Nutrient and microbial water quality of the upper Ganga River, India: Identification of pollution sources.

Michael J Bowes[1*], Daniel S Read[1], Himanshu Joshi[2], Rajiv Sinha[3], Aqib Ansari[3], Moushumi Hazra[2], Monica Simon[2], Rajesh Vishwakarma[2], Linda K Armstrong[1], David J E Nicholls[1], Heather D Wickham[1], Jade Ward[1,5], Laurence R Carvalho[4], H Gwyn Rees[1]

[1] UK Centre for Ecology & Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK

[2] Indian Institute of Technology, Roorkee, 247667, Uttarakhand, India

[3] Indian Institute of Technology, Kanpur, 208016, Uttar Pradesh, India

[4] UK Centre for Ecology & Hydrology, Bush Estate, Penicuik, Edinburgh, Midlothian, EH26 0QB, UK

[5] British Geological Survey, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK

- Corresponding author. Michael Bowes, mibo@ceh.ac.uk

Abstract

The Ganga River is facing mounting environmental pressures due to rapidly increasing human population, urbanisation, industrialisation and agricultural intensification, resulting in worsening water quality, ecological status, and impacts on human health. A combined inorganic chemical, algal and bacterial survey (using flow cytometry and 16S rRNA gene sequencing) along the upper and middle Ganga (from the Himalayan foothills to Kanpur) was conducted under pre-monsoon conditions. The upper Ganga had total phosphorus (TP) and total dissolved nitrogen concentrations of less than 100 μg l^{-1} and 1.0 mg l^{-1}, but water quality declined at Kannauj (TP = 420 μg l^{-1}) due to major nutrient pollution inputs from human-impacted tributaries (principally the Ramganga and Kali Rivers). The phosphorus and nitrogen loads in these two tributaries and the Yamuna were dominated by soluble reactive phosphorus and ammonium, with high bacterial loads and large numbers of taxa

indicative of pathogen and faecal organisms, strongly suggesting sewage pollution sources. The high nutrient concentrations, low flows, warm water and high solar radiation resulted in major algal blooms in the Kali and Ramganga, which greatly impacted the Ganga. Microbial communities were dominated by members of the Phylum Proteobacteria, Bacteriodetes and Cyanobacteria, with communities showing a clear upstream to downstream transition in community composition. To improve the water quality of the middle Ganga, and decrease ecological and human health risks, future mitigation must reduce urban wastewater inputs in the urbanised tributaries of the Ramganga, Kali and Yamuna Rivers.

1   Introduction

The Ganga River in northern India (known internationally as the Ganges) is one of the world's major rivers, with 40% of the Indian population (approximately 540 million people) living within its basin. India has experienced rapid economic development over recent decades, resulting in the Ganga River facing mounting environmental pressures due to increasing human population, urbanisation, industrialisation, water demand and agricultural intensification (Pandey et al., 2016; Yadav and Pandey, 2017a; Misra, 2011). To combat these growing pressures, the Ganga Action Plan (GAP) was introduced in 1985. The primary focus of the GAP was to intercept and treat raw effluent before reaching the Ganga, targeting urban wastes from 59 cities and major towns along the Ganga River (National River Conservation Directorate, 2009). To date, the building of environmental infrastructure in India such as sewage treatment works (STW), and the establishment of pollution permit regulations, regulatory monitoring and enforcement has not kept pace with the rate of urbanisation and industrialisation, resulting in continuing major water quality problems that affect both environmental and human health. Approximately 3000 million litres of urban wastewater still pour into the Ganga each day, with only 25 % of this treated by STWs (Jin et al., 2015).

Regulatory water quality analysis in India is predominantly limited to pH, dissolved oxygen (DO), biochemical oxygen demand (BOD) and faecal coliform counts, usually at monthly temporal resolution or less, from 14 sites along the Ganga, available from the Central Pollution Control Board (Mariya et al., 2019). These data provide a framework for classifying general levels of pollution across the Ganga catchment but there are some problems associated with these determinands. BOD and faecal coliform counts are time-consuming analytical techniques, and therefore costly. BOD only gives a general indication of gross organic loadings within the river (Jouanneau et al., 2014), but little information about the sources of this organic pollution. DO concentrations are known to vary greatly through the diurnal cycle, particularly in rivers with excessive phytoplankton biomass (Halliday et al., 2015), typically varying from oversaturated in the late afternoon to low DO concentrations in the early hours of the morning. Therefore, a single DO measurement during the day does not indicate potential

problems of anoxia overnight and is not a particularly effective means of assessing water quality status (Skeffington et al., 2015).

However, the fundamental data required to identify and quantify pollution sources have simply not been generally available at the spatial and temporal coverage required by Indian catchment managers, academics and legislators to allow them to manage and ultimately reduce this pollution loading in a cost-effective manner (Khan et al., 2016b; Yadav and Pandey, 2017a; Jin et al., 2015). These vital water quality parameters, routinely measured in more economically developed countries, such as phosphorus and nitrogen species, dissolved organic carbon, metals, pesticides, chlorophyll and major anion concentrations, are now beginning to be captured by recent Indian academic studies, but the data is usually aggregated to mean and range data, either across the seasons or from all monitoring sites along the river (Yadav and Pandey, 2017b; a; Khan et al., 2016b; Matta et al., 2017; Jin et al., 2015; Mariya et al., 2019; Agarwal et al., 2015), thereby losing spatial and temporal resolution. The presentation of water quality data from specific sampling dates are currently extremely rare in Indian river research, but starting to be presented in recent papers (Sen et al., 2018; Khan et al., 2016a; Pandey and Yadav, 2017). However, the situation in India is improving rapidly, with a wider range of water quality parameters now been measured by the regulatory authorities, including sub-daily data from in-situ water quality probes. Such data is vitally important for increasing system understanding, knowledge of how pollution sources change under a range of flow regimes (Bowes et al., 2014; Tappin et al., 2016), and how nutrients, water quality and ecology interact (Bowes et al., 2016; Hardenbicker et al., 2014).

Faecal coliform data is perhaps the most important water quality parameter in the current Indian context, as it has a direct link to human health, and Indian religious bathing customs (called Kumbh) bring people into contact with river water like nowhere else on Earth (Matta and Bisht, 2018). This routine regulatory data is now being augmented by the academic community in recent years, through the utilisation of 16S rRNA gene sequencing (Jani et al., 2018b; Jani et al., 2018a; Dixit et al., 2017; Zhang et al., 2018), but more of this in-depth bacterial research is urgently needed to better understand the composition of microbes in the Ganga and its tributaries. Despite the large phosphorus and nitrogen loadings to the Ganga River, from untreated urban wastewaters and agricultural practises, studies of the effects of eutrophication, particularly algal and cyanobacterial blooms, are also limited (Tare et al., 2003; Dixit et al., 2017; Yadav and Pandey, 2017b). Flow cytometry techniques have been developed in recent years to rapidly characterise phytoplankton communities and provide bacterial loadings in rivers (Read et al., 2014), and this could provide new insights into algal dynamics and bacterial pollution sources in the Ganga basin.

The key objectives of this study are to produce a detailed snapshot survey of inorganic chemical and microbial water quality along a >600 km stretch of this internationally important river and its major tributaries.

(1) This research will simultaneously evaluate both the inorganic chemical and ecological status of the rivers across this region, focussing on quantifying phosphorus and nitrogen species, and the characterisation of the phytoplankton community using flow cytometry. This will identify sites where excessive nutrient concentrations are resulting in ecological degradation associated with eutrophication, such as algal blooms and low dissolved oxygen concentrations.

(2) Bacterioplankton concentrations will be quantified using flow cytometry, and the bacterial community will be characterised using 16-S rRNA gene sequencing. This will establish whether flow cytometry could be used as a rapid and cheap proxy for determining pathogen loadings in India in the future.

(3) The survey will identify pollution hotspots across the catchment, indicating which river stretches are most heavily impacted and which tributaries or cities result in deterioration in the water quality of the Ganga. We will use the combined inorganic chemical, nutrient speciation and bacterial concentrations and characterisation data to identify potential pollution sources during the low-flow, pre-monsoon period when water quality is potentially at its worst. This knowledge should allow resources to be targeted to maximise the improvements in the water quality of the Ganga River.

(4) The monitoring data from his study will be made freely available, to promote future collaboration and data sharing within the Indian research community.

1.1   Study area

The Ganga basin covers an area of 1,086,000 km^2, and extends 2,525 km from its source at the Gangotri glacier in the western Himalayas, flowing south and then east through the Gangetic Plain to Bangladesh, where it discharges into the Bay of Bengal. This survey consists of 23 monitoring sites, and focusses on the stretch of the Ganga from Rishikesh to the city of Kanpur. The uppermost site (Site 1: Ganga upstream of Rishikesh) (Figure 1; Supplementary Figure 1a) is ca. 50 km downstream of the confluence of the Bhagirathi and Alaknanda rivers, from where the river is known as the Ganga. This site is situated at an elevation of ca. 400 m, in the foothills of the Himalayas, within steep sided valleys and gorges. Channel bed sediments are characterised by boulders, cobbles, pebbles and course sand (Sinha et al., 2017). The section between Rishikesh and Haridwar (Sites 2 and 3; Figure 1) represents the piedmont zone of the river, and is typified by the development of a narrow floodplain within this partly confined reach. The bed materials are principally coarse sands and silts, and some

alluvial islands and bars have formed within this reach. Haridwar (Site 3; Figure 1) marks the point at which the Ganga flows from the Himalayan foothills and onto the Gangetic plain (Sinha et al., 2017). Downstream of Haridwar to Kanpur (Sites 4 to 9; Figure 1; Supplementary Figure 1a), the Ganga occupies a wide active floodplain. Under the low flow conditions at the time of the survey, the channel is highly braided due to high sediment supply from the Himalayan sections, reduced gradient and lack of flow (mainly attributed to the diversion of much of the flow of the Ganga into the Upper Ganga Canal at Haridwar and over-extraction of groundwaters). Bed sediments mainly consist of coarse to fine sands. There are five major barrages along the upper and middle stretches of the Ganga, at Rishikesh, Haridwar, Bijnor, Narora and just upstream of Kanpur.

