Received: 8 January 2022 | Accepted: 21 March 2022

DOI: 10.1002/ecs2.4174

ECOSPHERE
AN ESA OPEN ACCESS JOURNAL

ARTICLE
Emerging Technologies

Multi-image flock size estimation with CountEm: A case study with half a million Common Eiders and Greater Snow Geese

Marcos Cruz[1] | **Javier González-Villa**[1] | **Josée Lefebvre**[2] | **Scott G. Gilliland**[3] | **Francis St-Pierre**[2] | **Matthew English**[3] | **Christine Lepage**[2]

[1]Department of Mathematics, Statistics and Computer Science, Universidad de Cantabria, Santander, Spain

[2]Canadian Wildlife Service—Québec Region, Environment and Climate Change Canada, Québec City, Québec, Canada

[3]Canadian Wildlife Service—Atlantic Region, Environment and Climate Change Canada, Sackville, New Brunswick, Canada

Correspondence
Marcos Cruz
Email: marcos.cruz@unican.es

Funding information
Proyecto Puente (Contrato Programa Gobierno de Cantabria—Universidad de Cantabria), Consejería de Universidades, Igualdad, Cultura y Deporte del Gobierno de Cantabria

Handling Editor: Tanya Berger-Wolf

Abstract

Many of the methods used for estimating population size from ecological surveys have limitations on precision, cost, and/or applicability. The CountEm method was proposed recently for estimating the number of individuals in large groups from single images. It is simple and efficient, and can be applied to any species. Here we present a case study by applying CountEm to a real ecological survey with 278 images of Greater Snow Geese (*Anser caerulescens atlanticus*) and Common Eiders (*Somateria mollissima*) flocks taken from fixed-wing aircraft in Eastern Canada. First, we evaluated the precision and counting time of CountEm on single images. Second, we developed and tested a new multi-image version of the CountEm software. We show that flock sizes of $N > 35{,}000$ can be estimated on single images in ~5 min, from counting a sample of ~200 birds, yielding relative SEs in the 5%−10% range. Processing times increased to 10–20 min when simultaneously processing large numbers of images that contained over half a million birds with only modest increases in relative SE (range: 10%−15%). Our results suggest that CountEm may be used to save time and resources if incorporated into monitoring programs that utilize imagery in the abundance estimates.

KEYWORDS
CountEm, ECA Flocks data set, flock size estimation, geometric sampling, population size estimation, quadrats

INTRODUCTION

Estimating population sizes is of great importance in ecology. Accurate estimates allow population trends to be detected (Seavy & Reynolds, 2007) and are useful in wildlife management (Zimmerman et al., 2012). The methods discussed in this article are applicable to any species, but the focus of this article is on birds.

Bird censuses are frequently based on visual estimation (Hagy et al., 2017; Sebastián-González & Green, 2016; Zhao et al., 2016; Zimmerman et al., 2012), which often become imprecise and biased for groups greater than

~1000 birds. Several studies have recommended using high-resolution photographs to achieve better estimates (Chabot & Bird, 2015; Hodgson et al., 2018; Lyons et al., 2019). However, for images with more than a few thousand individuals, exhaustive manual counting is laborious and slow (Hollings et al., 2018). Automated computer vision algorithms can work in some particular cases on regular patterns or non-overlapping birds with homogeneous backgrounds, see references in table 1 of Chabot and Francis (2016); however, these algorithms may perform poorly when these conditions are not met (Chabot & Francis, 2016; Hollings et al., 2018). In addition, automated approaches can be expensive and difficult to develop, especially when different types of objects (e.g., multiple species, sexes, and age classes) are being counted simultaneously.

An unbiased method for estimating the number of discrete objects in photographs (Cruz et al., 2015; hereafter CountEm) and the corresponding software (countem.unican.es; González-Villa & Cruz, 2020) were recently developed. The method is based on classical, well-established principles of geometric sampling (Gundersen, 1977) and has been previously applied in quantitative microscopy (Cruz-Orive, 2017; Howard & Reed, 2005). CountEm can be applied to any kind of discrete object or "particle" of interest (e.g., animals, humans, and trees) and only requires that all the particles must be unambiguously identifiable for manual counting in the specific image. To our knowledge, CountEm has not been used in ecological studies, and has only been tested in simulation studies and on single images (Cruz & González-Villa, 2018; 2021; González-Villa & Cruz, 2020).

Our objectives were as follows: (1) to investigate the precision and counting time of CountEm when working with large sets of images (i.e., hundreds) from ecological studies and (2) to implement and test a new functionality (hereafter "multi-image" mode) in the CountEm software, to efficiently estimate the number of birds on multiple images (e.g., Figure 1a). We addressed both objectives by applying CountEm to the Eastern Canada (ECA) Flocks data set (Cruz et al., 2021).

