2021

Modelling climatic and temporal influences on boating traffic with relevance to digital camera monitoring of recreational fisheries

Ebenezer Afrifa-Yamoah
*Edith Cowan University*

Stephen M. Taylor

Ute A. Mueller
*Edith Cowan University*

Follow this and additional works at: [https://ro.ecu.edu.au/ecuworkspost2013](https://ro.ecu.edu.au/ecuworkspost2013)

Part of the Aquaculture and Fisheries Commons, and the Environmental Sciences Commons

10.1016/j.ocecoaman.2021.105947

Afrifa-Yamoah, E., Taylor, S. M., & Mueller, U. (2021). Modelling climatic and temporal influences on boating traffic with relevance to digital camera monitoring of recreational fisheries. *Ocean & Coastal Management, 215*, article 105947. [https://doi.org/10.1016/j.ocecoaman.2021.105947](https://doi.org/10.1016/j.ocecoaman.2021.105947)

This Journal Article is posted at Research Online. [https://ro.ecu.edu.au/ecuworkspost2013/11419](https://ro.ecu.edu.au/ecuworkspost2013/11419)
Modelling climatic and temporal influences on boating traffic with relevance to digital camera monitoring of recreational fisheries

Ebenezer Afrifa-Yamoah a, *, Stephen M. Taylor b, Ute Mueller a

a School of Science, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA, 6027, Australia
b Department of Primary Industries and Regional Development, Western Australian Fisheries and Marine Research Laboratories, PO Box 20, North Beach, WA, 6920, Australia

ARTICLE INFO

Keywords: Temporal analysis Digital camera monitoring data Distributional regression Bayesian regression modelling Recreational fisheries

ABSTRACT

Digital camera monitoring data on recreational boating traffic are often manually interpreted and the reading cost can be expensive. Typically, these data are used along with other periodic survey information and camera data between these surveys may not be read, creating gaps in the time series. We predicted recreational boating traffic during these ‘gap’ periods using historical camera data and covariates to complete the time series data. Predictive models were built in a Bayesian regression modelling framework to determine the daily distribution of recreational boating traffic at two ramps in Western Australia based on climatic variables (temperature, humidity, wind speed, direction and gust, and sea level pressure) and some temporal classifications (month and day type). Two observed year-long datasets of boating traffic were used, with a year-long gap between them. One set was used to build models, and the other set was used for validation purposes. Models were developed using leave-one-out cross-validation, and ensemble prediction. Fitted models explained 50% (95% credible interval (CI) of $R^2$: 0.40–0.58) and 62% (95% CI of $R^2$: 0.58–0.66) of the variabilities in the daily number of boat launches at the two ramps. Subsequently, using data for the preceding period where camera data were read, we imputed plausible estimates for the period between readings. Imputed values generally aligned well with the observed data, with some temporal biases at the bulk and upper tail of the distributions. The 95% credible intervals adequately reflected the observed data at both ramps. Data for the constructed periods depicted the general trends for the observed periods. Our results provide useful insights into using climatic factors to predict boating traffic to ‘fill in the gaps’ between survey years which could assist in the ongoing monitoring to promote sustainable management of recreational fisheries.

1. Introduction

To achieve sustainable recreational fisheries, it is important for resource managers and researchers to anticipate current and future management needs. Information on recreational boating traffic at ramps could be useful in assessing the impact of fishing and development of appropriate protection and maintenance policies (Beaudreau and Whitney, 2016). For instance, monitoring trends in boating activity at boat ramps can assist in interpreting long-term trends in fishing effort (Hartill et al., 2019). However, the diverse nature and lack of mandatory requirements for recreational fishers to report the details of their fishing trips make it difficult to understand behaviours associated with recreational boaters (Cabanellas-Reboredo et al., 2014). Data acquisition has been a major challenge for recreational fisheries in accounting for the participation and frequency of boating effort among recreational fishers (Cabanellas-Reboredo et al., 2014; Farr et al., 2014; Smallwood and Beckley, 2008). In many regions, boat-based fishing originates from public boat ramps where launches and retrievals can be monitored in surveys, which serves as an important component in informing regulatory policies. Both on-site (e.g. access point, roving survey) and off-site (e.g. phone diary, mail) surveys have been used to collect recreational fishing data (Lai et al., 2019; Ryan et al., 2019; Smallwood et al., 2012; Viega et al., 2010). However, because of the logistical challenges, time constraints and relatively high cost associated with these survey methods, data collections are not continuous, which is in contrast to most commercial fisheries, where there is typically a time series of catch and effort information.

