Quantifying the causes and consequences of variation in satellite-derived population indices: a case study of emperor penguins

Authors: Sara Labrousse[1,2], David Iles[2,3], Lise Viollat[2], Peter Fretwell[4], Philip N. Trathan[4], Daniel P. Zitterbart[2,5], Stephanie Jenouvrier[2], Michelle LaRue[1,6]

1. Department of Earth and Environmental Sciences, University of Minnesota, Minneapolis, MN USA
2. Woods Hole Oceanographic Institution, Woods Hole, MA, USA
3. Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ON, CAN
4. British Antarctic Survey, Cambridge, UK
5. Department of Physics, Friedrich-Alexander University of Erlangen-Nürnberg, Erlangen, Germany
6. School of Earth and Environment, University of Canterbury, Christchurch, NZ

Corresponding author: Sara Labrousse, Mail Stop #50, Woods Hole Oceanographic Institution, 266 Woods Hole Road, Woods Hole, MA 02543-1050 U.S.A., sara.labrousse@gmail.com

Keywords: emperor penguin; intra-seasonal variability; population trend; satellite imagery.

Abstract

Very high-resolution satellite (VHR) imagery is a promising tool for estimating the abundance of wildlife populations, especially in remote regions where traditional surveys are limited by logistical challenges. Emperor penguins *(Aptenodytes forsteri*) were the first species to have a circumpolar population estimate derived via VHR imagery. Here we address an untested assumption from Fretwell et al. (2012) that a single image of an emperor penguin colony is a reasonable representation of the colony for the year the image was taken. We evaluated satellite-related and environmental variables that might influence the calculated area of penguin pixels to reduce uncertainties in satellite-based estimates of emperor penguin populations in the future. We focused our analysis on multiple VHR images from three representative colonies: Atka Bay, Stancomb-Wills (Weddell Sea sector) and Coulman Island (Ross Sea sector) between September and December during 2011. We replicated methods in Fretwell et *al.* (2012), which included using supervised classification tools in

ArcGIS 10.7 software to calculate area occupied by penguins (hereafter referred to as "population indices") in each image. We found that population indices varied from 2 to nearly 6-fold, suggesting that penguin pixel areas calculated from a single image may not provide a complete understanding of colony size for that year. Thus, we further highlight the important roles of: i) sun azimuth and elevation through image resolution and ii) penguin patchiness (aggregated versus distributed) on the calculated areas. We found an effect of wind and temperature on penguin patchiness. Despite intra-seasonal variability in population indices, simulations indicate that reliable, robust population trends are possible by including satellite-related and environmental covariates and aggregating indices across time and space. Our work provides additional parameters that should be included in future models of population size for emperor penguins.

Introduction

Very high-resolution (VHR; 0.3-0.6m spatial resolution) satellite imagery has been a disruptive technology for studying wildlife populations, especially in Antarctica (LaRue *et al.*, 2011; Fretwell *et al.*, 2012; Lynch and LaRue, 2014; McMahon *et al.*, 2014; Strycker *et al.*, 2020; Wege, Salas and LaRue, 2020). Emperor penguins *(Aptenodytes forsteri*), icons of the Antarctic, are a model species for direct, satellite-based investigation of their distribution and numbers: they leave a representative guano stain on the fast ice (i.e., sea ice fastened to the coastline) that indicates colony presence (Barber-Meyer, Kooyman and Ponganis, 2007; Fretwell and Trathan, 2009; Fretwell *et al.*, 2012); they are available for detection in austral spring when satellite images of the coastline are easily acquired; and good contrast (black penguins on white snow), makes their enumeration straight-forward.

Emperor penguins were the first species to have a circumpolar population estimate derived via VHR imagery (Fretwell *et al.*, 2012). Most emperor penguin colonies are difficult to access due to their location on remote sections of Antarctic fast ice, and very few of the 66 known colonies (Fretwell and Trathan, 2020) are available to survey using ground counts or aerial surveys (Ancel, Gendner, *et al.*, 1992; Barbraud and Weimerskirch, 2001; Kooyman and Ponganis, 2017; Richter *et al.*, 2018a). However, gaining empirical understanding of population change at multiple spatial scales is critical, as modeling studies suggest that most breeding colonies will be quasi-extinct by 2100 under 'business as usual' emissions scenarios (Jenouvrier *et al.*, 2014, 2020), resulting in dramatic declines in the global population size, even under optimistic dispersal scenarios (Jenouvrier et al. 2017). The ability to apply the baseline population provided by Fretwell *et al.* (2012) to monitor population trends will

improve our understanding and predictions of emperor penguin populations at multiple spatial scales, which is critical for conservation (Trathan *et al.*, 2020).

Emperor penguins breed on fast ice during total darkness in the winter when reproductive birds gather at the colony to mate, and raise and feed their chicks (Ancel, Kooyman, *et al.*, 1992; Kirkwood and Robertson, 1997). Strong winds (>130 km/h) combined with low temperatures (<40° C) favor huddling behavior of the males (Gilbert *et al.*, 2007) during incubation, and also to keep chicks warm through the winter and into the spring. Thus, the ideal time to estimate abundance of emperor penguins would be during austral winter, when only males are present at the colony, making enumeration straight-forward (counting of males in the huddle represents the number of breeding pairs). However, optical VHR imagery of the Antarctic coastline is only available between September and March, and emperor penguins spend January through April foraging away from their colonies. Thus, the only period when emperor penguin abundance can be estimated from VHR imagery is austral spring, during chick-rearing.

Furthermore, satellite-based estimates of emperor penguins during spring may be influenced by factors related directly to penguin behavior and by features of the satellite platform itself (i.e., the observation process). Breeding failure and foraging trips by adult penguins introduces variation into the number of birds available for detection by the satellite sensor at a colony (see an analogous discussion of this issue for surveys on King Penguins [*Aptenodytes patagonicus*] in Foley et *al.* (2020). Additionally, huddling behavior fluctuates during chick-rearing period and can introduce variation into satellite-based counts (Richter *et al.*, 2018b), particularly if birds are so densely huddled that the ability to distinguish individual birds becomes difficult (i.e., because multiple birds can potentially fit within a single VHR pixel). Additional variation in satellite-derived counts could be introduced by imprecision in the supervised classification, or by differences in the quality of images among successive counts (i.e., owing to differences in spatial resolution or sun angle). Given the remoteness of most emperor penguin colonies, satellite-based monitoring of population trends is currently the only viable method for monitoring this species across the species range and could play a central role in determining its conservation status. Thus, generating precise indices of annual abundance at individual colonies, and in turn, estimates of population trends, could heavily depend upon an ability to remove this “noise” in satellite-derived indices (i.e., observation error that is caused by the within-season huddling behavior, satellite-related covariates, or other factors described above). Conversely, an inability to sufficiently remove spurious observation error at individual emperor penguin colonies would suggest that either colonies must be monitored for many years to derive reliable trend

estimates, or that satellite-based monitoring will only be useful for estimating regional population trends in the short term (i.e., where observation error will "average out" across many colonies).

Here, we addressed an untested assumption from Fretwell et *al.* (2012) that a single VHR image of an emperor penguin colony would reasonably represent colony size for that year (calculated as number of breeding pairs; Fretwell et *al.*, 2012). Specifically, we aimed to understand satellite-related and environmental variables that might influence the calculated area of penguin pixels (hereafter referred to as the population indices) to reduce potential uncertainties associated with using only one image per year to assess colony size. We use the term "population indices" to refer to the calculated area of penguin pixels from each VHR image because the penguins available for detection on each image are a benchmark for colony status in that year. While our goals are not to complete the process of estimating populations, it is critical we test the representativeness of population indices calculated from a single VHR image because we already know that only one image per colony per year is available over the course of ~10 years; see available imagery via Maxar Technologies: discover.digitalglobe.com).

Finally, we conducted a series of simulations to evaluate the potential for covariates to improve estimates of population trends at a range of scales (i.e., from local populations to regional aggregations) and across different time horizons.

We hypothesized:

1. Satellite platform, e.g. spatial resolution of the panchromatic band will influence the area occupied by penguins in each image (i.e., population index) calculated from VHR images (i.e., lower resolution imagery will result in greater area of penguins, which could be interpreted as a higher population index);
2. Sun elevation angle and sun azimuth will influence the population index (i.e., lower sun elevation will cast more shadows resulting greater area of penguins; and sun azimuth could result in shadows being cast from surrounding features like ice cliffs would obscure penguins). Moreover, sun elevation is correlated with the day of the year and may integrate seasonal changes in penguin movements.
3. The spatial patchiness of penguins within a colony during a satellite survey (i.e. compactly-huddled versus widely-spread) will influence the population index given variation in density of birds; areas calculated from compact aggregations will be smaller than areas calculated from spread aggregations of birds.

4. Wind speed and temperature during the satellite survey will influence the population index, owing to the huddling behavior of emperor penguins during cold/windy conditions, which would result in compact groups that may lead to smaller population indices.
5. Population trends can be estimated more precisely at the colony level, and with fewer years of monitoring if these sources of spurious variation in counts are accounted for and removed. This hypothesis was tested using simulations to show how we improve population trends with those sources of variation; however, the translation of population indices (i.e., area of penguin pixel) to population size is not the goal of this research.

