CENTRE OF EXPERTISE FOR WATERS

Review of existing faecal indicator organism (FIO) models and future integration opportunities to support catchment-scale source apportionment in Scotland

Lisa Avery, Diana Tomczyk, Paul Quinn, David Oliver

The James Hutton Institute

UNIVERSITY of STIRLING

Scottish Government
gov.scot

CENTRE OF EXPERTISE FOR WATERS

Review of existing faecal indicator organism (FIO) models and future integration opportunities to support catchment-scale source apportionment in Scotland

Lisa Avery, Diana Tomczyk, Paul Quinn, David Oliver

Published by CREW – Scotland's Centre of Expertise for Waters. CREW connects research and policy, delivering objective and robust research and expert opinion to support the development and implementation of water policy in Scotland. CREW is a partnership between the James Hutton Institute and all Scottish Higher Education Institutes. The Centre is funded by the Scottish Government.

This document was produced by:
Lisa Avery[1], Diana Tomczyk[1], Paul Quinn[1], David Oliver[2]

[1]The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH
[2]Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA

CREW Project Manager: Anishka Cameron
CREW Communications and Impact Officer: Amy Cooper
CREW Manager: Nikki Dodd
CREW Director: Miriam Glendell

Project Steering Group: SEPA, Scottish Government, NatureScot and Scottish Water.

Please reference this report as follows:
Avery, L., Tomczyk, T., Quinn, P and Oliver D. (2026). *Review of existing faecal indicator organism (FIO) models and future integration opportunities to support catchment-scale source apportionment in Scotland.* Main Report. CD2025_02. Centre of Expertise for Waters (CREW).

Available online at: www.crew.ac.uk/publications/review-existing-faecal-indicator-organism-models-and-future-integration-opportunities

ISBN: 978-1-911706-53-3

Dissemination status: Unrestricted

Acknowledgements: With thanks to Jonathan Bowes, Rowena Diamond, Ian Dunhill and Chris Burgess and other stakeholders for discussing their experiences of FIO modelling with the research team.

AI statement: The content of this report and the literature review underpinning it was prepared by team members. Microsoft co-pilot was utilised in a limited editorial way to assist with re-phrasing and summarising.

Cover photograph courtesy of: Lisa Avery

Please contact enquiries@crew.ac.uk, to request report in an alternative format.

Contents

List of Tables

List of Figures

List of Abbreviations

EA	Environment Agency
CSO	Combined sewer overflow
FIO	Faecal indicator organism
HYPE	Hydrological Predictions for the Environment model
MST	Microbial source tracking
SAGIS	Source Apportionment GIS
SEPA	Scottish Environment Protection Agency
SIMCAT	SIMulation of CATchments model
STS	Septic tank systems
SUCCESS	System for Understanding Carrying Capacity, Ecological, and Social Sustainability model
SWAT	Soil and Water Assessment Tool
SWPAs	Scotland's Shellfish Water Protected Areas
ViPER	Visualising Pathogen & Environmental Risk model
WRc	Water Research Centre

Glossary

Agent-based model
A model that simulates behaviour and interactions of individual "agents" e.g. bacteria

Bayesian Network
A probabilistic graphical model that represents variables and dependencies

Calibration
Adjustment of model parameters to improve match of predicted vs. observed data

Decay rate/die off
The rate that microorganisms (e.g. FIOs) decrease

Escherichia coli *(E. coli)*
A commonly used faecal indicator organism

Empirical Model
A model based on observed relationships

Event-driven dynamics
Short term pollution events such as high rainfall, CSO overflow that drive inputs of FIOs to the water body

Exfiltration
Movement of water from underground storage into surrounding soil

Faecal indicator organism (FIO)
Microorganisms, commonly *E. coli*, total coliforms and Enterococci used as indicators of faecal pollution in waters

Hydrodynamic model
Models that simulate water movement and pollutant transport

Hydrological connectivity
pathways enabling water and pollutants to move through a catchment from source to water body

Hydrometeorological conditions
Environmental drivers influencing pollutant movement, such as rainfall, temperature and flow.

Infiltration
Movement of water from the soil surface to sub-surface

Machine Learning Models
Models driven by data (e.g. neural networks) that detect patterns without simulating processes

Monte Carlo methods
A computational algorithm that relies on repeated random sampling to solve computational problems.

Probabilistic framework
A modelling approach that incorporates uncertainty

Process-based model
A model that simulates physical, chemical and/or biological processes

Risk-based model
A simplified model used for screening and prioritisation

Seepage face
The visible emergence of groundwater from a saturated zone

Sequential coupling
Linking models inn sequence

Source apportionment
Identification and quantification of relative contributions of different pollution sources (e.g. agriculture, wastewater, wildlife)

Statistical Emulator
A simplified model approximating a complex model

Topographic wetness index (TWI)
An estimate of soil wetness based on slope and contributing area

z-score
How many standard deviations a value is from the average. In a well-fitting model, the z score is usually close to zero.

Executive Summary

Purpose of research

The purpose of this project was to review and synthesise existing modelling approaches for *E. coli* and faecal indicator organisms (FIOs) and to assess how they could support catchment scale source apportionment in Scotland. Specifically, the project aimed to clarify what FIO models are currently available, how they differ, and whether elements of these approaches could be combined to improve understanding of the most significant contamination pressures under different hydrological and seasonal conditions.

The review considered the extent to which existing models can be integrated to provide a high level, holistic representation of catchment behaviour, identified key gaps and limitations within the current modelling landscape, and evaluated the feasibility of addressing these gaps. A central focus was to assess whether a practical, high level catchment screening model could be developed to identify dominant *E. coli* sources and priority pressures, within the constraints of limited data availability and regulatory needs, to better support targeted management and decision making.

Background

Scotland's Shellfish Water Protected Areas, shellfish production zones and Bathing Waters are regulated through legislation derived from EU water quality directives. For Shellfish Water Protected Areas, compliance is assessed primarily using *E. coli* monitoring, particularly shellfish flesh data collected by Food Standards Scotland. While these data are essential for classification and regulatory decision making, they do not identify the sources of contamination. This presents a significant challenge in catchments where wastewater discharges, septic systems, livestock, wildlife and diffuse runoff interact to generate faecal indicator organism loading. Understanding the sources of faecal pollution is critical to inform decision making on interventions.

Modelling offers a means to address this gap. The occurrence and persistence of FIOs are strongly influenced by rainfall, hydrological connectivity, land use, seasonal management practices and short term human activities. Monitoring alone rarely captures this complexity or variability. Evidence from Scotland demonstrates that hydrometeorological conditions are a dominant control on FIO mobilisation and transport (Neill *et al.*, 2019, Tetzlaff *et al.*, 2012). Appropriately selected models can therefore help identify which sources are most influential under different conditions and support targeted, cost effective mitigation – critical for protecting shellfish waters and sustaining high microbiological quality.

A wide range of modelling approaches exists, from rapid screening tools and statistical methods to process based hydrological and water quality models. Each has strengths and limitations, particularly where monitoring data are sparse. Key knowledge gaps remain around FIO persistence in soils and manures, sub field scale mobilisation, sediment storage and resuspension, and model calibration under data constraints.

Key findings

- A wide range of models exists for *E. coli* and faecal indicator organisms FIOs, but no single model can adequately represent all sources, pathways and impacts across catchments at regulatory scale.
- Different model types serve different purposes:
    - o Process based models (e.g. HYPE, SWAT) represent hydrology and diffuse pollution well but are data and resource intensive.
    - o Risk based tools (e.g. SCIMAP FIO, SIMCAT–SAGIS, Bayesian Networks) are effective for high level screening and prioritisation in data limited contexts.
    - o Detailed research models (e.g. MAFIO) improve understanding of sub field mobilisation but do not scale to national or operational use.
- Hydrometeorological variability is a dominant control on FIO mobilisation and transport, with dominant sources changing between low flow and high flow conditions.
- Event driven processes, including storm runoff, combined sewer overflows and sediment resuspension, are poorly captured by many existing models despite driving most compliance failures.
- Source apportionment using FIOs remains a major challenge, as FIOs are non specific indicators and multiple sources often co occur and respond to the same drivers (e.g. rainfall).

- Model integration is technically possible but introduces significant challenges, including scale mismatches, high uncertainty propagation, calibration burden and long term maintenance risks.
- Experience shows that fully integrated "source to sea" models are difficult to operationalise and maintain for routine regulatory use.
- Long term sustainability, interoperability and usability are as critical as scientific capability when selecting or developing modelling tools.
- The evidence suggests greater benefit from restoring and strengthening national scale screening capability, supported by targeted detailed modelling where it adds clear management value.
- Future progress is likely to be driven more by better data, modular modelling frameworks and decision focused tools rather than by increasing model complexity.

Recommendations

This review shows that while many modelling approaches exist for FIOs, data limitations, system complexity and long term sustainability constrain what can realistically be delivered for routine regulatory use. Rather than pursuing fully integrated or highly complex models, future effort should focus on operational, transparent tools that directly support decision making.

Key recommendations

1. **Align modelling ambition with data reality**
   The Scottish Environment Protection Agency (SEPA) should begin with a national audit of data availability to distinguish between data that are routinely available and sustainable data (continued collection, continued availability in appropriate format, pertinent to multiple policy objectives) that could be strengthened through targeted investment, and data that are unlikely to be available at useful resolution. This audit should define the scope of any modelling framework, ensuring it is built around realistic and durable data inputs.
2. **Develop a simple national screening tool for operational use**
   The highest priority is to develop a modern, modular, national scale risk-based screening tool to identify relative FIO source risk under conditions most relevant to compliance (e.g. rainfall driven events). The tool should prioritise usability, transparency and maintainability to guide investigation and intervention.
3. **Establish a sustainable long term modelling framework**
   Modelling capability should be embedded within a long term operational framework that supports routine use, explicitly represents uncertainty, adopts open and modular software practices, and includes clear governance for maintenance and data updates. This will provide a basis for and flexibility to develop more complex modelling as needed in the future.

1.0 Introduction

1.1 Background and scope

Scotland's Shellfish Water Protected Areas (SWPAs), shellfish production zones, and Bathing Waters are governed by Scottish legislation that builds on longstanding EU water quality directives. For SWPAs, compliance assessments rely heavily on *E. coli* monitoring – particularly shellfish flesh data collected by Food Standards Scotland – to classify sites, identify areas at risk, and guide further investigation (SEPA, 2026). While these datasets are essential for regulatory purposes, they do not indicate which sources are responsible for contamination. The Scottish Environment Protection Agency (SEPA) is responsible for overseeing shellfish water quality under The Water Environment (Shellfish Water Protected Areas: Environmental Objectives etc.) (Scotland) Regulations 2013. In this role, SEPA must investigate pressures affecting SWPAs and take forward any required actions through the River Basin Management Plan 2021–2027 (RBMP). This creates a challenge in catchments where wastewater discharges, septic tanks, livestock, wildlife, and diffuse runoff all contribute to faecal indicator organism (FIO) loads. Understanding the sources of faecal pollution is critical to inform decision making on interventions.

Modelling provides a way to address this gap. Seasonal changes in rainfall, land use and management, spatial variability, climate-related stochasticity and hydrological connectivity can strongly influence when and where FIOs appear in receiving waters and for how long they persist (Krupska *et al*., 2024). This can also be affected by daily and weekly variations arising from, for example, variable wastewater flow patterns and livestock movements. Research in Scotland and elsewhere has shown that hydrometeorological variability plays a major role in FIO mobilisation and transport (Heal, 2025, Neill *et al*., 2020a, Neill *et al*., 2020b, Tetzlaff *et al*., 2012). This means that monitoring alone rarely captures the full pattern of contamination. If SEPA can draw on suitable modelling tools, it becomes possible to identify the most influential sources under different conditions and to target mitigation where it will have the greatest effect. This is particularly important for maintaining the integrity of SWPAs and supporting shellfish production in waters of consistently high microbiological quality.

A wide range of modelling approaches are available, each offering different strengths. Risk based tools such as SCIMAP FIO and VIPER have been used to highlight critical source areas in data limited catchments (Oliver, 2018, Porter, 2017) building on earlier diffuse pollution risk mapping frameworks (Reaney, 2011). Statistical approaches – including spatial stream network models (Neill *et al*., 2019), Bayesian modelling (Bertone *et al*., 2019, Glendell, 2021, Mzyece *et al.*, 2026, Pascual-Benito, 2020) and machine learning methods – have helped to characterise spatial and temporal patterns in FIO concentrations. Their transferability, however, is often constrained by the availability and quality of monitoring data (Heal, 2025, Oliver *et al*., 2016). The need for a screening model is driven by the cost, logistics and difficulty in generating sufficient monitoring data to quantify variability, therefore the approach of choice will need to function within these constraints.

Process based models offer a more detailed representation of hydrology, mobilisation, transport, and fate. Widely used frameworks include SWAT (Arnold, 2012, Kim *et al*., 2010), HYPE (Lindström, 2010, Srinivasan *et al*., 2021) and the agent based MAFIO model developed for Scottish catchments (Neill *et al*., 2020a, Neill *et al*., 2020b). Hydrodynamic water quality models have also been applied extensively in rivers, estuaries, and coastal waters (Cea *et al*., 2011, de Brauwere *et al*., 2011, Eregno *et al*., 2018, Huang *et al*., 2015). Sediment related processes – such as storage and resuspension – are increasingly recognised as important components of FIO behaviour (Gao *et al*., 2013, Wu, 2009). More recent work has explored surrogate approaches, including neural network and other nonlinear AI models, to reduce computational demands while retaining predictive skill (García-Alba, 2019, Lam and Ahmadian, 2023, Lin *et al*., 2008).

Despite this progress, several knowledge gaps remain. These include the survival and export of FIOs in the wider environment (for example, the survival of FIOs in livestock manures in the Scottish climate during dry periods which can then be washed into surface waters during high rainfall events – a risk which is likely to increase due to climate change), the influence of hydrological connectivity, sub field scale mobilisation processes, and the resuspension of streambed associated FIOs (Cho, 2016, Neill *et al*., 2020a, Neill *et al*., 2020b, Neill *et al*., 2019, Oliver *et al*., 2016). Additional challenges relate to model calibration under limited monitoring data (de Brauwere *et al*., 2009, Heal, 2025) the integration of microbial source tracking (MST) (Kelly, 2024, Myzece, 2025), and the need for open, interoperable modelling frameworks (Laniak, 2013, Oliver *et al.*, 2016). These issues are relevant in Scotland, where complex mixtures of diffuse and point sources interact with variable hydrology and land use.

Against this backdrop, SEPA requires clarity on which FIO models are currently available, how they differ, how they might be combined, and what developments would be needed to support catchment scale source apportionment. This project therefore aims to provide a concise, structured overview of existing FIO models for *E. coli* and related indicators, assess their potential for integration, identify gaps, and evaluate the feasibility of developing a high level catchment model capable of identifying the most significant pressures under varying hydrological and seasonal conditions.

This will provide key players including SEPA, Scottish Water, NatureScot, Food Standards Scotland, Scottish Government and the shellfish industry with a stronger evidence base to promote the development of models which enhance understanding of catchment to coast FIO dynamics. This in turn will help identify where to target investment and will improve risk-based decision making. Further, it will help to develop shared source to sea insights among key stakeholders to support regulatory management, promoting resilience and coordination within the water sector.

1.2 Project objectives

The overall aim of this project was to provide a concise overview of available FIO models for *E. coli*/FIOs and to explore where elements could potentially be integrated to support catchment-scale source apportionment.

The objectives were to:

1. Determine which *E. coli*/FIO models are already available
2. Determine whether these can be integrated to form a holistic high-level catchment understanding to identify most significant contributing sources under various conditions (seasonal land use changes, high and low flows)
3. Identify changes that would be required to allow these models to be combined
4. Identify the modelling gaps
5. Evaluate how feasible it would be to fill these gaps
6. Understand how feasible it is to create a high-level catchment model for *E. coli* to identify most significant pressures.

1.3 Structure of the report

The report introduces the approach applied (with further detail provided in the appendices) and first summarizes some of the main water quality models (in no specific order) that the research team considered to have merit for future application. These were either identified by the team and their networks or through the literature review or both. Other models identified in the literature review are then discussed briefly in terms of why they were not reviewed in further detail.