METHODS

The ECA Flocks data set

To test the precision and efficiency of CountEm, we assembled a data set with images of 278 flocks with over half a million, manually annotated bird positions of Greater Snow Geese (GSGO; *Anser caerulescens atlanticus*) and Common Eiders (COEI; *Somateria mollissima*). Publicly available annotated data sets are rare, and this data set has been published separately to provide opportunities to test other estimation approaches (ECA Flocks data set; Cruz et al., 2021). The ECA Flocks data set consists of aerial oblique images from 179 COEI and 99 GSGO flocks, which were photographed from fixed-wing aircraft using different cameras in the COEI (Bordage et al., 1998) and GSGO surveys (Bechet et al., 2004). The link to access the full data set, together with information on the images, such as camera model, date, and geographic coordinates, can be found in the metadata of Cruz et al. (2021).

All birds were counted manually with Cell Counter plugins of ImageJ's (Schneider et al., 2012), saving their Cartesian coordinates in pixel units. A total of 630,485 birds (123,806 COEI and 506,098 GSGO) were annotated. COEI and GSGO flock sizes ranged from 6 to 4154 and from 43 to 36,241, respectively. The 99 GSGO images and their corresponding point patterns are shown in Figure 1a,b. Each black dot of Figure 1b represents what is considered to be a bird by the observer that performed the manual count. Note that in some cases, bird overlaps, occlusions, or low resolution may make birds difficult to be identified without ambiguity. However, these issues are infrequent in the ECA Flocks data set. The data set presents a high variability in the number and spatial distribution of birds among images allowing us to test the precision of CountEm in diverse conditions.

CountEm can be used to estimate flocks of any size, although our early evaluations of the method suggested there were little benefit to use CountEm for small flocks ($N > 700$) as the CountEm and manual counting times were similar. However, a large portion of the birds captured in the imagery can occur in small flocks, and counting all the small flocks manually requires a significant effort. To apply CountEm to small flocks, we developed a multi-image mode. To test the performance of single and multi-image modes, we divided the ECA Flocks data set into groups based on flock size:

- Large flock group: 106 individual images (47 COEI and 59 GSGO), with flock size $N > 700$;
- COEI small flock group (hereafter COEIsg): A composite of 132 COEI images with flock sizes $N \leq 700$. The total number of COEI in this set is $N = 36{,}956$;
- COEI group (hereafter COEIg): A composite of all 179 COEI images, with $N = 123{,}806$;
- GSGO small flocks group (hereafter GSGOsg): A composite of 40 GSGO images with flock sizes $N \leq 700$. The total number of Snow Geese in this set is $N = 11{,}700$; and finally,
- GSGO group (hereafter GSGOg): A composite of all 99 GSGO images, with $N = 506{,}098$ (see Figure 1a).

