

Identifying and interpreting extreme rainfall events using image classification

Andrew Paul Barnes ^{*}, Nick McCullen and Thomas Rodding Kjeldsen

Department of Architecture and Civil Engineering, University of Bath, Bath, UK

*Corresponding author. E-mail: apb41@bath.ac.uk

 APB, 0000-0002-1709-5304

ABSTRACT

This study presents the first attempt to identify extreme rainfall events based on surrounding sea-level pressure anomalies, using neural network-based classification. Sensitivity analysis was also performed to identify the spatial importance of sea-level pressure anomalies. Three classification models were generated: the first classifies the patterns between extreme and regular rainfall events in the North West of England, the second classifies the patterns between extreme and regular rainfall events in the South East of England, and the third classifies between the patterns of extreme events in the North West and South East of England. All classifiers obtain accuracies between 60 and 65%, with precision and recall metrics showing that extreme events are easier to identify than regular events. Finally, a sensitivity analysis is performed to identify the spatial importance of the patterns across the North Atlantic, highlighting that for all three classifiers the local anomaly sea-level pressure patterns around the British Isles are key to determining the difference between extreme and regular rainfall events. In contrast, the pattern across the mid and western North Atlantic shows no contribution to the overall classifications.

Key words: classification, extreme events, image classification, rainfall extremes, sea-level pressure

HIGHLIGHTS

- Neural networks can distinguish between extreme and regular rainfall events.
- The sea-level pressure surrounding the UK is key to distinguishing extreme events.
- The western North Atlantic does not contribute to classifying extreme events.

1. INTRODUCTION

Flooding caused by extreme rainfall events can have severe social, environmental and economic consequences. Although variations in yearly flood trends have been studied extensively (e.g. Robson *et al.* 1998; Cox *et al.* 2002; Prosdocimi *et al.* 2019), such trends do not help in identifying the processes which lead to the extreme cases. For example, in February 2020 storms Ciara and Dennis passed over the UK, resulting in up to 177 mm of rainfall in a single 24-h period (Met Office 2020), and are estimated to have resulted in insured losses of up to £200 million (Finch 2020). To provide improved risk analysis and support flood management, the processes which cause these extreme rainfall events need to be better understood and differentiated from those of regular rainfall events.

The occurrence of extreme rainfall events in the UK has a strong dependence on the concurrent and prior meteorological conditions across North Western Europe and the North Atlantic. For example, Brown (2018) shows the dependence of extreme daily rainfall in the UK on large-scale meteorological indices: North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO), El Niño – Southern Oscillation (ENSO) and Atlantic Multidecadal Oscillation (AMO). These indices represent the difference in either sea-level pressure (NAO, PDO) or sea-surface temperature (ENSO, AMO) across their specified regions. Brown found the biggest impact was made by the NAO (the difference in sea-level pressure between Iceland and the Azores), a positive NAO increases the likelihood of extra-tropical cyclones developing over the North Atlantic. This relationship is further demonstrated by Richardson *et al.* (2017) and Schillereff *et al.* (2019) which both show a negative NAO correlates strongly with increased high-river flows.

Extra-tropical cyclones are known to be the main contributors to extreme precipitation across the globe (Pfahl & Wernli 2012) and have also been linked to the development of atmospheric rivers (ARs) (Gimeno *et al.* 2020). ARs are long plumes of

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

highly concentrated water vapour in the atmosphere which originate from the mid to lower latitudes, with those affecting the UK moving upwards towards North Western Europe from the Caribbean. [Lavers *et al.* \(2011\)](#) found ARs occurred during the 10 largest floods in the UK. Following this, [Lavers & Villarini \(2013\)](#) analysed the frequency and intensity of ARs with several climate change scenarios, concluding that both the intensity and frequency of the strongest ARs are expected to increase in the future. In contrast, [Champion *et al.* \(2015\)](#) found that less than 35% of winter and 15% of summer ARs are associated with an extreme rainfall event. This highlights the need to be able to determine the difference between atmospheric phenomena which cause extreme rainfall events and those which produce more moderate rainfall events of little or no interest in flood management.

Attempts to address this need have focussed on the identification of key meteorological patterns across the North Atlantic relating to extreme rainfall events derived from large-scale climate data. [Neal *et al.* \(2016\)](#) present 30 sea-level pressure anomaly (SLPA) patterns (MO-30) identified through the application of the *k*-means clustering algorithm ([Lloyd 1957](#)). These patterns represent the types of SLPA patterns which can be present over the North Atlantic on any day. The patterns were then combined subjectively into 8 SLPA patterns (MO-8), some of which are shown to strongly correlate with the NAO. [Richardson *et al.* \(2017\)](#) investigated the applicability of these cluster sets for identifying regional precipitation and drought climatology throughout the UK, finding that the smaller set of eight clusters do not aid in explaining precipitation variability. However, magnitude variation between patterns was observed, with some patterns producing consistently higher levels of median daily rainfall across all regions of the UK. However, this does not allow an easy distinction between extreme and regular rainfall SLPA patterns in the various regions.

[Ummerhofer *et al.* \(2017\)](#) attempted to cluster SLPA patterns using the anomaly precipitation spatial variation across Europe and found a dipole in SLPA across the UK, which can determine whether precipitation anomalies in the UK will be positive in the North West or South East. This relates to the findings by [Champion *et al.* \(2019\)](#), who found that summer extreme rainfall events in the North West are typically associated with a positive SLPA region over the UK. However, no such relationship was found when analysing extremes in the South East.

