Reconstructing Spatiotemporal Dynamics in Hydrological State Along Intermittent Rivers

Michael Eastman [1,*], **Simon Parry** [1], **Catherine Sefton** [1], **Juhyun Park** [2,3] and **Judy England** [4]

1   UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK; spar@ceh.ac.uk (S.P.); catsef@ceh.ac.uk (C.S.)
2   ENSIIE & LaMME, University of Paris-Saclay, 91025 Evry, France; juhyun.park@ensiie.fr
3   Department of Mathematics and Statistics, Lancaster University, Lancaster, Lancashire LA1 4YF, UK
4   Environment Agency, Wallingford, Oxfordshire OX10 8BD, UK; judy.england@environment-agency.gov.uk
*   Correspondence: miceas@ceh.ac.uk

Abstract: Despite the impact of flow cessation on aquatic ecology, the hydrology of intermittent rivers has been largely overlooked. This has resulted in a lack of monitoring projects, and consequently, datasets spanning a period of sufficient duration to characterise both hydrological extremes. This report documents an investigation into the potential for statistical modelling to simulate the spatiotemporal dynamics of flowing, ponded and dry hydrological states in an internationally rare hydrological state dataset. The models presented predict unrecorded hydrological state data with performance metrics exceeding 95%, providing insights into the relationship between ponding prevalence and the performance of statistical simulation of this ecologically important intermediate state between drying and flowing conditions. This work demonstrates the potential for hydrological intermittence to be simulated in areas where hydrological state data are often sparse, providing opportunities for quality control and data infilling. This further understanding of the processes driving intermittence will inform future water resource assessments and the influence of climate change on hydrological intermittence.

1. Introduction

Temporary flow cessation is common, with more than 50% of the global river network estimated to be intermittent [1]. Their dynamic hydrological behaviour, their prevalence across a range of physical and chemical conditions, and spatiotemporal variability in their habitat structure, mean that intermittent rivers and ephemeral streams (IRES) support diverse biological communities [2–7]. Despite this ecological importance and the ecosystem services they provide [8], they have been largely overlooked in governance and policy, resulting in a lack of legislative protection [9–11].

Processes driving hydrological regime in IRES include those relating to climate, geology and land cover [12]. Identification of drivers—and quantification of the relative sensitivity of the hydrological regime to each—is therefore an essential step towards the assessment of a river's response to hydrological extremes and to abstraction, land-use change and climate change pressures [1,13]. Such evidence would further our understanding of the ecological sensitivity of IRES and underpin the development of suitable management strategies for their protection.

However, the extent of the monitored IRES network is constrained both spatially and temporally, limiting the understanding that can be acquired from investigations into these datasets alone. Simulation of the hydrological regime of these rivers enables insights into the conditions beyond the monitored record, as well as inferences about the factors driving intermittence. Physically based modelling of intermittent headwaters

would require detailed understanding of the processes driving intermittence, but current knowledge of temporary headwater streams is extremely limited [14]. Statistical modelling, however, offers the potential to investigate these drivers and assess the impact of disturbances to the hydrological regime. Previous research demonstrated the potential of parametric approaches for wet/dry mapping [15], flow permanence simulation [16] and estimating probability of intermittence [17]. Furthermore, neural networks and random forest approaches have been demonstrated to be well suited for statistical simulation of intermittence [18]. Other recent studies have also taken a random forest approach to simulating intermittence [19–21], and some have compared approaches including neural networks, regression trees and multivariate adaptive regression splines [18,22].

Effective monitoring of IRES requires observations both of flow magnitude and of hydrological state, because the ponding conditions that exist as rivers transition between aquatic and terrestrial states are key to their ecological diversity [7,23]. Ponded reaches of IRES act as a refuge for many lotic species. This includes invertebrates which survive drying phases in ponds until flow resumes, as well as invertebrates that complete life cycle stages in ponds [24]. Ponds also act as refuge for some fish species which inhabit ponds until flow resumes [25,26]. Furthermore, a number of provisioning, regulating and cultural services have been attributed to temporary rivers, many of which apply to ponding phases, such as the deposit of sediment as flow ceases, which in combination with the colonisation of terrestrial plants binding sediment reduces erosion [27]. Modelling of ponding is rare, however, because data scarcity makes the calibration and testing of models challenging [28]. No studies have been found with a statistical modelling approach to the reconstruction of IRES dynamics in a hydrological state that includes ponding as well as flowing and dry conditions.

Data scarcity on the hydrological state of IRES pertains to spatiotemporal extent and resolution as well as to ponding, and existing studies have been limited in space to gauging stations [16], or to spatially distributed sites for a limited period of a single or a few years [18,21]. A UK. dataset across ten rivers in the East Chilterns [29] covers the full range of hydrological conditions (being twenty years in temporal extent), at good spatial resolution (approximately monthly) and across a range of flow regimes (especially given the limited spatial extent). However, there remain gaps because collecting data of this type is highly resource intensive. Modelling facilitates infilling of an extracted monthly time series of flowing, ponded and dry state that is vital for understanding the long-term behaviour of the rivers and quantifying their response to drivers and disturbances. Furthermore, modelling offers the potential to extrapolate spatially and temporally, both desirable given the cost of data collection. Thus, the spatial resolution and temporal extent of the East Chilterns dataset provide a unique opportunity for IRES research, identifying and quantifying the processes driving the hydrological state, as well as their relative sensitivities to environmental change.

In this study, ordinal regression models were trained which simulated the spatiotemporal dynamics of flowing, ponded and dry states using environmental drivers for the first time in the statistical modelling of IRES. The skill of the models in infilling gaps were tested across both hydrological extremes on ten intermittent rivers in the UK. The aim was to demonstrate the potential of statistical models for simulating a hydrological state, as well as improve understanding of the processes driving intermittence through an investigation into performance variation. This was achieved by investigating the following objectives:

- Determine how accurately hydrological intermittence in the East Chilterns rivers can be simulated.
- Identify the factors that drive variation in hydrological state simulation performance.

2. Materials and Methods

2.1. Study Area

The study area is located in the Chiltern Hills of lowland southeast England (30–250 m AOD–meters Above Ordnance Datum) with chalk bedrock partially overlain by superfi-
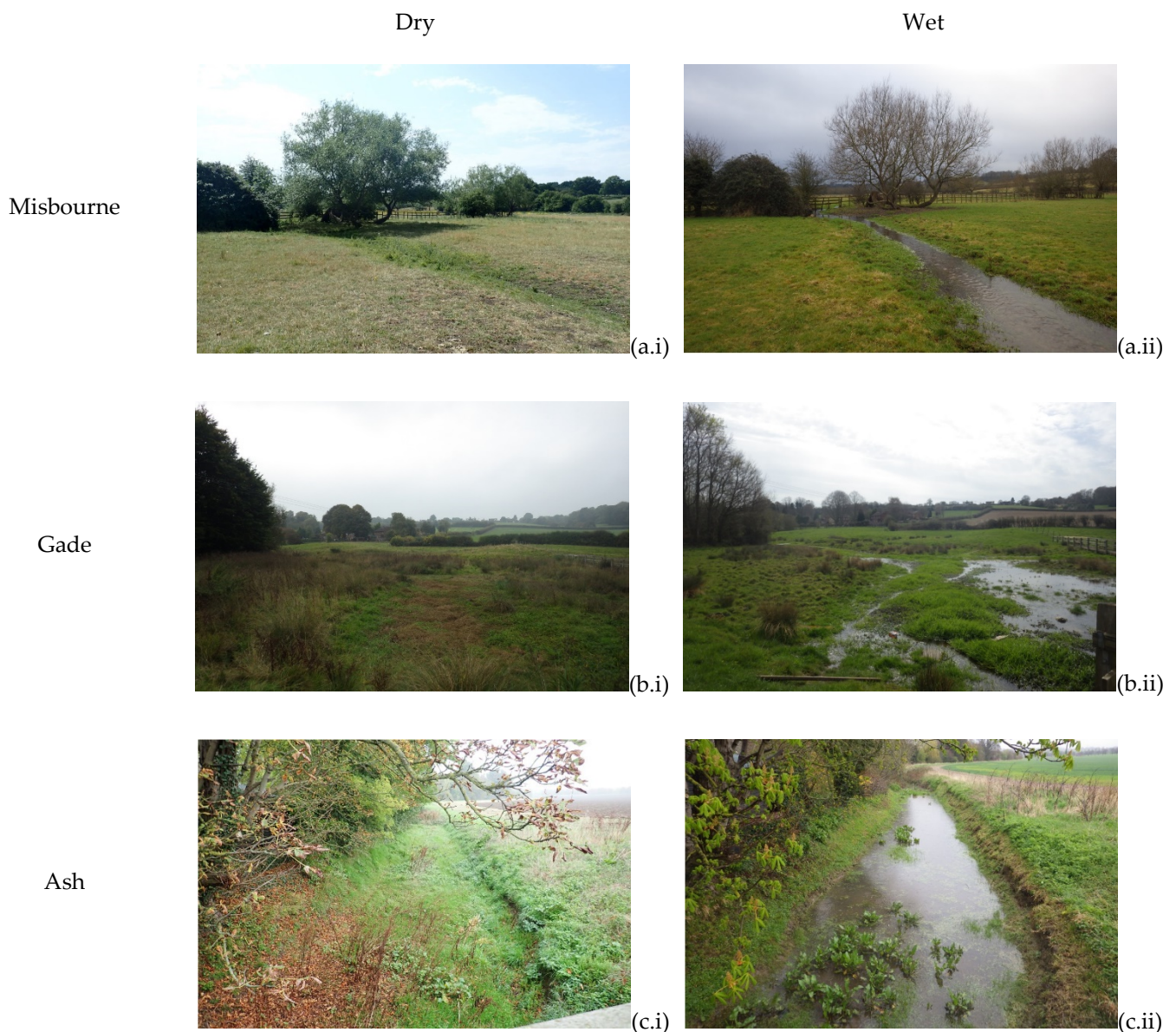
cial deposits and predominantly arable and grassland landcover, with some woodland
(8–21%) and urban cover (3–21%). The climate is temperate oceanic with 600–750 mm
annual average rainfall (1961–1990) and typically no strong seasonality in monthly rainfall
accumulations, but fewer rainy days in the summer months (June–August) than in the
winter (December–February).