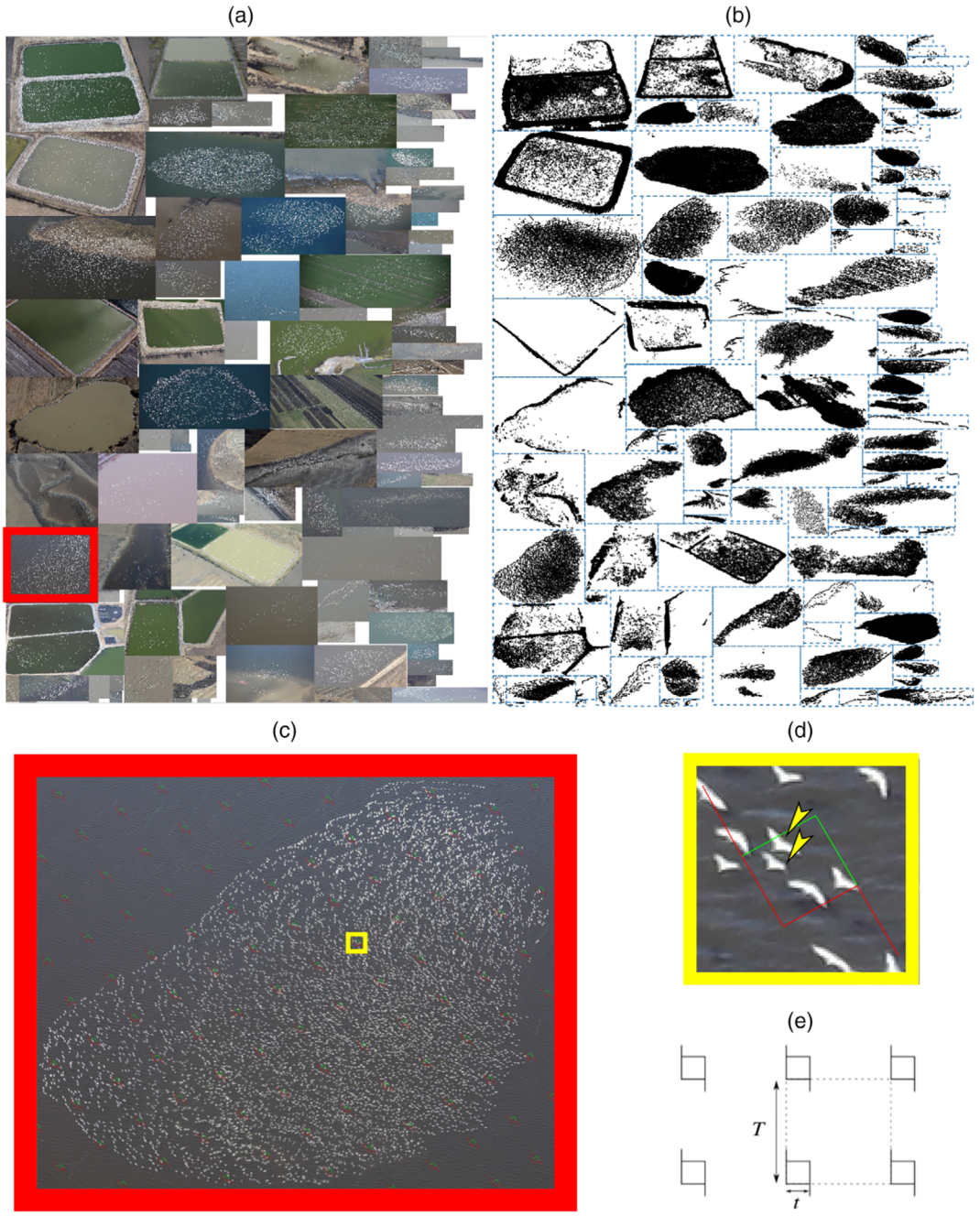

FIGURE 1 Group of the 99 Greater Snow Geese (GSGO) images of the Eastern Canada (ECA) Flocks data set (Cruz et al., 2021). The images have been packed to minimize the total area with no overlap. A single multi-image CountEm estimation yields $\widehat{N} = 527,790$ birds in ∼20 min. (b) Manually annotated positions of the 506,098 birds in (a). (c) Magnified version of the GSGO image marked in (a). $N = 8109$ GSGO were manually counted. A grid of quadrats of the type shown in (e), with $T = 324$ and $t = 36$ pixels ($n_0 = 100$, $f = 0.0123$), has been superimposed with a tilt of 30°. (d) Magnified version of the quadrat marked in (c). Only the two marked birds should be counted, applying the forbidden line counting rule (see text and Gundersen, 1977). (e) A portion of the grid of quadrats proposed in Cruz et al. (2015); Cruz and González-Villa (2018) for systematic sampling. The sampling fraction is t^2/T^2.

The CountEm method

The CountEm method is designed to estimate the number of individual features, elements, or units, generally called *particles* in an image (or group of images), N, by sampling and manually counting between 100 and 300 particles. In the present context, we call the relevant group of N particles a flock, where a particle is a planar projection of a bird or a clearly distinguishable fragment of it on an image. The estimation of the entire population including

birds that are not visible in the images will be addressed in future work. Our study constitutes a prerequisite to develop proper sample designs for population surveys.

Under the preceding mathematical definition of flock, the flock size, N, can be estimated with systematic sampling using a sampling grid (see Figure 1e). The grid is an infinite union of congruent quadrats of side length t, arranged as a square grid with a fixed distance, T, between quadrat centers. In practice, the grid will not be infinite but will encompass the whole flock in the image, as in Figure 1c. The position of the grid has to be uniform random with respect to the image.

The grid is determined by parameters $\{t, T\}$, or by the alternative pair $\{f, n_0\}$, where $f = t^2/T^2$ is the sampling fraction and $n_0 = B_x B_y/T^2$ is the initial number of quadrats, with B_x and B_y representing image width and height, respectively (Cruz & González-Villa, 2018). The latter parameters are the more intuitive, and they facilitate the choice of an adequate sampling grid that captures between 100 and 300 birds. The protocol to choose the parameters is described in Appendix S2.

With the preceding design, an unbiased flock size estimator, $\widehat{N}$, is (Cruz et al., 2015):

$$\widehat{N} = \frac{T^2}{t^2} \times Q = \frac{Q}{f}. \quad (1)$$

The sample size, Q, is the number of particles captured by the sampling grid and is counted manually by the observer. The error of any estimator can be decomposed into two sources, bias and variance. Flock size estimator $\widehat{N}$ has been proven to be unbiased (see appendix S1 in Cruz et al., 2015). Generally, a different value of $\widehat{N}$ is obtained every time the estimation is performed, since the grid position is random. Unbiasedness means that the mean of all the potential values of $\widehat{N}$ is the true value N. Because unbiasedness is a mathematical property, it holds for any image. Since the bias of the estimator is equal to zero, the only source of error comes from sampling variance, which is unknown and depends on the spatial distribution of the birds in the image, and on the chosen parameter values. However, the true sampling variance of $\widehat{N}$ for a given image with manually annotated bird positions can be estimated empirically with computer simulations (i.e., Monte Carlo resampling).

