Project LM0308 – Developing more effective models for managing water from the local to the national scale from a risk and pollution perspective

Current barriers to, and opportunities for, integrated water quality modelling: a review and recommendations

Andrew Wade[1], Matthew Fry[2], Richard Williams[2], Jack Cosby[3]

[1]Dept. of Geography and Environmental Science, University of Reading, Whiteknights, Reading, RG6 6DW.

[2]Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, OX10 8BB.

[3]Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, LL57 2UW

Executive Summary

This document reviews the barriers to, and opportunities for, integrated water quality modelling. Recent articles on the subject of integrated environment modelling highlight a need for a community approach where data, models and associated tools are openly accessible for all. The purpose of this is three-fold. The first is to make decision-making processes more transparent and efficient, with open evaluation of available models and modelled output. The second is to address the need to link or integrate models to provide a more complete description of the water environment and of the links between the drivers of change and the ecosystem response. The final opportunity is to capitalise on the vast new data sets that are available from multiple sources, so called 'Big Data', and take advantage of the information contained through new analytical techniques.

The main barriers to integrating data and water quality models are scientific, technical and cultural. The most significant barrier is that of Intellectual Property ownership. Scientific barriers will persist as decisions always have to be made on the best available knowledge to be incorporated in models, but these barriers will be overcome incrementally through ongoing research and model development. The need for more detailed models can be overcome by chaining models to represent multiple hydrological, hydrochemical and ecological processes. Technical barriers can be overcome, in the short-term, by community adoption of standards for data transfer between models and with visualisation tools and, in the longer term, by the implementation of integrated modelling interfaces. Overcoming the barriers due to Intellectual Property, perhaps more so with data than the models themselves, requires a cultural shift in UK science and its management.

Based on this review, seven recommendations are made:

Recommendation 1: A community approach is needed to bring together model users from the public, private and third sectors with model developers to share best practice and insights into model development, use, data sources, data processing and visualisation.

Recommendation 2: A single web-based platform where models are freely available and accessible should be created, ideally with data conditioned for model ingestion and licensing issues clarified.

Recommendation 3: A searchable database of models should be added to the platform, including summary information on the model structure, data requirements, outputs and access information.

Recommendation 4: A model usage database should also be added to store evaluations of model performance and pedigree. Such a database should include a report on the compliance of each model with the recommendations of the Macpherson Report.

Recommendation 5: The platform should allow data visualisation, both for ingestion into models and of modelled outcomes and, to help users identify and understand datasets, the platform should provide a comprehensive catalogue of the datasets available for use in water quality modelling.

Recommendation 6: Case studies should be identified to test the platform as a proof of concept.

Recommendation 7: Standards for data publishing and transfer between models and with visualisation tools should be established to promote ease of use.

Contents

1. Introduction

Questions of environmental management are typically complex and relate to achieving numerous objectives for multiple environmental policies and guidelines. For example, within the UK, there is a desire to achieve the optimum trade-off between efficient food and timber production and the cost-effective protection of soil function, water quality and biodiversity. In England, such a trade-off must be achieved within the context of at least 9 policy instruments spanning regulation, protected areas and pollution control (Arnell et al., 2015). To aid environmental management there is now a vast range of environmental models spanning issues from greenhouse gas emissions and risk from natural hazards, to ecosystem services and biodiversity loss. Environmental models are tools which can be used to help understand complex, catchment systems or, with less resolution, provide environmental assessments at the national scale. Importantly, once demonstrated as sufficiently robust through an ability to explain observations, they can be used to explore future scenarios of change and to evaluate management options, though as Bagstad et al. (2013) notes "*to enter widespread use, ecosystem service assessments [or environmental models] need to be quantifiable, replicable, credible, flexible and affordable"*.

The focus of this review is water quality modelling, as water quality is an important environmental problem given only 17% of UK surface water bodies currently meet the "good ecological quality" standard mandated by the EU Water Framework Directive. Moreover, water quality is a topic which has a range of well-developed models to inform policy, but the strengths and limitations of these cannot be readily identified, and presently these models are not applied consistently and efficiently.

