Title

Drones count wildlife more accurately and precisely than humans

Authors

Jarrod C. Hodgson[1*], Rowan Mott[2], Shane M. Baylis[2], Trung T. Pham[3],

Simon Wotherspoon[4, 5], Adam D. Kilpatrick[1], Ramesh Raja Segaran[1], Ian Reid[3],

Aleks Terauds[5] & Lian Pin Koh[1].

Affiliations

[1]School of Biological Sciences, University of Adelaide, Adelaide, SA 5005, Australia.

[2]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia.

[3]School of Computer Science, University of Adelaide, Adelaide, SA 5005, Australia.

[4]Institute of Marine and Antarctic Studies, University of Tasmania, Hobart, TAS 7001, Australia.

[5]Antarctic Conservation and Management, Australian Antarctic Division, Department of the Environment and Energy, Kingston, TAS 7050, Australia.

Corresponding author

Jarrod C. Hodgson

School of Biological Sciences, University of Adelaide, SA 5005, Australia.

Email: jarrod.hodgson@adelaide.edu.au

1   **Abstract**

2   Ecologists are increasingly using technology to improve the quality of data collected on wildlife,

3   particularly for assessing the environmental impacts of human activities. Remotely Piloted

4   Aircraft Systems (RPAS; commonly known as 'drones') are widely touted as a cost-effective

5   way to collect high quality wildlife population data, however, the validity of these claims is

6   unclear. Using life-sized seabird colonies containing a known number of replica birds, we show

7   that RPAS-derived data are, on average, between 43% and 96% more accurate than data from

8   the traditional ground-based collection method. We also demonstrate that counts from this

9   remotely sensed imagery can be semi-automated with a high degree of accuracy. The

10   increased accuracy and precision of RPAS-derived wildlife monitoring data provides greater

11   statistical power to detect fine-scale population fluctuations allowing for more informed and

12   proactive ecological management.

13   **Keywords**

14   bird, drones, ecology, population monitoring, RPAS, UAV, UAS, wildlife

15 **Introduction**

16 Wildlife populations are undergoing dramatic declines in response to a wide range of

17 human-induced threats (Dirzo *et al.* 2014; Tilman *et al.* 2017). High quality ecological data

18 are vital to monitor such changes. Emerging technologies, such as camera traps (Rowcliffe &

19 Carbone 2008) and radio telemetry (Hussey *et al.* 2015; Kays *et al.* 2015), have increasingly

20 been used to address this challenge (Moll *et al.* 2007; Hebblewhite & Haydon 2010; Pimm *et*

21 *al.* 2015), especially when wildlife populations are rapidly fluctuating, highly mobile or

22 located in remote habitats.

23 Remotely Piloted Aircraft Systems have been heralded as a game changer in ecology (Jones,

24 Pearlstine & Percival 2006; Watts *et al.* 2010; Koh & Wich 2012; Anderson & Gaston 2013;

25 Marris 2013; Chabot & Bird 2015; Linchant *et al.* 2015; Christie *et al.* 2016). They are used

26 for data collection in an increasingly diverse suite of ecological applications, including

27 identification of floristic biodiversity of understorey vegetation (Getzin, Wiegand &

28 Schöning 2012), monitoring for poaching activities (Mulero-Pazmany *et al.* 2014), and bird

29 surveys (Sarda-Palomera *et al.* 2012; Chabot, Craik & Bird 2015). However, there has been

30 little consideration of the quality of data obtained using RPAS compared to more

31 conventional methods (see Hodgson *et al.* (2016a) for an exception).

32 We assessed the accuracy of RPAS-facilitated wildlife population monitoring in comparison

33 with the traditional ground-based counting method. The task for both approaches was to

34 derive an estimate of the size (i.e. number of individuals) of 10 replica seabird colonies. Each

35 replica colony had a different known number of life-sized individuals. We hypothesised that

36 RPAS-derived counts would be more accurate and more precise than those generated using

37 the traditional approach, confirming RPAS-technology as revolutionary for ecological

38 monitoring.

39   **Materials and methods**

40   *Study site and simulated colony set-up*

41   Fieldwork (#epicduckchallenge) was completed at a metropolitan beach in South Australia

42   (Port Willunga, 35°15'33 S, 138°27'41 E) in accordance with relevant permits (Department of

43   Environment, Water and Natural Resources scientific research permit: M26523-1; City of

44   Onkaparinga location permit: 4138). The experimental design, including the majority of

45   anticipated statistical analyses, was pre-registered (Hodgson *et al.* 2016b).

46   Ten simulated Greater Crested Tern *Thalasseus bergii* breeding colonies were constructed

47   using commercial, life-size, plastic duck decoys (~ 25.5 x 11.3 cm, 185 cm^2 footprint).

48   Decoys provided a realistic representation of the nesting seabird stimuli observers encounter

49   in the field. Colonies were situated separately on the beach, above the high water mark, in

50   sandy areas that represented analagous nesting habitat. These were typically devoid of

51   vegetation but often contained natural beach debris.

52   As inter-indiviudal interactions are thought to influence colony layout, a model of nesting

53   pressure was applied to an underlying hexagonal grid to generate unique, unbiased colony

54   layouts (Hodgson *et al.* 2016b). The hexagonal grid was re-created in the field using a wire

55   mesh, upon which grid cell centres were marked (mean density: 11.39 m^{-2}). Pre-counted

56   wooden skewers were placed one per cell at a random location within all cells identified as

57   occupied in the colony layout map. The mesh was removed and each skewer was replaced

58   with a decoy facing approximately into the wind. The number of skewers retrieved was taken

59   to be the true number of individuals in the colony. Colony sizes were between 463 and 1017

60   individuals. One individual was placed in each occupied cell.

61   *Ground counting approach*

62   Ground counts were made by experienced seabird counters using a standard field technique

63   (Hodgson *et al.* 2016a). Counters used tripod-mounted spotting scopes or binoculars as

64   required. Hand-held tally counters were used to assist counting. Observation viewpoints at

65   similar altitude to the colony and which provided the optimum vantage were selected (Fig. 1).

66   Viewpoints were positioned 37.5 m from the nearest bird in the colony – this is the flight

67   initiation distance for Caspian Tern *Hydroprogne caspia* (Moller *et al.* 2014) and so is a

68   biologically-plausible minimum approach distance for a similar species in the field. Counts

69   ($n = 61$) were 7 ± 2.65 min (s.d.) in duration. Four to seven counters each made a single blind

70   count of the number of individuals in each colony. The numbers of counters were selected

4

71 based on a preliminary power analysis (Hodgson *et al.* 2016b) which investigated the sample

72 sizes necessary to detect small (~ 10%) differences in mean counts and count variances

73 between ground and RPAS-derived counts to high (80%, 90%, and 95%) power. Counters

74 had no knowledge of the true number of individuals in colonies or the colony set-up

75 technique. Counts were made between 0930 and 1645 on one day, resulting in variation in

76 illumination and shadows.