We then discuss strengths and complementarities of the models, including the opportunities and challenges of integration and lessons learned from previous integration efforts. Key modelling gaps are identified and suggested actions towards developing a coherent modelling framework for FIO screening in Scotland is presented. This is followed by specific recommendations and conclusions.

2.0 Research undertaken

A rapid evidence review was undertaken to identify relevant literature which was then interrogated to provide an evidence database (excel file) to allow the team to address the objectives. Inclusion and exclusion criteria were co-constructed with the project steering group (PSG). In addition to systematic database searches, the literature review was expanded by employing a “snowballing” approach, whereby references cited in key papers or papers otherwise identified by the team were followed iteratively to capture additional relevant studies and emerging insights that may not be identified through structured searches alone. The literature review was complemented by the research team’s prior work and research networks. Further, the research team engaged with key stakeholders who use or have been involved in development of relevant models. Detailed methodology is provided in Appendix 1.

3.0 Findings

3.1 Overview of Models

3.1.1 SIMCAT/SAGIS

This model was not identified in the literature review but has been widely used by regulators, primarily as a planning tool for water quality management and permitting. It has mainly been applied to water quality parameters other than FIOs, such as phosphorus and nitrate concentrations. The SIMulation of CATchments model (SIMCAT), developed by the Environment Agency (EA) and further advanced by the Water Research Centre (WRc), is primarily used to simulate river water quality and has been applied operationally for over two decades (Oliver, 2011). It allows multiple data layers to be overlayed so that geo-referenced spatial data can be incorporated into the framework (Crowther and and Douglass, 2016). SIMCAT is widely used alongside SAGIS (Source Apportionment GIS) to evaluate the impacts of discharges or interventions. SIMCAT is available as open source and is supported by comprehensive documentation, including a user manual (available at: https://github.com/DEFRA/SIMCAT/tree/main/docs).

SIMCAT is a one-dimensional, steady-state model with a deterministic structure and stochastic inputs (Darji, 2022, Ranjith, 2019, Warn, 2010). It adopts a simplified “data-matching” approach, whereby water quality processes are represented using empirical decay and reaeration rates rather than detailed mechanistic formulations. Model inputs are defined as statistical distributions (e.g. mean, standard deviation, and sample size), and simulations are performed using Monte Carlo methods to represent variability and uncertainty in observed data (Oliver, 2011, Ranjith, 2019). SIMCAT can use a monthly time-step, however this is not usually required and annual data is widely used (Warn, 2018). Although a relatively simple model, it is extensively used for source apportionment and targeting measures.

In SIMCAT, the river is divided into user-defined reaches, each assumed to be completely mixed with respect to pollutants. A mass balance is applied at the head of each reach, combining flows and pollutant loads from point and diffuse sources, and routing these downstream while applying first-order decay processes where appropriate (Oliver, 2011). Personal communication with the developers/users indicated that much time has been dedicated to calibrating SIMCAT for phosphate and this is where the model has been most significantly progressed. However, the model can simulate multiple pollutant types, including nutrients, dissolved oxygen (DO), biochemical oxygen demand (BOD), ammonia, and phosphate (Ranjith, 2019, Warn, 2010), distinguishing between conservative (e.g. chloride) and non-conservative (e.g. BOD, ammonia) substances, as well as dissolved oxygen dynamics influenced by reaeration and BOD decay (Oliver, 2011). Interactions with sediment are not explicitly represented (Cox, 2003).

The principal input data requirements are statistical summaries of river flow and water quality at headwaters, tributaries, and discharge points, typically derived from routine monitoring datasets (Oliver, 2011, Darji, 2022, Ranjith, 2019, Warn, 2010). Uncertainty in model outputs is primarily associated with variability and limitations in observed data, rather than structural representation of processes (McIntyre, 2004).

For example, the uncertainty with respect to septic tank loss estimations was noted by Crowther and Douglas (2016).

SAGIS quantifies pollutant loads to surface waters from a range of point and diffuse sources, including wastewater treatment works discharges, intermittent discharges from sewerage systems and runoff, agriculture, soil erosion, mine water drainage, septic tanks, and industrial inputs (Catchment Based Approach, 2026). These loads are converted into in-stream concentrations using the SIMCAT water quality model, which is embedded within SAGIS. This integration enables the contribution of individual sources to river concentrations to be quantified and supports the assessment of pollution control measures and monitoring strategies.

SIMCAT–SAGIS has been widely applied to evaluate nutrient dynamics and the impacts of mitigation measures; however, there is limited evidence of its application to FIOs. This is partly due to key limitations of the modelling framework, including its steady-state structure, lack of temporal resolution for dynamic or event-based processes, simplified representation of in-stream processes, and absence of sediment interactions (Cox, 2003). The input of annual statistics does not allow for an account of storm-based responses and while this can be improved through the incorporation of monthly statistics, to do so equates to a significant increase in data requirements while still not necessarily providing the event-level information sought for a predictive model of microbial pollution. As such, the model is less suitable for simulating microbial water quality, particularly where processes such as sediment association and storm-driven variability are important (Oliver *et al.*, 2016). The high variability of FIOs is not well represented by monthly or yearly averaged data. Developing a source apportionment framework for chemicals, Comber et al. demonstrate a typical cartographic output from SIMCAT–SAGIS modelling (Comber *et al.*, 2018) (Figure 1).

Personal communication with model users identified that an additional limitation is the significant cost of ArcGIS Pro licences and the challenges associated with testing SIMCAT SAGIS each time a new ArcGIS Pro update is released. The testing is undertaken by the EA Operations Services National Team each time a new update is required, and fixes are then applied by a consultant.

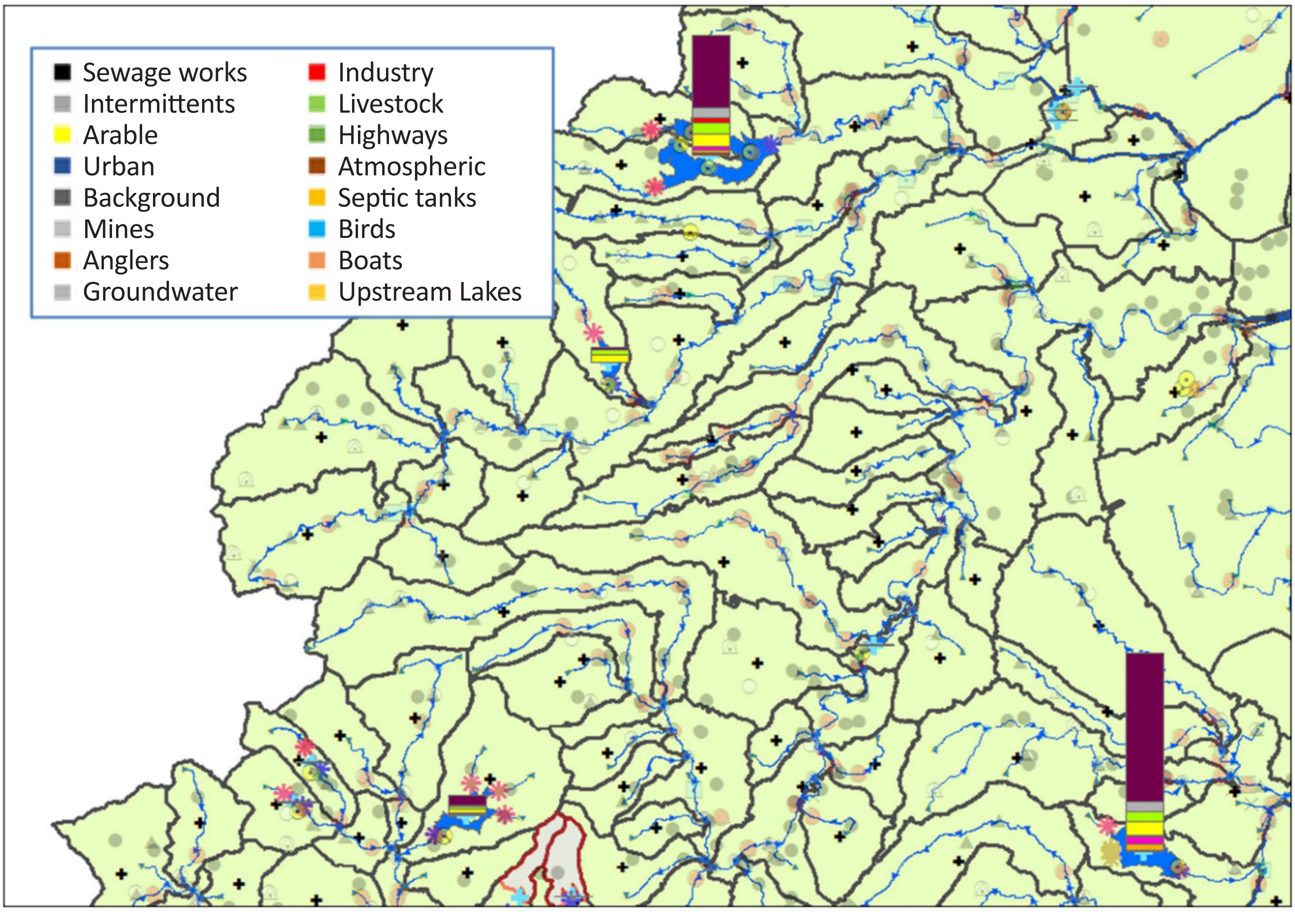

Figure 1. Cartographical output with sector bar charts demonstrating sector input loadings to lakes (Comber *et al.,* 2018).

It is the opinion of the research team that if a new national modelling approach were to be developed for Scotland, the SIMCAT–SAGIS framework provides a number of useful lessons. Its main strengths lie in doing a small number of things well: a clear and easily understood mass balance approach for routing loads through river networks; a consistent way of bringing together point and diffuse sources within the same framework; and close integration with GIS to produce spatially coherent outputs at national scale. Crucially, the model has been built around data that are routinely available for regulatory use, which has made it practical to operate, repeatable, and defensible over time. These characteristics have allowed SIMCAT–SAGIS to function effectively as a screening and prioritisation tool, even where data are limited (Comber, 2013). Any future redevelopment could sensibly retain this emphasis on simplicity, transparency and operational robustness, while addressing known weaknesses such as limited temporal resolution, poor representation of short term events and longer term software sustainability. Even so, updating and adapting the framework to meet current scientific and policy needs would still require a significant investment of effort.

3.1.2 MAFIO

MAFIO is an agent-based hydrological-environmental model (Figure 2), written in a Python 3.6 programming language and uses a PCRaster module. The model was developed at Aberdeen University (Neill *et al*., 2020a, Neill *et al*., 2020b). The MAFIO model for simulating FIOs in streams has been developed to overcome limitations in existing models. It is a spatially-distributed, process-based model designed to simulate the fate and transport of agents representing FIOs shed by livestock at the sub-field scale in small (<10 km^2) agricultural catchments. The model simulates FIO loading, die-off, detachment, surface routing, seepage, and channel routing on a regular spatial grid. It utilises a tracer-aided eco-hydrological model EcH$_2$O-iso which provides outputs to characterise the hydrological environment. The dynamics of stable isotopes (^{2}H and ^{18}O) yield information regarding velocities of water passing through a catchment to represent hydrological connectivity, storage and mixing. The open source code for EcH$_2$O-iso is available at: https://bitbucket.org/sylka/ech2o_iso/src/master_2.0/

MAFIO focuses on a catchment environment and a hydrological environment, using a daily timestep. The dynamic attributes of the spatial grid include probability of exfiltration, probability of infiltration,

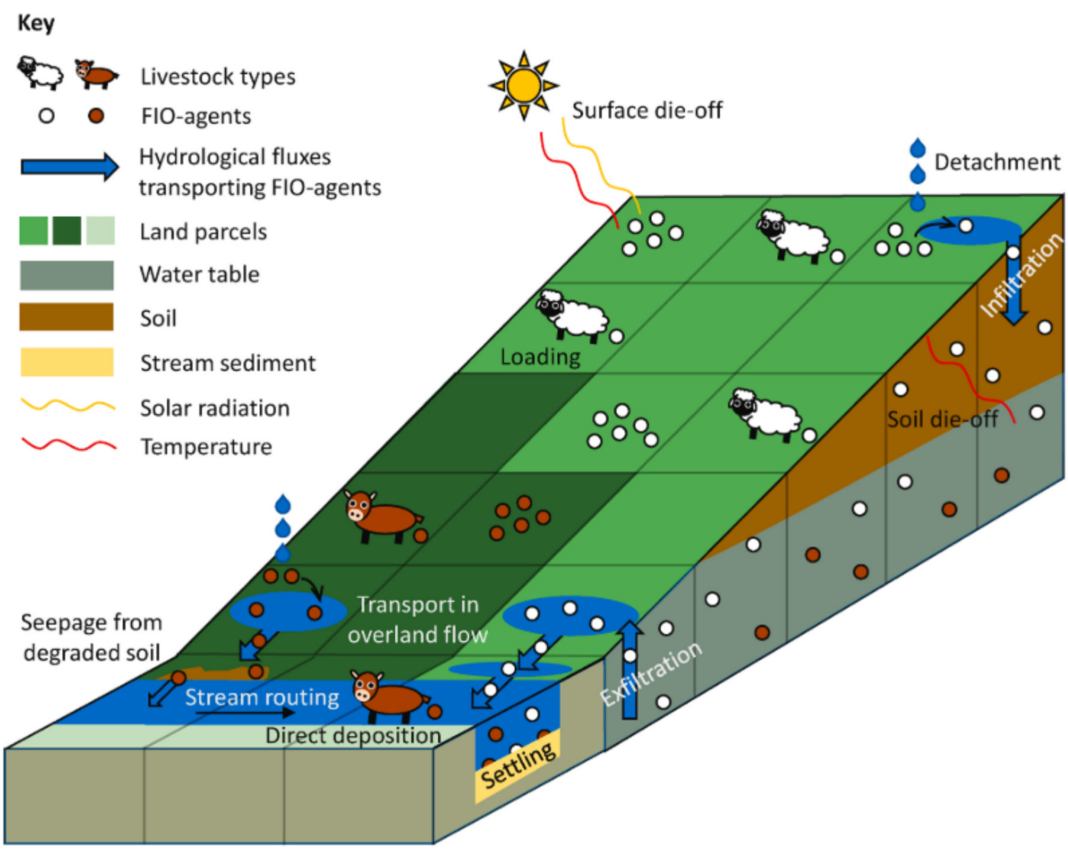

Figure 2. A conceptual diagram of MAFIO (Neill et al., 2020a).

probability of seepage and fraction of degraded soil (the latter influences seepage of FIOs) and are calculated for each timestep.

In Neil et al. (2020b), MAFIO was applied at a 0.42km^2 sub-catchment identified as a “hot spot” for *E. coli* contamination where livestock were considered the primary source of faecal contamination. Based on a summertime sampling campaign, the model reliably reproduced observed *E. coli* patterns, effectively capturing livestock derived FIO transfers to streams through direct deposition, overland flow, and seepage from degraded soils (Figure 3). The role of seepage faces and saturated zones is important and may need better representation in other models. The ability to map and predict flow connectivity and relative soil wetness could be important to identify management options and could be better represented in future models. For example, GIS based risk models may benefit from a small addition to reflect flow patterns seen in MAFIO. MAFIO highlighted important gaps in the understanding of *E. coli* behaviour in the environment, in particular the importance of sub-field scale processes such as where livestock congregate, microtopography and soil moisture patterns that strongly influence *E. coli* delivery to streams. It can improve representation of hydrological processes; however, its use is likely to be restricted to research purposes rather than as a screening tool for regulatory application, primarily because it requires high resolution, site-specific data (e.g. maps of catchment spatial extent, cells upslope of channel, cells upslope of degraded soil, channel location, channel width, degraded soil locations, distribution of land parcels, flow directions, livestock type and number over time, livestock stream access over time, land parcels associated with degraded soil, land parcels either side of stream) which is usually unavailable for most catchments.

