1 **Provenance of drinking water revealed through compliance sampling**

2

3

4 Matthew J. Ascott[a*], Marianne E. Stuart[a], Daren C. Gooddy[a], Ben M. Marchant[b], John C. Talbot[a], Ben

5 W.J Surridge[c] and David A. Polya[d]

6

7 [a] British Geological Survey, Maclean Building, Benson Lane, Crowmarsh, Oxfordshire, OX10 8BB, UK.

8 [b] British Geological Survey, Environmental Science Centre, Nicker Hill, Keyworth, Nottinghamshire,

9 NG12 5GG, UK.

10 [c] Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK

11 [d] School of Earth and Environmental Sciences and Williamson Research Centre for Molecular

12 Environmental Science, University of Manchester, Manchester, M13 9PL, UK

13

14

15

16 *Corresponding Author: matta@bgs.ac.uk, +44(0)1491 692408

17

18

19

20 **Table of Contents Entry**

21

22

23

24

25 Images copyright Peter Werkman (www.peterwerkman.nl) and US Department of Agriculture
26 (https://www.flickr.com/photos/usdagov/9404332949/in/photostream/). Images are licensed under
27 the Creative Commons Attribution 2.0 Generic Licence (CC BY 2.0, licence available at:
28 https://creativecommons.org/licenses/by/2.0/legalcode)

29

30

31

32

33

34

35

36

37

38

Abstract

Understanding drinking water hydrochemistry is essential for maintaining safe drinking water supplies. Whilst targeted research surveys have characterised drinking water hydrochemistry, vast compliance datasets are routinely collected but are not interrogated amidst concerns regarding the impact of mixed water sources, treatment, the distribution network and customer pipework. In this paper, we examine whether compliance samples retain hydrochemical signatures of their provenance. We first created and subsequently undertook the first hydrochemical analysis of a novel national database of publically available drinking water compliance analyses (n = 3,873,941) reported for 2015 across England and Wales. ~~Principal component analysis and~~ K-means cluster analysis revealed three spatially coherent clusters. Cluster 1 is dominated by groundwater sources, with high nitrate concentrations and mineralisation, and lower organic carbon, residual chlorine and THM formation. Cluster 2 was dominated by surface water sources and characterised by low mineralisation (low conductivity and major ion concentrations), low nitrate and high organic carbon concentrations (and hence residual chlorine and THM formation). Cluster 3 shows a mixture of groundwater overlain by confining layers and superficial deposits (resulting in higher trace metal concentrations and mineralisation) and surface water sources. These analyses demonstrate that, despite extensive processing of drinking water, at the national scale signatures of the provenance of drinking water remain. Analysis of compliance samples is therefore likely to be a helpful tool in the characterisation of processes that may affect drinking water chemistry. The methodology ~~presented~~ used is generic and can be applied in any area where drinking water chemistry samples are taken.

3

1   Introduction

Access to safe drinking water is a human right and a requirement for life [1]. In the developed world, the quality of water supplies has improved substantially in the past 25 years, largely through the introduction of regulation and advances in treatment[2]. In Europe, implementation of the European Union Drinking Water Directive (EUDWD, European Commission [3]) has resulted in compliance levels of over 99% in 2016[4]. Similar directives are also in place internationally (e.g. Australia[5], USA[6] and China[7]).

Against a backdrop of climate change and increased demand[8], water utilities are increasingly considering the use of raw and treated water transfers to supply customers[9]. Feasibility studies of local, small scale water transfers in the UK are required to establish the viability of a transfer in terms of environmental water resource availability and both drinking water and environmental water quality[10]. However, outside of the UK this is not always the case, as highlighted by the recent Flint Water Crisis[11]. In this case, the addition of highly corrosive surface water into a distribution system without corrosion control resulted in a significant public health incident[12]. Outside of the UK switching of supply water chemistry may be done without any systematic evaluation[13], and assessing the impacts of drinking water chemistry on potential future large scale raw and potable transfers is considered a significant research need[14].

The hydrochemical analyses required in order to support assessment of the water quality implications of transfers are complex. Changes in water quality associated with the mixing of raw water sources, treatment processes, passage through a utilities' distribution system and customer plumbing make unambiguous interpretation of drinking water chemistry data challenging[15]. Despite this, numerous studies have characterised drinking water hydrochemistry using specific sampling and laboratory analyses for research purposes[15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. A number of studies taking this approach have

87   shown a strong link between drinking water hydrochemistry and raw water sources. Dinelli, Lima [17]

88   and Demetriades [25] showed a clear influence of bedrock geology and aquifer composition on major

89   and trace elements in drinking waters in Italy and Greece respectively. Birke, Rauch [23] showed uranium

90   concentrations in drinking water to have a strong geological control. At the European scale, Banks,

91   Birke [21] and Flem, Reimann [15] showed that drinking water hydrochemistry can be interpreted in terms

92   of source water hydrogeology and land use, as these factors influence raw water chemistry. These

93   authors concluded that drinking water sampling is a highly cost-effective approach to characterise

94   controls on water chemistry at the European scale, with confident interpretation of numerous

95   parameters in terms of hydrogeochemical processes. Stable oxygen and hydrogen isotopes of drinking

96   water have also been shown to be a useful tracer of source waters and hydrological processes both at

97   the national [26, 27, 28] and city scale [29] in the USA and China. In the UK, national scale drinking water

98   trends broadly following the same spatial pattern as unconfined groundwaters [30].

99   There have been substantial reductions in funding for environmental regulators in recent years in

100   some developed countries [31, 32]. Consequently, environmental monitoring programmes have declined

101   [33]. In England and Wales the number of water chemistry measurements taken by the environmental

102   regulator has declined by 40% between 1993 and 2014 [34]. Environmental water chemistry monitoring

103   is typically devolved to a regional level which results in substantial spatial bias in sampling, as well as

104   both spatial and temporal variability in sampling methodologies, laboratory methods, standards,

105   reporting procedures and data quality assurance [35]. With a limited and reducing spatiotemporal extent

106   of environmental water chemistry monitoring, it is essential that other data sources are considered

107   for the characterisation of water chemistry required to assess the viability of raw and treated water

108   transfers. In addition to drinking water datasets collected specifically for research purposes, large

109   drinking water chemistry datasets have been and continue to be collected for regulatory compliance

110   across the developed world (e.g. Europe [4] and USA [36]). Under the EUDWD, around 100,000 water

111   supply zones are routinely sampled for regulatory compliance across Europe [3]. The need for data for

112   regulatory compliance results in consistent laboratory standards, extensive data quality assurance and

113   a large spatiotemporal sampling extent [3, 37]. These datasets have never been analysed in terms of their

114   hydrochemical characteristics and, potentially, represent a vast and powerful dataset that could

115   complement environmental water chemistry datasets and specific national [17, 25] and continental scale

116   drinking water research surveys [15, 21].

117

118   If water transfers are to be developed to meet future demand, it is essential that the hydrochemistry

119   of current the drinking water distribution is better understood. Moreover, beyond water quality

120   compliance reports, very little public information is available from water utilities on drinking water

121   sources and associated hydrochemistry. To this end, we examined whether drinking water samples

122   for regulatory compliance retain the hydrochemical signatures of their provenance? In this study we

123   present the first national-scale assessment of the hydrochemistry of drinking water based on

124   compliance sampling. Applied to England and Wales, we derived spatially distributed water chemistry

125   datasets based on published water company reports. We then undertook spatial and statistical

126   analyses to determine the likely factors controlling the spatial variation in drinking water chemistry.

127   Finally, we provide an outlook on the use of these datasets for future analysis of drinking water

128   hydrochemistry.

129

130

131   **2   Materials and Methods**

132   **2.1   Study area and regulatory context**

133

134   The countries of England and Wales were used as a study area for the research reported here (Figure

135   1). Drinking water supplies are obtained from both surface water and groundwater sources,

136   approximately in the ratio 60:40 overall [38], with raw water characteristics and treatment requirements

137   reflecting these different sources. Most water utilities supply water from both surface water and

138   groundwater sources, although in very different proportions depending on geographical location and

139   underlying geology. The most important aquifers used for water supply in the study area are the Chalk

140   and the Permo-Triassic rocks (referred to as Permo-Triassic or PT herein), are shown in Figure 1Figure 1Figure 1. At

141   one extreme in East Anglia, one utility draws drinking water supplies only from groundwater and

142   predominantly from the Chalk aquifer [39], whereas in Wales over 90% of water supplied is from surface

143   water sources [40].

144   As previously discussed, drinking water quality is regulated under the EUDWD. This is transposed into

145   UK law through primary legislation and regulations as the Water Supply (Water Quality) Regulations

146   [41]. Water is deemed to be wholesome if it does not contain substances which contravene the

147   concentrations listed in the Directive or National monitoring categories in Supplementary Table 1. A

148   further group of substances (indicator parameters) are also monitored and reported. Non-regulated

149   substances, such as calcium, magnesium and alkalinity, are measured less frequently and reporting of

150   results is not required.

