Journal of Environmental Management 270 (2020) 110903

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: http://www.elsevier.com/locate/jenvman

Research article

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

Aaron J. Neill [a,b,*], Doerthe Tetzlaff [c,d,a], Norval J.C. Strachan [e], Rupert L. Hough [b], Lisa M. Avery [b], Sylvain Kuppel [f,g,a], Marco P. Maneta [h,i], Chris Soulsby [a,c]

[a] Northern Rivers Institute, University of Aberdeen, Aberdeen, AB24 3UF, Scotland, United Kingdom
[b] The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, Scotland, United Kingdom
[c] IGB Leibniz Institute of Freshwater Ecology and Inland Fisheries, 12587, Berlin, Germany
[d] Department of Geography, Humboldt University Berlin, 10099, Berlin, Germany
[e] School of Biological Sciences, University of Aberdeen, Cruickshank Building, St Machar Drive, Aberdeen, AB24 3UU, Scotland, United Kingdom
[f] Institut de Physique du Globe de Paris, CNRS UMR 7154 - University of Paris, 75231, Paris, France
[g] INRAE, RiverLy, 69625, Villeurbanne, France
[h] Geosciences Department, University of Montana, Missoula, MT, 59812-1296, USA
[i] Department of Ecosystem and Conservation Sciences, W.A Franke College of Forestry and Conservation. Universtiy of Montana, Missoula, USA

ARTICLE INFO

Keywords:
Diffuse pollution
E. coli
EcH$_2$O-iso
Microbial water quality
Tracer-aided modelling
Water quality modelling

ABSTRACT

A new Model for the Agent-based simulation of Faecal Indicator Organisms (MAFIO) is developed that attempts to overcome limitations in existing faecal indicator organism (FIO) models arising from coarse spatial discretisations and poorly-constrained hydrological processes. MAFIO is a spatially-distributed, process-based model presently 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. Specifically, FIO loading, die-off, detachment, surface routing, seepage and channel routing are modelled on a regular spatial grid. Central to MAFIO is that hydrological transfer mechanisms are simulated based on a hydrological environment generated by an external model for which it is possible to robustly determine the accuracy of simulated catchment hydrological functioning. The spatially-distributed, tracer-aided ecohydrological model EcH$_2$O-iso is highlighted as a possible hydrological environment generator. The present paper provides a rationale for and description of MAFIO, whilst a companion paper applies the model in a small agricultural catchment in Scotland to provide a proof-of-concept.

1. Introduction

Transfer of faecal pathogens (e.g. *E. coli* O157) to water bodies represents a significant risk to public health, with the ingestion of contaminated water through drinking or recreational use having the potential to cause severe gastrointestinal illness in humans (Fewtrell and Kay, 2015; Oliver et al., 2005a). Faecal indicator organisms (FIOs), such as generic *E. coli*, are commonly used to monitor microbial water quality and indicate the potential presence of pathogens that are more difficult to quantify directly (Geldreich, 1996). Risk of impaired microbial water quality is often elevated in agricultural catchments due to the large range of FIO sources that are present (e.g. different livestock) and the multiple transport mechanisms (e.g. overland flow, seepage from areas of degraded soil) that can transfer FIOs to streams (Chadwick et al., 2008). Successful implementation of mitigation measures in such environments requires the spatio-temporal dynamics of sources and transfer mechanisms contributing FIOs to streams at the sub-field scale to be understood (Oliver et al., 2007, 2016; also c.f. Greene et al., 2015; Vinten et al., 2017). This will likely require the integration of empirical data with process-based FIO models (de Brauwere et al., 2014). However, uncertainties regarding the robustness of simulated hydrological processes and the common adoption of aggregative (i.e. simulated stores and fluxes of FIOs are integrated over spatial units), lumped to semi-distributed model structures may hinder the use of many existing models for understanding sub-field-scale drivers of in-stream FIO dynamics that emerge at the catchment scale (c.f. Rode et al., 2010; Wellen

* Corresponding author. Northern Rivers Institute, University of Aberdeen, Aberdeen, AB24 3UF, Scotland, United Kingdom.
E-mail address: aaron.neill@abdn.ac.uk (A.J. Neill).

https://doi.org/10.1016/j.jenvman.2020.110903
Received 31 October 2019; Received in revised form 29 May 2020; Accepted 31 May 2020

et al., 2015).

The overall aim of this paper is to outline a new Model for the Agent-based simulation of Faecal Indicator Organisms (MAFIO); it is structured as follows. Section 2 reviews the pertinent limitations of existing process-based FIO models before exploring how agent-based and tracer-aided modelling approaches have potential in helping these to be addressed. This provides the rationale for the development of MAFIO, which simulates the sub-field-scale fate and transport of agents representing FIOs in a spatially-distributed and process-based manner, underpinned by a hydrological environment generated by a robust hydrological model. A detailed model description is presented in Section 3. Finally, Section 4 introduces the tracer-aided ecohydrological model EcH$_2$O-iso (Kuppel et al., 2018a) as a potential example of a robust hydrological environment generator and details its coupling to MAFIO. A companion paper (Neill et al., 2020) provides a proof-of-concept application of MAFIO focusing on process representation and the model's potential for use in a management context.

2. Rationale for model development

2.1. Current limitations to simulating sources and transfer mechanisms contributing FIOs to streams at the sub-field scale

A prerequisite for using models to understand drivers of dynamics in water quality parameters is proper constraint of processes to which the parameters may be sensitive (Sokolova et al., 2018; Vaché and McDonnell, 2006). For microbial water quality, hydrological processes are often hypothesised as a major control on observed FIO dynamics (Kay et al., 2008; Oliver et al., 2005a; Tetzlaff et al., 2012). Catchment hydrological functioning in process-based FIO models can often only be constrained through calibration to discharge data from a catchment outlet (Cho et al., 2016). However, these data primarily capture the speed at which perturbations (e.g. rainfall) to the catchment system are transmitted to the outlet and provide limited information on internal catchment states and fluxes (Birkel et al., 2014a; McDonnell and Beven, 2014). Consequently, flow path dynamics and resulting hydrological connectivity are likely to be poorly constrained in FIO models, potentially undermining their ability to accurately represent hydrological mechanisms that could transfer FIOs to streams (c.f. Sokolova et al., 2018; Vaché and McDonnell, 2006; Wellen et al., 2015).

A further issue is coarse spatial discretisation in many process-based FIO models, with lumped (e.g. Haydon and Deletic, 2006; Neill et al., 2019) to semi-distributed (e.g. Sadeghi and Arnold, 2002; Whitehead et al., 2016) model structures generally being the norm. Whilst limited examples of fully-distributed FIO models do exist, they have often been implemented at coarse resolutions (e.g. the 1 km^2-resolution WATFLOOD model of Dorner et al., 2006). Consequently, the spatial discretisation adopted by many FIO models is often inconsistent with the scales at which heterogenity in hydrological processes and FIO fate and transport is expressed, and at which mitigation measures can be implemented (c.f. Fatichi et al., 2016; Rode et al., 2010; Wellen et al., 2015).

Recent progress in spatially-distributed, tracer-aided hydrological modelling has potential in simultaneously overcoming issues of scale and hydrological process realism in FIO models (see Section 2.3). However, spatially-distributed simulation of FIO dynamics alone is unlikely to be sufficient for fully understanding sources and transfer mechanisms contributing FIOs to streams; this requires explicit tracking of FIOs as they move through a catchment. As summarised by Reaney (2008), spatially-distributed models conventionally simulate fluxes into and out of individual grid cells, but no information is available on the spatial origin of the constituents (water/contaminants) composing those fluxes; consequently, it is often only possible to simulate the locations of active fluxes and what their magnitude is, but not actually whether those fluxes contribute water/contaminants to the stream. A possible solution, which could also have value in simulating the heterogenous behaviour of FIOs of different types or from different host animals (e.g. die-off kinetics: Avery et al., 2004), is the incorporation of agent-based methods into spatially-distributed modelling frameworks (Reaney, 2008).

2.2. The potential of agent-based modelling approaches

Agent-based models (ABMs) consist of three major elements: 1) agents representing autonomous individuals, each associated with a set of attributes and the ability to sense and process information regarding their surroundings; 2) a simulation environment that can be distributed in space in which the agents operate; 3) a set of rules, conditional or stochastic in nature, defining how agents interact with the environment and with each other based on their attributes (Abdou et al., 2012; Crooks and Heppenstall, 2012). Agent attributes can relate to any characteristic necessary to simulate the behaviours of the phenomena being studied or that is of interest for the problem being addressed (Macal and North, 2010). Whilst often more demanding than aggregative modelling approaches in terms of computation, data requirements and evaluation, the bottom-up approach of ABMs allows greater scope for representing heterogeneity amongst individual agents and their interactions with the simulation environment, permitting richer simulation of small-scale behaviours and how these cause higher-level system dynamics to emerge (O'Sullivan et al., 2012).

Use of ABMs for the process-based simulation of water quantity/quality dynamics has been relatively limited. However, there are promising examples of the development of such models to elucidate spatio-temporal patterns of surface hydrological connectivity through tracking locational attributes of agents representing water particles as they move through catchments (Reaney, 2008), and to simulate the hydrologically-induced erosion and transport of markers (similar to agents) representing sediment particles based on their unique characteristics (Cooper et al., 2012). Furthermore, several studies have successfully implemented closely-related yet simpler (Crooks and Heppenstall, 2012) cellular automata models to simulate hydrological processes (e.g. Cirbus and Podhoranyi, 2013; Hodge and Hoey, 2012; Ravazzani et al., 2011; Shao et al., 2015). Consequently, simulating the fate and transport of FIOs via agent-based approaches may hold significant potential for elucidating the sources and transfer mechanisms contributing FIOs to streams. In particular, attributes of agents representing FIOs could govern how they interact with different transfer mechanisms operating in a simulated environment and further permit the pathways taken by agents through a catchment to the stream to be tracked.

2.3. Progress in spatially-distributed, tracer-aided (eco)hydrological modelling

The dynamics of stable isotopes (^{2}H and ^{18}O) in streamwater contain information with respect to the velocities of water passing through a catchment, as influenced by factors such as hydrological connectivity, storage and mixing (Birkel and Soulsby, 2015; McDonnell and Beven, 2014). Consequently, incorporation of isotopes as tracers into hydrological models allows performance assessments to advance from how well a model fits an observed hydrograph, to also how consistent the model is with internal catchment processes giving rise to the velocity response (Birkel et al., 2014a; McDonnell and Beven, 2014). Improved confidence in the realistic simulation of catchment functioning afforded by such tracer-aided hydrological models has facilitated their use in providing valuable insights into how catchments store and release water (Birkel and Soulsby, 2015), and how these factors drive various water quality parameters including dissolved organic carbon (Birkel et al., 2014b; Dick et al., 2015) and FIOs (Neill et al., 2019). Recently, tracer-aided models have evolved from their early lumped conceptual structures towards ones that are fully-distributed and increasingly physically-based (e.g. Kuppel et al., 2018a; Remondi et al., 2018; van

Huijgevoort et al., 2016), reflecting cheaper isotope analysis (e.g. Berman et al., 2013) and recognition of the need for catchment systems to be modelled in this way to tackle issues of water quality and environmental change (Fatichi et al., 2016). Such models have permitted more nuanced simulation of catchment-scale discharge and tracer dynamics, and also facilitate specific interrogation of intra-catchment water and isotope dynamics to further increase confidence that internal catchment processes are being correctly represented (Ala-aho et al., 2017; Knighton et al., 2020; Kuppel et al., 2018a; Piovano et al., 2018, 2019; Remondi et al., 2018; Smith et al., 2019).

In addition to the value of tracers, the importance of resolving how water is partitioned between "green" (evaporation and transpiration) and "blue" (discharge and groundwater recharge) hydrological fluxes when representing the storage and release of water by catchments is increasingly recognised (Brooks et al., 2015; Falkenmark and Rockström, 2006). Most tracer-aided models only explicitly simulate "blue" fluxes, with "green" fluxes specified as a combined variable usually partitioned based on model parameters (Fatichi et al., 2016; Vivoni, 2012). However, tracer-aided ecohydrological models are emerging that explicitly resolve both "blue" and "green" fluxes whilst also simulating stable isotope dynamics to enable further evaluation of process consistency (Kuppel et al., 2018a; Maneta and Silverman, 2013). Such a model, EcH$_2$O-iso (Kuppel et al., 2018a), is introduced in Section 4 as an example of a hydrological environment generator for MAFIO. However, it is stressed that any model can be used, providing it can be robustly determined as to whether catchment hydrological functioning is being successfully captured.

3. Presentation of MAFIO

This presentation follows the ODD (Overview, Design Concepts, Details) protocol for describing agent-based models (Grimm et al., 2006, 2010). MAFIO is written in the Python 3.6 programming language and makes use of the PCRaster module (Karssenberg et al., 2010) for handling spatial inputs and outputs. Given that it can be coupled with any robust hydrological model that provides the necessary outputs, the following presentation of MAFIO is provided in a generic form, with Section 4 outlining the specific interfacing with EcH$_2$O-iso.

3.1. Purpose

The purpose of this initial version of MAFIO is to reveal the spatio-temporal dynamics of sources and transfer mechanisms contributing FIOs to streams at the sub-field scale in small (<10 km^2) agricultural catchments. This is achieved by simulating and tracking, in a spatially-distributed, process-based manner, the fate and transport of agents representing FIOs shed by different types of livestock (Fig. 1).

3.2. State variables and scales

MAFIO consists of four entities. Two environments drive the simulation: a catchment environment and a hydrological environment. These set the attributes of a spatial grid on which simulations take place over discrete timesteps. The behaviour of agents representing FIOs (FIO-agents) is then simulated based on their individual attributes and those of the spatial grid for each timestep.

