



PEARL

Development of a chemical source apportionment decision support framework for lake catchment management

Comber, Sean D.W.; Smith, Russell; Daldorph, Peter; Gardner, Michael J.; Constantino, Carlos; Ellor, Brian

Published in:

Science of the Total Environment

DOI:

[10.1016/j.scitotenv.2017.11.313](https://doi.org/10.1016/j.scitotenv.2017.11.313)

Publication date:

2018

Link:

[Link to publication in PEARL](#)

Citation for published version (APA):

Comber, S. D. W., Smith, R., Daldorph, P., Gardner, M. J., Constantino, C., & Ellor, B. (2018). Development of a chemical source apportionment decision support framework for lake catchment management. *Science of the Total Environment*, 0(0), 96-105. <https://doi.org/10.1016/j.scitotenv.2017.11.313>

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Wherever possible please cite the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

1 **This is a proof copy available for open access.**

2 **For published copy please see: Science of the Total Environment 622-623:96-**

3 **105 01 May 2018: doi.org/10.1016/j.scitotenv.2017.11.313**

4

5 **Development of a Chemical Source Apportionment Decision Support**

6 **Framework for Lake Catchment Management**

7 Sean D W Comber^{*}, Russell Smith¹, Peter Daldorph², Michael J Gardner², Carlos
8 Constantino², Brian Ellor³

9

10 ^{*} corresponding author: B525 Portland Square, Plymouth University, Drake Circus, Plymouth,
11 PL4 8AA, UK. Tel: +44(0)1752858974. Email: sean.comber@plymouth.ac.uk

12

13 ¹ Westcountry Rivers Trust, Rain-Charm House, Stoke Climsland, Cornwall, PL17 8PH

14 ² Atkins Limited, Western House, Peterborough Business Park, Peterborough, PE2 6FZ

15 ³ UK Water Industry Research (UKWIR), 3rd Floor, 36 Broadway, London SW1H 0BH.

16

17 **Research Article – Environmental Modelling**

18 **Abstract Art**

19 **Key words**

20 Chemicals, source apportionment, lakes, nutrients, metals, environmental modelling

21

22 **ABSTRACT**

23 Increasing pressures on natural resources has led to the adoption of water quality
24 standards to protect ecological and human health. Lakes and reservoirs are particularly
25 vulnerable to pressure on water quality owing to long residence times compared with rivers.

26 This has raised the question of how to determine and to quantify the sources of priority

27 chemicals (e.g. nutrients, persistent organic pollutants and metals) so that suitable measures

28 can be taken to address failures to comply with regulatory standards. Contaminants enter

29 lakes waters from a range of diffuse and point sources. Decision support tools and models

30 are essential to assess the relative magnitudes of these sources and to estimate the impacts

31 of any programmes of measures. This paper describes the development and testing of the

32 Source Apportionment Geographical Information System (SAGIS) for future management of

33 763 lakes in England and Wales. The model uses readily available national data sets to

34 estimate contributions of a number of key chemicals including nutrients (nitrogen and

35 phosphorus), metals (copper, zinc, cadmium, lead, mercury and nickel) and organic

36 chemicals (Polynuclear Aromatic Hydrocarbons) from multiple sector sources. Lake-specific

37 sources are included (groundbait from angling and bird faeces) and hydrology associated

38 with pumped inputs and abstraction. Validation data confirms the efficacy of the model to

39 successfully predicted seasonal patterns of all types of contaminant concentrations under a
40 number of hydrological scenarios. Such a tool has not been available on a national scale
41 previously for such a wide range of chemicals and is currently being used to assist with
42 future river basin planning.

43

44 **1. INTRODUCTION**

45

46 Lakes and reservoirs serve as vital sources of drinking water and support valuable
47 ecosystems across the globe. For example, there are more than 500,000 natural lakes larger
48 than 0.01 km² (1 ha) in Europe alone. Approximately 80 to 90% of these are small with a
49 surface area of between 0.01 and 0.1 km², with only around 16,000 having a surface area
50 exceeding 1 km² (EEA, 2015a). There are currently 7,000 large dams in Europe (defined as
51 having greater than 1,000,000m³ capacity) (EEA, 2015b). Reservoirs and lakes hold
52 approximately 32,000 million m³ of drinking water across Europe representing around 20%
53 of total supply (Leonard and Crouzet, 1998). In North America the Great Lakes (Erie,
54 Michigan, Huron, Ontario, Superior) contain one fifth of the world's freshwater and 84% of
55 the water supply for the United States of America (USA) and Canada (USEPA, 2017). Other
56 than drinking water, lakes and reservoirs supply power via hydroelectric schemes, augment
57 river flows, provide recreation and are valuable habitats for water fowl and species of fish
58 and invertebrates. There are significant pressures on water quality from numerous sources
59 including urban and highway runoff, industrial discharges, mining and in some cases
60 atmospheric deposition (Comber et al., 2013). Eutrophication is a major source of concern
61 regarding water quality, related mainly to inputs of nitrogen and phosphorus from agriculture
62 and sewage effluents. In Spain, for example 33% of reservoirs were identified as
63 mesotrophic, 27% eutrophic and 10% hypertrophic (Leonard and Crouzet, 1998). Toxic algal
64 blooms occur frequently with the example of suspension of water abstraction from Lake Erie
65 in 2014, affecting half a million people (USEPA, 2017). Loss of potential amenity value has
66 seen extensive steps being taken under European Union (EU) legislation to address these
67 issues.

68

69 Lakes and artificial water bodies such as reservoirs are particularly vulnerable ecosystems
70 for a number of reasons:

- 71 1) Residence times compared with rivers are often much longer so chemical inputs take
72 longer to be flushed out
- 73 2) Accumulation of contaminants is possible in the water column and sediments
- 74 3) The lack of flow leads to sediment accumulation within the lentic water body

75 4) Resident, fish, algae and invertebrates within the lake have few options regarding
76 avoiding the contamination present and therefore may be subject to bioaccumulation
77 and/or toxicity when exposed

78

79 To protect lake environments for both *in situ* ecology and human health via drinking water
80 abstraction, the European Union Water Framework Directive (WFD) (EU, 2000) sets criteria
81 (Environmental Quality standards – EQS) for all water bodies including lakes to meet a
82 defined status categorised as ‘Good’, for over 30 Priority and Priority Hazardous
83 Substances. In addition, the Drinking Water Directive (98/83/EC) places maximum
84 acceptable concentrations for a wide range of inorganic and organic chemicals which are
85 generally more stringent than environmental standards. Although, compliance can also be
86 controlled through treatment, options for many chemicals are exceedingly expensive and so
87 catchment solutions (source control) are preferable. Regulators therefore need tools to
88 apportion sources of contamination in order to guide future regulation and for known
89 pollution, plan remedial measures in a fair and proportionate way. For the UK this has been
90 achieved through the development of the Source Apportionment Geographical Information
91 System (SAGIS), originally developed for river catchments (Comber et al., 2013) but has
92 now been further developed for lakes. SAGIS combines a number of inputs including
93 modelled, measured and estimated loads from the main point and diffuse sources of metals,
94 organics and nutrients for catchments of England, Wales and Scotland.

95

96 Once discharged to a lake catchment, any given chemical will be subject to dilution and
97 undergo various biogeochemical processes, effects that might both be incorporated into a
98 model. However, whereas there have been published reports on load apportionment to lakes
99 and reservoirs and models to determine concentrations within such water bodies by taking
100 account of physico-chemical and biochemical processes (e.g. relationships between
101 biological growth rate and nutrient availability, sunlight and temperature, and phytoplankton
102 and the growth rate of zooplankton; Gough, 1969; Yih and Davison, 2008), few models have
103 attempted to combine the two. More recent water quality models which simulate lakes
104 specifically include the United States Environmental Protection Agency (USEPA) WASP
105 (Ambrose et al., 1988), and QUAL2E models (Shanahan et al., 1998), the MIKE3 model
106 developed by the Danish Hydraulics Institute (DHI), and the Systeme Hydrologique
107 European (SHE) (Abbott et al., 1986). At a larger spatial scale, catchment models include
108 BASINS (Nasr and Breun, 2004) and the Environmental Fluid Dynamics Code (EFDC) which
109 is a multifunctional surface water modelling system, which includes hydrodynamic, sediment-
110 contaminant, and eutrophication components. EFDC has been applied to over 100 water

111 bodies including lakes and reservoirs, and is a state-of-the-art hydrodynamic model that can
112 be used to simulate aquatic systems in one, two, and three dimensions (USEPA, 2007).