151

2.2 Water quality sampling

153

154

155   The 27 individual water utilities in England and Wales undertake water quality compliance sampling

156   to meet the requirements of the EUDWD. Measurements are made either at the customer's tap, at a

157   supply point (SP) or at the water treatment works (WTW) exit as set down in the regulations and

158   agreed with the UK Drinking Water Inspectorate (DWI). Monitoring at WTW and service reservoirs

159   (SR) is to quantify levels of residual disinfectant, and control of microbiological parameters and nitrite.

160   Substances can be monitored at designated SPs instead of taps where concentrations are not deemed

161 to change in the distribution network. Supplementary Table 1 shows both compliance and indicator

162 parameters and location of sampling points. Guidance on the analysis of samples to ensure

163 consistency is provided by the DWI, for a full range of aspects including analyst training, suitable

164 equipment and calibration, method specification, internal and external analytical quality control and

165 record retention [37]. Pesticides and microbiological parameters are not considered in this assessment.

166
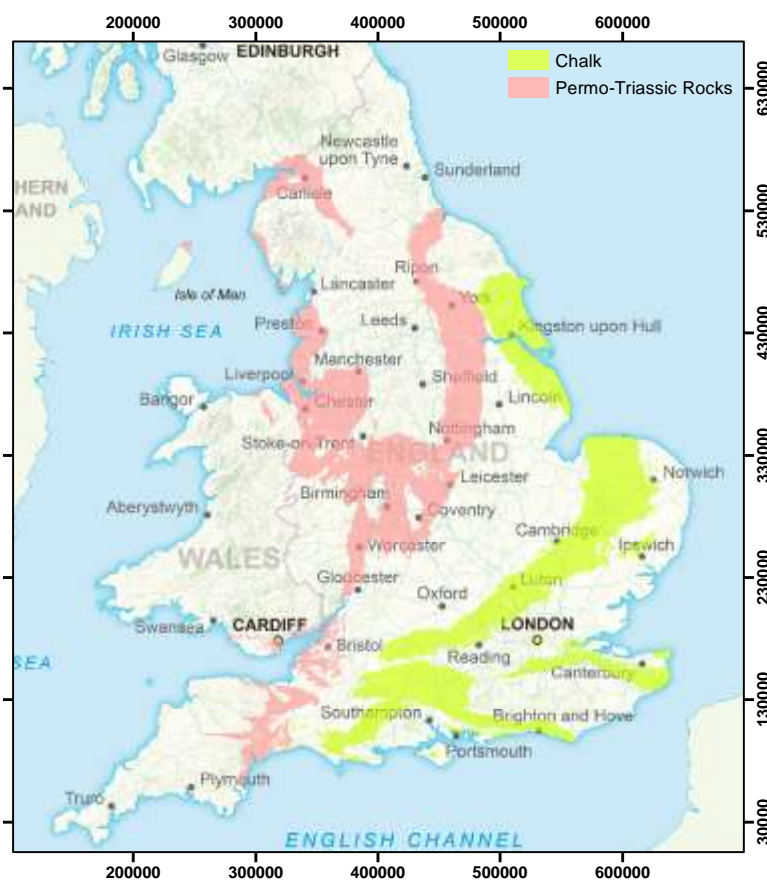
Figure 1 Location of the study area of England and Wales within the United Kingdom and the outcrop of the Chalk and Permo-Triassic rocks. Contains Ordnance Data © Crown Copyright and database rights 2017 . Ordnance Survey Licence no. 100021290

2.3 Data extraction, collation and statistical analysis

170

171

172 Under the Water Supply (Water Quality) Regulations [41], the water supply utilities in England and

173 Wales provide the results of the routine water quality sampling detailed above as PDF reports to

174 customers on their websites. Water utility supply areas are divided based on operational factors into

175 designated water supply zones (WSZ), which supply up to 100,000 people, have approximately

176 uniform quality and can comprise a combination of small communities in rural areas. Each water

177 quality report is for a defined WSZ and, under normal conditions, on request all customers within a

178 WSZ receive the same report. These reports can be downloaded using a postcode search. The

179 locations of WSZ boundaries are sometimes available but not consistently across the study area. We

180 downloaded all WSZ water quality reports for water companies in England and Wales for 2015. Where

181 WSZ boundary mapping was not available, we derived WSZ areas based on postcode data. We divided

182 England and Wales into a series of 1 km square grid cells. For each grid cell, the postcode in the centre

183 of the cell was extracted and the name of the corresponding WSZ recorded. We then merged the

184 areas returning the same WSZ report to derive the WSZ area outlines. The downloaded water quality

185 reports for each WSZ were then converted using the tabula software [42] and collated in a MS Access

186 database.

187

188 A large number of parameters are reported in the WSZ water quality reports as listed in

189 Supplementary Table 1. From this list we used the following criteria to exclude parameters which are

190 unlikely to reflect water provenance at the national scale:

191 • Copper, iron, aluminium, fluoride, lead and manganese, as these are all parameters that may

192   be significantly impacted by water treatment, the distribution network and customer

193   pipework.

9

194 • Phosphorus was not considered further due to the widespread practice of phosphate dosing

195   during water treatment [43]. Whilst chlorine and THMs are also artefacts of water treatment

196   processes, these parameters were included in the analysis as chlorination (and subsequent

197   THM formation) is more extensive in treatment of surface waters than groundwaters [15] and

198   thus may an indicator of provenance.

199 • No substantial data gaps at the national scale (<5% of water supply zones with missing data

200   for a certain parameter). As analysis for individual pesticides is assessed on a risk basis,

201   monitoring is not consistent across all WSZs so these were excluded.

202 • No datasets dominated by zero detects (bacterial counts, specific organic compounds (e.g.

203   benzene), radioactivity, taste/odour, pesticides)

204 Applying these criteria resulted in 17 parameters that are likely to reflect provenance, as shown in

205 Table 1. We then undertook further statistical analysis of these parameters. Some authors[44] have

206 advocated the use of compositional methods [45] to analyse water quality samples. These approaches

207 acknowledge that the concentrations of constituents in a sample sum to a whole and thus artefacts

208 can arise in standard analyses because an increase in the concentration of one constituent leads

209 directly to a decrease in the concentrations of the other constituents. Also, the sum of independent

210 predictions of each constituent do not generally sum to the whole. In a compositional approach these

211 artefacts are avoided since the concentrations are transformed to relative ratios of (often log-

212 transformed) constituents or products of constituents. We do not believe that such an approach is

213 required here for a number of reasons. First, quantities such as pH, turbidity and conductivity do not

214 form part of composition and could not be included in a compositional analysis. Second, the

215 compositional properties considered in this paper are only a subset of the constituents of a sample

216 and no not include water. Thus they sum to a tiny proportion of the whole and any artefacts in other

217 constituents resulting from an increase in one constituent will be negligible. Furthermore, the primary

218 purpose of compliance monitoring is to determine whether concentrations of individual constituents

219 are above pre-specified thresholds. Breaches of these thresholds will be harder to interpret if the

220 analysis is conducted in a transformed space which focuses on the ratio of concentrations of different

221 constituents of a sample rather than the magnitude of the concentrations.

222

223 ~~. Of the 17 parameters, data were missing for an average of 2.85% of water supply zones.~~ The

224 statistical analysis required measurements of all parameters in all water supply zones. ~~.~~ Of the 17

225 parameters, data were missing for an average of 2.85% of water supply zones. Where data were

226 missing we infilled using the median value of the same parameter at other sites. The median is a robust

227 measure of the expected value that is not unduly influenced by outliers, and the proportion of data

228 requiring infilling is very small. Thus this infilling is unlikely to introduce artefacts into the eventual

229 clusters. ~~Data were infilled for these supply zones using median values for each parameter.~~ The mean

230 and standard deviation was calculated for each determinand spilt up by aquifer type (Chalk, Permo-

231 Triassic rocks, Less productive and non-aquifers). The data were not suitable for a conventional

232 analysis of variance because they were spatially correlated and non-normally distributed. We

233 therefore followed the approach described by Lark and Cullis [46] to test the significance of any

234 differences in the mean values of each variable for each rock type. Briefly, we transformed the

235 observations of each variable to a normal distribution by a non-parametric (normal-scores) approach

236 and then estimated a linear mixed model of the transformed variable. The fixed effects of that linear

237 mixed model were categorical variables corresponding to the three rock types and the random effects

238 were assumed to have an exponential spatial covariance function. A series of Wald tests were then

239 applied to test for significant differences in the mean value of the transformed variable for each pair

240 of rock types. The spatial distribution of each parameter was assessed qualitatively by developing

241 national scale maps of the determinands with the outcrop of the principal aquifers overlain. These

242 maps show the raw data across the areal extent of WSZs, with no interpolation undertaken. The 17

243 parameters were standardised and ~~principal component analysis was applied to assess whether the~~

244 ~~distribution of concentrations of these parameters can be explained by a smaller number of~~

11

245 ~~determinands.~~ ~~W~~we then undertook k-means cluster analysis for k = 2 to k = 5 [47] using R [48]. As the

246 choice of an appropriate number of clusters is somewhat subjective, we developed a parsimonious,

247 rule based approach. We identified the smallest number of clusters which (1) produces spatially

248 coherent cluster membership at the national scale and~~,~~ (2) the spatial patterns of cluster membership

249 correspond to areas of groundwater and surface water supplies ~~and (3) shows coherent patterns with~~

250 ~~in the first 2 principal components~~. Using this approach, 3 clusters were identified as representing

251 drinking water provenance on the basis of groundwater and surface water at the national scale.