The ten study rivers form part of the Thames headwater network and are tributaries
of the Colne (Misbourne, Chess, Bulbourne, Gade and Ver) or the Lee (Mimram, Beane, Rib,
Ash and Stort) with stream lengths of 15 to 46 km (Figures 1 and 2). The Colne catchments
are a little higher and wetter than those of the Lee (30–250 m AOD and 700–750 mm;
30–200 m AOD and 600–650 mm, respectively) and the superficial deposits most extensive
in the east (75% in the Rib, Ash and Stort catchments). The hydrological regimes of the
Colne tributaries and also the Mimram are characteristic of chalk streams in the area,
heavily attenuated (mean Base Flow Index 0.88), and with flow typically increasing in
magnitude from October to March and decreasing from April to September. The Rib, Ash
and Stort have a greater responsive component (mean Base Flow Index 0.53). Daily flow
data for gauging stations on the study rivers and corresponding catchment descriptors are
openly available through the UK National River Flow Archive [30,31].

Figure 1. Location of the hydrological state observation sites, river flow gauging stations and groundwater boreholes in
the Colne (Misbourne, Chess, Bulbourne, Gade and Ver) and Lee (Mimram, Beane, Rib, Ash and Stort) catchments, with
underlying superficial geology types.

Dry | Wet

Misbourne

(a.i) (a.ii)

Gade

(b.i) (b.ii)

Ash

(c.i) (c.ii)

Figure 2. The Gade (**a**), Misbourne (**b**), and Ash (**c**) in dry (**i**) and wet (**ii**) phases. Photo credit: Environment Agency 2016–2019.

2.2. Data

2.2.1. Hydrological State

Hydrologists from England's regulatory authority, the Environment Agency, have monitored the hydrological state of the ten study rivers since 1997 [29]. The spatial extent of the surveys was designed to capture the hydrological state along the full intermittent reach of each river. Survey extent begins in the most upstream non-dry observation and extends to the downstream perennial reaches. This results in 8.1–34.4 km (38–99%) of the rivers upstream of their respective confluences being surveyed. Survey extent is divided into 18–32 sites, according to the hydrological behaviour and access restrictions. These observations are considered as connected reaches, with boundaries equidistant from sites [29]. Reach length varies from <0.1 km to 5.5 km, with a mean of 0.9 km.

Surveys were carried out approximately monthly, varying with hydrological conditions and resource availability. For example, the higher density of observations during and following the 2004–2006 drought, and the lower density in 2001–2004. The monitoring

approach was reviewed and subsequently standardised in 2004. Therefore, the models were trained on hydrological state data from 3 January 2004–30 April 2018.

The hydrological state observed was assigned one of eight categories according to predefined descriptions and photographic examples, which was reduced to three ecologically relevant states: dry, ponded and flowing [29,32]. A monthly hydrological state dataset was extracted to facilitate the quantification of intermittence variation with time. Each month was assigned with the hydrological state observed. In cases of two or more observations within a single month, the observation nearest to the middle of the month was used. If the intervals were equal, the observation that provided a more consistent gap between neighbouring month observations was used. This resulted in an average of 11.3% of months between 3 January 2004 and 30 April 2018 missing an observation, ranging from 4.7% in the Gade to 18.8% in the Beane.

2.2.2. Explanatory Data

Three environmental drivers, river flow, groundwater level and rainfall, were sourced from the Environment Agency and explored as model covariates. In addition, month of simulation, distance upstream of confluence and state from the preceding and following months were also used as covariates (Table 1). Monthly mean river flows were calculated from daily time series measured at a gauging station within the catchment (Table 2). Monthly mean groundwater level was calculated from available observations in each month (Table 2). Lumped catchment [33] average precipitation (areal rainfall) and percolation (modelled effective precipitation) for each month was derived from daily totals for the Colne and Lee catchments, respectively [34]. Frequencies of both precipitation and percolation were the proportions of non-zero days within each month. Where categorical and continuous data were missing, the modal class and mean over the study period (3 January 2004–30 April 2018), respectively, were used to infill observations corresponding to the site.

Table 1. Covariates used to simulate hydrological state.

Factor	Description
Flow	Mean flow in the month of the hydrological state being simulated (m^3s^{-1})
Groundwater	Mean level of water table in the month of the hydrological state being simulated (m AOD)
Precipitation	Mean precipitation in the month of the hydrological state being simulated (mm)
	Proportion of days with >0 mm precipitation in the month of the hydrological state being simulated
	Mean percolation in the month of the hydrological state being simulated (mm)
	Proportion of days with >0 mm percolation in the month of the hydrological state being simulated
Seasonality	Categorical month of hydrological state being simulated
Site	Distance downstream to the confluence (km)
State	Hydrological state (flowing, ponded, dry) in the previous and following months

Table 2. River and associated flow and groundwater data source.

River	Flow Gauging Station		Borehole	
	Name	National Grid Reference	Name	National Grid Reference
Misbourne	Little Missenden	SU9342198458	Woodlands Park	SP8846703308
Chess	Rickmansworth	TQ0657394809	Wayside	SP9477001065
Bulbourne	Two Waters	TL0550405809	Dudswell	SP9659009682
Gade	Bury Mill	TL0532607648	Dagnall OBH	SP9960015540
Ver	Redbourn	TL1091611904	River Hill Dip	TL0798015060
Mimram	Whitwell	TL1841021210	Lilley Bottom Dip	TL1569022760
Beane	Hartham Park	TL3250313143	Watton at Stone	TL2910220042
Rib	Wadesmill	TL3599017444	Lower Farm Buckland	TL3457033890
Ash	Wareside Mardock	TL3936014840	Much Hadham	TL4261218286
Stort	Stansted Springs	TL5003924731	Berden Hall	TL4669429548

2.3. Methods

Cumulative logit models (CLMs) [35], an ordinal regression approach, were selected for simulation of hydrological state at the observation sites due to the inherent ordering of the response variable (flowing, ponded and dry). This was performed using the Ordinal package in R [36].

The probability the ith response, Y_i, $i = 1, \ldots, N$, is smaller than or equal to category, j, $j = 1, \ldots, J$, given the set of ith covariates, x_i:

$$\Pr(Y_i \leq j \mid x_i) \tag{1}$$

Then the cumulative logits are defined as:

$$\text{logit}[\Pr(Y_i \leq j \mid x_i)] = log\frac{\Pr(Y_i \leq j \mid x_i)}{1 - \Pr(Y_i \leq j \mid x_i)}, j = 1, \ldots, J \tag{2}$$

The CLMs assume that:

$$\text{logit}[\Pr(Y_i \leq j \mid x_i)] = \theta_j - x_i^T \beta, \ j = 1, \ldots, J \tag{3}$$

This model satisfies the proportional odds property.

2.3.1. Partial Proportional Odds

The above proportional odds model does not allow the regression coefficients, β, to vary with the response category. Therefore, this model assumes that the effect of the covariates is not dependent on the response category, and thus a unit increase in any covariate would have an equal effect on the probability of an observation being in each of the response categories.

Whether that assumption was violated was assessed for each covariate individually by training partial proportional odds models, allowing the associated regression coefficient, β, to vary by state, j:

$$\text{logit}[\Pr(Y_i \leq j \mid x_i)] = \theta_j - x_i^T \beta_j \tag{4}$$

Differences between the proportional odds and partial proportional odds models were assessed using likelihood ratio tests. By default, proportional odds models were used to simulate hydrological state. However, if non-proportional odds were detected (defined as a significant difference, $p < 0.05$, between proportional odds and partial proportional odds models), partial proportional odds models were used to simulate hydrological state. This meant that coefficients for covariates that exhibited non-proportional odds could vary

with state, accounting for variability in their relationship with hydrological state. Whilst this increased the number of parameters that must be estimated, and therefore uncertainty, model performance was improved by better capturing the inherent variability of these dynamic rivers.

2.3.2. Parameter Estimation

The values of θ and β were estimated using Maximum Likelihood Estimation. The optimal set of covariates was estimated using five-fold cross validation. This means that the data is split evenly into five partitions. Four partitions were then used to train the model, and the remaining partition used to validate the model. This was performed iteratively five times until all five partitions had been used for validation. This enabled the models to be trained on the full dataset. Akaike's Information Criterion (AIC) scores were averaged across cross-validation scores, and the covariate set with the minimum resulting average AIC for all possible combinations (a total of 2,097,151 combinations) was selected for the final model, for each river. Thus, a total of 10,485,755 models were fitted to each river for validation.

Due to the considerable number of models trained, automated covariate selection was performed. All possible combinations of covariates were fitted, and the resulting AIC scores compared. The model with the lowest AIC was used.

2.3.3. Evaluation Metrics

The goodness-of-fit of each model was evaluated using five metrics to account for frequency imbalance in the response class:

$$\text{Probability of Detection} = POD = \frac{a}{a+c}, \tag{5}$$

$$\text{Probability of False Detection} = POFD = \frac{b}{b+d}, \tag{6}$$

$$\text{Precision} = p = \frac{a}{a+b}, \tag{7}$$

$$F_1 \text{ score} = F_1 = 2\frac{p \times POD}{p + POD}, \tag{8}$$

$$\text{Correct Classification Rate} = CCR = \frac{a+d}{N}, \tag{9}$$

where N is the sample size, and a, b, c, d are the number of true positives, true negatives, false positives and false negatives, respectively. The positive class was set to the predicted category, and all other categories assigned the negative class.