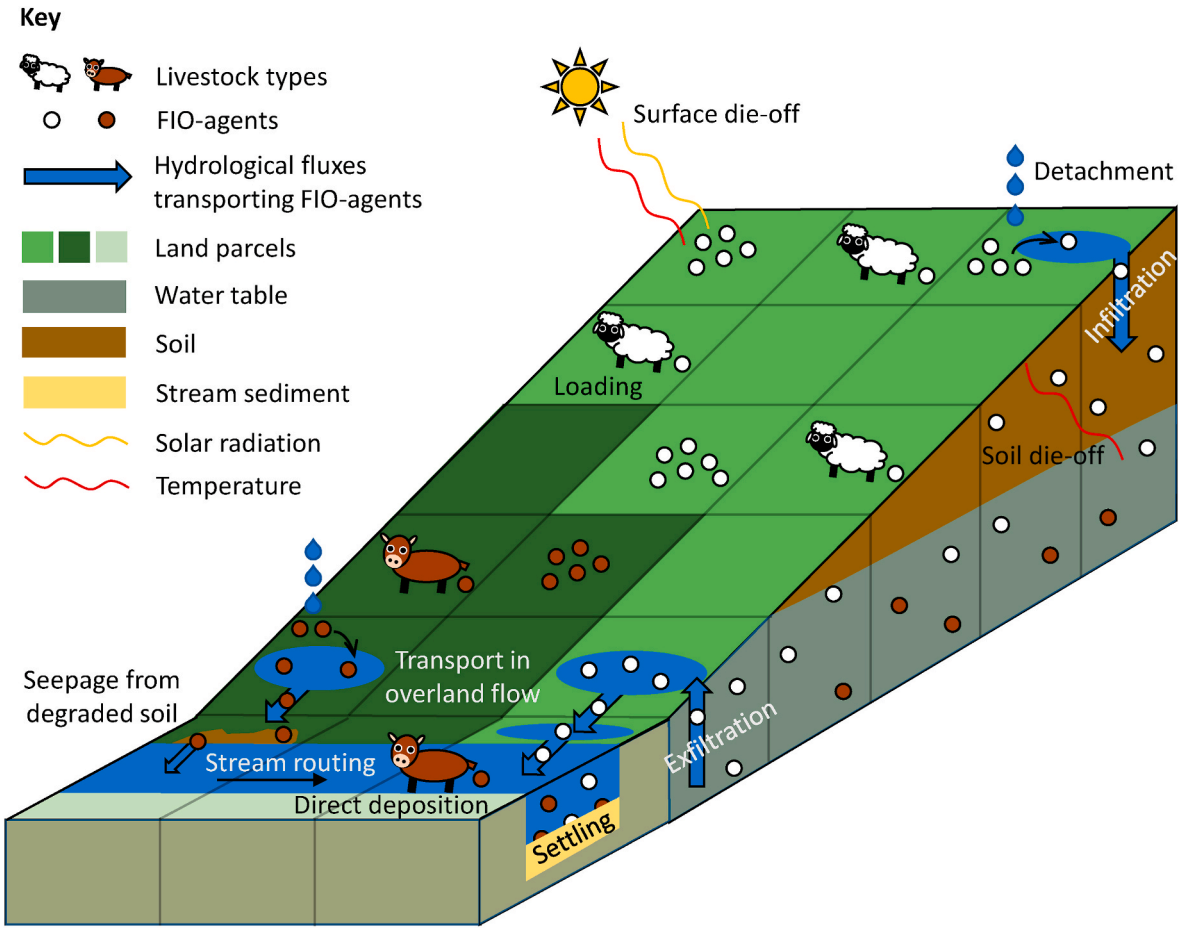

Fig. 1. A conceptual diagram of MAFIO showing the faecal indicator organism (FIO) behaviour and transport processes simulated by the model. Simulations take place on a regular spatial grid.

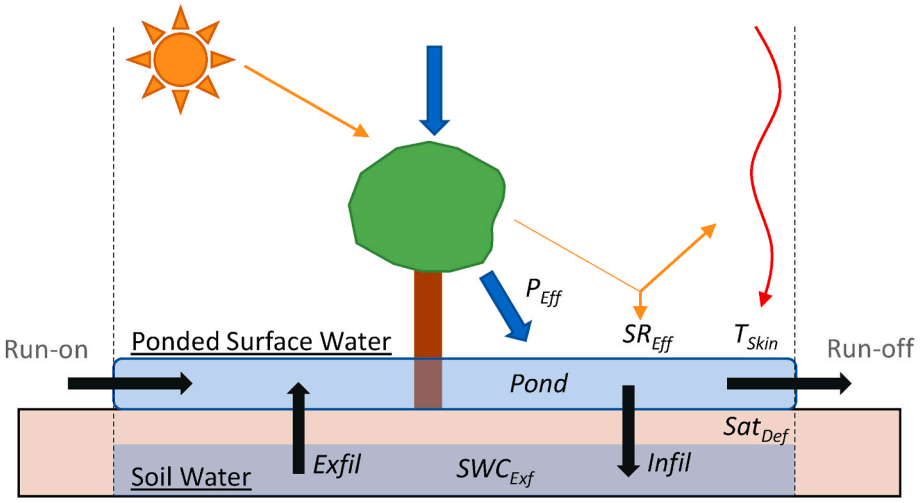

Fig. 2. The abstract catchment system represented by the variables of the hydrological environment. Variables in black italic font are those provided explicitly by the hydrological environment: Effective precipitation (P_{Eff}), effective solar radiation (SR_{Eff}), soil skin temperature (T_{Skin}), exfiltration (*Exfil*), infiltration (*Infil*), ponded water prior to infiltration (*Pond*), soil water content prior to exfiltration (SWC_{Exf}) and soil saturation deficit (Sat_{Def}). Grey, non-italicised font denotes fluxes capable of affecting FIO-agent behaviour (*Run-on* and *Run-off*) that are represented implicitly. Hydrological stores are underlined. Note that *Exfil* will be non-zero if Sat_{Def} is 0 – both are shown here for illustrative purposes only.

3.2.1. Catchment environment

The catchment environment defines, on a regular grid, the following fixed spatial characteristics of the catchment being modelled:

1) Spatial extent;
2) Distribution of land parcels, that may be delineated based on land uses (e.g. pastoral field, arable field) or covers (e.g. forest, moorland);
3) Local flow direction based on the D8 algorithm (O'Callaghan and Mark, 1984);
4) Location and width of the stream channel;
5) Locations where soil is degraded adjacent to the channel;
6) Grid cells immediately upslope of those containing a channel;
7) Grid cells immediately upslope of those containing degraded soil;
8) Vertical discretisation of the landscape, currently limited to the land surface and the first few centimetres of the upper soil profile (exact depth unspecified) where it is assumed most interactions of FIOs with the soil will occur (Stocker et al., 2015).

Additionally, counts of each livestock type (e.g. sheep, cattle) considered in the simulation for each land parcel, and which areas of the stream are accessible to livestock, are specified, both of which vary in time.

3.2.2. Hydrological environment

The hydrological environment (Fig. 2) is provided by a spatially-distributed hydrological model that can simulate the following on a regular spatial grid: effective precipitation (P_{Eff}), effective solar radiation (SR_{Eff}), ponded surface water prior to infiltration (*Pond*), infiltration flux (*Infil*), soil water content prior to exfiltration (SWC_{Exf}), exfiltration flux (*Exfil*), soil saturation deficit (scaled between 0 for fully saturated and 1 for only residual soil moisture present; Sat_{Def}) and soil skin temperature (T_{Skin}). Here, effective precipitation is total precipitation minus interception, whilst effective solar radiation is that absorbed by the ground after accounting for transmission and reflective losses. Whilst not explicitly part of the hydrological environment, run-on and run-off are accounted for implicitly by assuming that any water remaining on the surface following infiltration/exfiltration runs off to contribute to *Pond* of the next downslope grid cell (consistent with EcH_2O-iso). The hydrological environment gives the following abstract conceptualisation of catchment hydrological functioning from the perspective of MAFIO (Fig. 2). For a given grid cell, there are two hydrological stores: Ponded Surface Water and Soil Water. First, P_{Eff} and run-on from upslope cells enter Ponded Surface Water to define *Pond*. Water can then infiltrate (*Infil*) into Soil Water to increase SWC_{Exf} and reduce Sat_{Def}, after which water can exfiltrate (*Exfil*) back to Ponded Surface Water if the soil is now saturated (i.e. $Sat_{Def} = 0$). Finally, any

Table 1
The dynamic and fixed attributes of the spatial grid. In the Description section, italics and underlining are used to show where variables of the hydrological and catchment environments, respectively, are used in the derivation of attribute values.

Spatial grid attribute	Description	Units
Dynamic		
Effective precipitation (P_{Att})	*Effective precipitation* (P_{Eff}) that reaches the surface after accounting for interception losses	cm d^{-1}
Effective solar radiation (SR_{Att})	*Effective solar radiation* (SR_{Eff}) that is absorbed by the surface after accounting for transmission losses and reflection	Ly hr^{-1}
Probability of exfiltration (P_Exfil_{Att})	The proportion of *soil water content prior to exfiltration* (SWC_{Exf}) that constitutes the *exfiltration flux* (*Exfil*) to the surface	–
Probability of infiltration (P_Infil_{Att})	The proportion of *ponded surface water prior to infiltration* (*Pond*) that constitutes the *infiltration flux* (*Infil*) into the soil	–
Probability of seepage (P_Seep_{Att})	For cells with degraded soil adjacent to a channel, the probability of seepage to the channel as a function of *soil saturation deficit* (Sat_{Def}) and livestock counts for the land parcel(s) associated with the cell	–
Soil skin temperature (TS_{Att})	*Soil skin temperature* (T_{Skin})	°C
Fixed		
Cell size ($size_{Att}$)	The length of a grid cell	Metres
Channel cell ($chan_{Att}$)	Whether (>0) or not (0) a channel is present in the cell	–
Degraded soil in cell ($degSoil_{Att}$)	If the cell contains a channel, whether (>0) or not (0) there is an area of degraded soil surrounding the channel	–
Flow direction ($fDir_{Att}$)	Identifier for the direction of steepest decent	–
Land parcel(s) (LP_{Att})	Identifier for the land parcel(s) to which the cell belongs based on its location within the catchment environment	–
Proportion of cell area occupied by channel (P_Chan_{Att})	The channel area in the cell divided by the cell area	–
X–Y location (Loc_{Att})	The spatial co-ordinates of the cell	–

remaining water in Ponded Surface Water is routed to the next downslope cell unless a channel is present, in which case remaining water is routed to the stream and out of the catchment. It should be noted that this conceptualisation only includes stores and fluxes of water relevant to the transport of FIOs in naturally-drained environments; however, future iterations of MAFIO may look to allow for artificial drainage as a potentially-efficient FIO transport pathway (Oliver et al., 2005b).

3.2.3. Timestep

MAFIO operates on discrete timesteps consistent with the fixed temporal resolution of variables defined by the aforementioned environments. Here, a daily timestep is used. However, if needed to account for non-linearities in observed FIO dynamics (McKergow and Davies-Colley, 2010), sub-daily timesteps could be used with appropriate changes to parameter values and sufficient computing power.

3.2.4. Spatial grid

Simulations take place on a regular spatial grid (Fig. 1), with each grid cell having fixed and dynamic attributes (Table 1). The catchment environment determines values of fixed attributes. The position of a cell within the environment (*X–Y location* [Loc_{Att}]) determines values for *Channel cell* ($chan_{Att}$), *Flow direction* ($fDir_{Att}$), and *Land parcel(s)* (LP_{Att}). Multiple land parcels can be specified for a cell containing a channel to indicate those on either side of the channel. The attribute *Cell size* ($size_{Att}$) is determined by the spatial resolution of the catchment and hydrological environments. However, to allow for spatial resolutions too coarse to resolve small channels (Zhang and Montgomery, 1994) and processes occurring at the land-water interface that are possibly crucial for water quality (e.g. Karr and Schlosser, 1978), sub-grid heterogeneity can be represented in cells containing a channel. These cells have the additional fixed attributes *Proportion of cell area occupied by channel* (P_Chan_{Att}) and *Degraded soil in cell* ($degSoil_{Att}$) to allow for the channel not occupying the whole cell and for the presence of degraded soil in the vicinity of the channel due to livestock compaction, respectively. The latter may result in the chronic seepage of water and FIOs to the stream (Bilotta et al., 2007). P_Chan_{Att} is calculated by multiplying the width and straight-line length (*streamDist*, equal to either $size_{Att}$ *or* $[2\times size_{Att}^2]^{0.5}$ depending on $fDir_{Att}$) of the channel occupying a cell, then dividing by cell area.

The dynamic attributes of the spatial grid are calculated for each timestep. They are predominantly derived from the hydrological environment (Fig. 3) and are largely responsible for the behaviour of FIO-agents before entering the channel. *Effective precipitation* (P_{Att}), *Effective solar radiation* (SR_{Att}) and *Soil skin temperature* (TS_{Att}) are equal to P_{Eff}, SR_{Eff} *and* T_{Skin}, respectively. The attributes *Probability of exfiltration* (P_Exfil_{Att}) and *Probability of infiltration* (P_Infil_{Att}) are defined as:

$$P\_Exfil_{Att} = \frac{Exfil}{SWC_{Exf}} \tag{1}$$

$$P\_Infil_{Att} = \frac{Infil}{Pond} \tag{2}$$

Fig. 3a gives a visual example of how hydrological environment variables translate to attributes of the spatial grid. Since sub-grid heterogeneity is represented in MAFIO but not necessarily in the hydrological environment simulator, dynamic attributes of cells containing a channel may need derivation from variables simulated for neighbouring cells. This could involve taking average values from upslope cells which drain into the cell in question (Fig. 3a); conceptually, this implies that the land surface adjacent to the channel implicitly represented in MAFIO is an aggregated extension of the upslope cells.

