Received: 30 June 2020 | Revised: 19 January 2021 | Accepted: 11 February 2021
DOI: 10.1111/jfr3.12703

ORIGINAL ARTICLE

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management | WILEY

The effect of depth-duration-frequency model recalibration on rainfall return period estimates

Gianni Vesuviano[1] | **Elizabeth Stewart**[1] | **Peter Spencer**[2] | **James D. Miller**[1]

[1]UK Centre for Ecology & Hydrology, Wallingford, UK

[2]Environment Agency, Richard Fairclough House, Warrington, UK

Correspondence
Gianni Vesuviano, UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, OX10 8BB, UK.
Email: giaves@ceh.ac.uk

Funding information
Environment Agency; UKCEH National Capability

Abstract

In November 2009 and December 2015, two record-breaking 24-hr rainfalls occurred in Cumbria, UK, significantly changing the perception of flood risk for local communities. FEH13, the current UK rainfall depth-duration-frequency (DDF) model, estimated return periods of around 1,000 years for both events. The previous model, FEH99, received criticism from panel engineers responsible for making technical safety decisions relating to reservoirs for appearing to estimate relatively short return periods for extreme events. Although FEH13 is more consistent with current probable maximum precipitation (PMP) estimates, there is high uncertainty in both models due to the limited number of extremes captured by UK rain gauges. Furthermore, neither model included the 2009 or 2015 event in its calibration. Here, we re-calibrate FEH13 using additional gauged rainfall data collected in Cumbria during 2006–2016, including the record-breaking 2009 and 2015 storms. Using the updated calibration data set reduces the estimated return periods of the 2009 and 2015 events to approximately 140 years each. This case study illustrates the considerable uncertainty in short-sample records, demonstrates the importance of maximising the quantity of relevant calibration data, shows that perception of risk depends upon the method and data used, and illustrates the difficulty of separating trends and natural variability.

KEYWORDS

climate change, extreme events, mapping of hazard and risk, risk perception

1 | INTRODUCTION

From 0000 UTC on November 19, 2009, 316.4 mm of rainfall was measured in one 24-hr period at Seathwaite Farm in Cumbria, setting a new UK record and surpassing the previous record, from July 1955, by 37 mm (although that record was accumulated over a standard rainfall day, from 0900 UTC to 0900 UTC). The event at Seathwaite Farm was notable for its sustained character, with consistent hourly rainfall rates and no significant peaks (Met Office, 2012). In addition to the new 24-hr record, the Seathwaite Farm event formed the core of new three-day and four-day rainfall records, of 456.4 mm and 495.0 mm, respectively. This extreme rainfall affected 2,239 properties, 3,057 businesses, 250 farms, 25 bridges and 40 waste water treatment works, and resulted in one fatality (UK Government, 2013). It also occurred less than five years after the January 2005 floods, during which the flood level in Carlisle, Cumbria's largest town, reached at least 1 m higher than at any time since 1771 (Environment Agency, 2006).

From 1800 UTC on December 4, 2015, 341.4 mm of rainfall was observed in one 24-hr period at Honister Pass, less than 2 km away from Seathwaite Farm, breaking the UK 24-hr record by 25 mm just 6 years after it was last set. At nearby Thirlmere Reservoir, 405 mm was observed in one 38-hr period, breaking the previous two-day rainfall record. This extreme event directly flooded approximately 6,000 properties and 1,000 businesses, and impacted over 350 km of roads, 600 farms and 70 waste water treatment works (Cumbria County Council, 2018). The December 2015 floods in Cumbria occurred just over 3 years after a series of floods between June and November 2012, which themselves came less than 3 years after the November 2009 floods.

Both the 2009 and 2015 rainfalls were widely described as unprecedented, and rainfall totals within certain durations were genuinely record-breaking. Sediment analysis from Bassenthwaite Lake (all of the places mentioned in this paper are mapped in Figure 1) shows that both the peak flow into the lake in November 2009 and the frequency of large floods since 1990 is unparalleled at any time since at least 1,460 (Chiverrell et al., 2019). Miller, Kjeldsen, Hannaford, and Morris (2013) estimated a return period of over 50,000 years for the peak flow gauged in the River Derwent at Camerton, just upstream of Workington, in November 2009, falling to 771 years when that peak flow was included in the analysis. Both these analyses seem to suggest that floods of this extremity should only happen a couple of times every thousand years, implying that the extreme rainfalls driving these events are similarly rare. However,