Individual water quality models range in complexity and data requirements depending on the problem addressed, and typically new model applications tend to be time consuming to set up due to data collation and model calibration and testing efforts. Models range in complexity from relatively simple empirical models, to conceptual models, to more-complex physically based models. There are models designed for research and models designed for use by regulators and industry, and models that have moved from the research domain to become regularly used for policy and decision support. The number of models is confusing and metadata on their data requirements, pedigree, advantages and limitations has not been collated in one place. Furthermore, there is a tension in the discipline of water quality modelling. To date, general model complexity, a lack of access to models, and uncertainty in input data, model structure and parameter values have all led to scepticism in modelled outputs, and this is potentially holding back actions to improve the environment. From this perspective it is desirable to have simple, accessible models which are readily understood. Conversely, there is a need to consider water quality as one of many ecosystem services, and consider multiple pressures and pollutants across the air-soil-water continuum from source to sea to aid integrated management (Collins and McGonigle, 2008). This need means that the integration of water quality models with other models, such as on-farm nutrient budget or freshwater ecosystem models, is desirable. Whether the context requires a simple or more complex, integrated approach, there is a need for clarity over data inputs, assumptions, modelled outputs and transparency of model performance. These factors are now recognised as key if the modelled outputs are to be used to aid governance (Macpherson, 2013).

In parallel to these developments in water quality modelling, recent advances in remote sensing, sensor technologies and the growth in computing power, the internet, social media and in citizen science have led to the generation and availability of vast environmental datasets. A further challenge is develop the analytical and visualisation techniques to exploit these data, in this case for the purpose of improving water quality.

Contemporary projects have focussed on developing modelling frameworks and toolkits for the evaluation and integration of existing environmental models to support catchment policy development (e.g. the European funded projects: BMW; EUROHARP; HARMONIT and the NERC funded, EVO pilot project). Recently, a community-based approach to environmental modelling has been advocated to stop the reinvention of models and frameworks, which can happen due to organisational and technical barriers, and to increase the efficiency and re-use of past investment in the rich legacy of models already existing within the community (Mooij et al., 2010; Laniak et al., 2013; Janssen et al., 2015). Furthermore, a community-based approach can help build the integrated models that describe multiple ecosystems and their connectivity, and should also enable more transparent model use. This review aims to identify and consider opportunities from an improved accessibility to water quality models and associated data, and the barriers to realising those opportunities. To achieve this aim there are three objectives:

1. to identify the major barriers for integrated water quality modelling (section 2);
2. to identify the opportunities from linking data and models to address management questions related to water quality at the catchment and national scales, and from increased data access and visualisation (section 3);
3. to recommend how the identified barriers can be overcome within the specific context of this project (section 4).

By addressing the aim and three objectives, this document is essentially a high-level design document in terms of what is required within a community-based approach within the context of UK water quality modelling.

2. Current barriers for integrated water quality modelling

2.1 Scientific barriers

Models are simplifications of reality and therefore cannot provide precise answers. Model incompleteness in terms of process or spatial representation is often cited as a reason that the modelled outcomes are invalid, and the chaining or integration of multiple models is often considered to cause further uncertainty due to the accumulation of errors. In addition, models are often developed in relatively data-rich settings, but practitioners often want to apply models to instances with limited or no data. Despite these problems, if well documented, models provide a well-defined conceptual structure and sets of boundary conditions and assumptions, and are useful to explore the response to different inputs.

2.2 Technical barriers

For those unfamiliar with water quality modelling, a large amount of time can be spent trying to identify the right model for the task and then trying to obtain that model and the data to run it. Once a model is identified, which may be suitable or otherwise, collating data from multiple sources and dealing with licensing issues can be time consuming. In addition, different data sets are typically stored in a range of formats and often data processing is needed to provide the required model input data. Datasets do exist that have been processed into a more useable form for modelling, but these may be proprietary or restricted in resolution due to non-disclosure agreements, or the assumptions made in the processing steps are unavailable for scrutiny.

