UNIVERSITY^{OF} BIRMINGHAM University of Birmingham Research at Birmingham

A systematic approach to understand hydrogeochemical dynamics in large river systems

Richards, Laura A.; Fox, Bethany G.; Bowes, Michael J.; Khamis, Kieran; Kumar, Arun; Kumari, Rupa; Kumar, Sumant; Hazra, Moushumi; Howard, Ben; Thorn, Robin M.S.; Read, Daniel S.; Nel, Holly A.; Schneidewind, Uwe; Armstrong, Linda K.; Nicholls, David J.E.; Magnone, Daniel; Ghosh, Ashok; Chakravorty, Biswajit; Joshi, Himanshu; Dutta, Tapan K. *DOI:*

10.1016/j.watres.2022.118054

License: Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version Peer reviewed version

Citation for published version (Harvard):

Richards, LÁ, Fox, BG, Bowes, MJ, Khamis, K, Kumar, A, Kumari, R, Kumar, S, Hazra, M, Howard, B, Thorn, RMS, Read, DS, Nel, HA, Schneidewind, U, Armstrong, LK, Nicholls, DJE, Magnone, D, Ghosh, A, Chakravorty, B, Joshi, H, Dutta, TK, Hannah, DM, Reynolds, DM, Krause, S, Gooddy, DC & Polya, DA 2022, 'A systematic approach to understand hydrogeochemical dynamics in large river systems: development and application to the River Ganges (Ganga) in India', *Water Research*, vol. 211, 118054, pp. 118054. https://doi.org/10.1016/j.watres.2022.118054

Link to publication on Research at Birmingham portal

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.

•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.

•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?) •Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

1	A Systematic Approach to Understand Hydrogeochemical Dynamics in Large River Systems:
2	Development and Application to the River Ganges (Ganga) in India
3	
4	Laura A. Richards ^{1*} , Bethany G. Fox ² , Michael J. Bowes ³ , Kieran Khamis ⁴ , Arun Kumar ⁵ , Rupa Kumari ⁵ ,
5	Sumant Kumar ⁶ , Moushumi Hazra ⁷ , Ben Howard ⁴ , Robin M. S. Thorn ² , Daniel S. Read ³ , Holly A. Nel ⁴ ,
6	Uwe Schneidewind ⁴ , Linda K. Armstrong ³ , David J. E. Nicholls ³ , Daniel Magnone ⁸ , Ashok Ghosh ⁵ ,
7	Biswajit Chakravorty ⁹ , Himanshu Joshi ⁷ , Tapan K. Dutta ¹⁰ , David M. Hannah ⁴ , Darren M. Reynolds ² ,
8	Stefan Krause ⁴ , Daren C. Gooddy ¹¹ , David A. Polya ¹
9	
4.0	
10	¹ Department of Earth and Environmental Sciences and Williamson Research Centre for Molecular
11	Environmental Science, The University of Manchester, Williamson Building, Oxford Road,
12	Manchester, M13 9PL, UK
13	² Department of Applied Sciences, University of the West of England, Bristol, BS16 1QY, UK
14	³ UK Centre for Ecology & Hydrology, MacLean Building, Wallingford, Oxfordshire, OX10 8BB, UK
15	⁴ School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston,
16	Birmingham, B15 2TT, UK
17	⁵ Mahavir Cancer Sansthan and Research Centre, Phulwarisharif, Patna, 801505, Bihar, India
18	⁶ Groundwater Hydrology Division, National Institute of Hydrology Roorkee, Roorkee, 247667,
19	Uttarakhand, India
20	⁷ Department of Hydrology, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India
21	⁸ School of Geography, University of Lincoln, Lincoln, LN6 7TS, UK
22	⁹ National Institute of Hydrology, Phulwarisharif, Patna, 801505, Bihar, India

23	¹⁰ Bose Institute, Centenary Campus, P-1/12 C.I.T Scheme VII-M, Kolkata 700054, India
24	¹¹ British Geological Survey, Maclean Building, Wallingford, Oxfordshire OX10 8BB, UK
25	
26	*Corresponding author: laura.richards@manchester.ac.uk
27	
28	Submitted to: Water Research
29	
30	

31 Abstract

32 Large river systems, such as the River Ganges (Ganga), provide crucial water resources for the environment and society, yet often face significant challenges associated with cumulative impacts 33 34 arising from upstream environmental and anthropogenic influences. Understanding the complex 35 dynamics of such systems remains a major challenge, especially given accelerating environmental 36 stressors including climate change and urbanization, and due to limitations in data and process 37 understanding across scales. An integrated approach is required which robustly enables the 38 hydrogeochemical dynamics and underpinning processes impacting water quality in large river 39 systems to be explored. Here we develop a systematic approach for improving the understanding of 40 hydrogeochemical dynamics and processes in large river systems, and apply this to a longitudinal 41 survey (> 2500 km) of the River Ganges (Ganga) and key tributaries in the Indo-Gangetic basin. This 42 framework enables us to succinctly interpret downstream water quality trends in response to the 43 underpinning processes controlling major element hydrogeochemistry across the basin, based on 44 conceptual water source signatures and dynamics. Informed by a 2019 post-monsoonal survey of 81 45 river bank-side sampling locations, the spatial distribution of a suite of selected physico-chemical and 46 inorganic parameters, combined with segmented linear regression, reveals minor and major 47 downstream hydrogeochemical transitions. We use this information to identify five major 48 hydrogeochemical zones, characterized, in part, by the inputs of key tributaries, urban and agricultural

49 areas, and estuarine inputs near the Bay of Bengal. Dominant trends are further explored by 50 investigating geochemical relationships (e.q. Na:Cl, Ca:Na, Mg:Na, Sr:Ca and NO₃:Cl), and how water 51 source signatures and dynamics are modified by key processes, to assess the relative importance of 52 controls such as dilution, evaporation, water-rock interactions (including carbonate and silicate 53 weathering) and anthropogenic inputs. Mixing/dilution between sources and water-rock interactions 54 explain most regional trends in major ion chemistry, although localized controls plausibly linked to 55 anthropogenic activities are also evident in some locations. Temporal and spatial representativeness 56 of river bank-side sampling are considered by supplementary sampling across the river at selected 57 locations and via comparison to historical records. Limitations of such large-scale longitudinal 58 sampling programs are discussed, as well as approaches to address some of these inherent challenges. 59 This approach brings new, systematic insight into the basin-wide controls on the dominant 60 geochemistry of the River Ganga, and provides a framework for characterising dominant 61 hydrogeochemical zones, processes and controls, with utility to be transferable to other large river 62 systems.

63

Keywords: water quality, hydrogeochemical processes, Ganga River Basin, River Ganges, water-rock
 interaction, sampling design

66

67

68 **1. Introduction**

69 Understanding the dynamics of complex large river systems is a major environmental management 70 challenge, with important implications for global water, food and energy security. Water quality in 71 rivers is impacted by underpinning 'natural' hydrogeological and biogeochemical processes, as well as 72 human-environment interactions that are accelerating stress on water resources at unprecedented

73 rates (Best, 2019). Whilst remote sensing of major rivers offers increasing potential for large-scale 74 monitoring (Junqueira et al., 2021; Piégay et al., 2020) ground-truthing field measurements remains 75 essential to our understanding of hydrogeochemical trends and processes at appropriate scales, 76 especially given system complexity and spatial and temporal variability (Poudel et al., 2013; Varol et 77 al., 2012; Wilbers et al., 2014). Developing systematic approaches for field sampling of large river 78 systems is a critical research need; however, this remains challenging due to complexities around 79 coordination, logistics, resource constraints and monitoring fragmentation as well as conceptual 80 challenges in defining adequate sample/sampling strategies to ensure representativeness of 81 observations. Whilst a number of frameworks for water quality monitoring and management exist 82 (viz. Australian Government, 2018; Belitz et al., 2003; European Environment Agency, 2018 and 83 others), understanding longitudinal water quality patterns, particularly in large rivers, remains very 84 challenging due to multiple inputs and interactions, and associated inherent system heterogeneities.