There has been increasing use of digital cameras to monitor...
recreational boating traffic, and estimates of fishing effort can be generated from such data (Hartill et al., 2019; Taylor et al., 2018). A monitoring scheme provides opportunities to obtain reliable time series of recreational boating traffic. In practice, the purpose of digital camera monitoring usually is corroboration and validation and it is often used in conjunction with methods, including access-point and phone diary surveys (Lai et al., 2019; Ryan et al., 2017, 2019). As a result of management priorities and cost of manual data interpretation, data from digital camera monitoring between surveys may be recorded, but are often not read (Ryan et al., 2015, 2017). For periods that are read, monitoring may be interrupted due to power outages and other factors (e.g. vandalism, theft, unfavourable weather conditions) (Hartill et al., 2019). Analytical imputation methods have been developed to address such missing data problems (van Poorten et al., 2015; Hartill et al., 2016; Afrifa-Yamoah et al., 2020b). For instance, the methods described in Afrifa-Yamoah et al. (2020b) demonstrated satisfactory performance in imputing long duration outages (~1920 h). The value of information in Afrifa-Yamoah et al. (2020b) demonstrated satisfactory performance (e.g. vandalism, theft, unfavourable weather conditions) (Hartill et al., 2019).

Identifying appropriate physical drivers, such as climatic variables, along with predetermining recreational boaters’ behaviours can help describe the level of stochasticity exhibited in a modelling scheme (Arlinghaus et al., 2013; Bueno-Pardo et al., 2020; Soykan et al., 2014). Sea and air temperatures, waves, and tides can influence socio-economic values, and how they motivate human-mediated pressures on recreational fisheries, potentially driving temporal changes of fishing activities (Bryars et al., 2016). Wind speed and direction, type of day (weekday, weekend and public holidays) and time of day were among the significant predictors of boating traffic in Broome, Western Australia using remote camera monitoring data over a 12-month period (Desfosses and Beckley, 2015). Similarly, Widmer and Underwood (2004) found type of day and weather conditions to be significant drivers of the intensity of recreational boating traffic at Sydney Harbour, Australia. Predictors including month, sea surface temperature and height among other satellite-derived oceanographic variables and climate indices were identified as important for commercial fishing effort in the North Pacific Ocean (Soykan et al., 2014). In evaluating small lakes index management in British Columbia, Canada, Ashley et al. (2018) also found that type of day, day segments and hours of the day were significant predictors of boating traffic. Cabanellas-Reboredo et al. (2014) applied a hierarchical Bayesian model to determine the daily number of recreational boats and found that sea-bottom temperature negatively affected the number of boats, where greater number of boats were predicted at lower temperatures compared to higher temperatures in a recreational squid fishery in the Palma Bay (Balearic Islands).

In Western Australia, recreational boating is a popular outdoor activity and boating traffic is monitored using digital cameras at various fields of view including boat ramps, choke points and areas along the foreshore (Ryan et al., 2019). Reading of digital camera monitoring data coincides with 12-month state-wide surveys but for the period between surveys have not been read. Even though footage may exist, it was not fiscally possible to read more camera data for the purposes of this study, noting that the cost of analysing such footage would require a substantial investment (Afrifa-Yamoah et al., 2021). We sought to enrich the opportunities of predicting the distributions of recreational boating traffic between periodic surveys by describing their temporal distributions at two boat ramps in Western Australia which were located in different geographical regions with distinct climatic conditions. Climatic and temporal classification variables were used as potential drivers in a Bayesian Regression modelling framework to derive plausible numbers of boat launches for those periods where camera data were not read. Based on available data, we are proposing a possible solution for the data discontinuity problem of recreational fisheries monitoring. We have tested our approach for two boat ramps in different climate zones and demonstrated the fit of the formulated models using both in-sample and out-of-sample predictions and have provided model diagnostics to evaluate the approach. The evaluation comprised of several steps: first a cross-validation approach to evaluate the model fit, then a reconstruction of observed data based on the readings from the previous recording period and lastly the construction of values for the period between readings. In a practical sense, we have proposed a modelling framework which is adaptable to the level of stochasticity exhibited in recreational boating traffic and suggested possible auxiliary variables that contain some predictive information to support the model building process.

2. Methods

2.1. Data description and study area

This study analysed counts of recreational boat launches monitored using time lapse digital cameras, provided by the Department of Primary Industries and Regional Development and data on climatic variables including precipitation, temperature, humidity, wind speed, gust and direction and sea level pressure obtained from the Australian Government Bureau of Meteorology (BOM) (see Table 1 for a brief description).