RPAS description, flight characteristics and data collected by RPAS

78 A small, off-the-shelf quadcopter (Iris+, 3D Robotics) was used as a platform to image each

79 colony. After positioning the RPAS in the centre of the colony at 15 m above ground level, it

80 was piloted in altitude hold mode to make a vertical ascent without movement in other axes.

81 The RPAS was loitered for short periods (~ 10 seconds) to enable the capture of several

82 photographs at 30 m, 60 m, 90 m and 120 m above ground level (sample heights). Sampling

83 was restricted to a height of 120 m as this is a common maximum limit for standard RPAS

84 flight. Ground control station connection (Mission Planner, planner.ardupilot.com) was

85 utilised and total flight time for missions was 5-7 min. All missions were in accordance with

86 local regulations and flown by the same licenced pilot. Samples were collected within 40 min

87 of the completion of ground counts.

88 Imagery was captured using a compact digital camera (Cyber-shot RX100 III, Sony –

89 resolution: 5,472 x 3,648 px; sensor: CMOS; sensor size: 13.2 x 8.8 mm; lens: ZEISS

90 Vario-Sonnar T). Exposure time was set at 1/2000 seconds using shutter priority mode.

91 Photographs were captured successively (~ 1 sec intervalometer) using the Sony

92 PlayMemories Time-lapse application in jpeg format and at minimum focal length (8.8 mm).

93 The camera was mounted facing downward using a custom vibration dampening plate. The

94 footprint of a single image at each height encompassed the colony for all replicates. For

95 analysis, only the image captured closest to the middle of the loiter time period for each

96 sample height was used. These images (scenes; $n = 40$) were cropped (colony area < 50% of

97 footprint) so that the image footprint was identical for each sample height for a given colony.

98 High quality imagery was obtained for six of the ten colonies. Imagery for the remaining four

99 colonies was affected by vibration-blur caused by a failure of the sensor attachment, likely

100 due to wind speeds near the limit of the capability of the RPAS platform. Scenes are archived

101 online (Pham & Hodgson 2017).

102   The ground sample distance (GSD), being the distance between adjacent pixel centres on the

103   ground, for sample heights were 0.82 cm, 1.64 cm, 2.47 cm and 3.29 cm (Fig. 1). When

104   photographed at nadir, this approximated to 275, 69, 30 and 17 pixels per individual

105   respectively. The variance in GSDs was intended to represent the resolutions commonly

106   achieved in wildlife monitoring applications, which result from sensor and sampling height

107   variations.

108   *Manual RPAS image counting approach*

109   Manual counts of perceived individuals in digital imagery were completed following a

110   technique previously implemented for RPAS-derived monitoring of living seabirds (Hodgson

111   *et al.* 2016a). Systematic counts were made using the multi-count tool within an open source,

112   java-based scientific image processing program (ImageJ, http://imagej.net/). A grid plugin

113   was used to overlay a square matrix (cell sizes: 70,000, 15,000, 8,000 and 4,000 pixels for

114   each sample height) and counters were instructed to view the colony sequentially (gridcell-

115   by-gridcell: left to right, top to bottom). Counters were encouraged to zoom in to each cell as

116   they progressed and, upon completion, review their count at different levels of zoom until

117   they were satisfied they had counted all individuals. For each sample height, seven to nine

118   individuals counted each colony. Counters had no knowledge of the experimental setup and

119   only one had experience ground counting colonial birds.

120   *Semi-automated aerial image counting approach*

121   In each scene, digital bounding boxes were used to manually delimit a percentage of

122   individual birds (Supplementary Fig. 1a). Areas of background were also delimited. These

123   data were used to train a linear support-vector machine (a discriminative classifier; Cortes &

124   Vapnik 1995), which predicted the likelihood of each pixel being a bird or background when

125   applied to the corresponding scene (Supplementary Fig. 1b). Instead of relying on colour

126   intensities, we computed rotation-invariant Fourier histogram of oriented gradient (Liu et al.

127   2013) features for each pixel used in the training processes. This resulted in the classifiers

128   being trained to determine which features distinguished birds from the background. The

129   predicted likelihood (score) maps indicated the approximate locations of birds in the scenes,

130   and detections were generated by applying a threshold to the likelihood maps. This process

131   unavoidably resulted in redundant bird proposals (Supplementary Fig. 1c) and so the final

132   detection results were obtained by suppressing redundant proposals via minimising an energy

133   function (Pham *et al.* 2016; Supplementary Fig. 1d). This function encoded the spatial

134   distribution of objects and is informed by our knowledge of how the birds nest (e.g. two birds

135   cannot occupy the same location). The source code and dataset are archived online (Pham &

136   Hodgson 2017).

137   To determine the minimum amount of training data required for accurate detections relative

138   to manual image counts, we varied the percentage of individual birds used as training data

139   between 1% and 30% for each scene.

140   *Statistical methods*

141   All analyses were carried out in *R* version 3.2.2 (R Core Team 2016). Pre-registered analyses

142   were designed to investigate how within-colony absolute count error, within-colony

143   variability of counts, and within-colony bias of counts differed between count techniques

144   (Hodgson *et al.* 2016b).

145   For each test, a generalised linear mixed model was fit between the response (e.g. absolute

146   count error) and the technology used to make the count (e.g. ground-count, manually counted

147   RPAS at 30 m height, semi-automatically counted RPAS at 30 m height), with colony

148   included in the model as a random effect (Supplementary Information 1). To investigate

149   effects of counting technique on absolute count error, we defined the response as the absolute

150   difference between the true number of birds in a colony and the counted number of birds. To

151   investigate effects of counting technique on count-variability, we defined the response as the

152   absolute difference between each count and the mean of counts of the same colony taken

153   using the same method. Count variability was not estimated for semi-automated counts as

154   there was only a single semi-automated count per colony. To investigate the effect of

155   counting technique on relative count bias, we defined the response as the difference between

156   the true number of birds in the colony and the counted number of birds. For the absolute

157   count error model, we used a Quasipoisson distribution, and for the variability and bias

158   models, we used a Gaussian distribution. For each model, post-hoc Tukey tests were used to

159   test for differences in the response between all pairs of treatment levels.

160   Semi-automated count data were added to the experimental design subsequent to our

161   pre-registration of the analysis which necessitated minor analysis modification. The addition

162   of semi-automated count data, with a single replicate per colony, required fitting colony ID as

163   a random effect instead of as a fixed effect in each model.

7

164   Statements comparing the accuracy of RPAS-derived counts to ground-based counts are

165   based on the mean within-colony Root Mean Squared Error (RMSE) of that counting

166   approach, standardised as a proportion of the true count within each colony (Supplementary

167   Information 2). For instance, a statement that RPAS-derived counts are '95% more accurate

168   than ground-counts' means that, within-colony, the RMSE for RPAS-derived counts is 5%

169   the RMSE for ground-based counts, representing a 95% reduction in RMSE.

170   To investigate the probability of counting each individual correctly, we developed models

171   with a variety of possible counting outcomes for each object (Supplementary Information 3).

