Water Resources Research

RESEARCH ARTICLE
10.1029/2025WR041145

Using Environmental Tracers to Reduce Uncertainty in Natural Flood Management Modeling

Sarah L. Collins[1], **Leo Peskett**[2], **Andrew R. Black**[3], **Christopher R. Jackson**[4], **Andy Young**[3], and **Alan M. MacDonald**[1]

[1]British Geological Survey, Edinburgh, UK, [2]Heriot-Watt University, Edinburgh, UK, [3]University of Dundee, Dundee, UK, [4]British Geological Survey, Keyworth, UK

Key Points:

- The environmental tracer ANC reduced predictive uncertainty in modeled peak flows when applied to a woodland planting scenario by up to 39%
- The tracer also showed a high potential for reducing model uncertainty at low flows (Q60) and below
- The choice of how to represent woodland was the dominant source of uncertainty (>50%) in predictive models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. M. MacDonald,
amm@bgs.ac.uk

Citation:

Collins, S. L., Peskett, L., Black, A. R., Jackson, C. R., Young, A., & MacDonald, A. M. (2026). Using environmental tracers to reduce uncertainty in natural flood management modeling. *Water Resources Research*, *62*, e2025WR041145. https://doi.org/10.1029/2025WR041145

Received 23 MAY 2025
Accepted 31 MAR 2026

Author Contributions:

Conceptualization: Sarah L. Collins, Andy Young, Alan M. MacDonald
Data curation: Leo Peskett, Andrew R. Black
Formal analysis: Sarah L. Collins
Funding acquisition: Sarah L. Collins, Alan M. MacDonald
Investigation: Leo Peskett, Andrew R. Black
Software: Sarah L. Collins, Christopher R. Jackson
Writing – original draft: Sarah L. Collins
Writing – review & editing: Leo Peskett, Andrew R. Black, Christopher R. Jackson, Andy Young, Alan M. MacDonald

Abstract

Natural flood management (NFM) is a nature-based solution that has grown in importance within flood risk policy and management over the last two decades. There is limited evidence on nature-based solutions' effectiveness, and no accepted best practice on forecasting their performance. To explore NFM effectiveness, we built a hydrological model of a catchment in the UK uplands typical of areas targeted globally for NFM interventions. The model was calibrated on streamflow and groundwater contribution to streamflow, estimated from alkalinity data (ANC). We demonstrated this simple tracer can be a useful tool in model calibration, highlighting significant differences in performance between model runs that were hidden when analysing streamflow alone. In particular the use of the tracer helped identify models that better represented partitioning of flow between surface and sub-surface. The tracer reduced predictive uncertainty in peak flows when applied to a woodland planting scenario by upto 39% and showed even greater potential for reducing uncertainty (~50%) at low flows (below Q60). Further, by exploring three common representations of woodland we showed that the dominant remaining source of uncertainty (>50%) within the scenario modeling was the choice of how to represent the woodland planting on model parameters. This work underlines the value of using additional calibration data sets to improve process representation and prediction; the importance of long-term monitoring for improving the evidence for NFM effectiveness; and the need to further develop the representation of woodland planting in catchment models to improve forecasts of their impact on flow.

Plain Language Summary

Over the last two decades there has been increasing interest in using and enhancing natural landscapes to slow water flow, store water and reduce the risk of flooding. We built a hydrological model of a catchment in the UK uplands typical of those that have been targeted for such enhancements. Whereas hydrological models are usually calibrated on streamflow alone, we also calibrated the model based on percent groundwater contribution to streamflow, which was estimated from an environmental tracer. The environmental tracer proved to be a useful additional tool in model calibration and was found to reduce uncertainty in model predictions of how woodland planting would impact peak flows. Our results suggest the environmental tracer would also be beneficial when assessing the impact of landscape changes on low flows and droughts. Using these improved models, we found, that the primary source of uncertainty when modeling the impact of woodland planting on peak flows is likely to be the assumptions modelers make about how woodland planting affects catchment processes. This work highlights the importance of long-term monitoring and process understanding of nature-based schemes that seek to reduce flooding through landscape changes, as well as the increased value in introducing additional calibration data sets.

1. Introduction

Globally, unprecedented events are increasing the impact of flooding, despite improvements in risk management (IPCC, 2022; Kreibich et al., 2022). This is leading to growing concern that traditional "hard engineering" approaches to flood risk management are becoming ineffective or simply too expensive in the face of climate change (Vicarelli et al., 2024). Nature-based solutions, which work with natural processes to mitigate risk, are being increasingly promoted as an alternative or complementary approach for managing flood risk (Cohen-Shacham et al., 2016; Faivre et al., 2017; Maes & Jacobs, 2017; World Bank, 2018). In the UK, this has given rise to the field of natural flood management (NFM), which has been incorporated into policy over the last two decades (Flood and Water Management Act, 2010; Flood Risk Management (Scotland) Act, 2009). NFM is expected to not only reduce flood hazard, but also provide a variety of co-benefits, such as enhancements in biodiversity, carbon sequestration, and soil and water quality (Dadson et al., 2017).

At the heart of NFM is a catchment-wide approach that incorporates in-channel measures to slow river flow (e.g., woody debris dams, storage ponds, re-meandering) and land use or management measures, which seek to reduce runoff generation (e.g., woodland planting, livestock management, crop choice and rotation, and soil management) (CIRIA, 2022; SEPA, 2015). Woodland creation is popular among NFM schemes, as it provides a wide range of environmental and leisure benefits. Woodland alters the hydrology of a catchment in several ways, including increasing evapotranspiration, particularly wet-canopy evaporation, increasing infiltration into the soil, enhancing soil storage and inhibiting rapid overland flow (Burgess-Gamble et al., 2017; DeWalle, 2011; Martin et al., 2020).

The widespread adoption of nature-based solutions is being hampered by a lack of reliable comparative evaluations of effectiveness (Vörösmarty et al., 2021). For example, despite the development of approximately 100 NFM projects across the UK, catchment wide monitoring is usually limited and therefore efficacy is rarely quantified (Dadson et al., 2017; Kay et al., 2019). Studies are generally limited to plot and field scale experiments. Experiments have shown that soil permeability can be enhanced underneath some woodland but that woodland type and management are clearly important factors (Archer et al., 2013, 2016; Brickell et al., 2024; Chandler et al., 2018; Kingsbury-Smith et al., 2023; Monger et al., 2022; Murphy et al., 2021). Friction to overland flow is more difficult to measure, but Monger et al. (2022) were able to compare overland flow rates underneath different types of woodland and found that bracken in the understory significantly slowed flow. Whether these observed process changes at the field scale sum up to significant reductions in river flow during extreme events, however, remains an active topic of debate.

Due to the lack of catchment wide evidence, or even consistent evidence from plot scales, numerical modeling is heavily relied upon to provide evidence of effectiveness of NFM measures, particularly tree planting (Kay et al., 2019). However, in a review of the few observational studies that exist and modeling studies Stratford et al. (2017) found an even split between studies that found a positive impact from tree cover on flood peaks and those that found the opposite or no influence When only modeling studies were considered they tend to support the hypothesis that increasing tree cover reduces flooding (Stratford et al., 2017). This suggests a possible inherent bias in numerical modeling studies toward over-emphasizing the impact of tree planting, with wide reaching implications for emerging land use and flooding policy.

Conceptual lumped-parameter hydrological models are often applied in flood modeling owing to their simplicity and the ease with which a good calibration to observed events can often be obtained. It is, however, not straightforward to translate the effects of woodland planting into effective model parameters, and authors rely on speculative shifts in parameters representing routing or soil permeability (e.g., Ferguson & Fenner, 2020; Packman et al., 2004; Rose & Rosolova, 2007). Other authors have studied these "parameter shifts" for well-monitored micro-catchments before applying them at a wider basin scale (Goudarzi et al., 2021; Hankin et al., 2021; Kingsbury-Smith et al., 2023). However, there is a fundamental problem in scaling these shifts up to larger basins, given they are essentially an empirical fitting of equations that are not physically based, and it is yet more difficult to justify applying these results to other catchments, particularly where geology and climate vary. Process-based models, however, allow inclusion of woodland planting as a series of physical changes in measurable soil and vegetation properties (e.g., Buechel et al., 2022; Collins et al., 2023; Iacob et al., 2017). The limitation of this approach is that these models are data hungry and the soil parameters, in particular, may not be available. Moreover, while these process-based models provide a very detailed simulation of the 1D interactions between soil, water and vegetation, they often do not include routing through the catchment (e.g., Beuchel et al., 2022) and are therefore often inappropriate for flood studies.

If we are aiming to prevent floods by altering catchment processes, then it seems logical that we should develop a better understanding of where water is being routed and how to represent this in models. In other words, we should have confidence that our models are getting the "right answers for the right reasons" (Kirchner, 2006). Even where streamflow is well simulated at the catchment outlet, there can remain significant uncertainty in the magnitude and spatial patterns of internal model fluxes (Beven, 2006; Cao et al., 2006). Analyzing tracer data alongside streamflow data can provide new insights into catchment processes. Although tracer techniques are promoted by the hydrological community (e.g., Kirchner et al., 2023; McDonnell & Beven, 2014) and have been available since the 1960s, they are still rarely applied in practice or incorporated within models.

The largest insight gained through tracers to date has been the dominance of pre-event water in storm runoff, even at peak flow (Klaus & McDonnell, 2013), which has since led to questions about celerity (discharge) versus velocity (flow paths and connectivity) in catchment response (Bracken et al., 2013; McDonnell & Beven, 2014). The incorporation of tracers in model evaluation is a relatively recent development, with the use of stable isotopes, in particular, being boosted by improvements in reliability and cost of analysis. There are now several studies that show the utility of tracers in multi-criteria calibration, where they have been shown to constrain parameter sets and reduce equifinality (e.g., Ala-aho et al., 2017; Birkel et al., 2014; Holmes et al., 2020, 2022; Stevenson et al., 2021; Tunaley et al., 2017; Wu et al., 2023). However, having additional data does not necessarily make the task of the modeler easier, as trade-offs in performance among different data may be required (e.g., Fenicia et al., 2008; Scudeler et al., 2016) or errors in model structure or process conceptualization may be unearthed (McDonnell et al., 2007).

Within an NFM context, tracers have been used in two observational studies in a UK upland catchment. In the first, Peskett et al. (2023) used tracers to show that any effects of forest cover on increasing infiltration and storage in a UK upland catchment were masked by variations in soil hydraulic properties, even in small, flashy sub-catchments. In the second, Peskett et al. (2021) showed that plantation forest cover reduced the fraction of rapid rainfall runoff, but that soils and geology were the main control. To the authors' knowledge the only example of tracers being used in the modeling of nature-based solutions is Neill et al. (2021), who used a tracer-aided ecohydrological model to study the effects of rewilding on landscape restoration.