85 The River Ganges (known locally as the Ganga) represents one of the world's largest and most 86 important river systems, spanning > 2500 km in length through one of the most densely populated 87 areas of the world. As a major source of livelihood, the River Ganga is culturally very important and a 88 central focus for many social and religious traditions in the Indian-subcontinent (Kumar, 2017; 89 Lokgariwar et al., 2014). However, regardless of its importance, the River Ganga is facing increasing 90 environmental challenges associated with rapid development, climate change and increasing 91 urbanisation, population, water demand and agricultural intensity (Jain and Singh, 2020; Moors et al., 92 2011; Pandey et al., 2016; Trivedi, 2010; Whitehead et al., 2015). Indian Government initiatives, 93 including the Ganga Action Plan and the National Mission for Clean Ganga (Namami Gange) 94 Programme, have been established in an attempt to monitor, control and/or mitigate pollution in the 95 River Ganga (Ministry of Jal Shakti, 2021; Narain, 2014).

Surface water pollution in the River Ganga and tributaries has been widely reported (Bowes et al.,
2020; Central Pollution Control Board, 2019; Hamner et al., 2006; Khan et al., 2020; Lata et al., 2009;

98 Mariya et al., 2019; Paul, 2017; Satya and Narayan, 2018; Seth et al., 2013; Sharma et al., 2019; Sharma 99 et al., 2016; Sinha and Loganathan, 2015; Trivedi, 2010). However, most previous studies report only 100 a limited number of chemical parameters for specific smaller sub-sections of the Ganga, which 101 prevents detailed interpretation of underpinning hydrogeochemical drivers and controls of pollution 102 sources, transport and transformations. Government initiatives such as the National Water Quality 103 Monitoring Programme (NWMP) provide extensive historical records from 2002 (Central Pollution 104 Control Board, 2019). However, most of these records contain only summary statistics (e.g. annual 105 minimum and maximum) for temperature, dissolved oxygen (DO), pH, electrical conductivity (EC), 106 biochemical oxygen demand, nitrate, faecal coliform and total coliform (Central Pollution Control 107 Board, 2019), parameters which are largely aligned with regulatory requirements, but do not 108 necessarily allow for comprehensive evaluation of water quality or the underpinning 109 hydrogeochemical processes. More detailed studies of particular stretches have evaluated, for 110 example, nutrient and microbial water quality (Bowes et al., 2020), heavy metal pollution (Paul, 2017) 111 or emerging organic contaminants (Sharma et al., 2019). Systematic evaluation of overall 112 hydrogeochemical patterns and underpinning processes throughout the dynamic river system at the 113 basin scale remains a major research gap.

114 An integrated, comprehensive and basin-wide approach is required to understand the highly complex 115 nature of the River Ganga, arising both from natural environmental conditions and human-116 environment interactions. Here, our overall aim is to develop and demonstrate a systematic approach 117 for advancing the understanding of hydrogeochemical dynamics in large river systems such as the 118 River Ganga. This conceptual framework is a demonstration of a coordinated, comprehensive 119 approach to large-scale sampling, analysis and data interpretation which could be applied to a wide 120 variety of parameters and types of river systems for improved monitoring and/or process-based 121 understanding and to provide relevant information for water quality management. Our approach 122 brings new insight by: (i) improving the conceptual understanding of dominant longitudinal water 123 quality patterns and the underpinning hydrogeochemical processes in the Indo-Gangetic basin and

124 elsewhere, and (ii) addressing, in part, the inherent limitations of a large-scale longitudinal survey 125 spanning 1000s of kilometres. Our specific objectives are to (i) develop and test a conceptual 126 framework for the design and interpretation of large-scale river basin studies; (ii) determine the patterns and dynamics of a selected suite of dominant hydrogeochemical parameters across the 127 128 Ganga Basin; (iii) interpret the dominant hydrogeochemical processes across the basin; (iv) identify key hydrogeochemical zones across the basin and (v) evaluate the temporal and spatial 129 130 representativeness of longitudinal sampling. This systematic approach is adaptable, and could be 131 applied more widely, to develop process understanding of other large river systems across the world.

132

133 2. Methods

134 2.1 Conceptual Framework and Overall Approach

Our overall conceptual approach (**Figure 1**) comprises a number of stages, from planning to implementation, analysis and data interpretation to improve the understanding of hydrogeochemical dynamics in the River Ganga, enabling scope for integration across parameter-types, scales and datasets. Here we describe how the key stages of this framework have been developed and applied to a large-scale survey of the River Ganga although the approach could be adapted to other river systems.



Figure 1. Conceptual summary of approach to sampling, analysis and interpretation for understanding hydrogeochemical dynamics and processes in large river systems.

142

143

144

145 2.2 Study Area

146 The study area spans the Indo-Gangetic Basin in India from Devprayag (Uttarakhand) in the foothills 147 of the Himalayas to Noorpur (West Bengal) in the estuarine reaches south of Kolkata, covering a 148 distance of > 2500 km (Figure S1, sites also shown on Figure 5). Substantial environmental and 149 socioeconomic transitions occur along the course of the River Ganga (Bickle et al., 2003; Bickle et al., 150 2005; Dalai et al., 2003; Narain, 2014; Sharma et al., 2019; Tripathy and Singh, 2010); see 151 **Supplementary Information**. Sampling sites (n = 81) were spread longitudinally along the main Ganga 152 and Hooghly rivers (n = 64; noting the lower Ganga is known as the Hooghly downstream of approximately site G58) and key tributaries (n = 17). Sampling sites were selected to capture key 153 154 potential influences (e.q. tributary inputs and a range of geological/agricultural/urban conditions) and 155 to maintain regular sampling intervals (mean interval ~ 30 km) whilst balancing logistical constraints 156 such as site access and driving distance. Sampling sites names are coded as XYY where X is G (Ganga or Hooghly) or T (tributary) and YY is a sequential number increasing with downstream distance. 157 158 Approximate annual hydrological yields of the Ganga and key tributaries were obtained from the literature (Mariya et al., 2019); relative tributary contributions were estimated on the basis of total 159 160 Ganga yield.

161

162 2.3 Water Sampling Approach and Conditions

Longitudinal water sampling, the dominant focus of this study, was undertaken by three teams (coordinated to align sampling conditions) working along the "upper" (Devprayaj to Varanasi), "middle" (Varanasi to Begusarai) and "lower" (Begusarai to Noorpur) segments. Transition areas were 166 sampled by overlapping teams to compare data and ensure consistency. Surface water samples (n = 167 81) on the longitudinal survey were collected from accessible river bank-side locations (typically near 168 ghats or shallow sloping banks), using a bucket (~ 20 L). All sampling occurred during a period of 169 relatively stable conditions in the post-monsoon season in November 2019 (further details in 170 Supplementary Information and Figure S2). The post-monsoon sampling period is expected to be a 171 time of relatively high groundwater-surface water connectivity. Additional sampling was undertaken to assess the temporal and spatial representativeness of the main longitudinal survey (see 172 173 Supplementary Information). A brief overview of our approach was previously presented (Richards 174 and Team Saptanadi, 2020).

Water sub-samples for subsequent laboratory analysis of major and trace elements were filtered (0.45 μm, sterile ThermoFisher cellulose nitrate membrane filters) within ~ 5 minutes of sample collection and stored in acid-washed (20 % hydrochloric acid) Nalgene PTFE bottles (for primary analysis) and/or acid-washed (10 % nitric acid) and furnaced amber glass bottles (for limited secondary analysis). All re-useable sampling equipment was thoroughly rinsed between samples with sample water. Samples were stored in dark and cool conditions (as practicable under field conditions) prior to transport to the UK for further analysis.

