Stochastic simulation of daily streamflow sequences using a hidden Markov model

Douglas Pender, Sandhya Patidar, Gareth Pender and Heather Haynes

ABSTRACT

Estimation of daily streamflow time series is of paramount importance for the design and implementation of river engineering and management projects (e.g., restoration, sediment-transport modelling, hydropower). Traditionally, indirect approaches combining stochastic simulation of rainfall with hydrological rainfall–runoff models are used. However, these are limited by uncertainties in model calibration and computational expense. Thus, this paper demonstrates an alternative, direct approach, for stochastic modelling of daily streamflow data, specifically seeking to address well-known deficiencies in model capability to capture extreme flow events in the simulated time series. Combinations of a hidden Markov model (HMM) with the generalised extreme value (HMM-GEV) and generalised Pareto (HMM-GP) distributions were tested for four hydrologically contrasting catchments in the UK (Rivers Dee, Falloch, Caldew and Lud), with results compared to recorded flow data and estimations obtained from a simpler autoregressive-moving-average (ARMA) model. Results show that the HMM-GP method is superior in performance over alternative approaches (relative mean absolute differences (RMAD) of <2% across all catchments), appropriately captures extreme events and is generically applicable across a range of hydrological regimes. In contrast, the ARMA model was unable to capture the flow regime successfully (average RMAD of 14% across all catchments).

Key words | ARMA model, daily streamflow, hidden Markov model, stochastic modelling

Douglas Pender (corresponding author)
Sandhya Patidar
Gareth Pender
Heather Haynes
Institute for Infrastructure and Environment,
Heriot-Watt University,
Edinburgh EH14 4AS,
UK
E-mail: *d.pender@ymail.com*

INTRODUCTION

In recent years, ensuring the sustainability of the water resources and aquatic habitats of our river systems has become integral to legislative, management and engineering practices (European Commission 2000). Implicitly, this requires analysis of the long-term natural fluctuations (seasonality, floods, droughts and climate change) in streamflow records via backward-facing (historical data) and forward-facing (modelling) evidence-based approaches. This permits, for example: decisions on abstraction licensing or design of freshet release strategies to ensure appropriate flows for ecosystem productivity and pollution assimilation capacity; dynamic simulation of sediment/morphological response to variable flows for scheduling dredging intervention associated with flood risk management; feasibility assessments for design and operation of hydropower

schemes. As many UK rivers now offer flow gauge data over 50 years, informed sustainable river management is largely dependent on the development of robust tools which allow for the generation of synthetic streamflow sequences (past and projected) and capture the natural variability in flow. For many such projects, common practice is to provide stochastic representations of climate information (e.g., rainfall) and use a hydrological rainfall–runoff model to estimate the corresponding flows in the river channels. Although many of these models have been shown to produce reliable results with ever increasing computationally efficiency (e.g., Mellor *et al.* 2000; Dechmi *et al.* 2012; Zhang *et al.* 2013), depending on the catchment of interest, pre-existing models may be unavailable. Should this be the case, the computational time and data required for accurate

calibration and subsequent estimation of (potentially multiple) flow sequences can be extensive and, therefore, costly. An alternative methodology combines stochastic models and gauged flow data to provide a more efficient procedure for the generation of multiple daily streamflow sequences. This requires less data, is much quicker to implement at multiple sites and is, therefore, receiving increased attention from the hydrological research community.

One of the simplest, and most common, methods for stochastic simulation of river flows is the use of an autoregressive-moving-average (ARMA) model. While these have typically been applied to generate sequences of monthly- and annually-averaged flow (O'Connell 1974; Mohammadi *et al.* 2006; Sadek 2006; Musa 2013; O'Connell & O'Donnell 2014), the potentially large day-to-day variation in flow values means that the use of ARMA models for simulating daily streamflow sequences is far less common. Although there have been recent studies (Martins 2011; Can *et al.* 2012) these were focussed on international rivers with large (>20,000 km^2) catchments and a much stronger seasonality (wet and dry seasons) than occurs in the UK. To the best of the authors' knowledge their use for smaller catchments or UK rivers has not been robustly tested.

A more complex approach, also employed in modelling daily streamflow, is the use of Markov models or variations (Yapo *et al.* 1993, Xu *et al.* 2001, 2003, Szilagyi *et al.* 2006, Augustin *et al.* 2008). In their case study of the River Tees, UK, Augustin *et al.* (2008) proposed the use of a Markov model within a wider, multinomial logit, modelling framework to generate daily streamflow. However, while the underlying transition of the flow values was described by a Markov type model, the estimation of state transition probabilities is complex (with a heavy dependence on a range of associated climate variables observed at time t and in the past) and aligns more with regression modelling rather than a traditional Markov model (MM). Also, as the precise flow rates within each state were estimated by random sampling of values from the associated continuous distribution, sampling bias towards values of high probability is possible; hence introducing errors and uncertainties into the approach.

To overcome these shortfalls, the current paper uses a hidden Markov model (HMM) based solely on statistical properties of daily averaged streamflow data. This approach to flow series simulation implicitly includes a more robust sampling methodology, via systemic selection of all values based on their probabilities. The flexible structure of the model permits comparison of a traditional HMM combined with the generalised extreme value GEV (HMM-GEV) or generalised Pareto (HMM-GP) distributions to better replicate the entire range of flows. As all previous studies have been restricted to single test catchments or comparison of sites with low variability in the flow regime, a further objective of the paper is to provide a robust assessment of the HMM's generic suitability to a number of catchments exhibiting a range of hydrological regimes. To achieve this, development of a meaningful, non-subjective and common method for partitioning the flow records also requires consideration.

STUDY CATCHMENTS

The rivers chosen to develop and assess the modelling techniques are the Dee (Aberdeenshire), Falloch (Stirlingshire), Caldew (Cumbria) and Lud (Lincolnshire), with Figure 1 showing the location of the gauging stations, within the UK. The geographical spread, catchment characteristics, variation in hydrological regime and the fact that each site has at least 30 years of flow measurements, provides the reasoning behind their selection. The flow characteristics for each catchment are provided in Table 1, with Figure 1(a)–1(d) showing the monthly average flows and an example annual hydrograph from 1995.

River Dee

The Dee originates in the Cairngorm mountain range at an elevation of 1,309 m above ordnance datum (AOD) and flows east through Braemar, Ballater, Aboyne, Banchory and Aberdeen before discharging into the North Sea. The size and location of the catchment means that the geography and land use varies considerably from source to mouth; with regions of mountains, moorland, forestry and arable farm land.