MAFIO should be seen as useful research tool but not practical for upscaling. While it represents connectivity of FIOs to streams via soil and overland flow, it cannot be used alone for source attribution. However, it could contribute to an improved national scale model by strengthening understanding of processes influencing movement of *E. coli* from soils to water.

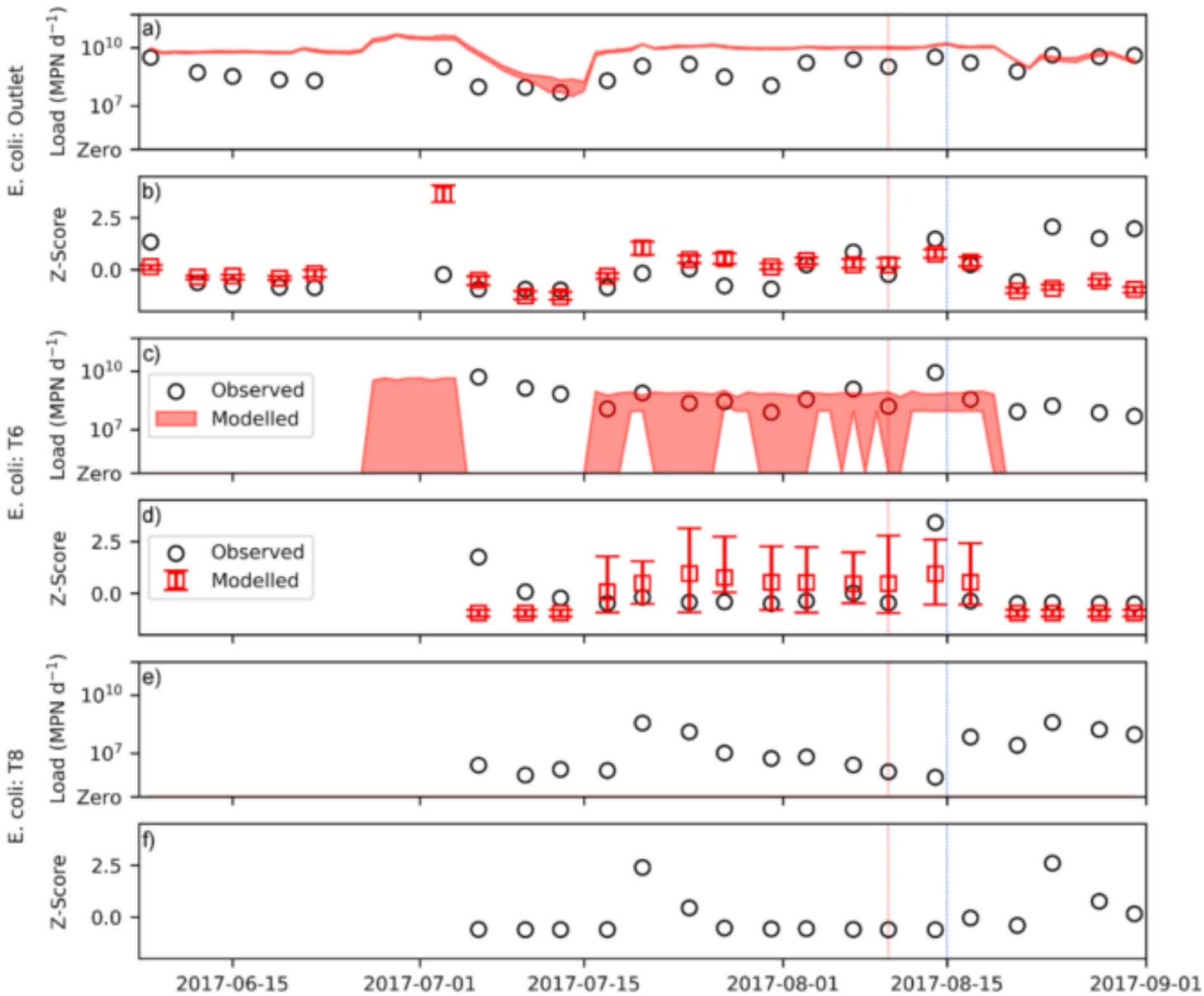

Figure 3. Observed vs. simulated *E. coli* loads and Z scores for a-b) catchment outlet; c-d) (T6 – shallow subsurface flow; e-f); T8 - deeper sub-surface/groundwater (Neill *et al.*, 2020b).

3.1.3 SCIMAP–ViPER (SCIMAP-FIO)

SCIMAP is a risk-based, spatial modelling framework designed to identify areas within a catchment that are likely to generate and hydrologically connect diffuse pollution sources to the river network, rather than to simulate in-stream water quality processes (Figure 4). It is therefore not a conventional water quality model (cf. SIMCAT), nor a rainfall–runoff model, but instead a tool to prioritise locations for management intervention based on relative risk. SCIMAP operates using readily available data, typically including high-resolution digital elevation models (DEMs) and land cover information, and produces spatially distributed maps of pollution risk within a catchment.

The underlying hydrological concepts in SCIMAP are based on digital terrain analysis. In particular, the model uses the topographic wetness index (a/tanβ, where a is the upslope contributing area and β is the local slope) to represent soil wetness propensity, coupled with a network index describing hydrological connectivity to the channel network (Lane, 2009, Reaney, 2011). Overland flow routing is represented by assigning a travel time to each grid cell based on flow path length and estimated velocity. Connectivity is determined using a network saturation concept, whereby a source area becomes hydrologically connected when a threshold condition is exceeded along the flow pathway. Each grid cell is therefore characterised by local saturation, network saturation, and travel time, which together define hydrologically similar zones and enable identification of critical source areas for diffuse pollution.

SCIMAP was originally developed to assess risks associated with fine sediment and phosphorus transfer, with more recent developments extending to nitrogen and flood risk. Model outputs are typically expressed as relative risk within a catchment, rather than absolute pollutant loads, although calibration approaches can be applied to derive more quantitative estimates.

Linking SCIMAP with the Visualising Pathogen & Environmental Risk (ViPER) model extends this framework to FIOs. In this adaptation, the hydrological connectivity component was retained unchanged, ensuring consistency with the original network index approach. However, the source risk layer based on land cover was replaced with a spatially distributed estimate of FIO loading, represented by *E. coli*. The *E. coli* source layer was derived from spatial data representing faecal loading potential, such as livestock distribution and faecal deposition rates, to estimate spatially distributed *E. coli* loading from grazing livestock, based on livestock density and assumed grazing regimes. These data were used to assign relative FIO source loading values to each grid cell. Oliver et al. (2018) detail how this data layer was generated. This forms the layer that then integrates with SCIMAP. The model simulates *E. coli* accumulation on pasture as a function of daily faecal inputs and persistence from previous days, represented using first-order die-off (Figure 4).

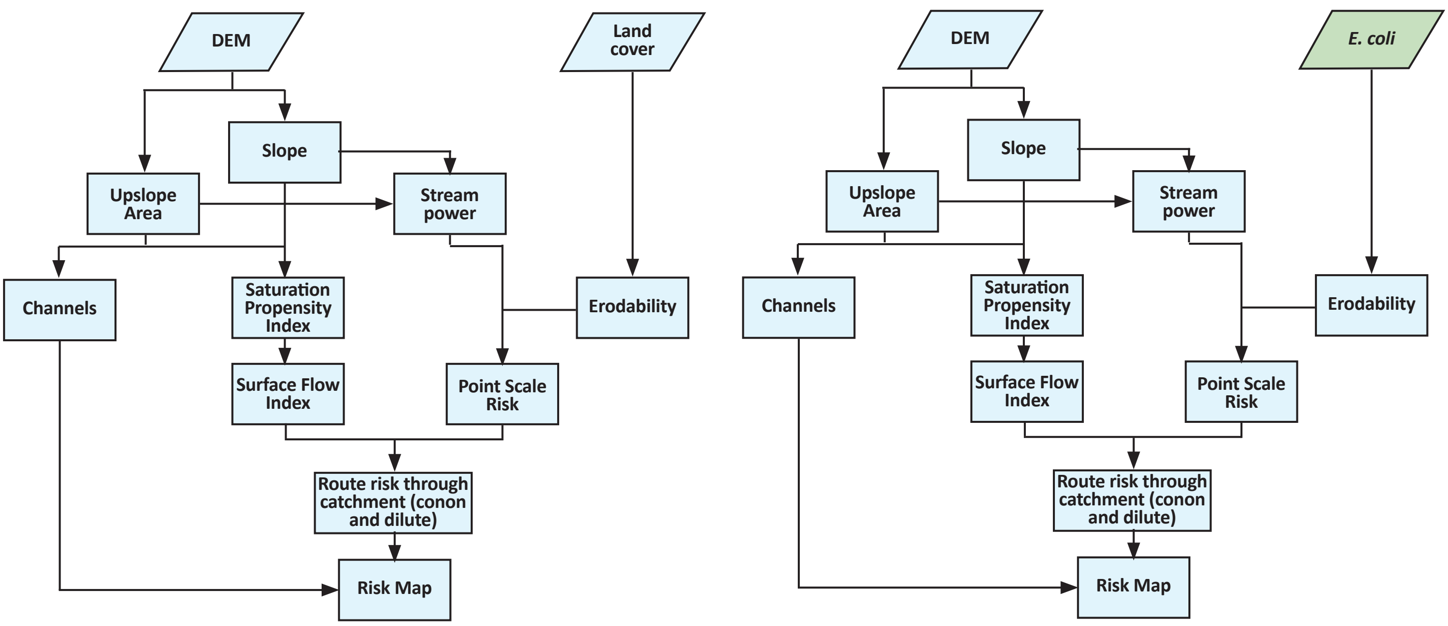

Figure 4. Conceptual model of SCIMAP-FIO (Oliver, D – pers. comm. 2026).

In application, SCIMAP provides the spatial structure of hydrological connectivity, identifying where pollutants are most likely to be mobilised and delivered to the river network, while ViPER quantifies the magnitude of microbial loading from land surfaces. This combined SCIMAP–ViPER framework therefore enables spatially explicit assessment of runoff-driven *E. coli* transfer across catchments (Figure 5).

In Oliver et al. (2018), the ViPER model was applied to two fully distributed catchments and run at a monthly timestep over one year. Results indicated significant seasonal variation in *E. coli* transfer, with higher predicted loads in summer and autumn. Standard SCIMAP-FIO model runs output monthly risk profiles rather than reflecting short-term or extreme events; however, it is possible to upload alternative datasets to replace the existing rainfall data layer, therefore the time-step could potentially be shortened.

The SCIMAP–ViPER approach offers a useful framework for identifying critical source areas and supporting risk-based management of microbial pollution. Strengths include its simplicity for end-users through lack of over-parameterisation, an easy-to-use web-interface (while any model can be communicated by a web interface, not all have such implementation or are intuitive) and the participatory model co-development with key stakeholders such as the EA, SEPA, Natural England, Rivers Trusts and Drinking Water Quality Regulator provided transparency and ensured development appropriate to end-user needs. However, the model relies on simplified hydrological connectivity metrics, does not explicitly represent in-stream processes and has limited ability to simulate event-driven dynamics. In addition, the requirement for appropriate livestock and land management data, and the need for statistical or empirical parameterisation of microbial die-off, introduce further uncertainty in predictions. Importantly, the model can represent uncertainty – the under-pinning ViPER model can include estimates of uncertainty around FIO loading to land, which can be propagated through the model.

Given some existing knowledge and expertise, the SCIMAP could be updated using existing GIS data, with some novel mobilisation and connectivity terms to set within a national scale framework. As for MAFIO, SCIMAP-FIO only accounts for agricultural sources of FIOs and therefore would require integration with additional models to support full source apportionment.

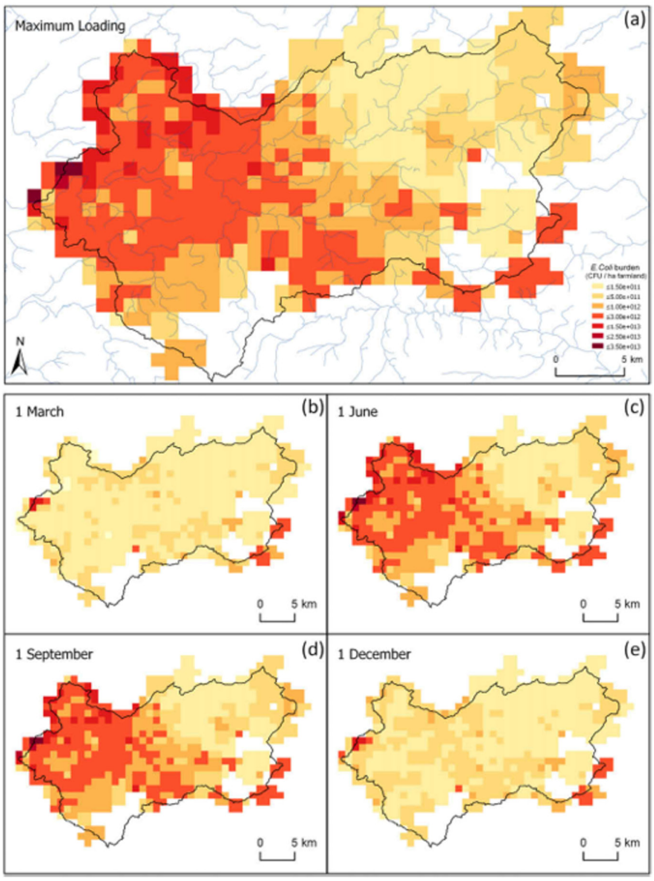

Figure 5. After Oliver et al. (2018). Modelled patterns of maximum *E. coli* burden per ha of farmland distributed across the River Ayr catchment at 1 km² grid resolution a) peak occurrence; b,c,d,e) snapshots in time.

3.1.4 Hydrological Predictions for the Environment (HYPE)

HYPE is a semi-distributed hydrological (rainfall-run-off) catchment model that simulates water flow and substance transport from precipitation through storage compartments and fluxes, to sea (SMHI, 2026a). The model has been applied globally and is most commonly used to predict flooding, drought, water quality and impacts of climate and hydro-morphological changes. It is used as a core model by the Environment Agency and was further developed by JBA consultants to specifically include FIOs and has been applied to evaluate catchment management measures in England (Hankin, 2019) The model was originally designed with a daily timestep, which was used by Hankin et al. (2019) for FIO application in England, but can be used for sub-daily time steps from 1–12 hours (SMHI, 2026b). HYPE uses a large number of input parameters, including rainfall, climate, soil characteristics, land use and irrigation and is calibrated for particular hydrological processes using data from gauged river basins. The model considers diffuse FIO sources including dairy, sheep, beef and deer grazing at catchment or sub-catchment scale and can handle up to three point sources from a possible eight, including wastewater, urban and stormwater inputs (SMHI, 2026c) (this captures the predominant livestock inputs to Shellfish Waters in Scotland from sheep and deer). Predictions are short term in nature. HYPE addresses multiple sources of animal-derived pollutants and can tackle multiple types of pollutant. For example, Srinivasan et al. (2021) applied HYPE as a national scale framework to characterise transfer of Nitrogen, Phosphorous and *E. coli* from land to water during the bathing season. The FIO-related output is *E. coli* concentrations in run-off and drainage.

