CCAMLR Science, Vol. 17 (2010): 229–241

IMPROVING ESTIMATES OF ADÉLIE PENGUIN BREEDING POPULATION SIZE: DEVELOPING FACTORS TO ADJUST ONE-OFF POPULATION COUNTS FOR AVAILABILITY BIAS

C. Southwell✉, J. McKinlay, L. Emmerson, R. Trebilco* and K. Newbery
Australian Antarctic Division
Department of Environment, Water, Heritage and the Arts
203 Channel Highway
Kingston, Tasmania 7050
Australia
Email – colin.southwell@aad.gov.au

*Current address: Biological Sciences
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6
Canada

Abstract

New methods are presented for the collection and development of data to adjust counts of Adélie penguins made at any time in the breeding season to an estimate of the breeding population. The development of 'availability adjustment factors' involves the collection of time-series counts of population objects, such as adults, nests or chicks, throughout the breeding season and standardisation of the time series to a reference point one week after the peak in egg laying, consistent with Standard Method A3 of the CCAMLR Ecosystem Monitoring Program (CEMP). Remotely operating cameras were used to obtain time-series counts of adults and nests, while time-series counts of chicks were obtained manually. Standardised time series were modelled using a generalised additive model framework to allow interpolation across days when no counts were available and to provide a formal estimate of uncertainty around the average function. The method is illustrated using data collected at the Béchervaise Island CEMP site in 2007/08. Estimates of the breeding population obtained by adjusting a one-off count using availability adjustment factors will not account for penguins that left the breeding site prior to the reference point because their breeding attempt failed, breeding-age penguins that failed to return to the breeding site in that breeding season, nor pre-breeding penguins. The abundance of these components of the population can only be estimated through alternative methods such as mark-recapture of individually tagged birds. Given the paucity of adjustment data in the literature, further collection of availability adjustment data across space and time is recommended.

Résumé

De nouvelles méthodes sont présentées pour la collecte de données et leur développement en vue d'ajuster les recensements de manchots Adélie effectués à différentes périodes pendant la saison de reproduction pour obtenir une estimation de la population reproductrice. Pour déterminer des « facteurs de correction en fonction de la disponibilité » il faut disposer de séries chronologiques de diverses composantes des populations telles que des adultes, des nids ou des jeunes, qui sont dénombrées tout au long de la saison de reproduction, et les normaliser par rapport à un point de référence une semaine après le pic de la ponte, selon les termes de la méthode standard A3 du Programme de contrôle de l'écosystème de la CCAMLR (CEMP). Des caméras télécommandées ont permis d'obtenir les séries chronologiques des décomptes d'adultes et de nids, alors que la série des décomptes des poussins a été obtenue *de visu*. Les séries chronologiques normalisées ont été modélisées grâce à un modèle additif généralisé dont la structure permet l'interpolation des jours sans décomptes et de fournir une estimation formelle de l'incertitude entourant la fonction moyenne. La méthode est illustrée à l'aide de données collectées au site du CEMP de l'île Béchervaise en 2007/08. Les estimations de la population reproductrice obtenues en appliquant des facteurs de correction en fonction de la disponibilité à un recensement unique ne tiennent compte ni des manchots ayant quitté le site de reproduction avant le point de référence en raison de l'échec de leur tentative de reproduction, ni des manchots qui, bien qu'étant en âge de se reproduire, ne sont pas retournés au site pendant la saison de reproduction en question, ni des pré-reproducteurs. L'abondance de ces composantes de la population ne peut être estimée que par des méthodes différentes telles que la recapture d'oiseaux marqués individuellement. Compte tenu de la rareté des données d'ajustement dans la littérature, il est recommandé de collecter d'autres données d'ajustement en fonction de la disponibilité en divers endroits et à divers moments.

Резюме

Представлены новые методы сбора и обработки данных с целью корректировки подсчетов пингвинов Адели, проводившихся в любое время в период сезона размножения, в соответствии с оценкой размножающейся популяции. Разработка "корректировочных коэффициентов наличия" включает сбор временных рядов подсчетов объектов в популяции, таких как взрослые особи, гнезда или птенцы, в течение всего сезона размножения и стандартизацию этих временных рядов в соответствии с контрольным значением через неделю после пика кладки яиц, согласно стандартному методу A3 Программы АНТКОМ по мониторингу экосистемы (CEMP). Для получения временных рядов подсчетов взрослых особей и гнезд использовались дистанционно управляемые камеры, а временные ряды подсчетов птенцов были получены вручную. Стандартизованные временные ряды моделировались с использованием системы обобщенной аддитивной модели, что позволяет провести интерполяцию в те дни, когда не было никаких подсчетов, и получить формальную оценку интервала неопределенности средней функции. Данный метод иллюстрируется с помощью данных, собранных на участке CEMP на о-ве Бешервез в 2007/08 г. Оценки размножающейся популяции, полученные путем корректировки одноразового подсчета с использованием корректировочных коэффициентов наличия, не учитывают пингвинов, покинувших гнездовья до контрольного момента времени из-за неудавшейся попытки размножения, пингвинов репродуктивного возраста, не вернувшихся в гнездовье в данном сезоне размножения, и пингвинов до начала размножения. Численность этих компонентов популяции можно определить только с помощью альтернативных методов, таких как мечение и повторная поимка отдельно помеченных птиц. Учитывая малое количество корректирующих данных в литературе, рекомендуется продолжать собирать корректирующие данные о наличии в пространственных и временных масштабах.

Resumen

Se presentan nuevos métodos de recolección y tratamiento de datos con el fin de ajustar los recuentos de pingüinos adelia efectuados en cualquier momento durante la temporada de reproducción y obtener una estimación de la población reproductora. El desarrollo de "factores de ajuste del sesgo por disponibilidad" requiere la recolección de series cronológicas de recuentos de sujetos de una población objetivo (como adultos, nidos o polluelos) durante toda la temporada de reproducción, y la normalización de las series cronológicas en relación con un punto de referencia una semana después de la fecha de máxima postura de huevos, de conformidad con el método estándar A3 del Programa de Seguimiento del Ecosistema de la CCRVMA (CEMP). Se utilizaron cámaras operadas por control remoto para obtener las series cronológicas del número de adultos y nidos, mientras que el número de polluelos fue contado a mano. Las series cronológicas normalizadas fueron modeladas mediante un modelo aditivo generalizado con el fin de permitir la interpolación para días sin recuentos y para proporcionar una estimación formal de la incertidumbre del promedio. Se ilustra el método con datos recogidos en el sitio CEMP de Isla Béchervaise en 2007/08. Las estimaciones de la población reproductora obtenidas aplicando factores de ajuste del sesgo por disponibilidad a un solo recuento no dan cuenta de los pingüinos que salieron de la colonia de reproducción antes del punto de referencia al no tener éxito en la reproducción, los pingüinos en edad de reproducirse que no regresaron a la colonia en dicha temporada de reproducción, o los pingüinos en edad pre-reproductiva. La abundancia de estos componentes de la población sólo puede ser estimada con otros métodos como el marcado y recaptura de aves marcadas individualmente. Dada la escasez de datos publicados sobre el ajuste, se recomienda continuar recogiendo datos para ajustar el sesgo debido a la disponibilidad, en distintas escalas espaciales y temporales.