The attribute *Probability of seepage* (P_Seep_{Att}) uses variables of the catchment and hydrological environments to encapsulate the combined effects of livestock grazing pressure and soil moisture on soil degradation (Drewry, 2006) and the associated potential for FIO seepage into streams (Bilotta et al., 2007; Tetzlaff et al., 2012). To indicate livestock grazing pressure, livestock count(s) for the present timestep for land parcels(s) associated with cells containing degraded soil are converted into livestock units per hectare (LU ha^{-1}; Natural England, 2013). This variable is then classified into a band ("low", "medium" or "high" - see Section 3.7.2.4) tied to a certain fraction of soil that could be degraded on the current timestep (D_{frac_new}; c.f. Sheath and Carlson, 1998); if no livestock are present, D_{frac_new} is 0. However, since soil takes time to recover once degraded, the final fraction of degraded soil (D_{frac}) used for the timestep is the maximum of D_{frac_new} and the damage fraction of the previous timestep (D_{frac_old}) subjected to exponential decay (Fig. 3b; c.f. Elliot et al., 2002); the latter occurs at a rate specified by the parameter k_d in units of T^{-1}, where T is the length of a timestep:

$$D_{frac} = \max\left\{D_{\text{frac\_new}}, D_{\text{frac\_old}} \cdot e^{-k_d T}\right\} \tag{3}$$

The value of D_{frac} is then scaled by the average Sat_{Def} of cells immediately upslope (as cells with degraded soil also contain a channel) to account for the impact of moisture content in the extent of soil degradation (Drewry, 2006), and to give the final probability of seepage (Fig. 3b):

$$P\_Seep_{Att} = D_{frac} \cdot \left(1 - Sat_{Def}\right) \tag{4}$$

3.2.5. FIO-agents

FIOs in MAFIO are represented by FIO-agents. Given likely large numbers of FIOs present within agricultural catchments, it is impossible for each FIO to be represented by a single FIO-agent. Consequently, MAFIO initially adopts the simple approach of assuming that the overall behaviour of a population can be inferred from simulating a suitably-sized sample (Parry and Bithell, 2012; Reaney, 2008). FIO-agents are characterised by dynamic, fixed and memory attributes (Table 2). Dynamic attributes can be updated multiple times within a timestep due to die-off and transport processes. *Attached-stream sediment* and *Detached-faeces* indicate whether an FIO-agent is attached to stream sediment or detached from a faecal deposit, respectively. The former determines how the FIO-agent is routed through the stream, whilst the latter specifies whether an FIO-agent is available for transport. *Domain type* accounts for the position of the FIO-agent at the sub-grid scale (either in a cell without a channel or otherwise its location within a cell containing a channel). *Infiltration stage* denotes where in the vertical discretisation of the catchment the FIO-agent is, which will affect the die-off and transport processes it can experience. *Life state* and *Movement tracker* follow whether the FIO-agent is alive or dead (to determine whether it is still tracked) and if the FIO-agent moved in the current timestep, respectively. Finally, *Timestep, X co-ordinate* and *Y co-ordinate* specify the current timestep, and X and Y co-ordinates of the FIO-agent; these allow recording of when changes to certain dynamic attributes occur in the simulation and where in the spatial grid the FIO-agent is situated. Memory attributes keep track of how dynamic attributes of the FIO-agents change over the simulation; this permits tracking of, for example, the origin of an FIO-agent that reaches the stream or the path that it took to the stream. Fixed attributes are assigned when the FIO-agent spawns and record its original host animal (*Livestock type)* and the land parcel into which it spawned (*Land parcel*). At present, MAFIO is configured to consider only a single type of FIO (*E. coli*); however, an additional fixed attribute for FIO-agents specifying their FIO type could easily allow representation of multiple FIO types with specific parameterisations.

3.3. Process overview and scheduling

Process scheduling during each timestep is shown in Fig. 4a. First, environmental variables define dynamic attributes of the spatial grid, which is loaded in. Next, the Defecation Sub-model (Section 3.7.1) simulates FIO loading by spawning FIO-agents in each land parcel in a set order, based on their associated livestock counts. When new FIO-

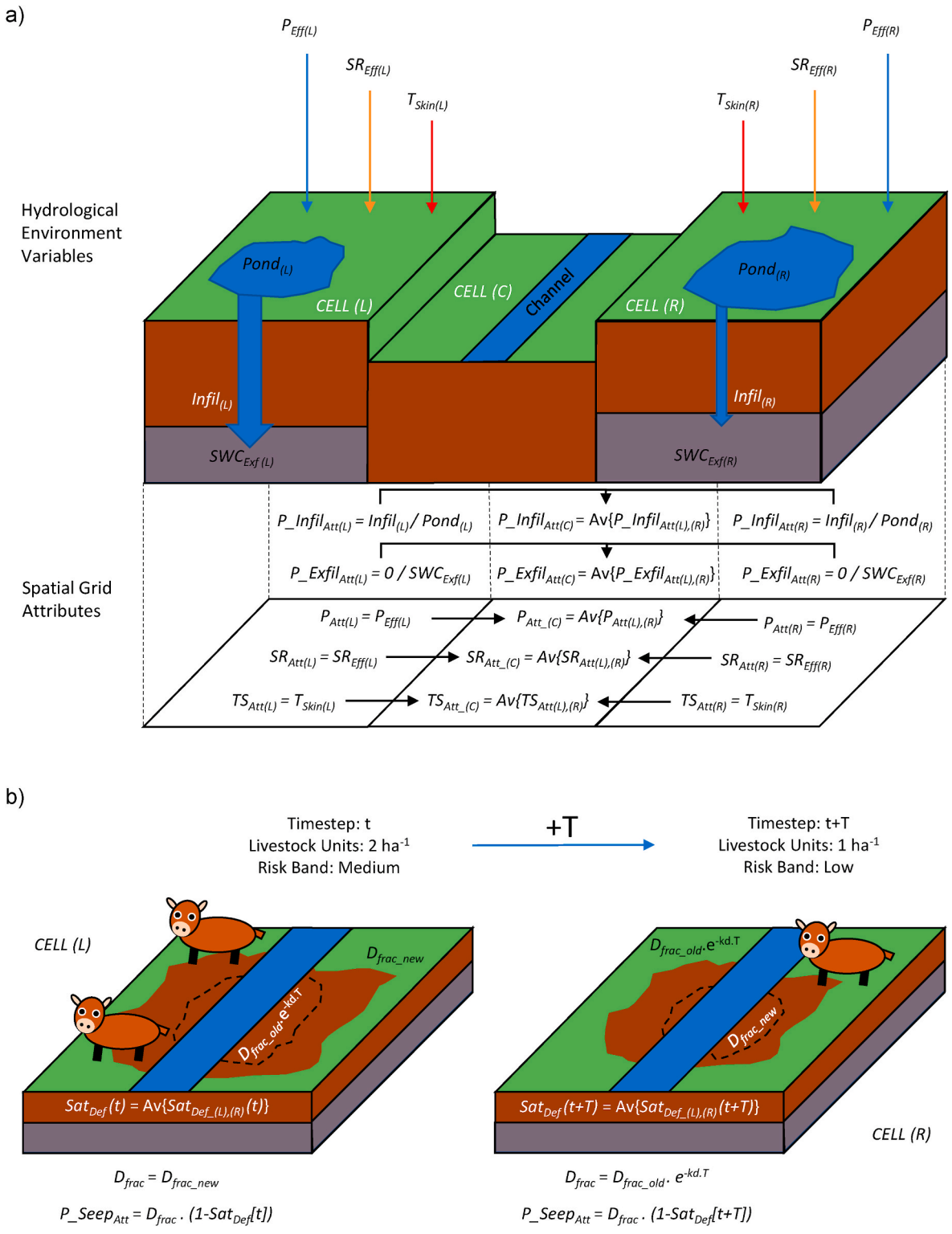

Fig. 3. Examples of how attributes of the spatial grid are calculated: a) Effective precipitation (P_{Eff}), effective solar radiation (SR_{Eff}), soil skin temperature (T_{Skin}), ponded surface water prior to infiltration (*Pond*), infiltration flux (*Infil*) and soil water content prior to exfiltration (SWC_{Exf}) for two cells (L and R) of the hydrological environment are used to calculate the attributes effective precipitation (P_{Att}), effective solar radiation (SR_{Att}), soil skin temperature (TS_{Att}), probability of infiltration (P_Infil_{Att}) and probability of exfiltration (P_Exfil_{Att}) for the corresponding cells of the spatial grid, whilst the average of these are used as the attributes of a channel cell (C) into which the cells drain; b) P_Seep_{Att} is calculated for two sequential timesteps – at timestep t, the new damage fraction (D_{frac_new}) is used in the calculation as the presence of a moderate number of livestock causes this to exceed the damage fraction from the previous timestep (D_{frac_old}) subject to exponential decay at rate k_d, whilst the reduction in livestock numbers at timestep $t + T$ causes the reverse to then be true.

Table 2
The dynamic and fixed attributes of FIO-agents.

FIO-agent attribute	Description	Possible values
Dynamic		
Attached-stream sediment	An abstract representation of whether the FIO-agent is attached to stream sediment once in the channel	• *0*: Not attached (*Initial value*) • *1*: Attached
Detached-faeces	An abstract representation of whether the FIO-agent is detached from a "faecal deposit"	• *0*: Not detached (*Initial value for Domain type = Land or Land_Channel*) • *1*: Detached (*Initial value for Domain type = Channel or Seepage*)
Domain type	The domain type that the FIO-agent currently occupies	• *Channel*: The FIO-agent is in a cell containing a channel, in the channel • *Land*: The FIO-agent is in a cell that does not contain a channel • *Land_Channel*: The FIO-agent is in a cell containing a channel, adjacent to the channel • *Seepage*: The FIO-agent is in a cell containing a channel, within an area of degraded soil
Infiltration stage	Denotes where the FIO-agent is within the vertical discretisation of the catchment environment	• *Surface*: The FIO-agent is on the land surface (*Initial value for Domain type = Land or Land_Channel*) • *Soil*: The FIO-agent is in the soil (*Initial value for Domain type = Seepage*)
Life state	Denotes if the FIO-agent is alive or dead	• *0*: Dead • *1*: Alive (*Initial value*)
Movement tracker	Tracker to record if the FIO-agent has moved on the current timestep	• *0*: The FIO-agent has not moved (*Initial value*) • *1*: The FIO-agent has moved
Timestep	The current timestep	• The current timestep
X co-ordinate	The X co-ordinate of the FIO-agent	• The current X co-ordinate of the FIO-agent
Y co-ordinate	The Y co-ordinate of the FIO-agent	• The current Y co-ordinate of the FIO-agent
Fixed		
Land parcel	The ID of the land parcel into which the FIO-agent initially spawned	• Land parcel ID
Livestock type	The ID of the host animal from which the FIO-agent came	• Livestock ID
Memory		
Location memory	The X–Y co-ordinates previously occupied by the FIO-agent	NA
Timestep memory	The timesteps during which the FIO-agent has been in the model domain	NA
Infiltration stage memory	The history of infiltration stages the FIO-agent has experienced	NA
Domain type memory	The history of domain types that the FIO-agent has occupied	NA

agents are spawned, their fixed attributes and initial values for dynamic attributes are set. Next, for each FIO-agent in turn (Fig. 4b), behaviours resulting from fate and transport processes operating at the sub-field scale during the current timestep are simulated via the small-scale behaviour sub-models (Section 3.7.2). Once the behaviour of all FIO-agents has been simulated, those with *Life state* = 0 are removed during a "clean-up" operation, before the spatio-temporal outputs for the current timestep are finally saved (Fig. 4a).

For each FIO-agent, the sequence of small-scale behaviour sub-models that runs depends on certain attribute values (Fig. 4b), and may include Die-off (Section 3.7.2.1; Fig. 5), Detachment (Section 3.7.2.2; Fig. 6), Surface Routing (Section 3.7.2.3; Fig. 7), Seepage (Section 3.7.2.4; Fig. 8) and Channel Routing (Section 3.7.2.5; Fig. 9). Different sub-models utilise various attributes of the spatial grid as input and have the potential to update the dynamic and associated memory attributes of the FIO-agent. Memory attributes are also updated for the start of a new timestep if the FIO-agent was already present in the model domain (Fig. 4b). There are six conditions that will cause MAFIO to move on to simulate the next FIO-agent: 1) the Die-off Sub-model results in the FIO-agent dying; 2) the FIO-agent is still attached to a "faecal deposit" after running the Detachment Sub-model and hence cannot move during the current timestep; 3) the FIO-agent is in the soil following the Surface Routing Sub-model running on the current timestep; 4) the FIO-agent remains in an area of degraded soil following the Seepage Sub-model being run on the current timestep; 5) the FIO-agent is deposited on the stream bed as it is routed through the channel by the Channel Routing Sub-model; or 6) the FIO-agent is successfully routed to the catchment outlet by the Channel Routing Sub-model (Fig. 4b).

3.4. Design concepts

Of the 11 design concepts outlined in the ODD protocol, only basic principles, emergence, sensing, interaction, stochasticity and observation are relevant to MAFIO.

3.4.1. Basic principles

The process-based, spatially-distributed nature of MAFIO reflects the need for such models to further quantitative process understanding of water quality dynamics (de Brauwere et al., 2014) and to overcome limitations associated with lumped/semi-distributed model structures (Section 2.1). The principles underlying the process-based simulation of FIO loading, die-off and transport within MAFIO are as follows (Fig. 1).

The Defecation Sub-model (Section 3.7.1) conceptualises FIO loading by livestock, assumed to be the main source of FIOs in agricultural areas (Chadwick et al., 2008). Loading by different livestock types can be simulated to account for differences in defecation characteristics (e.g. concentration of FIOs in faeces; Oliver et al., 2018) and FIO behaviour (e.g. die-off kinetics; Avery et al., 2004). FIO loading can be onto the land or directly into the channel via direct deposition (Chadwick et al., 2008). The potential for the latter depends on farm management, such as whether streams are fenced/gated off from the surrounding fields (Kay et al., 2007; Vinten et al., 2008). A sample of the FIOs loaded into the environment by individual animals is simulated in MAFIO by introducing one FIO-agent into the simulation for every specified number of FIOs loaded in reality (c.f. Parry and Bithell, 2012). This ensures that FIO-agents are introduced in correct proportions to account for different loadings by individual animals and livestock types. The accuracy with which the behaviour of the true population of FIOs can be inferred and resolved will vary with the number of FIOs for which

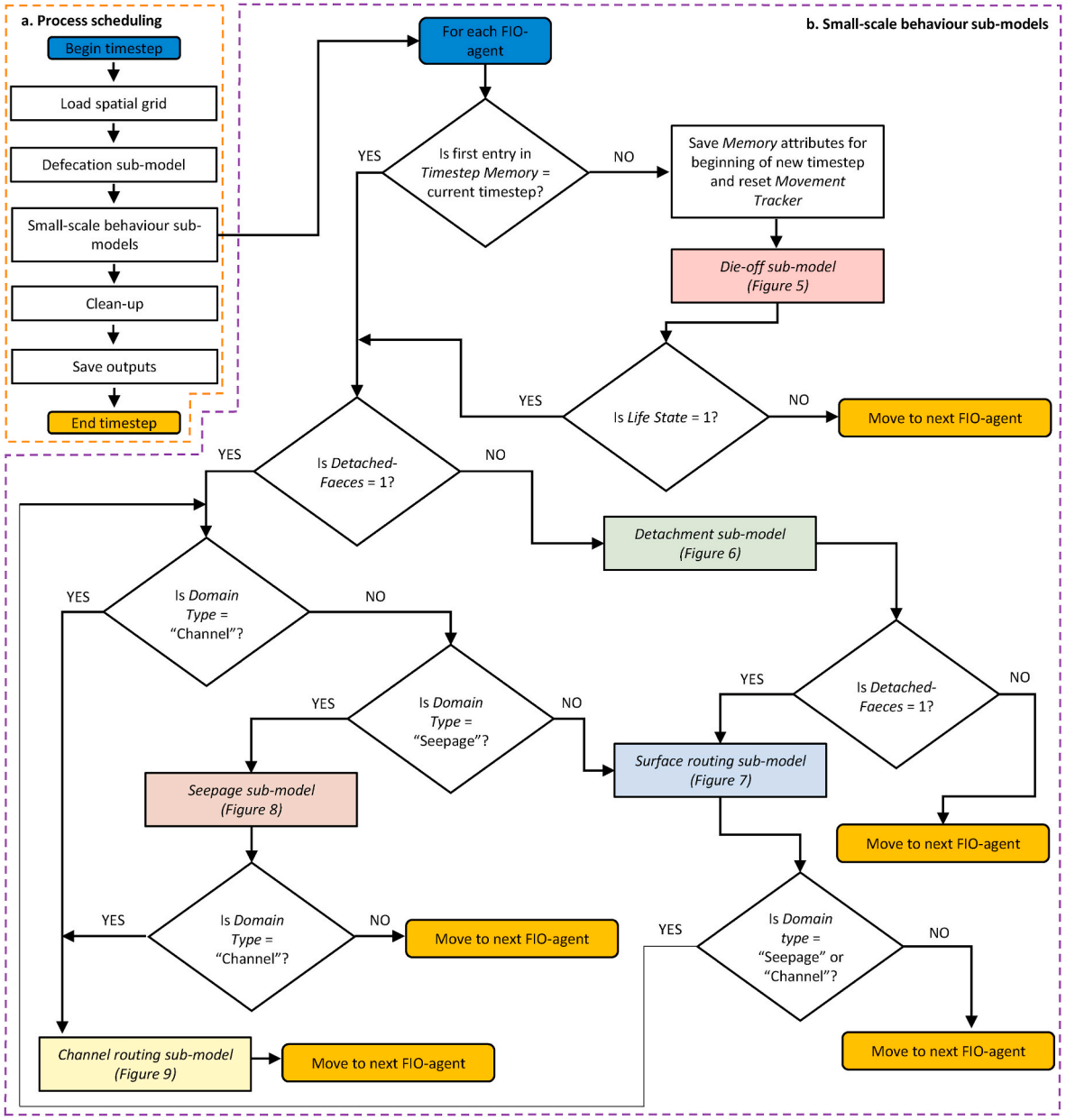

Fig. 4. Flowcharts showing a) The scheduling of processes during each timestep of MAFIO; b) The sequence of small-scale behaviour sub-models that can be run for FIO-agents conditional on the values of certain attributes. Detailed flowcharts showing the functioning of individual sub-models in b) are provided in Figs. 5–9.

an FIO-agent is introduced and the consequent total number of FIO-agents in the simulation (c.f. Parry and Bithell, 2012; Reaney, 2008).