This study also surveyed the major tributaries entering the Ganga within this stretch, plus the Yamuna River and some of its tributaries, which join with the Ganga ca. 200 km downstream of Kanpur at Prayagraj. The region is typified by a tropical monsoonal climate, with three distinct seasons; a cooler winter (November to February), hot dry summer or pre-monsoon (March to June) and wet monsoon period (July to October). The average annual rainfall for the Ganga basin is 1100 mm, of which 80 % falls within the monsoon period (National River Conservation Directorate, 2009).

2   Materials and methods

2.1   Sampling survey

The entire sampling campaign took place between 25th March and 1st April 2018. Water samples were taken for both inorganic chemical and microbial analyses at a total of 23 river monitoring sites. There were nine sites along the upper and middle reaches of the Ganga River itself, extending from upstream of Rishikesh (Uttarakhand) in the foothills of the Himalayas, to the city of Kanpur in the Gangetic plains of Uttar Pradesh (Figure 1). The larger tributaries within this study section of the Ganga Basin (the Song, Kali, Pandu, Rind and Garra Rivers) were also sampled. In addition, the two largest tributaries, the Ramganga and the Yamuna River, were each sampled at three points along their course (Figure 1), plus the Hindon River, which is a tributary of the Yamuna, known to be heavily polluted through sewage and industrial effluents (Jain et al., 2007). The Ganga Canal was also sampled at Roorkee (30 km from the point where it splits from the main Ganga) and in Kanpur (520 km downstream of Roorkee). The dates that each monitoring site was sampled are given in Supplementary Table 1.

Samples of water were taken from the margin of each river, from a site with good flow velocity to ensure mixing. A 2-litre polyethylene bottle was rinsed three times with river water, and then filled from 5 cm below the surface, taking care not to disturb the bed sediment of the river. The bulk sample was immediately subsampled into a 60 ml acid-washed (20 % hydrochloric acid) polyethylene bottle

for total phosphorus analysis. Other sub-samples were filtered immediately in the field through a 0.45 µm cellulose nitrate (Whatman WCN grade; Maidstone, UK) membrane filter into 60 ml acid-washed bottles, for total dissolved phosphorus, soluble reactive phosphorus, nitrogen species (nitrate, nitrite, ammonium, total dissolved nitrogen), dissolved organic carbon, dissolved reactive silicon, and major anions analysis. All chemical samples were stored in the dark in an ice box during field sampling. Subsamples of water were also collected for flow cytometry analysis of both bacterioplankton and phytoplankton, by pipetting 2 ml of water into sterile 5 ml Nalgene screw-top tubes (Thermo-Fisher, UK), and stored during transportation in an ice box in the dark. Bacterioplankton samples were preserved on the day of sampling by adding 0.25 % glutaraldehyde and 0.01% Pluronic F68 surfactant (Marie et al., 1997; Zubkov et al., 1998), and then freezing at -20 °C. Samples for DNA sequencing were filtered into 0.22 µm pore size Sterivex[TM] enclosed filters (Millipore, UK) manually, using sterile, single use 60 ml syringes. River water from the 2-litre acid-washed bottle was pushed through the filter until resistance caused by filter blockage meant that it was impractical to continue. The volume of water filtered at each site was recorded. Biomass on the filter surface was preserved by adding 2 ml of Zymo DNA/RNA shield[TM] (Zymo, UK) and storing at 4°C until DNA extraction. The water temperature of the bulk river water sample was measured in the field using an ATP Multi-Thermo digital thermometer (ATP Instrumentation Ltd. Ashby-de-la-Zouch, UK). Conductivity, pH and oxidation-reduction potential were measured immediately using an Ultrameter (Myron L, Carlsbad, California, USA).

2.2   Laboratory analytical methods

All chemical samples were stored in the dark at 4°C, prior to analysis at the CEH Nutrient Laboratories, UK. All analyses were carried out immediately on return to the UK on 3rd April, which was between two and eight days after samples were taken. Total phosphorus (TP) and total dissolved phosphorus (TDP) were determined by digesting an unfiltered and 0.45 µm filtered water sample (respectively) with acidified potassium persulphate in an autoclave at 121°C for 45 min. Acidified ammonium molybdate reagent was then added to the digested samples to produce a molybdenum–phosphorus complex. This intensely blue-coloured compound was then quantified spectrophotometrically at 880 nm (Eisenreich et al., 1975). Soluble reactive phosphorus (SRP) concentrations were determined on a filtered (0.45 µm WCN-grade cellulose nitrate membrane; Whatman, Maidstone, UK) sample, using the phosphomolybdenum-blue colorimetry method of Murphy and Riley (1962) as modified by Neal et al. (2000), using a Seal Auto Analyser 3 (Seal Analytical; Fareham, UK). Ammonium concentration was determined using an indophenol-blue colorimetric method (Leeks et al., 1997) using a Seal Auto Analyser 3. Dissolved organic carbon and total dissolved nitrogen were analysed by thermal oxidation using an Elementar Vario Cube (Elementar Ananlysensysteme GmbH; Langenselbold, Germany). Major dissolved anion (fluoride, chloride, nitrite, nitrate and sulphate) concentrations were determined by

ion chromatography (Dionex AS50, Thermo Fisher Scientific; Waltham, USA). All analyses were carried out alongside Aquacheck quality control standards (LGC Standards, Teddington, UK).

The flow cytometry analytical protocol for phytoplankton is outlined in detail in Read et al. (2014). The method was used to identify five broad groups of phytoplankton based on phenotypic characteristics (size and pigmentation) using a Gallios flow cytometer (Beckman Coulter, High Wycombe, UK) equipped with blue (488 nm) and red (638 nm) solid state diode lasers. These groups were diatom / large chlorophytes, meso-chlorophytes, nano- and pico-chlorophytes, cryptophytes and cyanobacteria. Samples were spiked with a set volume of FlowCount (Beckman Coulter) counting beads, and run for five minutes per sample at a high flow rate. For bacterioplankton analysis, an aliquot of 0.5 ml from each sample was stained with SYBR Green I (Sigma-Aldrich, UK) at a final concentration of 1:1000 for 30 min at room temperature in the dark. An addition of 2.5 µl of 1 µm diameter beads (Thermo-Fisher, UK) to each sample was used as a calibration and counting standard. Each sample was run for 1 min at a low flow rate (5 µl per min) using excitation with a 488 nm laser. Data was processed using the software Kaluza Analysis v1.5a (Beckman Coulter, UK).

DNA was extracted from the filters using the Qiagen PowerWater™ kit (Qiagen, UK) following manufacturer's instructions. Amplification of the 16S rRNA gene was done using primers 515f GTGYCAGCMGCCGCGGTAA and 806r GGACTACNVGGGTWTCTAAT (Walters et al., 2016). Amplicons were generated using a high fidelity DNA polymerase (Q5 Taq, New England Biolabs). After an initial denaturation at 95 °C for 2 minutes PCR conditions were: denaturation at 95 °C for 15 seconds; annealing at temperature 55 °C for 30 seconds with extension at 72 °C for 30 seconds; repeated for 25 cycles. A final extension of 10 minutes at 72 °C was included. PCR products were cleaned using Zymo ZR-96 DNA Clean-up Kit following manufacturer's instructions. MiSeq adapters and 8nt dual-indexing barcode sequences were added during a second step of PCR amplification. After an initial denaturation 95 °C for 2 minutes PCR conditions were: denaturation at 95 °C for 15 seconds; annealing at temperatures 55 °C; annealing times were 30 seconds with extension at 72 °C for 30 seconds; repeated for 8 cycles with a final extension of 10 minutes at 72 °C. Amplicon sizes were determined using an Agilent 2200 TapeStation system. Libraries were normalized using SequalPrep Normalization Plate Kit (Thermo Fisher, UK), quantified using Qubit dsDNA HS kit (Thermo Fisher, UK) and pooled together at equal concentrations. The pooled library was diluted to achieve 400 pM in a 20 µl volume after denaturation and neutralisation. Denaturation was achieved with 2 µl 2N NaOH for 5 minutes followed by neutralisation with 2 µl 2N HCl. The library was then diluted to its load concentration of 9pM with HT1 Buffer and 10% denatured PhiX control library. A final denaturation was performed by heating to 96°C for 2 minutes followed by cooling in crushed ice. Sequencing was performed on Illumina MiSeq using V2 500 cycle reagents.

All analyses of the sequencing data were carried out in R (R Core Team, 2013). Raw data was processed using the DADA2 pipeline (Callahan et al., 2016) on forward reads only to generate Amplicon Sequence variants (ASVs). Filtering settings were maximum number of Ns (maxN) = 0, maximum number of expected errors (maxEE) = (3,5), and amplicon primer sequences removed using trimLeft=c(20,20). Sequences were dereplicated and the DADA2 core sequence variant inference algorithm applied. Only forward reads were used for downstream analysis. ASV sequence tables were constructed and chimeric sequences were removed using removeBimeraDenovo default settings. ASVs were subject to taxonomic assignment using assignTaxonomy at default settings; training databases were GreneGenes v13.8 (DeSantis et al., 2006). A Non-Metric Multidimentional Scaling (NMDS) ordination plot was created using the package Vegan (Oksanen et al., 2017) to visually examine the relationship of the bacterial community composition across sites. The function 'envfit' in Vegan was used to statistically test the relationship between the measured physicochemical parameters and the bacterial community ordination.

Peer reviewed, government and health organisation literature was reviewed to identify pathogenic and faecal organisms that are known to be associated with freshwater. Taxa were selected to represent two categories of bacterial contamination of water; potential bacterial pathogens with a known association with water and commensal organisms associated with mammalian guts. The full list of groups and their references are shown in Supplementary Table 1. The species-level phylotype table was then searched for these taxa by name, and the relative abundance of these groups plotted using R (R Core Team, 2013).

3 Results and discussion

3.1 Chemical water quality along the Ganga River transect

The changes in nutrient and anion concentrations along the Ganga River from upstream of Rishikesh to Kanpur, alongside contributing tributaries within this study reach, are shown in Figure 2. Raw data for all study sites are available in Supplementary Table 2.

3.1.1 Nutrients

3.1.1.1 Phosphorus

The total phosphorus concentrations in the upper Ganga were < 100 μg l^{-1} for the *ca*. 440 km stretch from upstream of Rishikesh to Farrukhabad (Site 1-7; Figure 1), with SRP concentrations ranging from

2 to 46 µg P l^{-1} (Figure 2). These concentrations of SRP are likely to be limiting for phytoplankton and periphyton growth (Bowes et al., 2012; McCall et al., 2017; Dodds, 2006).