Keywords: abundance estimation, Adélie penguin, availability bias,
generalised additive modelling, parametric bootstrap, CCAMLR

Introduction

It is widely recognised that the interpretation of counts of penguin populations at their breeding sites during the breeding period is strongly dependent on the timing within a breeding season at which they are undertaken (Croxall and Kirkwood, 1979). Because of the population dynamics of birds present at any given breeding site throughout a breeding season, optimal times may exist for obtaining estimates of each of the different population components. Consequently, a count of some population object, such as adults, nests or chicks, outside the optimal time may need to be adjusted by some date- or stage-specific factor if the count is to form the basis of an estimate of some component of the total population, for example the breeding population (Lynch et al., 2009).

The need to adjust counts of penguins at their breeding sites with date- or stage-specific factors is related to the concept of 'availability bias'. This term is used in the abundance estimation literature (e.g. Marsh and Sinclair, 1989) to refer to counting bias that results from animals that are completely obscured from view, and hence are not available for counting. Terrestrial species taking shelter under vegetation, for example, may not be available for counting by an observer passing overhead in an aircraft, and marine species diving underwater may not be available for counting by an observer in a passing ship. The concept of availability can in principle be extended to surveys of colonial breeding species, such as penguins, where some breeding individuals may be foraging away from the breeding colony for periods during the breeding season, and hence are not available for counting at the breeding site (Southwell, 2004). For Antarctic penguins, including the Adélie penguin (*Pygoscelis adeliae*), the remoteness of breeding sites, and difficulty accessing them at predictable times of the year, means that population counts cannot always be obtained at an optimal time of a breeding season for obtaining an unbiased estimate of the breeding population. Consequently, one-off raw counts by themselves may be biased estimators of the breeding population and hence have low comparability with counts at other sites or times unless standardised to a common point of breeding chronology.

While the need to adjust raw penguin count data obtained at varying times within a breeding season to minimise bias in estimates of the breeding population has long been recognised, the considerable amount of work in obtaining one-off counts for population surveys has not been matched by the collection of adjustment data. This paper attempts to fill this gap by reporting on new methods for the collection and development of availability adjustment data for Adélie penguins in East Antarctica. Availability curves, such as those presented here, are further used as a critical data input to a parametric bootstrap model for bias-correcting historical and recent counts of Adélie penguins in Antarctica, as presented in a companion paper (McKinlay et al., 2010).

Material and methods

Relevant aspects of Adélie penguin population dynamics at the breeding site

Detailed studies have elucidated the dynamics of Adélie penguin populations at their breeding sites (e.g. Ainley, 2002). While these studies have generally not aimed to develop availability adjustment factors for population estimation, they provide important background information concerning patterns of attendance, described below. While the pattern presented here is based on data obtained from Béchervaise Island near Mawson Station, East Antarctica, it is applicable to other Adélie penguin populations.

Adélie penguins begin arriving at their breeding sites around mid-October, with adult numbers reaching a peak in early November (Figure 1). Many of these penguins form pairs, establish nests and subsequently produce eggs and chicks, but a proportion of the penguins either do not attempt to breed or fail in their initial breeding attempt. There is a subsequent decline in the number of birds at the breeding colony as females leave the breeding site shortly after laying their eggs around mid-November (Davis, 1988) to undertake a foraging trip of around 17 days duration (Emmerson et al., 2003). Non-breeders and failed-breeders also leave the breeding site around the time of egg-laying, leaving a population comprised almost entirely of single males incubating eggs at each nest. There can also be nest failure and the departure of attendant males in the time between first and last egg-lay. The number of occupied nests present after the last eggs have been laid, or equivalently around a week after the peak in egg laying, is considered the most representative count of the number of breeding pairs at the colony as defined by the CCAMLR Ecosystem Monitoring Program (CEMP) Standard Method A3 (CCAMLR, 2004) (indicated by a large circle in Figure 1). The incubation period then follows with a single adult on each nest, during which time females return from foraging and relieve the males from their incubating duties, who in turn leave the breeding sites to forage. Nest failure and departure of attendant adults may continue through the incubation period, resulting in a decline in both adults and occupied nests. Between late-December when

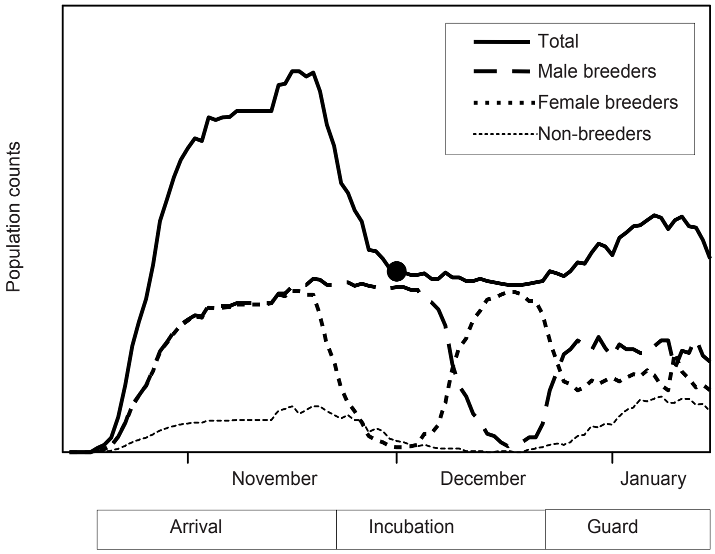

Figure 1: Population dynamics of Adélie penguins at a breeding site based on Béchervaise Island nest census and colony-wide counts in the 2002/03 breeding season. The circle indicates the time within the incubation period when a single adult is at each nest: this is about a week after peak clutch initiation.

eggs begin to hatch and mid-January when chicks become independent of the nest site, the chicks are alternately guarded by a single parent while the other undertakes a short foraging trip. During this guard phase, many non-breeding birds return to breeding sites (the re-occupation period) (Spurr, 1977; Ainley, 2002). From mid-January onwards, the number of adults and occupied nests declines as adults spend more time foraging to provision their chicks and for self-maintenance (Clarke et al., 2006), and chicks become independent of nest sites.