182

183 2.4 Chemical Analysis

184 2.4.1 In-Situ Analysis

- 185 Measurements of pH, oxidation-reduction potential (ORP), temperature (T) and electrical conductivity 186 (EC) were collected *in-situ* using a handheld meter (Myron L Ultrameter II, USA), and dissolved oxygen 187 (DO) was measured using an optical DO meter (Hach HQ10, USA).
- 188 **2.4.2.** Inorganic Laboratory Analysis

189 Sub-samples for the analysis of Ca, Na, Mg, Sr and Si were acidified (1 % w:v using analytical grade 190 HNO₃; J.T.Baker ULTREX II Reagent HNO₃) to pH < 2 upon return to laboratories in the UK ($\sim 2 - 14$ days after collection), and stored dark and chilled (~ 4 °C) until analysis. Inductively coupled plasma 191 192 optical emission spectroscopy (ICP-OES; Agilent 5110 with software ICP Expert version 7.4.2 10790) 193 was used for analysis of Ca, Na, Mg, Sr and Si at University of the West of England (Bristol, UK). Analysis 194 of major anions (Cl, NO₃ and SO₄) was undertaken using ion chromatography (IC; Dionex AS50, Thermo 195 Fisher Scientific) at the UK Centre for Ecology & Hydrology (CEH; Wallingford, UK). Inferred alkalinity 196 as HCO₃⁻ was estimated on the basis of charge balance. Independent secondary analysis was 197 undertaken for a subset of 24 samples using ICP-OES (Agilent 5800 for Ca, Na, Mg, Si, Sr) and IC (Dionex 198 ICS5000 for CI, NO₃ and SO₄) at the University of Manchester. See Supplementary Information and 199 Figure S3 for details on analytical quality assurance/quality control (QA/QC).

200

201 2.5 Data Interpretation: Software Packages and Analysis

202 OriginPro 2017 was used for linear regression, principal component analysis and data visualization. 203 Correlation statistics based on ordinary least squares linear regression are reported as " $t_{DF} = t$ value; 204 p = p value", with a significance level of $\alpha = 0.05$ and where DF = degrees of freedom. Principal 205 component analysis was undertaken on scaled and centred data. Segmented regression analysis was 206 undertaken using R (version 4.0.5 with RStudio 1.4.1106) and the segmented package to test for 207 threshold changes in the relationship between hydrogeochemical variables and distance downstream, 208 conceptually analogous to the serial discontinuity approach (Ward and Stanford, 1995). A Davies test 209 (Davies, 1987) was used to assess whether a significant breakpoint existed and segmented regression 210 models of increasing complexity were then fitted (Muggeo, 2008) and ranked using the Akaike 211 Information Criterion (AIC). Estimate(s) of the breakpoint(s) with a 95 % confidence interval and all 212 associated slopes were then extracted for the top-ranking model; see Supplementary Information for 213 further details on segmented regression. QGIS (version 3.12.2 București) was used to create maps

with layer details provided in the associated captions. Elevation was estimated using Google Earth
Pro (version 7.3.4.8248).

216

217 3. Results and Discussion

Following the structure of the interpretative stage of our framework, downstream patterns will be disentangled by firstly considering overall downstream trends and segmented regression to identify characteristic hydrogeochemical zones. Dominant hydrogeochemical processes will be identified using principal component analysis and bivariate relationships between key inorganic parameters. Finally, sampling representativeness and recommendations for future directions are discussed.

223 3.1 Overall downstream patterns in hydrogeochemical parameters

224 The downstream trends of a number of parameters indicate systematic shifts across the basin (Figure 225 2). A significant increase with downstream distance in the main Ganga/Hooghly river channel is 226 apparent for T (t_{62} = 12.7, p < 0.01), pH (t_{62} = 7.9, p < 0.01), EC (t_{62} = 2.5, p < 0.05), and the 227 concentrations of the dissolved ions Ca ($t_{61} = 5.1$, t < 0.01), Mg ($t_{61} = 2.4$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, p < 0.05), Sr ($t_{61} = 2.6$, $t_{61} = 0.05$), Sr ($t_{61} = 2.6$, $t_{61} = 0.05$), Sr ($t_{61} = 2.6$, $t_{61} = 0.05$), Sr ($t_{61} = 0.05$), Sr 0.05), and Si (t_{61} = 8.7, p < 0.01). The trends in concentrations of Cl (t_{62} = 1.5; p = 0.14), Na (p_{61} = 1.7, 228 229 p = 0.10) and NO₃ ($t_{62} = 1.0$, p = 0.31) generally increase downstream, albeit not statistically 230 significantly and concentrations are more variable than the other parameters. In contrast, there is a 231 significant decrease downstream for DO ($t_{62} = -5.4$, p < 0.01) and SO₄ ($t_{62} = -3.6$, p < 0.01). In 232 comparison to the upper reaches, the lower reaches of the basin are generally characterized by 233 relatively low concentrations of DO and SO₄, and relatively high T, pH, EC, Ca, Mg, Sr, Si, Na, Cl and NO_3 , consistent with the increasing cumulative inputs (e.g. urban centres, agricultural zones, tributary 234 235 influences) present along the river continuum. In some cases there are clear localized inputs (e.g. site 236 G14 at the city of Kanpur, Uttar Pradesh), near where previous studies have reported higher pollution 237 concentrations (Bowes et al., 2020; Trivedi, 2010). However, the general significance of downstream

trends indicates that regional rather than local controls are dominant for these selected parameters
on a basin-wide scale. It is important to note that downstream distance is an analogue for covariables
including elevation, temperature, flow velocity and discharge, and population density, and thus is not
an independent explanatory variable. A limited comparison of selected parameters to Indian
regulatory standards is in Supplementary Information.



Figure 2. Basin-wide downstream trends represented as bar plots of measurements/concentrations of selected parameters: (A) electrical conductivity (EC); (B) pH; (C) dissolved oxygen (DO); (D) temperature (T); (E) NO₃-N; (F) Cl; (G) SO₄; (H) Ca; (I) Na; (J) Mg; (K) Sr and (L) Si. No Ca, Na, Mr, Sr, Si data available for site G49. Grey arrows indicate statistically significant (p < 0.05) downstream trends (tributaries excluded). Bar colour indicates Ganga/Hooghly (peach) and tributaries (purple). Y-axis represents sequential order and does not quantify downstream distance.

245

246

247 3.2 Characteristic Hydrogeochemical Zones & Impact of Key Tributaries

Segmented regression enables further interpretation of spatial trends and possible underpinning controls and processes across the River Ganga (**Figure 3**). Although patterns are distinct and parameter-specific, there is a dominant grouping of parameters (*e.g.* EC, NO₃, Cl, SO₄, Na, Mg, Sr) displaying similar trends characterized by: (1) an increase to ~ 1500 km downstream; (2) a decrease until the major tributary inputs at ~ 1800 km; (3) a region of relative stability (~ 1800 – 2700 km); and (4) a sharp increase ~ 2700 km downstream.



Figure 3. Selected physico-chemical/hydrogeochemical parameters against distance from source for the main Ganga/Hooghly samples (tributaries excluded). Blue lines indicate the intersection

points of major tributaries in increasing distance downstream: Ramganga (~ 875 km; yield ~ 15,300 \cdot 10⁶ m³), Yamuna (~ 1270 km; yield ~ 76,000 \cdot 10⁶ m³), Ghagara (~ 1750 km; yield ~ 94,400 \cdot 10⁶ m³), Sone (~ 1770 km; yield ~ 22,420 \cdot 10⁶ m³), Gandak (~ 1810 km; yield ~ 52,200 \cdot 10⁶ m³) and Kosi (~ 2130 km; yield ~ 61,560 \cdot 10⁶ m³). Parameters shown are: (A) electrical conductivity (EC); (B) pH; (C) dissolved oxygen (DO); (D) temperature (T); (E) NO₃-N; (F) Cl; (G) SO₄; (H) Ca; (I) Na; (J) Mg; (K) Sr; (L) Sr:Ca; (M) Cl:Na; (N) Cl:NO₃-N; (O) Ca:Na and (P) Mg:Na. Green lines are modelled segmented linear regression outputs; red dots and error bars (95 % confidence interval) are estimated breakpoints calculated from segmented linear regression. Mean annual yield of River Ganga is ~ 525,000 \cdot 10⁶ m³; all yields as reported elsewhere (Mariya et al., 2019).

