Journal of Hydrology 628 (2024) 130449

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Quantifying the natural flood management potential of leaky dams in upland catchments, Part II: Leaky dam impacts on flood peak magnitude

Z.R. van Leeuwen, M.J. Klaar, M.W. Smith, L.E. Brown *

School of Geography & water@leeds, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom

A R T I C L E I N F O

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Szilágyi József, Associate Editor.

Keywords:
Nature based solutions
Large wood
Empirical
Hydrograph analysis
Ecosystem services
Transfer function noise model

A B S T R A C T

Leaky dams are an increasingly popular natural flood management measure, yet their impacts on flood peak magnitude have not yet been empirically quantified for a range of event types and magnitudes, even at the stream scale. In this study, the novel application of a transfer function noise modelling approach to empirical Before-After-Control-Impact stage data from an upland catchment allowed leaky dam effectiveness in reducing flood peak magnitude to be quantified. Flood peak stage and discharge magnitude changes were assessed from empirical data for 50 single and multi-peaked high flow events with return periods ranging from less than one year to six years. Overall, event peak magnitude was significantly reduced following the installation of eight leaky dams on the impact stream. Effectiveness was highly variable, but on average, flood peak magnitude was reduced by 10% for events with a return period up to one year. Some of the variability was explained by the size of the event and whether it was a single or multi-peaked event. This finding emphasises the need to manage expectations by considering both a range of event magnitudes and types when designing or assessing leaky dam natural flood management schemes.

1. Introduction

In response to increased flooding across Europe (Blöschl et al., 2019) there has been a shift towards a holistic, catchment wide approach to flood risk management (Commission of the European Communities, 2009). Managing flood hazard by working with natural processes, often referred to as Natural Flood Management (NFM), aims to restore or emulate the natural functioning of river catchments to increase infiltration, slow flows and store water (Forbes et al., 2015). NFM is used alongside traditional flood defences and is particularly suited to sparsely populated rural and upland areas where traditional flood defence schemes are less feasible (Sayers et al., 2002). NFM is increasingly prevalent, especially in the Global North (Sudmeier-Rieux et al., 2021), although there are examples of its implementation worldwide (Thaler et al., 2023; Iacob et al., 2012). In the UK, NFM is a popular flood risk management measure because of its relatively low cost (Burgess-Gamble et al., 2017; Kail et al., 2007; Bark et al., 2021), often community-led approach (Garvey and Paavola, 2022; Environment Agency, 2019), and multiple benefits, which range from ecological to cultural (Burgess-Gamble et al., 2017). NFM has been considered in the UK's approach to flood risk management for almost two decades (Wilby et al., 2008), and has been evident in flood risk management policy since 2005 (Defra, 2005); yet the efficacy of many NFM measures at reducing downstream flood risk remains unquantified (Burgess-Gamble et al., 2017; Lane, 2017; Wilkinson et al., 2019).

Leaky dams are one of many NFM measures which have been used in streams across the UK (Nisbet et al., 2015; National Trust, 2015; Uttley and Skinner, 2017; Dodd et al., 2016; Hester et al., 2016; Lavers et al., 2022; Black et al., 2021). In England and Wales, their installation for the purpose of flood risk management in uplands is incentivised through agricultural subsidies (Defra et al., 2016), and is likely to form part of the government's reformed agricultural and land management policy following its exit from the European Union (Klaar et al., 2020). The proposed Environmental Land Management Scheme (ELMS) provides financial compensation for landowners who provide public goods, including flood risk management (Defra, 2020). Successful implementation of the "public money for public goods" approach in the UK would lead to greater adoption of NFM features such as leaky dams.

Leaky dams consist of wood placed in the river channel and on the river banks to mimic the function of natural accumulations of large wood in rivers. Large wood has been widely used in river restoration for erosion control and its benefits to aquatic species (Kail et al., 2007; Abbe et al., 2003; Bernhardt et al., 2005; Addy and Wilkinson, 2016). Recently, leaky dams have been trialled as a measure for

* Corresponding author.
E-mail addresses: zoravanleeuwen@gmail.com (Z.R. van Leeuwen), L.Brown@leeds.ac.uk (L.E. Brown).

https://doi.org/10.1016/j.jhydrol.2023.130449
Received 12 May 2023; Received in revised form 25 September 2023; Accepted 17 October 2023
Available online 14 November 2023

NFM (Grabowski et al., 2019) and its design has been adapted for the primary purpose of delaying and reducing the magnitude of flood events. Leaky dams differ from typical river restoration large wood installations in that they are usually perpendicular to the direction of flow, span the channel width, extend onto floodplains, and are often raised above the channel bed to allow for fish passage (Nisbet et al., 2015; National Trust, 2015; Uttley and Skinner, 2017; Dodd et al., 2016; Hester et al., 2016; Abbe and Brooks, 2011). Leaky dams can range from a few to hundreds of interventions in a catchment (The Rivers Trust, 2021) and have varied designs, either replicating natural accumulations or taking a uniform, engineered approach (Kail et al., 2007).

Leaky dams are thought to delay the flood peak and reduce its magnitude because they are known to locally increase hydraulic roughness, decrease flow velocities and increase water levels (Curran and Wohl, 2003; Shields and Gippel, 1995) which has been shown to increase flood wave travel time (Black et al., 2021; Gregory et al., 1985; Kitts, 2010) and floodplain connectivity (Keys et al., 2018; Sear et al., 2010). However, there are few empirical studies which have successfully quantified the impacts of leaky dams in upland streams on flood peak magnitude (Addy and Wilkinson, 2019; Burgess-Gamble et al., 2017). Quantitative evidence of their impacts on flood peak magnitude is important for a number of reasons. Firstly, successful implementation of NFM schemes requires the buy-in of a wide range of stakeholders (Wingfield et al., 2019). Uncertainty in the effects of NFM measures on downstream flood risk currently undermines confidence in their uptake, which limits its adoption (Bark et al., 2021; Waylen et al., 2018; Wingfield et al., 2019). Secondly, without quantitative evidence of their impacts on flood peak magnitude, cost–benefit assessment, needed by the UK Government to fund flood risk management activities (Defra, 2009) is problematic (Lavers et al., 2022). Similarly, understanding the efficacy of NFM measures is important for the design of schemes to a desired level of protection (Defra, 2009). Finally, whilst communities are warned not to rely on NFM as a 'silver bullet' e.g. Dadson et al. (2017) and Wells et al. (2020), robust quantification of the effectiveness of NFM measures is needed to manage expectations of NFM efficacy. Managing expectations has been identified repeatedly as key to sustaining efforts to integrate NFM, and is essential to avoid placing communities inadvertently at greater risk of flooding (Collentine and Futter, 2018; Nisbet et al., 2015; Wells et al., 2020).

Empirical evidence of in-stream wood impacts on flood peak magnitude is limited to river restoration style in-stream wood in small streams (bankfull channel width <1 m) and for relatively few, artificial flood peaks, generated from reservoir releases (Wenzel et al., 2014; Keys et al., 2018). In the Ore Mountains in south eastern Germany, the placement of nine spruce tops (average length 8 m, average maximum trunk diameter 0.2 m) longitudinally in a 282 m first order, headwater stream reach (gradient 3.7%, width 0.8 m and average flow depth 0.3 m) reduced the flood peak by 2.2% (Wenzel et al., 2014). Three pieces of large wood with their rootwad facing upstream in a 50 m reach of a headwater stream in the mid-Atlantic region of the United States reduced peak magnitude by 8% for an artificial <1-in-1 year flood event (Keys et al., 2018). Although these studies provide some insights, they are limited in their applicability to larger streams and different types of in-stream wood structures. Moreover, it is essential to consider the range of flood peak magnitudes and the type of storm event (single or multi-peaked) to accurately assess the impact of leaky dams. Recent modelling efforts suggest that the effectiveness of leaky dams can vary significantly depending on the time required for the system to recover between peaks in multi-peaked events (Metcalfe et al., 2017). Therefore, it is necessary to examine the impact of leaky dams on a range of event types and magnitudes to understand their efficacy in reducing flood peak magnitude.

Whilst there is an emphasis on the need for gathering catchment scale evidence of NFM impacts e.g. Lane (2017) and Dadson et al. (2017), leaky dam impacts in small streams (catchment area <1 km^2) are relevant considering around half the flood risk in England is attributed to these streams (Beven et al., 2022; Hankin et al., 2021; Environment Agency, 2009). Accordingly, a multi-local scale modelling approach has been proposed to overcome the difficulties of assessing catchment scale impacts of NFM (Hankin et al., 2021; Beven et al., 2022). This approach is informed by evidence of leaky dam impacts at <1 km^2 sub-catchments, for small basin scales (<10 km^2) upstream of communities at risk of flooding (Hankin et al., 2021; Beven et al., 2022). Due to a lack of such empirical evidence, the representation of leaky dams in hydraulic and hydrological models at all spatial scales has thus far been heuristic, which undermines confidence in their outputs (Addy and Wilkinson, 2019). This further emphasises the need for empirical evidence of leaky dam impacts for small sub-catchments.

Robust Before After Control Impact (BACI) style empirical monitoring has been proposed to address the lack of empirical evidence of NFM impacts on flood peak magnitude (Ellis et al., 2021; Burgess-Gamble et al., 2017). However, the BACI approach has not yet been able to overcome difficulties associated with high levels of uncertainty in hydrological data, typically short periods of baseline data and the stochastic nature of flood events (Connelly et al., 2020; Ellis et al., 2021; Lane, 2017; Black et al., 2021). Even where large-scale, long-term monitoring studies have succeeded in quantifying impacts of leaky dams on flood peak timing, high levels of uncertainty have precluded assessment of their impacts on flood peak magnitude (Black et al., 2021; Kitts, 2010). To address this knowledge gap, van Leeuwen et al. (2023) demonstrated the opportunity offered by top-down, data-based time series modelling techniques to overcome the difficulties associated with detecting small changes in empirical hydrological data. This approach allowed accurate (±2cm) simulations of stream response to be made even when the models were based on relatively short and uncertain baseline data. van Leeuwen et al. (2023) details how these Transfer function Noise (TFN) models were developed and how confidence in their outputs was assessed. Whilst poorer model performance indicated a different class of model would be needed to appropriately represent baseline flow conditions on one of the three studied streams, there was a high level of confidence in the ability of the models to simulate baseline conditions on the remaining streams. This study aims to use these empirical data-based models to quantify, for the first time, the impact of upland leaky dams on the peak magnitude of a range of high flow events. To do this, simulations of stream baseline response, made using the data-based time series models developed in van Leeuwen et al. (2023), are compared to observations of the stream response after eight leaky dams were installed in the stream. Comparisons of the stream response were made for 50 storm events which exceeded the minimum stage threshold for interaction with the leaky dams.

2. Methods

Leaky dam impacts were derived from data collected during a BACI style monitoring study in a UK headwater catchment. The study consists of two parts: in part I, data-based time series models were developed which were able to accurately simulate the baseline stage response of two hydraulically similar streams (van Leeuwen et al., 2023). In this part of the study, simulations made using these models were compared to observations of stage after eight leaky dams were installed in one of the two streams. By comparing observed stage during high flow events to simulations of what the stage would have been during the event had no leaky dams been installed, the impact of the leaky dams on the flood hydrograph could be assessed in the same way that flow and river water temperature responses to perturbations were analysed by Watson et al. (2001), Gomi et al. (2006), Dickson et al. (2012) & O'Driscoll et al. (2016).