Time-series counts

Availability adjustment factors for correcting one-off counts of adults, nests or chicks can be obtained by utilising a time series of counts across a breeding season for each of those population objects. In this study, time-series counts of adults and nests were obtained using automated cameras (Newbery and Southwell, 2009), and time-series counts of chicks were obtained by direct observation.

Cameras were deployed approximately 20–50 m from the edge of colonies of breeding Adélie penguins and oriented to take in a fixed view of around 30–50 nest sites within the colonies. Ideally, cameras should be located at random locations with respect to sub-colonies, but for practical reasons consideration was given to topography (a slightly elevated position provided a better view of individual nests) and direction of view (wind direction can affect snow build-up and the orientation of nesting birds: a direction into the prevailing wind was preferred) when locating cameras. Data are presented here from three cameras operating at Béchervaise Island (67°35'S 62°49'E) near Mawson station during the 2007/08 breeding season. The cameras were programmed to take a single photo at solar midday each day from 10 October 2007 to 31 January 2008. Time-series counts of adult penguins and apparently occupied nests (defined as a nest structure of small rocks occupied by a penguin, either lying or standing; hereafter referred to as AO nests) were made from the sequence of images obtained from each camera. No counts were made in areas where visibility was limited by the presence of rocks, or in areas too distant for reasonable viewing. Purpose-built software written in Java was developed to facilitate rapid viewing of images, overlaying of boundaries and recording of counts.

While accurate time-series counts of chicks can be obtained from camera images for a restricted period when chicks are at their nests and are large enough not to be obscured by the attending adult, the reliability of counts from the fixed cameras is likely to decline after chicks become independent of the nest and move more widely over the breeding site. Time-series counts of chicks were therefore obtained by an observer who was free to move around the entire breeding site to maximise the period for obtaining reliable time-series counts. Chick counts were made in accordance with CEMP Standard Method A6a (CCAMLR, 2004), which specifies that chick counts be undertaken 'on the same day each year…when about two-thirds of chicks have entered crèches' in the same colonies used to assess breeding population size with CEMP Standard Method A3 (see below). Chick count data are presented here for the same site (Béchervaise Island) and breeding season (2007/08) as adult and AO nest data. In 2007/08, chick counts were undertaken for all sub-colonies on Béchervaise Island around the time of two-thirds crèching, and additional counts were made at regular intervals prior to and after this time. A total of 12 island-wide chick counts were made between 9 January and 16 February 2008.

Standardising time-series counts using a reference point count

In order for time-series counts to be used to adjust one-off counts for availability, they need to be transformed to a time series of availability adjustment factors by standardising them to a common point in the breeding chronology that represents a known component of the population. This requires the identification of a specific stage of the breeding chronology for use as a reference point.

The reference point detailed within CEMP Standard Method A3 for estimating breeding population size was used to standardise time-series counts. Method A3 specifies that counts of occupied nests should be undertaken a week after the peak in egg laying (CCAMLR, 2004). The method does not specify whether the 'peak in egg laying' refers to the first egg or all eggs, but the difference between the first and second egg-lay is just a few days (Ainley, 2002). At this stage of the breeding chronology, all or most breeding females, as well as the non-breeders, have left the breeding site, leaving a single male occupying each nest, and nest loss is minimal. Consequently, a count of occupied nests at this point provides the closest practical representation of the breeding population. Direct determination of the peak in either first or all egg laying requires detailed observation of individual nests over long periods of time, which was not undertaken at Béchervaise Island in 2007/08. However, as female Adélie penguins leave the breeding site to forage very shortly after laying their second egg (Penney, 1968), the peak in laying all eggs will coincide closely with the maximum rate of decline of adult numbers following their initial arrival. This point was identified by eye from the time series of adult counts from each camera, and the count of AO nests a week later was used as the reference point against which adult and AO nest time-series counts were standardised.

Camera counts of AO nests could not be used as a reference point for island-wide chick counts because of their restricted spatial coverage. However, CEMP Standard Method A3 (see above) was applied across Béchervaise Island as part of the ongoing CEMP program. This count was used as a reference point for chick time-series counts.

Modelling availability adjustment factor time series

In order to adjust one-off counts of adults, AO nests or chicks it is necessary to utilise a model of the availability of these population objects through time relative to an appropriate reference count. Modelling the average availability smoothes out small-scale irregularities in the series and allows interpolation across periods when no counts were available (e.g. no photographic counts due to blizzard events), and to provide a formal estimate of uncertainty around the average function. Since availability curves can show considerable variation, a flexible yet parsimonious modelling approach was needed. For this purpose, time series of availability adjustment factors were modelled using a generalised additive model (GAM) framework (Hastie and Tibshirani, 1990), using the penalised regression spline implementation of Wood (2006) in which the degree of smoothness is automatically selected by minimising a generalised cross-validation (GCV) criterion. All analyses were conducted in the R (v2.10.0) language for statistical computing (R Development Core Team, 2009) using the *mgcv* package of Wood (2006). Overfitting is occasionally observed when using GCV and this was controlled by increasing the smoothing parameter, γ, to 1.4 as recommended by Kim and Gu (2004) and Wood (2006).