252 Increasing the number of clusters above 3 resulted in incoherent patterns of cluster membership.

253 Such patterns are likely to represent more local scale hydrochemical processes effecting tap water

254 chemistry which are not the focus of this national scale study.

255

3 Results

3.1 Database statistics and regulatory compliance

258

259 The database developed covers 1539 supply zones across England and Wales. Based on the

260 downloaded water quality reports a total of 3,873,941 water chemistry samples were reported in

261 2015. There are 190 unique determinands within the database. For each determinand within a WSZ,

262 a maximum, minimum and mean concentration is reported, in addition to the number of samples

263 taken in the year and the number that exceeded the drinking water limit. For each water supply zone

264 the number of determinands varies substantially. The maximum and median number of determinands

265 reported for a WSZ was 272 and 75 respectively. This wide range in the number of determinands is

266 the result of different water supply zones having different reporting requirements associated with

267 different population levels. Water companies operating water supply zones which have experienced

268 water quality problems associated with certain parameters may have a regulatory obligation to report

12

269   these parameters. This is often the case with individual pesticides, which cover 111 of 190

270   determinands. The sample data, however, show a high level of compliance to DWI and EUDWD

271   standards, with 99.94% of samples compliant. This agrees well with the reported compliance statistics

272   presented by Drinking Water Inspectorate [2] for 2014 (99.96% for England).

273

3.2 Spatial distribution of determinands

275

276   In this section, the spatial distribution of concentration data for key parameters within drinking water

277   is presented. Determinands have been grouped based on similarity in their spatial distribution. Table

278   1 shows the mean and standard deviation of the determinands analysed split by principal aquifers

279   (Chalk and Permo-Triassic Rocks) and less productive aquifers and non-aquifers. Also shown are the

280   results of the significance test of Lark and Cullis [46]. Statistically significant differences were observed

281   between the rock types for 10 out of the 17 parameters (p < 0.001, for PT-Chalk, PT-Other and Chalk-

282   Other).

283

284

285

286

287

13

288 **Table 1 Mean and standard deviation for determinands for drinking water samples classified according to bedrock geology (principal aquifers (Permo-**
289 **Triassic (PT) and Chalk) and less productive aquifers and non-aquifers). Results of the significance test of Lark and Cullis [46] are shown in the last 6**
290 **columns. Positive sign indicates that the parameter is greater in the first rock type is greater than the second.**

| Determinand	Unit	PCV	Permo-triassic rocks		Chalk		Other rocks		PT - Chalk		PT - Other		Chalk - Other	
			Mean	SD	Mean	SD	Mean	SD	sign	p	sign	p	sign	p
Ammonium	mg NH4/l	0.5	0	0.01	0.01	0.03	0.03	0.06	-	0.029	-	<0.001	-	<0.001
Antimony	ug Sb/l	5	0.08	0.13	0.04	0.09	0.08	0.13	+	0.096	-	0.011	-	<0.001
Arsenic	ug As/l	10	0.57	0.95	0.27	0.44	0.4	0.48	+	0.117	-	0.221	-	0.003
Boron	mg B/l	1	0.01	0.02	0.02	0.05	0.03	0.03	-	<0.001	-	<0.001	-	0.011
Chloride	mg Cl/l	250	24.59	18.63	34.24	17.08	34.75	20.73	-	<0.001	-	<0.001	+	0.062
Chlorine	mg Cl2/l		0.38	0.23	0.28	0.12	0.38	0.21	+	<0.001	+	0.027	-	<0.001
Chromium	ug Cr/l	50	0.15	0.28	0.2	0.45	0.14	0.29	+	0.083	+	0.003	+	0.298
Conductivity	µS/cm @ 20 °C	2500	314.25	182.79	571.77	95.32	446.18	215.17	-	<0.001	-	<0.001	+	<0.001
pH	pH Units	6.50-9.50	7.5	0.26	7.41	0.16	7.55	0.25	+	<0.001	-	<0.001	-	<0.001
Nickel	ug Ni/l	20	0.79	1.04	1.07	1.69	1.22	1.11	-	<0.001	-	<0.001	-	0.01
Nitrate	mg NO3/l	50	10.83	10.68	25.26	11.07	13.45	10.58	-	<0.001	-	0.251	+	<0.001
Selenium	ug Se/l	10	0.18	0.3	0.5	0.61	0.28	0.43	-	<0.001	-	0.009	+	<0.001
Sodium	mg Na/l	200	18.3	12.65	19.2	10.98	23.12	13.06	-	<0.001	-	<0.001	-	0.078
Sulphate	mg SO4/l	250	38.32	25.86	34.51	24.04	49.61	30.99	+	0.431	-	<0.001	-	<0.001
Total Organic Carbon	mg/l		1.03	0.71	0.94	0.58	1.62	0.84	+	0.014	-	<0.001	-	<0.001
Total Trihalomethanes	ug/l	100	26.47	13.41	12.13	8.29	24.44	11.78	+	<0.001	+	0.897	-	<0.001
Turbidity	NTU	4	0.03	0.06	0.04	0.06	0.06	0.06	-	0.006	-	<0.001	-	0.004

291

292

293

294

3.2.1 Nitrate

Figure 2 shows the spatial distribution of nitrate concentrations in drinking waters in England and Wales. High nitrate concentrations are present in south and east England corresponding broadly to the outcrop of the Chalk aquifer and some parts of the Permo-Triassic rocks. Analyses of drinking waters from areas of the Chalk show a very different nitrate concentration distribution to those from the Permo-Triassic sandstones, with higher mean values (25.2 mg/L) and samples most frequently in the 20-40 mg/L range for Chalk compared to 10.8 mg/L and samples in the 0-10 mg/L range for the Permo-Triassic. Low concentrations are present where the Chalk is overlain by low-permeability Palaeogene and superficial deposits (primarily till) in East Anglia. Areas which are shown in white show returned no drinking water quality report. These areas can be considered to be where no mains supply is present and drinking water is obtained from local private supplies.
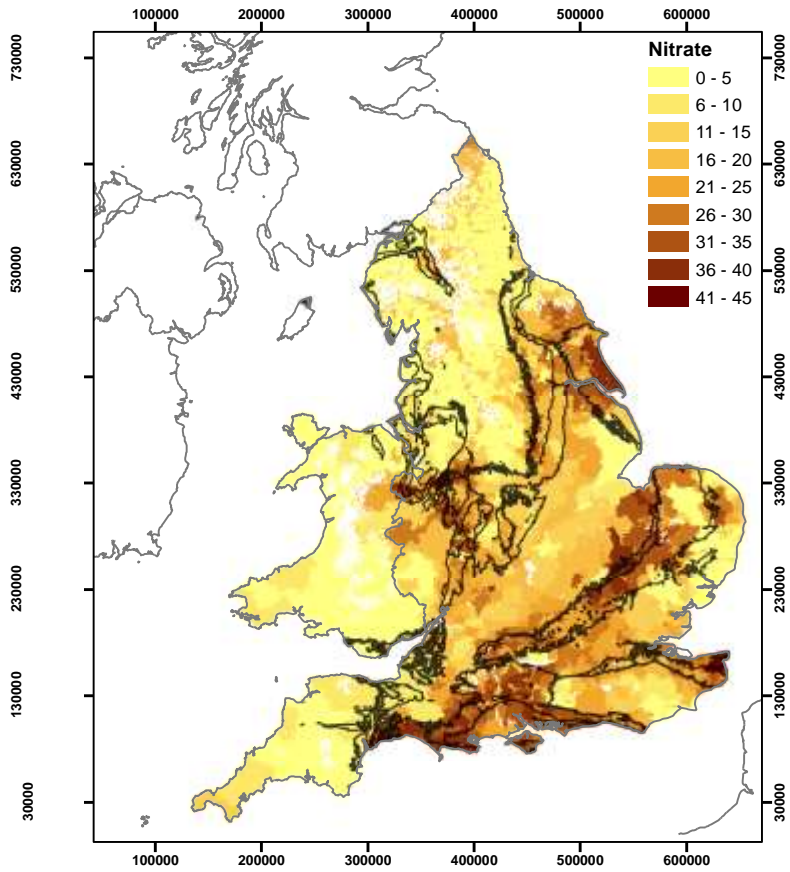
**Figure 2**2 **Nitrate concentrations (mg NO$_3$/L) in drinking water in England and Wales in 2015**