The Die-off Sub-model (Section 3.7.2.1) conceptualises the die-off of FIOs in response to temperature and solar radiation once in the environment; these factors are often considered the most important in determining FIO mortality (Blaustein et al., 2013; Hipsey et al., 2008). Both are used to determine the die-off of FIO-agents on the land surface, whilst only temperature is used for die-off of FIO-agents in the soil (Whitehead et al., 2016). Given the likely short (<1 day) travel-times of water in small agricultural streams, it is assumed that die-off of FIOs once in the stream is negligible. It may also be possible for FIOs to grow in the environment (e.g. Muirhead, 2009; Soupir et al., 2008). However, given the limited extent to which this process is understood and the consequent uncertainty (e.g. Hipsey et al., 2008; Oliver et al., 2010), it is not considered in this initial development of MAFIO.

The Detachment Sub-model (Section 3.7.2.2) simulates the release of FIO-agents from faecal deposits using a model based on effective precipitation depth and the assumption that this metric is sufficient to allow observed release kinetics to be captured (Blaustein et al., 2015b). Once released, lateral transport of FIOs can occur via overland flow, whilst infiltration may cause FIOs to move vertically into the soil where soil retention effects including straining and filtering may limit their further transport (Oliver et al., 2005a). Consequently, lateral movement of FIO-agents in MAFIO is only possible via overland flow; infiltrating FIO-agents are retained in the soil. However, FIOs are generally retained in the upper few centimetres of the soil, which may be within the "effective depth of interaction" of the soil with surface flows that could facilitate lateral movement (Stocker et al., 2015). To account for this, infiltrated FIO-agents can be exfiltrated back to the surface for transport

in overland flow. These processes of lateral movement, infiltration and exfiltration are conceptualised within the Surface Routing Sub-model (Section 3.7.2.3).

In agricultural settings, compaction in areas of animal congregation or locations where there is a high frequency of livestock traffic (e.g. gateways between fields) can cause soil degradation (Bilotta et al., 2007). If proximal to the channel, this can increase the likelihood that these low permeability areas will seep water and contaminants to the stream, potentially causing them to act as chronic sources of FIOs (Davies-Colley et al., 2004). For cells containing a channel where degraded soil is present, the Seepage Sub-model (Section 3.7.2.4) simulates this transfer to the stream.

Once in the stream, FIOs have an increased likelihood of settling to the channel bed if attached to sediment (Pachepsky and Shelton, 2011). Several attempts have been made to model the settling of sediment-associated FIOs (e.g. Collins and Rutherford, 2004; Hipsey et al., 2008; Jamieson et al., 2005; Kim et al., 2010). However, substantial uncertainty still surrounds exactly how the process of settling should be represented (Pachepsky and Shelton, 2011; Pandey et al., 2012). Consequently, a simple distance-decay approach is initially adopted in MAFIO, with the likelihood of an FIO-agent settling to the streambed increasing with distance travelled within the stream. The exact form of the relationship can be assumed to implicitly account for different factors (e.g. settling velocity) affecting the likelihood of settling (Kay and McDonald, 1980). The association of FIO-agents with stream sediment and their resulting transport in the stream is simulated by the Channel Routing Sub-model (Section 3.7.2.5). Whilst resuspension of FIOs has been found to contribute to stream FIO dynamics (Pachepsky and Shelton, 2011), this is not currently included in MAFIO as in the relatively low-energy streams characteristic of small agricultural catchments it is assumed that settling will be the dominant process except in the largest flow events (c.f. Nagels et al., 2002). Therefore, only settling was considered to assess whether this is sufficient to capture observed FIO dynamics (c.f. Clark et al., 2011).

3.4.2. Emergence

The model environments impose certain constraints on how simulations can unfold. Livestock counts, stream accessibility to livestock and locations of degraded soil determine into which land parcels FIO-agents can be spawned, where direct deposition may occur and where FIO-agents may chronically seep to the stream, respectively; thus, the field-scale distribution of *possible* FIO source areas is relatively imposed. Furthermore, the spatial distribution and magnitude of hydrological fluxes dictate where there are active flow paths capable of transporting FIO-agents. However, it is from simulating the small-scale behaviour of FIO-agents and their interaction with the variables of the model environments that the following simulated characteristics emerge: 1) The spatial distribution of FIO-agents stored on the surface and in the soil at the sub-field scale; 2) Fluxes of FIO-agents in the stream; 3) The pathways taken by FIO-agents from their source areas to the stream; 4) Time-varying contributions of FIO-agents to the stream made by different transfer mechanisms (overland flow, seepage and direct deposition) and host livestock; 5) The total number of FIO-agents active in the model domain at a given time.

3.4.3. Sensing and interaction

The behaviour of an FIO-agent is simulated based on it knowing its own attributes and those of the grid cell it currently occupies. Consequently, there are no interactions between FIO-agents, with all behaviours happening independently. A one-way interaction occurs between the FIO-agents and the spatial grid, with the FIO-agents responding to the conditions in the grid cell they occupy.

3.4.4. Stochasticity

MAFIO employs the following stochastic elements. In the Defecation Sub-model, uniform random sampling (c.f. Dorner et al., 2006; Haydon and Deletic, 2006) determines the spatial grid cells into which new FIO-agents are spawned to reflect natural variability (Benhamou, 2006, 2013) and uncertainty in livestock movement and consequent defecation locations at the sub-field scale. If an FIO-agent spawns in a cell containing a channel to which livestock have access, stochasticity is used to ensure that direct deposition occurs with a frequency that reflects the size of the channel relative to the total cell area. Whilst the number of FIOs shed by individual animals of different livestock types is currently deterministic (i.e. relies on fixed parameters to specify factors such as FIO concentrations in faeces, allowing MAFIO to be used in data-limited situations), variability and uncertainty in this variable could be accounted for stochastically where a basis exists for defining reasonable sampling distributions for relevant parameters (e.g. Schijven et al., 2015).

The Die-off, Detachment and Channel Routing Sub-models employ stochasticity to define the behaviour of individual FIO-agents based on governing equations that describe population-level behaviour. In the Surface Routing Sub-model, stochasticity conserves "concentrations" of FIO-agents in hydrological stores (i.e. Ponded Surface Water and Soil Water) and fluxes (e.g. infiltration, exfiltration). With the assumption that FIO-agents are fully and uniformly mixed within a hydrological store, probabilities define the behaviour of individual FIO-agents such

Table 3
The inputs required to characterise the catchment environment in MAFIO.

Input	Description
Maps	
Catchment spatial extent	A binary map (1 and NaN) indicating the cells which comprise the catchment.
Cells upslope of channel	A map defining the cells which are immediately upslope of each cell containing a channel. The value of each cell is the unique channel ID of the cell containing a channel into which it drains.
Cells upslope of degraded soil	A map defining the cells which are immediately upslope of each cell containing an area of degraded soil. The value of each cell is the unique seepage ID of the cell containing degraded soil into which it drains.
Channel location	A map defining cells containing a channel. Each cell with a channel has a unique channel ID.
Channel width	A map of the channel widths for each cell containing a channel.
Degraded soil locations	A map indicating the locations of where soil adjacent to the channel is degraded. Each cell with degraded soil has a unique seepage ID.
Distribution of land parcels	A map of the land parcels in the catchment. Each land parcel has a unique ID.
Local drainage direction	A map of flow directions for each cell in the catchment based on the D8 algorithm.
Timeseries	
Livestock counts	A timeseries detailing the count(s) of each livestock type for each land parcel.
Stream access	A timeseries detailing for each unique channel ID, whether livestock from each land parcel in its list of associated land parcels have access to the stream.
Lists	
Land parcel(s) associated with degraded soil	For each unique seepage ID, a list of the land parcel(s) from which livestock can contribute to soil degradation.
Stream-associated land parcel(s)	For each unique channel ID, a list of the land parcel(s) on either side of the stream.

Table 4
The parameters in the sub-models of MAFIO and their initial values based on simulating *E. coli* for sheep and cattle with a daily timestep.

Parameter	Description	Initial values*	Units	References*
Defecation sub-model				
*faecesConc***	The concentration of FIOs in livestock faeces	[1.73×10^6, 4.18×10^5]	$MPN^{+}\ g^{-1}$	Avery et al. (unpublished data)
FWeight	The mass of a single defecation	[58.3, 2300]	grams	Welch (1982)^
agentsRepresent	The number of FIOs for which an FIO-agent is introduced	4.18×10^5	MPN FIO-agent^{-1}	–
defecationsPerDay	The number of times an animal defecates per day	[16, 12]	day^{-1}	Welch (1982)
Die-off sub-model				
k_0**	Inactivation rate constant at a reference temperature of 20 °C	[0.242, 0.090]	day^{-1}	[Moriarty et al. (2011); Himathongkham et al. (1999)]
θ**	Temperature sensitivity parameter	[1.095, 1.069]	–	[Moriarty et al. (2011); Himathongkham et al. (1999)]
α_{SR}	Proportionality constant for die-off due to solar radiation	1	–	Thomann and Mueller (1987)
Detachment sub-model				
$k_{release}$**	Rate constant for detachment of FIO-agents	0.153	cm^{-1}	Blaustein et al. (2016)^
Seepage sub-model				
LUs	The number of livestock units represented by an individual animal	[0.12, 1]	LU^{++}	Natural England (2013)
riskBands	The boundaries of each risk band	$0 < Low \leq 1$ $1 < Medium < 4$ $High \geq 4$	$LU\ ha^{-1}$	Oliver et al. (2009)
dFrac_Bands	The damage fractions associated with each risk band	Low: 0.43 Medium: 0.53 High: 0.72	–	Sheath and Carlson (1998)
k_d	The rate constant for damage fraction decay	0.0063	day^{-1}	Elliot et al. (2002)^
Channel routing sub-model				
fracAttached	Fraction of FIO-agents attached to stream sediment	0.8	–	Hipsey et al. (2008)
λ_{sed}	Distance decay rate constant	0.00037	m^{-1}	Kay and McDonald (1980)

* Values in square brackets are for [sheep, cattle];+MPN: most-probable number; ++ LU: livestock units.
** Parameter values are for *E. coli*; ^ Derived from data presented in references.

that when a hydrological flux is generated, the proportion of water from the store that generates the flux is equal to the proportion of FIO-agents in the store that it transports. Finally, the Seepage Sub-model employs stochasticity to represent the general tendency for increased seepage potential as soils degrade in response to greater livestock compaction and wetness, without explicitly representing poorly-defined seepage processes (c.f. Collins and Rutherford, 2004).

3.4.5. Observation

For each timestep, the simulated flux of FIO-agents in the stream is recorded for each cell containing a channel. This facilitates comparison with spatially-distributed observations of stream FIO loads. For FIO-agents ultimately reaching the catchment outlet, the following attributes are observed: *Domain type memory*, *Livestock type*, *Location memory* (Table 2). These observations are fundamental to MAFIO fulfilling its stated purpose (Section 3.1) as they allow derivation of the transfer mechanisms, host animals and pathways contributing FIOs to the stream at each timestep. The number and spatial distribution of active FIO-agents in the model domain is also observed at each timestep so FIO-agent fluxes can be linked to storage dynamics. Simulated outputs should be observed for an ensemble of model runs using the same inputs and parameterisation to account for the effects of stochasticity in MAFIO (Abdou et al., 2012).

3.5. Initialisation

For the first timestep, an initial value of D_{frac} must be specified for each cell containing degraded soil to account for previous actions of livestock. Ideally, values should be based on livestock counts prior to the simulation start date; otherwise representative average counts (e.g. from the simulation period) may be used. In addition, each land parcel may be initialised with FIO-agents in the soil to allow for the possibility that FIOs from soil reservoirs reflecting past grazing patterns (which may persist for several months) contribute to FIO dynamics of the simulation period (Muirhead, 2009). Specifically, a number of FIO-agents can be randomly seeded in the soil of each land parcel at the start of the simulation based on estimates of FIOs in the soil derived from concentrations measured in the upper soil profiles of each land parcel, past livestock counts or literature values. Whilst initial values of D_{frac} and the number of FIO-agents to be seeded in the soil are consistent between simulations, the spatial locations of the latter are determined stochastically and therefore vary.

3.6. Inputs

MAFIO requires input data to characterise the catchment and hydrological environments. Table 3 documents those required for the catchment environment. The hydrological environment is characterised using inputs from an appropriate hydrological model, as detailed in Section 3.2.2. Spatially distributed inputs are provided as map files (.map) compatible with PCRaster (Karssenberg et al., 2010).

