Title: Changes in global groundwater organic carbon driven by climate change and urbanization

Authors: Liza K. McDonough[1, 2, *], Isaac R. Santos[3, 4], Martin S. Andersen[1, 5], Denis M. O'Carroll[1, 5], Helen Rutlidge[1, 5], Karina Meredith[6], Phetdala Oudone[1, 2], John Bridgeman[7], Daren C. Gooddy[8], James P. R. Sorensen[8], Dan J. Lapworth[8], Alan M. MacDonald[9], Jade Ward[8], Andy Baker[1, 2]

Affiliations:

[1] Connected Waters Initiative Research Centre, UNSW Sydney, NSW, 2052, Australia (correspondence: l.mcdonough@unsw.edu.au)

[2] School of Biological, Earth and Environmental Sciences, UNSW Sydney, NSW, 2052, Australia

[3] National Marine Science Centre, Southern Cross University, Coffs Harbour, NSW, 2450, Australia

[4] Department of Marine Sciences, University of Gothenburg, Sweden

[5] School of Civil and Environmental Engineering, UNSW Sydney, NSW, 2052, Australia

[6] Australian Nuclear Science and Technology Organisation (ANSTO), New Illawarra Rd, Lucas Heights, NSW, 2234, Australia

[7] University of Bradford, Bradford, West Yorkshire, BD7 1DP, UK

[8] British Geological Survey, Maclean Building, Wallingford, OX10 8BB, UK

[9] British Geological Survey, Lyell Centre, Research Avenue South ,Edinburgh EH14 6AJ, UK

[*]Correspondence to: l.mcdonough@student.unsw.edu.au

Abstract: Climate change and urbanization can increase pressures on groundwater resources, but little is known about how groundwater quality will change. Here, we rely on a global synthesis (n = 9,404) to reveal the drivers of dissolved organic carbon (DOC), which is an important component of water chemistry and substrate for microorganisms which

control many biogeochemical reactions. Groundwater ions, local climate and land use explained ~ 31% of observed variability in groundwater DOC, whilst aquifer age explained an additional 16%. We identify a 19% increase in DOC associated with urban land cover. We predict major groundwater DOC increases following changes in precipitation and temperature in key areas relying on groundwater. Climate change and conversion of natural or agricultural areas to urban areas will decrease groundwater quality and increase water treatment costs, compounding existing threats to groundwater resources.

Main Text: Groundwater is the largest global source of fresh water. The potability of groundwater is highly dependent upon the concentration of dissolved organic carbon (DOC) due to its ability to alter water chemistry and microbial abundances [1, 2, 3, 4, 5, 6]. Over 100,000 lifetime cancer cases in the United States (U.S.) can be attributed to contaminants in drinking water. A large proportion of the risk identified is associated with the presence of disinfection by products (DBPs) and arsenic [7], both of which are strongly linked to DOC [3, 8, 9, 10, 11]. Chlorination and ozonation used for water treatment can result in harmful by-products including 3-chloro-4-dichloromethyl-5-hydroxy-2(5H)-furanone, brominated acetic acid, trihalomethanes (THMs), formaldehyde, halogenated acetic acids, due to the presence of organic matter [12]. These by-products can be genotoxic, carcinogenic or result in tumors [12]. Since most of the health impacts caused by dissolved organic matter (DOM) are related to the formation of by-products and depend on the concentrations of other water chemical parameters, the World Health Organization [12] and many countries including Australia [13] do not regulate total organic carbon (TOC) or DOC concentrations in drinking water directly but many countries such as USA [14], Canada [15], France [16], China [17] and South Africa [18] highlight potential concerns related to THM formation, health effects and aesthetic quality in the broad DOC range of 0 - 5 mg L^{-1} during treatment.

In addition to health and aesthetic impacts, the presence of DOC in water can lead to membrane fouling after ozonation. In order to avoid this, a biological filtration step is advised

to be added to the water treatment process for water containing DOC concentrations > 1 mg L^{-1} [19]. This indicates that even relatively small DOC increases in raw groundwaters can have impacts not only on human health and water aesthetics, but also on the ease and cost of water treatment. High DOC concentrations can also increase the mobility of other contaminants in groundwater, including heavy metals and nutrients, by complex association with dissolved or colloidal organic matter (OM) [3].

Climate variables such as temperature and precipitation impact on net primary production and microbial activity in ecosystems [20, 21]. This drives availability of vegetation and its decomposition to DOC [21, 22]. Changed precipitation, increasing temperatures and evaporation rates and patterns under future climate change scenarios are expected to alter biomass, impact surface water quantity [23], and subsequently increase domestic and agricultural reliance on groundwater resources. Increasing reliance on groundwater due to climate change impacts may be compounded by urbanization and global population growth which is expected to increase groundwater contamination[24]. Recent research has focused on how climate change and urbanization will change groundwater quantities [25, 26], however understanding the impact of climate change and urbanization on the quality of our freshwater resources is also important [23, 27]. Establishing links between climate change and groundwater quality requires large datasets to produce meaningful global estimates.

Here, we quantify the change in groundwater DOC related to climate change and urban land cover. We present the largest global dataset of 9,404 published and unpublished groundwater DOC concentrations (Table S1) obtained from aquifers in 32 countries across 6 continents (Fig. 1). We provide an analysis of global groundwater DOC concentrations and quantify its key drivers. Specifically, we forecast changes in DOC concentrations due to projected changes in temperature and precipitation, as well as potential increases as a result of urban land use.

Global groundwater DOC distributions

Groundwater DOC concentrations vary spatially and are usually lower than surface water concentrations. The global mean, median and standard deviation of groundwater DOC concentrations are 2.7, 1.0 and 15.1 mg C L^{-1} respectively (Fig. 1). Fig 1A shows that most groundwater DOC concentrations fall within the 0 – 5 mg C L^{-1} range, with 84.1% of samples less than 5 mg C L^{-1}, with the dataset largely dominated by countries in low and mid latitudes.