Models are written in a variety of different programming languages and compiled for different platforms, and different models that do similar tasks may have different geometries or discretisation. User documentation is not always available and many models do not conform to any standards in software development. Often version control of models and data is limited and it is difficult for users to keep up with which versions do what. Practitioners need stable models that are robust and reliable with good documentation, but typically research sponsors do not wish to pay for such additions. The source code is often unavailable so it is difficult for users to make changes or test new ideas about processes or use new datasets. There are no common standards for data outputs and inputs, limiting the ease with which model outputs can be compared, analysed or visualised.

2.3 Organisational and cultural barriers

Organisational and cultural barriers prevent ready access to data and models. There is a tension as Higher Education Institutes, research institutes and consultants use data and models to generate income and therefore providing open access to key datasets, source code or executables will affect revenue and competitive advantage. As a result of this culture, much data and many models remain within institutional boundaries and this limits the consistency of different model runs, inhibits dissemination of expertise, training and confidence in modelled outputs. In addition, some data are only available at a coarser resolution than that of the original collection and datasets. For example, the Agricultural Census data is redacted or only made available at coarse resolution to prevent disclosure of the data from individual respondents.

This issue of intellectual property appears to be the hardest barrier to overcome, though a slow change in culture is being driven by UK Government to more open source models, data and publications, when the development is sponsored through some public funding streams. This welcome cultural change may make government data more accessible, but datasets which represent an investment of knowledge and money, and a source of income to the owner, are unlikely to be made open without provision of replacement sources of income or adequate financial incentive. It is usually the case that these restricted datasets are required, despite the costs, because developing a replacement dataset is not possible or is financially prohibitive. A significant amount of funding for the development of data products is undertaken through innovation streams that encourage commercial exploitation of outputs. It is likely therefore that whilst “raw” monitoring data will become more open and free-to-use, the need to pay for the next generation of derived data products that bring benefits to water quality modelling will remain. As such, there is an ongoing requirement to consider the cultural changes required to facilitate models of data use and payment that enable simple and flexible use of the best datasets available. Technology can help provide solutions, e.g. modelling platforms enabling authorised access to data, quantifying the data use, implementing licensing conditions, and charging appropriately. This could ensure key datasets remain available to researchers and are not prohibitive within minor modelling applications, which is essential to avoid large variations in the datasets used and the inconsistencies this brings to model results. Given the need to prioritise which data and models would be most useful to be open access, then strong leadership and buy-in from the user community is needed to help provide governance and a clear definition of what is needed.

3. The opportunities for integrated water quality modelling

Based on a review of existing projects and recent publications which are both detailed in the following sections, the main opportunities for integrated water quality modelling are: a community-based approach to share best practice; ready access to data and models; methods to help users identify the most appropriate model for the task in hand, and to understand if a model is well tried and tested and the right tool for the job; tools to help explore data prior to use in a model or as modelled output; new opportunities for linking models to better represent the cause and effect relationships between pressure and response; standards for model inputs and outputs to enable more efficient transfers between models and data visualisation tools, and the exploitation of new data sources.

3.1 A community-based approach

The development of a Community of Practice for integrated water quality modelling will allow model users from different backgrounds to gain help with specific model application problems (Laniak et al., 2013). A community consensus is required to build a data and model platform to help obtain data and models and to agree data standards (Saloranta et al., 2003; Mooij et al., 2010; Greene et al., 2015). Multi-disciplinary modelling teams are also required for the integration of models that cover a range of disciplines to understand how each component works and the assumptions behind it (Janssen et al., 2015). Such community-based approaches are already being developed in other topic areas, including hydrology (eWater, FEWS), natural resources management (ENSYM), environmental hazards (Hazard Manager UK), ecosystem biogeochemistry (GRAMP) and greenhouse gas emissions (Greenhouse Gas Platform). The main advantages of these developments are the provision of a forum for knowledge exchange around data and models, whilst the main limitation is that none of the tools yet focus on UK water quality (Table 1). The level of activity in allied areas demonstrates the desire to see such platforms become reality and also offers learning opportunities regarding best practice and what works well or otherwise. This project should seek were possible to collaborate with, and learn from, these allied projects.

Table 1. Cognate projects building web-based platforms to improve access to environmental data and models. The main limitations are viewed in the context of this project.