113

114 Source apportionment of nutrient loads within specific catchments is well developed (EEA,
115 2005) and a number of lakes have been modelled but sources have been aggregated and
116 classified as point, agriculture, and background only (e.g. MESAW model for Lake Peipsi
117 (Vassiljev and Stålnacke, 2003); Lake Mjøsa and Vättern (Nashoug, 1999); Lough Neagh
118 (Dardni, 2007 and Danish lakes (SFT, 2005)). Source apportionment models differ from
119 water quality models in that they have value in risk assessment by determining input loads
120 and to some degree, input locations. For accurate modelling of lentic water bodies a water
121 balance needs to be completed taking into account variability of flow and chemical load as
122 well as any abstractions in the case of reservoirs. Key physico-chemical parameters have to
123 be modelled including sedimentation of suspended solids. There are currently few models
124 which can do this on a local or regional scale and none on a national scale.

125

126 Examples of specific models which combine source loads and predicted concentrations
127 include the USEPA CMB8.2 model for has been used to apportion the sources of sediment-
128 bound polynuclear aromatic hydrocarbons (PAHs) in Lake Calumet, Chicago (Li et al.,
129 2003), and mass balance modelling for PAH distributions in Lac Saint Louis, Quebec
130 (MacKay and Hickie, 2000). For nutrients a nitrogen source apportionment model has been
131 developed which converts input loads, based on the surrounding land use, to concentrations
132 using the distributed HBV-N-D and Fyrismodel models for Swedish lake catchments. The
133 models used export coefficients and simple retention equations (Lingren et al., 2007). For
134 making decisions at a river basin or national level it is necessary to be able to predict loads
135 across many lakes and reservoirs and be able to predict if compliance can be achieved
136 based on predicted concentration data resulting from an identified load reduction mitigation
137 measure being applied.

138

139 Lakes and reservoirs are subject to additional and significant chemical sources compared
140 with rivers, particularly nutrients from birds and angling. Birds are recognised as significant
141 sources of nutrients to some lakes (Manny et al., 1994; Marion et al., 1994; Hahn et al.,
142 2008). Fishing is a popular pastime in the UK. Approximately 9% of the population in
143 England and Wales have been freshwater fishing as reported by the Environment Agency in
144 2010 (EA, 2010). Ground bait, comprising of ingredients such as maize, fish meal, milk
145 protein and semolina, is commonly used, particularly for coarse fishing and nutrient inputs to
146 water bodies from the use of ground bait may present a potential threat to water quality

147 where angling intensity is high (Arlinghaus and Mehner, 2003). These studies have
148 estimated German mean annual gross P-input of 1018 g P angler⁻¹ due to ground bait.

149 This paper describes the significant development of the existing SAGIS model for rivers
150 (Comber et al., 2013) to include 763 lakes and reservoirs within England and Wales. The
151 model utilises national datasets for multiple parameters including hydrology, rainfall,
152 modelled discharges of chemicals, reported discharge loads, and spatial datasets including
153 the locations of wastewater treatment works and smaller on-site works (often termed “septic
154 tanks”), combined sewer overflow locations, output from diffuse pollution risk models, road
155 and river system networks and lake specific inputs from birds and angling. It also allows
156 more detailed local data to be applied where this is available, for example pumped volumes
157 at intake locations.

158

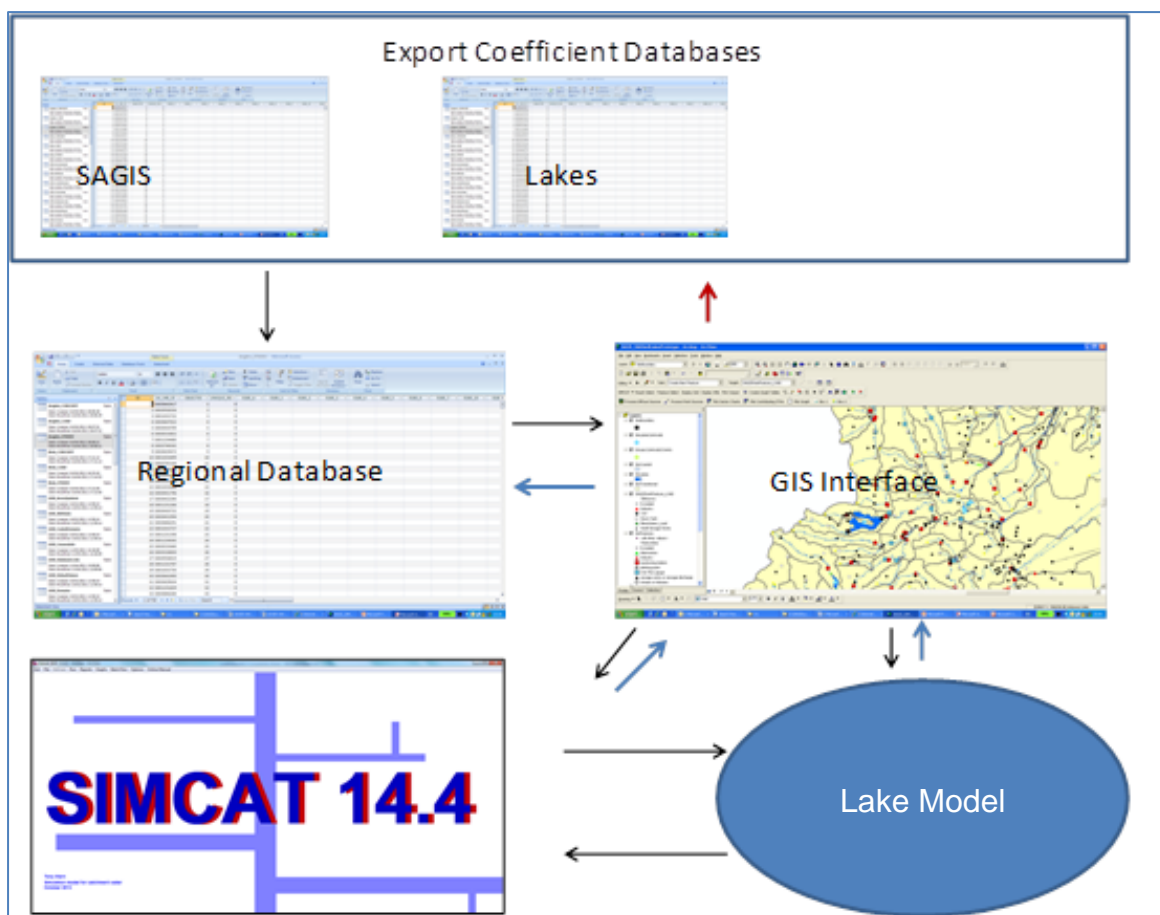
159 **2. METHODS**

160 The SAGIS modelling framework integrates data from multiple sources (Comber et al.,
161 2013), point source loads are expressed as mean and standard deviation of annual average
162 (or monthly if available) concentrations and flow, with diffuse sources input as mass per
163 year, or month, per km². Owing the large size of databases used, water bodies of England,
164 Wales and Scotland are broken up into 18 regional databases. A common map projection
165 was used for all databases and mapping based on a 1km² grid. Such GIS mapping
166 calibration and validation has been undertaken as part of previous projects associated with
167 the hydrological and diffuse source components for sources of chemicals to rivers (Comber
168 et al., 2013). Detailed information regarding the methodologies used to calculate loads for
169 each source is provided elsewhere (UKWIR, 2012). However, a brief description of the data
170 and method used to derived load estimates is provided in the following section.

171

172 For both diffuse and point sources loads of chemicals discharged into the water bodies were
173 either derived from an established model or were calculated as part of this research (See
174 Table S1 to S3 of the Electronic Supplementary Information (ESI)). Unlike estimates for river
175 catchments the lakes model includes 763 hydrologically connected lakes in England and
176 Wales. A number of additional sources are included which were considered significant for
177 these lentic water bodies. Figure 1 provides a schematic for the structure and key
178 components of the SAGIS decision support framework.

