

1 **The challenges of estimating the distribution of flight heights from**
2 **telemetry or altimetry data**

3 Guillaume Péron (1)# guillaume.peron@univ-lyon1.fr

4 Justin M. Calabrese (2,3) CalabreseJ@si.edu

5 Olivier Duriez (4) olivier.duriez@cefe.cnrs.fr

6 Christen H. Fleming (2,3) chris.h.fleming@gmail.com

7 Ruth García-Jiménez (5) ruth.garciajimenez@gmail.com

8 Alison Johnston (6) aj327@cornell.edu

9 Sergio Lambertucci (7) slambertucci@comahue-conicet.gob.ar

10 Kamran Safi (8) ksafi@orn.mpg.de

11 Emily L.C. Shepard (9) E.L.C.Shepard@swansea.ac.uk

12 (1) Univ Lyon, Université Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive UMR5558, F-
13 69622 Villeurbanne, France

14 (2) Smithsonian Conservation Biology Institute, National Zoological Park, Front Royal, Virginia 22630,
15 USA

16 (3) Department of Biology, University of Maryland, College Park, Maryland 20742, USA

17 (4) Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS-Université de Montpellier-EPHE-
18 Université Paul Valéry, 1919 Route de Mende, 34293 Montpellier cedex 5, France

19 (5) Department of Animal Science, Faculty of Life Sciences and Engineering, University of Lleida,
20 25198 Lleida, Spain

21 (6) Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA & Department of Zoology,
22 Conservation Science Group, University of Cambridge, Cambridge, UK

23 (7) Grupo de Investigaciones en Biología de la Conservación, Laboratorio Ecotono, INIBIOMA
24 (CONICET–Universidad Nacional del Comahue), Quintral 1250, 8400 Bariloche, Argentina

25 (8) Max Planck Institut für Ornithologie, Am Obstberg 1, 78315 Radolfzell, Germany

26 (9) Department of Biosciences, College of Science, Swansea University, Swansea SA2 8PP, United
27 Kingdom

28 # corresponding author

29 **Running title:** Sampling error in flight height data

30

31 **Abstract**

32 *Background.* Global positioning systems (GPS) and altimeters are increasingly used to monitor
33 vertical space use by aerial species, a key aspect of their niche that we need to know to understand
34 their ecology and conservation needs, and to manage our own use of the airspace. However, there
35 are various sources of error in flight height data ("height" above ground, as opposed to "altitude"
36 above a reference like the sea level): vertical error from the devices themselves, error in the ground
37 elevation below the tracked animals, and error in the horizontal position of the animals and thus the
38 predicted ground elevation below them.

39 *Methods.* We used controlled field trials, simulations, and the reanalysis of raptor case studies with
40 state-space models to illustrate the effect of improper error management.

41 *Results.* Errors of a magnitude of 20 meters appear in benign conditions (expected to be larger in
42 more challenging context). These errors distort the shape of the distribution of flight heights, inflate
43 the variance in flight height, bias behavioural state assignments, correlations with environmental
44 covariates, and airspace management recommendations. Improper data filters such as removing all
45 negative recorded flight height records introduce several biases in the remaining dataset, and
46 preclude the opportunity to leverage unambiguous errors to help with model fitting. Analyses that
47 ignore the variance around the mean flight height, e.g., those based on linear models of flight height,
48 and those that ignore the variance inflation caused by telemetry errors, lead to incorrect inferences.

49 *Conclusion.* The state-space modelling framework, now in widespread use by ecologists and
50 increasingly often automatically implemented within on-board GPS data processing algorithms,
51 makes it possible to fit flight models directly to raw flight height records, with minimal data pre-
52 selection, and to analyse the full distribution of flight heights, not just the mean. In addition to basic
53 research about aerial niches, behaviour quantification, and environmental interactions, we highlight
54 the applied relevance of our recommendations for airspace management and the conservation of
55 aerial wildlife.

56

57 Introduction

58 Describing the distribution of animals in environmental space is fundamental to understanding their
59 resource requirements, cognitive processes, energetic strategies, and ecological characteristics. The
60 distribution of animals in horizontal space has dominated ecological studies (Nathan et al. 2008),
61 however the vertical dimension is also important for flying animals, and for that matter also diving
62 and tree-climbing animals (Weimerskirch et al. 2005, Kunz et al. 2007, Bishop et al. 2015, Liechti et al.
63 2018). For example, flight height data could help documenting the vertical niche and community
64 ecology of aerial foragers (Arlettaz et al. 1997, Siemers and Schnitzler 2004). Flight height data
65 quantify the behaviour of flying animals and their flight strategies (Pirotta et al. 2018, Murgatroyd et
66 al. 2018), and their relationships with environmental factors (e.g., Péron et al. 2017). From an applied
67 perspective, we need an accurate, error-free description of the distribution of birds and other
68 animals in the aerosphere to avoid collisions with man-made structures, which is key to aircraft
69 safety and animal conservation, in the current context of increasing human encroachment into the
70 airspace (Lambertucci et al. 2015, Davy et al. 2017).