In this study, and for the first time in NFM modeling, we explore: how an environmental tracer can constrain uncertainty through a multi-criteria calibration; and how the representation of NFM in models impacts the assessment of effectiveness. To do this, we apply the widely used Dynamic TOPMODEL (Beven & Freer, 2001)—which is one of the most frequently used hydrological models in NFM studies (Hill et al., 2023)—to a well studied catchment in the Scottish Southern Uplands. The Eddleston Research Catchment is representative of many temperate upland areas combining grassland, forest and small scale agriculture (Spray et al., 2022). We derive groundwater contribution to streamflow from alkalinity data and combine it with streamflow to conduct a multi-criteria calibration. Using the typical methodology in the literature (Hill et al., 2023), we then apply an afforestation scenario to the calibrated model by altering the soil transmissivity. In doing so, we aim to answer the following questions:

1. Does the tracer provide additional information during model calibration that we do not get from evaluating flow alone?
2. How does our choice of speculative parameter shift from the literature affect our scenario predictions and assessment of effectiveness?
3. Can the tracer reduce predictive uncertainty in the context of NFM?

By answering these questions we discuss how predictions of the impact of natural flood management interventions such as tree-planting can be improved, and the wider implications of reducing uncertainty in numerical hydrological modeling by adapting distributed models to include better consideration of surface and sub surface routing, calibrated with tracers.

2. Methods

2.1. Study Area

The Eddleston Water, a tributary of the River Tweed, is a 69 km^2 catchment in the Scottish Southern Uplands, UK (Figure 1). It is the site of one of the largest and longest-running NFM studies in the UK. As such, a wide variety of NFM measures have been implemented and monitored in the catchment, from storage ponds and woody debris dams to river re-meandering and native woodland planting (e.g., Costaz-Puyou, 2022; Peskett et al., 2025; Spray et al., 2022).

The Eddleston Water is typical of upland catchments within the UK. Topographic elevation varies between 160 and 600 m asl (Figure 1a). The main river valley and lower hillslopes are dominated by improved grassland and, on higher ground, acid grassland is found in the east and coniferous plantation in the west, large areas of which were felled during the study period (Figure 1b). Mean annual precipitation (2012–2022) is ∼1,080 mm, which predominately falls as rainfall; monthly mean air temperatures are 3°C–13°C; and actual daily evapotranspiration increases from 0.2 mm in winter to 2.5 mm in summer (Peskett et al., 2023).

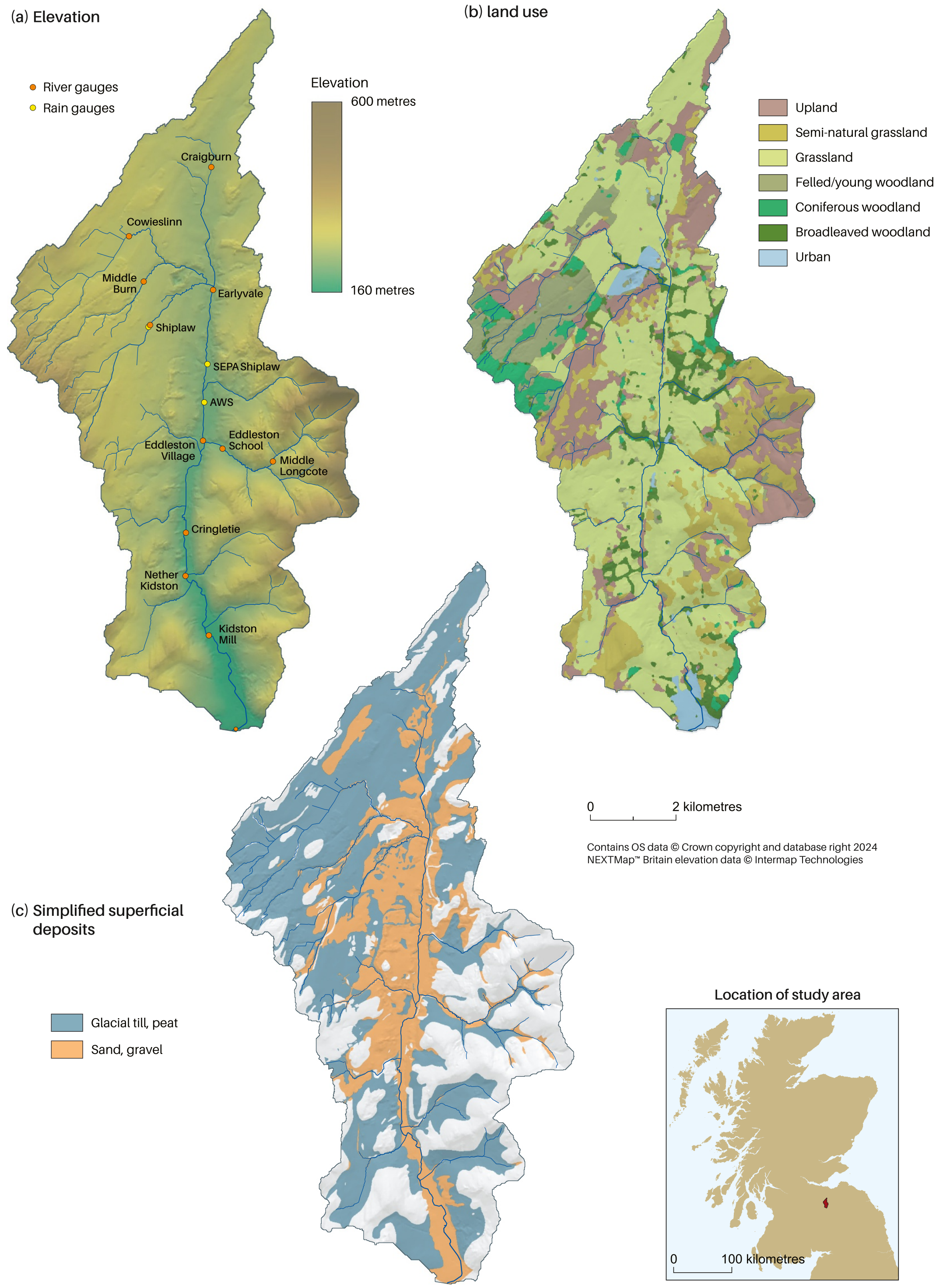

Figure 1. Eddleston Water location map showing (a) topography, gauge locations and river network. (b) Land use simplified and modified from NatureScot (2019) licensed under Open Government License v.3.0. (c) Superficial geology simplified from Auton (2011) BGS © UKRI 2011. AWS, automatic weather station.

There is a strong association between soils and geology in the Eddleston catchment. Overlying the poorly permeable glacial till of the western sub-catchments (Figure 1c), there are extensive areas of poorly permeable gley soils and peats. The subsurface thus provides little storage, resulting in flashy streamflows with rapid surface runoff (Black et al., 2021; Peskett et al., 2021, 2023). Through the central valley of the catchment, however, more freely draining brown soils are found, which allow recharge into the sands and gravels of the lower hillslopes and floodplain. The east is dominated by brown earth soils with some peaty and mineral podzols on hilltops and non-calcareous gley soils along the tributaries. Superficial deposits are largely absent in the east with soils overlying the low-permeability bedrock, Silurian greywacke (Ó Dochartaigh et al., 2015; Graham et al., 2009). Contrary to expectation, streamflows and hydrochemistry data from the eastern subcatchments indicate relatively high water storage and groundwater fraction (48%), suggesting the bedrock is weathered and fractured and allows significant groundwater flow (Peskett et al., 2023). The same behavior has been observed during storm events, for which peak discharge is lower and recession rates slower in the east of the catchment (Black et al., 2021; Peskett et al., 2021).

2.2. Hydrological Model

In this section, we describe the data inputs, set up and streamflow calibration of the hydrological model, including modifications that were made to the model. The second stage of the calibration based on groundwater fraction is described in the following Section 2.3.

2.2.1. Meteorological and Streamflow Data

Rainfall and meteorological variables for the calculation of evapotranspiration (ET) were obtained from an automatic weather station in the center of the catchment (Figure 1a). Gaps in the rainfall data were infilled with data from the two nearest rain gauges (Shiplaw Burn and SEPA Shiplaw, Figure 1a). The catchment has five rain gauge sites relatively equally dispersed across the catchment and at differing altitudes. Rainfall is monitored using both storage and Tipping Bucket Rain gauges at each of these sites. Rainfall data have been periodically quality checked, with <10% between storage gauges and TBRs in cumulative totals at each of the sites (Peskett et al., 2020). For comparison between sites, annual rainfall totals were compared—there is some variation between sites but this is not systematic. The standard deviation in annual totals between five rain gauge sites across 5 years of rainfall records including the time period of this study ranged from 2% to 9% of annual rainfall. Plots of rainfall accumulation for the five sites for each event are shown in Supporting Information S1.

We used streamflow data from seven of the gauging stations shown in Figure 1a—Eddleston School and Middle Longcote were excluded owing to high scatter in their rating curves, but were included in the tracer analysis (see Section 2.3) (Figure 1a). The gauges are operated on the velocity–area principle. No weirs were built (a) on cost grounds and (b) in keeping with the environmental ethos of the project. Typically flows are measured at eight calibration check gaugings per year, including flood conditions where practical, in order to define ratings and track changes which may arise from sediment movement and/or seasonal weed growth. Uncertainties in the flow data may be marked during high-flow episodes.

2.2.2. Hydrological Model Setup

The hydrological model DECIPHeR (Coxon et al., 2019), a version of Dynamic TOPMODEL (Beven & Freer, 2001) and referred to herein simply as "Dynamic TOPMODEL," was used to model the hydrology. As in the original Dynamic TOPMODEL, areas with similar characteristics are grouped together as hydrological response units (HRUs). For the Eddleston Water, we used superficial geology, land use and topography to define the HRUs. Superficial geology, which is strongly linked to soils within the catchment, was simplified into three classes: (a) low permeability superficial deposits, consisting primarily of glacial till; (b) medium to high permeability superficial deposits, comprising glaciofluvial and alluvial sands and gravels as well as head deposits; and (c) rock head, where superficial deposits are absent (i.e., white/blank in Figure 1c). Land use was derived from the Scotland Land Cover Map (NatureScot, 2019) and simplified into six categories (Figure 1b): upland, semi-natural grassland, grassland, felled/young woodland, coniferous woodland, broadleaved woodland and urban. Felled and young woodland were combined into a single category because they could not be differentiated on satellite imagery or existed together (i.e., young woodland planted/naturally regenerating among felled woodland). The land use classification was compared to another land cover map (CORINE, 2018) and satellite imagery checked to see whether any plantations had been recently felled.

The hydrological model was run for a period of ~10 years (March 2011–October 2021) at an hourly time step. Rainfall was not applied directly to the model, but instead used to drive an interception model, which produced

Table 1
Dynamic TOPMODEL Model Parameters

Parameter (unit)	Description	Lower bound	Upper bound
SR_{init} (m)	Initial storage in root zone	RAW[a]	RAW[a]
SR_{max} (m)	Maximum storage in root zone	[b]	[b]
SZM (m)	Form of exponential decline in transmissivity with depth	0.001	0.15
T_d (h/m)	Unsaturated zone time delay	0.1	40
CHV (m/h)	Channel routing velocity	100	4,000
Ln (T_0) (ln(m^2/h$^-$))	Lateral saturated transmissivity	−7	7
S_{max} (m)	Maximum effective deficit of saturated zone	0.3	3

[a]Set to the "readily available water" (RAW) content, but the parameter is insensitive owing to 2-year spin up. [b]Set according to land use using the CERF soil moisture accounting scheme (see Table S3 in Supporting Information S1).

throughfall to the land surface. The interception model is based on that of Rutter et al. (1975), which calculates a canopy water balance. We parameterized the model for different land uses according to Dunn and Mackay (1995), who applied the same model in a nearby catchment (River Tyne, NE England). The interception scheme and parameterization are detailed in Text S1, Figure S1, and Table S1 in Supporting Information S1.

In Dynamic TOPMODEL, actual ET is calculated by multiplying potential ET by the fraction of saturation of the root zone. We replaced this scheme with the CERF soil moisture accounting scheme (Griffiths et al., 2006), which explicitly accounts for the soil and vegetation type. The model is based on the widely applied FAO method (FAO, 1998) and is described in detail by Mansour et al. (2018) (see Tables S2 and S3 in Supporting Information S1).

2.2.3. Flow Calibration

The hydrological model was first calibrated based on streamflows using a Monte Carlo approach. Dynamic TOPMODEL includes seven parameters (Table 1), five of which were adjusted during calibration. The initial storage in the root zone was set to the "readily available water" content (i.e., the moisture level below which the vegetation will suffer water stress; FAO, 1998), but is insensitive to this parameter, as the first 2 years of the run were discarded as model spin up. The maximum storage in the root zone was set through the soil moisture accounting scheme (Section 2.2.2) to the difference between field capacity and wilting point (Table S3 in Supporting Information S1). The remaining five parameters were treated as calibration parameters and calibrated within a wide range used previously for UK catchments (Table 1; Coxon et al., 2019). As observational research conducted in the catchment shows that soils and geology provide a much stronger control on flow dynamics than land cover (Peskett et al., 2021, 2023), these five parameters were varied only according to the three classes of superficial geology (see Section 2.2.2), giving a total of 15 model parameters.

Owing to the long run time (8 min), the calibration was split into two stages. (a) 162,000 (pseudo)randomly generated parameter sets (Marsaglia & Zaman, 1987; with modification by James, 1988) were run through the model for the period September 2018–December 2020. The first 2 years were discarded as spin up and the performance of the model was assessed over the period August–December 2020. This acted as an initial "screening" to remove parameter sets that were obviously producing poor streamflow estimates during a wet period containing a number of peak flow events. (b) The parameter sets from stage 1 with a Kling-Gupta (KGE) efficiency >0.6 (13,195 models) were run for the entire period (March 2011–October 2021). In this second stage, the first 2 years were again discarded as spin up and the five largest peak flows over the remaining period analyzed (Table 2) as well as mean monthly flows. Although NFM is widely thought to be most effective for smaller events (Dadson et al., 2017), we used the five largest flows as they were known to have caused some degree of flooding in the settlements along the Eddleston Water. The return periods of these events were previously estimated by others (JBA, 2020), but the uncertainty in these estimates is large because there is insufficient historical data (Table 2). "Acceptable" models were identified as having a Nash Sutcliffe Efficiency (NSE) > 0.8 for all peaks and an NSE > 0.8 for mean monthly flows at Kidston Mill. The NSE—rather than KGE—was used because we were analyzing the performance of the peak, rather than a longer time series Selecting a value of NSE to represent a threshold of model fit is necessarily arbitrary, however a value of 0.8 represents the case that the unexplained

Table 2
Flow Events Used in Model Calibration and Scenario Analysis

Event no.	Date	Peak flow at Kidston Mill (m^3/s)	Estimated return period (JBA, 2020)
1	December 2014	28.9	5
2	December 2015	32.0	5
3	November 2016	36.7	10
4	December 2020	32.9	5
5	October 2021	36.1	10

variation is less than 20% of the total variation. We used only the Kidston Mill gauge for the calibration because it is the experimental catchment outlet (Figure 1a), the location of many of the highest flow calibration gaugings and also the location of the furthest downstream gran alkalinity measurements. We quote NSEs in Section 3 for all gauges except Eddleston School and Middle Longcote, owing to very high scatter on their ratings curves (Figure 1a). We used the generalized likelihood uncertainty estimation (GLUE) (Beven & Binley, 1992, 2014) method to weight each of the acceptable models and define 90% confidence bounds.

2.3. Alkalinity Data and Groundwater Fraction

Alkalinity data in the form of acid neutralizing capacity (ANC) were available across the whole catchment at a 2-weekly resolution over a 2-year period (May 2015–May 2017). The alkalinity data were used to estimate the groundwater fraction in streamflow with a two-component mixing model:

$$F_{\mathrm{gw}} = \frac{A_r - A_S}{A_r - A_{\mathrm{gw}}} = \frac{Q_{\mathrm{gw}}}{Q_t}$$

where F_{gw} is the groundwater fraction, Q_t is streamflow, Q_{gw} is groundwater contribution to streamflow, A_s is the ANC of streamflow, A_r is the ANC of the surface runoff endmember and A_{gw} is the ANC of the groundwater endmember. The groundwater endmember was defined as the mean ANC of the five lowest flows in each sub-catchment for the period September 2015-August 2016 based on weekly sampling conducted at this time; the surface runoff endmember was defined as 0 μeq l^{-1} given that this approximates the ANC in rainfall, and the soils comprise generally acidic brown earths, peaty podzols and non-calcareous gleys. Other endmember definitions for surface runoff were explored as part of a previous study (Peskett et al., 2023), resulting in variations in groundwater fraction estimates up to 25%, but all gave similar relative estimates between catchments. We compared the estimated groundwater fraction with the groundwater fraction simulated by the hydrological model. Splitting the observed streamflow into two distinct components, groundwater and surface water, clearly presents some limitations, as, in reality, flow paths will likely be more complex than this and there are three components to runoff generation in the model. In this version of Dynamic TOPMODEL, overland flow is comprised of precipitation and saturation excess flow (Figure 2). The main limitation is that the saturation excess component of overland flow may well have mixed to some degree with the saturated store. Despite the potential for saturated flow to become overland flow, the saturation excess flow is better described as surface runoff because it is generated in impermeable areas, where the transmissivity is insufficient to allow the water to be routed through the subsurface. Groundwater flow in the model occurs between neighboring HRUs and into the river channel and is controlled by the average saturation deficit across the HRU, the mean slope of the HRU, and an exponentially declining transmissivity with depth (Table 1, Figure 2; Coxon et al., 2019).

In this second phase of the multi-criteria calibration, we extracted the modeled groundwater fraction of streamflow at each gauge throughout the 2-year monitoring period for the acceptable models from the flow calibration (see Section 2.2.3). The root mean squared error (RMSE) between the modeled and estimated groundwater fraction according to alkalinity were calculated. The acceptable models were subjectively selected as those in which the RMSE at the catchment outlet, Kidston Mill, was below 25%. Although high, upon visual inspection, this seemed an appropriate threshold for separating those simulations that were roughly correct from those that were not.

10.1029/2025WR041145

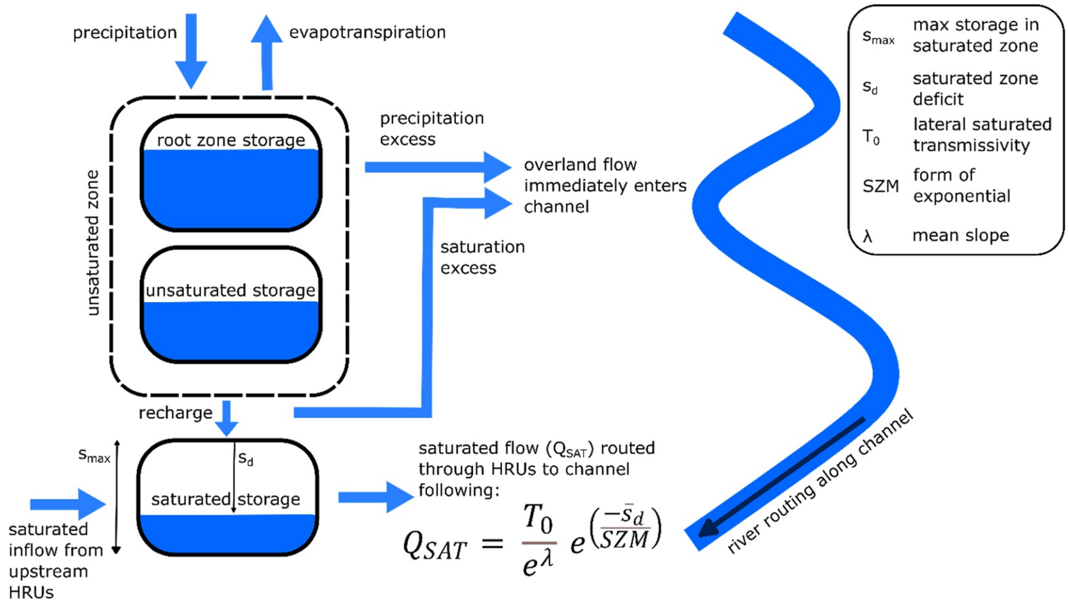

Figure 2. Description of the DECIPHeR version of Dynamic TOPMODEL (Coxon et al., 2019).

2.4. NFM Scenario Parameterization

A recent systematic review (Hill et al., 2023) highlights Dynamic TOPMODEL as one of the most commonly used models in assessing the impact of afforestation or land use change as an NFM intervention. Both interventions are typically represented as changes in shallow transmissivity, overland flow velocity and Manning's *n*, where the model is coupled with a hydraulic model. Here we concentrate on shifts to the parameters controlling transmissivity because, in the version of Dynamic TOPMODEL used, overland flow is not routed along the surface but instead added straight to the channel (Figure 2).

It is not straightforward to translate the effects of woodland planting into effective model parameter shifts in a conceptual hydrological model, and, even where some local observation data are available, these shifts will always be somewhat speculative. Hankin et al. (2016) provide a short review of how woodland affects the permeability of soils. They highlight three studies that show an increase in soil saturated conductivity of 1.81–3.40 times for deciduous trees versus grassland on gley soils and suggest this could be represented by a 1.5–3.5 factor increase in T_0 in Dynamic TOPMODEL. This effect is lower than that recorded in other observational studies on different soil types across the globe (Table 6.6 in Hankin et al., 2016). It is also significantly smaller than the 5−8 factor increase in soil permeability measured by Archer et al. (2013) in the Eddleston catchment between deciduous woodland and improved pasture on well drained brown soils. Brickell et al. (2024) measured saturated hydraulic conductivity of both gley and brown soils in the Eddleston catchment on different land uses and found permeabilities under both conifer plantation and rough grazing to be two orders of magnitude higher than under improved grassland (40−66 factor), which, although high, is not unheard of in the literature (Table 6.6 in Hankin et al., 2016). The large variation between field experiments, sometimes even within the same catchment, highlights the challenge in quantifying any change in soil permeability from woodland in a conceptual model.

In addition to an increase in saturated transmissivity, it is thought that the rate of exponential decline in conductivity (SZM or "*m*" in other versions of TOPMODEL) with depth is less steep underneath woodland. There is some observational evidence for this, including from the Eddleston catchment, where Peskett et al. (2020) attributed a lower water table under a forest strip to enhanced hydraulic conductivity in the soil and subsoils compared with adjacent grassland. There are several examples from the UK in the literature where either SZM and/or saturated transmissivity were enhanced to represent woodland planting. Ferguson and Fenner (2020) used Dynamic TOPMODEL to simulate the impact of woodland planting and used factor increases of 1.5 for T_0 and 1.2 for SZM. Coulthard et al. (2000) and Coulthard and Van De Wiel (2017) used TOPMODEL to simulate the impact of tree planting, changing only the SZM parameter by as much as a factor of 4 between sparse and very

Table 3
Parameter Shift Sets Used to Represent the Impact of Woodland Planting on Soil Permeability

Parameter	Set 1	Set 2	Set 3
Ln (T0)	1.5	3.5	1.0
SZM	1.2	1.7	4.0

dense vegetation. Iacob et al. (2017) used a multiplication factor of 8 for soil hydraulic conductivity to represent the impact of tree planting in Scotland; however, they used the physically based WaSiM-ETH model and thus changed the vertical conductivity rather than horizontal conductivity, which is modified in (Dynamic) TOPMODEL applications. Kingsbury-Smith et al. (2023) represented woodland planting in a slightly different version of Dynamic TOPMODEL with a 1.5 factor increase in SZM and a 0.9 decrease in saturated hydraulic conductivity. Whereas the change in SZM was estimated from the literature, the scaling factor for conductivity was based on local measurements, albeit from soil samples rather than in situ measurements. Finally, Hankin et al. (2021) studied two adjacent micro-catchments with similar gley soils (although with differing superficial geology), one of which was entirely rough grazing and the other one-third mature conifer plantation, with the remaining land use being felled/young woodland or rough grazing. They calibrated the SZM parameter separately for two storm events, finding SZM to be 1.53 and 1.78 times higher for the micro-catchment with woodland. They used SZM to encompass all changes between the two micro-catchments (i.e., no change in T_0 or overland flow velocity). If we accept the assumption that this shift in SZM is due to the land cover, rather than superficial geology, the effect appears quite large, especially as more recent plantations appear to have much lower soil permeability than older, more natural forests (e.g., Archer et al., 2013). A further issue with conifer plantations is that drainage ditches are often constructed to aid rapid tree growth, which we would expect to promote rapid runoff. In the Eddleston catchment, Peskett et al. (2021) detected only small reductions in rapid runoff from conifer plantations during storm events. Based on these previous studies we test three different parameter sets to explore the uncertainty deriving from how these parameter shifts are defined (Table 3). We use a single woodland planting scenario in which the entire catchment, aside from urban areas, is converted to broadleaved woodland.

3. Results and Discussion

The first part of this section describes the calibration of the model and the value of the tracer data (research question 1). In the second part, we explore the sensitivity of the model to how we define the parameter shift due to NFM (research question 2); and we determine to what degree the multi-criteria calibration has reduced the predictive uncertainty of the scenario model (research question 3).

3.1. Model Calibration

3.1.1. Streamflow

Calibration of the hydrological model produced 200 "acceptable" model runs with an NSE > 0.8 for all peaks and mean monthly flows at Kidston Mill. The models perform well at Kidston Mill, with median NSEs > 0.9 for events 3 and 4 and >0.84 for events 1, 2, and 3, which are slightly underestimated (Figure 3). A summary of the model performance at all gauges is given in Table 4 with the caveats given in Section 2.2.1 on the uncertainty in these observed flows. Along the main stem (Figure 1a), the models perform well at Nether Kidston and Eddleston Village, except for event 1 which is overpredicted at Eddleston Village (median NSE −0.56). The most upstream gauge on the main stem (Craigburn, catchment area ~3.6 km^2) is poorly simulated at both peak flows and mean monthly flows (Table 4; Figure 1a). The poor performance of the model in this small subcatchment cannot solely be attributed to uncertainty in observed flows, as the shape of the simulated hydrograph is also incorrect, with the model producing a response that is too flashy. The subcatchment hosts a number of NFM interventions—including river re-meandering and flow restrictors—which have been shown to increase lag times (Black et al., 2021), although these were mostly in place before event 1, which is well simulated. Alternatively, given that the topography of the subcatchment is very flat, the model assumption that all overland flow reaches the channel within a time step (1 hr) may be inappropriate in this part of the catchment (see Figure 2).

Cowieslinn, Shiplaw and Middle Burn are neighboring subcatchments to the west of the Eddleston Water with similar superficial geology, topography and land cover (Figure 1). The model performs significantly better at Shiplaw than at Cowieslinn or Middle Burn (Table 4), with simulated streamflows being consistently lower than observations at Middle Burn and consistently higher than observations at Cowieslinn. The reasons for this are likely manifold: from variations in rainfall (which has been applied uniformly across the catchment) and uncertainty in observed flows to potentially the topographic boundaries not representing closed water balances

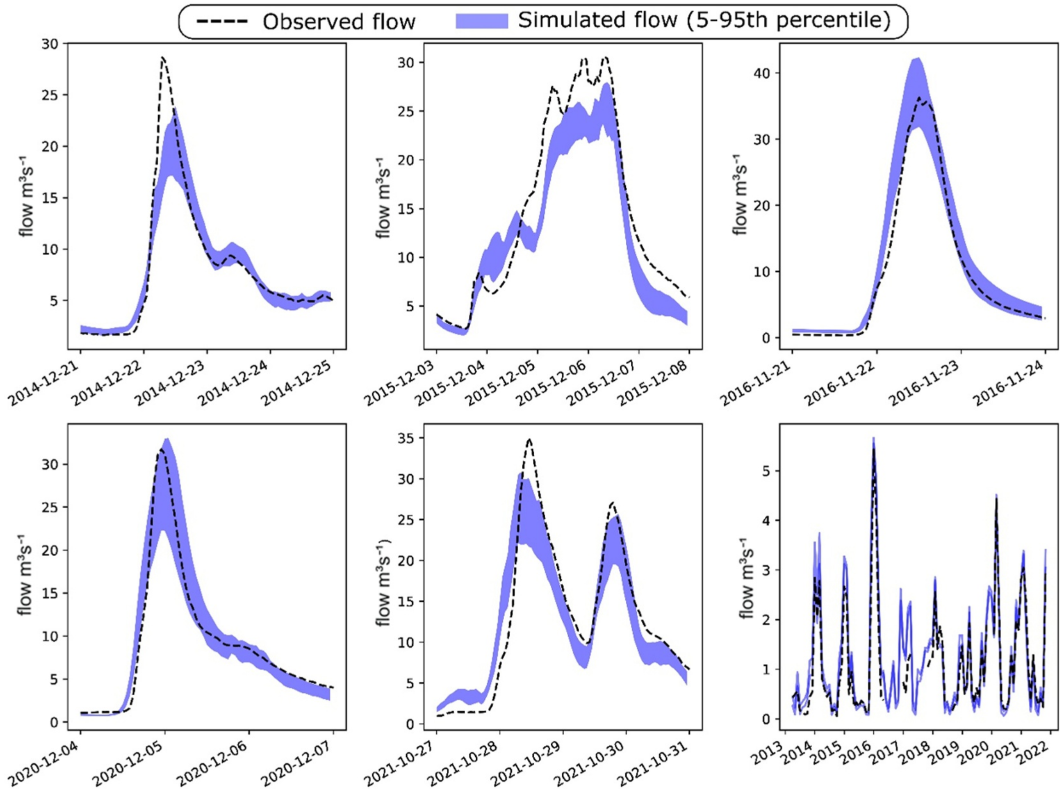

Figure 3. Simulated versus observed flow at Kidston Mill for the five events (Table 2) and mean monthly flows. The *x*-axis shows the time period.

(e.g., due to drains). For example, the observed flows at Middle Burn are also higher than rainfall inputs, which can be seen by calculating the empirical runoff coefficients for the five events (1.3, 2.9, 1.0, 1.4, 1.8). Similar to Craigburn, Middle Burn and Cowieslinn have been modified in recent years, with plantation woodland felling occurring during the modeling period and the construction of NFM interventions, predominately before but also during the modeling period. However, it is difficult to attribute the discrepancies in simulated versus observed peak flows to these modifications, primarily because the discrepancies are too large, but also for other reasons. For example, Middle Burn contains 35 channel flow restrictors, which should slow flow, but simulated flows are far too high; plantation felling took place during the simulation period, but the disparity between modeled and observed streamflow remains consistent for all of the peak flows (i.e., either too high or too low); and, in Cowieslinn, native tree planting is too recent and the capacity of NFM storage ponds too small (Peskett

Table 4
Median Nash Sutcliffe Efficiency (NSE) for the 200 Acceptable Models

		Median NSE					
Gauge	Position in catchment	Event 1	Event 2	Event 3	Event 4	Event 5	Monthly flows
Kidston Mill	Main stem	0.86	0.84	0.95	0.92	0.84	0.89
Nether Kidston		0.78	0.83	0.77	0.89	0.70	0.80
Eddleston Village		−0.56	0.80	0.72	0.83	0.79	0.86
Craigburn		0.68	−2.6	−2.9	−2.7	−3.9	0.57
Shiplaw	Western tributary	0.53	0.68	−0.39	0.53	0.62	0.88
Middle Burn		0.63	0.08	0.50	0.33	0.21	0.86
Cowieslinn		−1.9	−3.3	−6.6	−3.1	−0.26	0.69

Note. Gauge positions shown in Figure 1a.

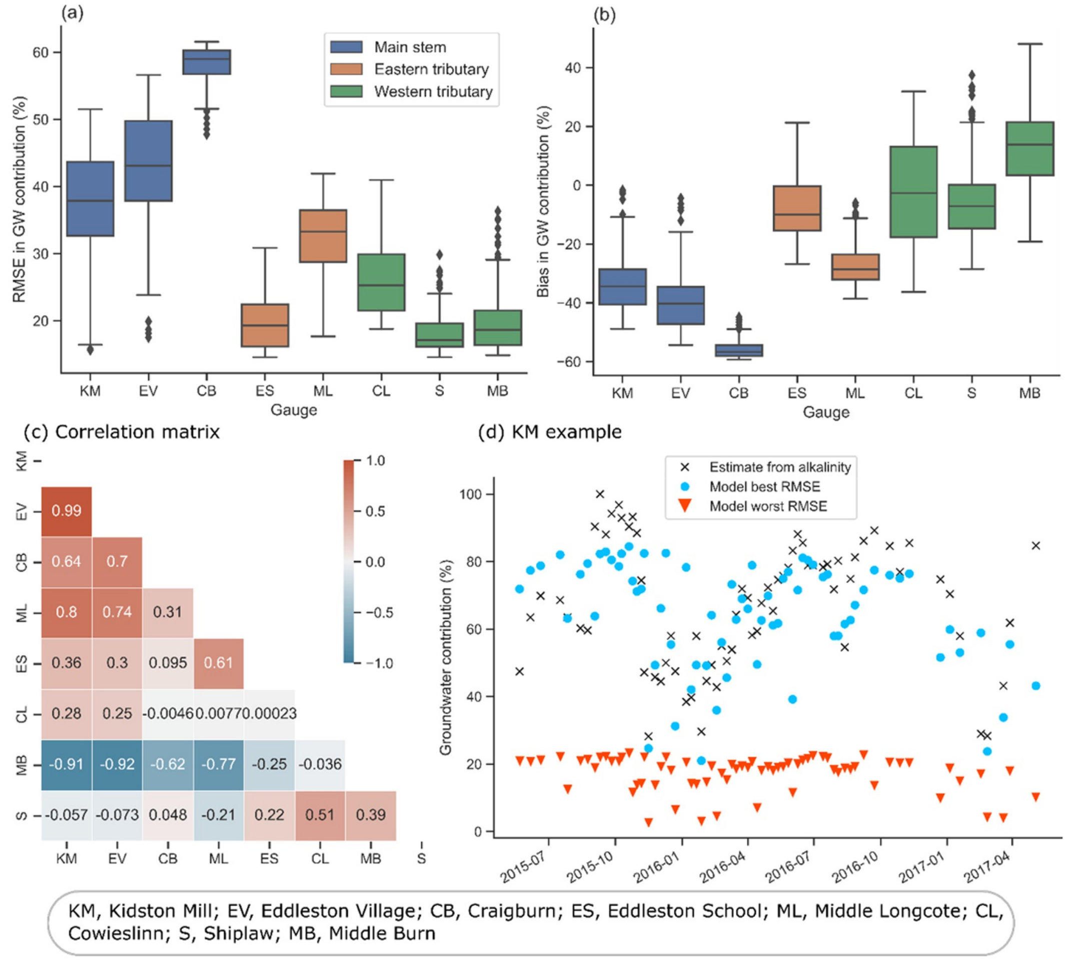

Figure 4. Groundwater representation of best models calibrated with streamflow alone (a) Root mean square error (RMSE) and (b) bias between the groundwater (GW) contribution estimated from alkalinity data and baseflow. (c) Correlation matrix for RMSE in GW contribution between different gauges for the 200 acceptable models. (d) Example plot for Kidston Mill gauge of the best performing and worst performing model with respect to groundwater contribution which were indistinguishable using streamflow calibration.

et al., 2025) to explain the overestimation of the model. Modeled mean monthly flows are reasonably consistent with observations at all three subcatchments (Table 4).

3.1.2. Groundwater Contribution

The groundwater contribution to streamflow estimated from alkalinity data (see Section 2.3) was compared against modeled baseflow for the 200 models deemed acceptable for their streamflow performance alone (Section 3.1.1). Figures 4a and 4b shows RMSE and bias in these models for groundwater contribution. There are models that perform well at each of the gauges except for Craigburn. Although these models produce accurate simulations of streamflow, they differ substantially in terms of predicted groundwater contribution to total flow, particularly for the gauges with the largest catchment areas on the main stem. There is no correlation between streamflow performance (NSE) and groundwater contribution performance (RMSE) ($r = 0.11$ at Kidston Mill), which shows that the alkalinity data are providing additional independent information that can be used to improve model calibration and ultimately provide a better representation of all hydrological budget components. The variation in model performance is illustrated in Figure 4d, which shows time series of groundwater contribution for the best and worst performing models at Kidston Mill, the catchment outlet. It is clear from the figure that, despite these two models meeting the calibration criteria for simulating peak flows and monthly mean flows, the mechanisms for generating streamflow within the model are very different.

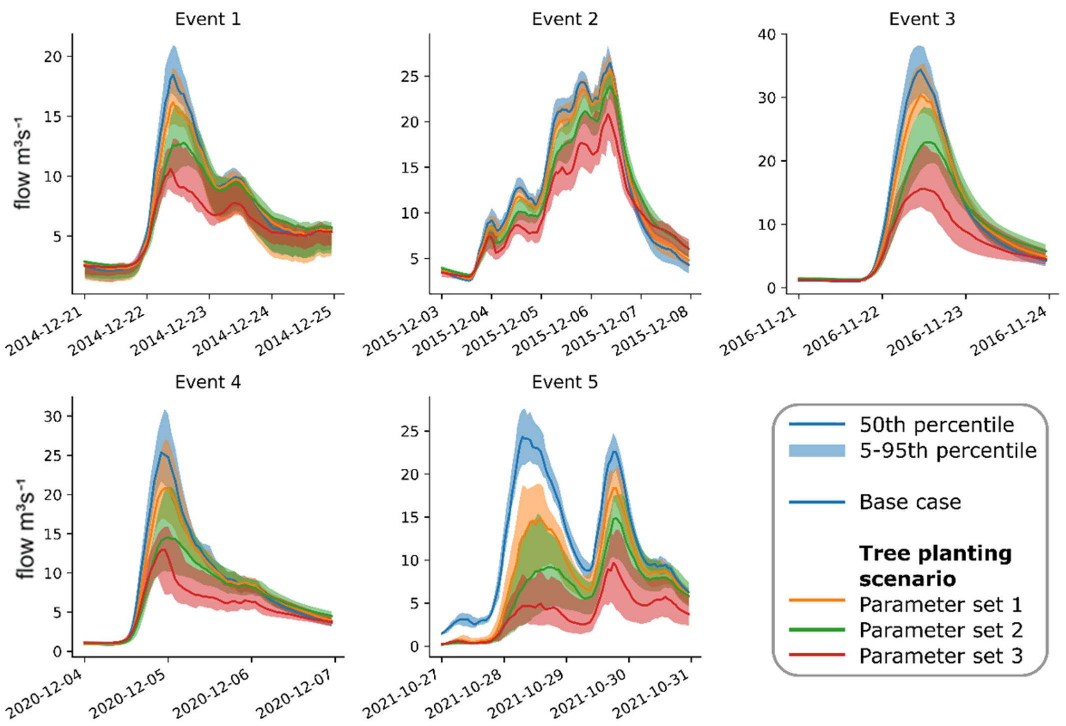

Figure 5. Impact of broadscale woodland planting for five flood events at the catchment outlet, Kidston Mill; *x*-axis is the date Parameter sets are given in Table 3.

In terms of bias, at the main stem gauges there is a tendency toward the models underestimating baseflow. This is particularly apparent at Craigburn, which also showed a poor match of peak and mean monthly flows, with simulated flows being too flashy (see Section 3.1.1). Model performance in terms of groundwater contribution is on average worse along the main stem than in the tributaries (Figures 4a and 4b). This is in contrast to streamflow performance, which is best at the outlet, Kidston Mill (Table 4). We can see from the correlation matrix (Figure 4c) that the models that perform well at Kidston Mill are also the best-performing models at the other main stem gauges, Craigburn and Eddleston Village, as well as at Middle Longcote, on the eastern tributary. These are the gauges with the largest groundwater contributions. In the western tributaries, where flows are flashy due to impermeable glacial till, there is little positive correlation with performance elsewhere in the catchment as well as among the three tributaries, with only Cowieslinn and Shiplaw having a weak positive correlation (0.51). Middle Burn is distinct from the other subcatchments in that it is the only gauge with a positive bias on average (i.e., simulated baseflow is too high) and a strong negative correlation with any of the other gauges—namely those along the main stem (−0.62 to −0.92) and the eastern tributary (−0.77) (Figure 4c). Middle Burn is the most forested of the subcatchments (93%), although a large portion (71%) of that was felled over the study period. The forest is coniferous plantation, and it could be that the drainage ditches constructed for the plantation are promoting a higher proportion of surface runoff than would be expected based on the superficial geology.

Carrying out model calibration based on streamflow and percent groundwater contribution (as determined by alkalinity measurement) versus on streamflow alone reduces the number of acceptable models from 200 to 18.

3.2. Evaluating Uncertainty

3.2.1. How Does the Choice of Speculative Parameter Shift Affect the Assessment of Effectiveness?

The impact of woodland planting can be assessed through changing model parameters controlling soil storage, overland flow velocity and soil permeability. In a sensitivity analysis, we tested three different parameter shifts—all within ranges found in the literature (see Section 2.4)—to see how these shifts affect the effectiveness of woodland planting as an NFM measure. Figure 5 shows the impact of the woodland planting scenario for the five largest peak flow events on record with the three different choices of parameter shifts (see Table 3).

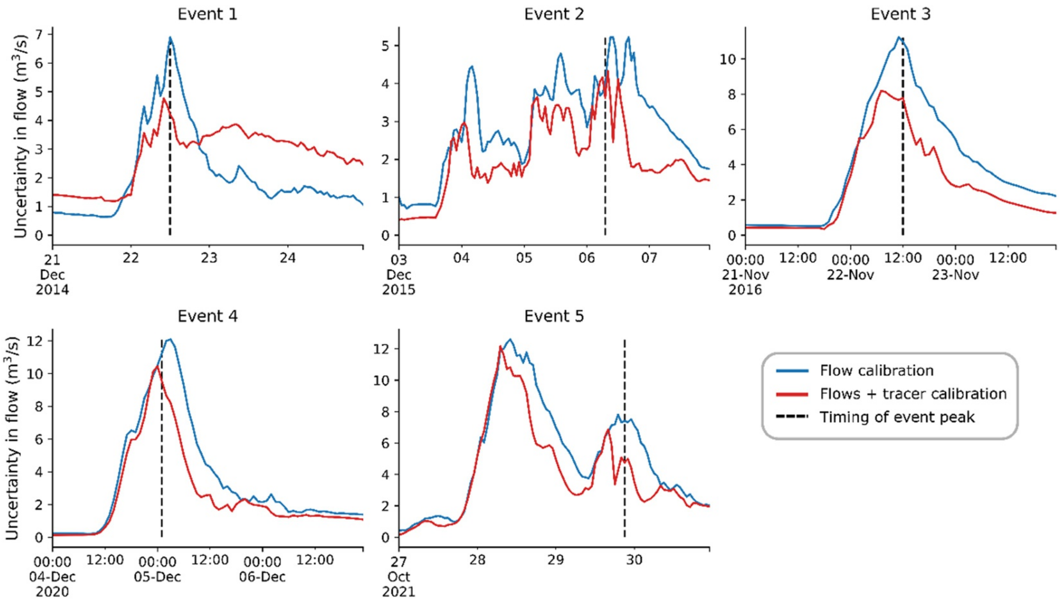

Figure 6. Predictive uncertainty in streamflow for the woodland planting scenario with parameter set 1 (see Table 3) where the acceptable models are selected based on streamflow and percent groundwater contribution performance versus on streamflow alone. The *x*-axis shows the time period.

For all five events, limiting the decline in hydraulic conductivity with depth and/or increasing the saturated transmissivity allows more water to infiltrate and flow through the saturated zone, leading to a less flashy response to heavy rainfall. The scenario is extreme, converting all land cover to broadleaved woodland, and therefore we see significant decreases in peak flows for most parameter sets and events. There is, however, a large variation in the predicted reduction in peak flow across the three parameter sets: for example, the reduction in the 50th percentile is 12%–43%, 3%−21%, 12%−55%, 18%−49%, and 24%−60% for events 1−5, respectively. These variations would likely be enough to alter a decision as to whether a tree planting scheme was worthwhile and cost effective or not. The uncertainty, defined as the difference between the 5th and 95th percentile, is large when considering all three possible parameter sets together: 10.1, 9.5, 22.6, 17.4, and 15.8 m^3/s for events 1−5, respectively (55%, 36%, 66%, 69%, 65% of median peak flow for the base case) (see Figure 5).

The uncertainty is lowest for event 2, which, while not the largest in terms of peak flow, is the largest event in terms of total volume of water, with flows being persistently high for several days. This reflects the widely held view (Dadson et al., 2017) that the impact of NFM measures reduces with the size of the event, as the catchment becomes overwhelmed with the volume of rainfall. For example, a higher rate of infiltration into the saturated zone may not be important if the saturated zone is full and the recharge is rejected. The model predicts the highest reductions in peak flow for event 5, which is likely a result of the timing of the event in October, when the broadleaved woodland still has its leaves and interception is high. The impact of season on the effectiveness of broadleaved woodland in reducing peak flows was also noted in the review by Burgess-Gamble et al. (2017). In this study we focus on the uncertainty derived from choosing the shift in soil parameters. There is, however, also significant uncertainty in how interception is parameterized (Page et al., 2020) and—perhaps to a lesser extent—in how overland flow is affected by land use change. Therefore, our results likely underestimate the uncertainty involved in modeling woodland planting as an NFM intervention.

3.2.2. Can Tracers Reduce Predictive Uncertainty in the Context of NFM?

Including the fraction of groundwater contribution to river flow as a calibration parameter results in a reduction in predictive uncertainty for the woodland planting scenario. Figure 6 compares the predictive uncertainty for both sets of acceptable models across the five highest peak flow events for the woodland planting scenario using the most modest shift in parameters (parameter set 1, Table 3). Introducing the tracer into the calibration reduces uncertainty at peak flow by 2.7 (39%), 0.37 (9%), 3.2 (29%), 1.7 (15%), and 2.7 m^3/s (36%) for events 1−5, respectively. Although encouraging, these reductions in uncertainty are smaller than the uncertainties arising

10.1029/2025WR041145

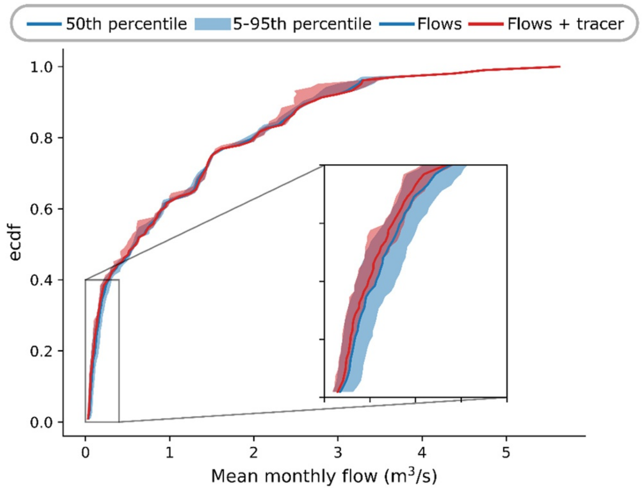

Figure 7. Empirical cumulative distribution function (ecdf) of mean monthly flows for the models calibrated only on streamflow versus on streamflow and percent groundwater contribution. Mean monthly flows are across the entire simulation period and for the woodland planting scenario using parameter set 1 (Table 3).

from how we represent woodland planting through parameter shifts (see Section 3.2.1), suggesting efforts to improve NFM scenario modeling should concentrate on how we represent NFM in models as much as on improving the models themselves. While predictive uncertainty in peak flow is reduced throughout the five events, we observe a large increase in predictive uncertainty in the falling limb of event 1. This is because some of the models show a very slow decline in the falling limb in event 1 and having a smaller number of acceptable models means that outliers have a stronger influence on the 5th and 95th percentiles of streamflow.

The models were calibrated not only to peak flows but also to mean monthly flows. Figure 7 shows the empirical cumulative distribution function of mean monthly flows over the entire simulation period for the woodland planting scenario using the most modest shift in parameters (parameter set 1, Table 3). It compares the 50th and 5−95th percentiles based on the calibration with streamflow and the tracer versus streamflow alone. For the highest ∼5% of monthly flows, although the difference in 5 to 95th percentiles is narrower for the multi-criteria calibration, the 50th percentiles are largely consistent. Thus, the multi-criteria calibration does not lead to a prediction of greater or lesser impact of the scenarios on the highest flows, but only a reduction in uncertainty. However, Figure 7 (inset) shows that the lowest ∼40% (Q60) of predicted monthly flows are consistently lower with the multi-criteria calibration as well as uncertainty being reduced (by ∼50%). This suggests that better constraining of the groundwater contribution to streamflow in hydrological models could have the largest impact in the study of low flows and droughts.

3.2.3. Wider Implications of the Study

The study has provided convincing evidence of the reduction in uncertainty in partitioning of flow in hydrological models which explore landcover change by including a second calibration data set. By evaluating the groundwater fraction using a model independent variable—ANC, we have shown that many of the models that were physically plausible and well calibrated using streamflow were giving the "right results for the wrong reasons" (Kirchner, 2006). By choosing models that also gave a plausible representation of the groundwater component, uncertainty was reduced by between upto 39% for woodland planting scenarios. Given the routine use of hydrological models for assessing the economic benefits from natural flood management (Hill et al., 2023), this reduction in uncertainty could be critical in deciding whether to proceed with a scheme and whether it was economically viable (Vörösmarty et al., 2021). The results have implications for modeling all scenarios that involve changes in landcover which essentially change the fraction of rainfall routed between surface and subsurface pathways. We did not aim to explore the value of alkalinity data in identifying better model structures or quantifying model structural uncertainty. However, this would be a valuable topic of further research. The application of alkalinity data could, for example, help to reject model structures within a framework for testing models as hypotheses such as that outlined by Beven (2018).

The uncertainty introduced by the different approaches used to represent woodland within the model also has wide ranging implications for uncertainty in catchment modeling. Given the lack of convincing direct evidence for tree planting on flood reduction (Carrick et al., 2019; Stratford et al., 2017) numerical modeling is often relied on to make predictions (Kay et al., 2019). However, the high sensitivity of results to how woodland is represented is troubling, with uncertainty in flood peak magnitude varying by more than 50%. This strongly suggests that more research is required to both gather evidence of the processes by which catchment tree-planting impacts flows and how these processes can be reliably incorporated in numerical models.

An interesting finding from the from the study was the improvements to the representation of low flows by using ANC as a secondary calibration. This is of particular relevance to current Natural Flood Management research where attention is beginning to grow on the potential for undesirable reductions in low flows due to Natural Flood Management (van Meerveld and Seibert, 2025). The increased use of tracers, such as ANC, that help identify groundwater fractions could therefore prove helpful in developing models that are appropriate to explore impacts at both high and low flows.

4. Conclusions

We built a hydrological model of a well-monitored catchment typical of temperate uplands. A multi-criteria calibration was applied to the model: first on streamflow and then on groundwater contribution to streamflow, which was estimated based on alkalinity (ANC) data. A broadscale woodland planting scenario was simulated by applying parameter shifts to the calibrated model. To investigate uncertainty, a sensitivity analysis was conducted whereby three different shifts representing enhanced soil permeability under woodland were applied, all within the range of values used in the literature. Finally, we quantified the extent to which the multi-criteria calibration reduced predictive uncertainty in the woodland scenario versus a calibration based only on streamflow.

We found that model runs indistinguishable with respect to peak and mean monthly flows could vary significantly in groundwater contribution to streamflow and how runoff is generated and routed. Therefore, including ANC data in model calibration improved model robustness and reduced model uncertainty in the context of NFM. Including ANC data also reduced predictive uncertainty, when applying the calibrated model to a woodland planting scenario by upto 39%. Although the focus of the work was on peak flows, our results also indicate that simple tracers such as alkalinity could be even more helpful when assessing low flows and droughts.

The three common assumptions we tested on how woodland planting alters soil properties significantly impacted our assessment of how effective woodland is as a measure against flooding. This introduced uncertainty of approximately 50% on peak flows. While we concentrate on woodland planting, the results are likely applicable to other forms of land-based NFM that aim to alter soil properties, such as changes to livestock or crop management. Therefore, while tracers significantly reduced uncertainty in scenario analysis, the uncertainty is likely dominated by how we choose to represent NFM in the model. This underlines the vital importance of long-term monitoring for improving the evidence for NFM, and process understanding to improve the representation of NFM in catchment models.

