



K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N

Less is more: on-board lossy compression of accelerometer data increases biologging capacity

Nuijten, Rascha; Gerritse, Theo; Shamoun-Baranes, Judy; Nolet, B.A.

published in

Journal of Animal Ecology
2020

DOI (link to publisher)

[10.1111/1365-2656.13164](https://doi.org/10.1111/1365-2656.13164)

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Nuijten, R., Gerritse, T., Shamoun-Baranes, J., & Nolet, B. A. (2020). Less is more: on-board lossy compression of accelerometer data increases biologging capacity. *Journal of Animal Ecology*, 89(1), 237-247. <https://doi.org/10.1111/1365-2656.13164>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

1 **Less is more: on-board lossy compression of**
2 **accelerometer data increases biologging**
3 **capacity**

Rascha J.M. Nuijten^{1*}

Theo Gerrits²

Judy Shamoun-Baranes³

Bart A. Nolet^{1,3}

4 *Corresponding author: r.nuijten@nioo.knaw.nl.

5 ¹ Department of Animal Ecology, Netherlands Institute of Ecology, Droevendaalsesteeg 10, 6708 PB
6 Wageningen, the Netherlands

7 ² madebytheo, Nijmegen, the Netherlands

8 ³ Department of Theoretical and Computational Ecology, Institute for Biodiversity and Ecosystem
9 Dynamics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands.

10

11 Running title: Lossy data compression in remote animal tracking

12

13 **Keywords**

14 Summary statistics – Bewick’s swan – GPS tracking – remote behavioural observations –
15 biotelemetry

16 **Author contributions**

17 RN and BN conceived the ideas and designed the methodology, technically supported by TG.

18 TG wrote the software installed in the tracking devices and provided the calculations on data

19 reduction. RN analysed the data, advised by JS and BN. RN led the writing of the manuscript.

20 All authors contributed critically to the drafts and gave final approval for publication.

21

22 Abstract

- 23 1. GPS-tracking devices have been used in combination with a wide range of additional
24 sensors to study animal behaviour, physiology and interaction with their environment. Tri-
25 axial accelerometers allow researchers to remotely infer the behaviour of individuals, at
26 all places and times. Collection of accelerometer data is relatively cheap in terms of energy
27 usage, but the amount of raw data collected generally requires much storage space and
28 is particularly demanding in terms of energy needed for data transmission.
- 29 2. Here we propose compressing the raw ACC data into summary statistics within the
30 tracking device (before transmission) to reduce data size, as a means to overcome
31 limitations in storage and energy capacity.
- 32 3. We explored this type of lossy data compression in the accelerometer data of tagged
33 Bewick's swans (*Cygnus columbianus bewickii*) collected in spring 2017. By using
34 software settings in which bouts of 2 s of both raw ACC data and summary statistics were
35 collected in parallel but with different bout intervals to keep total data size comparable, we
36 created the opportunity for a direct comparison of time budgets derived by the two data
37 collection methods.
- 38 4. We found that the data compression in our case yielded a 6 time reduction in data size
39 per bout, and concurrent, similar decreases in storage and energy use of the device. We
40 show that with the same accuracy of the behavioural classification, the freed memory and
41 energy of the device can be used to increase the monitoring effort, resulting in a more
42 detailed representation of the individuals' time budget. Rare and/or short behaviours such
43 as daily roost flights, were picked up significantly more when collecting summary statistics
44 instead of raw ACC data (but note differences in sampling rate). Such level of detail can
45 be of essential importance, for instance to make a reliable estimate of the energy budgets
46 of individuals.

47 5. In conclusion, we argue that this type of lossy data compression can be a well-considered
48 choice in study situations where limitations in energy and storage space of the device
49 pose a problem. Ultimately these developments can allow for long-term and nearly
50 continuous remote-monitoring of the behaviour of free-ranging animals.

51 **Introduction**

52 The use of biologging has increased enormously in ecology and allows for remote observation of
53 wild animals (Cooke et al., 2004; Wilmers et al., 2015). GPS-tracking devices have been used in
54 combination with heart rate measurements (Duriez et al., 2014; Wascher et al., 2018),
55 temperature sensors (Ryan et al., 2004; Sala et al., 2017), magnetometers (Laplanche et al.,
56 2015; Noda et al., 2014), accelerometers (Brown et al., 2013; Nathan et al., 2012) and even
57 cameras (Patel et al., 2017; Watanabe et al., 2003) to learn about animal behaviour and the
58 interaction of individuals with their environment. The remote tracking of individual animals has
59 solved many questions that were previously beyond reach (e.g. Mansfield et al., 2014; Williams
60 et al., 2014) and the observations, objective and undisturbed by the observer, are valuable for
61 both fundamental (Watanabe et al., 2014) and applied ecological research (Wilson et al. 2015).
62 Technological developments have made the devices increasingly smaller (Kays et al., 2015), so
63 that nowadays almost any mammal, bird or reptile species, and even amphibians and
64 invertebrates can be remotely observed to answer research questions about their biology
65 (Cagnacci et al., 2010; Kissling et al., 2014). Although this development has also reduced the
66 effects of a tracking device on the survival and behaviour of the animal, this can never be
67 completely excluded and should be monitored closely (Lameris et al. 2018). Practical limitations
68 regarding battery weight (and thus device weight) were reduced by the development and usage
69 of solar-energy to re-charge the battery while attached to the animal (Bouten et al., 2013;
70 Tomkiewicz et al., 2010). This reliable and predictable power source elongated deployment time
71 of devices in many environments apart from for example the marine domain (Adoram-Kershner
72 et al., 2017), under dense canopy cover (Kays et al., 2011), or in winter at high latitudes (Therrien
73 et al., 2012). Moreover, the use of remote download techniques such as Bluetooth, radio- and
74 GSM networks made re-catching of the individual redundant, allowing for increased data yield per
75 device deployment (Bouten et al., 2013; Tomkiewicz et al., 2010) and allowing more species to

76 be tracked (e.g. those that die during deployment or that do not return to accessible places for
77 tag retrieval). With these practical limitations being addressed, the road is paved for longer
78 deployment time and high(er) frequency measurements to answer more detailed research
79 questions about individual animal behaviour (Allan et al., 2018; Wilmers et al., 2015). For example
80 due to long deployment it was shown that migratory performance of Black Kite (*Milvus migrans*)
81 increases with age through a combination of individual improvement and selective mortality of
82 poor performers (Sergio et al., 2014). And thanks to frequent measurements the extraordinary
83 locomotor dynamics of hunting cheetahs (*Acinonyx jubatus*) were described (Wilson et al. 2013).

84

85 The high frequency required for answering detailed research questions comes at the cost of
86 storage space and energy use for data collection and transmission. The use of multiple sensors
87 or intensive use of a single sensor then becomes a trade-off: if additional sensor data is collected,
88 fewer fixes can be stored on the memory of the tracking device (Bouten et al., 2013; Wilson et al.,
89 2015). For devices that need to be retrieved to get the data, often storage space can be limiting
90 so that the research is restricted in either deployment time or frequency of sampling, for example
91 many seabird studies use tags in which individuals are followed for only several days (Dean et
92 al., 2012; Shaffer et al., 2017). Remote data download, on the other hand, mainly puts pressure
93 on the energy balance of the device since making connection with the download system and data
94 transmission requires a considerable amount of energy. In this case the speed of the network and
95 energy available for uploading become limiting with high frequency data collection.

96

97 With respect to the limitation in data transmission, there are broadly speaking two kinds of
98 solutions: 1) increasing the capacity of the network or 2) decreasing the amount of data that needs
99 to be transferred by clever data compression. The first solution is aimed at the bandwidth of a
100 certain system, i.e. the amount of data that can be transmitted per time interval through the
101 network. Improvements of this kind have indeed been implemented, for example in the Global

102 System for Mobile communications (GSM). Technological developments have advanced the
103 communication via this network from analog radio signals (1G) to digital radio signals (2G) and
104 then step-wise increased the bandwidth enormously (3G, 4G and consecutively 5G networks) so
105 that it can now support global telecommunication (Tondare et al., 2014). Although great profit can
106 be achieved from this type of advancement, changing the bandwidth of the network used for
107 biologging devices is often beyond the researcher's control. In case of the GSM network for
108 example, it depends on the availability of the network at the location of the animal. The second
109 solution on the other hand is within control, and already widespread in many aspects of digital
110 modern life. Files that are too big to send as an attachment are often compressed and then
111 extracted (for example in the ZIP file format), and for images, size can be reduced by storing it as
112 a Portable Network Graphic (PNG). These are both common examples of 'lossless' data
113 compression techniques, referring to the fact that no information is lost by the data compression.

