Heliyon 8 (2022) e08774

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Modeling digital camera monitoring count data with intermittent zeros for short-term prediction

E. Afrifa-Yamoah *, U.A. Mueller

School of Science, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA 6027, Australia

ARTICLE INFO

Keywords:
Digital camera monitoring
Time series modelling
Intermittent data
Integer-valued models
Imputation
Forecasting

ABSTRACT

Digital camera monitoring has revolutionised survey designs in many fields, as an important source of information. The extended sampling coverage offered by this monitoring scheme makes it preferable compared to other traditional methods of survey. However, data obtained from digital camera monitoring are often highly variable, and characterized by sparse periods of zero counts, interspersed with missing observations due to outages. In practice, missing data of relatively shorter duration are mostly observed and are often imputed using interpolation techniques, ignoring long-term trends leading to inherent estimation biases. In this study, we investigated time series forecasting methods that adequately handle intermittency and produced plausible estimates for imputation and forecasting purposes. The study utilised a yearlong digital camera monitoring data set of hourly counts of powerboat launches at three boat ramps in Western Australia. Several time series forecasting methods were evaluated and the accuracies of their point estimates of forecasts for various lead times in hours of up to one week were assessed using cross-validation techniques. Intermittent demand forecasting techniques, including Croston's method and Syntetos-Boylan Approximation (SBA) models, and count data forecasting methods including autoregressive conditional Poisson (ACP) models, integer-valued moving average (INMA) models, and integer-valued autoregressive (INAR) models were evaluated. ACP and INAR models performed better than intermittent demand forecasting techniques for short forecast horizons and provided some evidence of their sufficiency in predicting the dynamics in recreational boating activities. This result established that, in as much as intermittency may be a key feature for a given dataset, it should not override the systemic characteristics of data in the application of forecasting techniques. Our results provide plausible estimates for short-term missing data and forecasts for monitoring events, with applications in supporting proper tracking of usage of facilities, guiding resource allocations and providing insightful perspectives for management decisions.

1. Introduction

The use of digital cameras for events' monitoring is widespread and has become an integral source of data in many fields of application. For example, digital cameras are being used to monitor boating traffic at many recreational fishing sites worldwide for complementary and corroborative purposes in survey designs (Hartill et al., 2020; Lancaster et al., 2017; Steffe et al., 2017; van Poorten and Brydle, 2018). This monitoring scheme provides evidence of events unfolding in real time, in various formats including videos, and still images. For statistical analytical purposes, these data formats are additionally processed, referred to as data interpretation. Count data are common outputs from the data interpretation process. High-volume count data at sufficiently fine granularity are being observed to enable studies of both short- and long-term trends. Such data can reveal or obscure the various time series components at different data aggregation levels (Petropoulos and Kourentzes, 2014; Kourentzes et al., 2018). Particularly, disaggregated time series data of finer resolution will adequately reveal possible seasonality patterns in the data while components such as level, trends and cycles will be more clearly exhibited in aggregated time series (Petropoulos and Kourentzes, 2014). These components are vital in both short- and long-term operational forecasts. Additionally, data generated in real time are characterised by infrequent non-zero counts, often of variable size. The characteristics and structure of such data are comparable to those of intermittent demand data which are characterized by slow-moving items (Teunter and Duncan, 2009). Such data are often highly variable and difficult to forecast (Syntetos et al. 2015).

In retail, supply chain, sales and inventory control systems, intermittent demand data are common (Croston, 1972; Snyder et al., 2012;

* Corresponding author.
E-mail address: e.afrifayamoah@ecu.edu.au (E. Afrifa-Yamoah).

https://doi.org/10.1016/j.heliyon.2022.e08774
Received 18 August 2021; Received in revised form 17 October 2021; Accepted 12 January 2022

Kolassa, 2016). The exponential smoothing framework based on the normal distribution has been found to perform well in cases where larger counts, with less variability are observed, however for intermittent demand data, distributions tailored to count data must be considered (Snyder et al., 2012; Petropoulos and Kourentzes, 2014). For instance, there is an upward bias in the forecast directly after a non-zero count when the exponential smoothing technique is applied on intermittent demand data. Classical autoregressive integrated moving average (ARIMA) models allow negative integer values and are not suitable for application on intermittent demand data. Discrete autoregressive moving average (DARMA) models could be presented as an alternative modelling option because they are constrained to be non-negative, but they are limited to stationary intermittent demand data. However, intermittent demand data are often non-stationary (Syntetos and Boylan, 2005; Syntetos et al., 2015; Kolassa, 2016). Integer-valued moving average (INMA) and integer-valued autoregressive moving average (INARMA) models might be viable alternatives for modelling non-stationary series such as the number of transactions in stocks and insurance applications (Aleksandrov and Weiß, 2020; Brä nn ä s and Shahiduzzaman Quoreshi, 2010). The INARMA family uses a probabilistic operator called binomial thinning, developed by Steutel and van Harn (1979), as an alternative to the scalar multiplication used in the ARMA family, to avoid non-integer value forecasts (Weiß, 2008). Bourguignon et al. (2015) found that the point predictions of seasonal and non-seasonal integer-valued models were close to each other and they both seemed to provide reasonable estimates of the h-step ahead observations. In stationary count data of high resolution, seasonality is characterised by serial dependence. Weiß (2008) investigated serial dependence structures of stationary count processes and found the integer-valued models INAR (p), INMA (q) and INARMA (p, q) to be useful modelling techniques for such processes.

The work of Croston (1972) with a correction by Rao (1973) is often cited in the analysis of intermittent demand data. The method is based on an exponential smoothing scheme for updating the expected time gaps between non-zero demands and the expected demand for a period with the assumption that both events are statistically independent. Syntetos and Boylan (2005) found this method to produce biased forecasts. Subsequently, the authors proposed a combination of the Croston method with a Bayesian approach known as Syntetos-Boylan approximation (SBA), which estimates the probability of non-zero demands instead of interval size, using a Taylor series expansion. Other methods such as the Autoregressive Conditional Poisson model (ACP, Heinen (2003)), have been found to be adequate in modelling relatively rare events. This model deals explicitly with the discreteness and additional time series properties of the data, which if neglected could lead to a higher likelihood of misspecification. For instance, ACP efficiently handles autocorrelation and over dispersed time series count data (see Heinen 2003).

With the growing usage of digital camera monitoring in applications, it is imperative to provide statistical modeling support for data management and analysis. For instance, data obtained often have missing observations due to camera outages (Afrifa-Yamoah et al., 2020a). In practice, missing data of relatively shorter duration are mostly recorded and are often imputed using interpolation techniques (Lepot et al., 2017; Ryan et al., 2017; Wise and Fletcher, 2013), ignoring long-term trends leading to inherent estimation biases. More robust statistical modeling techniques for short-term predictions are required. This study evaluated time series forecasting methods, usually applied to intermittent demand and count data, to assess their suitability for data obtained from digital camera monitoring. A set of year-long hourly count data on recreational powerboat launch activities were obtained from digital camera monitoring of three boat ramps in Western Australia. Five forecasting methods were evaluated and the accuracy of their point estimates of forecasts for lead times of 12, 24, 48 and 168 h were compared using cross-validation techniques. Data were split into training and test data based on the forecast horizons, with the length of the test data given by the length of the forecasting horizon. In what follows, we first provide a data description, then provide background on the methods to be evaluated. We then present and discuss the cross-validation results and lastly provide concluding comments.

2. Methods

2.1. Data description and study area

Time lapse cameras have been installed together with other hardware at fields of view to record, store and transmit data on boat launches and retrievals to the Department of Primary Industries and Regional Development (DPIRD) in Western Australia. Counts and times of incidents of powerboat launches were recorded at three boat ramps (see Figure 1). Data from Leeuwin (in the West Coast bioregion) were collected between 1 March 2011 and 29 February 2012, whilst Broome (in the North Coast bioregion) and Denham (in the Gascoyne Coast bioregion) were observed between 1 May 2013 and 29 April 2014. There were 43.3% and 8.1% missing observations in the datasets from Broome and Denham respectively. Missing observations in the data were imputed at hourly resolution using the techniques described in Afrifa-Yamoah et al. (2019, 2020a, 2020b). More specifically, structural time series models with Kalman filters were used to impute missing data in climatic covariates, including temperature, humidity, precipitation, and wind speed (Afrifa-Yamoah et al., 2020b), which were then used in building an imputation model for the digital camera data. A generalized linear mixed effect model built on a fully conditional specification multiple imputation framework was used to impute the missing data in the digital camera data sets (Afrifa-Yamoah et al., 2019; 2020a). Subsequently, the data were aggregated at hourly resolution for analysis. The proportion of zeros was 47.0% and 75.9% for Broome and Denham respectively. The Leeuwin data set had 54.7% zero entries with no incident of missing data.

Figures 2, 3, and 4 present the data and some important features that are informative for modelling. The month-long time series insert in Figures 2, 3, and 4 illustrates the intermittency that exists in the dataset. The behaviours of the autocorrelation functions (ACF) and partial autocorrelation functions (PACF) (see Figures 2, 3, and 4) indicate that a non-negative integer-valued time series is stationary with serial dependence, which largely obscures the seasonality component of our data. The irregular serial dependence in the dataset stems from the nature of the boat launch activity cycle, in which launches in the early hours of the day dominate. Weiß (2008) found that non-seasonal integer valued models provide an appropriate fit for such data. In the case of Poisson autoregression models, the feedback mechanism terms in the models are observation-driven, implying that serial dependence is explained by past observations (Weiß, 2018).

2.2. Modelling techniques

This section presents simplified versions of the time series modelling techniques evaluated. More detailed information about the respective time series modelling techniques presented below can be found in Heinen (2003), Croston (1972), Syntetos and Boylan (2005), Al-Osh and Alzaid (1988), and Bourguignon et al. (2015). Also note that we are interested in comparing the point estimates from the models to true observed values, therefore uncertainties are not considered in evaluating how well the fit of the models aligns to observed values (although estimation information has been provided for completeness). The modelling techniques considered are special tailored for modelling integer-valued time series data, thus produced forecasts that are bounded below by 0 and are therefore generally coherent with the data (Freeland and McCabe, 2004).

2.2.1. Autoregressive Conditional Poisson (ACP)

Heinen (2003) proposed the autoregressive conditional Poisson (ACP) modelling framework to analyse count time series data exhibiting autoregressive properties. The model ACP(p, q) assumes that the counts are generated by a Poisson distribution

$$Y_t|Y_{t-1} \sim P(y_t, \mu_t) \tag{1}$$

with an autoregressive conditional intensity as in the Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998), such that

$$E(Y_t|Y_{t-1}) = \mu_t = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i \mu_{t-i} \tag{2}$$

with $\omega > 0$ and $\alpha_1, \cdots \alpha_p, \beta_1, \ldots, \beta_q \geq 0$. Provided that $\sum_{i=0}^{\max(p,\ q)} (\alpha_i + \beta_i) < 1$, a stationary solution of the unconditional mean is obtained as

$$E(Y_t) = \mu = \frac{\omega}{1 - \sum_{i=0}^{\max(p,\ q)} (\alpha_i + \beta_i)} \tag{3}$$

The unconditional variance is given by

$$Var(Y_t) = \frac{\mu\left(1 - \sum_{i=0}^{\max(p,\ q)} (\alpha_i + \beta_i)^2 + \sum_{i=0}^{p} \alpha_i\right)}{1 - \sum_{i=0}^{\max(p,\ q)} (\alpha_i + \beta_i)^2} > \mu \tag{4}$$

for positive parameters $\alpha_i, i = 1, \ldots, p$, $\beta_i, i = 1, \ldots, q$'s, where p and q are the orders of the autoregressive and feedback mechanism term (representing the hidden conditional means to enhance the memory of the model) components respectively (Weiß, 2018, Chapter 4).

Note that all the vectors of time-dependent covariates that influence the evolution of Eq. (1) is composed by the unobserved process μ_t. Therefore, we modelled the dynamics of the process using μ_t, which is a function of all information of Y up to $t-1$ and of the unknown regression parameters in Eq. (2). Using likelihood-based inference for linear Poisson autoregression, $\{\mu_t\}$ is regressed on past values of the observed process and past values of $\{\mu_t\}$ itself (see Heinen, 2003; Fokianos et al., 2009). Note that Eq. (2) is closely related to the GARCH (p, q) process (Bollerslev, 1986), since the Poisson distribution assumes equal value for the mean and variance. In the literature, the ACP modelling framework is sometimes referred to as INGARCH (Weiß, 2018).

The conditional median has been suggested for its robustness in analysing time series with heavy tails, as well as adherence to data coherency (Freeland and McCabe, 2004). However, the conditional median forecast can be misleading and may not be very informative, in that $P(X = 0) = 1 - P(X = 1) = 0.50$ has the same median (0) as $P(X = 0) =$

Figure 1. Study area showing the locations of the Leeuwin (in hot-summer Mediterranean climatic zone), Denham (in hot dessert) and Broome (in hot semi-arid climatic zone) boat ramps where digital camera data were analysed (Afrifa-Yamoah et al., 2021).

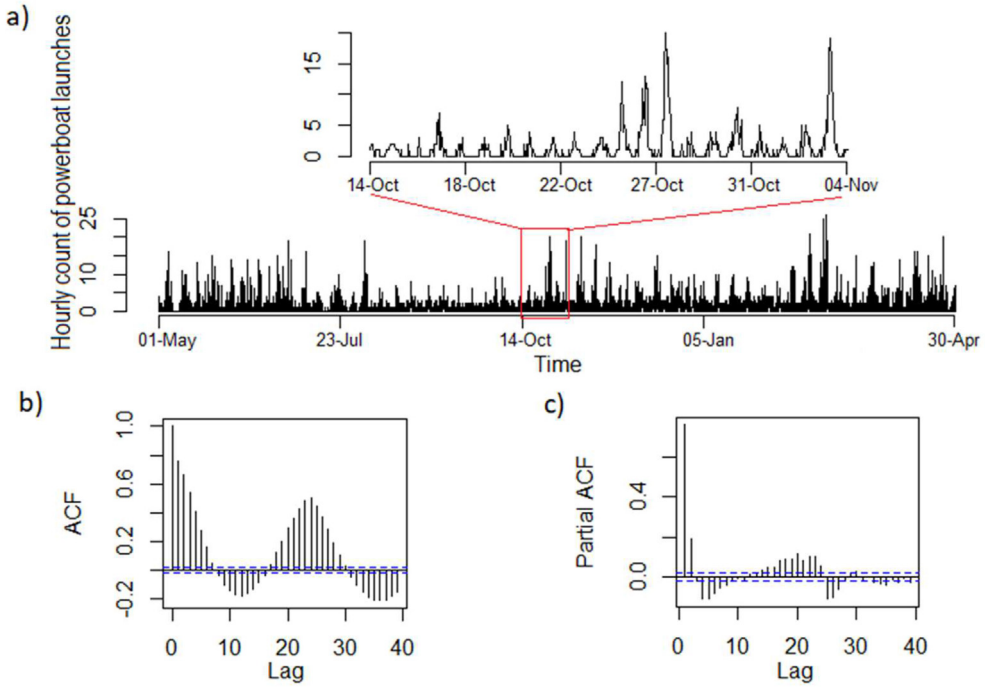

Figure 2. a) Time series plots b) ACF and c) PACF of the count of powerboat launches from digital camera monitoring observed at Broome boat ramp between 1 May 2013 and 30 April 2014. The inserted series details the behaviour of the process within the window outlined in red.

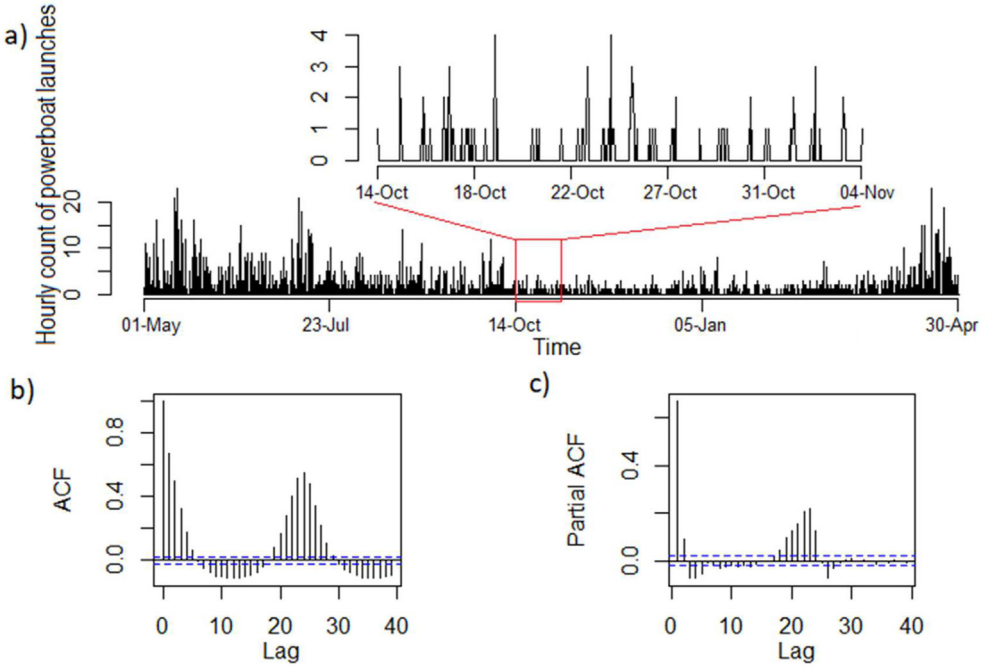

Figure 3. a) Time series plots b) ACF and c) PACF of the count of powerboat launches from digital camera monitoring observed at Denham boat ramp between 1 May 2013 and 30 April 2014. The inserted series details the behaviour of the process within the window outlined in red.

$1 - P(X = 1) = 0.90$. Meanwhile, in the second case, there is almost twice the probability of observing a zero (Freeland and McCabe, 2004). If we consider data of hourly distribution of boating retrievals, the upper tail distributional properties of boating activities are of interest for management purposes, as this information will help identify unusual growth patterns in boating activities and appropriately match to sustainability expectations. So long as the forecast estimates provided are above the zero threshold, meaningful inference can be deduced. For example, a mean estimate of 5.75 boating retrievals per day is meaningful and interpretable in practice. Consequently, the conditional mean was preferred over the conditional median.

2.2.2. *Croston's method*

For given time series data with many zeros, the method decomposes the original time series by constructing two new series using simple exponential smoothing forecasts, one for the time periods that observed

non-zero counts and the other for the inter-arrival times between non-zero counts (Croston, 1972). The models are constructed noting periods that contain zero and non-zero counts (Shenstone and Hyndman, 2005). Let Y_t, for $t = 1, \cdots, T$, be the count of powerboat retrievals occurring during period t, and I_t, an indicator variable for non-zero count periods, such that $I_t = 0$ when zero count occurs at time t and $I_t = 1$ otherwise. Further let k_t be the number of periods with non-zero counts between $[0, t]$, implying that $k_t = \sum_{i=1}^{t} I_i$. Then, let z_k represent the k^{th} non-zero count and a_k the interarrival time between y_{k-1} and y_k, and thus, we can write that $Y_t = I_t y_k$. Based on data up to count t, the one-step forecast for $(k+1)$ count and inter-arrival time from Croston method is given as

$$\widehat{z}_{k+1|k} = (1-\alpha)\widehat{z}_{k|k-1} + \alpha z_k \tag{5}$$

$$\widehat{a}_{k+1|k} = (1-\alpha)\widehat{a}_{k|k-1} + \alpha a_k \tag{6}$$

where $\alpha \in [0,\ 1]$ is a smoothing parameter, assumed to be the same for Eqs. (5) and (6). Let t be the time for last observed positive count in the data, then the h-step ahead forecast for the count at time $t+h$, is given by

$$\widehat{Y}_{t+h|T} = z_{t+1|t} / a_{t+1|t} \tag{7}$$

Shenstone and Hyndman (2005) found no algebraic results for the computation of the prediction interval around $\widehat{Y}_{T+h|T}$, as there is no statistical model that corresponds to the Croston method.

2.2.3. Syntetos-Boylan Approximation (SBA)

SBA applies a debiasing factor to the Croston's method to reduce the error in the final estimate. The resulting estimate for the h-step ahead forecast for the count at time $t+h$ from the SBA method is given by

$$\widehat{Y}_{t+h|T} = \left(1 - \frac{\alpha}{2}\right) \frac{y_{t+1|t}}{a_{t+1|t}} \tag{8}$$

where α is defined in Eqs. (5) and (6).

2.2.4. Integer-valued moving average (INMA)

An INMA (∞) process $(Y_t)_{\mathbb{N}}$ follows the recursive

$$Y_t = \sum_{i=0}^{\infty} (\theta_i \circ_t \varepsilon_{t-i}) \tag{9}$$

where $\circ_t$ is a binomial thinning operator at time t, which is a compound of Bernoulli i.i.d. random variables used (McKenzie 1988). The sequence $\{\varepsilon_t : t \in \mathbb{N}\}$ is integer-valued and i.i.d. with non-negative expected value $E(\varepsilon_t) = \mu$ and variance $Var(\varepsilon_t) = \sigma^2$. The parameters satisfy $\theta_0, \cdots, \theta_{q-1} \in [0, 1]$, $\theta_\infty \in (0, 1]$ and usually $\theta_0 = 1$. Assuming that $(Y_t)_{\mathbb{N}}$ is a stationary process, such that $\sum_{i=0}^{\infty} \theta_i < \infty$, then the unconditional mean and variance are expressed as

$$E(Y_t) = \lambda \left(1 + \sum_{i=1}^{\infty} \theta_i\right) \tag{10}$$

$$\text{Var}(Y_t) = \lambda \sum_{i=1}^{\infty} \theta_i (1-\theta_i) + \sigma^2 \left(1 + \sum_{i=1}^{\infty} \theta_i^2\right) \tag{11}$$

The unconditional moments for the process are

$$E(Y_t | Y_{t-1}) = \lambda + \sum_{i=1}^{\infty} \theta_i \varepsilon_{t-i} \tag{12}$$

$$\text{Var}(Y_t | Y_{t-1}) = \sigma^2 + \sum_{i=1}^{\infty} \theta_i (1-\theta_i) \varepsilon_{t-i} \tag{13}$$

For $h \geq 1$, the h-step ahead forecast is obtained as follows

$$\widehat{Y}_{t+h|T} = \lambda \sum_{i=1}^{h-1} \theta_i + \sum_{i=h}^{\infty} \theta_i \widehat{Y}_{t+h-i|T} \tag{14}$$

As $h \to \infty$, the limiting value of $\widehat{Y}_{T+h|T}$ is $\lambda \sum_{i=0}^{\infty} \theta_i$, which is the mean of the process.

2.2.5. Integer-valued autoregressive (INAR)

Let ε_t be an independent and identically distributed process with range $\mathbb{N}_0$ such that $E(\varepsilon_t) = \mu_\varepsilon$ and $Var(\varepsilon_t) = \sigma^2_\varepsilon$. Let $\alpha_1, \cdots, \alpha_p \in (0, 1)$, with $\alpha_\blacksquare = \sum_{i=1}^{p} \alpha_i < 1$. Then, an INAR($p$) process $(Y_t)_{\mathbb{N}}$ is defined recursively as

$$Y_t = \sum_{i=1}^{p} (\alpha_i \circ_t Y_{t-i}) + \varepsilon_t \tag{15}$$

where ε_t is independent of all Y_m and $\alpha_j \circ_{m+j} Y_m$ with $m < t$ and $j = 1, \cdots, p$, and all thinnings are performed independently of each other and ε_t and past observations of Y, where $\circ_t$ denotes the new thinning performed at time t. The mean and variance of a stationary $INAR$(p) process obtained via conditional maximum likelihood estimator (see Bourguignon et al., 2015) are given by

$$E(Y_t) = \mu_Y = \mu_\varepsilon / (1 - \alpha_\blacksquare) \tag{16}$$

$$Var(Y_t) = \mu_Y \cdot \left(\sum_{j=1}^{p} \alpha_j (1-\alpha_j) + \frac{\sigma^2_\varepsilon}{\mu_\varepsilon} \cdot (1-\alpha_\blacksquare)\right) \Big/ \left(1 - \sum_{i=1}^{p} \alpha_i . \rho(i)\right) \tag{17}$$

INAR(p) shows the typical AR(p) autocorrelation structure (Weiß, 2013) given by

$$\rho(k) = \sum_{i=1}^{p} \alpha_i \cdot \rho(|k-i|) \tag{18}$$

for. $k \geq 1$.

For practical purposes, the point forecast was taken to be the conditional mean, which is obtained by iteratively applying the law of total expectation. The h-step ahead forecast is estimated as follows

$$\widehat{Y}_{T+h|T} = \alpha^h \cdot Y_t + \mu(1 - \alpha^h) \tag{19}$$

for some $h \geq 1$. As the process $(Y_t)_{\mathbb{N}}$ is assumed to be a discrete-valued Markov chain, the conditional mean only depends on Y_t, but not on earlier observations (Weiß, 2018, Chapter 2). For strict coherent forecasting, Freeland and McCabe (2004) provide an explanation on how to estimate the transition probability distribution, whose corresponding conditional median or mode can be used.

The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used to select the best model among the competing integer-valued models.

2.3. Point forecast accuracy evaluation

Data were split into training and test sets to evaluate the performance of the models. In assessing forecast accuracy of single series, mean absolute error (MAE), root mean square error (RMSE) and mean absolute scaled error (MASE) were identified as the most appropriate performance metrics (Hyndman, 2014) for intermittent demand data.

$$MAE = \frac{\sum_{t=1}^{T} |Y_t - \widehat{Y}_t|}{T} \quad (20)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} (Y_t - \widehat{Y}_t)^2}{T}} \quad (21)$$

$$MASE = \frac{1}{K} \sum_{k=1}^{K} \left| Y_{T+k} - \widehat{Y}_{T+K|T} \right| / Q \quad (22)$$

where Q is a scaling factor defined using seasonal naïve forecast:

$$Q = \frac{1}{T-m} \sum_{t=m+1}^{T} |Y_t - Y_{t-m}|,$$

and $\widehat{Y}_{T+K|T}$ is an estimate of Y_{T+k} given the observations Y_t, for $t = 1, \cdots, T$, m is the seasonal index for the data and K is the number of step(s) for the forecasts made. We evaluated cases where. $K = 12, 24, 48$ and 168.

All models were implemented in R software (R Core Team, 2017), using packages '*acp* (version 2.1)' (Vasileios, 2015) for fitting the ACP models, '*tsintermittent* (version 1.9)' for fitting Croston and SBA models (Kourentzes and Petropoulos, 2016) and '*tscount* (version 1.4.2)' (Liboschik et al., 2017, 2020) for the integer-valued models and '*forecast* (version 8.7)' (Hyndman et al., 2019; Hyndman and Khandakar, 2008).

3. Results

We evaluated the point estimates of 5 different time series forecasting models to reconstruct recreational boating effort data that have high proportion of zeros observed at three different boat ramps. From Figures 2, 3, and 4, beyond the oscillating behaviour of the ACF for boating activity at each of the three locations, there is exponential decay in the autocorrelation for the peak hours of boat launches, suggesting a moving average of order 1. Also, the PACF highlights the prominence of lag-1 estimates in comparison to subsequent lags. The sudden drop in the PACF estimates at lag 1 and 2 is comparable to a tailing-off behaviour, making an autoregressive order of 1 a good initial choice. The search for the order (maximum order = 4) of the integer-valued models and Autoregressive Conditional Poisson model showed that for each of the 3 locations INMA (1), INAR (2) and ACP (2,1) models were the best ranked integer-valued models and Autoregressive Conditional Poisson among competing models based on AIC and BIC.

Table 1 compares the forecasting accuracy of the models for the respective performance metrics based on the predictions of the training data and *h*-steps test data. In cases where the metrics followed divergent paths in ranking the models, MASE was used in assessing the models (Hyndman 2006). For the training data, performance metrics diverged in the selection of the best forecasting method. ACP (2,1) and INAR (2) were the competing models across the forecasting horizons, with ACP (2,1) ranked best based on MASE in Denham and Leeuwin. However, for Broome INAR (2) was ranked the best by all performance metrics. The relatively low differences between the magnitude of error values for the two best models suggest that ACP (2,1) and INAR (2) produced similar point estimates of forecast values. The intermittent demand forecasting methods (Croston and SBA) were the two worst performing models.

Similar results were observed for the models in assessing their predictive power on the test data. It was observed that, the accuracy of the techniques investigated deceased with an increase in the forecast horizon (see Figure 5). While Croston, SBA and INMA (1) produced constant forecasts for the lead times, ACP (2,1) and INAR (2) produced forecasts that adequately captured the fluctuations in the original dataset, even at $h = 168$ (see Figure 4). Looking in more detail at Figure 5, ACP (2,1) and INAR (2) performed better for short forecast horizons and provided some evidence of their sufficiency in forecasting the dynamics in recreational boating activities. ACP (2,1) and INAR (2) were comparable (and best), with the ACP (2,1) model often ranked the best, especially for data with high proportion of zeros. Croston's model generally had the greatest errors with very slight improvements when SBA was used.

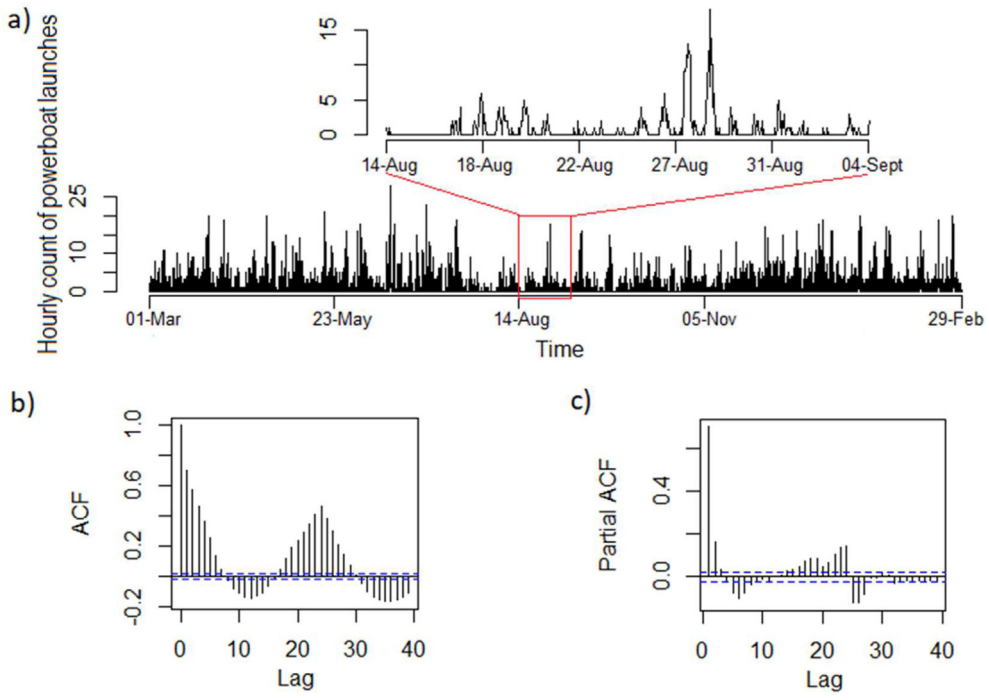

Figure 4. a) Time series plots b) ACF and c) PACF of the count of powerboat launches from digital camera monitoring observed at Leeuwin boat ramp between 1 May 2013 and 30 April 2014. The inserted series details the behaviour of the process within the window outlined in red.

Table 1. Mean absolute error (MAE), root mean square error (RMSE) and mean absolute scaled error (MASE) values for the performance evaluation of five techniques used to forecast lead times of 12-, 24-, 48- and 168-hours of recreational boating data for three study locations.

Location	h-forecast	Performance indicator	Training data					Test data				
			Croston	SBA	ACP(2,1)	INMA (1)	INAR(2)	Croston	SBA	ACP(2,1)	INMA (1)	INAR(2)
Broome	12-h	MAE	1.7209	2.3302	0.9357	1.5420	**0.9354**	1.7348	1.9779	1.2179	2.1464	**0.6687**
		RMSE	18558	2.5132	1.5268	2.3929	**1.5256**	2.4289	2.5504	1.7431	2.2097	**0.8521**
		MASE	1.9013	2.5745	0.6068	1.9694	**0.6066**	1.9167	2.1852	0.4383	1.7240	**0.4337**
	24-h	MAE	2.3317	2.0143	0.9361	1.5431	**0.9359**	1.9935	2.0992	1.3757	2.1174	**0.7076**
		RMSE	2.4815	2.1708	1.5276	2.3941	**1.5263**	2.5580	2.6181	1.9870	2.2495	**0.9495**
		MASE	1.9002	2.2539	0.6066	1.6989	**0.6065**	2.2011	2.3178	0.4915	1.7241	**0.4585**
	48-h	MAE	2.1773	2.1437	0.9367	1.5446	**0.9364**	1.8401	1.8081	1.0734	1.2544	**0.5348**
		RMSE	2.2654	2.2315	1.5287	2.3965	**1.5274**	2.4764	2.4606	1.5918	1.3278	**0.7570**
		MASE	2.4039	2.3668	0.6064	1.8969	**0.6063**	2.0315	1.9963	0.6949	1.8121	**0.3462**
	168-h	MAE	2.2578	2.2109	0.9350	1.5353	**0.9348**	1.8288	1.7732	1.0688	1.6283	**1.0457**
		RMSE	2.8099	2.7832	1.5226	2.3856	**1.5213**	2.4649	2.4383	1.8408	2.5567	**1.7478**
		MASE	2.4983	2.4464	0.6090	1.3989	**0.6089**	2.0236	1.9620	0.6961	1.0606	**0.6811**
Denham	12-h	MAE	3.1833	2.4157	**0.6094**	0.9112	0.6097	3.1453	2.2610	0.7711	0.7331	**0.6096**
		RMSE	3.3279	2.6152	1.2447	1.6692	**1.2411**	3.2581	2.3397	1.1530	0.8906	**0.9784**
		MASE	5.6825	4.3124	**0.6689**	3.4378	0.6691	5.6147	4.0362	**0.8463**	2.6265	1.0507
	24-h	MAE	2.6009	2.4175	**0.6096**	0.9117	0.6098	2.8717	2.7984	**0.7049**	2.2167	1.3405
		RMSE	2.7843	2.6169	1.2453	1.6702	**1.2416**	3.4997	3.4777	**0.9799**	3.7982	2.3423
		MASE	4.6421	4.3148	**0.6686**	2.7848	0.6688	5.1255	4.9947	**0.7732**	3.8767	1.8489
	48-h	MAE	2.1308	2.4137	**0.6090**	0.9120	0.6092	2.5901	2.6950	**0.7375**	1.9041	1.2047
		RMSE	2.3719	2.6138	1.2456	1.6717	**1.2419**	3.1338	3.1791	**1.0473**	3.2809	2.1211
		MASE	3.8083	4.3138	**0.6677**	2.9926	0.6680	4.6291	4.8166	**0.8086**	2.9878	1.5247
	168-h	MAE	2.7949	2.3168	**0.6002**	0.8913	0.6005	2.5071	2.3920	1.0392	1.6075	**1.0084**
		RMSE	2.9719	2.5190	1.2294	1.6337	**1.2257**	2.9697	2.8973	**1.8633**	2.8665	1.9371
		MASE	5.0592	4.1937	**0.6734**	1.9589	0.6737	4.5383	4.3299	**1.1658**	1.9906	1.3784
Leeuwin	12-h	MAE	2.1082	2.2078	**1.0923**	1.6776	1.0935	2.3053	2.3930	**0.6046**	1.1674	0.6567
		RMSE	2.7297	2.7817	1.8354	2.5942	**1.8348**	2.3882	2.4729	**0.7337**	1.2363	0.7794
		MASE	1.9682	2.0612	**0.6511**	0.9869	0.6518	2.1522	2.2341	**0.3604**	0.6958	0.3914
	24-h	MAE	2.1747	2.1387	**1.0931**	1.6792	1.0943	2.0429	2.0168	**0.6397**	1.1593	1.0134
		RMSE	2.7645	2.7457	1.8365	2.5957	**1.8360**	2.1860	2.1600	**0.7600**	1.3116	1.3258
		MASE	2.0288	1.9952	**0.6510**	0.7673	0.6517	1.9058	1.8815	**0.3809**	0.6904	0.6035
	48-h	MAE	2.0866	2.2098	**1.0934**	1.6819	1.0952	2.0279	1.9357	**0.7208**	1.0810	0.8149
		RMSE	2.7212	2.7846	1.8383	2.5986	**1.8377**	2.8484	2.1205	**0.8983**	1.2184	1.1040
		MASE	1.9456	2.0605	**0.6504**	0.8881	0.6512	1.7235	1.8049	**0.4286**	0.6427	0.4845
	168-h	MAE	2.3696	2.1380	**1.0580**	1.6622	1.0862	2.2543	2.1644	**1.1409**	1.6916	1.4139
		RMSE	2.8599	2.7216	1.8229	2.5592	**1.8227**	2.6099	2.5412	**1.5862**	2.4216	2.3609
		MASE	2.2257	2.0081	**0.6528**	0.9934	0.6535	2.1174	2.0329	**0.6864**	1.0177	0.8506

The best ranked technique with respect to locations have bold text.

4. Discussion

We compared intermittent demand and various count data time series modelling techniques based on forecasting accuracy of observed counts of recreational powerboat launches using lead times of 12, 24, 48 and 168 h. Here, we evaluated and compared the point estimates of forecasts obtained from these time series modelling techniques which are noted for modelling data with characteristics of significant intermittency and relative low counts of events. With such data, techniques that apply a continuum approximation fail (Czado et al., 2009). A key result of this study is that although intermittency is common in recreational boating effort data, intermittent demand forecasting models performed relatively badly as forecasting tools and were outperformed by integer-valued and conditional autoregressive models. Thus, although intermittency may be a key feature for a given dataset, it should not override the systemic characteristics of the data in the application of forecasting techniques. Integer-valued autoregressive (INAR) and autoregressive conditional Poisson (ACP) models were identified as useful for predicting short-term behaviour of recreational boating effort. Outcomes of this study provide robust short-term forecasting methods for imputing missing data in digital camera monitoring, and also provide time series modelling frameworks for forecasting purposes to guide resources allocation and management in many different areas of application, including, fisheries and wildlife management, tourism, and transportation.

In this study, ACP (2, 1) was identified as the best model for short-term forecasting of counts of recreational boating effort that exhibit properties of intermittent demand data. ACP models have been found to sufficiently model rare events, for example, assessing the volatility of the daily number of price changes in a stock (Heinen, 2003). ACP is a parametric approach, in which the marginal distribution for the count process is specified, such that the mean conditional on past observations is autoregressive. The model adequately addresses discreteness, over-dispersion and serial correlation. An advantage of using an ACP model is that it is estimated using maximum likelihood and this enables common tests to be performed, for instance, testing for autocorrelation and other standard hypothesis tests (Heinen, 2003). In addition, the autocorrelation and density are modelled explicitly allowing both point and density forecasting, allowing easy interpretation of results and flexibility in

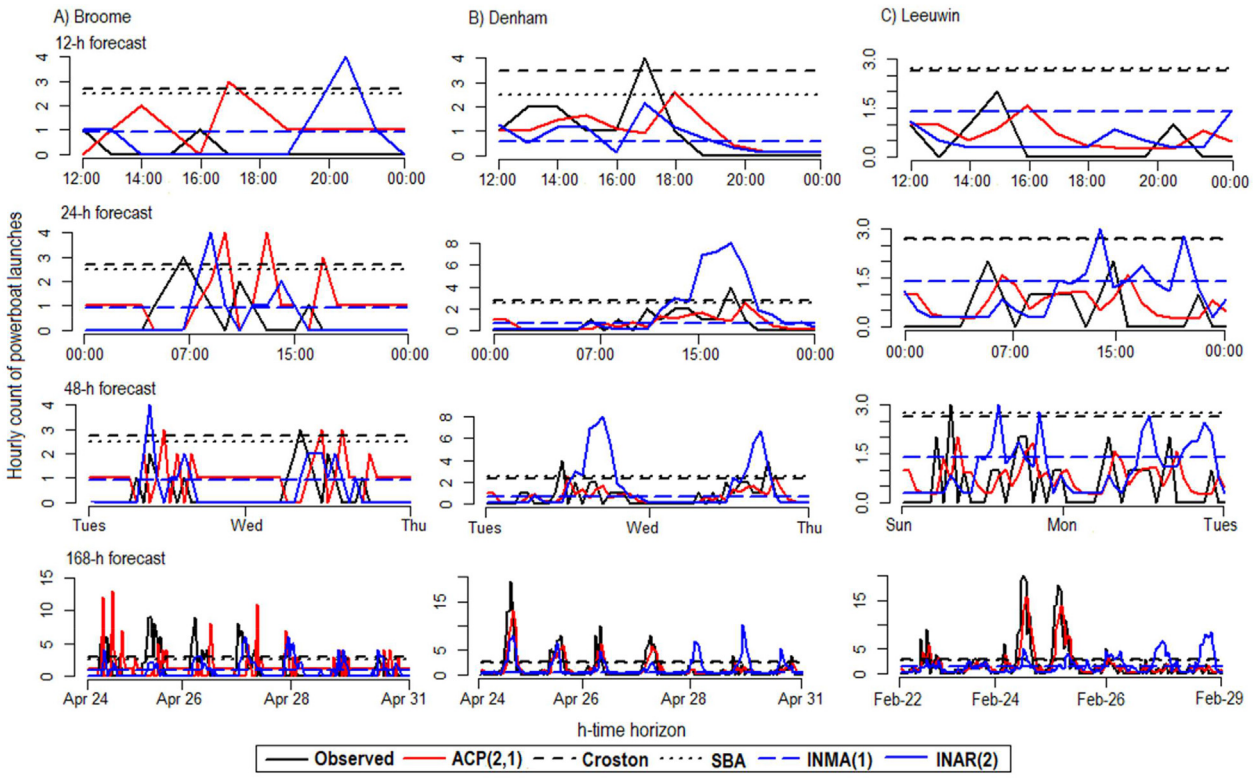

Figure 5. Reconstruction of test data based on the five time series models formulated using lead times of 12-, 24-, 48- and 168-hours at A) Broome, B) Denham and C) Leeuwin

application as ACP can be easily extended to most of the class of models in autoregressive conditional heteroskedasticity (ARCH) (Bollerslev et al. (1994); Engle, 1982; Heinen, 2003).

The INAR (2) model was also found to produce adequate forecast estimates for the recreational boating effort data. INAR models explicitly account for discreteness of data and have been identified to perform well in modelling datasets that are characterized by low counts, asymmetric distributions, excess zeros and overdispersion. These models are easy to apply and have wide areas of application. There are many approaches to fitting these models including models based on the theory of generalized linear models (Kedem and Fokianos, 2002), models where parameters are estimated on the premise that observed processes are driven by unobserved processes (Davis and Wu, 2009), models based on renewal processes for generating a correlated sequence using Bernoulli trials (Cui and Lund, 2009), and models that are observation-driven (Lu, 2018). Interestingly, they are linear-like for discrete time series with less complicated correlation structures and have likelihood functions and multi-step probabilistic forecasts that are numerically tractable (Lu, 2018; Silva, 2015).

Integer-valued moving average models (INMA) have often been used to model non-linear and/or non-stationary count data (Brännäs et al., 2002; Brä nn ä s and Shahiduzzaman Quoreshi, 2010). They typically produce parameters that take on values in narrower intervals compared to autoregressive moving average models (ARMA), satisfying the requirement that each parameter estimate of lagged variables must be found within a unit interval (Brä nn ä s and Shahiduzzaman Quoreshi, 2010). In this study, INMA (1) did not perform well compared to ACP (2, 1) and INAR (2) models, although it generally provided better forecasts than the Croston or SBA methods. In general, the estimation procedure for INMA models needs to be developed based on both known and unknown underlying distribution of the data (Quoreshi et al., 2019). Quoreshi et al. (2019) indicated that in the context of high frequency data, different applications may not have a known underlying distribution. Since the distribution of the counting series of high frequency data is generally unknown, the application of likelihood-based approaches for INMA models becomes restricted. The performance of forecasted values estimated by INMA (1) in this study indicates that an investigation into different estimation procedures may be required to establish a good fit for recreational boating effort data. Some inferential procedures for INMA models have been developed via conditional least squares (CLS), feasible generalized least square (FGLS) and generalized method of moments (GMM) approaches Aleksandrov and Weiß (2020); Brä nn ä s and Shahiduzzaman Quoreshi (2010); Martin et al. (2014); Quoreshi (2008). Additionally, the irregular nature of the correlational structure in the data may require the fit of a long-lag INMA model, which has been established to satisfy the modelling restriction of parameter estimates lying within a unit interval (Brä nn ä s and Shahiduzzaman Quoreshi, 2010). Furthermore, recreational boating effort datasets may also exhibit the long memory property, looking at the relatively slowly decaying autocorrelation during peak hours for boat launches (see Figures 2, 3, and 4). However, the long memory properties of parameters estimated from INMA models are generally assumed dependent and follow a gamma function and based on the cyclic nature of the recreational boat launch dataset, probably other functional forms with cyclical behaviour may be more appropriate.

The Croston model was originally designed to represent demand patterns for considerably slow-moving items which are dominant in service and inventory management industries. The model was developed to avoid errors inherent in exponential smoothing, which often leads to excess prediction of stock levels (Croston, 1972). This would ultimately produce forecasts that would enable reasonable control of stock system, thus eliminating or reducing the risk of obsolescence of intermittent items (Syntetos2015). A key component in the forecast estimation is ascertaining the inter-arrival intervals for the demand process, supposedly recognising the stochasticity of arrivals and differing sizes of demand. However, Shenstone and Hyndman (2005) showed that for this the technique there is no properly formulated underlying stochastic model. The original Croston model assumes that the distribution of

non-zero demand sizes is normal, the distribution of inter-arrival times is geometric, and that demand sizes and inter-arrival times are mutually independent. Revised assumptions have resulted in modified estimations techniques for the Croston model, including SBA (Syntetos and Boylan, 2005) and Shale-Boylan-Johnston (SBJ) (Shale et al., 2006). These techniques are widely reported as adequate tools for forecasting intermittent demand data (Gardner, 2006; Pennings et al., 2017). Although intermittent demand series usually contain a significant proportion of zero values, with non-zero values present randomly, it is important to note that the mechanism and complexity of the stochasticity of recreational boating activities within a given period may be different from the demand process of a stock system. In this study, we have established that intermittency should not override the overall systemic characteristics of a data in application of forecasting techniques.

5. Conclusion

Digital camera monitoring has become an established means of data collection in many fields of study, such as ecology, transportation and operation systems, and tourism. For example, the non-invasive property of digital cameras makes them more useful for monitoring rare events in application, where data generated are characterized by low counts with intermittencies, with some missing data due to outages. The ability to accurately predict the short-term temporal distribution would be useful support for ongoing monitoring programmes, particularly in imputing missing data. We have demonstrated that short-term forecast of such events could be performed using integer-valued and autoregressive conditional Poisson models. These modelling frameworks would be applicable for making necessary projections in anticipation of the current and future needs of short-term events.

Declarations

Author contribution statement

Ebenezer Afrifa-Yamoah: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Ute Mueller: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Ute Mueller was supported by the Government of Western Australia Department of Primary Industries and Regional Development (G1002222).

Data availability statement

The authors do not have permission to share data.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The authors express their sincere gratitude to the staff of DPIRD who spent much time reading the camera data. A special thanks to Dr. Stephen M. Taylor for his inputs and suggestions during the development of the manuscript. We are grateful to Dr. Ainslie Denham for her time and inputs during the internal review process by DPIRD. We are grateful for the insightful comments provided by the editor and the reviewers.

References

Afrifa-Yamoah, E., Mueller, U.A., Fisher, A.J., Taylor, S.M., 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)*, Guimarães, Portugal, 7-12 July.

Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M., Fisher, A.J., 2020b. Missing Data Imputation of High-Resolution Temporal Climate Time Series Data. Meteorological Applications, pp. 1–18.

Afrifa-Yamoah, E., Taylor, S.M., Fisher, A.J., Mueller, U., 2020a. Imputation of missing data from time-lapse cameras used in recreational fishing surveys. ICES (Int. Counc. Explor. Sea) J. Mar. Sci.

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 Coast Manag. 215, 105947.

Aleksandrov, B., Weiß, C.H., 2020. Parameter estimation and diagnostic tests for INMA (1) processes. Test 29, 196–232.

Al-Osh, M., Alzaid, A., 1988. Integer-valued moving average (INMA) process. Stat. Pap. 29, 281–300.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31, 307–327.

Bollerslev, T., Engle, R.F., Nelson, D.B., 1994. ARCH models. Handb. Econom. 4, 2959–3038.

Bourguignon, M., Vasconcellos, K.L.P., Reisen, V.A., Ispány, M., 2015. A Poisson INAR (1) process with a seasonal structure. J. Stat. Comput. Simulat. 86, 373–387.

Brännäs, K., Hellström, J., Nordström, J., 2002. A new approach to modelling and forecasting monthly guest nights in hotels. Int. J. Forecast. 18, 19–30.

Brännäs, K., Shahiduzzaman Quoreshi, A.M.M., 2010. Integer-valued moving average modelling of the number of transactions in stocks. Appl. Financ. Econ. 20 (18), 1429–1440.

Croston, J.D., 1972. Forecasting and stock control for intermittent demands. Oper. Res. Q. 23 (3), 289–303.

Cui, Y., Lund, R., 2009. A new look at time series of counts. Biometrika 96, 781–792.

Czado, C., Gneiting, T., Held, L., 2009. Predictive model assessment for count data. Biometrics 65, 1254–1261.

Davis, R., Wu, R., 2009. A negative binomial model for time series of counts. Biometrika 96, 735–749.

Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 50, 987–1008.

Engle, R.F., Russell, J., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica 66 (1127), 1162.

Fokianos, K., Rahbek, A., Tjøstheim, D., 2009. Poisson autoregression. J. Am. Stat. Assoc. 104, 1430–1439.

Freeland, R.K., McCabe, B.P.M., 2004. Forecasting discrete valued low count time series. Int. J. Forecast. 20 (3), 427–434.

Gardner, E.S., 2006. Exponential smoothing: the state of the art – part II. Int. J. Forecast. 22 (4), 637–666.

Hartill, B.W., Taylor, S.M., Keller, K., Weltersbach, M.S., 2020. Digital camera monitoring of recreational fishing effort: Applications and challenges. Fish and Fisheries 21, 204–215.

Heinen, A., 2003. Modeling Time Series Count Data: an Autoregressive Conditional Poisson Model. MPRA Paper No. 8113, Available online at. https://mpra.ub.uni-muenchen.de/8113/.

Hyndman, R.J., 2006. Another look at forecast-accuracy metrics for intermittent demand. Foresight 4, 43–46.

Hyndman, R., 2014. Time Series Forecasting Accuracy Measures: MAPE and MASE, URL (Version: 2014-10-03). https://stats.stackexchange.com/users/159/rob-hyndman. https://stats.stackexchange.com/q/117744.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeen, F., 2019. Forecast: Forecasting Functions for Time Series and Linear Models. R package version 8.7. http://pkg.robjhyndman.com/forecast.

Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. J. Stat. Software 26 (3), 1–22.

Kedem, B., Fokianos, K., 2002. Regression Models for Time Series Analysis. Wiley, Hoboken, NJ.

Kolassa, S., 2016. Evaluating predictive count data distributions in retail sales forecasting. Int. J. Forecast. 32, 788–803.

Kourentzes, N., Barrow, D., Petropoulos, F., 2018. Another look at forecast selection and combination: evidence from forecast pooling. Int. J. Prod. Econ. 1–10.

Kourentzes, N., Petropoulos, F., 2016. Tsintermittent: Intermittent Time Series Forecasting. R package version 1.9. https://CRAN.R-project.org/package=tsintermittent.

Lancaster, D., Dearden, P., Haggarty, D.R., Volpe, J.P., Ban, N.C., 2017. Effectiveness of shore-based remote camera monitoring for quantifying recreational Fisher compliance in marine conservation areas. Aquat. Conserv. Mar. Freshw. Ecosyst. 27, 804–813.

Lepot, M., Aubin, J.-B., Clemens, H.L.R., 2017. Interpolation in time series: an introductive overview of existing methods, their performance criteria and uncertainty assessment. Water 9, 796.

Liboschik, T., Fokianos, K., Fried, R., 2017. Tscount: an R package for analysis of count time series following generalized linear models. J. Stat. Software 82 (5), 1–51.

Liboschik, T., Fried, R., Fokianos, K., Probst, P., 2020. Tscount: Analysis of Count Time Series. R package version 1.4.2. https://CRAN.R-project.org/package=tscount.

Lu, Y., 2018. Exact Likelihood Estimation and Forecasting in Higher-Order INAR(p) Models. University Library of Munich, Germany. MPRA Paper 83682.

Martin, V.L., Tremayne, A.R., Jung, R.C., 2014. Efficient method of moments estimators for integer time series models. J. Time Anal. 35 (6), 491–516.

McKenzie, E., 1988. Some ARMA models for dependent sequences of Poisson counts. Adv. Appl. Probab. 20, 822–835.

Pennings, C.L.P., van Dalen, J., van der Laan, E., 2017. Exploiting elapsed time for managing intermittent demand for spare parts. Eur. J. Oper. Res. 258, 958–969.

Petropoulos, F., Kourentzes, N., 2014. Improving forecasting via multiple temporal aggregation. Foresight: Int. J. Appl. Forecast. 34, 12–17.

Quoreshi, A.M.M.S., 2008. A vector integer-valued moving average model for high frequency financial count data. Econ. Lett. 101, 258–261.

Quoreshi, A.M.M.S., Mamode Khan, N.A., Uddin, R., 2019. A Review of INMA Integer-Valued Model Class, Application and Further Development. Retrieved from. http://bth.diva-portal.org/smash/get/diva2:1316274/FULLTEXT01.pdf on 21/07/2020.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rao, A.V., 1973. A comment on: forecasting and stock control for intermittent demands. Oper. Res. Q. 24 (4), 639–640.

Ryan, K.L., Hall, N.G., Lai, E.K., Smallwood, C.B., Taylor, S.M., Wise, B.S., 2017. State-wide Survey of Boat-Based Recreational Fishing in Western Australia 2015/16. *Fisheries Research Report*. No. 287. Department of Primary Industries and Regional Development, Western Australia.

Shale, E.A., Boylan, J.E., Johnston, F.R., 2006. Forecasting for intermittent demand: the estimation of an unbiased average. J. Oper. Res. Soc. 57, 588–592.

Shenstone, L., Hyndman, R.J., 2005. Stochastic models underlying Croston's method for intermittent demand forecasting. J. Forecast. 24 (6), 389–402.

Silva, M.E., 2015. Modelling time series of counts: an INAR approach. Textos de Matemática 47, 107–121.

Snyder, R.D., Ord, K.J., Beaumont, A., 2012. Forecasting the intermittent demand for slow-moving inventories: a modelling approach. Int. J. Forecast. 28, 485–496.

Steffe, A.S., Taylor, S.M., Blight, S.J., Ryan, K.L., Desfossess, C., Tate, A., Smallwood, C.B., Lai, E.K., Trinnie, F.I., Wise, B.S., 2017. Framework for Integration of Data from Remotely Operated Cameras into Recreational Fishery Assessments in Western Australia. Fisheries Research Report No. 286. Department of Primary Industries and Regional Development (DPIRD), WA.

Steutel, F.W., van Harn, K., 1979. Discrete analogues of self-decomposability and stability. Ann. Probab. 7 (5), 893–899.

Syntetos, A.A., Boylan, J.E., 2005. The accuracy of intermittent demand estimates. Int. J. Forecast. 21, 303–314.

Syntetos, A.A., Zied Babai, M., Gardner Jr., E.S., 2015. Forecasting intermittent inventory demands: simple parametric methods vs, bootstrapping. J. Bus. Res. 68, 1746–1752.

Teunter, R.H., Duncan, L., 2009. Forecasting intermittent demand: a comparative study. J. Oper. Res. Soc. 60, 321–329.

van Poorten, B.T., Brydle, S., 2018. Estimating fishing effort from remote traffic counters: opportunities and challenges. Fish. Res. 204, 231–238.

Vasileios, S., 2015. Acp: Autoregressive Conditional Poisson. R Package Version 2.1. https://CRAN.R-project.org/package=acp.

Weiß, C.H., 2018. An Introduction to Discrete-Valued Time Series. John Wiley & Sons, Hoboken, NJ.

Weiß, C.H., 2008. Serial dependence and regression of Poisson INARMA models. J. Stat. Plann. Inference 138, 2975–2990.

Weiß, C.H., 2013. Integer-valued autoregressive models for counts showing under dispersion. J. Appl. Stat. 40 (9), 1931–1948.

Wise, B.S., Fletcher, W.J., 2013. Determination and Development of Cost-Effective Techniques to Monitor Recreational Catch and Effort in Western Australian Demersal Finfish Fisheries, Final Report for FRDC Project 2005/034 and WAMSI Subproject 4.4.3. Fisheries Research Report.