Materials and Methods

Study area

We focused our examination of variance in population indices (as calculated by area of penguin pixels on VHR images) on three emperor penguin colonies: the Stancomb-Wills (~5,455 breeding pairs) and Atka Bay (~9657 breeding pairs) colonies in the Weddell Sea sector, and the Coulman Island colony (~25,000 breeding pairs) in the Ross Sea sector (Fig. 1; Fretwell et al. 2012). These three colonies were chosen because they are each larger-than-average (average colony size in 2009 was ~4,300 breeding pairs; Fretwell et *al.*, 2012), they have been monitored by aerial or ground surveys on several occasions, are relatively stable in their annual occupancy, and were also unlikely to be impacted by confounding factors such as proximity to research stations, tourism, or pollution. Both the Weddell Sea and Ross Sea are characterized by wide bathymetric continental slopes, relatively cold waters, high primary productivity (particularly in the case of the Ross Sea, which is home to the largest open-ocean polynya in the Southern Ocean; Smith *et al.* (2014)), relatively stable sea ice regimes, and finally, both regions are likely to be refugia for emperor penguins in the future (Jenouvrier *et al.*, 2020). In other words, these colonies represent locations where human-induced variation is likely to be minimal, but where natural, intra-seasonal variation may be relatively high given the most-recent colony estimates (in number of breeding pairs of adults; Fretwell et *al.*, 2012; Kooyman and Ponganis 2017). Further, the colonies were sufficiently large, increasing the probability that any intra-seasonal changes could be detected. Changes or error in the estimation at small colonies are less consequential in understanding overall population status. In other words,

158 substantial intra-season fluctuations at large colonies are more consequential to estimating
159 populations than changes at smaller colonies.
160
161 ***VHR imagery and image processing***
162 We selected high-quality (i.e. cloud-free, no banding; Barber-Meyer et *al.*, 2007) VHR images
163 acquired for each of the three study colonies during spring 2011 (September through December), the
164 year with the highest number of repeat images acquired by DigitalGlobe, Inc. (now Maxar
165 Technologies) around the Antarctic coastline. Indeed, other than 2011, there are ~5 images at any
166 colony and in most cases there is one only useable image per colony. Images were primarily from
167 WorldView-2 (~0.46m panchromatic spatial resolution) and QuickBird-2 (~0.65m panchromatic
168 spatial resolution) satellites and were processed (e.g., pansharpened, orthorectified, and projected
169 to Antarctic Polar Stereographic) by the Polar Geospatial Center (PGC) at the University of Minnesota
170 (processing code on GitHub: https://github.com/PolarGeospatialCenter/).
171 To gain a population index of emperor penguins for each image and to test the assumption of
172 the representativeness of a single image per colony per year, we replicated methods first outlined in
173 Barber-Meyer et *al.* (2007) and built upon in Fretwell et *al.* (2012). Briefly, these methods involved
174 using ArcGIS software to first clip the image to our area of interest (the colony; Fig. 2A) and then
175 define three training classes (using a point shapefile with attribute classes of penguin, guano, and
176 snow, Fig. 2B-C; Barber-Meyer et *al.,* 2007) for a supervised classification on pansharpened images
177 of Antarctic fast ice. Notably, field tests of emperor penguin reflectance from satellite imagery have
178 not been conducted, let alone for various environmental scenarios (light cloud cover vs sunny
179 conditions) and therefore time-consuming, human interpretation was required in every step of the
180 process to ensure accuracy.
181 Once the training dataset was compiled, we then conducted a maximum likelihood
182 classification resulting in an output raster, which we converted to a polygon shapefile. Within the
183 polygon shapefile, we extracted only the penguin class (based on the grid value, which was defined
184 as aforementioned) since we were not interested in the amount of area of guano or snow (Fig. 2D-
185 E). Because of the simplicity of the maximum likelihood classifier, to ensure accuracy of results, and
186 to maintain one aim of Fretwell et *al.* (2012), which was to ensure this work could occur in fairly
187 accessible GIS software (e.g., ArcGIS rather than ENVI), we then visually reviewed each population
188 index on each image. Visual inspections of the resulting polygons included a combination of three
189 processes: 1. Accepting the results as-is; or 2. Retraining the supervised classification and re-running

the maximum likelihood classifier; and/or 3. Manually editing the population indices where minor adjustments were needed. Our final step was to then calculate the areas that comprise the penguins-only polygon to arrive at the calculated area of penguin pixels on each image, which represents the population index we report here, for each image date at each colony (Fig. 2E). This population index is the response variable for our statistical modeling (below).

Though one analyst was responsible for the majority of images analyzed here (largely due to the amount of time required for one person to conduct all analyses, let alone more people), independent analysis of one image per colony per year occurred, which we used as a basis for spot-checking results (please see bold data in Table 1).

Statistical modeling

We constructed a series of linear models to evaluate the factors that influence population indices of adult emperor penguins derived from satellites, which was our response variable. In all models, the population index was log-transformed to accommodate a normally-distributed error structure and to facilitate proportional comparisons among colonies of different mean sizes (according to Fretwell et *al.*, 2012). We included a fixed effect of colony in all models to account for differences in average colony size**.** To evaluate our primary hypotheses and thereby evaluate the factors that account for seasonal variation in satellite-derived estimates of penguin abundance, we constructed a series of alternative models containing different explanatory covariates. We describe this suite of models and justification for each explicit covariate below.

We were first interested in whether characteristics of the VHR image itself would influence the population index at each colony due to human interpretation of pixels classified as penguins versus other items on the landscape, such as shadows or guano (Hypotheses 1 and 2). In R (R Core Team, 4.0.1, 2020), we developed a linear model using the function *lm* from the package *stats*; our response variable was the population index (penguin area in meters) per image within a season (year 2011) for each colony. Our explanatory variables were effective panchromatic ground resolution (the spatial size of a pixel given the on-nadir band resolution for the platform combined with the actual off-nadir angle of the satellite platform; expressed in meters), the sun elevation angle, the sun azimuth (range: 0-360 degrees) and colony.

While breeding, emperor penguins remain within a larger area that encompasses the whole breeding site during a season, although the location of the actual colony at the micro-scale changes (Richter *et al.*, 2018a). To address hypothesis 3 (effects of colony patchiness on the population index),

we qualitatively categorized the colony patchiness on each image into “compact” and “spread”. We defined “compact” as when the birds were observed in discrete groups with little space between individuals (i.e., huddling behavior), and “spread” was defined as when there was obvious space between birds and the groups were more dispersed (Fig. 3). We developed a linear model in R with population index as the response variable, and patchiness (spread and compact) and colony as fixed effects.

To understand the variability of population indices related to environmental conditions (hypothesis 4), three different environmental variables likely influencing emperor penguins and their patchiness were tested (Richter *et al.*, 2018b): (i) the 10m zonal wind (U wind); (ii) the 10 m meridional wind (V wind); the 2m air temperature. We obtained these data from the European Centre for Medium-Range Weather Forecasts (ECMWF) “ERA5 hourly data on single levels from 1979 to present” dataset and computed for every hour. We extracted data from August 1st to December 31th 2011, with an hourly temporal resolution and a 0.25° x 0.25° spatial resolution (https://cds.climate.copernicus.eu/). We fit linear model in R with population index as the response variable and absolute wind speed derived from 10m meridional and zonal winds, 2 m temperature and colony as fixed effects.

We used Akaike Information Criterion (AIC) for model selection, combining both forward and backward selection (i.e. function *stepAIC* of the MASS package, R). For hypothesis 4, final models were developed for the environmental window of the date of image acquisition, and for 2-days, or 3-days prior to image acquisition. A comparison of AIC allowed us to choose the best environmental window. For all models, validations were checked by plotting Pearson residuals against fitted values, and against each explanatory variable, verifying homogeneity and normality of residuals (Zuur, Leno and Elphick, 2010). These models did not take into account temporal autocorrelation, but we checked temporal correlation of the residuals by plotting the residuals of the final model versus the Julian dates and checking the correlation (i.e. 0.0014).

Finally, to select the best covariates for accounting and removing sources of spurious variation in population indices, we used model selection to identify the most parsimonious model combining all satellite-related and environmental covariates. Two linear models combining all three colonies were fitted: one model was fitted with “colony” as fixed effects and the other included the satellite-related and environmental variables. We then calculated the proportion of variance explained by the covariates by comparing the R-squared from a model that included the satellite-related and environmental covariates to one that omitted them (but still retained the fixed effect of colonies).

The day of year was correlated (>0.5) with the sun elevation and the temperature so we did not include day of year in the models. However, we checked the absence of correlation between the final model residuals and the day of year.

Simulation to evaluate the effects of observation error on precision of trend estimates

Residual variance in our fitted models measures the magnitude of observation error among repeated surveys within a season. The null model includes the maximum amount of residual variance in surveys, while the "top" model indicates the degree to which covariates can reduce this variance by "correcting" for factors that influence the population index during a survey (e.g., weather conditions that cause penguins to densely huddle, resulting in a lower than expected count). To illustrate the potential effects of observation error on trend estimates, we focus our remaining analysis on comparisons between residual variance from these two models (i.e., the null and "top" model).

We conducted a series of simulations in which we introduced different magnitudes of observation error into population time series. We then evaluated the effects of this observation error on the resulting precision of trend estimates at multiple temporal and spatial scales. To achieve this, we simulated a known log-linear population trend of -0.037, resulting in ~30% population decline after 10 years and ~84% population decline after 50 years. While the magnitude of decline has no effect on estimates of trend precision, we included this trend for illustrative purposes and because it aligns with the IUCN Red List Criteria for "Vulnerable" Status. Parameter values for these simulations are described in Appendix S1. Our simulations assumed each colony was surveyed once per year (i.e., with a single VHR image), and observed population indices for each colony were subject to log-normal error ($\varepsilon_{i,t}$) with standard deviation equal to the residual standard error estimated from the statistical models described above. Using these simulated satellite observations as data, we then estimated population trends and annual expected population indices ($N_{i,t}$) independently for each of the *i* simulated colonies. The trend model for each colony was therefore:

$$log(Count_{i,t}) \sim Normal\left(log(N_{i,t}) - \frac{1}{2}\sigma_i^2, \sigma_i^2 \right),$$

$$log(N_{i,t}) = \alpha_i + \beta_i(t-1).$$

Accordingly, the log-linear trend for an individual colony is described by the parameter β_i and initial population index is equal to $exp(\alpha_i)$, while observed satellite counts represent normally distributed

deviations from the (log-scale) annual expected population index, with variance σ_i^2. The term $\frac{1}{2}\sigma_i^2$ corrects for asymmetries in estimating the mean of a log-normal distribution and ensures that aggregated population indices from multiple colonies are not artificially inflated.