The following steps were taken to assess the impacts of leaky dams on event peak magnitude: (1) Identification of high flow events in upstream, post-intervention stage series; (2) Simulation of downstream baseline stage response to upstream event stage time series using

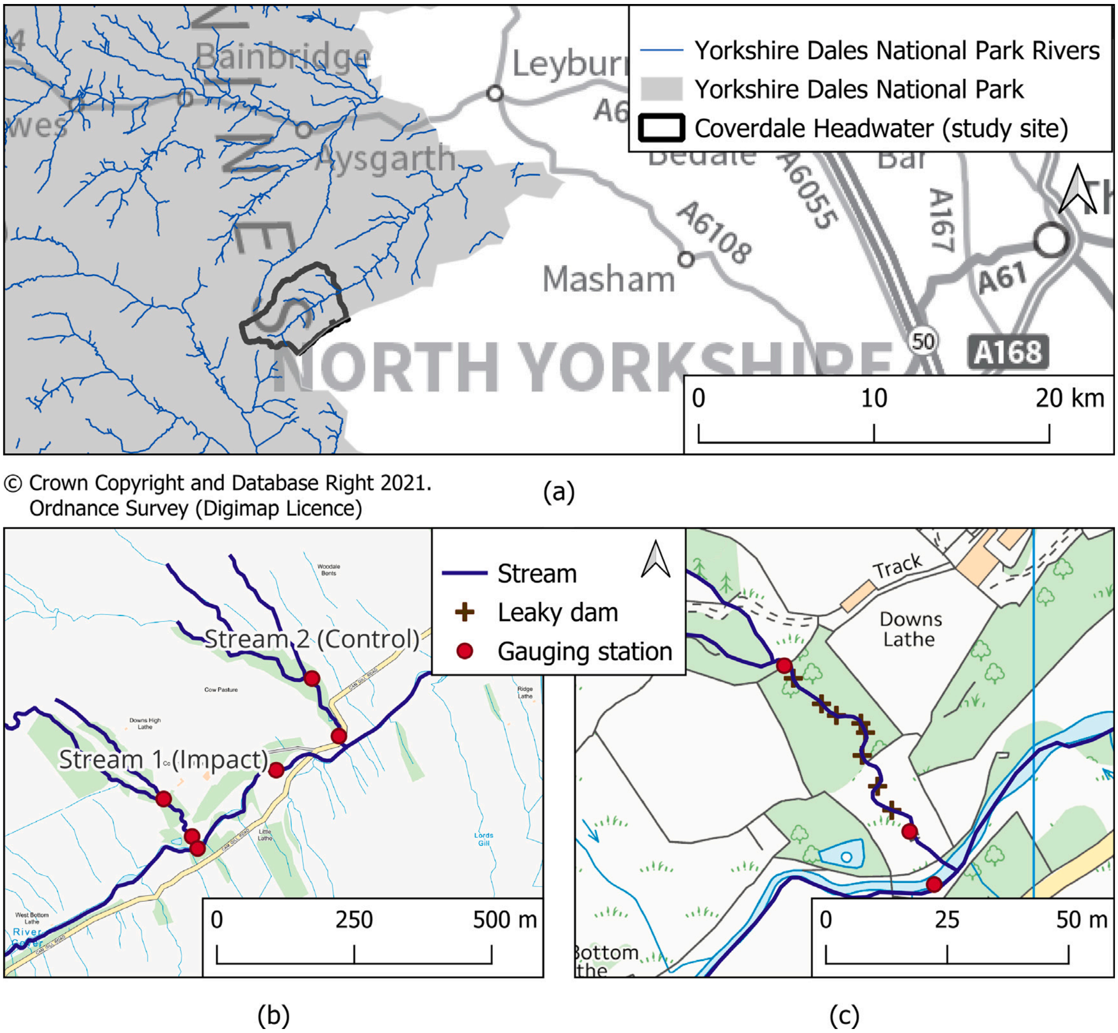

Fig. 1. (a) Location of study site in Coverdale, North Yorkshire, UK (b) Water level/flow gauging network in Coverdale catchment (c) Location of leaky dams on the impact stream (Stream 1).

models developed in van Leeuwen et al. (2023); (3) Calculation of differences between simulated and observed downstream stage for both baseline and post-intervention monitoring periods on the impact and control stream; (4) Testing for statistical differences between the simulated and observed response in the baseline and post-intervention periods; (5) Analysis of the impact of high flow event characteristics such as peak magnitude, duration and the number of peaks in the event on leaky dam effectiveness.

Confidence in attributing changes in flood peak magnitude to the leaky dams relied on the assumption that there were no other changes in the impact stream or its gauging stations. This assumption was backed up by the control stream, a thorough quality assurance process and local knowledge of changes which could affect the hydrology of individual streams, such as livestock density or harvesting of the commercial riparian forest on the streams, gained through frequent field observations and correspondence with the landowner.

2.1. Site description

The study site was located in the headwaters of the River Cover (54.20045°N, −1.98617°E), North Yorkshire, England (Fig. 1) and is described in van Leeuwen et al. (2023). This study focused on two small, parallel watercourses which form part of the headwaters of the River Cover; an impact stream (Stream 1) and a control stream (Stream 2) (Fig. 1) with similar hydrological characteristics (Table 1)

Table 1
Characteristics of the study streams.

Stream	Gradient (m/m)	Catchment area (km^2)	Monitored length (m)	Mean width (m)	leaky dams (count)
1 (impact)	0.13	1.1	280	2.6	8
2 (control)	0.11	1.9	260	3.0	0

and no established lateral inflows within the monitored reaches. The watercourses were of type A in the Rosgen classification; steep, partially entrenched and cascading with step/pool streams (Rosgen, 1994).

Eight leaky dams (Fig. 2) were built in stream 1 in October 2018 according to the guidance developed by local NFM practitioners (Yorkshire Dales Rivers Trust, 2018). Three types of flood water storage mechanisms were identified and opportunities were sought in the following priority order, designed to maximise flood storage volume: (1) Increased flood-plain connectivity with opportunities for re-routing of flood water to offline storage areas; (2) Increased flood-plain connectivity with in-line (floodplain) storage areas; (3) Increased in-channel storage. The dams were built from 2–5 locally felled tree stems with a minimum length of 1.5 times channel width. The stems were installed to span the channel perpendicular to the direction of flow. The dams had an average height of 0.8 m above the riverbed and were installed to provide approximately 0.3 m clearance from baseflow for fish passage.

(a) (b)

Fig. 2. Photographs of typical leaky dams on the impact stream: (a) elevation during baseflow conditions and (b) plan view during a high flow event.

2.2. Data collection

River stage was recorded at the upstream and downstream extent of both streams for 29 months, from September 2017 to February 2020, at one-minute intervals using In-Situ Rugged TROLL 100 (Redditch, UK) non-vented pressure transducers with 0.05% full scale accuracy (± 0.0045 m). The pressure transducers were placed in stilling wells and were corrected for atmospheric pressure using an In-Situ Rugged BaroTROLL (Redditch, UK) atmospheric pressure gauge (± 0.05% full scale accuracy). The leaky dams were built in the impact stream after 13 months of baseline monitoring, in October 2018. The streams were monitored for a further 16 month post-intervention monitoring period after the installation of the leaky dams in the impact stream.

Rating relationships (Fig. 3) were developed for the upstream and downstream gauging site on both streams by calibrating 1D hydraulic models of the sites, built in HEC-RAS 5.0.7 (USACE, 2020), to measurements of stage-discharge pairs at each site. Discharge was measured during a range of high flow events using slug-injection dilution methods (Moore, 2005), with salt pulses recorded using electrical conductivity as a proxy for concentration at one second intervals using a Campbell Scientific CR200 Data logger and conductivity probe. Between 10 and 26 stage-discharge pairs were collected at each of the four gauging sites. Confidence intervals to represent the degree of uncertainty in the rating relationships were calculated using the methodology presented by Lamb et al. (2003), and ranged from ± 0.03 m^3/s to ± 0.17 m^3/s at a discharge of 1.0 m^3/s. To avoid introducing this degree of uncertainty, the majority of the analysis in this study was conducted on the stage data, rather than the converted discharge data. The discharge data was used only to increase the comparability of the findings.

2.3. High flow events

High flow events were permitted to be single or multi-peaked. Discrete events were identified from the post-intervention stage time series using a similar rules-based methodology to Deasy et al. (2009) & Glendell et al. (2014) by requiring the following two criteria to be met: (1) A stage peak was considered a discrete high flow event if it was part of a defined flow event with duration >60 min and the upstream peak stage exceeded the mean stage recorded on the stream; (2) Events were classed as independent if they were separated by at least 15 min of stage below or within 10% of baseflow stage. Following the approach of Bezak et al. (2015) a consistent estimate of the baseflow stage series was obtained using the methods described in the World Meteorological Organisation's manual on low-flow estimation and prediction (WMO, 2009). The method identifies turning points based on minima found in defined time windows of daily time-series. To account for the flashy nature of the streams a three-day time window and turning point factor of 0.95 was used. The accompanying R package 'lfstat' v. 0.9.4 (Koffler et al., 2016) was used to implement the method. Finally, a visual inspection of the time series data was used to check that all events were extracted from the data and that the identified events were independent. The Hydrological Model Assessment and Development (HydroMAD) v.0.9-26 R package (Andrews and Guillaume, 2018) was used to identify discrete storm events using the above criteria.

Peak Over Threshold (POT) and Annual Maximum (AMAX) series (UK Centre for Ecology and Hydrology, 2021) from an Environment Agency maintained flow gauging station at Kilgram Bridge (station number F2206) on the River Ure, 6 km downstream of its confluence with the River Cover were used together with Met Office named storms (Met Office, 2021) to contextualise the events observed during the study period.

2.4. Event simulations

Transfer function noise (TFN) models were fitted to the stage time series collected in the baseline, pre-intervention monitoring period on the impact and control stream. Fitting and validation of the models is described in detail in the companion paper to this publication: van Leeuwen et al. (2023). In summary, the models given by Eqs. (1) and (2) were fitted to the baseline stage time series on each stream (September 2017–October 2018) using the upstream stage series as the predictor variable and the downstream stage series as the forecast variable. Both the upstream and downstream stage series were transformed to meet the requirements of stationarity by first order differencing. The fitted and validated models consisted of a dynamic regression and auto-regressive moving average (ARMA) noise component. The parsimonious form and coefficients of the models were inferred from the data. Blocked out of sample cross-validation showed that the models were able to simulate high flow event peak stage to within 0.02 m at the 80% confidence level on the control stream and at the 95% confidence level on the impact stream.

The transfer function noise models in Eqs. (1) and (2), with the parameter coefficients, υ, ϕ and θ given in Table 2, were used to simulate downstream baseline (pre-intervention) stage for high flow events in the post-intervention monitoring period.

$$D_t^* = U_t^* + \upsilon_1 U_{t-1}^* + \upsilon_2 U_{t-2}^* + \cdots + \upsilon_k U_{t-k}^* + N_t \tag{1}$$

$$N_t = \phi_1 N_{t-1} + \phi_2 N_{t-2} + \phi_3 N_{t-3} + \cdots + \phi_p N_{t-p} \\ +\theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} + \theta_q a_{t-q} + a_t \tag{2}$$