Separate models were developed for each type of count object (adults, AO nests or chicks). The linear predictor for each model was taken to be the availability factor $\hat{p}_a$ as a smooth function in time (t, days since 1 October), where $\hat{p}_a$ is defined as the time-series counts of a population object ($\hat{n}_t$) divided by the reference count of AO nests on the date one week after the peak in egg laying ($\hat{n}_{t_{adj}}$):

$$\hat{p}_a = \frac{\hat{n}_t}{\hat{n}_{t_{adj}}}.$$

Thus, a count of a population object at any date within a breeding season ($\hat{C}_{t=x}$) can be adjusted to the predicted number of AO nests ($\hat{N}$) present one week after the peak in egg laying by:

$$\hat{N} = \frac{\hat{C}_{t=x}}{\hat{p}_{a_{t=x}}}.$$

Results

The cameras provided clear images on 93% of camera-days. Instances when images were unsuitable for counting corresponded with the occurrence of blizzards, when blowing snow obscured all or most penguins and nests in the field of view. The longest gap in a time series of counts from any camera was three days. Outside these instances, it was possible to count adults in all images, although careful attention to counting was required during the first peak in attendance when some penguins can partially or largely obscure others from view. All adult time series showed an expected bimodal attendance profile (Figure 2). Maximum adult counts within the areas delineated as 'countable' at the first peak in attendance were 114, 96 and 135 for cameras 1–3 respectively. The dates of maximum decline in adults for camera locations 1–3 were 22, 20 and 22 November respectively (left-hand vertical dashed lines, Figure 2), and consequently the reference dates for adjusting counts were 29, 27 and 29 November (right-hand dashed vertical line, Figure 2). Adult availability adjustment factors (i.e. the ratio of adults to AO nests) at the reference dates ranged from 1.09–1.14, indicating the presence of a variable, but generally small, number of adults additional to the males occupying nests one week after the peak in egg laying. The reference point consistently occurred at or just after a transition in the slope of the adult time series (Figure 2), which would correspond closely to the time when egg laying was completed and all breeding females had left.

Accurate counts of AO nests were not possible during the first peak in attendance because most nests were obscured from view by attending adults. The number of AO nests declined slightly in the time between the reference date and late December, and thereafter declined quickly (Figure 3). AO nest counts on the reference date were 58, 51 and 62 for cameras 1–3 respectively. AO nest adjustment factors are by definition always unity at the reference point.

Direct counts of chicks were made on 12 days over the 39-day period from 9 January to 16 February 2008. The longest gap in the time series of counts was six days. Island-wide chick counts declined from 637 on 9 January to 39 on 16 February 2008 (Figure 4). The island-wide AO nest count on 30 November 2007 was 1 959.

Figures 2–4 show fitted models (solid line) and approximate 95% confidence intervals (CIs) (shaded bands) arising from GAM fits to availability adjustment factors. Fitted models for adults required 19.0, 21.5 and 21.7 effective degrees of freedom (e.d.f.), and accounted for 99, 98 and 99% of the deviance, for cameras 1–3 respectively. Fewer degrees of freedom were required for models of AO nests, which required 11.7, 10.8 and 6.5 e.d.f., and accounted for 99, 98 and 98% of the deviance, for cameras 1–3. The chick attendance curve was less complex than either the adult or AO nest curves, requiring only 3 e.d.f. and accounting for 99% of the deviance. Residual analysis from all models considered indicated a Gaussian error model to be satisfactory, and examination of the autocorrelation function of residuals showed no suggestion of serial dependence.

Discussion

The development of remotely operating cameras has opened up new possibilities for the widespread collection of availability adjustment data. Previously, this has required repeated visits by observers to breeding sites throughout the breeding period, or observers taking up residence at sites across the breeding period, in order to obtain time-series counts. However, travelling between sites in Antarctica can be logistically very difficult or impossible later in the breeding period after stable sea-ice disappears, and relying on resident observers at breeding sites limits data collection to the number of observers available. Once deployed, the cameras can operate independently over several years without the need for servicing or maintenance, and images can be downloaded at any time of the year when access is easiest. Further deployment of cameras at other sites and regions around Antarctica would allow improved assessment of the extent and nature of spatial variability in adjustment data, and continued operation over time will indicate the extent of temporal variability.

Knowledge of the magnitude of spatial and temporal variability in availability is important when accounting for uncertainty in broad-scale estimates of breeding abundance. While one-off counts may be obtainable over many breeding sites in any future broad-scale surveys, time-series counts for availability data will invariably be restricted to a sample of these sites, even if cameras are used. In the absence of data on availability at any site