255	The combination of breakpoints (<i>i.e.</i> approximate locations at which changes in trends are observed)
256	across parameters (Figure 4A) enables identification of approximate locations of minor and major
257	hydrogeochemical transitions which impact multiple parameters across the basin (Table S1). These
258	breakpoint transitions (± 95 % confidence interval) can be described with regard to distance from
259	source as follows: (i) ~ 530 \pm 100 km (near G08) minor shift in T and ORP; (ii) ~ 790 \pm 140 km (near
260	G09/G10) shift in NO ₃ ; (iii) ~ 1020 \pm 110 km (near G14/G15) minor shift in DO and Ca; (iv) ~ 1230 \pm 70
261	km (near G16/G17) shift in Sr:Ca; (v) \sim 1400 \pm 70 km (near G20 and downstream River Yamuna
262	tributary) minor shift in Ca:Na and Mg:Na; (vi) $^{\sim}$ 1570 \pm 70 km (near G24/G25) major shift in ORP, NO ₃ ,
263	Sr:Ca, Cl:Na, Mg:Ca, pH, EC, Cl, SO ₄ , Mg, Na, Sr; (vii) \sim 1850 ± 50 km (downstream of River Gandak
264	tributary and between G42/G43) major shift in ORP, NO ₃ , Sr:Ca, Mg:Na, Cl, Na; (viii) ~1930 \pm 80 km
265	(near G45/G46) minor shift in pH, EC, Sr; (ix) \sim 2120 \pm 80 km (downstream of River Koshi tributary near
266	G49/G51) major shift in DO, Ca:Na, Cl:Na, Mg:Ca, SO4, Mg; (x) ~ 2380 ± 80 km (near Ganga/Hooghly
267	transition and G61) minor shift in ORP, DO, pH; and (xi) \sim 2700 \pm 5 km (downstream of Kolkata near
268	estuary mouth and between G79/G80) major shift in Sr:Ca, Cl:Na, Mg:Ca, EC, Cl, Na and Cl:NO₃.



Figure 4. (A) Estimated breakpoints for selected physico-chemical/hydrogeochemical parameters as grouped by physico-chemical parameters (grey), anions (red), major/trace elements (blue) and elemental ratios (green). Small black dots represent sampling locations. (B) Modelled parameter trends based on segmented linear regression fits normalized to the mean value of each parameter. Line colours indicate the same groupings as shown on A. (C) Indicative patterns (no y-scale) of parameters with broadly similar behaviour, namely Type 1 (Cl, EC, Na, NO₃, Mg, Sr:Ca, Cl:Na, Cl:NO₃); Type 2 (Ca:Na, Mg:Na); Type 3 (Sr, SO₄); Type 4 (T, pH, Ca); and Type 5 (DO). (D) Elevation profile. Vertical dashed lines indicate modelled breakpoints of major hydrogeochemical transitions between zones (zones also correspond to background-coloured boxes and are defined as where there are overlapping breakpoints of \geq 4 parameters) and vertical dotted lines indicate minor hydrogeochemical transitions (2 or more parameters). Single-parameter shifts are not shown as vertical indicators.

269

270

This analysis identifies five key zones across the basin defined based on areas bounded by the estimated major breakpoints (**Figure 5** and **Table S1**). Zone 1 stretches from the upper reaches of the River Ganga downstream for ~ 1500 km and is generally characterized by increasing T, NO₃, DO, Sr:Ca. Ca:Na and Mg:Na, consistent both with increased agricultural and urban inputs as well as carbonate 275 weathering (Bickle et al., 2005). Zone 2 starts ~ 1570 ± 70 km downstream, near Ghazipur and downstream of the confluence with the River Yamuna tributary (annual yield ~ 76,000 \cdot 10⁶ m³ or ~ 14 276 277 % of that of the total Ganga yield) (Mariya et al., 2019). Zone 2 is characterized by significant decreases 278 of a number of parameters (EC, pH, DO, NO₃, Cl, SO₄, Mg, Na, Sr, Sr:Ca, Cl:Na and Mg:Ca) and an 279 increase in ORP, consistent with mixing and dilution from the Yamuna. Zone 3 starts at \sim 1850 ± 50 280 km, downstream of Patna and the confluence of three key tributaries, the Rivers Ghagara, Sone and 281 Gandak, of approximate annual yields of ~ $94,400 \cdot 10^6$ m³, ~ $22,420 \cdot 10^6$ m³ and ~ $52,200 \cdot 10^6$ m³, 282 or ~ 18 %, ~ 4 % and ~ 10 % of the yield of the Ganga, respectively (Mariya et al., 2019). Zone 3 is 283 characterized by an initial increase in NO₃, Cl and Na concentrations, and a decrease in ORP, Sr:Ca and 284 Mg:Na, likely reflecting combined contributions from both urban and tributary inputs. Continuing 285 downstream, pH, EC and Sr shift from decreasing to increasing values, perhaps from urban inputs 286 around Begusarai. Zone 4 begins ~ 2120 ± 80 km downstream, following the confluence with the Koshi 287 tributary (annual yield 61,560 \cdot 10⁶ m³ or ~ 12 % of that of the Ganga) (Mariya et al., 2019) and other 288 urban areas including Bhagalpur. This is initially characterized by an increase in DO, SO₄, Mg, Cl:Na 289 and Mg:Ca, and a decrease in Ca concentration. Within Zone 4, there is also a minor shift near the 290 Ganga/Hooghly transition, where DO and pH decrease and ORP increases. Zone 5, commencing ~ 291 2700 ± 5 km along the river and downstream of Kolkata, is strongly consistent with an estuariane 292 signature, with substantial increases in EC, Cl, Na, Sr:Ca, Cl:Na, Mg:Ca and Cl:NO₃ near the coast. 293 Characteristic trends across all zones are shown with normalized patterns for selected parameters and 294 for types/grouping of parameters which behave similarly (Figure 4B & C).



Figure 5. Map with indicative zones as divided by major hydrogeochemical transition (red bars; overlapping confidence intervals of estimated breakpoints of \geq 4 hydrogeochemical parameters), minor hydrogeochemical transition (orange bars; overlapping confidence intervals of estimated breakpoints of 2 – 3 hydrogeochemical parameters) and single parameter shift (yellow bars); see **Table S1** for details. Estimated breakpoints are calculated from segmented linear regression. Zones 1 – 5 are split by major hydrogeochemical transitions. Sites G01 and G02 are overlapping within symbol sizes shown. Underpinning geology (Wandrey and Law, 1998) with Q = Quaternary sediments; pC = undivided Precambrian rocks; N = Neogene sedimentary rocks; Jms = Jurassic metamorphic and sedimentary rocks; TrCs = Lower Triassic to Upper Carboniferous sedimentary rocks; MzPz = Mesozoic and Paleozoic intrusive and metamorphic rocks; Pz = undivided Paleozoic rocks; Ti = Tertiary igneous rocks; dark grey = other; light grey = no data. Cities and exaggerated river centrelines (which do not represent river width) are from Natural Earth (<u>https://www.naturalearthdata.com/</u>). Sample IDs are XYY where X is G (Ganga or Hooghly) or T (tributary) and YY is a sequential downstream number.

299 3.3 Dominant Hydrogeochemical Controls

300

301 3.3.1 Principal Component Analysis

302 Principal component (PC) analysis was undertaken to initially screen the hydrogeochemical processes 303 most likely to be dominant across the basin (Figure S4 and Table S2). Four components explained ~ 304 94 % of the variance, with PC1 (63.4 % of variance) strongly positively influenced by Mg, EC, Sr, Na, 305 Ca, NO₃, Cl, SO₄, pH and Si, consistent with dilution and mixing as major regional hydrogeochemical 306 controls. A number of samples, especially from Zones 1 and 2 are strongly dominated by PC1. PC2 307 (17.9 % of variance) instead has strong positive loadings for pH, Si and Ca, largely consistent with 308 controls on silicate and carbonate weathering, and a negative loading for SO₄, which is indicative of 309 sulfide weathering. Zone 1 samples trend towards the bottom left of the PC plot, consistent with the 310 lower pH and Si of the upper reaches. Zone 4, in particularly, is characterized by higher loadings of 311 PC2.