Therefore, two different sampling frames (i.e., sampling settings) (Gómez et al., 2019) were used in the present paper:

- *Real mode*: This mode was used to estimate flock size on single (Figure 1c) or grouped images (Figure 1a) with the aid of the CountEm software. The sampling unit (particle) is a projected bird, or a distinguishable part of it. Sample size Q was counted manually, using an unbiased counting rule, and we chose Gundersen's (1977) forbidden line rule to cope with edge effects: A particle was counted in a quadrat only if it has points in common with the quadrat but does not hit the extended, forbidden line of the quadrat (in red in Figure 1d). To aid observers, we developed a guided protocol that has been implemented in the CountEm software. The protocol provides a simple procedure for setting the initial grid parameter values and is described in Appendix S2. After completing all the steps of the guided protocol, the CountEm software returns the estimated flock size, $\widehat{N}$, obtained with Equation 1, and the predicted coefficient of error $\mathrm{CE}(\widehat{N})$ ($C2$ predictor described in Gómez et al., 2019).
- *Simulation mode*: The goal of *simulation mode* is to estimate the true sampling variance of $\widehat{N}$ in order to assess the precision of the method. This can be achieved via Monte Carlo resampling on point patterns (see Figure 1b) from manually annotated images with known flock size. Each particle in the images was replaced by an associated point, namely the one given by the corresponding manually annotated position, yielding a point pattern. The CountEm software is not designed for *simulation mode* analyses, hence we used the *spatstat* package (Baddeley et al., 2015) to generate a sampling grid and to automatically compute the number of point particles captured by the grid, Q. This setup allowed us to compute M different flock size estimations $\{\widehat{N}_1, \widehat{N}_2, \ldots, \widehat{N}_M\}$ from $M = 2000$ different grid positions, for a given point pattern. The empirical variance, $\mathrm{Var}_e(\widehat{N})$, is the variance among the M values of $\widehat{N}$. The empirical coefficient of error (namely the relative SE), $\mathrm{CE}_e(\widehat{N})$, is calculated by dividing the square root of the empirical variance by true flock size, N: $\mathrm{CE}_e(\widehat{N}) = \sqrt{\mathrm{Var}_e(\widehat{N})}/N$. The empirical values should be close to their respective true values if the number of replications, M, is large enough. Grid parameter selection was performed by simulating the guided protocol used in *real mode*. A more precise description of these calculations is given in Appendix S1.

To estimate the total number of birds in a large set of images (sets of about 100 images), we stitched all images together, resulting in just one big composite image, which was estimated in a single CountEm run. This approach should be more efficient than iteratively analyzing each image individually. This multi-image analysis required modifications to the CountEm software, since the fraction of empty area increased due to the gaps between images. We used the *rectangle-packer*

1.1.0 Python library in order to minimize these gaps (see Figure 1a). We empirically checked that the number of non-empty quadrats should be $n > 50$ when analyzing more than 10 images at the same time. Therefore, we defined a multi-image CountEm mode, modifying the software in Step 2 of the protocol, setting $n_0 = 500$, and increasing the required number of non-empty quadrats to 50 in Step 6 (see Appendix S2). We have implemented this multi-image mode in the new CountEm 1.5.0 software. The protocol is a result of our experience applying the CountEm method to the ECA Flocks data set.

RESULTS

Real mode precision and counting time assessment

We compared the CountEm *real mode* counting time, $t(N)_{\text{CountEm}}$, with the exhaustive manual counting time (i.e., annotating flocks with ImageJ), $t(N)_{\text{manual}}$, for a subset of 18 GSGO images with flocks ranging in size from 200 to 20,000, which is a representative range. Counting time varied between the two methods, with the time for the manual counting method significantly increasing with N: $t(N)_{\text{manual}} \sim 0.042N^{0.86}$, while the CountEm method was independent of N: $t(N)_{\text{CountEm}} \sim 5.1$ min (Figure 2). The central processing unit (CPU) time (i.e., the time needed by the software to generate the grid and process the images) was negligible for single images.

We used CountEm to perform a *real mode* flock size estimation of the five subsets described in the *Methods*, namely four multi-image subsets (COEIsg, COEIg, GSGOsg, and GSGOg) and 106 single images with $N > 700$. The single-image root mean squared error was 9.7% and 9.4% for the 47 COEI and 59 GSGO images with $N > 700$, respectively. These values are consistent with the *simulation mode* analysis shown below. Detailed single-image results can be accessed on Zenodo: https://doi.org/10.5281/zenodo.5819210.

The multi-image results are shown in Table 1. Note that the four groups of images were packed and stitched together as shown in Figure 1, and analyzed with one multi-image CountEm run. CPU time was dependent on the number of images analyzed, and ranged from a few seconds for GSGOsg (40 images) to about 17 min for COEIg (179 images). The relative deviation obtained with one *real mode* estimation, $100 \times (\widehat{N} - N)/N$, ranged from 4% to 11%. These values are consistent with the "true" relative standard deviation computed below in *simulation mode*.