HYPE-CSF is an extension of the HYPE model framework (Figure 6), developed with input from JBA Consulting, to better represent catchment-scale sources and pathways of FIOs. HYPE-CSF adds microbial source terms and die-off coefficients, representing a domain-specific extension layered onto an existing hydrological backbone. A series of desk studies were undertaken to refine and improve CSF-HYPE *E. coli* model (Hankin, 2020). In one of the studies, modelling *E. coli* concentrations in the bathing water at Ilkley (Yorkshire), the authors showed that under low flow conditions, continuous discharges contributed >60% of *E. coli* load and that modelled concentrations were <1000 CFU/100ml. In contrast, at high flows, concentrations in excess of 1000 CFU/100ml (the “sufficient” level of bathing

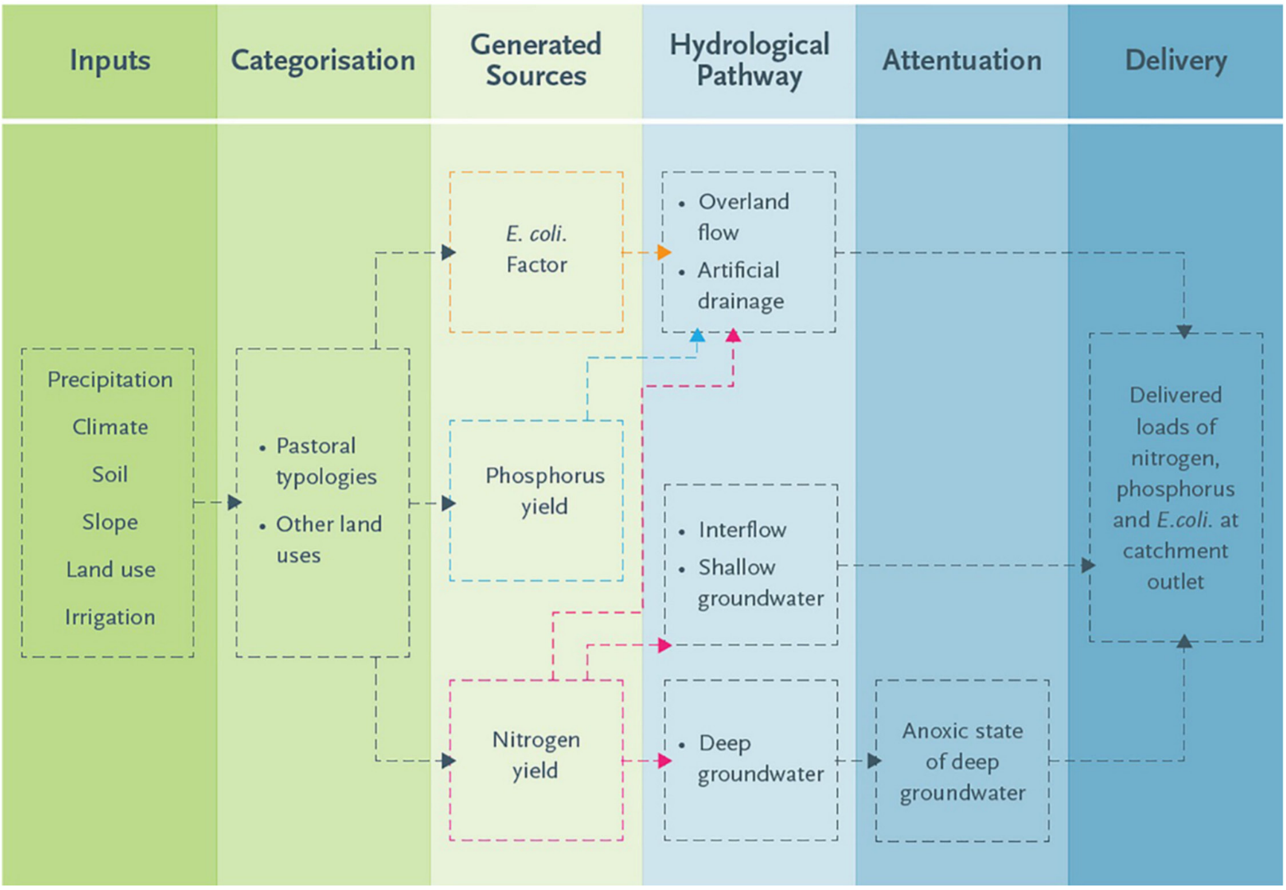

Figure 6. Schematic framework used to characterise transfers of N, P and *E. coli* from land to water, after Srinivasan et al. (2021).

water compliance standards is 500 CFU/100ml *E. coli* or fewer) occurred and agricultural sources were the greatest contributor to *E. coli* loads (mean of 65 %) (Hankin, 2023). This concurs with current understanding of source pressure changes with flow, and this trend is likely to be exaggerated in Scottish shellfish catchments where human sewage inputs are relatively low.

In contrast, in the Thames Basin, where human sewage inputs are more significant, dominating sources at the Wolvercote bathing water were sewage discharges. However, at very high flows, agricultural inputs increased by up to 25%. The authors noted that further sensitivity analyses and investigations of model parameters were required (Hankin, 2023). While HYPE-CSF is a useful tool, and is used by the EA, its application does further highlight the lack of available data on instream FIO concentrations within catchment and it is not particularly user-friendly. For example, a considerable amount of input files is required, resulting in large volumes of text files as outputs.

Following discussion with HYPE users, some experience of the research team using HYPE and interacting with the developers in the UK and Sweden, the research team concluded that, although a complex task, it could be possible to build a national scale HYPE model. Although development for England and Wales has taken a number of years by the EA, ongoing work at the James Hutton Institute is calibrating HYPE for low flows in 81 catchments in Scotland and this capability could, in the future, be aligned with developments in England (Glendell, pers. comm. 2026). Any development should augment and improve upon model weaknesses.

3.1.5 Soil and Water Assessment Tool (SWAT)

SWAT was developed by scientists at the United States Department of Agriculture and Agricultural Research Service (USDA-ARS) and Texas A & M University. It is a used to simulate surface and groundwater quality and quantity and to predict the impacts of land use, land management practices, and climate change, operating at scales from sub-catchment to river catchment scales (Soil and Water Assessment Tool, 2026). Multiple sub-basins (hydrologic response units) can also be joined up (Bieger *et al*., 2017). SWAT has been applied globally over the last 20 years to assess soil erosion, non-point source pollution control and regional catchment management. It has been modified into SWAT+ over recent years to facilitate model maintenance, collaborative integration of new science and future code modifications and incorporates a more flexible spatial representation of interactions and catchment processes (Soil and Water Assessment Tool, 2026). SWAT is a process-based, semi-distributed hydrological model operating at a daily time step and includes components such as weather, soil, hydrology, plant growth, nutrient cycles, and land management (Figure 7).

SWAT has been extensively applied in international research to investigate faecal indicator organisms, particularly across North America, parts of Europe and Asia (Bougeard *et al.*, 2011, Kim *et al.*, 2010, Shrestha *et al.,* 2014). In these settings, the SWAT bacteria module has been used in a substantial body of peer reviewed work to simulate catchment scale loads of *E. coli*, faecal coliforms and enterococci arising from agricultural, urban and mixed land use systems (Frey *et al.*, 2013).

Heal (2025) used the Soil and Water Assessment Tool (SWAT) to examine how FIOs, represented by Escherichia coli, vary across space and time within the Rivers Taw and Torridge catchments in Southwest England. Their approach coupled SWAT's bacterial fate and transport routines with catchment specific hydrological and land use data to simulate *E. coli* concentrations and loads arising largely from diffuse sources. Model outputs were compared with observed FIO monitoring data to evaluate how well SWAT could represent patterns of FIO pollution under UK conditions. The study demonstrated that while SWAT can provide useful insight into catchment scale FIO behaviour, its predictive performance is strongly constrained by the limited temporal resolution of routine microbial monitoring, particularly for capturing short lived but high magnitude contamination events that dominate overall FIO loads. An example of its application, Bougeard et al. (2011) combined SWAT with a hydrodynamic model (MARS 2D) to predict *E. coli* fluxes in sub-catchments impacting the Daoulas estuarine area in France. The models considered discharges from four wastewater treatment plants and *E. coli* arising from spreading of livestock manure (cattle, pigs, poultry.) Input data required were *E. coli* concentration in rivers, die-off rate at 20ºC, exposure time, flow, sub-basin run-off, topography with six land use classes, soil data (inc. texture) and rainfall. They found a strong relationship between modelled and measured *E. coli* concentrations for both river water and shellfish flesh.

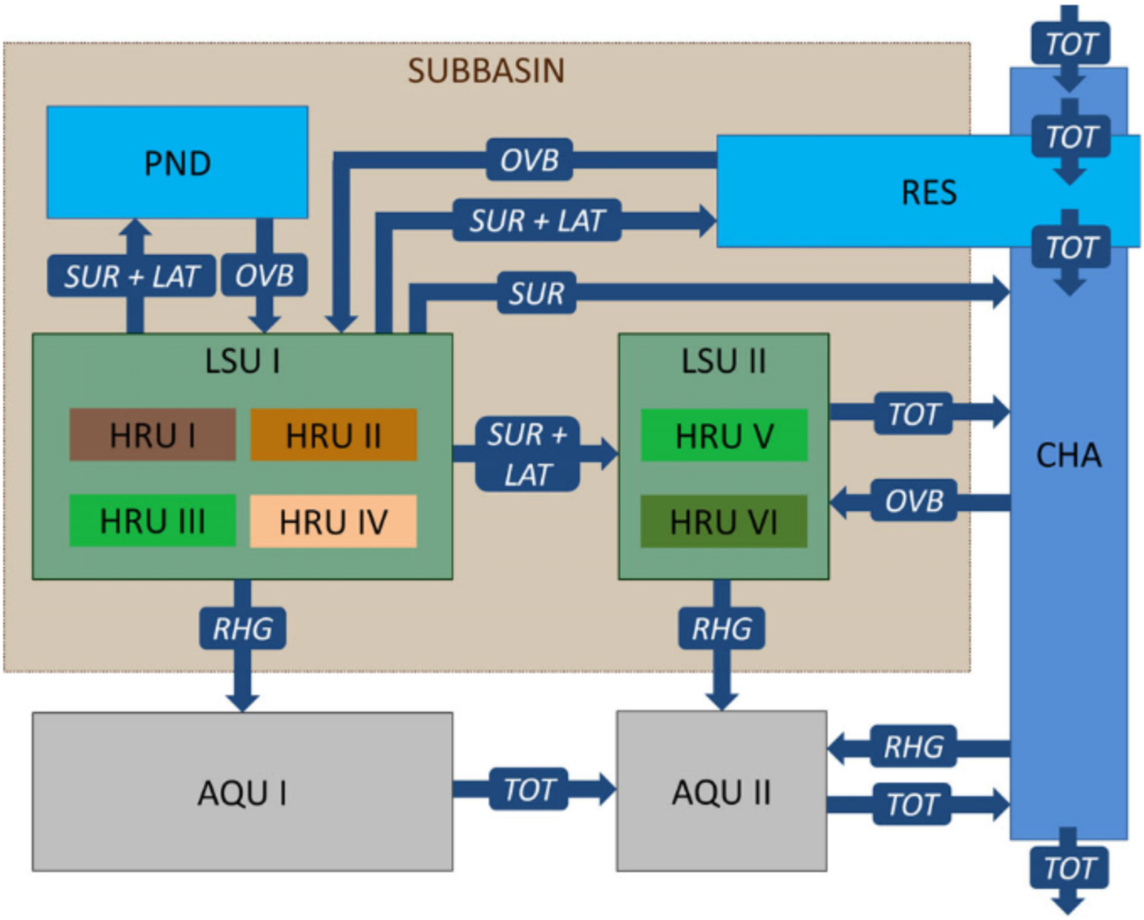

Figure 7. After Bieger et al. (2017). Conceptual Flow Paths of Water between the Most Important Spatial Objects in SWAT+. AQU, aquifer; CHA, channel; HRU, hydrologic response unit; LSU, landscape unit; PND, pond; RES, reservoir; LAT, lateral flow; OVB, overbank.

3.1.6 Coastal and Estuaries Models

In this review numerous and varied models were identified that had been applied to estuaries and coasts (Bougeard *et al.,* 2011, Cea *et al.*, 2011, de Brauwere *et al.,* 2011, Dumasdelage and Delestre, 2020, Gao *et al.,* 2013, Garcia-Armisen *et al.,* 2006, Wang *et al.,* 2026, Wolska *et al.,* 2022). Whilst movement and persistence of *E. coli* once they reach coastal waters is an important consideration, this is secondary to understanding the delivery of the bacteria to the coastal waters in the first place because it is the land-based apportionment and human faecal inputs that determine the loading. Thus, models predicting coastal or estuarine behaviour of FIOs must be coupled to catchment models. Where this is not the case, detailed monitoring of outflows into an estuary is required as demonstrated by Eregno et al. (2018; Figure 8) who integrated timing of flow and pollutant loadings into a bespoke Hydrodynamic model (Eregno *et al.,* 2018).

These hydrodynamic models use outputs from land-based and river network models as inputs (Huang *et al*., 2018, van Heijnsbergen, 2022, Wang *et al*., 2026). Each model has its own strength and different focus. For example, Wang et al. (2026) concentrated on the need to understand sediment recirculation to remobilise and extend the FIO problem in the estuaries. The most typical example is shown by Huang et al. where models are joined together (Huang *et al*., 2015, Huang *et al*., 2018) and applied to the Ribble catchment (Huang *et al*., 2015, Huang *et al*., 2018, Huang *et al*., 2017) In this case study, the Hydrological Simulation Program – FORTRAN (HSPF) (Bicknell, 1997) is used to simulate the land portion of the mobilisation. The Infoworks model represents pipes and sewerage networks including outfalls (Innovyze, 2022b, Innovyze, 2022a). The river network is simulated using a one-dimensional river network model (RMN1D) (Huang *et al*., 2017). Eventually these outputs feed into a hydrodynamic model called the modified EFDC 2D/3D model.

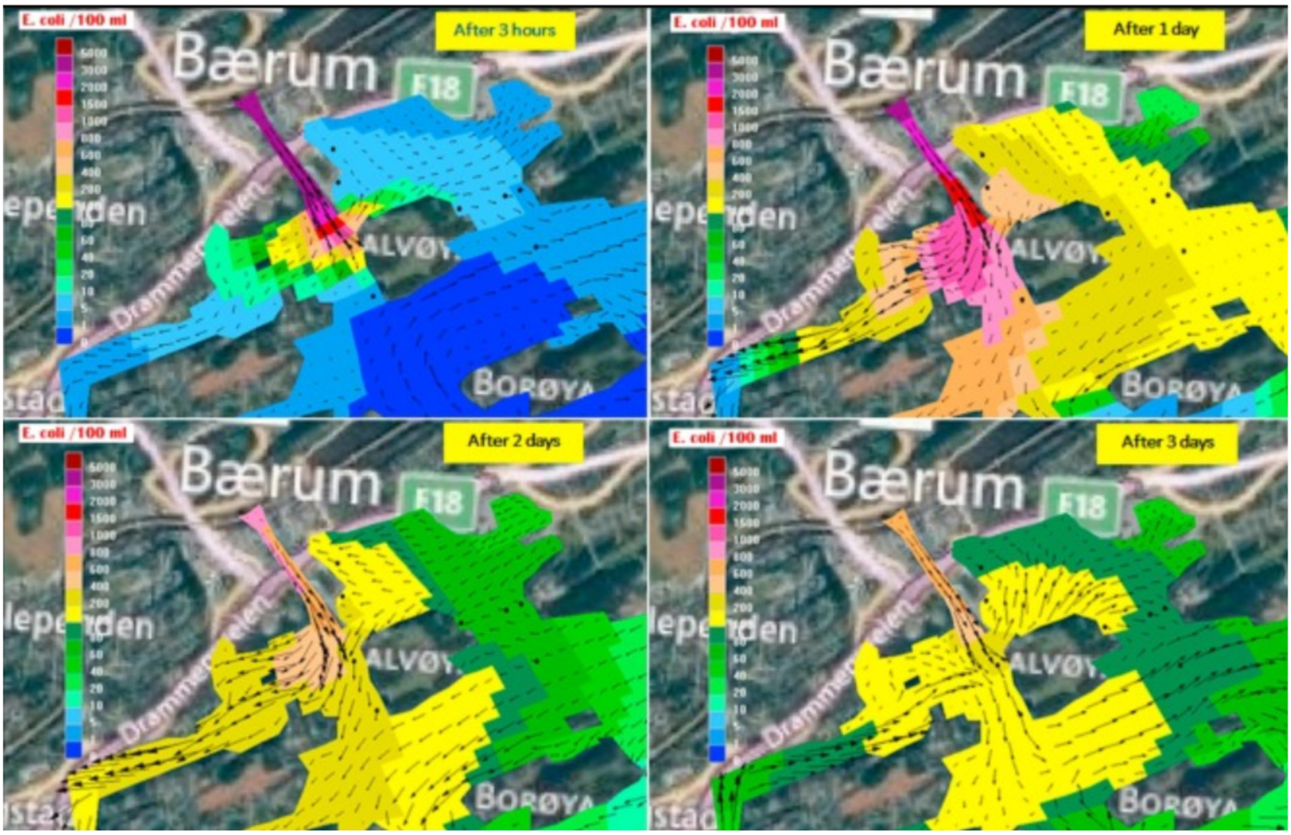

Figure 8. After Eregno et al. (2018). The top layer simulation graphics showing the spreading of *E. coli* at different time intervals after the rainfall event.