71 However, monitoring vertical airspace use by wildlife remains challenging. Ground-based surveys are
72 limited in their field of vision and time window. Airborne monitoring (e.g., from glider planes) is
73 logistically challenging and constrained by weather conditions. Radar-based methodologies are not
74 usually specific enough to assign records to species (but see Zaugg et al. 2008, Dokter et al. 2013).
75 Animal-borne tracking methodologies such as global positioning systems (GPS) and altimeters have
76 therefore become popular to monitor flying species (López-López 2016). They record data even when
77 the animals are out of sight for ground-based observers, over extensive, potentially uninterrupted
78 periods of time, and with no uncertainty about which species or individuals are being monitored. For
79 example, we can record raptors soaring over the high sea at night (Duriez et al. 2018). However, the
80 data that GPS and altimeters record are not error-free (D'Eon et al. 2002, Frair et al. 2004, Jerde and
81 Visscher 2005, Brost et al. 2015). Errors are particularly evident in the vertical axis because there are
82 unpassable barriers, e.g., the ground. Usually, a few unambiguously erroneous positions are
83 recorded beyond these barriers (Katzner et al. 2012, Ross-Smith et al. 2016, Weimerskirch et al.
84 2016, Péron et al. 2017, Krone and Treu 2018, Roeleke et al. 2018).

85 Most of the research into ways to deal with sampling errors in positioning data has focused on
86 horizontal animal movement (Freitas et al. 2008, Albertsen et al. 2015, Brost et al. 2015, Fleming et
87 al. 2016). There is very little guidance for ecologists about the challenges specific to vertical space
88 use data (Poessel et al. 2018). Many practitioners consider that vertical movement data need to be
89 “filtered” before analysis, i.e., they discard some records before proceeding with the analysis. They

90 may discard records that are too far from preceding ones (as often done for horizontal data; Freitas
91 et al., 2008), too far beyond impassable barriers (Katzner et al. 2012, Krone and Treu 2018), or
92 obtained from an unreliable configuration of the GPS satellite network (Poessel et al., 2018). Instead
93 of discarding the more erroneous records, researchers have also sometimes chosen to reset them to
94 plausible values (Weimerskirch et al. 2016, Roeleke et al. 2018). However, when applied improperly,
95 such filters can have undesirable consequences. We start by reviewing the sources of error in GPS
96 and altimeter flight height data (Part 1). In Part 2, we reanalyse case studies into the flight height of
97 three raptor species (Péron et al. 2017), and complement them with novel data from controlled field
98 trials and from simulations, in order to illustrate the stakes of proper error-handling in vertical
99 airspace use data.

100

101 **Part 1: Review of the sources of error in flight height data** 102 **from GPS and altimeters**

103 Throughout we refer to flight height h , which is the distance to the ground below the bird, different
104 from flight altitude z . The flight altitude denotes the distance to a reference altitude, often the
105 ellipsoid, i.e., a geometrically perfect (but simplistic) model of the sea level. Alternatively, some GPS
106 units may provide the altitude relative to the empirical sea level measured at a reference point over
107 a reference period (e.g., in France the “NGF-IGN 1969” norm means that altitude is measured
108 relative to the mean sea level in the port of Marseille between 1884 and 1896), or relative to the
109 geoid, which is a model of the sea level if it was only influenced by the local gravitational field and
110 the rotation of the Earth (Fowler 2005). There are databases and simple formulae to convert from
111 one system of reference to another, but this nevertheless represents a first potential source of error
112 in flight height data.

113 Flight height above the ground is computed as $h = z - z_{DEM}(x, y)$, where $z_{DEM}(x, y)$ is the ground
114 altitude predicted by a digital elevation model (DEM) at the recorded horizontal position (x, y) , in
115 the same system of reference as z . Errors in h can then be caused by errors in any of the three
116 components: z , z_{DEM} , or (x, y) (Fig. 1). Importantly, depending on the application, researchers might
117 want to study z not h (Pirootta et al. 2018, Murgatroyd et al. 2018). In the list below, sources of error
118 #3-#5 do not influence z .

119 1. Error in z when z is given by a GPS.

120 If recorded by a GPS, z is affected by the “user equivalent range error” (UERE) and the “vertical
121 dilution of precision” (VDOP) (Parkinson and Spilker 1996, Sanz Subirana et al. 2013).

122 The UERE stems from diffusion and diffraction in the atmosphere, reflection from obstacles, and
123 receiver noise (Parkinson and Spilker 1996, Sanz Subirana et al. 2013). The acronym UERE usually
124 directly refers to the root mean squared error, but here we will use the notation σ_{UERE} instead.
125 σ_{UERE} is usually in the order of a few meters and considered constant over time for a given device.
126 Some GPS manufacturers specify the horizontal σ_{UERE} , or alternatively it can be estimated from the
127 data (Johnson et al. 2008). The σ_{UERE} is however reputedly larger in the vertical axis than the
128 horizontal axes (D'Eon et al. 2002, Bouten et al. 2013), meaning that manufacturer-provided σ_{UERE}
129 should be considered conservative for vertical applications and should be used with appropriate
130 caution.

131 The vertical position dilution of precision factor (VDOP) quantifies the effect of changes in the size
132 and spatial configuration of the available satellite network on the precision of GPS records (Parkinson
133 & Spilker, 1996; Sanz Subirana et al., 2013; Fig. A1). The more satellites are available and the more
134 evenly spread apart they are, the more reliable the positioning is. Some GPS manufacturers do
135 provide a VDOP value for each record, but many only provide a more generic DOP value.