3.7. Sub-models

MAFIO consists of the Defecation Sub-model and several small-scale behaviour sub-models (Fig. 4). The required parameters and their initial values based on daily simulation of *E. coli* from sheep and cattle are given in Table 4. In the following equations, a sub-script *[LS]* denotes that the value of the variable or parameter varies based on the livestock type. Meanwhile, a sub-script *[xy]* indicates an attribute or derived variable of the spatial grid whose value varies depending on the cell.

3.7.1. Defecation Sub-model

The Defecation Sub-model runs once every timestep (Fig. 4a). The parameter *lsType* indicates which livestock are being considered in the simulation. For each livestock type, the number of FIO-agents introduced into the simulation for each defecation of an individual animal is calculated. First, the number of FIOs per gram of faeces is specified by

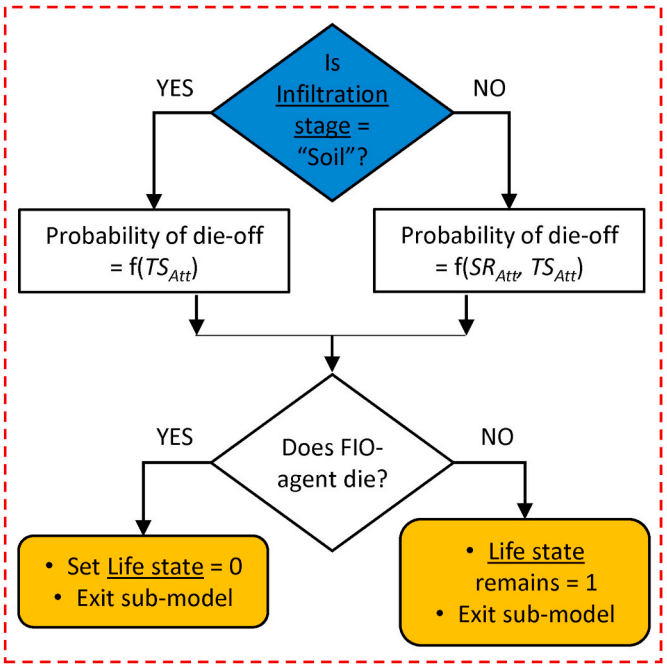

Fig. 5. Detailed flowchart for the Die-off Sub-model. Attributes of the spatial grid used by the sub-model are denoted by italic font (TS_{Att} = soil skin temperature, SR_{Att} = effective solar radiation) whilst attributes of FIO-agents that are used by the sub-model are underlined. The blue shape denotes the starting operation of the sub-model whilst orange shapes denote where the sub-model can exit. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

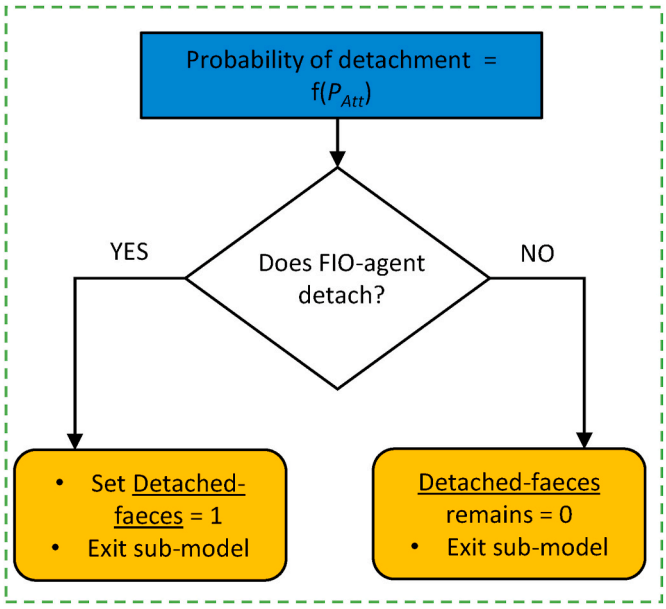

Fig. 6. Detailed flowchart for the Detachment Sub-model. Attributes of the spatial grid used by the sub-model are denoted by italic font (P_{Att} = effective precipitation) whilst attributes of FIO-agents that are used by the sub-model are underlined. The blue shape denotes the starting operation of the sub-model whilst orange shapes denote where the sub-model can exit. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the parameter *faecesConc*. Total FIO loading per defecation (*FIO_Load*) is then calculated as:

$$FIO\_Load_{[LS]} = faecesConc_{[LS]} \cdot FWeight_{[LS]} \quad (5)$$

where *FWeight* is the mass of a single defecation in grams. This value is then translated into a number of FIO-agents introduced into the simulation per defecation (*LS_Agents*), based on a specified number of FIOs to be represented by an FIO-agent (*agentsRepresent*):

$$LS\_Agents_{[LS]} = \text{ceil}\left(\frac{FIO\_Load_{[LS]}}{agentsRepresent}\right) \quad (6)$$

For each livestock type, the Defecation Sub-model then introduces FIO-agents to land parcels based on livestock counts for the current timestep. For each individual of a given livestock type, the number of defecations per timestep is specified by *defecationsPerDay*. For each defecation, a cell of the current land parcel is chosen using uniform random sampling. At this location, the number of FIO-agents specified by *LS_Agents* are spawned. If the cell contains a channel (specified by $Chan_{Att}$ of the cell) accessible to livestock, the probability that the FIO-agents spawn in the channel to simulate direct deposition is:

$$P(Channel) = P\_Chan_{Att_{[xy]}} \quad (7)$$

If they do not spawn in the channel or livestock do not have channel access, the FIO-agents either spawn in an area of degraded soil if present (specified by $degSoil_{Att}$), or on the land adjacent to the channel. Whilst the fixed nature of *faecesConc*, *FWeight*, and *defecationsPerDay* dictates that the number of FIO-agents introduced by individual animals of a given livestock type is deterministic, as noted in Section 3.4.4, stochastic treatment of this variable is possible if data exist to constrain appropriate sampling distributions for the aforementioned parameters.

3.7.2. Small-scale behaviour sub-models

Small-scale behaviour sub-models are applied every timestep to each FIO-agent in the model domain (Fig. 4a). The exact sequence of sub-models that runs depends on certain attributes of the FIO-agent in question (Fig. 4b). MAFIO will move on to simulate the behaviour of the next FIO-agent only once one of the six conditions in Section 3.3 are met.

3.7.2.1. Die-off Sub-model. The Die-off Sub-model (Fig. 5) only runs for FIO-agents that spawned before the current timestep (Fig. 4b), with die-off modelled based on first-order kinetics following Chick's Law (Chick, 1908):

$$\frac{N_{Alive}}{N_0} = e^{-k_{dead} \cdot T} \quad (8)$$

where N_{Alive}/N_0 is the fraction of the original population of FIOs (N_0) that is still alive (N_{Alive}) after being subject to exponential decay over a timestep, T, at a rate specified by the inactivation rate constant parameter k_{dead} with units of T^{-1}. To translate this population-level behaviour to individual FIO-agents, a probability that an FIO-agent dies is derived from Eq. (8) that is a function of the attributes of the grid cell it occupies and its host animal:

$$P(Dies) = 1 - e^{-k_{dead_{[LS][xy]}} \cdot T} \quad (9)$$

Eq. (9) assumes all FIO-agents from a given livestock type within a cell can be treated as a single population experiencing the same conditions and probability of die-off. The parameter k_{dead} is an overall inactivation rate constant obtained by summing the individual inactivation rate constants arising from factors (temperature and solar radiation in MAFIO) assumed to cause die-off (Chapra, 1997). The inactivation rate constant due to temperature (k_{temp}) is based on the equation proposed by Mancini (1978):

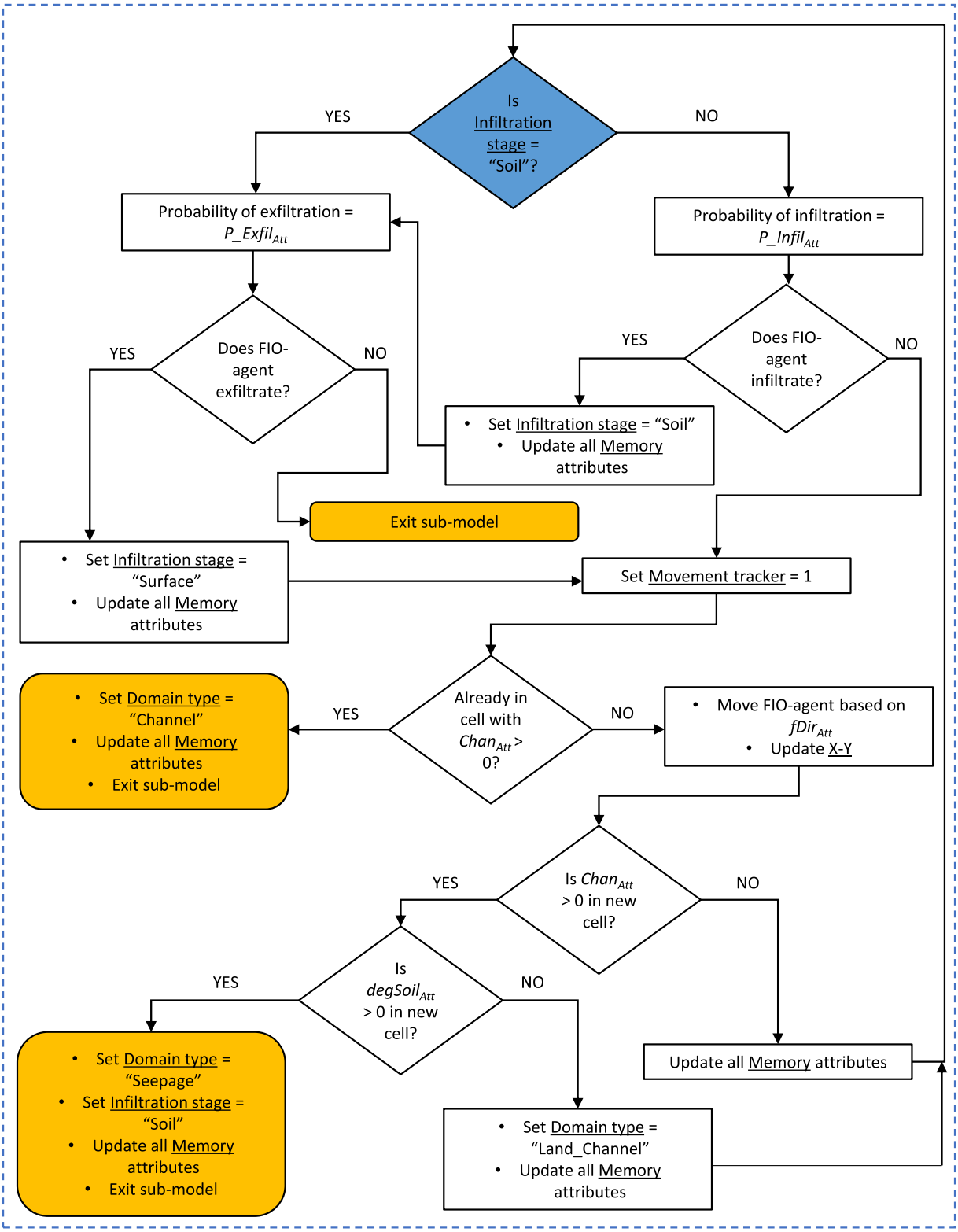

Fig. 7. Detailed flowchart for the Surface Routing Sub-model. Attributes of the spatial grid used by the sub-model are denoted by italic font (P_Infil_{Att} = probability of infiltration, P_Exfil_{Att} = probability of exfiltration, $Chan_{Att}$ = channel cell, $fDir_{Att}$ = flow direction, $degSoil_{Att}$ = degraded soil in cell) whilst attributes of FIO-agents that are used by the sub-model are underlined. The blue shape denotes the starting operation of the sub-model whilst orange shapes denote where the sub-model can exit. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

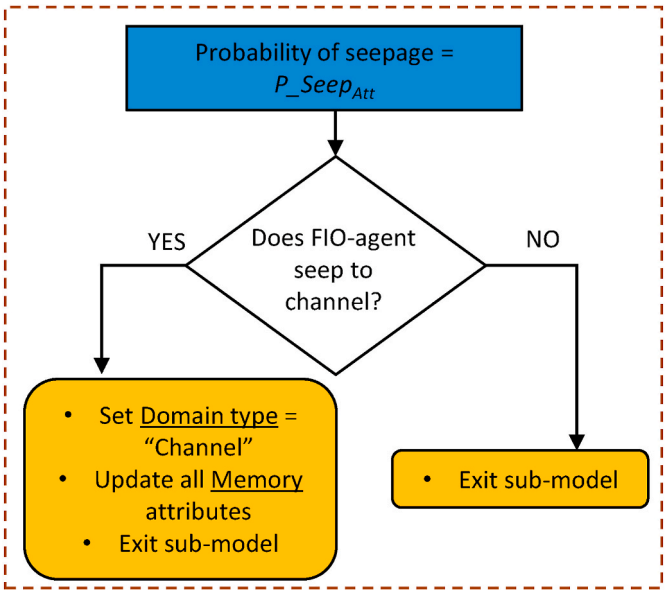

Fig. 8. Detailed flowchart for the Seepage Sub-model. Attributes of the spatial grid used by the sub-model are denoted by italic font (P_Seep_{Att} = probability of seepage) whilst attributes of FIO-agents that are used by the sub-model are underlined. The blue shape denotes the starting operation of the sub-model whilst orange shapes denote where the sub-model can exit. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

$$k_{temp_{[LS][xy]}} = k_{0_{[LS]}} . \theta_{[LS]}^{TS_{Att[xy]}-20} \quad (10)$$

where k_0 is the inactivation rate constant (units of T^{-1}) at a reference temperature of 20 °C, and θ is a temperature sensitivity parameter (dimensionless). Meanwhile, the inactivation rate constant due to solar radiation (k_{solar}) is calculated as (Thomann and Mueller, 1987):

$$k_{solar_{[LS][xy]}} = \alpha_{SR} . SR_{Att_{[xy]}} \quad (11)$$

where α_{SR} is a proportionality constant. When solar radiation is expressed in units of ly hr^{-1}, α_{SR} is approximately unity (Thomann and Mueller, 1987). The final value of k_{dead} depends on whether the FIO-agent is in the soil (Eq. (12a)) or on the surface (Eq. (12b)):

$$k_{dead_{[LS][xy]}} = k_{temp_{[LS][xy]}} \quad (12a)$$

$$k_{dead_{[LS][xy]}} = k_{temp_{[LS][xy]}} + k_{solar_{[LS][xy]}} \quad (12b)$$

If an FIO-agent dies, its attribute *Life state* is set to "0" so that it is no longer moved or tracked for the rest of the simulation.