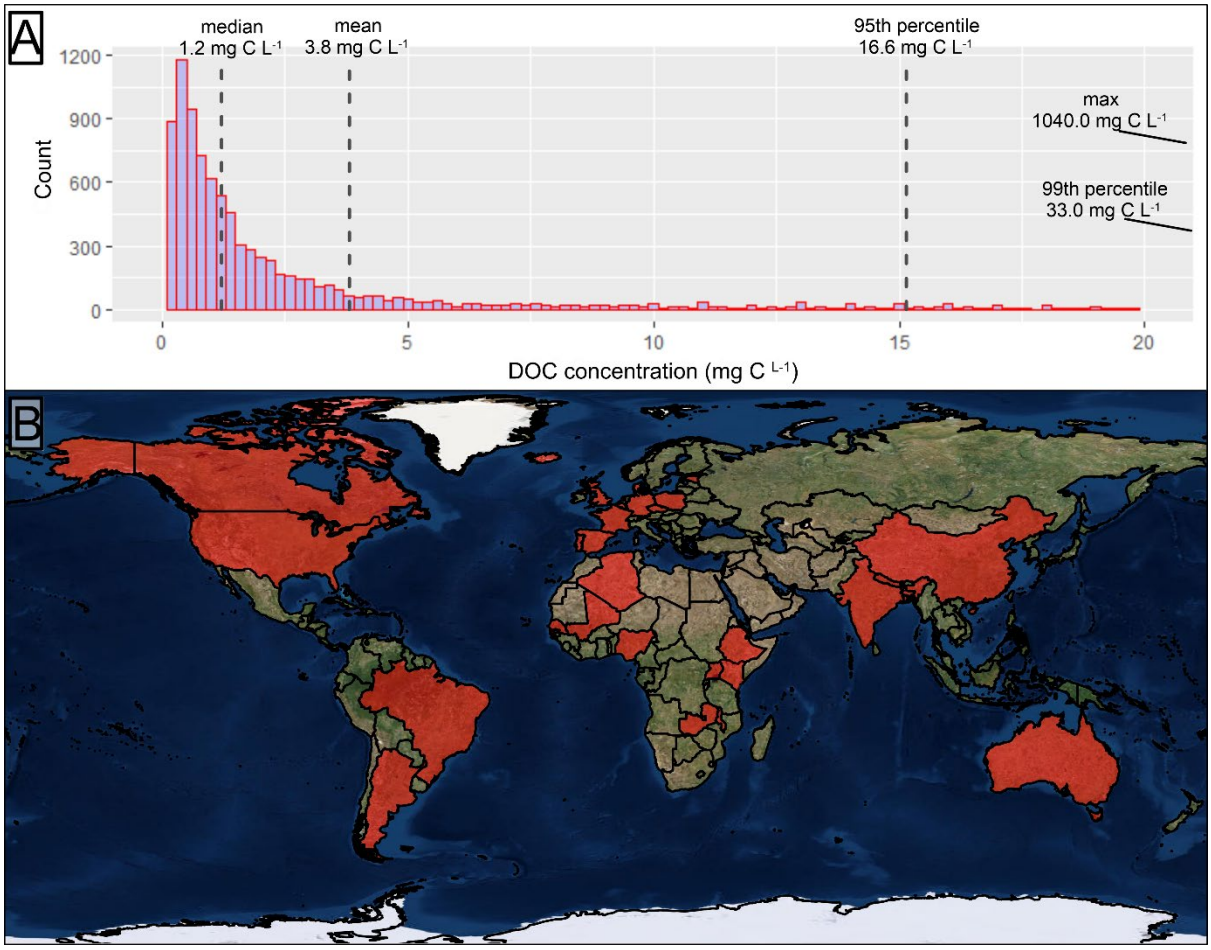


Fig. 1. A) histogram showing median global groundwater DOC concentrations (mg C L^{-1}). Sample sizes for individual countries ranged from 5 to 5,812, with 14 out of 32 countries having n < 30. We have therefore presented aggregated data. Samples above 20 mg C L^{-1} are not included in the graph for visual clarity (n = 337). The black dashed lines indicate the global median (1.2 mg C L^{-1}) and mean (3.8 mg C L^{-1}) and 95th percentile ((16.6 mg C L^{-1}) values respectively. Also shown are the 99th percentile value and the maximum value (33.0

and 1040.0 mg C L^{-1} respectively). B) countries from which groundwater DOC data was obtained.

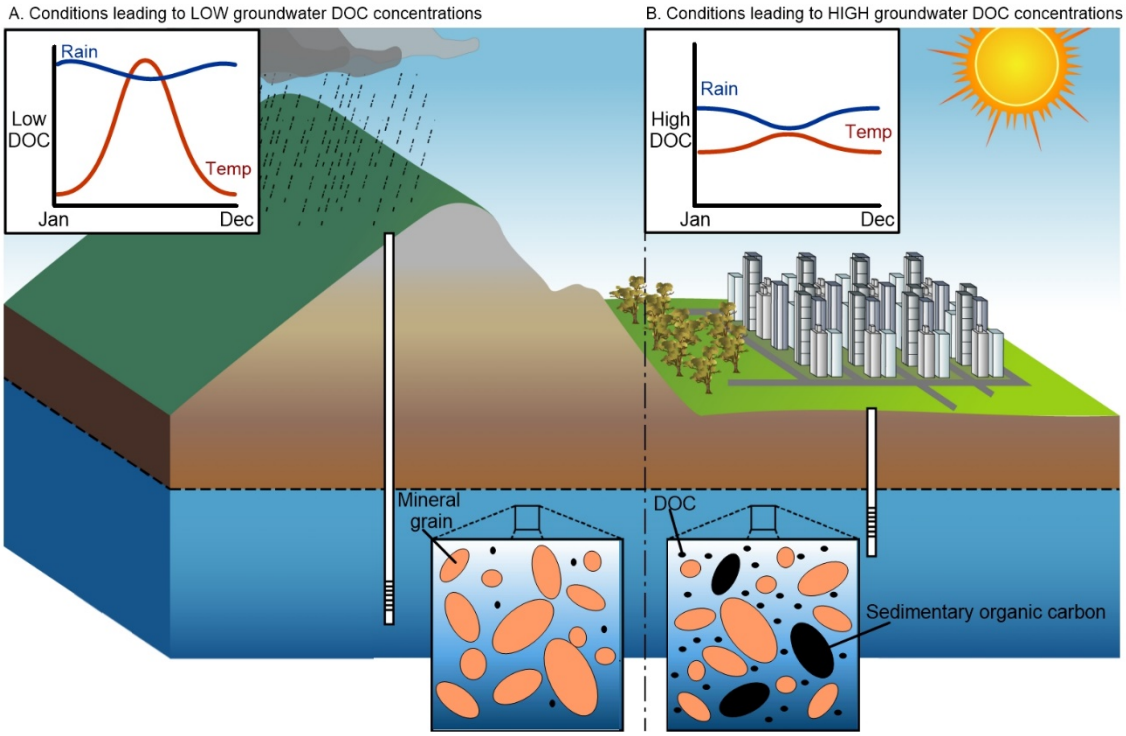
Variations in DOC concentrations between countries (Fig. 1) are likely to be related to recharge rates and aquifer types. World-wide Hydrological Mapping and Assessment Programme (WHYMAP) data [28] suggests that within the U.S. dataset (Fig. S1 [29]), major groundwater basins contain significantly lower DOC concentrations than local and shallow aquifers, and complex hydrogeological structures (both $p < 2.2 \times 10^{-16}$, Fig. S2). Therefore, groundwater age and depth seem to control groundwater DOC. There are also significantly higher DOC concentrations identified in aquifers with < 100 mm $year^{-1}$ recharge compared to those with high recharge rates (100 – 300 mm $year^{-1}$, $p = 2.342 \times 10^{-7}$) and very high recharge rates (> 300 mm $year^{-1}$, $p = 4.857 \times 10^{-5}$, Fig. S2), which could indicate a dilution effect.

Gorundwater DOC controls

To determine the drivers of global DOC concentrations in groundwater, we generated a linear mixed model (Table S2) for a large dataset (n = 2,196) collected by the National Water Quality Assessment (NWQA) program of the U.S. Geological Survey (USGS) [29]. This dataset was selected as it contained supplementary data including chemical parameters unavailable for other samples. This allowed us to extract supplementary climatic data [30], water table depth [31] and land use data [32, 33] for analysis in the model.

Overall, the model explained 47.7% of the variation in DOC concentrations, with 31.3% explained by the fixed factors alone (all fixed and random factors), and 16.3% explained by the random factor aquifer age (age of host rock). Our analysis (Fig. S3, Table S2) shows positive correlations between DOC and temperature in the wettest quarter of the year ($p < 2 \times 10^{-16}$), groundwater temperature ($p < 2 \times 10^{-16}$), and dissolved calcium (Ca) ($p < 2 \times 10^{-16}$), potassium (K) ($p = 2 \times 10^{-13}$) and iron (Fe) ($p < 2 \times 10^{-16}$). There was also a weaker

relationship between DOC and manganese (Mn) ($p < 0.039$). We also found negative relationships between DOC and temperature in the warmest quarter of the year ($p < 2 \times 10^{-16}$), precipitation in the driest month of the year ($p = 0.001$), silica (Si) ($p = 2 \times 10^{-6}$), pH ($p = 4.06 \times 10^{-5}$), sample depth below land surface ($p < 2 \times 10^{-16}$), land elevation ($p = 1 \times 10^{-6}$) and dissolved oxygen (DO) ($p < 2 \times 10^{-16}$). Our analysis also shows negative relationships ($p < 0.01$) between DOC and sodium (Na) ($p = 0.001$), and DOC and precipitation in the wettest month of the year ($p = 0.001$). Areas of urban land use were identified as having 19% higher groundwater DOC concentrations than natural or agricultural areas. Water table depth as a variable improved the overall model fit, however it is not a significant predictor of DOC concentrations ($p = 0.071$). The factors correlated with decreased and increased groundwater DOC concentrations are presented in Fig. 2. While the model represents large scale relationships between DOC and control variables, these relationships can vary on a site by site basis due to site specific variables. Our model implies that large scale groundwater DOC concentrations are determined by the interaction of four major controlling factors: (1) climate (2) urban land-use (3) water chemistry (redox controls), and (4) aquifer age and groundwater residence times (Fig. 3).