There was a change in phosphorus speciation as the Ganga left the Himalayas and entered the Gangetic Plain. Approximately 20 % of the phosphorus load was in soluble reactive form at the upper site, upstream of Rishikesh (Site 1), but this rapidly reduced to 3 % as the river flowed through the towns of Rishikesh and Haridwar (Site 2-3). This could indicate that SRP was being rapidly sequestered by the high load of suspended sediment that was observed to be entrained within the water column within this fast-flowing reach. It could also indicate that SRP was being bio-accumulated by river macrophytes and benthic algae along this river stretch (Jarvie et al., 2012; Bowes and House, 2001). The proportion of TP in SRP form in the Ganga River increased to approximately 45 % along the 120 km reach from Balawali and Brajghat (Sites 4-5). This could be due to the inputs of sewage effluents from Haridwar wastewater treatment works (5 km downstream of the Ganga at Haridwar, Site 3), and other untreated sewage effluent inputs from the towns along this reach of the Ganga, which will be predominantly in SRP form (Jarvie et al., 2006).

The phosphorus concentrations of the Ganga River increased markedly between the towns of Farrukhabad and Kannauj (Sites 7 and 8, Figure 1). Total phosphorus concentration increased three-fold, from 140 to 420 µg l^{-1}. This was due to the significant inputs of nutrients from the three major tributaries that enter this 70 km stretch of the Ganga between these two sampling sites. The Garra River had a TP concentration of 196 µg l^{-1}, and the Ramganga and Kali Rivers had TP concentrations of >300 µg l^{-1} which were twice as high as the TP concentration of the Ganga at Farrukhabad. These TP concentrations are a little higher than high population density catchments in northern Europe, with the Rivers Thames (UK), Elbe and Rhine (Germany) having annual average TP concentrations of 222 µg l^{-1} (Bowes et al., 2018), 120 µg l^{-1} and 180 µg l^{-1} (Hardenbicker et al., 2014) respectively, despite these European rivers having much higher levels of sewage treatment. The relative lack of impact of the Ganga may be a product of the level of dilution that such a huge river provides, and also the length of the river, giving time for particulate phosphorus to be deposited on the bed within this depositional environment.

The proportion of the total P load in SRP form in the Ganga increased from 3 % to 25 % between Farrukhabad and Kannauj. Again this was primarily due to the significant tributary inputs, principally from the Kali and Garra Rivers (sites 15 and 16), with SRP concentrations of 87 and 133 µg l^{-1} respectively. The proportion of the TP load in SRP form for the Garra was particularly high (68 %) but lower in the Kali (28 %) and the lower Ramganga River (Site 14) (3.7 %). These two tributaries, the

Ramganga and Kali, were very green in colour at the time of sampling and significant filamentous algal biomass was visible along the margins of each river (see Supplementary Figure 1 (b) inset photograph). This phytoplankton / periphyton bloom would result in SRP being bio-accumulated by the biota, and the excessive phytoplankton growth resulting in SRP being transformed into particulate P which was accounted for as TP.

3.1.1.2   Nitrogen

The nitrogen concentrations in the Ganga River were consistently <1 mg-N l^{-1} along the upper Ganga as far as Farrukhabad (Sites 1-7, Figure 1), but then significantly increased between Farrukhabad and Kannauj (Site 8), with TDN concentrations increasing from 0.6 to 2.6 mg-N l^{-1} (Figure 3). This was predominantly due to the TDN inputs from the Kali River (2.1 mg-N l^{-1}) and to a lesser extent, the Ramganga River (1.0 mg-N l^{-1}). The nitrogen speciation also changed along the Ganga River continuum. The proportion of the total dissolved nitrogen load in nitrate form increased from 30 to 40 % between Rishikesh and Haridwar, probably due to the large input of nitrate from the Song River tributary (Site 10) (Figure 3a). The high nitrate concentration in the Song is probably due to the river's groundwater spring source (Joshi et al., 1995). Between Haridwar and Farrukhabad (Site 3-7), the nitrogen load became dominated by dissolved organic nitrogen, indicating the high rates of biological processing that were occurring within the reach at this time. There was a large input of nitrate to the Ganga from the Kali River, which increased the proportion of nitrate from 8 % to 75 % of the total dissolved N load (Figure 3) between the Ganga monitoring sites at Farrukhabad and Kannauj. This nitrate input was rapidly consumed and biologically processed within the 20 km reach of the Ganga between Kannauj and Nanamau Bridge, which was experiencing a major algal bloom at this time. These high rates of microbial activity may result in high rates of denitrification within the stretch (Chen et al., 2012).

3.1.1.3   Dissolved organic carbon

DOC concentrations along the upper Ganga ranged from 0.7 to 3.0 mg-C l^{-1} (Figure 2). As with phosphorus and nitrogen, DOC concentration increased markedly downstream of Farrukhabad (Site 7), again due to the inputs from the tributaries within this stretch. Their inputs resulted in the Ganga DOC concentrations increasing from 2.5 mg-C l^{-1} to 6.0 mg-C l^{-1} between Farrukhabad and Kannauj (Site 8). All three tributaries had significantly higher DOC concentrations than the Ganga at Farrukhabad, with the Ramganga, Garra and Kali Rivers having DOC concentrations of 9.5, 5.5 and 3.9 mg-C l^{-1} respectively.

3.1.2   Anions and conductivity

The chloride, fluoride and sulphate concentrations were relatively low through the upper Ganga stretch (upstream of Rishikesh to Brajghat) (Sites 1-6, Figure 1) (Figure 2). Chloride and fluoride

concentrations increased along the reach by 240 % and 30 % respectively, and conductivity doubled from 138 to 264 µS cm^{-1}.

The anion concentrations and conductivity of the Ganga greatly increased between Farrukhabad and Kannauj (Sites 7 to 8). Again, concentrations in the three tributaries joining the Ganga within this stretch were higher than the Ganga at Farrukhabad, resulting in an increase in concentration downstream. Increases in fluoride and sulphate were relatively modest (54 and 68 % respectively), but the chloride concentration in the Ganga increased five-fold from 6.0 to 31.5 mg-Cl l^{-1}, mainly due to inputs from the Ramganga and Kali Rivers (Sites 14 and 15). The conductivity of the Ganga within this reach also increased from 303 to 421 µS cm^{-1}, also due primarily to inputs from the Ramganga and Garra. The large increase in chloride concentration, relative to fluoride (which will be primarily from geological sources) (Misra and Mishra, 2007), and the increase in conductivity alongside the high nutrient concentrations (especially high ammonium concentrations) suggests that there are significant untreated sewage sources entering the Ganga from the Ramganga and Kali Rivers. Both these rivers receive significant domestic and industrial wastewater effluents from the major cities of Bareilly, Moradabad and Rampur (Ramganga) and Meerut and Hapur (Kali). Similar conclusions about the urban sources of pollution in these tributaries have been made elsewhere (Indian Institutes of Technology, 2010).

3.2   Inorganic water quality along the Ramganga River transect

Phosphorus concentrations were highest immediately downstream of the city of Bareilly (Site 12) (population of approximately 1.3 million people), with a TP concentration of 778 µg P l^{-1} and SRP concentration of 445 µg P l^{-1}, equivalent to 57 % of the TP load in soluble reactive form (Figure 4). This probably indicates that there is a large sewage pollution source from Bareilly and possibly from the other upstream urban and industrial centres of Moradabad and Rampur (Pathak et al., 2018). The TP concentration decreased to 310 µg P l^{-1} at the downstream monitoring site near Farrukhabad (Site 14). As there were no major tributary inputs within this section, this suggests that the phosphorus was being naturally attenuated through deposition of suspended sediment and algal loads within the lower Ramganga during this low-flow period, and SRP within the water column was also being sequestered by the river bed sediments (Bowes and House, 2001). This is further supported by the observation that the proportion of the TP load in SRP form reduced from 57 % to <4 % in the lower Ramganga.

Total dissolved nitrogen concentrations in the Ramganga River also decreased markedly, from 4.41 mg N l^{-1} to 0.97 mg N l^{-1} near the confluence with the Ganga near Farrukhabad (Figure 4). The dissolved nitrogen speciation of the Ramganga River at Bareilly consisted of 34 % ammonium, 23 % nitrate and 13 % nitrite (Figure 3b), which suggests that the main source on nitrogen was raw sewage wastes from

11

the city. The nitrogen concentrations at the two lower Ramganga sampling sites at Kunauli and near Farrukhabad (Sites 13 and 14) were both dominated by the dissolved organic nitrogen fraction, indicating extremely high organic processing rates, as supported by the observation of high algal biomass within this stretch of the river, and the low concentrations of bioavailable phosphorus (SRP) (Wetz et al., 2017).

The dissolved organic carbon concentrations were also highest immediately downstream of Bareilly (13.39 mg C l^{-1}), again indicating a large organic input to the river at this point. Further downstream, the DOC concentration of the Ramganga decreased to 9.55 mg C l^{-1} at the monitoring site near Farrukhabad (Figure 4).

There was relatively little change in the concentration of chloride, fluoride and sulphate along the monitored reach of the Ramganga. These relatively conservative ions suggest that there are no significant pollution sources or inputs along this stretch of river downstream of Bareilly, and that the observed changes in the nutrient concentrations are predominantly due to within-channel chemical and biological processing and dynamics.

3.3   Inorganic water quality along the Yamuna River transect

The two Yamuna River monitoring sites within Delhi (Central Delhi, Site 21 and Okhla Barrage, Site 22) had by far the highest nutrient concentrations within the entire study. Phosphorus concentrations at the central Delhi monitoring site were 4410 µg TP l^{-1} and 3860 µg SRP l^{-1} (Figure 4), (equivalent to 87 % of the TP load) which are similar concentrations and SRP proportion to sewage influents from towns in the UK (Neal et al., 2005). The total dissolved N concentrations at these Delhi monitoring sites were also extremely high (up to 33.4 mg N l^{-1}) and over 85 % of this nitrogen was in the form of ammonium (Figure 3b). Again, these nitrogen concentrations and speciation are typical of raw domestic sewage. The strong smell of raw sewage at these sites, plus the extensive layer of foam covering the majority of the river surface at the Okhla Barrage monitoring site at the time of sampling (Supplementary Figure 1b, photo D), strongly suggests that the predominant source of pollution in this stretch of the Yamuna is untreated domestic sewage and industrial inputs from the cities of New Delhi and Ghaziabad. Other studies have calculated that this short stretch of river through New Delhi receives 2,871 Ml d^{-1} of municipal wastes from 22 drains, including 218 Ml d^{-1} of industrial effluents (Trisal et al., 2008).