3.2.2 Nickel and Selenium

Concentrations of trace substances (Ni, Se, As) are low over most of England and Wales. Elevated

concentrations of substances such as Ni and Se, are found in areas of East Anglia where the Chalk is

16

317   not at outcrop (Figure 3). Mean Ni and Se concentrations are very low from supplies on the Permo-

318   Triassic and approximately double from the Chalk (Table 2).
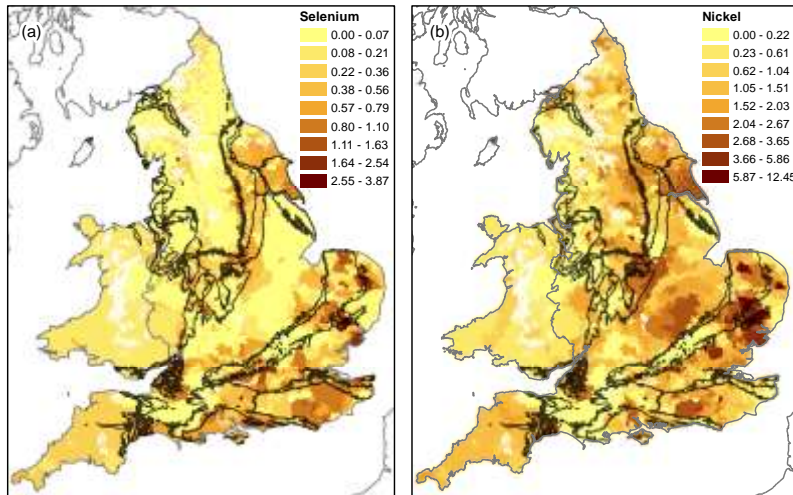
319

320   **Figure 33 Selenium (a) and Nickel (b) concentrations (µg/L) in drinking water in England and Wales**
321   **in 2015**

322

323

324   ### 3.2.3   TOC, Chlorine, THMs and Turbidity
325

326   Figure 4 shows TOC, chlorine, THMs and turbidity concentrations for drinking water in the study area.

327   Elevated TOC concentrations (of up to 3 mg/L) are measured in the northeast coast of England,
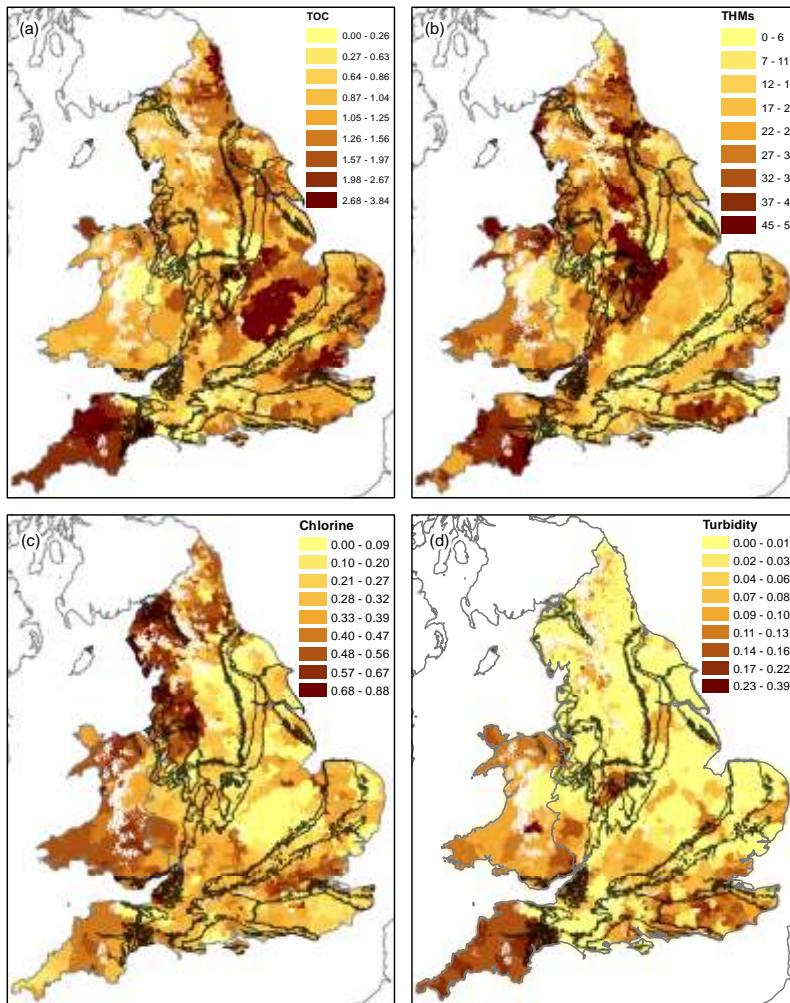
328   Anglesey, southwest England, Essex, and an area of central England around Bedford, Northampton

329   and Peterborough (Figure 4). Average concentrations in supplies located on the aquifers of the Chalk

330   and the Permo-Triassic are similar, about 1 mg/L, whereas the average for less productive aquifers

331   and non-aquifers is higher (1.62 mg/L, Table 3).

17

332   The highest residual chlorine concentrations are seen in northwest England (the Lake District Coast

333   and Cheshire) and parts of Wales and southwest England (Figure 4). Supplies from Chalk areas have

334   the lowest average residual chlorine (0.28 mg/L), with increasing concentrations on the Permo-Triassic

335   and on less productive aquifers and non-aquifers (0.38 mg/L, Table 3). Elevated THM concentrations

336   of up to 50 µg/L occur in south Wales and southwest England, the Weald, easterly East Anglia and the

337   Pennines (Figure 4). Average concentrations in supplies on the Chalk are 12.1 µg/L, whereas on the

338   Permo-Triassic and less productive aquifers and non-aquifers they are in the range 24 to 26 µg/L (Table

339   3). Turbidity values are higher in southwest England and parts of Wales (up to 0.3 NTU) than eastern

340   England (Figure 4). Average values are similar across the study area with the lowest for the Permo-

341   Triassic (0.03 NTU) and highest on less productive aquifers and non-aquifers (0.06 NTU) (Table 3).

Figure 44 TOC (mg/L, a), THMs (µg/L , b), Chlorine (mg/L, c) and Turbidity (NTU, d) in drinking water in England and Wales in 2015
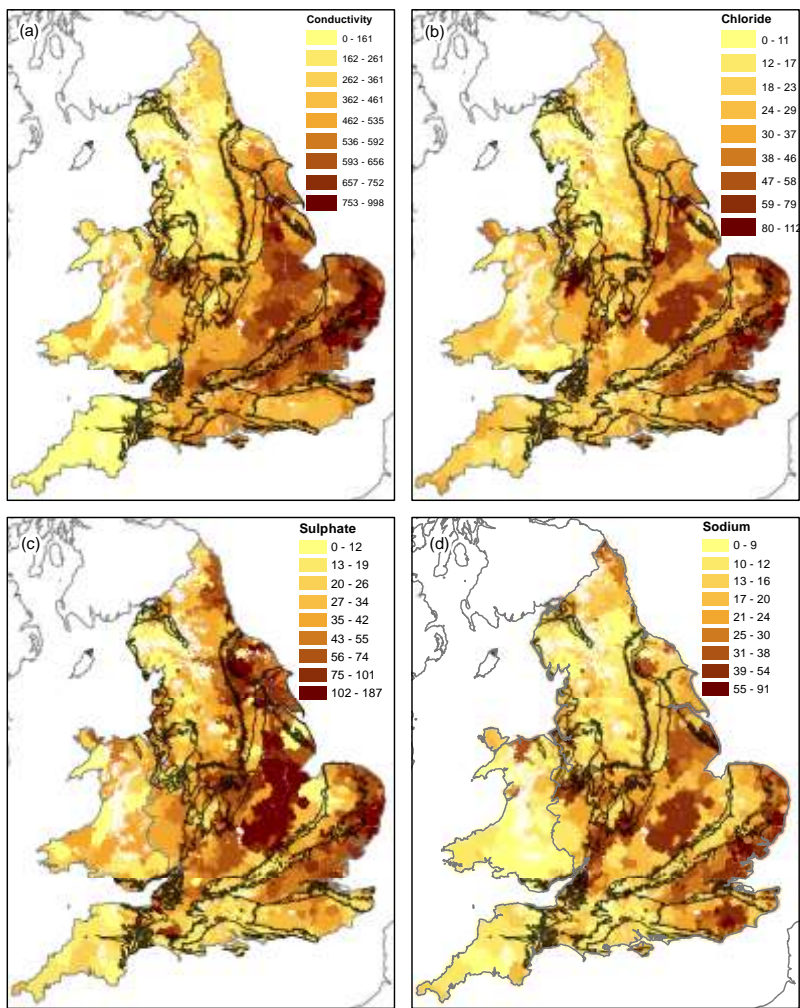
3.2.4 Conductivity, chloride, sodium and sulphate

Drinking water conductivity is lowest along the west coast and highest in eastern East Anglia where values of up to 900 µS/cm are recorded (Figure 5). Mean conductivity values are considerably higher from areas on the Chalk than on the Permo-Triassic or less productive aquifers and non-aquifers (Table 2). Chloride concentrations follows a similar pattern to conductivity but with additional elevated concentrations in Cheshire and the East Midlands (Figure 5). Mean chloride concentrations are higher on the Chalk and less productive aquifers and non-aquifers (34-35 mg/L) than on the Permo-Triassic (24.6 mg/L) (Table 2). Sodium also follows this pattern although average concentrations do not behave similarly. Mean sodium concentrations are higher on less productive aquifers and non-aquifers (23.1 mg/L) than on the Permo-Triassic and on the Chalk (18-19 mg/L). Sulphate is also similar with less obvious elevation of concentration in East Anglia and more in the East Midlands and Yorkshire. Average concentrations are in the range 20-30 mg/L. Like sodium, mean concentrations are considerably higher on less productive aquifers and non-aquifers (49.6 mg/L) than on the Permo-Triassic and the Chalk (34 -39 mg/L) (Table 3).