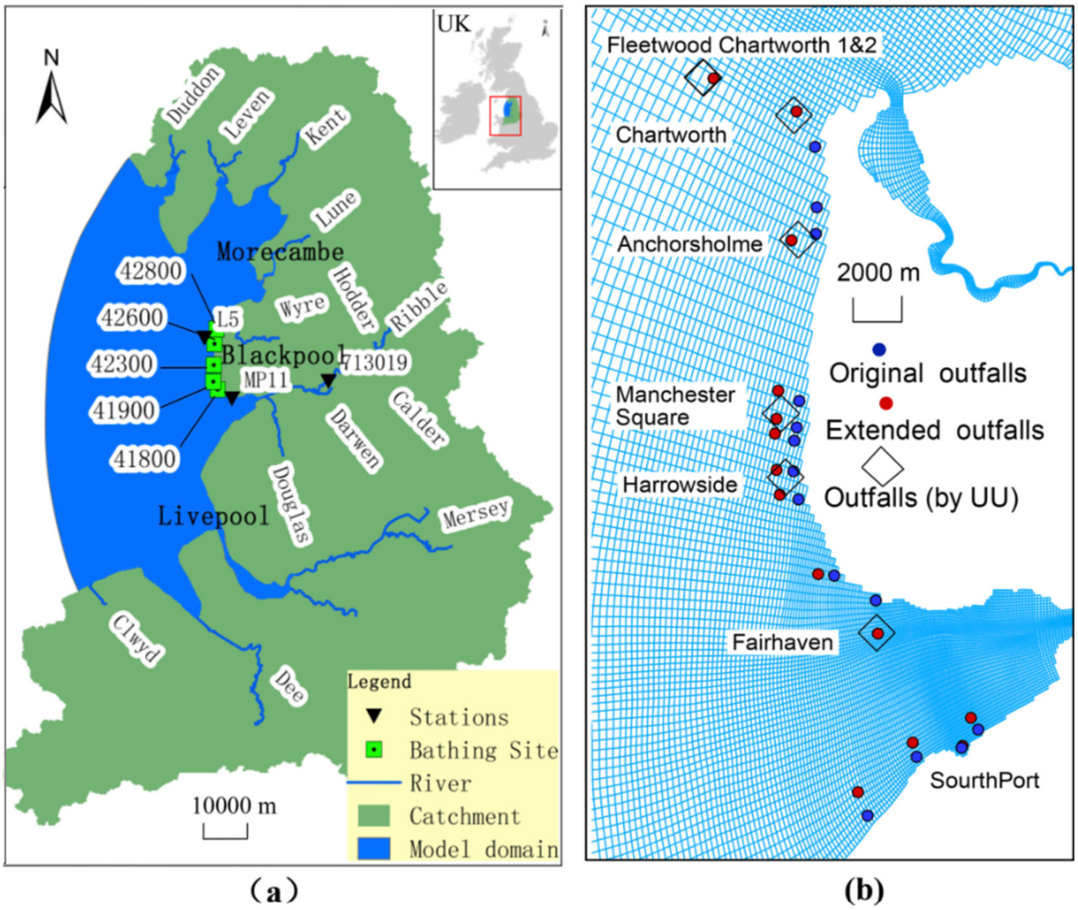

Figure 9. After Huang et al. (2017). (a) Map of the study area including: Monitoring sites and bathing water compliance sites for the rivers and coastal region; and (b) view of the nearshore region around the Ribble estuary, showing the original outfall locations.

This is an enormous but bespoke effort for this study site (Figure 9). Extensive work was undertaken to calibrate each component of the model, although some reservations about accuracy remain. In applying the model, the authors concluded that bathing water quality would be improved by extending sewer outfalls further out to sea.

The research team believe that it is important to improve the land and river network component of the models before addressing estuarine and coastal elements. The impact on downstream predictions will improve if these components are improved.

3.1.7 Bayesian approaches

Bayesian approaches did not feature strongly in the literature review, however they show promise as a framework to identify FIO risk hotspots within catchments (Schoen *et al*., 2010). Bayesian Networks (BN) are graphical models representing systems through networks of nodes (variables) and conditional dependencies (arrows). Probabilistic inferences can then be made, allowing uncertainty to be propagated directly in the modelling framework rather than treating it as a secondary issue (Troldborg, 2026, Schoen *et al*., 2010). Other advantages are that Bayesian approaches can be used to incorporate both qualitative and quantitative data including expert opinion. They can also be used in two ways - forward propagation to predict outcomes or backward propagation to determine variables which have the most influence. This latter capability is particularly useful for scenario analysis (Troldborg, 2026).

Schoen et al. (2010) demonstrated the use of Bayesian models to predict FIO concentrations in USA waterbodies with limited monitoring data. Two models were developed that predict FIO concentrations from flow. The approach differs from process-based and regression-based models in that the models presented by Schoen et al. (2010) make inferences about FIO concentration distributions directly from water quality observations rather than specifying parameters or explanatory variables to account for fate and transport processes.

Mzyece et al. (2026) developed a Bayesian Network (BN) framework to predict the risk of FIO pollution arising from septic tank systems (STS) in Scotland. Rather than attempting to simulate all hydrological and microbiological processes mechanistically, the approach integrates multiple lines of evidence probabilistically, including septic tank density, effluent movement risk, STS treatment levels and observed microbial monitoring data. The resulting framework provides a probabilistic estimate of pollution risk. Importantly, the model is explicitly designed to accommodate situations with sparse and uncertain data (Myzece, 2025).

There is substantial potential for the Mzyece et al. (2026) framework to be extended beyond septic tanks to incorporate additional FIO sources relevant to Scottish catchments. It is important to note that Bayesian network approaches are inherently modular (Chen, 2012). New datasets or source categories can be added incrementally without requiring redevelopment of an entire modelling framework. This is particularly attractive in an operational context where data availability and monitoring capability are expected to evolve over time.

3.2 Model strengths and complementarities

This review highlights that different modelling approaches exhibit distinct and complementary strengths. No single framework can represent the full source-pathway-receptor continuum that governs FIO dynamics in catchment systems. In practice, effective integration is less about achieving full system representation and more about selecting an appropriate level of complexity for the decision context. Catchment process models such as SWAT and HYPE are designed to represent hydrological processes, land-use driven runoff, and agricultural source mobilisation (Kim *et al*., 2010). However, they are not designed for resolving urban drainage and sewer network dynamics (Arnold, 2012, Lindström, 2010). In contrast, risk mapping and source targeting tools (e.g. SCIMAP-FIO, SAGIS) provide valuable insight into spatial risk weighting and source apportionment logic, but can lack fully dynamic hydrological representation. Depending on the model, these risk mapping tools may require adaptation to provide coverage of all source types (human, agricultural, wildlife). MAFIO is explicitly designed to track FIO loads from multiple sources and performs well at small landscape scales (Neill *et al*., 2020); however, its application to larger catchments is likely to be constrained by scalability and data requirements. Bayesian networks offer a conceptually different approach, particularly suited to data-poor systems, and enable the integration of both qualitative and quantitative information within a probabilistic framework (Mzyece *et al*., 2026).

3.2.1 Integrating model frameworks

There is potential to integrate modelling approaches, but this needs to be guided by a well-defined purpose and is accompanied by challenges. Model integration should not be viewed as a simple stacking exercise; rather, it requires deliberate architectural design aligned to the intended application (e.g., scenario testing, decision support, or process understanding). Several approaches to integration can be distinguished:

Sequential coupling:

In this approach, individual models are applied sequentially according to their strengths (Huang *et al*., 2018, Shrestha *et al*., 2014, Sokolova *et al*., 2012). For example, HYPE may be used to simulate runoff and agricultural loads; a septic tank model to estimate failure rates; a sewer model to represent combined sewer overflow (CSO) discharges; and SIMCAT (if further developed) to route FIO loads to compliance points. While conceptually straightforward, this approach presents several challenges, including mismatches in temporal resolution, incompatibilities in spatial discretisation, risks of double counting processes, and the propagation of uncertainty across model components. In addition, inconsistencies in parameter assumptions between models can reduce overall coherence. Sequential coupling is therefore often better suited to scenario exploration than to detailed mechanistic understanding.

Hybrid meta-model frameworks:

An alternative approach is to extract key functional components from existing models and embed them within a unified computational framework (e.g. in Python or R), rather than linking full model systems. For example, source loading functions, transport coefficients, decay rates, and hydrological drivers can be selectively integrated. This strategy reduces complexity relative to full process-based models while retaining essential system behaviour. It offers a pragmatic balance between realism and tractability, and may represent a more scalable route than full model integration (Razavi, 2012).

Bayesian network integration:

Bayesian networks can also act as an integration layer (meta model), structuring the system into nodes representing sources (e.g. agricultural, urban, wildlife), pathways (e.g. overland flow, drainage networks), and receptors (e.g. bathing waters). Evidence nodes (e.g. rainfall, seasonality, land use, population density) may inform probabilistic relationships between components (Bertone *et al*., 2019). This approach is particularly advantageous for handling uncertainty transparently, integrating heterogeneous data types, and updating predictions as new data become available (Glendell, 2021, Mzyece *et al*., 2026) . However, it is less well suited to representing high-frequency event dynamics and have previously used discretisation of variables into categorical states. One option as a way forward could be to include the agricultural pollution model from SCIMAP as a modified sub-module in the Bayesian Network model developed by Mzyece et al. (2026). FIO risk from sewage treatment works could be included as another sub-module, for example, modified from Troldborg et al. (2026). A CSO sub-module is due to be developed by Hutton in 2026 (M. Glendell, pers. comm. 2026). Hence, using attributes from SCIMAP and modifying existing Bayesian network models could facilitate a more rapid development of a risk based quantitative FIO model for Scotland at a monthly time step. Rainfall and runoff could be represented as a probability of daily rainfall intensity/total rainfall within each month. The modelling framework could be readily integrated in the SEPA SAGIS framework or served on a high-performance computing (HPC) cluster.

3.2.2 Technical challenges in integration

1. Structural and data integration

A central challenge in model integration is the reconciliation of temporal scales, both in terms of how models operate and the processes they are intended to represent. Differences in timestep resolution (e.g. hourly sewer dynamics versus daily catchment models) can lead to inconsistencies and information loss (Cho *et al*., 2010). Representation of decay and die-off processes introduces additional complexity, as different modelling approaches adopt simplified and often inconsistent assumptions regarding FIO survival (Cho, 2016). Estuarine and coastal processes are frequently the weakest component in source-to-sea modelling frameworks, despite their importance in determining exposure at receptor sites.

Uncertainty in source apportionment further complicates integration. Agricultural contributions are highly sensitive to management practices and stocking (Blaustein *et al*., 2016), septic tank failures are rarely monitored directly (Akoumianaki, 2022), and wildlife inputs remain poorly quantified (Afolabi, 2020, Oliver *et al*., 2016). Consequently, even well-integrated modelling frameworks may produce outputs that are dominated by uncertainty in source magnitude. Data harmonisation presents another significant barrier. Models often rely on differing land use classifications, hydrological routing schemes, and assumptions regarding connectivity. Achieving consistency across datasets and model structures can require substantial effort and may ultimately constrain integration feasibility.

2. Model sustainability, interoperability, and long-term operability

A critical, and often under-recognised, challenge in environmental modelling is not model capability but model sustainability. In many cases, the practical limitations of software architecture, interoperability, and long-term maintenance are as constraining as the scientific challenges of representing environmental processes. A key issue arises from differences in model architecture, programming languages, and underlying codebases (Laniak, 2013). For example, Shrestha et al. (2013) note that to work within the OpenMI platform, all models need to be OpenMI compliant. Effective integration requires that models are not only conceptually compatible but also technically able to communicate, which depends on interoperability at the software level. In practice, models are developed in diverse environments (e.g. Python, R, C++, and proprietary platforms), use different data structures, and rely on distinct dependency chains. These differences can make coupling difficult to achieve and even harder to sustain over time.

Even where integration is initially successful, maintaining functionality can be problematic. The SCIMAP-FIO framework provides a clear example: it successfully linked the SCIMAP risk mapping approach (Reaney, 2011) with the ViPER model (Oliver, 2018) to deliver a web-based interface for applied use, but subsequent licensing changes, software updates, and evolving security requirements ultimately disrupted communication between components and led to system failure. In particular, the need to comply with modern security standards, such as implementing HTTPS across web services (e.g. map servers and processing services), rendered parts of the system incompatible with institutional IT policies, resulting in withdrawal of hosting support.

More broadly, this reflects a systemic issue in environmental modelling. Many tools are developed within time-limited research projects and are not designed for long-term operational deployment (Neill *et al*., 2020b). As a result, they often rely on legacy dependencies, limited documentation, and tightly coupled code assembled from multiple sources. While such systems may appear functional at the user interface level, they are frequently difficult to maintain, extend, or transfer. Integrated modelling frameworks are particularly vulnerable in this respect, as they depend on multiple interacting components, external libraries, and, in some cases, proprietary software, increasing their fragility. In these situations, incremental maintenance can become inefficient, compared to redevelopment using modern, modular architectures (Lloyd, 2011). Approaches such as containerisation, improved version control, and clearer separation of model components can enhance resilience, but only when supported by adequate documentation and governance (Wilson, 2014). There is therefore a strong case for simplifying model implementations and embedding core functionality within widely used, well-supported platforms, such as GIS-based tools, rather than relying on complex, bespoke web-based model stacks. Ensuring that underlying datasets are openly accessible through curated repositories is equally important, reducing dependence on specific software environments and improving reproducibility (Laniak, 2013).

These issues also highlight important trade-offs between open and proprietary approaches. Open-source tools offer clear advantages in transparency, flexibility, and adaptability, particularly in a regulatory context where long-term use is required. However, they depend on sustained institutional support to remain operational. Proprietary systems may provide stability and formal support in the short term, but can introduce constraints in interoperability, transparency, and long-term flexibility. These challenges suggest that institutional and technical sustainability may be as limiting as scientific capability in determining the success of environmental modelling frameworks. Addressing these issues requires explicit consideration of software design, maintenance, and governance from the outset, rather than treating them as secondary concerns.

3.2.3 Lessons from previous integration efforts of FIO models (beyond sequential coupling)

1. Integrated deterministic modelling: Cloud2Coast

Previous large-scale research programmes have attempted to develop integrated modelling frameworks for FIO dynamics. A notable example is the Cloud-to-Coast (C2C) project (2011–2015), which aimed to predict pathogen exposure and associated health risks in nearshore waters.

Although the resulting framework (Huang *et al*., 2015) represented a significant technical achievement, its uptake has been limited. This highlights a key lesson: highly integrated, end-to-end models are often difficult to operationalise due to their complexity, data demands, and lack of flexibility. Future progress may therefore depend on adopting more modular, adaptable integration strategies that balance scientific rigour with usability and decision-making relevance. While full integration of different models is technically possible, in practice it can become a research programme in itself rather than a model tweak, and timelines for delivery can extend over multiple years. Key risks to manage include: (i) over parameterisation; (ii) data quality limitations; (iii) inconsistent spatial scales; (iv) uncertainty compounding; (v) model defensibility to regulators. In addition, maintenance and calibration burden increases dramatically.