172   We assumed that there are $N = n0 + n1 + n2$ objects that are counted by an observer, of which

173   n1 are counted correctly, n0 are missed and n2 are double counted. We assumed that $n = (n0,$

174    $n1, n2)$ is multinomially distributed with probability $p = (p0, p1, p2)$. In this structure, n0, n1,

175   n2 are latent and the observer can only report the total count $M = n1 + 2n2$. Allowing each

176   object to be at most double counted constrains n considerably, and the probability mass

177   function can then be formulated. We adopted a Dirichlet prior for p ($p \sim Dirichlet(a)$) making

178   the conditional distribution of p: $p|n(k) \sim Dirichlet(a+n(k))$.

179   We then ran a Gibbs sampling routine for p by alternately sampling n(k) and p from these

180   two distributions. Since there was considerably more variation between ground counts

181   compared to manual RPAS counts, we ran the analyses for each ground counter separately,

182   and pooled data across counters for manual RPAS counts. The RPAS analyses were run

183   separately for each sample height, and two sets of analyses were undertaken, one with data

184   from all colonies, and the other with data from the subset of colonies with high quality

185   imagery. This statistical approach does not account for objects being mistakenly identified as

186   birds (i.e. false positives). Furthermore, as this model assumes each individual is counted

187   independently, which may not always be true (particularly for ground counters who tend to

188   count in clusters), care needs to be taken in the interpretation of the estimated probability

189   values (and their variance).

190   To compare the semi-automated counts to that of the people counting the images, we first

191   took the semi-automated count after 10% of training data had been used for each scene. Ten

192   percent of training data was consistently identified as a threshold over which little

193   improvement in counts occurred for all scenes. We compared this count to each of the manual

194   counts of the same image using ANOVA for all scenes, and also for those scenes of high

195   quality. We also used Poisson generalized linear models to make more quantitative

196   comparisons of the two approaches.

197   *Code availability*

198   R scripts used for analyses are available in the Supplementary Information.

199   *Data availability*

200   The pre-registered experimental design is available via the Open Science Framework (*URL

201   to be made public on publication*). The count data, scenes and code for the image-analysis

202   algorithm are archived online (*URL to be made public on publication*).

203   **Results**

204   *Manual RPAS-derived image counting versus ground counts*

205   On average across all colonies, RPAS-derived counts were between 43% and 96% more

206   accurate than ground counts, depending on the sample height (between 92% and 98% for the

207   colonies with high quality imagery; Supplementary Table 1). The mean absolute error was

208   significantly smaller for RPAS-derived counts at all heights compared to ground counts (all

209    $P < 0.001$; Fig. 2a).

210   No significant increase in count accuracy was achieved by obtaining imagery from heights

211   lower than or equal to 90 m. Using data only from colonies with high quality imagery, there

212   was no significant change in count accuracy across the range of heights. The lower accuracy

213   of ground counts was due to significant underestimations of the true number of individuals in

214   colonies (Fig. 2b). RPAS-derived counts from imagery obtained at 30 m and 60 m did not

215   significantly under- or overestimate the true number of individuals in a colony, and there was

216   no evident bias in RPAS-derived counts at any height for colonies with high quality imagery

217   (Fig. 2b).

218   Using data from all colonies, RPAS-derived counts from 30 m and 60 m had a much higher

219   probability (90% and 50%, respectively) of correctly counting an individual than counts from

220   ground observers ($< 10\%$) (Fig. 3). However, 90 m and 120 m probabilities were largely

221   indistinguishable from the ground count probabilities, with a slightly higher likelihood of

222   missing individuals compared to counting them twice (Supplementary Fig. 3, 4). Colony

223   counts made from high quality imagery had a much higher probability of individuals being

224   counted correctly, with $> 85\%$ probability of correctly counting an individual at all heights

225   (Fig. 3). By contrast, ground counts had a low probability of counting an individual correctly,

226   with the probability of double counting and missing an individual varying considerably

227   between observers (Fig. 3 and Supplementary Fig. 3, 4).

228   RPAS-derived counts were more precise (i.e. had lower inter-observer variability) than

229   ground counts, regardless of the height at which imagery was obtained ($t_{4,560}$ -10.21 to -13.37,

230   all $P < 0.001$; Supplementary Fig. 5). RPAS-derived counts were more precise for imagery

231   obtained at 30 m compared to those obtained from 120 m ($P = 0.01$), however, there were no

232   significant differences in precision among RPAS-derived counts at different heights for

233   colonies with high quality imagery (all $P > 0.98$).

234   *Semi-automated RPAS approach*

235   By increasing the percentage (1 – 30%) of individuals used as training data for the image-

236   analysis algorithm, 10% training data was consistently identified as a threshold above which

237   little improvement in count accuracy was achieved in this semi-automated approach

238   (Supplementary Fig. 2). There was no significant difference between counts that were made

239   with 10% training data and those made by manual counting of RPAS-imagery across all

240   scenes. The semi-automated results were 94% similar to manual counts across all scenes

241   (98% for the colonies with high quality imagery; see also Supplementary Table 1).

242   **Discussion**

243   RPAS-derived data were more accurate and more precise than the traditional data collection

244   method validating claims that RPAS will be a revolutionary tool for ecologists. Never has the

245   importance of accurate wildlife population monitoring data been greater than at present given

246   the alarming population declines observed in animal species across the globe (Dirzo *et al.*

247   2014). By facilitating accurate census, RPAS will provide ecologists with confidence in

248   population estimates from which management decisions are made. Furthermore, the superior

249   precision of RPAS-derived counts increases statistical power to detect population trends,

250   owing to the lower type II error rate in statistical analysis that comes with comparing

251   measures with smaller variance (Gerrodette 1987). The improved precision of wildlife

252   population census using RPAS has been demonstrated for free-living seabird colonies

253   (Hodgson *et al.* 2016a) suggesting our results are generalizable to natural settings.

254   Differences in accuracy and precision between RPAS-facilitated and traditional survey

255   methods can be attributed to the sources, and magnitude, of variance introduced into the two

256   approaches which are strongly affected by the different vantages of the two methods

257   (Hodgson *et al.* 2016a).

258   We have conducted two independent analyses of how count error differs across count

259   approaches: a Frequentist analysis which estimates mean absolute count error, and a Bayesian

260   analysis which estimates the probability of double-counting or missing individual animals.

261   The two analyses are in agreement on the broad patterns: RPAS-derived counts are estimated

262   to have lower error than ground counts in both analyses, and the error-rate is fairly insensitive

263   to sample height for RPAS-derived counts. The Bayesian analysis makes restrictive

264   assumptions about the process by which counting errors occur, and these assumptions may

265   not fully reflect real-world counting processes. Nevertheless, the Bayesian analyses provide a

266   first estimate of the extent to which overall count accuracy is dependent on the

267   double-counting of some individuals cancelling out the effect of missing others. As an

268   extreme example, our analysis suggests that < 10% of animals are correctly classified as a

269   single animal in typical ground counts.

270   Manual counting of RPAS-derived imagery returned high quality data, but also involved

271   substantial labour investments. Recent advances in digital sensors and image-analysis

272   techniques have been increasingly employed to streamline the detection process (Chabot &

273   Francis 2016). By applying a semi-automated image-based object detection algorithm to each

274   scene, we vastly improved efficiency compared to the manual RPAS-derived census.