312

313 3.3.2 Mixing and Dilution

314 The relationship between Na and Cl (Figure 6A) is consistent with mixing and dilution as dominant 315 hydrogeochemical controls across the basin. Fresh (low Na, low Cl) sources are observed in upstream 316 reaches of Zone 1, consistent with Himalayan runoff (Bisht et al., 2018). Most samples are parallel to, 317 or near, the equimolar line representing differing degrees of mixing between fresher sources (e.g. 318 near G01) and high Na-Cl end-members (e.g. relatively polluted site G14 and estuarine site G81) 319 and/or halite dissolution. The tributaries clearly have mixed inputs, with the Ghaghara, Sone, Gandak 320 and Koshi characterized by relatively fresh sources, likely to influence mixing/dilution at the 321 confluences, whereas the Yamuna has relatively high Na and Cl concentrations, potentially due to high 322 sewage effluent inputs from the cities of New Delhi, Ghaziabad and Agra (Mandal et al., 2010).

323 Concentrations of Na in slight excess of what would be expected from halite dissolution, especially at 324 the upstream reaches, are likely to derive from other water-rock interactions. The general trend towards higher Na and Cl downstream is consistent with evaporative concentration (especially given 325 consistency across parameters noted in PC1) and/or cumulative inputs from halite dissolution. Halite 326 327 dissolution has been observed to be an important process in other circum-Himalayan basins including the Three Rivers of Eastern Tibet (Noh et al., 2009). Although Cl is commonly used in hydrological 328 329 studies as a relatively conservative tracer, competing processes (e.g. mineral dissolution, diffusion), 330 can also influence CI concentration, introducing uncertainty in in-depth interpretation (Horner et al., 2017; McArthur et al., 2012; McArthur et al., 1989). 331



Figure 6. Basin-wide (A) Na versus CI; (B) NO₃₋N versus CI; (C) Ca:Na versus Mg:Na (red dash line indicates linear regression); and (D) Sr:Ca versus Ca:Na all samples (n = 81). The silicate end-member indicated on (C) is defined as Ca:Na = 0.35 ± 0.15 and Mg:Na = 0.24 ± 0.12 (Gaillardet et al., 1999). Colour scale indicates downstream distance from source and symbol shape indicates Ganga/Hooghly (square) and tributaries (triangle); all data is on a molar basis.

333

334 3.3.3 Human-environment interactions

335 The relationship between NO₃ and Cl is more variable (Figure 6B). At upper reaches of Zone 1, Cl 336 increases whereas NO₃ remains roughly constant, suggesting water-rock interactions (e.g. halite dissolution) increase the concentration of Cl. Throughout the rest of the catchment, NO₃ inputs 337 338 increase, approximately following the 1:1 line with Cl. Elevated NO₃ concentrations are likely to arise 339 from both urban and agricultural inputs, with Zone 2 and parts of Zone 1 generally having the highest 340 NO₃:Cl ratios. Additional high concentrations of Cl in Zone 5 without a proportional increase in NO₃ 341 are likely from estuarine inputs (Kaul and Froelich Jr, 1984). Detailed interpretation of sources and processing of nutrients in this system, including with regard to land use/land cover, is the subject of 342 343 ongoing investigation by co-authors.

344

345 3.3.4 Water-rock interactions

346 The relationship between Ca:Na and Mg:Na (Figure 6C) suggests that there are substantial 347 contributions associated with carbonate weathering particularly in Zone 1 as well as from the 348 northerly Ghaghara and Gandak tributaries. These geological controls suggest that sources 349 representing the end-member ratios of Ca:Na and Mg:Na are likely to be the dominant controls on 350 overall major element water chemistry. Upstream reaches generally trend towards the calcite (CaCO₃) 351 end-member whereas the estuarine samples approach the silicate end-member. The Ca:Na of 352 Himalayan silicate rocks is always < 1 (Bickle et al., 2005 and references within), with an estimated 353 granitic silicate end-member of Ca:Na = 0.35 ± 0.15 (Gaillardet et al., 1999), consistent with the 354 furthest downstream samples. The slope of the linear regression between Ca:Na and Mg:Na is ~ 3.2 355 \pm 0.1, slightly higher than previously reported (~ 1.1 \pm 0.4) for the upper reaches of the basin (Bickle 356 et al., 2005).

357 Sr:Ca values (Figure 6D) are dependent on rock type; with Sr:Ca ratios previously reported to be higher 358 for limestones ($\sim 0.22 \pm 0.02$) than dolomites ($\sim 0.15 \pm 0.02$) in the Ganga headwaters (Bickle et al., 359 2005). The ratios of Sr:Ca to Ca:Na tend to be clustered particularly in Zone 2, 3 and 4 suggesting 360 common rock types in those zones. Both of the cation plots Ca:Na to Mg:Na (Figure 6C) and Sr:Ca to 361 to Ca:Na (Figure 6D) show reasonable approximations of two-component mixing of carbonate and 362 silicate-derived components and are consistent with trends previously reported (Bickle et al., 2005). Deviations from dominant two-component mixing are likely to derive from additional water-rock 363 364 interactions related to, for example, evaporites, phosphorites and siliceous lithologies (Bickle et al., 365 2005).

366 The downstream trends of Ca are strongly in agreement with inferred HCO₃⁻, with very similar 367 breakpoints identified ~ 1000 km downstream (Figure S5), consistent with trends expected from the 368 weathering of carbonates and silicates. In the middle zones where many parameters decrease 369 significantly, carbonate dissolution is likely to contribute to sustaining Ca and pH, resulting in the only 370 slightly decreasing trends for these parameters. Sulfide weathering also appears to be an important 371 process impacting concentrations of SO₄ and indeed sulfide minerals are very prevalent in the Ganges-372 Brahmaputra basin (Fendorf et al., 2010; Galy and France-Lanord, 1999). Interestingly, although the 373 acidity produced by sulfide weathering may influence carbonate dissolution (Bufe et al., 2021), the 374 relative importance and co-variation of these processes appears to be spatially-dependent as Ca 375 appears to be continually produced within Zone 3 even when SO₄ is decreasing. Selected important 376 chemical reactions expected to impact the hydrogeochemistry are shown in Supplementary 377 Information.

378

379 3.3.5 Groundwater-surface water interactions

380 Groundwater has been demonstrated to substantially contribute to the water in the River Ganga, in 381 addition to glacial melt and surface runoff (Mukherjee et al., 2018). Groundwater-surface water

382 interactions are largely controlled by the relative difference between river stage and groundwater 383 level as well as regional geology, with higher permeability sediments in the upper reaches of the Ganga 384 Basin (Bonsor et al., 2017) leading to increased groundwater-surface water connectivity. During most 385 of the year, the dominant groundwater flow direction is towards the River Ganga constituting its 386 baseflow (Mukherjee et al., 2018), although this is expected to be strongly seasonally dependent as 387 hydraulic gradients change with the monsoon (Lu et al., submitted). Relatively high degrees of 388 groundwater-surface water connectivity are expected in the post-monsoon season, both for the 389 Ganga and key tributaries. The observed chemical signatures are likely to thus inherently reflect this 390 groundwater-surface water continuum, as well as the water-rock interactions which may have 391 influenced the composition of the contributing groundwater. Particularly in zones with minimal 392 influence from tributaries (e.g. Zone 1), a strong influence from groundwater is expected to be 393 reflected in the observed surface water signatures. In particular, the observed freshening of river 394 water (e.g. decreased EC, Na, Cl, NO₃, SO₄) at the transition to Zone 2 begins ~ 1570 km, which is 395 upstream of the contributions from most of the major tributaries ~ 1700 – 1800 km. A further possible 396 explanation for this is the release of stored fresh water (e.g. monsoonal flood water) from bank 397 storage zones which are influenced both by the permeability of surrounding sediments (Bonsor et al., 398 2017) as well as the depositional patterns related to the energy of the river (Rhodes et al., 2017). The 399 upper reaches of the Ganga have a steeper elevation gradient (Figure 4D) and are higher energy, which 400 may lead to relatively high release of bank-stored water back to the river (Rhodes et al., 2017), as 401 compared to lower energy zones further downstream. Localized biogeochemical processes in the 402 hyporheic zone may also influence chemical signatures, although potential impacts are expected to 403 be small because major patterns are reasonably explained by dominant processes including mixing, 404 dilution, water-rock interactions and, to a lesser extent, human-environment interactions.