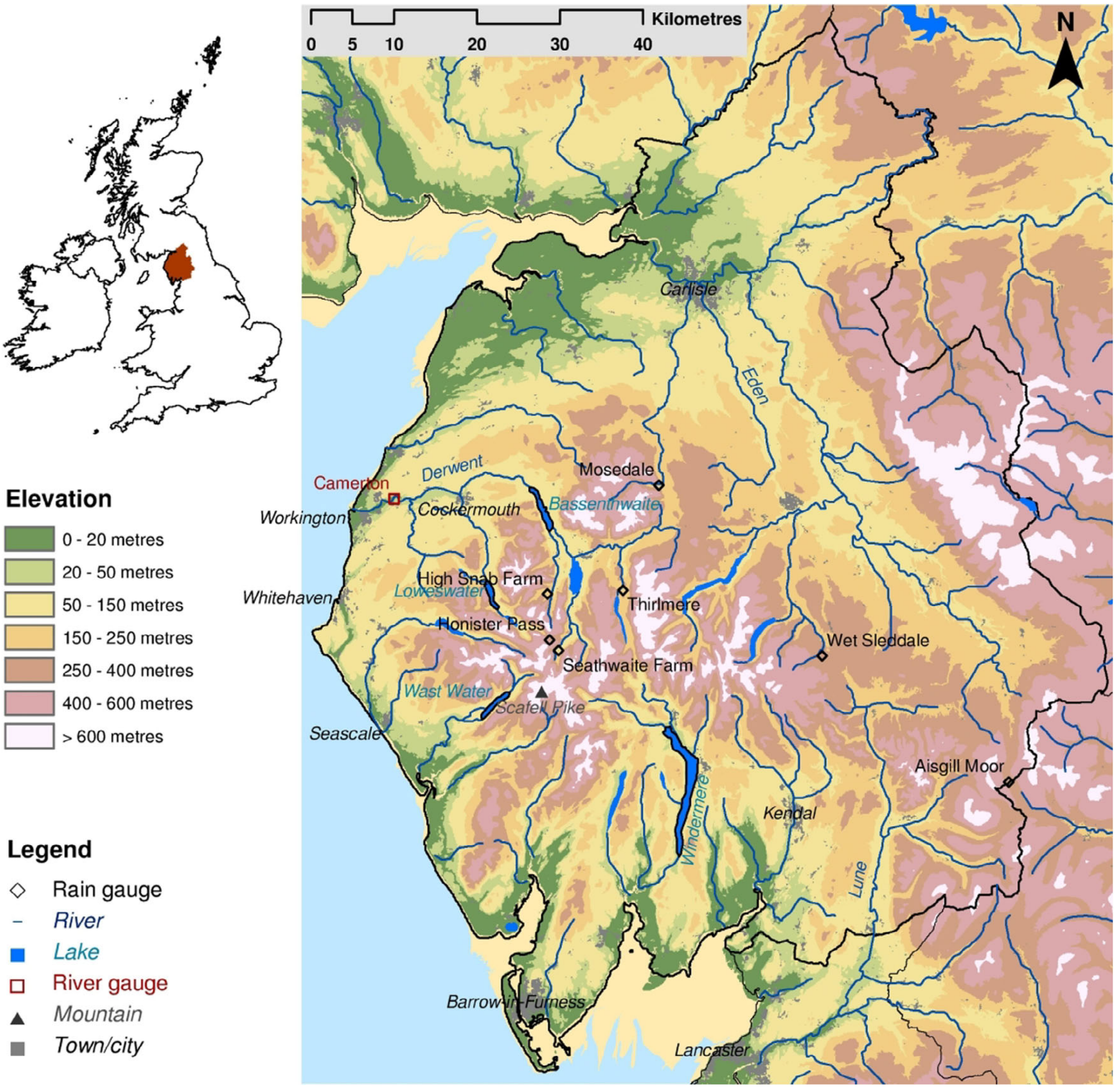

FIGURE 1 Map of the British Isles, with Cumbria highlighted in orange (top left), and map of Cumbria, outlined in black, including all named features mentioned in this paper (right). Contains OS data © Crown copyright and database right 2018

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

the occurrence of two record-breaking rainfalls in 6 years has led to questions about whether the assumed rarity of such events has changed recently.

In the UK, the return period (or average recurrence interval) of extreme rainfall events is estimated using a national rainfall depth-duration-frequency (DDF) model. The current UK DDF model is one of the suite of methods making up the Flood Estimation Handbook (FEH: Institute of Hydrology, 1999 and subsequent updates) and is known as FEH13 (Stewart et al., 2013). However, the rain gauge data on which the FEH13 model is based only go up to 2006 and therefore do not include either of the record-breaking November 2009 or December 2015 events.

In light of these more recent extreme events, and given evidence that such events may become more common in the future due to climate change (Osborn & Maraun, 2008), it was deemed a priority to understand how rainfall frequency relationships in flood-affected Cumbria are impacted by the inclusion of these and other recent rainfall data in FEH13 model calibration. A small study on this topic was conducted, funded jointly by the UK Centre for Ecology & Hydrology and the Environment Agency (England). This paper describes the study, discussing the background to rainfall frequency modelling methods used in the UK and presenting details of the study area and the available data. The key results of the recalibration are outlined and the flood risk implications for communities in Cumbria are discussed. This paper is a summary of a longer and more detailed report (Vesuviano & Stewart, 2021). The rainfall data collected for the recalibration will be pooled with additional rainfall data collected since the end of this study and used for a UK-wide recalibration of the FEH13 model, the results of which will be made available through the FEH Web Service (https://fehweb.ceh.ac.uk).

2 | RAINFALL FREQUENCY ESTIMATION IN THE UNITED KINGDOM

2.1 | History

The UK has a long history of rainfall DDF modelling for use principally in hydrological design applications and for post-event analysis. The first national model of rainfall frequency was published in 1975 in the Flood Studies Report (FSR: NERC, 1975), presenting a method for deriving depth-duration-frequency (DDF) estimates for rainfall durations from 1 min to 25 days and return periods from 6 months to 10,000 years (throughout this paper, return periods denoted in years measure the average interval between years containing events of the given magnitude, and are the exact reciprocal of the annual exceedance probability, AEP). The FSR also provided estimates of probable maximum precipitation (PMP) for the same range of durations.

The FSR DDF model was criticised for being over-generalised, masking important local and regional variations (Bootman & Willis, 1981) and was superseded by the Flood Estimation Handbook (FEH) DDF model (Faulkner, 1999), now referred to as FEH99. Soon after release of the FEH99 method, it was found that estimated rainfall depths increased rapidly when return periods were extrapolated beyond 1,000 years and that, in many cases, the FEH99 model exceeded FSR estimates of PMP at return periods below 10,000 years (Babtie Group et al., 2000; MacDonald & Scott, 2001). This caused concern for the UK's reservoir engineering community, as PMP is considered a higher reservoir design standard than the 10,000-year event (ICE, 2015). An independent assessment of the FEH99 methodology by Cox (2003) formed the basis for development of an updated FEH DDF model (see Stewart et al., 2013 for details of the study), now known as the FEH13 model (UKCEH, 2015).

2.2 | FEH13 model overview

The FEH13 model was based on an analysis of over 170,000 station-years of data from daily rain gauges throughout the UK, together with about 17,000 station-years of hourly data, substantially more data than were used in the FSR and FEH99 studies. FEH13 was developed to allow the estimation of rainfall depths falling over durations from 1 hr to 192 hr (8 days) for return periods from 2 years to over 10,000 years. Following initial development of the FEH13 model (Stewart et al., 2013), the range of return periods was extended to cover 12 months to 100,000 years, and a procedure was added to estimate rainfalls for durations as short as 5 min, for drainage design applications. Development of FEH13 retained the basic index-flood approach of the FEH99 model, whereby the rainfall frequency curve is obtained by multiplying an index variable specific to the site of interest with a regionally derived growth curve. Its main advances were:

- increased availability of rainfall maxima, particularly for sub-daily durations,
- a revised standardisation that uses standard-period average annual rainfall (*SAAR*) and northing in addition to the index variable *RMED* (the median annual maximum rainfall) to remove more of the

location-dependent variation in rainfall before combining maxima from networks of rain gauges,
- a revised spatial dependence model,
- improvements to the FEH99 FORGEX method of deriving growth curves,
- a more flexible depth-duration-frequency (DDF) model structure.

A number of stages were used in the development of the FEH13 rainfall estimates for the UK:

- Abstraction of annual maxima (AMAX) of different durations from continuous hourly and daily rain gauge data,
- Estimation of *RMED* (equivalent to the 2-year rainfall) for the same durations across the whole UK, using the median values of the abstracted AMAX series,
- Standardisation of all AMAX values, using *SAAR*, northing and estimated *RMED*,
- Pooling of standardised AMAX records, and fitting segmented lines to the maximum values of progressively larger pools (the Revised FOcused Rainfall Growth EXtension [FORGEX] methodology),
- Fitting a consistent DDF model to the Revised FORGEX lines, to ensure a monotonic relationship between rainfall depth, duration and rarity,
- Spatial smoothing of modelled rainfalls, to avoid sudden changes from point-to-point.

The DDF model is applied at each point on a regular 1-km grid, for seven durations based on hourly rainfall accumulations (1–24 hr) and four durations based on daily rainfall accumulations starting at 0900 UTC (1–8 days). Only an overview of the method is given here; more detail can be found in Vesuviano and Stewart (2021).