114

115 An alternative is the so-called 'lossy' data compression, when some data is lost. A popular lossy
116 compression method for images is JPEG, where the image visualization is stored as block-wise
117 quantized discrete cosine transform coefficients (Fridrich et al., 2015). This reduces the quality of
118 the image, but the content is mostly still clear enough for its purpose. Losing information may
119 sound unwanted, but often there is quite some redundant information in a large data-file that can
120 be lost without compromising the output. For example in a video, background features shot by
121 the same camera often do not change for the duration of a scene. These less complex 'chunks'
122 of the video, in terms of motion and detail, can be encoded separately and with a lower bitrate,
123 thus reducing the data size of that part. This 'chunk-based' encoding allows for high quality video
124 streaming even in low-bandwidth internet connections (De Cock et al., 2016; Norkin et al., 2016).

125

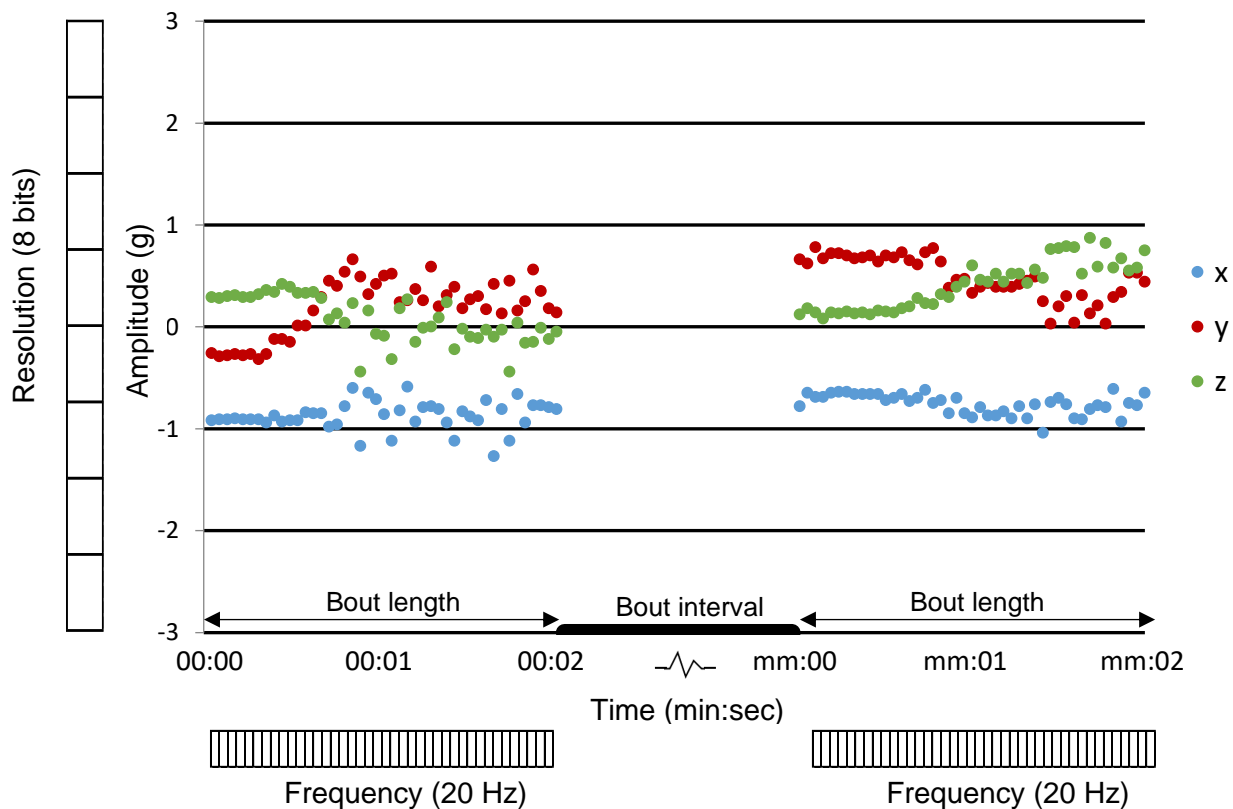
126 Similar solutions of lossy data compression could be advantageous in biologging. In the
127 challenging marine environment of the Antarctic, in terms of tag retrieval and data transmission,

128 such solutions have already been applied to study behaviour in seals. To be able to collect data
129 on prey catch dives in these animals, an abstract from peaks in acceleration indicative of rapid
130 head movements was calculated on-board the data logger (Cox et al., 2018; Heerah et al., 2019).
131 Also in a less challenging environment, compressing acceleration data can be advantageous to
132 overcome storage- and bandwidth- limitations. Liechti et al. (2018) recently showed that only
133 storing a summary of acceleration data in the z-axis enabled the collection of data on the full
134 migratory journey of small trans-Saharan migrants, something that was not possible before. One
135 example of lossy data collection is a conditional sampling regime, where the frequency of
136 sampling is not continuously the same. The exact frequency of sampling can then for example be
137 determined by the researcher (Bouten et al. 2013), based on the energy level of the device (Dokter
138 et al., 2018; see Appendix A), or the inferred behaviour of the focal animal (e.g., flight detection,
139 based on GPS-ground speed (Harel et al., 2016) or the overall activity level (Brown et al., 2012)).
140 Although this can reduce the data size over the study period, it is still a compromise as continuous
141 and high frequency long-term sampling is not achieved, and one has to choose beforehand, which
142 time-periods or behaviours will be monitored with high frequency and which are of less interest
143 (and thus 'lost').

144

145 A biologging sensor that may be particularly suitable for data compression is the accelerometer
146 (ACC) as has recently been suggested in the technological literature (le Roux et al., 2018). Tri-
147 axial accelerometer sensors are becoming an increasingly common addition to GPS-tracking
148 devices. Tri-axial accelerometers measure the rate of change in directional speed along three
149 orthogonal axes, traditionally called x or 'surge', y or 'sway' and z or 'heave' (Yoda et al., 2001).
150 The first reported use of accelerometer data in ecology was in Adélie Penguins (*Pygoscelis*
151 *adeliae*), where ACC data enabled the researchers to distinguish seven types of behaviour (Yoda
152 et al., 1999). ACC data has been measured in two ways, either continuously for short deployments
153 of several days (Chimienti et al., 2016; Wilson et al., 2008) or for longer deployments up to several

154 years, in short bouts (Flack et al., 2018; Yoda et al., 1999). Collection of ACC data is relatively
 155 cheap in terms of energy usage, however, the storage of the data requires a lot of space and the
 156 data is particularly demanding in terms of energy needed for data-transmission (Wilson et al.,
 157 2008). For example, if the ACC sensor collects tri-axial data at a resolution of 1 byte with a
 158 duration (referred to as bout-length in the remainder of this study; for a full explanation of terms
 159 see Fig. 1) of 2 s, and a signal frequency of 20 Hz, it means that 120 bytes are stored in the
 160 device (2 s x 20 Hz x 3 axes) per bout. For species that live in remote areas and are therefore not
 161 easy to reach or observe, and that one would like to follow long-term (preferably year-round, if
 162 not multiple years; (Wikelski et al., 2007)), this amount of data can altogether easily become
 163 problematic and compromise either deployment time or the number of measurements taken.
 164



165
 166 *Fig. 1. Generic example of accelerometer output with time (min:sec) on the horizontal axis and rate of*
 167 *change (Amplitude) in speed, in the three acceleration axes (x, y and z), on the vertical axis. The three*

168 *coloured data series represent the data from the three orthogonal axes of the accelerometer: x (blue), y*
169 *(red) and z (green). Bout length and bout interval represent the duration of the measurement and the interval*
170 *between the ACC measurements, respectively, and are graphically represented in the figure just above the*
171 *horizontal axis. In our study, bout length was 2 seconds and bout interval 1.58 min for the summary statistic*
172 *measurements (thus one measurement of two seconds was taken every two minutes), and 2 seconds and*
173 *14.58 min, respectively, for the raw ACC measurements (one measurement of two seconds every 15*
174 *minutes). On the x-axis in the figure 'mm' equals '02' for the SS method and '15' for the raw ACC method.*
175 *Frequency (graphically represented just below horizontal axis) and Resolution (graphically represented left*
176 *of the vertical axis) depict the settings for the accelerometer sensor as they were used in this study.*
177 *Frequency refers to the sampling frequency in Hertz (Hz), in this case 20 Hz. Resolution refers to the level*
178 *of detail of each x, y or z measurement within a bout in bits (b). Resolution in this study is 8 b, which*
179 *corresponds to one byte and results in $2^8 = 256$ different potential 'levels' for every measurement of x, y*
180 *and z between -3 and 3 g. One bout of 2 s and 20 Hz thus comprises of 120 bytes (see main text). For*
181 *graphical purposes we did not visualize all 256 levels here, therefore each unit on the Resolution scale in*
182 *the figure represents 32 levels. For the summary statistics, the recorded x, y and z values within a bout are*
183 *summarized to statistics in the device before being stored and transmitted.*