Project	Description	Link/Reference	Key advantages	Main limitations
eWater	Web-based access to hydrological and water quality models	http://www.toolkit.net.au/	Includes data, models, forums, guidelines on model choice and best practice.	Does not include models commonly used in UK.
ENSYM	Environmental Systems Modelling Platform	https://ensym.dse.vic.gov.au/cms/	Includes models, data export in multiple formats and compatible with GIS.	Does not include models commonly used in UK.
CSDMS (Community Surface Dynamics Modelling System)	Land surface modelling community platform, with model and dataset repository, and information on accessing HPC resources.	https://csdms.colorado.edu/wiki/Models_all	Features a well-populated model repository, user derived content is moderated, provides information on setting up models to be run, over the web, on super-computers	Focussed on this specific research community.
FEWS	An open shell system for managing flow forecasting processes and/or handling time series data. Written by DELTARES	https://publicwiki.deltares.nl/display/FEWSDOC/Home	Data handling utilities for visualisation and analysis	Visualisation tool kit rather than data and modelling platform.
FluidEarth	Software development kit allowing users to link their models to data and other models	http://fluidearth.net/default.aspx	Openly available. Based on OpenMI. Includes case studies and examples of model linkage.	Unclear if project has ceased as no updates to website since 2014.
GRAMP	Global Research Alliance Modeling - Platform Web-based modelling platform	http://gramp.org.uk/	Sign-posts to models, source code, performance summaries, publications	Focused on greenhouse gas emissions and soil carbon and nitrogen.
Greenhouse Gas Platform	The aim is to provide the evidence for a UK specific method of calculating methane and nitrous oxide emissions that will reflect the adoption of mitigation practices by the industry.	http://www.ghgplatform.org.uk/Home.aspx	Collation of project outputs	No collation of models.

Table 1. Cognate projects building web-based platforms to improve access to environmental data and models.

Hazard Manager (UK)	Hazard Manager is a one-stop information source for the emergency response community.	http://www.metoffice.gov.uk/publicsector/hazardmanager	Interactive web portal using maps which can be integrated with weather and incident related information.	Visualisation tool kit rather than data and modelling platform.
NERC Workbench	The Environmental Research Workbench enables NERC scientists to easily exploit the underpinning JASMIN infrastructure, providing an intuitive, user-friendly interface and a range of tools and services that will help them analyse large and heterogeneous environmental datasets and accelerate science discovery.	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/285075/bis-exceptions-to-spending-controls-ict-spending-approvals-october-to-december-2013.csv/preview	Help overcome a major hurdle for the many scientists who lack the technical knowledge to directly exploit high performance and cloud computing facilities	Only available within NERC.
AgriMetrics	A computational platform for data integration and analysis and modelling across the whole food chain, and the use of smart analytics (statistics, visualisation and modelling).	https://www.gov.uk/government/news/new-agrimetrics-centre-will-boost-food-and-farming-industries	Exploitation of ‘Big Data’ for Agri-Environmental issues	Just started
Nexus Tools Platform	Interactive platform for inter-model comparison of existing modelling tools related to the water-soil-waste nexus. The NTP provides detailed model information and advanced filtering based on real-time visualizations and will continuously grow with the input and feedback from the model developers and model users. Currently, the NTP database consists of 60 models from around the world.	https://data.flores.unu.edu/projects/ntp/#/dashboard/elasticsearch/Nexus%2520Tools%2520Platform%2520(alpha)	Tools database that allows the interactive comparison of those tools.	No sign-posts to data.

Table 1. Cognate projects building web-based platforms to improve access to environmental data and models.

Policy Support Systems	Website for download of range of policy support tools including those for water resources (WaterWorld) and ecosystem services (Co$ting Nature). Includes visualisation tools, model descriptions, examples and evaluations.	http://www.policysupport.org/	Collation of models, visualisation tools to explore example modelled outputs, training videos	Limited range of models at this time. No signposts to data.