405

406 **3.4. Representativeness of Longitudinal Sampling**

407 Temporal and spatial representativeness can be inherent limitations in longitudinal sampling of large
408 river systems. Here we have integrated a number of steps to assess and quantify sampling
409 representativeness; specific recommendations follow in Section 3.5.

410 3.4.1 Short-term Variations

411 The extent of short-term variability (e.g. hourly to weekly time-scales) in hydrogeochemical conditions 412 has been estimated using various approaches (see Supplementary Information and Figure S6). 413 Substantial short term (e.g. hourly) shifts were observed over 15 minute sampling intervals (Figure 414 **S6A - D**), indicating that short-term impacts can clearly be important. Variation in pH and T is broadly 415 consistent with other studies of diurnal patterns in surface waters (Nimick et al., 2011; Pokrovsky and 416 Shirokova, 2013) and likely reflects a combination of diurnal changes arising from natural daily shifts 417 in temperature (and associated parameters), river metabolism (Cohen et al., 2013), shifting flow 418 regimes of water within the river, and variable urban, anthropogenic and other upstream inputs. 419 Although short-term variability may have significant impact on local biogeochemical processes, the 420 extent of temporal variability on these timescales is much less than the spatial variability observed 421 longitudinally.

422

423 **3.4.2** Seasonal/Annual Variations & Comparison to Historical Data

The magnitude and spatial trends in our post-monsoon data show strong general agreement with historical records from India's National Water Quality Monitoring Programme (NWQM) (Central Pollution Control Board, 2019) (**Figure S8**). The relative similarity and consistency between T, pH and EC across years corroborates that the major hydrogeochemical controls (including concentrations of dominant inorganic ions) are likely largely controlled by regional factors including hydrogeological setting and climate. Importantly, however, temporal representativeness may be parameter-

430 dependent, with wider variability expected particularly for parameters associated with localized
431 anthropogenic activities (*e.g.* nutrients, microplastics).

432

433 **3.4.3** Representativeness of Bank-side Sampling

River cross-sectional variability is discussed in Supplementary Information and Figure S9. In brief, the
 cross-sectional variability is generally less than the diurnal variations, and considerably less than the
 longitudinal variability observed.

437

438 **3.5 Adaptability of Approach and Future Directions**

439 Although we report here a hydrogeochemical investigation across the River Ganga in India, our 440 approach could easily be adapted to improve the understanding of dynamics and processes of other 441 large river systems and/or for systematic investigation of other types of parameters. In large river 442 systems which span diverse environmental conditions (e.g. of varying geological, climatic, topographic 443 and anthropogenic characteristics), it is imperative to identify the dominant drivers impacting regional 444 water quality. In this case, statistical analysis including segmented regression and principal 445 component analysis allows distinct zones and key hydrogeochemical controls (e.g. dilution, mixing, 446 water-rock interactions) to be identified across the basin. Once a baseline study has been developed 447 and implemented, such as is reported here, future studies could further expand the spatial and/or 448 temporal resolution and/or types of parameters considered. Selection of appropriate methods for 449 data analysis and interpretation are key design considerations for transferability and depend on the 450 nature of specific research aims. To our knowledge, a systematic, coordinated basin-wide dataset for 451 hydrogeochemical understanding of the River Ganga at this large scale has not been established in 452 India. This approach creates a platform to investigate other parameters and to establish comparisons 453 to other large river basins internationally in the future.

454 Whilst large-scale longitudinal surveys enable highly valuable spatial information to be obtained, the 455 limitations inherent in longitudinal surveys (e.g. capturing complexities of localized inputs, diurnal 456 hydrogeochemical controls or seasonal variability) must be considered. Recommendations to 457 address, mitigate and/or quantify the impact of some of these inherent limitations whilst undertaking 458 a large-scale longitudinal survey include: (i) conduct sampling ideally at the same time of day to 459 mitigate the impact of diurnal variability (although this is not always feasible in studies across large 460 river systems); (ii) undertake systematic time-series sampling throughout an entire day, ideally both at urban and rural locations, to quantify the extent of diurnal variability; (iii) undertake supplementary 461 462 cross-sectional surveys at selected key locations, including near the extreme ends of longitudinal 463 surveys, noting potential logistical constraints such as practicalities of small boat access in a busy 464 shipping/transport hub such as Kolkata; (iv) compare data to historical records to assist in 465 contextualizing annual representativeness; (v) repeat longitudinal sampling of all or selected sites in 466 contrasting flow conditions (e.g. post- and pre-monsoonal conditions) to understand how 467 underpinning hydrogeochemical processes may change throughout the year. Whilst many of these 468 measures were integrated in our study here, it would be a recommended target for future work to 469 repeat a similar longitudinal survey in the Ganga Basin under relatively low-flow conditions to expand 470 temporal resolution of the dataset.

471

472 4. Conclusions

Here we develop and apply a systematic approach for advancing understanding of hydrogeochemical dynamics and processes in large river systems, as demonstrated via a longitudinal survey (> 2500 km) of the River Ganga and key tributaries from the Himalayas to the Bay of Bengal. The application of our framework evidences that overall, the lower reaches of the basin are characterized by higher T, pH, EC and concentrations of Ca, Mg, Sr, Si, Na, Cl and NO₃, and lower DO and SO₄, which is largely consistent with expected cumulative downstream inputs arising from both water-rock interactions

479 and human-environment interactions over 1000s of kilometres. Although there are localized inputs 480 in some cases, regional controls are likely dominant for most of these parameters on a basin-wide 481 scale. Segmented regression enables the identification of estimated downstream breakpoints and 482 five associated hydrogeochemical zones. The framework has been used to reveal that mixing and 483 dilution are the most important hydrogeochemical controls across the basin, as well as carbonate and 484 silicate weathering, strongly influencing the major element composition of surface water. The 485 magnitude and spatial trends in our data are generally in strong agreement with historical 486 governmental records. This agreement suggests that the major hydrogeochemical controls are largely 487 controlled by regional factors (e.g. hydrogeological setting), although annual representativeness may 488 not extend to all water quality parameters, especially those directly related to anthropogenic 489 activities. The magnitude of short-term temporal variability in water quality parameters was found to 490 be less than their longitudinal diversity. Selected cross-sectional surveys indicated some cross-491 channel variability, particularly near tributary inputs, likely attributed to differences in stream depth, 492 flow and mixing. Rigorous interpretation of our results applying our framework to the River Ganga 493 allows limitations of large-scale longitudinal sampling programs to be identified, including for instance 494 challenges in assessing localized inputs and temporal/seasonal controls, along with some strategies to 495 mitigate these impacts. In addition to providing new insight to the dominant hydrogeochemical 496 processes impacting surface water composition in the River Ganga, our systematic approach is 497 adaptable to other parameters and/or similar coordinated surveys of other large river systems across 498 the world.

499

500 Acknowledgements

501 This research was supported by two NERC-DST Indo-UK Water Quality Programme projects 502 (NE/R003386/1 and DST/TM/INDO-UK/2K17/55(C) & 55(G) to DAP *et al*; see <u>www.farganga.org</u> and 503 NE/R003106/1 and DST/TM/INDO-UK/2K17/30 to DR *et al*.), NE/R000131/1 to Jenkins *et al*. and a

504 Dame Kathleen Ollerenshaw Fellowship (LAR). This work is a contribution to the UNESCO UNITWIN in 505 Ecohydrological Interfaces by partners at the University of Birmingham (including the UNESCO Chair 506 in Water Sciences) and IIT-Roorkee. Aman Gaurav and Siddhu Kumar (both Mahavir Cancer Sansthan), 507 Sanveev Kumar (IIT-Roorkee), Sam Addison (The University of Manchester) and Daniela Mewes 508 (formerly University of Birmingham) are thanked for contributions to fieldwork. Alun Owen 509 (University of the West of England, Bristol), Ilya Strashnov and Roseanna Byrne (both The University 510 of Manchester) are thanked for analytical support. DCG publishes with the permission of the Director, 511 British Geological Survey (NERC). We thank the anonymous reviewers for their very constructive 512 suggestions to improve the manuscript. The views expressed here do not necessarily represent those of the institutions, funders or individuals whose support is acknowledged. 513

514

515 Author Contributions

516 LAR: investigation, formal analysis, methodology, visualization, supervision, writing – original draft. 517 BGF: investigation, methodology, data curation. MJB: investigation, conceptualization, methodology, 518 funding acquisition, resources. KK: investigation; software, methodology. AK: investigation, 519 methodology. RK: investigation. SuK: investigation, methodology. MH: investigation. BH: 520 **RMST**: investigation, methodology, funding acquisition. investigation, methodology. DSR: 521 methodology, resources. HAN: interpretation. US: interpretation. LAK: analysis, resources. DJEN: 522 analysis, resources. DM: software, interpretation. DMH: interpretation, funding acquisition. AG: 523 conceptualization, funding acquisition, methodology, resources. BC: funding acquisition. HJ: funding 524 acquisition. TKD: funding acquisition. DMR: conceptualization, funding acquisition, resources. StK: 525 conceptualization, funding acquisition. DCG: conceptualization, funding acquisition. DAP: 526 conceptualization, funding acquisition, resources. All authors have contributed to manuscript 527 reviewing and editing.