184

185 In ecology, there is reluctance towards the idea of lossy data compression, because of the loss
186 of raw data and potentially important information in the process. Here we propose and test a
187 method for lossy data compression by reducing the raw ACC data to summary statistics per ACC-
188 bout and discuss its advantages and disadvantages. This type of data compression reduces the
189 amount of data that needs to be stored, and thus the amount of bytes that need to be transmitted
190 by the device. By using this type of data compression, the monitoring coverage of data collection
191 (either by reducing the bout interval between measurements, increasing the frequency or
192 increasing the bout length) can be greatly improved, by enabling higher frequency monitoring or
193 longer tracking periods.

194

195 This study presents a methodological approach to compress the raw accelerometer data within
196 the device to summary statistics and simultaneously decrease the bout interval between sampling
197 bouts. We calculated and compared time budgets of free ranging Bewick's swans (*Cygnus*
198 *columbianus bewickii*) derived from both raw and summary statistic ACC data collected in parallel.

199

200 **Materials and Methods**

201 *Study species*

202 The Bewick's swan is a long-distance migratory bird, which in the western part of its range, winters
203 in North-Western Europe and breeds at the European Russian tundra (Rees, 2006). The
204 migration route and breeding area of this population is well-known due to extensive tracking
205 efforts with PTT transmitters and GPS loggers in the past (Beekman et al., 2002; Nuijten et al.,
206 2014). In the summer of 2016 and 2017 observations were carried out in three zoos of captive
207 Bewick's swans equipped with GPS/GSM tracking devices to ground-truth the accelerometer data
208 and build a behavioural classification model (Nuijten et al., unpublished). In the winter of
209 2016/2017 30 free-ranging Bewick's swans were equipped with these tracking devices in the
210 province of Noord-Brabant (The Netherlands). Tracking data from spring 2017 (1 Feb – 31 May)
211 of 10 individuals in which both raw ACC and ACC summary statistics were collected at high rate
212 was used to apply the behavioural classification model and create individual time budgets (this
213 study).

214 *Device and settings*

215 We used custom designed, 3D-printed GPS/GSM neck-collars with a weight of 70 g, an inner
216 diameter of 51 mm and a height of 80 mm. The weight of the collar (including the tag with sensors)
217 represented 1.1% and 1.2% of the average weight of adult and non-adult Bewick's swans
218 respectively, based on a dataset of 295 Bewick's swans caught in the Netherlands between 2005

219 and 2017. During previous observations of captive Bewick's swans with such collars, the swans
220 with the collars preened more at first but no effect on the behaviour of the swans was found after
221 four weeks (Nuijten et al., 2014). The collar contained, apart from the GPS sensor, a tri-axial
222 accelerometer and a water sensor, and sent its data remotely via the GSM network. The
223 accelerometer collected data with a bout length of 2 s and a frequency of 20 Hz (Fig. 1).

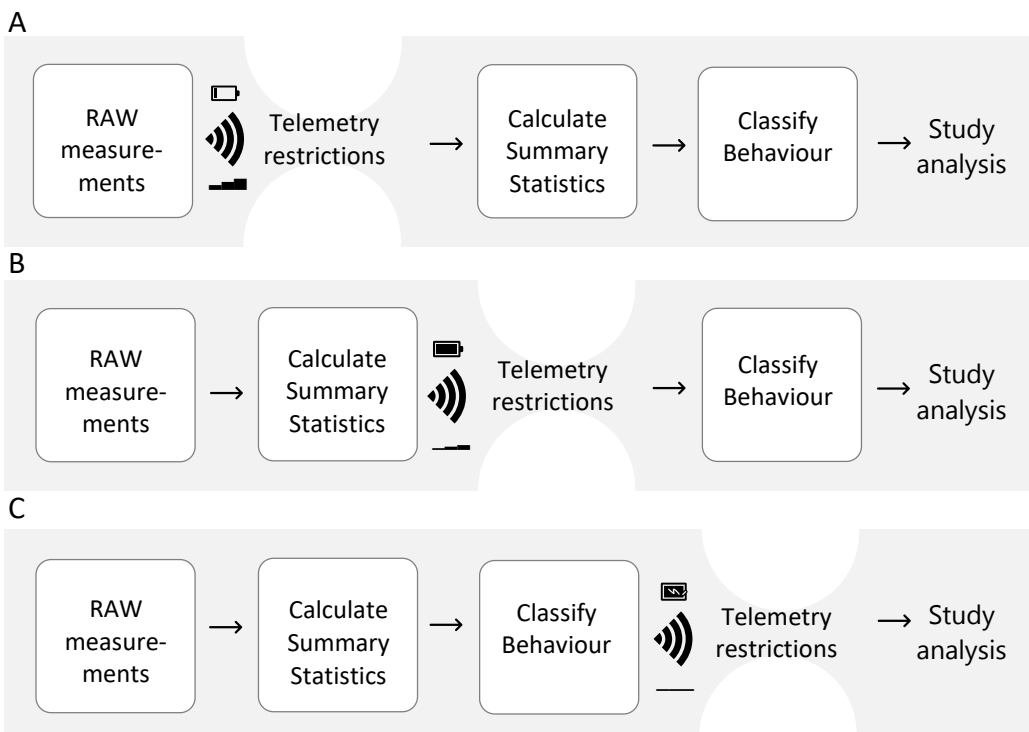
224 The accelerometer and water sensor sampling were programmed separately rather than
225 simultaneously with the GPS-fixes, to be able to maintain a fixed sampling scheme for the water
226 sensor, while the GPS and accelerometer settings were made dependent on battery voltage of
227 the device (see Appendix A for an overview of all settings). In the spring season, the period of
228 which we used the data in this study, GPS fixes were collected every 15 minutes (Appendix A).
229 Raw ACC data was stored also every 15 minutes, and ACC data summarized to summary
230 statistics every two minutes (see *Raw ACC and summary statistics data collection*). The water
231 sensor recorded water (1) or no water (0) every second. The collar connected to the GSM network
232 once a day to transmit the data. The settings could not be changed after deployment.

233 *Raw ACC and summary statistics data collection*

234 The accelerometer, as mentioned before, is a very demanding sensor in terms of energy needed
235 for transmission of the data. Combining the frequency, the axes and bout length for the raw ACC
236 data in this study, every bout adds up to 120 bytes per bout ($20 \text{ Hz} * 3 \text{ axes} * 2 \text{ seconds}$). In our
237 study (i.e. 1 bout every 15 minutes) this equals 480 bytes per hour and 11520 bytes per day for
238 the raw acceleration data only. That is excluding metadata such as date, time, individual ID,
239 column labels etcetera.

240 Compressing the amount of data already within the tracking device by reducing the raw ACC
241 numbers to summary statistics (SS) such as average x or Overall Dynamic Body Acceleration
242 (ODBA; see Appendix B for an overview of all SS used in this study) over the bout length reduces

243 the amount of data that needs to be stored and transmitted per bout. Here we used 20 summary
 244 statistics to compress the raw ACC data per ACC bout, which equals 20 bytes per bout (excluding
 245 metadata), a reduction of $(120 / 20)$ 6 times when compared to a raw ACC bout. To keep the total
 246 amount of data approximately the same between the two data collection methods (for the purpose
 247 of comparison), we increased the number of ACC bouts per time unit for the SS method
 248 accordingly. We therefore programmed the ACC sensor of each collar to collect SS ACC every
 249 two minutes (excluding the time points when raw ACC was collected), to be able to compare two
 250 datasets collected with the same storage and energy capacity of the device.