3.2 Accessibility to data, models and useful tools – a web-based platform

A single web-based platform will facilitate having data and models readily available and well documented in one place, with the models evaluated in terms of performance and pedigree, and the input data signposted and its veracity authenticated. This will save time on data collation and quality checks, and allow users to more easily find the most appropriate model for a particular water quality problem. Third party data, models and associated tools for pre- and post-processing data and visualisation could be efficiently located from the platform via Uniform Resource Locators (URLs). Ideally such a platform would provide access to a model executable (or source code), have data to hand and allow example model applications to be run for model inter-comparison and exploration by non-expert users. Tools for sensitivity and uncertainty analysis embedded within the platform would also provide a means to help assess the key mechanism operating and the uncertainty in the modelled outcomes. Everyone has their own favourite method and tool for investigating parameter uncertainty, but users could initially select some common approaches for integration on the platform, or sign-posting. There is now more ready access to both commercial and research cloud computing facilities (e.g. NERC's JASMIN facility which can be accessed via a collaborative platform, the Environmental Research Workbench). Ultimately, the use of cloud computing would allow model users to run models without the need to download the model and data to their own computer which may be limited in its processing power, memory and data storage capacity. For example, the Majic system (majic.ceh.ac.uk) implements the JULES Land-surface model providing a simple interface to run a complex model across a high-performance-computing environment and simply visualise and access results. This is perhaps too ambitious for the scope of this project but should be an ultimate goal.

There are also benefits of such a platform for model developers. More people will be able to gain access to models which should then become more used. This is beneficial for continued model enhancement and will help demonstrate the non-academic impact of the model development. For example, the ready availability of the SWAT water quality model has led to widespread, international application. The development of a data and model platform will also provide security in terms of investment by the major sponsors of model developments through ready access to the product and long-term curation of data, models, and the outputs of important model runs. These could be stored on the platform directly or in other long-term repositories and sign-posted. It is anticipated that community buy-in would start to evolve a new 'market place' where models absent from such a platform will be seen as only for research or less transparent or tested.

The platform would help disseminate expertise and could help with training by: aiding decisions on model choice and data availability; allowing users to download example data and models and run example applications on their desktop, and by providing a forum to discuss data and model issues. A new user, with confidence built that they can achieve same result as an example, could then run their own application. Once done, the new results and model performance evaluation could be uploaded to the platform to serve as a further case study and also to build parameter and modelled output libraries.

It is an intended consequence of greater access to data and models, and a better understanding of models and the modelling process, that confidence in modelled outputs will improve and therefore can be used to help enact land management change. The lack of confidence in models could, in part, be overcome by bench marking. More specifically, a model could be run by a user and compared with modelled output already stored on platform to check that nothing has gone wrong in data transfer or model setup. There are opportunities to do this by using existing models with a strong pedigree in a selection of well-monitored observatory catchments covering the major land cover types in the UK.

3.3 Finding the best model for the job

The development of a web-based portal should host a searchable database of model metadata that, in turn, would permit users to find the model(s) most appropriate for their management questions. The metadata should include links to the latest version of a model and to a stable release, most likely on external websites. In addition, clarification of model geometry, time step, data input and calibration requirements, modelled outputs and links to documentation would be useful. Ultimately, it would be useful to compile a database of parameter sets for models including both 'factory settings' and library of model parameters for different geographical settings By drawing together model evaluations then it will be possible to better establish overall model robustness. Collation of metadata will also help users identify suitable candidate models for linking or running in ensembles.

3.4 Enhancing confidence in models: pedigree and performance

Practitioners need stable, well-documented models that have a good pedigree. Bagstad et al. (2013) suggest eight criteria for model evaluation which are a useful starting point for evaluating model performance and utility: quantification and uncertainty assessment, time needed for model application, capacity for independent application, the level of development and documentation, scalability between different spatial scales (or the ability to link multiple models to cover the scale range needed), generalisability (the ability of a model to be applied for a range of different landscape types), incorporation of nonmonetary and cultural perspectives (for ecosystem valuation), and affordability and integration with existing datasets and models (to allow efficient use of time and money). The platform should offer a repository for model pedigree assessments for models for stable versions. After each model application, a short report evaluating the model should be added to repository to build a picture over time of what instances a particular model was useful for (or not). At present, assessments of model performance are often buried in the literature with the results, once found, difficult to interpret in terms of whether the model would be useful for a different user for a different situation. This evaluation report will incorporate the recommendations of Bagstad et al. (2013).