We intentionally omitted inter-annual temporal process variance from our simulations (i.e., variance in β_i from year to year), given that we were unable to estimate this quantity from a single year of surveys (our study), and there are currently insufficient data to evaluate its likely magnitude from other studies. However, we note that process variance is a strong determinant of precision in trend estimates and is distinct from observation error (the focus of this study). Thus, our simulations represent a "best case scenario" that illustrate the potential improvement in precision that could be attained by accounting for environmental covariates during surveys, if process variance is zero. In practice, improvements in precision will be lower if process variance is high.

We examined how the precision of trend estimates changed with an increasing number of survey years by refitting the trend model to different lengths of simulated data (t = 10 to 40 years for each colony). Additionally, to examine the potential to improve trend precision by aggregating annual population estimates for multiple colonies, we selected different numbers of colonies (ranging from I = 2 to 40) and summed their annual indices to generate an estimated "regional" index where $R_t = \sum_{i=1}^{I} N_{i,t}$. We calculated the temporal trend for the regional population as $\frac{log(R_t) - log(R_1)}{(t-1)}$. Further details of simulation and trend analyses, including model fitting procedures, are described in Appendix S1. In all simulations, we quantified the precision associated with trend estimates as the width of the 95% equal-tailed credible interval. We repeated this simulation exercise 100 times for each variance scenario (residual variance based on either the null or "top covariate" model), and each combination of monitoring length (10-40 years) and colony aggregation (1-40 colonies aggregated). We report mean trend precision for the repeated simulations. We considered trends to be estimated with "high precision" if the width of the confidence interval was less than 0.035 (i.e., a change of approximately 3.5% per year). This threshold is consistent with the high precision category for other large-scale avian monitoring programs (e.g., Status of Birds in Canada; Environment Canada 2019), but we note that any categorical threshold is somewhat arbitrary and mainly used for illustrative purposes.

Results

We analyzed a total of 44 images across three colonies during spring 2011 and found that the population index (again, area of penguin pixels in m^2) calculated by VHR imagery within a single season varied among repeated surveys at all three emperor penguin colonies (Tables 1 and 2; Figs. 1 and 4). Both colonies in the Weddell Sea varied by a factor of ~5 and Coulman Island (in the Ross Sea) varied by a maximum factor of two throughout spring 2011. Dates of minimum population indices occurred in September across all three colonies but the date of maximum population indices varied (Table 2, Fig. 1). We failed to support hypothesis 1, as satellite resolution during a survey was not correlated with the population index on that survey. However, in support of hypothesis 2, sun elevation and sun azimuth had a significant positive effect on the population index within a season at these colonies (Table 3).

In our test of hypothesis 3, all three colonies were categorized as both "spread" and "compact", roughly equally through the season, with no tendency toward one or the other at any point (i.e., colonies were not necessarily defined "compact" in early season versus later). We did find, however, that patchiness (i.e., compact vs. spread) had a significant effect on the population index across all colonies (Table 4): when penguins were spread out, the population indices were approximately 1.7% bigger (i.e. median of 10781 m^2 for spread and 6283 m^2 for compact) than when the colony was categorized as "compact". Thus, colonies fluctuate between "compact" and "spread" patterns throughout the spring survey period (September through December), which influences the resulting index of population on any given survey.

Population indices were negatively correlated with strong wind speeds and low temperatures on the day of the survey (hypothesis 4; Table 5). Environmental conditions in the 2- and 3-day period leading up to a survey were also correlated with population indices but received less support in our models than a 1-day environmental window.

To examine the overall effect of accounting for these covariates, we constructed a final model that included additive combinations of the covariates from our hypothesis tests. We again included a fixed effect of colony in all models to account for differences in the mean index among colonies. After model selection, we retained variables: wind speed for the date of VHR image acquisition, sun elevation, sun azimuth, and satellite resolution (though this effect was not significant using a p-value threshold of 0.05). In combination, these covariates explained 46% of the variance in population indices among surveys within a colony (Table 6). This reflects the variance in population indices explained among repeated surveys within colonies, and is independent from the variance explained among colonies by the fixed colony effect.

With regard to our simulations, residual observation error led to uncertainty in estimates of population trend (Appendix S1; Fig. 5). As expected, trend estimates were more precise (95% credible interval widths smaller) when colonies were monitored for a longer duration and when annual estimates were aggregated for multiple colonies. Trend precision was also considerably higher after accounting for survey covariate effects (compare Fig. 5B to 5A). On average, trends at individual colonies could be estimated with "precision" (i.e., 95% credible interval width < 0.035) after 24 years of monitoring if survey covariates were accounted for. In contrast, 31 years of monitoring were required to achieve precision if survey covariates were not accounted for. Population trends for aggregations of multiple colonies could be estimated with high precision with fewer years of monitoring. For example, when accounting for environmental covariates, high precision in trend estimates could be achieved after only 10 years of monitoring if approximately 18 colonies were aggregated. Conversely, without accounting for environmental covariates, approximately 33 colonies must be aggregated to achieve high precision in trend estimates after 10 years of monitoring. Accounting for the environmental and behavioral drivers of observation error can substantially improve confidence in population trends.

Discussion

Our analysis is the first to i) address the intra-seasonal variability in VHR-derived population indices at three emperor penguin colonies, and to ii) identify covariates that can correct for these sources' observation error. In the first study to estimate the global population of emperor penguins using VHR surveys, Fretwell et *al.* (2012) assumed that area of penguin pixels (our "population indices" here) derived from a single image within a season would reasonably represent colony size for that year. This assumption appears to be valid for coarse comparisons among colonies that differ substantially in size; VHR-derived surveys can readily distinguish a colony of many thousands of individuals (e.g., Coulman Island) from a colony of several hundred (e.g., Beaufort Island, Fretwell et *al.*, 2012). However, our study revealed that VHR-derived population indices vary substantially among repeated surveys throughout a single season at each of our three colonies. Further, we showed that satellite-related and environmental variables can describe intra-season variation in area of penguin pixel at a colony, which is essential for calculating robust estimates of population size and trends in the future, especially when only one satellite image is typically available per year. This work has major implications for the future assessment of emperor penguin responses to climate change.

Overall, population indices range from 2326 to 11748 m^2 for Atka Bay, from 14964 to 31005 m^2 for Coulman Island and from 3155 to 18724 m^2 for Stancomb-Wills. Variation in population indices among repeated surveys arises from intrinsic behavior of the birds (e.g., foraging trips by adults that cause temporary fluctuations in colony abundance throughout a season, or huddling behavior that obscures individuals from view) and counting errors owing to imprecision in the observation process (e.g., differences in satellite position, or other factors that cannot be controlled during surveys). Collectively, this "within-season" observation error causes surveys to deviate from a seasonal expected count at the colony. Encouragingly, our study demonstrates that covariates can be used to "correct" for several important drivers of observation error, such as sun angles and weather during a survey. Large-scale monitoring programs routinely correct for variables known to influence counts during surveys. For example, the North American Breeding Bird Survey corrects for observer experience (Sauer, Peterjohn and Link, 1994), and numerous covariates are used to correct for phenological and environmental effects during harbor seal (*Phoca vitulina*) surveys (Hoef, 2003). Recently, Foley et *al.* (2020) developed a phenological correction model for King Penguins that accounts for the seasonal timing of surveys and corrects for attrition of multiple life cycle stages. This was a necessary step to "standardize" surveys collected in many different years, often in different stages of the species' life cycle. In the present study, a large proportion of observation error remains to be explained, and some may in fact be unexplainable (i.e., controlled by a combination of factors that are irreducibly complex, for example the movements of adults to and from the colony on foraging trips). Nevertheless, improvements to VHR-derived population indices described here are an important step toward any future research and monitoring and are therefore critical for the conservation of the species (Trathan *et al.*, 2020).

Emperor penguin colonies are highly dynamic within a season (Figs. 1, 3 and 4). Depending upon the prevailing conditions, penguins may disperse and spread out, or they may cluster and aggregate forming compact groups in response to local weather conditions (Richter *et al.*, 2018b). Our results confirmed that compact huddling behavior was detectable with VHR imagery and was more likely to occur in cold and windy conditions. This makes sense because penguins form huddles to conserve energy (Le Maho, 1977; Gilbert *et al.*, 2009), and huddling increases with colder temperatures and stronger wind speed (Gilbert *et al.*, 2006, 2007). Importantly, this behavior affected the resulting population index during a survey. Cold and windy conditions resulted in fewer pixels classified as "penguin", likely because multiple huddling individuals fit within a single pixel. As

a result, population indices were ~0.6% smaller (i.e. based on medians) when colonies were categorized as "compact".

