suggesting shared microbial species. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The plot provides a dimensionality-reduced visualization of microbial composition based on 16S metagenomic sequencing.

Sample-Wise Interpretation:

- ***S-11*** → Expected to be far from other samples, indicating a unique microbial structure.
- ***S-6 and S-10*** → Likely appear closer together, suggesting some common microbial diversity.
- ***S-7*** → May be moderately distant from S-6 and S-10, reflecting partial similarity.
- ***S-5*** → Likely an outlier, appearing far from others due to the dominance of Leptospira and low microbial diversity.

PCoA Plot (Principal Coordinate Analysis – Distance-Based Microbial Comparisons)

Principal Coordinate Analysis (PCoA) is a powerful method used in 16S metagenomics to visualize and explore the relationships between microbial communities based on their composition. The PCoA plot provides a graphical representation of the similarities or dissimilarities between samples, allowing to identify patterns, clusters, and trends within the data. PCoA is similar to PCA but uses distance-based metrics (like UniFrac) to measure microbial composition differences. The plot helps visualize how microbial communities differ based on genetic relationships.

Sample-Wise Interpretation:

- ***S-11*** → Expected to be distinct from all other samples, confirming its unique microbial diversity.
- ***S-6 and S-10*** → Likely to appear close together, indicating moderate microbial similarity.
- ***S-7*** → Expected to be moderately separated from S-6 and S-10.

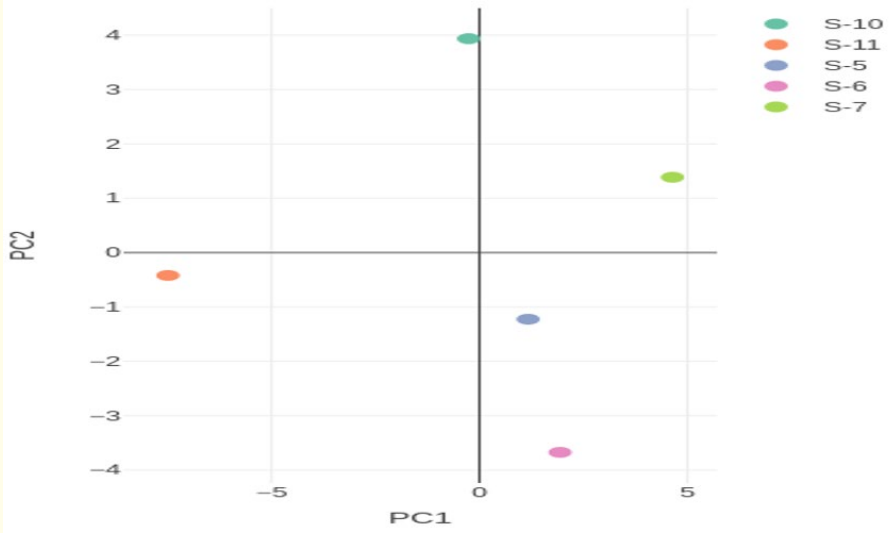


Figure 9: PCA plot visualizing microbial diversity variations across samples. The position of each sample reflects similarities or differences in microbial community composition. Samples containing Leptospira are expected to show distinct clustering based on their bacterial diversity. S-11 is distinct due to its unique bacterial community, while S-5 is an outlier due to Leptospira dominance. S-6, S-7, and S-10 show moderate similarity. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6).

- ***S-5*** → Will likely be far from all others, reflecting its low microbial diversity with Leptospira dominance.

Alpha diversity analysis

Alpha diversity is a fundamental concept in microbial ecology, particularly in 16S metagenomics, where it provides insight into the diversity of microbial communities within a single sample. An alpha diversity plot is a graphical representation of alpha diversity metrics calculated from the abundance of microbial taxa in each sample. These metrics include richness, which measures the number of unique microbial taxa present, and evenness, which assesses how evenly distributed these taxa are within the community.

Alpha diversity assesses microbial diversity within each sample, using different indices that focus on richness (number of species) and evenness (how evenly species are distributed). The following indices provide different insights into microbial diversity:

1. ***Gini-Simpson Index*** – Measures microbial evenness (distribution balance).
2. ***Inverse-Simpson Index*** – Measures richness and evenness together.
3. ***Shannon Index*** – Evaluates both species count and their abundance balance.
4. ***Unit Index*** – General assessment of microbial diversity across samples.

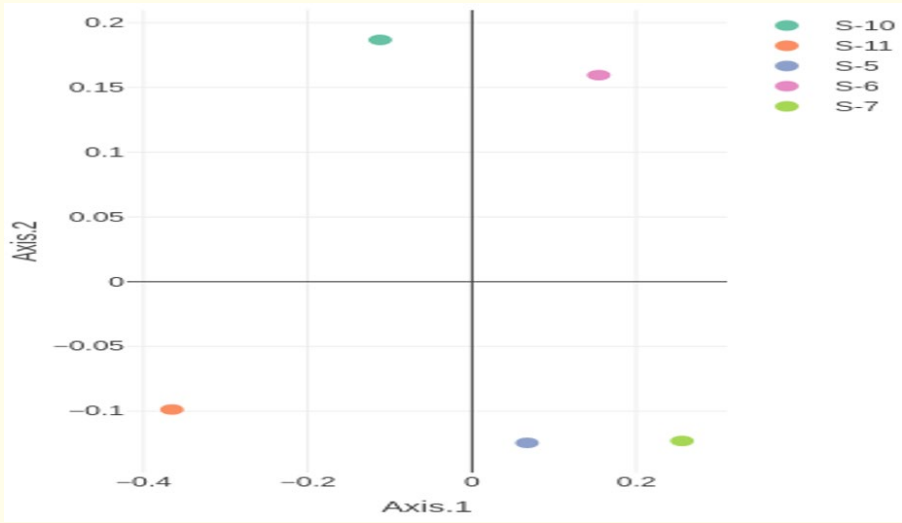


Figure 10: PCoA plot representing microbial community distances across samples. The distance between points represents the level of microbial composition similarity, with Leptospira-containing samples expected to cluster separately from those without. S-11 appears far from other samples due to its unique microbiome, while S-5 is an extreme outlier, dominated by Leptospira. S-6, S-7, and S-10 show moderate clustering, indicating shared microbial compositions. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The PCoA plot visualizes differences in microbial composition, highlighting patterns, clusters, and relationships between the samples based on their taxonomic diversity.

Gini-Simpson Index (Evenness-Focused)

It shows:

- Higher values indicate a more balanced microbial community with evenly distributed species.
- Lower values indicate dominance by a few species in the sample.

