2018-05-01

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

Comber, Sean

http://hdl.handle.net/10026.1/10409

10.1016/j.scitotenv.2017.11.313
Science of the Total Environment
Elsevier BV

For published copy please see: Science of the Total Environment 622-623**:96-105 01 May 2018: doi.org/10.1016/j.scitotenv.2017.11.313**

Development of a Chemical Source Apportionment Decision Support Framework for Lake Catchment Management

Sean D W Comber*, Russell Smith[1], Peter Daldorph[2], Michael J Gardner[2], Carlos Constantino[2], Brian Ellor[3]

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

[1] Westcountry Rivers Trust, Rain-Charm House, Stoke Climsland, Cornwall, PL17 8PH
[2] Atkins Limited, Western House, Peterborough Business Park, Peterborough, PE2 6FZ
[3] UK Water Industry Research (UKWIR), 3rd Floor, 36 Broadway, London SW1H 0BH.

Research Article – Environmental Modelling

Abstract Art

Key words

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

ABSTRACT

Increasing pressures on natural resources has led to the adoption of water quality standards to protect ecological and human health. Lakes and reservoirs are particularly vulnerable to pressure on water quality owing to long residence times compared with rivers. This has raised the question of how to determine and to quantify the sources of priority chemicals (e.g. nutrients, persistent organic pollutants and metals) so that suitable measures can be taken to address failures to comply with regulatory standards. Contaminants enter lakes waters from a range of diffuse and point sources. Decision support tools and models are essential to assess the relative magnitudes of these sources and to estimate the impacts of any programmes of measures. This paper describes the development and testing of the Source Apportionment Geographical Information System (SAGIS) for future management of 763 lakes in England and Wales. The model uses readily available national data sets to estimate contributions of a number of key chemicals including nutrients (nitrogen and phosphorus), metals (copper, zinc, cadmium, lead, mercury and nickel) and organic chemicals (Polynuclear Aromatic Hydrocarbons) from multiple sector sources. Lake-specific sources are included (groundbait from angling and bird faeces) and hydrology associated with pumped inputs and abstraction. Validation data confirms the efficacy of the model to

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

1. INTRODUCTION

Lakes and reservoirs serve as vital sources of drinking water and support valuable ecosystems across the globe. For example, there are more than 500,000 natural lakes larger than 0.01 km^2 (1 ha) in Europe alone. Approximately 80 to 90% of these are small with a surface area of between 0.01 and 0.1 km^2, with only around 16,000 having a surface area exceeding 1 km^2 (EEA, 2015a). There are currently 7,000 large dams in Europe (defined as having greater than 1,000,000m^3 capacity) (EEA, 2015b). Reservoirs and lakes hold approximately 32,000 million m^3 of drinking water across Europe representing around 20% of total supply (Leonard and Crouzet, 1998). In North America the Great Lakes (Erie, Michigan, Huron, Ontario, Superior) contain one fifth of the world's freshwater and 84% of the water supply for the United States of America (USA) and Canada (USEPA, 2017). Other than drinking water, lakes and reservoirs supply power via hydroelectric schemes, augment river flows, provide recreation and are valuable habitats for water fowl and species of fish and invertebrates. There are significant pressures on water quality from numerous sources including urban and highway runoff, industrial discharges, mining and in some cases atmospheric deposition (Comber et al., 2013). Eutrophication is a major source of concern regarding water quality, related mainly to inputs of nitrogen and phosphorus from agriculture and sewage effluents. In Spain, for example 33% of reservoirs were identified as mesotrophic, 27% eutrophic and 10% hypertrophic (Leonard and Crouzet, 1998). Toxic algal blooms occur frequently with the example of suspension of water abstraction from Lake Erie in 2014, affecting half a million people (USEPA, 2017). Loss of potential amenity value has seen extensive steps being taken under European Union (EU) legislation to address these issues.