275   Importantly, the reduction in person-hours provided by this semi-automated approach did not

276   diminish data quality. This will be of particular interest in today's research environment

277   where funding for conservation is limited (Waldron *et al.* 2013) and researchers are under

278   ever more pressing time commitments (Fischer, Ritchie & Hanspach 2012).

279   The capture quality and resolution of RPAS-derived imagery heavily influenced the results of

280   both human and semi-automated detection. Consequently, ecologists should determine the

281   minimum required GSD for their context, and optimise their sensor accordingly (e.g.

282   resolution, focal length) relative to sample height. When determining an appropriate sample

283   height, best practice protocols should be considered to minimise potential disturbance to

284   wildlife (Hodgson & Koh 2016), while complying with relevant local aviation legislation and

285   achieving an acceptable sample area within the possible survey time period.

286   The ability to collect data with higher accuracy, higher precision, and less bias than the

287   existing approach confirms that RPAS are a scientifically rigorous data collection tool for

288   wildlife population monitoring. This approach also produces a permanent record, providing

289   the unique opportunity to error-check, and even recount with new detection methods, unlike

290   ground count data. As RPAS platforms, sensors and computer vision techniques continue to

291   develop, it is likely that the accuracy and cost effectiveness of RPAS-based approaches will

292   also continue to improve.

300   **Author contributions**

301   J.C.H., R.M., S.M.B., A.T. and L.P.K. designed the study, analysed the data and wrote the

302   manuscript. A.D.K. assisted with designing the study. J.C.H., R.M., S.M.B., A.D.K., R.R.S.

303   and L.P.K. collected the data. T.T.P., J.C.H., L.P.K. and I.R. developed the semi-automated

304   detection technique. S.W. and A.T. completed the multinomial analyses. All authors

305   contributed to drafting the manuscript.

306   **References**

307   Anderson, K. & Gaston, K.J. (2013) Lightweight unmanned aerial vehicles will revolutionize

308   spatial ecology. *Frontiers in Ecology and the Environment,* **11,** 138-146.

309   Chabot, D. & Bird, D.M. (2015) Wildlife research and management methods in the 21st

310   century: Where do unmanned aircraft fit in?1. *Journal of Unmanned Vehicle Systems,*

311   **3,** 137-155.

312   Chabot, D., Craik, S.R. & Bird, D.M. (2015) Population census of a large common tern

313   colony with a small unmanned aircraft. *PloS One,* **10,** e0122588.

314   Chabot, D. & Francis, C.M. (2016) Computer-automated bird detection and counts in high-

315   resolution aerial images: a review. *Journal of Field Ornithology,* **87,** 343-359.

316   Christie, K.S., Gilbert, S.L., Brown, C.L., Hatfield, M. & Hanson, L. (2016) Unmanned

317   aircraft systems in wildlife research: current and future applications of a transformative

318   technology. *Frontiers in Ecology and the Environment,* **14,** 241-251.

319   Cortes, C. & Vapnik, V. (1995) Support-vector networks. *Machine Learning,* **20,** 273-297.

320   Dirzo, R., Young, H.S., Galetti, M., Ceballos, G., Isaac, N.J.B. & Collen, B. (2014)

321   Defaunation in the Anthropocene. *Science,* **345,** 401-406.

322   Fischer, J., Ritchie, E.G. & Hanspach, J. (2012) Academia's obsession with quantity. *Trends*

323   *in Ecology & Evolution,* **27,** 473-474.

324   Gerrodette, T. (1987) A power analysis for detecting trends. *Ecology,* **68,** 1364-1372.

325   Getzin, S., Wiegand, K. & Schöning, I. (2012) Assessing biodiversity in forests using very

326   high-resolution images and unmanned aerial vehicles. *Methods in Ecology and*

327   *Evolution,* **3,** 397-404.

328   Hebblewhite, M. & Haydon, D.T. (2010) Distinguishing technology from biology: a critical

329   review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the*

330   *Royal Society of London. Series B: Biological Sciences,* **365,** 2303-2312.

331   Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. & Clarke, R.H. (2016a) Precision wildlife

332   monitoring using unmanned aerial vehicles. *Scientific Reports,* **6,** 22574.

333   Hodgson, J.C., Baylis, S.M., Mott, R. & Koh, L.P. (2016b) A comparison of the accuracy of
334   simulated animal counts using traditional and UAV-assisted methods. (*URL to be
335   made public on publication*). Open Science Framework.

336   Hodgson, J.C. & Koh, L.P. (2016) Best practice for minimising unmanned aerial vehicle
337   disturbance to wildlife in biological field research. *Current Biology,* **26,** R404-405.

338   Hussey, N.E., Kessel, S.T., Aarestrup, K., Cooke, S.J., Cowley, P.D., Fisk, A.T., Harcourt,
339   R.G., Holland, K.N., Iverson, S.J., Kocik, J.F., Mills Flemming, J.E. & Whoriskey,
340   F.G. (2015) Aquatic animal telemetry: A panoramic window into the underwater world.
341   *Science,* **348,** 1255642.

342   Jones, G.P., Pearlstine, L.G. & Percival, H.F. (2006) An assessment of small unmanned aerial
343   vehicles for wildlife research. *Wildlife Society Bulletin,* **34,** 750-758.

344   Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015) Terrestrial animal tracking as an
345   eye on life and planet. *Science,* **348,** aaa2478.

346   Koh, L.P. & Wich, S.A. (2012) Dawn of drone ecology: low-cost autonomous aerial vehicles
347   for conservation. *Tropical Conservation Science,* **5,** 121-132.

348   Linchant, J., Lisein, J., Semeki, J., Lejeune, P. & Vermeulen, C. (2015) Are unmanned
349   aircraft systems (UASs) the future of wildlife monitoring? A review of
350   accomplishments and challenges. *Mammal Review,* **45,** 239-252.

351   Liu, K., Skibbe, H., Schmidt, T., Blein, T., Palme, K., Brox, T. & Ronneberger, O. (2013)
352   Rotation-invariant HOG descriptors using fourier analysis in polar and spherical
353   coordinates. *International Journal of Computer Vision,* **106,** 342-364.

354   Marris, E. (2013) Fly, and bring me data. *Nature,* **498,** 156-158.

355   Moll, R.J., Millspaugh, J.J., Beringer, J., Sartwell, J. & He, Z. (2007) A new 'view' of ecology
356   and conservation through animal-borne video systems. *Trends in Ecology & Evolution,*
357   **22,** 660-668.

358   Moller, A.P., Samia, D.S., Weston, M.A., Guay, P.J. & Blumstein, D.T. (2014) American
359   exceptionalism: population trends and flight initiation distances in birds from three
360   continents. *PloS One,* **9,** e107883.

361   Mulero-Pazmany, M., Stolper, R., van Essen, L.D., Negro, J.J. & Sassen, T. (2014) Remotely
362   piloted aircraft systems as a rhinoceros anti-poaching tool in Africa. *PloS One,* **9,**
363   e83873.