Figure 5 Conductivity (μS/cm, a), chloride (mg/L, b), sulphate (mg/L, c) and sodium (mg/L, d) in drinking water in England and Wales in 2015

3.2.5   Other factors

A small group of the 17 parameters only provide limited insight into hydrochemical processes. Ammonium concentrations are slightly elevated in confined areas of the Chalk in the London area and in East Anglia with some concentrations above 0.05 mg/L. Average concentrations range from 0.03 mg/L on less productive aquifers and non-aquifers to <LOD in the Permo-Triassic. Arsenic concentrations are elevated in a few localities, in Cheshire, and the Bristol area. Average values are highest in the Permo-Triassic where it can be naturally occurring and lowest in the Chalk (Table 3). Average antimony concentrations are very low (0.04-0.08 µg/L but also exhibit locally higher concentrations in Cheshire. Boron concentrations are also very low (0.01-0.03 mg/L) with highest concentrations in the Weald and in southern East Anglia.

3.3   Statistical analysis

Figure 6 shows the results of the ~~principal component and~~ cluster analysis; three spatially coherent clusters can be identified. Cluster 1 comprises WSZs in the south east of England and some parts of the Midlands, with significant areas overlapping the outcrop of Chalk and Permo-Triassic aquifers. Cluster 2 WSZs are located in Wales and the southwest and the north of England, where there are limited groundwater resources. Cluster 3 is more spatially variable, covering parts of East Anglia and the southeast, the East Midlands and northeast England. In these areas there is a combination of groundwater resources (including the

(Figure 7~~Figure 6 (c)~~) show the differences between clusters for key determinands. Cluster 1 has high nitrate concentrations and conductivity, low organic carbon, chlorine and THM concentrations in comparison to cluster 2. Cluster 2 has low nitrate concentrations, conductivity, sodium and chloride concentrations and higher chlorine and THM concentrations. Cluster 3 has higher conductivity,

22

393 sodium, chloride and sulphate concentrations in addition to higher boron, antimony, nickel and

394 selenium concentrations. Cluster 3 also has relatively low chlorination and THMs, despite higher TOC
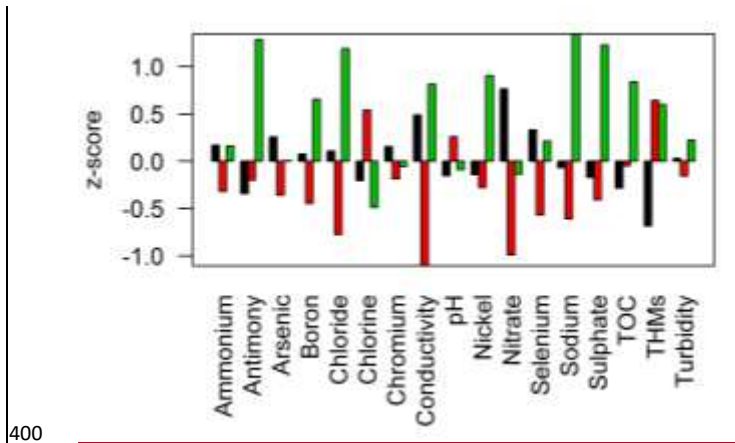
395 concentrations.

396
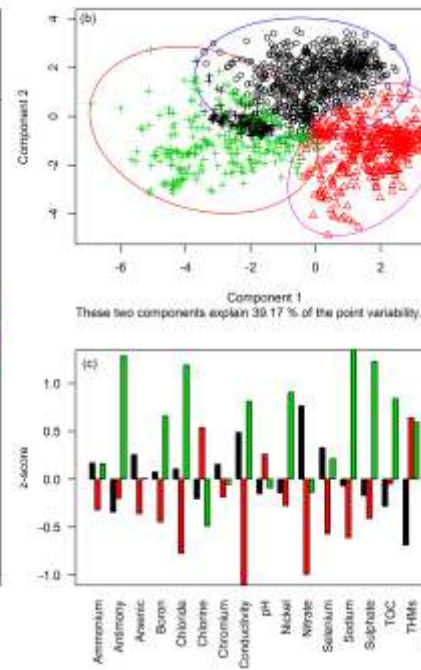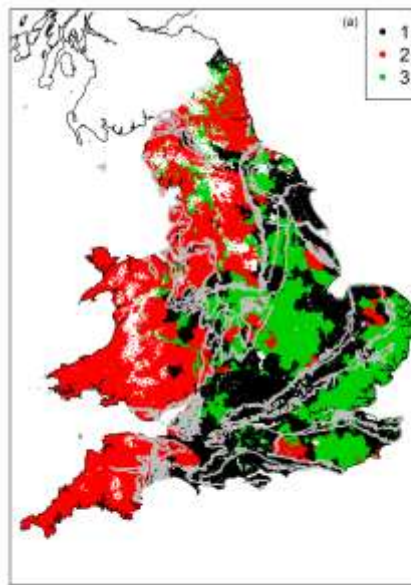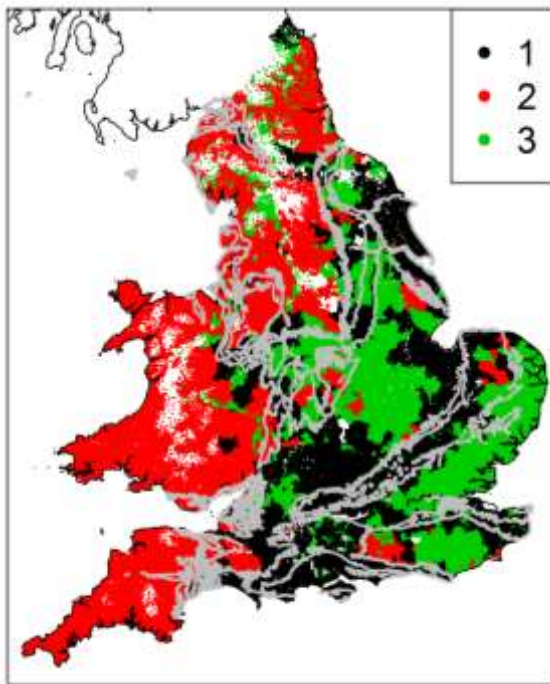
397

398 **Figure 66 Location of water supply zone clusters**

399

23

Figure 7 ~~(a), cluster bivariate plot (b) and z~~Z-scores for the cluster centroids ~~(c)~~

5.~~1~~4.1   Hydrogeochemical controls on drinking water typologies

In this section we relate the spatial distributions presented in section 3.2 to potential controlling factors in water provenance. It should be noted that water utilities use a number of options for ensuring that drinking water is compliant with the water quality regulations. These can include removal/reduction of determinands by water treatment which can result in regulated substances exhibiting a truncated distribution of concentrations.- In this analysis, it is assumed that water comes either from groundwater or surface water. However, in the future drinking water may also be obtained by desalination. In England currently there is only one plant used to desalinate water using the reverse osmosis (RO) process for public supply, on the Thames Estuary, and which has operated since 2010 providing up to 150 ML/day during peak times [49]. Drinking water derived from this source will differ

416   significantly in terms of hydrochemistry compared to that from groundwater or surface water sources,

417   because it is derived from the tidal zone of the Thames and has undergone demineralisation [50].

418

419   The spatial distribution of nitrate concentrations (Figure 2) shows a clear influence of both underlying

420   hydrogeology and land use, identifiable in cluster 1 (Figure 6). Large areas of southern and eastern

421   England obtain the majority of their supplies from groundwater [51]. The high nitrate concentrations in

422   drinking waters derived from the Chalk may reflect the storage of nitrate in the thick Chalk

423   unsaturated zone and slower flushing of nitrate following changes in agricultural management

424   practices [52, 53, 54, 55]. This assessment does not include areas of the Chalk where it is not at outcrop, e.g.

425   the eastern part of East Anglia where some elevated values are shown in Figure 2. Drinking water

426   chemistry demonstrates a residual land use/geology signature despite treatment of water for elevated

427   nitrate [56]. This is unsurprising given that nitrate removal by ion exchange is unlikely to be undertaken

428   on raw waters where concentrations are below 50 mg NO_3/L. It would be anticipated that phosphate

429   would be similarly useful were its distribution not obscured by treatment for plumbosolvency [43].