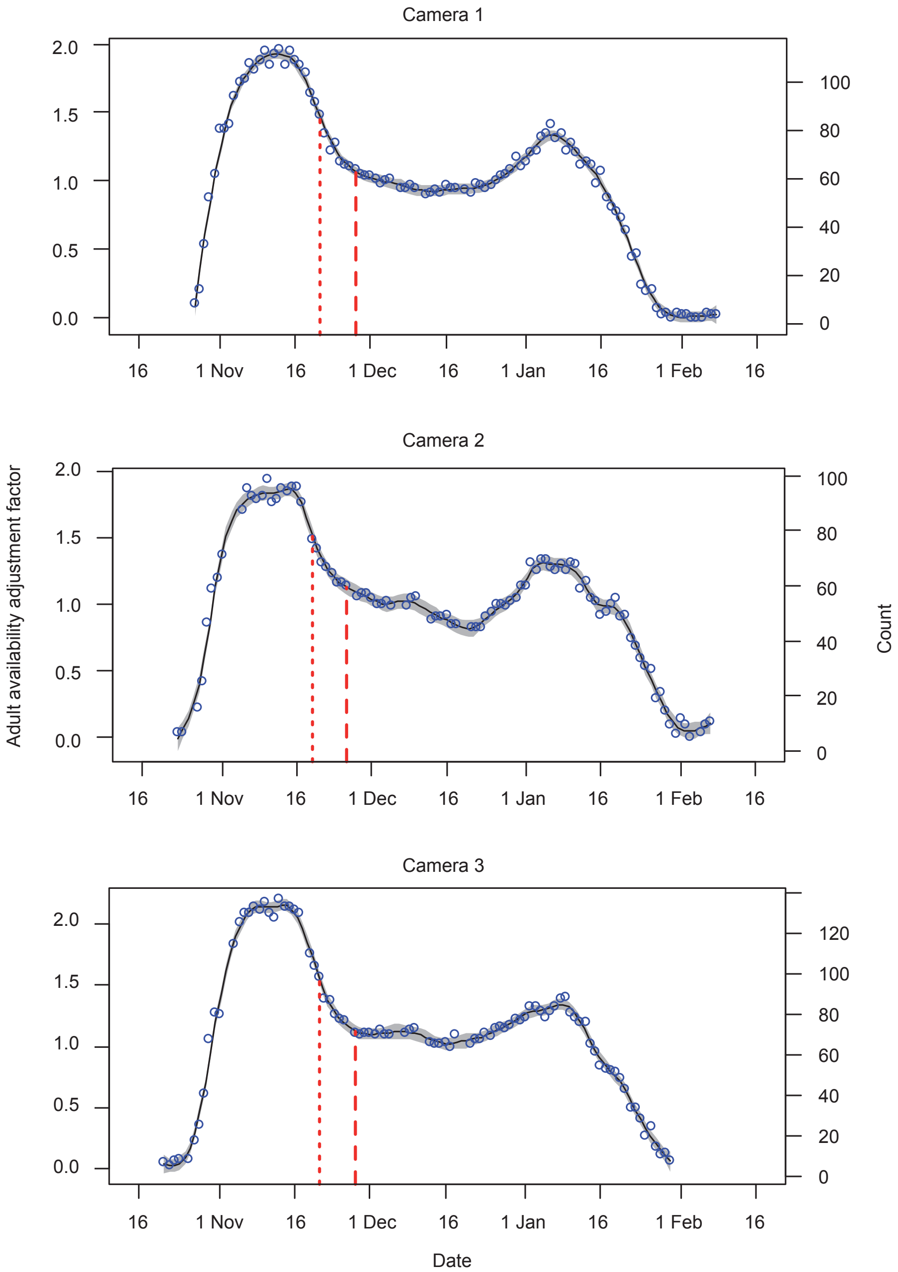

Figure 2: Time-series counts and modelled time series (showing approximate 95% CI as shaded bands) of availability adjustment factors ($\hat{p}_a$) for adults from images taken by three remotely operating cameras on Béchervaise Island in the Mawson region during the 2007/08 breeding season. The left vertical dashed line indicates the date of maximum rate of decline in adults (determined by eye), taken as the approximate time of peak egg-lay. The right vertical dashed line indicates the date a week after the maximum decline, which was used as the reference point to standardise counts ($n_{t_{adj}}$).

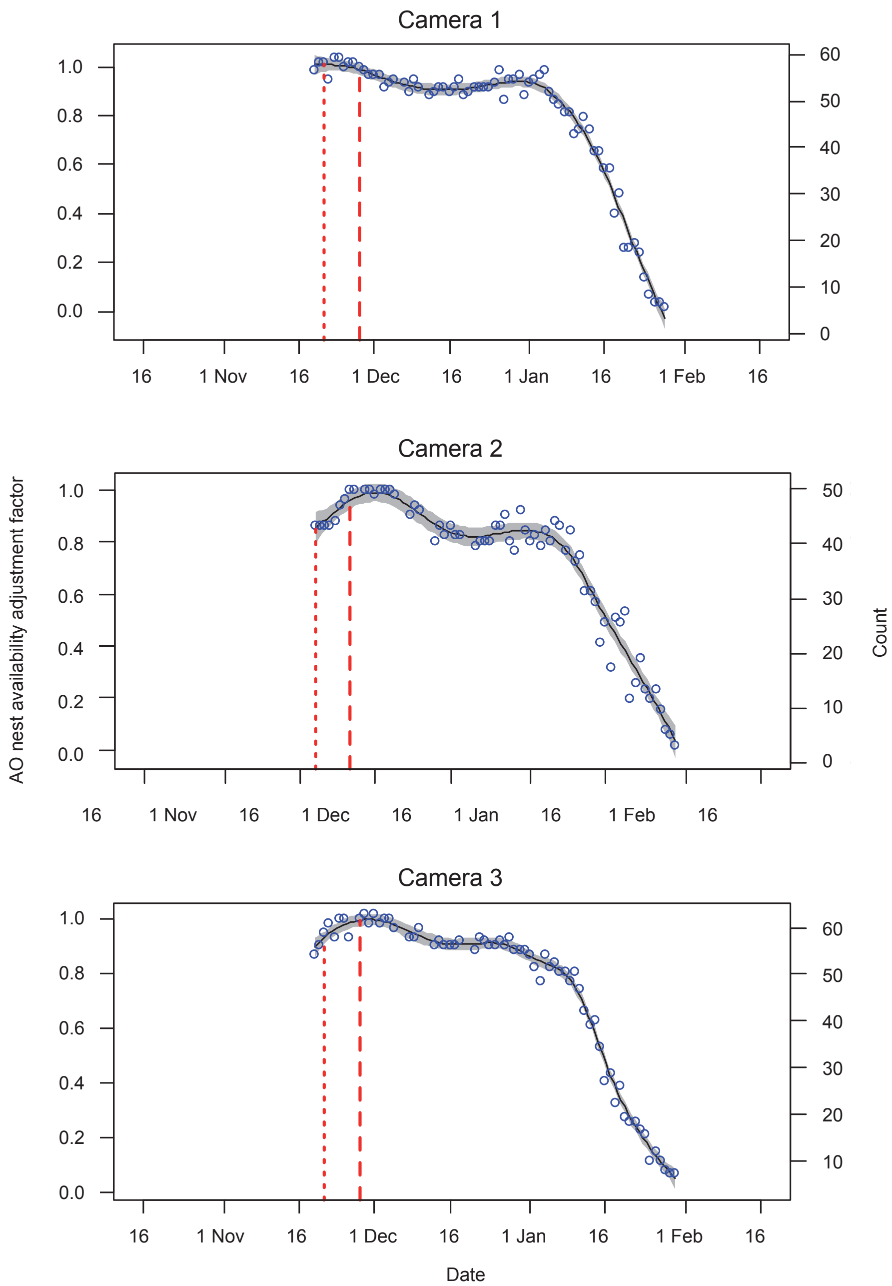

Figure 3: Time-series counts and modelled time series (showing approximate 95% CI as shaded bands) of availability adjustment factors ($\hat{p}_a$) for AO nests from images taken by three remotely operating cameras on Béchervaise Island in the Mawson region during the 2007/08 breeding season. The left vertical dashed line indicates the date of maximum rate of decline in adults (determined by eye), taken as the approximate time of peak egg-lay. The right vertical dashed line indicates the date a week after the maximum decline in adults, which was used as the reference point to standardise counts ($n_{t_{adj}}$).