Counts of recreational boat launches at designated ramps were recorded to the nearest minute (see Fig. 1). The use of recreational boat launches information was motivated by the fact that the covariates predetermine recreational boaters’ behaviours. Moreover, recreational boat launches are highly correlated with retrievals (which are more useful for survey design) in WA, because most boaters retrieve their boats at launching sites such as boat ramps (Afrifa-Yamoah, 2021). The type of vessel launched was recorded as either commercial, powerboat, jet-ski, kayak or other. The analysis in this paper focuses on powerboats as this is the common vessel type used for boat-based recreational fishing activities and the data were analysed at daily resolution (i.e. the number of launches over a 24-h period). A technical overview of the camera monitoring scheme can be found in Blight and Smallwood (2015).

Data collected between March 1, 2011 and February 29, 2012 were used for the model fitting and the predictive capabilities were verified via data collected between May 1, 2013 and April 30, 2014 at the Hillarys (Lat 31.822, Long 115.739) (see Fig. 2) and Broome (Lat 18.008, Long 122.208) boat ramps (see Fig. 3), respectively. The models were used to construct the data for the unobserved periods between March 01, 2012 to April 30, 2013 and May 01, 2014 to July 31, 2015 respectively. These locations were chosen firstly because their designated ramps are monitored and form part of ongoing research on recreational fishing (Steffe et al., 2017). The densities of boating traffic are distinct: Hillarys is one of the busiest boat ramps in the state while

| Table 1 | Study variables and their attributes. |
|---------|-------------------------------------|
| Variable         | Type      | Description                                      |
| Launches         | Count     | daily aggregated counts of powerboat retrievals.  |
| Precipitation    | Continuous| average daily amount of rainfall (mm)            |
| Temperature      | Continuous| average daily air temperature (°C)               |
| Humidity         | Continuous| average daily levels of humidity (%)             |
| Wind speed       | Continuous| average daily wind speed (km h⁻¹)                |
| Wind direction   | Categorical| average daily wind direction was categorized into northerly winds (N) ~ [315°, 45°], easterly winds (E) ~ [45°, 135°], southerly winds (S) ~ [135°, 225°] and westerly winds (W) ~ [225°, 315°] |
| Wind gust        | Continuous| average hourly maximum wind speed (km/h)         |
| Sea level pressure| Continuous| average daily sea level pressure (hPa)            |
| Day type         | Categorical| weekday or weekend/public holiday                 |
| Month            | Categorical| month of the year (January–December)             |
Broome experiences a comparatively lower level of traffic. Annual estimates of the frequency of recreational boating traffic in Hillarys is approximately five times more than that of Broome. A second motivation for their choice was their bioregional locations. There are four marine bioregions off the coast of Western Australia: North, Gascoyne, West and South Coasts (Ryan et al., 2015). Hillarys, located in the Perth metropolitan region, is on the West Coast region whereas Broome is on the North Coast. These locations are 2800 km apart (along the WA coast) and experience different climatic conditions, Hillarys is in a hot-summer Mediterranean climatic zone and Broome is in a hot semi-arid zone according to the Köppen-Geiger climate classification.

Characteristics of some of the climatic variables are presented in Fig. 4 to highlight some of the unique differences across locations. For example, with respect to precipitation, Broome observed sporadic heavy downpours for both training and test data in the summer months, whereas Hillarys observed more consistent levels of rain in the winter months. Also, generally higher temperatures were observed at Broome compared to Hillarys. These unique characteristics are expected to result in differing marginal contributions of the variables in predicting the number of powerboat launches across locations.

In both data components, there were incidences of missing data. Missing data in the climatic and camera data were imputed via the methods described in Afrifa-Yamoah et al. (2019, 2020a, 2020b) respectively.

2.2. Bayesian Regression modelling

Let $Y$ represent the total number of daily powerboat launches at the

---

Fig. 1. Study area showing the locations of the Hillarys (in hot-summer Mediterranean climatic zone), and Broome (in hot semi-arid climatic zone) boat ramps where digital camera data were analysed.
ramp with realizations $y_i, i = 1, \ldots, n$, where $n$ is the length of the series. We denote the covariates by $X$ with realizations $x_{1i}, \ldots, x_{pi}$, where $p$ denotes the number of covariates. We specify

$$y_i \sim D(g(\eta_i), \alpha),$$

where $D$ denotes the family of distributions for $Y$, $\eta_i$ represents a linear combination of covariates transformed by the inverse link function $g$ based on $D$ (which is dependent on the conjugate prior of $D$), and $\alpha$ describes any additional family specific parameters.