430

431   The spatial distribution of nickel and selenium (Figure 3) reflects geochemical processes occurring as

432   recharge occurs through overlying superficial deposits. For example, Ander, Shand [57] showed that

433   oxidation of sulphide minerals (e.g. pyrite) in overlying till deposits in East Anglia is the primary source

434   of high nickel concentrations in Chalk groundwater.

435

436

437   Total organic carbon and other associated parameters (Figure 4) shows a clear influence of surface

438   water, identifiable in cluster 2 (Figure 6) . Higher concentrations of total organic carbon (TOC) would

439   be expected to occur in areas of hard-fractured rocks or sandstones where superficial deposits may

440   be peaty and/or supplies may be predominantly from surface water [58]. These areas correspond to the

441   predominance of surface water supply. Trihalomethanes (THMs) are a long-recognised by-product of

442   water disinfection by chlorine and result from reaction of chlorine with organic carbon [59]. The reaction

443   is enhanced in the presence of bromide [60, 61]. Higher dosing of chlorine is required in water with a

444   higher TOC content to obtain an acceptable residual chlorine concentration. In this dataset, the spatial

445   distribution of THMs shows a qualitative relationship to that of TOC (Figure 4). Although quantitatively

446   the relationship has substantial scatter (R^2 = 0.21), this is broadly in agreement with the findings of

447   Valdivia-Garcia, Weir [62] which showed dissolved organic carbon to be an important predictor variable

448   in the spatial distribution of THMs. Together these substances (TOC, chlorine and THMs) provide a

449   clear indication where water derived from surface water predominates in drinking water.

450

451   Conductivity and associated parameters (Figure 5) show a strong east-west spatial trend likely to be

452   associated with recharge processes. Rainfall for England and Wales is predominantly from the

453   southwest with highest amounts recorded on upland areas of Wales and the Lake District and low

454   values in Eastern England, including London, East Anglia and Lincolnshire. The distribution of

455   conductivity values appears to be inversely related to recharge [63] and therefore predominantly reflects

456   meteorological setting. High chloride concentrations in Cheshire may be associated with halite

457   deposits in the Mercia Mudstone and related salt mining activity [64]. Conductivity and the major ions

458   included in regulatory monitoring are likely to be little affected by drinking water treatment [65] and

459   therefore retain their hydrological signature of the raw waters. Chloride could be augmented by

460   treatment for nitrate by ion-exchange (see section 4.2).

461

462   **5.2 Impact of interventions to ensure compliance**

26

463 ~~There is also the possibility that side effects of interventions could be the modification of other components in the raw water Some types of water treatment are very widely applied~~

464

465

466 **5.3~~4~~.2** **Drinking water compliance sample data for hydrochemical**

467 **characterisation: An outlook**

468 **5.3.1~~4~~.2.1** **Benefits and limitations** ~~of the methodology~~

469

470 As previously discussed, the interpretation of drinking water datasets for hydrochemistry has been

471 shown to be challenging due to mixing of water sources, treatment, the distribution network and

472 sampling point location. Nevertheless, the cluster analysis and the data discussed above clearly shows

473 that compliance samples do reveal drinking water provenance in terms of the raw water sources that

474 dominate water supply in the study area. For a number of these determinands, a relatively confident

475 interpretation of the environmental controls on the spatial trends can be made. Flem, Reimann [15]

476 suggested that sampling and centralised analysis of drinking water may be an effective low cost

477 method for gaining insights into processes effecting drinking water chemistry. Building on this, here

478 we suggest that significant further understanding into these processes can be gained from analysis of

479 compliance samples. Uniform analytical, sampling and reporting standards mean that datasets from

480 different water companies can be compared. The use of compliance water company samples for

481 hydrochemical characterisation over specific centralised sampling [15, 21] for research has both

482 advantages and disadvantages. Compliance samples cover a much denser sampling network both

483 spatially and temporally than specific samples. However, the parameter range for routine samples is

484 restricted to determinands which are of concern for human health. Consequently, there are a

485 significant number of parameters which are not consistently reported which would be of significant

486 hydrogeochemical interest (e.g. alkalinity, dissolved oxygen, calcium, magnesium, potassium). As a

487 result, it is unlikely that data from compliance sampling could be used in conventional

27

488 hydrogeochemical analyses and modelling (e.g. development of Piper/Durov diagrams, PHREEQC

489 modelling). For example, Shand, Edmunds [66] report on baseline groundwater chemistry for England

490 and Wales focussing on major and minor aquifers and Smedley [67] examined UK bottled water

491 chemistry, which tends to reflect the relatively minor aquifers. These studies, to which this work is

492 complementary, discuss primarily major ion chemistry and a range of trace elements not necessarily

493 represented in drinking water regulatory monitoring.

494

495 ~~5.3.2~~4.2.2   **Applications and further work**

496

497 Drinking water compliance data have been used extensively in regulatory reporting. Detailed

498 hydrochemical analysis and interpretation of this data has never before been reported. We consider

499 there to be a wide range of potential applications of ~~both~~ the dataset ~~and the analytical methodology~~

500 used in the research reported here.

501 The data could be used to support management decisions regarding the potential water chemistry

502 implications of raw and treated water transfers. ~~Figure 9~~Figure 8 shows the location of the clusters identified

503 in this study and suggested raw and treated water transfers [14]. Where transfers are between clusters,

504 addition of water of different hydrochemical typologies may have significant implications for both

505 human and environmental health. Without further water treatment, transfers of corrosive surface

506 waters into areas previously supplied by groundwater may result in dissolution of metals from water

507 mains. Where mains water leakage is significant, transfers may result in a flux of water that is

508 hydrochemically different to the water in the environment. Recent work has shown mains water

509 leakage to be a significant source of phosphorus (P) to the environment [68, 69, 70, 71]. Transfer of

510 phosphorus dosed mains water into an area without historic P dosing and subsequent leakage into

511 the environment could represent a significant additional source of P. ~~The methodology~~Application of

512 the data presented in this study would be an ideal high-level screening tool to evaluate the water

28

513 quality implications of water transfers at the national scale. At the level of individual transfers,

514 substantial additional work would be required considering the water quality of both the transferred

515 water and the current water in a supply zone, the distribution network age, material type and location.

516
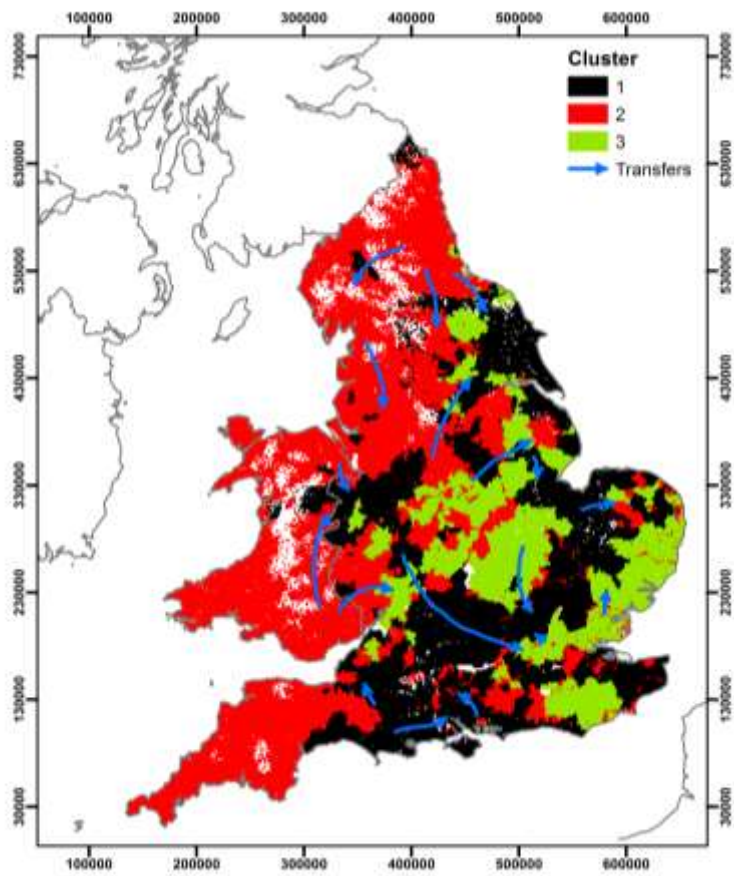
517

518 **Figure 88 Location of water supply zone clusters and suggested [14] large scale water transfers**

519

520

521   The datasets could be reviewed in the context of national scale health datasets. The Environment and

522   Health Atlas [72] provides detailed maps of both environmental agents and health conditions in England

523   and Wales. This already includes trihalomethanes but could be extended to consider other potential

524   environmental agents which are reported in the drinking water dataset. Drinking water in England and

525   Wales are compliant with current regulations but such an approach could perhaps provide evidence

526   to be used in future drinking water quality reviews.