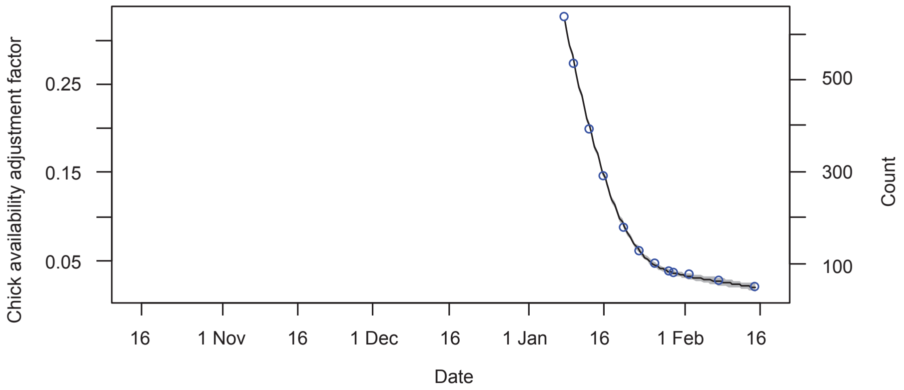

Figure 4: Time-series counts and modelled time series (showing approximate 95% CI as shaded bands) of availability adjustment factors ($\hat{p}_a$) for chicks from direct counts on Béchervaise Island in the Mawson region during the 2007/08 breeding season. A count of AO nests on 30 November 2007 was used as the reference point to standardise counts ($n_{t_{adj}}$).

where a one-off count has been made, it would be precautionary to apply adjustment data from all those sites in the region with availability data. If spatial variation in availability adjustment data exists, there will be some inferential cost or uncertainty associated with the application of multiple adjustment factors. For the same reason, temporal variability would also lead to some uncertainty in adjusting counts if the adjustment data are collected for years other than the time of the count data, as will be the case when interpreting most historical count data. Thus, while this paper only addresses the development of adjustment factors for a single site and year, it is expected that in most situations adjustment data from several sites and/or years would be applied to any one-off count datum. The bootstrapping method presented in McKinlay et al. (2010) quantifies the uncertainty that is propagated into the abundance estimate due to this process. As the amount of adjustment data increases across space and time in the future, it may be possible to develop predictive models relating variation in adjustment time series to environmental covariates (e.g. Lynch et al., 2009), and these models might be used to minimise the resulting uncertainty in abundance estimates.

Our application of the concept of developing adjustment factors to take account of availability bias in relation to chick counts includes both chick mortality at the breeding site before fledging and chicks leaving the breeding site at the time of fledging. While including chick mortality may not strictly be in the spirit that the concept of availability was originally developed, its inclusion is a practical means of accounting for overall declines in the number of chicks produced at the beginning of the breeding season. This is important because adjustment factors have been used to determine the number of breeding pairs from chick counts. For the reasons discussed above, in most cases, application of chick availability factors to adjust one-off count data would use chick availability data from several years, and because chick mortality is highly variable from year to year (Ainley, 2002), the associated uncertainty would be much larger than from the single year shown in Figure 4.

The potential for bias in adult time-series counts during the first peak in attendance due to some penguins partially or largely obscuring others can be minimised or eliminated if cameras are located at an elevated vantage point. A zoom facility in the purpose-built Java software used for counting penguins from photographs also helped to minimise any bias. Time-series counts from photographs were strongly coincident with island-wide manual time-series counts by a free-ranging observer, suggesting that any bias in camera counts during peak attendance was trivial.

The use of manual survey methods for obtaining chick time-series counts recognised the potential limitation in using fixed cameras for this purpose, given that chicks become independent of nest sites and move around the breeding site from mid-January onwards. Nevertheless, further investigation of the utility of cameras for obtaining chick counts is warranted. It is possible, for instance, that reliable counts might be obtained if multiple images are taken each day during the time when chicks are independent of the nest. This is easily achieved by re-programming the cameras and is a focus of current work at Béchervaise Island. In the absence of automated data collection, the continued application of manual counts at Béchervaise Island and other CEMP sites will provide a valuable source of data for adjustment of chick count data.

The complexity of availability curves differed markedly between the three types of penguin count objects (adults, AO nests and chicks). Across a large range of availability datasets, in addition to those presented here, adult models consistently required greatest complexity (typically 20–25 e.d.f.), followed by nest models (8–11 e.d.f.), with chick models requiring the least (6 or fewer e.d.f.). Due to the relatively long nature of adult and AO nest time series derived from automated cameras (typically greater than 90 and 50 time points respectively), the sometimes complex shape of the attendance function for those count objects has been able to be adequately captured in all models considered so far. Similarly, the modelling approach presented has been able to successfully accommodate the regular 4–6 day gaps in chick time series, largely due to the simpler structure of the chick attendance function between crèche and fledging stages compared with adult and AO nest attendance functions.

The types of models presented naturally invites the question of whether there might be some minimum amount of data required before fitting attendance functions by this method. While clearly some lower limit will apply to the number of data points required for successfully capturing adjustment function shape, the answer will depend on the count object considered and the structure of missing values in the series. It is therefore difficult to provide prescriptive guidance, except perhaps to say that there should be sufficient data to effectively capture the shape of the function. To this end, expert knowledge, in the form of an *a priori* understanding of the general expected form of the function, is required. For example, it may be possible to fit a linear relationship to a small number of widely spaced data points, however, few would argue that a straight line might provide a reasonable representation of an attendance function over the course of a season. Normal principals of regression modelling should be observed, such as residual analysis and outlier investigation, and these may help inform when particular series are unsuitable for modelling in this way.

At present, the reference point for calculating adjustment factors for adults is chosen as occurring seven days after the date judged by eye to correspond to the peak in egg laying (i.e. seven days after the date with maximum rate of decline of adult numbers following their initial arrival). An automated, objective method for selection of this point of maximum rate of decline is desirable, but attempts to do so have so far been unsuccessful. One method examined involved calculating rates of decline over the key period to find a maximum, but two issues were immediately apparent. First, the timing of peak attendance may shift considerably depending on site-specific or seasonal considerations (Ensor and Bassett, 1987; Woehler et al., 1991), making it difficult to formulate a single automated method capable of operating over all space and time. Second, breeding failure at a site may result in an uninterrupted decline in adult numbers from the time of peak attendance, making determining the point of maximum decline ambiguous. While work to determine an objective method for automatically determining adjustment dates continues, for the present, expert opinion is the preferred method to judge the date of peak egg laying.

The approach to modelling availability curves adopted here treats subsamples (i.e. multiple locations measured at a single site in a season) as independent data, with separate models developed for each subsample. Other approaches are possible and may be preferred in some circumstances. For instance, subsamples could be modelled as random effects in a generalised additive mixed model of overall site attendance. For a single modelling exercise this would almost certainly be a preferred approach. However, the analysis presented here is consistent with allied work incorporating these models into a parametric bootstrap model for broad-scale adjustment and aggregation of multiple one-off counts from Eastern Antarctica. In that work, site-specific adjustment factors are determined by appropriately weighting subsamples in the bootstrap process (McKinlay et al., 2010). An alternate approach, previously considered for data of this type, is quantile regression incorporating smoothing splines (Koenker et al., 1994). Emmerson et al. (2004) used this method to model different regions of Antarctica by fitting a single model to multiple time series simultaneously by considering them as a single series. This approach proved an attractive one for capturing the shape and variability of multiple series, since it inherently captured non-constant variance over the modelled time-period. However, the degree of variance heterogeneity with time is closely linked to the inclusion of multiple series in a single model, and the issue of non-constant variance is naturally accommodated by fitting separate models and combining them in a weighted bootstrap setting.