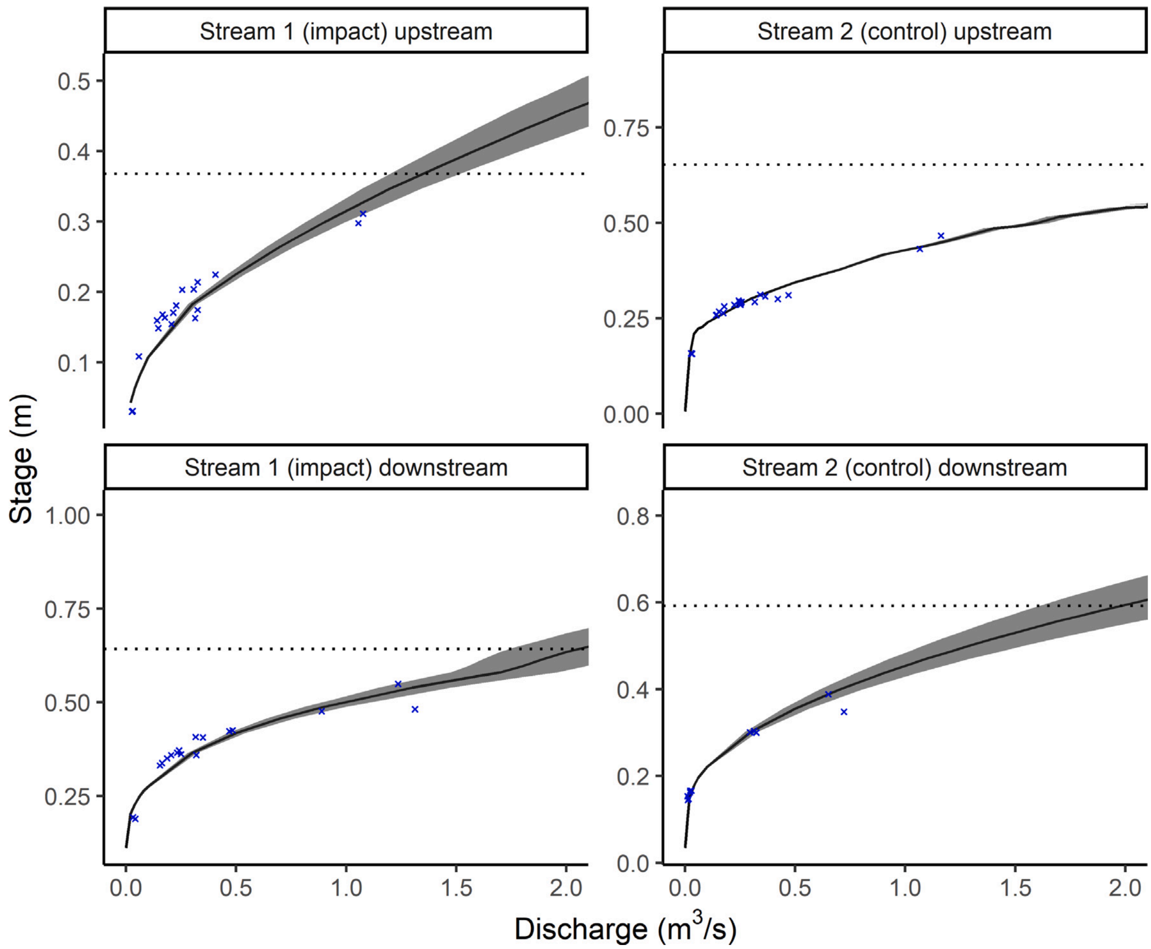

Fig. 3. Stage-Discharge rating relationships and their 95% confidence intervals (grey shading) for the upstream and downstream gauge on the impact and control streams. The dotted line indicates the maximum recorded stage during the study and the blue crosses indicate observations of stage-discharge pairs.

Table 2
Parsimonious TFN model parameter coefficients; coeff. = coefficient, s.e. = standard error in metres.

	Parameter	Stream 1 (impact) coeff.	Stream 2 (control) coeff.
Transfer Function Parameters υ	U^*_{t}	0.687	0.430
	U^*_{t-1}	0.332	0.198
	U^*_{t-2}	0.165	0.114
	U^*_{t-3}	0.102	0.083
	U^*_{t-4}	–	−0.042
	U^*_{t-11}	−0.014	–
	U^*_{t-20}	−0.010	–
AR terms ϕ	N_{t-1}	0.893	1.035
	N_{t-2}	0.179	−0.196
	N_{t-3}	−0.437	0.173
	N_{t-4}	0.288	−0.062
	N_{t-5}	−0.162	–
MA terms θ	a_{t-1}	−0.554	−0.739
	a_{t-2}	−0.654	−0.253
	a_{t-3}	0.453	–

Simulations of baseline stage were made for every discrete high flow event in the post-intervention monitoring period for the impact and control stream. The forecast variable was downstream stage, D^*_t (m), at time t, transformed to be stationary by first order differencing. The upstream stage series at time t formed the predictor variable, U^*_t (m), which was also transformed by first order differencing. N_t was the autocorrelated noise term which was modelled using the ARMA model (Eq. (2)), and a_t was a random noise term which was independently and identically distributed. The package 'forecast' v. 8.12 (Hyndman and Khandakar, 2008) was used in R v. 4.0.2 (R Core Team, 2020) to make simulations using the TFN models. The simulations were made for 20-hour windows of upstream stage data at 15-minute timesteps centred on the peak of the event. Simulations of baseline peak stage and empirical prediction intervals were taken from the blocked out of sample cross validation procedure described in van Leeuwen et al. (2023).

2.5. Treatment effect

The difference in peak stage with and without leaky dams was assessed by calculating the difference between the simulated peak stage and the observed peak stage in metres and was called the treatment effect, (T_e),

$$T_e = D_{peak} - \hat{D}_{peak} \tag{3}$$

where D_{peak} is the observed downstream stage in metres at the peak of the event (i.e., with leaky dams on the impact stream), and $\hat{D}_{peak}$ is the simulated baseline downstream stage in metres at the peak of the event (i.e., without leaky dams). This approach was adapted for peak stage from that of Dickson et al. (2012) and Gomi et al. (2006), who calculated treatment effect to assess differences in stream temperature response.

The model's empirical prediction intervals were used to assess whether the treatment effect was greater, less or within the same magnitude as the expected model error at the peak of the event. The prediction intervals were estimated in van Leeuwen et al. (2023), based on the distribution of peak stage error in the out of sample blocked cross-validation and were estimated at the 95% and 80% confidence levels.

Like Gomi et al. (2006) & Dickson et al. (2012) an approximate assessment of statistical significance was given by adapting the methods of Watson et al. (2001). Each high flow event for which T_e was calculated was assumed to be independent, given the criteria for storm separation (van Leeuwen et al., 2023), thus every value of T_e was assumed to be independent and can be described as the model error at the peak of the event. To test for significance in the difference between the distribution of treatment effect in the pre-and post-intervention monitoring periods the non-parametric two-sample Kolmogorov–Smirnov test was applied (Dickson et al., 2012; Gomi et al., 2006). The null hypothesis of 'no treatment effect' would be accepted if the distributions of the disturbances, or treatment effect, in the pre- and post-intervention monitoring periods were the same.

To aid in comparability between sites and studies, the treatment effect was also expressed in terms of the absolute and percentage change in peak discharge, T_{eQ}. The observed and simulated peak stage were converted to discharge in $\mathrm{m^3/s}$, Q_p and $\hat{Q}_p$, respectively using stage-discharge rating relationships in Fig. 3. The treatment effect was calculated as the absolute difference ($\mathrm{m^3/s}$) (Eq. (4)), and percentage difference (Eq. (5)) in event peak discharge,

$$T_{eQ} = (\hat{Q}_p - Q_p) \tag{4}$$

$$T_{eQ} = 100\left(\frac{(\hat{Q}_p - Q_p)}{\hat{Q}_p}\right) \tag{5}$$

2.6. Variation in treatment effect

To assess whether there was a relationship between an event's characteristics and the effectiveness of the leaky dams (treatment effect) a number of event characteristics were collected for each of the discrete storm events. The metrics used to describe the magnitude of the event were peak stage, S_p in metres, peak discharge (Q_p) in $\mathrm{m^3/s}$, duration, D in hours, and total stage, S_t in metres. The peak stage was defined as the maximum stage between the start and end of the event and was therefore based on the largest peak of multi-peaked events. Event duration was calculated as the time from the start to the end of the event and total stage was the sum of stage from the start to the end of the event, which was used as a proxy for event volume. After Potter (1991) the hydrograph rise time, T_{rise} in hours, defined as the time between the start and peak of the event, was calculated to describe the peakedness of the event. The time since the previous event, D_a in hours, was calculated for each event to give an approximation of antecedent conditions. Finally, time since the interventions were installed on the impact stream T_{int} in days was included to account for changes in the effectiveness of the leaky dams over time.

Scatter plots were initially used to determine whether an assumption of linearity between treatment effect and each of the characteristics was appropriate. Association between treatment effect and each of the event characteristics was assessed for normally distributed variables by calculating the Pearson's product moment correlation (Freedman et al., 2007). Where a significant association between treatment effect and peak magnitude was found, the assumption of linearity was reasonable, and errors were approximately normally distributed and homoskedastic, a linear regression was performed to assess the form and strength of the relationship. The null hypothesis was that the slope of the regression line was equal to zero. The coefficient of determination (R^2) was used to assess what proportion of the variation of treatment effect was explained by the event characteristic. Where appropriate, the regression relationship and its 95% prediction intervals were used to make predictions of treatment effect dependent on the event characteristic.

Where linearity between treatment effect and an event characteristic could not be assumed, or the assumption of normally distributed, homoskedastic errors could not be satisfied, an appropriate generalised linear model (GLM) was chosen. Both single and multiple variable GLM's were trialled to assess whether including one or more additional (uncorrelated) predictors improved the model fit. Gamma regression and standard linear regression with transformed variables were considered but none of the models provided a satisfactory fit to the data and were therefore not further pursued. Instead, to assess whether the characteristics affected the magnitude of T_e, the events were grouped according to whether they decreased peak magnitude (T_e > 95% prediction interval), increased peak magnitude (T_e < 95% prediction interval), or had an insignificant impact on peak magnitude (T_e within prediction interval). Binary logistic regression was used to test whether any of the event characteristics, or combinations of event characteristics, were significant predictors of whether treatment effect decreased peak magnitude or was insignificant. Events during which the dams increased peak magnitude were excluded from the analysis because of the low number of occurrences (n=3). The analysis was performed using R v. 4.0.2 (R Core Team, 2020).

2.7. Peak order

The effect of leaky dams on event peak magnitude is thought to differ depending on whether a peak is the first peak of an event (including single-peaked events) or subsequent peaks of multi-peaked events (Metcalfe et al., 2017). As numerous multi-peaked events were observed, the impact of peak order on treatment effect was examined. Individual peaks were deemed peaks of the same event if the criteria of the rules-based approach used to separate the flood events were not met. This meant that the stage could fall below the level of the leaky dams (0.3 m) between peaks of the same event, as long as they did not return to within 10% of baseflow for 15 min or more. The errors of multiple step ahead forecasts were additive, therefore, each peak of multi-peaked events was simulated at an equal number of time steps (10 h) from the beginning of the simulation. T_e was calculated for every peak and grouped according to whether they were the first, second, third or subsequent peak of the event. Significance of the differences in T_e before and after the dams were installed for the groups of first, second, third and subsequent peaks of events were assessed using Welch's t-test (Moser and Stevens, 1992).

3. Results

3.1. Characterisation of hydrological events

The post-intervention monitoring period was generally wetter than the baseline monitoring period, with higher total and annual maximum daily rainfalls recorded at the nearest Environment Agency (EA) operated rainfall gauge (Table 3). The annual maximum (AMAX) flows recorded at a downstream flow gauge during the post-intervention monitoring period were the third and ninth largest events since records began in 1966. In total, the discharge exceeded the peak over threshold (POT) for the gauge 10 times during the post-intervention monitoring period, compared to three times during the baseline monitoring period. The POT series is a record of all flows which exceed a certain threshold. The threshold is chosen for the whole dataset so that, on average, it is exceeded five times per year.