The large-scale meteorological patterns being considered by the studies above are represented by images, 2D matrices with pixel colour representing the numerical value of the meteorological variable in question (e.g. SLPA). Neural networks have proved to be effective at image classification across various domains, from the classification of YouTube videos ([Karpathy *et al.* 2014](#)) to tumour feature extraction ([Yang *et al.* 2019](#)) but, up until now, they have not been used effectively to classify meteorological patterns such as distinguishing between SLPA of extreme events. This study SLPA applies neural networks to identify the differences in SLPA patterns across the North Atlantic for extreme and regular rainfall events, applied to both the North West and South East of England. In particular, the study demonstrates the differences in SLPA between:

1. Daily extreme and regular rainfall events in the North West of England.
2. Daily extreme and regular rainfall events in the South East of England.
3. Daily extreme rainfall events in the North West and South East of England.

Following this, a sensitivity analysis was conducted to compare which regions of the North Atlantic are most important to determining between the above classifications.

2. DATA

2.1. Rainfall events

Extreme and regular daily rainfall series were extracted from the administrative regions of the North West and South East of England. The rainfall series used was the UK Centre of Ecology and Hydrology's CEH-GEAR data set ([Tanguy *et al.* 2019](#)). This data set combines data from the Met Office database of observed precipitation data ([Met Office 2014](#)) using a natural neighbour interpolation method to generate a regular grid of rainfall covering the British National Grid with daily rainfall values from 1890 to 2017. The application of the natural neighbour interpolation allows grid cells to be generated based on the closest viable rainfall monitors. To extract representative rainfall for each region (North West and South East England), a regular grid of coordinates is generated for each region at 30 km intervals, bounded by the North West and South East administrative regions. [Figure 1](#) shows the position of these points in both the North West (b) and the South East (a). For each day, the rainfall magnitude for a region is given by the total rainfall at each of the points identified. The days on which the rainfall total is greater than a trace level (>0.5 mm at each location) is considered a rainfall day and is added to the non-trace rainfall series for the region. Following this, each non-trace event is standardised based on the

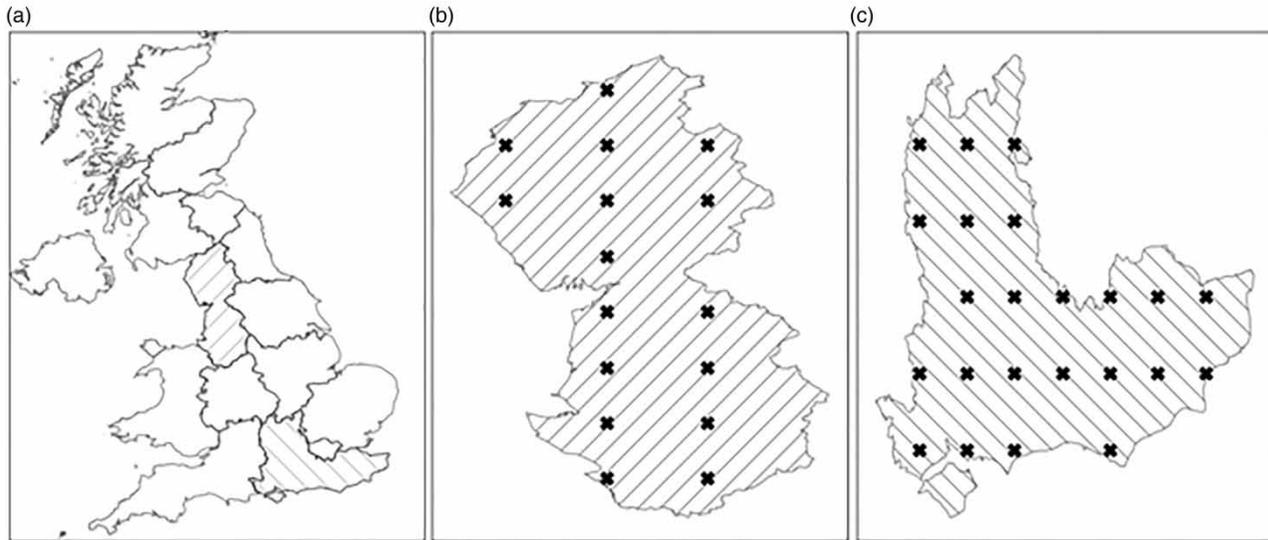


Figure 1 | Locations of the regions (a) and points from which rainfall is extracted for both the North West region (b) and the South East region (c).

mean and standard deviation of the month in which it occurs, using the following equation:

$$r_{l,m} = \frac{r_{l,m} - \overline{r_{l,m}}}{std(r_{l,m})}, l = 1, \dots, 12 \quad (1)$$

where $r_{l,m}$ represents the non-trace rainfall series for region l for all non-trace rainfall events in month m , $\overline{r_{l,m}}$ is the mean and $std(r_{l,m})$ is the standard deviation of the $r_{l,m}$ series. Following standardisation, the magnitude of each event represents the deviation of that event's magnitude from the mean of its respective location and month. Selecting extreme rainfall days from a rainfall series can be achieved by taking the maximum rainfall day in each time period (e.g. annual maximum rainfall); alternatively, a threshold can be applied such that all days with a daily rainfall quantity greater than the threshold are selected. In this study, the latter approach is used, as maxima approaches require a rainfall day to be selected in each time-period which can lead to the selection of rainfall days which in the context of the rainfall series are not extreme events such as in a particularly dry year. Selecting the events with the highest 10% of standard deviations from the mean provides a set of extreme rainfall events, the 10% of events closest to 0 can then be chosen as the regular rainfall data set as these represent the expected or mean amounts of rainfall in a given month for a given location.