Like many rivers in this part of the UK, the Dee is well-known for salmon fishing. This attracts anglers from all over the world and provides a valuable source of income for the
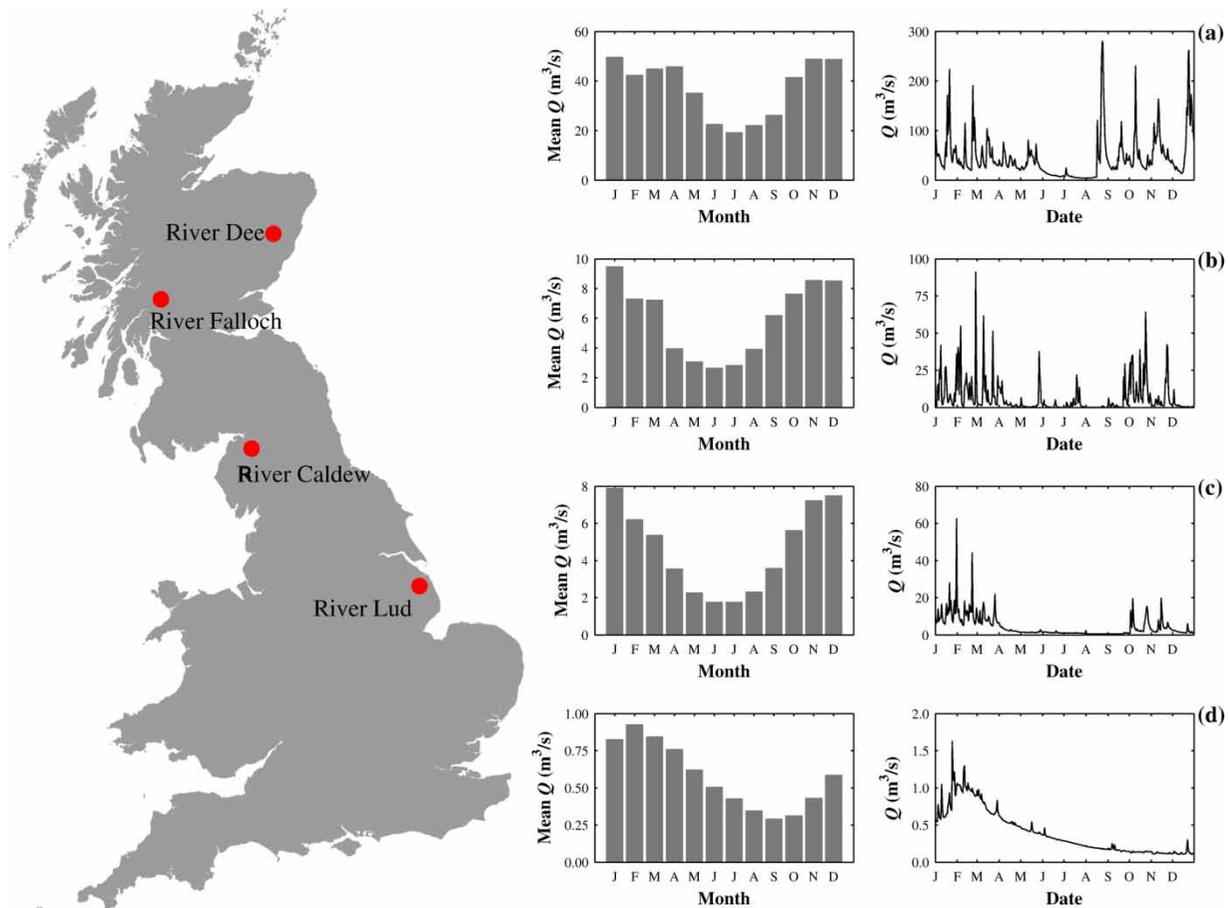
Figure 1 │ Location of gauging stations used in the study within the UK and mean monthly flows (L panel) and 1995 hydrographs (R panel) recorded on (a) River Dee, (b) River Falloch, (c) River Caldew and (d) River Lud.

Table 1 │ Characteristics of the recorded flow data used in the study

River	Record length (years)	Base flow index	$Q_{max\%}$ (m³/s)	$Q_{10\%}$ (m³/s)	$Q_{50\%}$ (m³/s)	$Q_{70\%}$ (m³/s)	$Q_{95\%}$ (m³/s)
Dee	83	0.53	648.50	73.91	26.52	18.49	8.29
Falloch	42	0.19	123.60	16.22	2.17	1.00	0.27
Caldew	32	0.49	96.39	10.10	2.69	1.66	0.75
Lud	46	0.90	5.64	0.89	0.35	0.24	0.13

local economy. As such, it is important that the timing of spates, maintenance of flows and wider competition for water resources are managed in a manner to ensure long-term sustainability of the industry. The ability to generate long-term synthetic flow sequences, on rivers such as the Dee, can therefore be used in conjunction with habitat models to allow for the creation and implementation of future management plans.

The flow gauge is located at Woodend to the west of Banchory (OS grid reference NO634956) and has been active since 1929, with the data used in this study being recorded between 1929 and 2012. At this location the catchment area is 1,370 km². The Flood Estimation Handbook (FEH) catchment descriptors indicate that the mean annual rainfall is 1,108 mm and 0.05% of the catchment is considered urban.

River Falloch

The Falloch is located near Crianlarich in Glen Falloch, Stirlingshire. It is the main river that drains the surrounding mountains (maximum elevation 1,120 m AOD), before discharging into Loch Lomond at its northern end. Owing to its location, the catchment is mainly mountainous characterised by moorland and small amounts of forestry at the lower levels.

With the requirement for sustainable energy sources becoming increasingly important, numerous hydropower schemes have been constructed within the West Highlands of the UK. Within Glen Falloch itself, there has been recent development of schemes on some of the smaller watercourses (Loch Lomond & Trossachs National Park 2012), with a run-of-the-river scheme proposed for the Falloch itself (The Scottish Government 2009); although this has yet to be commissioned. The generation of synthetic flow sequences, for rivers like the Falloch, can be used by engineers to provide a more detailed assessment of future hydropower schemes, ensuring their long-term sustainability.

The gauge is located in Glen Falloch (OS grid reference NN3121194) and has been active since 1970, with the flow data used in this study being recorded between 1970 and 2012. The FEH catchment descriptors show that the mean annual rainfall is 2,848 mm and 0.00% of the 80 km^2 catchment is considered urban.

River Caldew

The Caldew is located in Cumbria and has its source on Skiddaw Peak (926 m AOD) in the Lake District. The river then flows north into the town of Carlisle where it reaches its confluence with the River Eden. The catchment is characterised by heath and moorland in the headwaters, with both arable farming and urban centres in the lower reaches.

The Caldew is a relatively steep (channel bed gradient > 1:500) gravel bed river with a highly active sediment transport and morphological regime. Historically, it has been responsible for flooding parts of Carlisle and, as such, has been the subject of numerous flood risk studies (Neal et al. 2009, 2013; Mason et al. 2009; Horritt et al. 2010). After the floods in 2005, an alleviation scheme was constructed within Carlisle to protect the town from future flood events on the rivers Eden, Caldew and Petteril. Owing to the active morphological regime of the Caldew, a sediment transport modelling study was commissioned by the Environment Agency (EA) (Jacobs 2007), and some gravel bars are now regularly monitored to ensure levels do not undermine the flood defences in Carlisle. The ability to simulate long-term synthetic flow sequences, combined with up-to-date morphological modelling, would allow for more comprehensive sediment management plans to be implemented for rivers where sediment transport affects flood risk, such as the Caldew.

The gauge is located at Holm Hill approximately 10.5 km upstream of Carlisle (OS grid reference NY378468) and was active between 1968 and 2000. Since the closure of the gauge in 2000 a new one has now opened further downstream at Cummersdale (OS grid reference NY394527). As the historic gauge at Holm Hill has a longer flow record (31 years compared to 14 years) these measurements are used in the study. The FEH catchment descriptors show that the mean annual rainfall is 1,213 mm and 0.13% of the 147 km^2 catchment is considered urban.