Future application of these satellite- and environmental-based corrections will need to account for sources of observation error that are likely to differ among colonies. Some sites may be less exposed to winds and cold temperatures (e.g., sheltered colonies located in the lee of islands or peninsulas, or within ice creeks), which could affect the probability a colony will be densely huddled during a survey. Factors that affect the supervised classification process may also differ among colonies. Clouds, shadows, and dense guano stains make images more difficult to interpret (Barber-Meyer et *al.*, 2007), resulting in a less precise classification and a potential overestimate of abundance. Here we showed that lower sun elevation will cast more shadows and increase the number of pixels classified as "penguin". Similarly, sun azimuth values that result in shadows being cast from surrounding features like ice cliffs could obscure penguins that would otherwise be visible. Unfortunately, in practice we do not have the option to choose which date range(s) have the highest quality cloud-free images at a colony. In the rare cases where multiple high-quality images exist within a season, we strongly advocate for the approach we adopted in this study (i.e., leveraging information from *all* available images and statistically accounting for factors that introduce sampling variation). Ongoing efforts to identify these sources of spurious variation (and bias) in surveys are required for improved monitoring of this species.

The application of these methods and use future results has implications for Research and Monitoring Plans, which are a prerequisite for marine protected areas (MPA) designated by the Commission on the Conservation of Antarctic Marine Living Resources (CCAMLR). To advance our understanding of emperor penguins status within current MPAs (e.g., the largest MPA in the world, Ross Sea) and future MPAs, our work would facilitate the development of such a framework. Our simulations found that several emperor penguins colonies need to be aggregated to detect real metapopulation changes as detailed in Kooyman and Ponganis (2017); this suggests the need for a regional network of monitoring and is instructive in the context of the creation of marine protected areas based on ecoregions (Brooks *et al.*, 2020). Given that a primary tenet of the CAMLR Convention is to ensure "maintenance of the ecological relationships between harvested, dependent and related populations of Antarctic marine living resources" – and that emperor penguins are dependent and related populations – it is possible that we would not be able to detect alterations to the ecosystem with monitoring tools at present. Our results therefore support a regional network of emperor penguin colony monitoring, which could take the form of a network of MPAs.

Acknowledgements

Geospatial support for this work provided by the Polar Geospatial Center under NSF-OPP awards 1043681 and 1559691. NCAR- PPC visitor funds and Ian Nisbet that supported the internship of LV. WWF-UK supported PNT and PTF under grant GB095701. DZ was supported by The Penzance Endowed Fund and The Grayce B. Kerr Fund in Support of Assistant Scientists. To SJ, ML, SL, LV, NSF OPP 1744794. Special thanks go to Rose Nichol-Foster for analyzing one image of Coulman Island for comparison and to Frédéric Le Manach for figure designs.

Authors' contributions

SJ, SL, DI and ML conceived the ideas and designed methodology. SJ and ML obtained funding for the analyses; SL and LV analysed the satellite VHR with the help of ML; SL and LV performed the statistical analysis and DI developed the population model and performed the simulation study with the help of SJ. SL led the writing of the manuscript with DI and ML. All authors contributed critically to the final draft and gave final approval for publication.

Data availability statement

Data and data products related to the paper will be available on the following repository http://dx.doi.org/10.17632/kz6j2m4m7t.1 upon acceptance.

References

Ancel, A., Kooyman, G. L., *et al.* (1992) 'Foraging behaviour of emperor penguins as a resource detector in winter and summer', *Nature*, 360(6402), pp. 336–339. doi: 10.1038/360336a0.

Ancel, A., Gendner, J. P., *et al.* (1992) 'Satellite radio-tracking of emperor penguins walking on sea-ice to refeed at sea', *Wildlife Telemetry Remote Monitoring and Tracking of Animals. Ellis Horwood, New York, NY*, pp. 201–202.

Barber-Meyer, S. M., Kooyman, G. L. and Ponganis, P. J. (2007) 'Estimating the relative abundance of emperor penguins at inaccessible colonies using satellite imagery', *Polar Biology*, 30(12), pp. 1565–1570. doi: 10.1007/s00300-007-0317-8.

Barbraud, C. and Weimerskirch, H. (2001) 'Emperor penguins and climate change', *Nature*, 411, pp. 183–186. doi: 10.1038/35075554.

Brooks, C. M. *et al.* (2020) 'Progress towards a representative network of Southern Ocean protected areas', *PLOS ONE*. Edited by Y. Ropert-Coudert, 15(4), p. e0231361. doi: 10.1371/journal.pone.0231361.

Foley, C. M., Fagan, W. F. and Lynch, H. J. (2020) 'Correcting for within-season demographic turnover to estimate the island-wide population of King Penguins (Aptenodytes patagonicus) on South Georgia', *Polar Biology*, 43(3), pp. 251–262. doi: 10.1007/s00300-020-02627-0.

Fretwell, P. T. *et al.* (2012) 'An Emperor Penguin Population Estimate: The First Global, Synoptic Survey of a Species from Space', *PLoS ONE*. Edited by A. Chiaradia, 7(4), p. e33751. doi: 10.1371/journal.pone.0033751.

Fretwell, P. T. and Trathan, P. N. (2009) 'Penguins from space: faecal stains reveal the location of emperor penguin colonies', *Global Ecology and Biogeography*, 18, pp. 543–552.

Fretwell, P. T. and Trathan, P. N. (2020) 'Discovery of new colonies by Sentinel2 reveals good and bad news for emperor penguins', *Remote Sensing in Ecology and Conservation*. Edited by K. Scales and P. Bouchet, p. rse2.176. doi: 10.1002/rse2.176.

Gilbert, C. *et al.* (2006) 'Huddling behavior in emperor penguins: Dynamics of huddling', *Physiology & Behavior*, 88(4–5), pp. 479–488. doi: 10.1016/j.physbeh.2006.04.024.

Gilbert, C. *et al.* (2007) 'How do weather conditions affect the huddling behaviour of emperor penguins?', *Polar Biology*, 31(2), pp. 163–169. doi: 10.1007/s00300-007-0343-6.

Gilbert, C. *et al.* (2009) 'One for all and all for one: the energetic benefits of huddling in endotherms', *Biological Reviews*, p. no-no. doi: 10.1111/j.1469-185X.2009.00115.x.

Hoef, J. M. V. (2003) 'A Bayesian Hierarchical Model for Monitoring Harbor Seal Changes in Prince William Sound, Alaska', *Environmental and Ecological Statistics*, 10(2), pp. 201–219. doi: 10.1023/A:1023626308538.

Jenouvrier, S. *et al.* (2014) 'Projected continent-wide declines of the emperor penguin under climate change', *Nature Climate Change*, 4(8), pp. 715–718. doi: 10.1038/nclimate2280.

Jenouvrier, S. *et al.* (2020) 'The Paris Agreement objectives will likely halt future declines of emperor penguins', *Global Change Biology*, 26(3), pp. 1170–1184. doi: 10.1111/gcb.14864.

Kirkwood, R. and Robertson, G. (1997) 'Seasonal change in the foraging ecology of emperor penguins on the Mawson Coast, Antarctica', *Marine Ecology Progress Series*, 156, pp. 205–223.

Kooyman, G. L. and Ponganis, P. J. (2017) 'Rise and fall of Ross Sea emperor penguin colony populations: 2000 to 2012', *Antarctic Science*. 2016/12/05 edn, 29(3), pp. 201–208. doi: 10.1017/S0954102016000559.

LaRue, M. A. *et al.* (2011) 'Satellite imagery can be used to detect variation in abundance of Weddell seals (Leptonychotes weddellii) in Erebus Bay, Antarctica', *Polar Biology*, 34(11), pp. 1727–1737. doi: 10.1007/s00300-011-1023-0.

Le Maho, Y. (1977) 'The emperor penguin: A strategy to live and breed in the cold: Morphology, physiology, ecology, and behavior distinguish the polar emperor penguin from other penguin species, particularly from its close relative, the king penguin', *American Scientist*, 65(6), pp. 680–693.

Lynch, H. J. and LaRue, M. A. (2014) 'First global census of the Adélie Penguin', *The Auk*, 131(4), pp. 457–466. doi: 10.1642/AUK-14-31.1.

McMahon, C. R. *et al.* (2014) 'Satellites, the All-Seeing Eyes in the Sky: Counting Elephant Seals from Space', *PLoS ONE*. Edited by Y. Ropert-Coudert, 9(3), p. e92613. doi: 10.1371/journal.pone.0092613.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richter, S. *et al.* (2018a) 'A remote-controlled observatory for behavioural and ecological research: A case study on emperor penguins', *Methods in Ecology and Evolution*, 9(5), pp. 1168–1178. doi: 10.1111/2041-210X.12971.

Richter, S. *et al.* (2018b) 'Phase transitions in huddling emperor penguins', *Journal of Physics D: Applied Physics*, 51(21), p. 214002. doi: 10.1088/1361-6463/aabb8e.

Sauer, J. R., Peterjohn, B. G. and Link, W. A. (1994) 'Observer Differences in the North American Breeding Bird Survey', *The Auk*, 111(1), pp. 50–62. doi: 10.2307/4088504.

Smith, W. O. *et al.* (2014) 'The Oceanography and Ecology of the Ross Sea', *Annual Review of Marine Science*, 6(1), pp. 469–487. doi: 10.1146/annurev-marine-010213-135114.

Strycker, N. *et al.* (2020) 'A global population assessment of the Chinstrap penguin (Pygoscelis antarctica)', *Scientific Reports*, 10(1), p. 19474. doi: 10.1038/s41598-020-76479-3.

Trathan, P. N. *et al.* (2020) 'The emperor penguin - Vulnerable to projected rates of warming and sea ice loss', *Biological Conservation*, p. 108216. doi: 10.1016/j.biocon.2019.108216.