364   Pham, T.T. & Hodgson, J.C. (2017) Semi-automated bird detection from drone-derived
365   imagery. (*URL to be made public on publication*).

366   Pham, T.T., Rezatofighi, S.H., Reid, I. & Chin, T.-J. (2016) Efficient point process inference
367   for large-scale object detection. *CVPR.*

368   Pimm, S.L., Alibhai, S., Bergl, R., Dehgan, A., Giri, C., Jewell, Z., Joppa, L., Kays, R. &
369   Loarie, S. (2015) Emerging technologies to conserve biodiversity. *Trends in Ecology &*
370   *Evolution,* **30,** 685-696.

371   R Core Team (2016) R: A language and environment for statistical computing. Vienna,
372   Austria. *http://www.R-project.org*.

373   Rowcliffe, J.M. & Carbone, C. (2008) Surveys using camera traps: are we looking to a
374   brighter future? *Animal Conservation,* **11,** 185-186.

375   Sarda-Palomera, F., Bota, G., Vinolo, C., Pallares, O., Sazatornil, V., Brotons, L., Gomariz,
376   S. & Sarda, F. (2012) Fine-scale bird monitoring from light unmanned aircraft systems.
377   *Ibis,* **154,** 177-183.

378   Tilman, D., Clark, M., Williams, D.R., Kimmel, K., Polasky, S. & Packer, C. (2017) Future
379   threats to biodiversity and pathways to their prevention. *Nature,* **546,** 73-81.

380   Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts,
381   J.T. & Gittleman, J.L. (2013) Targeting global conservation funding to limit immediate
382   biodiversity declines. *Proceedings of the National Academy of Sciences,* **110,** 12144-
383   12148.