River Lud

The Lud is located in Lincolnshire and flows east then northeast from its source in the Wolds hill range (maximum elevation of 153 m AOD). During the 1700s the river was canalised just downstream of the town of Louth, with the majority of the flow being diverted into this new canal system. Although the canal is no longer used for inland navigation, it is still an integral part of the drainage system for the region. As such, the EA maintains a gauging station to measure the discharge in the canal.

Apart from the town of Lough, the catchment is mainly rural, dominated by arable farming. It is predominantly (c. 70%) chalk, meaning that a high proportion of the river flow is generated from stored sources; resulting in a higher baseflow index (BFI) than the other rivers (0.90 compared to 0.53, 0.19 and 0.49). This means that, compared to the other catchments, a relatively high baseflow is maintained during dry periods. This, combined with the good chemical quality and low turbidity of the water, makes them notable for providing important habitats for a wide range of fish,

invertebrate and plant species (Wilby *et al.* 1998; Mainstone 1999; Cranston & Darby 2004).

The gauge is located in the town of Louth (OS grid reference TF337879) and has been active since 1968, with the flow data used in this study being recorded between 1968 and 2013. The FEH catchment descriptors show that the mean annual rainfall is 698 mm and 2.56% of the 55 km^2 catchment is considered urban.

Site comparison

Comparing Figure 1(a)–1(d) it is evident that each catchment experiences the same seasonal variation of drier summer and wetter winter months; with the average summer (May–August) flow being 50%, 35% and 27% of that of the winter (November–February) for the Dee, Falloch and Caldew, respectively. Owing to the Falloch catchment being very mountainous, the hydrograph has a flashier nature than those of the other rivers; in evidence, Figure 1(b) exhibits sharper transitions (i.e., steeper rising limbs to peak flows) and appears most prone to summer convective storms (demonstrating a number of high flow events in mid-summer, June–August). This is also confirmed by the Falloch having the lowest BFI (0.19), which indicates little groundwater storage of the rainfall. In addition, the higher altitude and latitude of the Dee and Falloch catchments give rise to more notable snowmelt, in March and April specifically.

To better understand the difference in the natural hydrographs and provide an effective comparison between the sites, the recorded flows have undergone a unity based normalisation (Equation (1)) and log-transformation (log$_e$). This allows for easier visual comparison between the data sets. Comparing the probability densities for the recorded flows (Figure 2), the different hydrological nature of each catchment becomes evident. Owing to the predominant, rain inducing, low pressure weather systems originating in the Atlantic Ocean, there are considerable differences between the meteorology and climate in eastern and western regions (Kingston *et al.* 2006). This is visible in Figure 3, with the Dee (north east Scotland) and Lud (eastern England) having much narrower distributions than the Falloch and Caldew (south west Scotland and north west England, respectively). This indicates that the hydrographs for rivers in the west of the country are more variable in nature,
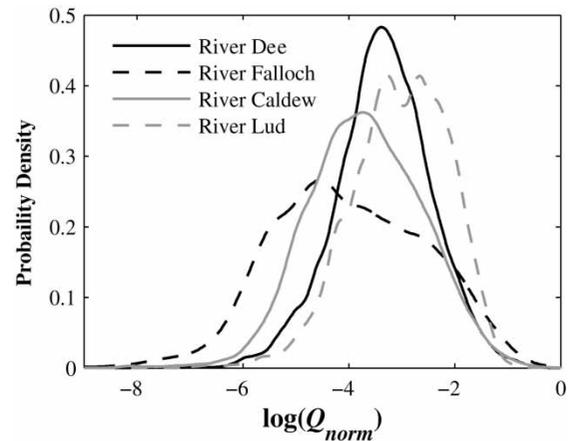

Figure 2 │ Comparison of the normalised and log-transformed flow records for the four study catchments.

which is confirmed by comparing the 1995 hydrographs in Figure 1.

$$Q_{\mathrm{norm},i} = \frac{Q_i - \min(Q)}{\max(Q) - \min(Q)} \tag{1}$$

MODELLING TECHNIQUES

ARMA modelling

An ARMA model is a method of analysing a stationary time series of data using an autoregressive and moving-average polynomial (Box *et al.* 2008). The model is generally referred to as an ARMA(p,q) where p and q are the orders of the autoregressive and moving-average polynomials, respectively. The general form of an ARMA(p,q) model is described by Equation (2).

$$X_t = c + \varepsilon_t + \sum_{t=1}^{p} \varphi_i X_{t-i} + \sum_{t=1}^{q} \theta_i \varepsilon_{t-i} \tag{2}$$

where X_t is a time series of estimated data; c is a constant; ε_t is a time series of white noise generated from a normal distribution (N(0,1)); and φ and θ are estimated parameters of the model.

Equation (2) allows for the generation of a synthetic time series from a fitted ARMA(p,q) model.
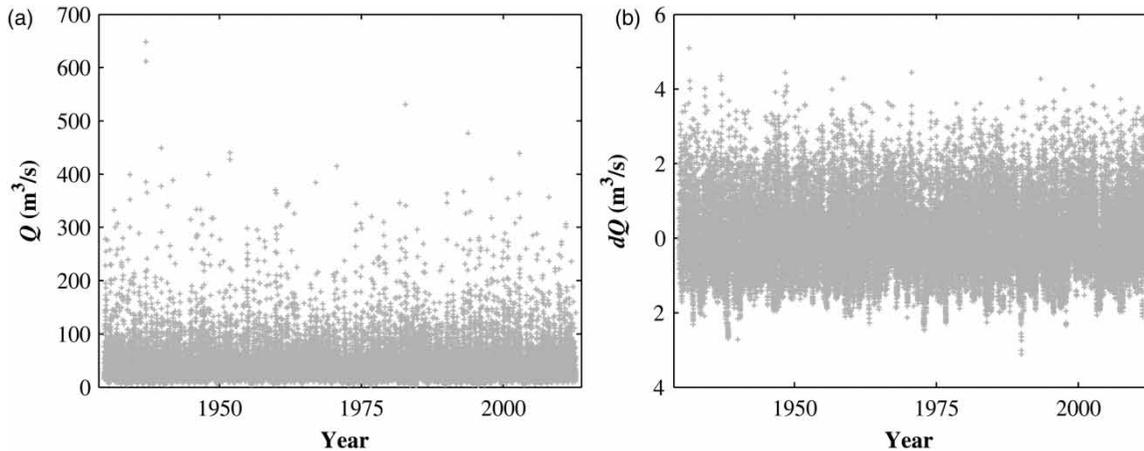
Figure 3 | (a) Recorded daily mean flow at the Woodend gauge on the River Dee and (b) the corresponding de-seasonalised flows.