Sample-Wise Interpretation

- ***S-11 (Varhaldevi lake soil, pumping site)*** → Highest evenness; microbial species are evenly distributed.
- ***S-6 (Varhaldevi lake water, pumping site)*** → Moderate evenness; species distribution is balanced.
- ***S-7 (Kamvari Nadi, Eidgah site)*** → Lower evenness; a few dominant species exist.
- ***S-5 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge water)*** → Lowest evenness, suggesting a few species (e.g., *Leptospira*) dominate.
- ***S-10 (Varhaldevi Lake Commissioner Bungalow Inlet Discharge soil)*** → Moderate evenness but still lower than S-11.

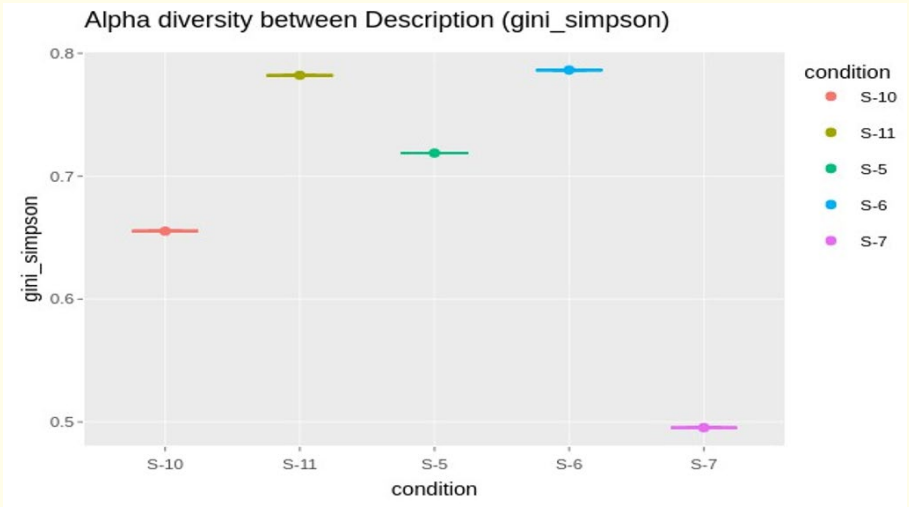


Figure 11: Alpha diversity analysis using the Gini-Simpson index analysis of microbial evenness across environmental samples. S-11 shows the highest evenness, indicating a well-balanced microbial community, while S-5 has the lowest, suggesting dominance by a few species, including Leptospira.The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The plot visualizes richness and evenness of microbial taxa within each sample, providing insights into the community composition and distribution.

Inverse-Simpson Index (Richness & Evenness Combined)

It Shows:

- Higher values indicate greater species diversity and better distribution.
- Lower values indicate fewer species or uneven microbial distribution.

Sample-Wise Interpretation:

- ***S-11*** → Highest diversity, with a variety of species distributed evenly.
- ***S-6*** → Moderate diversity; species are present in a well-balanced way.
- ***S-7*** → Lower diversity; some species dominate while others are less frequent.
- ***S-5*** → Lowest diversity; few bacterial species dominate, reducing overall microbial richness.
- ***S-10*** → Moderate diversity, but not as high as S-11.

Shannon Index (Richness & Abundance Balance)

It shows:

- Higher values indicate more species and even distribution.
- Lower values indicate fewer species, with some dominating the sample.

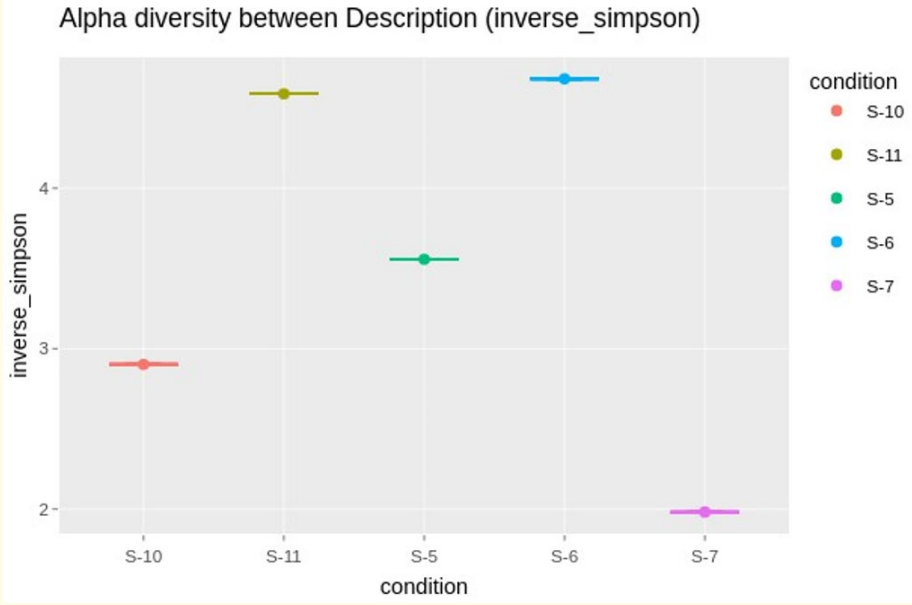


Figure 12: Alpha diversity analysis using the Inverse-Simpson index showing microbial richness and evenness across samples. S-11 has the highest diversity, while S-5 has the lowest, suggesting limited species variety and possible dominance by Leptospira. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Inverse-Simpson index accounts for both richness and evenness, providing a robust measure of microbial diversity, where higher values indicate greater diversity.

Sample-Wise Interpretation:

- ***S-11*** → Highest richness and abundance balance, indicating a highly diverse microbial community.
- ***S-6*** → Moderate richness and evenness, showing a diverse microbial community.
- ***S-7*** → Lower richness; microbial diversity is not as high as S-11 and S-6.
- ***S-5*** → Lowest richness, meaning a few bacterial species dominate this sample.
- ***S-10*** → Moderate diversity, but not as high as S-11.

Unit Index (Overall Microbial Diversity Measure)

It shows:

- A general measure of microbial diversity across all samples.

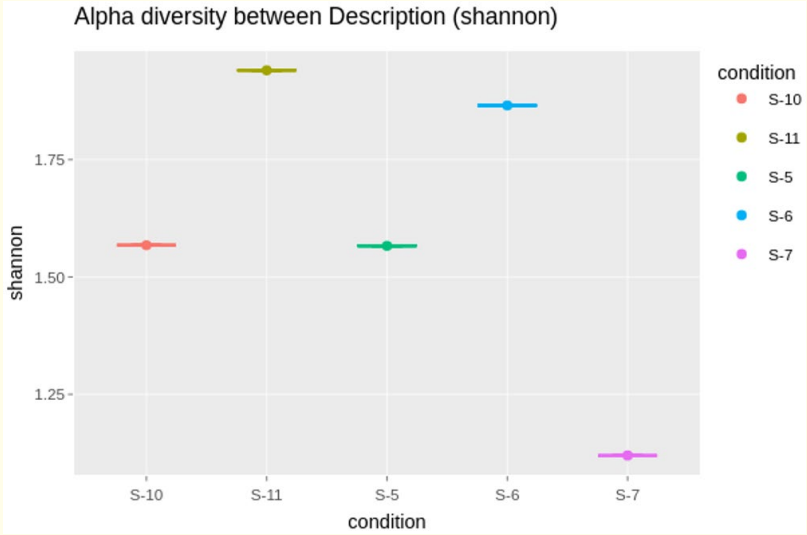


Figure 13: Alpha diversity analysis using the Shannon method, Shannon index analysis representing microbial richness and evenness across samples. S-11 has the highest diversity, while S-5 shows the lowest, suggesting microbial dominance by specific taxa like Leptospira. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Shannon index measures both species richness and evenness, providing insights into microbial community diversity, where higher values indicate a more diverse and evenly distributed community.