528

529

530 References

- Australian Government 2018 Charter: National Water Quality Management Strategy, Department of
 Agriculture and Water Resources,
- 533 (https://www.waterquality.gov.au/sites/default/files/documents/nwqms-charter_0.pdf,
- 534 accessed Dec 2021), Canberra.
- Belitz, K., Dubrovsky, N.M., Burow, K., Jurgens, B. and Johnson, T. 2003 Framework for a Ground Water Quality Monitoring and Assessment Program for California: U. S. Geological Survey
 Water-Resources Investigations Report 03-4166
- 538 (https://pubs.usgs.gov/wri/wri034166/pdf/wri034166.pdf, accessed Dec 2021).
- 539 Best, J. 2019. Anthropogenic stresses on the world's big rivers. Nature Geosciece 12, 7 21.
- Bickle, M.J., Bunbury, J., Chapman, H.J., Harris, N.B.W., Fairchild, I.J. and Ahmad, T. 2003. Fluxes of
 Sr into the headwaters of the Ganges. Geochemica et Cosmochimica Acta 67(14), 2567 2584.
- Bickle, M.J., Chapman, H.J., Bunbury, J., Harris, N.B.W., Fairchild, I.J., Ahmad, T. and Pomiès, C. 2005.
 Relative contributions of silicate and carbonate rocks to riverine Sr fluxes in the headwaters of
 the Ganges. Geochim. Cosmochim. Acta 69(9), 2221 2240.
- Bisht, H., Arya, P.C. and Kumar, K. 2018. Hydro-chemical analysis and ionic flux of meltwater runoff
 from Khangri Glacier, West Kameng, Arunachal Himalaya, India. Environmental Earth Sciences
 77(598).
- Bonsor, H.C., MacDonald, A.M., Ahmed, K.M., Burgess, W.G., Basharat, M., Calow, R.C., Dixit, A.,
 Foster, S.S.D., Gopal, K., Lapworth, D.J., Moench, M., Mukherjee, A., Rao, M.S., Shamsudduha,
 M., Smith, L., Taylor, R.G., Tucker, J., van Steenbergen, F., Yadav, S.K. and Zahid, A. 2017.
 Hydrogeological typologies of the Indo-Gangetic basin alluvial aquifer, South Asia.
 Hydrogeology Journal 25, 1377 1406.
- Bowes, M.J., Read, D.S., Joshi, H., Sinha, R., Ansari, A., Hazra, M., Simon, M., Vishwakarma, R.,
 Armstrong, L.A., Nicholls, D.J.E., Wickham, H.D., Ward, J., Carvalho, L.R. and Rees, H.G. 2020.
 Nutrient and microbial water quality of the upper Ganga River, India: identification of
 pollution sources. Environ. Monit. Assess. 192, 533.
- Bufe, A., Hovius, N., Emberson, R., Rugenstein, J.K.C., Galy, A., Hassenruck-Gudipati, H.J. and Chang,
 J.-M. 2021. Co-variation of silicate, carbonate and sulfide weathering drives CO2 release with
 erosion. Nat. Geos. 14, 211 216.
- 560 Central Pollution Control Board 2019 Water Quality Database National Water Quality Monitoring
 561 Programme (NWMP) h<u>ttp://www.cpcbenvis.nic.in/</u> (accessed June 2021).
- 562 Cohen, M.J., Kurz, M.J., Heffernan, J.B., Martin, J.B., Douglass, R.L., Foster, C.R. and Thomas, R.G.
 563 2013. Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large
 564 spring-fed river. Ecological Monographs 83, 155 176.

- Dalai, T.K., Krishnaswami, S. and Kumar, A. 2003. Sr and ⁸⁷Sr/⁸⁶Sr in the Yamuna River System in the
 Himalaya: sources, fluxes, and controls on Sr isotope composition. Geochemica et
 Cosmochimica Acta 67(16), 2931 2948.
- Davies, R.B. 1987. Hypothesis testing when a nuisance parameter is present only under the
 alternative. Biometrika 74, 33 43.
- European Environment Agency 2018 European waters: Assessment of status and pressures 2018,
 EEA Report No 7/2018 (https://www.eea.europa.eu/publications/state-of-water, accessed
 Dec 2021), Copenhagen.
- Fendorf, S., Michael, H.A. and van Geen, A. 2010. Spatial and temporal variations of groundwater
 arsenic in south and southeast Asia. Science 328, 1123-1127.
- Gaillardet, J., Dupré, B., Louvat, P. and Allègre, C.J. 1999. Global silicate weathering and CO2
 consumption rates deduced from the chemistry of large rivers. Chem. Geol. 159(1 4), 3 30.
- 577 Galy, A. and France-Lanord, C. 1999. Weathering processes in the Ganges–Brahmaputra basin and 578 the riverine alkalinity budget. Chem. Geol. 159(1 - 4), 31 - 60.
- Hamner, S., Tripathi, A., Mishra, R.K., Bouskill, N., Broadaway, S.C., Pyle, B.H. and Ford, T.E. 2006.
 The role of water use patterns and sewage pollution in incidence of water-borne/enteric
 diseases along the Ganges river in Varanasi, India. International Journal of Environmental
 Health Research 16(2), 113 132.
- Horner, K.N., Short, M.A. and McPhail, D.C. 2017. Chloride and bromide sources in water:
 Quantitative model use and uncertainty. J. Hydrol. 549, 571 580.
- Jain, C.K. and Singh, S. 2020. Impact of climate change on the hydrological dynamics of River Ganga,
 India. Water & Climate Change 11, 274 290.
- Junqueira, A.M., Mao, F., Mendes, T.S.G., Simões, S.J.C., Balestieri, J.A.P. and Hannah, D.M. 2021.
 Estimation of river flow using CubeSats remote sensing. Sci. Total Environ. 788, 147762.
- Kaul, L.W. and Froelich Jr, P.N. 1984. Modeling estuarine nutrient geochemistry in a simple system.
 Geochemica et Cosmochimica Acta 48(7), 1417 1433.
- Khan, M.Y.A., Hu, H., Tian, F. and Wen, J. 2020. Monitoring the spatio-temporal impact of small
 tributaries on the hydrochemical characteristics of Ramganga River, Ganges Basin, India.
 International Jrounal of River Basin Management 18(2), 231 241.
- Kumar, D. 2017. River Ganges Historical, cultural and socioeconomic attributes. Aquatic Ecosystem
 Health & Management 20(1 2), 8 20.
- Lata, P., Ram, S., Agrawal, M. and Shanker, R. 2009. Enterococci in river Ganga surface waters:
 Propensity of species distribution, dissemination of antimicrobial-resistance and virulence markers among species along landscape. BMC Microbiology 9(140).
- Lokgariwar, C., Chopra, R., Smakhtin, V., Bharati, L. and O'Keeffe, J. 2014. Including cultural water
 requirements in environmental flow assessment: an example from the upper Ganga River,
 India. Water International 39, 81 96.