Individual metrics, M_i, $i = 1, \ldots, N$ were averaged according to the number of observations within each river:

$$W = \frac{\sum_{i=1}^{N} M_i \, n_i}{\sum_{i=1}^{N} n_i}, \tag{10}$$

where n_i is the number of observations available for river i.

2.3.4. Data Visualisation

Data was visualised in a series of heatmaps displaying the observed, simulated and infilled record. Observed heatmaps display observed data when available, and columns of "missing" data when unavailable, whereas the simulated heatmaps display simulated hydrological state for all months. Infilled heatmaps display observed hydrological state when available, and simulations when observations were unavailable. Reaches were displayed as spatially contiguous due to the high spatial resolution of the data, evidenced in the tick-marks demonstrating the location of each observation site along the observed reach.

Model performance was also demonstrated in monthly time series indicating whether flowing, ponded and dry observations were simulated correctly, and when and where each

state was observed was also presented (Figures 5, 7 and 9). Base Flow Index (BFI) at the linked flow gauging stations sourced from the UK National River Flow Archive [31] was also plotted against the evaluation metrics described in Section 2.3.3.

3. Results

3.1. Overall Performance

CLMs were able to accurately simulate the hydrological regime in the ten chalk streams explored here. Weighted f-scores range from 0.759–0.955, with six out of ten chalk streams achieving scores above 0.935 (Table 3). Of these six chalk streams, the Mimram and the Colne tributaries were the most westerly streams in this dataset.

Table 3. Cumulative logit model performance metrics for each river. POD—Probability of Detection, POFD—Probability of False Detection, CCR—Correct Classification Rate. The rivers are ordered west to east.

	POD	POFD	f-Score	CCR
Misbourne	0.961	0.088	0.942	0.942
Chess	0.968	0.083	0.954	0.954
Bulbourne	0.958	0.116	0.935	0.935
Gade	0.973	0.093	0.955	0.955
Ver	0.964	0.093	0.944	0.949
Mimram	0.954	0.183	0.941	0.939
Beane *	0.852	0.168	0.848	0.909
Rib *	0.775	0.203	0.759	0.849
Ash *	0.780	0.170	0.768	0.857
Stort *	0.908	0.288	0.855	0.861

* Partial proportional odds cumulative logit models.

This spatial pattern is replicated in both the associated Probability of Detection (POD) scores, and Correct Classification Rates (CCRs), which range from 0.775–0.973, and 0.849–0.955, respectively (Table 3). The highest POD scores were associated with the Mimram and Colne tributaries. This is consistent with the CCRs, which ranged from 0.935–0.955 in the aforementioned rivers, and from 0.849–0.909 in the most easterly rivers— Rib, Ash and Stort.

Probability of False Detection (POFD) scores did not exhibit as clear a pattern, ranging from 0.083–0.288 (Table 3). The Colne tributaries had the lowest associated POFD scores of 0.083–0.116, with a POFD score of 0.083 associated with the Chess. However, the Mimram had a POFD score of 0.183, higher than those of the Ash and Beane, 0.170, and 0.168, respectively.

Figure 3 suggests that this apparent spatial pattern may be related to the BFI of each river. Models trained on rivers with relatively high BFI exhibited relatively high performance.

The proportion of ponding observations within each dataset ranged from 2.81–16.78%. The Colne tributaries and Mimram had the lowest proportion of ponding observations (2.81–4.50%). This suggests the performance of the CLMs may be related to the proportion of ponding observations within the data. This is supported by state-specific POD across all 10 rivers: flowing = 0.964, ponded = 0.265, dry = 0.829.

The same spatial pattern was evident in the validity of the proportional odds assumption. No violations were found in the Colne tributaries or the Mimram, whilst further east, violations were found on the Beane, Rib, Ash and Stort. These four rivers were those with the highest proportions of ponded observations (13.2%, 14.2%, 16.8%, and 11.1%, respectively). The variables that violated the proportional odds assumption varied between these rivers (Table 4). According to the likelihood ratio tests, the assumption was violated for flow and hydrological state in both the previous and following months on all four of them, whilst there were violations for distance to confluence on the Ash and the Stort, and across all explanatory variables on the Beane. This means that the coefficients on these

variables had been allowed to vary with hydrological state and suggests variability in their relationship with hydrological state.
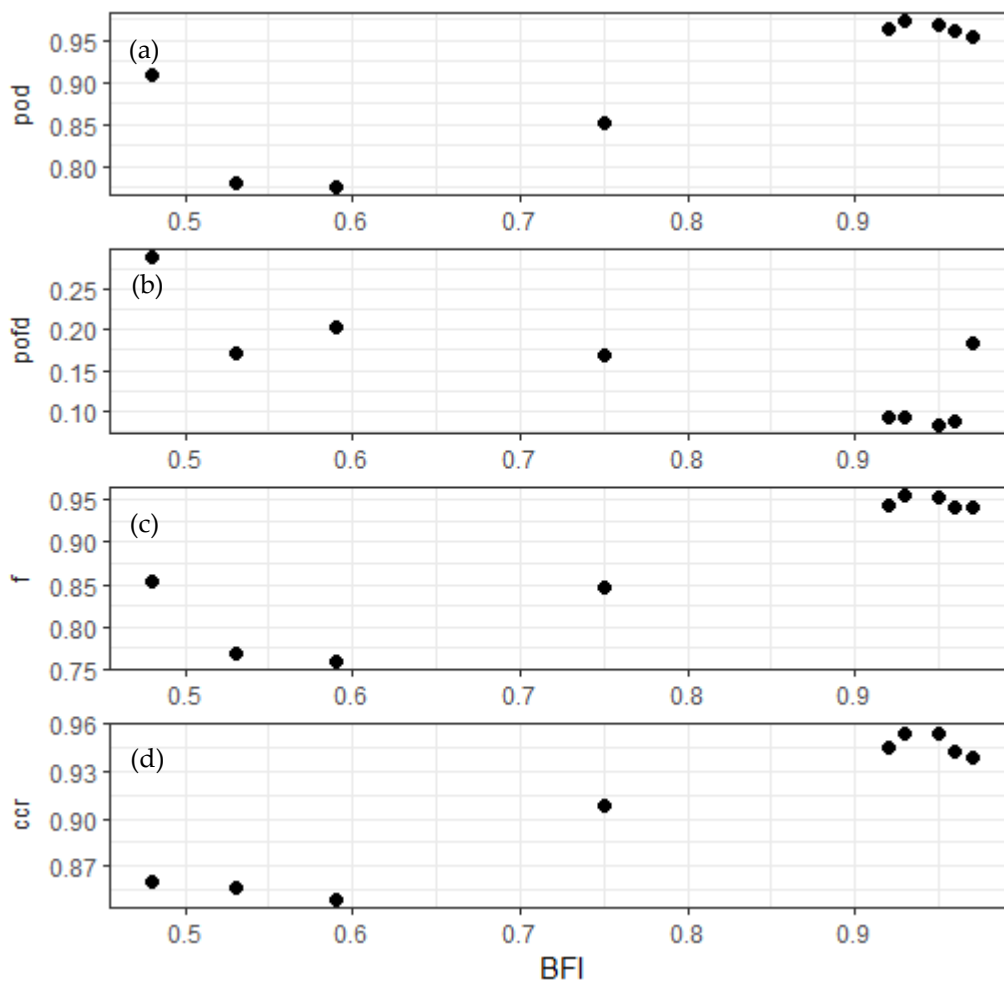

Figure 3. Relationships between Base Flow Index and weighted evaluation metrics. (**a**) POD; (**b**) POFD; (**c**) f; (**d**) CCR.

Table 4. Coefficients that were identified (x) as significant predictors of hydrological state in their respective models. * denotes non-proportional odds variables.

River	Misbourne	Chess	Bulbourne	Gade	Ver	Mimram	Beane	Rib	Ash	Stort
Flow average	X	X	X	X	X		X *	X *	X *	X *
Average water table height	X	X	X	X	X	X	X *	X	X	X
Average precipitation				X			X *	X		X
Precipitation frequency	X			X		X			X	X
Average percolation	X	X		X					X	X*
Percolation frequency	X	X	X				X *	X	X	X
Distance upstream of confluence	X	X	X	X	X	X	X *	X	X *	X *
Month	X			X		X	X	X *		X
Previous month state	X	X	X	X	X	X	X *	X *	X *	X *
Following month state	X	X	X	X	X	X	X *	X *	X *	X *

Water table height, distance upstream of confluence and previous and following month state variables were used to train all 10 models. However, not all explanatory variables were used in each model (Table 4). Precipitation average, precipitation frequency and percolation average were used in 4–5 models each. At least one rainfall explanatory

variable was used in nine models. The Mimram model did not use a rainfall variable to simulate hydrological state, instead using water table height, distance upstream of confluence, month and previous and following state.

Three catchments were selected for further inspection of model performance: one river with highly attenuated hydrological regime (Gade), one more responsive (Ash) and one featuring spatial fragmentation of flow (Misbourne).

3.2. Gade

The Gade observation record was 69.1%, 3.0% and 27.9% flowing, ponded and dry observations, respectively. This was well approximated by the model simulations, which were 70.1%, 0.0% and 29.0% flowing, ponded and dry, respectively. The Gade model simulated hydrological state with the highest associated f-score (0.955) and POD (0.973) (Table 1) using mean flow, mean water table height, mean precipitation, mean percolation, distance upstream of confluence, previous and following month state (Table 3). Figures 4 and 5 demonstrate the accurate simulation of the flowing/non-flowing division across the full time series. The timing of drying and rewetting is accurately reconstructed, and the extent of dry reaches is simulated well within the model. Despite the lack of ponded simulation, the infrequency of ponded observations (2.97% of observations were ponded) meant that this had minimal impact on the ability of this model to accurately reconstruct dry/wet dynamics, source migration and stream profile (Figures 4 and 5).