142

143 **Fig. 2.** Factors and processes leading to low (A) or high (B) groundwater DOC
144 concentrations. Factors negatively correlated with groundwater DOC concentrations include
145 increasing pH, DO, Na and Si, precipitation in the wettest and driest months, temperature in
146 the warmest quarter, sample depth, elevation and aquifer age. Factors positively correlated
147 with groundwater DOC concentrations include Mn, Ca, Fe, and K, groundwater temperature
148 and temperatures in the wettest quarter. Urban land use was found to be 18% and 19%
149 higher in groundwater DOC concentrations than agricultural and natural land uses
150 respectively.

151

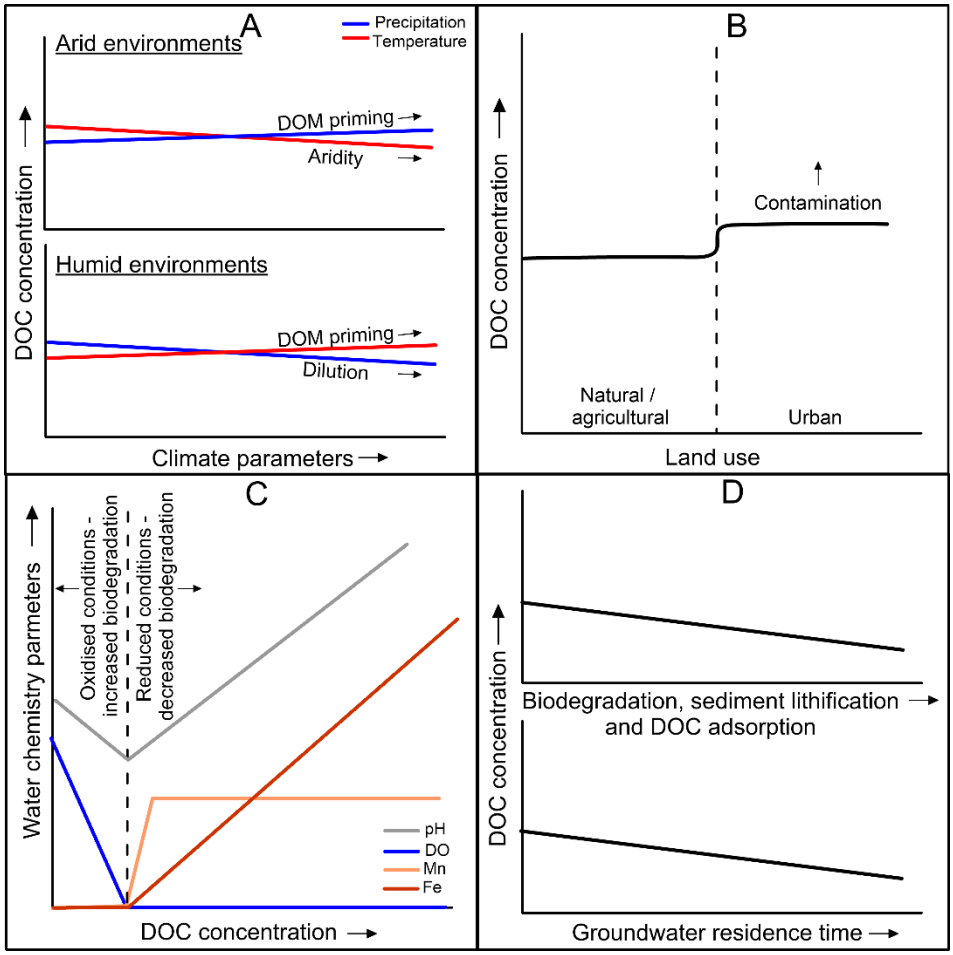


Fig. 3. Conceptual model showing the mechanisms for change in groundwater DOC concentrations. (A) Climate parameters: in arid climates, groundwater DOC concentrations increase with increased precipitation due to the priming of organic matter by microbes under warm and increasingly wet conditions. Increased temperatures in arid environments lead to decreased groundwater DOC due to increasing aridity. Increased precipitation in humid environments decreases groundwater DOC concentrations due to dilution whilst increased temperatures increase DOM priming by microbes. (B) Urban land use contributes to groundwater DOC through contamination, for example through leaking septic and sewer systems. (C) Water quality parameters and groundwater DOC concentrations are linked and are largely controlled by redox conditions (NB: C shows variables where DOC is the driver for the observed changes in water chemistry). (D) Increasing aquifer age results in a decline in groundwater DOC due to sediment lithification and a depletion of sedimentary organic matter over time. Increasing groundwater residence times lead to decreasing DOC

166 by a combination of filtration of DOC through smaller aquifer pore sizes and adsorption
167 (where residence times correspond to longer flow paths), and increased exposure to
168 biodegradation over time.
169
170 **Climate controls**
171 Temperature and precipitation play an important role in predicting groundwater DOC
172 concentrations. Overall, DOC decreases by 9.5 ± 1.1% for every 10 mm increase in
173 precipitation in the driest month of the year and decreases by 2.5 ± 0.8% for every 10 mm
174 increase in precipitation in the wettest month of the year. This is likely due to a dilution effect
175 whereby accumulated soil DOM infiltrates the aquifer during initial rainfall and is later diluted
176 by additional rainfall [34]. In arid climates, some of these trends may be reversed (Fig. 3). For
177 example, a decrease in aridity represented by decreased temperature and increased
178 precipitation, would increase groundwater DOC concentration since the precipitation in the
179 wettest month of the year is not enough to cause significant dilution. Furthermore, in arid
180 climates, groundwater DOC concentrations would be low due to the high temperatures and
181 low rainfall leading to limited vegetation cover and bioavailable DOM [35]. We observed these
182 trend reversals in linear analyses for the smaller Australian dataset (n = 79 after removing
183 incomplete data).
184
185 The model shows an overall groundwater DOC concentration increase by 3.4 ± 0.3% for
186 every 1 °C increase in average air temperatures in the wettest quarter of the year and 4.6 ±
187 0.5% for every 1 °C increase in sample groundwater temperature. In contrast, groundwater
188 DOC concentrations decrease by 8.9 ± 1.1% for every 1 °C increase in temperatures in the
189 warmest quarter of the year. The source of DOC is dependent upon availability of water. In
190 humid climates, increases in mean surface temperature in the wettest quarter of the year
191 and increased groundwater temperatures are likely to cause increased temperatures in the
192 soil zone. Under conditions of increased soil moisture, warm temperatures can stimulate
193 biological activity, DOM priming [36], and an increase in groundwater DOC.