The predominance of ammonium and virtual absence of nitrate (making up only 0.1 % of the TDN load) resulted in extremely reducing conditions, with these sites having oxidation-reduction potential (ORP) values of ca. -300 mV (Supplementary Table 2). It is likely that carbon load within the Yamuna would be converted to methane under such reducing conditions, which may partially explain the 25 % reduction in DOC concentration between the two New Delhi monitoring sites.

The Hindon River enters the Yamuna approximately 32 km downstream of the Okhla Barrage monitoring site. At the monitoring site in Ghaziabad (Site 20, Figure 1), the Hindon was enriched with nutrients, with 690 μg l^{-1} of TP, and 7.8 mg N l^{-1} of total dissolved nitrogen (Figure 4). The site had the highest DOC concentration in the survey (20.3 mg C l^{-1}), which could be due to the high density of sugar factories and distilleries within the catchment (Jain et al., 2007).

The nutrient concentrations in the Yamuna River at Kalpi (approximately 650 km downstream of New Delhi) (Site 23, Figure 1) were much lower than those observed upstream in New Delhi and the River Hindon. Much of this reduction could be due to the major tributary inputs from the Chambal and Sind Rivers (Figure 1), but unfortunately it was not possible to monitor these rivers during this particular study. The major anion concentrations and conductivity at Kalpi were also greatly reduced (Figure 4), which does imply that there was a dilution of the Delhi pollution load by these tributaries and contributing groundwaters. However, the reductions in the relatively conservative chloride, fluoride, sulphate and conductivity between Okhla Barrage and Kalpi (48 %, 57 %, 5 % and 39 % respectively) were much less than the reductions in TP (85 %), SRP (90 %), TDN (96 %) and DOC (68 %). This strongly suggests that the major reduction in nutrient concentrations along this reach of the Yamuna River was primarily due to biological uptake and processing of P, N and C, and geochemical processes such as sediment deposition and sorption of SRP to river bed sediments.

3.4   Inorganic water quality at the additional monitoring sites

The Pandu and Rind Rivers (Sites 18 and 19, Figure 1) were moderately enriched with TP (166 and 130 μg P l^{-1} respectively), but SRP concentrations were at potentially limiting concentrations (40 and 41 μg P l^{-1}) (O'Hare et al., 2018). TDN concentrations were also low (<1 mg N l^{-1}), although there were relatively high concentrations of ammonium of 0.14 mg N l^{-1} at both sites (approximately 20 % of the total nitrogen load), which strongly suggests significant inputs of sewage, fertilisers or animal wastes. Similar conclusions about P and N loadings from fertilisers and urban wastes have been reached in the study of the Pandu River by Sen et al. (2018).

The Upper Ganga Canal at Roorkee (Site 11, Figure 1) had relatively good nutrient water quality, with a TP and SRP concentration of 46 and 2 μg l^{-1} respectively, and TDN concentration of 0.96 mg N l^{-1} (Figure 4). Most of the total phosphorus is probably predominantly particulate-bound P associated with the high sediment load within this fast moving waterbody. Water quality parameters were very similar to the Ganga River at Haridwar, as the canal branches off the Ganga River at this monitoring site (30 km upstream). The Ganga Canal at Kanpur (520 km downstream of Roorkee; Site 17) had higher TP and SRP concentrations (138 and 20 μg P l^{-1} respectively), but SRP concentrations were still at potentially limiting concentrations, and the TDN concentration had halved to 0.43 mg N l^{-1}. Nitrate

and ammonium concentrations were extremely low (0.03 and 0.01 mg N l^{-1}). This indicates that there were no significant nutrient pollution sources entering the canal, and the increase in total P and conductivity (from 160 to 250 μS cm^{-1}) are probably due to small diffuse-source inputs (i.e. possibly from atmospheric dust deposits and agricultural inputs), in conjunction with an increase in solute concentrations due to evaporation along the length of the canal. The nitrogen speciation changed along the course of the canal (Figure 3b), with the proportion of TDN load in dissolved organic form increasing from 53 % at Roorkee to 91 % at Kanpur, indicating biological processing of nutrients was occurring.

3.5 Phytoplankton concentrations and community composition

3.5.1 Ganga transect

Phytoplankton densities were relatively low in the upper Ganga from upstream of Rishikesh to Balawali (Sites 1-4) (<4,000 cells ml^{-1}), increasing to 30,584 cells ml^{-1} at the Ganga at Brajghat monitoring site (Site 6) (Figure 5; Supplementary Table 3). These are much lower than the phytoplankton densities observed in the middle River Thames in southern England, using the same flow cytometry methodology, which consistently had >700,000 cells ml^{-1} through the spring – autumn growing period of 2011 (Read et al., 2014) and between 50,000 and 200,000 cells ml^{-1} in 2015 (Moorhouse et al., 2018).

The community composition was dominated by cyanobacteria (equivalent to between 45 to 70 % of the total phytoplankton cell abundance) in the Ganga reach from upstream of Rishikesh to Haridwar (Sites 1-3). The increase in phytoplankton cell density in the river reach from Balawali to Brajghat (Sites 4-6) was primarily due to an increase in the population of the smaller meso- and pico-chlorophytes. By the time the Ganga reached Farrukhabad (Site 7), the phytoplankton community was again dominated by cyanobacteria, which comprised *ca.* 80 % of the total cell numbers. Cyanobacteria favour high nutrient concentrations and warm temperatures (Paerl and Huisman, 2008), and so this shift in community structure could be related to increasing higher TP and TDN concentrations and water temperatures along the Ganga transect, with a 1.2 $^{\circ}$C increase between Balawali and Farrukhabad (from 26.4 to 27.6 $^{\circ}$C).

There were major inputs of phytoplankton to the Ganga between Site 6 and 7 from two of the incoming tributaries (Figure 5a). The Kali and Ramganga Rivers had total phytoplankton concentrations of 518,000 and 411,000 cells ml^{-1} respectively, confirming visual observations made at the monitoring sites that these rivers were very green and experiencing a major algal bloom at the time of sampling (Supplementary Figure 1b, photo B). The blooms in both rivers were dominated by meso- and pico-chlorophytes with significant numbers of diatoms (28,770 diatom cells per ml,

equivalent to 7 % of total phytoplankton population) in the Ramganga. These major phytoplankton biomass inputs from the Ramganga and Kali Rivers caused a 26% increase in total phytoplankton cell density in the Ganga at Kannauj (Site 8) (to 112,000 cells ml^{-1}), but the amount of increase was mitigated due to the dilution effect of the third tributary joining the Ganga within this short stretch; the Garra River (Site 16), which had a relatively low phytoplankton concentration of 46,000 cells ml^{-1}. However, the major inputs of meso- and nano-chlorophytes from the Kali and Ramganga rivers caused a major shift in phytoplankton community structure within the Ganga, reducing the cyanobacterial composition from 80 % of the total phytoplankton cell count at Farrukhabad to 4 % at Kannauj.

The highest phytoplankton biomass in the Ganga occurred in the final 18 km stretch of the study reach, between Kannauj and Nanamau Bridge (Sites 8-9) (Figure 5a). The phytoplankton cell densities doubled to 247,000 cells ml^{-1}, primarily due to rapid growth of diatoms, meso- and nano-chlorophytes. The large phytoplankton biomass in the lower Ramganga River was very likely to be severely nutrient limited, with nitrate concentrations below detection limits (<0.1 mg NO_3 l^{-1}) and SRP concentrations of only 12 μg P l^{-1}. When this biomass entered the Ganga River, excess nutrients immediately became available (with the Ganga at Kannauj having SRP and nitrate-N concentrations of 107 P μg l^{-1} and 8.56 mg NO_3-N l^{-1}), thereby eliminating the P and N limitation and resulting in the rapid growth rates observed in this reach.

3.5.2   Tributaries

The highest phytoplankton concentrations were observed in the Ramganga (reaching a peak of 760,000 cells ml^{-1} at Kunauli) and Kali Rivers (518,000 cells ml^{-1}). These populations predominantly consisted of meso- and pico-chlorophytes. The longer tributaries (Pandu, Ganga Canal at Kanpur, Yamuna at Kalpi) had between 80,000 to 100,000 cells per ml^{-1}, with the majority of the phytoplankton communities of the Pandu and Ganga Canal consisting of cyanobacteria. The remaining tributaries (Song, upper Yamuna, Rind and Hindon) had phytoplankton cell densities <25,000 cells ml^{-1}.

3.6   Bacterial concentrations by flow cytometry

The highest bacterial concentrations were observed in the Yamuna River at Central Delhi (Site 21) and Okhla Bridge (Site 22) (59 million and 66 million bacterial cells ml^{-1} respectively), followed by the Hindon River (Site 20) (29 million cells ml^{-1}), and the Ramganga at Bareilly, Kali and Rind Rivers (Sites 12, 15 and 19), all with concentrations of *ca*. 20 million cells ml^{-1} (Figure 6; Supplementary Table 3). The sampling sites on these rivers are all downstream of major cities. This strongly suggests that the primary source of bacterial loadings across the Ganga basin were influenced by wastewater from the major cities and urbanised areas. A recent study of the Ganga and its tributaries using regulatory faecal coliform most probable number counts (Milledge et al., 2018) also highlighted the close relationship

between microbial pollution and upstream population density. The bacterial concentrations in these tributaries are much higher than those typically found in UK rivers, which range between one million to seven million cells ml^{-1} across the highly-populated Thames basin (Read et al., 2015). This suggests that the bacterial loadings are much higher in the Indian context, either due to a lack of sewage treatment, or due to higher water temperatures and nutrient concentrations being able to maintain a higher bacterial biomass.