- Lu, C., Richards, L.A., Wilson, G., Krause, S., Lapworth, D.J., Gooddy, D.C., Chakravorty, B., Polya, D.A.
 and Niasar, V.J. submitted. Quantifying the impacts of groundwater abstraction on Ganges
 river water ingress into shallow aquifers under the rapidly developing city of Patna, India. J.
 Hydrol.
- Mandal, P., Upadhyay, R. and Hasan, A. 2010. Seasonal and spatial variation of Yamuna River water
 quality in Delhi, India. Environ. Monit. Assess. 170, 661 670.
- Mariya, A., Kumar, C., Masood, M. and Kumar, N. 2019. The pristine nature of river Ganges: its
 qualitative deterioration and suggestive restoration strategies. Environ. Monit. Assess. 2019,
 542.
- McArthur, J.M., Sikdar, P.K., Hoque, M.A. and Ghosal, U. 2012. Waste-water impacts on
 groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal
 Basin and Red River Basin, Vietnam. Sci. Total Environ. 437, 390 402.
- McArthur, J.M., Turner, J., Lyons, W.B. and Thirlwall, M.F. 1989. Salt sources and water-rock
 interaction on the Yilgarn Block, Australia: isotopic and major element tracers. Appl. Geochem.
 4(1), 79 92.
- Ministry of Jal Shakti 2021 National Mission for Clean Ganga (Namini Gange Programme)
 (https://nmcg.nic.in/NamamiGanga.aspx, accessed June 2021).
- Moors, E.J., Groot, A., Biemans, H., van Scheltinga, C.T., Siderius, C., Stoffel, M., Huggel, C., Wiltshire,
 A., Mathison, C., Ridley, J., Jacob, D., Kumar, P., Bhadwal, S., Gosain, A. and Collins, D.N. 2011.
 Adaptation to changing water resources in the Ganges basin, northern India. Environmental
 Science & Policy 14(7), 758 769.
- Muggeo, V.M.R. 2008. Segmented: an R package to fit regression models with broken-line
 relationships. R news 9, 20 25.
- Mukherjee, A., Bhanja, S.N. and Wada, Y. 2018. Groundwater depletion causing reduction of
 baseflow triggering Ganges river summer drying. Sci. Rep. 8.
- Narain, S. 2014 Ganga: The River, Its Pollution and What We Can Do to Clean it. A Centre for
 Science and Environment Briefing Paper. (https://cdn.cseindia.org/userfiles/ganga-the-river pollution.pdf, accessed June 2021), New Delhi.
- Nimick, D.A., Gammons, C.H. and Parker, S.R. 2011. Diel biogeochemical processes and their effect
 on the aqueous chemistry of streams: A review. Chem. Geol. 283(1 2), 3 17.
- Noh, H., Huh, Y., Qin, J. and Ellis, A. 2009. Chemical weathering in the Three Rivers region of Eastern
 Tibet. Geochemica et Cosmochimica Acta 73, 1857 1877.
- Pandey, J., Tripathi, S. and Pandey, U. 2016. Anthropogenically Induced Shifts in N:P:Si
 Stoichiometry and Implications in Ganga River. Air, Soil and Water Research 9, 35 43.
- Paul, D. 2017. Research on heavy metal pollution of river Ganga: A review. Annals of Agrarian
 Science 15, 278 286.
- Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., Dufour, S., Liébault, F., RuizVillanueva, V. and Slater, L. 2020. Remotely sensed rivers in the Anthropocene: state of the
 art and prospects. Eart. Surf. Proc. Land. 45(1), 157 188.

- Pokrovsky, O.S. and Shirokova, L.S. 2013. Diurnal variations of dissolved and colloidal organic
 carbon and trace metals in a boreal lake during summer bloom. Water Res. 47(2), 922 932.
- Poudel, D.D., Lee, T., Srinivasan, R., Abbaspour, K. and Jeong, C.Y. 2013. Assessment of seasonal and
 spatial variation of surface water quality, identification of factors associated with water quality
 variability, and the modeling of critical nonpoint source pollution areas in an agricultural
 watershed. Journal of Soil and Water Conservation 68(3), 155 171.
- 647 Rhodes, K.A., Proffitt, T., Rowley, T., Knappett, P.S.K., Montiel, D., Dimova, N., Tebo, D. and Miller,
 648 G.R. 2017. The Importance of Bank Storage in Supplying Baseflow to Rivers Flowing Through
 649 Compartmentalized, Alluvial Aquifers. Water Resour. Res. 53(12), 10539 10557.
- Richards, L.A. and Team Saptanadi 2020 Water Quality across the River Ganga Basin in India: Trends,
 Dominant Geochemical Processes and Impacts (EGU21-10642;
 https://doi.org/10.5194/egusphere-egu21-10642).
- Satya, K. and Narayan, C. 2018. Study of Physico-chemical and Biological Characteristics of the
 Water of River Ganga at Patna, India Current World Environment 13(3), 374 379.
- Seth, R., Singh, P., Mohan, M., Singh, R. and Aswal, R.S. 2013. Monitoring of phenolic compounds
 and surfactants in water of Ganga Canal, Haridwar (India). Applied Water Science 3, 717 720.
- Sharma, B.M., Bečanová, J., Scheringer, M., Sharma, A., Bharat, G.K., Whitehead, P.G., Klánová, J.
 and Nizzetto, L. 2019. Health and ecological risk assessment of emerging contaminants
 (pharmaceuticals, personal care products, and artificial sweeteners) in surface and
 groundwater (drinking water) in the Ganges River Basin, India. Sci. Total Environ. 646, 1459 1467.
- Sharma, B.M., Bharat, G.K., Tayal, S., Larssen, T., Bečanová, J., Karásková, P., Whitehead, P.G., Futter,
 M.N., Butterfield, D. and Nizzetto, L. 2016. Perfluoroalkyl substances (PFAS) in river and
 ground/drinking water of the Ganges River basin: Emissions and implications for human
 exposure. Environ. Pollut. 208, 704 713.
- Sinha, R.K. and Loganathan, B.G. (2015) Water Challenges and Solutions on the Global Scale, pp. 129
 159, American Chemical Society.
- Tripathy, G.R. and Singh, S.K. 2010. Chemical erosion rates of river basins of the Ganga system in
 the Himalaya: Reanalysis based on inversion of dissolved major ions, Sr, and ⁸⁷Sr/⁸⁶Sr.
 Geochem. Geophys. Geosyst. 11(3), 1 20.
- Trivedi, R.C. 2010. Water quality of the Ganga River An overview. Aquatic Ecosystem Health &
 Management 13(4), 347 351.
- Varol, M., Gökot, B., Bekleyen, A. and Şen, B. 2012. Spatial and temporal variations in surface water
 quality of the dam reservoirs in the Tigris River basin, Turkey. CATENA 92, 11.
- Wandrey, C.J. and Law, B.E. 1998 Maps Showing Geology, Oil and Gas Fields and Geological
 Provinces of South Asia. Open-File Report 97-470C (https://pubs.usgs.gov/of/1997/ofr-97470/OF97-470C/ofr97470C.pdf, with associated data in geo8ag package "Geologic map of
 South Asia" downloadable from https://catalog.data.gov/dataset/geologic-map-of-south-asiageo8ag, both accessed 27 June 2021), p. 10, U.S. Department of the Interior Geological Survey.

- Ward, J.V. and Stanford, J.A. 1995. The serial discontinuity concept: Extending the model to
 floodplain rivers. Regulated Rivers: Research & Management 10(2 4), 159 168.
- Whitehead, P.G., Barbour, E., Futter, M.N., Sarkar, S., Rodda, H., Caesar, J., Butterfield, D., Jin, L.,
 SInha, R., Nicholls, R. and Salehin, M. 2015. Impacts of climate change and socio-economic
 scenarios on flow and water quality of the Ganges, Brahmaputra and Meghna (GBM) river
 systems: low flow and flood statistics. Environmental Science Processes & Impacts 17, 1057.
- Wilbers, G.-J., Becker, M., Nga, L.T., Sebesvari, Z. and Renaud, F.G. 2014. Spatial and temporal
 variability of surface water pollution in the Mekong Delta, Vietnam. Sci. Total Environ. 485 486, 653 665.

689