194

195 **Water chemistry**

196 In the saturated zone, redox conditions and pH are strongly related to DOC concentration,
197 with DOC concentrations 9.2 ± 2.4% lower for each unit increase in pH, and 6.8 ± 0.6%
198 lower with every 1 mg L^{-1} increase in DO. We also observe a 4.5 ± 0.4% increase in DOC
199 associated with a 10 mg L^{-1} increase in Ca. The smaller Australian dataset (n = 79) also was
200 consistent with the larger U.S. dataset (n = 2,916) (see Supplementary Methods). The
201 mineralization of DOC via biodegradation consumes DO, produces CO_2 and hence carbonic
202 acid, which causes a decrease in pH and related calcite dissolution and dissolved Ca
203 production. Once conditions become anoxic and biodegradation rates are reduce, pH levels
204 increase. The relationship between DOC and Ca, as well as microbial respiration by-
205 products such as ammonium has been observed in regional-scale studies [3].
206
207 We also show that reduced dissolved species of Fe(II) and Mn(II) are positively correlated
208 with DOC concentrations. This trend may be explained by microbial use of Mn- and Fe-
209 oxides as alternative electron acceptors to DO in the presence of DOC under anoxic
210 conditions [37]. Aerobic microbes metabolize carbon at a faster rate than anaerobic microbes,
211 therefore a lack of DO limits DOC biodegradation rates. Fe can accumulate within oxic
212 sediment layers due to oxidation and precipitation of dissolved Fe in young sediments [38].
213 This Fe can become coated with OM and re-dissolve under reduced conditions releasing the
214 OM, which can increase Fe with increased DOC. Decreased sulphate (SO_4) and chloride
215 (Cl) deposition due to recent emission regulations cause increases in DOC concentrations in
216 surface waters [39] which could also lead to increased DOC concentrations in shallow
217 groundwaters.
218

219 **Aquifer age and groundwater evolution**

220 The age of the geological formation, or aquifer age, explained 16.3% of variability in
221 groundwater DOC. Groundwater in younger aquifers of Cenozoic sediments contained 41%

higher DOC concentrations than older Mesozoic and Paleozoic Era aquifers which support previous observations in smaller datasets [4]. Despite site specific observations of high groundwater DOC associated with older aquifers [40], the model suggests sedimentary OM in young aquifers is more likely to be mobilized than in older, lithified aquifers. Other studies have also reported higher groundwater DOC concentrations originating from the matrix of younger aquifers [41, 42].

We also observed a decrease in groundwater DOC concentrations of 7.7 ± 0.6% for every 10 m increase in sample depth. As DOC moves through porous media it undergoes filtration and oxidation to DIC. Consolidated sediment pore-throat sizes can occur in sizes much smaller than DOC, which is defined as the fraction of total organic carbon being able to pass through a membrane with pores typically between 0.2 to 0.7 μm. For example some pore-throat sizes in Permo-Triassic sandstones have been determined to be as low as 0.01 μm [43]. Additionally, deeper groundwaters often have longer residence times [44]. This is implied by a positive relationship between the mineral weathering product Si and sample depth ($p < 2 \times 10^{-06}$). We found a negative correlation between groundwater DOC and Si, with DOC decreasing by 6.3 ± 1.3% with every 10 mg L^{-1} increase in Si. This relationship has also been observed in surface waters in the U.S. [45]. The main source of dissolved Si is silicate mineral dissolution [37]. The negative relationship between DOC and Si is explained by the dissolved solids accumulated due to water-rock interaction in older groundwaters [46]. In surface waters, lakes and streams with short water residence times have also been shown to be biogeochemical hotspots where DOC is rapidly produced and consumed [47]. These deeper and older groundwaters are more likely to be depleted in DOC due to oxidation processes, biodegradation and adsorption to soil and aquifer mineral surfaces [48].

Land use

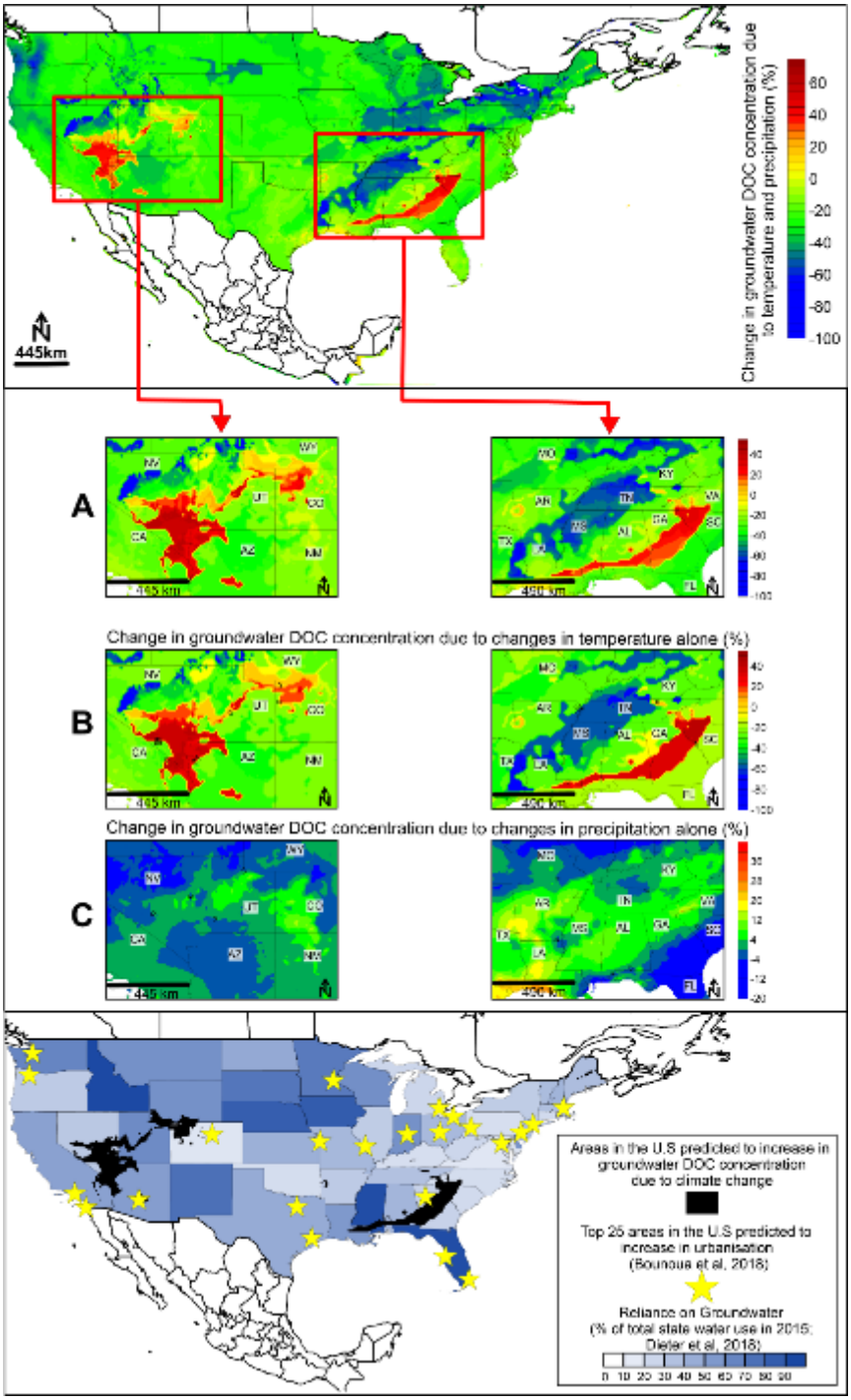
There is a significant increase of 19% in groundwater DOC concentrations in urban areas compared to natural land. Urban land use contributes DOC to groundwater through leaking