This extraction resulted in 3,008 individual events (1,504 extreme and regular rainfall, respectively) in the North West and 2,290 events (1,145 extreme and regular events) in the South East. The disparity here is due to there being fewer non-trace rainfall days in the South East, with the North West having 15,046 non-trace rainfall days and the South East having only 11,450.

2.2. Meteorological patterns

To classify each event, the SLPA pattern is required across the North Atlantic. For each of the regular and extreme rainfall days for both the North West and South East regions, SLPA patterns are extracted across the North Atlantic. The patterns are extracted from the 2.5 gridded NCEP/NCAR Reanalysis 1 data set (Kalnay *et al.* 1996), with each pattern bounded between 15 and 70 latitudes and -80 and 15 longitudes.

Each of these patterns is then standardised by the monthly mean and standard deviation patterns of the given variable across the bounded region, such that:

$$\mathbf{P}_{i,m} = \frac{\mathbf{P}_{i,m} - \overline{\mathbf{P}_m}}{std(\mathbf{P}_m)}, m = \{1, \dots, 12\} \quad (2)$$

where $\mathbf{P}_{i,m}$ represents the 2D pattern for event i in month m which is standardised using element-wise operations using $\overline{\mathbf{P}_m}$ and $std(\mathbf{P}_m)$. The monthly mean $\overline{\mathbf{P}_m}$ and standard deviation $std(\mathbf{P}_m)$ are defined using all daily patterns in month m , each resulting in a 2D matrix with each cell representing a single pixel across the North Atlantic.

3. CLASSIFICATION

To classify the rainfall events, a neural network-based classification method is used, the classifiers are trained using the SLPA patterns for each event over the North Atlantic. Three classifiers are required; the first classifies the North Western extreme and regular event patterns, the second classifies the South Eastern extreme and regular patterns and finally, the third classifier classifies the extreme event patterns from both the North West and the South East. The exclusion of a fourth model distinguishing between North West and South East regular events is intentional, as the focus of this paper is on the identification and comparison of extreme events. Table 1 describes each of these classifiers (M_{NW} , M_{SE} and M_{comp}) and indicates which data sets are used in each model and which class they represent. This section introduces the neural network classification method and the optimisation procedure including how the data is split for training.

3.1. Neural network classification

A neural network consists of at least two layers of nodes connected through edges. Figure 2 shows an example architecture with three layers: an input layer, a hidden layer and an output layer. Assigned to each edge in the graph is a weight which is optimised through training and is indicated by w_x where $x = \{1, \dots, 8\}$ is the edge label.

When propagating an input vector, the input nodes in the neural network take on the values of their relevant features, and in the case of Figure 2, there are only two input nodes and hence two features in the input vector. Following this, the value propagated out of each edge can then be calculated by multiplying the value of the input node by the weight assigned to the edge, and in the case of edge 1, the value can be calculated as follows:

$$edge_1 = Aw_1 \quad (3)$$

After calculating the values of each edge, the next layer (the hidden layer in the case of Figure 2) sums the values of the edges which link to it. Following this, an activation function is used, which in this study is the rectified linear unit (ReLU) function (Equation (4)), outputting the maximum value of 0 or the sum of the edge values. ReLU is chosen as it both increases the speed of convergence in neural networks and can help improve accuracy as shown in text-based classification (Cui *et al.* 2017). The value at C can then be calculated as follows:

$$C = \max(0, edge_1 + edge_3) = \max(0, Aw_1 + Bw_3) \quad (4)$$

Table 1 | Purpose of each model and the data sets that are used

Model	Description	North West		South East	
		Extreme	Regular	Extreme	Regular
M_{NW}	Distinguishes between North Western extreme and regular events	1	0		
M_{SE}	Distinguishes between South Eastern extreme and regular events			1	0
M_{comp}	Distinguishes between North Western and South Eastern extremes	1		0	

The numbers indicate the class of the given data set in the given model.

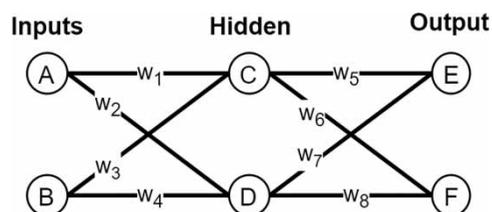


Figure 2 | An example neural network structure consisting of three layers where each node is connected via weights (as given by w_x) to the nodes in the layers both before and after.

A similar calculation can be done for node **D** and then the values at nodes **C** and **D** can be propagated forwards in the same manner to nodes **E** and **F** along their respective edges. In [Figure 2](#), two output nodes are given, indicating a binary classification, a choice of two class nodes: either **E** or **F**. For classifiers M_{NW} and M_{SE} , the outputs nodes will represent whether the input pattern is an extreme event or a regular event, and in the case of M_{comp} , these will indicate whether the pattern is a South Eastern or North Western extreme. To make the process of selection easier a softmax function (Equation (5)) is applied to the output nodes such that the output from node **E** is,

$$\text{softmax}(E) = \frac{\exp(E)}{\exp(E) + \exp(F)} \quad (5)$$

where E is the output from node **E**, F is the output from node **F** and $\exp(\dots)$ is the exponential function. Softmax forces the output to be between 0 and 1 giving a probability of the given pattern being in classes **E** and **F**.