179



180

181 **Figure 1. Schematic diagram for SAGIS tool structure**

182

183 **2.1 Inputs already present within the SAGIS rivers model**

184 Sources are represented within a Microsoft Access™ database either as a point source with
 185 an X and Y, UK national grid location coordinate or as an individual 1 km² grid
 186 (approximately 150,000 for England and Wales). The main database is split into regional
 187 Access databases (see Figure S1 of the ESI) which form the attribute tables behind the
 188 features in ArcMAP GIS software. Lists of the key datasets used to derive the exported loads
 189 from each point and diffuse source and on how these are used to derive the calculated loads
 190 to water bodies are provided in the ESI.

191

192 The functionality with ArcMap and bespoke macros developed in Visual Basic are then used
 193 to extract the necessary data and generate the text file required to run SIMCAT, a stochastic
 194 water quality model. SIMCAT can be run from within SAGIS and provide outputs (total and
 195 dissolved concentrations and loads for metals, total concentrations and loads for nutrients
 196 and organics) which are fed back into ArcGIS to provide cartographic, graphical and

197 spreadsheet outputs. The existing river model was validated using monitoring data provided
198 by the Environment Agency (EA) and provided a good fit with predicted values (Comber et
199 al., 2013). The EA's Water Quality Archive provides data on water quality measurements
200 from across England for over 100 determinands, analysed within accredited laboratories.
201 Samples include coastal or estuarine waters, rivers, lakes, ponds, canals or groundwaters.
202 They are taken for a number of purposes including compliance assessment against
203 discharge permits, investigation of pollution incidents or environmental monitoring. The
204 archive provides data on measurements and samples dating from 2000 to present day.

205

206 **2.2 Additional sources to lakes**

207 **Data sources**

208 Additional data was gathered to support the development of the SAGIS Lakes tool (Table 1)
209 including where available, operational data on water company drinking water supply
210 reservoir.

211

212 **Table 1. Summary of databases utilised for the lake and reservoir source**
 213 **apportionment model**

Information	Source
WFD Lake shape files	Provided by Environment Agency (EA) and Scottish Environment Protection Agency (SEPA)
Lake Volume and area	Derived from EA and SEPA, WFD lake classification data and water company survey
Estimated Catchment Area	Derived from EA lake classification data
Lake Abstraction Location and Abstraction Quantities	Compiled from EA water resources GIS and locations and licence numbers identified from SEPA
Intake Location and Abstraction Quantities	Compiled from EA water resources GIS and locations and licence numbers identified from SEPA
Estimated bird numbers	UK wide summary data from Wetland Bird Survey data (WeBS, 2014)
Export of nutrients per bird (related to each species)	Waterbirds v1.1, 2007. Hahn et al, (2007; 2008)
Estimated angler numbers of visits	National (England and Wales) summary (EA, 2010)
Estimated load associated with each angler	Published data by Arlinghaus and Mehner (2003); Arlinghaus and Niesar (2005)
Water quality monitoring data for validation purposes	EA Water Information Management System database (WIMS) and SEPA equivalent
Discharges to lakes	SAGIS databases Comber et al., (2013)

214
 215 The key elements of the existing databases used within the model are listed and described
 216 elsewhere (Comber et al., 2013). Table S6 provides details of updated structure for lakes
 217 and reservoirs.

218
 219 **Angling ground bait (nutrients) inputs to lakes**

220 A detailed description of the methodology is provided in S1 of the ESI.

221 To derive an export load for the SAGIS model, the following data was obtained from
 222 scientific literature:

- 223 1. The number of lakes in the UK classified under the WFD that are fished and allow
 224 ground baiting;
- 225 2. The surface area and boundary length of these lakes;
- 226 3. The number of coarse fishing anglers per year in the UK; and
- 227 4. Nitrogen and phosphorus content of ground bait.

228 Annual loads to lakes were based on the lake boundary length as a proportion of the
229 boundary length of all lakes in England and Wales and the estimated national input from all
230 angling based on the number of licences, estimated number of fishing trips per licence and
231 the bait usage per trip. Loads were distributed across the year based on a typical distribution
232 of angling trips, taking into account the closed season for coarse fishing (Table S7).

233 Other factors that might be important such as the distance of the lake from urban areas and
234 level of active management of the fishery, e.g. fish stocking, are not taken into account but
235 the model user interface allows the user to specify numbers of anglers based on local
236 knowledge.

237 **Bird nutrient excretion (nutrients)**

238 A detailed description of the methodology is provided in S2 of the ESI.

239 Briefly, to estimate loads from this source, the following data was sought from scientific
240 literature, existing models and general internet searches:

- 241 1. Bird populations on lakes in the UK; and
- 242 2. Guano composition and excretion rates from literature for key species.

243 **Assumptions included:**

- 244 • Bird population data of which it was assumed that 50% of the population were based
245 on lakes. This can be replaced by detailed local information by the model user if
246 available.
- 247 • Nutrient inputs estimated from Netherlands data which takes account of food intake,
248 foraging behaviour and digestive performance of water birds (Table S8) (Hahn et al.,
249 2007; 2008).
- 250 • Total bird populations (Table S9) and seasonal distribution (Table S10) combined
251 with survey data to pro rata distribute loads to individual lakes and reservoirs (based
252 on size). Again, default values for number of birds can be replaced by better local
253 data from bird counts and the estimated nutrient input recalculated accordingly.

254

255 **2.3 Allocation of chemical loads to lakes**

256 Processing tools within SAGIS have been developed to allocate chemical loads from the
257 existing and new sectors to each lake are described below.

258 **Inputs from rivers**

259 Where rivers feed directly into lakes and reservoirs, monthly model output from SIMCAT at
260 defined input locations are used to derive riverine loads to these water bodies (Comber et
261 al., 2013). SIMCAT output files are processed by the tools in SAGIS to populate tables in the
262 regional databases with estimated annual and monthly chemical loads and flows which input
263 to each lake water body.

264 **Pumped inputs**

265 Pumped inputs to lakes are calculated from simulated chemical concentrations from
266 SIMCAT at the intake along with information on the licensed abstraction rate and 'hands off
267 flow'. Either annual average or distributed monthly inputs can be calculated. Alternatively, a
268 non parametric distribution for each month can be input which the model then samples.

269 **Direct inputs from the catchment.**

270 In addition to inputs from rivers, there are additional inputs from the local catchment that are
271 not included within the river inputs. Export loads for each chemical substance and sector to
272 the surrounding water bodies are taken from the existing SAGIS load databases. A
273 proportion of these water body loads is allocated to a lake on a *pro rata* basis in relation to
274 the length of river within the lake area compared to the length of river within the water body
275 as a whole as defined by input and output nodes as shown in Figure S2. Any lake inputs
276 such as sewage works in these reaches are excluded from the water quality simulation in
277 SIMCAT to avoid double counting.

278 Lakes which are 'offline' and not connected to rivers, are identified so and a local catchment
279 area for them defined. Diffuse inputs are then calculated *pro rata* based on the proportion of
280 the overall area of the water body that forms the local lake catchment whilst point sources
281 within the area are allocated to the lake

282

283 **Groundwater inputs**

284 In the absence of a national database to provide groundwater inputs, the model allows the
285 user to specify estimated inputs of groundwater to lakes and associated water quality where
286 considered significant.

287 **2.4 SAGIS and SIMCAT model updates**

288 Greater detail on the lake quality modelling is provided in S3 of the ESI. Briefly, Visual Basic
289 for Applications (VBA) was used to model concentrations in lakes, using input data for
290 inflows and chemical loads derived from the export load databases and outputs from the
291 SIMCAT water quality model. Monte Carlo simulation is used to generate distributions of
292 concentrations (Comber et al., 2013). Monthly statistics for water quality are simulated for
293 each chemical substance for each source. In contrast to the SIMCAT river quality model
294 used in SAGIS for rivers, the Monte Carlo simulation is driven by a hydrologically based time
295 sequence of river flows so that wet and dry periods can be taken into account in affecting
296 lake storage and the accumulation of chemicals. This sequence is defined for each regional
297 database based on long term naturalised times series for river flow. The lake simulation is
298 then carried out on a time series basis for a period specified by the user.