sewage systems, landfill leaching, animal waste, fertilizer run-off and industrial and residential waste washed into stormwater drains [27, 49, 50, 51]. Within natural and agricultural areas there is a significant decrease in median groundwater DOC concentrations where the subsoil clay fraction is > 30% (n = 2,127) compared to subsoils with clay fraction ≤ 30% (n = 2,372, Fig. S4). In contrast, low (< 1%, n = 4,382) and high (> 1%, n = 106) soil organic carbon content within natural and agricultural areas do not appear to influence groundwater DOC concentrations (p = 0.472, Fig. S4) suggesting that adsorption in interlamellar spaces or complexation with Fe[52] may play a more important role in determining groundwater DOC concentrations than overlying soil organic carbon content. Our global dataset supports local scale observations showing that urban land use increases DOC in surface and groundwaters [53, 54, 55] and shows that the impact of urban land use also extends to groundwater systems on a broader scale. We found no significant difference (p = 0.841) in groundwater DOC concentrations between natural and agricultural areas. It is noted that the finding of increased DOC in urban areas from our space-for-time analysis, as well as a previous space-for-time analysis [55], cannot be used to infer how this increase has occurred over time. A search for available groundwater TOC and DOC timeseries data in urban areas produced two datasets from Florida, U.S., and Perth, Australia. These data, collected from the 1980's to present in largely-residential urban areas, show no clear trend in groundwater DOC concentration (Fig. S5) over this timescale. Longer groundwater DOC timeseries could help elucidate processes occurring. For example, fluvial DOC concentrations in the Thames Basin since 1883 [56] have attributed increases in DOC concentrations since World War 2 to sewerage inputs and changes in land management. We suggest that further groundwater DOC research is required to investigate the impact of urban area expansion, for example into lowland regions where DOC might be high, mobilization of previously stable soil DOC as a result of development, and legacy contamination of groundwater in urban areas.

Implications

277 Continental-scale changes to groundwater DOC concentrations respond to changing
278 temperature and precipitation patterns. Using our results and $IPPC_5$ ($CMIP_5$) climate
279 projections (www.worldclim.org), we identify more extreme groundwater DOC concentration
280 changes associated with changing temperatures modulated by changing precipitation rates
281 and patterns (Fig. 4). We identify hotspots of high groundwater DOC concentration
282 (increases of up to 45%) associated largely with increased temperatures in the wettest
283 quarter of the year in a number of south eastern U.S. states under the “business-as-usual”
284 Intergovernmental Panel on Climate Change (IPCC) climate change prediction scenario
285 RCP8.5 (Fig. 4). Increasing temperatures stimulate phenol oxidase activity [57], which
286 increases surface water DOC by 5.4% per year in the United Kingdom [58]. Importantly,
287 relatively recalcitrant phenolic compounds [59] are selectively released as a result of this
288 process. Therefore, under warmer conditions, more DOC may persist along a flow path and
289 ultimately enter groundwater systems. This increased carbon loading in groundwaters can
290 change redox conditions and terminal electron acceptor availability for microorganism use.
291 This could drive changes in groundwater microbial communities [60] and potentially enhance
292 survival rates for microbial communities in groundwater systems.
293
294 The areas most at risk of future groundwater DOC concentration increases are those facing
295 increased DOC due to climate change and where urbanization is predicted to occur. It is
296 likely that DOC leaching will increase due to urbanization and population pressure on waste
297 disposal networks. This will be a particularly significant issue for regions which have a large
298 or increasing reliance on groundwater as a source of fresh water. For example areas in the
299 U.S. predicted to be impacted by DOC increases associated with climate have between 13%
300 to 87% reliance on groundwater as their source of fresh water (Fig. 4). Some of those
301 regions were also projected to experience large increases in urbanization by 2020 [61]. Our
302 analysis suggest that this could lead to increased groundwater DOC concentrations,
303 however it is noted that these results are based on a space-for-time analysis.
304

On a global scale, 54% of the worlds' population live in urban areas. By 2050 the world's urban population will increase to 66% [62]. By 2030, urban area is predicted to increase by 1.2 million km^2 [63]. This presents a significant issue when combined with the current lack of adequate sanitation services [64] which result in the pollution of groundwater systems. Areas likely to see urban expansion and population growth include eastern China, India and parts of Africa [63]. These areas are already facing high urban growth, high rates of the population without even basic sanitation levels (25% and 56% and up to 64% respectively) [64], and severe groundwater contamination issues that threaten local livelihoods [65, 66]. Groundwater quality issues in south-eastern China may be further compounded by groundwater DOC increases associated with large predicted increases in temperature (up to 10 °C) in the wettest quarter of the year by 2050 [30].



317

318 **Fig. 4.** Changes in groundwater DOC concentrations by year 2050 in the U.S. due to
319 temperature and precipitation changes. Panel A shows the two areas in the U.S predicted to
320 experience the largest increases in DOC concentration due to temperature and precipitation
321 changes by 2050. Panels B and C show changes in DOC concentrations in these areas
322 caused by temperature variables (temperature in the wettest and warmest quarters of the
323 year) and precipitation variables (precipitation in the wettest and driest months of the year)

alone. Groundwater DOC concentrations changes are calculated using model results and IPPC$_5$ (CMIP$_5$) climate projections from the end of the 20th century (average of values from the period 1960 – 1990) to year 2050 (average of predicted values for the period 2041 – 2060) [30] for a "business-as-usual" climate change scenario (RCP8.5) as outlined in IPCC [67]. Lowermost map shows U.S. state reliance on groundwater as for total water use [68] overlain with areas predicted to experience increases in groundwater DOC concentrations due to climate change variables and urbanization.

In some locations, increased groundwater DOC concentrations associated with climate change and urban land cover will lead to an increase in the scale and hence capital and operational costs of water treatment facilities for groundwater DOC removal. Besides the direct implications of increased DOC, groundwater organic matter degradation in Holocene and Pleistocene aquifers in countries including Vietnam and Bangladesh have also been shown to correlate with an increase in Fe, NH_4 and As due to reductive dissolution of Fe-oxides in sediments where As is associated with Fe-oxides [3, 69]. An increase in these dissolved species reduces groundwater quality and affects human health.

One common method of DOC removal from drinking water is via adsorption onto granular activated carbon. Increasing groundwater DOC concentrations in certain locations will increase the need for water treatment facilities to implement granular activated carbon as a second stage filter for DOC removal. The current cost of water for a family of four is approximately US$845 per year [70]. The implementation of granular activated carbon filtration methods by a 6.6 mega gallon per day facility (approximately 25 mega liter per day) would increase monthly water costs for a family of four by US$134 per year. This equates to a 16% increase in annual household water costs in areas of Nevada, Georgia and South Carolina.

Overall, our investigation reveals that changes in climate and urban land cover are likely to impact groundwater DOC concentrations globally. These impacts on groundwater DOC will

not be evenly distributed. Increases in temperatures in the warmest quarter of the year will decrease groundwater DOC concentrations due to aridity, whilst increased temperatures in the wettest quarter of the year will increase groundwater DOC concentrations due to the stimulation of microbial activity. We identify hotspots of high groundwater DOC concentrations in areas that will undergo future urbanization and population growth. This could substantially increase the treatment costs to DOC from groundwater in many locations. Our results indicate that climate change and urban land cover will not only impact the quantity of the groundwater resource [23], but can also decrease groundwater quality and increase water treatment costs.