136 When σ_{UERE} and VDOP are known, the error-generating process can then be approximated by a
137 Gaussian process with time-varying standard deviation $\sigma_z(t) = \text{VDOP}(t) \cdot \sigma_{\text{UERE}}$ (Eq. 6.45 in Sanz
138 Subirana et al., 2013). Therefore, the DOP is not a direct index of precision. The spread of the error
139 distribution increases with the DOP, but the error on any given record is stochastic. The DOP is
140 therefore not intended to be used as a data filter (e.g., discard any data with DOP above a given
141 threshold), but instead it should be used to model the error-generating process.

142

143 2. Error in z when z is given by an altimeter

144 If recorded using an altimeter, z is computed from the barometric pressure, using the formula
145 $z = c \cdot T \cdot \log(P_{\text{REF}}/P)$ (Monaldo et al. 1986, Crocker and Jackson 2018). c is a calibration constant
146 that mostly depends on the composition of the air (e.g., percentage of vapour) and on the
147 gravitational field. T is the air temperature in Kelvin, P is the air pressure, and P_{REF} is the air pressure
148 at an elevation of reference (both pressures in mbar or in Pascal). However, this formula only holds
149 when the atmosphere is at equilibrium. Changes in temperature, pressure, and air composition, i.e.,
150 the weather, alter the link between z and P . These influences are difficult to control fully because one
151 would need to measure the weather variables both where the bird is, and at the reference elevation
152 immediately below the bird. In other words, altimeters can be more accurate than GPS to monitor
153 flight height, but only over short periods of time when the weather can be considered constant and
154 the altimeter is calibrated for that weather. One should ideally regularly re-calibrate the altimeters

155 using direct observations of flight height and accurate measures of P_{REF} and T . Unfortunately, field
156 calibrations are rarely feasible in practice (but see Shepard et al., 2016; Borkenhagen et al., 2018).
157 The consequence is that altimeters are often miscalibrated. The degree of miscalibration depends
158 mostly on the weather. This generates temporally autocorrelation in the error time series. Over a
159 restricted time period, the error patterns are thus more akin to a bias (a systematic over- or under-
160 estimation of flight height) than to an error in the statistical sense of a zero-mean, identically and
161 independently distributed random process. Importantly, altimeter data still allow one to compute
162 the derivative of flight height, i.e., climb rate, because the amount of bias can be considered constant
163 over short periods of time. In Part 2.1, we will directly compare the errors from GPS and altimeters
164 using controlled field experiments.

165 3. GPS horizontal error.

166 (x, y) is also affected by a user equivalent range error and a dilution of precision (Fig. 1). The
167 horizontal error in (x, y) can thus also be described as a Gaussian process with time-varying standard
168 deviation: $\sigma_{xy}(t) = 1/\sqrt{2} \cdot HDOP(t) \cdot \sigma_{USERE}$. Note that we use here a horizontal dilution of
169 precision factor, HDOP. An often-overlooked consequence of errors in the horizontal position is that
170 they introduce flaws in the link to spatially-explicit environmental covariates (Hays et al. 2001,
171 Bradshaw et al. 2007). In particular, the ground elevation z_{DEM} is extracted from a location (x, y)
172 that is slightly different from the true location (Katzner et al. 2012). If the terrain is very rough, then
173 the ground elevation at the recorded location (x, y) may be significantly different from the ground
174 elevation below the actual location of the bird. In Part 2.2 we will use simulations to quantify the
175 influence of horizontal errors.

176 4. Interpolation error in z_{DEM} .

177 z_{DEM} is interpolated from discrete ground elevation measurements (Gorokhovich and Voustianiouk
178 2006, Januchowski et al. 2010). The ground elevation is measured at a few select locations, but it is
179 interpolated between them. The result of the interpolation is then rasterized at a set resolution, and
180 the result is the DEM. This process can be quite imprecise (Gorokhovich and Voustianiouk 2006,
181 Januchowski et al. 2010). At a cliff, for example, the ground elevation may drop by several hundred
182 meters within a single pixel of the DEM.

183 5. Errors in DEM base data.

184 The original measurements from which DEMs are interpolated are not necessarily error free either.
185 These errors are assumed small relative to the other sources, however, there is, to our knowledge,

186 not much information available about the base datasets from which DEM are interpolated and their
187 precision.

188

189 **Part 2: Field trials, simulations, and reanalysis of raptor** 190 **data**

191 **Material and Methods**

192 **Controlled field trials**

193 To quantify the magnitude of the vertical error in altimeters and GPS devices, we conducted three
194 controlled trial experiments.

195 First, we attached an “Ornitrack 25” GPS-altimeter unit (Ornitela) to a drone. We then flew the drone
196 above the rooftop of the Max-Planck institute in Radolfzell, Germany at heights ranging from 0
197 (drone landed on the rooftop) to 90m. We conducted 6 flight sessions over two days, each lasting
198 between 15 and 140min, collecting one record every ten minutes for a total of 30 records. We also
199 monitored the air pressure and temperature on the rooftop, which we used to recalibrate the
200 altimeter post-hoc. Lastly, the drone carried a separate, on-board, altimeter.