For the successful use of ARMA models, the time series they are fitted to must be representative of a stationary process (i.e., constant mean and variance). Owing to the nature of the climate in the UK, river flows are not stationary and they have an annual variation based on the seasons (i.e., wetter winter and drier summer), as indicated in Figure 1(a)–1(d). To use an ARMA model to describe the behaviour of daily streamflow, the original data must be deseasonalised to remove these monthly fluctuations in flow magnitude and ensure a stationary time series for model fitting. The deseasonalisation of a flow time series is achieved using the relationship described by Equation (3). Figure 3 provides an example of how the seasonality of the recorded data is removed by the application of Equation (3) to the data from the River Dee.

$$Q_{\mathrm{deseason},i} = \frac{Q_i - \mathrm{mean}(Q_i)}{\mathrm{stdev}(Q_i)} \qquad (3)$$

where Q_{deseason} is the de-seasonalised flow; Q is the log-transformed recorded flow; and i is the month of the year.

Upon de-seasonalisation of the recorded flows, the most suitable ARMA model is fitted to Q_{deseason} and the model parameters (φ and θ) estimated. The fitting and estimation procedure was conducted using the R statistical software package (www.r-project.org). The most suitable model was chosen following the method of Can *et al.* (2012), where a range of stationary ARMA(p,q) models are fitted to the data and the one that minimises the Akaike information criterion (AIC) identified as the most suitable. The AIC

(Akaike 1974) is a widely used (Jones 1980; Findley 1985; Jones & Vecchia 1993; Ji & Chee 2011; Can *et al.* 2012) tool for model selection that identifies the most suitable parameter set by balancing the goodness-of-fit of the model with its complexity, to avoid over-parameterisation. The individual parameters of the model were then estimated using a maximum likelihood approach. Once the parameters of the model have been estimated, synthetic flow sequences can be generated using Equation (2). These flow sequences are then re-seasonalised to give physical values using the inverse of Equation (3). The order of the best-fit ARMA(p,q) models (Table 2) were determined to be an ARMA(4,1), ARMA(2,2), ARMA(3,2) and ARMA(2,2) for the Dee, Falloch, Caldew and Lud, respectively. Table 3 provides the parameter estimates for each model using a maximum likelihood estimation procedure.

Hidden Markov modelling

HMM was first introduced by Baum & Petrie (1966) and, since then, has been applied successfully to a range of complex problems where the standard MM technique appears limited. For effective modelling of such processes, in addition to the time-evolution of observed states (as in a simple MM), a HMM allows the evolution of the process through underlying un-observed (hidden) states. Thus, HMM are theoretically more advanced.

The model is essentially a dynamic Bayesian network that allows for the modelling of time series using probability

Table 2 | Fitted ARMA models and their corresponding AIC values

River Dee		River Falloch		River Caldew		River Lud	
Model	AIC	Model	AIC	Model	AIC	Model	AIC
ARMA(3,1)	43,617	ARMA(1,1)	30,201	ARMA(2,3)	15,659	ARMA(1,1)	6,339
ARMA(4,0)	43,656	ARMA(1,2)	30,075	ARMA(3,3)	15,675	ARMA(1,2)	4,898
ARMA(4,1)	**43,566**	**ARMA(2,2)**	**30,070**	**ARMA(3,2)**	**15,658**	**ARMA(2,2)**	**4,783**
ARMA(4,2)	43,619	ARMA(2,3)	30,072	ARMA(4,4)	15,659	ARMA(2,3)	4,786
ARMA(4,3)	43,574	ARMA(3,2)	30,072	ARMA(3,5)	15,660	ARMA(3,2)	4,786

The bold values indicate the best-fit models for each data set.

Table 3 | Parameter estimates from ARMA model fittings

River	φ_1	φ_2	φ_3	φ_4	θ_1	θ_2
Dee	1.787 ± 0.025	-0.979 ± 0.026	0.256 ± 0.014	1.787 ± 0.025	-0.078 ± 0.08	NA
Falloch	1.067 ± 0.078	-0.180 ± 0.064	NA	NA	-0.320 ± 0.080	-0.107 ± 0.012
Caldew	1.146 ± 0.080	-0.098 ± 0.135	-0.090 ± 0.070	NA	-0.318 ± 0.081	-0.229 ± 0.081
Lud	1.260 ± 0.023	-0.263 ± 0.023	NA	NA	-0.535 ± 0.022	-0.212 ± 0.013

distributions (Ghahramani 2001). Statistically, the HMM is distinct from ARMA as it is based on accounting for evolution probabilities of the observed state at time *t* from one previous state (first order Markov model) at *t-1*, through a set of un-observed states; in comparison, ARMA models use only auto-correlation of observations at time *t*, with a few subsequent observations in the past. Thus, the general make up of HMMs is described by Jenkins *et al.* (2014) using the following five elements: (a) a set of distinct observed states; (b) a set of unobserved states within the observed states; (c) a matrix to describe the probability of transition between the states – *state transitional probability matrix*; (d) a matrix of the probabilities of unobserved states – *emission probability matrix*; and (e) a set of initial probabilities for the observed states. To conduct modelling of a time series using a HMM, these five elements have to be obtained from the recorded data.

The implementation of a HMM to simulate daily streamflow is summarised below:

1. Identify discrete states (*S*) in the flow record.
2. Define the set of *N* unobserved states to account for all possible values between the discrete state limits. For example, if state A were between 5 and 15 m^3/s the

unobserved states correspond to 5, 6, 7, 8, ... 15 m^3/s and so on. Each discrete state has *N* unobserved states.
3. Define the state transition probability matrix. For *S* states this is an $S \times S$ matrix with the value in row *i* and column *j* corresponding to the probability of flow transition from state S_i to state S_j.
4. Define the emission probability matrix that contains the emission probabilities of all unobserved states that correspond to each observed state, i.e., an $S \times N$ matrix.
5. Define a set of *S* initial probabilities for each discrete state following the method of Rabiner (1989).

As one of the objectives of this paper is to assess the general applicability of a HMM approach to UK catchments, it is important that the partitioning of the recorded flow time series into distinct states is consistent across all catchments; this process is well-known to suffer from subjectivity (Augustin *et al.* 2008) and hence standard hydrological practices of using percentiles of flow is employed (Table 4). Each flow record is divided into 11 states, using increments of 10%, which ensures that each is sampled approximately the same number of times throughout the simulation and provides a transferrable methodology between test catchments. This is important, so as to ensure that the underlying flow values

Table 4 | Discrete flow states used in the HMM for each site

State	Percentile range (%)	Dee *Q* range (m³/s)	Falloch *Q* range (m³/s)	Caldew *Q* range (m³/s)	Lud *Q* range (m³/s)
A	0–10	<11.0	<0.4	<1.0	<0.15
B	10–20	11.1–15.0	0.5–0.7	1.1–1.3	0.16–0.19
C	20–30	15.1–19.0	0.8–1.0	1.4–1.7	0.20–0.24
D	30–40	19.1–22.0	1.1–1.5	1.8–2.1	0.25–0.29
E	40–50	22.1–27.0	1.6–2.2	2.2–2.7	0.30–0.35
F	50–60	27.1–32.0	2.3–3.3	2.8–3.5	0.36–0.44
G	60–70	32.1–40.0	3.4–5.3	3.6–4.6	0.45–0.54
H	70–80	40.1–51.0	5.4–9.0	4.7–6.5	0.55–0.68
I	80–90	51.1–74.0	9.1–16.2	6.6–10.1	0.69–0.89
J	90–99	74.1–182.0	16.3–45.0	10.2–27.4	0.90–1.52
K	99–100	>182.0	>45.0	>27.4	>1.52

are well represented, allowing for an accurate emission probability matrix to be formulated. Owing to the difference in magnitude of the flows in the rivers, all recorded flows were rounded to zero decimal places for the Dee, one decimal place for the Falloch and Caldew and two decimal places for the Lud.