The flows in the study streams reflected the UK's typical rainfall seasonality (Met Office, 2020), with more high flow events in the autumn and winter months, although high flow events were observed throughout the year (Fig. 4, Table 4). Storms named by the Met Office (i.e. those which were accompanied by an amber or red weather warning for rain, snow or wind) resulted in some of the highest flows observed in the study streams. Similar discharges were recorded during these events in the baseline and post-intervention monitoring periods, despite more severe rainfall occurring in the post-intervention monitoring period (Fig. 4, Table 4). The study site discharges in Fig. 4, and

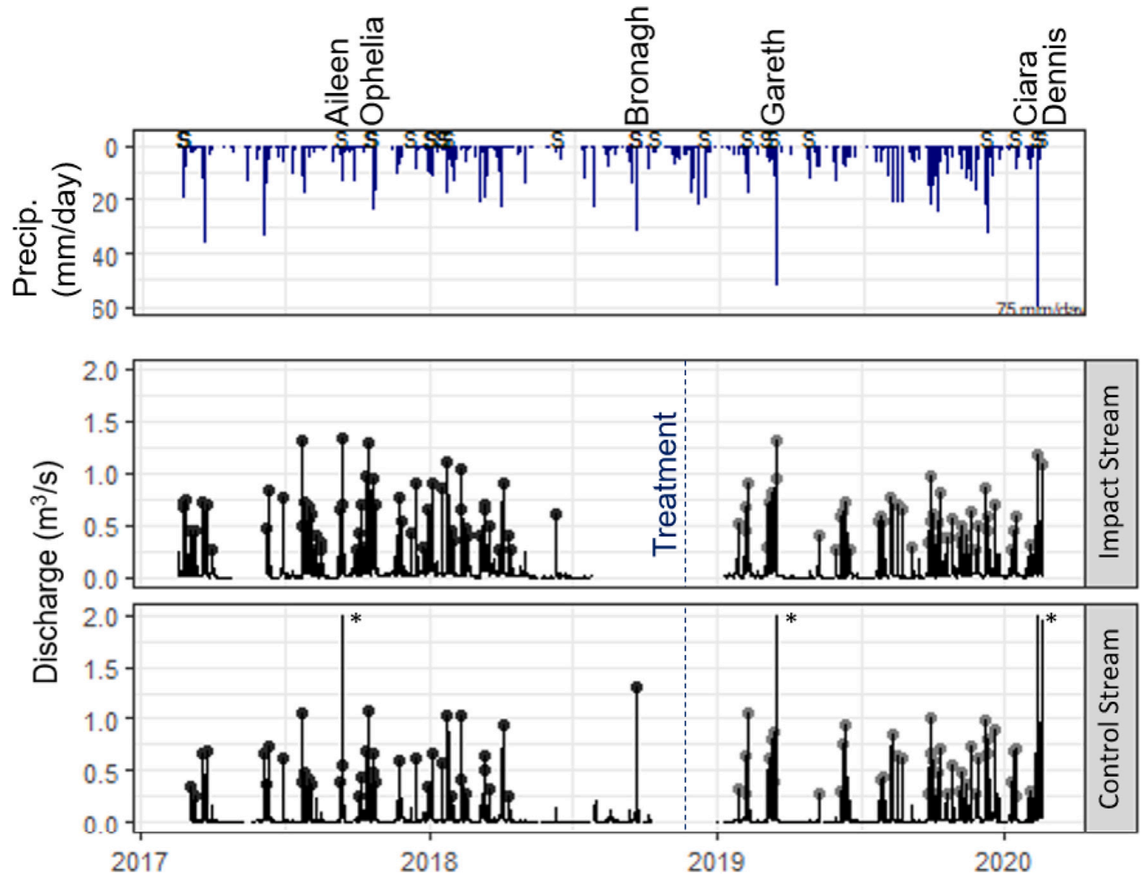

Fig. 4. Study site rainfall (EA gauge 7426) and discharge recorded at upstream extent of study streams, 'S' indicates storms named by the UK Met Office, and those which affected the study site are labelled along the upper margin. Points indicate high flow events identified in the data, asterisks indicate discharge >2 m^3/s which were deemed erroneous due to stage datum errors. Note that storm Ciara and storm Dennis were not included in the analysis because significant damage was caused to the gauging equipment during these events.

Table 3
Climate data at Scar House and Kilgram Bridge gauge (*to end of monitoring period, February 2020), water year is defined as 1st October to 30th September.

Water Year	Baseline		Post-intervention	
	2016–2017	2017–2018	2018–2019	2019–2020
Total precipitation (mm)	1153	1204	1466	1183*
Max. daily rainfall (mm)	50.4	32.4	52.2	92.4*
AMAX flow (m^3/s)	3.1	3.9	4.7	5.4
AMAX rank	48	30	9	3
Number of POT events	1	2	5	5

Table 4 display the effects of stage datum errors, with unrealistically high discharges recorded on the control stream during three of the named storms. Although not always as directly visible in the data, such errors occurred frequently, both on the impact and control stream and were most likely caused by blockage of the gauging stations with both coarse and fine material. These errors, which presented as changes in the stage datum, were avoided in the remaining analysis by first order differencing of the data, which made the stage datum arbitrary for each event. Errors in the data were identified through visual inspection of the data, after Crochemore et al. (2019) as well as single variable inspection, multi-variable inspection of correlated variables and detailed relationship examination, after Pastorello et al. (2014).

3.2. Statistical significance of treatment effect

After the leaky dams were installed on the impact stream, treatment effect was greater than the model 95% prediction interval for a third of high flow events (Fig. 5). The null hypothesis that the distributions of

Table 4
Top ten high flow events recorded on the impact stream during the monitoring period based on peak stage. The numbers in brackets indicate the rank of the equivalent POT event on the downstream Kilgram Bridge gauge, where relevant. Total rainfall indicates the total rainfall recorded in the 72 h up to and including the event date.

Rank	Date	Storm name	Impact peak stage (m)	Control peak stage (m)	Kilgram Bridge peak flow (m^3/s)	Scar House total rainfall (mm)
Baseline monitoring period						
1	13/09/2017	Aileen	0.58	0.50	171 (226)	35.4
2	14/10/2017	Ophelia	0.54	0.45	173 (222)	42.2
3	08/02/2018		0.52	0.40	104	13.0
4	23/01/2018	Georgina	0.51	0.41	159	31.4
5	03/04/2018		0.50	0.40	130	36.2
6	11/10/2017		0.49	0.36	162	27.2
7	21/10/2017	Brian	0.48	0.38	103	50
8	13/12/2017		0.48	0.37	92	21.2
9	03/01/2018	Eleanor	0.47	0.39	131	50.8
10	15/01/2018	Fionn	0.47	0.38	118	24.0
Post-intervention monitoring period						
1	09/02/2020	Ciara	0.63	0.62	361 (3)	113.0
2	15/02/2020	Dennis	0.60	0.73	292 (14)	56.2
3	16/03/2019	Gareth	0.58	0.59	303 (10)	72.4
4	18/12/2018	Deirdre	0.55	0.42	118	41.4
5	07/12/2018		0.55	0.38	130	37.2
6	09/02/2019	Erik	0.54	0.40	165 (251)	55.2
7	29/09/2019		0.53	0.44	179 (192)	76
8	05/12/2019	Atiyah	0.52	0.39	121	22.8
9	09/08/2019		0.51	0.42	174 (215)	36.8
10	19/12/2019		0.51	0.42	100	27

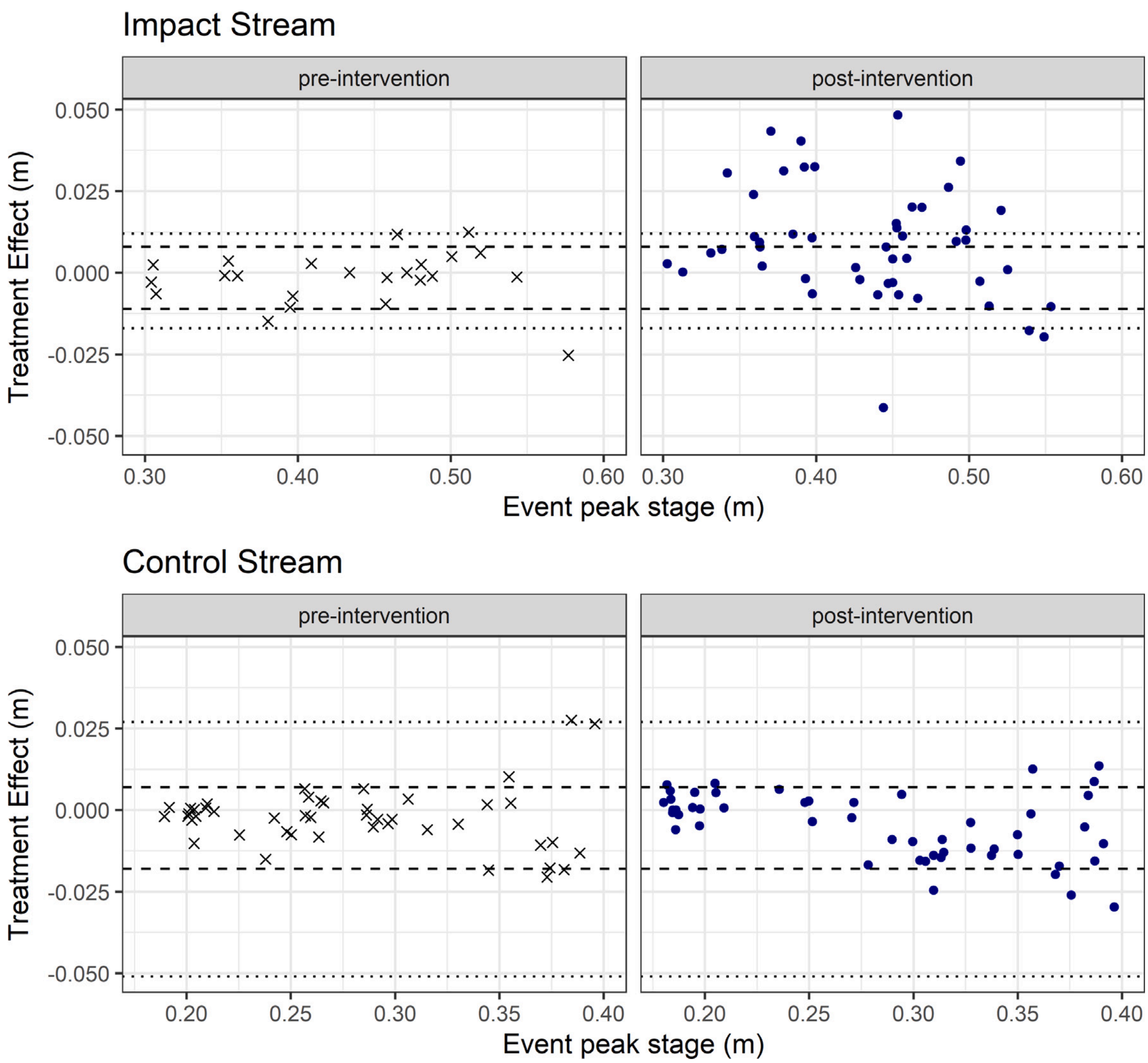

Fig. 5. Treatment effect on peak stage in metres for high flow events observed in baseline and post-intervention monitoring periods on the impact and control stream. The dashed and dotted lines indicate the empirical 95% and 80% prediction intervals respectively.

T_e came from the same population before and after the leaky dams were installed was rejected on the impact stream ($p < 0.01$) indicating that, overall, the leaky dams significantly reduced event peak magnitude.

On the control stream, the null hypothesis was not rejected ($p > 0.05$) indicating that downstream peak stage response of the control stream was not significantly different in the two monitoring periods.

3.3. Magnitude of treatment effect

The magnitude of the treatment effect (T_{eQ}) varied with the peak magnitude (Q_p) of the high flow events (Fig. 6). Peak magnitude was reduced by 8% on average for events with a peak magnitude of up to 1.2 m^3/s and by 10% on average for high flow events with a return period of up to 1 year (peak discharge 1.0 m^3/s). The linear relationship between treatment effect and peak discharge (Fig. 6) was significant but explained little of the variation in treatment effect (R^2=0.13, Root Mean Square Error (RMSE) = 15.9, p = 0.02).

The treatment effect was greatest for the most frequent, smaller events (peak discharge $\leq$0.3 m^3/s) which were reduced by 16% ($\pm$8%) (Table 5). As the peak magnitude increased the percentage reduction in peak magnitude decreased (Table 5).

Likewise, on average, absolute peak discharge reduction was greater for smaller events (Fig. 6); however, the largest reductions in absolute peak discharge (>0.2 m^3/s) for individual events were observed for three events with a peak magnitude of 0.7–0.9 m^3/s. For events on the impact stream with a peak discharge below 0.3 m, which flowed through the gap beneath the leaky dams, the observed and simulated peak magnitudes were similar in the pre- and post-intervention monitoring periods (data not presented). On the control stream there was a treatment effect of −2% on average in the pre-intervention monitoring period and -10% on average in the post-intervention monitoring period. However, unlike the difference on the impact stream, the difference in the control stream was not significant ($p > 0.05$).

3.4. Variation of treatment effect

The treatment effect expressed as the reduction in peak discharge (T_{EQ}) varied between −30% and 51% for individual events on the impact stream (Fig. 6) reflecting that, for the range of flows in which the interventions were observed, they could increase, decrease, or have a negligible impact on the event peak discharge compared to the baseline scenario. On the impact stream, the leaky dams reduced peak magnitude in almost a third (32%) of events at the 95% prediction intervals, and almost half (46%) of events at the 80% prediction interval (Table 6, Fig. 5). On the control stream, treatment effect of 87% of events was within the 80% prediction intervals.

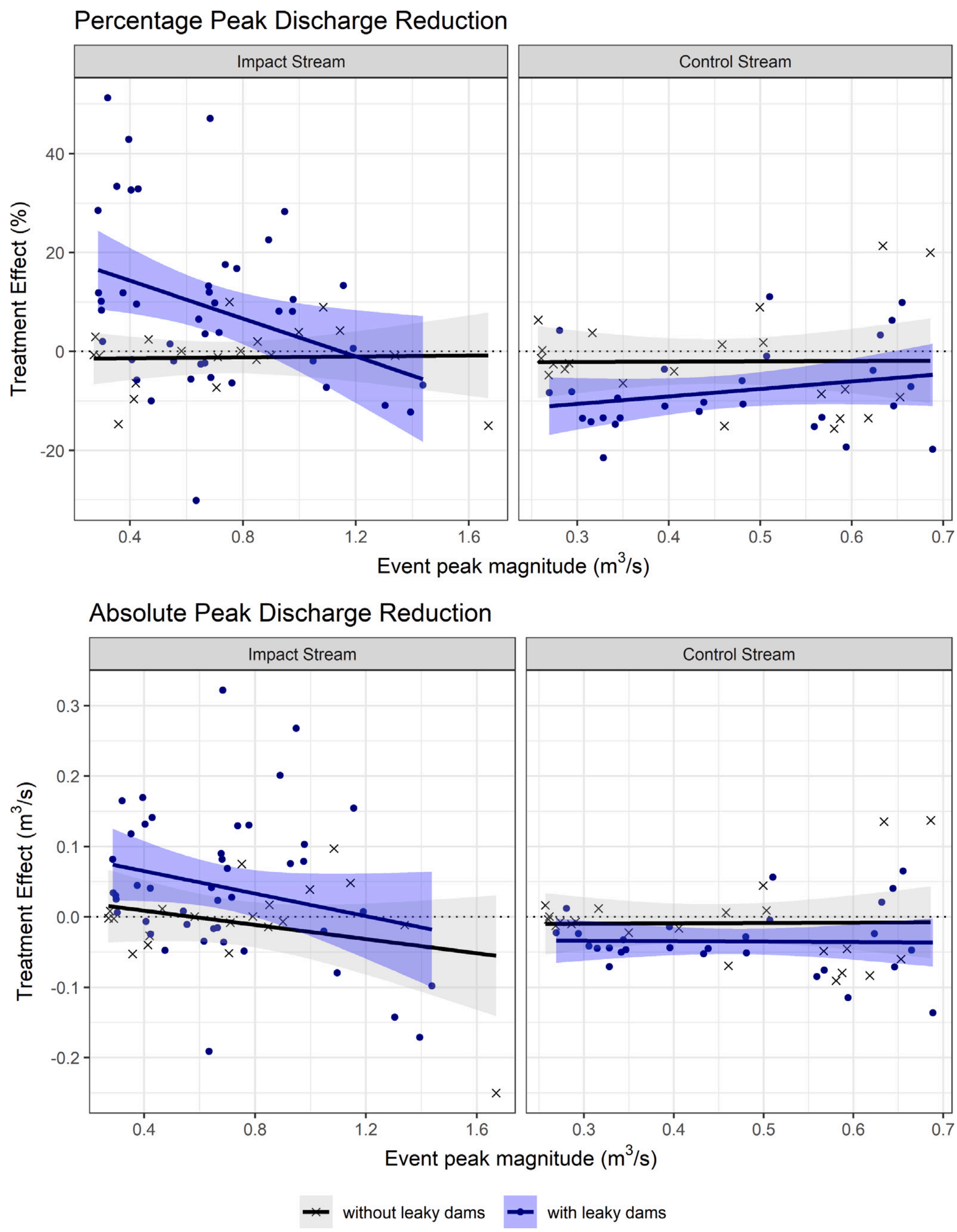

Fig. 6. Percentage and absolute Treatment effect on peak discharge for high flow events observed in baseline and post-intervention monitoring periods on the impact and control stream. A positive treatment effect indicates that the peak magnitude was reduced compared to the baseline scenario.

A treatment effect which reduced flood peak magnitude (i.e., points above the model prediction intervals) was predominantly observed for events with a shorter duration (D), lower total stage (S_t) and shorter time to rise (T_{rise}) (Fig. 7). Based on the upper and lower quartile of event characteristics within each group, the majority of events during which peak magnitude was reduced had short durations (D = 37-52 h) and had a total stage between 241 m and 346 m whilst the majority of events during which the treatment effect was negligible spanned a wider range of durations (D = 32-77 h) and total stage (S_t = 186-550 m) (Fig. 7). Time since previous event (D_a) did not affect the magnitude of treatment effect. The linear relationship between peak magnitude and treatment effect was not evident in the stage data as it had been when transformed to discharge. The number of days since the interventions were installed in the impact stream (T_{int}) shows that a peak magnitude reductions were observed only after 250 days of the dams being installed, although this is also when the majority of events occurred and could represent a seasonal effect. Logistic regression (Fig. 8) showed, however, that none of the event characteristics were significant predictors of whether the dams had a treatment effect which reduced peak magnitude, or not, during an event ($p > 0.05$).

3.5. Impact of peak order on treatment effect

During the combined baseline and post-intervention monitoring periods there were 144 peaks within 75 events, with a peak magnitude greater than 0.3 m, on the impact stream and 115 peaks in 54 events on the control stream. During the baseline monitoring period the treatment effect was similar, and close to zero, for the first, second and third peak of events on both the impact and control streams (Fig. 9), as expected. Following the installation of leaky dams on the impact

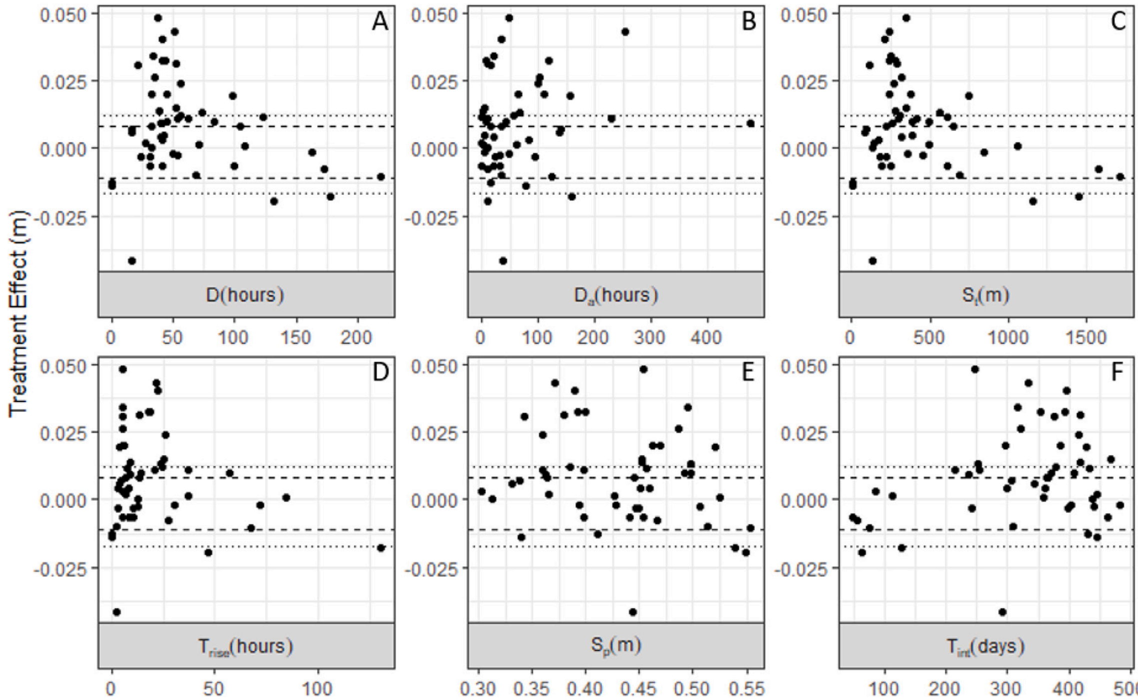

Fig. 7. Effect of event characteristics on treatment effect, where (A) D is event duration, (B) D_a is time since previous event, (C) S_t is the total stage (as a proxy for event volume), (D) T_{rise} is the time taken for the rising limb to reach the peak of the event, (E) S_p is the peak stage and (F) T_{int} is the time since the interventions were installed in the impact stream. Dashed lines represent the 95% prediction interval, and the dotted lines represent the 80% prediction interval, points above the upper lines indicate events during which the leaky dams had a positive treatment effect.

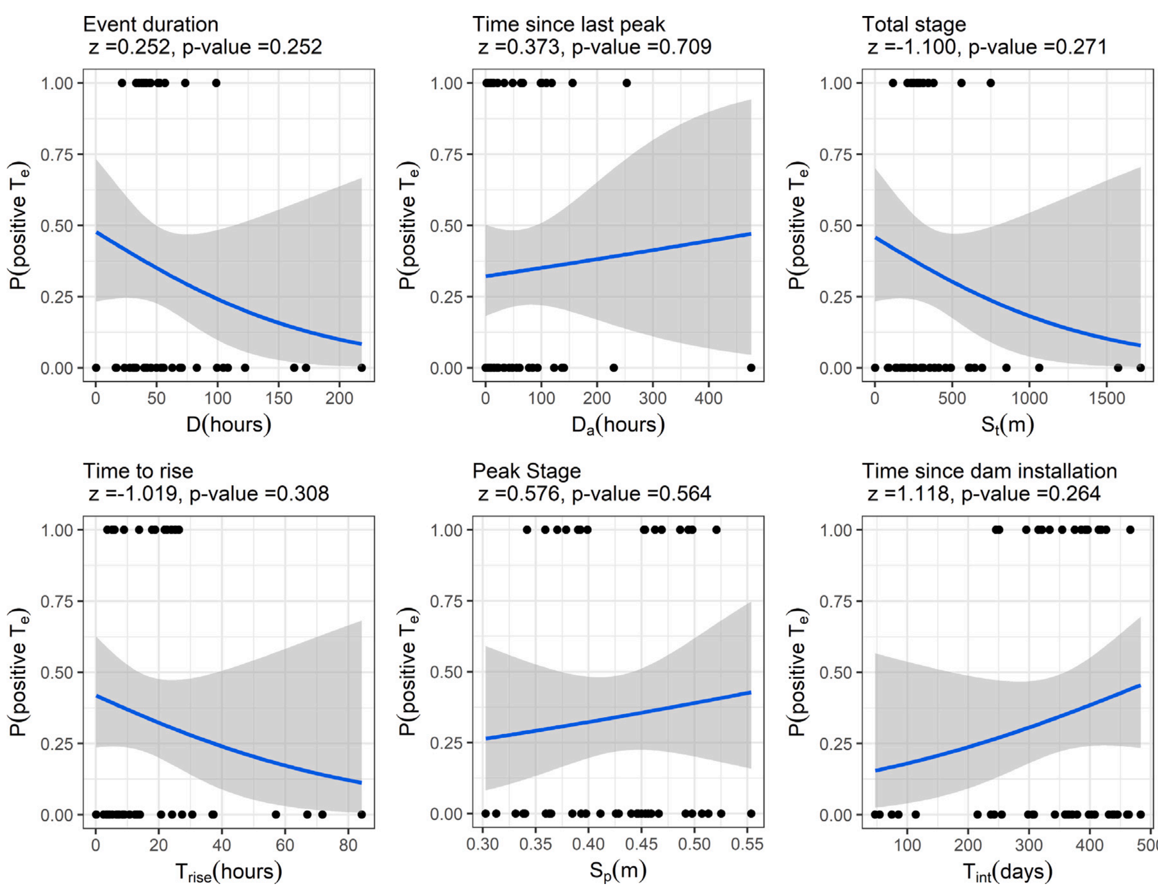

Fig. 8. Binary logistic relationships of event characteristics, where the line represents the model prediction of the probability of a positive treatment effect between 0 (insignificant treatment effect) and 1 (positive treatment effect). Shading is the 95% confidence interval of the regression and the points are the observations to which the relationship was fitted.