3.5 Data availability and visualisation

Once a decision to run a particular model is made, then the biggest task is often data collation and reformatting for model ingestion. Data is often distributed across many sources and provided in many different formats which often require processing prior to ingestion. A dedicated catalogue would help in identification and access to data, and description of datasets using standard metadata would aid understanding of their provenance (e.g. ISO19115 used widely for spatial data and the data.gov.uk catalogue). Use of Digital Object Identifiers (DOIs) would ensure long-term availability of web-resources. Use of standards to make datasets accessible, and simple tools for transforming them to appropriate forms for use in models would avoid the often daunting task of tackling unfamiliar and often complex data formats.

There are likely opportunities in terms of more efficient licensing of data sets, whereby data sets procured from the same source could be covered by a single licence agreement. The licencing stage often leads to delays in the model application, and may lead to use of unlicensed or lower quality data as time pressure increases. With open access to data then the need for data repositories and long-term curation of data increases, as does the need to secure and protect personal data on which open-access data are based.

The increase in data volume means that new tools are needed to handle interpretation and visualisation. In addition, data visualisation tools are needed to explore data sets prior to ingestion into models and to review modelled outcomes. A key challenge is to deal with both spatial and time-series data and the host of formats that these data are collected and stored in. User interfaces for prioritised models could be enhanced to increase usability including user friendly interfaces for model graphing tools. It would be beneficial to be able to compare modelled outputs and observations of different resolutions by re-mapping. This could be aided by the inclusion of map co-ordinates as standard in datasets to allow Geographical Information System (GIS) presentation.

3.6 Model linkages: some examples of what could be possible

To deal with complex catchment management problems it is useful to be able to soft-link, or even integrate, two or more models to represent difference aspects of an issue, for example for consideration of multiple agri-environmental measures, multiple ecosystem services or the response of the freshwater ecology to upstream changes in land management. The enhancement of models and modelling frameworks is envisaged for models that are already industry standards (e.g. Farmscoper and SAGIS) or newly emerging with unique capabilities (e.g. LUCI) to: embed alternative specialist models; extend the models to cover additional pollutants, to include additional ecosystem services (e.g. ecological impacts), or run an ensemble of models (Mooij et al., 2010). Example case studies with potential benefits to industry, regulation and government might consider:

- the potential trade-offs and co-benefits for a suite of ecosystem services at the farm to catchment scale that may be potential ‘by-products’ of an agri-environment scheme designed to reduce total multi-pollutant loads entering watercourses;
- the potential impact of an agri-environment scheme on total pollutant loads (from all sectors) entering watercourses at the national scale, and any additional consequences of scheme implementation for national greenhouse gas emissions.
- a comparison water treatment costs and those of upstream pollution control measures for achieving compliance with environmental regulations and drinking water standards.

- the effectiveness of nutrient pollution reduction measures given future projections of climate and land cover change at the catchment scale.
- the attribution of error sources and their effects in estimating the ecological response to mitigation methods aimed at nutrient reduction.
- the sensitivity / uncertainty in outputs of models at various scales resulting from differences or uncertainties in input driving data.

3.7 Efficient integration of data and models: data standards

To link multiple models, it is beneficial to have standards for model input and output data. This may be something simple to enable data ingestion between models and with visualisation tools, or something more complex, such as OpenMI (Open Modelling Interface, http://www.openmi.org/) which allows data exchange between models during run time (Gregersen et al., 2007). Already there are common file formats, such as netcdf for spatial data that is used widely in the climate science community and in Geographical Information Systems, which could become more commonplace in water quality modelling. Also, there are new standards being developed to allow multiple models to be linked such as those of the Open Geospatial Consortium (OGS). OpenMI and OGS are future options for cross-linkage of models and data though given the need for expert programming skills then adoption of these standards in modelling is not straightforward.