Sample-Wise Interpretation

- ***S-11*** → Highest microbial diversity overall.
- ***S-6*** → High microbial diversity but slightly lower than S-11.
- ***S-7*** → Moderate microbial diversity.
- ***S-10*** → Moderate microbial diversity, slightly lower than S-7.
- ***S-5*** → Lowest microbial diversity, dominated by only a few species.

The above table reflects:

1. ***S-11 (Varhaldevi lake soil, pumping site) has the highest microbial diversity and balance***, meaning more species and an evenly distributed community.
2. ***S-5 (Inlet Discharge water) has the lowest microbial diversity***, meaning fewer species dominate, likely due to Leptospira.
3. ***S-6 and S-10 have moderate diversity***, indicating a well-populated microbial community but not as diverse as S-11.
4. ***S-7 has lower diversity***, suggesting dominance by specific microbial species.

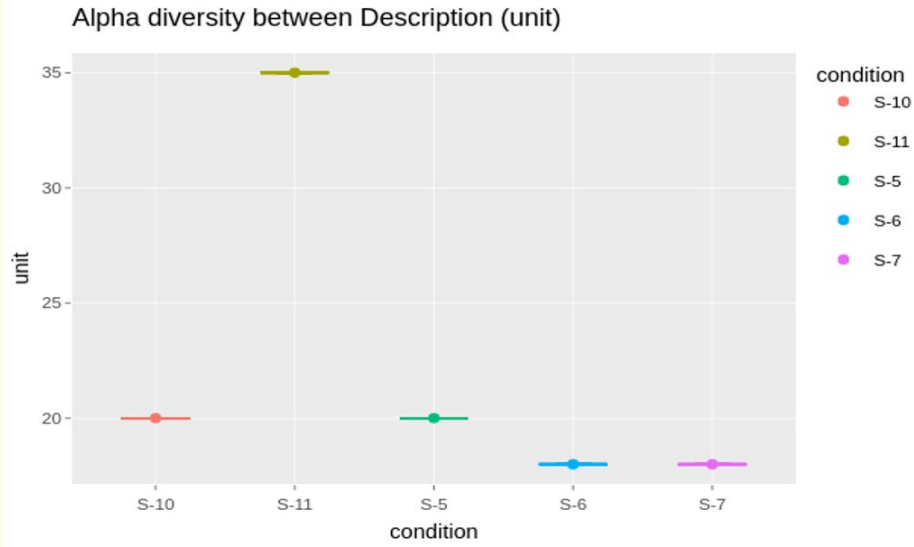


Figure 14: Alpha diversity analysis using the Unit method analysis representing overall microbial diversity across samples. S-11 has the highest microbial diversity, while S-5 has the lowest, likely due to the dominance of a few bacterial species, including Leptospira. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Unit method assesses microbial diversity within each sample, offering insights into species richness and distribution patterns.

Sample ID	***Richness (Chao1, Shannon, Inverse-Simpson)***	***Evenness (Gini-Simpson, Shannon)***	***Overall Diversity (Unit Index)***
S-11	Highest richness (most diverse microbial community)	Most evenly distributed species	Highest microbial diversity
S-6	Moderate richness	Well-balanced evenness	High microbial diversity
S-7	Lower richness	Uneven microbial distribution	Moderate diversity
S-10	Moderate richness	Some dominant species present	Moderate diversity
S-5	Lowest richness (few species)	Most uneven distribution (dominated by few species)	Lowest diversity

Table 12: Final Summary of Alpha Diversity Across All Samples.

Conclusion of alpha diversity

- S-11 is the most diverse and well-balanced microbial community.
- S-5 is the least diverse, likely due to the dominance of Leptospira.
- These diversity metrics help understand microbial composition and environmental impact in different sites.

Beta Diversity Analysis Across All Samples

Beta diversity is a crucial aspect of ecological studies, especially in microbial ecology, as it measures the differences in species composition between different communities. It measures or assesses microbial differences between samples by comparing microbial compositions. Different methods measure these variations:

- ***Unweighted UniFrac (Presence/Absence-Based Comparison)*** – Measures microbial differences by focusing on the presence or absence of species in different samples, ignoring abundance. Measures microbial differences between samples based only on the species present, without considering abundance.
- ***Weighted UniFrac (Abundance-Based Comparison)*** – Considers both species presence and abundance, giving more weight to dominant species. Samples with completely different microbial compositions will be far apart, while similar ones will be close.

Sample-Wise Interpretation

- ***S-11 (Varhaldevi lake soil, pumping site)*** → Most distinct sample, indicating a unique microbial composition compared to others.
- ***S-6 (Varhaldevi lake water, pumping site)*** → Moderately different; shares some species with other samples but has unique ones too.
- ***S-7 (Kamvari Nadi, Eidgah site)*** → More similar to S-6 and S-10, meaning they share microbial species.
- ***S-5 (Inlet Discharge water)*** → Least similar to other samples; its microbial composition is very different due to the presence of Leptospira.
- ***S-10 (Inlet Discharge soil)*** → Shows some similarity to S-6 and S-7, but still distinct from S-5 and S-11.

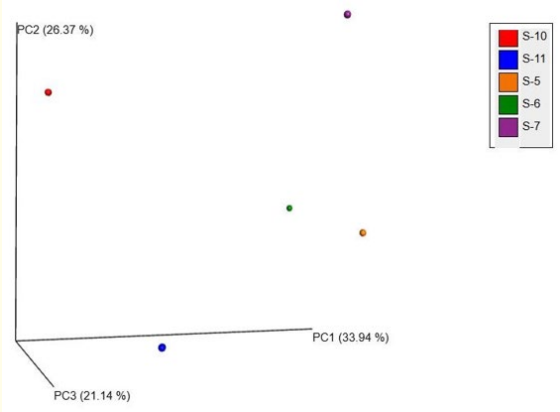


Figure 15: Beta diversity analysis using Principal Coordinate Analysis (PCoA) with the Unweighted UniFrac distance method, representing microbial community differences across various environmental samples. Beta diversity analysis using the Unweighted UniFrac method, comparing microbial compositions between samples based on species presence or absence. S-11 appears most distinct, indicating a unique microbial composition, while S-5 is the least similar to others, likely due to Leptospira dominance. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Unweighted UniFrac distance method considers phylogenetic differences between microbial communities, emphasizing the presence or absence of taxa rather than their abundance, allowing for a detailed comparison of community composition across different environments.