3.3. Misbourne

Flowing, ponded and dry states comprised 64.6%, 3.7% and 31.7% of the Misbourne observation record, respectively. The simulation dataset had similar proportions with 66.4%, 0.0% and 33.6% flowing, ponded and dry simulations, respectively. The Misbourne is also one of the westerly chalk streams, however, its hydrological regime (Figure 6) contrasts with that of the Gade. Whereas the Gade typically exhibits a clear division between flowing and non-flowing reaches, the Misbourne often has contiguous dry reaches downstream of flowing reaches. However, despite this behaviour, the CLM trained on the Misbourne was able to accurately simulate the hydrological regime, with comparable performance to the Gade CLM. The Misbourne CLM had a weighted f-score of 0.942 and demonstrated ability to simulate the division between where flowing reaches end and drying reaches begin (Figures 6 and 7). Whilst 3.71% of Misbourne observations were ponded, the CLM did not make any ponded estimations of state (Figures 6 and 7). However, the spatial extent, duration and timing of non-flowing reaches were accurately simulated by the CLM.

3.4. Ash

Flowing, ponded and dry states comprised 59.9%, 16.8% and 23.3% of the Ash observation record, respectively. The simulation dataset consisted of 65.1%, 11.4% and 23.5% flowing, ponded and dry simulations, respectively. The Ash is more easterly than both the Gade and Misbourne, lying between the Rib and Stort. This is reflected in the flashiness observed in Figure 8, resulting, in part, due to the clay and glacial drift deposits increasing the drainage densities in these areas. Similarly, the BFI on the Ash at Mardock is 0.53, whereas the Gade at Bury Mill has a BFI of 0.93 [30].

Likelihood ratio tests suggested the Ash data violated the proportional odds assumption. Accordingly, a partial proportional odds CLM was trained, using flow average, water table height average, precipitation frequency, percolation average, percolation frequency, distance upstream of confluence and previous and following month states to simulate hydrological state. Specifically, flow average, distance upstream of confluence, and previous and following month hydrological state violated the proportional odds assumption and therefore the coefficients for these variables were allowed to vary between states.

The model simulated hydrological state with a weighted average f-score of 0.77. This is, in part, due to challenges estimating ponding. The Ash had the highest proportion