2.2.1 | RMED estimation

The estimation of *RMED* is a key requirement for the standardisation of AMAX data prior to pooling and involves a three-stage process. First, gauged values of *RMED* are increased to estimates of what they would be if the gauges recorded rainfall continuously, rather than between fixed intervals. These adjusted estimates are known as "fully sliding" rainfall depths. Second, a regression equation is used to estimate broad-scale variation in *RMED*. Model coefficients are expressed as functions of rainfall duration and fitted jointly at all durations; failure to do this could cause shorter-duration event depths to exceed longer-duration event depths. The regression explains 23% of variation in 1-hr *RMED*, increasing to 60% of variation in 4-hr *RMED* and 91% of variation in 8-day *RMED*. Third, the broad-scale estimates are multiplied by a correction layer, which brings the estimated *RMED* values closer to the gauged values where they exist. As the gauged *RMED* values are not necessarily the "true" values of median annual maximum rainfall at that site, the correction layer does not force exact equivalence between gauged *RMED* estimates and final gridded *RMED* estimates at any site. However, as longer records are subject to lower uncertainty, the maximum permitted deviation is proportional to the number of gauged years at a site.

2.2.2 | Revised FORGEX method

The Revised FORGEX (FOcused Rainfall Growth EXtension) method is used to estimate the rainfall depth-frequency relationship for a given duration at a given focal point, using annual maxima pooled from increasingly wider circles to define the relationship for longer return periods. It was initially described in Faulkner (1999) and extended by Stewart et al. (2013).

First, all valid annual maxima from all gauges within a circle are divided by the gridded *RMED* value relevant to their gauge and adjusted by a scaling factor. This has the effect of standardising the distributions of annual maxima between gauges. Then, for each circle, a single annual maximum record is created, which contains only the largest (standardised) event from any gauge for each year. Return periods are assigned to each event according to the number of gauges and years in a circle, and their interdependence. Valid points are plotted on a frequency plot, with standardised growth factor on the y-axis and Gumbel reduced variate on the x-axis. Finally, a segmented line is optimised that minimises squared distance to the plotted points and specifies a growth factor of 1 at an AEP of 50%. This is done for every duration, with 24 hr and 1 day considered separately.

2.2.3 | DDF model

The DDF model serves several purposes:

- To ensure that rainfall depths for any duration increase with increasing return period, and that rainfall depths for any return period increase with increasing duration.
- To allow estimates for durations in between those fitted in the Revised FORGEX method.
- To allow indicative extrapolation to longer and shorter durations that could compromise model performance

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

at the main durations of interest (1–192 hr) had they all been calibrated in a single model structure.
- To allow extrapolation to longer return periods than those estimated for the rarest events in the calibration data.

The model structure uses the 24-hr *RMED* as an index rainfall, while the growth curve consists of a weighted total of two Gamma distributions raised to a power. Both Gamma location parameters are set to give 24-hr *RMED* a nominal value of 100 at every 1-km grid location, while the scale and shape parameters are functions of event duration, giving the DDF model 11 parameters that vary spatially. Although 11 parameters may seem complex for a stationary rainfall model, the more-than 187,000 calibration points revealed a variety of rainfall depth-duration-frequency relationships that were required to be modelled. Expressing the model parameters as functions of duration ensures that rainfall depth always increases for increasing duration or return period and allows all FORGEX outputs for a grid location to be fitted jointly. Setting the DDF model output to exactly 100 for the 24-hr *RMED* means that it is modelled exactly and that the growth curve expresses rainfall depths for other durations and return periods as a "percentage of the 24-hr *RMED*". Both 24-hr and 1-day FORGEX lines contribute to fitting the model at 24-hr duration.

2.2.4 | Post-processing

Post-processing of the model outputs consists of spatial smoothing and durational smoothing. Spatial smoothing ensures that no standardised point rainfall estimate is excessively larger or smaller than one made nearby. Durational smoothing ensures that rainfall depth estimates for similar durations, made at the same point and return period, are similar. In practice, 86% and 99.94% of smoothed combinations of depth, duration, frequency and location were <1% and <5% different from their unsmoothed values, respectively.

2.3 | Uncertainty and non-stationarity

Consideration of uncertainty is an important part of risk estimation, particularly for very rare events for which calibration and verification data are very limited. While it is important and increasingly common to assign a range of values to an event (either a range of rainfall depths to one specified return period or a range of return periods to one specified depth), modifying the several stages of the modelling method to allow this was outside the scope of this project. Uncertainty is considered to some extent in the estimation of gridded *RMED*, where the gridded estimate of the true value is allowed to deviate further from the gauged estimates at rain gauges with shorter records. The asymptotic variance in any quantile of a sample drawn from a continuous probability density function is inversely proportional to the length of the sample (Mosteller, 1946), so uncertainty is also mitigated through the concatenation of records from gauges within 300 m of each other (unless it is known that doing so would not be valid).

The existing FEH13 method cannot account for potential non-stationarity in the input rainfall data, and modifying the method to allow this was also outside the scope of this project. While there is evidence (Blöschl et al., 2019; Osborn & Maraun, 2008) for trends in some properties of UK rainfall, Luke, Vrugt, AghaKouchak, Matthew, and Sanders (2017) find no conclusive evidence to recommend non-stationary models over regularly recalibrated stationary models, that is, "updated stationarity", independently of what trends are present in flood peak data in the USA. Serinaldi, Kilsby, and Lombardo (2018) conclude that no non-stationary model can be justified or specified correctly without first identifying a clear, physical and deterministic cause for the potential presence of non-stationarity. From a practical perspective, non-stationary models always have more parameters than their stationary equivalents, hence there is additional scope for inaccurate parameterisation (Faulkner, Warren, Spencer, & Sharkey, 2020). This is a very real risk, given that the stationary FEH13 DDF model has 11 parameters.

3 | STUDY AREA

Cumbria is an upland county of north-west England, with a total area of 6,767 km^2 and a population of approximately 500,000, making it one of the most-sparsely populated places in England (Cumbria County Council, 2017). Carlisle, in the north of the county, is the only city.

The Lake District National Park lies entirely inside Cumbria, occupying 2,362 km^2 or 35% of the total area of Cumbria. The Lake District is very mountainous, containing England's highest mountain, Scafell Pike (978 m), all of the land in England above 914 m (3,000 ft), and the largest and deepest lakes in England: Windermere and Wast Water respectively (Lake District National Park Authority, 2018). Bassenthwaite Lake and all of Seathwaite Farm, Honister Pass, Thirlmere, and High Snab Farm rain gauges are inside the Lake District. Cumbria also contains significant portions of the Yorkshire

Dales National Park. The mountainous topography of Cumbria brings extreme orographic enhancement to rainfall; each of the January 2005, November 2009, and December 2015 extreme events was driven by warm, moist, westerly or south-westerly airstreams associated with deep Atlantic low-pressure systems being forced onto high ground upon reaching Cumbria (Met Office, 2019; Stewart, Morris, Jones, & Gibson, 2012). Winter floods are exacerbated by generally wet winter ground conditions (Blöschl et al., 2019; Met Office, 2019).