Methods

Literature survey and spatial coverage

Google Scholar, Scopus as well as public data sources were searched using terms “DOM”, “DOC”, “dissolved organic carbon”, “dissolved organic matter”, “groundwater quality” for datasets presenting original (i.e. non-summarized) values of DOC. A number of authors, government departments and colleagues also provided original published and unpublished datasets. A total of 7,849 unique groundwater DOC observations were obtained from published and unpublished datasets (Table S1) after eliminating samples with a negative concentration value (n = 36), or those flagged with a “V” in the National Water Quality Assessment (NWQA) data (n = 461). The authors were advised that samples flagged with a “V” were suspected of being contaminated by methanol used in cleaning of the sampling apparatus and therefore they were excluded from the analysis. The data represents DOC concentrations for 31 countries on six continents with samples measured between 1992 and 2018. The authors excluded datasets reported to have been sampled from aquifers known to be heavily contaminated. A large proportion of the data come from samples obtained in U.S. (n = 5,704), followed by Australia (n = 780), Scotland (n = 270), England and Wales (n = 113), Zambia (n = 110) and Czech Republic (n = 104), with a lesser proportion of samples obtained from Malawi (n = 89), India (n = 79), Uganda (n = 71), Canada (n = 52), Ethiopia (n

= 44), Nepal (n = 40), Poland (n = 40), Kenya (n = 36), Nigeria (n = 35), China (n = 34), Brazil (n = 30), Portugal (n = 28), Iceland (n = 24), Senegal (n = 22), Denmark (n = 20), Estonia (n = 19), Belgium (n = 18), Cook Island (n = 17), Argentina (n = 15), Bangladesh (n = 13), Mali (n = 12), Spain (n = 10), Malta (n = 8), France (n = 7) and Algeria (n = 5). We used these data to determine average DOC concentrations globally. 5,459 samples from the U.S. dataset which represent the data collected by the National Water Quality Assessment (NAWQA) program of the U.S. Geological Survey (USGS) [29], were then used to generate the linear mixed model. This dataset was selected due to its large number of samples, ancillary data (inorganic water quality parameters) and availability of coordinates which allowed for the extraction of climatic, land use and unsaturated zone thickness data.

Global groundwater DOC distributions

DOC concentrations for each country were collected from the sources shown in Table S1. Non-parametric unpaired one tailed Wilcoxon tests were used to identify differences in the groundwater DOC concentrations between continents, aquifer types and recharge rates using the function wilcox.test() in RStudio.

Worldclim climate data

High resolution (30 arc second) global ESRI grids were obtained for bioclimatic variables from www.worldclim.org (v1.4) [30]. Bioclimatic variables including annual mean temperature and precipitation, mean temperature and precipitation of the driest and wettest quarters, mean temperature and precipitation of the warmest and coldest quarters, as well as annual temperature range, isothermality, temperature and precipitation seasonality were extracted to each sample location using ArcGIS (v10.4.1). Where Worldclim data output showed that data was unavailable, as indicated by a -999.9 or -9999 value, these were removed and left blank.

Land use data

Land use data was obtained for the U.S. dataset using a shapefile developed by the University of Maryland, Department of Geography and NASA [32, 33]. Land uses were assigned to each sample location coordinate using ArcMAP (v10.4.1). Nineteen land use classifications are used in the file. Land use classifications were reassigned to “agricultural” (n = 3047) where the land use type included the word “cropland”. Areas were assigned as “urban” (n = 956) for any area listed as “urban/built up”. Areas were assigned as “wilderness” (n = 257) for any samples containing the keywords “forest”, “shrublands”, “wetlands”, “marsh”, “water” and “savannahs”.

Unsaturated zone thickness data

Unsaturated zone thickness data [31] and was downloaded through GLOWASIS, the European Union collaborative project of Global Water Scarcity Information Service, at https://glowasis.deltares.nl/thredds/catalog/opendap/opendap/Equilibrium_Water_Table/catalog.html. Data was extracted to each sample location using ArcGIS (v10.4.1).

Model development and statistics for U.S. dataset

A mixed linear model was developed using climatic data, land use data and unsaturated zone thickness data as well as parameters available in the U.S dataset available at https://doi.org/10.1594/PANGAEA.896953. Parameters included DOC, dissolved oxygen (DO), dissolved iron (Fe), sulphate (SO_4), magnesium (Mg), manganese (Mn), calcium (Ca), potassium (K), silica (Si), sodium (Na), fluoride (F), chloride (Cl), pH, sample temperature, sample depth below land surface, depth to the water table, land elevation, northing, precipitation in the wettest, driest, coldest and warmest quarters, maximum temperature in the warmest month, minimum temperature in the coldest month, temperature in the wettest, driest, coldest and warmest quarters of the year, mean diurnal temperature range, temperature seasonality, annual temperature range, annual precipitation, precipitation seasonality, annual average temperature, precipitation in the driest month, aquifer age, and land use type. Aquifer age was selected as a random effect, with all other variables applied

as fixed effects which were selected using the manual Akaike Information Criteria (AIC) based backward selection using the drop1() function in RStudio. This function allows for the identification of the variable with the lowest AICs so that they can be removed from the model.

Quality assurance procedures

Prior to data analysis, the data set was screened for "<X" values, which indicate a limit of detection in the analysis. Where these were identified, the value was replaced with a randomized value between 0 and X to ensure that bias associated with assigning these data points as either 0 or 1/2X is eliminated. A number of data points were flagged as potentially being contaminated with methanol and these samples were removed from the dataset.

Accuracy of sample coordinates were checked by adding an XY map of sample coordinates to a world map using Golden Software Surfer® (v 13.6.618). Any samples that were not located in the correct area as indicated by their ID label was investigated for typological error in the assigned coordinate and corrected (n = 3).

Prior to model generation, the response variable "DOC" was log transformed to normalize the data. Any sample with a missing value for one of the variables was removed using the na.omit function in R (v 3.3.1). This resulted in a final n of 2,916 complete sample points used in the model. Predictor variables were then individually fitted to a simple linear model with DOC as the response variable to check assumptions. Standardized residuals versus fitted value plots (Fig. S6), Q-Q plots (Fig. S7) and boxplots of residuals (Fig. S8) were examined for each quantitative model variable to check that the assumption of constant variance and normality held true for residuals. Collinearity was checked through a regression matrix using R (v 3.3.1), which confirmed the presence of multicollinearity between some variables. The variance inflation factors (VIFs) for each variable were also checked. Typically, a variable is considered collinear with another variable when the VIF is greater

than 10 [71]. Some literature however recommends removing variables with VIFs greater than 4 or 5 [72]. Some variables were known to covary, and thus were removed from the list of variables. These including Mn (due to covariance with Fe), EC (due to covariance with ions), temperature in the coldest and warmest months (covariance with temperatures in the coldest and warmest quarters), temperature and precipitation seasonality (due to covariance with temperature and precipitation range). All remaining variables with VIFs greater than 4 were then removed, with the variable with the highest VIF removed first before re-running the code each time until all remaining variables had VIFs > 10. The variables with VIFs > 10 were removed in the following order; annual temperature, temperature in the coldest quarter, precipitation in the driest quarter, annual precipitation, isothermality, precipitation in the wettest quarter, precipitation in the warmest quarter, precipitation in the coldest quarter, chloride, temperature in the driest quarter, northing, mean diurnal range. The results are reported as a percent change in DOC concentration, with standard error reported where possible (i.e. for continuous fixed effects variables).

Change in groundwater DOC concentration due to climate change in 2050

Contours for changes in groundwater DOC concentration due to climate change in 2050 were developed by using current climate grid files from Worldclim v1.4 [30] and future IPPC$_5$ (CMIP$_5$) climate projections (www.worldclim.org). Current climate values for climate variables used in the model (temperature in the wettest quarter of the year, temperature in the warmest quarter of the year, precipitation in the wettest month of the year and precipitation in the driest month of the year) were subtracted from future IPPC$_5$ (CMIP$_5$) climate projection values (www.worldclim.org) for the same variables for the year 2050 under a “business-as-usual” representative concentration pathway (RCP8.5) using Surfer® v.11.0.642. As DOC units in the model are log natural concentration values, the difference from zero for the exponents of the intercepts for the four variables were multiplied by the difference in current and future values using a positive or negative sign at the front of the exponent depending upon whether the variable was positively or negatively correlated with

DOC concentration. This provides a percent change in groundwater DOC concentration due to change in each climate variable predicted for 2050. The four grid files containing change in DOC concentration associated with each of the four variables were then added together into a single grid file to get a total change in DOC (%) associated with climate change in 2050.