Beta Diversity - Weighted UniFrac (Focuses on Presence & Abundance)

- Measures microbial differences between samples by considering both species presence and abundance.
- Dominant species have a stronger influence on the results.

Sample-Wise Interpretation

- ***S-11*** → Most diverse, but closer to samples with abundant species in common (e.g., S-6).
- ***S-6*** → Shares microbial profiles with other samples, but its microbial balance differs slightly.
- ***S-7*** → More similar to S-10, suggesting they have some common species in abundance.
- ***S-5*** → Highly distinct due to Leptospira dominance; fewer shared species with other samples.
- ***S-10*** → Moderately diverse, sharing characteristics with both S-6 and S-7.

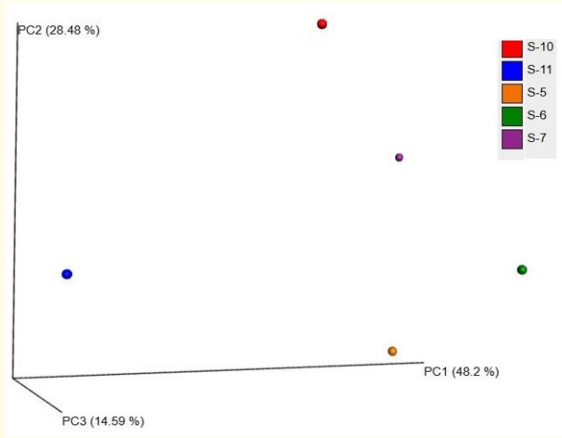


Figure 16: Beta diversity analysis using Principal Coordinate Analysis (PCoA) with the Weighted UniFrac distance method, representing microbial community differences across various environmental samples. Beta diversity analysis using the Weighted UniFrac method, comparing microbial compositions based on both presence and abundance. S-11 shows the highest diversity, while S-5 remains distinct due to the dominance of specific bacterial species like Leptospira. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Weighted UniFrac distance method considers both the phylogenetic differences and the relative abundance of taxa, providing a comprehensive comparison of microbial community structures across different environmental sites.

Sample ID	*Similarity (Unweighted UniFrac)*	*Similarity (Weighted UniFrac)*
S-11	Most unique microbial composition	Highly diverse, but shares species with S-6
S-6	Moderately similar to other samples	Shares species with S-11 and S-10
S-7	Similar to S-10, but some distinct species	Moderate diversity, close to S-10
S-10	Shares some species with S-6 & S-7	Moderately diverse
S-5	Most distinct, very different from others	Least similar due to Leptospira dominance

Table 13: Beta Diversity Summary.

Key observations

1. S-11 has the most unique microbial community, followed by S-6.
2. S-5 is highly distinct, dominated by Leptospira.
3. S-7 and S-10 are more similar to each other than to S-5 or S-11.
4. Weighted UniFrac highlights species abundance differences, while Unweighted UniFrac focuses on composition shifts.

Rarefaction plot

A rarefaction plot in the context of 16S data analysis is a graphical representation used to explore and assess the richness of microbial diversity within a sample or a set of samples. It is a valuable tool in metagenomics and microbiome studies, helping to understand how effectively they have sampled the microbial community and whether additional sequencing effort is necessary.

It help to determine if enough sequencing was done to capture microbial diversity. Different richness metrics assess completeness:

1. ***PD Whole Tree (Phylogenetic Diversity)*** – Measures evolutionary relationships among species.
2. ***Observed OTUs (Species Count)*** – Counts the number of unique species.
3. ***Shannon Index (Diversity & Evenness)*** – Measures both richness and evenness.
4. ***Chao1 Index (Predicted Species Richness)*** – Estimates total species, including undetected ones.
1. PD Whole Tree (Evolutionary Diversity Measurement) show
    - Higher values indicate greater evolutionary diversity in microbial communities.

Sample-Wise Interpretation

- ***S-11*** → Highest evolutionary diversity, meaning a wide range of microbial species.
- ***S-6*** → Moderate phylogenetic diversity.
- ***S-7 & S-10*** → Lower diversity compared to S-11.
- ***S-5*** → Lowest diversity, as a few species (e.g., Leptospira) dominate.

Observed OTUs (Actual Species Count): Counts the number of unique microbial species detected in each sample.

Sample-Wise Interpretation

- ***S-11*** → Highest species richness.
- ***S-6 & S-10*** → Moderate species count.
- ***S-7*** → Lower than S-6 and S-10.
- ***S-5*** → Fewest species, dominated by specific bacteria.

Shannon Index (Richness & Evenness Together)

- A combined measure of species richness and their distribution balance.

Sample-Wise Interpretation

- ***S-11*** → Most diverse and evenly distributed species.
- ***S-6*** → Moderate richness and evenness.
- ***S-7 & S-10*** → Lower than S-6 but still diverse.
- ***S-5*** → Least diverse, with microbial dominance by a few species.

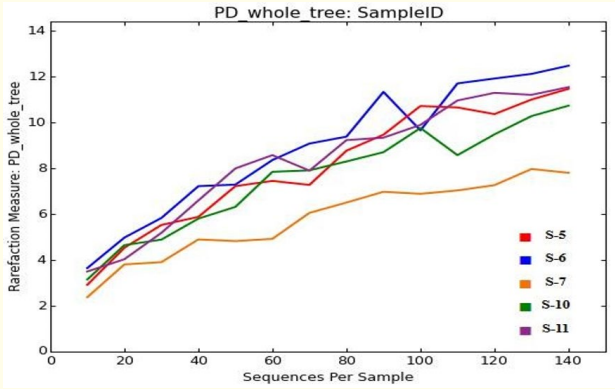


Figure 17: Rarefaction plot depicting PD whole tree richness estimates categorized by Sample ID. Rarefaction plot based on PD Whole Tree index, showing phylogenetic diversity across samples. S-11 has the highest evolutionary diversity, while S-5 shows the lowest, indicating a lack of phylogenetic variety. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The PD_whole_tree (Phylogenetic Diversity) measure evaluates the richness of microbial taxa based on the evolutionary tree, offering a view of the sample's phylogenetic diversity. The plot helps assess the completeness of sequencing efforts and whether additional sequencing would further capture microbial diversity within sample.