201 In a second, separate experiment, we attached two “Gipsy 5” GPS units (Technosmart) to an ultra-
202 light aircraft, with a vertical distance of 1.8m between the two units. We then flew the aircraft near
203 Radolfzell while the two units simultaneously tracked its flight height, collecting one record per
204 second for a total of 11.5 hours over 5 days.

205 Third, we compared the vertical positions recorded by 4 different units from 3 different
206 manufacturers: Technosmart (AxyTrek and Gipsy 5), Microwave (GPS-GSM 20-70), and Ornitela (GPS-
207 GSM Ornitrack 85). We (RG and OD) carried these units to 21 known geodesic points, of which the
208 altitude was precisely documented by the French National Geographic Institute. The units recorded
209 their position once every minute for a total of 894, 934, 560, and 563 data points, keeping only the
210 unit * location combinations that yielded more than 25 fixes. We computed the bias and root mean
211 squared error of the vertical measurement by comparing these data to the actual, known altitudes of
212 the geodesic points. Importantly, the manufacturers do not use the same reference to compute the
213 altitude: Microwave uses the geoid (WGS 84 EGM-96 norm), whereas the others use the mean sea
214 level (assumed to correspond to the local reference, meaning the NGF-IGN 1969 norm, but sea
215 below). We expressed all altitudes in the same norm before computing biases and errors, and
216 accounted for sampling effort (number of fixes) and location when comparing the performance of
217 different units.

218 **Simulations of flight tracks**

219 We simulated flight tracks that followed Ornstein-Uhlenbeck processes (Dunn and Gipson 1977). This
220 is a class of continuous-time stochastic models, which is not specific to vertical movement or even to
221 movement (Dunn and Gipson 1977). In the case of vertical movement, the parameters of the
222 Ornstein-Uhlenbeck processes control the mean flight height, the variance in flight height, and the
223 temporal autocorrelation in the flight height time series. We transformed the raw Ornstein-
224 Uhlenbeck simulations using an atanh link as described by Péron et al. (2017) to enforce positive
225 flight height. The time unit was arbitrary. An attractive feature of simulations in the context of this
226 study is that we know both the *true* flight height and the *recorded* flight height, which is the true
227 flight height plus an independent and identically distributed zero-mean Gaussian error.

228 **Simulations of synthetic landscapes**

229 The objective was to quantify the influence of horizontal errors. We generated synthetic landscapes
230 of varying complexity and roughness (Fig. A2). We then transposed the flight track of a lesser kestrel
231 *Falco naumanni* over these synthetic landscapes. The individual originally flew over extremely flat
232 terrain (the Crau steppe in France). The data (Pilard and OD, unpublished) were collected every 3
233 minutes using a Gipsy 5 GPS unit from Technosmart, and processed through the state-space model of
234 Péron et al. (2017) to account for real sampling errors before use. We then added simulated random
235 telemetry noise of controlled standard deviation.

236 **Raptor case studies**

237 We reanalysed the data from Péron et al. (2017), where the field procedure, data selection, and data
238 analysis procedures are described in full. Briefly, we studied three species of large soaring raptors:
239 Andean condors *Vultur gryphus* (five juveniles, 1,692 individual.days of monitoring, 15 minute
240 interval), Griffon vultures *Gyps fulvus* (eight adults, 2,697 individual.days, 1-5 minute interval), and
241 Golden eagles *Aquila chrysaetos* (six adults, 3,103 individual.days, 6-10 minute interval). After
242 applying the analytical procedure, for each data point, we could compare the *corrected* position, an
243 estimate of the *true* position, to the *recorded* position, which was affected by the sources of errors
244 we listed in Part 1.

245 We selected the period between 11:00 and 15:00, which concentrates condor activity and therefore
246 flight time, and discarded other records. For the vultures, we selected the period between 09:00 and
247 16:00. For the eagles, we selected the period between 08:00 and 17:00 and, because a lot of time is
248 spent motionless in this species even during their core activity period, we further removed all the
249 records that were less than 15 meters from the previous record. We acknowledge the arbitrary
250 nature of this data selection and emphasize that it is not necessary or even recommended to apply

251 such filters before analysis. We however stress that in the context of the present study, the case
252 studies perform an illustrative function, meaning that we use them to highlight the effect of
253 improper error-handling, at least during the particular time periods that we selected for analysis
254 because we consider them relevant for biological inference, and that the same analytical procedures
255 can indiscriminately be applied to other time frames.

256 **Collision risk**

257 In several instances, we will illustrate the potential effect of improper data-handling on management
258 recommendations by estimating the risk of collision with wind turbines as the proportion of records
259 between 60 and 180m above ground (assuming no behavioural adjustment in the presence of wind
260 turbines). Collision risk estimated from GPS tracks is increasingly used to make recommendations
261 about the choice of locations for new turbines, or to schedule the operation of existing ones. We
262 expected that the estimated collision risk would depend on flight parameters (mean flight height,
263 variance in flight height), on the magnitude of errors, and on error-handling. For example, a large
264 variance in flight height might lead to a high collision risk even if the mean flight height is beyond the
265 collision zone. Improperly handled errors may lead to positions being erroneously recorded in the
266 collision zone when the birds actually flew outside of it, and *vice versa*. The same type of thinking
267 could be applied to other types of collision risk, e.g., antennas, utility lines, buildings with bay
268 windows, except that the collision zone would be at a different height.