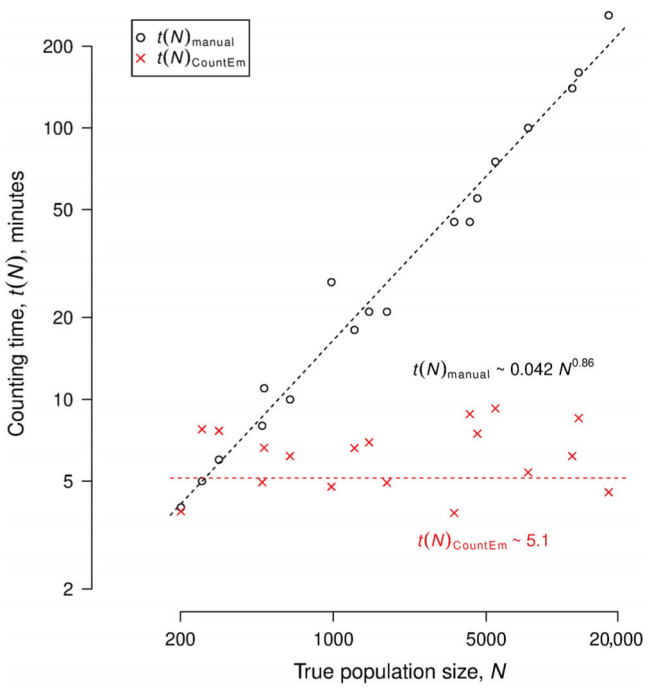

FIGURE 2 Manual (black circles) and CountEm (red crosses) counting time in minutes versus true flock size for 18 images of Greater Snow Goose flocks. Note that the axes are logarithmic.

Real mode–simulation mode consistency test

Real mode and *simulation mode* flock size estimations are plotted in Figure 3 to evaluate their consistency. Figure 3a shows the consistency of the variance within a single run of each mode. The agreement holds when comparing *real mode* data with 100 simulated estimations (Figure 3b). Moreover, a visual verification of the unbiasedness of the guided protocol is provided, since the mean of the simulated (crosses) and real (circles and triangles) estimations lies close to the corresponding true value (line). This test validated the use of *simulation mode* to assess the precision of the CountEm method.

Simulation mode precision assessment

We calculated the "true" relative SE, that is, the empirical coefficient of error, $\text{CE}_e(\widehat{N})$, for single images and the four multi-image groups described above, in order to evaluate the precision of the method. The results for the 106 single flock images and the four multi-image sets (marked in green) are shown in Figure 4. Eighty percent of the coefficients of error are below 10%, including the large GSGO group shown in Figure 1a. The predicted coefficient of error $\text{CE}(\widehat{N})$ obtained by applying

TABLE 1 Multi-image results.

Group name	$\widehat{N}$	Deviation	%	No. images	CPU time (s)	Q	Counting time (s)
COEIsg[a]	33,004	−3952	10.7	132	390	223	959
COEIg[b]	135,415	11,609	9.3	179	1040	371	1329
GSGOsg[a]	10,422	−1278	10.9	40	26	216	666
GSGOg[b]	527,790	21,692	4.3	99	135	219	1067

Abbreviations: COEI, Common Eiders; CPU, central processing unit; GSGO, Greater Snow Geese.
[a]COEI identifies images of Common Eider, and GSGO, images of Greater Snow Goose. Letters "sg" identify a group of images of all the small flocks (≤700 birds per image).
[b]Letters "g" identify a group of all the images in the data set, without limit in flock size.