527

528   The data collated in this study could also be compared against raw untreated water samples. This has

529   been undertaken at a continental scale in Europe by Flem, Reimann [73] but only using a small sample

530   of drinking waters analysed centrally rather than routine compliance samples. This would give an

531   indication of the efficiency of treatment processes. Comparison with groundwater and surface water

532   data would also give an indication of whether water lost through leakage would be significantly

533   different from the water in the environment. In some cases (e.g. phosphorus addition), leakage may

534   be a source of nutrients to the environment. In contrast, in cases where treatment has removed

535   contaminants from the water, leakage may dilute the concentration of pollutants already existing

536   within groundwater or surface water.

537

538   In addition to the parameters reported here, there are a large number of other non-standard

539   parameters reported on a case by case basis. The majority of these parameters (58%) are pesticides.

540   Reporting for pesticides is risk-based and thus some determinands may only be reported for a small

541   number of supply zones. The sporadic nature of these reports would make a statistical analysis such

542   as the methodology presented here challenging, but an overall qualitative interpretation would be

543   possible.

544   Further work could also explore changes in drinking water chemistry through time. The dataset

545   reported in this study is for 2015. Historically water utilities have reported similar datasets to

546   regulators back to 1993 [74]. A wide range of factors are likely to be controlling changes in water quality

547   through time such as changes in source water quality, treatment processes and water source blending.

548   Consequently unambiguous interpretation of such time series data is likely to be challenging.

549   The ~~methodology~~ use of compliance samples ~~used~~ to characterise drinking water provenance ~~based~~

550   ~~on compliance sampling~~ is likely to be broadly applicable across much of the developed world. In

551   Europe, the EUDWD [3] requires member states to report a number of determinands. High level

552   compliance summaries are reported by the European Commission e.g. [4]. In the USA, national

553   databases [75, 76] are available which report compliance failures. Whilst a few countries hold publically

554   accessible national scale databases for drinking water quality data (e.g. France, [77]), in both the USA

555   and large parts of Europe water quality data are held at the water company level. Given the high level

556   of fragmentation in the water sector in both USA and Europe (>50,000 utilities in USA[78], >6,200 in

557   Germany alone [79] data collation from individual companies would be an extremely labour intensive

558   task. Given that water utilities already report compliance data to regulators, it would be helpful if

559   regulators consistently provided these reports to the public in addition to high-level compliance

560   summaries.

561   ## ~~6~~5 Conclusions

562

563   This study has shown that compliance samples reveal the hydrogeochemical provenance of drinking

564   waters for the first time at the national scale. Despite extensive modification of source waters through

565   treatment, blending and pipework, compliance data still show a hydrochemical signature of the source

566   waters. The ~~integrated~~ use of ~~principal component and~~ cluster analysis reveals a distinct groundwater-

567   surface water split. The spatial distribution of a number of parameters which control this cluster

568 partition (nitrate, nickel and selenium, TOC, THMs, conductivity) can be interpreted relatively

569 unambiguously in terms of the source water hydrogeology. The approach ~~developed~~ used in this study

570 is low cost and utilises existing datasets. It is highly generic and can be applied anywhere where

571 compliance drinking water sampling is undertaken. The limited range of determinands measured

572 during compliance sampling make this approach complementary to targeted hydrochemical

573 investigations. The datasets developed have a wide range of applications including high level

574 screening of the hydrochemical impacts of future water transfers, assessment of the impacts of water

575 mains leakage on nutrient fluxes into the environment and comparison with national public health

576 datasets.

577

578   ## ~~7~~6 Conflicts of Interest
579

580 There are no conflicts of interest to declare.

581   ## ~~8~~7 Acknowledgements
582

583

591

98 References

1. World Health Organisation. Guidelines for Drinking-water Quality 4th Edition. (ed^(eds) (2011).

2. Drinking Water Inspectorate. Drinking water quality in England: The position after 25 years of regulation. (ed^(eds) (2015).

3. European Commission. Drinking Water Directive (Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption). (ed^(eds) (1998).

4. European Commission. Synthesis Report on the Quality of Drinking Water in the Union examining Member States' reports for the 2011-2013 period, foreseen under Article 13(5) of Directive 98/83/EC. COM/2016/0666 final. (ed^(eds) (2016).

5. NRMMC. Australian Drinking Water Guidelines Paper 6 National Water Quality Management Strategy. (ed^(eds). National Health and Medical Research Council, National Resource Management Ministerial Council, Commonwealth of Australia (2011).

6. United States Environmental Protection Agency. Safe Drinking Water Act (SDWA). (ed^(eds) (2017).

7. Qu W, Zheng W, Wang S, Wang Y. China's new national standard for drinking water takes effect. *The Lancet* **380**, e8 (2012).

8. Simpson P. Water stewardship in the twenty-first century. *Nature Climate Change* **4**, 311-313 (2014).

9. Committee on Climate Change. CCRA2: Updated projections for water availability for the UK. Final Report. MAR5354- RT002-R05-00. (ed^(eds) (2016).

10. Environment Agency. Water Resources Planning Guideline: The Technical Methods and Instructions. (ed^(eds). Environment Agency (2012).

11. Butler LJ, Scammell MK, Benson EB. The Flint, Michigan, Water Crisis: A Case Study in Regulatory Failure and Environmental Injustice. *Environmental Justice* **9**, 93-97 (2016).

12. Flint Water Advisory Task Force. Final Report (ed^(eds) (2016).

13. Liu G*, et al.* Potential impacts of changing supply-water quality on drinking water distribution: A review. *Water Res* **116**, 135-148 (2017).

33

632
633   14.   Water UK. Water resources long-term planning framework. (ed^(eds) (2016).

634
635   15.   Flem B, Reimann C, Birke M, Banks D, Filzmoser P, Frengstad B. Inorganic chemical quality of
636   European tap-water: 2. Geographical distribution. *Appl Geochem* **59**, 211-224 (2015).

637
638   16.   Smedley PL, Cooper DM, Lapworth DJ. Molybdenum distributions and variability in drinking
639   water from England and Wales. *Environmental Monitoring and Assessment* **186**, 6403-6416
640   (2014).

641
642   17.   Dinelli E*, et al.* Major and trace elements in tap water from Italy. *Journal of Geochemical*
643   *Exploration* **112**, 54-75 (2012).

644
645   18.   Le Bot B, Lucas J-P, Lacroix F, Glorennec P. Exposure of children to metals via tap water
646   ingestion at home: Contamination and exposure data from a nationwide survey in France.
647   *Environment International* **94**, 500-507 (2016).

648
649   19.   Brima EI. Physicochemical properties and the concentration of anions, major and trace
650   elements in groundwater, treated drinking water and bottled drinking water in Najran area,
651   KSA. *Applied Water Science* **7**, 401-410 (2017).

652
653   20.   M. Carstea E, Levei EA, Hoaghia M-A, Savastru R. Quality assessment of Romanian bottled
654   mineral water and tap water. *Environmental Monitoring and Assessment* **188**, 521 (2016).

655
656   21.   Banks D, Birke M, Flem B, Reimann C. Inorganic chemical quality of European tap-water: 1.
657   Distribution of parameters and regulatory compliance. *Appl Geochem* **59**, 200-210 (2015).

658
659   22.   Patterson KY, Pehrsson PR, Perry CR. The mineral content of tap water in United States
660   households. *Journal of Food Composition and Analysis* **31**, 46-50 (2013).

661
662   23.   Birke M, Rauch U, Lorenz H, Kringel R. Distribution of uranium in German bottled and tap
663   water. *Journal of Geochemical Exploration* **107**, 272-282 (2010).

664
665   24.   Cidu R, Frau F, Tore P. Drinking water quality: Comparing inorganic components in bottled
666   water and Italian tap water. *Journal of Food Composition and Analysis* **24**, 184-193 (2011).

667
668   25.   Demetriades A. Quality of hellenic bottled mineral and tap water as a proxy for groundwater
669   geochemistry. *European Geologist* **35**, 9-16 (2013).

670
671   26.   Bowen GJ, Ehleringer JR, Chesson LA, Stange E, Cerling TE. Stable isotope ratios of tap water
672   in the contiguous United States. *Water Resources Research* **43**, (2007).

673

34

674 27. Landwehr JM, Coplen TB, Stewart DW. Spatial, seasonal, and source variability in the stable
675   oxygen and hydrogen isotopic composition of tap waters throughout the USA. *Hydrological*
676   *Processes* **28**, 5382-5422 (2014).

677
678 28. Zhao S*, et al.* Divergence of stable isotopes in tap water across China. *Scientific Reports* **7**,
679   43653 (2017).

680
681 29. Jameel Y, Brewer S, Good SP, Tipple BJ, Ehleringer JR, Bowen GJ. Tap water isotope ratios
682   reflect urban water system structure and dynamics across a semiarid metropolitan area.
683   *Water Resources Research* **52**, 5891–5910 (2016).