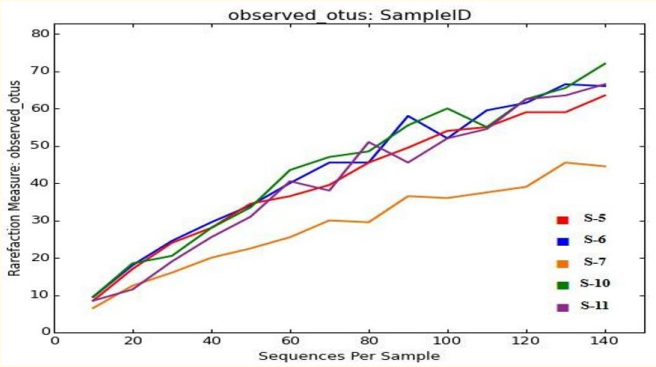


Figure 18: Rarefaction plot depicting Observed OTUs richness estimates categorized by SampleID. Rarefaction plot showing Observed OTUs across samples. S-11 has the highest species richness, while S-5 shows the lowest due to a few dominant species like Leptospira. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Observed OTUs measure captures the number of distinct operational taxonomic units (OTUs) detected in each sample, providing insights into the richness and diversity of microbial communities. The rarefaction plot helps assess the adequacy of sequencing depth and compares the microbial richness across the different sample.

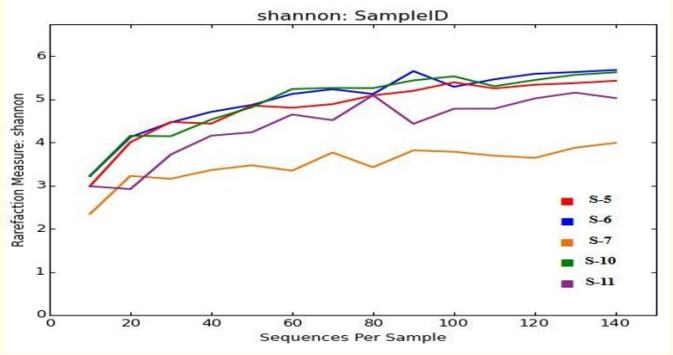


Figure 19: Rarefaction plot depicting Shannon richness estimates categorized by SampleID. Rarefaction plot using Shannon index, measuring microbial richness and evenness across samples. S-11 has the highest diversity, while S-5 has the lowest, indicating microbial imbalance. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Shannon index measures microbial diversity by considering both the richness and evenness of the species in the sample. The rarefaction plot helps visualize how the Shannon diversity index changes with sequencing depth, providing insights into the completeness of microbial diversity sampling for each environmental sample.

Chao1 Index (Estimated Total Species)

- Predicts the total number of species, including undetected ones.

Sample-Wise Interpretation

- ***S-11*** → Highest predicted species richness.
- ***S-6 & S-10*** → Moderate species estimates.
- ***S-7*** → Lower than S-6 and S-10.
- ***S-5*** → Lowest predicted richness.

PICRUSt analysis

PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) is a computational tool commonly used in microbiome research to predict the functional potential of microbial communities based on 16S rRNA gene sequencing data or other marker gene data.

STAMP analysis

STAMP (Statistical Analysis of Metagenomic Profiles) is a powerful bioinformatics tool commonly used for the analysis and interpretation of 16S rRNA gene sequencing data and other metagenomic datasets. STAMP provides a wide range of statistical and visualization techniques to explore and gain insights from microbial community composition and functional profiles.

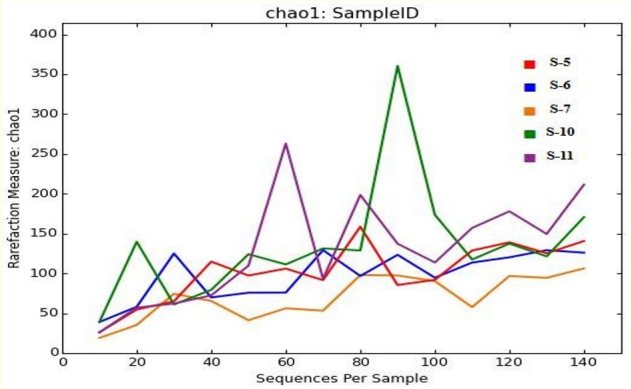


Figure 20: Rarefaction plot depicting Chao1 richness estimates categorized by SampleID. Rarefaction plot using Chao1 index, estimating total species richness in each sample. S-11 has the highest predicted microbial diversity, while S-5 has the lowest. The samples include Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge soil sample (S-10), Varhaldevi Lake (Water Pumping Site) soil sample (S-11), Kamvari Nadi (Eidgah Site) (S-7), Varhaldevi Lake Commissioner Bungalow Site Inlet Discharge water sample (S-5), and Varhaldevi Lake water (Water Pumping Site) sample (S-6). The Chao1 index estimates the total number of species (OTUs) present in a sample based on observed species and their frequencies, offering a prediction of microbial richness, especially for under-sampled communities. The rarefaction plot helps assess whether further sequencing could uncover additional species and ensures sufficient sampling depth for each sample.

The extended error bar plot generated by the STAMP tool serves as a valuable visualization for understanding the distribution and significance of microbial taxa or pathways across different conditions. By effectively conveying both mean values and variability, it enhances the interpretation of metagenomic data and supports the identification of biologically relevant patterns that warrant further exploration.

The overall findings of the Varal Devi and Kamvari river study is summarized below.

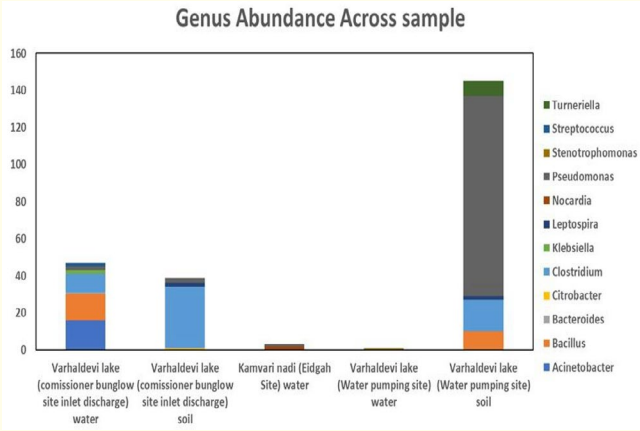


Figure 21: Genus abundance across the samples shows dominance of Leptospira.

Leptospira-Specific breakdown is evident with the findings. Leptospira, a bacterium from the phylum Spirochaetes, is known for its pathogenic potential and environmental persistence in water and soil. This section provides a detailed taxonomic breakdown, abundance across samples, and insights into its ecological and health significance.

Taxonomic Level	*Leptospira Classification*	*Presence in Samples*	*Key Observations*
Phylum	Spirochaetes	S-5, S-11	Detected at low relative abundance (~0.1%-0.2%).
Class	Leptospirae	S-5, S-11	Low abundance suggests localized presence.
Order	Leptospirales	S-5, S-11	Environmental persistence seen in S-11 (soil).
Family	Leptospiraceae	S-5, S-11	Found in water (S-5) and soil (S-11), potential contamination indicator.
Genus	Leptospira, Turneriella	S-5: Leptospira (1 count), S-11: Turneriella (8 counts), Leptospira (1 count)	S-11 shows a stronger presence of Turneriella, an environmental relative of Leptospira.