To give the best representation of the flow regime, it is preferable that the values within each state have a continuous distribution. The location of the first 10 states (A–J) ensured this for all catchments but this was not possible for the upper state (K) due to the sparseness of data at the extreme limit. To ensure that the extreme limits of the synthetic sequences are not constrained by the measured data, a continuous distribution is required to describe the flows within this state. Although any number of probability distributions could be sampled from, due to the extreme nature of the data within this state, it makes sense that an extreme value distribution is used. To overcome this limitation, Augustin *et al.* (2008) used the GEV distribution to estimate flows in their upper state. As the GEV model combines the traditional extreme value families (Weibull, Fréchet and Gumbel) it is considered the most suitable for inclusion (Coles 2001). Thus, a similar approach has been adopted here (HMM-GEV).

To implement this HMM-GEV framework, the GEV is fitted to recorded flow data above the 99th percentile using a maximum likelihood approach following the procedure of Coles (2001). The cumulative distribution function of the GEV is provided in Equation (4) with details of the parameter

estimates for each catchment given in Table 5. Figure 4 provides the quantile plots in order to assess the goodness-of-fit of the model, for the estimated parameters.

$$\Pr\{X \leq x\} = \begin{cases} \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} & \text{for } \xi \neq 0 \\ \exp\left\{\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} & \text{for } \xi = 0 \end{cases} \quad (4)$$

where μ is the location parameter; σ is the scale parameter; and ξ is the shape parameter.

Similar to the GEV is the GP distribution, which is a threshold-based model that has been readily applied for modelling environmental extremes (Coles & Tawn 1991; Cañellas *et al.* 2007; Callaghan *et al.* 2008; Pender & Karunarathna 2013; Karunarathna *et al.* 2014). Although this GP model only fits the distribution to the extreme tail of the data, the HMM-GP framework was implemented using the same >99th percentile fitting approach as the HMM-GEV. The selection of the thresholds was made from exploratory

Table 5 | Estimated GEV parameters for extreme (>99th percentile) flows on each river

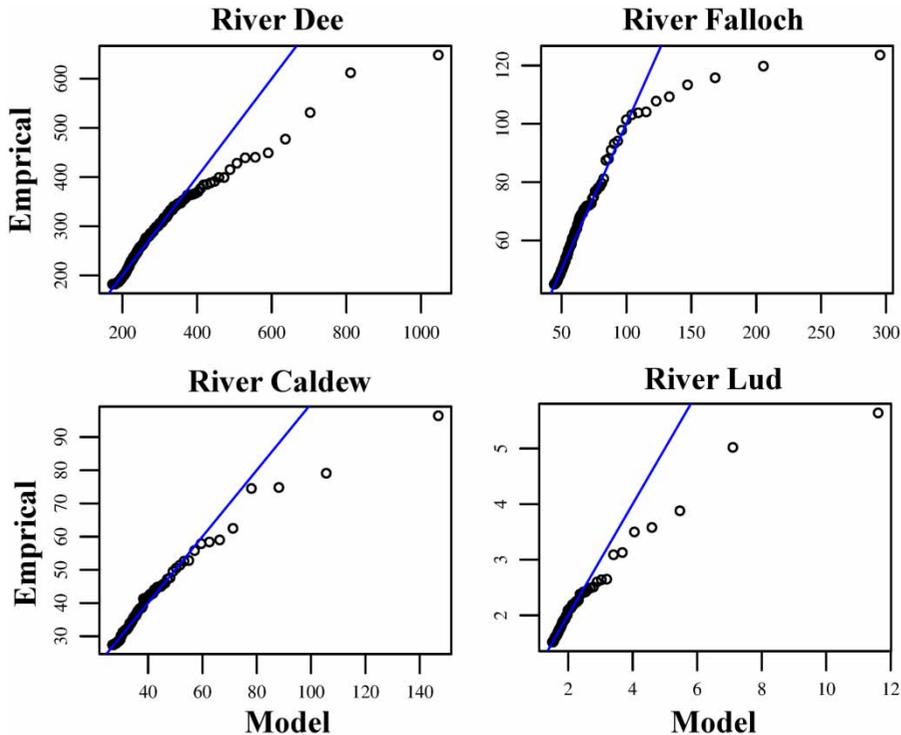| River	μ	σ	ξ
Dee	214.30	33.23	0.43
Falloch	51.33	6.79	0.62
Caldew	31.49	4.49	0.59
Lud	1.61	0.11	0.84

Figure 4 | Quantile plots demonstrating the goodness-of-fit of the GEV model to flows >99th percentile for all four rivers.

analysis of the data following the procedure recommended by Coles (2001). The cumulative distribution function of the GP is provided in Equation (5) with details of the parameter estimates for each catchment given in Table 6. Figure 5 provides the quantile plots in order to assess the goodness-of-fit of the model, for the estimated parameters.

$$\Pr\{X \le x\} = \begin{cases} 1 - \zeta_u \left[1 + \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} & \text{for } \xi \ne 0 \\ 1 - \zeta_u \left[\exp\left(-\frac{x - \mu}{\sigma} \right) \right] & \text{for } \xi = 0 \end{cases} \quad (5)$$

where u is the threshold level; σ is the scale parameter; and ξ is the shape parameter; $\zeta_u = \Pr\{X > u\}$.

Table 6 | Estimated GP parameters for extreme (>99th percentile) flows on each river

River	u	σ	ξ
Dee	250	62.52	0.04
Falloch	50	15.95	0.06
Caldew	35	10.88	0.07
Lud	2	0.50	0.26

Comparison of the quantile plots for both fitted models (Figures 4 and 5) shows that the use of the GP provides a better representation of the extreme tail of the flow data, for all catchments. This indicates that, for these data records, there are too many data above the 99th percentile to ensure a good fit with the GEV. As such, it is expected that, although improving on the ARMA model, the HMM-GEV will provide less realistic estimations of the upper extremes, compared to the HMM-GP.

As with the ARMA modelling, both HMM approaches were implemented using the *R* statistical software package to allow for efficient generation of synthetic flow sequences for each site.