Major rivers in Cumbria include the Eden, draining approximately 2,300 km^2 through Carlisle, and the Derwent, draining approximately 700 km^2 through Cockermouth and Workington. About half of the Lune catchment, which drains approximately 1,000 km^2 through Lancaster in Lancashire, is in Cumbria (all catchment areas from UKCEH, 2015). All of these rivers experienced extreme peak flows during the aforementioned floods of the early 21st century.

4 | DATA

The FEH13 model is calibrated with rainfall data from two types of gauge. The majority of contributing gauges are storage gauges, read daily and emptied at 0900 UTC, to produce a record of 1-day accumulations. A minority of gauges are recording gauges, which only hold a small quantity of rainfall, automatically empty when they are filled and record the times at which this occurs. The recording gauges are used to derive 1-hr accumulation records, each sample of which starts on the clock-hour.

All storage gauge records were supplied by the UK Met Office and covered the period up to 2016. The vast majority of records began in or after 1961 although some began as early as 1853. These data were quality controlled by the Met Office, according to Met Office (2001), and again by UKCEH according to Svensson, Folwell, Dempsey, Dent, and Fung (2009).

Recording gauge records were supplied by the Met Office for the whole UK for the period 2006–2013 and the Environment Agency for north-west England for the period 2005–2016. These were combined with the hourly data used in the previous calibration of the FEH13 model. The previous data were quality controlled in UKCEH according to Svensson et al. (2009) while the validity of large accumulations in the recent data was also checked.

To reduce uncertainty at sites, records from pairs of gauges of the same type (storage or recording) separated by less than 300 m were concatenated unless there was a known reason why doing so would be invalid.

Maximum accumulations for each calendar year were extracted from each concatenated record for various durations, specifically: 1, 2, 4, 6, 12, 18, and 24 hr for hourly records, and 1, 2, 4, and 8 days for daily records. Gaps in the abstracted hourly annual maximum series were filled using tabulated data from the FEH99 project wherever possible. Records with six or more annual maxima were used in the estimation of gridded *RMED* and records with nine or more were used for frequency analysis. The number(s) of records used for each stage is shown in Table 1, where ranges indicate that not all records were suitable for all accumulation durations.

Although most of these gauges are located outside Cumbria, the method requires that *RMED* be mapped simultaneously for the whole of the UK and that gauges up to 200 km away from the edge of the study area be included in the frequency analysis.

5 | RESULTS

In this study, we re-applied the FEH13 method, from *RMED* estimation to post-processing, to give recalibrated FEH13 rainfall estimates for Cumbria and a buffer zone around it. In this section, we first consider the results for two case study points that show contrasting behaviours. Next, we consider the changes that recalibration brings to the estimated return periods of the two record-breaking rainfalls. Finally, we present maps showing the changes in rainfall depths associated with common design standards in the UK.

5.1 | Case study

The two contrasting sites selected for the case study are Honister Pass and Aisgill Moor (Figure 1). Honister Pass is the site of the record-breaking 24-hr rainfall in December 2015 and is less than two kilometres from the previous record-holding site, Seathwaite Farm. Aisgill Moor is at the eastern edge of Cumbria and has never recorded more than 96 mm of rainfall in one rain day.

TABLE 1 Number of gauge records of each type available for each modelling stage

Record type	*RMED* estimation	Frequency analysis
Daily (storage)	8,102–8,106	7,651–7,678
Hourly (recording)	1,213–1,309	951–1,036

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

For shorter duration rainfalls at both sites, FORGEX outputs tend to show a typical arc-shaped relationship between growth factor and return period, where the line becomes steeper as the Gumbel reduced variate increases (Figure 2). For longer durations, FORGEX lines at Aisgill Moor follow the same arc-shape but become less steep as duration increases. However, FORGEX lines at Honister Pass tend to follow an S-shape, where the line initially becomes steeper with increasing Gumbel reduced variate, before becoming shallower. While a relationship between shorter durations and steeper rainfall growth curves has been observed in previous UK-wide analyses (Faulkner, 1999), no S-shaped rainfall growth relationships were identified during the original calibration of the FEH13 method. They can be explained at Honister Pass by the pooling radius structure of the FORGEX method. For sites near Seathwaite Farm and Honister Pass, the 2009 and 2015 events are included in the smallest or second-smallest network and assigned relatively short return periods. Due to their record-breaking characteristics, there are few larger events (after standardisation), so by the time the FORGEX radius is expanded enough to reach an event that is only slightly larger, the number of valid gauge-years inside the radius is so great that the slightly larger event is assigned a much longer return period. The main reason that S-shaped lines are not observed for durations shorter than 1 day is that no hour, or few hours, of either the 2009 or 2015 events was particularly intense.

In general, the DDF model structure is most appropriate for rainfall at Aisgill Moor, even though a compromise must be made where the 18-hr FORGEX line crosses the 1-day FORGEX line at return periods above 100 years. For Honister Pass, the DDF model can fit arc-shaped curves at shorter durations and S-shaped curves at longer durations, but it must assume a smooth transition between the two shapes for intermediate durations (e.g., 12 hr). This compromise also forces the 48-hr and 96-hr curves to flatten less than the FORGEX outputs.

For both sites, the vertical spacing between different DDF curves, representing different durations, does not always reflect the vertical spacing between different FORGEX outputs. The relationship between FEH13 rainfall depths of different durations is controlled by duration-dependent parameters in the DDF model. If an apparently inconsistent relationship between durations, such as that shown by the closeness of the 2-day and 4-day FORGEX lines, is a true feature of the rainfalls and not sampling error, then the form or parameters of the DDF model may need to be re-investigated.

Smoothing has little effect on the DDF curves fitted at Aisgill Moor, as they are fairly "typical" and consistently arc-shaped (Figure 3). Conversely, smoothing at Honister Pass has the effect of attenuating the differences between