Wege, M., Salas, L. and LaRue, M. (2020) 'Citizen science and habitat modelling facilitates conservation planning for crabeater seals in the Weddell Sea', *Diversity and Distributions*. Edited by A. Grech, 26(10), pp. 1291–1304. doi: 10.1111/ddi.13120.

Zuur, A. F., Leno, E. N. and Elphick, C. S. (2010) 'A protocol for data exploration to avoid common statistical problems: Data exploration', *Methods in Ecology and Evolution*, 1(1), pp. 3–14. doi: 10.1111/j.2041-210X.2009.00001.x.

Tables and Figures

Table 1. List of the images used in the study for the three colonies and their estimated penguin areas (expressed in m^2). In bold is indicated the areas calculated from two different analysts for comparisons, the replicated images indicated with a star were not used in the analysis.

Colony	Image ID	Date	Satellite	Area (m^2)	Analysts
Coulman Island	101001000E224A00	09/17/2011	QB02	23985.05	Lise Viollat
Coulman Island	101001000E23DB00	09/18/2011	QB02	24899.15	Lise Viollat
Coulman Island	101001000E283100	09/21/2011	QB02	15047.01	Lise Viollat
Coulman Island	101001000E311E00	09/27/2011	QB02	29729.95	Lise Viollat
Coulman Island	101001000E32A400	09/28/2011	QB02	20738.58	Lise Viollat
Coulman Island	101001000E357300	09/30/2011	WV02	14964.65	Lise Viollat
Coulman Island	101001000E36F100	10/01/2011	QB02	18488.57	Lise Viollat
Coulman Island	101001000E418600	10/08/2011	QB02	23490.58	Lise Viollat
Coulman Island	**101001000E59A900**	**10/24/2011**	**QB02**	**23490.84**	**Lise Viollat**
Coulman Island	**101001000E59A900**	**10/24/2011**	**QB02**	**25274.66**	**Rose Nichol***
Coulman Island	101001000E686700	11/03/2011	QB02	23718.13	Lise Viollat
Coulman Island	103001000F7F8B00	11/19/2011	WV02	31005.29	Lise Viollat
Atka Bay	103001000D400A00	09/03/2011	WV02	6001.51	Lise Viollat
Atka Bay	103001000D023100	09/04/2011	WV02	6565.007	Lise Viollat
Atka Bay	103001000D5A8100	09/06/2011	WV02	2326.293	Lise Viollat
Atka Bay	103001000D295800	09/15/2011	WV02	11748.12	Lise Viollat
Atka Bay	101001000E24BD00	09/19/2011	QB02	5558.519	Lise Viollat
Atka Bay	101001000E262100	09/20/2011	QB02	8367.776	Lise Viollat
Atka Bay	**103001000D63FD00**	**09/20/2011**	**WV02**	**8533.22**	**Lise Viollat**
Atka Bay	**103001000D63FD00**	**09/20/2011**	**WV02**	**8449.75**	**Peter Fretwell***
Atka Bay	103001000DD35500	09/21/2011	WV02	6506.849	Lise Viollat
Atka Bay	101001000E291500	09/22/2011	QB02	8774.873	Lise Viollat
Atka Bay	103001000D965F00	09/22/2011	WV02	5108.962	Lise Viollat
Atka Bay	101001000E2BBA00	09/24/2011	QB02	3450.174	Lise Viollat
Atka Bay	103001000E2B6200	09/25/2011	WV02	9401.531	Lise Viollat
Atka Bay	101001000E3F5200	10/07/2011	QB02	7760.415	Lise Viollat
Atka Bay	101001000E526C00	10/19/2011	QB02	7480.21	Lise Viollat
Atka Bay	103001000FD05F00	11/21/2011	WV02	10422.23	Lise Viollat
Stancomb-Wills	103001000DB9F900	09/13/2011	WV02	4685.562	Lise Viollat
Stancomb-Wills	103001000E7F2B00	09/14/2011	WV02	3864.594	Lise Viollat
Stancomb-Wills	**101001000E21CB00**	**09/17/2011**	**QB02**	**8232.159**	**Lise Viollat**
Stancomb-Wills	**101001000E21CB00**	**09/17/2011**	**QB02**	**6132**	**Peter Fretwell***
Stancomb-Wills	103001000D081200	09/18/2011	WV02	4727.548	Lise Viollat
Stancomb-Wills	103001000D1DD300	09/19/2011	WV02	5626.239	Lise Viollat
Stancomb-Wills	103001000DA3AC00	09/20/2011	WV02	4895.73	Lise Viollat
Stancomb-Wills	103001000D01A900	09/25/2011	WV02	3155.904	Lise Viollat
Stancomb-Wills	101001000E2DBA00	09/25/2011	QB02	3441.86	Lise Viollat
Stancomb-Wills	103001000EC82200	10/04/2011	WV02	8727.082	Lise Viollat
Stancomb-Wills	101001000E3ACA00	10/04/2011	QB02	11141.14	Lise Viollat
Stancomb-Wills	103001000E6D7F00	10/05/2011	WV02	9964.59	Lise Viollat
Stancomb-Wills	101001000E43EE00	10/10/2011	QB02	12289.29	Lise Viollat
Stancomb-Wills	101001000E458800	10/11/2011	QB02	11540.82	Lise Viollat
Stancomb-Wills	101001000E510200	10/18/2011	QB02	9488.866	Lise Viollat
Stancomb-Wills	101001000E7D1500	11/19/2011	QB02	15267.16	Lise Viollat
Stancomb-Wills	101001000E874200	11/27/2011	QB02	9863.186	Lise Viollat
Stancomb-Wills	103001000F3B6500	12/11/2011	WV02	18724.48	Lise Viollat
Stancomb-Wills	1030010010C89E00	12/14/2011	WV02	13013.86	Lise Viollat

Table 2. Range of "penguin estimated area" (i.e. population index) calculated via supervised classification on VHR imagery at three emperor penguin colonies in Antarctica, including the average area over the season, number of images analyzed per colony, minimum area calculated (m^2), date of the image when minimum area was calculated, maximum area calculated (m^2), date of the image when the maximum area was calculated, and the ratio between the maximum and minimum area calculations per colony to exemplify the magnitude of intra-season change.

Colony Name	Avg area	# images	Min. Area	Date Min. Area	Max. Area	Date Max. Area	Max:min
Atka Bay	7,200	15	2,326	Sept 6, 2011	11,748	Sept 15, 2011	5.05
Coulman Island	22,687	11	14,965	Sept 30, 2011	31,005	Nov 19, 2011	2.07
Stancomb-Wills	8,814	18	3,156	Sept 25, 2011	18,724	Dec 11, 2011	5.93

Table 3. Results of the linear model to determine whether attributes of the satellite platform (resolution (expressed in meters), sun elevation and sun azimuth angles (expressed in degrees); hypotheses 1 and 2) influenced the emperor penguin population indices (expressed in log scale of the area in m^2) calculated from VHR imagery. Adjusted $r^2 = 0.7191$. The colony effect and values are not displayed on the table.

	Value	Std. Error	DF	t-value	p-value
Intercept	7.79	0.31	36	24.903	<2e-16
Panchromatic resolution	0.42	0.42	36	1.014	0.317
Sun elevation angle	**0.028**	**0.008**	**36**	**3.622**	**0.000895**
Sun azimuth angle	**0.0078**	**0.002**	**36**	**3.707**	**0.000702**

Table 4. Results of the linear model to address whether patchiness (i.e., "compact" or "spread") influenced the emperor penguin population indices calculated (expressed in log scale of the area in m^2) from VHR imagery (hypothesis 3). Adjusted $r^2 = 0.7647$. The colony effect and values are not displayed on the table.