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

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

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

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

Once discharged to a lake catchment, any given chemical will be subject to dilution and undergo various biogeochemical processes, effects that might both be incorporated into a model. However, whereas there have been published reports on load apportionment to lakes and reservoirs and models to determine concentrations within such water bodies by taking account of physico-chemical and biochemical processes (e.g. relationships between biological growth rate and nutrient availability, sunlight and temperature, and phytoplankton and the growth rate of zooplankton; Gough, 1969; Yih and Davison, 2008), few models have attempted to combine the two. More recent water quality models which simulate lakes specifically include the United States Environmental Protection Agency (USEPA) WASP (Ambrose et al., 1988), and QUAL2E models (Shanahan et al., 1998), the MIKE3 model developed by the Danish Hydraulics Institute (DHI), and the Systeme Hydrologique European (SHE) (Abbott et al., 1986). At a larger spatial scale, catchment models include BASINS (Nasr and Breun, 2004) and the Environmental Fluid Dynamics Code (EFDC) which is a multifunctional surface water modelling system, which includes hydrodynamic, sediment-contaminant, and eutrophication components. EFDC has been applied to over 100 water bodies including lakes and reservoirs, and is a state-of-the-art hydrodynamic model that can be used to simulate aquatic systems in one, two, and three dimensions (USEPA, 2007).

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

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

Lakes and reservoirs are subject to additional and significant chemical sources compared with rivers, particularly nutrients from birds and angling. Birds are recognised as significant sources of nutrients to some lakes (Manny et al., 1994; Marion et al., 1994; Hahn et al., 2008). Fishing is a popular pastime in the UK. Approximately 9% of the population in England and Wales have been freshwater fishing as reported by the Environment Agency in 2010 (EA, 2010). Ground bait, comprising of ingredients such as maize, fish meal, milk protein and semolina, is commonly used, particularly for coarse fishing and nutrient inputs to water bodies from the use of ground bait may present a potential threat to water quality where angling intensity is high (Arlinghaus and Mehner, 2003). These studies have estimated German mean annual gross P-input of 1018 g P angler^{-1} due to ground bait.

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

2. METHODS

The SAGIS modelling framework integrates data from multiple sources (Comber et al., 2013), point source loads are expressed as mean and standard deviation of annual average (or monthly if available) concentrations and flow, with diffuse sources input as mass per year, or month, per km^2. Owing the large size of databases used, water bodies of England, Wales and Scotland are broken up into 18 regional databases. A common map projection was used for all databases and mapping based on a 1km^2 grid. Such GIS mapping calibration and validation has been undertaken as part of previous projects associated with the hydrological and diffuse source components for sources of chemicals to rivers (Comber et al., 2013). Detailed information regarding the methodologies used to calculate loads for each source is provided elsewhere (UKWIR, 2012). However, a brief description of the data and method used to derived load estimates is provided in the following section.