384   Watts, A.C., Perry, J.H., Smith, S.E., Burgess, M.A., Wilkinson, B.E., Szantoi, Z., Ifju, P.G.
385   & Percival, H.F. (2010) Small unmanned aircraft systems for low-altitude aerial
386   surveys. *Journal of Wildlife Management,* **74,** 1614-1619.
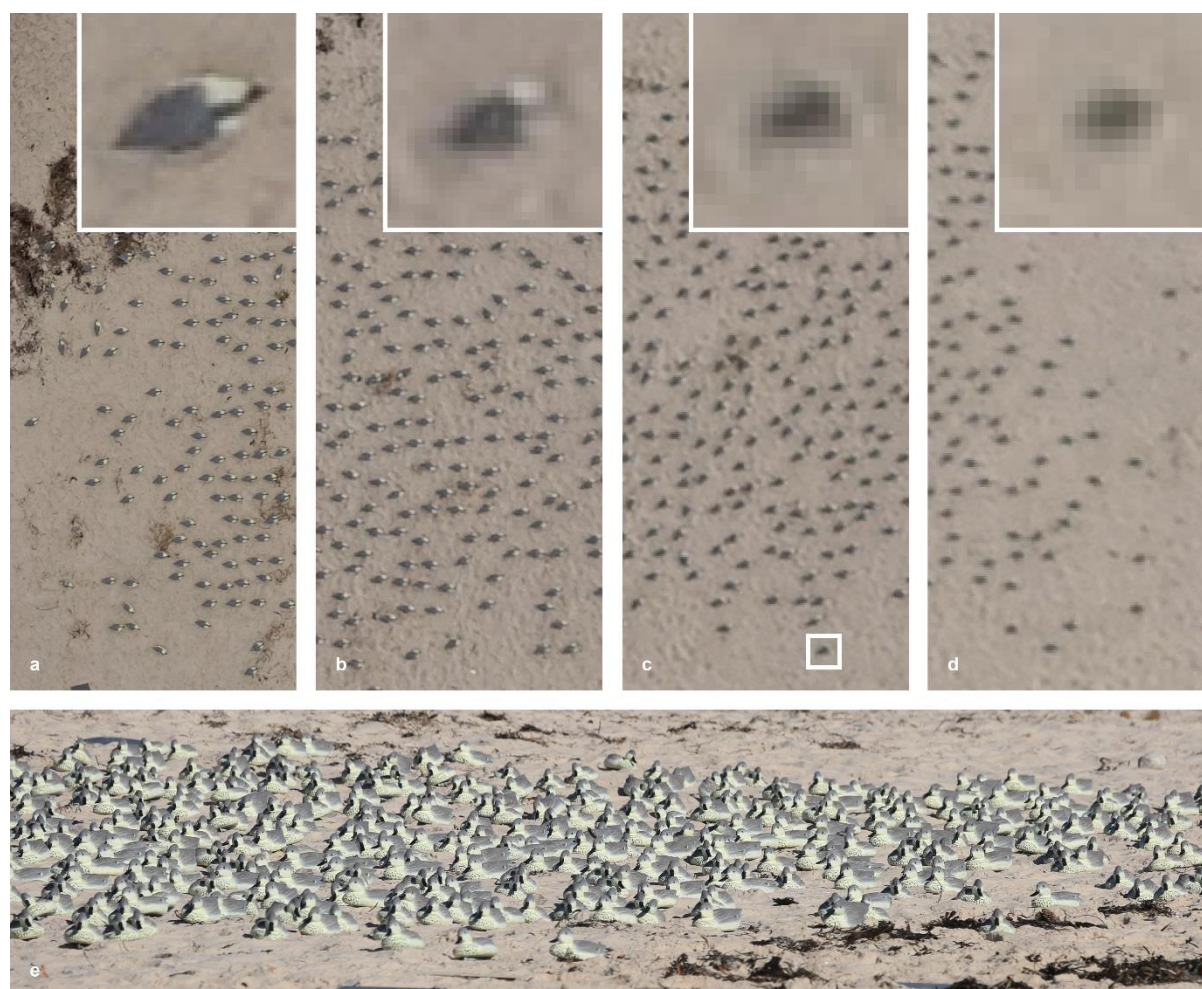
Figure 1: Aerial vantage of a replica seabird colony compared with the ground counter's viewpoint. One colony represented by a mosaic of images (**a-d**) photographed from a RPAS-mounted camera at varying heights (30 m, 60 m, 90 m and 120 m) and resulting ground sample distances (GSD; 0.82 cm, 1.64 cm, 2.47 cm and 3.29 cm). Insets are of the same individual (square; **c**) at each height, displaying the decrease in resolution relative to an increase in GSD. **e**, View of the colony from a ground counter's standing position.
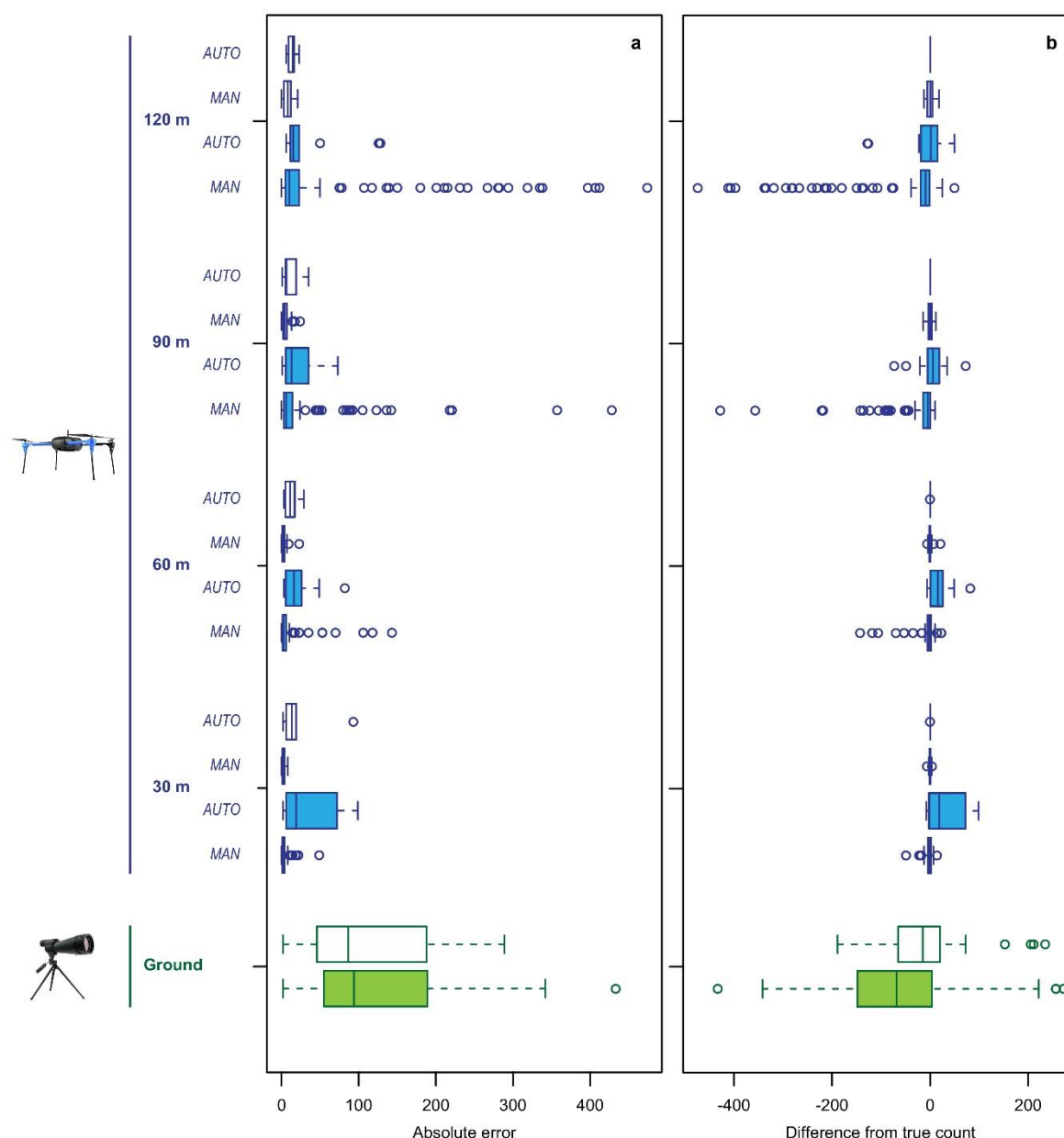
Figure 2: Accuracy and bias of RPAS and traditional wildlife monitoring approaches.
The absolute error (**a**) and difference from the true count (**b**) of each method. Data from all colonies (*n* = 10; shaded) and also for the subset of colonies with high quality imagery (*n* = 6; unshaded) are presented for both RPAS-derived (blue) and ground (green) counts. RPAS-derived manual-human (Man) and semi-automated (Auto) counts are displayed and data are grouped by height which reflects ground sample distance (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm).
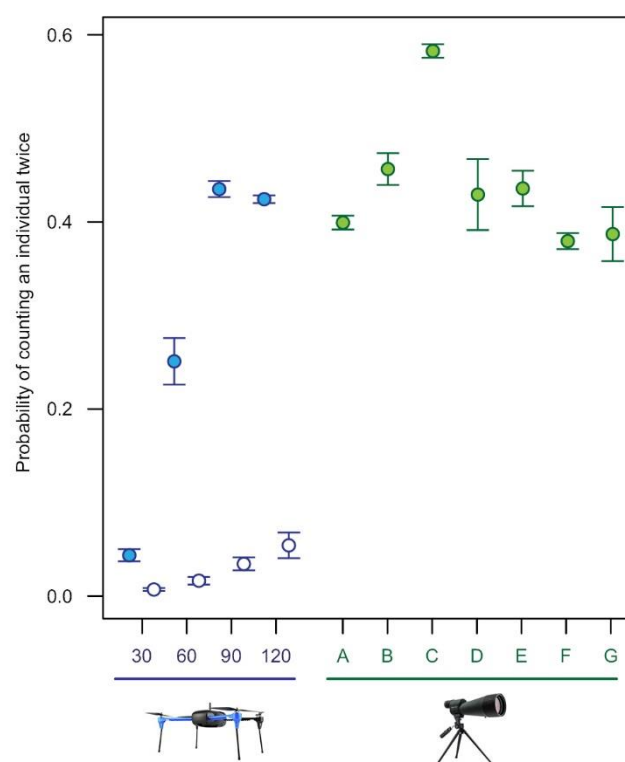
402
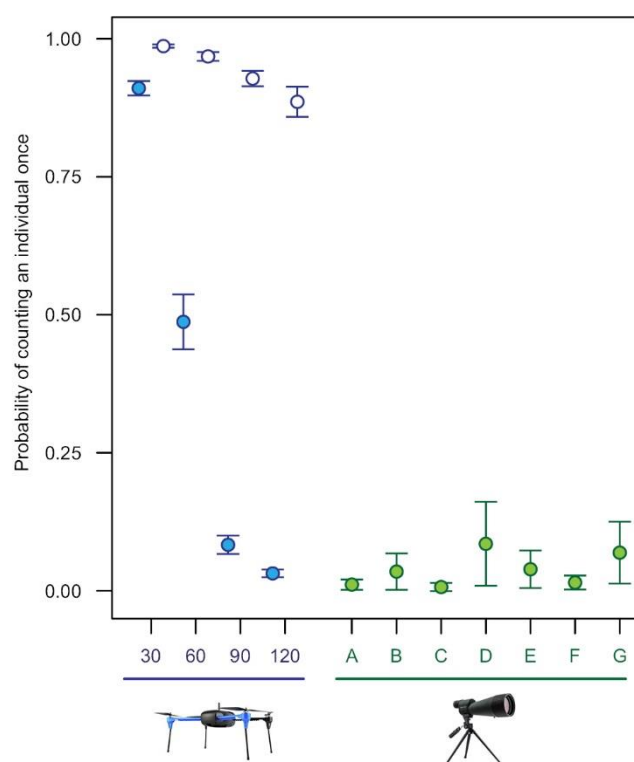
Figure 3: Probability of counting an individual in a colony once (correctly). Data from all colonies ($n = 10$; shaded) and also for the subset of colonies with high quality imagery ($n = 6$; unshaded) are presented for RPAS-derived (blue) manual counts. These data are grouped by height (m) which reflects ground sample distance (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm). Probabilities from ground count data (green) for all colonies are estimated for each counter individually (A-G). Error bars represent standard deviation.