3.7.2.2. Detachment Sub-model. The Detachment Sub-model (Fig. 6) only runs for FIO-agents with the attribute *Detached-faeces* = "0" and *Life state* = "1" (Fig. 4b). In MAFIO, the exponential model of Bicknell et al. (1997) simulates the release of FIOs from faeces:

$$\frac{N_{Release}}{N_{Feaces}} = 1 - e^{-k_{release} . P} \quad (13)$$

where $N_{Release}/N_{Faeces}$ is the fraction of FIOs in the faeces (N_{Faeces}) released ($N_{Release}$), $k_{release}$ is a rate constant with units of cm^{-1} and P is precipitation (centimetres). This is used because it is a single-parameter model that performs comparably to alternative multi-parameter models (Blaustein et al., 2015a, 2016). From Eq. (13), the probability that an FIO-agent is detached from its faecal deposit is derived as:

$$P(Detached_{Faeces}) = 1 - e^{-k_{release} . P_{Att[xy]}} \quad (14)$$

Despite the value of $k_{release}$ likely varying with livestock type (e.g. reflecting differently-composed faeces; c.f. Hodgson et al., 2009), a current lack of information (c.f. Sokolova et al., 2018) dictates this parameter be specified for the type of FIO being modelled. If the FIO-agent is detached, then *Detached-faeces* is set to "1" so that the Surface Routing Sub-model can run (Figs. 4b and 6). Otherwise, the FIO-agent cannot move for the current timestep.

3.7.2.3. Surface Routing Sub-model. The Surface Routing Sub-model (Fig. 7) only runs for FIO-agents with the attributes *Life state* = "1", *Detached-faeces* = "1" and *Domain type* = "Land" or "Land_Channel" (Fig. 4b). The sub-model assumes full and immediate mixing of FIO-agents within hydrological stores and fluxes (c.f. Dorner et al., 2006); consequently, as soon as an FIO-agent enters a new hydrological store (i.e. Ponded Surface Water or Soil Water), it is immediately available for transport from that store in a hydrological flux (e.g. overland flow, exfiltration). The validity of this assumption varies with cell size, being more valid for smaller grid cells (Reaney, 2008).

For each timestep, the Surface Routing Sub-model proceeds as follows for each applicable FIO-agent (Fig. 7). If the attribute *Infiltration stage* of the FIO-agent is not equal to "Soil", its infiltration into the soil is determined based on a probability of infiltration:

$$P(Infiltrate) = P\_Infil_{Att_{[xy]}} \quad (15)$$

If the FIO-agent infiltrates, its attribute *Infiltration stage* is set to "Soil". Whether it exfiltrates back to the surface then depends on a probability of exfiltration:

$$P(Exfiltrate) = P\_Exfil_{Att_{[xy]}} \quad (16)$$

If the *Infiltration stage* attribute of the FIO-agent was already "Soil" to begin with, only the potential for exfiltration is evaluated. If the FIO-agent exfiltrates, its *Infiltration stage* attribute is set to "Surface". If after infiltration and exfiltration the FIO-agent has *Infiltration stage* equal to "Surface", then it is assumed to be transported laterally in overland flow, and its attribute *Movement tracker* is set to "1". If the FIO-agent is already in a cell containing a channel (cell with $Chan_{Att} > 0$), then its attribute *Domain type* is set to "Channel" to denote that overland flow has transported the FIO-agent from the land adjacent to the channel into the channel itself; the sub-model then exits and the Channel Routing Sub-model runs (Fig. 4b). If not already in a cell with $Chan_{Att} > 0$, the FIO-agent moves to the downslope cell based on $fDir_{Att}$ of the cell it currently occupies. If the FIO-agent moves into a cell with $Chan_{Att} > 0$, what happens next depends on the presence of degraded soil (specified by $degSoil_{Att}$). If present, the *Domain type* and *Infiltration stage* attributes of the FIO-agent are changed to "Seepage" and "Soil", respectively, and the sub-model exits so the Seepage Sub-model can run (Fig. 4b). If not present, then *Domain type* is set to "Land_Channel" and the Surface Routing Sub-model runs again to determine whether the FIO-agent infiltrates into the soil next to the channel or is routed into the channel. If the latter, *Domain type* is set to "Channel" and the sub-model exits so the Channel Routing Sub-model can run; otherwise MAFIO moves on to simulate the next FIO-agent (Fig. 4b). If the FIO-agent did not move into a cell with $Chan_{Att} > 0$, the Surface Routing Sub-model continues until the FIO-agent infiltrates into the soil but does not exfiltrate, reaches an area of degraded soil or enters the channel; an FIO-agent cannot, therefore, remain on the surface after the Surface Routing Sub-model has run.

3.7.2.4. Seepage Sub-model. The Seepage Sub-model (Fig. 8) runs for FIO-agents with *Life state* and *Domain type* = "1" and "Seepage", respectively (Fig. 4b). It determines whether an FIO-agent in an area of degraded soil adjacent to the channel seeps to the stream based on a probability of seepage:

$$P(Seep) = P\_Seep_{Att_{[xy]}} \quad (17)$$

If the FIO-agent seeps to the stream, its attribute *Domain type* is set to

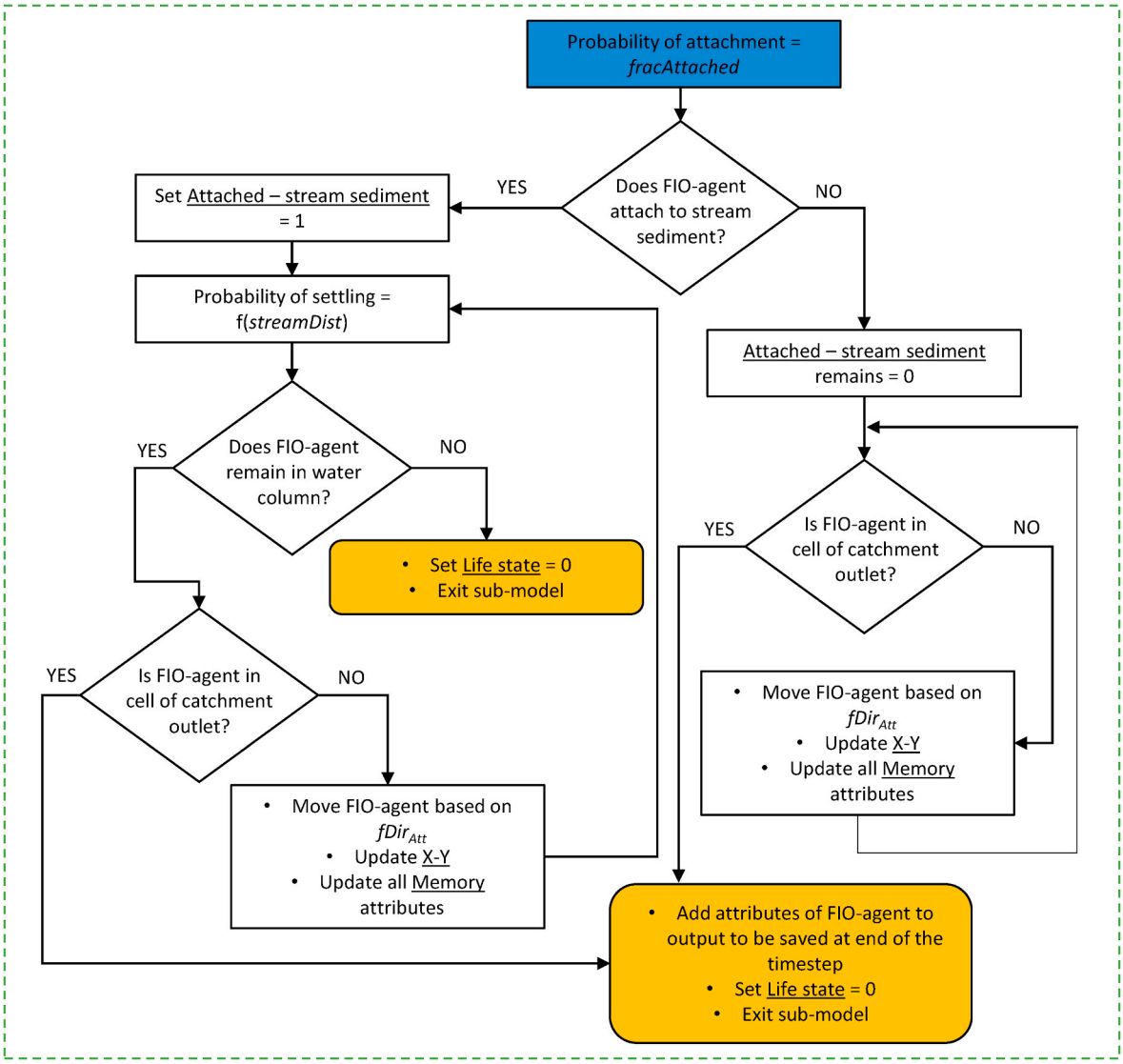

Fig. 9. Detailed flowchart for the Channel Routing Sub-model. Attributes of the spatial grid used by the sub-model are denoted by italic font ($fDir_{Att}$ = flow direction); exceptions are *streamDist* (straight-line length of the channel), which is derived from the flow direction and cell size attributes of the spatial grid, and *fracAttached*, which is a global model parameter denoting the fraction of FIOs assumed to be attached to sediment once in the stream. Attributes of FIO-agents that are used by the sub-model are underlined. The blue shape denotes the starting operation of the sub-model whilst orange shapes denote where the sub-model can exit. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

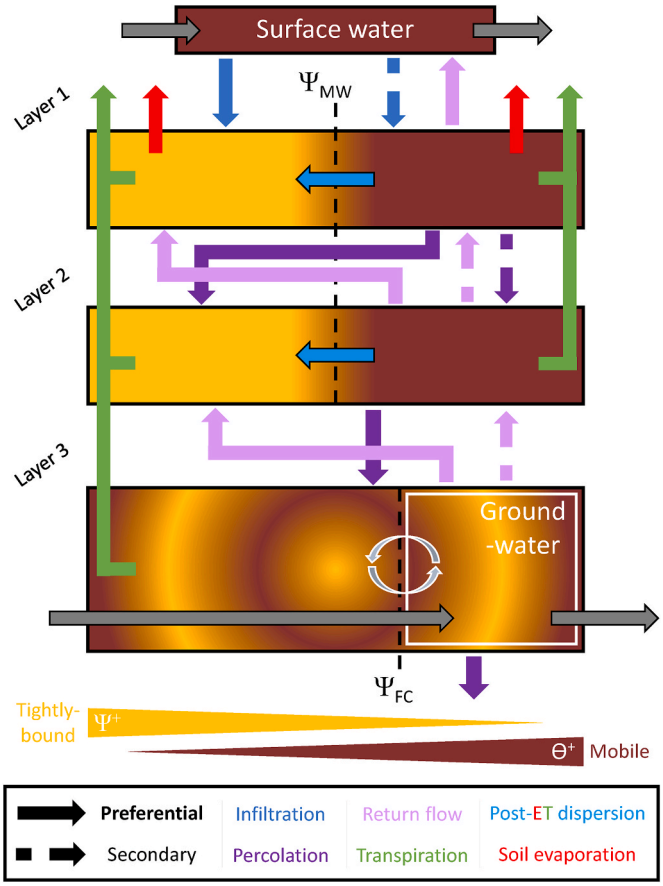

Fig. 10. A conceptual graphic of the below-canopy liquid stores and new water tracking scheme in each pixel of EcH$_2$O-iso. The two-pore domain implementation (tightly-bound [TB] and mobile [MW]) in Layers 1 and 2 uses a tension threshold for mobile water (Ψ_{MW}). Liquid water exchanges between layers (infiltration, percolation and return flow) follow the preferential pathways (solid arrows) unless the recipient layer/domain is full, in which case secondary flow paths (dashed arrows) are triggered. In Layers 1 and 2, soil evaporation and plant transpiration draw from both domains in proportion to their respective water contents. Dispersion replenishes TB after evaporative losses, potentially emptying MW. The groundwater store is defined using a field capacity tension threshold (Ψ_{FC}), and tracking assumes full mixing with the non-saturated Layer 3.

"Channel" so the Channel Routing Sub-model can run (Fig. 4b). Otherwise, the FIO-agent remains in the degraded soil for the current time-step. Whilst the Seepage Sub-model does not utilise any parameters, it is necessary to specify the number of livestock units each animal of a given livestock type represents (*LUs*), the values of livestock units per hectare that constitute the "low", "medium" and "high" risk bands (*riskBands*), the damage fraction associated with each risk band (*dFrac_Bands*) and the rate constant for the exponential decline in damage fraction over time (k_d) used in the calculation of the spatial grid attribute P_Seep_{Att} (Section 3.2.4).

3.7.2.5. Channel Routing Sub-model. The Channel Routing Sub-model (Fig. 9) runs for FIO-agents with *Domain type* = "Channel" (Fig. 4b). First, the sub-model determines whether the FIO-agent attaches to stream sediment when entering the channel based on the parameter *fracAttached*:

$$P(Attached_{StreamSed}) = fracAttached \quad (18)$$

If the FIO-agent does not attach, it is assumed it cannot settle to the streambed and the FIO-agent is routed through the channel network to the catchment outlet. If the FIO-agent attaches, its attribute *Attached-stream sediment* is set to "1", and routing occurs as follows. The probability that the FIO-agent settles to the streambed in its current cell is determined based on length of the channel in the cell (*streamDist*) and the distance-decay model of Kay and McDonald (1980):

$$P(Settle) = 1 - 10^{-\lambda_{sed} \cdot streamDist_{[xy]}} \quad (19)$$

where λ_{sed} is the rate constant parameter in units of m^{-1}. If the FIO-agent does not settle, it is routed to the next downstream cell, and the probability is calculated again. This continues until the FIO-agent settles to the streambed or is routed out of the catchment. Once the FIO-agent reaches the cell representing the catchment outlet, it may still settle to account for the length of the stream in that cell. As with the Surface Routing Sub-model, the Channel Routing Sub-model assumes full and immediate mixing of FIO-agents within each cell. In addition, it assumes that all FIO-agents passing through a cell can be considered as a single population experiencing the same decay with distance, similar to the Die-off Sub-model.

4. A hydrological environment generator: EcH$_2$O-iso

4.1. Overview of EcH$_2$O-iso

EcH$_2$O-iso is a process-based, spatially-distributed ecohydrological model that simulates coupled energy, hydrological and vegetation dynamics (Kuppel et al., 2018b; Maneta and Silverman, 2013) with the capacity to track the stable isotope (^{2}H and ^{18}O) compositions and water ages of simulated hydrological stores and fluxes (Kuppel et al., 2018a). EcH$_2$O-iso has successfully been applied to temperate (Kuppel et al., 2018a) and cold (Smith et al., 2019) regions, and at the plot- (Douinot et al., 2019) to catchment-scale (Knighton et al., 2020; Kuppel et al., 2018a; Smith et al., 2019), demonstrating the versatility of the model in different environments and process domains.