There was a seven-fold increase in bacterial cell densities in the Ganga between Brajghat and Farrukhabad (Site 6–7). The major increases in pollution and algal loadings consistently observed within this study in the stretch downstream of Farrukhabad, due to inputs from the Kali and Ramganga rivers, did not have an impact on bacterial numbers, as the concentrations of bacteria in these tributaries was lower than in the Ganga River. Bacterial concentrations from upstream of Rishikesh to Bijnor (Sites 1-5) were relatively low (<3 million cells ml^{-1}), suggesting that the bacterial inputs from the cities of Rishikesh and Haridwar were diluted by the high flows in the upper Ganga. The extremely high bacterial concentrations in the Yamuna within Delhi reduced by >80 % by the time it reached Kalpi (Site 23). This will be partially due to dilution from the Chambal and Sind River tributaries (Figure 1), but is also likely to be due to natural attenuation and the rapid dying off of human gut microbes in raw sewage inputs, once within the riverine environment. The spatial pattern of bacterial concentrations in this study was similar to that of faecal coliform concentrations monitored between 2002 and 2012 (Milledge et al., 2018), which also showed a marked increase in microbial pollution in the Ganga between the Ramganga and Kali confluences and Kanpur, with the highest concentrations of faecal coliforms in the Yamuna River within Delhi.

3.7 Bacterial community composition

At the phylum level, the bacterial communities observed in the Ganga River and its tributaries were similar to those reported from freshwater rivers worldwide, with Proteobacteria, Bacteroidetes, Cyanobacteria and Actinobacteria making up >80% of the community in most samples (Figure 7A). At the ASV level (identified to Genus), commonly occurring taxa were typical of those occurring in freshwaters, including representatives of Polynucleobacter (Burkholderiaceae), Limnohabitans (Burkholderiaceae), Flavobacterium (Flavobacteriaceae) and Rhodobacter (Rhodobacteraceae) (Staley et al., 2013; Read et al., 2015; Simonin et al., 2019). The NMDS ordination plot shows that there were distinct changes in microbial community composition along the river transect (Figure 7B), with samples collected from the oligotrophic upstream sites (Sites 1-3) clustered at the top (e.g. Rishikesh, Haridwar and Treveni Ghat) and the communities from the eutrophic downstream sites and tributaries clustered in the lower half of the ordination (e.g. The Ganga at Nanamau Bridge and Kannauj (Sites 8 and 9), as well as the Ramganga and Garra rivers). An envfit analysis (Supplementary Table 4 and

Supplementary Figure 2) showed that the sites clustered in the lower NMDS ordination were associated with higher levels of total phosphorus, dissolved organic carbon, temperature, chloride, fluoride and conductivity, as well as higher counts of bacterioplankton, phytoplankton and chlorophyte algae. An additional envfit analysis of the phyla level table against the NMDS ordination (Supplementary Figure 3) showed that the phyla Fusobacteria and Nitrospirae were associated with the more oligotrophic upstream sites, whereas the phyla Cyanobacteria, Verrucomicrobia and Chloroflexi were associated with the eutrophic downstream sites. Both Fusobacteria and Nitrospirae have previously been identified in high abundance in groundwater, (Hershey et al., 2018) whereas planktonic Cyanobacteria are generally associated eutrophic rivers with long residence times (Paerl and Paul, 2012), and Verrucomicrobia have previously been identified as increasing downstream in a UK river (Read et al., 2015). The sample collected from the Song River (Site 10) was an outlier in terms of microbial community composition, forming a distinct point on the NMDS ordination. This site was associated with elevated levels of total dissolved nitrogen, sulphate and conductivity (Figure 2), indicating that sewage pollution from the nearby urban areas and historic mining activity (Joshi et al., 1995) may influence water chemistry. Supporting this observation was the fact that the bacterial community from the Song River was moderately enriched in bacteria taxa that are commonly associated with faecal or sewage contamination, including faecally-associated bacterial families, Enterobacteriaceae and Prevotellaceae (Figure 7C and D). The Ramganga River had by far the highest concentration of both pathogen- and faecal-indicating bacteria, with the Kali and Garra rivers also contributing significant inputs that contaminated the downstream sites along the transect, indicating again that it is these highly urbanised tributary inputs that are having major impacts on chemical and microbiological water quality.

The pathogen- and faecal-indicating bacterial abundances determined by 16-s sequencing (Figure 7) produced a similar pattern to the flow cytometry total bacterial numbers (Figure 6) along the Ganga, with low levels in the upper section and the highest concentrations found at Farrukhabad and Kannauj. However, the flow cytometry technique did not detect the particularly high total bacterioplankton concentrations in the Ramganga, and so the technique may be better suited as an initial and rapid screening tool for estimating pathogen loadings in the Ganga basin.

4   Conclusions

This study has shown that the water quality of the Ganga River is relatively good in the upper stretches (Sites 1 to 6), with TP concentrations below 100 µg l^{-1}, SRP at potentially-limiting concentrations, and TDN <1 mg l^{-1}, despite the quantities of untreated sewage, industrial effluents and fertiliser use within the catchment. This indicates that the huge volumetric flow of the Ganga is able to dilute these inputs

sufficiently to prevent them getting to hazardous concentrations, and also that the long length and residence time of the river allows nutrient pollution to be removed by sedimentation and biological uptake.

The decline in water quality of the middle sections of the Ganga River upstream of Kanpur during the pre-monsoon season is mainly due to major pollution inputs from human-impacted tributaries (principally the Ramganga and Kali Rivers). The high nutrient concentrations and P and N speciation in these tributaries and the Yamuna River, plus high chloride concentrations strongly suggest that these excessive nutrient loads are principally derived from urban wastewaters. Similar conclusions have been made in previous studies (Jin et al., 2015; Trivedi, 2010). This is further supported by the bacterioplankton community composition from these tributaries, which had large concentrations of pathogens and faecal-indicating bacterial families. The impact of these polluted tributary inputs is compounded by a reduction in flow within the Ganga in the March to April pre-monsoon period, due to the diversion of a large proportion of its water to support agriculture via the irrigation canals. At the time of sampling, the high nutrient concentrations, low flows, warm water conditions and high solar radiation resulted in algal blooms in the Kali and Ramganga, which greatly impacted the Ganga below Farrukhabad. This observation highlights the importance of conducting simultaneous chemical and biological surveying, and shows the potential for using flow cytometry as a means of rapidly and cheaply monitoring the concentrations of algae, cyanobacteria and bacteria in the future. However, total bacterioplankton concentration by flow cytometry was not a good proxy for pathogens and faecal-indicating bacteria at some study sites, but could still be a useful rapid-screening tool for identifying potentially-contaminated water bodies.

To protect and improve the water quality and ecology of the Ganga upstream of Kanpur; -

- The major domestic and industrial effluents from the cities along the Kali, Ramganga, Garra and Yamuna Rivers need to be targeted, by intercepting and treating these effluents to the standards adopted by more-developed industrial countries, thereby reducing nutrient, pollution and bacterial loadings to the Ganga itself.
- This effluent treatment should be supported by regular, quality-controlled analytical testing of effluents and the receiving rivers. Regulation of wastewater treatment plants and industry by imposing consents would help to ensure that pollution loads were controlled and ultimately reduced.
- Augmentation of the flow in the Ganga during this dry, pre-monsoon period could be potentially achieved by diverting less water through the irrigation canals and moving away from crops with high water demand, such as sugar cane (Sapkota et al., 2013). Releasing water

from the barrages along the Ganga / Ramganga would also help to dilute wastewater inputs and provide ecological flows that could help to support key aquatic species, such as river dolphins, gharials and mahseer (O'Keefe et al., 2012; Shah et al., 2018).

Future chemical and biological surveys at different times of the year may capture greater agricultural diffuse inputs of nutrients from the catchment and particularly the inundated floodplains of the larger rivers, such as the Ganga and Yamuna, which are intensively farmed during the dry period. These surveys will start to build up an understanding of how nutrient and pollution sources to the Ganga change seasonally, and how these impact on the aquatic ecology. Ultimately, there must be a programme of water quality testing at regular intervals across this region by the Indian Regulators, which covers both rivers and sewage / industrial point sources. The samples must be analysed for a wide range of chemical and biological indicators, so that accurate and reliable source apportionment can be carried out and pollution incidents can be fully investigated. These intensive, multidisciplinary data sets are vital for determining the most effective mitigation measures for halting and ultimately reversing the degradation of water quality of the Ganga and its tributaries, which would potentially transform ecological status and human health within this internationally-important catchment.

Acknowledgements

List of Figures

Figure 1. Map of the upper Ganga River catchment, indicating location of sampling sites

Figure 2. Physicochemcial data along Ganga River transect (blue bars), including contributing tributaries to the Ganga River (orange bars), including total phosphorus (μg l^{-1}), soluble reactive phosphorus (μg l^{-1}), total dissolved nitrogen (mg l^{-1}), dissolved organic carbon (mg l^{-1}), chloride (mg l^{-1}), fluoride (mg l^{-1}), sulphate (mg SO$_4$ l^{-1}) and conductivity (μS cm^{-1}).

Figure 3. Nitrogen speciation, including dissolved organic N, Nitrite-N, Ammonium-N and Nitrate-N (a) along the Ganga River transect, including contributing tributaries to the Ganga transect (marked with *) and (b) at all monitored tributary sites.

Figure 4. Water quality data from tributary study sites, including total phosphorus (µg l^{-1}), soluble reactive phosphorus (µg l^{-1}), total dissolved nitrogen (mg l^{-1}), dissolved organic carbon (mg l^{-1}), chloride (mg l^{-1}), fluoride (mg l^{-1}), sulphate (mg SO$_4$ l^{-1}) and conductivity (µS cm^{-1}). (ND = No data due to missing sample).

Figure 5. Phytoplankton abundances and community composition along (a) the Ganga River transect and (b) major tributaries, measured by flow cytometry. The * symbols indicate contributing tributaries to the Ganga River transect.

Figure 6. Bacterial cell densities measured by flow cytometry (a) along the Ganga River transect and (b) in the tributary sites.

Figure 7. Changes in bacterial community composition along the River Ganga transect and contributing tributaries. (A) Changes in phyla. (B) NMDS ordination plot of sites in relation to bacterial community structure. (C) Pathogenic bacteria abundances. (D) Faecally-associated bacteria abundances.

5   References

Agarwal, A., Prajapati, R., Singh, O. P., Raza, S. K., & Thakur, L. K. (2015). Pesticide residue in water-a challenging task in India (Article). *Environmental Monitoring and Assessment*, 187(2), 21. doi:10.1007/s10661-015-4287-y.

Bowes, M. J., Armstrong, L. K., Harman, S. A., Wickham, H. D., Nicholls, D. J. E., Scarlett, P. M., et al. (2018). Weekly water quality monitoring data for the River Thames (UK) and its major tributaries (2009-2013): the Thames Initiative research platform. *Earth System Science Data*, 10(3), 1637-1653. doi:10.5194/essd-10-1637-2018.

