Water Research 128 (2018) 82–91

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Population density controls on microbial pollution across the Ganga catchment

D.G. Milledge [a, *], S.K. Gurjar [b], J.T. Bunce [c], V. Tare [b], R. Sinha [b, d], P.E. Carbonneau [a]

[a] Institute of Hazard, Risk, and Resilience and Department of Geography, Durham University, Durham, UK
[b] Department of Civil Engineering, Indian Institute of Technology Kanpur, India
[c] School of Engineering, Newcastle University, Newcastle, UK
[d] Department of Earth Sciences, Indian Institute of Technology Kanpur, India

ARTICLE INFO

Article history:
Received 17 May 2017
Received in revised form
12 October 2017
Accepted 17 October 2017
Available online 27 October 2017

Keywords:
Faecal coliform
River network
Population density
Catchment-scale

ABSTRACT

For millions of people worldwide, sewage-polluted surface waters threaten water security, food security and human health. Yet the extent of the problem and its causes are poorly understood. Given rapid widespread global urbanisation, the impact of urban versus rural populations is particularly important but unknown. Exploiting previously unpublished archival data for the Ganga (Ganges) catchment, we find a strong non-linear relationship between upstream population density and microbial pollution, and predict that these river systems would fail faecal coliform standards for irrigation waters available to 79% of the catchment's 500 million inhabitants. Overall, this work shows that microbial pollution is conditioned by the continental-scale network structure of rivers, compounded by the location of cities whose growing populations contribute c. 100 times more microbial pollutants per capita than their rural counterparts.

1. Background

Rising demands on water resources raise concerns about the sustainable provision of clean water worldwide. Unclean water poses significant risks of diarrhoea, opportunistic infections, and consequent malnutrition accounting for ~1.7 million deaths annually, of which >90% are in developing countries and almost half are children (Prüss-Ustün et al., 2014). These deaths are primarily due to ingestion of faecal pathogens from humans or animals (Ashbolt, 2004; Kotloff et al., 2013; Prüss-Ustün et al., 2014).

India's growing population and economy are driving rapid urbanisation (30% of the population now live in urban areas (Census of India, 2011a)) and exerting increased pressure on surface and groundwater availability. In rural areas ~67% of the population defecate in the open (Census of India (2011b)), a practice that poses severe risk to health and safety (Clasen et al., 2010; Mara et al., 2010; Ziegelbauer et al., 2012; Kotloff et al., 2013). In urban areas ~80% of the population have access to a toilet (Census of India (2011b)), but only ~30% are connected to a sewage pipeline and few pipelines are connected to a treatment plant (Narain, 2012). The impact of these sanitation problems on surface water quality has been documented for many years at individual sample locations or river reaches across India (Bhargava, 1983; Mukherjee et al., 1993; Baghel et al., 2005; Mishra et al., 2009; Central Pollution Control Board, 2010). However, there has been no catchment-wide quantification of the problem and limited indication of what is driving it. The former is essential to fully understand the scale of intervention required, while the latter might inform decision-making on 'what to do where'. Urban areas often dominate the microbial pollution signal in rivers (Tchobanoglous et al., 1991; Kay et al., 2008; McGrane et al., 2014) but there is little consensus on the extent to which this reflects an increased impact per capita or simply a larger population and thus source. This difference is important since a higher per capita impact indicates reduced attenuation, perhaps due to more efficient delivery to the river system or less efficient treatment. If the difference can be attributed to per capita contribution this will define the extent to which urban or rural focused interventions will improve surface water quality.

We address this question using archival water quality data from across the Ganga (Ganges) catchment and show the pattern of microbial pollution in its major rivers. We compare instream concentrations of a pollution proxy with upstream densities of the two

* Corresponding author. Department of Geography, Durham University, Science Laboratories, Durham DH1 3LE, UK.
E-mail address: d.g.milledge@durham.ac.uk (D.G. Milledge).

https://doi.org/10.1016/j.watres.2017.10.033

major sources of faecal pathogens (humans and livestock) at 100 sites spanning an approximate surface area of 10^6 km^2.

Faecal pathogens are difficult to measure; however thermotolerant coliforms, which originate in faeces (i.e. faecal coliforms, FC), are easily detectable and routinely monitored as indicator organisms (Ashbolt et al., 2001). FCs are not a perfect predictor of human pathogen presence, rather they establish connectivity between defecation and some receiving environment which could be contributed to by a pathogen carrier. New host-specific tracing techniques allow more precise tracking of microbial pollution sources that can help to better assess risks to human health (Harwood et al., 2014, Field and Samadpour, 2007). However, such techniques are not used within routine monitoring in India and thus do not have the spatial coverage required for our analysis. Furthermore, the use of FCs for monitoring pollution is still regarded as a viable measure of drinking and irrigation water quality (WHO, 2017).