EcH$_2$O-iso has assumed full and immediate mixing for tracer and age tracking within each liquid hydrological store to minimize parameterisation and computing time (Kuppel et al., 2018a). However, limitations to this approach have been highlighted, even in wet, energy-limited catchments where heterogeneous mixing and flow is reduced compared to other ecoclimatic settings (Geris et al., 2015; Kuppel et al., 2018a). More generally, there is a need to consider partial mixing in tracer-aided models (Cain et al., 2019; Knighton et al., 2017) to accommodate possible ecohydrological separation of water in the subsurface – that is its partitioning between water that is tightly-bound to the soil matrix in small pores and mobile water which may predominantly contribute to groundwater recharge and streamflow (Brooks et al., 2010; Goldsmith et al., 2012; Sprenger et al., 2018). To the authors' knowledge, two-pore conceptualisations of the subsurface in physically-based models are only just emerging in catchment studies (Hopp et al., 2020), with efforts largely having been limited to the plot scale (Jackisch and Zehe, 2018; Sprenger et al., 2018; Stumpp and Maloszewski, 2010; Vogel et al., 2010).

To incorporate heterogeneous subsurface mixing, EcH$_2$O-iso has been updated so that the tracked isotopic composition and age of water considers a two-pore domain in the upper two layers of the three-layer soil profile (Fig. 10). The two domains (tightly-bound [TB] and mobile [MW]) are distinguished using a tension threshold for mobile water (Ψ_{MW}), a parameter shared by both layers, with a corresponding water content threshold (SWC_{MW}) determined by the Brooks-Corey pedotransfer functions used in EcH$_2$O-iso (Maneta and Silverman, 2013). In each layer, the MW domain only exists if total water storage in the layer (SWC_t) exceeds the storage capacity of TB ($SWC_t > SWC_{\mathrm{MW}}$). Liquid water fluxes downwards (infiltration and percolation) and upwards (return flow) draw from MW and primarily refill TB, except for surface water considered as MW (Fig. 10). Secondary flow paths (MW to MW) occur if the recipient TB reaches its maximum storage capacity. Soil evaporation and plant transpiration draw from both domains in proportion to their respective water contents. After evaporative losses, dispersion replenishes TB from MW (potentially emptying MW), assuming continuity in water contents between the two domains.

The two-pore domain conceptualisation only affects tracer and age tracking; computation of water fluxes between layers (infiltration/percolation, return flow) and from the subsurface system (soil evaporation, transpiration) is unchanged from previous versions of EcH$_2$O-iso. The groundwater store is also still defined using a field capacity tension threshold (Ψ_{FC}, fixed to 3.36 m), and tracking assumes full mixing within the 3rd soil layer even if unsaturated (Fig. 10). The remaining workings of EcH$_2$O-iso are as described in Maneta and Silverman (2013) and Kuppel et al. (2018a, b), with the further addition of species-dependent root profiles as implemented in Douinot et al. (2019) and Knighton et al. (2020).

4.2. Interfacing EcH$_2$O-iso with MAFIO

EcH$_2$O-iso provides all outputs required to characterise the hydrological environment of MAFIO. However, EcH$_2$O-iso simulates infiltration twice within a grid cell, first following the addition of effective precipitation to ponded surface water and then again after the addition of snowmelt and run-on from upslope cells (see Fig. 1 in Kuppel et al., 2018a). For processes in MAFIO to be consistent with this, two probabilities of infiltration need to be assigned to each cell of the spatial grid (P_Infil1_{Att} and P_Infil2_{Att}). Then, FIO-agents which have been detached on a given timestep and are still in their initial cell may first infiltrate with probability P_Infil1_{Att} and then, if still on the surface, with probability P_Infil2_{Att}, whilst FIO-agents which move into a cell in run-on from upslope can only infiltrate with probability P_Infil2_{Att}. This aside, the process conceptualisation of EcH$_2$O-iso is compatible with the catchment conceptualisation represented by the variables of the hydrological environment of MAFIO (Section 3.2.2). As EcH$_2$O-iso conceptualises three hydrological layers in the soil, $Exfil$, SWC_{Exf} and Sat_{Def} are taken from the upper layer to be consistent with the representation of the upper soil profile in MAFIO.

5. Conclusions

This paper has reported the rationale for and development of MAFIO, a new agent-based model for simulating the behaviour and transport of agents representing FIOs. The model is intended to elucidate the sources and transfer mechanisms contributing FIOs to streams at the sub-field scale in small agricultural catchments, with improved confidence in hydrological process representation afforded by the use of a robust external model for simulating the hydrological environment. Here, the spatially-distributed, tracer-aided ecohydrological model EcH$_2$O-iso was introduced as an exemplar hydrological environment generator. A companion paper (Neill et al., 2020) provides a proof-of-concept application of MAFIO focusing on process representation and the potential the model has for use in a management context.

Software availability

The source code for MAFIO as outlined in this work is available via the University of Aberdeen PURE repository: https://doi.org/10.20392/66f74663-ece3-4a52-8bed-f0cf52d0831a.

The source code for EcH$_2$O-iso is available at: https://bitbucket.org/sylka/ech2o_iso/src/master_2.0/

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Aaron J. Neill: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Doerthe Tetzlaff:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Norval J.C. Strachan:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Rupert L. Hough:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Lisa M. Avery:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Investigation. **Sylvain Kuppel:** Methodology, Software, Writing - original draft, Writing - review & editing. **Marco P. Maneta:** Methodology, Software. **Chris Soulsby:** Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

Funding for this work from the Scottish Government's Hydro Nation Scholars Programme is gratefully acknowledged. Many thanks go to Sim Reaney and John Wainwright for initial discussions regarding environmental agent-based modelling approaches and methods for their implementation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2020.110903.

References

Abdou, M., Hamill, L., Gilbert, N., 2012. Designing and building an agent-based model: 141-166. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (Eds.), Agent-Based Models of Geographical Systems. Springer, Dordrecht, Heidelberg, London, New York.

Ala-aho, P., Tetzlaff, D., McNamara, J.P., Laudon, H., Soulsby, C., 2017. Using isotopes to constrain water flux and age estimates in snow-influenced catchments using the STARR (Spatially distributed Tracer-Aided Rainfall-Runoff) model. Hydrological Earth Systems Science 21, 5089–5110.

Avery, S.M., Moore, A., Hutchison, M.L., 2004. Fate of *Escherichia coli* originating from livestock faeces deposited directly onto pasture. Lett. Appl. Microbiol. 38, 355–359.

Benhamou, S., 2006. Detecting an orientation component in animal paths when the preferred direction is individual-dependent. Ecology 87, 518–528.

Benhamou, S., 2013. Of scales and stationarity in animal movements. Ecol. Lett. 17, 261–272.

Berman, E.S.F., Levin, N.E., Landais, A., Li, S., Owano, T., 2013. Measurement of $\delta^{18}O$, $\delta^{17}O$, and ^{17}O-excess in water by off-Axis integrated cavity output spectroscopy and isotope ratio mass spectrometry. Anal. Chem. 85, 10392–10398.

Bicknell, B.R., Imhoff, J.C., Kittle Jr., J.L., Donigan Jr., A.S., Johanson, R.C., 1997. Hydrological Simulation Program – FORTRAN User's Manual for Version 11. Environmental Protection Agency Report No: EPA/600/R-97/080. U.S. Environmental Protection Agency, Athens, Ga.

Bilotta, G.S., Brazier, R.E., Haygarth, P.M., 2007. The impacts of grazing animals on the quality of soils, vegetation and surface waters in intensively managed grasslands. Adv. Agron. 94, 237–280.

Birkel, C., Soulsby, C., 2015. Advancing tracer-aided rainfall-runoff modelling: a review of progress, problems and unrealised potential. Hydrol. Process. 29, 5227–5240.

Birkel, C., Soulsby, C., Tetzlaff, D., 2014a. Developing a consistent process-based conceptualization of catchment functioning using measurements of internal state variables. Water Resour. Res. 50, 3481–3501.

Birkel, C., Soulsby, C., Tetzlaff, D., 2014b. Integrating parsimonious models of hydrological connectivity and soil biogeochemistry to simulate stream DOC dynamics. J. Geophys. Res. 119, 1030–1047.

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

Blaustein, R.A., Pachepsky, Y., Hill, R.L., Shelton, D.R., 2015a. Solid manure as a source of fecal indicator microorganisms: release under simulated rainfall. Environ. Sci. Technol. 49, 7860–7869.

Blaustein, R.A., Pachepsky, Y., Shelton, D.R., Hill, R.L., 2015b. Release and removal of microorganisms from land-deposited animal waste and animal manures: a review of data and models. J. Environ. Qual. 44, 1338–1354.

Blaustein, R.A., Pachepsky, Y., Hill, R.L., Shelton, D.R., Whelan, G., 2013. *Escherichia coli* survival in waters: temperature dependence. Water Res. 47 (2), 569–578.

Brooks, J.R., Barnard, H.R., Coulombe, R., McDonnell, J.J., 2010. Ecohydrologic separation of water between trees and streams in a Mediterranean climate. Nat. Geosci. 3, 100–104.

Brooks, P.D., Chorover, J., Fan, Y., Godsey, S.E., Maxwell, R.M., McNamara, J.P., Tague, C., 2015. Hydrological partitioning in the critical zone: recent advances and opportunities for developing transferable understanding of water cycle dynamics. Water Resour. Res. 51, 6973–6987.

Cain, M.R., Ward, A.S., Hrachowitz, M., 2019. Ecohydrologic separation alters interpreted hydrologic stores and fluxes in a headwater mountain catchment. Hydrol. Process. 33, 2658–2675.

Chadwick, D., Fish, R., Oliver, D.M., Heathwaite, L., Hodgson, C., Winter, M., 2008. Management of livestock and their manure to reduce the risk of microbial transfers to water – the case for an interdisciplinary approach. Trends Food Sci. Technol. 19, 240–247.

Chapra, S.C., 1997. Surface Water-Quality Modelling. McGraw-Hill, New York.

Chick, H., 1908. Investigation of the laws of disinfection. J. Hyg. 8, 655.

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

Cirbus, J., Podhoranyi, M., 2013. Cellular automata for the flow simulations on the Earth surface, optimisation computation process. Applied Mathematics and Information Sciences 7, 2149–2158.

Clark, M.P., Kavetski, D., Fenicia, F., 2011. Pursuing the method of multiple working hypotheses for hydrological modeling. Water Resour. Res. 47, W09301.

Collins, R., Rutherford, K., 2004. Modelling bacterial water quality in streams draining pastoral land. Water Res. 38, 700–712.

Cooper, J.R., Wainwright, J., Parsons, A.J., Onda, Y., Fukuwara, T., Obana, E., Kitchener, B., Long, E.J., Hargrave, G.H., 2012. A new approach for simulating the redistribution of soil particles by water erosion: a marker-in-cell model. J. Geophys. Res. 117, F04027.

Crooks, A.T., Heppenstall, A.J., 2012. Introduction to agent-based modelling: 85-108. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (Eds.), Agent-Based Models of Geographical Systems. Springer, Dordrecht, Heidelberg, London, New York.

Davies-Colley, R.J., Nagels, J.W., Smith, R.A., Young, R.G., Phillips, C.J., 2004. Water quality impact of a dairy cowherd crossing a stream. N. Z. J. Mar. Freshw. Res. 38 (4), 569–576.

de Brauwere, A., Ouattara, N.K., Servais, P., 2014. Modeling fecal indicator bacteria concentrations in natural surface waters: a review. Crit. Rev. Environ. Sci. Technol. 44, 2380–2453.

Dick, J.J., Tetzlaff, D., Birkel, C., Soulsby, C., 2015. Modelling landscape controls on dissolved organic carbon sources and fluxes to streams. Biogeochemistry 122, 361–374.

Dorner, S.M., Anderson, W.B., Slawson, R.M., Kouwen, N., Huck, P.M., 2006. Hydrologic modeling of pathogen fate and transport. Environ. Sci. Technol. 40, 4746–4753.

Drewry, J.J., 2006. Natural recovery of soil physical properties from treading damage of pastoral soils in New Zealand and Australia: a review. Agric. Ecosyst. Environ. 144, 159–169.

Douinot, A., Tetzlaff, D., Maneta, M., Kuppel, S., Schulte-Bisping, H., Soulsby, C., 2019. Ecohydrological modelling with EcH2O-iso to quantify forest and grassland effects of water partitioning and flux ages. Hydrol. Process. 33, 2174–2191.

Elliot, A.H., Tian, Y.Q., Rutherford, J.C., Carlson, W.T., 2002. Effect of cattle treading on interrill erosion from hill pasture: modelling concepts and analysis of rainfall simulator data. Aust. J. Soil Res. 40, 963–976.

Falkenmark, M., Rockström, J., 2006. The new blue and green water paradigm: breaking new ground for water resources planning and management. J. Water Resour. Plann. Manag. 132, 129–132.

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., Tarboton, D., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. J. Hydrol. 537, 45–60.

Fewtrell, L., Kay, D., 2015. Recreational water and infection: a review of recent findings. Current Environmental Health Reports 2 (1), 85–94.

Geldreich, E.E., 1996. Pathogenic agents in freshwater resources. Hydrol. Process. 10, 315–333.

Geris, J., Tetzlaff, D., McDonnell, J., Anderson, J., Paton, G., Soulsby, C., 2015. Ecohydrological separation in wet, low energy northern environments? A preliminary assessment using different soil water extraction techniques. Hydrol. Process. 29, 5139–5152.

Goldsmith, G.R., Munoz-Villers, L.E., Holwerda, F., McDonnell, J.J., Asbjornsen, H., Dawson, T.E., 2012. Stable isotopes reveal linkages among ecohydrological processes in a seasonally dry tropical montane cloud forest. Ecohydrology 5, 779–790.

Greene, S., Johnes, P.J., Bloomfield, J.P., Reaney, S.M., Lawley, R., Elkhatib, Y., Freer, J., Odoni, N., Macleod, C.J.A., Percy, B., 2015. A geospatial framework to support integrated biogeochemical modelling in the United Kingdom. Environ. Model. Software 68, 219–232.

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: a review and first update. Ecol. Model. 221, 2760–2768.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R.A., Vabø, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual and agent-based models. Ecol. Model. 198, 115–126.

Haydon, S., Deletic, A., 2006. Development of a coupled pathogen-hydrologic catchment model. J. Hydrol. 328, 467–480.