Bowes, M. J., & House, W. A. (2001). Phosphorus and dissolved silicon dynamics in the River Swale catchment, UK: a mass-balance approach. *Hydrological Processes*, 15(2), 261-280. <Go to ISI>://000167211600006.

Bowes, M. J., Ings, N. L., McCall, S. J., Warwick, A., Barrett, C., Wickham, H. D., et al. (2012). Nutrient and light limitation of periphyton in the River Thames: implications for catchment management. *Sci Total Environ*, 434, 201-12. doi:10.1016/j.scitotenv.2011.09.082.

Bowes, M. J., Jarvie, H. P., Naden, P. S., Old, G. H., Scarlett, P. M., Roberts, C., et al. (2014). Identifying priorities for nutrient mitigation using river concentration–flow relationships: The Thames basin, UK. *Journal of Hydrology*, 517(0), 1-12. doi:http://dx.doi.org/10.1016/j.jhydrol.2014.03.063.

Bowes, M. J., Loewenthal, M., Read, D. S., Hutchins, M. G., Prudhomme, C., Armstrong, L. K., et al. (2016). Identifying multiple stressor controls on phytoplankton dynamics in the River Thames (UK) using high-frequency water quality data. *Science of The Total Environment*, 569–570, 1489-1499. doi:http://dx.doi.org/10.1016/j.scitotenv.2016.06.239.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data (Article). *Nature Methods*, 13(7), 581-+. doi:10.1038/nmeth.3869.

Chen, X., Yang, L., Xiao, L., Miao, A., & Xi, B. (2012). Nitrogen removal by denitrification during cyanobacterial bloom in Lake Taihu. *Journal of Freshwater Ecology*, 27(2), 243-258. doi:10.1080/02705060.2011.644405.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069-5072. doi:10.1128/aem.03006-05.

Dixit, R. B., Patel, A. K., Toppo, K., & Nayaka, S. (2017). Emergence of toxic cyanobacterial species in the Ganga River, India, due to excessive nutrient loading (Article). *Ecological Indicators*, 72, 420-427. doi:10.1016/j.ecolind.2016.08.038.

Dodds, W. K. (2006). Eutrophication and trophic state in rivers and streams (Article). *Limnology and Oceanography*, 51(1), 671-680. <Go to ISI>://WOS:000241296700027.

Eisenreich, S. J., Bannerman, R. T., & Armstrong, D. E. (1975). A simplified phosphorus analytical technique. *Environmental Letters*, 9(4), 5-53.

Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., et al. (2015). High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. *Hydrological Processes*, 29, 3388-3407. doi:10.1002/hyp.10453.

Hardenbicker, P., Rolinski, S., Weitere, M., & Fischer, H. (2014). Contrasting long-term trends and shifts in phytoplankton dynamics in two large rivers. *International Review of Hydrobiology*, 99(4), 287-299. doi:10.1002/iroh.201301680.

Hershey, O. S., Kallmeyer, J., Wallace, A., Barton, M. D., & Barton, H. A. (2018). High Microbial Diversity Despite Extremely Low Biomass in a Deep Karst Aquifer (Original Research). *Frontiers in Microbiology*, 9(2823). doi:10.3389/fmicb.2018.02823.

Indian Institutes of Technology (2010) 'Ganga River basin environmental management plan' *River Ganga at a glance: Identification of issues and priority actions for restoration*. December 2010.

Jain, C. K., Singhal, D. C., & Sharma, M. K. (2007). Estimating nutrient loadings using chemical mass balance approach (Article). *Environmental Monitoring and Assessment*, 134(1-3), 385-396.

Jani, K., Dhotre, D., Bandal, J., Shouche, Y., Suryavanshi, M., Rale, V., et al. (2018a). World's Largest Mass Bathing Event Influences the Bacterial Communities of Godavari, a Holy River of India. *Microbial Ecology*, 76(3), 706-718. doi:10.1007/s00248-018-1169-1.

Jani, K., Ghattargi, V., Pawar, S., Inamdar, M., Shouche, Y., & Sharma, A. (2018b). Anthropogenic Activities Induce Depletion in Microbial Communities at Urban Sites of the River Ganges. *Current Microbiology*, 75(1), 79-83. doi:10.1007/s00284-017-1352-5.

Jarvie, H. P., Neal, C., & Withers, P. J. A. (2006). Sewage-effluent phosphorus: A greater risk to river eutrophication than agricultural phosphorus? *Science of the Total Environment*, 360(1-3), 246-253. <Go to ISI>://000237870500019

Jarvie, H. P., Sharpley, A. N., Scott, J. T., Haggard, B. E., Bowes, M. J., & Massey, L. B. (2012). Within-River Phosphorus Retention: Accounting for a Missing Piece in the Watershed Phosphorus Puzzle. *Environmental Science & Technology*, 46(24), 13284-13292. doi:10.1021/es303562y.

Jin, L., Whitehead, P. G., Sarkar, S., Sinha, R., Futter, M. N., Butterfield, D., et al. (2015). Assessing the impacts of climate change and socio-economic changes on flow and phosphorus flux in the Ganga river system (Article). *Environmental Science-Processes & Impacts*, 17(6), 1098-1110. doi:10.1039/c5em00092k.

Joshi, H., Shishodia, S. K., Kumar, S. N., Saikia, D. K., Nauriyal, B. P., Mathur, R. P., et al. (1995). Ecosystem studies on upper region of Ganga River, India. *Environmental Monitoring and Assessment*, 35(3), 181-206. doi:10.1007/bf00547631.

Jouanneau, S., Recoules, L., Durand, M. J., Boukabache, A., Picot, V., Primault, Y., et al. (2014). Methods for assessing biochemical oxygen demand (BOD): a review. *Water Res*, 49(0), 62-82. doi:10.1016/j.watres.2013.10.066.

Khan, M. Y. A., Gani, K. M., & Chakrapani, G. J. (2016a). Assessment of surface water quality and its spatial variation. A case study of Ramganga River, Ganga Basin, India. *Arabian Journal of Geosciences*, 9(1). doi:10.1007/s12517-015-2134-7.

Khan, M. Y. A., Khan, B., & Chakrapani, G. J. (2016b). Assessment of spatial variations in water quality of Garra River at Shahjahanpur, Ganga Basin, India (Article). *Arabian Journal of Geosciences*, 9(8), 10. doi:10.1007/s12517-016-2551-2.

Leeks, G. J. L., Neal, C., Jarvie, H. P., Casey, H., & Leach, D. V. (1997). The LOIS river monitoring network: strategy and implementation. *Science of the Total Environment*, 194-195, 101-109. http://www.sciencedirect.com/science/article/B6V78-3SVYTV1-10/2/0e1f34b9e5f7ef62185848a8c8b0b6f7.

Marie, D., Partensky, F., Jacquet, S., & Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I (Article). *Applied and Environmental Microbiology*, 63(1), 186-193. <Go to ISI>://WOS:A1997WA16800029.

Mariya, A., Kumar, C., Masood, M., & Kumar, N. (2019). The pristine nature of river Ganges: its qualitative deterioration and suggestive restoration strategies (journal article). *Environmental Monitoring and Assessment*, 191(9), 542. doi:10.1007/s10661-019-7625-7.

Matta, G., Srivastava, S., Pandey, R. R., & Saini, K. K. (2017). Assessment of physicochemical characteristics of Ganga Canal water quality in Uttarakhand. *Environment Development and Sustainability*, 19(2), 419-431. doi:10.1007/s10668-015-9735-x.

Matta, N., & Bisht, G. R. S. (2018). Detection and Enumeration of Coliforms in Ganga Water Collected from Different Ghats. *Journal of Bioprocessing & Biotechniques*, 08(02), 10. doi:10.4172/2155-9821.1000320.

McCall, S. J., Hale, M. S., Smith, J. T., Read, D. S., & Bowes, M. J. (2017). Impacts of phosphorus concentration and light intensity on river periphyton biomass and community structure (journal article). *Hydrobiologia*, 1-16. doi:10.1007/s10750-016-3067-1.

Milledge, D. G., Gurjar, S. K., Bunce, J. T., Tare, V., Sinha, R., & Carbonneau, P. E. (2018). Population density controls on microbial pollution across the Ganga catchment. *Water Research*, 128(Supplement C), 82-91. doi:https://doi.org/10.1016/j.watres.2017.10.033.

Misra, A. K. (2011). Impact of Urbanization on the Hydrology of Ganga Basin (India) (journal article). *Water Resources Management*, 25(2), 705-719. doi:10.1007/s11269-010-9722-9.

Misra, A. K., & Mishra, A. (2007). Study of quaternary aquifers in Ganga Plain, India: Focus on groundwater salinity, fluoride and fluorosis. *Journal of Hazardous Materials*, 144(1), 438-448. doi:https://doi.org/10.1016/j.jhazmat.2006.10.057.

Moorhouse, H. L., Read, D. S., McGowan, S., Wagner, M., Roberts, C., Armstrong, L. K., et al. (2018). Characterisation of a major phytoplankton bloom in the River Thames (UK) using flow cytometry and high performance liquid chromatography. *Science of The Total Environment*, 624, 366-376. doi:https://doi.org/10.1016/j.scitotenv.2017.12.128.

Murphy, J., & Riley, J. P. (1962). A modified single solution method for the determination of phosphorus in natural waters. *Analytica chemica acta*, 12, 31-36.

National River Conservation Directorate (2009) 'Status paper on River Ganga: State of environment and water quality'. August 2009. Ministry of Environment and Forests. , pp. 1-31.

National_River_Conservation_Directorate (2009) 'Status paper on River Ganga: State of environment and water quality'. August 2009. Ministry of Environment and Forests. , pp. 1-31.

Neal, C., Jarvie, H. P., Neal, M., Love, A. J., Hill, L., & Wickham, H. (2005). Water quality of treated sewage effluent in a rural area of the upper Thames Basin, southern England, and the impacts of such effluents on riverine phosphorus concentrations. *Journal of Hydrology*, 304(1-4), 103-117. <Go to ISI>://000227881700008

Neal, C., Neal, M., & Wickham, H. (2000). Phosphate measurement in natural waters: two examples of analytical problems associated with silica interference using phosphomolybdic acid

methodologies (Article). *Science of the Total Environment*, 251, 511-522. <Go to ISI>://000087253500034

https://www.sciencedirect.com/science/article/pii/S0048969700004022?via%3Dihub.