Two key issues that must be addressed are: 1) the extent to which the FC signal that we observe reflects human sources; and 2) the potential impact of FC die-off in our pollution tracer. Upstream livestock and human population densities are strongly correlated at the catchment scale, limiting our capacity to identify the source of the pollution signal. To address this, we seek to de-correlate the predictor variables by using a mixing model to estimate contributions from each non-overlapping segment of the catchment (our sub-catchments). To address the impact of die-off in our pollution tracer we adjust the population and livestock densities using a distance decay function then seek decay parameters that will maximise performance of our statistical model.

In the sections that follow we first introduce our null hypothesis that pollution should be linearly related to source density (both with and without accounting for die-off). We then detail our data sources and methods for their analysis, and introduce the mixing model that we use to calculate effective FC concentrations and source densities for each sub-catchment (the non-overlapping segments of the catchment).

2. Theory: expected relationship between FC concentration and upstream source density with and without die-off

The FC concentration (C_{FC}) at a given location is defined by the ratio of the FC flux (Q_{FC}) to the water flux (Q_w):

$$C_{FC} = \frac{Q_{FC}}{Q_w} \tag{1}$$

Under the assumption that there is no die-off in FCs over time, the FC flux is calculated from:

$$Q_{FC} = (P_h N_h + P_a N_a) = (P_h \rho_h + P_a \rho_a) A \tag{2}$$

where: P_h is the production rate of FCs per human head [MPN #$^{-1}$ T^{-1}]; P_a is the production rate per head of livestock [MPN #$^{-1}$ T^{-1}]; N_h and N_a are the total upstream populations of humans and livestock respectively [#]; ρ_h and ρ_a are the upstream population densities of humans and livestock respectively [# L^{-2}]; and A is the catchment area [L^2]. Under the assumption of spatially uniform and time invariant runoff R_w [L T^{-1}] the water flux Q_w [L^3 T^{-1}] is calculated from:

$$Q_w = R_w A \tag{3}$$

Substituting equations (2) and (3) into equation (1) gives the following equation for FC concentration at each measurement point as a function of upstream population density.

$$C_{FC} = \frac{(P_h \rho_h + P_a \rho_a)}{R_w} = k_h \rho_h + k_a \rho_a \tag{4}$$

where: $k_h = P_h/R_w$ and $k_a = P_a/R_w$. It is clear from this relationship that under these assumptions C_{FC} should be a linear function of upstream population and livestock density with the gradients defined by the ratio of production rate, P, to runoff, R_w.

The assumption of no FC die-off is unlikely to be true but controls on die-off remain poorly understood. Given the uncertainties, die-off is most often represented using an exponential decay based on first order kinetics (Crane and Moore, 1986; Sadeghi and Arnold, 2002; Cho et al., 2012):

$$Q_{FC} = Q_0 e^{-k_1 t} \tag{5}$$

where: Q_0 is the FC flux at time t_0 (the time of exit from the gut) [MPN T^{-1}], t is time since exit [T], k_1 is a decay coefficient [T^{-1}]. Assuming uniform time invariant FC velocity from source to measurement point the FC flux Q_{FC} can be expressed as a function of distance:

$$Q_{FC} = Q_0 e^{-k_1 \left(\frac{x}{v}\right)} \tag{6}$$

where: x is the travel distance from source to measurement point [L] and v is the characteristic velocity [L T^{-1}]. Changing population (of people or livestock) with distance x upstream of the sampling point can be calculated as the derivative of N(x):

$$n(x) = -\frac{dN}{dx} = -\rho(x)\frac{dA}{dx} - A(x)\frac{d\rho}{dx} \tag{7}$$

Assuming that FC production rates are time invariant and incorporating characteristic velocity into the decay coefficient to express decay in terms of distance, the FC flux can be calculated by combining equations (2), (6) and (7) and integrating over the range of travel distances from the measurement point to the furthest point upstream:

$$Q_{FC} = \int_0^{x_{max}} \left((P_h n_h(x) + P_a n_a(x)) e^{-kx} \right) dx \tag{8}$$

where change in population (for both humans and livestock) and area are a function of travel distance; and $k = k_1/v$ the distance decay coefficient [L^{-1}]. Substituting equations (3) and (8) into equation (1) gives the following equation for FC concentration:

$$C_{FC} = \int_0^{x_{max}} \left(\frac{(P_h n_h(x) + P_a n_a(x)) e^{-kx}}{R_w A} \right) dx \tag{9}$$

This can be implemented in discrete form by summing over the ncells upslope of the measurement point where for each cell the flow path lengths and routes are derived from digital elevation data, and human and livestock population data from the sources described below.

$$C_{FC} = \sum_{i=1}^{ncells} \left(\frac{(P_h \rho_{hi} + P_a \rho_{ai}) A_i e^{-k x_i}}{R_w A_i} \right) \tag{10}$$

where: ρ_{hi} and ρ_{ai} are the density of human and animal populations respectively in cell i; A_i is the area of cell i; and x_i is the average flowpath length from cell i to the measurement point. Rearranging and simplifying equation (10) gives: