DOI: 10.1002/hvp.14418

IMPACTS OF OBSERVATIONAL UNCERTAINTY ON ANALYSIS AND MODELLING OF HYDROLOGICAL PROCESSES

Using micro-catchment experiments for multi-local scale modelling of nature-based solutions

Barry Hankin¹ | Trevor J. C. Page¹ | Nick A. Chappell¹ | Keith J. Beven¹ | Paul J. Smith² | Ann Kretzschmar¹ | Rob Lamb^{1,3}

¹Lancaster Environment Centre, Lancaster University, Lancaster, UK ²Waternumbers, Lancaster, UK ³JBA Trust, Skipton, UK

Correspondence

Barry Hankin, Lancaster Environment Centre, Lancaster University, Lancaster, UK. Email: barry.hankin@jbaconsulting.com

Funding information Natural Environment Research Council, Grant/ Award Number: NE/R004722/1

Abstract

The Q-natural flood management project has co-developed with the Environment Agency 18 monitored micro-catchments (~1 km²) in Cumbria, UK installing calibrated flumes aimed at quantifying the potential shift in observed flows resulting from a range of nature-based-solutions installed by local organizations. The small-scale reduces the influence of variability characterizing larger catchments that would otherwise mask any such shifts, which we attempt to relate to a shift in model parameters. This paper demonstrates an approach to applying donor-parameter-shifts obtained from modelling two of the paired micro-catchments to a much larger scale, in order to understand the potential for improved distributed modelling of nature-based solutions in the form of additional tree-planting. The models include a rainfall-runoff model, Dynamic Topmodel, and a 2D hydrodynamic model, JFlow, permitting analysis of changes in hillslope processes and channel hydrodynamics resulting from a range of distributed measures designed to emulate natural hydrological processes that evaporate, store or infiltrate flows. We report on attempts to detect shift in hydrological response using one of the paired-micro-catchment moorland versus forestry sites in Lorton using Dynamic Topmodel. A donor-parameter-shift approach is used in a hypothetical experiment to represent new woodland in a much larger catchment, although testing all combinations of spatial planting strategies, responses to multiple-extremes, failure-modes and changes to synchronization becomes intractable to support good decision making. We argue that the problem can be re-framed to use donor-parameter-shifts at multi-localscale catchments above communities known to be at risk, commensurate with most of the evidence of NbS impacts being effective at the small scale (ca. 10 km²). This might lead to more effective modelling to help catchment managers prioritize those communities-at-risk where there is more evidence that NbS might be effective.

KEYWORDS

Dynamic Topmodel, hydrology, natural flood management, nature-based solutions, uncertainty

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Hydrological Processes published by John Wiley & Sons Ltd.

WILEY

^{2 of 17} WILEY-

1 | INTRODUCTION

In recent years, there has been a marked international interest in restoring natural capital to seek multiple benefits from a range of nature-based solutions, NbS (e.g., Bridges et al., 2018; European Commission, 2020), with an accompanying scientific interest to assess efficacy (in theoretical trials) and effectiveness (in practice) of environmental benefits including potential to provide flood risk reduction (Burgess-Gamble et al., 2017; Dadson et al., 2017; Lane, 2017). The specific NbS of reducing flood hydrographs is known as natural flood management (NFM) in the United Kingdom.

Quantifying the effectiveness of range of different NFM measures aimed at storing, infiltrating and evaporating flood water, has mainly been possible with current trials only at the small scale of the order 10 km² (Dadson et al., 2017). However, at larger scales, the observed change in hydrological response resulting from a large number of distributed NFM measures (impacting multiple hydrological processes) soon gets 'drowned out' by inevitable environmental variability. This is exacerbated in complex networks where synchronization issues come into play (Ferguson & Fenner, 2020; Metcalfe et al., 2018; Pattison et al., 2014), and for which NFM asset failure can add to the complexity. For example, failure and cascade failure of leakybarriers is explored in Hankin et al. (2020), or the interdependencies of distributed NFM in combination with an urban drainage network is considered in Ferguson and Fenner (2020).

Coupled with variable hydrological extremes of different durations and spatial patterns within large catchments, the testing framework and advice on spatial deployment of NbS at larger scales, becomes almost intractable. An alternative solution, which stems from engaged environmental NGO partners (Hankin, 2020) and the Environment Agency, is to focus on NbS placement in small subcatchments above communities at risk of frequent flooding (in locations where the flow accumulation is also smaller than 10 km²), where the supporting experimental evidence is greater and where the necessary intensive deployment of measures to reduce risk may be achieved more quickly.

1.1 | Research questions

Q-NFM is one of three UKRI Natural Environment Research Council investigations (NERC, 2020) quantifying the effectiveness of NFM. NFM is specifically about flood hydrograph reductions and so one aspect of the wider NbS for environmental, social and economic gain. The project has set up 18 micro-basins in Cumbria (UK), with accurately calibrated flows and local raingauges to try and measure any changes in hydrological response resulting from a range of NbS measures. One aspect of the Q-NFM project is to detect any shift in hydrological response at the micro-basin scale (1 km²) due to NbS and map this onto model parameters, so this shift could be applied at greater scales up to the macro-scale (>1000 km²). The key research questions addressed here are answered in-part through a number of modelling experiments based on the paired micro-basins and observational data from a macro-scale catchment, making this a demonstration paper. These modelling experiments are subject to measurement and knowledge uncertainties, and it is acknowledged from the outset that many more such modelling experiments will be needed across multiple micro-basins and through before-after intervention experiments for a range of storms and antecedent conditions to support the target research questions as follows:

- Can changes in hydrological responses due to distributed NbS be characterized with fuzzy parameter shifts in a distributed rainfall runoff models using micro-basin comparisons? For example, are there dominant (but uncertain) changes seen in the effective parameters in a calibration of a wooded catchment as opposed to a moorland catchment for the same event, that are scalable to other locations?
- 2. Can these fuzzy parameter-shifts be attributed to differing landscape scale (spatial) characteristics of the micro-catchments
- 3. Can these be used to quantify change associated with particular hydrological (temporal) changes arising from nature-based solutions (NbS)?
- 4. If these shifts are implemented in a model of a much larger scale, such as the Eden (2300 km²), is it possible to reduce uncertainties with new types of data, such as satellite imagery?
- 5. At high stream-orders in large catchments, is it possible to appraise risk by characterizing an 'average response' (Hankin et al., 2017a), or key modes of behaviour (Hankin, 2020) stemming from the superposition of many small-scale changes to hydrological response resulting from many small-scale distributed NbS measures?
- 6. Given a distributed hydrological model of a large system (>100 km²) with a strong global measure of performance at the outlet, should we focus the assessment of flood risk-reduction due to NbS at the *multi-local scale*? In this way, the donor parameter shifts are applied at a similar scale to the micro-basins they are derived from, and the scaling-issue associated with non-linear hydrological processes is partly avoided.

1.2 | Overview of donor parameter shift experiment

The focus in this paper is to test the potential for detecting of parameter shifts, given the uncertainties in the modelling process, in a pair of these micro-basins in a tributary of the River Cocker catchment (Cumbria, UK), one with moorland grazing and the other with a commercial forestry plantation (see Figure 1). These changes are then used as 'donor parameter shifts' to represent similar, future hydrological changes as a result of woodland planting in this instance, or more generally for other types of NbS interventions.

The broader experimental approach used here uses the rainfall runoff model, Dynamic Topmodel (Beven & Freer, 2001; Metcalfe et al., 2016) using a recently improved open-source code (Smith, 2020), calibrated to the two micro-catchments for two storms



FIGURE 1 Overview of Q-natural flood management paired micro-basins

in December 2019 (named Storm Atiyah) and then in February 2020 (named Storm Ciara). Dynamic Topmodel is fundamentally suited to the idea of identifying changes to hydrological responses since it groups areas of similar hydrological response in the landscape (hydrological response units or HRUs), typically based on classes of wetness index and for example average rainfall.

As with many applications of Dynamic Topmodel, the greatest sensitivity is in the *m* parameter, defined as the rate of decline in the downslope transmissivity with depth below the ground surface (see Beven, 2001a), and strongly influences the rate of hydrograph recession. In this paper, *m* has been calibrated for the two particular winter storms and antecedent conditions. Myers et al. (2021) investigated the impact of the calibration period and how it can have a large impact on derived parameter values and stress the need to consider effects of non-stationarity in parameter estimates. Here, the donor parameter-shift identified in the calibration events is applied to one of the same storms in a different larger catchment to understand possible changes to hydrological response at a much larger scale in the 2300 km² Eden catchment.

Linking recession parameters to broad land use/cover types is not new, and has been demonstrated in a range of studies (see Bulygina et al., 2012; Heuvelmans et al., 2004; Karvonen et al., 1999), including Bogaart et al. (2016), where it was concluded that dynamic recession parameters can be linked to landscape evolution. Beven (2001b) also discusses mapping parameter-landscape mapping in an uncertainty framework, whereas here we are interested in contrasting behaviour of two adjacent catchments. Rather than aiming to transpose 'calibration parameters from similar, gauged catchments' to address the ungauged catchment problem (Blöschl, 2006), the approach here is to 'transpose observed contrasts in calibration parameters to reflect internal hydrological differences', with respect to particular magnitude storms and for particular types of NbS. The approach is also more aligned with the 'uncertain or fuzzy landscape space to model space mapping' discussed by Beven (2001a). Whilst the approach of donor parameters is used commonly in hydrological analyses of ungauged catchments, it does not address the scale problem on how parametershifts might change with size of catchment, the size and duration of storms and antecedent conditions, due to the non-linearity of hydrological processes. However, this is partly addressed through the recommended approach of considering the multi-local scale, where we step away from trying to understand the whole-system response to many small distributed interventions at very large scale, where observations have only been made at the micro-basin scale, and instead focus in multiple small scale watersheds above communities at risk.

The paper explores the fuzzy change observed in calibrated *m* parameters in Dynamic Topmodel between wooded and non-wooded

hillslopes, having similar size, aspect and hydrological characteristics (e.g., the baseflow index and standard percentage runoff are the same) for two different winter storms, and the same relative shift was then applied at sites at a much larger scale in the Eden catchment (2300 km²) in Cumbria, where tree planting would be expected to have the greatest impact on reducing the likelihood of saturation overland flow for one of the same storms (Storm Ciara). However, whilst the paired catchments were selected carefully for similar scale, aspect, slope and topography, it is acknowledged that the donor-shifts can in part be due to other differences, not just as a result of woodland, including differences in micro-topography, drainage and geology. Eliminating such differences in a non-laboratory system is very difficult, so further replication or apportionment of change due to other differences is needed.

2 | DATA AND METHODS

The paired micro-basins in the Cocker catchment are first described, along with the much larger Eden catchment in Section 2.2. Two key storm periods were investigated, for named-storm Atiyah in December, 2019 and named-storms Ciara and Dennis in February 2020, which impacted some communities badly (e.g., Appleby in the Eden catchment). The first two sections introduce the catchment characteristics, followed by details of the independent fuzzy calibrations of the three catchments, along with detailed of the modelling experiments.

2.1 | Micro-basin data: Micro-basins in the cocker catchment

Figure 1 presents the two small paired Cumbrian micro-basins, with moorland characterizing the Sware Gill micro-basin to the west in red, and the commercial conifer forest of the Darling How micro-basin to the east. The different characteristics have been obtained from the UK Flood Estimation Handbook (Centre for Ecology & Hydrology, 1999) and the BGS (2021), and are summarized below.

2.1.1 | Sware Gill

Land-use: *rough grazing*; Area: 0.24 km², Superficial geology—absent; solid geology: Loweswater formation dominated by relatively low permeability greywacke beneath the lower half of the basin, and Kirk Stile Formation of mudstone and siltstone beneath the upper half of the basin; altitude: 404 mAOD; base flow index: 0.57, specific percentage runoff: 33.3, standard annual average rainfall: 1916 mm).

2.1.2 | Darling How

Land-use: working confer plantation (33% mature trees with some larch and pine); Area: 0.46 km², Superficial geology–glacial till diamicton

covers the centre of the catchment; solid geology: Kirk Stile Formation, predominantly of permeable sandstone; altitude: 403 mAOD; base flow index: 0.57, specific percentage runoff: 33.1, standard annual average rainfall: 1786 mm).

In summary, the Sware Gill and Darling How catchments contrast markedly in the surficial geology with slowly permeable till being present beneath parts of Darling How and absent beneath Sware Gill. The Kirk Stile Formation has been shown to be permeable elsewhere in Cumbria, and its extent differs between the two catchments. However, the hydrological descriptors of base flow index and percentage runoff are very similar. There is a difference in the annual average rainfall based on the Flood Estimation Handbook (Institute of Hydrology, 1999) given their proximity, although this is likely to be an artefact of the strong local hydrological gradients due to the mountainous terrain, and unlikely to be significant, especially when it is considered that the estimates are based on an interpolation of a 1 km grid. All of these differences highlight epistemic uncertainties and illustrate how difficult it is to make inferences in hydrology when no two catchments, even neighbouring micro-catchments within 1 km of each other are the same.

This hydrological gradient could also be responsible for the lower specific runoff generation in Darling How compared with Sware Gill for the two storms generated, although there is no rainfall measurement in Darling How, with the model being driven by the Sware Gill rain-gauge only 1 km to the south west. This results in a key knowledge uncertainty (in the absence of rainfall measurement in both micro-catchments), although for the specific storms there is no knowing if this is realistic, or whether it simply adds additional uncertainty.

2.2 | Macro-basin data for Eden catchment

The much larger Eden catchment investigated here, rises in Black Fell moss, Mallerstang valley and flows south to Carlisle (shown as red circle in Figure 2). Stagnosol and Gleysol soils are extensive in the catchment and derived from the underlying glacial till diamicton. The valley floor is also underlain by deep permeable sandstone aquifers. There is also a strong NE to SW positive gradient in the rainfall between 868 mm to 1548 mm based on the Met Office 1961–1990 averages. There have been numerous recent floods impacting especially Appleby town and Carlisle city, including in 2005, 2009, 2015, 2020, where the peak flow for named storm in (2015) achieved a record level of approximately 1740 m³/s, while Storm Ciara, modelled here, had a peak of 1000 m³/s.

Previous modelling of the Eden catchment has highlighted the potential synchronization issues (Pattison et al., 2014) and studies to understand the resilience of integrated flood risk management measures that include NbS across multiple hydrological extremes (Hankin et al., 2017a). Prioritizing spatial configurations of NbS such as tree-planting in such a large catchment is difficult, given the different pressures and localized conditions. One approach identified here has been to identify all communities at risk from frequent flooding (3.33% annual exceedance probability–AEP) in small catchments (<10 km²)

FIGURE 2 Whole Eden rendered as a projection with complexity of network, the woodland planting scenario (green) and Environment Agency river gauge at Sheepmount shown as a green triangle



Contains OS data © Crown copyright and database right 2021

following co-creation work with environmental NGOs (Hankin, 2020). These watersheds are then used to intersect a national set of maps set up to identify areas with potential for NbS for the Environment Agency (see Hankin et al., 2017b), in particular for advantageous wider woodland planting, which is identified by the presence of tilldiamicton. Strong spatial correlations between till and gleyed soils were found in comparisons with those mapped in detailed 'soil series' maps by the Soil Survey of England and Wales (Hankin et al., 2018), showing till to be synonymous with the presence of slowly-permeable soils, more likely to generate saturation overland flow.

The areas that this approach suggests should be targeted for planting (shown as green in Figure 2) are also constrained by the presence of existing woodland, which to avoid double-counting was removed based on Forestry Inventory (Forest Research, 2020) and OS open woodland data (Ordnance Survey, 2021).

The new 'modelled' woodland is assumed to give rise to changes in a number of hydrological processes from increasing friction to overland and near surface flows (Goudarzi et al., 2021), enhancing infiltration rates due to roots, and enhanced through-storm wet canopy evaporation (this latter has been studied in detail by Page et al., 2020). Rather than modifying multiple Dynamic Topmodel parameters (listed in Section 2.3) as in the approach of Ferguson and Fenner (2020), all these effects are lumped into the shift in the key sensitive parameter, *m*, as a percentage of the independently calibrated value for the whole of the Eden. A 10 m resolution digital terrain model (DTM) was used to process the topographic index for the Eden (the finest scale it was

^{6 of 17} | WILEY-

possible given computational limitations). Whilst only 200 simulations were undertaken for the much larger Eden due to run-time constraints, some strong performance measures were obtained (Nash-Sutcliffe Efficiency—NSE up to 0.9), and a sub-set of these were used for which NSE \geq 0.7 in the larger scale modelling experiment to understand change in hydrological responses.

2.3 | Micro-basin modelling method: Calibrating Dynamic Topmodel in Sware Gill and Darling How to explore parameter shift

Two independent Dynamic Topmodel models were set up for the neighbouring micro-catchments of Sware Gill and Darling How, using 20 classes of topographic index based on an analysis of 2 m resolution DTM, and two spatial classes for rainfall to allow the potential for variable inputs. The seven key Dynamic Topmodel parameters are:

- Lateral saturated transmissivity, In(T0) (m²/h)
- Form of exponential decline in transmissivity, m (m)
- Maximum root storage, srz_{max} (m)
- Initial root storage, srz₀ (m)
- Unsaturated zone time delay, t_d (h/min)
- Overland flow routing velocity, v_{of} (m/h)
- Channel flow routing velocity, v_{ch} (m/h)

The ranges were selected to be relatively wide, uniform priordistributions (table 1 in Metcalfe et al., 2018) and 10 000 Monte-Carlo simulations were undertaken independently for each storm, calibrating against the data from the two micro-catchments collected for the storms in December 2019 and February 2020.

Local rainfall was measured with an RG3 raingauge connected to an RX3000 telemetry unit (Onset Computer Corporation, Bourne, USA). This was telemetered from Sware Gill and compared using double-mass plots with a local Environment Agency gauge 5 km away in the main Cocker valley near Crummock Water, and whilst a reasonable relationship was obtained for the December storm, the Q-NFM rain-gauge was damaged after Storm Ciara, resulting in under-recording and meaning that the subsequent Storm Dennis was not modelled. Furthermore, for both storms a consistent adjustment was made to the rainfall and the minimum potential evapotranspiration (PET min), which is typically estimated at 2 mm for this latitude, and whilst different for the two events (due to malfunction of the single raingauge), these adjustments were the same for both micro-catchments, such that the same net rainfall was used in each event (suspecting that the rainfall measurements had a positive mean bias in the first storm). The maximum potential evapotranspiration (PET max) was set at 9 mm, typical of the maximum at this latitude.

Details of the fuzzy-calibration are as follows, with the corrections to the rainfall and the evapotranspiration, plus the runoff coefficients for the two storms indicating a reasonable mass balance, which is notably stronger for second storm for Darling How.

- December 2019-named Storm Atiyah
 - 5 days warm-up; no correction to rainfall
 - Sware Gill
 - PETmin = 5.5 mm; PETmax = 9 mm
 - Runoff Coefficient = 0.48
 - Darling How
 - PETmin = 5.5 mm; PETmax = 9 mm,
 - Runoff Coefficient = 0.62
- February 2020-named storm Ciara
 - 3 days warm-up; rain factor of 1.8 required due to underrecording for both models
 - Sware
 - PETmin = 2 mm; PETmax = 9 mm
 - Runoff Coefficient = 0.67
 - Darling How
 - PETmin = 2; PETmax = 9
 - Runoff Coefficient = 0.95

Figure 3 shows the range of predictions for a higher-performing model for Sware Gill to the micro-flume telemetered data for the December event. The more complete results are shown in the results section.

Following independent calibration of both catchments a range of 'dotty plots' (where the performance in terms of the NSE is plotted against the parameter value) were developed for the higher performing models, and the differences in these were explored.

A further modelling experiment was then undertaken whereby the resulting contrast in *m* was then applied to a hypothetical planting of the lower third of the Sware Gill, to understand the potential shift in the hydrological response based on the differences in response between the two basins is driven by the difference in landuse, and not due to intrinsic geological differences. This is only a hypothetical assumption to demonstrate the subsequent upscaling of example results from 1 km^2 basins to multiple 10 km^2 basins (watersheds above communities at risk) distributed across a basin >2000 km².

2.4 | Macro-basin modelling method: Calibrating Dynamic Topmodel and model cascade for the Eden

The macro-scale Eden catchment Dynamic Topmodel was set up and calibrated against discharge data from the Environment Agency Sheepmount gauge (NRFA, 2021) for storm Ciara. The rainfall data was based on weighted average of the Great Asby (upper headwaters in south-east), Penrith (mid-catchment), Ullswater (mid catchment, mountainous) and Carlisle (lower western) raingauges. This makes the assumption that there was no hydrological gradient for the large storm under investigation, which could be improved in the future with distributed rainfall.

A 10 m resolution DTM of the Eden was used to divide the large catchment into hydrological similarity units (HSUs) based on:



FIGURE 3 Predicted and observed hydrographs for Sware Gill, storm Atiyah (2019) with 10 and 90 percentile of higher performing model predictions

- 20 classes of topographic wetness index;
- 1 km² grid of distributed rainfall (held at the mean in this experiment);
- An indexing grid of locations for hypothetical conifer planting in areas of slowly permeable soil in small catchments above communities at risk (Metcalfe et al., 2017a).

Monte-Carlo analysis was used at this much larger scale to identify strongly fitting models for the storm sequence that included storms Ciara (as for the micro-catchments) and Dennis and resulted in a peak flow of 1000 m^3 /s the Sheepmount gauge. Strongly performing base-line models with NSE up to 0.9 were identified (see Figure 8), and the *m* parameters HSUs falling within the hypothetical tree-planting grid were later increased based on the shift from the paired basin results for Storm Ciara alone. This makes the tacit assumption that the better performing *m* parameters are specific for this storm and can only be used as a donor for this storm with associated antecedent conditions (and for the same change in landscape).

The model was re-run and the change to the hydrograph at the Sheepmount gauge was explored with reference to models having NSE ≥0.7 with and without hypothetical tree-planting on areas of slowly permeable soils. The strongly performing models typically represent the peak flows well, but with underestimate of the recession, implying a greater groundwater response was needed—which could be relatively significant due to sandstone geology in the lower part of the catchment. The distributed changes in runoff response at the local, river reach scale were also investigated to understand the significant changes in the vicinity of hypothetical new tree-planting.

In a further modelling experiment the distributed streamflow from the top-performing model was fed into a 2D hydrodynamic

model following the hybrid approach of Hankin, Kretzschmar, et al. (2019); Hankin et al. (2019), in part, to enable comparisons of flood extent with new types of data. In this approach, Dynamic Topmodel is used to predict sub-surface and overland flows emerging to every pre-defined river segment (~100 m segments) and the emerging flows are used in an internal inflow boundary to drive the 2D hydrodynamic model, JFlow. The predicted runoff components for all 21 595 river segments were input to the 2D Shallow Water Equation solver, JFlow (Environment Agency, 2013) and routed to allow a better understanding of attenuation, and to permit modelling of hydrodynamic interactions with instream and floodplain including the effects of hypothetical tree planting.

The JFlow model was used to generate maximum depth rasters at 10 m resolution, from which vector flood outlines were derived to permit comparison with satellite remotely sensed flood extents for Storm Ciara. The satellite-based flood extent was based on a changedetection algorithm applied to Sentinel-1 C band synthetic aperture radar (SAR) data. The approach used the ratio of pre-flood (5th February 2020) and post-flood (10th February) images, with a threshold applied to attain the greatest contrast.

To enable further benchmarking of the hybrid modelling approach, the predicted flood extent was also compared with Flood Foresight (Bevington et al., 2019). This system is based on a realtime indexing approach that links the peak flow across a network of flow gauges to the volume on the floodplain computed for a range of pre-simulated scenarios (five design events ranging from 20 year to 1000 year return periods). It has been set up for use in real-time and forecasting modes and produces 30 m resolution depth grids spanning the whole of Great Britain. It therefore produces an interpolated flood-footprint based on linking the rarity of the flow for a

8 of 17 WILEY-

network of telemetered flow gauges (a network of over 1000 gauges across England operated by the Environment Agency are available on-line).

3 | RESULTS

3.1 | Micro-basin modelling results

Using Dynamic Topmodel (Smith, 2020), the new experimental data is used here to independently calibrate the paired catchments

using 10 000 Monte-Carlo simulations, with wide effective parameter ranges, similar to those reported in Metcalfe et al. (2017b) for two winter storms. For the 50 top-performing parameter combinations (having NSE between 0.5 and 0.8), the range of the most sensitive parameter, *m*, were then compared between the two micro-basins (Figure 4). A more constrained superset of results could also be compared for example NSE >0.7, giving a similar results, but all the values are shown in Figure 4 are used in this instance. All the parameters exhibited equifinality, apart from *m* and to a much smaller degree the maximum root-zone storage (*SRmax*).



FIGURE 4 Contrasts in '*m*' parameter between Darling How (conifer) and Sware Gill (moorland) micro-catchments specifically for storm Atiyah (2019) and Ciara (2020) storms



FIGURE 5 Shifts in '*m*' parameter averaged across the two calibration events for Sware Gill catchment (red) and forested Darling How catchment (blue)

The contrasts in the mean value of *m* for both storms are shown in Figure 4 and an average contrast is shown in Figure 5. The NSE values do tend to be better on average for Sware Gill, which could be as a result of the rain gauge being within the Sware Gill catchment and being used by proxy for Darling How. The percentage difference is in relation to the Sware Gill parameter set. To reiterate, if we make a hypothetical assumption that the contrast in the *m* parameter is caused by the woodland enhancing sub-surface movement of water (rather than intrinsic differences arising from contrasting geologies) then we make take the apparent difference and apply to hypothetical target areas for tree planting in the Eden catchment. Storm Ciara was used to illustrate the approach, it being the larger storm in the Eden catchment.

The average percentage change (in the mean value of *m*) was not applied in the larger Eden model, but rather the shift of 53% observed in Figure 4 for Storm Ciara, the damage-causing storm with a peak of 1000 m³/s at Sheepmount. This is somewhat larger than for example the value of 20% increase used in modelling experiments undertaken in Ferguson and Fenner (2020), although none of the other Dynamic Topmodel parameters have been changed here as was the case in that study. Here only the *m* parameter was found to be significantly sensitive when calibrated independently for the same storm in the paired catchments.

However, the shift in the *m* parameter was first applied in another micro-scale modelling experiment as a demonstrator, whereby the lower third of Sware Gill was hypothetically planted with conifers (green area in Figure 6). The HSUs were split to include this area, and the baseline simulation without planting and with planting were then simulated for a top-performing simulation to demonstrate the effect.

The addition of this green area as a realistic planting scenario for this catchment, with a greater m parameter value compared to the

calibrated m for Sware Gill results, as expected in a small reduction in the peak flow (Figure 7).

Clearly this is a hypothetical storm-mitigation scenario for conditions similar to Ciara in this specific location. Figure 7 gives an indication of the potential scale of the reduction of peak flow (17%) for this hypothetical scenario (which also happens to be a similar proportion to the 33% of current conifer extent in Darling How).

3.2 | Macro-scale modelling results

At the macro-scale, the Dynamic Topmodel performance is strong (Figure 8), with the best scenario yielding NSE = 0.9, shown individually to illustrate clearly what happens when the distributed runoff flows in the detailed river reach network are fed in to the 2D hydrodynamic JFlow model (the spread of ensemble of better performing models are shown more clearly in Figure 9 to illustrate uncertainty). There is a clear reduction in the peaks when the hybrid modelling approach is used, only partly accounted for by increased attenuation not included in the routing implicit to Dynamic Topmodel. However, this does not account for the full volume difference, and it is considered that surface water stores/floodplain storage (in areas of concavity) are also being filled and that water is not draining down again in the 2D model domain. This was in part alleviated through setting an initial warm-up or feeder flow in the main river, this representing the base flow in the 2D model domain, and the two significant bodies of water (Haweswater reservoir and the natural lake Ullswater) were modified such that levels were approximately full before the simulation begins. It was not possible to apply the same approach to the numerous smaller lakes and waterbodies so hence some of the missing water will be surface water stores at these locations.

10 of 17 WILEY-



FIGURE 6 Map of the two catchments showing land-cover and inset below, the modification of hydrological similarity units. Sware Gill catchment in red; Darling How forested catchment in blue; telemetered gauges as green triangles, green shaded areas denotes hypothetical tree planting

The 10 m resolution JFlow model maximum depth grid is compared in Figure 9 with the Sentinel-based remotely sensed flood outline and the interpolated footprint generated using Flood Foresight (Bevington et al., 2019).

317500

318000

317000

The left-hand panel of Figure 9 shows the hybrid model of the Eden set up here (Dynamic Topmodel coupled with JFlow) and suggests that the model is over-predicting the flood extent in the vicinity of Carlisle city (at the location of the Sheepmount river gauge), although the satellite image processing is known not to be accurate where there is a lot of vegetation.

Comparing with the right-hand panel, the hybrid model is very similar to the footprint predicted by Flood Foresight (Bevington et al., 2019). Both models are over-predicting the remotely sensed flooding, although it should be noted the more localized flooding that occurred in Appleby 50 km upstream, was not picked up by the remote sensing analysis. Without this level of fine-scale detection (potentially as a result of greater tree-cover), it is concluded that Sentinel data is currently too coarse to place additional significant spatial constraints on the whole-system behaviour for this particular storm at this resolution, but that it may be possible in the future.

Eight of the 200 Monte-Carlo simulations were considered acceptable (NSE \geq 0.7) and were taken forward to the macroscale hypothetical tree-planting experiment. For the tree-planting scenario the calibrated *m* parameter was then shifted and increased by 53% based on the microscale paired findings. The difference between the simulated discharge at Sheepmount and that following the application of an ensample of *m* parameter shifts is given in Figure 10. The first peak represents Storm Ciara, having a monitored peak flow of 1000 m³/s, implying a range of 0.5–5% peak flow reduction (1%

FIGURE 7 Reduction in hydrograph peak from hypothetical tree-planting scenario





FIGURE 8 Sample calibration for Dynamic Topmodel and hybrid models for the whole Eden

median) across the ensemble, arising from the targeted planting in small watersheds above communities at risk. A full GLUE analysis (Beven & Binley, 1992, 2014), that would include allowing for

uncertainty in shifted parameter values, has not been undertaken here, but the range in predicted change hydrographs gives a good indication of uncertainty across the ensemble.

12 of 17 WILEY-



FIGURE 9 Spatio-temporal calibration: Comparison of hybrid model outputs with flood foresight and the Sentinel flood footprint



FIGURE 10 Difference between baseline and with-trees scenario

WILEY 13 of 17

This range demonstrates that the model uncertainty is large, but potentially also that the distributed changes to *m* are sometimes interacting 'constructively' or 'destructively' in this very large network of 21 595 river segments. This is for a single damage-causing storm, and a single tree-planting strategy summarized as planting on slowly permeable soils in small catchments above communities at risk of frequent flooding, where based on open-data there is not already planting. The huge number of combinations of policy (spatial deployment strategy)/storm size and shape/model uncertainty leads to the conclusion that it may instead be more effective to focus on the multi-local scale-to those small at-risk communities (shown in pink in Figure 11), where the modelled changes are relatively significant. The modelling experiment indicates that that it is going to be very difficult, even with modern remotely sensed data, to objectively constrain a model of a large catchment sufficiently to demonstrate the integrated impact of many distributed NFM interventions.

Figure 11 shows how the peak runoff during storm Ciara changes spatially at the river segment scale, and segments with an increase in peak flow of >0.25 m³/s are highlighted in red. These are significant changes between the scenarios at this scale.

There also appears to be some interaction here between the scale of HSUs and the resolution of the woodland planting shown in green, for which the *m* values have been uplifted for Storm Ciara.

4 | DISCUSSION

The discussion section refers back to the numbered research questions in parenthesis. The main premise of the paper has been to understand the potential for upscaling changes in effective parameters calibrated for micro-basins to much larger scales using a process of donor-parameter-shift (1). Changes to multiple processes are 'wrapped-up' in the shift in the Dynamic Topmodel *m* parameter, yet without detailed investigations it is impossible to know what has caused the contrasting between the two study basins (2). The modelling therefore is a hypothetical experiment in order to demonstrate the donor-parameter-shift approach, and to explore more of the implications when modelling much larger scales. The approach was also used in Hankin et al. (2017) in a more complex uncertainty framework whereby stratified sampling was undertaken of the fuzzy shift applied to a fuzzy set of better performing baseline (pre-NbS) scenarios. With the large run-time of the Eden model (48 h), the mean donor-shift (as opposed to sampling from a range of weighted shifts) has been applied across an ensemble of eight strongly performing models for a single large flood event (3).

The complexity of the interactions and possible changes in synchronization, resulting from different modelled distributed hydrological responses interacting differently, becomes apparent at this large scale especially from Figure 10 where very different changes are predicted in the ensemble of top performing models. This is despite seeing relatively significant changes at the local scale (0.25 m³/s in many river segments draining catchments 10 km²), which without destructive interference,

would add to reductions in peak flows many times greater (estimated 50 m³/s) than the median peak reduction (15 m³/s) seen in Figure 10 for storm Ciara (the first peak). The large-scale parameterizations are difficult to constrain even using modern satellite detected data (4), but potentially represents a problem domain where local changes to the modelled scenario could be considered in more detail to reduce risk at the impacted communities. This multi-local scale seems to offer a better posed optimisation problem, such that if peak flows are observed to be reduced locally then it is already known that there will be a local benefit (5), and there is little point looking beyond this at the whole catchment scale where modelling the synchronization effects becomes intractable and difficult to constrain even with current advances in remote sensing. The argument requires further evidence, which will hopefully stem from the 18 micro-catchments under investigation, where it is hoped more evidence for parameter shifts from NFM measures can be identified (1, 2, 3), and such evidence is crucial for taking the donor parametershift approach further

In an attempt to delineate some of the process uncertainty here. a modelling experiment was also investigated using a spatially more simplistic rainfall-runoff model, based on the probability moisture distribution approach of Moore (1985) called HYMOD (e.g., see Quan et al., 2015), which also differs from Dynamic Topmodel as it is lumped, rather than spatially distributed. This was used to help understand whether a pure increase in hydrological losses (e.g., due to enhanced wet-canopy evaporation) from planting trees could also result in a shift to the key parameter controlling recession (m). HYMOD was calibrated to storm Ativah (2019) for Sware Gill, and the resulting fitted flow response was then reduced by 20% to emulate increased losses associated with woodland for high elevation, nonsaturated conditions (see Page et al., 2020). This reduced response was then used to re-calibrate the model with the same rainfall, to see if the HYMOD parameter controlling recession shifted significantly. Figure 12 shows this is the case based on 10 000 simulations.

This suggests that the effects of a pure loss can manifest itself as a significant shift to the key parameters controlling recession. However, on investigating the same approach applied to Dynamic Topmodel for the same Sware Gill micro-catchment, no significant shift in the *m* parameter was detected, in fact there was no trend and this implied that the shift is more complex than just a change to hydrological losses (due to wet-canopy evaporation), and impacts the soil storage and recession in a more complex way. Furthermore, the larger flood runoff coefficients for the conifer-covered Darling How basin presented in Section 2.3 is not consistent with it having a larger rate of wet-canopy evaporation compared to the moorland Sware Gill catchment. This again points to intrinsic geological differences between the two catchments giving contrasting water storage dynamics in the surficial and solid geology of the micro-catchments.

The modelled shift in *m* was applied to the macro-scale, and the model outputs compared through time at the furthest downstream gauge and then spatially using new Sentinel C-band data (4). The ensemble of predicted hydrograph differences is very wide, and the median reduction in the peak is relatively small. In reality, neither



FIGURE 11 Differences between baseline and with-trees scenario-Red is a significant increase in runoff generation >0.25 m³/s per segment

dataset provides satisfactory 'additional' constraint to the model parameterisation, although using multiple datasets can help improve confidence in the baseline model.

The problem at the large scale is that to properly appraise the net impact of millions of permutations of: distributed NbS interventions and spatial strategy; variability of extreme weather; and performance



FIGURE 12 Shift in exponent controlling recession in HYMOD– Black is baseline, red includes losses

failure—requires enormous computational resources, and the resulting model uncertainty cannot necessarily be constrained even considering just key modes of whole-system behaviour (demonstrated in Hankin et al., 2019). More local, distributed measurements closer to the interventions are needed to identify the donor-shifts and constrain which of the ensemble model predictions are realistic.

The conclusion is that analysis of effectiveness should focus on the multi-local scale where observed contrasts in behaviour might be observed in response to targeted interventions and evidence for shifts at scales commensurate with the micro-catchments where the parameter shifts are evidenced (6). This could in turn help with the scaling issue, and intrinsic hydrological non-linearities whereby modelled shifts in parameters at different scales will likely be different. There will of course be a need to model larger parts of a catchment to develop larger-scale storage solutions, such as broad-scale floodplain re-connection, but the risk for these can also be understood at the reach-scale with a detailed hydrodynamic investigation.

The multi-local scale could be an important principle for the many applications of NbS that are being currently being implemented (e.g., see 60 case studies in Burgess-Gamble et al., 2017), without significant 'model-driven design', informed by credible distributed hydrological modelling. This can partly be addressed by focussing initially on these local communities at risk, but also if more donor-shifts from accurate paired plot experiments can be generalized more. An example of this are the changes to model effective roughness (using Mannings n) due to local re-vegetation (e.g., see Addy & Wilkinson, 2019; Shuttleworth et al., 2018) across a variety of landscapes.

5 | CONCLUSIONS

This paper demonstrates applying donor-parameter-shifts from effective model parameters based on paired micro-catchments set-up to quantify changes in hydrological responses from different types of NbS measures, in this case specifically for the deliverable of flood hydrograph reductions (NFM). The detected shifts are place- and storm- specific and based on the independent fitting of the distributed hydrological rainfall-runoff model, Dynamic Topmodel to new, accurate monitoring data. Applying shifts in the parameter controlling the rate of decline of the downslope transmissivity controlling sub-surface flows to a much larger catchment and exploring the outputs using modelling experiments has led to a number of conclusions which will require further evidence.

- Using models to scale micro-scale findings on how NbS shifts hydrological response at the macro-scale is still limited by the large uncertainties in non-linear scaling of hydrological processes, and the complexity of whole system interactions as the river network increases in size, and the signal from our changes is lost in the inevitable environmental variability.
- It is not surprising that there is some evidence to suggest NbS measures can be effective at the small scale (<10 km²), but much less at larger scales (>10 km²), simply because testing and constraining the whole system response with large input-errors and knowledge uncertainties becomes intractable, and because we lack direct, high quality observations of hydrological change in the few global locations where wholescale landscape change has taken place.
- Models of natural systems will always be both imperfect descriptions of reality and predictors of environmental behaviour (Beck et al., 1993), but they are better and more constrained by directly observed evidence, more widely available at smaller catchment scales.
- Local deployment of NbS in very few locations of a large catchment is by definition only going to reduce flood risk locally for communities affected by small streams. Therefore, we should have a more nuanced focus at the multi-local scale, not at the macroscale, partially to side-step the scaling problem, and partially to remove the need for fully testing whole system response in larger catchments for huge numbers of permutations. This moves the focus into communities at risk in small catchments where there is strong evidence for NbS being more effective.

ACKNOWLEDGEMENTS

This project has been supported by NERC grant NE/R004722/1 and has been a collaborative effort between JBA Consulting and Lancaster University in the UK. The Environment Agency is thanked for the use of rainfall and terrain data under licence CL77737MG. The authors thanks West Cumbria Rivers Trust, David Kennedy and colleagues at the Environment Agency; lain Craigen, Andrew Fielding for JFlow modelling, Doug Pender for Sentinel C band analysis, Gareth McShane for his work on the hydrometry in these difficult environments.

DATA AVAILABILITY STATEMENT

The data presented here will be made available on the NERC datastore, along with models and predictions. Key datasets used and

the versions of Dynamic Topmodel have been referenced, although the open-source model development is still on-going (Smith, 2020).

ORCID

Barry Hankin D https://orcid.org/0000-0001-7315-3321 Trevor J. C. Page D https://orcid.org/0000-0002-1684-6049 Keith J. Beven D https://orcid.org/0000-0001-7465-3934

REFERENCES

- Addy, S., & Wilkinson, M. (2019). Representing natural and artificial inchannel large wood in numerical hydraulic and hydrological models. WIREs Water, 6(6), 1–20. https://doi.org/10.1002/wat2.1389
- Beck., M. B., Jakeman, A. J., & McAleer, M. J. (1993). Construction and evaluation of models of environmental systems. In A. J. Jakeman, M. B. Beck & M. J. McAleer (Eds.), *Modeling change in environmental systems*. John Wiley and Sons Ltd. ISBN 0-471-94263-4
- Beven, K. (2001b). How far can we go in distributed hydrological modelling? Hydrology and Earth System Sciences, 5(1), 1–12. https://doi.org/ 10.5194/hess-5-1-2001
- Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6, 279–298.
- Beven, K. J. (2001a). Rainfall-runoff modelling: the primer. Wiley.
- Beven, K. J., & Binley, A. (2014). GLUE: 20 years on. *Hydrological Processes*, 28, 5897–5918.
- Beven, K. J., & Freer, J. (2001). A dynamic topmodel. Hydrological Processes, 15(10), 1993–2011.
- Bevington, J., Hankin, B., Berry, R., Gubbin, A., Lowis, A., Pinnell, M., Revilla-Romero, B., Shelton, K., & Wood, E. (2019). Flood Foresight: Global Flood Impact Forecasting. *Geophysical Research Abstracts*, 21.
- BGS (2021). Retrieved from http://mapapps.bgs.ac.uk/geologyofbritain/ home.html
- Blöschl, G. (2006). Rainfall-runoff modeling of ungauged catchments. In M. G. Anderson & J. J. McDonnell (Eds.), *Encyclopedia of hydrological sciences*. Wiley. https://doi.org/10.1002/0470848944.hsa140
- Bogaart, P. W., van der Velde, Y., Lyon, S. W., & Dekker, S. C. (2016). Streamflow recession patterns can help unravel the role of climate and humans in landscape co-evolution. *Hydrology and Earth System Sciences*, 20(4), 1413–1432.
- Bridges, T. S., Bourne, E. M., King, J. K., Kuzmitski, H. K., Moynihan, E. B., & Suedel, B. C. (2018). *Engineering with nature: An atlas. ERDC/EL SR-18-8*. U.S. Army Engineer Research and Development Center. https://doi.org/10.21079/11681/27929
- Bulygina, N., Ballard, C., McIntyre, N., O'Donnell, G., & Wheater, H. (2012). Integrating different types of information into hydrological model parameter estimation: Application to ungauged catchments and land use scenario analysis. *Water Resources Research*, 48(6), W06519. https://doi.org/10.1029/2011WR011207
- Burgess-Gamble, L., Ngai, R., Wilkinson, M., Nisbet, T., Pontee, N., Harvey, R., Kipling, K., Addy, S., Rose, S., Maslen, S., Jay, H., Nicholson, A., Page, T., Jonczyk, J., & Quin, P. (2017). Working with natural processes—Evidence directory. Environment Agency Project. SC150005.

Centre for Ecology & Hydrology. (1999). Flood Estimation Handbook.

- Dadson, S., Hall, J. W., Murgatroyd, A., Acreman, M., Bates, P., Beven, K., Heathwaite, L., Holden, J., Holman, I. P., Lane, S. N., O'Connell, E., Penning-Rowsell, E., Reynard, N., Sear, D., Thorne, C., & Wilby, R. (2017). A restatement of the natural science evidence concerning catchment-based natural flood management in the UK. *Proc. R. Soc. A.*, 473, 20160706. https://doi.org/10.1098/rspa.2016.0706
- Environment Agency. (2013). Benchmarking the latest generation of 2D hydraulic modelling packages, Report No. SC120002. Retrieved from https://consult.environment-agency.gov.uk/engagement/

bostonbarriertwao/results/appendix-6-neelz-s.-pender-g.-2013-bench marking-the-latest-generation-of-2d-hydraulic-modelling-packages.-br istol_environment-agency.pdf