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

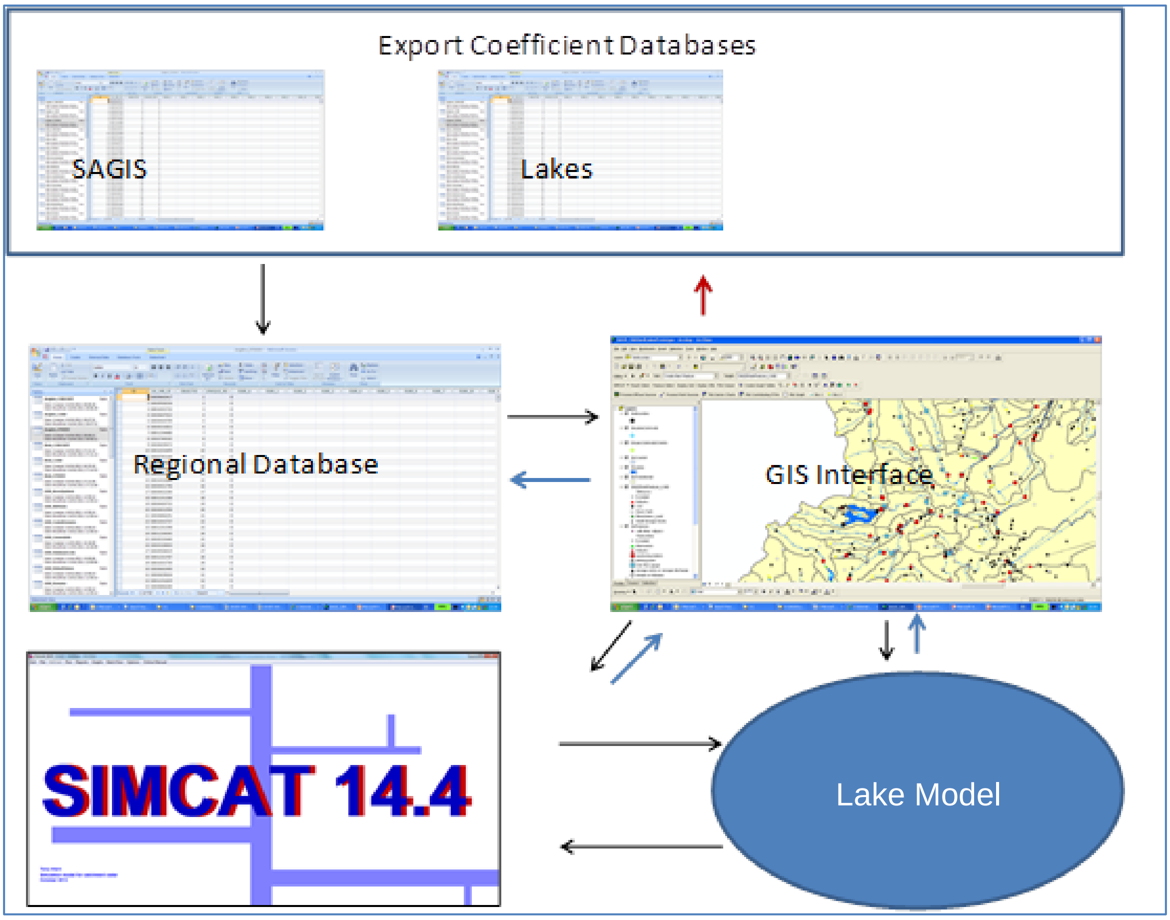

Figure 1. Schematic diagram for SAGIS tool structure

2.1 Inputs already present within the SAGIS rivers model

Sources are represented within a Microsoft Access™ database either as a point source with an X and Y, UK national grid location coordinate or as an individual 1 km^2 grid (approximately 150,000 for England and Wales). The main database is split into regional Access databases (see Figure S1 of the ESI) which form the attribute tables behind the features in ArcMAP GIS software. Lists of the key datasets used to derive the exported loads from each point and diffuse source and on how these are used to derive the calculated loads to water bodies are provided in the ESI.

The functionality with ArcMap and bespoke macros developed in Visual Basic are then used to extract the necessary data and generate the text file required to run SIMCAT, a stochastic water quality model. SIMCAT can be run from within SAGIS and provide outputs (total and dissolved concentrations and loads for metals, total concentrations and loads for nutrients and organics) which are fed back into ArcGIS to provide cartographic, graphical and spreadsheet outputs. The existing river model was validated using monitoring data provided by the Environment Agency (EA) and provided a good fit with predicted values (Comber et al., 2013). The EA's Water Quality Archive provides data on water quality measurements from across England for over 100 determinands, analysed within accredited laboratories. Samples include coastal or estuarine waters, rivers, lakes, ponds, canals or groundwaters. They are taken for a number of purposes including compliance assessment against discharge permits, investigation of pollution incidents or environmental monitoring. The archive provides data on measurements and samples dating from 2000 to present day.

2.2 Additional sources to lakes

Data sources

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

Table 1. Summary of databases utilised for the lake and reservoir source 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)

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

Angling ground bait (nutrients) inputs to lakes

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

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

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

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

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

Bird nutrient excretion (nutrients)

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

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

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

Assumptions included:

- Bird population data of which it was assumed that 50% of the population were based on lakes. This can be replaced by detailed local information by the model user if available.

- Nutrient inputs estimated from Netherlands data which takes account of food intake, foraging behaviour and digestive performance of water birds (Table S8) (Hahn et al., 2007; 2008).

- Total bird populations (Table S9) and seasonal distribution (Table S10) combined with survey data to pro rata distribute loads to individual lakes and reservoirs (based on size). Again, default values for number of birds can be replaced by better local data from bird counts and the estimated nutrient input recalculated accordingly.

2.3 Allocation of chemical loads to lakes

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

Inputs from rivers

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

Pumped inputs

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

Direct inputs from the catchment.

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

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

Groundwater inputs

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

2.4 SAGIS and SIMCAT model updates

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

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

Lake process parameterisation

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

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

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

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

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

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

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

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

1. New boundary features called lake inflow and lake outflow and
2. A new input feature downstream of the lake.

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

2.5 Model validation

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

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.