251

252 *Fig. 2 Graphical representation of data collection schemes in this study. (A) Raw ACC scenario. Raw*
 253 *acceleration data are collected by the biologging device and sent via the network to a server. The raw data*
 254 *represent a large amount of data and transmission is costly in terms of energy usage. Only after*
 255 *transmission the ACC data will be summarized and classified to behaviour. (B) Summary statistics scenario.*

256 *Raw acceleration data are collected by the biologging device and summary statistics per bout from these*
257 *ACC data, such as average x or Overall Dynamic Body Acceleration (ODBA) are calculated on board. The*
258 *summary statistics comprise less data and thus take less energy to be transmitted to the server. From the*
259 *summary statistics, behaviour can be classified for further analysis. (C) Behaviour indicator scenario. Raw*
260 *acceleration data are collected by the biologging device, summary statistics are calculated and the*
261 *behaviour is classified on board. This results in only a single indicator being sent via the network to the*
262 *server, using only very little energy. Scenarios (A) and (B) are compared in this study, scenario (C)*
263 *represents a next step in the developments of accelerometer research.*

264

265 *Behavioural classification and statistical analysis*

266 We used an ensemble learning decision tree method (random forest, Liaw & Wiener, 2002) to
267 build a classification tree from the annotated acceleration data obtained in the zoo, complemented
268 with flight data from free-ranging Bewick's swans as flapping flight is very easy to distinguish from
269 other behaviours (Nuijten et al., unpublished; Bishop et al., 2015; Shamoun-Baranes et al., 2016).
270 When working with raw ACC data it is common practice to reduce this data to classifiers (i.e.
271 summary statistics) before applying the classification model (Bom et al., 2014; Shamoun-Baranes
272 et al., 2012). We used 21 statistics (20 ACC summary statistics + the information from the water
273 sensor) to classify the behaviours in this study (Appendix B). The same 20 summary statistics
274 were calculated in the SS and raw ACC bouts, the sole difference between the datasets being
275 the moment of calculation (i.e. before and after transmission respectively; cf. Fig. 2A and 2B).
276 Five-minute aggregates of the water sensor data (i.e. 300 seconds) were aligned to the ACC data
277 based on the satellite timestamps of both measurements. If for ≥ 30 s within this 5-min aggregate
278 water was recorded, the overlapping bouts were assigned a '1', otherwise a '0'.

279 The behavioural classification for both the raw ACC data and the SS data from the free-ranging
280 individuals was performed with the same classification tree which had an overall classification

281 accuracy of 91% (recall: 0.89; precision: 0.92) and included the behaviours sleeping, resting,
282 terrestrial active (combination of terrestrial foraging and preening behaviour), swimming, aquatic
283 foraging and flying (Nuijten et al., unpublished). The classified data was used to visualize daily
284 time budgets for free-ranging individual swans in spring, once for the raw ACC and once for the
285 SS dataset over the same time period. Additionally, proportions of each behaviour per day were
286 calculated for both the raw ACC and the SS data. Sample sizes for the daily proportions were
287 maximally 96 per day for the raw ACC data (one ACC bout every 15 min) and 672 per day for the
288 SS data (one ACC bout every two minutes, excluding the time points when raw ACC data was
289 collected).

290 To assess whether the different datasets yielded different time budgets, we calculated the mean
291 difference between raw ACC and SS-based daily proportions per behaviour, and calculated the
292 probability of this observed mean difference originating by chance using a non-parametric
293 permutation test. We did this by randomizing of the sign of the difference between raw ACC-
294 based proportions and SS-based proportions per day, and taking the mean of these differences.
295 By repeating this 10,000 times we created a distribution of randomized mean differences between
296 raw ACC and SS proportions against which the observed mean difference was tested.

297

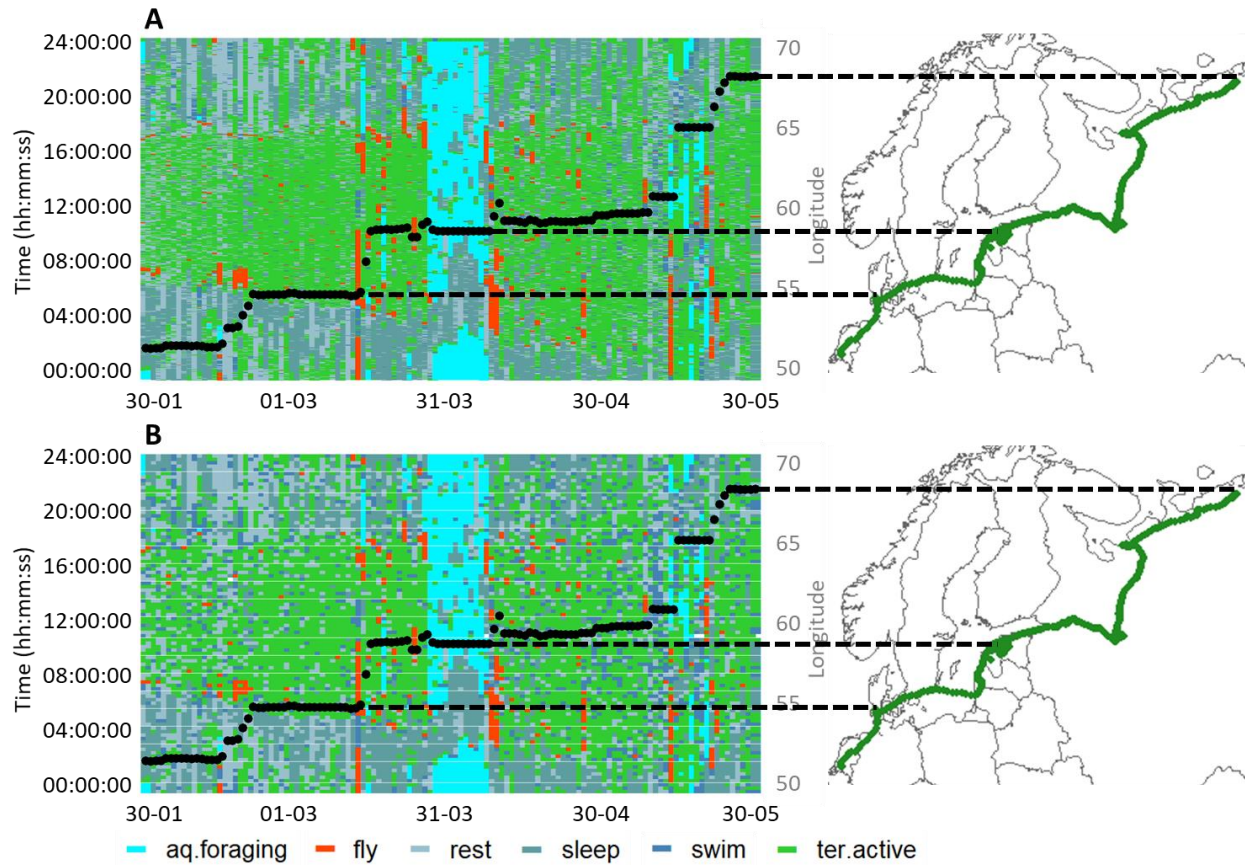
298 **Results**

299 On board calculation of summary statistics greatly reduced the amount of data per bout to be
300 transmitted by the biologging devices. Concerning the accelerometer data only, we reduced our
301 data size per 2 s bout six times from 120 bytes ($2 \text{ s} * 20 \text{ Hz} * 3 \text{ axes}$) to 20 bytes, by storing 20
302 summary statistics on board the biologgers rather than the raw tri-axial accelerometer data.
303 Including metadata such as individual ID and timestamp, we realized a 4.7x reduction in the
304 amount of data per bout (127 versus 27 bytes, respectively). This resulted in a similar decrease
305 in energy needed for transmission of the data. Transmission of the raw ACC data over the network

306 took approximately 5 min and 2639 μ Wh for all data of one day. Transmitting the SS data took
307 roughly 1 minute and required 528 μ Wh from the collar. The extra energy needed for the
308 calculation of the SS within the device was only 0.239 μ Wh by which 1344 SS bouts of 27 bytes
309 were created (i.e. one day worth of SS data). So by 'paying' 0.239 μ Wh as a cost for calculation,
310 and with similar circumstances in terms of bandwidth and connection with the network for both
311 data collection methods, a 5x reduction ($= 2639 / (528+0.239)$) in energy use for transmission
312 was realized.