3.2. Classifier training

Each of the three models introduced at the beginning of this section (M_{NW} , M_{SE} and M_{comp}) requires different data sets, and hence three data sets are produced consisting of a labelled set of input vectors. Each data set is represented by a matrix of n rows and m columns where n is determined by the number of events and m is the number of cells (or pixels) in a given pattern. The row which represents a given pattern is generated through the flattening of its matrix, which is done through the concatenation of each row in the matrix to each other resulting in a vector. Through this operation, each 22×38 cell pattern is converted to a single vector of size 836. Each set of events is then split into two subsets: a training and a test set. The training set consists of 80% of the input vectors and is used to train the neural network, the remaining 20% is the testing data set and is used to validate the accuracy of the network on unseen data. The actual selection of events as belonging to either the training or testing data sets is randomised.

Training the network involves the optimisation of the weights between the nodes in the network, this optimisation is done using stochastic gradient descent ([Bottou 1998](#)) and the backpropagation algorithm ([Rumelhart et al. 1986](#)). Backpropagation is a way to propagate the error calculated at the output nodes back through the weights of a neural network ([Laug & Haykin 1991](#)). To do this, an objective function is required, and in this study, the cross-entropy loss function ([Goodfellow et al. 2016](#)) is used to determine the error between the output nodes and the intended classification label (extreme or regular). The cross-entropy error will take the log difference between the true label and its calculated probability, resulting in the two output nodes containing the probability of each class being correct.

To further provide refinements to the final model, we trial models with a varying number of hidden nodes. In the example given in [Figure 2](#), only two hidden nodes are used; however, in the SLPA classifiers, the number of hidden nodes is trialled from 10 to 100 in increments of 10. This will enable the selection of a suitable number of hidden nodes for the final representative model.

4. RESULTS

The accuracy and interpretation of the three SLPA classifiers are given in this section. The first part of this section discusses the accuracy of each classifier and then the sensitivity of the classifiers to regions of the North Atlantic's SLPA.

4.1. Model accuracy

4.1.1. North west classifier (model M_{NW})

The results presented in [Figure 3\(c\)](#) show that the testing accuracies during training plateau at approximately 60%; in contrast, the training accuracies continue to increase close to 100% for classifiers with 90 and 100 hidden nodes. The classifier with 30 hidden nodes ends the 100 epochs of training with the highest testing accuracy of 62% and hence is selected for further investigation.

Investigating the $M_{NW,30}$ classifier, precision between the extreme and regular event classifications differs at 68 and 55%, respectively. This indicates the classifier is better at identifying only extreme events, with fewer false positives. However, when comparing the recall scores of each classification, which are 60% for extremes and 63% for regular events, this shows the classifier is equally good at labelling positive extreme and regular events alike.

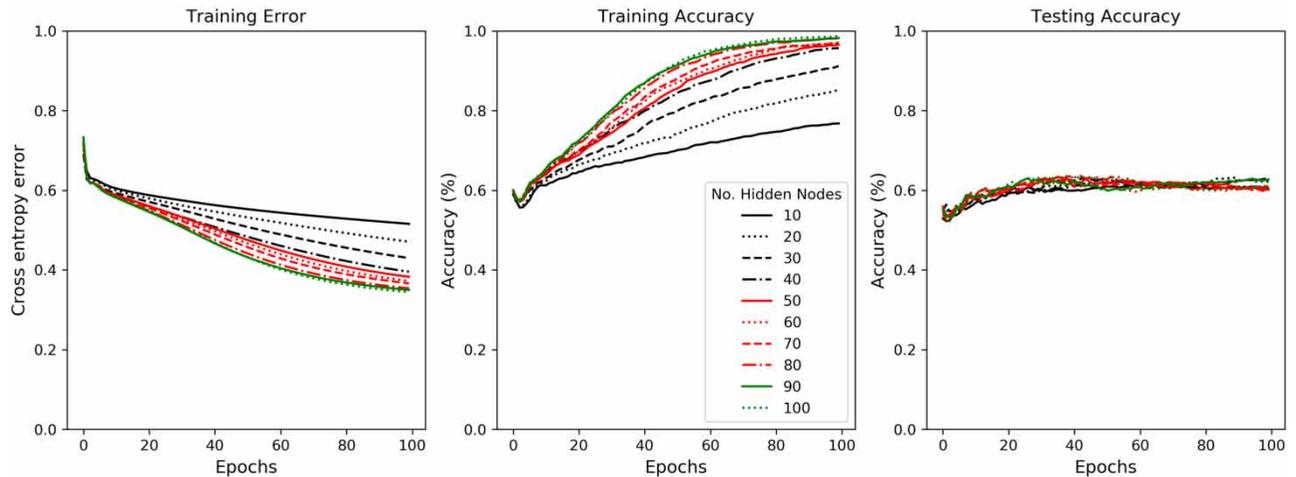


Figure 3 | Loss and accuracies for models classifying between SLPA patterns relating to North West extreme and regular rainfall events.