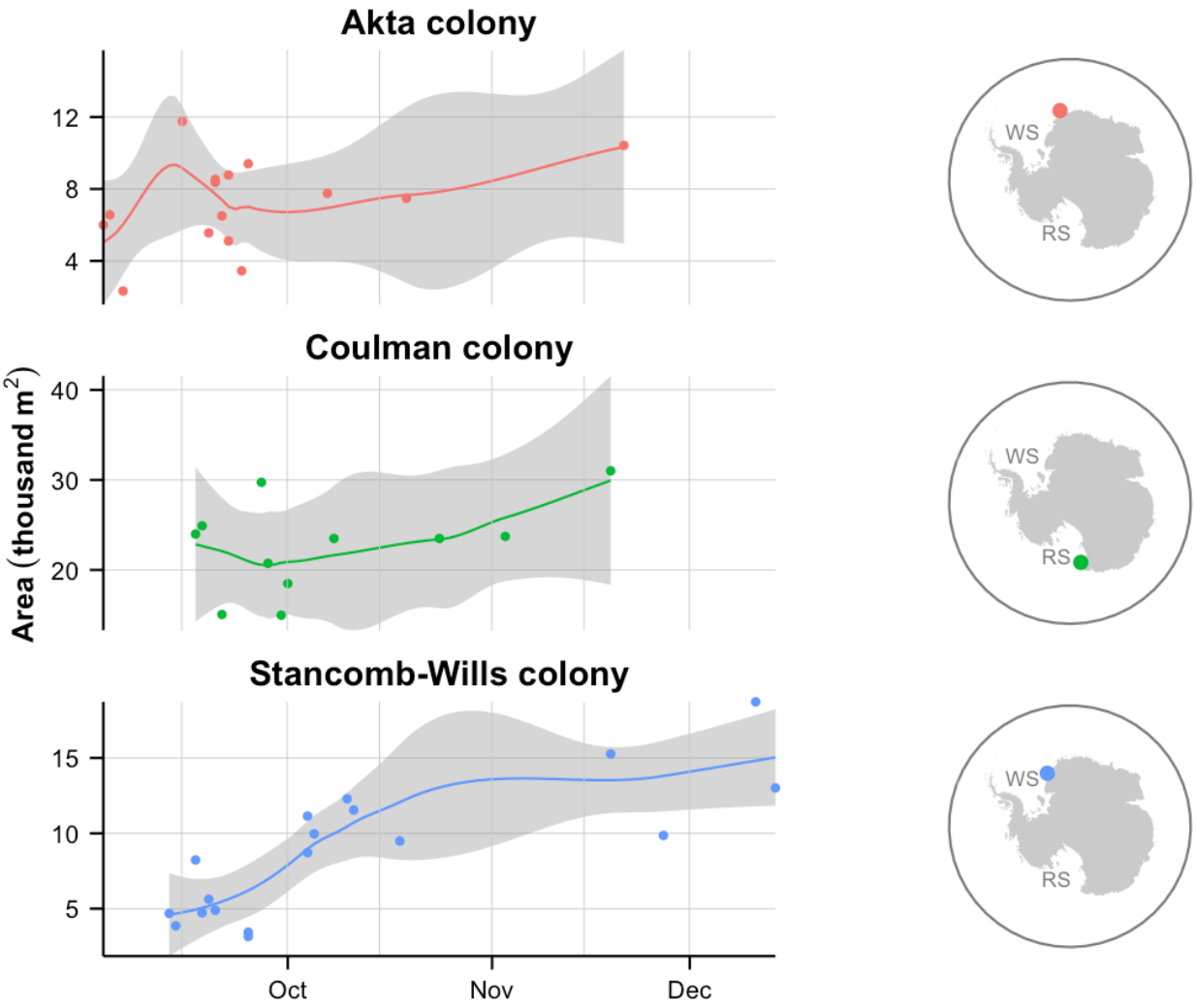
	Value	Std. Error	DF	t-value	p-value
Intercept	9.62	0.16	38	60.385	<2e-16
patchiness	**-0.61**	**0.098**	**38**	**-5.825**	**9.91e-07**

Table 5. Results of the best linear model to address hypothesis 4; testing the influence of environmental covariates (wind (expressed in $m.s^{-1}$) and temperature (expressed in degrees Celsius)) on the emperor penguin population indices (expressed in log scale of the area in m^2) calculated from VHR imagery. Adjusted r^2 = 0.7245.

	Value	Std. Error	DF	t-value	p-value
Intercept	0.71	2.68	37	0.264	0.79320
Absolute wind speed 10m	**-0.067**	**0.021**	**37**	**-3.258**	**0.00241**
Surface temperature	**0.034**	**0.011**	**37**	**3.206**	**0.00277**

Table 6. Results of the best linear model linking survey covariates (resolution (meters), sun elevation and sun azimuth angles (degrees) and wind (in $m.s^{-1}$)) with emperor penguin population indices (expressed in log scale of the area in m^2) to account for observation error during surveys of emperor penguins using VHR imagery. Adjusted r^2 = 0.75 (compared to 0.54 for a "null" model that only included a fixed effect of colony but no survey covariates). The proportion of variance within colonies explained by the survey covariates is 46% (note this is distinct from the proportion of variance among colonies, explained by the colony fixed effect).

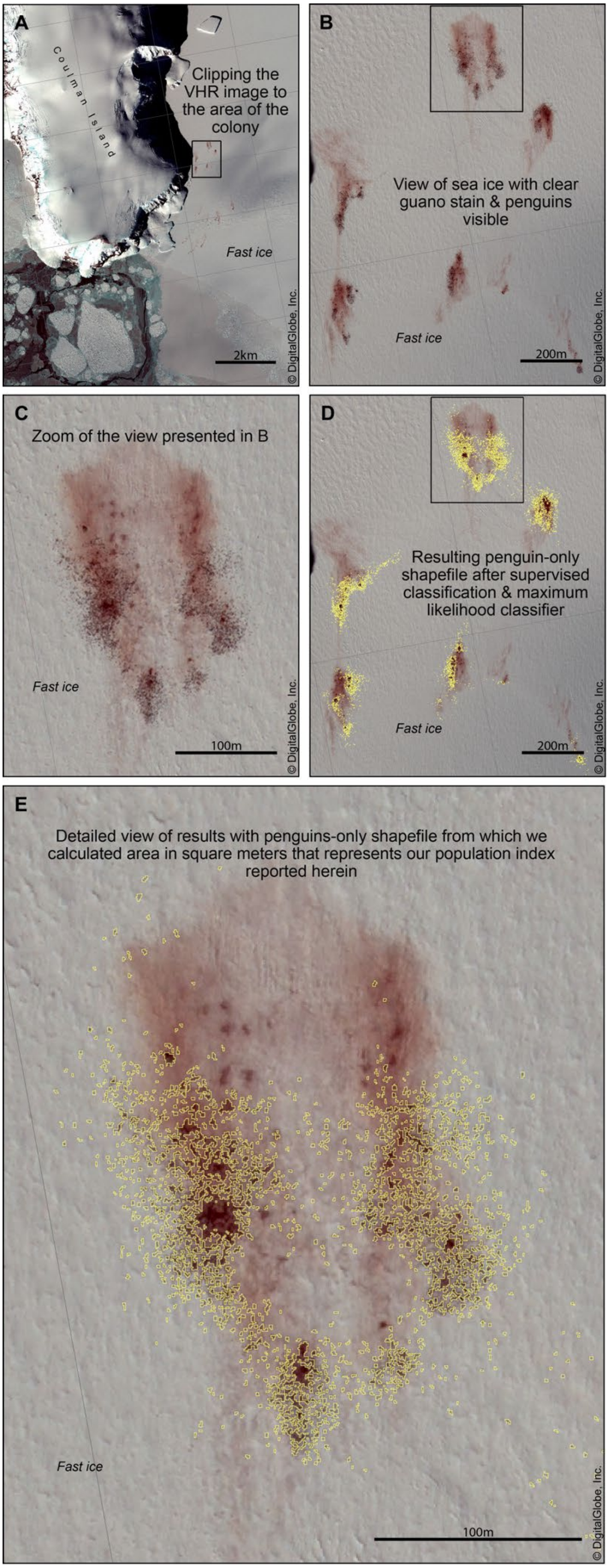
	Value	Std. Error	DF	t-value	p-value
Intercept	8.19	0.33	35	24.552	< 2e-16
Panchromatic resolution	0.51	0.39	35	1.320	0.19527
Sun elevation angle	**0.024**	**0.007**	**35**	**3.215**	**0.00280**
Sun azimuth angle	**0.006**	**0.002**	**35**	**2.955**	**0.00556**
Absolute wind speed 10m	**-0.052**	**0.021**	**35**	**-2.496**	**0.01742**



587

588 **Figure 1.** Time-series of the population indices (in thousand m^2) for the three emperor penguin
589 colonies (left panels), with the location of these colonies given in the right panels (WS = Weddell Sea;
590 RS = Ross Sea). A 'locally weighted smoothing' (loess) regression was applied to each time-series
591 (degree = 1 and default span of 0.75) using R's loess function; vertical bars indicate the 1st and 15th
592 day of each month between 15 September 2011 and 1 December 2011.

593


A
Coulman Island
Clipping the VHR image to the area of the colony
Fast ice
2km
© DigitalGlobe, Inc.
B
View of sea ice with clear guano stain & penguins visible
Fast ice
200m
© DigitalGlobe, Inc.
C
Zoom of the view presented in B
Fast ice
100m
© DigitalGlobe, Inc.
D
Resulting penguin-only shapefile after supervised classification & maximum likelihood classifier
Fast ice
200m
© DigitalGlobe, Inc.
E
Detailed view of results with penguins-only shapefile from which we calculated area in square meters that represents our population index reported herein
Fast ice
100m
© DigitalGlobe, Inc.

594

Figure 2. Diagram outlining the method protocol for the VHR image processing using an example for the Coulman Island colony. The VHR image is a Quickbird-02 image of Coulman Island emperor penguin colony acquired on October 24, 2011 (catalog ID: 101001000E59A900). Imagery copyright DigitalGlobe, Inc.