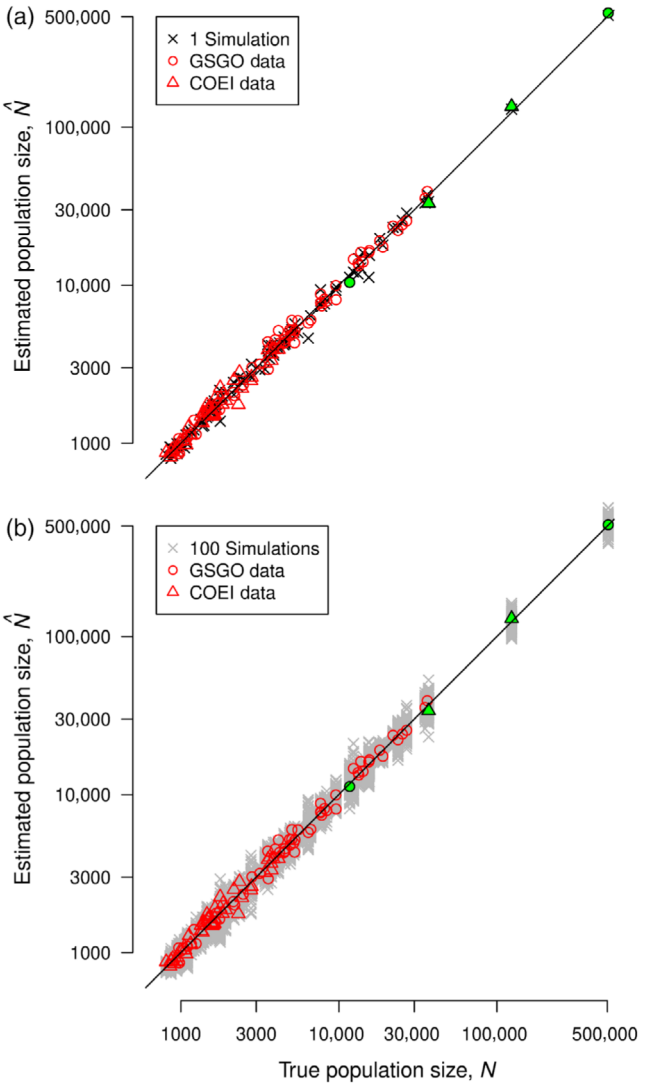

FIGURE 3 (a) Flock size estimations obtained with CountEm versus true flock size. *Real mode* data obtained from manually applying CountEm are marked in red, whereas one simulated estimation per image, obtained in *simulation mode*, is shown in black. The data marked in green correspond to the four multi-image groups. (b) Same as (a) with 100 simulated estimations plotted in gray. The black line represents the 1:1 line of equality. COEI, Common Eiders; GSGO, Greater Snow Geese.

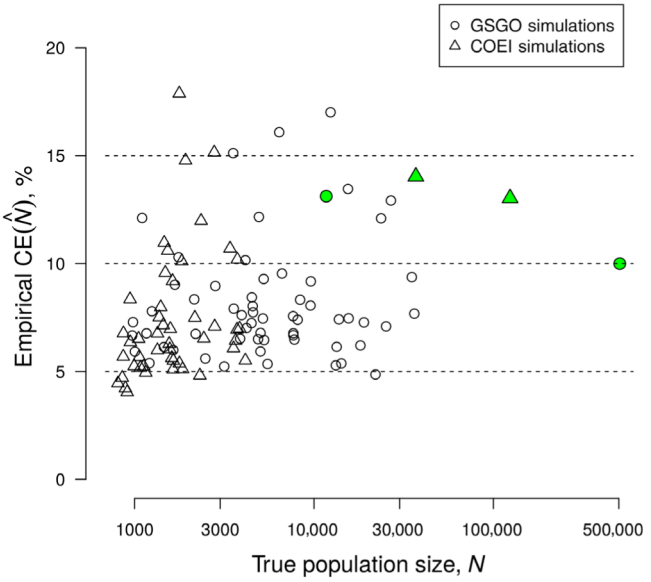

FIGURE 4 Empirical coefficient of error (CE) obtained with 2000 simulated measurements of the guided protocol in *simulation mode* versus true flock size. The data marked in green correspond to the four multi-image groups. COEI, Common Eiders; GSGO, Greater Snow Geese.

CountEm to these images is plotted against the empirical coefficient of error $\mathrm{CE_e}(\widehat{N})$ in Figure 5. The outlier with $\mathrm{CE}(\widehat{N}) > 25\%$ corresponds to image COEI_175 (Cruz et al., 2021), where the spatial distribution of the birds was particularly inhomogeneous.

DISCUSSION

The CountEm software has been empirically tested with a large set of images of flocks of birds and extended to multi-image analysis. The ECA Flocks data set allowed us to check the method across a wide range of flock sizes and spatial patterns. We evaluated the precision of CountEm in *real mode* (i.e., manually using the software

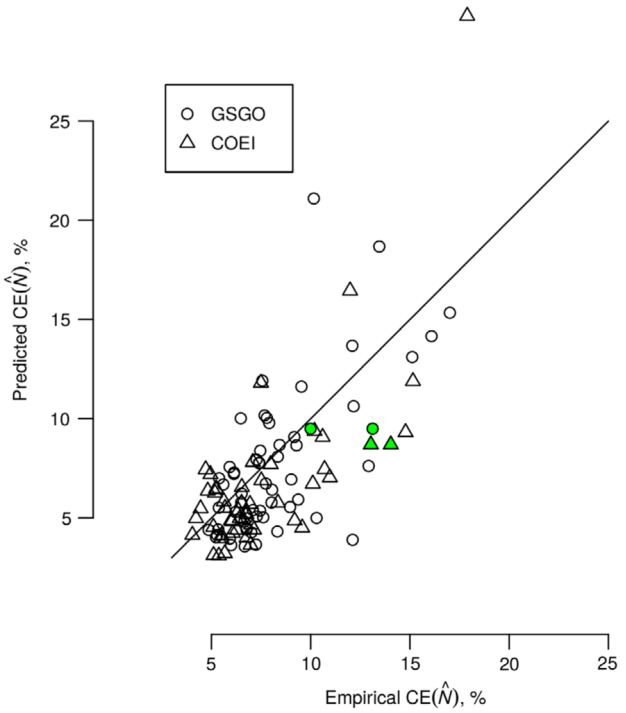

FIGURE 5 Estimated relative SE obtained from manually applying CountEm versus empirical coefficient of error (CE) (as in Figure 4). The data marked in green correspond to the four multi-image groups. The black line represents the 1:1 line of equality. COEI, Common Eiders; GSGO, Greater Snow Geese.

with the images) and in *simulation mode* (i.e., via Monte Carlo resampling of the point patterns that are obtained from the annotated positions). The typical relative SEs from CountEm on single images were in the 5%−10% range, with an average counting time of ∼5 min per image. The relative errors for large multi-image groups ranged between 10% and 15%, with counting times between 10 and 20 min. These results are convincing in terms of counting time and precision.