of ponding observations, with 16.78% of states observed to be ponded. This resulted in 11.36% of Ash hydrological state estimations being ponded. This under-estimation of the prominence of ponding is compounded by challenges surrounding its timing, reducing the performance of the Ash model (Figures 8 and 9). However, grouping ponding and flowing observations into a 'wet' category increased the weighted f-scores from 0.77 to 0.86, demonstrating the ability of this model to simulate wetting and drying dynamics in the Ash.
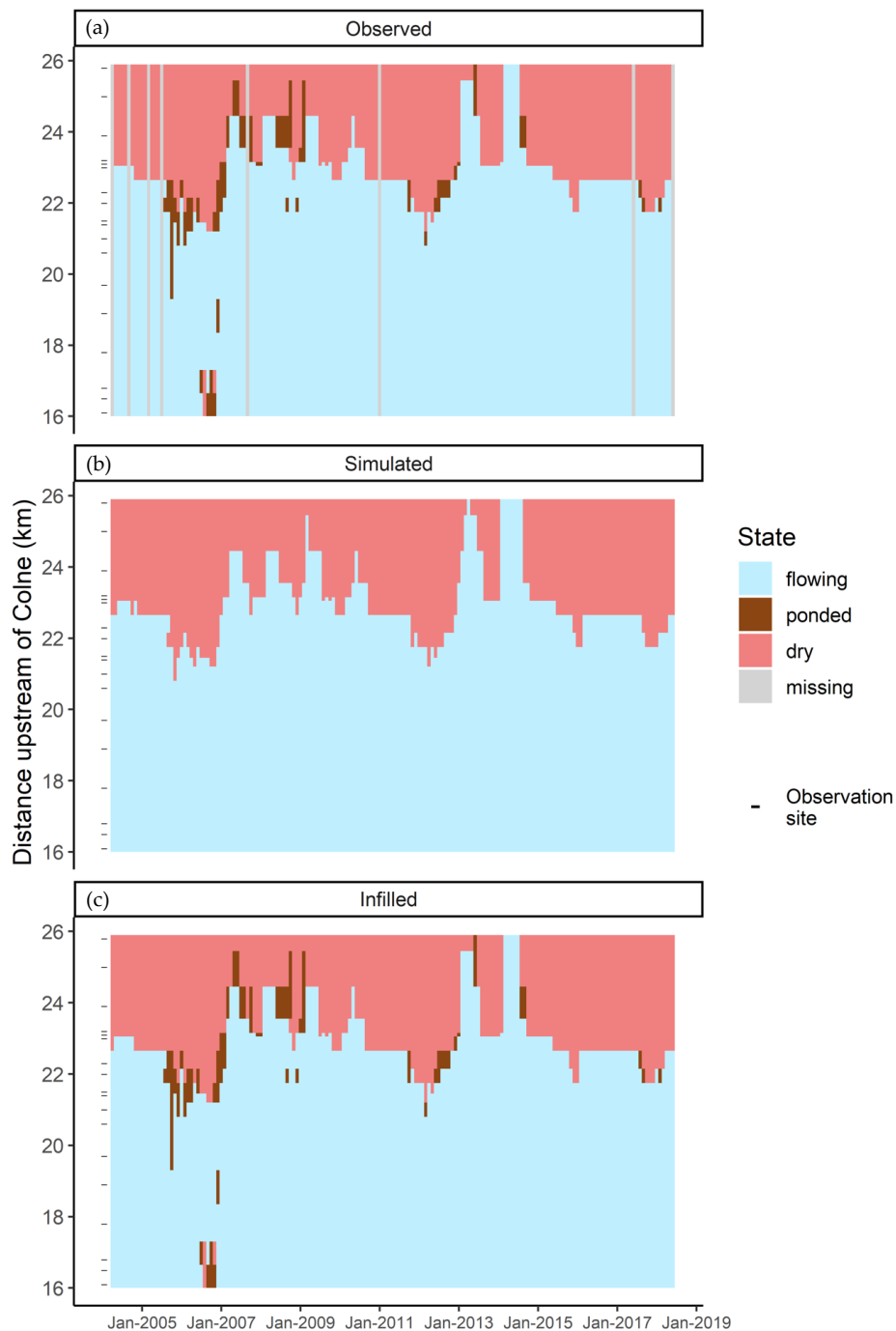
Figure 4. Monthly time series of observed (**a**), simulated (**b**), and infilled (**c**) hydrological state on the Gade, where infilled states are observations where available and simulations where observations are unavailable. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.
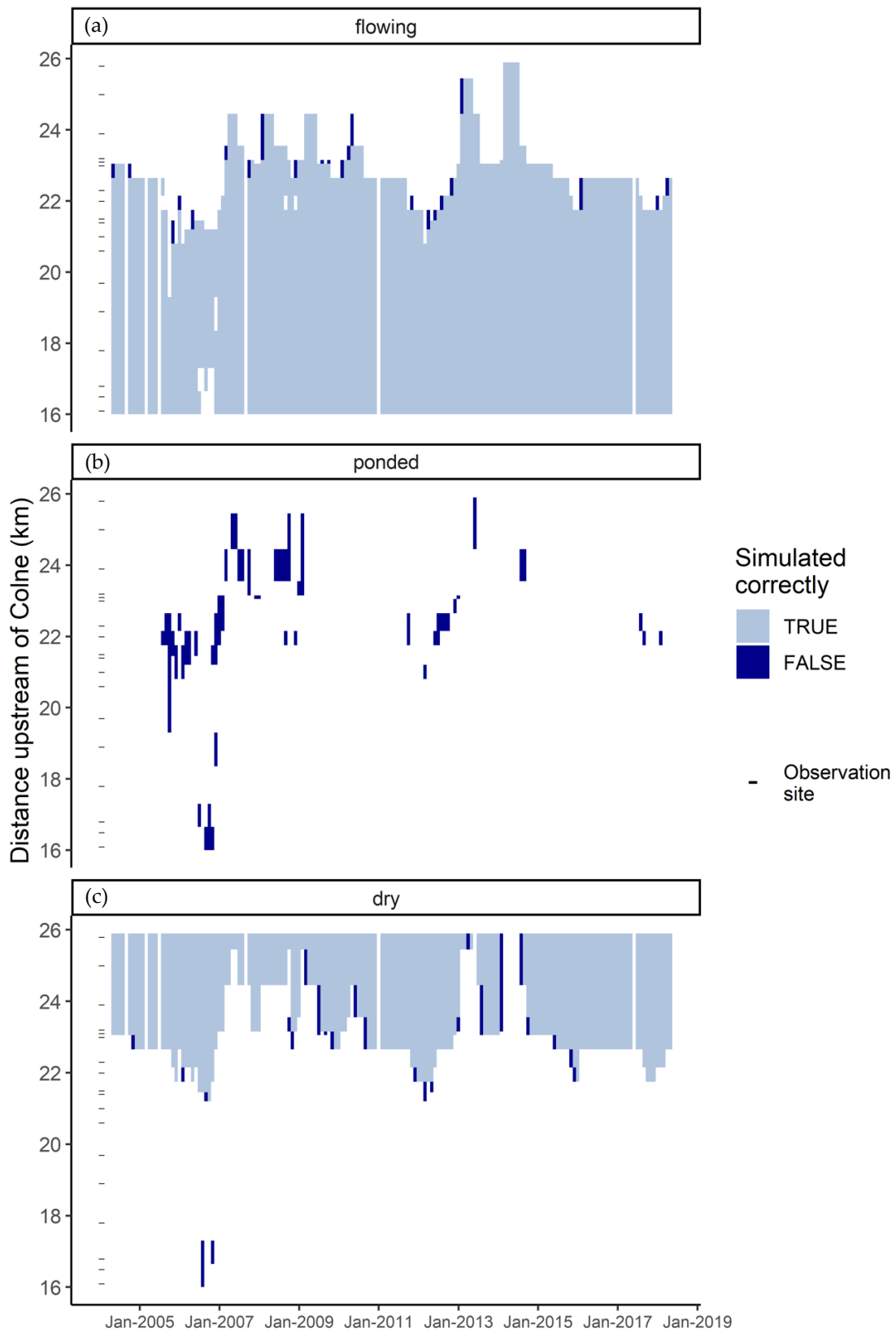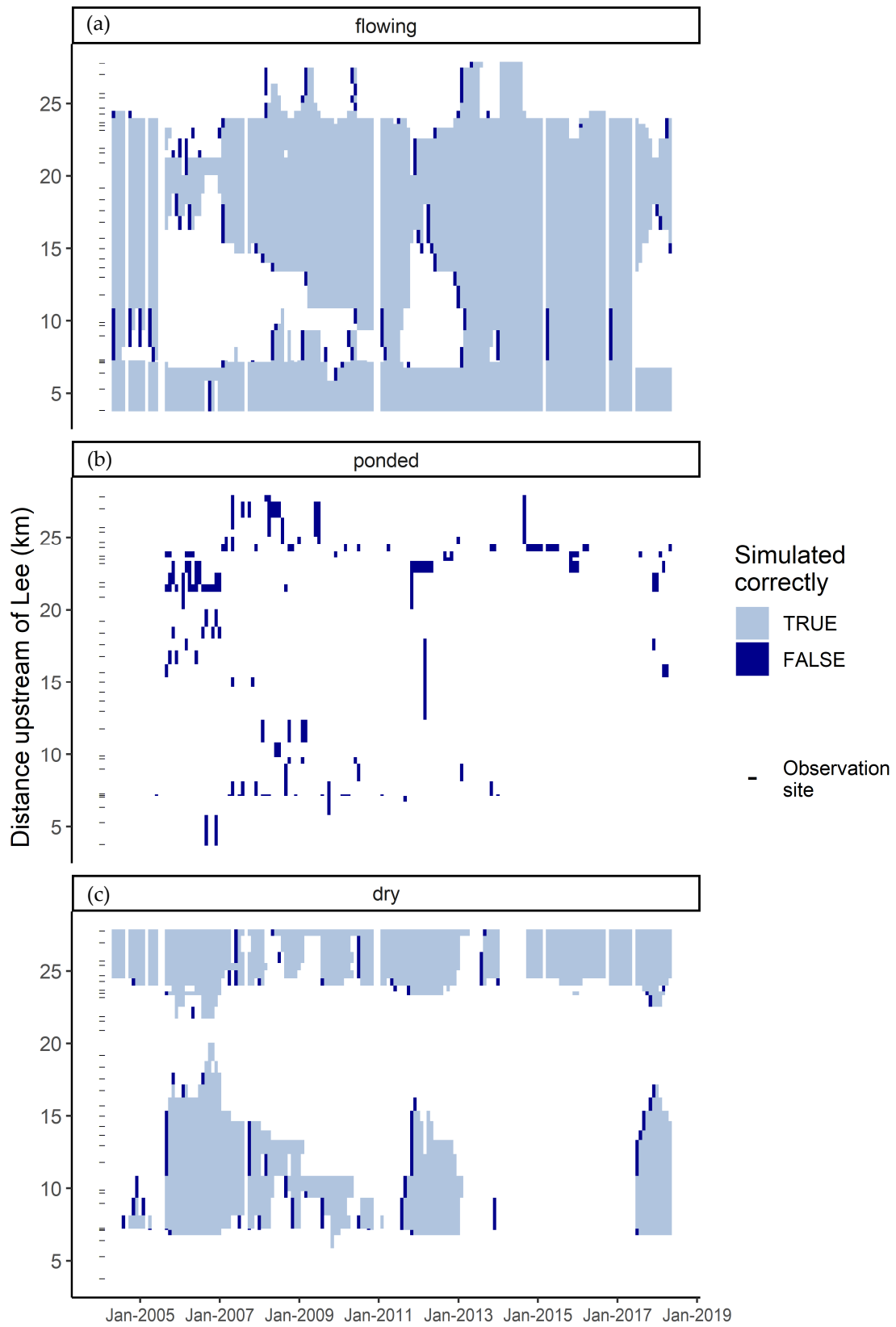
Figure 5. Monthly time series indicating whether flowing (**a**), ponded (**b**), and dry (**c**) states were simulated correctly and when and where the states were observed on the Gade. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.
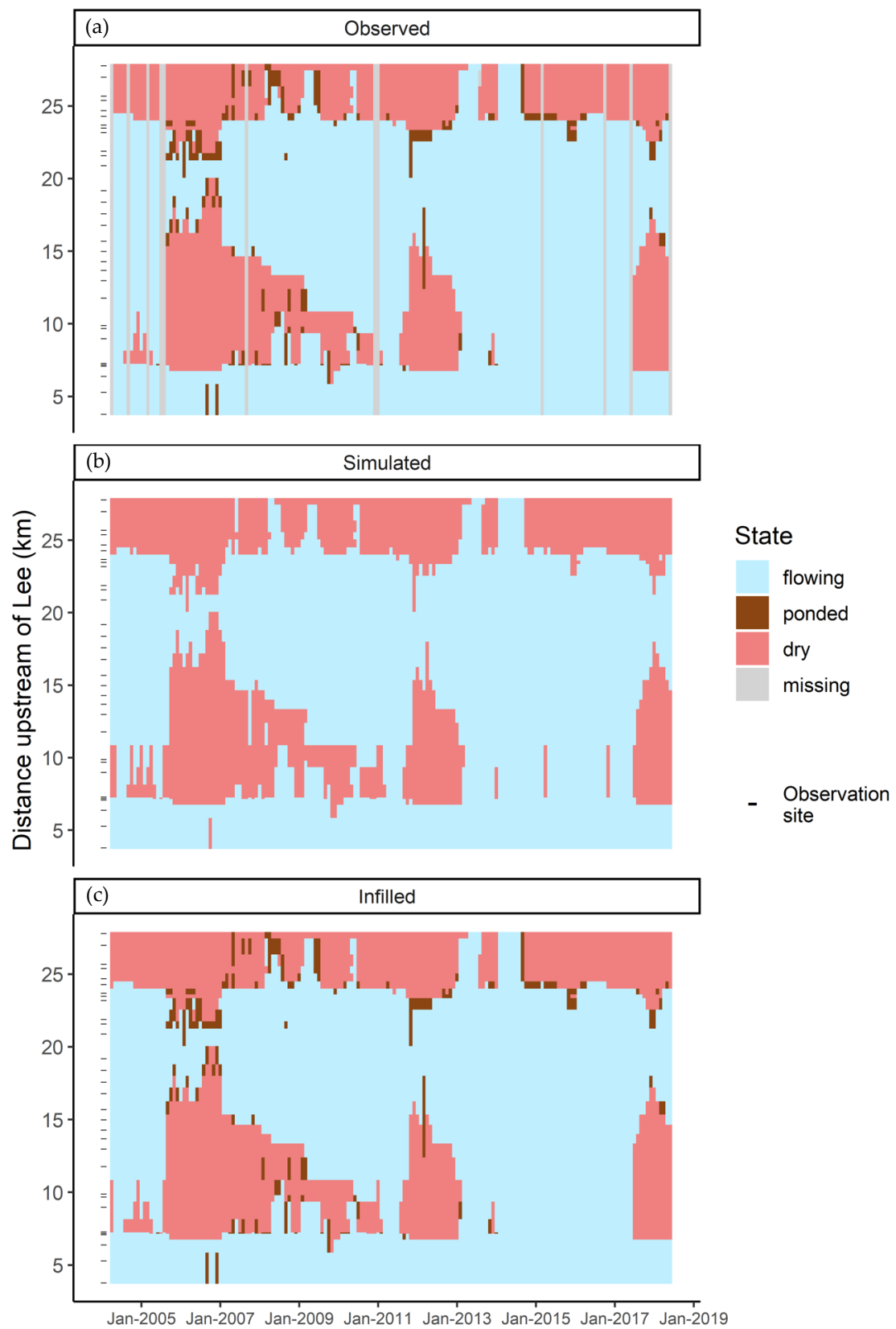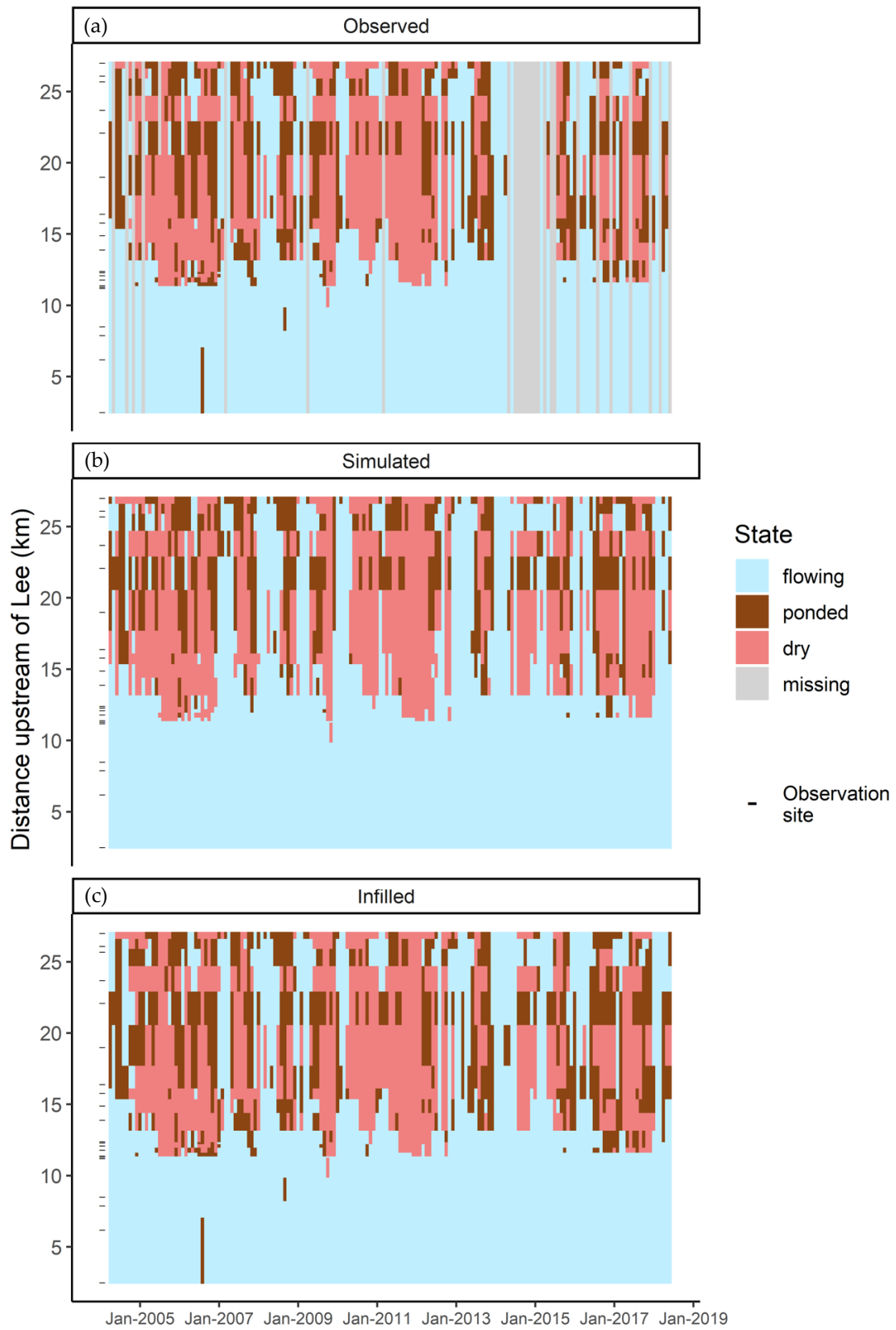
Figure 6. Monthly time series of observed (**a**), simulated (**b**), and infilled (**c**) hydrological state on the Misbourne, where infilled states are observations where available and simulations where observations are unavailable. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.
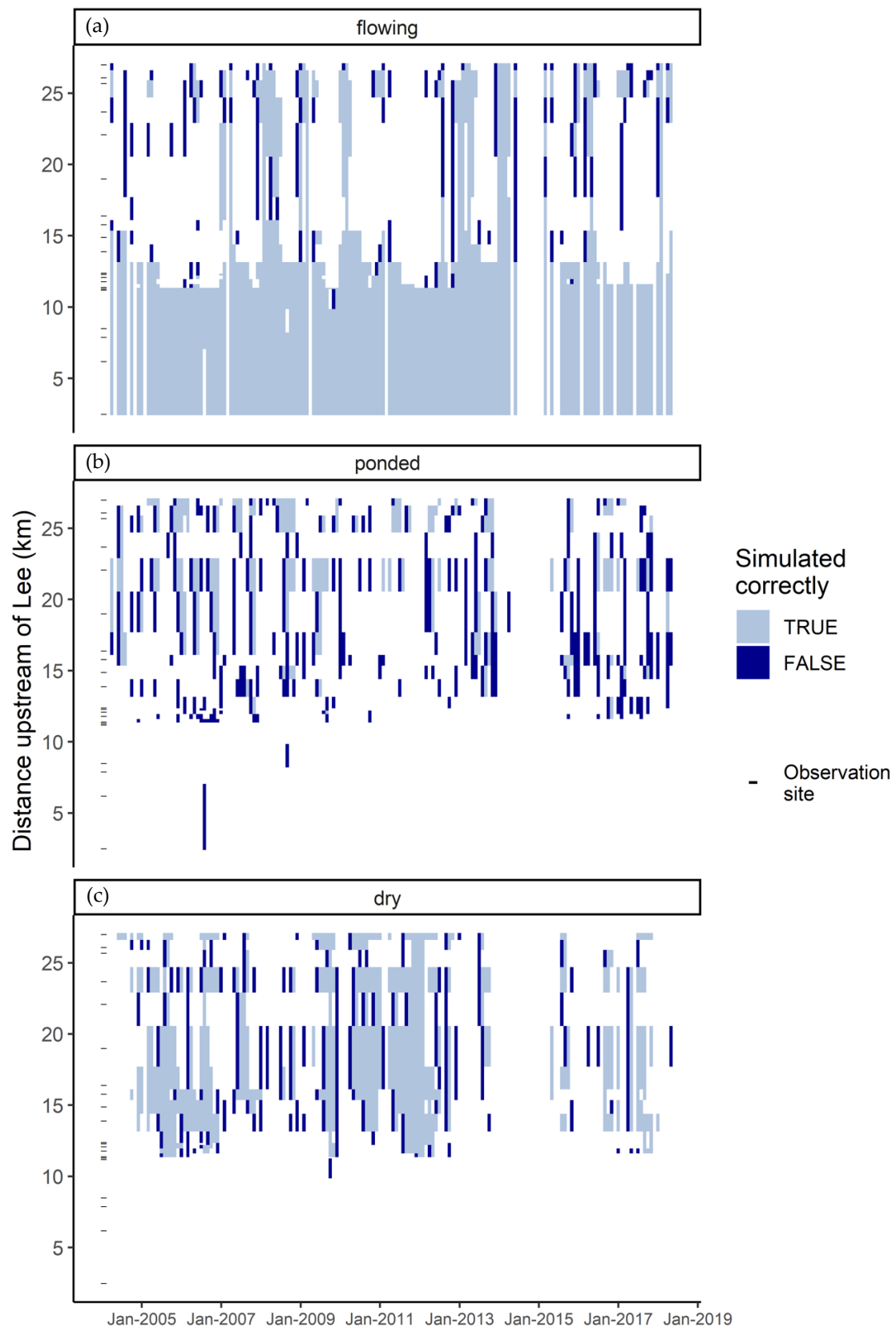
Figure 7. Monthly time series indicating whether flowing (**a**), ponded (**b**), and dry (**c**) states were simulated correctly and when and where the states were observed on the Misbourne. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.