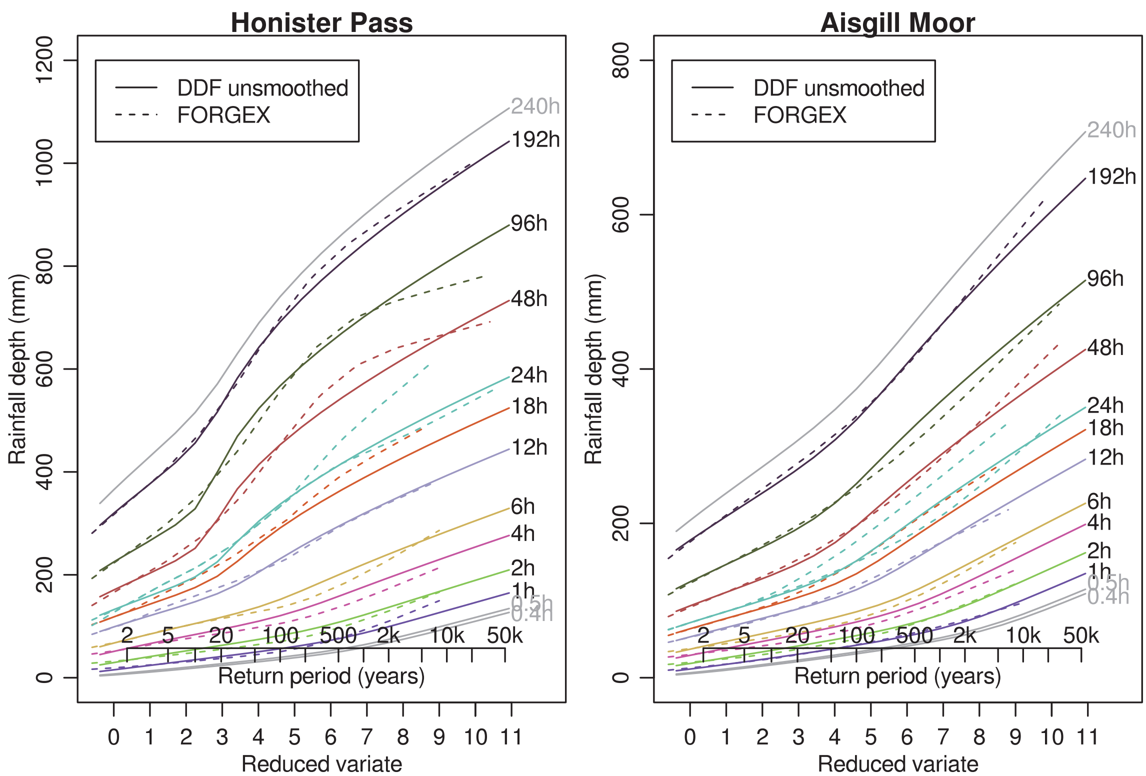

FIGURE 2 Unsmoothed DDF curves and FORGEX outputs for Honister Pass and Aisgill Moor

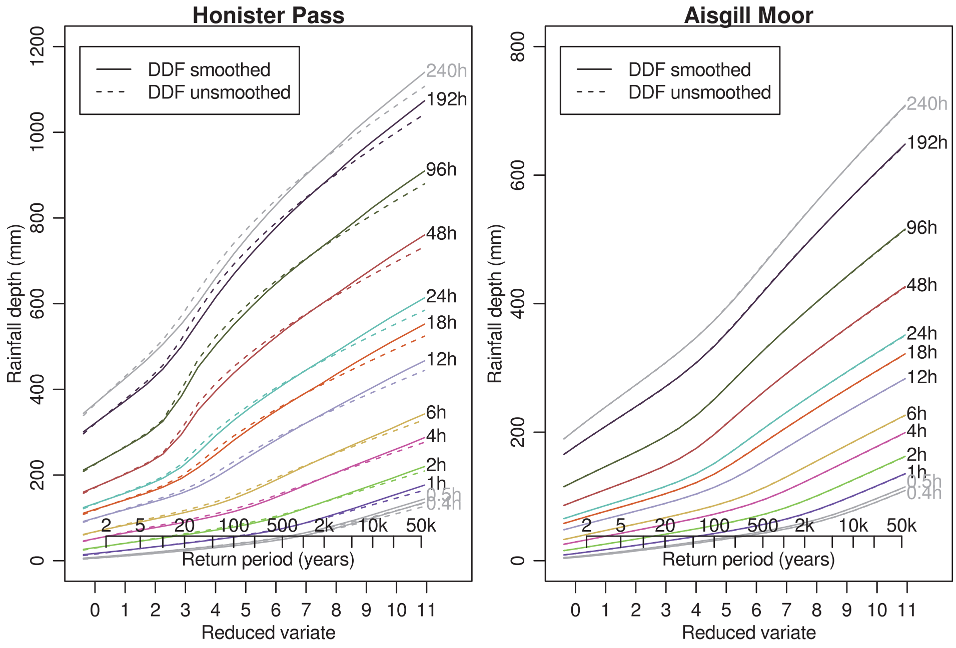

FIGURE 3 Smoothed and unsmoothed DDF curves for Honister Pass and Aisgill Moor

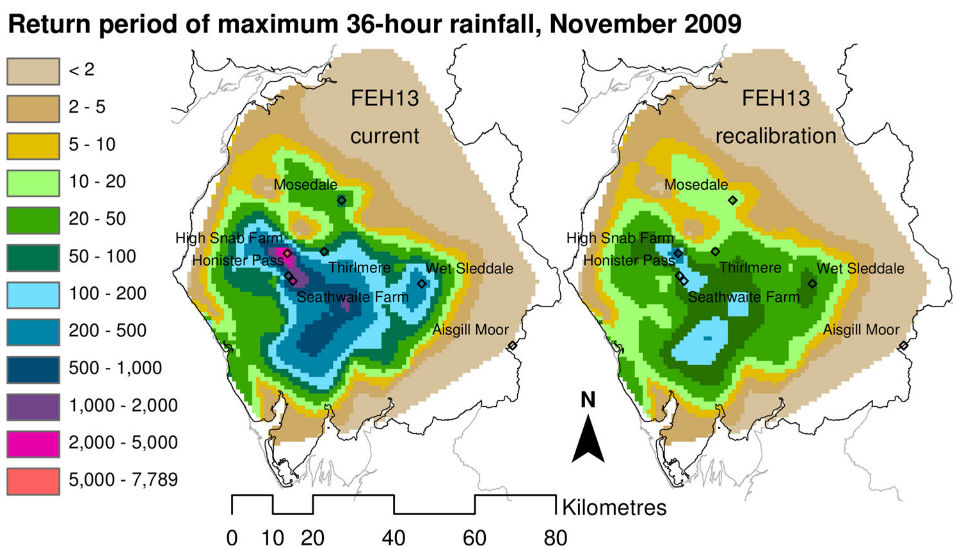

FIGURE 4 Spatial variation in return period of maximum 36-hr rainfall for November 2009 event according to current FEH13 model (left) and recalibrated FEH13 model (right). Contains OS data © Crown copyright and database right 2018

the arc-shaped and S-shaped DDF curves, further reducing the rapid rise and flattening of the longer-duration DDF curves. It is important to note that the smoothed DDF curves at Aisgill Moor and Honister Pass are highly dissimilar, meaning that differences in the pooled rainfall maxima at each site are preserved.

5.2 | Application to historical extreme events

Figure 4 shows the spatial variation in return period estimated by the new and currently available calibrations of the FEH13 model for the maximum 36-hr rainfall occurring in each grid cell during the November 2009 event, using the same estimated rainfall spatial field as Stewart et al. (2012). The overall spatial patterns are similar, with the peak return period occurring in the same grid square as the High Snab Farm rain gauge in both cases. However, the actual value is reduced 16-fold, from 7,789 to 502 years. Around the periphery, specific squares only begin to diverge for return periods longer than about 10 years. It is noted that the grid square with the longest return period (containing the High Snab Farm rain gauge) does not have the greatest rainfall accumulation over any duration; it does however have the greatest standardised rainfall accumulation over 36 hr.