The Bayesian modelling framework allows greater flexibility in the choice of model distribution allowing better adaptability to the varying characteristics of outcome variable. Here, the total number of daily powerboat launches was assumed to follow a negative binomial distribution, as the estimates for the mean and variance were different across the ramps. The advantage of negative binomial distribution in modelling is the opportunity to adjust the variance estimates independently from

Fig. 2. Time series and distribution of the number of daily powerboat launches at Hillarys between March 1, 2011 and February 29, 2012 (training set) and May 1, 2013 and April 30, 2014 (test set).

Fig. 3. Time series and distribution of the number of daily powerboat launches at Broome March 1, 2011 and February 29, 2012 (training set) and May 1, 2013 and April 30, 2014 (test set).
the mean. The model setup was as follows:

\[ E(y_i) = \mu_i = g^{-1}(\eta_i), \]

(2)

where \( \eta_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}. \) The negative binomial regression model for an observation \( i \) is given by

\[ P(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(\gamma_i + \alpha^{-1})}{\Gamma(\gamma_i + 1)} \left( \frac{1}{1 + \mu_i} \right)^{\gamma_i} \left( \frac{\mu_i}{1 + \mu_i} \right)^\alpha. \]

(3)

We further assumed flat priors for the \( \beta \)s and shape parameter \( \alpha \sim \Gamma(k,k) \), where \( k \in \mathbb{Q} \). The model choice was informed by the fact that the variance of our dataset was greater than the mean.

Generally, the Occam’s razor principle makes variable selection an important objective in explaining data with parsimonious models. However, there may be instances where saturated models may be preferred especially when the relationship between response and predictor variables is weak and/or when underlying processes are not well understood. In such cases, the collective contribution of all covariates, irrespective of the statistical significance of their coefficients would be useful in describing their relationship. Bayesian regression is generally suitable for assessing distributional relationships between response and predictors. It is important to note that there is a weak relationship between the climatic variables considered and recreational boating activity, and that means variable selection would lead to less efficient models (Afrifa-Yamoah et al., 2019, 2020b). Here it is more important to make inferences from these relationships in order to establish their potential for describing the distributional behaviour of recreational boating activities. In this modelling framework, probability distributions were used to formulate the linear regression instead of point estimates used in the frequentist settings. For model parameters \( \beta \) and \( \alpha \) conditioned on the data, the posterior probabilities were obtained as;

\[ P(\beta|y,X) = \frac{P(y|\beta,X)P(\beta|X)}{P(y|X)} \]

(4)

\[ P(\alpha|y,X) = \frac{P(y|\alpha,X)P(\alpha|X)}{P(y|X)} \]

(5)

where \( P(y|\beta,X) \) and \( P(y|\alpha,X) \) are likelihoods of the data with respect to \( \beta \) and \( \alpha \), \( P(\beta|X) \) and \( P(\alpha|X) \) are prior distributions of the model parameters and \( P(y|X) \) is the normalizing constant. The posterior distributions of the model parameters were approximated using a variant of the Markov Chain Monte Carlo sampling algorithm known as No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014; Bürkner, 2017). NUTS proposed is an extension to the Hamilton Monte Carlo algorithm, which eliminates the requirement for a user-defined number of desired steps, and even provides a method based on primal-dual averaging for adjusting the step size parameter while in process. The sampler spans the phase space of the target parameter space using recursive algorithm and automatically stops sampling when the algorithm starts to retrace it steps.

2.3. Model specifications and implementation

All models were implemented in R software (R Core Team, 2017), using packages such as ‘brms (v 2.11.1)’ for fitting the models (Bürkner, 2017), ‘rstanarm (v 2.19.2)’ serving as an interface for running models in Stan for full Bayesian inference (Goodrich et al., 2018; Brilleman et al., 2018), ‘Rtools’ for the additional C++ compiler required to compile Stan codes created internally by ‘brms’ (Bürkner, 2017; R Core Team, 2017), ‘sjstats (v 0.17.9)’ for setting auto priors to model parameters (Lüdecke, 2020), and ‘coda (v 0.19-3)’ for visualization diagnostic tool for model convergence (Plummer et al., 2006).
In this study, models were fitted using 4 chains, each with 10,000 iterations of which the first 1000 were burn-in to tune the sampler. Flat prior distributions were assumed on the population-level effect parameters, as we were particularly interested in assessing the posterior effect of the covariates on the response. Gamma priors were set to the shape parameter using the default setting of ‘ijstats’. The distribution of the response variable was set using ‘family = negbinomial’. The sampling behaviour in Stan was adjusted with ‘control = list (adapt_delta = 0.97)’. This was done to improve the efficiency of sampling convergence, thus decreasing the number of divergent transitions that can affect the validity of our posterior samples (Bürkner, 2017).