O'Keefe, J., Kaushal, N., Bharati, L., & Smakhtin, V. U. (2012) 'Assessment of environmental flows for the Upper Ganga Basin' *Project report of the environmental flows assessment done under the Living Ganga Program*. New Delhi, India.: World Wide Fund for Nature - India), p. 161p.

O'Hare, M. T., Baattrup-Pedersen, A., Baumgarte, I., Freeman, A., Gunn, I. D. M., Lázár, A. N., et al. (2018). Responses of Aquatic Plants to Eutrophication in Rivers: A Revised Conceptual Model (Review). *Frontiers in Plant Science*, 9(451). doi:10.3389/fpls.2018.00451.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlin, D., et al. (2017). vegan: Community Ecology Package. R package version 2.4-3. 2017 [accessed 2016 Jan 1].

Paerl, H. W., & Huisman, J. (2008). Blooms Like It Hot. *Science*, 320(5872), 57-58. doi:10.1126/science.1155398.

Paerl, H. W., & Paul, V. J. (2012). Climate change: Links to global expansion of harmful cyanobacteria. *Water Research*, 46(5), 1349-1363. doi:https://doi.org/10.1016/j.watres.2011.08.002.

Pandey, J., Tripathi, S., & Pandey, U. (2016). Anthropogenically Induced Shifts in N:P:Si Stoichiometry and Implications in Ganga River (Article). *Air Soil and Water Research*, 9, 9. doi:10.4137/aswr.s32780.

Pandey, J., & Yadav, A. (2017). Alternative alert system for Ganga river eutrophication using alkaline phosphatase as a level determinant (Article). *Ecological Indicators*, 82, 327-343. doi:10.1016/j.ecolind.2017.06.061.

Pathak, D., Whitehead, P. G., Futter, M. N., & Sinha, R. (2018). Water quality assessment and catchment-scale nutrient flux modeling in the Ramganga River Basin in north India: An application of INCA model. *Science of The Total Environment*, 631-632, 201-215. doi:https://doi.org/10.1016/j.scitotenv.2018.03.022.

R Core Team (2013) 'R: a Language and Environment for Statistical Computing' R. F. f. S. Computing. Vienna, Austria. Available at: http://www.R-project.org/.

Read, D. S., Bowes, M. J., Newbold, L. K., & Whiteley, A. S. (2014). Weekly flow cytometric analysis of riverine phytoplankton to determine seasonal bloom dynamics. *Environ Sci Process Impacts*, 16(3), 594-603. doi:10.1039/c3em00657c.

Read, D. S., Gweon, H. S., Bowes, M. J., Newbold, L. K., Field, D., Bailey, M. J., et al. (2015). Catchment-scale biogeography of riverine bacterioplankton (Original Article). *ISME J*, 9(2), 516-26. doi:10.1038/ismej.2014.166.

Sapkota, P., Bharati, L., Gurung, P., Kaushal, N., & Smakhtin, V. (2013). Environmentally sustainable management of water demands under changing climate conditions in the Upper Ganges Basin, India. *Hydrological Processes*, 27(15), 2197-2208. doi:10.1002/hyp.9852.

Sen, I. S., Boral, S., Ranjan, S., & Tandon, S. K. (2018). Small But Important: The Role of Small Floodplain Tributaries to River Nutrient Budgets. *ACS Earth and Space Chemistry*, 2(1), 64-71. doi:10.1021/acsearthspacechem.7b00112.

Shah, T., Ray, C., & Lele, U. (2018). How to clean up the Ganges? *Science*, 362(6414), 503-503. doi:10.1126/science.aav8261.

Simonin, M., Voss, K. A., Hassett, B. A., Rocca, J. D., Wang, S.-Y., Bier, R. L., et al. (2019). In search of microbial indicator taxa: shifts in stream bacterial communities along an urbanization gradient. *Environmental Microbiology*, 21(10), 3653-3668. doi:10.1111/1462-2920.14694.

Sinha, R., Mohanta, H., Jain, V., & Tandon, S. K. (2017). Geomorphic diversity as a river management tool and its application to the Ganga River, India (Article). *River Research and Applications*, 33(7), 1156-1176. doi:10.1002/rra.3154.

Skeffington, R. A., Halliday, S. J., Wade, A. J., Bowes, M. J., & Loewenthal, M. (2015). Using high-frequency water quality data to assess sampling strategies for the EU Water Framework Directive (Article). *Hydrology and Earth System Sciences*, 19(5), 2491-2504. doi:10.5194/hess-19-2491-2015.

Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., et al. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *Journal of Applied Microbiology*, 115(5), 1147-1158. doi:10.1111/jam.12323.

Tappin, A. D., Comber, S., & Worsfold, P. J. (2016). Orthophosphate-P in the nutrient impacted River Taw and its catchment (SW England) between 1990 and 2013 (Article). *Environmental Science-Processes & Impacts*, 18(6), 690-705. doi:10.1039/c6em00213g.

Tare, V., Yadav, A. V. S., & Bose, P. (2003). Analysis of photosynthetic activity in the most polluted stretch of river Ganga (Article). *Water Research*, 37(1), 67-77. doi:10.1016/s0043-1354(01)00385-2.

Trisal, C., Tabassum, T., & Kumar, R. (2008). Water quality of the river Yamuna in the Delhi stretch: Key determinants and management issues. *Clean-Soil Air Water*, 36(3), 306-314. doi:10.1002/clen.200700044.

Trivedi, R. C. (2010). Water quality of the Ganga River – An overview. *Aquatic Ecosystem Health & Management*, 13(4), 347-351. doi:10.1080/14634988.2010.528740.

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., et al. (2016). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems*, 1(1), e00009-15. doi:10.1128/mSystems.00009-15.

Wetz, M. S., Cira, E. K., Sterba-Boatwright, B., Montagna, P. A., Palmer, T. A., & Hayes, K. C. (2017). Exceptionally high organic nitrogen concentrations in a semi-arid South Texas estuary susceptible to brown tide blooms (Article). *Estuarine Coastal and Shelf Science*, 188, 27-37. doi:10.1016/j.ecss.2017.02.001.

Yadav, A., & Pandey, J. (2017a). Contribution of point sources and non-point sources to nutrient and carbon loads and their influence on the trophic status of the Ganga River at Varanasi, India (Article). *Environmental Monitoring and Assessment*, 189(9), 19. doi:10.1007/s10661-017-6188-8.

Yadav, A., & Pandey, J. (2017b). Water Quality Interaction with Alkaline Phosphatase in the Ganga River: Implications for River Health. *Bulletin of Environmental Contamination and Toxicology*, 99(1), 75-82. doi:10.1007/s00128-017-2108-4.

Zhang, S. Y., Tsementzi, D., Hatt, J. K., Bivins, A., Khelurkar, N., Brown, J., et al. (2018). Intensive allochthonous inputs along the Ganges River and their effect on microbial community composition and dynamics. *Environmental Microbiology*, 0(ja). doi:10.1111/1462-2920.14439.