RESULTS AND DISCUSSION

Analysis methods

To test the suitability of the modelling techniques at reproducing realistic flow sequences, the statistics of synthetic sequences were compared to those of the recorded values.
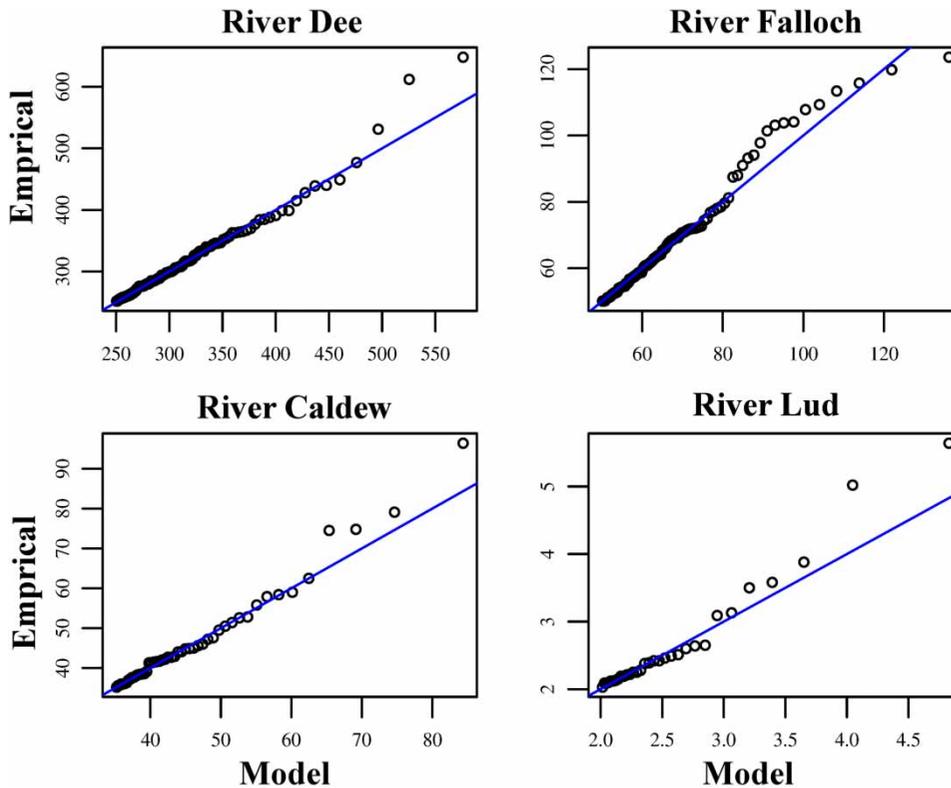
Figure 5 │ Quantile plots demonstrating the goodness-of-fit of the GP model to flows >99th percentile for all four rivers.

For each river, 100 sequences, the same length as the recorded flow data, were generated. This permits statistical analysis of correlations between real and simulated data in order to assess confidence in realistic sequences being output. Also, it allows for an upper and lower bound to be placed on the results, thus providing a range of possible values for comparison to the recorded data. The assessment comprises a comparison of the probability densities, between the recorded and synthetic sequences, and a more specific comparison at a range of percentiles. The probability density comparison uses log-transformed flow as it allows for better visual assessment. Detailed evaluation at individual percentiles consists of the determination of percentage differences between the mean value from the 100 synthetic sequences and the corresponding recorded value (Equation (6)):

$$\%\text{Difference} = \frac{\text{mean}(Q_{p,\text{synthetic}}) - Q_{p,\text{recorded}}}{Q_{p,\text{recorded}}} \quad (6)$$

where $Q_{p,\text{synthetic}}$ and $Q_{p,\text{recorded}}$ are the synthetic and recorded flow values at percentile p, respectively.

To provide a comparison between the ARMA, HMM-GEV and HMM-GP models, and each river, the overall relative mean absolute difference (RMAD) across the percentile range was determined as a percentage (Equation (7)):

$$\text{RMAD} = \sum_{p=1}^{n} \left| \frac{Q_{p,\text{synthetic}} - Q_{p,\text{recorded}}}{Q_{p,\text{recorded}}} \right| \times \frac{100\%}{n} \quad (7)$$

where $Q_{p,\text{synthetic}}$ and $Q_{p,\text{recorded}}$ are the synthetic and recorded flow values at percentile p, respectively.

ARMA model

The probability density and percentage difference plots for the ARMA model are provided in Figure 6(a)–6(d) and 6(m)–6(p), respectively. The RMAD were determined to be 5.6%, 6.0%, 38.5% and 5.1% for the Dee, Falloch, Caldew and Lud, respectively.

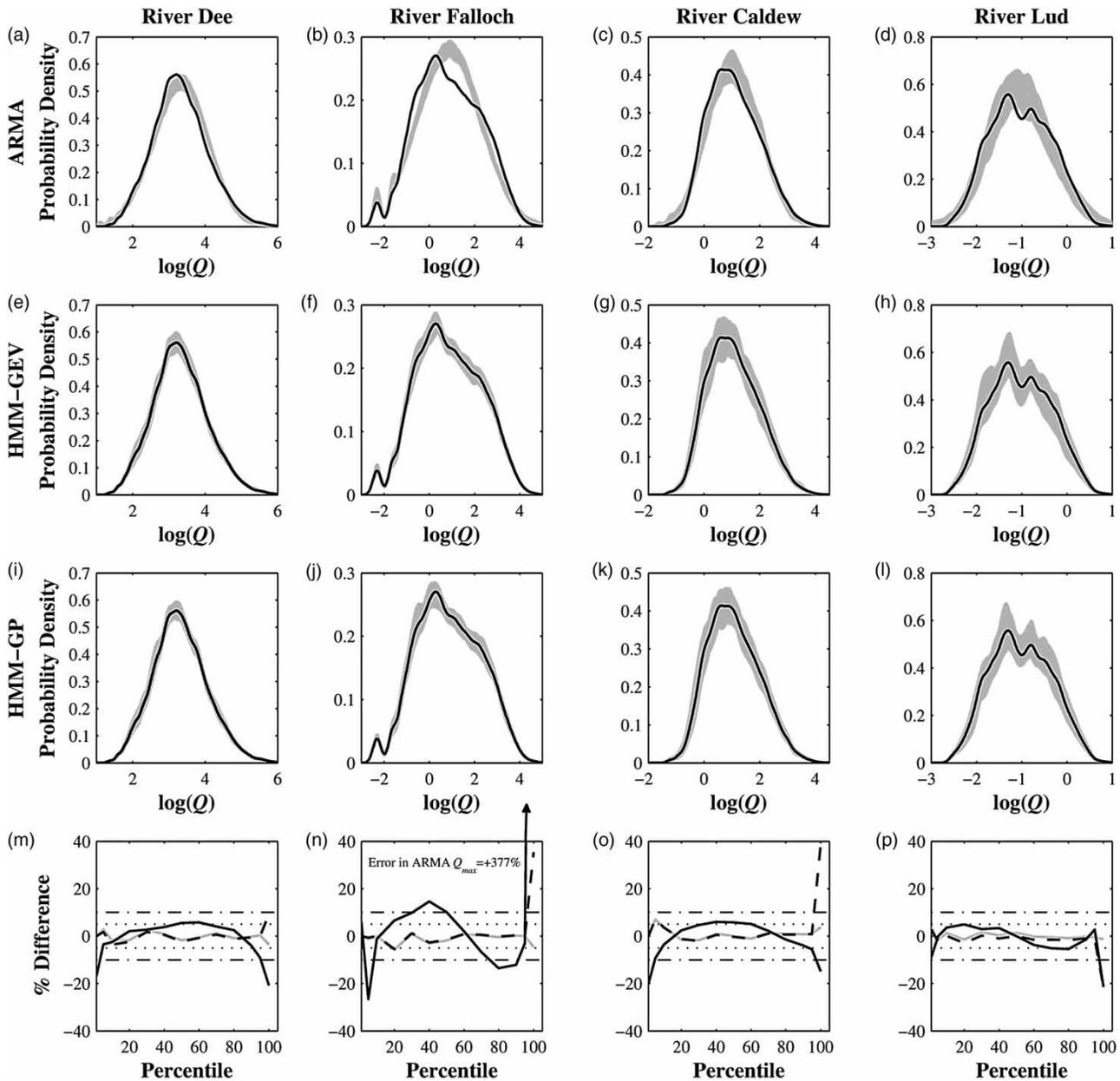The results from the ARMA modelling show that generating synthetic flow sequences is more successful on the