Acknowledgments: The authors would like to thank various ANSTO personnel such as Robert Chisari, Alan Williams, Barbara Gallagher, Kelly Farrawell, Chris Dimovski and Jennifer Van Holst for sample preparation and analysis for a portion of the Australian data. The authors thank Wuhui Duan from UNSW Sydney for his assistance in the collection and translation of data published in Chinese. We also thank graphic designer Anna Blacka from UNSW Sydney's Water Research Laboratory for her assistance with the preparation of Fig. 2 and Professor Richard Kingsford at UNSW Sydney for his comments on the manuscript. Thanks to statisticians Peter Geelan-Small, Eve Slavich and Nancy Briggs from UNSW Stats Central for their advice and guidance in the model development and interpretation.

Funding: This work was supported by the Australian Research Council [DP160101379], the Australian Government Research Training Program Scholarship a scholarship provided by ANSTO. A large portion of the Australian DOC data was also made possible by grants from the Federal Government initiative National Collaborative Research Infrastructure Strategy (NCRIS), the Cotton Catchment Community (CRC-CCC), the NSW Department of Primary Industries Office of Water and support for students via the National Centre for Groundwater Research and Training (NCGRT). D.J.L and D.C.G were funded through the LOCATE NERC grant NE/N018087/1. Ethiopia data used in the global dataset comparison was funded by Hidden Crisis 2 NE/M008606/1. Iceland data was funded through BGS-NERC Earth Hazards and Systems Directorate. Kenya data collection was funded through funded through HyCRISTAL NE/M020452/1. Mali and Nigeria data collection were funded through UK

519 Department for International Development, GA/09F/094, Groundwater resilience to climate
520 change in Africa. Malawi data collection was funded through TRIGR, REACH programme
521 and Hidden Crisis 2 NE/M008606/1. Nepal data collection was funded through UK
522 Department for International Development, Groundwater Resources in the Indo-Gangetic
523 Basin, Grant 202125-108. Senegal data collection was funded through AfriWatSan, Royal
524 Society DFID Africa Capacity Building Initiative, ref: AQ140023. Uganda data collection was
525 funded through HyCRISTAL NE/M020452/1 and Hidden Crisis 2 NE/M008606/1. Brazil,
526 Argentina and a number of Australian data used in the global analysis were collected with
527 funding from the Australian Research Council (FT170100327).

528

529 Author contributions: L.K.M. drafted the main manuscript and was responsible for model
530 development and figure preparation. A.B, M.A., D.M.O., H.R. and K.M. supervised the
531 project and provided manuscript comments and guidance. I.R.S provided unpublished data,
532 guidance on the direction of the paper and manuscript comments. P.O. assisted with
533 sourcing of data and provided advice on DOC sources and sinks. J.B. provided water
534 treatment costs associated with groundwater DOC removal and provided calculations on
535 increases in household water costs associated with future temperature and precipitation
536 changes. D.C.G, J.P.R.S, D.J.L., A.M.M. and J.W. provided published and unpublished
537 groundwater datasets and manuscript comments.

538

539 Data and materials availability: The U.S. dataset analyzed during the study is available on
540 data repository PANGAEA (https://doi.pangaea.de/10.1594/PANGAEA.896953). Sources of
541 published datasets used in the global groundwater DOC analysis are available in the
542 Extended Data - Tables supporting document. Unpublished datasets that support the
543 findings of the global groundwater DOC analysis will be provided by the corresponding
544 author upon reasonable request and with permission of the owner(s) of the data.

Code availability: Code will be provided upon request by the authors.

References

1. Judd KE, Crump BC, Kling GW. Variation in dissolved organic matter controls bacterial production and community composition. **87**, 2068-2079 (2006).

2. Romera-Castillo C, Pinto M, Langer TM, Álvarez-Salgado XA, Herndl GJ. Dissolved organic carbon leaching from plastics stimulates microbial activity in the ocean. *Nature Communications* **9**, 1430 (2018).

3. Harvey CF*, et al.* Arsenic Mobility and Groundwater Extraction in Bangladesh. *Science* **298**, 1602 (2002).

4. Jakobsen R, Postma D. In situ rates of sulfate reduction in an aquifer (Rømø, Denmark) and implications for the reactivity of organic matter. *Geology* **22**, 1101-1106 (1994).

5. Richards LA*, et al.* Dissolved organic matter tracers reveal contrasting characteristics across high arsenic aquifers in Cambodia: A fluorescence spectroscopy study. *Geoscience Frontiers* **10**, 1653-1667 (2019).

6. Richards LA*, et al.* Dual in-aquifer and near surface processes drive arsenic mobilization in Cambodian groundwaters. *Science of The Total Environment* **659**, 699-714 (2019).

7. Evans S, Campbell C, Naidenko OV. Cumulative risk analysis of carcinogenic contaminants in United States drinking water. *Heliyon* **5**, e02314 (2019).

8. Sadiq R, Rodriguez MJ. Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review. *Science of The Total Environment* **321**, 21-46 (2004).

9. Lawson M, Polya DA, Boyce AJ, Bryant C, Ballentine CJ. Tracing organic matter composition and distribution and its role on arsenic release in shallow Cambodian groundwaters. *Geochimica et Cosmochimica Acta* **178**, 160-177 (2016).

10. Fleck JA, Bossio DA, Fujii R. Dissolved Organic Carbon and Disinfection By-Product Precursor Release from Managed Peat Soils. **33**, 465-475 (2004).

11. Bauer M, Blodau C. Mobilization of arsenic by dissolved organic matter from iron oxides, soils and sediments. *Science of The Total Environment* **354**, 179-190 (2006).

12. WHO. Guidelines for drinking-water quality. Fourth Edition.) (2017).

13. NHMRC. Australian Drinking Water Guidelines 6.). Australian Government (2018).

14. US EPA. 40 CFR Parts 9, 141, and 142 National Primary Drinking Water Regulations: Stage 2 Disinfectants and Disinfection Byproducts Rule; Final Rule) (2006).

15. Health Canada. Guidance on natural organic matter in drinking water. Document for Public Consultation. (2019).

16. Government of France. Order of 11 January 2007 relating to the limits and references for the quality of raw water and water intended for human consumption mentioned in Articles R. 1321-2, R. 1321-3, R. 1321-7 and R. 1321-38 of Public Health Code. *Journal Officiel de la République Francaise* **31**, 2180 (2007).

17. Ministry of Health (China). National Standard of the People's Republic of China: Standards for Drinking Water Quality.). Minitstry of Health of the People's Republic of China (2006).

18. DWAF. South African water quality guidelines. Volume 1: Domestic Water Use, Second Edition.) (1996).

19. von Gunten U. Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Research* **37**, 1469-1487 (2003).

20. Nemani RR*, et al.* Climate-Driven Increases in Global Terrestrial Net Primary Production from 1982 to 1999. *Science* **300**, 1560 (2003).

21. Davidson EA, Janssens IA. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* **440**, 165-173 (2006).

22. Jenkinson DS, Adams DE, Wild A. Model estimates of CO2 emissions from soil in response to global warming. *Nature* **351**, 304-306 (1991).

23. Lipczynska-Kochany E. Effect of climate change on humic substances and associated impacts on the quality of surface water and groundwater: A review. *Science of The Total Environment* **640-641**, 1548-1565 (2018).

24. Foster SSD, Chilton PJ. Groundwater: the processes and global significance of aquifer degradation. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **358**, 1957-1972 (2003).

25. Taylor RG*, et al.* Ground water and climate change. *Nature Climate Change* **3**, 322 (2012).

26. Rodell M*, et al.* Emerging trends in global freshwater availability. *Nature* **557**, 651-659 (2018).

27. Lapworth DJ*, et al.* Urban groundwater quality in sub-Saharan Africa: current status and implications for water security and public health. *Hydrogeology Journal* **25**, 1093-1116 (2017).