269 **Part 2.1: The magnitude of vertical errors in GPS and altimeters**

270 During the first controlled field trial (with the drone), DOP values between 1.2 and 1.6 indicated that
271 the configuration of the satellite network was reliable throughout. Nevertheless, 6.7% of the GPS
272 flight height records were below the rooftop height, i.e., obviously erroneous. For the altimeter, with
273 default settings, 10% of the records were below the rooftop height. The default settings of the
274 altimeter therefore did not correspond to the atmospheric conditions during the experiment. The
275 standard deviation of the difference between the recalibrated altimetry and the GPS data was 22m,
276 between the recalibrated altimetry and default-setting altimetry it was 14m, and between the
277 recalibrated altimetry and the on-board drone altimeter it was 19m. This means that, with default
278 settings, the altimeters had approximately the same precision as the GPS.

279 During the second controlled field trial (with two GPS units attached to the same aircraft), in 35% of
280 cases, the lower unit was erroneously recorded above the higher unit. The standard deviation of the
281 difference between the height recorded by the two units was 7.1m. The highest of the two units
282 recorded 3% of negative flight heights. The lowest unit recorded 13% of negative flight heights.

283 During the third controlled field trial (with GPS units carried to a geodesic point of precisely known
284 altitude), the mean absolute bias of the vertical measurement was 27m on average across units and
285 locations. The root mean squared error ranged from 14m to 42m depending on the unit, with a small
286 effect of location. However, the within-session standard deviation ranged only to 28m, suggesting
287 that a bias in the sea level reference point (probably incorrectly assumed to follow the French norm)
288 inflated the RMSE. The average bias ranged between -17m and +12m depending on the unit, after
289 correcting for significant location effect, but without effect of altitude. This means that different
290 brands of GPS unit yield different rate of error in their altitude measurements, which can impair the
291 comparison of datasets collected by different units. Further investigation or communication with
292 manufacturers should decipher whether this stems from different fix acquisition procedures (e.g.,
293 satellite detection) or different post-processing algorithms, and should also make clear which sea
294 level reference point different manufacturers are using.

295 These controlled field trials, along with other similar reports (Bouten et al., 2013; Ross-Smith et al.,
296 2016), highlight that even in benign conditions, GPS and altimeter data are sufficiently error-prone to
297 tamper with ecological inference in many cases (range of the standard deviation of the error: 4 –
298 50m). The issue is only suspected to be more acute in operational conditions when the DOP is larger,
299 the terrain rougher, the weather more variable, and there are more obstacles to signal diffusion than
300 in controlled field trials. Furthermore, the rate of error depended on the brand of the unit and on the
301 location, which can be of importance when comparing across studies.

302 **Part 2.2: Horizontal errors can cause vertical errors**

303 In the synthetic landscape simulations, the frequency of negative flight height records increased with
304 the standard deviation of both the horizontal and vertical telemetry error (Fig. A2a), and with the
305 landscape roughness and complexity (Fig. A2b). However, the various sources of errors acted in a
306 multiplicative way, so that even when the telemetry noise was small (SD of 1m), the error in h could
307 be large (SD of 20m; Fig. A2c; darkest grey curve). Perhaps unexpectedly, when the horizontal error
308 was large, the error in the height above ground tended to be independent of the vertical error in the
309 GPS (on average across all simulations; Fig. A2c; lightest grey curve). This means that the effect of the
310 horizontal error in the GPS can supersede the effect of the vertical error, if the terrain is rough. Even
311 in the absence of any vertical error, the horizontal error was indeed routinely sufficient to cause 10-
312 20% of the data points to be below ground (Fig. A2a).

313 **Part. 2.3: Errors inflate the recorded variance in flight height**

314 In the simulations of flight tracks, errors in h inflated the variance in the distribution of recorded
315 flight heights, i.e., the variance in the true flight height was consistently lower than the variance in

316 the recorded flight height (Fig. 2). In the raptor case studies, we obtained the same result, with the
317 caveat that we did not access to the true flight height, but we could instead use the corrected flight
318 heights (Fig. 2).

319 Indeed, if the movement and error processes are independent, the total variance in flight height is
320 theoretically exactly the sum of the movement and sampling variances (e.g, Auger-Méthé et al.,
321 2016; see also Gould & Nichols, 1998 and references therein). If the movement and error processes
322 are not independent, the total variance is still larger than the movement variance. Yet, what we need
323 for biological inference is the movement variance. In a naïve analysis of the raptor case studies that
324 would confound telemetry errors with rapid movements, the birds would therefore have appeared
325 more vertically mobile and with a more spread-out distribution in the aerosphere than they actually
326 were. This type of issue is potentially quite widespread in movement ecology, e.g., in behavioural
327 assignment exercises that use movement variances (daily displacements, turning angles, etc.) to
328 determine the behavioural state of animals.