Figure 6 | Results from the modelling of synthetic flow sequences for the River Dee (first column), River Falloch (second column), River Caldew (third column) and River Lud (fourth column). The first three rows provide a comparison of the log-transformed flows, for the recorded values (solid line), and the range of synthetic sequences (grey shaded region) for the ARMA, HMM-GEV and HMM-GP, respectively. The bottom row shows the percentage difference between the mean of the synthetic sequences (ARMA – black, HMM-GEV – dashed black, HMM-GP – solid grey) and the recorded values for a range of percentiles. The dash-dot and dotted lines indicate the ±10% and ±5% error ranges, respectively.

Dee, Caldew and Lud (RMAD 5.6%, 6.0% and 5.1%) compared to the Falloch (RMAD 38.5%). This reflects differences in the geographic and catchment characteristics, indicating poor model performance for the highly variable flow regime of the runoff-dominated (Figure 1(b)) mountainous catchment of the Falloch. In explanation, the ARMA

model is based on a random white noise component; this fails to account for the probability of change in the flow regime and therefore induces poorer model performance where flows exhibit a rapid change in state.

Thus, only for the more attenuated hydrological regimes of the Dee, Caldew and Lud can the ARMA model capture

and replicate the flow regime within the ±10% range of acceptability, between approximately the 5th and 95th percentiles (Figure 6(m)–6(p)). At <5th and >95th percentiles the ARMA model estimates fail to adequately simulate flow, indicating weakness in the modelling approach at extreme flows. On average, sequences for the Dee and Lud underestimate the most extreme values, with the percentage difference at the peak flow (100th percentile) being −11.1% and −25.1%, respectively. Although the ARMA model has been fitted to the data to give the most appropriate parameter set-up, it is important to note that this fitting only provides the best representation of the average changes in the flow regime, not extremes. This explains its inability to effectively capture the extreme flows.

HMM extreme value models

The probability density and percentage difference plots for the HMM-GEV model are provided in Figure 6(e)–6(h) and 6(m)–6(p), respectively. The modelling resulted in a RMAD of 1.9%, 3.8%, 4.4% and 2.9% for the Dee, Falloch, Caldew and Lud, respectively. Equivalent plots for the HMM-GP model are provided in Figure 6(i)–6(l) and 6(m)–6(p), with the RMAD determined to be 1.4%, 1.5%, 1.9% and 1.1% for the Dee, Falloch, Caldew and Lud, respectively.

The HMM-GEV model is shown to produce more realistic results than that of the ARMA model. For the HMM-GEV, all data between the 0 and 95th percentiles lie within the ±10% acceptability limit, with the majority falling within ±5% (Figure 6(m)–6(p)) and the overall mean RMAD are <5% for all catchments. These results show significant improvement compared to the simple ARMA model, particularly in simulating the flashier hydrograph of the Falloch, where the RMAD is now only 3.8%. However, examining the fit to the upper extremes (Figure 6(m)–6(p)) indicates the inadequacy of the GEV method for >99th percentile. While these extreme events on the highly attenuated flow regime of the River Dee do fall within the acceptability range (±10%), the flashier regimes of the Caldew and Falloch are very poorly represented. Reviewing the GEV methodology, this is due to there being too many data above the 99th percentile to result in

an accurate fit for the tail of the distribution, as discussed previously and demonstrated by Figures 4 and 5.

The use of the alternative HMM-GP approach is shown to produce a more realistic representation of the entire flow regime, including the upper extremes. From Figure 6(m)–6(p) it is clear that all simulated flows fall within the acceptability limits, with the GP method yielding the lowest overall RMAD of <2% for all catchments. All flow extremes are fully captured by this approach, with all high flows simulated well within the ±10% acceptability range, such that the HMM-GP approach appears generically applicable for simulating a wide range of hydrographic regimes across the geographic regions of the UK. The main reason behind the improvement, compared to the HMM-GEV model, is the improved representation of the extreme tail of the flow data, as evident from Figures 4 and 5.

CONCLUSIONS

This paper has compared and contrasted a simple ARMA approach to more complex HMM methods for stochastic generation of daily streamflow sequences. In developing an appropriate HMM methodology, the use of GEV and GP extreme value distributions has been analysed and a generic methodology for the partitioning of the flow record has been developed in line with standard flow percentile descriptors. All models were tested across four UK catchments of geographic diversity and varying hydrologic regime.

The ARMA model has the advantages of: being much simpler to implement; being unconstrained by estimates to a number of decimal places; and having no requirement for an emission probability matrix. These results showed that the ARMA model is reasonably capable of producing acceptable (within ±10% of recorded values) realisations within the 5th to 95th percentile range, but unable to effectively capture extreme values. This limits the use of an ARMA model to catchments with flow regimes with highly attenuated hydrographs and, therefore, it has been determined that ARMA is unsuitable for accurate representation of the entire flow regime for the range of UK catchments modelled.

As the alternative HMM method makes use of the actual probabilities of transitioning between flow states and probabilities of occurrence within these states, a HMM has been shown to be more suitable for modelling a range of flow time series with different characteristics. As the traditional HMM is limited for streamflow data sets due to the long data tail and sparseness of recorded values in this part of the distribution, it has been proven that accurate simulation of extreme flows requires the HMM to be combined with a GP extreme value distribution, to sample values above the 99th percentile. The HMM-GP model output accurately (RMAD < 2%) simulated the full range of flows for all hydrological regimes of the UK case studies tested, showing robustness of methodology and confidence in generic applicability. It is therefore advocated that, provided that a sufficient historic data set exists, use of a HMM-GP approach, combined with a consistent percentile-based method for partitioning of the flow record, makes for an accessible and straightforward modelling tool that can be readily applied in practice. The use of the outputs of such an approach, in conjunction with more detailed hydraulic models, will significantly benefit a range of detailed sustainability assessments paramount to the future security of water resource, flood risk and ecosystem management in UK river systems.

ACKNOWLEDGEMENTS

REFERENCES

Akaike, H. 1974 A new look at the statistical mode identification. *IEEE Trans. Automatic Control* **19** (6), 716–723.

Augustin, N. H., Beevers, L. & Sloan, W. T. 2008 Predicting river flows for future climates using an autoregressive multinomial logit model. *Water Resour. Res.* **44** (7), W007403.

Baum, L. E. & Petrie, T. 1966 Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.* **37** (6), 1554–1563.

Box, G. E. P., Jenkins, G. & Reinsek, C. 2008 *Time Series Analysis, Forecasting and Control.* 4th edn. John Wiley & Sons, New York.

Callaghan, D., Nielsen, P., Short, A. & Ranasinghe, R. 2008 Statistical simulation of wave climate and extreme beach erosion. *Coastal Eng.* **55** (5), 375–390.