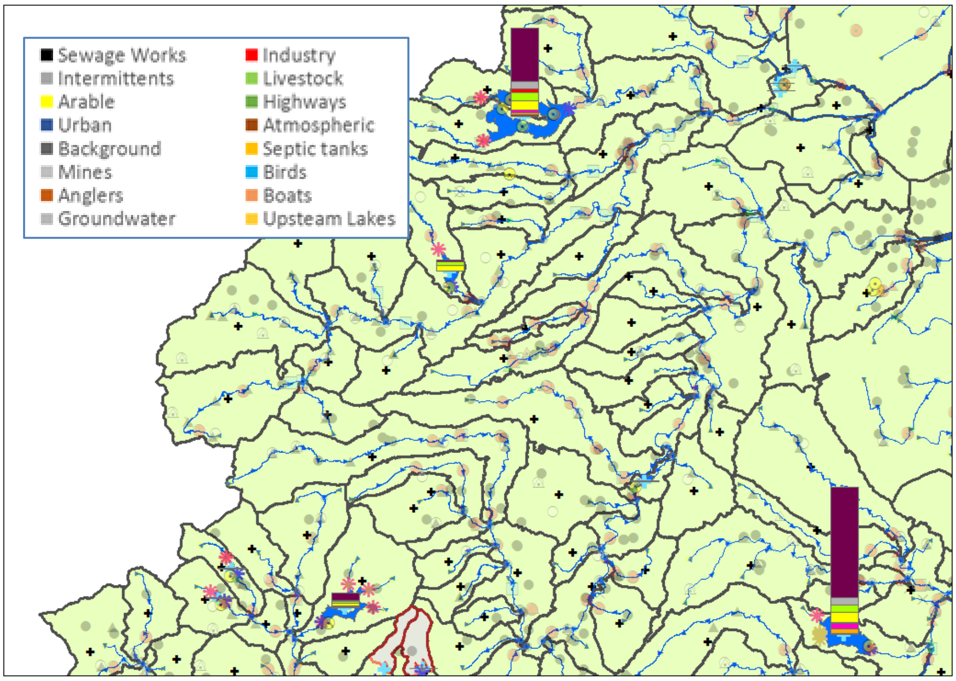

Figure 2. **Sector bar charts for sector input loads to lakes**

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

- **Pumped Storage:** Rutland Water (Anglian Water)
- **Natural Refill:** Ormsby Broads (Trinity Broads, Anglian Water)
- **Bankside Storage:** Costessey Pits (Anglian Water)

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

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

- Total Phosphorus
- Nitrate
- Copper
- Benzo-ghi-perylene

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

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

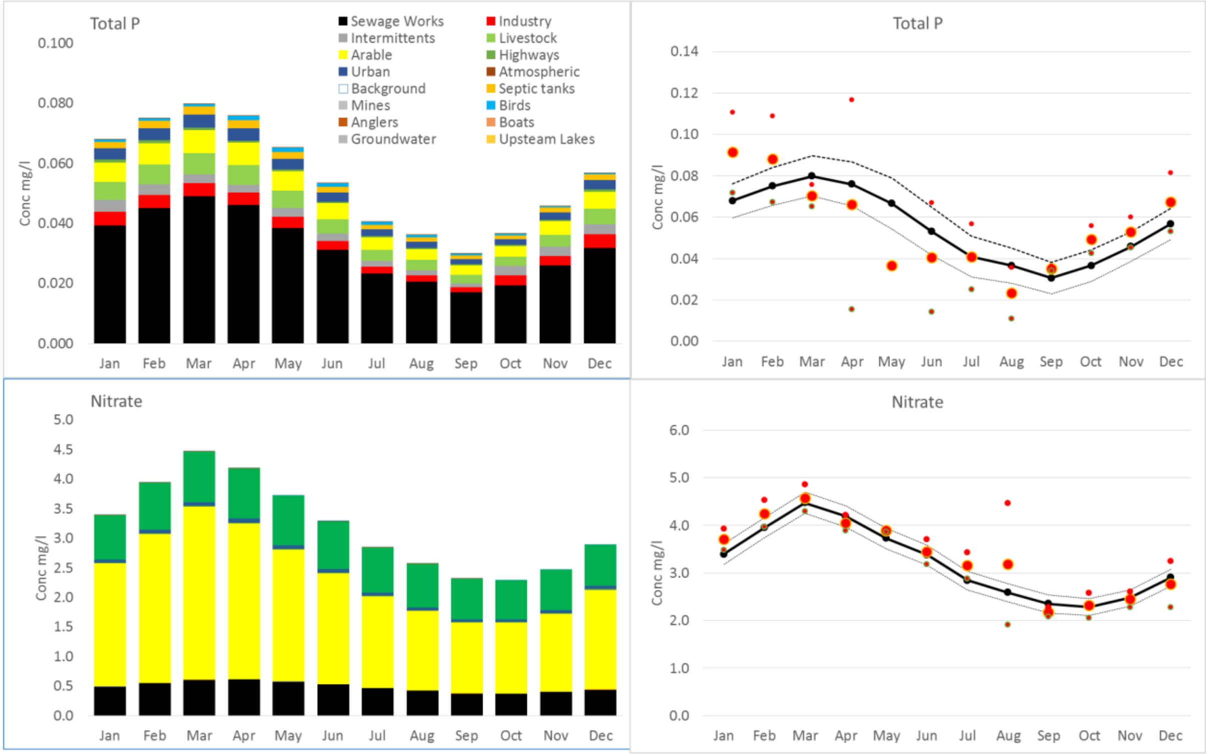

Figure 3. Modelled (black line) and observed (orange dots) monthly average total phosphorus and nitrate concentrations (+/- standard deviation) in Rutland Water with monthly source apportion

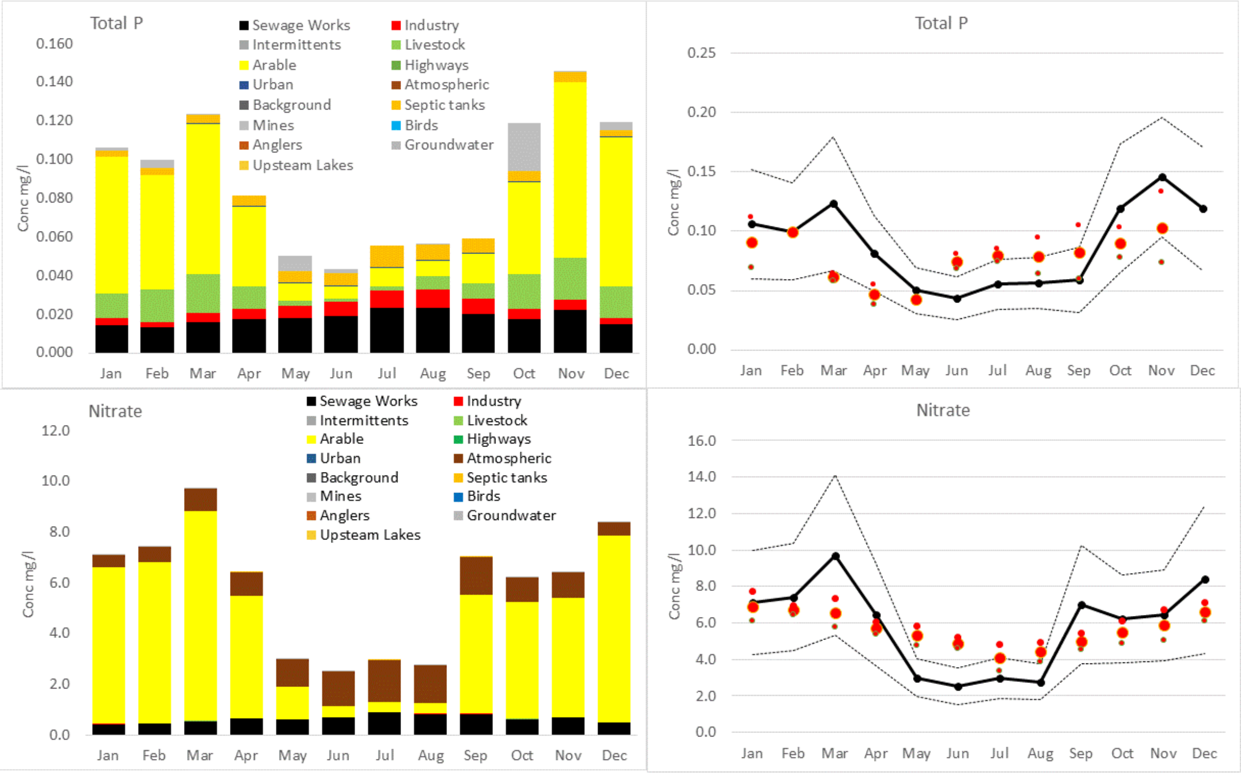

Figure 4. Modelled (black line) and observed (orange dots) monthly average total phosphorus and nitrate concentrations (+/- standard deviation) in Cotessey Pits with monthly source apportionment