Conclusion

The adjustment factors presented here allow estimates of Adélie penguin breeding populations to be made from counts of adults, AO nests or chicks at any time of the breeding season. Specifically, in this context the breeding population comprises penguins that were attempting to breed at the reference point of one week after the peak in egg laying. This does not include penguins that left the breeding site prior to the reference point because their breeding attempt failed, breeding-age penguins that declined to return to the breeding site in that breeding season, nor pre-breeding penguins. The abundance of these components of the population can only be estimated through alternative methods, such as mark-recapture analyses of detection histories collected over multiple years of individually tagged birds.

The methods proposed here should have general applicability to species with similar biology to the Adélie penguin, which has a highly constrained and synchronised breeding phenology. For species with lower synchrony, such as the gentoo penguin (*Pygoscelis papua*) (Croxall, 1984; Lynch et al., 2009), only a proportion of all birds breeding in a summer may be present at any one time. In this case, the methods outlined here will overestimate availability, and therefore underestimate breeding abundance, unless some auxiliary information, such as mark-recapture data, are available.

Acknowledgements

We thank Rhonda Pike and David Wilson for downloading and processing camera images and collecting CEMP A6a and A3 data, and Steve Candy for discussing statistical issues related to modelling adjustment time-series data. We would also like to thank Jaume Forcada, Heather Lynch and one anonymous reviewer for many helpful suggestions on an earlier draft of this manuscript.

References

Ainley, D. 2002. *The Adélie Penguin – Bellwether of Climate Change*. Columbia University Press, New York: 416 pp.

CCAMLR. 2004. *CCAMLR Ecosystem Monitoring Program: Standard Methods for Monitoring Studies*. CCAMLR, Hobart, Australia.

Clarke, J., L.M. Emmerson and P. Otahal. 2006. Environmental conditions and life history constraints determine foraging range in breeding Adélie penguins. *Mar. Ecol. Prog. Ser.*, 310: 247–261.

Croxall, J.P. 1984. Seabirds. In: Laws, R.M. (Ed.). *Antarctic Ecology*, 2. Academic Press, London: 533–619.

Croxall, J.P. and E.D. Kirkwood. 1979. The distribution of penguins on the Antarctic Peninsula and islands of the Scotia Sea. British Antarctic Survey, Cambridge: 186 pp.

Davis, L.S. 1988. Coordination of incubation routines and mate choice in Adélie penguins (*Pygoscelis adeliae*). *Auk*, 105: 428–432.

Emmerson, L.M., J. Clarke, K.R. Kerry and C. Southwell. 2003. Temporal variability and the interrelationships between CEMP parameters collected on Adélie penguins at Béchervaise Island. *CCAMLR Science*, 10: 75–90.

Emmerson, L., B. Raymond and C. Southwell. 2004. Modelling availability bias using existing time series count data: Adélie penguins as a case study. Document *WG-EMM-04/54*. CCAMLR, Hobart, Australia.

Ensor, P.H. and J.A. Bassett. 1987. The breeding status of Adélie penguins and other birds on the coast of George V Land, Antarctica. *ANARE Res. Notes*, 50. Australian Antarctic Division, Hobart, Australia.

Hastie, T.J. and R.J. Tibshirani. 1990. *Generalized Additive Models*. Chapman and Hall, London: 335 pp.

Kim, Y.-J. and C. Gu. 2004. Smoothing spline Gaussian regression: more scalable computation via efficient approximation. *J. Roy. Stat. Soc. B*, 66 (2): 337–356.

Koenker, R., P. Ng and S. Portnoy. 1994. Quantile smoothing splines. *Biometrika*, 81: 673–680.

Lynch, H.J., W.F. Fagan, R. Naveen, S.G. Trivelpiece and W.Z. Trivelpiece. 2009. Timing of clutch initiation in *Pygoscelis* penguins on the Antarctic Peninsula: towards an improved understanding of off-peak census correction factors. *CCAMLR Science*, 16: 149–165.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *J. Wildl. Manage.*, 53 (4): 1017–1024.

McKinlay, J., C. Southwell and R. Trebilco. 2010. Integrating Count Effort by Seasonally Correcting Animal Population Estimates (ICESCAPE): a method for estimating abundance and its uncertainty from count data using Adélie penguins as a case study. *CCAMLR Science*, this volume.

Newbery, K.B. and C. Southwell. 2009. An automated camera system for remote monitoring in polar environments. *Cold Reg. Sci. Technol*., 55 (1): 47–51.

Penney, R.L. 1968. Territorial and social behavior in the Adélie penguin. *Antarct. Res. Ser*., 12: 83–131.

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Version 2.10.0. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, www.R-project.org.

Southwell, C. 2004. Developing and applying a general abundance estimator for land-based predator surveys: Adélie penguins as a case study. Document *WG-EMM-04/56*. CCAMLR, Hobart, Australia.

Spurr, E.B. 1977. Adaptive significance of the reoccupation period of the Adélie penguin. In: Llano, G.A. (Ed.). *Proceedings of the Third SCAR Symposium on Antarctic Biology*. Washington, DC: 605–618.

Woehler, E.J., D.J. Slip, L.M. Robertson, P.J. Fullagar and H.R. Burton. 1991. The distribution, abundance and status of Adélie penguins *Pygoscelis adeliae* at the Windmill Islands, Wilkes Land, Antarctica. *Mar. Ornithol*., 19: 1–18.

Wood, S.N. 2006. *Generalized Additive Model: an Introduction with R*. Chapman and Hall, London: 391 pp.

Liste des figures

Figure 1: Dynamique de la population de manchots Adélie sur un site de reproduction, à partir du recensement des nids et des dénombrements à l'échelle de la colonie, à l'île Béchervaise pendant la saison de reproduction 2002/03. Le cercle indique le temps de présence au nid d'un seul adulte pendant la période d'incubation : soit environ une semaine après le début du pic de la couvée.