20

410
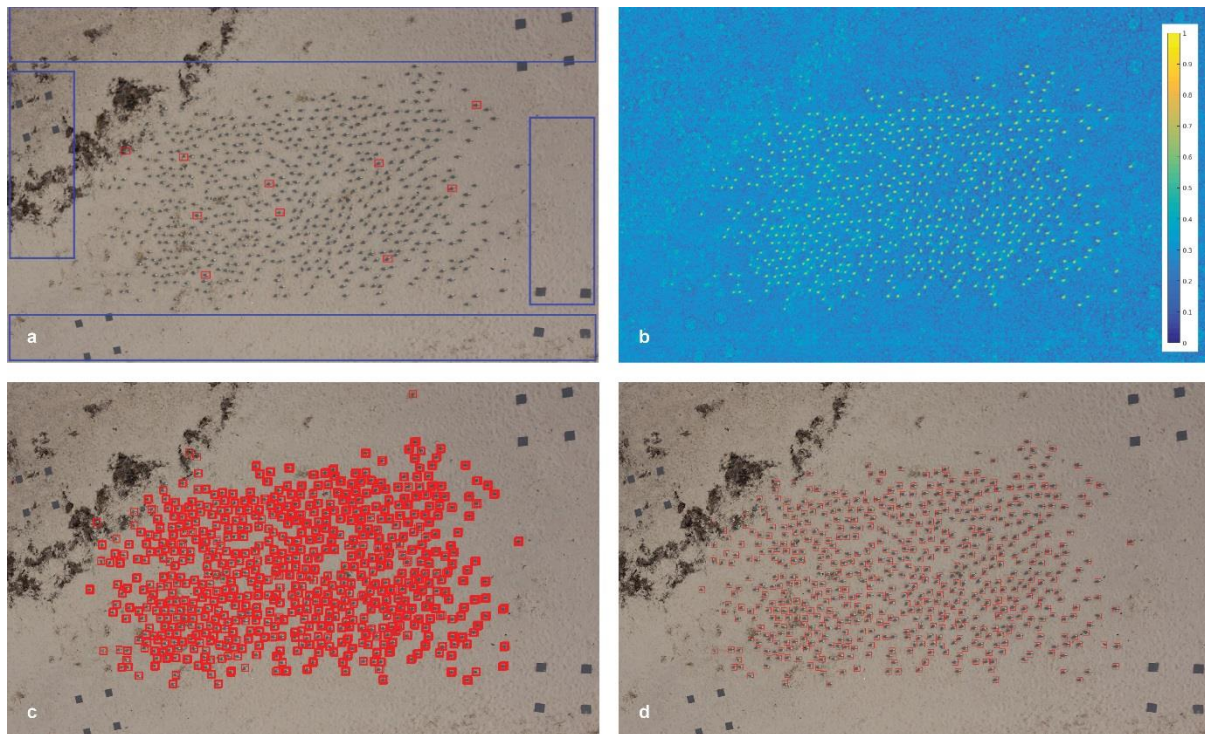
Supplementary Figure 1: Semi-automated detection and counting of wildlife using computer vision techniques. (**a**) User annotation of perceived target objects (red) and background (blue). (**b**) score map generated by the trained classifier which has automatically determined which image features distinguish objects from background, independent of scale and orientation. Warmer colours indicate increasing likelihood of the pixel being a target object. (**c**) target object proposals (red) computed by thresholding the score map. Object size is estimated from the annotations. (**d**) final output (which includes a total count and detection co-ordinates) where detected individuals are delineated (red) after redundant detections have been automatically suppressed.
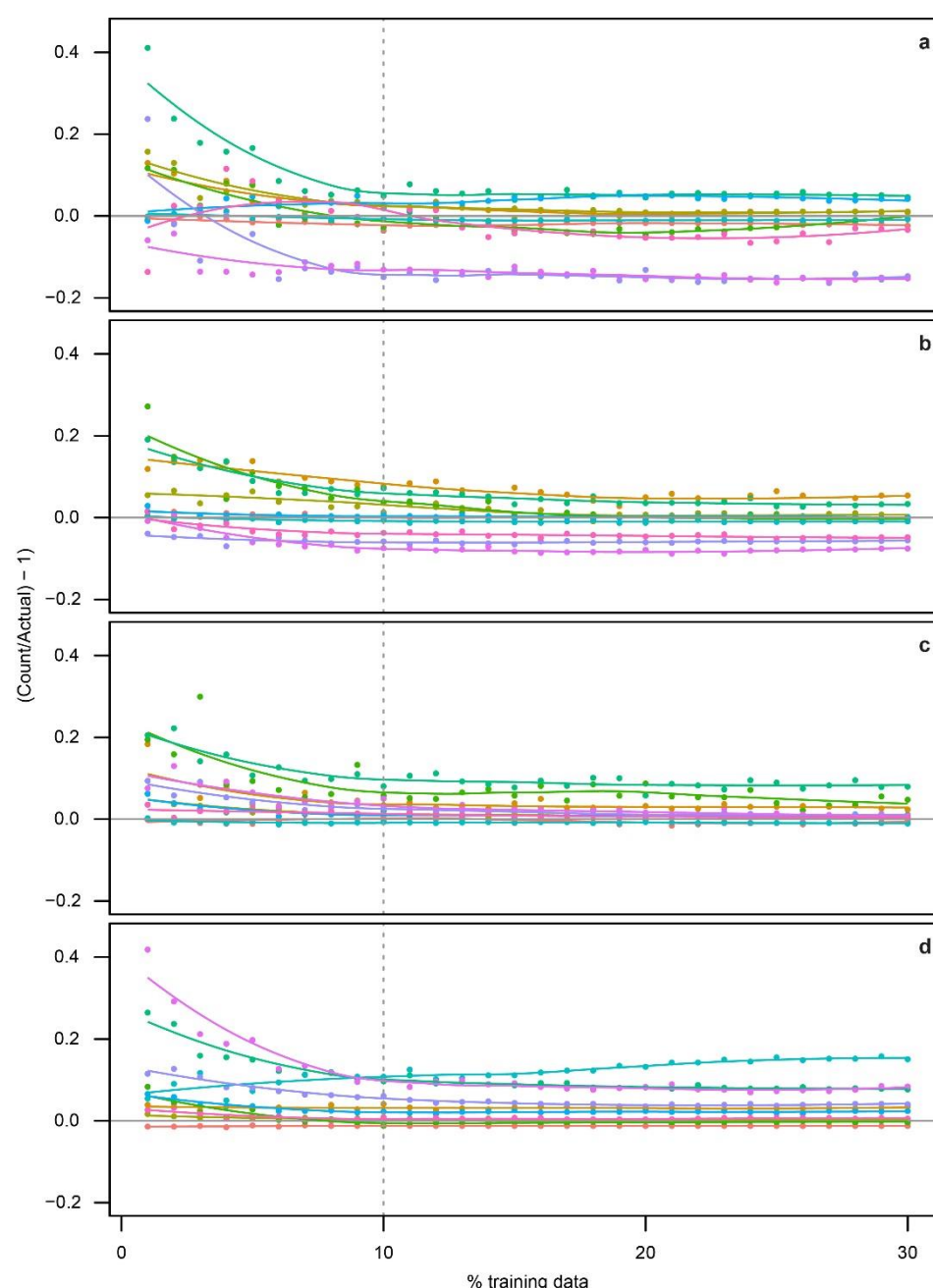
Supplementary Figure 2: Improvement in accuracy of semi-automated detection counts with increasing training data. Colonies ($n = 10$) are represented by individual colours at each height which reflects ground sample distance (GSD): 120 m height = 3.29 cm GSD (**a**); 90 m = 2.47 cm (**b**); 60 m = 1.64 cm (**c**); 30 m = 0.82 cm (**d**). Lowess smoothed trendlines are displayed. Analyses were computed using count estimates generated from 10 % training data (dashed line).