684
685 30. Darling W, Bath A, Talbot J. The O & H stable isotopic composition of fresh waters in the
686   British Isles. 2. Surface waters and groundwater. *Hydrology and Earth System Sciences* **7**,
687   183-195 (2003).

688
689 31. Scientific American. Trump Wants Deep Cuts in Environmental Monitoring. (ed^(eds)
690   (2017).

691
692 32. The Independent. Massive job cuts at the Environment Agency. (ed^(eds) (2013).

693
694 33. Acevedo M. *Real-Time Environmental Monitoring: Sensors and Systems*. CRC Press (2016).

695
696 34. Environment Agency. Water Quality Archive. (2017).

697
698 35. Stuart M, Chilton P, Kinniburgh D, Cooper D. Screening for long-term trends in groundwater
699   nitrate monitoring data. *Quarterly Journal of Engineering Geology and Hydrogeology* **40**,
700   361-376 (2007).

701
702 36. United States Environmental Protection Agency. Drinking Water Tools. (ed^(eds) (2016).

703
704 37. Drinking Water Inspectorate. Guidance on the implementation of the Water Supply (Water
705   Quality) Regulations (as amended) in England. (ed^(eds) (2012).

706
707 38. Department for Environment Food & Rural Affairs. Water abstraction from non-tidal surface
708   water and groundwater: England and Wales. (ed^(eds) (2016).

709
710 39. Cambridge Water Company. Your Supply. (ed^(eds) (2016).

711
712 40. Dŵr Cymru. Water resources. (ed^(eds) (2016).

713

714 41. Water Supply (Water Quality) Regulations. In: *S.I. 2016/614* (ed^(eds). Stationery Office Ltd
715   (2016).

716
717 42. Tabula. Tabula 1.1.1. (ed^(eds) (2017).

718
719 43. Hayes C. *Best practice guide on the control of lead in drinking water*. IWA Publishing (2010).

720
721 44. Buccianti A, Pawlowsky-Glahn V. New Perspectives on Water Chemistry and Compositional
722   Data Analysis. *Mathematical Geology* **37**, 703-727 (2005).

723
724 45. Aitchison J. The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical*
725   *Society: Series B (Methodological)* **44**, 139-160 (1982).

726
727 46. Lark R, Cullis B. Model-based analysis using REML for inference from systematically sampled
728   data on soil. *European Journal of Soil Science* **55**, 799-813 (2004).

729
730 47. Webster R, Oliver MA. *Statistical methods in soil and land resource survey*. Oxford University
731   Press (OUP) (1990).

732
733 48. R Development Core Team. R: A language and environment for statistical computing.
734   (ed^(eds). R Foundation for Statistical Computing (2016).

735
736 49. Thames Water. Thames Gateway Water Treatment Works. (ed^(eds) (2016).

737
738 50. Water Projects Online. Thames Gateway Water Treatment Plant. (ed^(eds) (2010).

739
740 51. Ascott MJ. Groundwater resources in the UK. (ed^(eds) (2017).

741
742 52. Ascott MJ, Wang L, Stuart ME, Ward RS, Hart A. Quantification of nitrate storage in the
743   vadose (unsaturated) zone: a missing component of terrestrial N budgets. In: *European*
744   *Geosciences Union 2015* (ed^(eds) (2015).

745
746 53. Wang L*, et al.* Prediction of the arrival of peak nitrate concentrations at the water table at
747   the regional scale in Great Britain. *Hydrol Processes* **26**, 226-239 (2012).

748
749 54. Wang L, Butcher A, Stuart M, Gooddy D, Bloomfield J. The nitrate time bomb: a numerical
750   way to investigate nitrate storage and lag time in the unsaturated zone. *Environ Geochem*
751   *Health* **35**, 667-681 (2013).

752

36

753 55. Wang L, *et al.* The changing trend in nitrate concentrations in major aquifers due to
754   historical nitrate loading from agricultural land across England and Wales from 1925 to 2150.
755   *Sci Total Environ* **542**, 694-705 (2016).

756
757 56. UK Water Industry Research Ltd. Implications of Changing Groundwater Quality for Water
758   Resources and the UK Water Industry- Phase 3: Financial and Water Resources Impact.
759   (ed^(eds). UKWIR (2004).

760
761 57. Ander EL, Shand P, Griffiths KJ, Lawrence AR, Hart P, Pawley J. Baseline Report Series: 13.The
762   Great Ouse Chalk aquifer, East Anglia. (ed^(eds). British Geological Survey Commissioned
763   Report CR/04/236N (2004).

764
765 58. Worrall F, *et al.* Trends in Dissolved Organic Carbon in UK Rivers and Lakes. *Biogeochemistry*
766   **70**, 369-402 (2004).

767
768 59. Cotruvo JA. Trihalomethanes (THMs) in drinking water. *Environmental Science & Technology*
769   **15**, 268-274 (1981).

770
771 60. Chowdhury S, Champagne P, McLellan PJ. Investigating effects of bromide ions on
772   trihalomethanes and developing model for predicting bromodichloromethane in drinking
773   water. *Water Research* **44**, 2349-2359 (2010).

774
775 61. Cooper WJ, Zika RG, Steinhauer MS. Bromide-oxidant interactions and THM formation: a
776   literature review. *Journal of the American Water Works Association*, 116-121 (1985).

777
778 62. Valdivia-Garcia M, Weir P, Frogbrook Z, Graham DW, Werner D. Climatic, Geographic and
779   Operational Determinants of Trihalomethanes (THMs) in Drinking Water Systems. *Scientific*
780   *Reports* **6**, 35027 (2016).

781
782 63. Mansour MM, Wang L, Whiteman M, Hughes AG. Estimation of spatially distributed
783   groundwater potential recharge for the United Kingdom. *Quarterly Journal of Engineering*
784   *Geology and Hydrogeology*, (2018).

785
786 64. Griffiths KJ, Shand P, Ingram J. Baseline report series. 8, the Permo-Triassic sandstones of
787   Manchester and east Cheshire. (ed^(eds). Environment Agency (2003).

788
789 65. Asenbaum A, *et al.* Influence of various commercial water treatment processes on the
790   electric conductivity of several drinking waters. *J Mol Liq* **160**, 144-149 (2011).

791
792 66. Shand P, Edmunds WM, Lawrence AR, Smedley P, Burke S. The natural (baseline) quality of
793   groundwater in England and Wales. (ed^(eds). British Geological Survey Research Report
794   RR/07/06 and Environment Agency Technical Report NC/99/74/24 (2007).

795

796   67.   Smedley PL. A survey of the inorganic chemistry of bottled mineral waters from the British
797   Isles. *Applied Geochemistry* **25**, 1872-1888 (2010).

798
799   68.   Ascott MJ, Gooddy DC, Lapworth DJ, Stuart ME. Estimating the leakage contribution of
800   phosphate dosed drinking water to environmental phosphorus pollution at the national-
801   scale. *Science of the Total Environment* **572**, 1534-1542 (2016).

802
803   69.   Lapworth DJ, Gooddy DC, Allen D, Old GH. Understanding groundwater, surface water, and
804   hyporheic zone biogeochemical processes in a Chalk catchment using fluorescence
805   properties of dissolved and colloidal organic matter. *Journal of Geophysical Research:*
806   *Biogeosciences* **114**, G00F02 (2009).

807
808   70.   Gooddy DC, Lapworth DJ, Ascott MJ, Bennett SA, Heaton TH, Surridge BW. Isotopic
809   fingerprint for phosphorus in drinking water supplies. *Environmental Science & Technology*
810   **49**, 9020-9028 (2015).

811
812   71.   Gooddy DC*, et al.* Mains water leakage: Implications for phosphorus source apportionment
813   and policy responses in catchments. *Sci Total Environ* **579**, 702-708 (2017).

814
815   72.   UK Small Area Health Statistics Unit. Environment and Health Atlas. (ed^(eds) (2016).

816
817   73.   Flem B, Reimann C, Fabian K, Birke M, Filzmoser P, Banks D. Graphical statistics to explore
818   the natural and anthropogenic processes influencing the inorganic quality of drinking water,
819   ground water and surface water. *Appl Geochem*, (2017).

820
821   74.   European Commision. SYNTHESIS REPORT ON THE QUALITY OF DRINKING WATER IN THE
822   MEMBER STATES OF THE EUROPEAN UNION IN THE PERIOD 1993-1995 (ed^(eds). European
823   Commission, (1996).

824
825   75.   United States Environmental Protection Agency. Enforcement and Compliance History
826   Online (ECHO). (ed^(eds) (2016).

827
828   76.   United States Environmental Protection Agency. Analyse Trends: Drinking Water Dashboard.
829   (ed^(eds) (2016).

830
831   77.   Ministry of Social Affairs and Health. Qualité de l'eau potable. (ed^(eds) (2017).

832
833   78.   Global Water Intelligence. Water Market: USA. (ed^(eds) (2016).

834
835   79.   Deloitte. Water Country Profiles. (ed^(eds). Deloitte Touche Tohmatsu Ltd (2014).

836

38

837

838

839