CIWEM Chartered Institution of Water and Environmental Management Journal of Flood Risk Management—WILEY

TABLE 2 Return periods estimated by recalibrated FEH13 ("recal"), current FEH13 and FEH99 models for extreme events in Cumbria

				Return period (years)		
Location	**Date**	**Depth (mm)**	**Duration**	**recal**	**FEH13**	**FEH99**
Honister Pass	Dec 2015	341.4	24 clock hours	131	988	1,118
Thirlmere	Dec 2015	405.0	38 clock hours	8,293	>100 k	4,017
Thirlmere	Dec 2015	405.0	2 rain days	7,020	>100 k	4,751
Seathwaite Farm	Nov 2009	316.4	24 clock hours	150	980	160
Seathwaite Farm	Nov 2009	392.6	36 clock hours	192	2,604	172
Seathwaite Farm	Nov 2009	456.4	3 rain days	132	3,224	133
Seathwaite Farm	Nov 2009	495.0	4 rain days	113	2,847	109

Table 2 compares point rainfall return periods estimated by the recalibrated and original FEH13 calibrations and the FEH99 model over several durations for the grid squares containing the record-breaking rainfall depths of the November 2009 and December 2015 events.

Recalibration of FEH13 has greatly increased the estimated probability of the November 2009 event. However, the new return periods are not unprecedented, being similar to those estimated by the FEH99 model. Recalibration has also increased the estimated probability of the December 2015 event. However, the new return period for the maximum 24-hr accumulation at Honister Pass is far lower than that estimated by either the FEH99 or original FEH13 model. The similar return periods for this and the Seathwaite Farm event are consistent with both having standardised growth factors of approximately 2.5.

The accumulation at Thirlmere, despite being part of the same event, has a longer return period as its standardised growth factor is over four, because *RMED* and *SAAR* at Thirlmere are considerably lower than at Honister Pass (2,147 mm vs. 3,389 mm for *SAAR*, about 40% lower for all *RMED* between 6 hr and 4 days).

Despite both being 405 mm, the return period of the 38 clock-hour event is greater than that of the 2 clock-day event at Thirlmere when using the FEH13 model, while the opposite is true when using the FEH99 model. This is due to the discretisation conversion factors used by each model to scale up the depths recorded with fixed start and end times to "fully sliding" estimates of what the equivalent rainfall in an unrestricted 38-hr or two-day period could be.

Given that Cumbria has recently experienced two record-breaking rainfalls in 6 years, it is at least possible that the new return period estimates, which start around 100 years, are more plausible than the original FEH13 estimates, which start around 1,000 years and extend beyond 100,000 years.

5.3 | Design events

Figure 5 compares recalibrated rainfall depths at a range of durations and return periods typically used in reservoir design with those from the original FEH13 model. For the very shortest durations, rainfalls estimated by the recalibrated model can be smaller than those estimated by the original FEH13 model, especially in northern coastal Cumbria, and more so for longer return periods. This is because few intense short-duration events (e.g., thunderstorms) were gauged in and around Cumbria between 2006 and 2016. As a result, the largest events from the pre-2006 record are now perceived as being rarer, since they were not significantly exceeded during the 11 additional years of data collection. As duration increases, recalibrated FEH13 rainfall depth estimates begin to exceed original FEH13 estimates across the whole of Cumbria, although still more so for shorter return periods. This reflects the S-shapes of some of the recalibrated DDF relationships, in which the growth rate of rainfall depth with return period first increases and then reduces. The greatest proportional increases over the original FEH13 model are centred over three locations: Honister Pass/Seathwaite Farm, Mosedale, and Wet Sleddale. The point analysis of Honister Pass (Figures 2 and 3) shows that the 12-hr DDF curve and FORGEX lines are very similar and therefore that changes to 12-hr rainfalls are due to changes in the calibration data rather than the DDF model structure. However, there are certain mapped durations and return periods where the DDF model output is considerably below the expected position of the FORGEX output, particularly the 1-hr, 10,000-year and 36-hr, 1,000-year events. This implies that the increase in those events after recalibration would be higher if not for the DDF model attempting to unify all durations, including both arc-shaped and S-shaped FORGEX outputs in some cases.