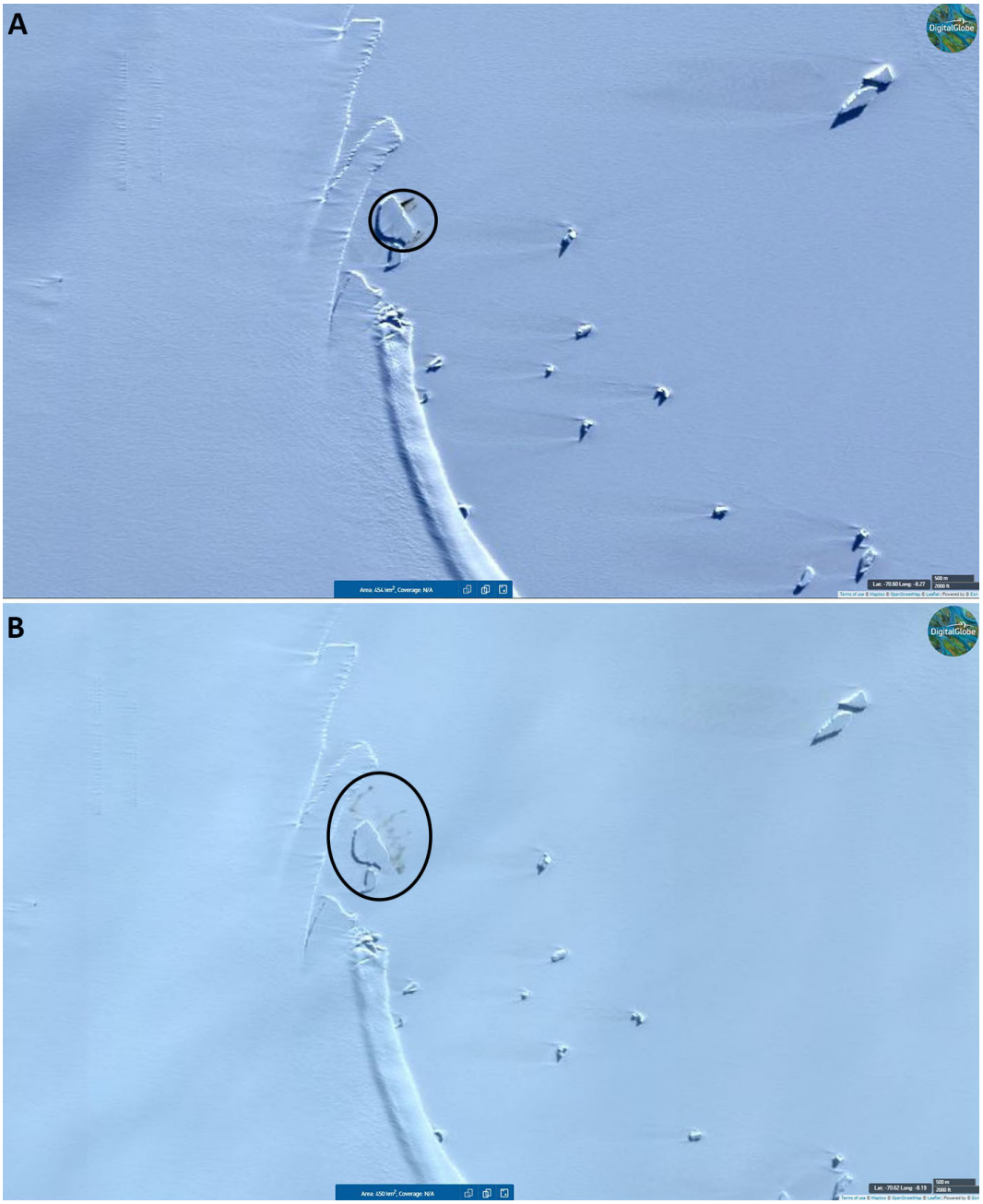
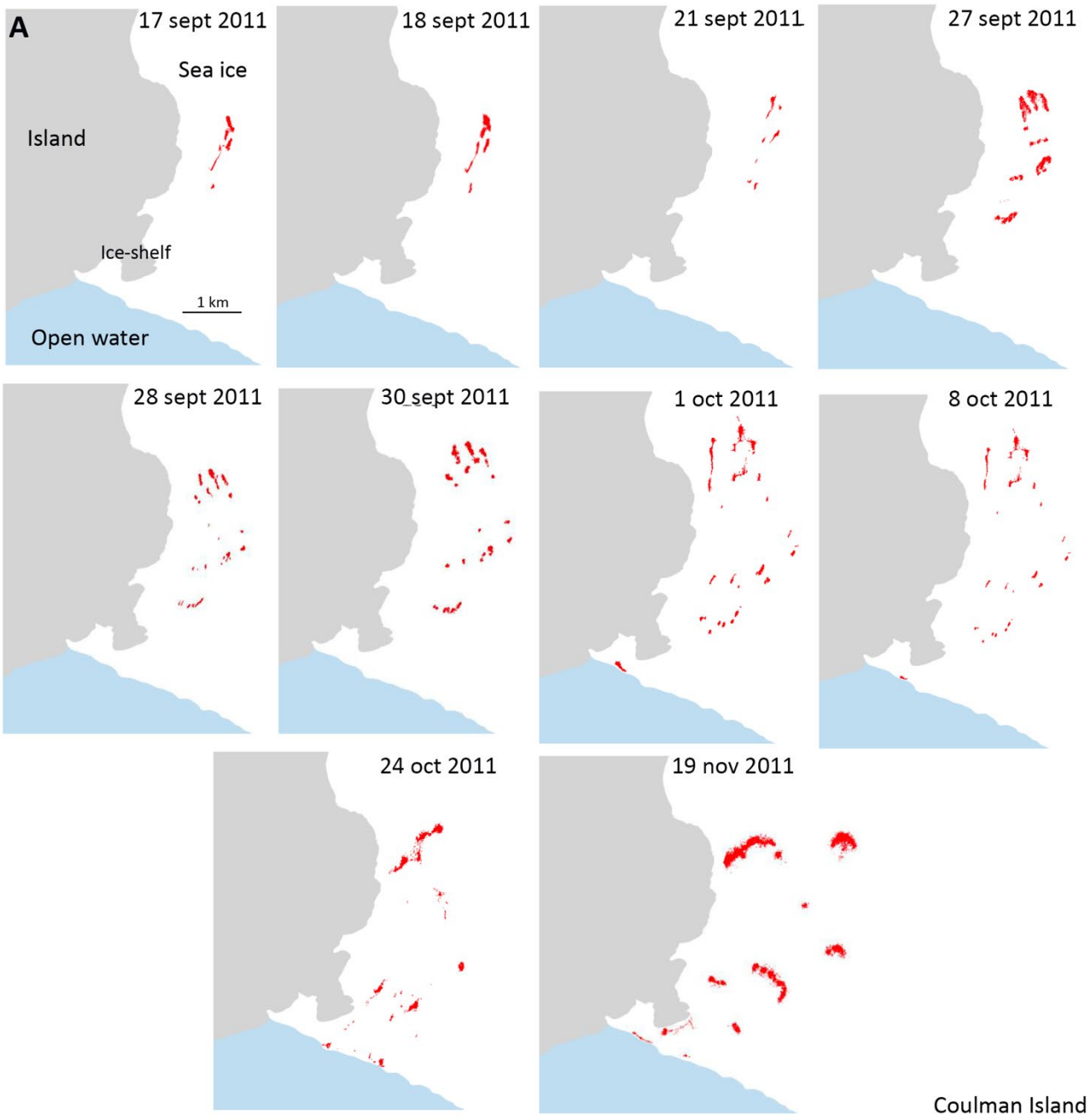
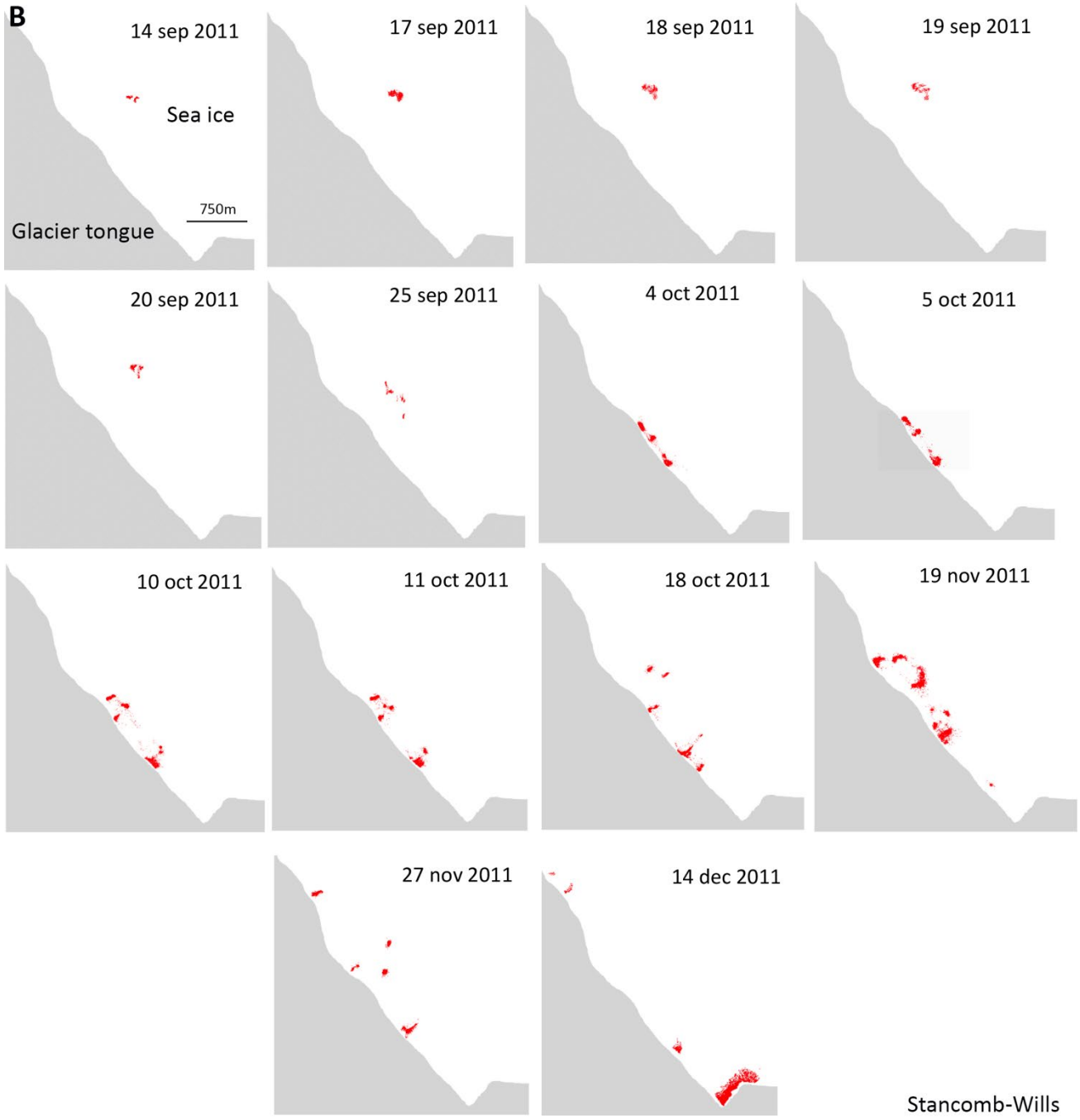


Figure 3. WorldView-2 satellite image from Atka Bay emperor penguin colony for September 3rd 2011 (A), exemplifying "compact" patchiness (group of birds is circled) and WorldView-2 satellite image from Atka Bay later in the season, on September 25th 2011 (B), showing an example of "spread"

patchiness (the group of birds is circled again, and the guano stain spread out over a much larger area). Image courtesy DigitalGlobe, Inc. (Maxar Technologies) and scale bars on bottom right of each image are 500m and 2,000 feet.