Furthermore, detailed frameworks for integrated modelling already exist and these provide templates for environmental model application to help best practice in modelling and a commonality in approach to help build confidence in modelled outcomes by applying models in a manner accepted by the community (Refsgaard et al., 2006). An improved reliability in modelled outcomes through consistency in approach and application of specific steps (calibration, testing, sensitivity and uncertainty analysis) and quality assurance measures will lead to a greater trust in modelled outputs and a more robust evidence base for decision making. These motives are at the centre of the Macpherson Report whose recommendations detail the quality assurance needed for all models used to aid government decision making (Macpherson, 2013). These recommendations have been adopted in this project and, furthermore, it would be appropriate to have standard templates for uploading metadata to describe new models and model evaluations.

3.8 Opportunities from new data sources

In recent years, remote sensing capabilities have increased with new missions launched to improve, the imagery of vegetation, soil and water cover, inland waterways, coastal areas, land-surface temperature, ocean colour and land colour, for example, through the European Space Agency's Copernicus Sentinel programme. In addition, in-situ sensor technology has improved. It is now possible to routinely monitor water quality at sub-daily time steps and to make some biological measurements, such as microalgae presence and abundance, at weekly intervals. The growth of the internet, social media and smart devices has enabled data to be captured at multiple locations, collated and analysed, with the results stored in cloud-based platforms. As discussed in section 2.3 project, government is intending to further make public data available for the greater benefit using on-line platforms, in reusable formats and under open licence. The availability of higher spatial and temporal resolution data is an exciting prospect with which to develop the next generation of water quality and ecosystem service models through better representation of climate inputs, land management, catchment geology and soil characteristics and water and chemical stores and transport pathways.

4. Conclusion and recommendations

4.1 Recommendations

Recommendation 1: A community approach is needed to bring together model users from the public, private and third sectors with model developers to share best practice and insights into model development, use, data sources, data processing and visualisation.

Recommendation 2: A single web-based platform where models are freely available and accessible should be created, ideally with data conditioned for model ingestion and licensing issues clarified. The platform should form a basis for community interactions to help with the education of users and knowledge transfer, and provide security in terms of investment by the major sponsors of model developments through ready access to the product and long-term curation of the models. There must be continual feedback between users and developers to build the site to foster the community approach.

Recommendation 3: A searchable database of models should be added to the platform, including summary information on the model structure, data requirements, outputs and access information.

Recommendation 4: A model usage database should also be added to store evaluations of model performance and pedigree. Such a database should include a report on the compliance of each model with the recommendations of the Macpherson Report.

Recommendation 5: The platform should allow data visualisation, both for ingestion into models and of modelled outcomes. Furthermore, to help users identify and understand datasets, the platform should provide a comprehensive catalogue of the datasets available for use in water quality modelling, particularly by the most widely used models.

Recommendation 6: Case studies should be identified to test the platform as a proof of concept. The case studies should focus on key user questions identified through canvassing of opinions of the practitioner community. Together the case studies should represent model applications at different scales, the soft-linking of different models, and assessments of structural and data uncertainty. The recommendation is to run the models first without the platform and then run the models using the platform to determine and check on benefits, efficiency improvements, self-help to run model for case study and for own application (on desktop). These case studies should also provide opportunity for those unfamiliar with the models to run the model application to learn about the model, data requirements and potential benefits and limitations.

Recommendation 7: Standards for data publishing and transfer between models and with visualisation tools should be established to promote ease of use. In addition, there should be templates to collate model metadata and evaluation of model performance and pedigree to ensure a consistency in information.

4.2 What remains out of scope?

Models will not be run on the platform and at present there is no access to cloud resources, so this aspect is out of scope for the project. The analysis of 'big data' analytics is being addressed in environmental science by new centres such as Agrimetrics and the Institute for Environmental Analytics and the NERC Environmental Research Workbench. New data standards are also beyond the scope of this project.

4.3 What of the next 10 years?

Ultimately it is desirable to link process-based models with 3D visualisation toolsets for the display of environmental data, in particular maps. This would enable simultaneous exploration of the landscape aesthetic and environmental impacts, and the creation of management scenarios that consider both. Ideally such visualisation tools and models would be coupled with sensor networks to enable the current state of the environment to be continuously assessed and modelled, thereby providing a basis for continuously updated short-term forecasting, for example of micro-organic pollution being transported along a river network, and the longer term projections of the impacts of management and environmental change.