Himathongkham, S., Bahari, S., Riemann, H., Cliver, D., 1999. Survival of *Escherichia coli* O157:H7 and Salmonella typhimurium in cow manure and cow manure slurry. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett. 178, 251–257.

Hipsey, M.R., Antenucci, J.P., Brookes, J.D., 2008. A generic, process-based model of microbial pollution in aquatic systems. Water Resour. Res. 44, W07408.

Hodge, R.A., Hoey, T.B., 2012. Upscaling from grain-scale processes to alluviation in bedrock channels using a cellular automaton model. J. Geophys. Res. 117, F01017.

Hodgson, C.J., Bulmer, N., Chadwick, D.R., Oliver, D.M., Heathwaite, A.L., Fish, R.D., Winter, M., 2009. Establishing relative release kinetics of faecal indicator organisms from different faecal matrices. Lett. Appl. Microbiol. 49, 124–130.

Hopp, L., Glaser, B., Klaus, J., Schramm, T., 2020. The relevance of preferential flow in catchment scale simulations: calibrating a 3D dual-permeability model using DREAM. Hydrol. Process. 34 (5), 1237–1254.

Jackisch, C., Zehe, E., 2018. Ecohydrological particle model based on representative domains. Hydrol. Earth Syst. Sci. 22, 3639–3662.

Jamieson, R., Joy, D.M., Lee, H., Kostaschuk, R., Gordon, R., 2005. Transport and deposition of sediment-associated *Escherichia coli* in natural streams. Water Res. 39, 2665–2675.

Karr, J.R., Schlosser, I.J., 1978. Water resources and the land-surface interface. Science 201, 229–234.

Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K., Bierkens, M.F.P., 2010. A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. Environ. Model. Software 25, 489–502.

Kay, D., McDonald, A., 1980. Reduction of coliform bacteria in two upland reservoirs: the significance of distance decay relationships. Water Res. 14, 305–318.

Kay, D., Crowther, J., Stapleton, C.M., Wyer, M.D., Fewtrell, L., Anthony, S., Bradford, M., Edwards, A., Francis, C.A., Hopkins, M., Kay, C., McDonald, A.T., Watkins, J., Wilkinson, J., 2008. Faecal indicator organism concentrations and catchment export coefficients in the UK. Water Res. 42, 2649–2661.

Kay, D., Aitken, M., Crowther, J., Dickson, I., Edwards, A.C., Francis, C., Hopkins, M., Jeffrey, W., Kay, C., McDonald, A.T., McDonald, D., Stapleton, C.M., Watkins, J., Wilkinson, J., Wyer, M.D., 2007. Reducing fluxes of faecal indicator compliance parameters to bathing waters from diffuse agricultural sources: the Brighouse Bay study, Scotland. Environ. Pollut. 147, 138–149.

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

Knighton, J., Saia, S.M., Morris, C.K., Archibald, J.A., Todd Walter, M., 2017. Ecohydrological considerations for modeling of stable water isotopes in a small intermittent watershed. Hydrol. Process. 31, 2438–2452.
Knighton, J., Kuppel, S., Smith, A., Spenger, M., Soulsby, C., Tetzlaff, D., 2020. Using isotopes to incorporate tree water storage and mixing dynamics into a distributed hydrologic modeling framework. Ecohydrology 13, e2201.
Kuppel, S., Tetzlaff, D., Maneta, M.P., Soulsby, C., 2018a. EcH2O-iso 1.0: water isotopes and age tracking in a process-based, distributed ecohydrological model. Geosci. Model Dev. (GMD) 11, 3045–3069.
Kuppel, S., Tetzlaff, D., Maneta, M.P., Soulsby, C., 2018b. What can we learn from multi-criteria calibration of a process-based ecohydrological model? Environ. Model. Software 101, 301–316.
Macal, C.M., North, M.J., 2010. Tutorial on agent-based modelling and simulation. J. Simulat. 4 (3), 151–162.
Mancini, J.L., 1978. Numerical estimates of coliform mortality rates under various conditions. J. Water Pollut. Control Fed. 50, 2477–2484.
Maneta, M.P., Silverman, N.L., 2013. A spatially distributed model to simulate water, energy, and vegetation dynamics using information from regional climate models. Earth Interact. 17, 1–44.
McDonnell, J.J., Beven, K., 2014. Debates – the future of hydrological sciences: a (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. Water Resour. Res. 50, 5342–5350.
McKergow, L.A., Davies-Colley, R.J., 2010. Stormflow dynamics and loads of *Escherichia coli* in a large mixed land use catchment. Hydrol. Process. 24, 276–289.
Moriarty, E.M., Mackenzie, M.L., Karki, N., Sinton, L.W., 2011. Survival of Escherichia coli, enterococci and Campylobacter spp. in sheep feces on pastures. Appl. Environ. Microbiol. 77, 1797–1803.
Muirhead, R., 2009. Soil and faecal material reservoirs of *Escherichia coli* in a grazed pasture. N. Z. J. Agric. Res. 52, 1–8.
Nagels, J.W., Davies-Colley, R.J., Donnison, A.M., Muirhead, R.W., 2002. Faecal contamination over flood events in a pastoral agricultural stream in New Zealand. Water Sci. Technol. 45, 45–52.
Natural England, 2013. Higher Level Stewardship: Environmental Stewardship Handbook, fourth ed. http://publications.naturalengland.org.uk/file/2819648.
Neill, A.J., Tetzlaff, D., Strachan, N.J.C., 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. J. Hydrol. 570, 423–435.
Neill, A.J., Tetzlaff, D., Strachan, N.J.C., Hough, R.L., Avery, L.M., Maneta, M.P., Soulsby, C., 2020. 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. J. Environ. Manag. https://doi.org/10.1016/j.jenvman.2020.110905.
O'Callaghan, J.F., Mark, D.A., 1984. The extraction of drainage networks from digital elevation data. Comput. Vis. Graph Image Process 28, 323–344.
O'Sullivan, D., Millington, J., Perry, G., Wainwright, J., 2012. Agent-based models – because they're worth it? In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (Eds.), Agent-Based Models of Geographical Systems. Springer, Dordrecht, Heidelberg, London, New York, pp. 109–123.
Oliver, D.M., Clegg, C.D., Haygarth, P.M., Heathwaite, A.L., 2005a. Assessing the potential for pathogen transfer from grassland soils to surface waters. Adv. Agron. 85, 125–180.
Oliver, D.M., Heathwaite, L., Haygarth, P.M., Clegg, C.D., 2005b. Transfer of Escherichia coli to water from drained and undrained grassland after grazing. J. Environ. Qual. 34, 918–925.
Oliver, D.M., Heathwaite, A.L., Hodgson, C.J., Chadwick, D.R., 2007. Mitigation and current management attempts to limit pathogen survival and movement within farmed grassland. Adv. Agron. 93, 95–152.
Oliver, D.M., Page, T., Heathwaite, A.L., Haygarth, P.M., 2010. Re-shaping models of *E. coli* population dynamics in livestock faeces: increased bacterial risk to humans? Environ. Int. 36, 1–7.
Oliver, D.M., Bartie, P.J., Heathwaite, A.L., Reaney, S.M., Parnell, J.A.P., Quilliam, R.S., 2018. A catchment-scale model to predict spatial and temporal burden of *E. coli* on pasture from grazing livestock. Sci. Total Environ. 616–617, 678–687.
Oliver, D.M., Fish, R.D., Hodgson, C.J., Heathwaite, A.L., Chadwick, D.R., Winter, M., 2009. A cross-disciplinary toolkit to assess the risk of faecal indicator loss from grassland farm systems to surface waters. Agric. Ecosyst. Environ. 129, 401–412.
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., Carbonneau, P., McGrane, S.J., Quilliam, R.S., 2016. Predicting microbial water quality with models: over-arching questions for managing risk in agricultural catchments. Sci. Total Environ. 544, 39–47.
Pachepsky, Y.A., Shelton, D.R., 2011. *Escherichia coli* and fecal coliforms in freshwater and estuarine sediments. Crit. Rev. Environ. Sci. Technol. 41 (12), 1067–1110.
Pandey, P.K., Soupir, M.L., Rehmann, C.R., 2012. A model for predicting resuspension of *Escherichia coli* from streambed sediments. Water Res. 46, 115–126.
Parry, H.R., Bithell, M., 2012. Large scale agent-based modelling: a review and guidelines for model scaling: 271-308. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (Eds.), Agent-Based Models of Geographical Systems. Springer, Dordrecht, Heidelberg, London, New York.
Piovano, T.I., Tetzlaff, D., Ala-aho, P., Buttle, J., Mitchell, C.P.J., Soulsby, C., 2018. Testing a spatially distributed tracer-aided runoff model in a snow-influenced catchment: effects of multicriteria calibration on streamwater ages. Hydrol. Process. 30, 3089–3107.
Piovano, T.I., Tetzlaff, D., Carey, S.K., Shatilla, N.J., Smith, A., Soulsby, C., 2019. Spatially distributed tracer-aided runoff modelling dynamics of storage and water ages in a permafrost-influenced catchment. Hydrol. Earth Syst. Sci. 23, 2507–2523.
Ravazzani, G., Rametta, D., Mancini, M., 2011. Macroscopic cellular automata for groundwater modelling: a first approach. Environ. Model. Software 26, 634–643.
Reaney, S.M., 2008. The use of agent based modelling techniques in hydrology: determining the spatial and temporal origin of channel flow in semi-arid catchments. Earth Surf. Process. Landforms 33, 317–327.
Remondi, F., Kirchner, J.W., Burlando, P., Fatichi, S., 2018. Water flux tracking with a distributed hydrological model to quantify controls on spatio-temporal variability of transit time distributions. Water Resour. Res. 54, 3081–3099.
Rode, M., Arhonditsis, G., Balin, D., Kebede, T., Krysanova, V., van Griensven, A., van der Zee, S.E.A.T.M., 2010. New challenges in integrated water quality modelling. Hydrol. Process. 24, 3447–3461.
Sadeghi, A., Arnold, J., 2002. A SWAT/microbial sub-model for predicting pathogen loadings in surface and groundwater at watershed and basin scales. In: Total Maximum Daily Load (TMDL): Environmental Regulations, Proceedings of 2002 Conference. American Society of Agricultural and Biological Engineers, p. 56.
Schijven, J., Derx, J., de Roda Husman, A.M., Blaschke, A.P., Farnleitner, A.H., 2015. QMRAcatch: microbial quality simulation of water resources including infection risk assessment. J. Environ. Qual. 44, 1491–1502.
Shao, Q., Weatherly, D., Huang, L., Baumgartl, T., 2015. RunCA: a cellular automata model for simulating surface runoff at different scales. J. Hydrol. 529, 816–829.
Sheath, G.W., Carlson, W.T., 1998. Impact of cattle treading on hill land: 1. Soil damage patterns and pasture status. N. Z. J. Agric. Res. 41 (2), 271–278.
Smith, A.A., Tetzlaff, D., Laudon, H., Maneta, M., Soulsby, C., 2019. Assessing the influence of soil freeze-thaw cycles on catchment water storage-flux-age interactions using a tracer-aided ecohydrological model. Hydrol. Earth Syst. Sci. 23, 3319–3334.
Sokolova, E., Lindström, G., Pers, C., Strömqvist, J., Lewerin, S.S., Wahlström, H., Sörén, K., 2018. Water quality modelling: microbial risks associated with manure on pasture and arable land. J. Water Health 16 (4), 549–561.
Soupir, M.L., Mostaghimi, S., Lou, J., 2008. Die-off of *E. coli* and Enterococci in dairy cowpats. Transactions of the ASABE 51 (6), 1987–1996.
Sprenger, M., Tetzlaff, D., Buttle, J., Laudon, H., Leistert, H., Mitchell, C.P.J., Snelgrove, J., Weiler, M., Soulsby, C., 2018. Measuring and modeling stable isotopes of mobile and bulk soil water. Vadose Zone J. 17, 170149.
Stocker, M.D., Pachepsky, Y.A., Hill, R.L., Shelton, D.R., 2015. Depth-dependent survival of *Escherichia coli* and Enterococci in soil after manure application and simulated rainfall. Appl. Environ. Microbiol. 81, 4801–4808.
Stumpp, C., Maloszewski, P., 2010. Quantification of preferential flow and flow heterogeneities in an unsaturated soil planted with different crops using the environmental isotope $\delta^{18}O$. J. Hydrol. 394, 407–415.
Tetzlaff, D., Capell, R., 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. Sci. Total Environ. 434, 110–122.
Thomann, R.V., Mueller, J.A., 1987. Principles of Surface Water Quality Modeling and Control. Harper and Row, New York.
Vaché, K.B., McDonnell, J.J., 2006. A process-based rejectionist framework for evaluating catchment runoff model structure. Water Resour. Res. 42, W02409.
van Huijgevoort, M.H.J., Tetzlaff, D., Sutanudjaja, E.H., Soulsby, C., 2016. Using high resolution tracer data to constrain water storage, flux and age estimates in a spatially distributed rainfall-runoff model. Hydrol. Process. 30, 4761–4778.
Vinten, A., Sample, J., Ibiyemi, A., Abdul-Salam, Y., Stutter, M., 2017. A tool for cost-effectiveness analysis of field scale sediment-bound phosphorus mitigation measures and application to analysis of spatial and temporal targeting in the Lunan Water catchment, Scotland. Sci. Total Environ. 586, 631–641.
Vinten, A.J.A., Sym, G., Avdic, K., Crawford, C., Duncan, A., Merrilees, D.W., 2008. Faecal indicator pollution from a dairy farm in Ayrshire, Scotland: source apportionment, risk assessment and potential of mitigation measures. Water Res. 42, 997–1012.
Vivoni, E.R., 2012. Diagnosing seasonal vegetation impacts on evapotranspiration and its partitioning at the catchment scale during SMEX04–NAME. J. Hydrometeorol. 13 (5), 1631–1638.
Vogel, H.-J., Weller, U., Ippisch, O., 2010. Non-equilibrium in soil hydraulic modelling. J. Hydrol. 393, 20–28.
Welch, D., 1982. Dung properties and defecation characteristics in some Scottish herbivores, with and evaluation of the dung-volume method of assess occupance. Acta Theriol. 27, 191–212.
Wellen, C., Kamran-Disfani, A.R., Arhonditsis, G.B., 2015. Evaluation of the current state of distributed watershed nutrient water quality modeling. Environ. Sci. Technol. 49, 3278–3290.
Whitehead, P.G., Leckie, H., Rankinen, K., Butterfield, D., Futter, M.N., Bussi, G., 2016. An INCA model for pathogens in rivers and catchments: model structure, sensitivity analysis and application to the River Thames catchment, UK. Sci. Total Environ. 572, 1601–1610.
Zhang, W., Montgomery, D.R., 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations. Water Resour. Res. 30, 1019–1028.