Further research should focus on: (a) the use of secondary data sets for calibration, such as ANC, stable isotopes or organic carbon reduction to reduce uncertainty at both high and low flows and choice of model structure; (b) gathering systematic evidence for how woodland planting and other catchment NFM measures impact soil permeability and partitioning between runoff and sub surface flow to improve representation in catchment models.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The model code is available via the Zenodo repository (Coxon & Dunne, 2019). The model inputs and outputs are freely available at http://www.hydroshare.org/resource/432d6a253ea3482cb1dc123120979f18.

Acknowledgments
We would like to thank the Tweed Forum for supporting this study as part of the Eddleston Water Project. The project is a partnership initiative led by Tweed Forum, with the Scottish Government, SEPA and University of Dundee, and is funded by the Scottish Government. We also wish to thank Gemma Coxon for her help and advice on the use of the hydrological model, DECIPHeR. Sarah Collins, Christopher Jackson and Alan MacDonald publish with the permission of the Executive Director of the British Geological Survey (UKRI).

References

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. *Hydrology and Earth System Sciences*, *21*(10), 5089–5110. https://doi.org/10.5194/hess-21-5089-2017

Archer, N. A. L., Bonell, M., Coles, N., MacDonald, A., Auton, C., & Stevenson, R. (2013). Soil characteristics and landcover relationships on soil hydraulic conductivity at a hillslope scale: A view towards local flood management. *Journal of Hydrology*, *497*, 208–222. https://doi.org/10.1016/j.jhydrol.2013.05.043

Archer, N. A. L., Otten, W., Schmidt, S., Bengough, A. G., Shah, N., & Bonell, M. (2016). Rainfall infiltration and soil hydrological characteristicsbelow ancient forest, planted forest and grassland in a temperate northern climate. *Ecohydrology*, *9*, 585–600. https://doi.org/10.1002/eco.1658

Auton, C. (2011). *Eddleston water catchment, superficial geology, 1: 25 000 scale*. British Geological Survey.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Beven, K., & Binley, A. (2014). Glue: 20 years on. *Hydrological Processes*, *28*(24), 5897–5918. https://doi.org/10.1002/hyp.10082

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *Journal of Hydrology*, *249*(1–4), 11–29. https://doi.org/10.1016/s0022-1694(01)00421-8

Beven, K. J. (2018). On hypothesis testing in hydrology: Why falsification of models is still a really good idea. *Wiley Interdisciplinary Reviews. Water*, *5*(3), e1278. https://doi.org/10.1002/wat2.1278

Birkel, C., Soulsby, C., & Tetzlaff, D. (2014). Developing a consistent process-based conceptualization of catchment functioning using measurements of internal state variables. *Water Resources Research*, *50*(4), 3481–3501. https://doi.org/10.1002/2013WR014925

Black, A., Peskett, L., MacDonald, A., Young, A., Spray, C., Ball, T., et al. (2021). Natural flood management, lag time and catchment scale: Results from an empirical nested catchment study. *Journal of Flood Risk Management*, *14*(3), e12717. https://doi.org/10.1111/jfr3.12717

Bracken, L. J., Wainwright, J., Ali, G., Tetzlaff, D., Smith, M., Reaney, S., & Roy, A. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, *119*, 17–34. https://doi.org/10.1016/j.earscirev.2013.02.001

Brickell, J., MacDonald, A., Collins, S., Peskett, L., Ball, T., Black, A., et al. (2024). *Investigations into the effect of different land use on field-saturated hydraulic conductivity in the Eddleston Water catchment. CR/24/046N*. British Geological Survey. Retrieved from https://nora.nerc.ac.uk/id/eprint/537616/

Buechel, M., Slater, L., & Dadson, S. (2022). Hydrological impact of widespread afforestation in Great Britain using a large ensemble of modelled scenarios. *Communications Earth & Environment*, *3*, 1–10. https://doi.org/10.1038/s43247-021-00334-0

Burgess-Gamble, L., Ngai, R., Wilkinson, M., Nisbet, T., Pontee, N., Harvey, R., et al. (2017). *Working with natural processes–evidence directory* Report No. SC150005. Environment Agency.

Cao, W., Bowden, W. B., Davie, T., & Fenemor, A. (2006). Multi-variable and multi-site calibration and validation of swat in a large mountainous catchment with high spatial variability. *Hydrological Processes: International Journal*, *20*(5), 1057–1073. https://doi.org/10.1002/hyp.5933

Carrick, J., Abdul Rahim, M. S. A. B., Adjei, C., Ashraa Kalee, H. H. H., Banks, S. J., Bolam, F. C., et al. (2019). Is planting trees the solution to reducing flood risks? *Journal of Flood Risk Management*, *12*(S2), e12484. https://doi.org/10.1111/jfr3.12484

Chandler, K., Stevens, C., Binley, A., & Keith, A. (2018). Influence of tree species and forest land use on soil hydraulic conductivity and implications for surface runoff generation. *Geoderma*, *310*, 120–127. https://doi.org/10.1016/j.geoderma.2017.08.011

CIRIA. (2022). The natural flood management manual (C802F). Retrieved from https://www.ciria.org/ItemDetail?iProductCode=C802F

Cohen-Shacham, E., Walters, G., Janzen, C., & Maginnis, S. (2016). *Nature-based solutions to address global societal challenges* (Vol. 97, pp. 2016–2036). IUCN.

Collins, S. L., Verhoef, A., Mansour, M., Jackson, C. R., Short, C., & Macdonald, D. M. J. (2023). Modelling the effectiveness of land-based natural flood management in a large, permeable catchment. *Journal of Flood Risk Management*, *16*(2), e12896. https://doi.org/10.1111/jfr3.12896

CORINE. (2018). Land cover, Europe, 6-yearly. Retrieved from https://land.copernicus.eu/en/products/corine-land-cover?tab=overview

Costaz-Puyou, I. (2022). *Assessing the potential of channel re-meandering for flood attenuation: A detailed case study of the restoration of the Eddleston Water, Scotland* Doctoral Thesis. University of Dundee. https://doi.org/10.15132/20000201

Coulthard, T. J., Kirkby, M. J., & Macklin, M. G. (2000). Modelling geomorphic response to environmental change in an upland catchment. *Hydrological Processes*, *14*(11-12), 2031–2045. https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<2031::aid-hyp53>3.0.co;2-g

Coulthard, T. J., & Van De Wiel, M. J. (2017). Modelling long term basin scale sediment connectivity, driven by spatial land use changes. *Geomorphology*, *277*, 265–281. https://doi.org/10.1016/j.geomorph.2016.05.027

Coxon, G., & Dunne, T. (2019). DECIPHeR-v.10. Code. *Zenodo*. https://doi.org/10.5281/zenodo.2604120

Coxon, G., Freer, J., Lane, R., Dunne, T., Knoben, W. J., Howden, N. J., et al. (2019). Decipher v1: Dynamic fluxes and connectivity for predictions of hydrology. *Geoscientific Model Development*, *12*(6), 2285–2306. https://doi.org/10.5194/gmd-12-2285-2019

Dadson, S. J., Hall, J. W., Murgatroyd, A., Acreman, M., Bates, P., Beven, K., et al. (2017). A restatement of the natural science evidence concerning catchment-based 'natural' flood management in the UK. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2199), 20160706. https://doi.org/10.1098/rspa.2016.0706

DeWalle, D. R. (2011). *Benchmark papers in forest hydrology: Introduction, selection and commentary* (p. 474). IAHS Press.

Dunn, S. M., & Mackay, R. (1995). Spatial variation in evapotranspiration and the influence of land use on catchment hydrology. *Journal of Hydrology*, *171*(1–2), 49–73. https://doi.org/10.1016/0022-1694(95)02733-6

Faivre, N., Fritz, M., Freitas, T., De Boissezon, B., & Vandewoestijne, S. (2017). Nature-based solutions in the EU: Innovating with nature to address social, economic and environmental challenges. *Environmental Research*, *159*, 509–518. https://doi.org/10.1016/j.envres.2017.08.032

FAO. (1998). Crop evapotranspiration; guidelines for computing crop water requirements. In *FAO Irrigation and Drainage Paper 56*. FAO.

Fenicia, F., McDonnell, J. J., & Savenije, H. H. (2008). Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resources Research*, *44*(6). https://doi.org/10.1029/2007wr006386

Ferguson, C., & Fenner, R. (2020). The impact of natural flood management on the performance of surface drainage systems: A case study in the Calder Valley. *Journal of Hydrology*, *590*, 125354. https://doi.org/10.1016/j.jhydrol.2020.125354

Flood and Water Management Act. (2010). Retrieved from http://www.legislation.gov.uk/ukpga/2010/29/contents

Flood Risk Management (Scotland) Act. (2009). Retrieved from https://www.legislation.gov.uk/asp/2009/6/contents

Goudarzi, S., Milledge, D. G., Holden, J., Evans, M. G., Allott, T. E., Shuttleworth, E. L., et al. (2021). Blanket peat restoration: Numerical study of the underlying processes delivering natural flood management benefits. *Water Resources Research*, *57*(4), e2020WR029209. https://doi.org/10.1029/2020wr029209

Graham, M. T., Ball, D. F., Ó Dochartaigh, B. É., & MacDonald, A. M. (2009). Using transmissivity, specific capacity and borehole yield data to assess the productivity of Scottish aquifers. *The Quarterly Journal of Engineering Geology and Hydrogeology*, *42*(2), 227–335. https://doi.org/10.1144/1470-9236/08-045

Griffiths, J., Young, A. R., & Keller, V. (2006). *Model scheme for representing rainfall interception and soil moisture*. Environment Agency Environment Agency R & D Project W6-101 Continuous Estimation of River Flows (CERF).