2.4. Model evaluation

Effective sample size (ESS) is a heuristic method for checking convergence among samplers in approximating $P(X|\beta, \alpha)P(\beta, \alpha)$. It represents the number of independent samples from the posterior distribution that would be expected to yield the same standard error of the posterior mean as is obtained from the dependent samples (doing local stepping around posterior space) returned by an MCMC algorithm (Bürkner, 2017). In other words, it measures the cost of dependent sampling. For the same number of samples, the informational value of independent samples is more valuable than that of dependent samples. Thus, ESS for a predictor depicts its informational value in calculating the posterior mean.

Potential scale reduction statistics ($\hat{R}$), also known as the Gelman-Rubin statistic, is the ratio of the variance of a parameter when the posterior mean as is obtained from the dependent samples (doing local stepping around posterior space) returned by an MCMC algorithm (Gelman and Rubin, 1992). It measures the extent to which chains are reaching different conclusions. The further the value of this statistic from 1, the worse the convergence of the chains.

The coefficient of determination ($R^2$) expresses the amount of variation in the response variable that the model explains. The value also indicates how well the model predicts future outcomes of the response variable. This information generally reflects the strength of the relationships between the predictor variables and the response variable.

A comparison of the predicted values and their 95% credible intervals for an ensemble of draws from the model’s posterior distribution at the observed values of the covariates to the observed response data was performed using various visualization tools. The credible interval is the Bayesian analogue of the frequentist confidence interval and represents the interval within which an unobserved parameter value falls with a particular probability, here taken to be 95%. The model fit was further examined using the posterior predictive plot which compares the density of the original response values with a defined number of draws (we evaluated up to 1000 draws) from the posterior distribution of the model. A good fit is achieved when the density of the samples from the model’s posterior distribution at the observed values of the covariates aligns with that of the observed response data. Also, diagnostic plots using the leave-one-out cross-validated probability integral transform (LOO-PIT) technique were used to assess the fit of the models (Gebry et al., 2019). PIT is a tool for assessing the probabilistic calibration of the predictive distribution (see Gneiting et al., 2007), which will follow a uniform distribution if the predictive distribution is correct. The

| Table 2 | Summary statistics for the model fit and posterior distributions of the population level effects and family specific parameter. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Population level effects | Hillarys | | | | | | | | | |
| Intercept | Estimate | Est. error | 95% CI | u-95% CI | Rhat | Bulk ESS | Tail ESS | | | |
| Climatic variables | | | | | | | | | | |
| Temperature | -0.02 | 0.02 | -0.07 0.02 | 1.00 2225 | 5129 | 0.08 0.02 | 0.03 0.12 | 1.00 677 | 1781 |
| Humidity | -0.01 | 0.01 | -0.02 0.00 | 1.00 4302 | 9630 | 0.00 0.00 | -0.00 0.01 | 1.00 1441 | 3442 |
| Wind speed | 0.01 | 0.06 | -0.12 0.13 | 1.00 3703 | 6912 | -0.06 0.08 | -0.23 0.11 | 1.00 1087 | 2164 |
| Sea level pressure | -0.01 | 0.01 | -0.04 0.01 | 1.00 2439 | 4823 | 0.01 0.02 | -0.03 0.05 | 1.01 483 | 1001 |
| Precipitation | 0.00 | 0.02 | -0.04 0.04 | 1.00 6312 | 11,429 | -0.05 0.02 | -0.09 -0.01 | 1.00 904 | 1821 |
| Wind gust | 0.00 | 0.04 | -0.09 0.65 | 1.00 3612 | 6828 | -0.04 0.06 | -0.16 0.07 | 1.00 1083 | 2213 |
| Wind direction | | | | | | | | | | |
| Easterly winds | -0.08 | 0.10 | -0.28 0.13 | 1.00 5428 | 11,237 | 0.15 0.11 | -0.06 0.37 | 1.00 1300 | 2716 |
| Westerly winds | -0.99 | 0.18 | -1.34 -0.64 | 1.00 5923 | 11,980 | 0.54 0.11 | 0.33 0.76 | 1.00 1165 | 2625 |
| Northerly winds | Ref | | | | | | | | | |
| Southernly winds | -0.04 | 0.14 | -0.30 0.32 | 1.00 5755 | 11,589 | 0.65 0.12 | 0.42 0.89 | 1.00 1404 | 2730 |
| Day type | | | | | | | | | | |
| Weekday | Ref | | | | | | | | | |
| Weekned | Ref | | | | | | | | | |
| Month | Ref | | | | | | | | | |
| January | 0.10 | 0.22 | -0.33 0.54 | 1.00 2063 | 7159 | -0.18 0.15 | -0.47 0.11 | 1.01 800 | 2065 |
| March | 0.04 | 0.21 | -0.38 0.46 | 1.00 2673 | 8201 | -0.05 0.15 | -0.36 0.25 | 1.00 662 | 1288 |
| April | 0.19 | 0.25 | -0.28 -0.68 | 1.00 1544 | 4882 | 0.19 0.18 | -0.17 0.54 | 1.00 410 | 755 |
| May | -0.21 | 0.21 | -0.38 0.46 | 1.00 2673 | 8201 | 0.02 0.23 | -0.44 0.49 | 1.00 347 | 531 |
| June | -0.29 | 0.24 | -0.76 0.18 | 1.00 1558 | 5615 | 0.77 0.27 | 0.22 1.31 | 1.00 335 | 578 |
| July | -0.56 | 0.28 | -1.11 -0.01 | 1.00 1395 | 4943 | 0.97 0.25 | 0.48 1.46 | 1.00 326 | 467 |
| August | -0.41 | 0.31 | -1.01 0.19 | 1.00 1325 | 4415 | 0.69 0.24 | 0.22 1.17 | 1.00 330 | 483 |
| September | -0.64 | 0.31 | -1.24 -0.03 | 1.00 1408 | 4394 | 0.38 0.22 | -0.04 0.81 | 1.01 328 | 487 |
| October | -0.52 | 0.31 | -1.11 0.08 | 1.00 1395 | 4361 | 0.50 0.17 | 0.17 0.84 | 1.00 409 | 873 |
| November | -0.25 | 0.30 | -0.83 0.35 | 1.00 1280 | 4190 | 0.17 0.16 | -0.15 0.48 | 1.00 500 | 1089 |
| December | 0.67 | 0.26 | 0.17 1.18 | 1.00 1447 | 4621 | 0.18 0.16 | -0.15 0.50 | 1.00 821 | 1777 |
| Family specific parameter | | | | | | | | | | |
| Shape | 1.68 | 0.13 | 1.44 1.95 | 1.00 7669 | 12,270 | 4.71 0.51 | 3.80 5.79 | 1.00 882 | 3776 |
| $R^2$ | 0.50 | 0.05 | 0.40 0.58 | 0.62 0.02 | 0.58 0.66 |