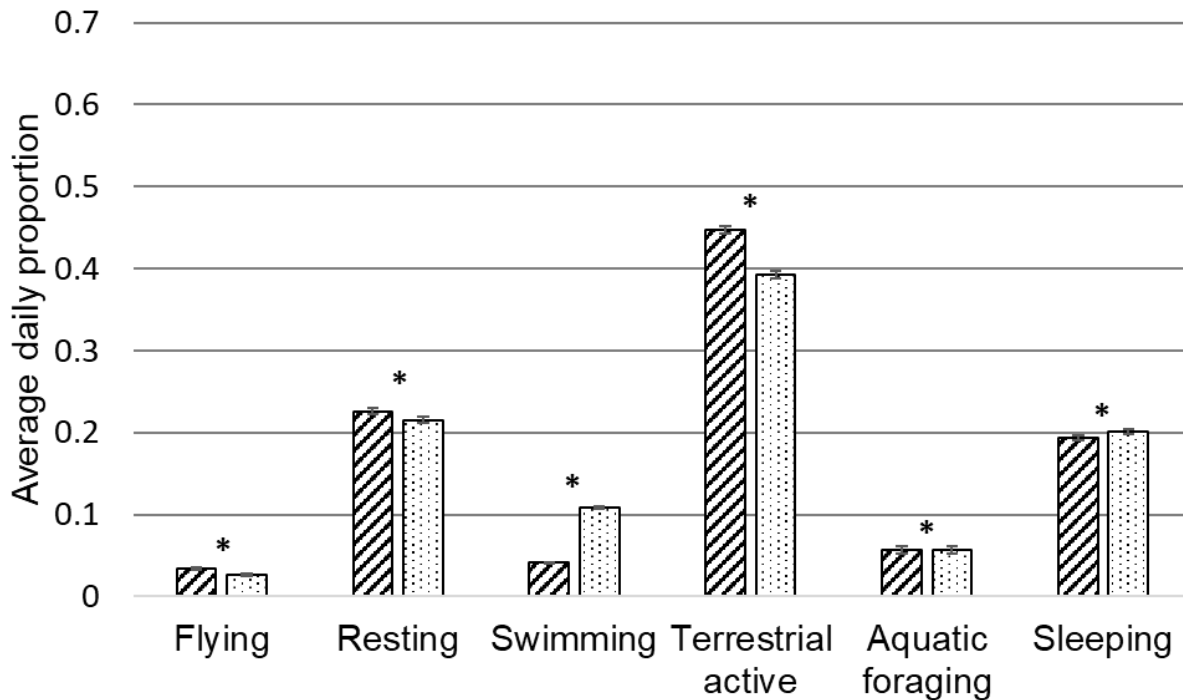
313 Both the raw ACC and SS data were used to create time budgets for each individual (see Fig. 3
314 for an example). Within individuals, the difference in bout interval between the two methods is
315 clearly visible in the time budget graphs (Fig. 3). This difference in bout interval resulted in some
316 biologically relevant behaviours to be better represented by the SS compared to the raw ACC-
317 based data. For instance, roost flights, a twice-daily behaviour of relatively short duration that
318 Bewick's swans perform to travel between sleeping and foraging areas in the morning and
319 evening, was detected on significantly more occurrences (paired t-test: $N = 14$ days; $t = 4.8963$;
320 $p = 0.001$) in the SS data than the raw ACC data at the end of the winter season (i.e. the first 14
321 days of our study period; SS: 20.11 ± 2.1 and raw ACC: 10.6 ± 2.2 days (mean \pm se)) when the
322 swans are known to perform this behaviour.

323 We found a significant difference between the raw ACC and SS-based average daily proportions
324 for all behaviours over the study period ($p \ll 0.0001$). All permutation tests had a sample size of
325 1200 (120 days * 10 individuals). The proportion of flying (observed mean difference -0.007),
326 standing (-0.011), terrestrial active (-0.056) and aquatic foraging (-0.001) was higher when based
327 on SS data when compared to raw ACC data, while the proportion for swimming (observed mean
328 difference 0.068) and sleeping (0.007) was lower (Fig. 4).



329

330 *Fig. 3 Example of a daily time budget for the spring migration of individual 233E with time on the y-axis and*
 331 *date (1 Feb – 31 May) on the x-axis. (A) is based on the summary statistics data, collected every two*
 332 *minutes, (B) is based on the raw ACC data, collected every 15 min. The different colours indicate the*
 333 *following behaviours: sleeping (blue-grey), resting (grey), terrestrial active (green), aquatic foraging (aqua),*
 334 *swimming (dark blue) and flying (red). The black dots represent the latitudinal location of the swan on that*
 335 *specific day as collected by the GPS of the neck-collar. For reference the trajectory of the swan, for the*
 336 *same period, is plotted in the map adjacent to the time budget graph. Dotted lines connect the latitudinal*
 337 *location of the swan on the graph to that on the map. Time budgets for all 10 individuals are presented in*
 338 *Appendix C.*



339

340 *Fig. 4 Daily proportions per behaviour for spring 2017 (1 Feb – 31 May), mean \pm se, N = 1200. Bars with*
 341 *diagonal lines represent the data based on summary statistic data (i.e. the short bout interval), bars with*
 342 *dots represent the data based on the raw ACC data (i.e. the long bout interval). For all behaviours the*
 343 *difference between the proportions of both methods were statistically significant ($P < 0.0001$, indicated by*
 344 ** in the figure).*

345 Discussion

346 We explored the use of lossy data compression in biologging devices as a solution to overcome
 347 limitations in energy capacity of the device, specifically with regard to the accelerometer sensor.
 348 By using ACC data collected in free-ranging Bewick's swans as an example we show that lossy
 349 data compression reduces the size of the ACC data that needs to be stored and transmitted by
 350 the tracking device without loss of biological information. The exact reduction factor depends on
 351 the settings of the accelerometer (bout length, frequency, amplitude and resolution; Fig. 1) and
 352 the number of SS stored (Appendix D). The freed capacity of the device by using the SS data

353 collection method instead of raw ACC can be used to decrease bout interval (as was done in this
354 study), or increase the frequency or resolution of the ACC measurements during the setup of the
355 study which will lead to an increased level of detail in the output data (see for example Bom et al.,
356 2014; Broell et al., 2013). Alternatively, the freed capacity can be used to increase the frequency
357 or resolution of another sensor, to elongate the deployment time of the device or by including
358 other (data-rich) sensors such as a heart-rate sensor or sensors that measure features from the
359 environment. These latter scenarios were not considered in this study, but can have huge
360 advantages in studies where ACC data transmission is currently limiting.

361 Both the raw ACC and SS data were classified with the same behavioural model that was built
362 based on the zoo observations (Nuijten et al., unpublished). In such a supervised classification
363 model, raw ACC data is commonly reduced to summary statistics before classification can be
364 done (Bom et al., 2014; Shamoun-Baranes et al., 2012), so our method does not differ from
365 classical ACC analyses in that respect. In the classification we used the same summary statistics
366 as were calculated for the SS bouts within the device as these represent such a broad range of
367 statistics that all behaviours should be represented by one or a combination of several of them.
368 This was confirmed by the high performance of the model (91% correct classification overall).
369 Also, our final behavioural classification model only used four out of the 21 statistics that were
370 collected (ODBA, maximum z-value, mean z-value and the water sensor; Nuijten et al.
371 unpublished), so the selection of summary statistics before deployment of the devices could have
372 been more restrictive, resulting in a more than ten-fold reduction in data size per bout. This shows
373 that the use of SS in accelerometer data collection can even increase the biologging capacity of
374 this sensor more than we demonstrate in this study.

375 Application of the model on both SS and raw ACC datasets yielded the classified datasets that
376 were used to create the time budgets for the individual swans for spring 2017. When testing for
377 differences between raw ACC and SS-based daily proportions of behaviour, we found a small but

378 statistically significant difference for all behaviours. Although it is generally assumed that a
379 discontinuous but structured ACC sampling can be used validly as a proxy for continuous
380 measurement of behaviour (Brown et al., 2013), we found a small but statistically significant
381 difference for all behaviours when testing for differences between raw ACC and SS-based daily
382 proportions of behaviour. Although both methods collect the same type of data every bout (20 Hz
383 ACC data of 2 s duration), there are two differences that could have caused the differences that
384 we found. First, the SS method takes more samples of ACC data in the same time interval (raw
385 1: SS 7 bouts in 15 min; Fig. 1). This leads to a higher monitoring coverage in the SS method.
386 Second, due to this higher monitoring coverage the SS bouts are taken at different time points
387 than the raw ACC bouts. The differences in the proportions are not unidirectional (i.e. that SS is
388 always higher or always lower than raw ACC), and cannot be, because the behaviours are
389 proportional and thus not independent from each other (Appendix E). For example, when a swan
390 increases the time spent foraging there is less time for other activities (for example sleeping). This
391 is a property of proportional data, as all proportions together must sum to 1. We found that
392 especially both foraging behaviours were negatively correlated (so when more time was spent on
393 aquatic foraging, less time was spent on terrestrial foraging (classified as terrestrial active in this
394 study); Pearson correlation coefficient -0.58; Appendix E). Due to the higher monitoring coverage
395 of the SS bouts (i.e. more samples to represent the continuum of an animal's behaviour), we
396 believe that the proportions and time budgets calculated based on these data give a better
397 representation of the real behaviour of the swans than the proportions and time budgets based
398 on the raw ACC. And although significantly different, the actual differences between the two
399 datasets is so small that it can be questioned whether this implies a biologically relevant
400 difference.