Supplementary Figure 3: Probability of counting an individual in a colony twice (double counting). Data from all colonies (*n* = 10; shaded) and also for the subset of colonies with high quality imagery (*n* = 6; unshaded) are presented for RPAS-derived (blue) manual counts. These data are grouped by height (m) which reflects ground sample distance (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm). Probabilities from ground count data (green) for all colonies are estimated for each counter individually (A-G). Error bars represent standard deviation.

435
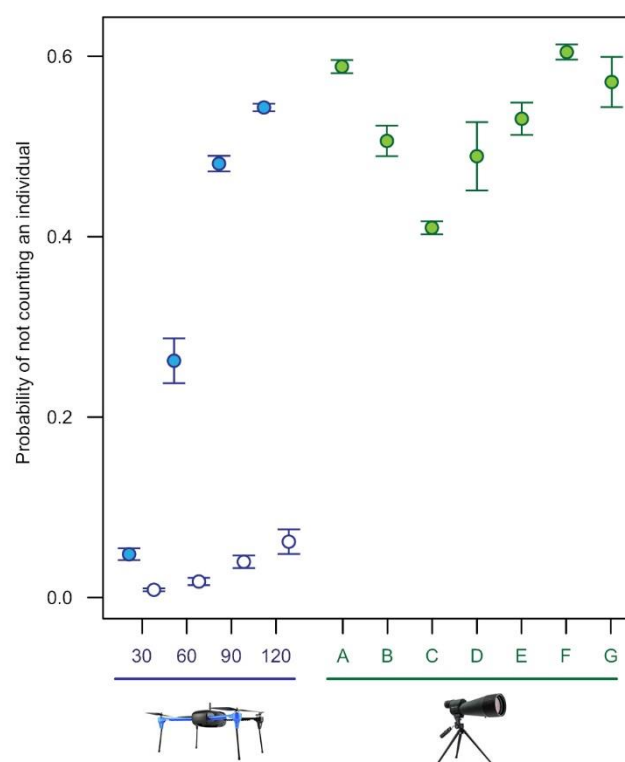
Supplementary Figure 4: Probability of not counting (missing) an individual in a colony. Data from all colonies ($n = 10$; shaded) and also for the subset of colonies with high quality imagery ($n = 6$; unshaded) are presented for RPAS-derived (blue) manual counts. These data are grouped by height (m) which reflects ground sample distance (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm). Probabilities from ground count data (green) for all colonies are estimated for each counter individually (A-G). Error bars represent standard deviation.

443
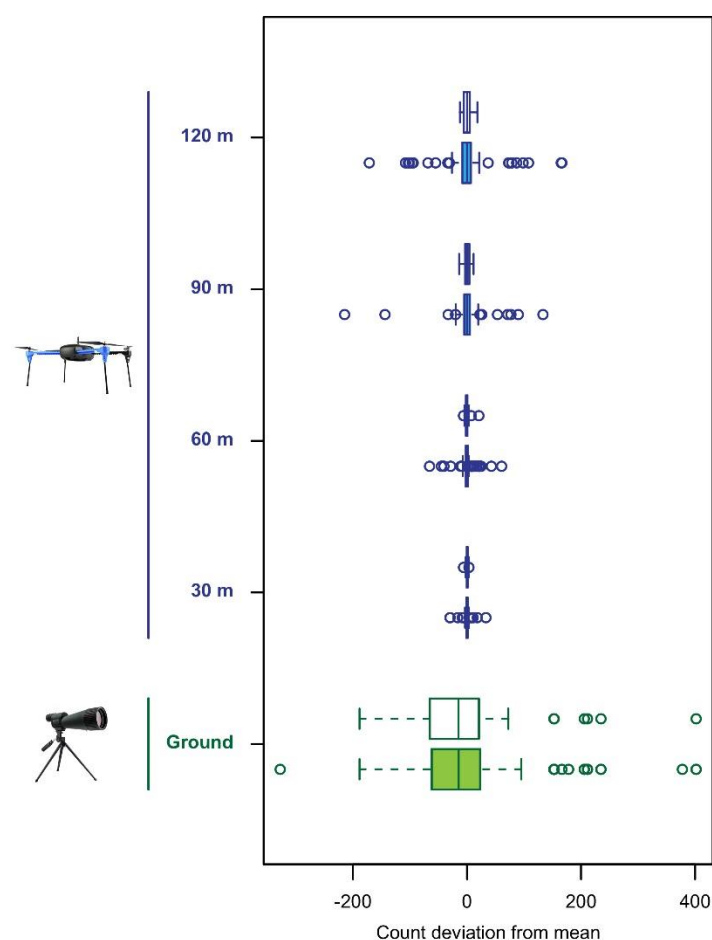
Supplementary Figure 5: Precision of RPAS and traditional wildlife monitoring approaches. Data from all colonies (*n* = 10; shaded, lower box in each course) and also for the subset of colonies with high quality imagery (*n* = 6; unshaded, upper box in each course) are presented for both RPAS-derived manual-human (blue) and ground (green) counts. Data are grouped by height which reflects ground sample distance (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm).

450 **Supplementary Table 1: Mean percentage increase in accuracy of RPAS wildlife**

451 **monitoring approaches compared with the traditional ground count approach.**

452 Percentages are calculated for RPAS-derived human manual (Man) and semi-automated

453 (Auto) counts using data from all colonies ($n = 10$) as well as the subset of colonies with high

454 quality imagery ($n = 6$). Data are grouped by height which reflects ground sample distance

455 (GSD; 30 m height = 0.82 cm GSD, 60 m = 1.64 cm, 90 m = 2.47 cm, 120 m = 3.29 cm).

Height (m)	All colonies (%)		Colonies with high quality imagery (%)	
	Man	Auto	Man	Auto
30	96	77	98	88
60	90	84	97	88
90	74	80	94	89
120	43	77	92	85

456

457