Hankin, B., Craigen, I., Chappell, N. A., Page, T., & Metcalfe, P. (2016). *Strategic investigation of natural flood management in Cumbria*. The Rivers Trust, Technical Report, UK. https://doi.org/10.13140/RG.2.2.31539.53287

Hankin, B., Page, T. J., Chappell, N. A., Beven, K. J., Smith, P. J., Kretzschmar, A., & Lamb, R. (2021). Using micro-catchment experiments for multi-local scale modelling of nature-based solutions. *Hydrological Processes*, *35*(11), e14418. https://doi.org/10.1002/hyp.14418

Hill, B., Liang, Q., Bosher, L., Chen, H., & Nicholson, A. (2023). A systematic review of natural flood management modelling: Approaches, limitations, and potential solutions. *Journal of Flood Risk Management*, *16*(3), e12899. https://doi.org/10.1111/jfr3.12899

Holmes, T., Stadnyk, T. A., Kim, S. J., & Asadzadeh, M. (2020). Regional calibration with isotope tracers using a spatially distributed model: A comparison of methods. *Water Resources Research*, *56*(9), e2020WR027447. https://doi.org/10.1029/2020WR027447

Holmes, T. L., Stadnyk, T. A., Asadzadeh, M., & Gibson, J. J. (2022). Variability in flow and tracer-based performance metric sensitivities reveal regional differences in dominant hydrological processes across the Athabasca River Basin. *Journal of Hydrology: Regional Studies*, *41*, 101088. https://doi.org/10.1016/j.ejrh.2022.101088
Iacob, O., Brown, I., & Rowan, J. (2017). Natural flood management, land use and climate change trade-offs: The case of Tarland Catchment, Scotland. *Hydrological Sciences Journal*, *62*(12), 1931–1948. https://doi.org/10.1080/02626667.2017.1366657
IPCC. (2022). In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, et al. (Eds.), *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 3056). Cambridge University Press. https://doi.org/10.1017/9781009325844
James, F. (1988). *A review of pseudorandom number generators*. CERN Data Handling Division.
JBA. (2020). Eddleston water hydrologic and hydraulic modelling of NFM: Phase 2 report. Retrieved from https://tweedforum.org/eddleston-project-database/
Kay, A., Old, G., Bell, V., Davies, H., & Trill, E. (2019). An assessment of the potential for natural flood management to offset climate change impacts. *Environmental Research Letters*, *14*(4), 044017. https://doi.org/10.1088/1748-9326/aafdbe
Kingsbury-Smith, L., Willis, T., Smith, M., Boisgontier, H., Turner, D., Hirst, J., et al. (2023). Evaluating the effectiveness of land use management as a natural flood management intervention in reducing the impact of flooding for an upland catchment. *Hydrological Processes*, *37*(4), e14863. https://doi.org/10.1002/hyp.14863
Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, *42*(3). https://doi.org/10.1029/2005wr004362
Kirchner, J. W., Benettin, P., & van Meerveld, I. (2023). Instructive surprises in the hydrological functioning of landscapes. *Annual Review of Earth and Planetary Sciences*, *51*(1), 277–299. https://doi.org/10.1146/annurev-earth-071822-100356
Klaus, J., & McDonnell, J. (2013). Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, *505*, 47–64. https://doi.org/10.1016/j.jhydrol.2013.09.006
Kreibich, H., Van Loon, A. F., Schröter, K., Ward, P. J., Mazzoleni, M., Sairam, N., et al. (2022). The challenge of unprecedented floods and droughts in risk management. *Nature*, *608*(7921), 80–86. https://doi.org/10.1038/s41586-022-04917-5
Maes, J., & Jacobs, S. (2017). Nature-based solutions for Europe's sustainable development. *Conservation Letters*, *10*(1), 121–124. https://doi.org/10.1111/conl.12216
Mansour, M. M., Wang, L., Whiteman, M., & Hughes, A. G. (2018). Estimation of spatially distributed groundwater potential recharge for the United Kingdom. *The Quarterly Journal of Engineering Geology and Hydrogeology*, *51*(2), 247–263. https://doi.org/10.1144/qjegh2017-051
Marsaglia, G., & Zaman, A. (1987). *Towards a universal random number generator* Report FSU-SCRI-87-50. Florida State University.
Martin, E. G., Costa, M. M., & Máñez, K. S. (2020). An operationalized classification of nature based solutions for water-related hazards: From theory to practice. *Ecological Economics*, *167*, 106460. https://doi.org/10.1016/j.ecolecon.2019.106460
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 Resources Research*, *50*(6), 5342–5350. https://doi.org/10.1002/2013wr015141
McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., et al. (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, *43*(7). https://doi.org/10.1029/2006wr005467
Monger, F., V Spracklen, D., J Kirkby, M., & Schofield, L. (2022). The impact of semi-natural broadleaf woodland and pasture on soil properties and flood discharge. *Hydrological Processes*, *36*(1), e14453. https://doi.org/10.1002/hyp.14453
Murphy, T. R., Hanley, M. E., Ellis, J. S., & Lunt, P. H. (2021). Native woodland establishment improves soil hydrological functioning in UK upland pastoral catchments. *Land Degradation & Development*, *32*(2), 1034–1045. https://doi.org/10.1002/ldr.3762
NatureScot. (2019). Scotland habitat and land cover map—2019. Retrieved from https://data.gov.uk/dataset/6afd78f1-1cfe-4735-9169-6a3de6fee958/scotland-habitat-and-land-cover-map-2019
Neill, A. J., Birkel, C., Maneta, M. P., Tetzlaff, D., & Soulsby, C. (2021). Structural changes to forests during regeneration affect water flux partitioning, water ages and hydrological connectivity: Insights from tracer-aided ecohydrological modelling. *Hydrology and Earth System Sciences*, *25*(9), 4861–4886. https://doi.org/10.5194/hess-25-4861-2021
Ó Dochartaigh, B. E., MacDonald, A. M., Fitzsimons, V., & Ward, R. (2015). *Scotland's aquifers and groundwater bodies* (p. 76). British Geological Survey Open Report, OR/15/028.
Packman, J., Quinn, P., Hollis, J., & O'Connell, P. (2004). Review of impacts of rural land use and management on flood generation. In *Short term improvement to the FEH rainfall−runoff model: Technical Background*. DEFRA.
Page, T., Chappell, N. A., Beven, K. J., Hankin, B., & Kretzschmar, A. (2020). Assessing the significance of wet-canopy evaporation from forests during extreme rainfall events for flood mitigation in mountainous regions of the United Kingdom. *Hydrological Processes*, *34*(24), 4740–4754. https://doi.org/10.1002/hyp.13895
Peskett, L., MacDonald, A., Heal, K., McDonnell, J., Chambers, J., Uhlemann, S., et al. (2020). The impact of across-slope forest strips on hillslope subsurface hydrological dynamics. *Journal of Hydrology*, *581*, 124427. https://doi.org/10.1016/j.jhydrol.2019.124427
Peskett, L. M., Collins, S. L., Black, A., Arran, M., MacDonald, A. M., & Young, A. (2025). Using storage ponds in natural flood management schemes in practice: The need for fine-tuning and upscaling. *Journal of Flood Risk Management*, *18*(2), e70059. https://doi.org/10.1111/jfr3.70059
Peskett, L. M., Heal, K. V., MacDonald, A. M., Black, A. R., & McDonnell, J. J. (2021). Tracers reveal limited influence of plantation forests on surface runoff in a UK natural flood management catchment. *Journal of Hydrology: Regional Studies*, *36*, 100834. https://doi.org/10.1016/j.ejrh.2021.100834
Peskett, L. M., Heal, K. V., MacDonald, A. M., Black, A. R., & McDonnell, J. J. (2023). Land cover influence on catchment scale subsurface water storage investigated by multiple methods: Implications for UK natural flood management. *Journal of Hydrology: Regional Studies*, *47*, 101398. https://doi.org/10.1016/j.ejrh.2023.101398
Rose, S., & Rosolova, Z. (2007). Ripon land management project. In *JBA consulting*.
Rutter, A. J., Morton, A. J., & Robins, P. C. (1975). A predictive model of rainfall interception in forests. II. Generalization of the model and comparison with observations in some coniferous and hardwood stands. *Journal of Applied Ecology*, *12*(1), 367–380. https://doi.org/10.2307/2401739
Scudeler, C., Pangle, L., Pasetto, D., Niu, G.-Y., Volkmann, T., Paniconi, C., et al. (2016). Multiresponse modeling of variably saturated flow and isotope tracer transport for a hillslope experiment at the landscape evolution observatory. *Hydrology and Earth System Sciences*, *20*(10), 4061–4078. https://doi.org/10.5194/hess-20-4061-2016
SEPA. (2015). *Natural flood management handbook*. SEPA. Available at: sepa-natural-flood-management-handbook1.pdf.

Spray, C., Black, A., Bradley, D., Bromley, C., Caithness, F., Dodd, J., et al. (2022). Strategic design and delivery of integrated catchment restoration monitoring: Emerging lessons from a 12-year study in the UK. *Water*, *14*(15), 2305. https://doi.org/10.3390/w14152305

Stevenson, J. L., Birkel, C., Neill, A. J., Tetzlaff, D., & Soulsby, C. (2021). Effects of streamflow isotope sampling strategies on the calibration of a tracer-aided rainfall-runoff model. *Hydrological Processes*, *35*(6), e14223. https://doi.org/10.1002/hyp.14223

Stratford, C., Miller, J., House, A., Old, G., Acreman, M., Dueñas-Lopez, M., et al. (2017). Do trees in UK-relevant river catchments influence fluvial flood peaks. In *A systematic review* (Vol. 1).

Tunaley, C., Tetzlaff, D., Birkel, C., & Soulsby, C. (2017). Using high-resolution isotope data and alternative calibration strategies for a tracer-aided runoff model in a nested catchment. *Hydrological Processes*, *31*(22), 3962–3978. https://doi.org/10.1002/hyp.11313

van Meerveld, I., & Seibert, J. (2025). Reforestation effects on low flows: Review of public perceptions and scientific evidence. *Wiley Interdisciplinary Reviews. Water*, *12*(1), e1760. https://doi.org/10.1002/wat2.1760

Vicarelli, M., Sudmeier-Rieux, K., Alsadadi, A., Shrestha, A., Schütze, S., Kang, M. M., et al. (2024). On the cost-effectiveness of nature-based solutions for reducing disaster risk. *Science of the Total Environment*, *947*, 174524. https://doi.org/10.1016/j.scitotenv.2024.174524

Vörösmarty, C. J., Stewart-Koster, B., Green, P. A., Boone, E. L., Flörke, M., Fischer, G., et al. (2021). A green-gray path to global water security and sustainable infrastructure. *Global Environmental Change*, *70*, 102344. https://doi.org/10.1016/j.gloenvcha.2021.102344

World Bank. (2018). *Nature-based solutions for disaster risk management* (p. 24). World Bank. Retrieved from http://documents.worldbank.org/curated/en/253401551126252092/pdf/134847-NBS-for-DRM-booklet.pdf

Wu, S., Tetzlaff, D., Yang, X., Smith, A., & Soulsby, C. (2023). Integrating tracers and soft data into multi-criteria calibration: Implications from distributed modeling in a riparian wetland. *Water Resources Research*, *59*(11), e2023WR035509. https://doi.org/10.1029/2023WR035509