401 The added value of the SS ACC collection method, through a decreased bout interval in our case,
402 is especially visible in rare behaviours or short duration behaviours, since a sensor with a longer
403 bout interval is more likely to miss these behaviours. Five of the behaviours tested here are not

404 considered rare nor of short duration (aquatic foraging, terrestrial active, swimming, standing
405 resting and sleeping). Flight, however, might be considered rare, especially in non-migratory
406 seasons, when flight is mainly used to get to and from the roost site (i.e. roost flights), a behaviour
407 that tends to last less than 10 min (Nolet et al., 2002). We indeed found a significant difference
408 between the two methods in the number of days that these roost flights were detected. For such
409 an important behaviour in terms of energy expenditure (Nolet et al., 2002), even small differences
410 in duration can have important consequences. Because flight is a biologically relevant and
411 expensive behaviour in terms of energy use, accurate estimation of its occurrence and duration
412 is valuable. For detailed questions with potential management implications an underestimation of
413 flight behaviour can have important consequences. For example geese that are 'scared' 5 times
414 a day as part of a damage control management, fly more and need to compensate for this extra
415 energy expenditure by eating 12-16% more grass (Nolet et al., 2016). This compensational
416 feeding could cause more damage to agricultural fields, while the scaring was actually meant to
417 decrease the damage (Nolet et al., 2016). To obtain accurate model input for such predictions
418 and link them to the feeding and reproductive ecology of the species it is important to be able to
419 estimate the time spent on each behaviour as precise as possible.

420 Reductions of ACC data size such as by using summary statistics as we show here, can be
421 advantageous for future biologging studies. For example in a study of migratory dark-bellied
422 geese (*Branta b. bernicla*) the amount of ACC data that was collected within the limits of collar
423 storage and data transmission only allowed for a very rough behavioural classification into the
424 categories 'active' and 'inactive' (Dokter et al., 2018). Although this yielded interesting results in
425 combination with the GPS data of the same tags, more ACC measurements could have increased
426 the understanding of the behavioural patterns of these geese in their fueling and migration
427 periods.

428 Despite the clear advantage of a decrease in data size and the accompanying possibility to
429 elongate the deployment time or reduce the interval of measurements to obtain a more detailed
430 dataset, the method described here might not be suitable for all study systems. Proper use of
431 summary statistics requires a thorough understanding of the study system and a priori annotation
432 of the behaviour so that the summary statistics can be chosen wisely. Only then will these
433 predictors be useful in classifying the behaviour of interest after collection of the data. When no
434 prior knowledge on behavioural patterns is present, or the behaviour of interest is difficult to
435 capture with commonly used summary statistics or might differ significantly among individuals it
436 is recommended to collect raw ACC data.

437 If the data compression is used to increase the monitoring coverage (this study), the level of detail
438 obtained by using SS opens up the opportunity to study specific research questions that are out
439 of reach with the data yield from raw ACC, such as the example of the roost flights in this study.
440 By using lossy data compression as a means to elongate deployment time, one could answer a
441 whole different set of questions by potentially tracking individuals for several years and compare
442 their time budgets or (migratory) performance (see Harel et al., 2016; Sergio et al., 2014) across
443 seasons or developmental or life history stages. A higher monitoring coverage by using SS not
444 only means a more accurate representation of the time budgets, but also allows for a more in
445 depth study of causal factors and drivers of change. However, the data on these (ecological)
446 drivers then also needs to be very fine-scale which is often not available (Wilmers et al., 2015). A
447 solution is to use the animals themselves to collect valuable data on their environment by including
448 extra sensors in the tracking devices (Kays et al., 2015). This is already successfully done in
449 some marine animals (Evans et al., 2013; Fedak, 2004; Sala et al., 2017). For example, elephant
450 seals (*Mirounga leonina*) equipped with oceanographic sensors collected data on ocean structure
451 and salinity that enabled researchers to map the ice front south of 60°S and calculate the sea ice
452 formation rate from upper ocean salinity levels on rarely observed sites (Charrassin et al., 2008).

453 Collection of environmental data by animal-borne sensors is providing very time- and space-
454 specific information that can be dependent on preferences of the animal, but at the same time
455 this gives a very accurate look inside the lives of these animals and the conditions they encounter.
456 The collection of environmental variables by tracking devices is facilitated by using SS to store
457 the data from the ACC sensor, since the freed storage space and bandwidth can be used for this
458 purpose.

459 The field of biotelemetry is continuously developing. Just as computational developments for the
460 processing of large amounts of biologging data produced by sensors like the accelerometer (see
461 for example Wilson et al. (2018)), the methodology in this study can be seen as a part of this
462 development. Especially in a well-studied system, the behaviours of importance are generally
463 known and these can be reliably classified by using familiar summary statistics. A next step is to
464 use all known information to not only summarize but to classify behaviour on board already (Fig.
465 2C). This might not be possible for all behaviours, but for some very common or easily
466 recognizable behaviours such as sleeping or flying in this study it is feasible. The biologging
467 device could be programmed in such a way that it would attempt to recognize the behaviour
468 performed through time-series classification of raw sensor output (see for example Wilson et al.
469 2018). If it does recognize the behaviour, it can suffice with storing and transmitting a single
470 number or letter for that bout, indicating the specific behaviour. The device could even be
471 programmed in such a way that settings (bout interval and bout duration for example) are
472 dependent on which behaviour is performed (see Harel et al. 2016 for an example of flight
473 detection). If the algorithm does not recognize the behaviour either the SS or the raw ACC data
474 can be stored and (later) sent to the researcher (combination between Fig. 2B and 2C). Often the
475 behaviours that can be classified with very high accuracy together make up a large part of the
476 daily time budget, so this can potentially yield large reductions in data size. With such a 'smart'
477 sampling schedule, prior knowledge about the species is used optimally and the storage space

478 and available bandwidth are used for collecting new information about the study species and
479 behaviours of interest. This makes the proposed lossy data collection method a very lucrative way
480 of reducing data size. Because the behaviours, classification and summary statistics will vary
481 greatly per species and research question, a close collaboration with system developers is
482 necessary to make the proposed progress in remote animal observation. These developments
483 can pave the way for continuous remote monitoring of animal behaviour in the future.

484

485 *Acknowledgements*

486

487 We thank all persons that have contributed to the work in the field or during data analysis: Gerard
488 Müskens, Youri van der Horst, Erik Kleyheeg, Peter de Vries, Thomas Lameris, Anna Hermsen,
489 Jan Vegelin, Fred Cottaar, Sibrand Rinzema, Stefan Vriend, Anne-Lieke Knaven, Nina Thierij.
490 We acknowledge the Royal Burgers' Zoo in Arnhem, Avifauna in Alphen aan de Rijn and GaiaZoo
491 in Kerkrade for their permission to test the GPS/GSM collars and observe the captive Bewick's
492 swans in their institutions. We thank anonymous referees and the associate editor for their
493 constructive comments to earlier versions of this manuscript. The catching and tagging of the wild
494 Bewick's swans was carried out under licences 2016518 of the Centrale Commissie Dierproeven
495 and FF/75/2016/044 of the Flora- en faunawet. RJMN was supported by NWO-NPP grant
496 866.15.206.

497

498 Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.6djh9w0x9>
499 (Nuijten et al., 2019).