(\text{l-95% CI} – lower bound of credible interval; \text{u-95% CI} – upper bound of credible interval; \text{ESS} – Effective Sample Size).
leave-one-out cross-validation procedure was carried out 300 times.

3. Results

3.1. Model fit and cross-validation

For the Hillarys ramp, the coefficient of determination ($R^2$) for the model fit was 0.50 (95% credible interval (CI): 0.40–0.58). The model achieved convergence, as indicated by the potential scale reduction statistics for the variables (see Table 2). Based on ESS scores, wind direction, precipitation and humidity were the most influential predictors for boating activities. These variables were often sampled in predicting the posterior distribution of the number of daily powerboat launches. Among the temporal classifications, weekend was the most influential variable (see Table 2). For the training data, the distribution of the predicted data overlapped with the distribution of the observed data, with some temporal biases at the bulk and upper tail and a moderate similarity index (see Figure SM1 (A-B) of supplementary material). Predicted data aligned well with the observed data, with the 95% credible intervals of the predicted data containing the observed data (Figure SM1 C of supplementary material). Marginal effects of the predictors are shown in Figure SM2 of supplementary material. For instance, temperature, humidity, sea level pressure showed a negative effect on the number of launches (that is, more boating traffic was recorded at lower values, implying that a larger number of boat counts were observed at lower humidity). Also, fewer boating activities were recorded during westerly winds (see Figure SM2 of supplementary material). The PP-plot for the leave-one-out cross validation indicated consistency between the predicted and observed data (see Fig. 5-A1). Also, the distributions of 1000 ensemble draws of predicted data from the formulated posterior distribution overlapped with the distribution of the observed data (see Fig. 5-A2).

For the Broome ramp, the coefficient of determination ($R^2$) for the model fit was 0.62 (95% CI: 0.58–0.66). Humidity and wind direction were among the most influential predictors for boating activities in this region, with weekend often sampled among the temporal classification based on the ESS scores (see Table 2). Model predictions for the training data, showed least temporal bias at the bulk, as depicted by the plot of predicted against observed data (see Figure SM5 A-B). The marginal effects of the predictors indicated that humidity and sea level pressure showed positive effect on the number of launches (larger number of boat counts were observed at higher values). Precipitation showed a negative effect on the number of launches (more boating traffic was recorded at lower values). Southerly winds were comparatively more favourable for boating activities, whilst the least activities were recorded during easterly winds (Figure SM6 of the supplementary material). PP-plot for the leave-one-out cross validation revealed adequate alignment between the predicted and observed data (see Fig. 5-B1). The distributions of 1000 ensemble draws of predicted data from the posterior distribution overlapped with the distribution of the observed data (see Fig. 5-B2).