2. The SUCCESS framework

A more recent example of an integrated modelling framework is the SUCCESS (System for Understanding Carrying Capacity, Ecological, and Social Sustainability) model, which has been applied in Dundrum Bay (Ferreira, 2023). SUCCESS adopts a catchment-to-coast approach, linking terrestrial loading, hydrological transport, coastal hydrodynamics, biogeochemical processes, and ecological responses within a unified modelling environment (Ferreira *et al*., 2023). The framework attempts the simulation of multiple interacting pressures, including land-based nutrient and microbial inputs, short-term discharge events, and their impacts on coastal water quality and shellfish systems. A key strength of SUCCESS lies in its ability to represent cross-domain interactions, such as the influence of catchment-derived loads on coastal eutrophication and aquaculture productivity, effectively functioning as a “digital twin” of the land-sea continuum. Further, it can be used to explore scenarios and can handle event-based dynamics. The framework uses the SWAT model as the driving hydrological model. This is in itself a large, complex model with many assumptions and some hydrological modelling problems (Quinn, P – pers. comm., 2026, Tan *et al*., 2020)

However, this level of integration introduces several challenges. The framework is inherently complex, requiring the coupling of multiple sub-models (e.g. hydrological, hydrodynamic, ecological and microbiological components), each with distinct parameterisations, spatial discretisation, and data requirements. This increases calibration burden and can obscure process understanding due to the high degree of internal model interaction. In addition, the computational and data demands associated with such fully integrated systems can limit their transferability to other catchments. While SUCCESS is capable of representing high-impact, short-term events (e.g., pulse discharges), capturing these dynamics reliably depends on the availability of high-resolution input data, which is often lacking. More broadly, the framework illustrates a key limitation of end-to-end integration: although scientifically powerful, such models can be difficult to operationalise for routine management and decision-making, reinforcing the trade-off between system completeness and practical usability. The SUCCESS framework represents the upper end of the integration spectrum, highlighting both the potential and the practical limitations of fully coupled, source-to-sea modelling systems.

3. SCIMAP-FIO

The SCIMAP and ViPER models were successfully integrated in the development of SCIMAP-FIO and this functioned as a web-based interface, allowing easy engagement with the underpinning models. As discussed above, SCIMAP-FIO encountered issues with functionality; however, the underlying model can still operate through script. A simple approach to rebuild SCIMAP-FIO for wider access and engagement would be to develop a QGIS plugin capable of linking all the data processing or to extend the current SAGA-GIS plugin. The key factor being to ensure that the underlying datasets are available to users. That could be as simple as archiving the monthly and maximum FIO burden data (the burden layer) onto a data archive, e.g., Zenodo, with accompanying how-to videos on making the risk maps in SAGA-GIS. To overcome the technical challenges around the reinstatement of the web-based interface would require significantly greater effort.

In addition, currently SCIMAP-FIO focuses primarily on agricultural sources of *E. coli* risk, with risk propagated through the landscape based on hydrological connectivity. Extending this framework to include human-derived sources (e.g., wastewater treatment works, combined sewer overflows, and septic tanks) would require the incorporation of point and near-point source inputs within the same connectivity-based routing structure. One approach would be to represent these sources as discrete nodes within the drainage network, each assigned a relative risk weighting based on discharge magnitude, frequency, and treatment performance. These inputs could then be propagated downstream using the existing SCIMAP framework, with appropriate attenuation factors applied to reflect in-stream decay and dilution. Temporal variability could be incorporated by conditioning source activation on drivers such as rainfall (for CSOs), system performance (for wastewater treatment works), and failure probability (for septic systems), thereby introducing a semi-dynamic component within an otherwise static framework.

However, several challenges arise in implementing such an extension. Data availability remains a key limitation, particularly for CSO spill frequency, septic tank location and failure rates, and high-resolution microbial monitoring at wastewater treatment works. There are also inherent scale mismatches between the grid-based representation of hydrological connectivity in SCIMAP and the network-based structure of sewer systems, requiring simplification or translation between spatial frameworks. The incorporation of temporally dynamic human sources into a predominantly static risk mapping framework introduces additional uncertainty, especially in representing episodic high-magnitude events that can dominate downstream microbial risk. This extension would move SCIMAP-FIO towards a more holistic representation of faecal pollution sources, while still operating within a relative risk framework; however, careful treatment of temporal dynamics and data uncertainty would be essential to avoid overinterpretation of results.

Recent implementation of real-time monitoring of combined sewer overflows by Scottish Water (Scottish Water, 2024a; 2024b) offers new opportunities to better characterise CSO dynamics, particularly in terms of spill timing and duration. Such datasets could be used to condition the activation of CSO sources within modelling frameworks, reducing reliance on assumed rainfall–spill relationships (Stapleton *et al*., 2008) and enabling more realistic representation of episodic pollution events. For example, (Locatelli *et al*., 2020) observed that while just millimeters of rainfall can trigger CSO flows, multiple rainfall events can merge into a single CSO-derived pollution episode, and pollution persists well beyond the cessation of CSO flows. However, the absence of concurrent microbial concentration data and uncertainties in discharge magnitude mean that significant challenges remain in translating these observations into quantitative FIO load estimates.

Emerging data-driven and AI-based approaches

In recent years there has been growing interest in data-driven and artificial intelligence approaches for predicting FIO concentrations (Vidal *et al.,* 2024). For example, initiatives such as River Deep Mountain AI (RDMAI; developed via the Water Innovation Fund) have explored the use of machine learning techniques to predict *E. coli* levels at bathing waters using environmental (land cover, weather, satellite) and monitoring data (Environment Agency Water Quality Archive – WIMS; *E. coli* samples from Natural Resources Wales and SEPA (RDMAI, 2025). These approaches offer several potential advantages. They can capture complex, non-linear relationships between drivers such as rainfall, river flow, and water quality without requiring explicit representation of underlying processes. They are also well suited to high-frequency datasets and can provide rapid, site-specific predictions, making them attractive for operational forecasting and public information systems; however, there are important limitations.

AI-based models are typically data-hungry and perform best where long, high-quality time series are available. They also tend to function as “black box” models, providing limited insight into underlying source contributions, which constrains their usefulness for source apportionment and management planning. In addition, their transferability between sites is often limited, as relationships learned in one catchment may not hold elsewhere. This was apparent during benchmarking of RDMAI against Wessex Water’s single site *E. coli* model (Slaughter, 2026). The authors also note that at this stage, RDMAI has relatively limited accuracy with a high proportion of false negatives (19-26% for the UK-wide dataset and 37-66% for the single-site model). As such, AI approaches are best viewed as complementary tools within a broader framework. They can enhance prediction and forecasting at specific receptor sites, particularly for bathing water compliance, but are unlikely to replace process-based or risk-based approaches for understanding catchment-scale sources and pathways. In the research team’s view, integrating these methods alongside traditional models and monitoring data may offer the greatest benefit, particularly where they can inform short-term risk management while more mechanistic approaches support longer-term intervention planning.

Graphic User Interface

A well-developed graphic user interface (GUI) is key and should be part of a web-based platform for any emerging model/decision support tool development. A web-based client server approach would be a priority, the rationale being: accessibility across different desktop operating systems (Windows, Linux, OSX), no installation of programmes being required, the underpinning model always being up-to-date because data are kept centrally, and potential to share data with other users via online permissions.

3.3 Gaps in modelling

1. Multiple, interacting sources:

Most impacted waters receive mixed inputs from different sources simultaneously. These sources can co-occur and are spatially and temporally variable, making attribution inherently difficult.

Modelling Gap 1: Generally, models struggle to disentangle overlapping FIO pollution signals when sources are co-linear (e.g., rainfall drives both CSO spills and agricultural runoff).

2. Source apportionment

There is a reliance on FIOs as non-specific indicators. FIOs are not source specific and this creates ambiguity. The same concentration could arise from human sewage or livestock. This is further complicated by mismatches in scale of understanding.

Modelling Gap 2: Traditional models using FIO loads cannot robustly partition human versus animal sources, limiting management decisions – upscaling introduces large uncertainty; fine-scale source signals get diluted or masked.

3. Process representation gaps in models

This includes event-driven dynamics, with FIO transfer recognised to be highly episodic via, e.g., rainfall-runoff mobilisation, CSO spills, sediment resuspension inputs; but also sediment-bacteria interactions and die-off variability.

Modelling Gap 3: Many models still rely on averaged conditions, failing to capture episodic peaks that can drive compliance failures or timing relative to bathing events or shellfish harvesting. Likewise, sediment-FIO interactions are often oversimplified, omitted or poorly parameterised. Models typically use constant decay rates, whereas real decay is highly variable and source-dependent.

3.4 Towards a Plan of Action for a Coherent Modelling Framework for FIO screening in Scotland

Effective management of faecal pollution affecting bathing waters and shellfish harvesting areas in Scotland requires a shift in emphasis from seeking a single integrated modelling solution towards building a coherent, evidence-led framework that combines models, monitoring, and regulatory practice. Here, we propose a ten-point plan to progress development of this approach.

Table 1: A ten-point plan for strengthening source apportionment modelling of faecal pollution in Scotland.	
1	**Align modelling outputs with management needs** Modelling must directly support regulatory and operational decision making. The priority is to identify dominant pollution sources under different conditions, highlight when and where risks are greatest, and assess the likely effectiveness of interventions. Models should provide value in guiding practical action.
2	**Prioritise interoperability and long term maintainability of modelling systems** Future modelling efforts should avoid overly complex, tightly coupled systems that are difficult to sustain. Scotland would benefit from modular, flexible modelling architectures that can evolve as new data, tools, and regulatory requirements emerge. Long term maintainability should be treated as a core design principle.
3	**Re-establish a national spatial risk screening capability** Scotland needs a robust, modernised national tool for identifying hydrologically connected, high risk areas—particularly for agricultural sources. Reinvestment in a platform such as SCIMAP-FIO and/or coupling it with Bayesian networks would establish a consistent national screening framework. This can be complemented by more detailed modelling in priority catchments where it will directly inform management actions. Updated datasets (e.g. high resolution DEMs, improved land use mapping) make this an efficient and future proof starting point. A version of SCIMAP-FIO incorporating erosion, pollution and natural flood management (NFM) may make this coupled model a good starting point as one of the easiest to innovate.
4	**Strengthen data sharing and governance across organisations** Stronger coordination in data sharing and governance is essential. Improved integration of datasets across SEPA, Scottish Water, local authorities, and other stakeholders, alongside greater standardisation of data formats and assumptions, will be fundamental to enabling more effective source apportionment. A broader discussion is needed now, about what Scotland wants in term of modelling capability, covering numerous pollution parameters, floods and droughts set within a climate change framework.
5	**Expand event based and high frequency monitoring** Short duration, high magnitude pollution events are a major driver of Bathing Water compliance failures, yet they are poorly captured by routine sampling. Increasing rainfall triggered, high frequency monitoring – supported by flow and turbidity data – will substantially improve model calibration and understanding of pollution dynamics.
6	**Integrate emerging wastewater monitoring data** New wastewater datasets, particularly real time information on combined sewer overflow activity from Scottish Water, offer a step change in understanding urban pollution pathways. These data should be systematically incorporated into risk assessments and modelling frameworks to improve accuracy and responsiveness.
7	**Improve the evidence base for poorly characterised sources** Key knowledge gaps remain for septic tanks, livestock management practices, and wildlife contributions – especially in sensitive coastal catchments. Better national information on septic tank distribution and performance, alongside improved understanding of livestock management impacts in saturated or riparian areas (as identified through MAFIO), will strengthen model inputs and intervention planning.
8	**Enhance representation of the land–sea interface** Estuarine and coastal processes often determine whether upstream pollution leads to regulatory failures under the Water Environment (Shellfish Water Protected Areas) (Scotland) Regulations 2013, yet they remain the weakest component of current modelling. For Scotland, an initial focus on the land component and river network feeding the estuaries will be a major area of improvement. Better integration of estuarine and coastal process could follow.
9	**Develop a national hydrological backbone model** Scotland should explore adopting a national scale application of the HYPE CSF model, already used by the Environment Agency, to provide a consistent hydrological foundation for runoff, flow pathways, and diffuse pollution mobilisation. Its role should be clearly defined as complementary – recognising that it will not resolve urban drainage or coastal processes – and the investment required should be carefully assessed. Selecting HYPE-CSF over SIMCAT and SWAT would be favourable in terms of model confidence and potential.
10	**Increase the use of probabilistic and statistical evidence integration tools** Approaches such as Bayesian networks can combine diverse datasets, explicitly represent uncertainty, and support adaptive, risk based decision making. These methods can help integrate behavioural, environmental, and operational factors, supporting management strategies that reduce agricultural mobilisation and address wastewater contributions.

4.0 Recommendations

Final Recommendations: A pragmatic and operational approach

The preceding ten-point plan sets out a range of options for improving modelling approaches. However, given the constraints identified, particularly around data availability, system complexity, and long-term sustainability, it is necessary to define a more focused and deliverable path forward achievable in the near to medium term.

A key starting point is to clarify the primary purpose of any modelling approach. Rather than seeking comprehensive representation of all sources and processes, the priority should be to support routine decision-making, specifically the identification of the most significant sources of faecal pollution under conditions relevant to compliance risk. This implies a shift away from attempting full source-pathway-receptor integration, towards developing a tool that is fit-for-purpose, transparent, and operational over the long term. Tools must be designed for routine operational use, not one-off research applications, and should be capable of supporting screening, prioritisation, and progression to more detailed investigations where required. This also requires explicit consideration of software sustainability, including maintainability, interoperability, and compatibility with evolving IT and security environments. Within this context, three priority actions are identified.

Recommendation 1 (R1): Undertake a national audit of data availability to define modelling scope

As a first step, a structured audit of available data should be undertaken to categorise:

1. data that are routinely available and sustainable long term (e.g. rainfall, river flow, land use, regulatory monitoring, wastewater asset data from Scottish Water)
2. data that could be improved through targeted investment (e.g. septic tank inventories, high-frequency microbial monitoring)
3. data that are unlikely to be available at sufficient resolution (e.g. detailed wildlife contributions).

This audit should then be used to define the scope and ambition of any modelling approach, ensuring that development is aligned with realistic data availability rather than theoretical capability. Thus, any modelling framework is designed around the first category, with optional extensions for the second, and without reliance on the third. This ensures that the tool remains operational and scalable.

Feasibility and cost:

High feasibility. This represents a logical next phase of work and can be delivered relatively quickly through coordination across organisations. Costs are low to moderate, primarily associated with staff time and data collation.

Recommendation 2 (R2): Develop a simple, modular screening tool for operational use

Based on the outcomes of the data audit, the priority should be to develop (or redevelop) a national-scale screening tool to support routine use by SEPA. Given the challenges with monitoring previously discussed, this tool could help with the investigation into causes of impact, particularly for shellfish waters where MST monitoring is of limited use. This investigation into impacts in SWPAs is required by The Water Environment (Shellfish Water Protected Areas: Environmental Objectives etc.) (Scotland) Regulations 2013. This tool should:

- focus on identifying relative source risk
- be designed using a modular architecture, allowing components to be updated as new data become available
- prioritise usability, transparency, and maintainability over model complexity
- focus on high-risk scenarios (e.g. rainfall-driven events) rather than long-term averages.

The tool should be framed explicitly as **decision-support**, guiding where to target investigations, monitoring, and interventions. While it can draw on established concepts, it should be rebuilt using modern, sustainable software practices to ensure national consistency and future resilience.

Feasibility and cost:

High feasibility. Development could be achieved within 18–24 months using existing datasets and established approaches. Costs are expected to be significantly lower than those associated with fully integrated modelling systems.

Recommendation 3 (R3): Establish a sustainable, long-term operational framework for modelling and data integration

To ensure longevity and value, modelling capability should be embedded within a broader operational framework that:

- supports routine use in decision-making, rather than one-off analyses
- incorporates uncertainty explicitly, enabling risk-based interpretation of outputs
- adopts open, modular, and well-documented systems to ensure maintainability and adaptability over time
- engages with users and stakeholders to understand any wider, future wider operational needs.

This should be supported by ongoing data improvement (where feasible) and clear governance arrangements for maintaining and updating both data and tools.

Feasibility and cost:

Moderate feasibility. While technically achievable, this requires sustained organisational commitment and coordination. Costs represent an investment in long-term capability and avoided rework.