329

330 **Part 2.4: Negative flight height records provide useful information**

331 In this section we focus on negative records, i.e., unrealistically low records, but the same logic can
332 be applied to unrealistically high records. Negative flight height records are more likely to occur
333 when animals are near the ground, either perched or flying. If we remove the negative records
334 (Poessel et al. 2018), perching and low flight are under-sampled in the final dataset (Roeleke et al.
335 2018). To illustrate this point, we used a flight track from a migrating juvenile osprey (*Pandion*
336 *haliaetus*) as it crossed the sea between the Italian mainland and Corsica (Duriez et al. 2018). During
337 a portion of that sea crossing, its Ornitela GPS unit recorded flight heights that oscillated between -
338 2m and -7m below the sea level (Fig. A3, inset). The amplitude of the oscillation suggested that the
339 bird followed the swell of the waves. The complete sequence (Fig. A3) depicts a progressive loss of
340 altitude as the bird glided towards firm ground, and a period of active flapping flight (as per the
341 accelerometry record) very low above the waves once the bird had lost all of its accumulated
342 potential energy before reaching firm ground. These negative flight height records documented a
343 critical time period. First, the risk of having to make a sea landing were clearly much greater in the
344 few minutes when the osprey was flying low over the waves, compared to the rest of the sea
345 crossing when the bird was often soaring high (Duriez et al. 2018). In addition, when flying low, the
346 bird had no other choice than to flap and therefore expend energy; whereas when higher above the
347 sea, the bird had the option to soar and therefore spare energy. It is critical that negative flight
348 height records like these are maintained, even if, instead of a fully interpretable high-resolution

349 sequence like in this example, there are just a few isolated negative flight height records in a low-
350 resolution dataset.

351 In addition, if we only kept the records with positive flight height, we would obtain a biased sample
352 of the distribution of flight height. Both in simulations and in the raptor case studies, discarding
353 negative flight height records led to the overestimation of the mean flight height in the remaining
354 dataset, the underestimation of the variance in flight height, the introduction of a right skew in the
355 distribution of flight height, and the overestimation of the collision risk (Fig. 3). The latter result was
356 because negative records mostly occurred when the bird was flying below the collision zone, and
357 thus removing negative records led to under-sample safe periods of time. Note that this particular
358 result pertains to the wind turbine application case only; in other types of collision risk, e.g., buildings
359 and utility lines, the collision zone starts closer to the ground.

360 The simulations nicely complemented the raptor case studies by 1) eliminating any debate about
361 whether the corrected flight heights in the raptor case studies were trustworthy or not (in the
362 simulations, the true flight heights are exactly known) and 2) increasing the range of flight
363 behaviours, since the raptors tended to exhibit lower percentage of time near the ground (in part
364 because we purposely tried to exclude time spent perched) and different distributions of the
365 sampling error. The amount of bias appeared highly dependent on the underlying flight behaviour
366 and error distribution, and therefore not easy to predict and account for without appropriate error-
367 handling methodology.

368 Additionally, there are many other major consequences of discarding negative flight heights. One is
369 the disruption of the expected balance of positive and negative errors in the remaining data.
370 Negative flight height records only arise when the error is negative, and so removing them
371 introduces a bias towards positive errors, thereby disrupting the shape of the distribution of errors in
372 the remaining data. Yet, we need the full range of errors to fit the models in Part 3. Another,
373 unrelated consequence is the disruption of the sampling schedule of the remaining data. Many
374 movement analyses are critically sensitive to the sampling schedule, and therefore their outcome will
375 not be the same after removing the negative records. Lastly, and perhaps most importantly, negative
376 flight height records can help fit the models that separate the error and movement processes,
377 because they are unambiguously erroneous and can be informed as such in the model-fitting
378 procedure (cf. Part 3). Some authors have applied less stringent filters, such as removing only the
379 most negative flight height records and removing an equal amount of extremely positive flight height
380 records. While the effect on the remaining distribution, and on the balance of negative and positive
381 errors is supposedly weaker than if removing all of the negative records, we warn that the remaining

382 records are still affected by the same error process that generated the records that were deemed too
383 erroneous to keep, thus the issues in Part 2.3 still need to be addressed. In addition, these extremely
384 erroneous records are potentially the most informative regarding the shape of the error distribution
385 (cf. Part 3).

386

387 **Part 2.5: The mean flight height is not sufficient to describe the** 388 **distribution of flight heights**

389 Flight height datasets are often reduced to a single summary metric, the mean flight height and its
390 variation with environmental and individual covariates (Walter et al. 2012, Cleasby et al. 2015,
391 Poessel et al. 2018, Tikkanen et al. 2018, Balotari-Chiebao et al. 2018). This decision is mostly based
392 on the ease of implementing spreadsheets, linear models, moving averages, or spline models. In this
393 section we instead call for approaches that describe the full distribution of flight heights in the
394 aerosphere, not only the mean flight height. To justify this call, we again focus on collision risk
395 estimation. Indeed, if the variance in flight height is large enough, a proportion of time may be spent
396 in the collision zone even if the mean flight height is outside the collision zone. In simulations, the
397 proportion of time spent in the collision zone indeed depended on both the mean and the variance in
398 flight height (Fig. 4a-b). In the raptor datasets, the estimated probability of flying in the collision zone
399 did not decrease much for the individuals whose mean flight height was estimated above the
400 collision zone (Fig. 4c). Similarly, the individuals that had an estimated mean flight height well below
401 the collision zone were predicted to spend about 20% of their time in the collision zone (Fig. 4c). We
402 strongly recommend that collision risk forecasts should not be based on the fixed effects of linear
403 models, but instead on the full distribution of flight heights – a recommendation that will likely hold
404 for all studies into vertical airspace use.