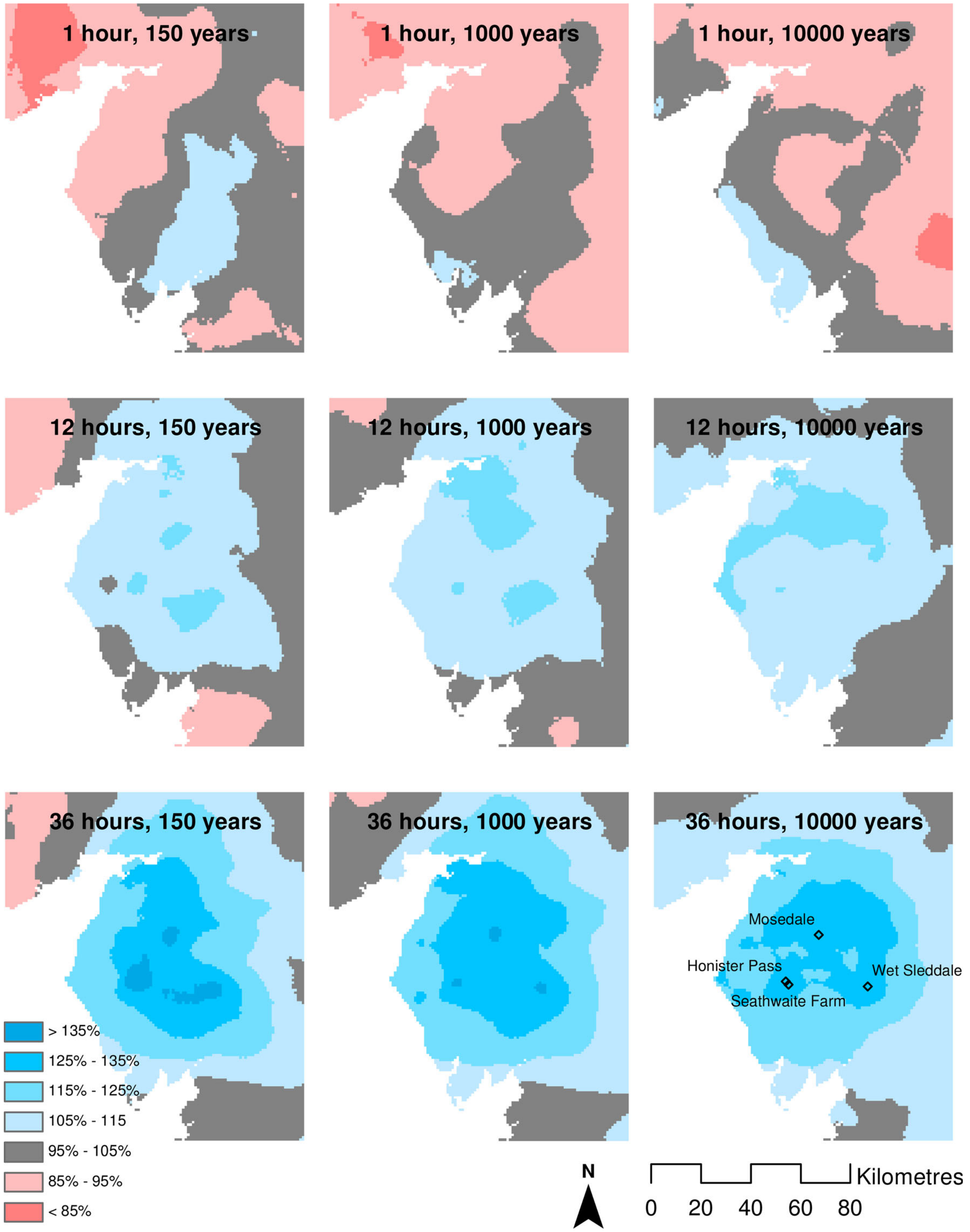

FIGURE 5 FEH13 recalibrated rainfall depths as a percentage of FEH13 current rainfall depths for 1-, 12- and 36-hr duration, 150-, 1,000- and 10,000-year return period

6 | IMPLICATIONS

Cumbria is a region of the UK that has experienced severe, extreme rainfall-driven flooding four times over the period 2005–2015. In the two most severe cases, the flooding was clearly driven by rainfall totals accumulated over 1–2 days. These are then the durations of rainfall events that are the most increased by model recalibration. Considering the depth-duration-frequency relationships given by recalibration, it is estimated that the rainfall depths associated with somewhat rare events (up to 10-year return periods) of these durations are largely unchanged, while rarer events are becoming larger – or, conversely, large events are becoming more common (e.g., Figure 4). The 150-, 1,000- and 10,000-year rainfall events, normally used for Category D, Category C and Category B reservoirs respectively, could be increased by 35% or more when 36-hr rainfall accumulations are considered (Figure 5). However, for 12-hr accumulations, more relevant to smaller (or further upstream) catchments, the proportional increase may only be 5–15%, and for 1-hr accumulations, there may be a small decrease in design rainfall. The three areas of Cumbria with the greatest proportional rainfall increases for durations of 12 hr or more are all on higher land, upstream of Cockermouth (Honister Pass/Seathwaite Farm and Mosedale) and Carlisle (Wet Sleddale), towns that have been affected severely in recent history.

This recalibration, by augmenting the original dataset with 11 extra years, should in theory be more accurate than the original calibration. However, if the occurrence of two 300+ mm daily rainfall totals in 6 years is not representative of the current average Cumbrian climate, then the extended dataset is less practical in the long-term than the original. The FORGEX and DDF fitting methods can defend somewhat against over-influence from one or two extreme events by pooling in data from a 200 km radius to put these extreme events in a wider context. However, the event standardisation against *SAAR*, northing and estimated *RMED* suggests that the Seathwaite Farm and Honister Pass events were not as rare as their record-breaking depths suggest. Unfortunately, it is impossible to know whether the return periods of such events are 100 years, 1,000 years, or longer, without maintaining rainfall records for several times the length of the return period, over which the climate is almost certain to change considerably and repeatedly. In stationary analyses, uncertainty in the return periods of extreme rainfall events, or the depths associated with extreme return periods, remains high for as long as the collection period does not exceed the return period of interest several times over. However, extending the data collection period in time threatens the assumption of a stationary climate that is required to perform stationary analyses, and the increased quantity of calibration data only reduces uncertainty due to natural variation if the climate is acceptably stationary over the whole data collection period. Simulations of the climate using numerical weather prediction models could also be used to generate very long, climate-controlled rainfall records from which extreme events can be extracted. However, developing, verifying, evaluating and running such models are all highly resource-intensive activities.

Uncertainty due to model structure is more difficult to quantify. However, recording more extreme events and using them to inform the model structure at long return periods can help to increase confidence, although this is also impossible to do without maintaining rainfall records for long enough to obtain several events of the magnitude of interest. As the model structure was unchanged before and after recalibration, this study did not make use of the increased calibration data to inform the underlying model structure.

Climate change has been shown to have a considerable role in altering the hydro-climate of the Cumbrian region, resulting in total winter precipitation increasing considerably over the 20th century (Osborn & Maraun, 2008). Otto et al. (2018) found that the effects of climate change made the December 2015 event 40% more likely to occur than under a simulated pre-industrial climate. A European-scale assessment of flood risk found increasing river flood discharges driven by increasing autumn and winter rainfall in north-west Europe, with a particular hotspot in northern England and southern Scotland (Blöschl et al., 2019). All such studies suggest that the climate of the region is changing relatively quickly, raising questions over the suitability of stationary models for rainfall DDF estimates. Using trend tests and non-stationary analysis, Faulkner et al. (2020) found that non-stationary flow estimates in north-west England were up to 55% higher than stationary estimates. This could suggest that estimated return periods for the events considered here could become shorter and therefore that similar future events could occur more frequently than stationary analysis would suggest. However, fitting a non-stationary model requires assumptions to be made about how the effects of climate change translate into specific model parameterisation and can greatly increase the complexity of model fitting and risks of equifinality, which may lead to wildly different results when extrapolating. Even assuming stationarity, Griffin, Vesuviano, and Stewart (2019) highlight the large impact that one additional year of data collection can have on flood frequency analysis, and this finding is equally applicable to rainfall frequency analysis. Taken together with the emerging evidence on regional climate change, we demonstrate the importance of routinely updating the calibration of the FEH13 DDF model to ensure that rainfall DDF estimates always use the most up-to-date and reliable data. This is of considerable importance for ensuring suitable drainage design and dam safety in the region.

7 | SUMMARY AND CONCLUSIONS

In this study, we recalibrated the UK's standard rainfall depth-duration-frequency (DDF) model, FEH13, in

Cumbria, using more up-to-date rainfall data that included the extreme rainfall events that led to severe flooding during 2005–2015. This recalibration can be considered an implementation of "updated stationarity" as the calibration data were augmented but the method was essentially unchanged. We did not attempt to compare the relative advantages and disadvantages of implementing updated stationarity methods in preference to non-stationary methods.

Inclusion of the more recent data resulted in instances where the relationship between rainfall depth and return period (specifically Gumbel reduced variate) followed an S-shape, whereas typical extreme value plots in hydrology tend to show points following either an arc or a relatively straight line. S-shaped relationships occurred at sites where the recent record-breaking rainfalls were assigned relatively low return periods, such as Honister Pass. As the S-shaped relationships are most pronounced for rainfall durations of 2 and 4 days, the data suggest that it is these durations of storm that are becoming more extreme. This is supported by maps comparing common design events across Cumbria before and after model recalibration; these show essentially no overall change in rainfall depths for short durations but increases of 35% or more for 36-hr events.When applied to the record-breaking November 2009 and December 2015 events, the recalibrated model estimates shorter return periods than does the current FEH13 model. It estimates similar return periods to the FEH99 model at Seathwaite Farm (~100–200 years) and broadly so at Thirlmere (~4,000–8,000 years), but not at Honister Pass, where the FEH99 and FEH13 models agree on a return period around 1,000 years but the recalibration estimates a return period of 131 years for the 24-hr rainfall total. Spatial analysis of the November 2009 events shows that the recalibration has little effect on the DDF relationships for somewhat rare (<10-year return period) 36-hr events but that the effect of recalibration accelerates for increasing return periods beyond 10 years.