- European Commission (2020). Nature-based solutions: State of the art in EU-funded projects. European Commission B-1049 Brussels. https://doi.org/10.2777/236007 KI-04-20-168-EN-N.
- Ferguson, C. R., & Fenner, R. A. (2020). The potential for natural flood management to maintain free discharge at urban drainage outfalls. *Journal of Flood Risk Management*, 13(3), e12617. https://doi.org/10. 1111/jfr3.12617
- Forest Research. (2020). National Forest Inventory electronic data download. Retrieved from https://www.forestresearch.gov.uk/tools-andresources/national-forest-inventory/.
- Goudarzi, S., Milledge, D. G., Holden, J., Evans, M. G., Allott, T. E. H., Shuttleworth, E. L., Pilkington, M., & Walker, J. (2021). Blanket Peat Restoration: Numerical Study of the Underlying Processes Delivering Natural Flood Management Benefits. *Water Resources Research*, 57, 4. https://doi.org/10.1029/2020WR029209
- Hankin, B. (2020). NbS Modelling: Developing collaborative end user scenarios. Rivers Trust Catchment Data & Evidence Forum. Retrieved from https://catchmentbasedapproach.org/learn/catchment-dataand-evidence-forum-2020/
- Hankin, B., Chappell, N.A., Page, T.J.C., Kipling, K., Whitling, M., & Burgess-Gamble, L. (2018). Mapping the potential for working with natural processes—Technical report. SC150005/R6. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/ uploads/attachment_data/file/677592/Working_with_natural_proces ses_mapping_technical_report.pdf.
- Hankin, B., Hewitt, I., Sander, G., Danieli, F., Formetta, G., Kamilova, A., Kretzschmar, A., Kiradjiev, K., Wong, C., Pegler, S., & Lamb, R. (2020). A risk-based, network analysis of distributed in-stream leaky barriers for flood risk management. *Natural Hazards and Earth System Sciences*, 20, 2567–2584. https://doi.org/10.5194/nhess-20-2567-2020
- Hankin, B., Kretzschmar, A., Page, P., Chappell, N., Lamb, R., Beven, K. Coxon, G., & Freer, J. (2019). Efficient cascade modelling of naturebased solutions scaled to larger catchments to model extremes taking account of uncertainties. Retrieved from https://meetingorganizer. copernicus.org/EGU2019/EGU2019-1024.pdf
- Hankin, B., Lamb, R., Craigen, I., Page, T., Chappell, N. & Metcalfe, P. (2017a). A whole catchment approach to improve flood resilience in the Eden. Winning entry to the Defra Flood Modelling Competition 2016 [Internet]. Retrieved from https://consult.defra.gov.uk/waterand-flood-risk-management/flood-risk-management-modelling-compe tition/results/jba_defra_winning_entry_full_report.pdf
- Hankin, B., Metcalfe, P., Craigen, I., Page, T., Chappell, N., Lamb, R., Beven, K., & Johnson, D. (2017b). Strategies for testing the impact of natural flood risk management measures. *Flood Risk Management*. IntechOpen.
- Hankin, B. G., Metcalfe, P., Beven, K., & Chappell, N. (2019). Integration of hillslope hydrology and 2D hydraulic modelling for natural flood management. *Hydrology Research*, 50, 1535–1548. https://doi.org/10. 2166/nh.2019.150
- Heuvelmans, G., Muys, B., & Feyen, J. (2004). Evaluation of hydrological model parameter transferability for simulating the impact of land use on catchment hydrology. *Physics and Chemistry of the Earth, Parts a*/*B*/*C*, 29(11–12), 739–747.
- Karvonen, T., Koivusalo, H., Jauhiainen, M., Palko, J., & Weppling, K. (1999). A hydrological model for predicting runoff from different land use areas. *Journal of Hydrology*, 217(3–4), 253–265.
- Lane, S. N. (2017). Natural flood management. Wiley Interdisciplinary Reviews: Water, 4(3), e1211.
- Metcalfe P., Beven, K., & Freer, J. (2016). Dynatopmodel: Implementation of the dynamic TOPMODEL hydrological model. R package version 1.1. Retrieved from https://CRAN.R-project.org/package= dynatopmodel

- Metcalfe, P., Beven, K., Hankin, B., & Lamb, R. (2017a). Simplified representation of runoff attenuation features within analysis of the hydrological performance of a natural flood management scheme. *Hydrology* and Earth System Sciences, 22, 2589–2605. https://doi.org/10.5194/ hess-22-2589-2018
- Metcalfe, P., Beven, K., Hankin, B., & Lamb, R. (2017b). A modelling framework for evaluation of the hydrological impacts of nature-based approaches to flood risk management, with application to in-channel interventions across a 29km² scale catchment in the United Kingdom. *Hydrological Processes*, 31, 1734–1748. https://doi.org/10.1002/hyp. 11140
- Metcalfe, P., Beven, K., Hankin, B., & Lamb, R. (2018). A new method, with application, for analysis of the impacts on flood risk of widely distributed enhanced hillslope storage. *Hydrology and Earth System Sciences*, 22(4), 2589–2605, 17 p. https://hess.copernicus.org/articles/22/2589/2018/
- Moore, R. J. (1985). The probability-distributed principle and runoff production at point and basin scales. *Hydrological Science Journal*, 30(2), 273–297.
- Myers, D. T., Ficklin, D. L., Robeson, S. M., Neupane, R. P., Botero-Acosta, A., & Avellaneda, P. M. (2021). Choosing an arbitrary calibration period for hydrologic models: How much does it influence water balance simulations? *Hydrological Processes*, 35(2), e14045.
- NERC. (2020). Retrieved from https://nerc.ukri.org/research/funded/ programmes/NbS/
- NRFA. (2021). Retrieved from https://nrfa.ceh.ac.uk/data/station/info/ 76007.html.
- Ordnance Survey. (2021). Retrieved from https://www.ordnancesurvey. co.uk/business-government/tools-support/open-data-support.
- Page, T., Beven, K., Chappell, N., Kretzschmar, A., & Hankin, B. (2020). Assessing the significance of wet-canopy evaporation from forests

during extreme rainfall events for flood mitigation in mountainous regions of the UK. *Hydrological Processes. Aug*, 27, 13895. https://doi. org/10.1002/hyp.13895

- Pattison, I., Lane, S. N., Hardy, R. J., & Reaney, S. M. (2014). The role of tributary relative timing and sequencing in controlling large floods. *Water Resources Research*, 50, 5444–5458. https://doi.org/10.1002/ 2013WR014067
- Quan, Z., Teng, J., Sun, W., Cheng, T., & Zhang, J., (2015). Evaluation of the HYMOD model for rainfall-runoff simulation using the GLUE method in Remote Sensing and GIS for Hydrology and Water Resources (IAHS Publ. 368, 2015). (Proceedings RSHS14 and ICGRHWE14, Guangzhou, China, August 2014). https://doi.org/10. 5194/piahs-368-180-2015
- Shuttleworth, E. L., Evans, M. G., Pilkington, M., Spencer, T., Walker, J., Milledge, D., & Allott, T. E. H. (2018). Restoration of blanket peat moorland delays stormflow from hillslopes and reduces peak discharge. *Journal of Hydrology*, 2, 100006. https://doi.org/10.1016/j.hydroa.2018.100006
- Smith, P. (2020). Retrieved from https://waternumbers.github.io/ dynatopGIS/index.html.

How to cite this article: Hankin, B., Page, T. J. C., 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. <u>https://doi.org/10.1002/hyp.</u> 14418