299
300 A water balance and chemical mass balance is calculated for inputs and outputs of water
301 and chemical loads, respectively, i.e. pumped inputs, abstractions, groundwater inflows,
302 upstream inflow, rainfall and evaporation. The water quality simulation taking account of
303 sediment interactions is based on a simple lake model developed by Chapra (1997) (see S3
304 of ESI). Apparent settling, re-suspension rates rate and burial rates, are applied to simulate
305 chemical exchanges between the water column and sediment and determine within-lake
306 losses. These can be specified as a range of values that are sampled in the Monte Carlo
307 process and also specified monthly if required (e.g. to represent widely observed seasonal
308 changes in sediment exchange and nitrification). Parameterisation and simulation of nitrogen
309 and phosphorus are described below and represented schematically in the Figure S3 of the
310 supporting information.

311

312 **Lake process parameterisation**

313 To represent within lake processes for phosphorus and nitrogen, parameters for settling rate,
314 sediment release rate and permanent burial rate were modified to produce a better fit with
315 observed data. Removal rates of these chemicals in lakes are known to vary considerably
316 (by orders of magnitude), so this step is necessary as part of the modelling process (Table

317 2). Parameter values were derived based on data obtained for lakes used for the validation
 318 process. For example, for the east of England, Ormesby Broads monthly release rates were
 319 applied to simulate sediment release in the summer. This process is consistent with the
 320 observed data and summer release of sediment phosphorus is known to be important in the
 321 Norfolk Broads (E of England). Similar approaches could be applied to metal and organic
 322 contaminants if such data were available

323

324 **Table 2. Lake processes parameterisation**

Lake	Phosphorus (m/d)			Nitrate (m/d)
	Settling	Release	Burial Rate	Settling Rate
Rutland Water	0.3	0.00175	0.0000175	0.05
Hollowell Reservoir	0.25	0.00175	0.0000175	0.05
Grafham Water	0.125	0.00175	0.0000175	0.05
Costessey Pits	0.25	0.00075	0.00003	0
Ormesby Broad	0.15	0.0007*	0.00003	0.05
Queen Mary Reservoir	0.25	0.00075	0.00003	0
Lower Shustoke Reservoir	0.035	0.00125	0.00003	0.05
Lake Windermere	0.25	0.000375	0.00003	0
Lake Vyrnwy	0.11	0.00075	0.00003	0.025
Sutton Bingham Reservoir	0.15	0.000375	0.00003	0.075
Clatworthy Reservoir	0.15	0.000375	0.00003	0
Kielder Water	0.25	0.00075	0.00003	0

325 *Seasonal pattern of release applied to simulate summer phosphorus sediment release (average
 326 value shown)

327 Settling rates for phosphorus and nitrate tended to be lower in upland reservoirs with lower
 328 levels of nutrient loading. No parameter values were beyond the range of those expected for
 329 normal lake processes. For example, Chapra (1997) gives a range of settling of 0.05 to 0.6
 330 m/day. Parameter values will also reflect inaccuracies in the inputs from the river model,
 331 such as if nitrate concentrations are over predicted by the river model, the nitrate loss rate in
 332 the lake will be increased to compensate. Future development of the tool may make it
 333 possible to derive default rate values for different types of lakes e.g. upland reservoirs,
 334 pumped storage, lowland etc.

335 Using the above parameters, each model simulation can be carried out for each month for a
336 number of Monte Carlo ‘shots’ with different starting conditions for lake volume also specified
337 by the user, i.e. based on historical patterns of lake volume. Alternatively, in the absence of
338 this information, the lake is assumed to be full at the start of each month. Within this period
339 input flows and loads are sampled from distributions generated from averages and standard
340 deviations in loads taken from SIMCAT output, i.e. for river inflows, and the export load
341 databases. The outputs of each month define the starting conditions from the next month
342 along with a new starting volume. The simulation is continued for a number of years defined
343 by the user until a steady state seasonal pattern of concentrations is reached. For larger
344 lakes with longer retention times, a longer simulation period is required before stable
345 conditions are achieved. Settling rates, release rates and starting sediment concentrations
346 can also be sampled between bounds set by the user, initially based on literature values.

347

348 Monthly outputs for total simulated concentrations and concentrations associated with each
349 sector are generated for plotting either as monthly averages or long term monthly time series
350 results. The latter provides information on how long the lake takes to reach a new steady
351 state of seasonal pattern in concentrations (Figure S4).

352

353 A number of changes were carried out on the existing SIMCAT water quality model to
354 support the lake simulations:

355 1. New boundary features called lake inflow and lake outflow and

356 2. A new input feature downstream of the lake.

357 The SIMCAT simulation effectively ends at the inflow boundary and resumes at the outflow
358 after collecting output from the lake model.

359 **2.5 Model validation**

360 To validate the model, SAGIS-Lake output was compared with observed water quality data
361 for 11 lakes, selected to be representative of different lake types (pumped storage, natural
362 refill and bankside storage) and geographical regions. A combination of chemical
363 substances were selected for the comparison (nutrients, one metal and one organic); Total
364 Phosphorus, nitrate, benzo-ghi-perylene and copper. In the results below we show outputs
365 from three representative lakes with each of the main types of hydrological control; pumped
366 storage, natural refill and bankside storage.

367
368
369
370
371
372
373
374
375
376
377
378

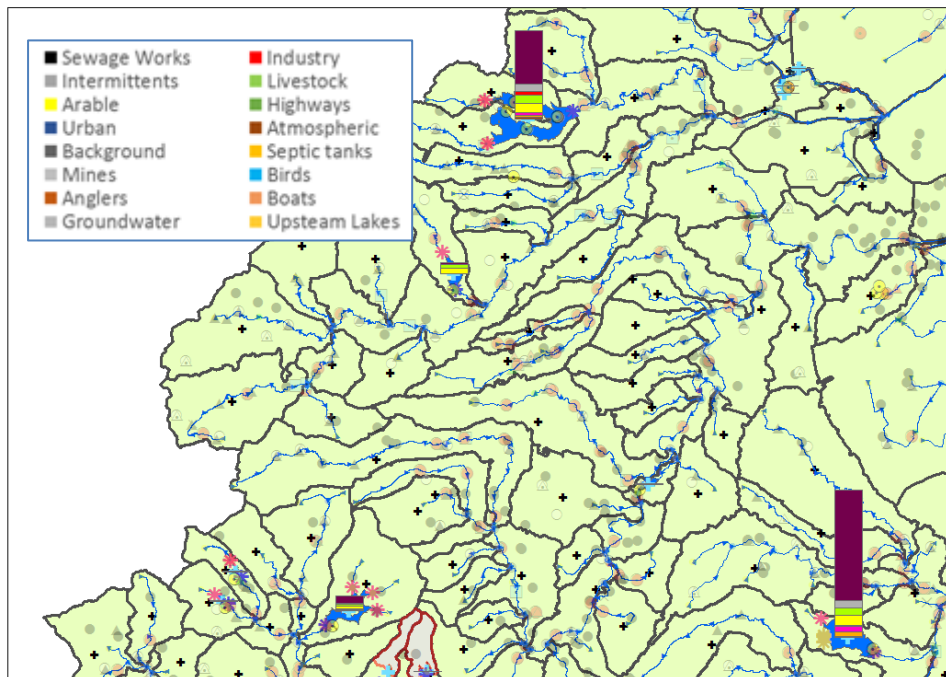
379
380
381
382
383
384
385

386

3. RESULTS

In general the lake model performed well with regard to comparison between model output and observed data and consistency of the source apportionment output with the nature of the lake catchments and upstream chemical sources for the 11 lakes used for validation purposes, across all types of hydrology (UKWIR 2013). For illustrative purposes one of each hydrological type is provided here, Rutland Water (pumped storage), Cotessey Pits (bankside storage) and Ormesby Broad (natural refill). Model performance was shown to be influenced by the accuracy of the SAGIS river model that provides inputs to the lake model and accuracy of the representation of within lake processes. The river model is most accurate for phosphorus and less accurate for the other chemicals.