With more widespread access to the internet, a growing awareness of data availability, a move to greater data access by governments, and applications to integrate diverse datasets then individuals, businesses and educators and applications will expect to be able to utilise this information. Conversely there will also be a need to protect the privacy of the individual in the collation and publication of those data.

5. References

Arnell, N. W., Halliday, S. J., Battarbee, R. W., Skeffington, R. A. and Wade, A. J. 2015. The implications of climate change for the water environment in England. *Progress in Physical Geography,* 39**,** 93-120.

Bagstad, K. J., Semmens, D. J., Waage, S. and Winthrop, R. 2013. A comparative assessment of decision-support tools for ecosystem services quantification and valuation. *Ecosystem Services,* 5**,** E27-E39.

Collins, A. L. and Mcgonigle, D. F. 2008. Monitoring and modelling diffuse pollution from agriculture for policy support: UK and European experience. *Environmental Science & Policy,* 11**,** 97-101.
Greene, S., Johnes, P. J., Bloomfield, J. P., Reaney, S. M., Lawley, R., Elkhatib, Y., Freer, J., Odoni, N., Macleod, C. J. A. and Percy, B. 2015. A geospatial framework to support integrated biogeochemical modelling in the United Kingdom. *Environmental Modelling & Software,* 68**,** 219-232.
Gregersen, J. B., Gijsbers, P. J. A. and Westen, S. J. P. 2007. OpenMI: Open modelling interface. *Journal of Hydroinformatics,* 9**,** 175-191.
Janssen, A. B. G., Arhonditsis, G. B., Beusen, A., Bolding, K., Bruce, L., Bruggeman, J., Couture, R.-M., Downing, A. S., Elliott, J. A., Frassl, M. A., Gal, G., Gerla, D. J., Hipsey, M. R., Hu, F., Ives, S. C., Janse, J. H., Jeppesen, E., Joehnk, K. D., Kneis, D., Kong, X., Kuiper, J. J., Lehmann, M. K., Lemmen, C., Oezkundakci, D., Petzoldt, T., Rinke, K., Robson, B. J., Sachse, R., Schep, S. A., Schmid, M., Scholten, H., Teurlincx, S., Trolle, D., Troost, T. A., Van Dam, A. A., Van Gerven, L. P. A., Weijerman, M., Wells, S. A. and Mooij, W. M. 2015. Exploring, exploiting and evolving diversity of aquatic ecosystem models: a community perspective. *Aquatic Ecology,* 49**,** 513-548.
Laniak, G. F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M., Peckham, S., Reaney, S., Gaber, N., Kennedy, R. and Hughes, A. 2013. Integrated environmental modeling: A vision and roadmap for the future. *Environmental Modelling & Software,* 39**,** 3-23.
Macpherson, N. 2013. Review of quality assurance of government models. *In:* TREASURY, H. (ed.). HM Treasury.
Mooij, W. M., Trolle, D., Jeppesen, E., Arhonditsis, G., Belolipetsky, P. V., Chitamwebwa, D. B. R., Degermendzhy, A. G., Deangelis, D. L., Domis, L. N. D. S., Downing, A. S., Elliott, J. A., Fragoso, C. R., Jr., Gaedke, U., Genova, S. N., Gulati, R. D., Hakanson, L., Hamilton, D. P., Hipsey, M. R., T Hoen, J., Huelsmann, S., Los, F. H., Makler-Pick, V., Petzoldt, T., Prokopkin, I. G., Rinke, K., Schep, S. A., Tominaga, K., Van Dam, A. A., Van Nes, E. H., Wells, S. A. and Janse, J. H. 2010. Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquatic Ecology,* 44**,** 633-667.
Refsgaard, J. C., Van Der Sluijs, J. P., Brown, J. and Van Der Keur, P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources,* 29**,** 1586-1597.
Saloranta, T. M., Kamari, J., Rekolainen, S. and Malve, O. 2003. Benchmark criteria: A tool for selecting appropriate models in the field of water management. *Environmental Management,* 32**,** 322-333.