4.1.2. South east classifier (model M_{SE})

Similarly to the North West classifier a plateau occurs when attempting to train a neural network to classify between South Eastern extreme and regular rainfall event SLPA patterns across the North Atlantic. As presented in Figure 4, the training accuracy increases, as the training error of these classifiers decreases; however, as found in the training for M_{NW} , the testing accuracy remains consistent at around 60%.

For M_{SE} , the optimal classifier contains 10 hidden nodes ($M_{SE,10}$), this classifier finishes the 100 epochs with a 60% testing accuracy and 84% training accuracy. Breaking this into the precision and recall values, a similar trend again is seen to $M_{NW,30}$ with precision values for both regular and extreme events being similar at 55 and 59%, respectively. Further to this, the recall values also show a similar trend with the extreme event patterns having a recall higher than that with the regular event patterns (61 and 53%, respectively).

4.1.3. Extreme event classifier

The final classifier (M_{comp}) looked to distinguish between extreme event patterns for North West and South East, similarly again to M_{NW} and M_{SE} a plateau of testing accuracy occurs, as shown in Figure 5, but with a slight variability depending

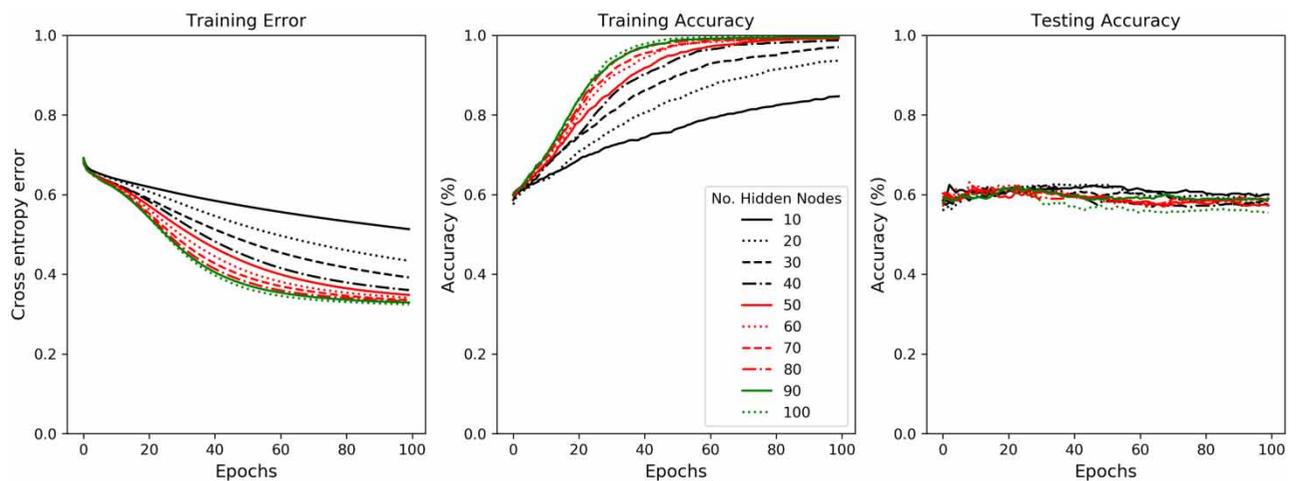


Figure 4 | Loss and accuracies for classifiers classifying between the SLPA patterns relating to South East extreme and regular rainfall events.

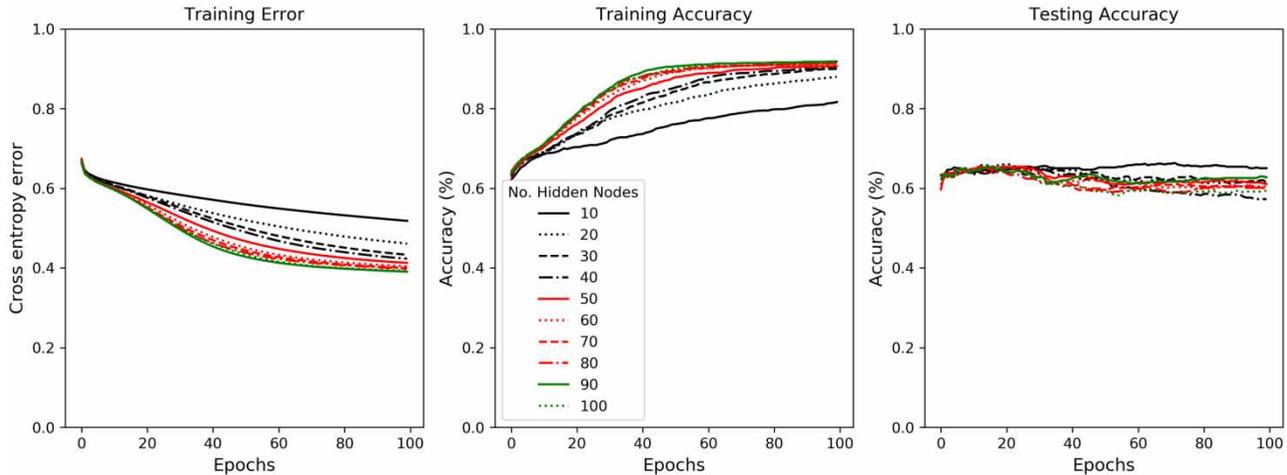


Figure 5 | Loss and accuracies for models classifying between SLPA patterns relating to extreme events in the North West and South East.

on the number of hidden nodes used. The spread of training accuracies and training errors is also marginally lower than those presented in the previous classifiers.