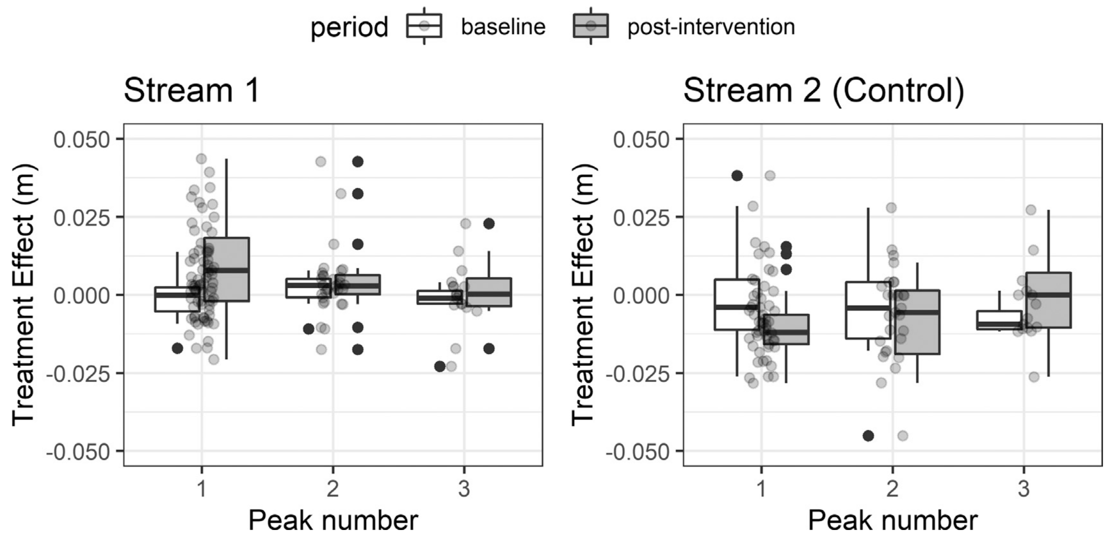

Fig. 9. Impact of peak order on treatment effect, jittered points illustrate the number of event peaks in each group, solid black points represent outliers from the range represented by the whiskers of the boxplots.

Table 5
Percentage reduction in event peak discharge for events with peak discharge between 0.3 m^3/s and 1.0 m^3/s based on linear relationship between treatment effect and event peak magnitude.

Peak Discharge (m^3/s)	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	Mean
Reduction in peak discharge (%)	16	14	12	10	9	7	5	3	1	1	8
lower 95% confidence limit (%)	8	8	7	5	4	1	−1	−4	−7	−10	1
upper 95% confidence limit (%)	24	21	18	16	13	12	11	10	9	8	14

Table 6
Number of events with during which the treatment effect increased, decreased, or had no significant effect on flood peak magnitude (PI= Prediction interval).

Stream	Te > PI peak reduced	PI < Te > PI no effect	Te < PI peak increased	Total
95% prediction interval				
Impact	16	31	3	50
Control	0	78	0	78
80% prediction interval				
Impact	23	22	5	50
Control	6	68	4	78

stream, the treatment effect for the first peak increased significantly ($p < 0.01$ Welch's t-test) but remained similar for subsequent peaks ($p > 0.05$). On the control stream, treatment effects in the post-intervention monitoring period were not significantly higher or lower than in the baseline monitoring period for the first, second, third or subsequent peaks ($p > 0.05$).

4. Discussion

This study has shown that leaky dams installed for the purpose of NFM in Pennine upland streams can statistically significantly reduce the peak magnitude of a range of frequent high flow events at the stream scale. Changes in peak magnitude were assessed for 50 high flow events with a return period ranging from < 1 year to 6 years and up to eight event peaks. Treatment effect-size was highly variable, but on average, flood peak magnitude was reduced by 10% for events up to a 1-in-1 year return period. Some of the variability in treatment effect was explained by the order of event peaks: the leaky dams were effective during single peaked events and for the first peak of multi-peaked events, but not for the second or subsequent peaks of multi-peaked events.

The study quantified the impact of leaky dams on larger, steeper streams than previous research, which focused on in-stream wood installed for the purpose of river restoration, rather than flood risk management (Wenzel et al., 2014; Keys et al., 2018). Although leaky dam impacts on flood peak magnitude have not previously been quantified for leaky dams installed for the purpose of NFM, the results can be compared to two studies which quantified the impact on river restoration style in-stream wood. The average reduction in event peak magnitude was similar to the findings of Keys et al. (2018) who found just three pieces of large wood in a 50 m section of a steep, headwater stream in the Mid-Atlantic region of the US decreased flood peak magnitude by 8% for a < 1-in-1 year event. Similarly, the findings were in agreement with Wenzel et al. (2014), who assessed the impact of nine spruce tops placed in a small, steep stream in the Ore mountains, Germany, that for a more extreme event (1-in-3.5 year) the impact on flood peak magnitude was negligible (2.3%). Unlike Wenzel et al. (2014) & Keys et al. (2018) who studied just one type and magnitude of event, the data-based time series modelling approach taken in this study allowed the variability of leaky dam impacts in the stream to be studied for 50 different storm events with differing characteristics including a range of peak magnitudes, storm durations and number of event peaks.

4.1. Leaky dam effectiveness during flood events

Previous research has shown that the effectiveness of leaky dams at delaying flood peaks reduced as event magnitude increased (Kitts, 2010; Gregory et al., 1985). This study has shown that this is also the case for the effectiveness of leaky dams at reducing peak magnitude. Like Gregory et al. (1985) and Kitts (2010) the study found that leaky dam effectiveness decreased as event peak magnitude increased. The leaky dam impacts became negligible at the stream scale for > 1-in-1 year events, supporting reviews of the NFM approach (Burgess-Gamble et al., 2017; Dadson et al., 2017; Lane, 2017) in questioning leaky dam impacts at higher event magnitudes. This conclusion could be a product of the location, design, number or size of interventions tested in this study and elsewhere (Ellis et al., 2021), although modelling has shown that adding more leaky dams in steep streams does not necessarily mean their storage will be utilised (Hankin et al., 2020).

Recent modelling of extensive hillslope storage bunds across a catchment has shown that individual features with a small storage capacity fill up during the rising limb of larger events, which impedes their impact on downstream flooding (Beven et al., 2022). The leaky dams in the study site were placed in narrow sections of the stream with little opportunity for out of bank flows. This effect was, therefore, observed in the study site; the leaky dams had a finite amount of storage volume, which limited their effectiveness during > 1-in-1 year events. It follows, that if it was possible to increase or remove the limit on storage volume, leaky dams could have the potential to reduce the peak magnitude of larger events. The availability of increased flood plain connectivity and expandable field storage during higher return period events has been shown to increase leaky dam impacts in some (sub-)catchments (Black et al., 2021; Hankin et al., 2020; Kay et al., 2019; Thomas and Nisbet, 2012). By empirically demonstrating the limitations of finite storage capacity, the results of this study emphasise the importance of designing leaky dam schemes which take advantage of expandable field storage, so that they have the potential to reduce peak flows of larger storm events.

Although the interventions had a negligible local impact on peak magnitude during events which caused significant flooding of properties and transport links in North Yorkshire (flood return period > 1 year), it remains to be seen what impact reduced peak flows of < 1-in-1 year events in headwater streams have further downstream. The more frequent, small-scale flood events which were impacted by the interventions can cause localised flooding which can result in waterlogging, bank erosion and debris deposition on downstream agricultural land (Posthumus et al., 2008). The vulnerability of agricultural land depends on land use and the frequency, duration, depth and seasonality of the flood event (Morris and Hess, 1988). Whilst grassland used for livestock can tolerate winter flooding, summer floods can destroy an entire harvest (Morris and Brewin, 2014; Posthumus et al., 2009). Flooding and waterlogging of agricultural land was estimated to have an average economic cost of £12,000 per hectare, or £90,000 per farm, albeit during the extreme 2007 UK floods (Posthumus et al., 2009). Greater understanding of the impacts of leaky dams during different types of events, such as typical summer and winter storms, could, therefore, play an important role in the reduction of economic costs of flooding to farmers. Additionally, placing leaky dams in upland watercourses could reduce the reliance on flooding of productive agricultural land as temporary flood storage areas which would reduce the costs associated with compensation payments under schemes such as the UK's proposed Environmental Land Management Scheme (ELMS).

4.2. Variability in leaky dam effectiveness

Although there was a relationship between treatment effect and peak magnitude, the variability in effectiveness of the leaky dams in the study site was not explained by peak magnitude alone. Whilst leaky dams did reduce event peak magnitude of some events, their effectiveness was highly variable. Variability in leaky dam scheme effectiveness was found in a catchment scale model of leaky dam impacts (Dixon et al., 2016), but was attributed to varying interactions of sub-catchments, which did not play a role at the stream scale of this study. None of the flood event characteristics (event duration, time since previous event, time to rise, total event stage, peak stage or time since the leaky dams were installed) determined during which events reductions in peak magnitude were or were not observed ($p > 0.05$). Continued monitoring to obtain a larger number of events would be needed to provide the data required to perform additional statistical analyses to determine which event characteristics are useful predictors of leaky dam effectiveness.

The data did show that leaky dams were significantly more effective during single-peaked and the first peak of multi-peaked events, than during the second or subsequent peaks of multi-peaked events. Although not significant ($p > 0.05$), a difference was also found for the control stream; the first peaks of events were higher than expected in the post-intervention monitoring period. It is possible that this difference is due to limitations of the TFN model of the control stream, which was less good at predicting peak magnitude of larger events (van Leeuwen et al., 2023) than the model of the impact stream. Larger events were more prevalent in the post-intervention monitoring period, which could explain this unexpected difference. Limitations of the TFN models are discussed in detail in van Leeuwen et al. (2023). The findings on the impact stream, nonetheless, provide empirical evidence to support the conclusion drawn from catchment scale modelling that multi-peaked events should be considered in the design and assessment of leaky dam efficacy (Metcalfe et al., 2017), and suggests this is an important consideration even at the stream scale. Had the study considered only events with well-defined peaks, which is common practice (Gregory et al., 1985; Grayson et al., 2010; Metcalfe et al., 2017), it may have concluded considerably higher average leaky dam impacts despite the majority of events being multi-peaked. Failing to do so can be misrepresentative and may distort expectations of leaky dam effectiveness. The assessment of leaky dam impacts on a range of events, therefore, helps to manage expectations. Managing expectations is crucial to avoid over-reliance of communities on NFM measures (Wells et al., 2020), and to sustain confidence in NFM in the long term (Vira and Adams, 2009).

4.3. Recommendations for next steps

Funding of NFM schemes through traditional mechanisms is held back by lack of quantification of their flood risk management benefits, which is required to justify the spending of public funds (Defra, 2009). The benefits of flood risk management schemes are usually designed and assessed in hydraulic and hydrological models. The lack of quantitative evidence of leaky dam impacts means that their representation in such models has not been validated, leading to low confidence in their outputs (Addy and Wilkinson, 2019). By presenting one of the most comprehensive quantifications of leaky dams impacts to date, spanning a range of peak magnitudes and event types, this study has provided one of the first diverse validation datasets with which to assess the representation of leaky dams in hydraulic and hydrological models in steep upland streams. Furthermore, applying the data-based time series modelling approach to new and existing BACI data from other NFM sites could provide the validation datasets needed to gain confidence in the representation of leaky dams in hydraulic and hydrological models more generally.

Increased confidence in the representation of leaky dams in hydraulic and hydrological models is particularly important to be able to address questions about the impacts of NFM measures at larger spatial scales (Burgess-Gamble et al., 2017; Dadson et al., 2017; Ellis et al., 2021; Lane, 2017). Although leaky dam impacts may be small at the stream scale, by desynchronising tributary flows, downstream flood risk during large floods could be significantly reduced (Pattison et al., 2014). Modelling has shown that leaky dams, in combination with reforestation, could have considerable impacts on catchment scale flood risk (Dixon et al., 2016). There are few empirical tests of these effects due to the time and financial investment needed to install large numbers of NFM features and monitoring equipment across multiple upland headwater sub-basins (Black et al., 2021). A multi-local scale modelling approach has been proposed to address scaling issues, in which empirical data from small sub-catchments (<1 km^2) is used to assess NFM impacts in multiple larger catchments (<10 km^2) upstream of communities at risk of flooding (Hankin et al., 2021). Data from more sub-catchments, for a range of high flow events and antecedent conditions is needed to support the development of methods which advance understanding of catchment scale impacts of NFM (Hankin et al., 2021). This study has demonstrated that the data-based time series modelling approach can be invaluable in obtaining such evidence from small catchments.