28. Richts A, Struckmeier WF, Zaepke M. WHYMAP and the Groundwater Resources Map of the World 1:25,000,000. In: *Sustaining Groundwater Resources: A Critical Element in the Global Water Crisis* (ed Jones JAA). Springer Netherlands (2011).

29. NWQMC. Water Quality Data.). National Water Quality Monitoring Council (2017).

30. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. **25**, 1965-1978 (2005).

31. Fan Y, Li H, Miguez-Macho G. Global Patterns of Groundwater Table Depth. *Science* **339**, 940 (2013).

32. Friedl MA*, et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment* **114**, 168-182 (2010).

33. Channan S, Collins K, Emanuel WJUoM, the Pacific Northwest National Laboratory CP, Maryland, USA. Global mosaics of the standard MODIS land cover type data. **30**, (2014).

34. Smith TE, Kolka RK, Zhou X, Helmers MJ, Cruse RM, Tomer MD. Effects of native perennial vegetation buffer strips on dissolved organic carbon in surface runoff from an agricultural landscape. *Biogeochemistry* **120**, 121-132 (2014).

35. Chapelle FH, Bradley PM, Journey CA, McMahon PB. Assessing the Relative Bioavailability of DOC in Regional Groundwater Systems. *Groundwater* **51**, 363-372 (2013).

36. Grieve IC. Seasonal, hydrological, and land management factors controlling dissolved organic carbon concentrations in the loch fleet catchments, Southwest Scotland. *Hydrological Processes* **4**, 231-239 (1990).

37. Appelo C, Postma D. *Geochemistry, Groundwater and Pollution*, 2nd edn. Balkema (2005).

38. Schwertmann U. Inhibitory Effect of Soil Organic Matter on the Crystallization of Amorphous Ferric Hydroxide. *Nature* **212**, 645-646 (1966).

39. Monteith DT*, et al.* Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. *Nature* **450**, 537-540 (2007).

40. MacDonald AM, O Dochartaigh BE, Smedley PL. Baseline groundwater chemistry in Scotland's aquifers.). British Geological Survey (2017).

41. Rivett MO, Smith JWN, Buss SR, Morgan P. Nitrate occurrence and attenuation in the major aquifers of England and Wales. *Quarterly Journal of Engineering Geology and Hydrogeology* **40**, 335-352 (2007).

42. Stuart ME, Lapworth DJ. Macronutrient status of UK groundwater: Nitrogen, phosphorus and organic carbon. *Science of The Total Environment* **572**, 1543-1560 (2016).

43. Bloomfield J, Gooddy D, Bright M, Williams P. Pore-throat size distributions in Permo-Triassic sandstones from the United Kingdom and some implications for contaminant hydrogeology. *Hydrogeology Journal* **9**, 219-230 (2001).

44. Weissmann GS, Zhang Y, LaBolle EM, Fogg GE. Dispersion of groundwater age in an alluvial aquifer system. **38**, 16-11-16-13 (2002).

45. Aulenbach BT*, et al.* Approaches to stream solute load estimation for solutes with varying dynamics from five diverse small watershed. *Ecosphere* **7**, (2016).

46. Peters NE, Burns DA, Aulenbach BT. Evaluation of High-Frequency Mean Streamwater Transit-Time Estimates Using Groundwater Age and Dissolved Silica Concentrations in a Small Forested Watershed. *Aquatic Geochemistry* **20**, 183-202 (2014).

47. Evans CD, Futter MN, Moldan F, Valinia S, Frogbrook Z, Kothawala DN. Variability in organic carbon reactivity across lake residence time and trophic gradients. *Nature Geoscience* **10**, 832 (2017).

48. Shen Y-H. Sorption of natural dissolved organic matter on soil. *Chemosphere* **38**, 1505-1515 (1999).

49. Utton AE. The Development of International Groundwater Law. *Natural Resources Journal* **22**, 95-118 (1982).

50. Foley JA*, et al.* Global Consequences of Land Use. *Science* **309**, 570 (2005).

51. Lapworth DJ, Gooddy DC, Butcher AS, Morris BL. Tracing groundwater flow and sources of organic carbon in sandstone aquifers using fluorescence properties of dissolved organic matter (DOM). *Applied Geochemistry* **23**, 3384-3390 (2008).

52. Kodama H, Schnitzer M. Evidence for interlamellar adsorption of organic matter by clay in a podzol soil. *Canadian Journal of Soil Science* **51**, 509-512 (1971).

53. Johnson L, Richards C, Host G, Arthur J. Landscape influences on water chemistry in Midwestern stream ecosystems. **37**, 193-208 (1997).

54. Meierdiercks KL, Kolozsvary MB, Rhoads KP, Golden M, McCloskey NF. The role of land surface versus drainage network characteristics in controlling water quality and quantity in a small urban watershed. **31**, 4384-4397 (2017).

55. Lawrence A, Gooddy D, Kanatharana P, Meesilp W, Ramnarong VJHJ. Groundwater evolution beneath Hat Yai, a rapidly developing city in Thailand. **8**, 564-575 (2000).

56. Noacco V, Wagener T, Worrall F, Burt TP, Howden NJK. Human impact on long-term organic carbon export to rivers. *J Geophys Res Biogeosci* **122**, 947-965 (2017).

57. Freeman C, Ostle N, Kang H. An enzymic 'latch' on a global carbon store. *Nature* **409**, 149-149 (2001).

58. Freeman C, Evans CD, Monteith DT, Reynolds B, Fenner N. Export of organic carbon from peat soils. *Nature* **412**, 785-785 (2001).

59. Wetzel RGJH. Gradient-dominated ecosystems: sources and regulatory functions of dissolved organic matter in freshwater ecosystems. **229**, 181-198 (1992).

60. Danczak RE, Johnston MD, Kenah C, Slattery M, Wilkins MJ. Microbial Community Cohesion Mediates Community Turnover in Unperturbed Aquifers. *mSystems* **3**, e00066-00018 (2018).

61. Bounoua L, Nigro J, Thome K, Zhang P, Fathi N, Lachir A. A Method for Mapping Future Urbanization in the United States. *Urban Science* **2**, (2018).

62. UN. *World urbanization prospects The 2014 revision - highlights* (2014).

63. Seto KC, Güneralp B, Hutyra LR. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences* **109**, 16083 (2012).

64. WHO, UNICEF. Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines.). World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) (2017).

65. Rodríguez-Lado L*, et al.* Groundwater Arsenic Contamination Throughout China. *Science* **341**, 866 (2013).

66. Sahu S, Saha D. Role of shallow alluvial stratigraphy and Holocene geomorphology on groundwater arsenic contamination in the Middle Ganga Plain, India. *Environmental Earth Sciences* **73**, 3523-3536 (2015).

67. IPCC. Climate Change 2014: Synthesis Report Contribution of Working Groups I,II and III to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change. (eds Pachauri RK, Meyer LA). IPCC (2014).

68. Dieter CA*, et al.* Estimated use of water in the United States county-level data for 2015 (ver. 2.0, June 2018): U.S. Geological Survey data release.). U.S. Geological Survey (2018).

69. Postma D*, et al.* Arsenic in groundwater of the Red River floodplain, Vietnam: Controlling geochemical processes and reactive transport modeling. *Geochimica et Cosmochimica Acta* **71**, 5054-5071 (2007).

70. Walton B. Price of Water 2017: Four Percent Increase in 30 Large U.S. Cities.). Circle of Blue (2018).

71. Kleinbaum DG, Kupper LL. *Applied Regression Analysis and Other Multivariable Methods*. PWS-Kent Publishing Company (1988).

72. Rogerson PA. A Statistical Method for the Detection of Geographic Clustering. *Geographic Analysis* **33**, 215-227 (2001).