Figure 8. Monthly time series of observed (**a**), simulated (**b**), and infilled (**c**) hydrological state on the Ash, where infilled states are observations where available and simulations where observations are unavailable. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.

Figure 9. Monthly time series indicating whether flowing (**a**), ponded (**b**), and dry (**c**) states were simulated correctly and when and where the states were observed on the Ash. Hydrological state is assumed to be consistent along observed reaches due to the high spatial resolution of the data.

4. Discussion

4.1. Controls on Model Performance

Variations in the complexity of catchments in the Chilterns partially determine model performance. Whilst the evaluation of the CLMs demonstrated their generally strong performance in simulating hydrological state, the performance of models generally decreases with distance east. This is likely driven by variations in the complexity in hydrological regime. Easterly rivers exhibit higher spatial fragmentation [37] and more frequent transitioning between states-upstream/downstream compared to groundwater-dominated catchments in the west. Furthermore, whilst observed in all rivers, seasonal network contraction and expansion is more evident on the flashier rivers [29]. Due to the uncertainty associated with estimating parameters, rivers that are dependent on fewer, simpler factors (i.e., groundwater-fed catchments in the west) often have less uncertainty, and thus higher model performance.

BFI estimates the influence of stored water on the hydrological regime of a river [30]. The positive relationships between model performance and BFI (Figure 3) suggest that model performance is dependent on the influence of groundwater on hydrological regime, with model performance being better in rivers that are more influenced by groundwater (high BFI). However, whilst all 10 catchments are underlain by chalk, the permeability of superficial hydrogeology varies between rivers, resulting in spatial fragmentation variation [29]. The eastern rivers (Beane, Rib, Ash and Stort) are more frequently underlain by superficial deposits and have higher spatial variation in their distribution. Furthermore, intra-river variation in superficial deposit distribution likely generates higher spatial fragmentation, which is more challenging to simulate. Despite the non-uniform distribution of superficial deposits, the models herein account for intra-river site variation linearly with distance from the confluence. Improved accountancy of intra-river geological variation within models may increase model performance in rivers that exhibit localised hydrogeology. Previous research found associations between hydrological intermittence and a range of catchment characteristics, including catchment area, altitude, geology, land cover [17,19,38,39], and total precipitation and forest cover [19,20]. However, recent research performed in France did not identify either catchment area or altitude as drivers of flow intermittence [18]. This underlines the challenges in determining the drivers of hydrological intermittence observed in the chalk streams presented herein.