Annual and monthly input loads to each lake and reservoir from all sectors and source types, i.e. river, direct, local and pumped, are compiled in tables in the regional databases. Mapping tools have been developed in SAGIS to present source apportionment plots for these inputs in the form of pie and stacked bar charts (e.g. Figure 2.) as well as being able to show the relative contribution of each upstream sewage works to the total phosphorus loads (e.g. Figure S5 for Rutland Water). This is based on the functionality in SIMCAT to 'track' the contribution of individual sources to sector loads.



398 **Figure 2. Sector bar charts for sector input loads to lakes**

399

400 Model output for simulated water quality was compared with observed data for the selected
401 lakes listed below.

- 402
- **Pumped Storage:** Rutland Water (Anglian Water)

403

 - **Natural Refill:** Ormsby Broads (Trinity Broads, Anglian Water)

404

 - **Bankside Storage:** Costessey Pits (Anglian Water)

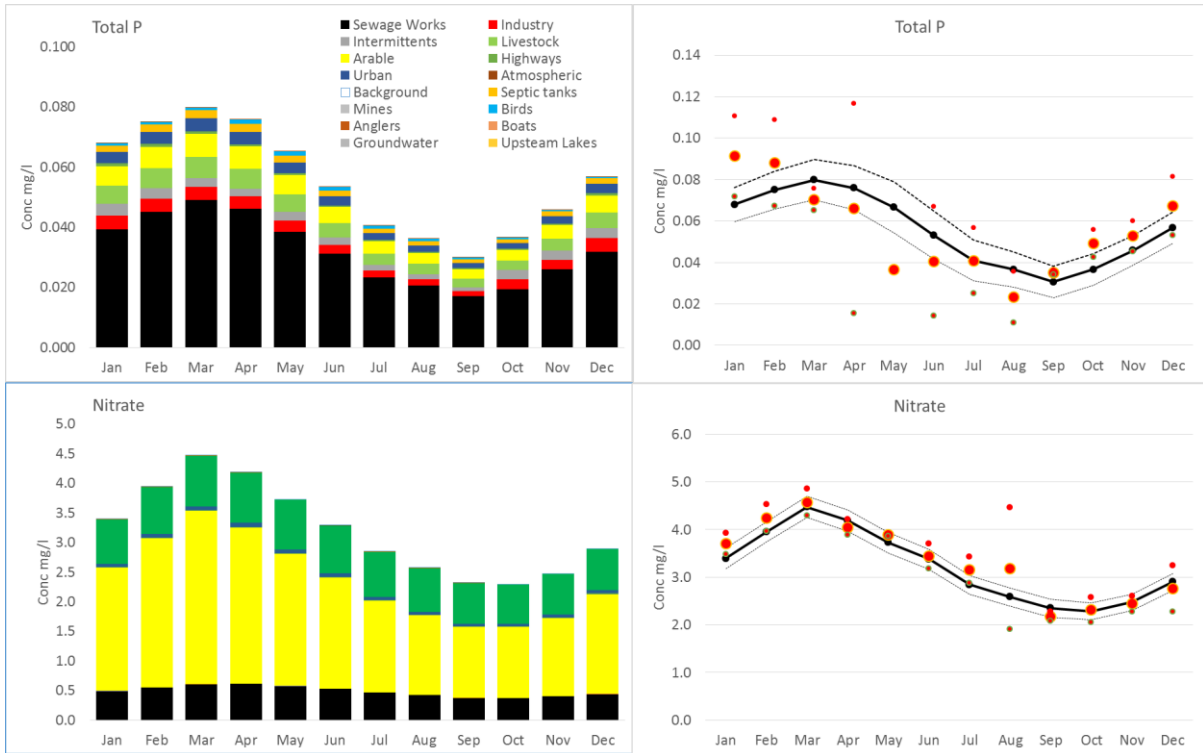
405 These lakes were chosen with the aim of representing a variety of lake types with regards to
406 location, hydrology and the nature of the chemical inputs. For illustrative purposes, one
407 example of each is provided here where monitoring data were available. Further figures are
408 provided in the Electronic Supporting Information (Figures S6 to S11) and a full validation
409 report elsewhere (UKWIR, 2013).

410 The following chemical substances were selected for the comparison (nutrients, one metal
411 and one organic).

- 412 • Total Phosphorus
- 413 • Nitrate
- 414 • Copper
- 415 • Benzo-ghi-perylene

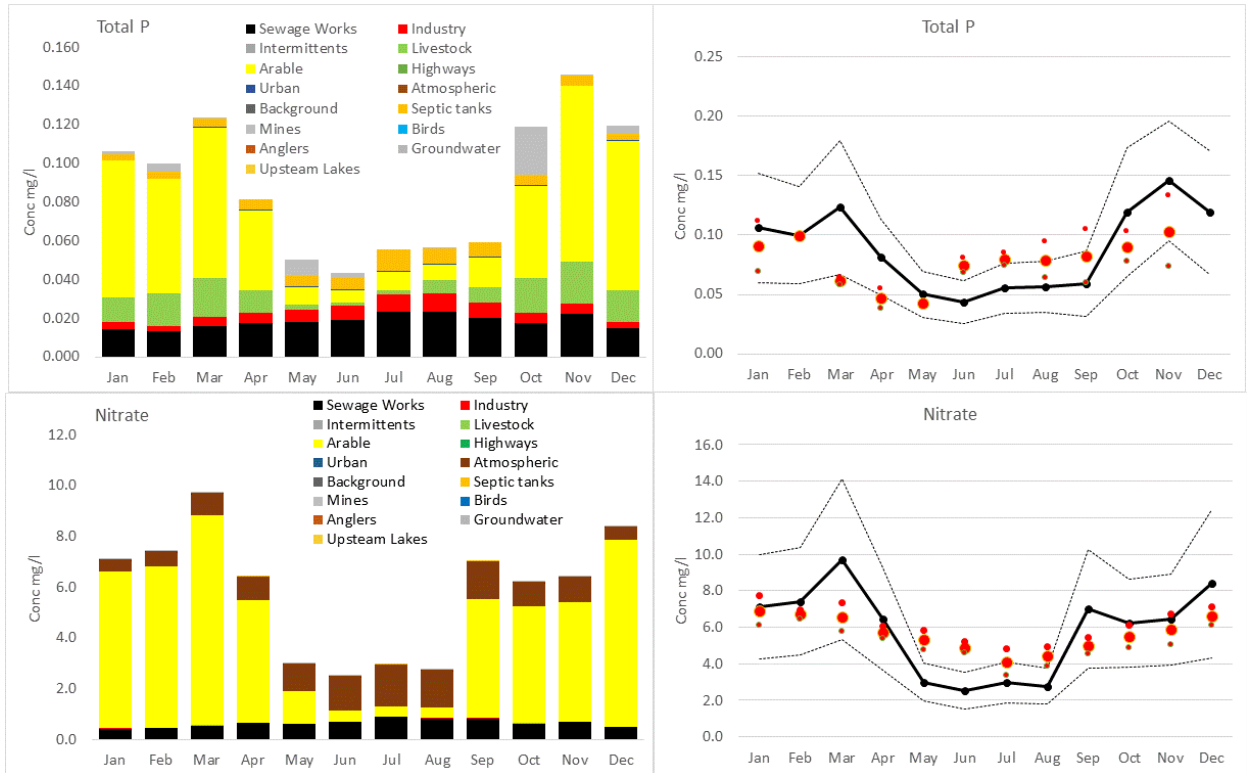
416 Source apportionment information and a comparison between model output and observed
417 data are provided in Figures 3 and 4. For phosphorus and nitrate, lake model parameters
418 were modified to best represent within lake processes and improve the fit with observed data
419 (Table 2) but other chemicals were assumed to behave conservatively and no loss
420 parameters were applied from the water column. For Ormesby Broad, groundwater inflows
421 are based on outputs from regional EA groundwater models and monitoring data. Source
422 apportionment for phosphorus at Rutland Water shows a wide range of sources, dominated
423 by WwTW effluent. Bird and angler inputs of nutrients, new to the lake model, were shown to
424 be only a relatively small contribution to the overall load. Nitrate loads on the other hand
425 were made up almost entirely of arable sources and atmospheric deposition. Lowest
426 concentrations were predicted and observed during the drier summer months. For Cotessey
427 Pits a similar seasonal trend in nutrient concentration is observed, but agricultural inputs
428 dominate both nitrate and phosphorus reflecting a more rural catchment. Ormesby Broad
429 nutrient sources were dominated by groundwater inputs (Figure S11). Unlike the river SAGIS
430 version, the model has the capacity to include bird and angling inputs for nitrogen and
431 phosphorus from roosting and over wintering fowl and the use of nutrient-rich groundbait
432 used by anglers. Based on the assumptions used these sources appear to be insignificant
433 for these case studies sites, although it is not possible to rule out greater loads where the
434 density of angling and bird populations are greater. Given the lack of detailed databases
435 available regarding lake specific activities it is anticipated that users would be reliant on
436 overwriting the default values for actual angling statistics and bird populations where they
437 may be considered significant.