The return periods given by the recalibrated FEH13 model are arguably more accurate as they contain more, more recent and more relevant calibration data, theoretically reducing the uncertainty resulting from natural variations in weather. However, a high level of uncertainty still remains in defining return periods in the hundreds or thousands of years using records that are typically 20–30 years long. Uncertainty in the FEH13 model can only be reduced further by obtaining more calibration data, whether gauged or simulated, updating the calibration with more extreme events as and when they occur, and using the growing dataset of these events to evaluate and inform the structure of the model at extreme return periods.

ACKNOWLEDGMENTS

This study was funded jointly by the Environment Agency and UK Centre for Ecology & Hydrology National Capability funding. The authors thank the editor, associate editor and two anonymous reviewers for their helpful comments on previous drafts of this manuscript.

DATA AVAILABILITY STATEMENT

The rainfall data provided for this study are the property of the UK Met Office and/or Environment Agency. A UK-wide re-calibration of the FEH13 model, pooling the data collected for this study with updated rainfall data for the whole UK, will be made available through the FEH Web Service (https://fehweb.ceh.ac.uk).

ORCID

Gianni Vesuviano https://orcid.org/0000-0003-2157-8875

REFERENCES

Babtie Group, CEH Wallingford and Rodney Bridle Ltd. (2000). Reservoir safety—Floods and reservoir safety: Clarification on the use of FEH and FSR design rainfalls. Final report. Glasgow: Babtie Group.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., … Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*, 108–111.

Bootman, A. P., & Willis, A. (1981). Discussion on papers 4-6. In *Flood studies report—Five years on*. London: Thomas Telford Ltd.

Chiverrell, R. C., Sear, D. A., Warburton, J., Macdonald, N., Schillereff, D. N., Dearing, J. A., … Bradley, J. (2019). Using lake sediment archives to improve understanding of flood magnitude and frequency: Recent extreme flooding in northwest UK. *Earth Surface Processes and Landforms*, *44*, 2366–2376.

Cox, D. R.. (2003). Some comments on 10,000 year return period rainfall. Report for Defra.

Cumbria County Council. (2017). *Population density, Mid-2016*. Carlisle: Cumbria County Council.

Cumbria County Council. (2018). *Flooding in Cumbria, December 2015 impact assessment*. Carlisle: Cumbria County Council.

Environment Agency. (2006). *Cumbria floods technical report: Factual report on meteorology, hydrology and impacts of January 2005 flooding in Cumbria*. Bristol: Environment Agency.

Faulkner, D. S. (1999). *Rainfall frequency estimation (flood estimation handbook volume 2)*. Wallingford: Institute of Hydrology.

Faulkner, D., Warren, S., Spencer, P., & Sharkey, P. (2020). Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. *Journal of Flood Risk Management*, *13*, e12582.

Griffin, A., Vesuviano, G., & Stewart, E. (2019). Have trends changed over time? A study of UKpeak flow data and sensitivity to observation period. *Natural Hazards and Earth System Sciences*, *19*, 2157–2167.

ICE. (2015). *Floods and reservoir safety* (4th ed.). London: Institution of Civil Engineers.

Institute of Hydrology. (1999). *Flood estimation handbook (five volumes)*. Wallingford: Institute of Hydrology.

Lake District National Park Authority. (2018). *Lake District facts and figures*. Kendal: Lake District National Park Authority.

Luke, A., Vrugt, J. A., AghaKouchak, A., Matthew, R., & Sanders, B. F. (2017). Predicting nonstationary flood frequencies: Evidence supports an updated stationarity thesis in the United States. *Water Resources Research*, *53*(7), 5469–5494.

MacDonald, D. E., & Scott, C. W. (2001). FEH vs FSR rainfall estimates: An explanation for the discrepancies identified for very rare events. *Dams & Reservoirs*, *11*(2), 28–31.

Met Office. (2001). *Quality control of rainfall observations 01/0329*. Bracknell: Met Office Observations Supply.

Met Office. (2012). *Heavy rainfall/flooding in the Lake District, Cumbria—November 2009*. Exeter: Met Office.

Met Office. (2019). *Flooding in Cumbria December 2015*. Exeter: Met Office.

Miller, J. D., Kjeldsen, T. R., Hannaford, J., & Morris, D. G. (2013). A hydrological assessment of the November 2009 floods in Cumbria, UK. *Hydrology Research*, *44*(1), 180–197.

Mosteller, F. (1946). On some useful "inefficient" statistics. *Annals of Mathematical Statistics*, *17*(4), 377–408.

NERC. (1975). *Flood studies report (five volumes)*. London: Natural Environment Research Council.

Osborn, T., & Maraun, D. (2008). *Changing intensity of rainfall over Britain (climatic research unit information sheet 15)*. Norwich: University of East Anglia.

Otto, F. E. L., van der Wiel, K., van Oldenborgh, G. J., Philip, S., Kew, S. F., Uhe, P., & Cullen, H. (2018). Climate change increases the probability of heavy rains in northern England/southern Scotland like those of storm Desmond—A real-time event attribution revisited. *Environmental Research Letters*, *13*, 024006.

Serinaldi, F., Kilsby, C. G., & Lombardo, F. (2018). Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources*, *111*, 132–155.

Stewart, E. J., Morris, D. G., Jones, D. A., & Gibson, H. S. (2012). Frequency analysis of extreme rainfall in Cumbria, 16-20 November 2009. *Hydrology Research*, *43*(5), 649–662.

Stewart, E. J., Jones, D. A., Svensson, C., Morris, D. G., Dempsey, P., Dent, J. E., … Anderson, C. W. (2013). *Reservoir safety—Long return period rainfall*. London: Department for Environment, Food and Rural Affairs.

Svensson, C., Folwell, S., Dempsey, P., Dent, J., & Fung, C. F. (2009). *Report on data consolidation, and appendices. Report No. 4 to Defra, Contract WS 194/2/39. Revised April 2009*. Unpublished.

UKCEH. (2015). *Flood estimation handbook web service*. Wallingford: UK Centre for Ecology & Hydrology.

UK Government. (2013). *Economic and business recovery—Cumbria floods 2009*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/61901/EconomicandBusinessRecovery-CumbriaFloods2009.pdf.

Vesuviano, G., & Stewart, E. (2021). *Recalibration of FEH13 rainfall model for Cumbria*. Wallingford: UK Centre for Ecology & Hydrology.

How to cite this article: Vesuviano G, Stewart E, Spencer P, Miller JD. The effect of depth-duration-frequency model recalibration on rainfall return period estimates. *J Flood Risk Management*. 2021; 14:e12703. https://doi.org/10.1111/jfr3.12703