Table 14: Leptospira Taxonomic Breakdown Across Samples.

Sample ID	*Leptospira Presence*	*Relative Abundance*	*Key Observations*
S-5 (Inlet Discharge Water)	Present	0.2%	Leptospira detected at low levels, suggesting potential contamination.
S-11 (Lake Soil, Pumping Site)	Present	0.2%	Turneriella (8 counts) & Leptospira (1 count) indicate Leptospiraceae persistence in soil.
S-6, S-7, S-10	Absent	0%	No Leptospira detected, suggesting contamination is localized to S-5 and S-11.

Table 15: Leptospira Abundance Per Sample.

Interpretation of Leptospira Findings

Why is Leptospira Important?

Leptospira is associated with leptospirosis, a zoonotic disease transmitted through contaminated water, often linked to sewage runoff, agricultural activity, or animal waste. Even at low levels, its presence indicates a potential contamination source.

Key Observations from the Data

1. Leptospira is detected only in two samples (S-5, S-11), suggesting localized contamination rather than widespread distribution.
   - S-5 contains Leptospira in water, a potential risk indicator for human exposure.
   - S-11 contains both Leptospira and Turneriella, are both Leptospiraceae family showing soil acts as a natural reservoir for this bacterial family.
2. Absence in S-6, S-7, and S-10 suggests limited environmental spread.
   - S-7 (Kamvari Nadi), a flowing water source, does not contain Leptospira, suggesting no direct sewage contamination.
   - S-10 (Inlet Discharge Soil) lacks Leptospira, meaning contamination is not uniform across soil samples.
3. Turneriella (related to Leptospira) is more abundant in S-11, suggesting soil stability.

- Turneriella is not a known pathogen but shares a taxonomic relationship with Leptospira, indicating environmental adaptation rather than active pathogenic contamination.

4. NSTI scores suggest moderate taxonomic confidence in Leptospira detection.
   - NSTI score for S-5 = 0.196, indicating a moderate match with reference genomes.
   - NSTI score for S-11 = 0.244, suggesting a potential presence of novel or less-characterized Leptospira species.

Potential Environmental and Health Risks

1. ***Water Contamination Risk (S-5)***
   - Leptospira detected in S-5 (Inlet Discharge Water) could indicate runoff from contaminated sources (e.g., sewage, animal waste, agricultural runoff).
   - Even at low abundance, Leptospira can multiply under suitable conditions, leading to increased public health risks.
2. ***Soil Reservoir Potential (S-11)***
   - Leptospira found alongside Turneriella in S-11 suggests that the soil environment could act as a long-term microbial reservoir.
   - Turneriella's higher abundance (8 counts vs. 1 count for Leptospira) suggests environmental adaptation rather than an active outbreak.
3. ***Lack of Detection in Other Samples Suggests Localized Contamination***
   - No Leptospira was found in S-6, S-7, or S-10, suggesting contamination is site-specific rather than widespread.
   - This limits the risk of large-scale contamination but highlights the need for targeted monitoring at S-5 and S-11.

Sr. No	***Test***		***Result***
	Microbial Contamination	***Media Used***	
1.	Total Aerobic Count CFU/ML	Nutrient Agar	$17*10^6$
2.	Total Coliform	CLED Media	Present
3.	*E.Coli*	EMB Media	Present++
4.	*Salmonella*	Salmonella- Shigella Media	Absent
5.	*Pseudomonas*	Cetrimide Agar	Present
6.	*Enterococci Spp.*	CLED Media	Present
7.	*Faecal Coliform*	MAC Media	Present
8.	*Candida Spp.*	Chloramphenicol-Yeast mannitol	Absent
9.	*Staphylococcus Spp.*	NA+7.5% NACL	Absent
10.	*Fungal Spp.*	Sabouraud's Agar	Absent
11.	Facultative aerobes and Anaerobic	Fluid thio-glycolate broth	Aerobic
12	MPN/100ML	Lactose Broth	9/100ml

Table 16: Tap water analysis of Varaldevi lake water supply.

Genus	***S-10 (Soil)***	***S-11 (Soil)***	***S-7 (Water)***	***S-5 (Water)***	***S-6 (Water)***
Methanobacterium	2	12	0	0	0
Methanolinea	1	4	0	0	0

Table 17: Hydrogenotrophic Methanogens (Sample-Wise Presence & Abundance).

- Methanobacterium and Methanolinea were detected only in soil samples (S-10, S-11).
- No hydrogenotrophic methanogens were found in water samples (S-7, S-5, S-6).

Sample	*Syntrophic Organism Percentage (%)*
S-10 (Soil, Inlet Discharge)	0.000%
S-11 (Soil, Pumping Site)	0.021%
S-7 (Water, Kamvari Nadi)	0.000%
S-5 (Water, Inlet Discharge Water)	0.014%
S-6 (Water, Pumping Site)	0.000%

Table 18: Syntrophic Organisms Percentage per Sample.

- Syntrophic organisms (methanogenic archea bacteria) that perform anaerobic fermentation were detected at very low percentages in S-11 (0.021%) and S-5 (0.014%).
- The syntrophic bacteria detected in the samples include: *Syntrophobacter, Syntrophomonas, Syntrophus, Pelotomaculum, Desulfovibrio, Desulfotomaculum, Desulfobacter, Desulfobulbus, Desulfuromonas.*
- No syntrophic organisms were found in S-10, S-7, or S-6.

Risk Assessment of Pathogenic Bacteria Across Samples

This risk assessment categorizes each sample based on the type and abundance of pathogenic bacteria, their potential health and environmental risks, and suggested remediation actions.

Risk Categories:

- 🔴 High Risk (Significant health hazards, likely contamination).
- 🟠 Moderate Risk (Pathogens present, but lower risk for direct harm).
- 🟢 Low Risk (Minimal pathogenic bacteria, safer environment).