438 The seasonal trends in the lakes are related to the changes in input concentrations, input
439 flows and retention time in the lake through the year. Nutrient uptake and release from the
440 sediment is modelled as part of the simulation and can be varied through the year if there is
441 an indication of phosphorus sediment release in the summer. Apart from Ormesby Broad
442 where this is the case, the loss rates are constant throughout the year in the examples
443 shown.



444

445 **Figure 3. Modelled (black line) and observed (orange dots) monthly average total**
 446 **phosphorus and nitrate concentrations (+/- standard deviation) in Rutland Water with**
 447 **monthly source apportion**

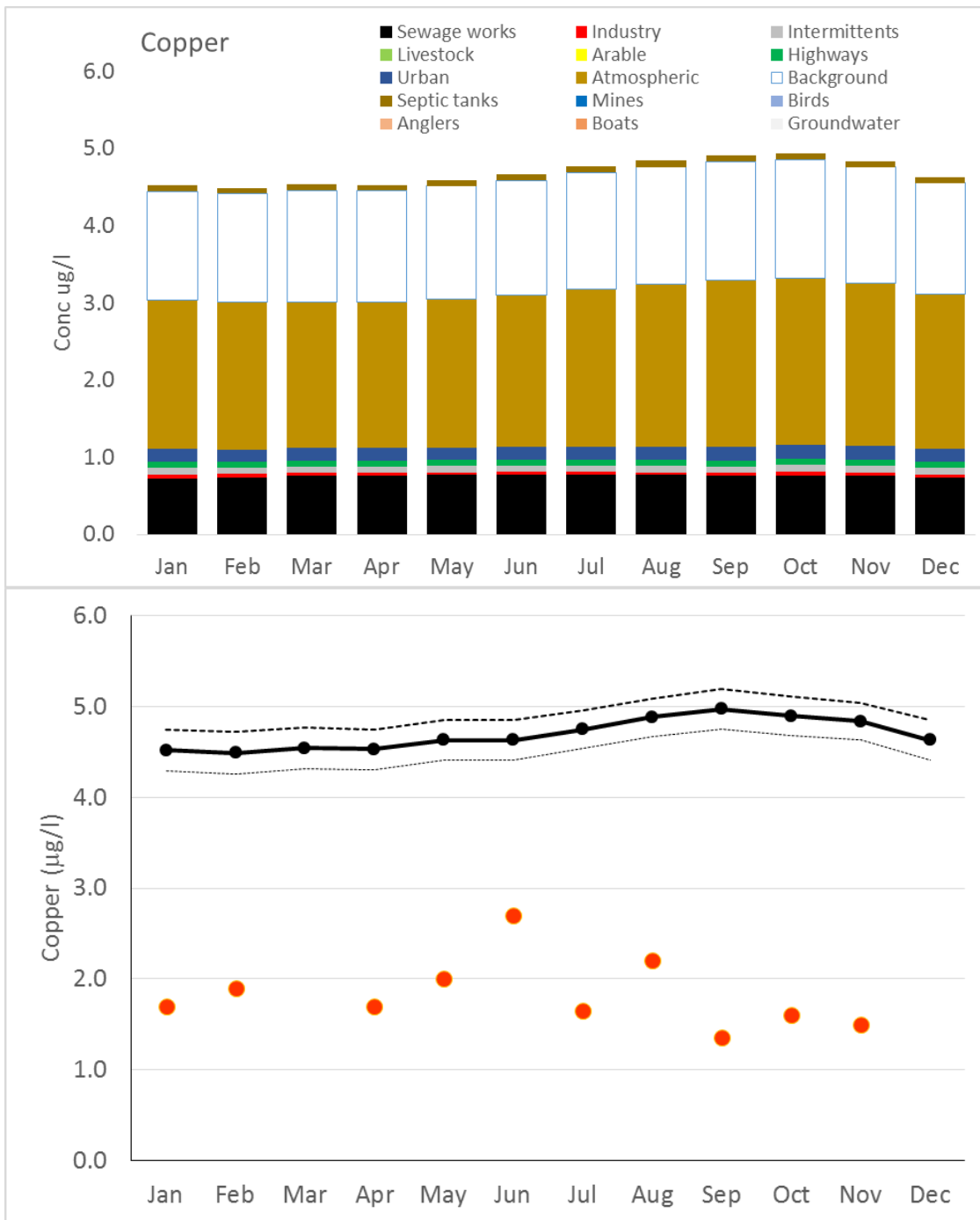


448

449 **Figure 4. Modelled (black line) and observed (orange dots) monthly average total**
 450 **phosphorus and nitrate concentrations (+/- standard deviation) in Cotessey Pits with**
 451 **monthly source apportionment**

452
453 For copper in Rutland Water (Figure 5) concentrations are predicted to be between 4 and 5
454 µg/l. There are no recent data with which to compare the simulations so for illustrative
455 purposes data from 2006 (one or two per month with the exception of February and
456 December) are provided. Observed concentrations are of the same order but lower; in the
457 order of 2 µg/l, dominated by WwTW effluent, atmospheric deposition and background
458 geology. Neither observed or predicted data suggested any significant seasonal variations in
459 copper concentrations reflecting the low levels of inputs and broad range of sources.
460 Performance of the river model for copper was generally good but background geology
461 inputs for copper were derived from relatively sparse FOREGS data (Comber et al., 2013)
462 potentially impacting on model accuracy for static water bodies. Furthermore, it should also
463 be noted that Rutland's catchment is large, most of the county of Northamptonshire, and so
464 a factor of two difference between predicted and observed copper data may be considered
465 acceptable based on a lack of calibration, uncertainty about transport and limited datasets.

466 Observed data for benzo(ghi)perylene were absent, which means it is not possible to
467 validate the model, but predictions suggest the presence of low ng/l levels with little seasonal
468 variation, with urban runoff, atmospheric deposition and background contributions from
469 contaminated soil being the main sources as would be expected from such a persistent
470 contaminant largely released through combustion processes over hundreds of years (Figure
471 S7).



472

473

474 **Figure 5. Modelled (black line) and observed (for 2006 data - orange dots) monthly**
 475 **average total copper concentrations (+/- standard deviation) in Rutland**
 476 **Water and monthly source apportionment.**

477

478

479 **4. DISCUSSION**

480 A validation report comparing model results is provided elsewhere (UKWIR, 2013). The
481 validation process aimed to identify any: i) calculation errors; ii) systematic errors that require
482 modification related to the methodologies and their associated assumptions; iii) underlying
483 uncertainties that may affect the performance of the tool and iv) possible performance
484 improvements. Model outputs were compared with observed data for 11 lakes, selected to
485 be representative of different lake types (pumped storage, natural refill and bankside
486 storage) and geographical regions. Total phosphorus, nitrate, benzo(ghi)perylene and
487 copper were selected to represent the main type of chemicals for which the model would
488 likely to be used and which have widely differing sources and chemical characteristics. The
489 lake model performed well with regard to predicted versus observed concentrations and
490 consistency of the source apportionment output, with the nature of the lake catchments and
491 upstream chemical sources. Model outputs were very much controlled by the accuracy of the
492 SAGIS river model, that provides inputs to the lake model, and accuracy of the
493 representation of within lake processes. The river model was most accurate for nutrients and
494 less accurate for metals and organic substances because the latter groups have more
495 limited data that form the basis of the source loads.