Zubkov, M. V., Sleigh, M. A., Tarran, G. A., Burkill, P. H., & Leakey, R. J. G. (1998). Picoplanktonic community structure on an Atlantic transect from 50 degrees N to 50 degrees S (Article). *Deep-Sea Research Part I-Oceanographic Research Papers*, 45(8), 1339-1355. doi:10.1016/s0967-0637(98)00015-6.
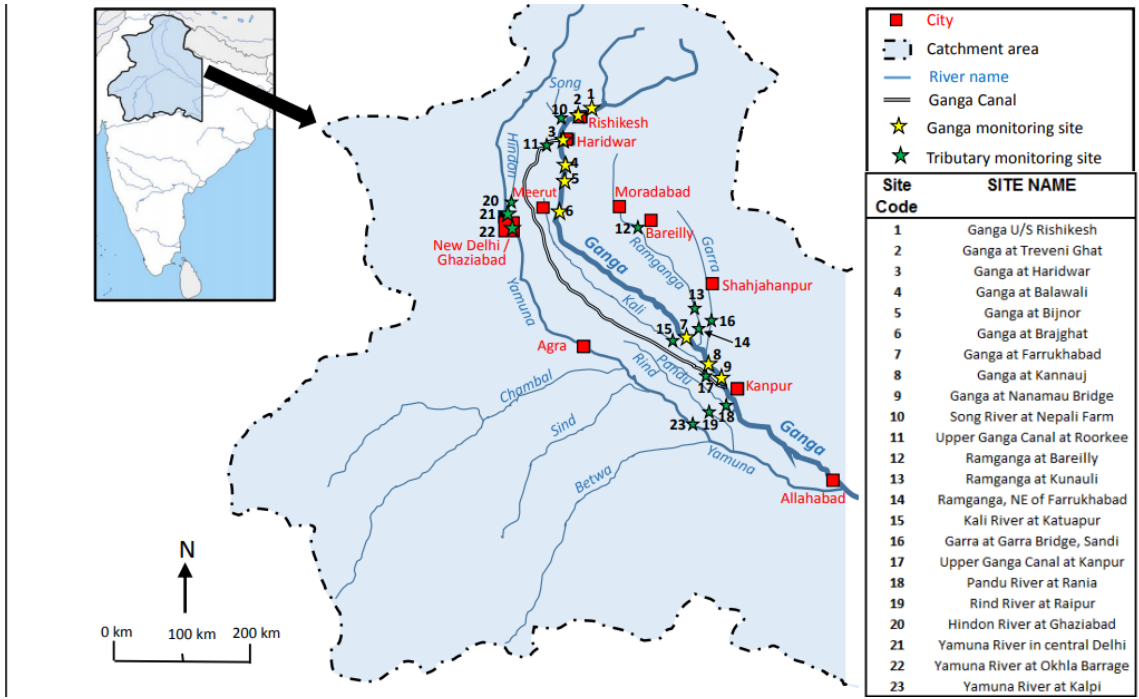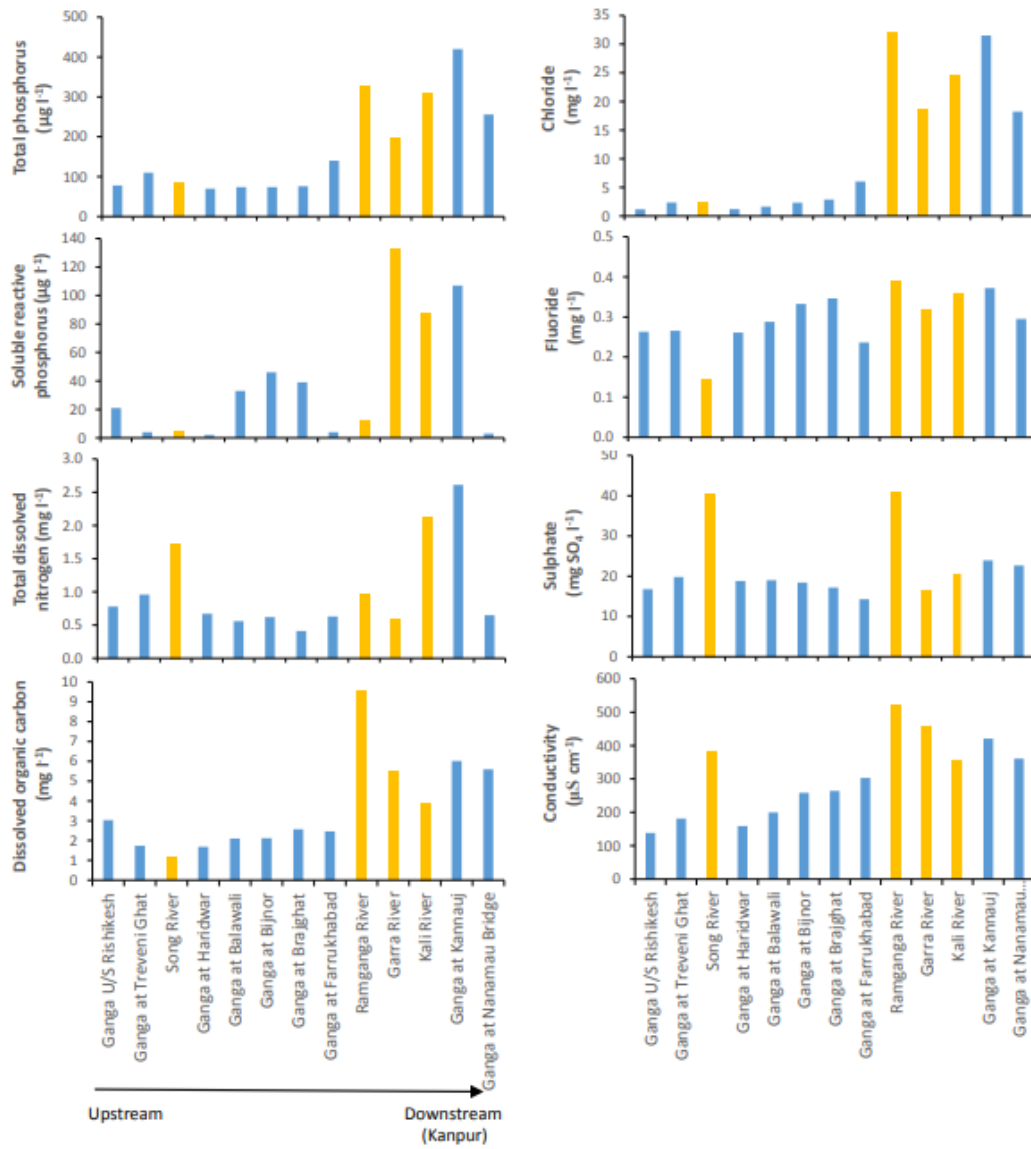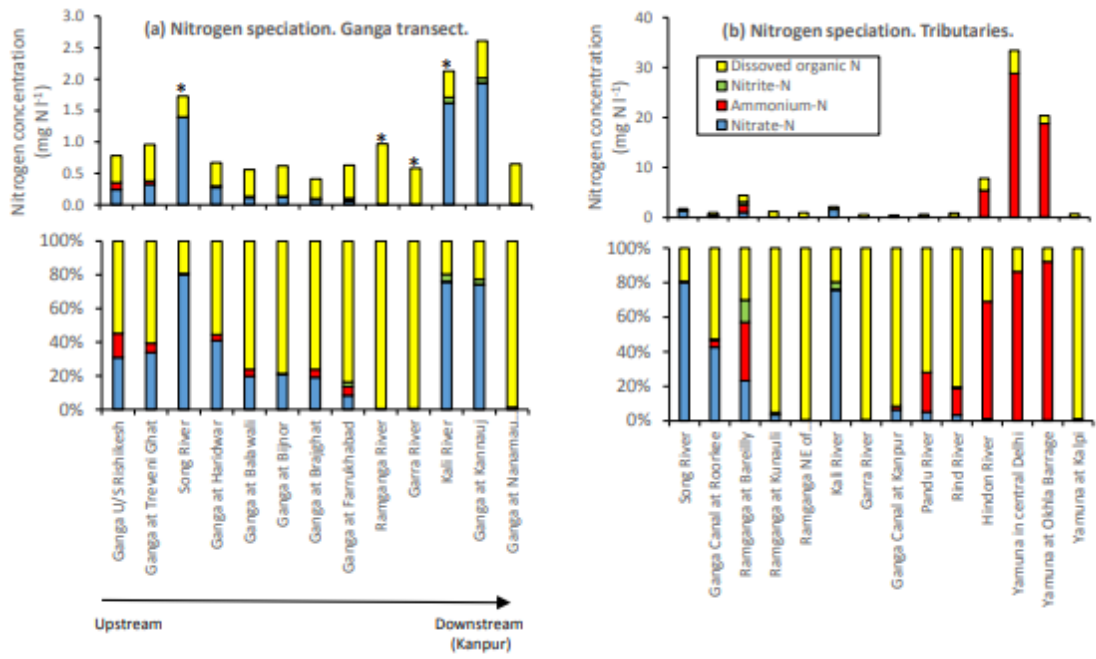
Figure 1

Figure 2

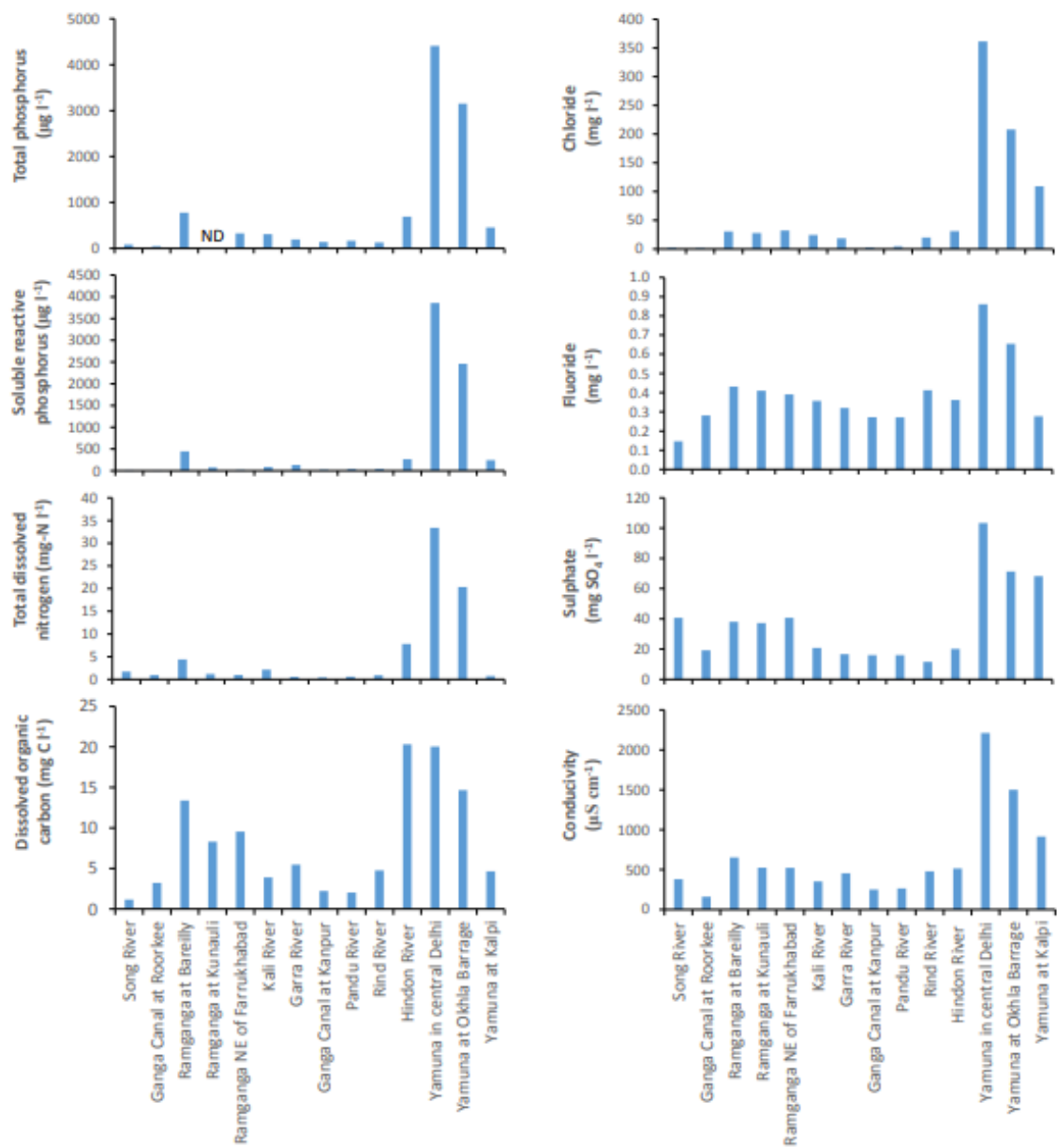(a) Nitrogen speciation. Ganga transect.

(b) Nitrogen speciation. Tributaries.

Figure 3

Figure 4

Figure 5

(a) Ganga River transect

Bacterial cells ml⁻¹

Ganga U/S Rishikesh, Ganga at Treveni Ghat, Ganga at Haridwar, Ganga at Balawali, Ganga at Bijnor, Ganga at Brajghat, Ganga at Farrukhabad, Ganga at Kannauj, Ganga at Nanamau Bridge

(b) Tributaries

Bacterial cells ml⁻¹

Song River, Ganga Canal at Roorkee, Ramganga at Bareilly, Ramganga at Kunauli, Ramganga near Farrukhabad, Kali River, Garra River, Ganga Canal at Kanpur, Pandu River, Rind River, Hindon River, Yamuna in central Delhi, Yamuna at Okhla Barrage, Yamuna at Kalpi

Figure 6

Figure 7