Visual inspection revealed that images with particularly inhomogeneous bird distributions lead to higher errors due to a lower number of non-empty quadrats. Some adjustments can be made depending on the spatial homogeneity of the data set, the desired coefficient of error, and counting time. For instance, inhomogeneous data would be required to increase the sample size, Q, and the number of non-empty quadrats, n. This can be achieved by manually modifying the initial number of quadrats, n_0, and sampling fraction, f, in the standard mode of the CountEm software. However, exact statements regarding the estimation variance under systematic sampling are generally not available.

Multi-image counting time depends on image quality, sample size, Q, and the number of images. A high number of images increase the difficulty of parameter selection and evaluation. For instance, GSGOg and GSGOsg have very similar image quality and sample size. However, the number of images and also counting time is larger for GSGOg. Low image quality for some images in COEIsg could be the cause of the higher counting time. However, the multi-image mode allows a reduction in total processing time (CPU time + counting time) by a factor of ∼20 with respect to iteratively analyzing single images in a data set of about 100 images.

Based on our results, CountEm is an attractive option for estimating the numbers of birds in flocks from imagery. Alternative methods exist, namely manual counting and computer vision methods. However, manual counts are time-consuming (Figure 2) and false detections and a relatively low accuracy across large spatial extents are currently limitations for automated methods (Hollings et al., 2018). Imagery collections such as ECA Flocks that contain large numbers of oblique aerial photos are challenging for computer vision methods, since these are sensitive to large lighting variations, varying resolutions, superimposed birds, and inhomogeneous backgrounds. Furthermore, they need specific remodeling if there is more than one cohort of interest in the image (e.g., multiple species, age, or sex). CountEm is not affected by these issues (see Figure 4), and it can be applied to any kind of particles, with the only requirement that all the particles in the flock should be unambiguously identifiable for counting. This requirement can be met when the flock sits in an open area (e.g., on water), which is often the case in the ECA Flocks data set. In addition, the CountEm software allows saving the results and the images, with the quadrats facilitating a fast verification of the result. In addition, the software provides an error estimation. As expected from results obtained with human crowd images (Gómez et al., 2019), the predicted error often underestimates the empirical, "true" value.

When making manual counts from images, observers can use other features in the image to make decisions on what is being counted. For example, manual counts are often used to estimate the numbers of breeding pairs from large colonies of nesting birds where the counting unit is not a single bird but occupied nesting sites (see Chardine et al., 2013). These conditions are challenging for computer vision methods as the counting unit (occupied nest sites) may consist of one or two birds close together within the breeding area of colonies, and birds that occur outside the nesting areas must be discarded. CountEm retains the human observer, and with the new multi-image mode, allows fast estimation for large groups of images, which is more efficient than iteratively analyzing each image individually. For instance, the GSGO group with 99 images (Figure 1a) had a total of

$N = 506{,}098$ manually counted birds and was estimated with CountEm, $\widehat{N} = 527{,}790$, in 20 min with ~200 bird counts and relative SE of ~10%. This performance is made possible because CountEm exploits the strength of systematic sampling. CountEm can be applied in any situation where large aggregations of animals can be captured in imagery. Its applications allow for estimating colony size, and CountEm can be particularly useful when an unbiased estimate of animals is needed rapidly to respond to an environmental emergency such as an oil spill (see Robertson et al., 2014).

ACKNOWLEDGMENTS

Marcos Cruz acknowledges funding from Proyecto Puente (Contrato Programa Gobierno de Cantabria—Universidad de Cantabria), Consejería de Universidades, Igualdad, Cultura y Deporte del Gobierno de Cantabria.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The ECA Flocks data set is publicly available at the Dryad repository: https://doi.org/10.5061/dryad.98sf7m0hx, as described in the data paper by Cruz et al. (2021). The CountEm source code and software are publicly available at countem.unican.es. Single-image CountEm results can be accessed on Zenodo: https://doi.org/10.5281/zenodo.5819210.

ORCID

Marcos Cruz https://orcid.org/0000-0002-4767-530X

REFERENCES

Baddeley, A., E. Rubak, and R. Turner. 2015. *Spatial Point Patterns: Methodology and Applications with R*. Boca Raton, FL: CRC Press.

Bechet, A., A. Reed, N. Plante, J.-F. Giroux, and G. Gauthier. 2004. "Estimating the Size of the Greater Snow Goose Population." *The Journal of Wildlife Management* 68: 639–49.

Bordage, D., N. Plante, A. Bourget, and S. Paradis. 1998. "Use of Ratio Estimators to Estimate the Size of Common Eider Populations in Winter." *The Journal of Wildlife Management* 62: 185–92.

Chabot, D., and D. M. Bird. 2015. "Wildlife Research and Management Methods in the 21st Century: Where Do Unmanned Aircraft Fit in?" *Journal of Unmanned Vehicle Systems* 3: 137–55.