500 **References**

- 501 Adoram-Kershner, L., Bruce, T., Morris, C., Holser, R., Costa, D., Flikkema, P. G., & Shafer, M. W.
502 (2017). Modeling and testing solar power for globally migrating submarine systems. *Oceans-*
503 *Anchorage IEEE.*, 1–9.
- 504 Allan, B. M., Nimmo, D. G., Ierodiaconou, D., VanDerWal, J., Koh, L. P., & Ritchie, E. G. (2018).
505 Futurecasting ecological research: the rise of technoecology. *Ecosphere*, 9(5), e02163.
- 506 Beekman, J. H., Nolet, B. A., & Klaassen, M. (2002). Skipping swans: fuelling rates and wind conditions
507 determine differential use of migratory stopover sites of Bewick's Swans *Cygnus bewickii*. *Ardea*,
508 90, 437–460.
- 509 Bishop, C. M., Hawkes, L. A., Chua, B., Frappell, P. B., Milsom, W. K., Natsagdorj, T., Newman, S. H.,
510 Scott, G. R., Takekawa, J. Y., Wikelski, M., & Butler, P. J. (2015). The roller coaster flight strategy of
511 bar-headed geese conserves energy during Himalayan migrations. *Science*, 347, 250–254.
- 512 Bom, R. A., Bouten, W., Piersma, T., Oosterbeek, K., & van Gils, J. A. (2014). Optimizing acceleration-
513 based ethograms: The use of variable-time versus fixed-time segmentation. *Movement Ecology*,
514 2(1), 1–8.
- 515 Bouten, W., Baaij, E. W., Shamoun-Baranes, J., & Camphuysen, K. C. J. (2013). A flexible GPS tracking
516 system for studying bird behaviour at multiple scales. *Journal of Ornithology*, 154(2), 571–580.
- 517 Broell, F., Noda, T., Wright, S., Domenici, P., Steffensen, J. F., Auclair, J.-P., & Taggart, C. T. (2013).
518 Accelerometer tags: detecting and identifying activities in fish and the effect of sampling frequency.
519 *Journal of Experimental Biology*, 216(7), 1255–1264.
- 520 Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable
521 through acceleration logging of animal behavior. *Animal Biotelemetry*, 1(1), 1–20.
- 522 Brown, D. D., LaPoint, S., Kays, R., Heidrich, W., Kümmeth, F., & Wikelski, M. (2012). Accelerometer-
523 informed GPS telemetry: Reducing the trade-off between resolution and longevity. *Wildlife Society*
524 *Bulletin*, 36(1), 139–146.
- 525 Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets GPS-based
526 radiotelemetry: A perfect storm of opportunities and challenges. *Philosophical Transactions of the*
527 *Royal Society B: Biological Sciences*, 365(1550), 2157–2162.
- 528 Charrassin, J.-B., Hindell, M., Rintoul, S. R., Roquet, F., Sokolov, S., Biuw, M., ... Guinet, C. (2008).
529 Southern Ocean frontal structure and sea-ice formation rates revealed by elephant seals. *PNAS*,
530 105(33), 11634–11639.

- 531 Chimienti, M., Cornulier, T., Owen, E., Bolton, M., Davies, I. M., Travis, J. M. J., & Scott, B. E. (2016). The
532 use of an unsupervised learning approach for characterizing latent behaviors in accelerometer data.
533 *Ecology and Evolution*, 6(3), 727–741.
- 534 Cooke, S. J., Hinch, S. G., Wikelski, M., Andrews, R. D., Kuchel, L. J., Wolcott, T. G., & Butler, P. J.
535 (2004). Biotelemetry: a mechanistic approach to ecology. *Trends in Ecology & Evolution*, 19(6),
536 334–343.
- 537 Cox, S. L., Orgeret, F., Gesta, M., Rodde, C., Heizer, I., Weimerskirch, H., & Guinet, C. (2018).
538 Processing of acceleration and dive data on-board satellite relay tags to investigate diving and
539 foraging behaviour in free-ranging marine predators. *Methods in Ecology and Evolution*, 9(1), 64–
540 77.
- 541 Daniel Kissling, W., Pattemore, D. E., & Hagen, M. (2014). Challenges and prospects in the telemetry of
542 insects. *Biological Reviews*, 89(3), 511–530.
- 543 De Cock, J., Li, Z., Manohara, M., & Aaron, A. (2016). Complexity-based consistent-quality encoding in
544 the cloud. *Image Processing (ICIP), 2016 IEEE International Conference*, 1484–1488.
- 545 Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R. A., Perrins, C. M., & Guilford, T. (2012).
546 Behavioural mapping of a pelagic seabird: combining multiple sensors and a hidden Markov model
547 reveals the distribution of at-sea behaviour. *Journal of the Royal Society Interface* 10(78).
- 548 Dokter, A. M., Fokkema, W., Ebbinge, B. S., Olf, H., van der Jeugd, H. P., & Nolet, B. A. (2018).
549 Agricultural pastures challenge the attractiveness of natural saltmarsh for a migratory goose.
550 *Journal of Applied Ecology*, 55(6), 2707–2718.
- 551 Duriez, O., Kato, A., Tromp, C., Dell’Omo, G., Vysotski, A. L., Sarrazin, F., & Ropert-Coudert, Y. (2014).
552 How cheap is soaring flight in raptors? A preliminary investigation in freely-flying vultures. *PLoS*
553 *ONE*, 9(1).
- 554 Evans, K., Lea, M.-A., & Patterson, T. A. (2013). Recent advances in bio-logging science: Technologies
555 and methods for understanding animal behaviour and physiology and their environments. *Deep Sea*
556 *Research Part II: Topical Studies in Oceanography*, 88–89, 1–6.
- 557 Fedak, M. (2004). Marine animals as platforms for oceanographic sampling: a “win/win”
558 situation for biology and operational oceanography. *Memoirs of National Institute of Polar Research*,
559 (58), 133–147.
- 560 Flack, A., Nagy, M., Fiedler, W., Couzin, I. D., & Wikelski, M. (2018). From local collective behavior to
561 global migratory patterns in white stork. *Science*, 360(May), 911–914.

- 562 Fridrich, J., Goljan, M., & Du, R. (2015). Lossless data embedding for all image formats. *Security and*
563 *Watermarking of Multimedia Contents IV*, 4675, 572–584.
- 564 Harel, R., Horvitz, N., & Nathan, R. (2016). Adult vultures outperform juveniles in challenging thermal
565 soaring conditions. *Scientific Reports*, 6(1), 27865.
- 566 Heerah, K., Cox, S. L., Blevin, P., Guinet, C., & Charrassin, J.-B. (2019). Validation of Dive Foraging
567 Indices Using Archived and Transmitted Acceleration Data: The Case of the Weddell Seal. *Frontiers*
568 *in Ecology and Evolution*, 7(February), 1–15.
- 569 Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). ECOLOGY. Terrestrial animal tracking as an
570 eye on life and planet. *Science*, 348(6240), aaa2478.
- 571 Kays, R., Tilak, Sa., Crofoot, M., Fountain, T., Obando, D., Ortega, A., Kuemmeth, F., Mandel, J.,
572 Swenson, G., Lambert, T., Hirsch, B., & Wikelski, M. (2011). Tracking Animal Location and Activity
573 with an Automated Radio Telemetry System in a Tropical Rainforest. *The Computer Journal*, 54(12),
574 1931–1948.
- 575 Lameris, T. K., Müskens, G. J., Kölzsch, A., Dokter, A. M., Van der Jeugd, H. P., & Nolet, B. A. (2018).
576 Effects of harness-attached tracking devices on survival, migration, and reproduction in three
577 species of migratory waterfowl. *Animal Biotelemetry*, 6(1), 7.
- 578 Laplanche, C., Marques, T. A., & Thomas, L. (2015). Tracking marine mammals in 3D using electronic tag
579 data. *Methods in Ecology and Evolution*, 6(9), 987–996.
- 580 le Roux, S. P., Wolhuter, R., Stevens, N., & Niesler, T. (2018). Reduced Energy and Memory
581 Requirements by On-Board Behavior Classification for Animal-Borne Sensor Applications. *IEEE*
582 *Sensors Journal*, 18(10), 4261–4268.
- 583 Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- 584 Liechti, F., Bauer, S., Dhanjal-Adams, K. L., Emmenegger, T., Zehtindjiev, P., & Hahn, S. (2018).
585 Miniaturized multi-sensor loggers provide new insight into year-round flight behaviour of small trans-
586 Sahara avian migrants. *Movement Ecology*, 6(1), 1–10.
- 587 Mansfield, K. L., Wyneken, J., Porter, W. P., & Luo, J. (2014). First satellite tracks of neonate sea turtles
588 redefine the “lost years” oceanic niche. *Proceedings of the Royal Society. Biological Sciences.*, 281.
- 589 Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. (2012). Using tri-axial
590 acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools
591 illustrated for griffon vultures. *Journal of Experimental Biology*, 215, 986–996.