4.2. Influence of Modelling Approach

The modelling approach may also influence how well inter- and intra-river variation has been accounted for by alternative modelling approaches, including random forest [21], and both random forests and neural networks [18]. Each of these approaches offers ability to characterise non-linear relationships, meaning they are well-suited to simulating complex systems, particularly appropriate when estimating the relationships driving intermittence across a wide range of environmental conditions, such as those in France and Queensland. Disadvantages of these approaches are the need to tune a relatively high number of parameters, particularly challenging when data availability is limited. The decision to fit models to each river in this study constrained the data availability but allowed relationships to be river-specific without accounting for the relationships driving variation between rivers. The advantages of CLMs (reduced parameterisation required for model fitting) outweighed the limitations and resulted in the strong performance observed.

4.3. Suitability for Infilling

A key aim of this research was to develop capability for infilling hydrological state datasets. Broadly speaking, the strong performance of the models (Table 3) demonstrates the potential of this approach in applications requiring infilling of datasets. The mean f-score across the 10 study rivers was 0.89, and as high as 0.955 on the most accurately simulated rivers. In addition to high values of performance metrics, visual inspection of observed and simulated heatmaps (Figures 4, 6 and 8) suggests appropriate reconstruction

of the spatial and temporal dynamics of hydrological intermittence, including source migration and partitioning between flowing and non-flowing. Nevertheless, in rare instances, the methodology applied herein encounters some challenges. For example, the unique hydrology and hydrogeology of the downstream drying section of the Misbourne (Figure 6) presents some challenges for reconstructing data. The importance of hydrological state at time step $(t - 1)$ for simulating state at time step t in the Misbourne CLM, and the use of modal state of that site in place of missing data at $(t - 1)$, produces some seemingly erroneous values for hydrological state (e.g., ponding amongst a considerable amount of flowing data; Figure 6). Potential solutions to this include using the most recent hydrological state as a simple approximation for missing data, or an increased reliance on upstream/downstream sites. Nevertheless, this rare example should not detract from the general success of the methodology in reconstructing the hydrological state, and within an operational context such erroneous data could easily be identified and quality checked. Complete data are necessary for a wide range of applications, including those concerning the duration of flowing and dry phases, an important influence on ecological responses to drying and rewetting dynamics [40,41], and gaps in the observation record present challenges when they coincide with points of interest in time, such as an invertebrate survey [42]. The methodology presented herein provides clear potential to address such concerns.

4.4. Challenges and Limitations

The variability in superficial deposit, and thus permeability, and distribution is reflected in localised ponding [29]. Ponding is an ecologically important intermediate state in the drying and rewetting of sites. The ability for the models to simulate ponding dynamics varied between rivers. Generally, ponding simulation was more accurate in rivers with higher proportions of ponding observations (e.g., the Ash). Elsewhere, limited occurrence of ponding in observed data presented challenges in the calibration of models for simulation. Data scarcity poses challenges for accurately characterising the relationships that drive hydrological state dynamics. Limited data on ponding has also proved challenging for validation of state estimation in other studies, for example [28]. Nevertheless, in many of the study catchments (e.g., the Misbourne), the relative infrequency of ponding compared to dry and flowing states means that ponding was never simulated in some instances, but with few observations to compare against, the CLMs performed strongly under evaluation. In cases where accurate simulation of ponding is a priority, the use of an alternative cost function that weights ponding simulation accuracy more heavily may improve ponding simulation.

Whilst the lack of ponded observations posed considerable challenges for simulating this state on most rivers, this was not the only factor limiting performance on the Ash. The proportions of flowing, ponded and dry states observed on the Ash were 0.6, 0.17 and 0.23, respectively. However, the model was able to simulate dry more accurately than ponded observations (Figures 8 and 9). The onset and termination of hydrological states also posed a challenge for simulation. This is most evident in Figures 5 and 7, where despite the timing of hydrological states being well approximated, the majority of incorrect classifications occur at the beginning or ending of a period of consistent hydrological state. Therefore, the performance of these models is limited on flashier rivers due to rapidly changing hydrological states. This is likely in part driven by the relatively high weighting of lagged hydrological state variables, due to this generally being an accurate predictor of hydrological state. However, the complexity of the relationships driving state onset and termination also likely limits the performance. Therefore, improved understanding and accounting for these relationships in partnership with constrained weighting of lagged state variables or alternative evaluation metrics may further improve the performance of these models.

A particular challenge for simulating hydrological state is ensuring consistency of qualitative observations where thresholds between categories may not be clear. However,

consistency and accuracy of observations is ensured by a defined protocol followed by trained hydrologists. This was further accounted for by only training models on observations taken since 2004, following the review and standardisation of the monitoring approach.

Despite this subsetting of the dataset, such an extensive period of record (2004–2018) is rare for hydrological state data compared to previous studies, for example [18,21], although these studies benefit from a wider spatial extent than the dataset used herein. Whilst they would need to be re-trained for the specific application, the modelling approach presented herein could be applied within other intermittent river contexts internationally in order to accurately simulate the hydrological regime, providing there is sufficient data availability. As such, this study provides a valuable addition to existing demonstrations of the potential for statistical models to simulate hydrological intermittence [15–22].

4.5. Potential for Future Applications

Extensions to this research would facilitate investigations into the potential impacts of climate, hydrogeology and land cover on these systems. Specifically, accurate reconstruction of hydrological state dynamics also suggests that temporal extrapolation beyond the observational record is possible. However, this is dependent on the relationships driving the intermittence exhibiting stationarity. Selecting covariate sets according to minimum AIC may result in overfitting, and therefore reduced extrapolation potential. As the temporal extent of the dataset grows, and/or with the incorporation of data from other sources, such as citizen science initiatives [43], further investigation into the suitability of the models for extrapolation would evaluate, and potentially improve, the potential of this methodology for temporal extrapolation. This includes analysing the impact of adding random effects to the models, which would demonstrate the potential to accurately extrapolate these results both spatially and temporally. If satisfied, temporal extrapolation would enable further research into the nature of potential climate change impacts on intermittent rivers, which previous research demonstrates has the potential to influence the hydrological regime of intermittent rivers [44–46]. Furthermore, accurate temporal extrapolation would facilitate operational management with wide utility in forecasting, drought tracking, and assessing the influence of abstraction pressures on the future of these valuable ecosystems.

5. Conclusions

The understanding of intermittent rivers, and the processes driving them, is limited despite their widespread prevalence and importance. However, statistical simulation of the hydrological regime of these rivers by estimating the relationships driving hydrological state provides ample opportunity for developing this understanding further and improving their protection and appreciation. As demonstrated here, it is possible to simulate hydrological state with unprecedented performance using CLMs. This not only provides insight into what occurred in the record when observations are unavailable, but also enabled further investigation into the processes driving spatial variation in performance. The clear distinction in performance between rivers in the east and west, as well as the association between BFI and performance, demonstrates a critical relationship between groundwater influence and the ability to simulate the hydrological regime in an intermittent river.

This work enables further understanding to be garnered from future research. In its simplest form, this includes linking data with newly available infills, however, this should be expanded into further investigations into how climate change and projected abstraction pressures may influence intermittent rivers in the future. As well as temporal extrapolation, there is potential in improving the understanding of intermittent rivers that do not have as extensive of a record. Infilled datasets may be used to improve the accuracy of metric estimation and calculate metrics that are dependent on complete data, such as duration of flowing and dry phases. However, it would also enable the tracking of hydrological state when observations are not possible. This work presents a key step in improving

the characterisation of intermittent streams, understanding how they interact with the surrounding environment and the protection of these key ecosystems.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AIC	Akaike's Information Criterion
BFI	Base Flow Index
CCR	Correct Classification Rate
CLM	Cumulative Logit Model
IRES	Intermittent Rivers and Ephemeral Streams
m AOD	meters Above Ordnance Datum
POD	Probability of Detection
POFD	Probability of False Detection

References

1. Datry, T.; Larned, S.T.; Tockner, K. Intermittent Rivers: A Challenge for Freshwater Ecology. *Bioscience* **2014**, *64*, 229–235. [CrossRef]
2. Smith, H.; Wood, P.J.; Gunn, J. The influence of habitat structure and flow permanence on invertebrate communities in karst spring systems. *Hydrobiologia* **2003**, *510*, 53–66. [CrossRef]
3. Stubbington, R. The hyporheic zone as an invertebrate refuge: A review of variability in space, time, taxa and behaviour. *Mar. Freshw. Res.* **2012**, *63*, 293. [CrossRef]
4. Datry, T.; Larned, S.T.; Fritz, K.M.; Bogan, M.T.; Wood, P.J.; Meyer, E.I.; Santos, A.N. Broad-scale patterns of invertebrate richness and community composition in temporary rivers: Effects of flow intermittence. *Ecography* **2014**, *37*, 94–104. [CrossRef]
5. Leigh, C.; Datry, T. Drying as a primary hydrological determinant of biodiversity in river systems: A broad-scale analysis. *Ecography* **2017**, *40*, 487–499. [CrossRef]
6. Stubbington, R.; England, J.; Wood, P.J.; Sefton, C.E.M. Temporary streams in temperate zones: Recognizing, monitoring and restoring transitional aquatic-terrestrial ecosystems. *WIREs Water* **2017**, *4*, e1223. [CrossRef]
7. Hill, M.J.; Milner, V.S. Ponding in intermittent streams: A refuge for lotic taxa and a habitat for newly colonising taxa? *Sci. Total Environ.* **2018**, *628–629*, 1308–1316. [CrossRef] [PubMed]
8. Datry, T.; Pella, H.; Leigh, C.; Bonada, N.; Hugueny, B. A landscape approach to advance intermittent river ecology. *Freshw. Biol.* **2016**, *61*, 1200–1213. [CrossRef]
9. Leibowitz, S.G.; Wigington, P.J.; Rains, M.C.; Downing, D.M. Non-navigable streams and adjacent wetlands: Addressing science needs following the Supreme Court's Rapanos decision. *Front. Ecol. Environ.* **2008**, *6*, 364–371. [CrossRef]
10. Fritz, K.; Cid, N.; Autrey, B. Governance, legislation, and protection of intermittent rivers and ephemeral streams. In *Intermittent Rivers and Ephemeral Streams*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 477–507. ISBN 9780128038352.