405 **Part 3: Statistical solutions**

406 The state-space model framework (de Valpine & Hastings, 2002; Fig. 5) has a structure that is
407 naturally aligned with the challenges of sampling errors in vertical space-use data. A state-space
408 model is a stochastic model describing the changes over time in a state variable (here, the true flight
409 height), when that variable is imperfectly observed (here, the recorded flight height). There is a
410 “state process”, separated from an “observation process” (Fig. 5). State-space models are routinely
411 used to correct for positioning errors in satellite-tracking data (chap. 6 in Sanz Subirana et al., 2013),
412 including in wildlife studies (Patterson et al. 2008, Johnson et al. 2008, Albertsen et al. 2015, Brost et
413 al. 2015, Buderman et al. 2015, Fleming et al. 2017). Importantly, these applications are not to be
414 confused with another application of state-space models to movement data, when the focal state

415 variable is a “behavioural state” whose Markovian transitions drive changes in movement rates
416 (Gurarie et al. 2016, Pirootta et al. 2018, Murgatroyd et al. 2018). Instead, when the objective is to
417 correct for positioning errors, the state variable is the position itself.

418 In studies of flight height, the movement model can be set up such that the state variable always
419 stays above zero. Then, if the recorded flight height is -7m, the model “knows” that the error was at
420 least 7m (Ross-Smith et al., 2016; cf. Part 2.4). Actually, the presence of unambiguously erroneous
421 records makes flight height studies better-suited to apply state-space models than many studies into
422 horizontal space use by animals. Indeed, even when in theory the model is estimable, sometimes
423 only a subset of the parameters of a state-space model are separately estimable, a phenomenon
424 called “weak identifiability” that occurs when the sampling variance largely exceeds the process
425 variance. An example of weak identifiability is when the difference between two classes of
426 individuals are larger than the differences within the classes (Garrett and Zeger 2000). In addition,
427 there are large statistical correlations between variance parameters in a movement model (Fleming
428 et al. 2017), making it extra difficult to accurately separate movements and errors in sparse datasets.
429 In that context, unambiguously erroneous records, such as negative flight heights, represent an
430 additional source of information (Brost et al. 2015). They can help separate the process and sampling
431 variances (Péron et al. 2017) and solve issues of weak identifiability.

432 As a perspective, we stress that there are also ways to obtain unambiguously correct records. These
433 records could in theory perform a role similar to that of unambiguously erroneous records. For
434 example, sometimes the position of the animals can be confirmed, e.g., at a documented feeding
435 site, a nest, or by an incidental ground-based sighting. Those records can then be matched to the
436 GPS track, yielding an exact measure of the local error. Animal-borne devices may also include a
437 transponder designed to signal passage near strategically-placed emitters (e.g., Rebke, Coulson,
438 Becker, & Vaupel, 2010). This type of validation data is routinely used in other applications of the GPS
439 technology (Sanz Subirana et al. 2013). Lastly, the state-space framework is naturally conducive to
440 the joint analysis of multiple sources of error-prone data (e.g., Péron, Nicolai, & Koons, 2012). In
441 flight height studies, it is therefore possible to jointly analyse GPS and altimeter data, or multiple GPS
442 streams coming from the same animal. This double-data approach is expected to help with statistical
443 covariance issues, but cannot be expected to fully resolve all identifiability issues (Besbeas & Morgan,
444 2017), which only error-free validation data can do.

445 We should eventually stress that several wildlife GPS manufacturers already use a state-space model
446 as part of the onboard data pre-processing algorithm, i.e., the released data have already been
447 corrected by a proprietary state-space algorithm which may furthermore rely on proprietary

448 validation data (Ornitela staff, pers. comm.). From our experience, in wildlife applications, these pre-
449 processing algorithms are only applied during “bursts” of high-frequency data acquisition, not when
450 the users request a more traditional low-frequency data acquisition schedule. Importantly, the data
451 may not be pre-processed *across* bursts. The error from the first location of a burst is then carried
452 over the entire burst sequence. Flight height tracks affected by this issue would exhibit a staircase-
453 shaped profile. Overall, this type of data pre-processing trades a lower error variance against a larger
454 error autocorrelation. Additional state-space modelling of the released pre-processed data can deal
455 with this type of error autocorrelation, but the models need to be custom-made, i.e., are not
456 routinely implemented in software. Perhaps more worryingly, some commercially-available GPS units
457 apparently simply truncate the recorded height at zero above sea level (pers. obs.). We call for a
458 more open approach to these data manipulations, including making the raw, unprocessed GPS
459 records available, in addition to any pre-processed data, and with a formal description of the pre-
460 processing algorithm.

461 We also acknowledge that the fitting of state-space models to vertical space use data still requires
462 relatively rare statistical skills. Nevertheless, there are already several free, open-source computing
463 environments to fit state-space models to vertical (and horizontal) movement data, and thereby
464 estimate the most likely movement track as a by-product of the estimated parameters, similarly to
465 how the individual values would be computed in a generalized mixed model with individual random
466 effects:

467 - The *crawl* (Johnson et al. 2008) and *ctmm* (Calabrese et al. 2016) packages for R. These compute
468 the likelihood of the state-space model using a Kalman filter. This algorithm is fast but requires all the
469 model processes to be Gaussian or approximately Gaussian (no truncation or constraint, no excess
470 extreme values, no excess kurtosis or skew).