B
14 sep 2011
Sea ice
750m
Glacier tongue
17 sep 2011
18 sep 2011
19 sep 2011
20 sep 2011
25 sep 2011
4 oct 2011
5 oct 2011
10 oct 2011
11 oct 2011
18 oct 2011
19 nov 2011
27 nov 2011
14 dec 2011
Stancomb-Wills

610

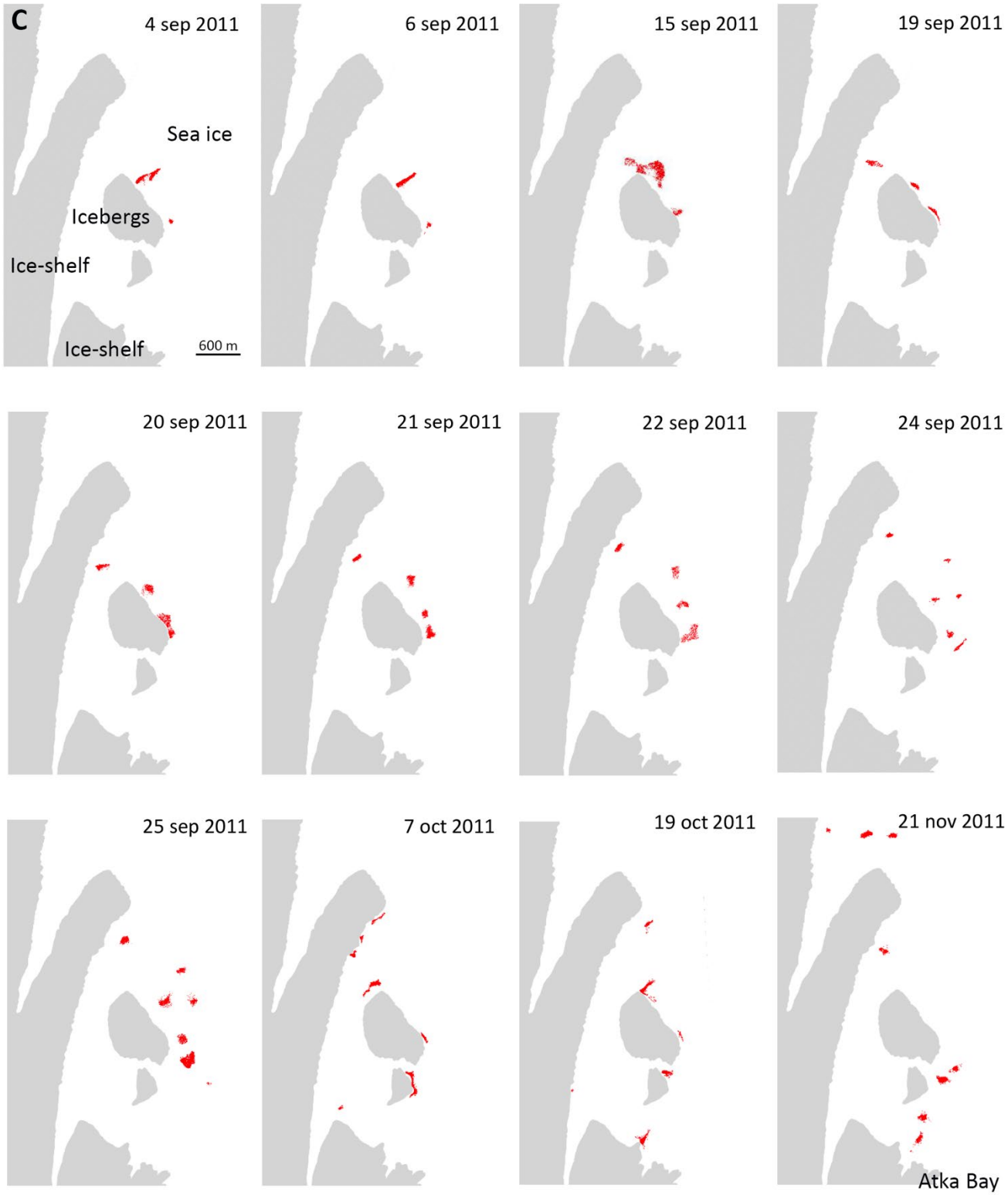
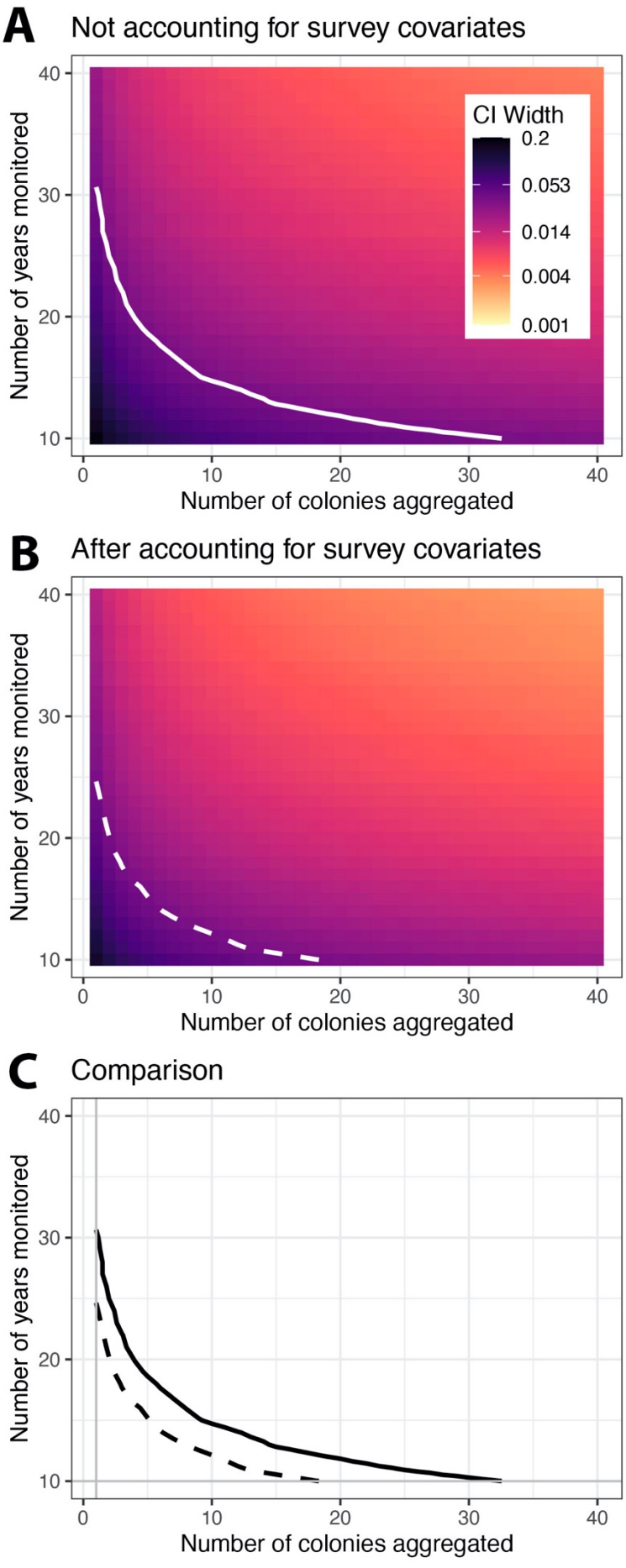


Figure 4. Emperor penguin estimated areas (i.e. population indices) at Coulman Island (panel A), Stancomb-Wills (panel B) and Atka Bay (panel C) colonies during the breeding season. Gray shapes represent the island, ice-shelves, icebergs or glacier tongue near the colony, blue shapes the open water and red shapes the emperor penguin estimated surface areas. All images are not represented, please see the list of images in Table 1.


A
Not accounting for survey covariates
CI Width
0.2
0.053
0.014
0.004
0.001
Number of years monitored
Number of colonies aggregated
B
After accounting for survey covariates
Number of years monitored
Number of colonies aggregated
C
Comparison
Number of years monitored
Number of colonies aggregated

617

Figure 5. Precision associated with trend estimates resulting from simulations that incorporate residual variance in annual population indices from a null model (panel A) and a model that accounts for the effects of environmental drivers on daily population indices (panel B). X-axis denotes the number of colonies that are aggregated; Y-axis denotes the number of years each colony is monitored for. Shading indicates the resulting precision of estimated log-linear trend, measured as width of the 95% equal-tailed credible interval associated with trend estimate. Solid and dashed contour lines in each panel denote the boundary at which trends can be estimated with "high precision" (using a threshold where credible interval width is equal to 0.035). Panel C represents the comparison of panel A & B.