5. Conclusion

Leaky dam impacts on flood peak magnitude were negligible for flood peaks with > 1 year return period, including events during which downstream flooding of properties and transport links was observed. Leaky dams have the potential, however, to decrease the magnitude of frequent flood peaks (up to 1 year return period) on high gradient streams by 10% on average, although their effects are highly variable. The data-based time series modelling approach allowed the impacts of leaky dams during a large number and range of event types to be assessed, for the first time. The results have important implications for the design and assessment of leaky dam schemes. Whilst event peak magnitude is an important factor to consider when designing leaky dam schemes, the conditions required for the system to recover between event peaks are also important. Whether assessing the impact of leaky dams empirically or using numerical models, the results show that an assessment of leaky dam impacts is not complete without considering a range of event types as well as event peak magnitudes. Leaky dam schemes which are assessed using single-peaked design storms only are likely to overpromise and underdeliver on flood risk management benefits. By supporting the BACI approach with data-based time series modelling, the challenges associated with quantifying NFM effectiveness in a range of environments and for a range of events can be overcome.

CRediT authorship contribution statement

Z.R. van Leeuwen: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **M.J. Klaar:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **M.W. Smith:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **L.E. Brown:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank the Yorkshire Dales Rivers Trust for assistance with building leaky dams and the Mawle family for allowing this research to be completed on their land. This work was funded by the Natural Environment Research Council with Industrial CASE support from JBA Trust and the Yorkshire Dales Rivers Trust, NE/N008065/1. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

References

Abbe, T., Brooks, A., 2011. Geomorphic, engineering, and ecological considerations when using wood in river restoration. Geophys. Monogr. Ser. 194, 419–451. http://dx.doi.org/10.1029/2010GM001004.

Abbe, T.B., Brooks, A.P., Montgomery, D.R., 2003. Wood in river rehabilitation and management. In: Gregory, S.V., Boyer, K.L., Gurnell, A.M. (Eds.), The Ecology and Management of Wood in World Rivers. American Fisheries Society, pp. 1–26. http://dx.doi.org/10.47886/9781888569568.

Addy, S., Wilkinson, M., 2016. An assessment of engineered log jam structures in response to a flood event in an upland gravel-bed river. Earth Surf. Process. Landf. 1670 (April), 1658–1670. http://dx.doi.org/10.1002/esp.3936.

Addy, S., Wilkinson, M.E., 2019. Representing natural and artificial in-channel large wood in numerical hydraulic and hydrological models. WIREs Water 6 (6), 1–20. http://dx.doi.org/10.1002/wat2.1389.

Andrews, F., Guillaume, J., 2018. hydromad: Hydrological model assessment and development. URL: http://hydromad.catchment.org/.

Bark, R.H., Martin-Ortega, J., Waylen, K.A., 2021. Stakeholders' views on natural flood management: Implications for the nature-based solutions paradigm shift? Environ. Sci. Policy 115, 91–98. http://dx.doi.org/10.1016/j.envsci.2020.10.018, URL: https://linkinghub.elsevier.com/retrieve/pii/S1462901120313678.

Bernhardt, E.S., Palmer, M.A., Allan, J.D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G.M., Lake, P.S., Lave, R., Meyer, J.L., Donnell, T.K.O., Pagano, L., Powell, B., Sudduth, E., 2005. Synthesizing U.S. river restoration efforts. Science 308, 636–637.

Beven, K., Page, T., Hankin, B., Smith, P., Kretzschmar, A., Mindham, D., Chappell, N., 2022. Deciding on fitness-for-purpose-of models and of natural flood management. Hydrol. Process. 36 (11), http://dx.doi.org/10.1002/hyp.14752, URL: https://onlinelibrary.wiley.com/doi/10.1002/hyp.14752.

Bezak, N., Horvat, A., Šraj, M., 2015. Analysis of flood events in Slovenian streams. J. Hydrol. Hydromech. 63 (2), 134–144. http://dx.doi.org/10.1515/johh-2015-0014.

Black, A., Peskett, L., MacDonald, A., Young, A., Spray, C., Ball, T., Thomas, H., Werritty, A., 2021. Natural flood management, lag time and catchment scale: Results from an empirical nested catchment study. J. Flood Risk Manag. 14 (3), 1–16. http://dx.doi.org/10.1111/jfr3.12717, URL: https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12717.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G.B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T.R., Kohnová, S., Koskela, J.J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J.L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K., Živković, N., 2019. Changing climate both increases and decreases European river floods. Nature 573, 108–111. http://dx.doi.org/10.1038/s41586-019-1495-6.

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., Quinn, P., 2017. Working with Natural Processes - Evidence Directory. Technical Report, Environment Agency, Bristol, URL: https://assets.publishing.service.gov.uk/.

Collentine, D., Futter, M.N., 2018. Realising the potential of natural water retention measures in catchment flood management: trade-offs and matching interests. J. Flood Risk Manag. 11 (1), 76–84. http://dx.doi.org/10.1111/jfr3.12269.

Commission of the European Communities, 2009. White Paper - Adapting to Climate Change - Towards a European Framework for Action. Technical Report, European Commission, Brussels, URL: https://eur-lex.europa.eu/legal-content/.

Connelly, A., Snow, A., Carter, J., Lauwerijssen, R., 2020. What approaches exist to evaluate the effectiveness of UK-relevant natural flood management measures? A systematic map protocol. Environ. Evid. 9 (1), 1–13. http://dx.doi.org/10.1186/s13750-020-00192-x.

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., Arheimer, B., 2019. Lessons learnt from checking the quality of openly accessible river flow data worldwide. Hydrol. Sci. J. 65 (5), 699–711. http://dx.doi.org/10.1080/02626667.2019.1659509, URL: https://www.tandfonline.com/doi/full/10.1080/02626667.2019.1659509.

Curran, J.H., Wohl, E.E., 2003. Large woody debris and flow resistance in step-pool channels, Cascade Range, Washington. Geomorphology 51 (1–3), 141–157. http://dx.doi.org/10.1016/S0169-555X(02)00333-1, URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-0037456686&partnerID=40&md5=ad5dbf8c02bb0ea869fce7d0508f16e8, https://linkinghub.elsevier.com/retrieve/pii/S0169555X02003331.

Dadson, S.J., 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 (2199), 20160706. http://dx.doi.org/10.1098/rspa.2016.0706, URL: https://royalsocietypublishing.org/doi/10.1098/rspa.2016.0706.

Deasy, C., Brazier, R.E., Heathwaite, A.L., Hodgkinson, R., 2009. Pathways of runoff and sediment transfer in small agricultural catchments. Hydrol. Process. 23 (1), 1349–1358.

Defra, 2005. Making Space for Water:Taking Forward a New Government Strategy for Flood and Coastal Management in England. Technical Report, DEFRA, London.

Defra, 2009. Appraisal of Flood and Coastal Erosion Risk Management. A Defra Policy Statement. Technical Report June, Department for Environment, Food and Rural Affairs, London, p. 47, URL: https://www.gov.uk/government/publications/.

Defra, 2020. The Environmental Land Management Scheme: Public Money for Public Goods. Technical Report, Department for Environment, Food and Rural Affairs, London, pp. 1–4, URL: https://www.gov.uk/government/publications/.

Defra, EAFRD, Natural England, 2016. Countryside Stewardship - How can it help reduce flooding? URL: https://www.gov.uk/government/uploads/.

Dickson, N.E., Carrivick, J.L., Brown, L.E., 2012. Flow regulation alters alpine river thermal regimes. J. Hydrol. 464–465, 505–516. http://dx.doi.org/10.1016/j.jhydrol.2012.07.044.

Dixon, S.J., Sear, D.A., Odoni, N.A., Sykes, T., Lane, S.N., 2016. The effects of river restoration on catchment scale flood risk and flood hydrology. Earth Surf. Process. Landf. 41 (7), 997–1008. http://dx.doi.org/10.1002/esp.3919, URL: https://onlinelibrary.wiley.com/doi/10.1002/esp.3919.

Dodd, J.A., Newton, M., Adams, C.E., 2016. The Effect of Natural Flood Management In-Stream Wood Placements on Fish Movement in Scotland, CD2015_02. Technical Report, CREW, URL: crew.ac.uk/publications.

Ellis, N., Anderson, K., Brazier, R., 2021. Mainstreaming natural flood management: A proposed research framework derived from a critical evaluation of current knowledge. Prog. Phys. Geogr. Earth Environ. 1–23. http://dx.doi.org/10.1177/0309133321997299, URL: http://journals.sagepub.com/doi/10.1177/0309133321997299.

Environment Agency, 2009. Flooding in England: A National Assessment of Flood Risk. Technical Report, Environment Agency, Bristol, p. 36, URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/292928/geho0609bqds-e-e.pdf.

Environment Agency, 2019. Natural Flood Management Programme : Interim Lessons Learnt. Technical Report January, Environment Agency, Bristol.

Forbes, H., Ball, K., McLay, F., 2015. Natural Flood Management Handbook. Technical Report, SEPA, Stirling, p. 142.

Freedman, D., Pisani, R., Purves, R., 2007. Statistics (International Student Edition), fourth ed. WW Norton and Company, New York, pp. 37–40.

Garvey, A., Paavola, J., 2022. Community action on natural flood management and the governance of a catchment-based approach in the UK. Environ. Policy Gov. 32 (1), 3–16. http://dx.doi.org/10.1002/eet.1955.

Glendell, M., Extence, C., Chadd, R., Brazier, R.E., 2014. Testing the pressure-specific invertebrate index (PSI) as a tool for determining ecologically relevant targets for reducing sedimentation in streams. Freshwater Biol. 59 (2), 353–367. http://dx.doi.org/10.1111/fwb.12269.

Gomi, T., Moore, R.D., Dhakal, A.S., 2006. Headwater stream temperature response to clear-cut harvesting with different riparian treatments, coastal British Columbia, Canada. Water Resour. Res. 42 (8), 1–11. http://dx.doi.org/10.1029/2005WR004162.

Grabowski, R.C., Gurnell, A.M., Burgess-Gamble, L., England, J., Holland, D., Klaar, M.J., Morrissey, I., Uttley, C., Wharton, G., 2019. The current state of the use of large wood in river restoration and management. Water Environ. J. 33, 366–377. http://dx.doi.org/10.1111/wej.12465, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/wej.12465 https://onlinelibrary.wiley.com/doi/10.1111/wej.12465.

Grayson, R., Holden, J., Rose, R., 2010. Long-term change in storm hydrographs in response to peatland vegetation change. J. Hydrol. 389 (3–4), 336–343. http://dx.doi.org/10.1016/j.jhydrol.2010.06.012.

Gregory, K.J., Gurnell, a.M., Hill, C.T., 1985. The permanence of debris dams related to river channel processes. Hydrol. Sci. J. 30 (3), 371–381. http://dx.doi.org/10.1080/02626668509491000.

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. Nat. Hazards Earth Syst. Sci. 20 (10), 2567–2584. http://dx.doi.org/10.5194/nhess-20-2567-2020.

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. Hydrol. Process. 35 (11), http://dx.doi.org/10.1002/hyp.14418, URL: https://onlinelibrary.wiley.com/doi/10.1002/hyp.14418.

Hester, N., Rose, S., Worrall, P., 2016. Case Study 3 - Holnicote Demonstration Project - Somerset. Technical Report, Environment Agency, Bristol, URL: https://www.gov.uk/government/publications/.

Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: The forecast package for R. J. Stat. Softw. 27 (3), 22, URL: http://www.jstatsoft.org/v27/i03/paper.

Iacob, O., Rowan, J., Brown, I., Ellis, C., 2012. Natural flood management as a climate change adaptation option assessed using an ecosystem services approach. In: Hydrology for a Changing World, No. 2011. British Hydrological Society, pp. 01–06. http://dx.doi.org/10.7558/bhs.2012.ns26, URL: http://www.hydrology.org.uk/assets/2012papers/Iacob_26.pdf.