5.0 Conclusions

No single modelling framework can currently fully resolve faecal contamination sources across the entire catchment to coast system. The most effective regulatory strategy is to prioritise **Recommendation 2:** developing a simple, 'smart' national screening tool built on the improved datasets identified in Recommendation 1. With appropriate investment, this approach offers the greatest near term benefits. It also provides a practical foundation from which Scotland could progress towards the more ambitious longer term goal outlined in Recommendation 3.

In practice, Scotland is likely to require a **tiered, hybrid approach** that combines spatial risk mapping, targeted process based modelling, and increasing use of real time monitoring data. These components should be supported by statistical or probabilistic frameworks that integrate multiple evidence sources and explicitly quantify risk and uncertainty.

Current limitations stem less from the modelling tools themselves and more from gaps in source characterisation, event scale understanding, representation of urban systems, and land-sea connectivity. Addressing these gaps – alongside improving data integration, operational usability, and transparency – is likely to deliver greater benefits than further attempts to build fully integrated, highly complex modelling systems.

There is also a broader need to strengthen Scotland's modelling capability, with well trained, adaptable teams able to maintain and evolve national tools. Scotland already has strong datasets and considerable expertise in faecal indicator organism modelling. This provides a solid basis for developing a **scale appropriate, user driven, semi distributed modelling tool** that can underpin regulatory decision making. The detailed design of such a tool will require further discussion.

Finally, future modelling should aim to represent the impacts of land management options, agronomic practices, runoff and erosion control, and nature based solutions such as buffer strips and afforestation. Improving the ability of models to guide **better source management and mobilisation reduction** may be more effective than focusing solely on improving the precision of source tracking.

References

Abdallah, A. M., Rheinheimer, D.E., Rosenberg, D.E., Knox, S. and Harou, J.J. 2022. An interoperable software ecosystem to store, visualize, and publish water resources systems modelling data. *Environmental Modelling & Software*, 151, 105371.

Afolabi, E. O., Quilliam, R.S. and Oliver, D.M. 2020. Impact of Freeze–Thaw Cycles on Die-Off of *E. coli* and Intestinal Enterococci in Deer and Dairy Faeces: Implications for Landscape Contamination of Watercourses. *Internationa; Journal of Environmental Research and Public Health*, 17, 6999.

Akoumianaki, I. A. I., A. 2022. Understanding problems associated with small-scale Private Sewage Systems (PSS) from regulators' perspectives. Scotland's Centre of Expertise for Waters (CREW).

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., Van Griensven, A., Van Liew, M. W., Kannan, N., and Jha, M. K. 2012. SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55, 1491–1508.

Bertone, E., Kozak, S. and Roiko, A. 2019. Understanding and Modeling the Occurrence of *E. coli* Blooms in Drinking Water Reservoirs. *Water Resources Research*, 55, 10518–10526.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L., Donigian, A. S., and Johanson, R. C. 1997. Hydrological Simulation Program – FORTRAN (HSPF): User's Manual for Release 11.

Bieger, K., Arnold, J. G., Rathjens, H., White, M. J., Bosch, D. D., Allen, P. M., Volk, M. and Srinivasan, R. 2017. Introduction to SWAT+, A Completely Restructured Version of the Soil and Water Assessment Tool. *JAWRA Journal of the American Water Resources Association*, 53, 115–130.

Blaustein, R. A., Hill, R. L., Micallef, S. A., Shelton, D. R. and Pachepsky, Y. A. 2016. Rainfall intensity effects on removal of fecal indicator bacteria from solid dairy manure applied over grass-covered soil. *Science of The Total Environment*, 539, 583–591.

Bougeard, M., Le Saux, J., Pérenne, N., Baffaut, C., Robin, M. and Pommepuy, M. 2011. Modeling of Escherichia coli Fluxes on a Catchment and the Impact on Coastal Water and Shellfish Quality. *Journal of the American Water Resources Association*, 47, 350–366.

Catchment Based Approach. 2026. *Catchment-based approach* [Online]. Available: https://catchmentbasedapproach.org/learn/source-apportionment-gis-tool/ [Accessed 04-05 2026].

Cea, L., Bermúdez, M. and Puertas, J. 2011. Uncertainty and sensitivity analysis of a depth-averaged water quality model for evaluation of *Escherichia Coli* concentration in shallow estuaries. *Environmental Modelling & Software*, 26, 1526–1539.

Chen, S. H. A. P., C.A. 2012. Good practice in Bayesian network modelling. *Environmental Modelling & Software*, 37, 134–145.

Cho, K., Pachepsky, Y., Kim, J., Guber, A., Shelton, D. and Rowland, R. 2010. Release of *Escherichia coli* from the bottom sediment in a first-order creek: Experiment and reach-specific modeling. *Journal of Hydrology*, 391, 322–332.

Cho, K. H., Pachepsky, Y. A., Oliver, D. M., Muirhead, R. W., Park, Y., Quilliam, R. S., and Shelton, D. R. 2016. Modeling fate and transport of fecally-derived microorganisms at the watershed scale: State of the science and future opportunities. *Water Research*, 100, 38–56.

Choi, Y.-D., Goodall, J. L., Sadler, J. M., Castronova, A. M., Bennett, A., Li, Z., Nijssen, B., Wang, S., Clark, M. P., Ames, D. P., Horsburgh, J. S., Yi, H., Bandaragoda, C., Seul, M., Hooper, R. and Tarboton, D. G. 2021. Toward open and reproducible environmental modeling by integrating online data repositories, computational environments, and model Application Programming Interfaces. *Environmental Modelling & Software*, 135, 104888.

Choi, Y. D., Goodall, J.L., Sadler, J.M., Castronova, A.M., Bennett, A., Li, Z., Nijssen, B., Wang, S., Clark, M.P., Ames, D.P. and Horsburgh, J.S 2021. Toward open and reproducible environmental modeling by integrating online data repositories, computational environments, and model Application Programming *Environmental Modelling & Software*, 135, 104888.

Comber, S. D., Smith, R., Daldorph, P., Gardner, M. J., Constantino, C., and Ellor, B. 2013. Development of a Chemical Source Apportionment Decision Support Framework for Catchment Management. *Environmental Science & Technology*, 47, 9824–9832.

Comber, S. D. W., Smith, R., Daldorph, P., Gardner, M. J., Constantino, C. and Ellor, B. 2018. Development of a chemical source apportionment decision support framework for lake catchment management. *Science of The Total Environment*, 622-623, 96–105.

Cox, B. A. 2003. A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers. *Science of the Total Environment*, 314-316, 335–377.

Crowther, J., Kay, D., Anthony, S., Gooday, R., Burgess C., and Douglass, J. 2016. Developing a methodology for screening and identifying potential sources of bacteria to improve bathing, shellfish and drinking water quality. 1 Phase 1: To design and scope a suitable methodology. Scotland's Centre of Expertise for Waters (CREW).

Darji, J., Lodha, P., and Tyagi, S. 2022. Assimilative capacity and water quality modeling of rivers: A review. *AQUA – Water Infrastructure, Ecosystems and Society*, 71, 1127–1147.

De Brauwere, A., De Brye, B., Servais, P., Passerat, J. and Deleersnijder, E. 2011. Modelling Escherichia coli concentrations in the tidal Scheldt river and estuary. *Water Research*, 45, 2724–2738.

De Brauwere, A., De Ridder, F., Gourgue, O., Lambrechts, J., Comblen, R., Pintelon, R., Passerat, J., Servais, P., Elskens, M., Baeyens, W., Kärnä, T., De Brye, B. and Deleersnijder, E. 2009. Design of a sampling strategy to optimally calibrate a reactive transport model: Exploring the potential for *Escherichia coli* in the Scheldt Estuary. *Environmental Modelling & Software*, 24, 969–981.

Dumasdelage, R. and Delestre, O. 2020. Simulating coliform transport and decay from 3D hydrodynamics model and in situ observation In Nice Area. *Sn Applied Sciences*, 2.

Eregno, F., Tryland, I., Tjomsland, T., Kempa, M. and Heistad, A. 2018. Hydrodynamic modelling of recreational water quality using *Escherichia coli* as an indicator of microbial contamination. *Journal of Hydrology*, 561, 179–186.

Ferreira, J. G., Bernard-Jannin, L., Cubillo, A., E Silva, J.L., Diedericks, G.P., Moore, H., Service, M. and Nunes, J.P. 2023. From soil to sea: An ecological modelling framework for sustainable aquaculture. *Aquaculture*, 577, 739920.

Frey, S. K., Topp, E., Edge, T., Fall, C., Gannon, V., Jokinen, C., Marti, R., Neumann, N., Ruecker, N., Wilkes, G. and Lapen, D. R. 2013. Using SWAT, Bacteroidales microbial source tracking markers, and fecal indicator bacteria to predict waterborne pathogen occurrence in an agricultural watershed. *Water Research*, 47, 6326–6337.

Gao, G., Falconer, R. and Lin, B. 2013. Modelling importance of sediment effects on fate and transport of enterococci in the Severn Estuary, UK. *Marine Pollution Bulletin*, 67, 45–54.

Garcia-Armisen, T., Thouvenin, B. and Servais, P. 2006. Modelling faecal coliforms dynamics in the Seine estuary, France. *Water Science and Technology*, 54, 177–184.

García-Alba, J., Bárcena, J. F., Ugarteburu, C., and García, A. 2019. Artificial neural networks as emulators of process-based models to analyse bathing water quality in estuaries. *Water Research*, 150, 283–295.

Glendell, M., Gagkas, Z., Richards, S., Halliday, S. CRW2018_12. Scotland’s Centre of Expertise for Waters (Crew). 2021. Developing a probabilistic model to estimate phosphorus, nitrogen and microbial pollution to water from septic tanks.

Hankin, B., Strömqvist, J., Burgess, C., Pers, C., Bielby, S., Revilla-Romero, B. and Pope, L. 2019. A new national water quality model to evaluate the effectiveness of catchment management measures in England. *Water*, 11, 1612.

Hankin, B., Stromqvist, J., Warren, S. Wood, N. and Stunell, J. 2020. CSF-Hype Improvements Project 2020/2021- Technical Summary.

Hankin, B., Wood, N. and Robinson, P. 2023. Bathing Water Modelling Investigation Final Report. Environment Agency.

Heal, R., Rostant, W. and Posen, P. 2025. Challenges and Limitations of Using Monitoring Data in Catchment-Based Models – A Case Study of Rivers Taw and Torridge, UK. *Hydrology*, 12, 212.

Huang, G., Falconer, R. and Lin, B. 2015. Integrated River and Coastal Flow, Sediment and *Escherichia coli* Modelling for Bathing Water Quality. *Water*, 7, 4752–4777.

Huang, G., Falconer, R. and Lin, B. 2018. Evaluation of *E.coli* losses in a tidal river network using a refined 1-D numerical model. *Environmental Modelling & Software*, 108, 91–101.

Huang, G., Falconer, R. A. and Lin, B. 2017. Integrated hydro-bacterial modelling for predicting bathing water quality. *Estuarine, Coastal and Shelf Science*, 188, 145–155.

Innovyze. 2022a. InfoWorks ICM Key Concepts [Online]. Available: https://help2.innovyze.com/infoworksicm/Content/HTML/ICM_IN_ILCM/InfoWorks_Key_Concepts.htm [Accessed 27th May 2026].

Innovyze. 2022b. Water Quality Model [Online]. Available: https://help2.innovyze.com/infoworksicm/Content/HTML/ICM_ILCM/Water_Quality_Model.htm [Accessed 27th May 2026].

Kelly, L. T., Sissons, J., Thompson, L., and Pearman, J. K. 2024. Faecal source apportionment using molecular methods: A proof of concept using the FEAST algorithm. *Water Research*, 266.

Kim, J., Pachepsky, Y., Shelton, D. and Coppock, C. 2010. Effect of streambed bacteria release on E. coli concentrations: Monitoring and modeling with the modified SWAT. *Ecological Modelling*, 221, 1592–1604.

Krupska, K. U., Speight, L., Robinson, J. S., Gilbert, A. J. and Cloke, H. 2024. Forecasting bathing water quality in the UK: A critical review. *WIREs Water*, 11, e1718.

Lam, M. and Ahmadian, R. 2023. Predicting Fecal-Indicator Organisms in Coastal Waters Using a Complex Nonlinear Artificial Intelligence Model. *Journal of Environmental Engineering*, 149.

Lane, S. N., Reaney, S. M., and Heathwaite, A. L. 2009. Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research*, 45.

Laniak, G. F., Olchin, G., Goodall, J., Voinov, A., Hill, M., Glynn, P., Whelan, G., Geller, G., Quinn, N., Blind, M. and Peckham, S. 2013. Integrated environmental modeling: a vision and roadmap for the future. *Environmental modelling & software*, 39, 3–23.

Lin, B., Syed, M. and Falconer, R. 2008. Predicting faecal indicator levels in estuarine receiving waters – An integrated hydrodynamic and ANN modelling approach. *Environmental Modelling & Software*, 23, 729–740.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and Arheimer, B. 2010. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology Research*, 41, 295–319.

Lloyd, W., David, O., Ascough Ii, J.C., Rojas, K.W., Carlson, J.R., Leavesley, G.H., Krause, P., Green, T.R. and Ahuja, L.R. 2011. Environmental modeling framework invasiveness: Analysis and implications. *Environmental Modelling & Software*, 26, 1240–1250.

Locatelli, L., Russo, B., Oliete, A., Catalán, J., Martínez-Gomariz, E. and Martínez, M. 2020. Modeling of E. coli distribution for hazard assessment of bathing waters affected by combined sewer overflows. *Natural Hazards and Earth System Sciences*, 20, 1219–1232.

McIntyre, N. R. 2004. *Analysis of uncertainty in river water quality modelling*. PhD, Imperial College London.

Myzece, C. 2025. *Faecal pollution from rural septic tank systems: Modelling environmental risk and understanding human perceptions*. PhD, University of Stirling.

Mzyece, C., Glendell, M., Gagkas, Z., Troldborg, M., Negri, C., Pagaling, E., Jones, I. and Oliver, D. 2026. Validating a Bayesian network model to characterise faecal indicator organism loss from septic tank systems in rural catchments. *Water Research*, 288.

Neill, A., Tetzlaff, D., Strachan, N., Hough, R., Avery, L., Kuppel, S., Maneta, M. and Soulsby, C. 2020A. An agent-based model that simulates the spatio-temporal dynamics of sources and transfer mechanisms contributing faecal indicator organisms to streams. Part 1: Background and model description. *Journal of Environmental Management*, 270.

Neill, A., Tetzlaff, D., Strachan, N., Hough, R., Avery, L., Maneta, M. and Soulsby, C. 2020B. An agent-based model that simulates the spatio-temporal dynamics of sources and transfer mechanisms contributing faecal indicator organisms to streams. Part 2: Application to a small agricultural catchment. *Journal of Environmental Management*, 270.

Neill, A., Tetzlaff, D., Strachan, N. and Soulsby, C. 2019. To what extent does hydrological connectivity control dynamics of faecal indicator organisms in streams? Initial hypothesis testing using a tracer-aided model. *Journal of Hydrology*, 570, 423–435.

Oliver, D., Crowther, J., Haygarth, P., Heathwaite,. L., Kay, D. and Page, T. 2011. Catchment Modelling Strategies for Faecal Indicator Organisms:

Options Review and Recommendations – Interim Report on Literature Review Listing Sources Accessed. DEFRA.

Oliver, D. M., Bartie, P. J., Heathwaite, A. L., Reaney, S. M., Parnell, J. A. Q., and Quilliam, R. S. 2018. A catchment-scale model to predict spatial and temporal burden of *Escherichia coli* on pasture from grazing livestock. *Science of the Total Environment*, 616–617, 678–687.

Oliver, D. M., Porter, K. D. H., Pachepsky, Y. A., Muirhead, R. W., Reaney, S. M., Coffey, R., Kay, D., Milledge, D. G., Hong, E., Anthony, S. G., Page, T., Bloodworth, J. W., Mellander, P.-E., Carbonneau, P. E., Mcgrane, S. J. and Quilliam, R. S. 2016. Predicting microbial water quality with models: Over-arching questions for managing risk in agricultural catchments. *Science of The Total Environment*, 544, 39–47.