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

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

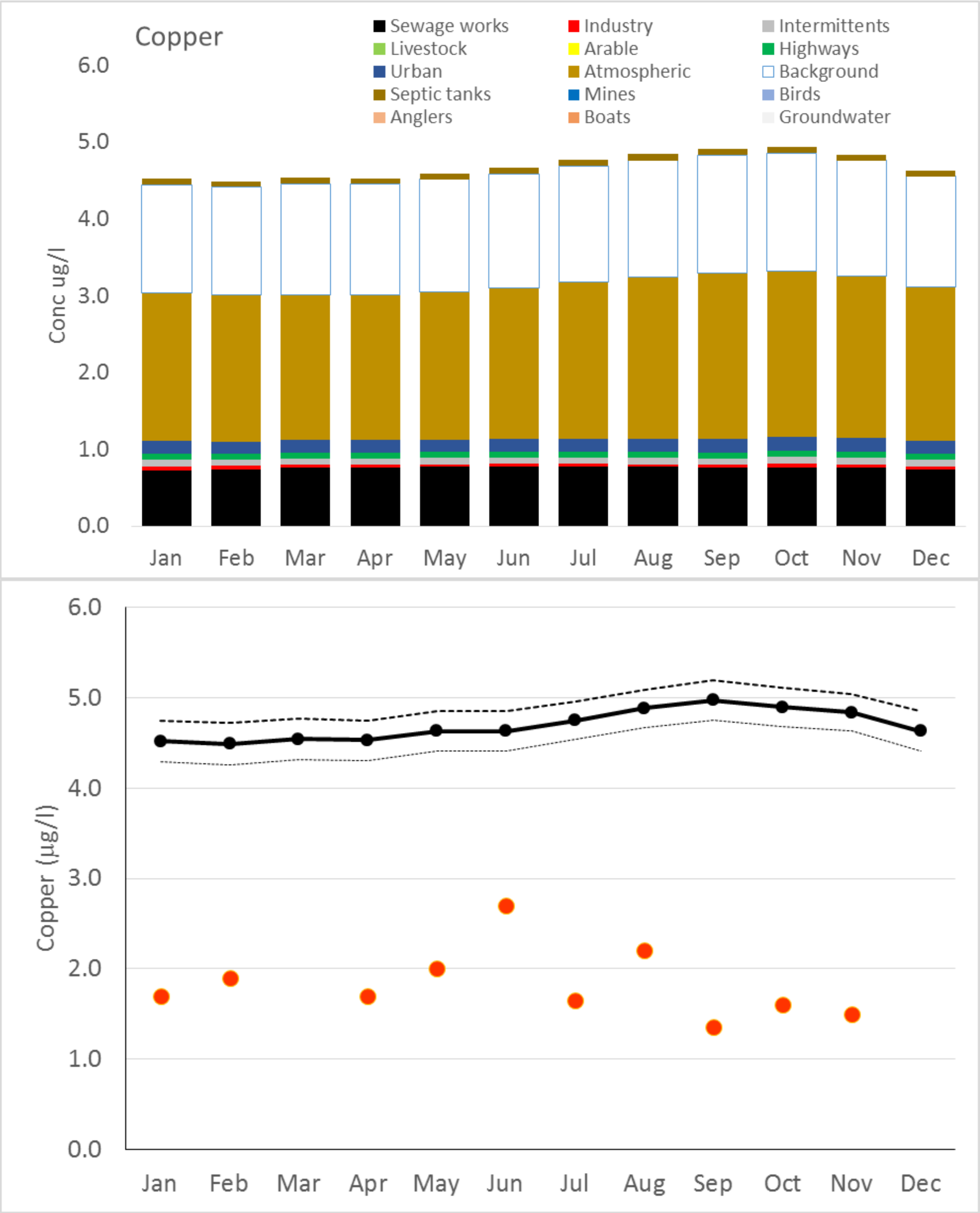

Figure 5. Modelled (black line) and observed (for 2006 data - orange dots) monthly average total copper concentrations (+/- standard deviation) in Rutland Water and monthly source apportionment.

4. DISCUSSION

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

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

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

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

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

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

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

Table 3. Source of uncertainty in the model

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.