Kail, J., Hering, D., Muhar, S., Gerhard, M., Preis, S., 2007. The use of large wood in stream restoration: Experiences from 50 projects in Germany and Austria. J. Appl. Ecol. 44 (6), 1145–1155. http://dx.doi.org/10.1111/j.1365-2664.2007.01401.x.

Kay, A.L., Old, G.H., Bell, V.A., Davies, H.N., Trill, E.J., 2019. An assessment of the potential for natural flood management to offset climate change impacts. Environ. Res. Lett. 14 (4), http://dx.doi.org/10.1088/1748-9326/aafdbe.

Keys, T.A., Govenor, H., Jones, C.N., Hession, W.C., Hester, E.T., Scott, D.T., 2018. Effects of large wood on floodplain connectivity in a headwater Mid-Atlantic stream. Ecol. Eng. 118 (March), 134–142. http://dx.doi.org/10.1016/j.ecoleng.2018.05.007.

Kitts, D.R., 2010. The hydraulic and hydrological performance of large wood accumulation in a low-order forest stream (Ph.D. thesis). University of Southampton, p. p. 368, URL: http://eprints.soton.ac.uk/185791/.

Klaar, M.J., Carver, S., Kay, P., 2020. Land management in a post-Brexit UK: An opportunity for integrated catchment management to deliver multiple benefits? Wiley Interdiscip. Rev. Water 7 (5), 1–6. http://dx.doi.org/10.1002/wat2.1479.

Koffler, D., Gauster, T., Laaha, G., 2016. lfstat. URL: https://cran.r-project.org/package=lfstat.

Lamb, R., Zaidman, M., Archer, D., Marsh, T., Lees, M., 2003. River Gauging Station Data Quality Classification (GSDQ). R&D Technical Report W6-058/TR, for The Environment Agency, Technical Report, Environment Agency, Bristol, UK.

Lane, S.N., 2017. Natural flood management. WIREs Water 4 (3), 1–14. http://dx.doi.org/10.1002/wat2.1211, URL: https://onlinelibrary.wiley.com/doi/10.1002/wat2.1211.

Lavers, T., Charlesworth, S.M., Lashford, C., Warwick, F., Fried, J., 2022. The performance of natural flood management at the large catchment-scale: A case study in the Warwickshire Stour Valley. Water 14 (23), http://dx.doi.org/10.3390/w14233836.

Met Office, 2020. Met Office: Climate Averages 1971- 2000. Met Office, URL: http://www.metoffice.gov.uk/climate/.

Met Office, 2021. UK Storm Centre. URL: https://www.metoffice.gov.uk/weather/.

Metcalfe, P., Beven, K., Hankin, B., Lamb, R., 2017. 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 29-km2 scale catchment in the United Kingdom. Hydrol. Process. 31 (9), 1734–1748. http://dx.doi.org/10.1002/hyp.11140, URL: https://onlinelibrary.wiley.com/doi/10.1002/hyp.11140.

Moore, R.D., 2005. Introduction to salt dilution gauging for streamflow measurement: Part 1. Streamline 8 (4), 1–6. http://dx.doi.org/10.1592/phco.23.9.1S.32890.

Morris, J., Brewin, P., 2014. The impact of seasonal flooding on agriculture: The spring 2012 floods in Somerset, England. J. Flood Risk Manag. 7 (2), 128–140. http://dx.doi.org/10.1111/jfr3.12041.

Morris, J., Hess, T.M., 1988. Agricultural flood alleviation benefit assessment - a case study. J. Agric. Econ. 39 (3), 402–412. http://dx.doi.org/10.1111/j.1477-9552.1988.tb00600.x, URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1477-9552.1988.tb00600.x.

Moser, B.K., Stevens, G.R., 1992. Homogeneity of variance in the two-sample means test. Amer. Statist. 46 (1), 19. http://dx.doi.org/10.2307/2684403, URL: https://www.jstor.org/stable/2684403?origin=crossref.

National Trust, 2015. From Source to Sea - Natural Flood Management - The Holnicote Experience. Technical Report, National Trust, p. 28.

Nisbet, T., Thomas, H., Roe, P., 2015. Case Study 12 . Slowing the Flow at Pickering. Technical Report, Defra, Environment Agency, pp. 1–11.

O'Driscoll, C., O'Connor, M., Asam, Z., de Eyto, E., Brown, L.E., Xiao, L., 2016. Forest clearfelling effects on dissolved oxygen and metabolism in peatland streams. J. Environ. Manag. 166, 250–259. http://dx.doi.org/10.1016/j.jenvman.2015.10.031, URL: https://linkinghub.elsevier.com/retrieve/pii/S0301479715303297.

Pastorello, G., Agarwal, D., Papale, D., Samak, T., Trotta, C., Ribeca, A., Poindexter, C., Faybishenko, B., Gunter, D., Hollowgrass, R., Canfora, E., 2014. Observational data patterns for time series data quality assessment. In: Proceedings - 2014 IEEE 10th International Conference on eScience, Vol. 1, No. October. eScience 2014, pp. 271–278. http://dx.doi.org/10.1109/eScience.2014.45.

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 Resour. Res. 50 (7), 5444–5458. http://dx.doi.org/10.1002/2013WR014067, URL: http://doi.wiley.com/10.1002/2013WR014067.

Posthumus, H., Hewett, C.J., Morris, J., Quinn, P.F., 2008. Agricultural land use and flood risk management: Engaging with stakeholders in North Yorkshire. Agricult. Water Manag. 95 (7), 787–798. http://dx.doi.org/10.1016/j.agwat.2008.02.001.

Posthumus, H., Morris, J., Hess, T.M., Neville, D., Phillips, E., Baylis, A., 2009. Impacts of the summer 2007 floods on agriculture in England. J. Flood Risk Manag. 2 (3), 182–189. http://dx.doi.org/10.1111/j.1753-318X.2009.01031.x.

Potter, K., 1991. Hydrological impacts of changing land management practices in a moderately sized agricultural catchment. Water Resour. Res. 27 (5), 845–855.

R Core Team, 2020. A Language and Environment for Statistical Computing, Vol. 2. R Foundation for Statistical Computing, Vienna, Austria, URL: http://www.r-project.org.

Rosgen, D.L., 1994. A classification of natural rivers. Catena 22 (3), 169–199. http://dx.doi.org/10.1016/0341-8162(94)90001-9.

Sayers, P.B., Hall, J.W., Meadowcroft, I.C., 2002. Towards risk-based flood hazard management in the UK. Proc. Inst. Civ. Eng. Civ. Eng. 150 (5), 36–42. http://dx.doi.org/10.1680/cien.2002.150.5.36, URL: http://www.icevirtuallibrary.com/doi/10.1680/cien.2002.150.5.36.

Sear, D.A., Millington, C.E., Kitts, D.R., Jeffries, R., 2010. Logjam controls on channel:floodplain interactions in wooded catchments and their role in the formation of multi-channel patterns. Geomorphology 116 (3–4), 305–319. http://dx.doi.org/10.1016/j.geomorph.2009.11.022.

Shields, Jr., F.D., Gippel, C.J., 1995. Prediction of effects of woody debris removal on flow resistance. J. Hydraul. Eng. 121 (4), 341–354.

Sudmeier-Rieux, K., Arce-Mojica, T., Boehmer, H.J., Doswald, N., Emerton, L., Friess, D.A., Galvin, S., Hagenlocher, M., James, H., Laban, P., Lacambra, C., Lange, W., McAdoo, B.G., Moos, C., Mysiak, J., Narvaez, L., Nehren, U., Peduzzi, P., Renaud, F.G., Sandholz, S., Schreyers, L., Sebesvari, Z., Tom, T., Triyanti, A., van Eijk, P., van Staveren, M., Vicarelli, M., Walz, Y., 2021. Scientific evidence for ecosystem-based disaster risk reduction. Nat. Sustain. 4 (9), 803–810. http://dx.doi.org/10.1038/s41893-021-00732-4, URL: https://www.nature.com/articles/s41893-021-00732-4.

Thaler, T., Hudson, P., Viavattene, C., Green, C., 2023. Natural flood management: Opportunities to implement nature-based solutions on privately owned land. Wiley Interdiscip. Rev. Water http://dx.doi.org/10.1002/wat2.1637.

The Rivers Trust, 2021. The rivers trust NFM project monitoring & evaluation tool. URL: https://www.arcgis.com/apps/.

Thomas, H., Nisbet, T.R., 2012. Modelling the hydraulic impact of reintroducing large woody debris into watercourses. J. Flood Risk Manag. 5 (2), 164–174. http://dx.doi.org/10.1111/j.1753-318X.2012.01137.x.

UK Centre for Ecology and Hydrology, 2021. National river flow archive. URL: https://nrfa.ceh.ac.uk/.

USACE, 2020. HEC-RAS 5.0.7. URL: http://www.hec.usace.army.mil/.

Uttley, C., Skinner, A., 2017. CIEEM Webinar: Natural Flood Management: Policy and Practical Implementation. CIEEM, URL: http://www.cieem.net/events/.

van Leeuwen, Z.R., Klaar, M.J., Smith, M.W., Brown, L.E., 2023. Quantifying the natural flood management potential of leaky dams in upland catchments, Part I: a data-based modelling approach. J. Hydrol. http://dx.doi.org/10.1016/j.jhydrol.2023.130448.

Vira, B., Adams, W.M., 2009. Ecosystem services and conservation strategy: beware the silver bullet. Conserv. Lett. 2 (4), 158–162. http://dx.doi.org/10.1111/j.1755-263x.2009.00063.x.

Watson, F., Vertessy, R., McMahon, T., Rhodes, B., Watson, I., 2001. Improved methods to assess water yield changes from paired-catchment studies: Application to the Maroondah catchments. Forest Ecol. Manag. 143 (1–3), 189–204. http://dx.doi.org/10.1016/S0378-1127(00)00517-X.

Waylen, K.A., Holstead, K.L., Colley, K., Hopkins, J., 2018. Challenges to enabling and implementing Natural Flood Management in Scotland. J. Flood Risk Manag. 11, 1078–1089. http://dx.doi.org/10.1111/jfr3.12301.

Wells, J., Labadz, J.C., Smith, A., Islam, M.M., 2020. Barriers to the uptake and implementation of natural flood management: A social-ecological analysis. J. Flood Risk Manag. 13 (Suppl. 1), 1–12. http://dx.doi.org/10.1111/jfr3.12561, URL: https://onlinelibrary.wiley.com/doi/10.1111/jfr3.12561.

Wenzel, R., Reinhardt-Imjela, C., Schulte, A., Bölscher, J., 2014. The potential of in-channel large woody debris in transforming discharge hydrographs in headwater areas (Ore Mountains, Southeastern Germany). Ecol. Eng. 71, 1–9. http://dx.doi.org/10.1016/j.ecoleng.2014.07.004.

Wilby, R.L., Beven, K.J., Reynard, N.S., 2008. Climate change and fluvial flood risk in the UK: more of the same? Hydrol. Process. 22 (1), 2511–2523. http://dx.doi.org/10.1002/hyp, arXiv:arXiv:1011.1669v3.

Wilkinson, M.E., Addy, S., Quinn, P.F., Stutter, M., 2019. Natural flood management: small-scale progress and larger-scale challenges. Scott. Geogr. J. 135 (1–2), 23–32. http://dx.doi.org/10.1080/14702541.2019.1610571.

Wingfield, T., Macdonald, N., Peters, K., Spees, J., Potter, K., 2019. Natural Flood Management: Beyond the evidence debate. Area 51 (4), 743–751. http://dx.doi.org/10.1111/area.12535, URL: http://doi.wiley.com/10.1111/area.12535, https://onlinelibrary.wiley.com/doi/10.1111/area.12535.

WMO, 2009. Manual on Low-flow Estimation and Prediction. Operational Hydrology Report No. 50 Technical Report, WMO, Geneva.

Yorkshire Dales Rivers Trust, 2018. Naturally Resilient - Slowing the Movement of Water - Leaky Dams. Technical Report, Interreg, pp. 2–4, URL: https://www.ydrt.org.uk/wp-content/.