Figure 2: Série chronologique des dénombrements et série chronologique modélisée (montrant dans une bande grise un intervalle de confiance à environ 95%) des facteurs de correction ($\hat{p}_a$) en fonction de la disponibilité des adultes, à partir d'images prises par trois caméras télécommandées sur l'île Béchervaise dans la région de Mawson pendant la saison de reproduction 2007/08. Le trait vertical de gauche en pointillés indique la date du taux maximal de déclin des adultes (déterminé *de visu*), correspondant approximativement au pic de la ponte. Le trait vertical de droite en pointillés indique la date, une semaine après le déclin maximal, ayant servi de point de référence pour normaliser les dénombrements ($n_{t_{adj}}$).

Figure 3: Série chronologique des décomptes et série chronologique modélisée (montrant dans une bande grise un intervalle de confiance à environ 95%) des facteurs de correction ($\hat{p}_a$) des nids apparemment occupés, à partir d'images prises par trois caméras télécommandées sur l'île Béchervaise dans la région de Mawson pendant la saison de reproduction 2007/08. Le trait vertical de gauche en pointillés indique la date du taux maximal de déclin des adultes (déterminé *de visu*), correspondant approximativement au pic de la ponte. Le trait vertical de droite en pointillés indique la date, une semaine après le déclin maximal des adultes, ayant servi de point de référence pour normaliser les dénombrements ($n_{t_{adj}}$).

Figure 4: Série chronologique des décomptes et série chronologique modélisée (montrant dans une bande grise un intervalle de confiance à environ 95%) des facteurs de correction ($\hat{p}_a$) des jeunes, à partir de dénombrements directs sur l'île Béchervaise, dans la région de Mawson, pendant la saison de reproduction 2007/08. Un décompte des nids apparemment occupés, daté du 30 novembre 2007, a servi de point de référence pour normaliser les dénombrements ($n_{t_{adj}}$).

Список рисунков

Рис. 1: Динамика популяции пингвинов Адели на участке размножения по данным учета гнезд и подсчетов по всей колонии на о-ве Бешервез в сезоне размножения 2002/03 г. Кружок показывает время в течение инкубационного периода, когда в каждом гнезде находится одна взрослая птица: это происходит примерно через неделю после начала пикового периода кладки.

Рис. 2: Временные ряды подсчетов и смоделированные временные ряды (заштрихованными полосками показаны приблизительные 95% ДИ) корректировочных коэффициентов наличия ($\hat{p}_a$) взрослых птиц по фотографиям, полученным с помощью трех дистанционно управляемых камер на о-ве Бешервез в районе станции Моусон в сезоне размножения 2007/08 г. Левая вертикальная пунктирная линия показывает дату максимальных темпов сокращения численности взрослых птиц (определено на глаз), происходящего приблизительно в пиковый период откладывания яиц. Правая вертикальная пунктирная линия показывает дату через неделю после максимального сокращения численности, которая использовалась как контрольное значение для стандартизации подсчетов ($n_{t_{adj}}$).

Рис. 3: Временные ряды подсчетов и смоделированные временные ряды (заштрихованными полосками показаны приблизительные 95% ДИ) корректировочных коэффициентов наличия ($\hat{p}_a$) явно занятых гнезд по фотографиям, полученным с помощью трех дистанционно управляемых камер на о-ве Бешервез в районе станции Моусон в сезоне размножения 2007/08 г. Левая вертикальная пунктирная линия показывает дату максимальных темпов сокращения численности взрослых птиц (определено на глаз), происходящего приблизительно в пиковый период откладывания яиц. Правая вертикальная пунктирная линия показывает дату через неделю после максимального сокращения численности, которая использовалась как контрольное значение для стандартизации подсчетов ($n_{t_{adj}}$).

Рис. 4: Временные ряды подсчетов и смоделированные временные ряды (заштрихованными полосками показаны приблизительные 95% ДИ) корректировочных коэффициентов наличия ($\hat{p}_a$) птенцов по непосредственным подсчетам на о-ве Бешервез в районе станции Моусон в сезон размножения 2007/08 г. Подсчет явно занятых гнезд 30 ноября 2007 г. использовался как контрольное значение для стандартизации подсчетов ($n_{t_{adj}}$).

Lista de las figuras

Figura 1: Dinámica de la población de pingüinos adelia en una colonia de reproducción, en base al recuento de nidos y de toda la colonia realizado en Isla Béchervaise durante la temporada de reproducción 2002/03. El círculo indica la fecha del período de incubación cuando sólo se encuentra un adulto en cada nido: esto ocurre aproximadamente una semana después de la fecha de máxima postura de huevos.

Figura 2: Serie cronológica de recuentos y serie cronológica modelada (mostrando el IC del 95% en bandas sombreadas) de los factores de ajuste del sesgo por disponibilidad ($\hat{p}_a$) para los adultos, de imágenes tomadas por tres cámaras operadas con control remoto en Isla Béchervaise de la región de Mawson durante la temporada de reproducción 2007/08. La línea entrecortada vertical a la izquierda indica la fecha cuando la tasa de disminución de adultos fue máxima (determinada al ojo), que se toma como la fecha aproximada de la máxima postura de huevos. La línea entrecortada vertical a la derecha indica la fecha una semana después de la máxima disminución de adultos, que se utiliza como punto de referencia para normalizar los recuentos ($n_{t_{adj}}$).

Figura 3: Serie cronológica de recuentos y serie cronológica modelada (mostrando el IC del 95% en bandas sombreadas) de los factores de ajuste del sesgo por disponibilidad ($\hat{p}_a$) para nidos AO, de imágenes tomadas por tres cámaras operadas por control remoto en Isla Béchervaise de la región de Mawson durante la temporada de reproducción 2007/08. La línea entrecortada vertical a la izquierda indica la fecha cuando la tasa de disminución de adultos fue máxima (determinada al ojo), que se toma como la fecha aproximada de la máxima postura de huevos. La línea entrecortada vertical a la derecha indica la fecha una semana después de la máxima disminución de adultos, que se utiliza como punto de referencia para normalizar los recuentos ($n_{t_{adj}}$).

Figura 4: Serie cronológica de recuentos y serie cronológica modelada (mostrando el IC del 95% en bandas sombreadas) de los factores de ajuste del sesgo por disponibilidad ($\hat{p}_a$) para los polluelos, de recuentos directos efectuados en Isla Béchervaise de la región de Mawson durante la temporada de reproducción 2007/08. Se utilizó el recuento de nidos AO efectuado el 30 de noviembre de 2007 como punto de referencia para normalizar los recuentos ($n_{t_{adj}}$).