Pascual-Benito, M., Nadal-Sala, D., Tobella, M., Ballesté, E., García-Aljaro, C., Blanch, A. R., and Lucena, F. 2020. Modelling the seasonal impacts of a wastewater treatment plant on water quality in a Mediterranean stream using microbial indicators. *Journal of Environmental Management*, 261, 261, 110220.

Porter, K. D. H., Reaney, S. M., Quilliam, R. S., Burgess, C., and Oliver, D. M. 2017. Predicting diffuse microbial pollution risk across catchments: The performance of SCIMAP and recommendations for future development. *Science of the Total Environment*, 609, 456–465.

Ranjith, S., Shivapur, A. V., Kumar, P. S. K., Hiremath, C. G., and Dhungana, S. 2019. Water quality model for streams: A review. *Journal of Environmental Protection*, 10, 1612–1648.

Razavi, S., Tolson, B. A., and Burn, D. H. 2012. Review of surrogate modelling in water resources. *Water Resources Research*, 48, W07401.

RDMAI 2025. Open *E. coli* model Report.

Reaney, S. M., Lane, S. N., Heathwaite, A. L. and Dugdale, L. J. 2011. Risk-based modelling of diffuse land use impacts from rural landscapes upon salmonid fry abundance. *Ecological Modelling*, 4, 1016–1029.

Schoen, M. E., Small, M. J. and Vanbriesen, J. M. 2010. Bayesian model for flow-class dependent distributions of fecal-indicator bacterial concentration in surface waters. *Water Research*, 44, 1006–1016.

Scottish Water. 2024a. *Annual Performance Report 2023/24* [Online]. Dunfermline, UK: Scottish Water. Available: https://www.scottishwater.co.uk/-/media/ScottishWater/Document-Hub/Key-Publications/Annual-Reports/Scottish-Water-Annual-Report-2024.pdf [Accessed].

Scottish Water. 2024b. *Overflow Map – Near Real-Time CSO Monitoring*. [Online]. Glasgow, UK: Scottish Water. Available: https://www.scottishwater.co.uk/Your-Home/Your-Waste-Water/Overflows/Live-Overflow-Map [Accessed].

SEPA. 2026. *Shellfish Waters Protected Areas* [Online]. Available: https://www.sepa.org.uk/environment/water/shellfish-water-protected-areas/ [Accessed 13-01-26].

Shrestha, N., Leta, O., De Fraine, B., Garcia-Armisen, T., Ouattara, N., Servais, P., Van Griensven, A. and Bauwens, W. 2014. Modelling *Escherichia coli* dynamics in the river Zenne (Belgium) using an OpenMI based integrated model. *Journal of Hydroinformatics*, 16, 354–374.

Slaughter, A., Allan, D., and Battersby, N. 2026. *Model Benchmarking Report: E. coli* [Online]. https://github.com/Cognizant-RDMAI/Open-Ecoli-Models/blob/main/Supporting-documents/RDMAI%20Model%20Benchmarking%20Report%20-%20BB1B%20E.%20coli.pdf. Available: https://github.com/Cognizant-RDMAI/Open-Ecoli-Models/blob/main/Supporting-documents/RDMAI%20Model%20Benchmarking%20Report%20-%20BB1B%20E.%20coli.pdf [Accessed BB1B].

SMHI. 2026a. *HYPE - our hydrological model* [Online]. Available: https://www.smhi.se/en/research/research-units/hydrology/hype---our-hydrological-model [Accessed 08-04-2026 2026].

SMHI. 2026b. *Tutorial sub-daily time steps in HYPE [HYPE Model Documentation]* [Online]. Available: http://hype.smhi.net/wiki/doku.php?id=start:hype_tutorials:subdaily_timesteps&s%5b%5d=time&s%5b%5d=step [Accessed 27 May 2026].

SMHI. 2026c. *Water Management [HYPE Model Documentation]* [Online]. Available: http://hype.smhi.net/wiki/doku.php?id=start:hype_model_description:hype_human_water#point_sources [Accessed 27th May 2026].

Soil and Water Assessment Tool. 2026. *Soil and Water Assessment Tool* [Online]. Available: https://swat.tamu.edu/ [Accessed 04-05-2026 2026].

Sokolova, E., Åström, J., Pettersson, T., Bergstedt, O. and Hermansson, M. 2012. Decay of Bacteroidales Genetic Markers in Relation to Traditional Fecal Indicators for Water Quality Modeling of Drinking Water Sources. *Environmental Science & Technology*, 46, 892–900.

Srinivasan, M., Muirhead, R., Singh, S., Monaghan, R., Stenger, R., Close, M., Manderson, A., Drewry, J., Smith, L., Selbie, D. and Hodson, R. 2021. Development of a national-scale framework to characterise transfers of N, P and *Escherichia coli* from land to water. *New Zealand Journal of Agricultural Research*, 64, 286–313.

Stapleton, C., Wyer, M., Crowther, J., Mcdonald, A., Kay, D., Greaves, J., Wither, A., Watkins, J., Francis, C., Humphrey, N. and Bradford, M. 2008. Quantitative catchment profiling to apportion faecal indicator organism budgets for the Ribble system, the UK's sentinel drainage basin for Water Framework Directive research. *Journal of Environmental Management*, 87, 535–550.

Tan, M. L., Gassman, P. W., Yang, X. and Haywood, J. 2020. A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes. *Advances in Water Resources*, 143, 103662.

Tetzlaff, D., Cape, R. and Soulsby, C. 2012. Land use and hydroclimatic influences on Faecal Indicator Organisms in two large Scottish catchments: Towards land use-based models as screening tools. *Science of the Total Environment*, 434, 110–122.

Troldborg, M., Glendell, M., Gagkas, Z., Adams, K., Negri, C., Taylor, P., Zhang, Z., Cooper, P., Brown, A., May, L., Corrochano-Fraile, Beevers, L., and Tyler, A. 2026. Probabilistic modelling of pharmaceutical pollution risk from sewage treatment work discharges using a Bayesian Network: application to a Scottish river catchment *arthArXiv*, Non-peer reviewed preprint.

Van Heijnsbergen, E., Van Kessel, T., Van Maren, D. S., and Winterwerp, J. C. 2022. A process-based modelling framework for coastal water quality: integrating hydrodynamics, sediment transport and microbial contamination. *Journal of Marine Systems*, 229.

Vidal, V., Sampognaro, L., De León, F., Kruk, C., Perera, G., Crisci, C. and Segura, A. M. 2024. A critical review of model construction and performance for nowcast systems for faecal contamination in recreational beaches. *Science of The Total Environment*, 954, 176233.

Wang, H., Blauw, A., Van Gils, J., Boelee, E., Sylvestre, É. and Medema, G. 2026. Development and evaluation of a physico-biochemical model for *Escherichia coli* in bathing waters. *Water Research*, 291.

Warn, A. 2010. *SIMCAT—A river quality model. Environment Agency internal technical documentation and user guidance*. [Online]. DEFRA. Available: https://github.com/DEFRA/SIMCAT [Accessed].

Warn, T. 2018. *SIMCAT 15.7. The SIMCAT Manual* [Online]. Available: https://github.com/DEFRA/simcat/blob/main/docs/181118%20SIMCAT%20Manual.docx [Accessed].

Wilson, G., Aruliah, D.A., Brown, C.T., Chue Hong, N.P., Davis, M., Guy, R.T., Haddock, S.H., Huff, K.D., Mitchell, I.M., Plumbley, M.D. and Waugh, B. 2014. Best practices for scientific computing. *PLoS Biology*, 12, 1001745.

Wolska, L., Kowalewski, M., Potrykus, M., Redko, V. and Rybak, B. 2022. Difficulties in the Modeling of E. coli Spreading from Various Sources in a Coastal Marine Area. *Molecules*, 27.

Wu, J., Rees, P., Storrer, S., Alderisio, K., and Dorner, S. 2009. Fate and transport modeling of potential pathogens: The contribution from sediments. J*ournal of the American Water Resources Association*, 45, 35–44.

Appendices

Appendix 1. Literature Review Methods

1. Purpose

This appendix documents the methods used to identify, screen, and synthesise published evidence on **faecal indicator organism modelling** to support catchment scale source apportionment in Scotland. The approach combined structured database searching, iterative refinement of search terms, and expert input from the **Project Steering Group (PSG)**.

2. Development of Search Strategy

2.1 Initial keyword generation

The research team developed an initial long list of keywords relating to **FIO models**, catchment modelling frameworks, hydrodynamics, and decision support systems. This list was presented to the PSG during an online meeting, where members contributed additional terms, model names, and thematic areas relevant to Scottish catchments.

2.2 Iterative testing of search strings

Multiple combinations of search strings were trialled in **Web of Science** to balance sensitivity (capturing all relevant studies) and specificity (avoiding unrelated microbiology or laboratory studies). The target was to identify a search string returning **fewer than 200 papers** for manageable initial screening.

2.3 Final search string

The final search string selected after iterative testing was:

"E. coli OR Escherichia OR faecal indicator OR enterococci OR fecal indicator OR coliforms OR FIO" (Title)*** AND "model OR decision support OR framework" (Title) AND "catchment OR watershed OR river OR drainage network OR estuary OR bathing water OR shellfish" (Topic)**

This search returned **171 articles** (including two review papers).

3. Search Within Results Checks

To confirm that key model families were captured, the Web of Science **search within results** function was used to test for expected model names and acronyms. The following were detected:

- **VIPER** (2)
- **AN** (2)
- **Integrated modelling** (2)
- **INCA** (1)
- **SWAT** (Soil and Water Assessment Tool) (14)
- **FIO BBN** (1)

No references to **SAGIS** were detected using either acronyms or full model names.

4. Screening Process

4.1 Title and abstract screening

Articles were screened for relevance based on titles and abstracts. Papers were excluded if they:

- Focused on unrelated uses of *E. coli* (e.g., laboratory model organism studies)
- Fell outside the agreed geographical scope
- Did not involve modelling or decision support applications

After this stage, **139 papers** remained.

4.2 Geographical scope

The PSG and research team jointly defined the geographical scope as countries with **climates broadly comparable to the UK**, to ensure transferability of modelling approaches. A climate comparison identified the following regions as in scope:

Table 1A. Countries and regions considered climatically comparable to the UK	
Country/Region	**Rationale**
Ireland	Very similar maritime climate
NW France	Marine climate zone
Netherlands	Mild, wet, maritime
Belgium	Maritime climate
Denmark	Cfb climate, cool summers
NW Germany	Temperate oceanic
New Zealand	Similar variability and rainfall
N. coastal Japan	Comparable weather patterns
Tasmania / S. coastal Australia	Temperate oceanic
Southern Chile	Temperate oceanic
USA (Pacific Northwest)	Cool, wet maritime

4.3 Full text screening

Full texts were obtained via EndNote. A second, more detailed screen removed 71 papers due to insufficient geographical relevance or lack of modelling focus. A further subset was flagged as partially relevant. 51 papers were ultimately classified as fully relevant.

5. Flow Diagram

Code

Identification

Web of Science search results 171

Additional sources (known models, expert input) .... —

Duplicates removed 0

Screening

Title/abstract screening 139

Excluded (out of scope / irrelevant).......................32

Eligibility

Full-text articles assessed 139

Excluded after full-text review 71

Reasons: geographical scope, non-modelling focus

Included

Studies included in synthesis 51

Partially relevant studies .... —

----

6. Summary of Identified Literature

Across the 139 papers initially identified:

- Publication years ranged from **1978 to 2026**
- Peak publication year: **2018 (14 papers)**
- Second peak: 2015 **(13 papers)**
- Rapid expansion from **2010 onwards**

Methodological evolution shows a progression from:

1. Empirical models
2. Process based hydrodynamic and catchment models (≈60% of studies)
3. Statistical predictive models
4. Machine learning and hybrid approaches, particularly post 2015

7. Additional Knowledge Sources

Alongside the structured search, the research team compiled a list of models already known through professional experience and prior projects (provided separately in Excel).

A supplementary search targeting review papers was also conducted:

"E. coli OR Escherichia OR faecal indicator OR enterococci OR fecal indicator OR coliforms OR FIO" (Title)* **AND** "*model* OR decision support OR framework" (Title)* **AND** "water" (Topic)

This returned **517 articles**, including **seven review papers**, which were examined for additional insights and cross references.

Appendix 2. Stakeholder Engagement and "Snowballing" Approach

1. Purpose

This appendix summarises the stakeholder engagement activities and the **snowballing literature identification** undertaken alongside the structured database search described in Appendix 1. These activities were designed to ensure that the review captured practitioner knowledge, emerging modelling tools, and relevant grey literature not always indexed in academic databases.

2. Stakeholder Engagement

2.1 Identification of key informants

The research team identified a small group of individuals known to be involved in the development, application, or evaluation of FIO modelling tools across the UK water sector. These individuals were selected based on:

- direct involvement in model development
- operational use of catchment or bathing water models
- experience with integrated modelling frame-works
- knowledge of emerging or unpublished tools

A total of six individuals were approached by email.

2.2 Engagement outcomes

Of the six individuals contacted:

- **five** participated in online video conference discussions
- **one** did not respond to email invitations

The discussions focused on:

- models currently in operational use
- perceived strengths and limitations of existing tools
- data availability and integration challenges
- emerging approaches and research gaps
- opportunities for future development

2.3 Contribution to the review

Insights from these conversations directly informed:

- practical application of models and modelling frameworks
- interpretation of findings from the academic literature
- understanding of practical constraints relevant to Scottish catchments

These contributions were incorporated into the main report and helped ensure that the review reflected both academic and practitioner perspectives.

3. Snowballing Approach

3.1 Rationale

To complement the structured Web of Science search, the research team applied a **snowballing** approach, whereby additional relevant literature was identified through:

- references cited within papers already included in the review
- forward citation searches
- informal searches conducted during reading
- documents recommended by stakeholders

This method is widely used to ensure that important studies not captured by database searches – particularly grey literature, technical reports, and emerging modelling work – are not overlooked.

3.2 Process

During the review period, any paper, report, or resource encountered through reading or stakeholder discussion that appeared relevant to **catchment scale FIO** modelling was:

- added to a separate EndNote reference library

3.3 Contribution to the evidence base

Snowballing contributed:

- additional modelling case studies
- grey literature (e.g., utility reports, regulatory documents)
- context for interpreting model performance and applicability

This ensured that the final evidence base was comprehensive and reflective of both academic and practitioner domains.

4. Summary

Stakeholder engagement and snowballing provided essential complementary evidence to the structured literature search. Together, these activities strengthened the review by:

- incorporating practitioner insights into model use and limitations
- expanding the evidence base to include grey literature and emerging work

These contributions helped ensure that the findings presented in the main report are grounded in both scientific evidence and real world operational experience.

Appendix 3. Evidence Mapping Database

1. Overview

This appendix provides access to the Excel file containing the **evidence mapping database** developed as part of the structured **literature review** described in Appendix 1. The database includes **only the studies that passed the full screening process** and were assessed as relevant or partially relevant to catchment scale FIO modelling.

It does not include:

- papers identified through stakeholder engagement
- papers identified through the **snowballing process**
- grey literature or additional informal sources.

These were tracked separately.

2. Access to the Database

A link to the Excel file is provided here:

www.crew.ac.uk/publications/review-existing-faecal-indicator-organism-models-and-future-integration-opportunities

This file contains the final set of screened studies used in the evidence mapping task.

3. File Contents

The Excel file includes the following sheets:

- **Evidence Mapped Studies** – all papers that passed full screening
- **Relevance Classification** – relevant vs. partially relevant.

CENTRE OF EXPERTISE FOR WATERS

Centre of Expertise for Waters

James Hutton Institute
Craigiebuckler
Aberdeen AB15 8QH
Scotland UK

www.crew.ac.uk

CREW publications can be accessed here:
www.crew.ac.uk/publications

CREW-scotland

Bluesky
@crew-waters

CREW is a partnership between the James Hutton Institute
and all Scottish Higher Education Institutes and Research Institutes.
The Centre is funded by the Scottish Government.

Scottish Government
Riaghaltas na h-Alba
gov.scot