496 It is anticipated the SAGIS Lake model will be used in the following ways:

- 497 • **Water quality planning for lakes** - The lakes tools provide the first national
498 modelling platform for water quality in lakes linked to a national river planning tool.
499 This will now allow the extension of river basin planning to include lakes. Where non-
500 compliance with water quality objectives are observed within lentic water bodies, it is
501 now possible to identify the main sources contributing to this exceedance.
- 502 • **Improve river quality simulations** - By taking into account the influence of lakes on
503 river quality and flow by modifying flow and providing natural purification of
504 chemicals.
- 505 • **Reporting** - SAGIS provides a range of visualisation options for chemical inputs and
506 predicted within-stream and lake concentrations which could be readily used for
507 reporting of pressure characterisation and compliance.
- 508 • **Testing of measures** - SAGIS provides a means to assess the efficacy of
509 remediation options related to each source and sector.

- 510 • **Catchment management stakeholder engagement** – The variety of sources and
511 chemicals covered, linked to GIS mapping tools enhances the options for stakeholder
512 engagement.
- 513 • **Identify further monitoring and research** – The use of national datasets highlights
514 areas of uncertainty in the estimation of source apportionment, thereby providing a
515 focus for targeting resources on improving source data or the methodologies used to
516 create the export coefficient databases.

517 As a national tool, no calibration or model conditioning has been carried out on the tool at
518 present and default values have been used in many cases; for example for river travel times
519 and decay rates and lake settling and release parameters. Depending on the intended
520 purpose of the model, it is important to review the accuracy of the initial output in relation to
521 the questions being asked and consider the value of improving the input data and
522 undertaking model calibration. For some chemicals and catchment conditions, complex
523 hydrological interactions and dynamics (e.g. stratification) and active management
524 (pumping) of lakes have a strong influence on water quality. In these circumstances, other
525 models such as time series models or hydrodynamic mixing models may be more suitable
526 than SAGIS or SAGIS output may be used to provide improved inputs to complex models.

527 In considering output from the tool, it is important to understand uncertainty in the input data
528 and representation of processes that will inevitably result in errors in the model output and
529 differentiate this from systematic errors that may result from calculation or assumption based
530 errors that can be corrected. Effects of uncertainty in the input data on the outputs are likely
531 to be more evident at the local scale because the sample size of the data will be smaller.
532 SAGIS in its current form provides outputs based on national data sets and for subsequent
533 enhancement it will be important to identify key areas that would benefit from improvement in
534 the underlying data and model refinement at the local scale.

535 Although uncertainty between observed and predicted concentrations can be generated for
536 individual lakes, owing to the variation in sources of the uncertainty it is not possible to
537 provide quantify them as a whole. Consequently key sources of uncertainty related to input
538 data, process representation and potential improvements are discussed below in Table 3
539 (uncertainties in the existing river model, as reported previously (Comber et al., 2013) and
540 available in the supporting information Table S11).

541

542

543 **Table 3. Source of uncertainty in the model**
 544

Uncertainty Source	Description of uncertainty	Likely impact on model outputs and ways to reduce uncertainty
Direct Data Inputs		
Groundwater	Groundwater flows and water quality are provided as user defined inputs. Information is rarely available to quantify groundwater inputs so a default of no inputs is applied for most lakes.	Groundwater inputs are likely to be underestimated (more significant in lowland than upland areas).
Pumped inflows and outflows to lakes	Pumped flows are based on the abstraction data in SIMCAT and information in the Water Resources GIS on the abstraction locations that operate as intakes to reservoirs. This information is incomplete and some lake abstractions will not be included in the SIMCAT databases e.g. if lakes are not located on the river polyline.	Review of abstraction data for lakes and input of better local data required where likely to be significant.
Inputs to SIMCAT		
Lake modelling	The lake model used in SAGIS-Lakes is a simple zero dimensional tank model with no horizontal or vertical spatial differentiation. Uneven mixing between different areas of a lake or vertical stratification are, therefore, not taken into account so performance is likely to be worse in lake where these processes are important.	
History of lake loading	Within lake concentrations are a reflection of historical loading of chemicals, e.g. phosphorus, whereas the SAGIS model only accounts for current loading.	Assess history of loading when interpreting outputs
Inputs from other models		
Wildfowl	National wildfowl numbers and distributed across the lakes based on the perimeter length. Lakes will however vary and local data is required if wildfowl is likely to contribute significantly to input loads.	Inputs will tend to be overestimated in lakes with small bird populations and underestimated if bird populations are large.
Anglers	National angler numbers and distributed across the lakes based on the perimeter length. Again, local data is required if angling is likely to contribute significantly to input loads, which is most likely to be the case in small lakes.	Inputs will tend to be overestimated in smaller, heavily fish stocked lakes.

545

546 Such a large spatial model such as SAGIS clearly requires national datasets with the
 547 aim of providing a consistent approach to national and regional water quality
 548 planning. This is obviously at the expense of localised sources which may impact on
 549 water quality. Furthermore, hydrological connectivity between lakes and the
 550 catchment is often complex and operational management may substantially influence

551 water quality, e.g. the timing of pumping to reservoirs. The export load databases for
552 birds and anglers are, for example, derived from regional information and local more
553 detailed information will may be available, e.g. bird numbers on lakes which could be
554 utilised in future. This is particularly important for lakes because of their hydrological
555 complexities. Furthermore, lakes often have relatively small local catchments that
556 strongly influence water quality. At this spatial scale it is often essential to apply
557 detailed local knowledge to develop a reliable understanding of key influences and
558 processes. The functionality of the lakes model in SAGIS ensures that more detailed
559 local information can be easily accommodated by the model user.

560 **5. CONCLUSIONS**

561 The SAGIS model represents the first comprehensive source apportionment tool to
562 be developed on a national scale for such a wide variety of chemicals and sources to
563 lakes and reservoirs. To meet ever more stringent standards multiple interventions
564 will be required to reduce discharges from point and diffuse sources. SAGIS will
565 assist regulators in making effective decisions regarding how best to meet
566 challenging water quality targets by identifying the predominant source of a chemical
567 which can now be extended to management of lakes and reservoirs.

568

569 SAGIS provides a flexible framework and research to improve the model datasets
570 and representation of processes is ongoing. This process is supported by the
571 regulators and industry within the UK to drive consistent, fair attributed, cost-effective
572 water quality improvements.

573

574 **Acknowledgements**

575 The authors would like to thank UKWIR and Mr Roger Trengrove who acted as
576 project manager, the Environment Agency including Mr Tony Heaney, the Scottish
577 Environmental Protection Agency and the Water Companies of the UK for funding
578 this project and supplying the data required to build the model. Other organisations
579 who have contributed data include the Highways Agency, Defra, ADAS, CEH and
580 FOREGS. We would also like to thank Dr Tony Warn for providing an updated
581 SIMCAT model.

582

583 **Supporting Information**

584 Tables of databases, default values and information used to develop the model are
585 provided in the supporting information in addition to comparison data for model
586 outputs versus observed data.