Chabot, D., and C. M. Francis. 2016. "Computer-Automated Bird Detection and Counts in High-Resolution Aerial Images: A Review." *Journal of Field Ornithology* 87: 343–59.

Chardine, J. W., J.-F. Rail, and S. Wilhelm. 2013. "Population Dynamics of Northern Gannets in North America, 1984–2009." *Journal of Field Ornithology* 84: 187–92.

Cruz, M., D. Gómez, and L. M. Cruz-Orive. 2015. "Efficient and Unbiased Estimation of Population Size." *PLoS One* 10: 1–14.

Cruz, M., and J. González-Villa. 2018. "Simplified Procedure for Efficient and Unbiased Population Size Estimation." *PLoS One* 13: 1–11.

Cruz, M., and J. González-Villa. 2021. "Unbiased Population Size Estimation on Still Gigapixel Images." *Sociological Methods & Research* 50: 627–48.

Cruz, M., J. González-Villa, J. Lefebvre, S. Gilliland, F. St-Pierre, M. English, and C. Lepage. 2021. "Eastern Canada Flocks: Images and Manually Annotated Bird Positions." *Ecology* 102: e03444.

Cruz-Orive, L. 2017. "Stereology: A Historical Survey." *Image Analysis & Stereology* 36: 153–77.

Gómez, A. I., M. Cruz, and L. M. Cruz-Orive. 2019. "Variance Prediction for Population Size Estimation." *Image Analysis & Stereology* 38: 131–9.

González-Villa, J., and M. Cruz. 2020. "The CountEm Software: Simple, Efficient and Unbiased Population Size Estimation." *Ecography* 43: 251–5.

Gundersen, H. J. G. 1977. "Notes on the Estimation of the Numerical Density of Arbitrary Profiles: The Edge Effect." *Journal of Microscopy* 111: 219–23.

Hagy, H. M., C. S. Hine, M. M. Horath, A. P. Yetter, R. V. Smith, and J. D. Stafford. 2017. "Waterbird Response Indicates Floodplain Wetland Restoration." *Hydrobiologia* 804: 119–37.

Hodgson, J. C., R. Mott, S. M. Baylis, T. T. Pham, S. Wotherspoon, A. D. Kilpatrick, R. Raja Segaran, I. Reid, A. Terauds, and L. P. Koh. 2018. "Drones Count Wildlife More Accurately and Precisely Than Humans." *Methods in Ecology and Evolution* 9: 1160–7.

Hollings, T., M. Burgman, M. van Andel, M. Gilbert, T. Robinson, and A. Robinson. 2018. "How Do you Find the Green Sheep? A Critical Review of the Use of Remotely Sensed Imagery to Detect and Count Animals." *Methods in Ecology and Evolution* 9: 881–92.

Howard, C. V., and M. G. Reed. 2005. "Unbiased Stereology." In *Three-Dimensional Measurement in Microscopy*, 2nd ed. Oxford: Bios/Taylor & Francis.

Lyons, M. B., K. J. Brandis, N. J. Murray, J. H. Wilshire, J. A. McCann, R. T. Kingsford, and C. T. Callaghan. 2019. "Monitoring Large and Complex Wildlife Aggregations with Drones." *Methods in Ecology and Evolution* 10(7): 1024–35.

Robertson, G. J., S. G. Gilliland, P. C. Ryan, J. Dussureault, K. Power, and B. C. Turner. 2014. "Mortality of Common Eider, *Somateria mollissima* (Linnaeus, 1758), and Other Water Birds during Two Inshore Oiling Events in Southeastern Newfoundland, 2005 and 2006." *The Canadian Field-Naturalist* 128: 235–42.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. "Nih Image to Imagej: 25 Years of Image Analysis." *Nature Methods* 9: 671–5.

Seavy, N. E., and M. H. Reynolds. 2007. "Is Statistical Power to Detect Trends a Good Assessment of Population Monitoring?" *Biological Conservation* 140: 187–91.

Sebastián-González, E., and A. J. Green. 2016. "Reduction of Avian Diversity in Created Versus Natural and Restored Wetlands." *Ecography* 39: 1176–84.

Zhao, Q., E. Silverman, K. Fleming, and G. S. Boomer. 2016. "Forecasting Waterfowl Population Dynamics under Climate Change-Does the Spatial Variation of Density

Dependence and Environmental Effects Matter?" *Biological Conservation* 194: 80–8.

Zimmerman, G. S., J. R. Sauer, W. A. Link, and M. Otto. 2012. "Composite Analysis of Black Duck Breeding Population Surveys in Eastern North America." *The Journal of Wildlife Management* 76: 1165–76.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cruz, Marcos, Javier González-Villa, Josée Lefebvre, Scott G. Gilliland, Francis St-Pierre, Matthew English, and Christine Lepage. 2022. "Multi-Image Flock Size Estimation with CountEm: A Case Study with Half a Million Common Eiders and Greater Snow Geese." *Ecosphere* 13(8): e4174. https://doi.org/10.1002/ecs2.4174