Such a large spatial model such as SAGIS clearly requires national datasets with the aim of providing a consistent approach to national and regional water quality planning. This is obviously at the expense of localised sources which may impact on water quality. Furthermore, hydrological connectivity between lakes and the catchment is often complex and operational management may substantially influence water quality, e.g. the timing of pumping to reservoirs. The export load databases for birds and anglers are, for example, derived from regional information and local more detailed information will may be available, e.g. bird numbers on lakes which could be utilised in future. This is particularly important for lakes because of their hydrological complexities. Furthermore, lakes often have relatively small local catchments that strongly influence water quality. At this spatial scale it is often essential to apply detailed local knowledge to develop a reliable understanding of key influences and processes. The functionality of the lakes model in SAGIS ensures that more detailed local information can be easily accommodated by the model user.

5. CONCLUSIONS

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

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

Acknowledgements

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

Supporting Information

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

5. REFERENCES

Abbott M.B., Bathurst J.C., Cunge J.A., O'Connell P.E., and Rasmussen J. An introduction to the European hydrological system: Systeme Hydrologique European, 'SHE.' 2: Structure of a physically-based, distributed modeling system. *J. Hydrol.*, 1986;87:61 –77.

Ambrose, R.B., Wool, T.A., and Connolly J.P. WASP4, A Hydrodynamic And Water Quality Model. Model Theory, User's Manual, and Programmer's Guide, US Environmental Protection Agency, Athens, Ga, USA, 1988.

Arlinghaus, R. and Mehner, T. Socio-economic characterisation of specialised common carp (Cyprinus carpio L.) anglers in Germany, and implications for inland fisheries management and eutrophication control. *Fish. Res*. 2003;61:19-33.

Arlinghaus, R. and Niesar M. (2005) Nutrient digestibility of angling baits for carp, Cyprinus carpio, with implications for groundbait formulation and eutrophication control. *Fish. Man. & Ecol*. 2005;12(2):91–97.

Chapra S.C. (1997) Surface Water Quality Modelling. McGraw Hill International. ISBN 13: 9780070113640.

Comber S., Smith R., Daldorph P., Gardner M., Constantino C. and Ellor B. Development of a chemical source apportionment decision support framework for catchment management. *Env. Sci. and Tech*, 2013;47(17):9824-9832.

Dardni. 'Phosphorus inputs to Lough Neagh'. Presentation. Department of Agriculture and Rural Development, Northern Ireland. http://www.afsni.ac.uk/Research/P_sources/sld012.htm. 2007.

Davison, P. S., Withers, P. J. A., Lord, E. I., Betson, M. J. & Stromqvist, J. PSYCHIC - A process-based model of phosphorus and sediment mobilisation and delivery within agricultural catchments. Part 1: Model description and parameterisation. *J. Hydrol*. 2008;350:290-302.

EA (2010) Environment Agency. Public Attitudes to Angling 2010 - A survey of attitudes and participation in England & Wales. http://publications.environment-agency.gov.uk/PDF/GEHO0610BSOS-E-E.pdf

EEA (2005) European Environment Agency. Source apportionment of nitrogen and phosphorus inputs into the aquatic environment. European Environment Agency, Report No7/2005. ISSN 1725-9177, 2005.

EEA (2015a) EEA European Environment Agency, Lake information. http://www.eea.europa.eu/themes/water/european-waters/lakes. 2015, Accessed 24th November 2015.

EEA (2015b) EEA European Environment Agency, Lake information. http://www.eea.europa.eu/themes/water/european-waters/reservoirs-and-dams. 2015b, Accessed 24th November 2015.

EU (1976) European Union: Directive 76/464/EEC of 4 May 1976 on pollution caused by certain dangerous substances discharged into the aquatic environment of the Community. Codified as 2006/11/EC.

EU (2000) European Union Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy.

EU (2008) European Union: Directive 2008/105/EC of the European Parliament and the Council on environmental quality standards in the field of water policy, amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, 86/280/EEC and amending Directive 2000/60/EC. Published in the Official Journal on 24 December 2008.

Gough D.I. Incremental stress under a two-dimensional artificial lake, *Can. J. of Earth Sci.* 1969;6(5):1067–1075.

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

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

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

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

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

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

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

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

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

Nashoug, Nashoug (ed.) Vannkvaliteten i Mjøsa—før og nå Mjøsovervåkingen gjennom 25 år. Report from Styringsgruppa for overvåking av Mjøsa.1999. 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**.

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

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

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

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

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

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

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

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

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