Can, İ., Tosunoğlu, F. & Kahya, E. 2012 Daily streamflow modelling using autoregressive moving average and artificial neural networks models: case study of Çoruh basin, Turkey. *Water Environ. J.* **26** (4), 567–576.

Cañellas, B., Orfila, A., Méndez, F. J., Menéndez, M. & Tintoré, J. 2007 Application of a POT model to estimate the extreme significant wave height levels around the Balearic Sea (Western Mediterranean). *J. Coast. Res.* **50**, 329–333.

Coles, S. 2001 *An Introduction to Statistical Modelling of Extreme Values*, 1st edn. Springer, London.

Coles, S. G. & Tawn, J. A. 1991 Modelling extreme multivariate events. *J. Roy. Stat. Soc. Series B (Methodological)* **53** (2), 377–392.

Cranston, E. & Darby, E. 2004 *Ranunculus in Chalk Rivers: Phase 2.* Environment Agency Science Report W1-042/TR.

Dechmi, F., Burguete, J. & Skhiri, A. 2012 SWAT Application in intensive irrigation systems: model modification, calibration and validation. *J. Hydrol.* **470–471**, 227–238.

European Commission 2000 Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *CELEX-EUR Official Journal* **L327**, 1–72.

Findley, D. F. 1985 On the unbiasedness property of AIC for exact or approximating linear stochastic time series models. *J. Time Series Anal.* **6** (4), 229–252.

Ghahramani, Z. 2001 An introduction to hidden Markov models and Bayesian networks. *Int. J. Pattern Recog. Artif. Intell.* **15** (01), 9–42.

Horritt, M. S., Bates, P. D., Fewtrell, T. J., Mason, D. C. & Wilson, M. D. 2010 Modelling the hydraulics of the Carlisle 2005 flood event. *Proc. ICE Water Manage.* **163** (6), 273–281.

Jacobs 2007 *Caldew and Carlisle City Flood Alleviation Scheme – Geomorphology and Sediment Modelling.* Jacobs, Glasgow, UK.

Jenkins, D. P., Patidar, S. & Simpson, S. A. 2014 Synthesising electrical demand profiles for UK dwellings. *Energy Buildings* **76**, 605–614.

Ji, W. & Chee, K. C. 2011 Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. *Solar Energy* **85** (5), 808–817.

Jones, R. H. & Vecchia, A. V. 1993 Fitting continuous ARMA models to unequally spaced spatial data. *J. Am. Stat. Assoc.* **88** (423), 947–954.

Jones, R. H. 1980 Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics* **22** (3), 389–395.

Karunarathna, H., Pender, D., Ranasinghe, R., Short, A. D. & Reeve, D. E. 2014 The effects of storm clustering on beach profile variability. *Marine Geol.* **348**, 103–112.

Kingston, D. G., Lawler, D. M. & McGregor, G. R. 2006 Linkages between atmospheric circulation, climate and streamflow in the northern North Atlantic: research prospects. *Prog. Phys. Geogr.* **30** (2), 143–174.

Loch Lomond, Trossachs National Park 2012 2 Megawatt Hydro Scheme opens in National Park [Online]. http://www.lochlomond-trossachs.org/looking-after/2-megawatt-hydro-scheme-opens-in-national-park-19/07/2012/menu-id-522.html (accessed 24 April 2014).

Mainstone, C. P. 1999 *Chalk Rivers: Nature Conservation and Management.* Produced on behalf of English Nature and the Environment Agency. English Nature, Peterborough.

Martins, O. Y. 2011 ARMA Modelling of Benue River flow dynamics: comparative study of PAR model. *Open J. Modern Hydrol.* **1**, 1–9.

Mason, D. C., Bates, P. D. & Dall'Amico, J. T. 2009 Calibration of uncertain flood inundation models using remotely sensed water levels. *J. Hydrol.* **368** (1–4), 224–236.

Mellor, D., Sheffield, J., O'Connell, P. E. & Metcalfe, A. V. 2000 A stochastic space-time rainfall forecasting system for real time flow forecasting I: development of MTB conditional rainfall scenario generator. *Hydrol. Earth Syst. Sci.* **4** (4), 603–615.

Mohammadi, K., Eslami, H. R. & Kahawita, R. 2006 Parameter estimation of an ARMA model for river flow forecasting using goal programming. *J. Hydrol.* **331** (1–2), 293–299.

Musa, J. J. 2013 Stochastic modelling of Shiroro River stream flow process. *Amer. J. Eng. Res.* **2** (6), 49–54.

Neal, J. C., Bates, P. D., Fewtrell, T. J., Hunter, N. M., Wilson, M. D. & Horritt, M. S. 2009 Distributed whole city water level measurements from the Carlisle 2005 urban flood event and comparison with hydraulic model simulations. *J. Hydrol.* **368** (1–4), 42–55.

Neal, J., Keef, C., Bates, P., Beven, K. & Leedal, D. 2013 Probabilistic flood risk mapping including spatial dependence. *Hydrol. Process.* **27** (9), 1349–1363.

O'Connell, P. E. 1974 Stochastic modelling of long-term persistence in stream flow sequences. PhD Thesis, Imperial College of Science and Technology, London.

O'Connell, P. E. & O'Donnell, G. 2014 Towards modelling flood protection investment as a coupled human and natural system. *Hydrol. Earth Syst. Sci.* **18** (1), 155–171.

Pender, D. & Karunarathna, H. 2013 A statistical-process based approach for modelling beach profile variability. *Coastal Eng.* **81**, 19–29.

Rabiner, L. R. 1989 A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77** (2), 257–286.

Sadek, N. 2006 River Nile Flood Forecasting and its Effect on National Projects Implementation. In: *10th International Water Technology Conference*, 23–25 March, Alexandria, Egypt.

Szilagyi, J., Balint, G. & Csik, A. 2006 Hybrid, Markov chain-based model for daily streamflow generation at multiple catchment sites. *J. Hydrol. Eng.* **11** (3), 245–256.

The Scottish Government 2009 River Falloch Hydro – Scoping Option. http://www.scotland.gov.uk/Topics/Business-Industry/Energy/Infrastructure/Energy-Consents/Applications-Database/Hydro/River-Falloch-Scoping (accessed 24 April 2014).

Wilby, R. L., Cranston, L. E. & Darby, E. J. 1998 Factors governing macrophyte status in Hampshire chalk streams: implications for catchment management. *Water Environ. J.* **12** (3), 179–187.

Xu, Z., Schumann, A., Brass, C., Li, J. & Ito, K. 2001 Chain-dependent Markov correlation pulse model for daily streamflow generation. *Adv. Water Resour.* **24**, 551–564.

Xu, Z., Schumann, A. & Li, J. 2003 Markov cross-correlation pulse model for daily streamflow generation at multiple sites. *Adv. Water Resour.* **26** (3), 325–335.

Yapo, P., Sorooshian, S. & Gupta, V. 1993 A Markov chain flow model for flood forecasting. *Water Resour. Res.* **29** (7), 2427–2436.

Zhang, R., Santos, C. A. G., Moreira, M., Freire, P. K. M. M. & Corte-Real, J. 2013 Automatic calibration of the SHETRAN hydrological modelling system using MSCE. *Water Resour. Managet.* **27** (11), 4053–4068.