588 **5. REFERENCES**

- 589 Abbott M.B., Bathurst J.C., Cunge J.A., O'Connell P.E., and Rasmussen J. An
 590 introduction to the European hydrological system: Systeme Hydrologique European,
 591 'SHE.' 2: Structure of a physically-based, distributed modeling system. *J. Hydrol.*,
 592 1986;87:61 –77.
- 593 Ambrose, R.B., Wool, T.A., and Connolly J.P. WASP4, A Hydrodynamic And
 594 Water Quality Model. Model Theory, User's Manual, and Programmer's Guide, US
 595 Environmental Protection Agency, Athens, Ga, USA, 1988.
- 596 Arlinghaus, R. and Mehner, T. Socio-economic characterisation of specialised
 597 common carp (*Cyprinus carpio* L.) anglers in Germany, and implications for inland
 598 fisheries management and eutrophication control. *Fish. Res.* 2003;61:19-33.
- 599 Arlinghaus, R. and Niesar M. (2005) Nutrient digestibility of angling baits for
 600 carp, *Cyprinus carpio*, with implications for groundbait formulation and eutrophication
 601 control. *Fish. Man. & Ecol.* 2005;12(2):91–97.
- 602 Chapra S.C. (1997) Surface Water Quality Modelling. McGraw Hill International.
 603 ISBN 13: 9780070113640.
- 604 Comber S., Smith R., Daldorph P., Gardner M., Constantino C. and Ellor B.
 605 Development of a chemical source apportionment decision support framework for
 606 catchment management. *Env. Sci. and Tech.*, 2013;47(17):9824-9832.
- 607 Dardni. 'Phosphorus inputs to Lough Neagh'. Presentation. Department of
 608 Agriculture and Rural Development, Northern Ireland.
 609 http://www.afsni.ac.uk/Research/P_sources/sld012.htm. 2007.
- 610 Davison, P. S., Withers, P. J. A., Lord, E. I., Betson, M. J. & Stromqvist, J.
 611 PSYCHIC - A process-based model of phosphorus and sediment mobilisation and
 612 delivery within agricultural catchments. Part 1: Model description and
 613 parameterisation. *J. Hydrol.* 2008;350:290-302.
- 614 EA (2010) Environment Agency. Public Attitudes to Angling 2010 - A survey of
 615 attitudes and participation in England & Wales. [http://publications.environment-](http://publications.environment-agency.gov.uk/PDF/GEHO0610BSOS-E-E.pdf)
 616 [agency.gov.uk/PDF/GEHO0610BSOS-E-E.pdf](http://publications.environment-agency.gov.uk/PDF/GEHO0610BSOS-E-E.pdf)
- 617 EEA (2005) European Environment Agency. Source apportionment of nitrogen
 618 and phosphorus inputs into the aquatic environment. European Environment Agency,
 619 Report No7/2005. ISSN 1725-9177, 2005.
- 620 EEA (2015a) EEA European Environment Agency, Lake information.
 621 [http://www.eea.europa.eu/themes/water/european-waters/lakes. 2015](http://www.eea.europa.eu/themes/water/european-waters/lakes.2015), Accessed 24th
 622 November 2015.
- 623 EEA (2015b) EEA European Environment Agency, Lake information.
 624 <http://www.eea.europa.eu/themes/water/european-waters/reservoirs-and-dams>.
 625 2015b, Accessed 24th November 2015.
- 626 EU (1976) European Union: Directive 76/464/EEC of 4 May 1976 on pollution
 627 caused by certain dangerous substances discharged into the aquatic environment of
 628 the Community. Codified as 2006/11/EC.
- 629 EU (2000) European Union Directive 2000/60/EC of the European Parliament
 630 and of the Council of 23 October 2000 establishing a framework for Community
 631 action in the field of water policy.
- 632 EU (2008) European Union: Directive 2008/105/EC of the European Parliament
 633 and the Council on environmental quality standards in the field of water policy,
 634 amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC,
 635 84/156/EEC, 84/491/EEC, 86/280/EEC and amending Directive 2000/60/EC.
 636 Published in the Official Journal on 24 December 2008.
- 637 Gough D.I. Incremental stress under a two-dimensional artificial lake, *Can. J. of*
 638 *Earth Sci.* 1969;6(5):1067–1075.

639 Hahn, S., Bauer, M. And Klaassen, M., 2007. Quantification of allochthonous
640 nutrient input into freshwater bodies by carnivorous waterbirds. *Freshwat. Biol.*
641 2007;52(12):2421-2433.

642 Hahn, S., Bauer, M. And Klaassen, M. Quantification of allochthonous nutrient
643 input into freshwater bodies by herbivorous waterbirds. *Freshwat. Biol.*
644 2008;53(1):181-193.

645 Leonard J. and Crouzet P. Lakes and reservoirs in the EEA region. Topic Report
646 1/1999. European Environment Agency, November 1998.

647 Li A., Jank J.K. and Scheff P.A. Application of EPA CMB8.2 model for source
648 apportionment of sediment PAHs in Lake Calumet, Chicago. *Env. Sci Tech.*
649 2003;37(13):2958-65.

650 Lindgren G.A., Wrede S., Seibert J. and Wallin M. Nitrogen source
651 apportionment modeling and the effect of land-use class related runoff contributions.
652 *Nordic Hydrol.* 2007;38(4-5):317-331.

653 Lord, E.I. & Anthony, S.G. NEAP-N: A modeling framework for evaluating nitrate
654 losses at national and catchment scales. *Soil Use & Man.* 2000;16:167-174.

655 MacKay D. and Hickie B. Mass balance model of source apportionment,
656 transport and fate of PAHs in Lac Saint Louis, Quebec. *Chemosphere* 2000;5:681-
657 92.

658 Manny, B.A., Johnson, W.C. and Wetzel, Nutrient additions by waterfowl to lakes
659 and reservoirs: predicting their effects on productivity and water quality.
660 *Hydrobiologia* 1994;279/280:121-132.

661 Marion, L., Clergeau, P., Brient, L. and Bertru, G., The importance of avian
662 contributed nitrogen (N) and phosphorus (P) to Lake Grandlieu, France.
663 *Hydrobiologia* 1994;279/280:133-147.

664 Nashoug, Nashoug (ed.) Vannkvaliteten i Mjøsa—før og nå Mjøsovervåkingen
665 gjennom 25 år. Report from Styringsgruppa for overvåking av Mjøsa.1999.
666 [http://www.fylkesmannen.no/fmt_liste_kort.asp?tgid=5374&gid=5383&amid=1003083](http://www.fylkesmannen.no/fmt_liste_kort.asp?tgid=5374&gid=5383&amid=1003083&g5375=x&g5374=x&Nasr A and Bruen M. Physically-based, distributed, catchment modelling for estimating sediment and phosphorus loads to rivers and lakes: issues of model complexity, spatial and temporal scales and data requirements. National Hydrology Seminar, 2004.)
667 &g5375=x&g5374=x&Nasr A and Bruen M. Physically-based, distributed, catchment
668 modelling for estimating sediment and phosphorus loads to rivers and lakes: issues
669 of model complexity, spatial and temporal scales and data requirements. National
670 Hydrology Seminar, **2004**.

671 SFT (2005). Home page: Miljøstatus i Norge, Fosfor til Mjøsa.
672 http://www.miljostatus.no/oppland/tema/Vannkvalitet/fosfor_til_mjosa.htm.

673 Shanahan P., Henze M., Koncsos L., Rauch W., Reichert P., Somlyody L.,
674 Vanrolleghem P. River water quality modeling: II. Problems of the art. *Wat. Sci. Tech.*
675 **1998**;38(11):245 –252.

676 UKWIR (2012) Chemical Source Apportionment under the WFD (12/WW/02/3).
677 Final report for UK Water Industry Research, 1 Queen Annes Gate, London, ISBN: 1
678 84057 637 5.

679 USEPA (2007). Environment Fluid Dynamics Code (EFDC) water quality model.
680 EPA version 1.01; released September **2007**. U.S. Environmental Protection Agency,
681 Ecosystems Research Division, Center for Exposure Assessment Modeling (CEAM),
682 Athens, USA.

683 UKWIR (2013) WFD Requirements for lakes, transitional and coastal waters
684 source apportionment for nutrients and priority chemicals. Final Report with
685 Appendices. Report Ref. No. 13/WW/02/6.

686 USEPA (2017) United States Environmental Protection Agency;
687 <https://www.epa.gov/greatlakes/great-lakes-facts-and-figures>; **2017**, accessed
688 13/3/17.

689 Vassiljev and Stålnacke. Statistical modelling of riverine nutrient sources and
690 retention in the Lake Peipsi drainage basin'. Paper presented at the Diffuse Pollution
691 Conference, Dublin, **2003**. http://www.ucd.ie/dipcon/docs/theme10/theme10_09.PDF.

692 WeBS (2014) The Wetland Bird Survey, 2013/14. Available at:
693 <http://app.bto.org/webs-reporting/>; accessed 24 Nov 2015.

694 Yih S.M. and Davidson B. Identification in non-linear distributed parameters
695 water quality models. *Water Resour. Res* 1975;11(5):693-704.
696
697