3.2. Reconstructing observed data

In assessing the predictive power of the models, we fitted them to complete sets of covariates and reconstructed the daily number of powerboat launches for a year-long period from May 1, 2013 to April 30, 2014 for the two ramps. Figs. 6–7 (A-C) present visual summaries comparing predicted values (and 95% credible intervals) and observed data for Hillarys and Broome ramps respectively. Generally, forecast data aligned well to the observed data at both ramps. The similarity index between the forecast and observed data was moderately good, with significant correlation coefficients of 0.68 and 0.71 for Hillarys and Broome ramps, respectively. The models adequately captured the dynamics in both data, by observing the time series plots for the observed data.
and forecast data (including the 95% credible intervals). However, the density of the predictions for the Hillarys ramp was more peaked than that of the true data, although the two distributions were reasonably overlapped. For the Broome ramp, there seems to be a slight shift in the mode, but overall, the shapes were similar.

3.3. Imputation estimates for unobserved periods

Using the data for the preceding period where camera data were read, Bayesian regression models were used to construct data for the periods where camera data were not read (see Fig. 8). Overall, the time series trends for the unobserved periods aligned with the observed periods. In contrast to the forecasts at the Hillarys ramp, the forecasted values of powerboat launches for periods where camera data was not read were relatively higher than the observed values for the Broome ramp. However, the 95% credible intervals suggest that data constructed were plausible given the range of the observed data.

4. Discussion

Understanding the usage patterns at boat ramps is useful, because knowledge of the temporal distribution of recreational boating traffic promotes effective survey and sampling designs, and support management practices. We have explored the effectiveness of incorporating climatic and some temporal classifications as covariates in models that describe the temporal distribution of boat launches from two boat ramps in Western Australia. The count of daily boat launches recorded via digital camera monitoring was used as the dependent variable to build the models in a Bayesian regression modelling framework. The results showed that the covariates contained adequate predictive abilities to reveal patterns and trends in recreational boating traffic. The Bayesian regression modelling scheme has the potential to uncover the relationships between these covariates and daily recreational boating traffic. The model has the ability to describe meaningful characteristics for building time series data and to corroborative estimates of fishing effort obtained from other surveys that are ongoing in Western Australia. Advancements in technology and the application of statistical methods have enabled reliable forecasts that predict the temporal distribution of recreational boating effort. This could assist in the management of fisheries and urban planning relating to ramp usage.

The results reinforce three things: that predictors contributed differently with respect to boat ramp, prior distributions and modelling specifications may not be uniform across ramps and forecasting long horizons (such as 12-months prediction) would require much longer time series data. Regarding the contributions of climatic predictors, while wind direction and precipitation were the most important drivers for the metropolitan Hillarys ramp in the south, humidity and wind direction were for the northern Broome ramp. Additionally, from the marginal effect plots (see Fig SM2 and SM6 supplementary materials) whilst more boating activities were recorded in high temperature and humidity conditions in the semi-arid zone (Broome), such conditions led to lesser boating activities in the hot-summer Mediterranean (Hillarys). Soykan et al. (2014) found similar patterns of differing contributions of their predictors with respect to location. In effect, potential predictors for recreational boating effort may be location-driven, highlighting the need for extended exploration of additional predictors unique to locations. The results for Broome confirm earlier findings of Desfosses and Beckley (2015) where wind direction among other variables was...
identified to significantly influence recreational boating activity. Comparatively, the imputed values for Hillarys ramp aligned better to the observed data than those from the Broome ramp. The relatively high imputed values for the Broome ramp may suggest that the choice of an informative prior distribution on some or all of the covariates may improve the fit of the model. Also, it is known that the relative correlations between the predictors and response variable are very weak (Afrifa-Yamoah et al., 2020b), and thus additional known information may serve as better support for the model estimation process. Additionally, the length of the training data was shorter than the duration of the imputed time series and possibly may account for inadequate modelling memory and more erratic predictions especially for locations where variables’ relationships were much weaker. The models presented here did not account for the yearly effect on boating effort because of the nature of the observed data (i.e., annual with gaps between years). The results observed suggest that model estimation and predictions would be enhanced with a longer training dataset.