- 592 Noda, T., Kawabata, Y., Arai, N., Mitamura, H., & Watanabe, S. (2014). Animal-mounted
593 gyroscope/accelerometer/magnetometer: In situ measurement of the movement performance of
594 fast-start behaviour in fish. *Journal of Experimental Marine Biology and Ecology*, 451, 55–68.
- 595 Nolet, B. A., Bevan, R. M., Klaassen, M., Langevoord, O., & Van Der Heijden, Y. G. J. T. (2002). Habitat
596 switching by Bewick's swans: maximization of average long-term energy gain? *Journal of Animal
597 Ecology*, 71(6), 979–993.
- 598 Nolet, B. A., Kölzsch, A., Elderenbosch, M., & van Noordwijk, A. J. (2016). Scaring waterfowl as a
599 management tool: how much more do geese forage after disturbance? *Journal of Applied Ecology*,
600 53(5), 1413–1421.
- 601 Norkin, A., De Cock, J., Mavlankar, A., & Aaron, A. (2016). More Efficient Mobile Encodes for Netflix
602 Downloads. The Netflix Tech Blog.
- 603 Nuijten, R. J. M., Gerrits, T., Shamoun-Baranes, J. & Nolet, B. A. (2019). Data from: Less is more: on-
604 board lossy compression of accelerometer data increases biologging capacity. Dryad Digital
605 Repository, <https://doi.org/10.5061/dryad.6djh9w0x9>
- 606 Nuijten, R. J. M., Kölzsch, A., van Gils, J. a., Hoyer, B. J., Oosterbeek, K., de Vries, P. P., Klaassen, M., &
607 Nolet, B. a. (2014). The exception to the rule: retreating ice front makes Bewick's swans *Cygnus
608 columbianus bewickii* migrate slower in spring than in autumn. *Journal of Avian Biology*, 45(2), 113–
609 122.
- 610 Patel, A., Stocks, B., Fisher, C., Nicolls, F., & Boje, E. (2017). Tracking the Cheetah Tail Using Animal-
611 Borne Cameras, GPS, and an IMU. *IEEE Sensors Letters*, 1(4), 1–4.
- 612 Rees, E. (2006). *The Bewick's Swan*. T & AD Poyser.
- 613 Ryan, P. G., Petersen, S. L., Peters, G., & Grémillet, D. (2004). GPS tracking a marine predator: The
614 effects of precision, resolution and sampling rate on foraging tracks of African Penguins. *Marine
615 Biology*, 145(2), 215–223.
- 616 Sala, J. E., Pisoni, J. P., & Quintana, F. (2017). Three-dimensional temperature fields of the North
617 Patagonian Sea recorded by Magellanic penguins as biological sampling platforms. *Estuarine,
618 Coastal and Shelf Science*, 189, 203–215.
- 619 Sergio, F., Tanferna, A., De Stephanis, R., Jiménez, L. L., Blas, J., Tavecchia, G., Preatoni, D., & Hiraldo,
620 F. (2014). Individual improvements and selective mortality shape lifelong migratory performance.
621 *Nature*, 515(7527), 410–413.
- 622 Shaffer, S. A., Cockerham, S., Warzybok, P., Bradley, R. W., Jahncke, J., Clatterbuck, C. A., Lucia, M.,

- 623 Jelincic, J. A., Cassell, A. L., Kelsey, E. C., & Adams, J. (2017). Population-level plasticity in
624 foraging behavior of western gulls (*Larus occidentalis*). *Movement Ecology*, 5(1), 27.
- 625 Shamoun-Baranes, J., Bom, R., van Loon, E. E., Ens, B. J., Oosterbeek, K., & Bouten, W. (2012). From
626 sensor data to animal behaviour: An oystercatcher example. *PLoS ONE*, 7(5), 28–30.
- 627 Shamoun-Baranes, J., Bouten, W., van Loon, E. E., Meijer, C., & Camphuysen, C. J. (2016). Flap or
628 soar? How a flight generalist responds to its aerial environment. *Philosophical Transactions of the*
629 *Royal Society B: Biological Sciences*, 371.
- 630 Therrien, J. F., Gauthier, G., & Bêty, J. (2012). Survival and reproduction of adult snowy owls tracked by
631 satellite. *Journal of Wildlife Management*, 76(8), 1562–1567.
- 632 Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., & Bates, K. K. (2010). Global positioning system and
633 associated technologies in animal behaviour and ecological research. *Philosophical Transactions of*
634 *the Royal Society of London. Biological Sciences*, 365(1550), 2163–2176.
- 635 Tondare, S. M., Panchal, S. D., & Kushnure, D. . (2014). Evolutionary steps from 1G to 4.5 G.
636 *International Journal of Advanced Research in Computer and Communication Engineering*, 3(4).
- 637 Wascher, C. A. F., Kotrschal, K., & Arnold, W. (2018). Free-living greylag geese adjust their heart rates
638 and body core temperatures to season and reproductive context. *Scientific Reports*, 8(1), 2142.
- 639 Watanabe, Y., Mitani, Y., Sato, K., Cameron, M. F., & Naito, Y. (2003). Dive depths of Weddell seals in
640 relation to vertical prey distribution as estimated by image data. *Marine Ecology Progress Series*,
641 252, 283–288.
- 642 Watanabe, Y. Y., Ito, M., & Takahashi, A. (2014). Testing optimal foraging theory in a penguin-krill
643 system. *Proceedings of the Royal Society. Biological Sciences*, 281(1779), 20132376.
- 644 Wikelski, M., Kays, R. W., Kasdin, N. J., Thorup, K., Smith, J. A., & Swenson, G. W. (2007). Going wild:
645 what a global small-animal tracking system could do for experimental biologists. *Journal of*
646 *Experimental Biology*, 210(2), 181–186.
- 647 Williams, T. M., Wolfe, L., Davis, T., Kendall, T., Richter, B., Wang, Y., Bryce, C., Elkaim, G. H., &
648 Wilmers, C. C. (2014). Instantaneous energetics of puma kills reveal advantage of felid sneak
649 attacks. *Science*, 346(6205), 81–85.
- 650 Wilmers, C. C., Nickel, B., Bryce, C. M., Smith, J. A., Wheat, R. E., & Yovovich, V. (2015). The golden
651 age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology. *Ecology*,
652 96(7), 1741–1753.

- 653 Wilson, A. D. M., Wikelski, M., Wilson, R. P., & Cooke, S. J. (2015). Utility of biological sensor tags in
654 animal conservation. *Conservation Biology*, 29(4), 1065–1075.
- 655 Wilson A. M., Lowe J. C., Roskilly K., Hudson P. E., Golabek K. A., McNutt J. W. (2013) Locomotion
656 dynamics of hunting in wild cheetahs. *Nature* 498:185-189.
- 657 Wilson, R. P., Holton, M. D., di Virgilio, A., Williams, H., Shepard, E. L. C., Lambertucci, S., Quintana, F.,
658 Sala, J. E., Balaji, B., Lee, E. S., Srivastava, M., Scantlebury, D. M., & Duarte, C. M. (2018). Give
659 the machine a hand: A Boolean time-based decision-tree template for rapidly finding animal
660 behaviours in multisensor data. *Methods in Ecology and Evolution*, 9(11), 2206–2215.
- 661 Wilson, R. P., Shepard, E. L. C., & Liebsch, N. (2008). Prying into the intimate details of animal lives: use
662 of a daily diary on animals. *Endangered Species Research*, 4(1–2), 123–137.
- 663 Yoda, K., Naito, Y., Sato, K., Takahashi, A., Nishikawa, J., Ropert-Coudert, Y., Kurita, M., & Le Maho, Y.
664 (2001). A new technique for monitoring the behaviour of free-ranging Adélie penguins. *Journal of*
665 *Experimental Biology*, (204), 685–690.
- 666 Yoda, K., Sato, K., Niizuma, Y., Kurita, M., Bost, C. A., Le Maho, Y., & Naito, Y. (1999). Precise
667 monitoring of porpoising behaviour of Adélie penguins determined using acceleration data loggers.
668 *The Journal of Experimental Biology*, (202), 31021–33126.

669

670 **Supporting information**

671 The following Supporting Information is available for this article online.