Sample ID	*Pathogenic Bacteria*	*Abundance (%)*	*Risk Level*	*Potential Risks & Implications*	*Possible Causes*
S-11 (Varhaldevi Lake, Water Pumping Site, Soil)	Leptospira (0.2%) Campylobacter (0.2%) Pseudomonas (2.8%) Xanthomonas (2.1%) Legionella (0.3%)	High 🔴	Severe Risk: - *Leptospira:* Zoonotic disease, kidney/liver failure - *Campylobacter:* Severe diarrhea, foodborne illness - *Legionella:* Legionnaires' disease - *Pseudomonas:* Opportunistic infections	Contaminated water, sewage runoff, animal waste	🚨 Urgent Action Needed - Restrict public access - Test water for E. coli, coliforms - Implement wastewater treatment & filtration - Monitor for animal/human contamination
S-5 (Varhaldevi Lake Commissioner Bungalow, Inlet Discharge, Water)	Leptospira (0.1%) Xanthomonas (14.0%) Pseudomonas (1.1%) Enterobacter (0.2%)	Moderate-High 🔴🟠	Increased Contamination Risk: - *Leptospira:* Zoonotic risks - *Xanthomonas:* Agricultural contamination - *Pseudomonas & Enterobacter:* Waterborne infections	Sewage contamination, agricultural runoff, wastewater discharge	⚠️ Moderate Intervention Required - Improve wastewater filtration - Reduce industrial/agricultural runoff - Increase monitoring of waterborne pathogens

S-6 (Varhaldevi Lake, Water Pumping Site, Water)	Legionella (0.4%) Pseudomonas (1.1%) Aeromonas (0.9%)	Moderate 🟠	Health Risk from Aerosolized Water: - *Legionella:* Risk of pneumonia (Legionnaires' disease) - *Aeromonas:* GI infections, wound infections	Biofilm formation in pipes, stagnant water, warm temperatures	🚰 Precautionary Actions Needed - Disinfect water systems - Monitor for Legionella outbreaks - Improve water circulation & aeration
S-10 (Varhaldevi Lake Commissioner Bungalow, Inlet Discharge, Soil)	Hydrogenophilales (8.2%) Pseudomonas (3.4%) Enterobacter (0.3%)	Moderate 🟠	Wastewater Contamination Risk: - *Pseudomonas & Enterobacter:* Soil contamination, potential human infections	Industrial pollution, sewage discharge	🏭 Environmental Clean-Up Needed - Soil bioremediation - Industrial wastewater monitoring
S-7 (Kamvari Nadi, Eidgah Site, Soil)	Pseudomonas (0.2%) Sphingomonas (0.1%)	Low 🟢	Minimal Pathogen Load: - *Pseudomonas:* Potentially opportunistic infections in immunocompromised individuals	Minimal contamination, relatively clean	✅ Safe for General Use - Continue monitoring - Prevent future contamination

Table 19: Sample-Specific Risk Evaluation.

Overall Risk Summary

🔴 ***High-Risk Sites***

- ***S-11 (Varhaldevi Lake, Soil)*** → High contamination from wastewater & animal waste
- ***S-5 (Inlet Discharge Water)*** → High agricultural/industrial contamination

🟠 ***Moderate-Risk Sites***

- ***S-6 (Water Pumping Site)*** → Presence of Legionella (respiratory risk)
- ***S-10 (Commissioner Bungalow, Soil)*** → Sewage/industrial waste contamination

🟢 ***Low-Risk Site***

- ***S-7 (Kamvari Nadi, Soil)*** → No major contaminants detected

Action Plan for Risk Mitigation

Risk Level	***Recommended Actions***
🔴 High Risk	🚨 Urgent intervention needed: - Improve wastewater treatment - Restrict access to contaminated areas - Test for E. coli, coliforms, heavy metals - Remove biofilms from water systems

🟠 Moderate Risk	⚠ Precautionary actions needed: - Increase monitoring of microbial load - Improve sewage & industrial waste disposal - Treat biofilm formation in pipes
🟢 Low Risk	✅ Maintain safety measures: - Continue periodic water quality checks - Prevent future contamination

Conclusion

Metagenomics tool was first time used to study sewage dynamics in contaminated Varal devi lake and Kamvari river in Bhiwandi city, Thane dist. MS, India as Varal devi lake supply 5 MLD potable water and may have adverse consequences. This lake has 4 inlet discharge of raw sewage. Findings of soil and water samples metagenomics study revealed potential environmental and health risk as the lake and river water showed presence of novel or less-characterized Leptospira and Turneriella species dominant in the system localized in Varal devi lake pumping and inlet discharge sites. Hydrogenotrophic Methanogens were detected in soil samples, where as it was absent in water. A detailed microbiome taxonomic breakdown, abundance across samples, and insights into its ecological and health significance is comprehended with the recommendations to take appropriate measures of nature based strategies for ecological restoration of the lake and riverine ecosystem. Based on the findings of Varaldevi lake or Kamvari river water and sediment samples, water quality is declared to be unfit for drinking or agriculture purposes.

Sewage is a climate linked microbial science to protect public health. Any talk of water and nutrient recovery is a proof of sewage illiteracy. Faeces (animal, humans, gut refuse) is a product of an evolutionary design for energy harvest following the redox potential staircase, scripted in (1) molecular language as per drawings (2) prepared by microbial genome.

Both display the highest level of complexity and synergy. In fact, the rules laid by the inorganic chemistry, animal gut provided the sanctuary for ancient microbial communities. For any successful STP, hardware employed must recognize and facilitate these communities to be effective. Unfortunately, world over STPs do not provide. Over dependence on hardware sophistication even if successful (?) and claim sustainable treatment/ development is a MYTH.

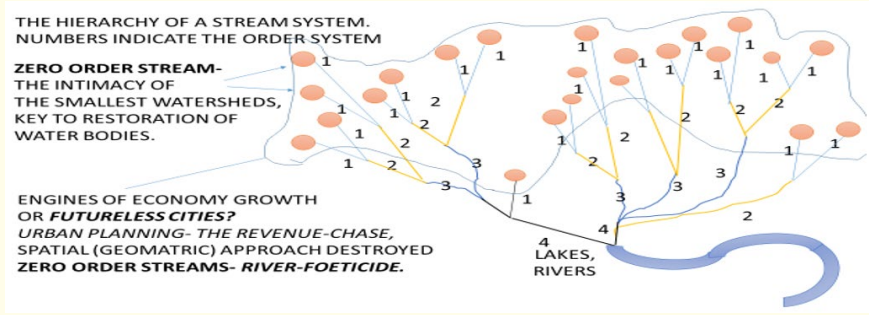
Conventional urban sewage management is founded on gravity (sedimentation) and geometry, an outcome of neurons at work- intelligence. Boulders (apples, almonds- high ORP)) to gravel to sand (acetate- gut discard sludge, low ORP status) a degradation following 2nd law of Thermodynamics. Defying this law, ASP (activated sludge process) variants convert those sand (acetate) particles to New Empire state buildings (far more complex molecules with high ORP ie acetate + aminoglycans). In short, conventional treatment is excreta acetate to faecal sludge (acetate + aminoglycans) ultimately transferred to landfills for final breakdown. That is trillions of litters of water as poop carriage, billions of dollars of expenditure, and millions of KW of energy is conventional SEWAGE MANAGEMENT. On top of it no pathogen elimination in sight. The ultimate result is repeated pandemics; a coincidence after coincidence in name of a coincidence IS NOT A COINCIDENCE. An ideal mechanism to render human society eternally poor.