11. Skoulikidis, N.T.; Sabater, S.; Datry, T.; Morais, M.M.; Buffagni, A.; Dörflinger, G.; Zogaris, S.; Del Mar Sánchez-Montoya, M.; Bonada, N.; Kalogianni, E.; et al. Non-perennial Mediterranean rivers in Europe: Status, pressures, and challenges for research and management. *Sci. Total Environ.* **2017**, *577*, 1–18. [CrossRef]
12. Costigan, K.H.; Jaeger, K.L.; Goss, C.W.; Fritz, K.M.; Goebel, P.C. Understanding controls on flow permanence in intermittent rivers to aid ecological research: Integrating meteorology, geology and land cover. *Ecohydrology* **2016**, *9*, 1141–1153. [CrossRef]
13. Larned, S.T.; Datry, T.; Arscott, D.B.; Tockner, K. Emerging concepts in temporary-river ecology. *Freshw. Biol.* **2010**, *55*, 717–738. [CrossRef]
14. Van Meerveld, H.J.I.; Sauquet, E.; Gallart, F.; Sefton, C.; Seibert, J.; Bishop, K. Aqua temporaria incognita. *Hydrol. Process.* **2020**. [CrossRef]
15. Jensen, C.K.; McGuire, K.J.; Shao, Y.; Andrew Dolloff, C. Modeling wet headwater stream networks across multiple flow conditions in the Appalachian Highlands. *Earth Surf. Process. Landf.* **2018**, *43*, 2762–2778. [CrossRef]
16. Yu, S.; Bond, N.R.; Bunn, S.E.; Xu, Z.; Kennard, M.J. Quantifying spatial and temporal patterns of flow intermittency using spatially contiguous runoff data. *J. Hydrol.* **2018**, *559*, 861–872. [CrossRef]
17. Kaplan, N.H.; Blume, T.; Weiler, M. Predicting probabilities of streamflow intermittency across a temperate mesoscale catchment. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 5453–5472. [CrossRef]
18. Beaufort, A.; Carreau, J.; Sauquet, E. A classification approach to reconstruct local daily drying dynamics at headwater streams. *Hydrol. Process.* **2019**. [CrossRef]
19. González-Ferreras, A.M.; Barquín, J. Mapping the temporary and perennial character of whole river networks. *Water Resour. Res.* **2017**, *53*, 6709–6724. [CrossRef]
20. Jaeger, K.L.; Sando, R.; McShane, R.R.; Dunham, J.B.; Hockman-Wert, D.P.; Kaiser, K.E.; Hafen, K.; Risley, J.C.; Blasch, K.W. Probability of Streamflow Permanence Model (PROSPER): A spatially continuous model of annual streamflow permanence throughout the Pacific Northwest. *J. Hydrol. X* **2019**, *2*, 100005. [CrossRef]
21. Yu, S.; Bond, N.R.; Bunn, S.E.; Kennard, M.J. Development and application of predictive models of surface water extent to identify aquatic refuges in Eastern Australian temporary stream networks. *Water Resour. Res.* **2019**, *55*, 9639–9655. [CrossRef]
22. Bond, N.R.; Kennard, M.J. Prediction of hydrologic characteristics for ungauged catchments to support hydroecological modeling. *Water Resour. Res.* **2017**. [CrossRef]
23. Bonada, N.; Cañedo-Argüelles, M.; Gallart, F.; von Schiller, D.; Fortuño, P.; Latron, J.; Llorens, P.; Múrria, C.; Soria, M.; Vinyoles, D.; et al. Conservation and management of isolated pools in temporary rivers. *Water* **2020**, *12*, 2870. [CrossRef]
24. Drummond, L.R.; McIntosh, A.R.; Larned, S.T. Invertebrate community dynamics and insect emergence in response to pool drying in a temporary river. *Freshw. Biol.* **2015**, *60*, 1596–1612. [CrossRef]
25. Kerezsy, A.; Gido, K.; Magalhaes, M.F.; Skelton, P.H. The biota of intermittent rivers and ephemeral streams: Fishes. In *Intermittent Rivers and Ephemeral Streams: Ecology and Management*; Datry, T., Bonada, N., Boulton, A.J., Eds.; Elsevier, Inc.: Cambridge, MA, USA, 2017; pp. 273–298.
26. Arthington, A.H.; Balcombe, S.R. Extreme flow variability and the 'boom and bust' ecology of fish in arid-zone floodplain rivers: A case history with implications for environmental flows, conservation and management. *Ecohydrology* **2011**, *4*, 708–720. [CrossRef]
27. Stubbington, R.; England, J.; Acreman, M.; Wood, P.J.; Westwood, C.; Boon, P.; Mainstone, C.; Macadam, C.; Bates, A.; House, A.; et al. *The Natural Capital of Temporary Rivers: Characterising the Value of Dynamic Aquatic-Terrestrial Habitats*; Valuing Nature Natural Capital Synthesis Report VNP12; 2018; Available online: https://valuing-nature.net/TemporaryRiverNC (accessed on 11 February 2021).
28. Querner, E.P.; Froebrich, J.; Gallart, F.; Cazemier, M.M.; Tzoraki, O. Simulating streamflow variability and aquatic states in temporary streams using a coupled groundwater-surface water model. *Hydrol. Sci. J.* **2016**, *61*, 146–161. [CrossRef]
29. Sefton, C.E.M.; Parry, S.; England, J.; Angell, G. Visualising and quantifying the variability of hydrological state in intermittent rivers. *Fundam. Appl. Limnol.* **2019**. [CrossRef]
30. Gustard, A.; Bullock, A.; Dixon, J.M. *Low Flow Estimation in the United Kingdom (IH Report)*; Centre for Ecology & Hydrology: Wallingford, UK, 1992; p. 292. ISBN 0948540451.
31. National River Flow Archive. Available online: https://nrfa.ceh.ac.uk/ (accessed on 18 September 2020).
32. Gallart, F.; Prat, N.; García-Roger, E.M.; Latron, J.; Rieradevall, M.; Llorens, P.; Barberá, G.G.; Brito, D.; De Girolamo, A.M.; Lo Porto, A.; et al. A novel approach to analysing the regimes of temporary streams in relation to their controls on the composition and structure of aquatic biota. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 3165–3182. [CrossRef]
33. Environment Agency: Groundwater and Hydrology. Monthly Water Situation Report: Hertfordshire and North London. 2021. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/951151/Hertfordshire_and_North_London_Water_Situation_Report_December_2020.pdf (accessed on 21 January 2021).
34. Hyoms, C.M. Development of a Soil Moisture Model for the Estimation of Percolation. Ph.D. Thesis, The City University, London, UK, 1980.
35. Agresti, A. *Categorical Data Analysis*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; pp. 274–282. ISBN 0471249688.
36. Christensen, R.H.B. Regression Models for Ordinal Data. R Package Version 2019.12-10. Available online: https://CRAN.R-project.org/package=ordinal (accessed on 18 September 2020).

37. Hargis, C.D.; Bissonette, J.A.; David, J.L. *The Behavior of Landscape Metrics Commonly Used in the Study of Habitat Fragmentation*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1998. [CrossRef]

38. Snelder, T.H.; Datry, T.; Lamouroux, N.; Larned, S.T.; Sauquet, E.; Pella, H.; Catalogne, C. Regionalization of patterns of flow intermittence from gauging station records. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 2685–2699. [CrossRef]

39. D'Ambrosio, E.; De Girolamo, A.M.; Barca, E.; Ielpo, P.; Rulli, M.C. Characterising the hydrological regime of an ungauged temporary river system: A case study. *Environ. Sci. Pollut. Res. Int.* **2017**, *24*, 13950–13966. [CrossRef]

40. Steward, A.L.; von Schiller, D.; Tockner, K.; Marshall, J.C.; Bunn, S.E. When the river runs dry: Human and ecological values of dry riverbeds. *Front. Ecol. Environ.* **2012**, *10*, 202–209. [CrossRef]

41. Leigh, C.; Boulton, A.J.; Courtwright, J.L.; Fritz, K.; May, C.L.; Walker, R.H.; Datry, T. Ecological research and management of intermittent rivers: An historical review and future directions. *Freshw. Biol.* **2016**, *61*, 1181–1199. [CrossRef]

42. Sarremejane, R.; England, J.; Sefton, C.E.M.; Parry, S.; Eastman, M.; Stubbington, R. Local and regional drivers influence how aquatic community diversity, resistance and resilience vary in response to drying. *Oikos* **2020**. [CrossRef]

43. CrowdWater. Available online: https://crowdwater.ch/ (accessed on 12 November 2020).

44. Tramblay, Y.; Rutkowska, A.; Sauquet, E.; Sefton, C.; Laaha, G.; Osuch, M.; Albuquerque, T.; Alves, M.H.; Banasik, K.; Beaufort, A.; et al. Trends in flow intermittence for European rivers. *Hydrol. Sci. J.* **2021**, *66*, 37–49. [CrossRef]

45. De Girolamo, A.M.; Bouraoui, F.; Baffagni, A.; Pappagallo, G.; Lo Porto, A. Hydrology under climate change in a temporary river system: Potential impact on water balance and flow regime. *River Res. Appl.* **2017**, *33*, 1219–1232. [CrossRef]

46. Pumo, D.; Caracciolo, D.; Viola, F.; Noto, L.V. Climate change effects on the hydrological regime of small non-perennial river basins. *Sci. Total Environ.* **2016**, *542*, 76–92. [CrossRef] [PubMed]