Despite this, the most accurate classifier has 10 hidden nodes ($M_{\text{comp},10}$) and gives a testing accuracy of 65% and a training accuracy of 81%. In contrast to the $M_{\text{NW},30}$ and $M_{\text{SE},10}$ models, the precision and recall values of identifying South Eastern extremes (66 and 67%) are both higher than those of identifying North Western extremes (55 and 54%). This is counter-intuitive as the model was trained using 31% more examples of North Western extremes, indicating that the model would have more experience classifying these types of extremes.

4.2. Spatial sensitivity

To identify the regions of interest to each classifier, a saliency map is created, representing the relative contribution of each cell (input feature) to the overall classification. To calculate the contribution of each cell to the overall classification, the back-propagation algorithm is used on a baseline image which is a pattern consisting of only 0 and a given classification, for example, an extreme event in M_{NW} . The error generated by the network is then propagated back through the network; the weights of edges will show a stronger difference if they are important to the given classification, whereas those with little relevance will not change by a comparatively large amount (Simonyan *et al.* 2014). When the errors reach the input nodes, they can be rank-ordered to identify which pixels were contributing the most to the given classification, in this study this is achieved by normalising the contribution values between 0 and 1.

Figure 6 shows the spatial contribution patterns for both $M_{\text{NW},30}$ (left) and $M_{\text{SE},10}$ (right). Both maps show little contribution from cells in the mid and western regions of the North Atlantic; however, a strong contribution is present closer to the British Isles. $M_{\text{NW},30}$ presents higher levels of contribution from both the Irish and North Seas, whereas $M_{\text{SE},10}$ presents relatively weak contributions from these regions but a higher level of contribution from the coast of Brittany in North Western France. Ummenhofer *et al.* (2017) show the difference in SLPA across the North Atlantic for various precipitation anomaly patterns; the key difference in the patterns presented is the SLPA just west and South West of the British Isles (a positive SLPA leads to negative precipitation anomalies in the North West and a negative SLPA indicates positive precipitation anomalies). Similarly, $M_{\text{SE},10}$ shows interest in the South West of the UK which continues to match the findings of Ummenhofer *et al.* (2017).

Next, the saliency map for $M_{\text{comp},10}$, which distinguishes between extreme events in the North West and extreme events in the South East, is shown in Figure 7 and presents a high level of contributions across the North of England, the Irish Sea and the coast of Brittany. This reinforces the case of local meteorological conditions creating the difference not only between extremes and regular rainfall events but also the difference between North West and South East extreme events. Further to this, both Figures 6 and 7 indicate on the day of occurrence the conditions across the rest of the North Atlantic are not contributing to the resulting classifications. This raises the question of how these contributions could change if the classifiers were trained using patterns of the days prior to an event, as it is known that certain prior meteorological conditions are common to some extreme rainfall events (Allan *et al.* 2019).

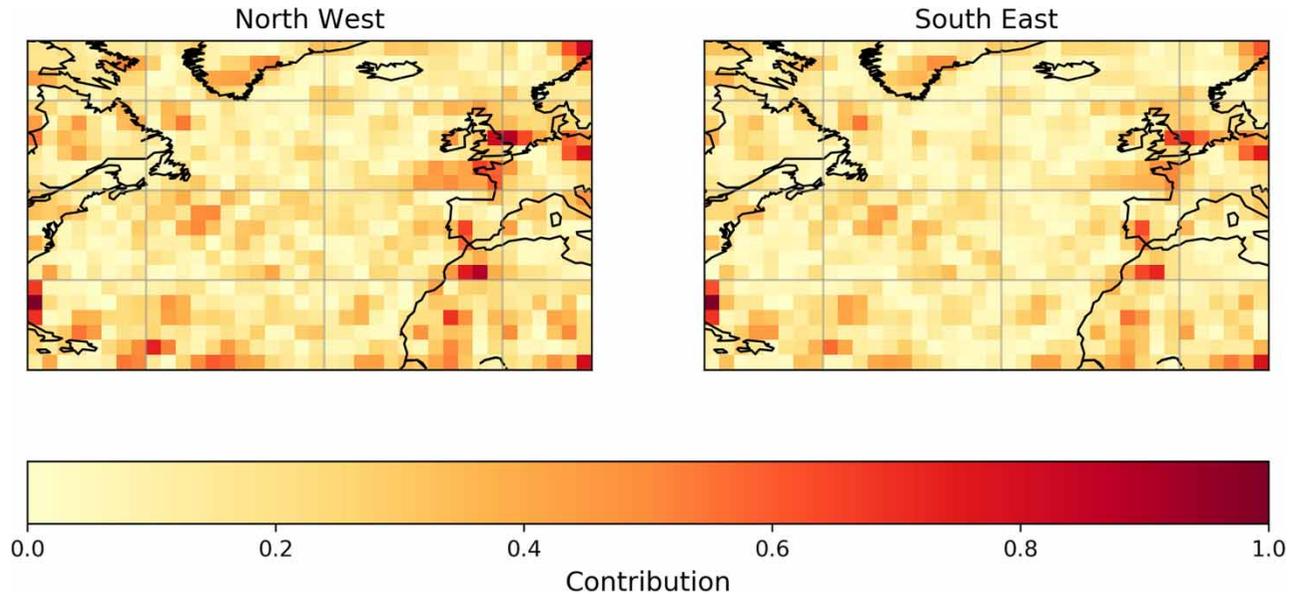


Figure 6 | Saliency map showing the spatial contributions of SLPA patterns to classifying extreme and regular rainfall events for both $M_{NW,30}$ (left) and $M_{SE,10}$ (right).