Sewage not entering in the STP is THE CHALLENGE its management is facing. SBR, MBR, MMBR, all are on a flawed path as they use common SEEDs organisms, which are not colonic in origin. Prokaryotes are the major layers in sewage degradation (Metcalf Eddy, 2017) and Although Oxygen is poorly soluble in water, aeration is being done unnecessarily costing heavily on the operations and maintenance and obstruction in degradation process as well. Thus, a decentralized sewage treatment by septic tanks interception and Advanced Soil biotechnology (graded gravel system) /constructed wetlands is suggested, simultaneously with Kamvari river rehabilitation from zero order stream onwards. As these methods most importantly reduces sludge production which causes transfer of pathogens in the landfills and aeration infrastructure cost, as occurring in the conventional system. Restoration of flood plains and

buffer zones are also necessary by action on the encroachments, for self-regulation mechanism of the nature to prevail. This paper is an attempt to create awareness among the stakeholders about the waterbodies with river basin approach and strengthen scientific communication for good water governance. Engineering infrastructures like conventional STPs lack understanding and design of biological concepts and operationalization, this has led to the deterioration of almost all waterbodies. Therefore, this paper emphasizes the imperatives to work with integrated approach for future demand and water resources management by handholding between the community and government agencies. This study recommends and explains that as an anaerobic digester, a Septic Tank provides an extension of colon to complete the process. Expressed at its best and easily demonstrable the microbial synergy in nature's most sophisticated anaerobic digester is a septic tank working on the photon- independent energy source. Genome based microbial dark matter in colon/ST (only BW) known as the 'black box' holds the secret of pulverizing complex organic molecules. Hence its study is recommended in future research work and to plan and design Nature based solutions.

Recommendations for Nature Based Solutions

1. Billions and trillions of micro-climate droplets make a zero order stream. Hence they need to be protected.



The connection between soil and rain water need to be understood. Water in pipe destroys the hydrology that maintains the homeostasis of this planet. Knowing hydraulics will never offer a solution for disturbed hydrology. In name of flood mitigation pushing rain water to sea and install desalination plant to obtain fresh water is not green, certainly not sustainable.

Zero order streams with each rain drop, a nano-reactor, have been destroyed by urban proliferation turning cities into a graveyard of human civilization. For those blinded by technology, zero order streams are those neonatal forms preceding 1st order streams not recorded by satellite imaging. Covering them with Asphalt lining is a river foeticide. As a result, we end up losing rivers, floodplains, lakes, ponds. Of course, pipes can divert all that water to a sea. But the wetland-block shown in a diagram should be able to explain the rate of reaction for the water mass of flood (that is 2, 3 feet of water logging over a few square miles of a city). And that is the critical part of the whole process which is beyond an AutoCAD generated drawing. A century ago London storm water disposal line was on 6 mm rain per hour. Climate change proposition (2012) was for 80 mm rain per hour, @ £ 19 billion. It was focused only on transfer of floodwaters with no respect for reactions between water, soil, and other components involved in the process. Thus resilient smart cities with green design is a myth.

Knowledge is inversely proportional to problems not solutions in today's world. Direction also none have.

2. Proper characterization of sewage and industry effluents before treatment and design sewage size of reactor. Mixed sewage characterisation and pre-treatment of chemicals before constructed wetland or any other method of treatment.
3. Understand climate linked microbiology, principles of anaerobic reaction from pyruvate to lactic acid.
4. Hydrology and geology affect ecology; hence this needs to be studied thoroughly.

References

1. Bolzonella David and Pavan Paolo. Biohydrogen (second Edition) Hydraulic Retention Time (HRT) Circular Economy and Sustainability (2019).
2. Georgios A Pavlopoulos., et al. "Unraveling the functional dark matter through global metagenomics". Nature 622 (2023): 594-602.
3. Javier Mateo-Sagasta, Liqa Rschid-Sally and Anne Thebo. "Global wastewater and sludge production, treatment and use". Table of Municipal wastewater production and treatment with the largest urban population data source AQUASTAT 2014, GWI 2014 (2015).
4. Kaiser Gary. https://bio.libretexts.org 7.8: The Endosymbiotic Theory – Biology (2023).
5. Robert Seviour and Per H Nielsen. "Microbial Ecology of Activated Sludge". IWA (2010).
6. https://www.business-standard.com/india-news/even-originating-point-of-ganga-polluted-by-stp-discharge-govt-to-ngt-124111000408_1.html
7. Marina Albert. "The dynamics of Urban ecosystem. Advanced in Urban ecology". Springer (2008): 8.
8. George Tchobanoglous, Franklin L Burton and H David Stensel. "Wastewater Engineering- Treatment and Reuse". Metcalf & Eddy Inc. Forth Edition McGraw Hill publication (2017): 555.
9. Si Chen and Zhiyou Wen. Advances in bioenergy (2021).
10. Robert Sevious and Per Nielsen. IWA publishing. Microbial Ecology of Activated Sludge (2010): 95 & 120.
11. Ronald Droste and Ronald Gehr. Theory and Practice of Water and Wastewater Treatment Second publication. 2019. Read pg 48. Activated sludge process, pg. 571, about BOD COD TOC pg 573 Thymidine tagging and pg 727 Anaerobic Wastewater treatment (2019).
12. Yulin Zhang., et al. "The microbial dark matter and "wanted list" in worldwide wastewater treatment plants". Springer 11.59 (2023).

Important links

1. What is eutrophication? Causes, effects and control. http://www.eniscuola.net/en/2016/11/03/what-is-eutrophication-causes-effects-and-control/
2. https://www.toledoblade.com/local/environment/2024/09/22/algal-blooms-are-now-in-all-50-states-as-well-as-other-parts-of-the-world/stories/20240919036
3. https://www.earthisland.org/journal/index.php/articles/entry/ten-years-on-algal-blooms-continue-to-plague-lake-erie
4. Enzyme and aeration solutions are discussed in this HABS FAQ: http://bit.ly/FAQHABS Harmful algal blooms.
5. Gene regulation study reports surprising results: Extensive regions of DNA belong to multiple gene switches, 2024 by University of Bonn. https://phys.org/news/2024-11-gene-results-extensive-regions-dna.html
6. Meet the Microbes That Breathe Nitrate Instead of Oxygen – And They're Everywhere By Max Planck Institute for Marine Microbiology (2025). https://scitechdaily.com/meet-the-microbes-that-breathe-nitrate-instead-of-oxygen-and-theyre-everywhere/
7. Plant-based alternative for harmful algal bloom mitigation discovered by Clarkson University. Using Moringa oleifera to combat the cyanobacterium that causes HABs—known as Microcystis aeruginosa (2025). https://phys.org/news/2025-01-based-alternative-algal-bloom-mitigation.html

Volume 8 Issue 6 June 2025