In understanding the complexity of recreational boaters’ behaviour and decision making on participation and frequency of boating, previous studies had used different covariates to estimate recreational boating effort from digital camera monitoring Farr et al. (2014); Lancaster et al. (2017); Askey et al., (2018). For instance, social and economic indicators of individuals including marital status, household income, employment status and type of job have been identified as potential predictors of recreational fishing participation and frequency (Farr et al., 2014). Their study confirmed that where people fish and why they fish are strongly tied to personal reasons. Hunt (2005) also identified environmental quality as an important determinant of boating effort. The wide range of predictors for recreational boating effort make modelling such phenomenon a challenge, notably, a search for parsimonious model may not be feasible. The modelling approach adopted in this study is more apt as it draws inference from the collective contribution of potential predictors to describe the posterior distribution of recreational boating activity. This would provide fisheries managers and researchers the opportunity to have greater anticipation of trends in recreational boating effort.

4.1. Strength and limitations

Bayesian modelling provides a statistical modelling approach with greater flexibility and power to uncover complex relationships between variables. This modelling framework can effectively handle collinearity, interactions between variables and non-linear relationships (characteristics which are typical among climatic variables) between covariates and response variables. For example, non-linear relationships could be easily resolved by the incorporation of prior knowledge about parameters in model building (Bürkner, 2018). Another major advantage is that the modelling framework used in this study allows for the prediction of all response parameters at the same time, often referred to as a distributional model or a model for location, and shape (Rigby and Stasinopoulos, 2005; Bürkner, 2018). Also, the modelling framework can derive probability information for all model inputs. The No-U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014) ensured that the parameter space of the model was well sampled and converged much more quickly than other Markov-Chain Monte-Carlo (MCMC) algorithms (Betancourt 2017). There may be issues raised of increased computing power,
time-consuming and error prone process associated with this modelling framework, but software packages such as Stan (Stan Development Team 2017) provide a useful platform upon which other simple, formula-like modelling syntax packages such as brms (Bürkner, 2017) are built on to facilitate easy model building.

5. Conclusion

Digital camera monitoring provides a unique opportunity to obtaining accurate data on recreational boat traffic for a field of view, but comes at a substantial cost of manual data interpretation in the long run, especially if multiple fields of view are monitored. We have demonstrated the ability to use climatic and temporal classification to satisfactorily predict the temporal distribution of boating traffic for periods between surveys, for instance 12-month state-wide surveys in WA which are carried out triennially. Also, the relatively cheaper cost involved in obtaining information on the predictors used in the models makes this study easily applicable. The proposed modelling framework can impute plausible data for long missing gaps to support the monitoring of trends in recreational boating traffic. This would provide innovative and cost saving opportunities of obtaining continuous estimates on boating traffic. Additionally, it provides support for dealing with short-term missing observation issues in digital camera monitoring data. The modelling framework advocated is easily implemented in statistical software to support fisheries scientists, as well as researchers in allied fields, such as ecology, transportation and tourism, to understand complex relationships between recreational boating and the environment.

Declaration of competing interest

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

Acknowledgements

This study was funded and supported by the Government of Western Australia Department of Primary Industries and Regional Development (DPIRD) and Edith Cowan University. The authors are grateful to the Australian Government Bureau of Meteorology for providing the data for the study. The authors express their sincere gratitude to the staff of DPIRD who spent much time reading the camera data. The authors would like to thank Stuart Blight and Cameron Desfosses for maintaining the network of cameras. A special thanks to Ainslie Denham, Karina Ryan, and Claire Smallwood for their time and input during the internal review process by DPIRD.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ocecoaman.2021.105947.

References

Afrifa-Yamoah, E., 2021. Imputation, Modelling and Optimal Sampling Design for Digital Camera Data in Recreational Fisheries Monitoring. Doctoral dissertation, Edith Cowan University, Perth, Australia. Retrieved from. https://ro.ecu.edu.au/theses/2387.
Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M., Fisher, A.J., 2019. Fixed versus random effects models: an application in building imputation models for missing data in remote camera surveys. In: The Proceedings of the 34th International Workshop on Statistical Modelling (IWSM) (Volume II). Guimarães, Portugal, pp. 7–12. July 2019.
Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M., Fisher, A.J., 2020a. Missing data imputation of high-resolution temporal climate time series data. Meteorol. Appl. 1–18. https://doi.org/10.1002/met.1873.
Afrifa-Yamoah, E., Taylor, S.M., Fisher, A.J., Mueller, U.A., 2020b. Imputation of missing data from time-lapse cameras used in recreational fishing surveys. ICES (Int. Counc. Explor. Sea) J. Mar. Sci. https://doi.org/10.1093/icesjms/fsaa180.