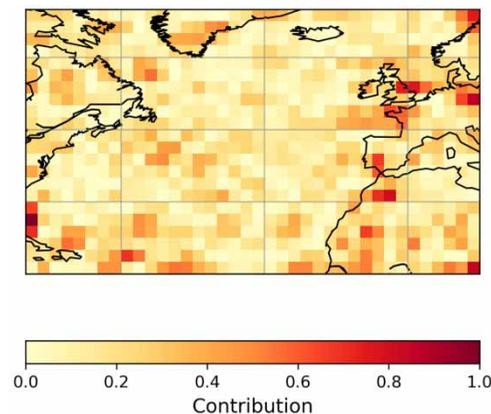


Figure 7 | Spatial contributions of SLPA patterns in determining the difference between North West and South East extreme events.

4.3. Limitations

The neural network-based approach used in this study relies heavily on the data used to train and test the models. As neural networks were trained to differentiate between extreme/regular rainfall events, hence, the training process is sensitive to how extreme events were selected. In this study, a threshold method is used, which selects the top 10% of standardised rainfall days in each region to represent extreme events. The reasons for this selection are outlined in Section 2.1; however, the use of an alternative threshold (e.g. 5%) or a maxima-based method would result in a set of models which vary greatly from the models presented here.

Furthermore, the method used for calculating the representative daily rainfall total for each region and day is the result of a trade-off, between computational time and accuracy. As presented in Section 2.1, the representative sample for each day is collated from a regular grid of points at 30 km intervals, the CEH-GEAR data set is provided at 1 km intervals; however, increasing the resolution of the regular grid would substantially increase the computational time required to calculate the average daily regional rainfall total. This trade-off was considered reasonable, as the present study only uses the magnitudes to determine the extreme and regular events within each region. If, however, we were to compare the magnitudes between regions, then a different approach using higher resolution representations would be necessary.

Finally, the sensitivity analyses shown in Section 4.2 highlight some interesting disparities between regular and extreme events in both regions. However, the interpretability and reliability of the sensitivity analysis are tied to how effective the optimisation of the neural network's parameters was during training. Hence, with further fine-tuning of the network's parameters, the resulting sensitivity analyses should reveal clearer disparities and offer further insight into the differences.

5. CONCLUSIONS

Neural network-based image classification was used to identify the concurrent, and SLPA differences across the North Atlantic between the following types of daily rainfall events for two homogenous rainfall regions in the UK:

1. Extreme and regular rainfall events in the North West of England.
2. Extreme and regular rainfall events in the South East of England.
3. Extreme rainfall events in the North West and South East of England.

Through the generation and optimisation of several neural network classifiers to represent each of the above scenarios, the following conclusions can be drawn:

1. The differences in SLPA for extreme and regular rainfall events in both the North West and South East of England are close enough to make numerical classification difficult.
 - a. Classifying between North Western extremes and regular rainfall events has an accuracy of 2% higher than classifying between South Eastern extremes and regular events (62 and 60%, respectively).
 - b. In both regional classifiers, the precision and recall of extreme events are higher than those of the regular rainfall patterns, indicating that extreme event SLPA patterns are more defined than regular events.
 - c. In determining the differences between extreme events in the North West and those in the South East, both recall and precision are higher for South Eastern extremes, indicating that the patterns relating to these events are more defined than those in the North West.
2. Saliency maps have been used to identify the spatial regions of SLPA which contribute to the classifications.
 - a. The local SLPA patterns across the British Isles are key to determining the difference between extreme and regular rainfall events in both the North West and the South East.
 - b. The mid and western North Atlantic, however, has been shown not to provide any substantial contribution to any of the classifications developed in this study.

Finally, the patterns presented in the sensitivity analysis indicate the potential for this method to be used to identify the spatial importance of meteorological variables in the days prior to extreme or regular events. This will aid meteorologists to target the regions of importance without the need to incorporate global data, reducing computational requirements. Furthermore, opening this method for the application of other meteorological variables such as precipitable water, sea-surface temperature and geopotential height may enable further inference to be gained on why extreme events are different.

ACKNOWLEDGEMENTS

Andrew Barnes acknowledges funding as part of the Water Informatics Science and Engineering Centre for Doctoral Training (WISE CDT) under the National Productivity Investment Fund (grant no. EP/R512254/1).

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories (<https://catalogue.ceh.ac.uk/documents/ee9ab43d-a4fe-4e73-afd5-cd4fc4c82556> and <https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.surface.html>).

REFERENCES

- Allan, R. J., Blenkinsop, S., Fowler, H. J. & Champion, A. J. 2019 Atmospheric precursors for intense summer rainfall over the United Kingdom. *Int. J. Climatol.* **40**(8), 1–19.
- Bottou, L. 1998 *Online Algorithms and Stochastic Approximations*. Cambridge University Press, Red Bank, NJ.
- Brown, S. J. 2018 The drivers of variability in UK extreme rainfall. *Int. J. Climatol.* **38**, 119–130.
- Champion, A. J., Allan, R. P. & Lavers, D. A. 2015 Atmospheric rivers do not explain UK summer extreme rainfall. *J. Geophys. Res.* **120** (14), 6731–6741.

- Champion, A. J., Blenkinsop, S., Li, X. & Fowler, H. J. 2019 Synoptic-scale precursors of extreme U.K. summer 3-hourly rainfall. *J. Geophys. Res. Atmos.* **124**, 4477–4489.
- Cox, D., Hunt, J., Mason, P., Wheater, H., Wolf, P. & Robson, A. J. 2002 Evidence for trends in UK flooding. *Philos. Trans. Royal Soc. A* **360**, 1327–1343.
- Cui, J., Qiu, S., Jiang, M., Pei, Z. & Lu, Y. 2017 Text classification based on ReLU activation function of SAE algorithm. *Int. Symp. Neural Netw.* **10261**, 44–50.
- Finch, J. 2020 *Storm Ciara Expected to Cost Up to £200 m in Insurance Claims*. The Guardian.
- Gimeno, L., Vazquez, M., Eiras-Barca, J., Sori, R., Stojanovic, M., Algarra, I., Nieto, R., Ramos, A. M., Duran-Quesada, A. M. & Dominguez, F. 2020 Recent progress on the sources of continental precipitation as revealed by moisture transport analysis. *Earth Sci. Rev.* **201**. <https://doi.org/10.1016/j.earscirev.2019.103070>.
- Goodfellow, I., Bengio, Y. & Courville, A. 2016 *Deep Learning*. MIT Press. Available from: <https://www.deeplearningbook.org/>
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. & Joseph, D. 1996 The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–472.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. & Fei, L. 2014 Large-scale video classification with convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1725–1732.
- Laung, H. & Haykin, S. 1991 The complex backpropagation algorithm. *IEEE Trans. Signal. Process.* **39** (9), 2101–2104.
- Lavers, D. A. & Villarini, G. 2013 The nexus between atmospheric rivers and extreme precipitation across Europe. *Geophys. Res. Lett.* **40** (12), 3259–3264.
- Lavers, D. A., Allan, R. P., Wood, E. F., Villarini, G., Brayshaw, D. J. & Wade, A. J. 2011 Winter floods in Britain are connected to atmospheric rivers. *Geophys. Res. Lett.* **38** (23). <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011GL049783>
- Lloyd, S. P. 1957 *Least Squares Quantization in PCM*. Technical Report RR5497, Bell Lab.
- Met Office. 2014 *Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853–Current)*. Available from: <https://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0> (accessed 12 July 2021).
- Met Office. 2020 *Storm Ciara*. Available from: <https://www.metoffice.gov.uk/weather/warnings-and-advice/uk-storm-centre/storm-ciara> (accessed 18 May 2020).
- Neal, R., Fereday, D., Crocker, R. & Comer, R. E. 2016 A flexible approach to defining weather patterns and their application in weather forecasting. *Meteorol. Appl.* **38**, 119–130.
- Pfahl, S. & Wernli, H. 2012 Quantifying the relevance of cyclones for precipitation extremes. *J. Clim.* **33** (12), 6770–6780.
- Prosdocimi, I., Dupont, E., Augustin, N. H., Kjeldsen, T. R., Simpson, D. P. & Smith, T. R. 2019 Areal models for spatially coherent trend detection: the case of British peak river flows. *Geophys. Res. Lett.* **46** (22), 13054–13061.
- Richardson, D., Fowler, H. J., Kilsby, C. G. & Neal, R. 2017 A new precipitation and drought climatology based on weather patterns. *Int. J. Climatol.* **38** (2), 630–648.
- Robson, A., Joes, T., Reed, D. & Bayliss, A. 1998 A study of national trend and variation in UK floods. *Int. J. Climatol.* **18**, 165–182.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. 1986 Learning representations by back-propagating errors. *Nature* **323**, 533–536.
- Schillereff, D. N., Chiverrell, R. C., Macdonald, N., Hooke, J. M., Welsh, K. E., Piliposian, G. & Croudace, I. W. 2019 Convergent human and climate forcing of late-Holocene flooding in Northwest England. *Glob. Planet. Change* **182**. <https://www.sciencedirect.com/science/article/pii/S0921818118307252?via%3Dihub>
- Simonyan, K., Vedaldi, A. & Zisserman, A. 2014 Deep inside convolutional networks: visualising image classification models and saliency maps. In *Proc. of the 2nd International Conference on Learning Representations*.
- Tanguy, M., Dixon, H., Prosdocimi, I., Morris, D. G. & Keller, V. D. J. 2019 *Gridded Estimates of Daily and Monthly Areal Rainfall for the United Kingdom (1890–2017) [CEH-GEAR]*. NERC Environmental Information Data Centre.
- Ummenhofer, C. C., Seo, H., Kwon, Y., Parfitt, R., Brands, S. & Joyce, T. M. 2017 Emerging European winter precipitation pattern linked to atmospheric circulation changed over the North Atlantic region in recent decades. *Geophys. Res. Lett.* **44**, 8557–8566.
- Yang, A., Yang, X., Wu, W., Liu, H. & Zhaunsun, Y. 2019 Research on feature extraction of tumor image based on convolutional neural networks. *IEEE Access* **7**, 24204–24213.

First received 1 March 2021; accepted in revised form 11 August 2021. Available online 27 August 2021