471 - The *TMB* package for R (Kristensen et al. 2014) approximates the likelihood of the state-space
472 model using the automatic differentiation algorithm with Laplace approximation. That approach
473 makes computing times shorter than the next option, while still allowing for flexible modelling such
474 as non-Gaussian errors (Albertsen et al. 2015), custom link functions (Péron et al. 2017), or multiple
475 data streams.

476 - The Monte Carlo Markov Chain Bayesian framework (Plummer 2003, Spiegelhalter et al. 2003,
477 Csilléry et al. 2010) generates parameter distributions that iteratively converge towards the solution.
478 This option is the most flexible in terms of nonlinearities and non-Gaussian features, such as
479 truncated distributions (Brost et al. 2015), but the computing time can be prohibitive large for
480 datasets.

481

482 **Conclusion**

483 Improper error-handling methodologies yield a flawed picture of aerial niches. For example,
484 discarding negative flight height records artificially truncates the observed distribution of flight
485 heights (Fig. 3), and focusing on the mean flight height alone (for example when using linear models)
486 does not fully describe the aerial niche (Fig. 4). While these observations are quite intuitive, bad
487 practices remain common enough that it was important to stress these issues and illustrate them
488 thoroughly. On the other hand, not addressing the occurrence of errors at all would artificially
489 spread-out the observed distribution of flight heights (Fig. 2), leading for example to increased
490 observed vertical overlap between species and individuals, which can modify the inference about
491 community processes. Improper error handling procedures would also tamper with the
492 quantification of behaviour and flight strategies, by increasing or decreasing the observed vertical
493 velocity, and interfere with behavioural state assignments. Lastly, errors may covary with
494 environmental covariates such as terrain roughness and wind speed, e.g., GPS positioning precision
495 decreases with terrain roughness (D'Eon et al. 2002) and wind speed decreases near the ground
496 (Sachs 2005). Thereby, selectively discarding records based on the number of available satellites or
497 the dilution of precision would lead to imbalanced sampling of terrain roughness, and discarding
498 negative flight height records (that predominantly occur near the ground) would lead to
499 misrepresent the relationship to wind speed.

500 Regarding applied consequences, we focused on demonstrating how improper methods would
501 imperfectly quantify the time spent by GPS-tracked raptors in the rotor-swept zone of wind turbines
502 (Fig. 3b). There are many other human-wildlife conflicts for the use of the aerosphere, for example
503 bird strikes near airports and disturbance of wildlife by drones and other recreational aircraft.
504 Regarding bird strikes, GPS-based predictive models of bird flight height (e.g., Péron et al. 2017)
505 might help plan ahead the operation of airports. The state-space class of model that we advocate is
506 actually already used, in real time, to exploit bird activity data from radar monitors and generate a
507 warning system for airport managers (Bruder 1997). Regarding recreational aircraft and drones,
508 analysing bird-borne GPS tracks may help reveal the effect of the disturbance, which is expected to
509 increase in frequency as drones in particular become more popular (Rebolo-Ifrán et al. 2019). The
510 recommendations we made about the effect of errors on the estimation of aerial niche overlaps and
511 the quantification of behaviours seem particularly relevant in this context.

512 In conclusion, the issue of properly handling errors in flight height data is key to any aeroecology
513 study. We strongly advise against ad-hoc “data quality” filters, and against statistical tools that only

514 document variation in the mean flight height instead of the full distribution of flight height. Our
515 proposed statistical framework based on state-space models and the analysis of the full distribution
516 of flight heights requires interdisciplinary work between experts in flight behaviour and experts in
517 data analysis, and the emergence of interface specialists, but the insights and the applied decisions
518 based on those insights are expected to be more reliable.

519

520 **List of abbreviation:** *h*: flight height above ground; *z*: flight altitude (relative to the same reference as
521 the DEM, e.g., the ellipsoid); DEM: digital elevation model; UERE: user equivalent range error; DOP:
522 dilution of precision; SD: standard deviation.

523

524 **Declarations**

525 * Ethics approval and consent to participate

526 Permissions to trap and tag griffon vultures and lesser kestrels were given by the Centre de
527 Recherche sur la Biologie des Populations d'Oiseaux, Museum National d'Histoire Naturelle, Paris to
528 O. Duriez (Personal Program PP961) and to P. Pilard (PP311). Permissions to trap and tag osprey were
529 given by Instituto Nazionale per la Protezione e la Ricerca Ambientale (ISPRA) under the
530 authorization of Law 157/1992 [Art. 4 (1) and Art. 7 (5)] and under authorization no. 2502
531 05.05.2016 and no. 4254 27.03.2018 issued by Regione Toscana. Permissions to trap and tag condors
532 were given by the Dirección de Fauna de Río Negro (DFRN), the Administración de Parques
533 Nacionales (APN) from Argentina and the owners and managers of local farms. All condor procedures
534 were approved by the DFRN, RN132.730-DF, and APN, 14/2011.

535 * Consent for publication

536 Not applicable

537 * Availability of data and material

538 The data have been uploaded to MoveBank (www.movebank.org)

539 * Competing interests

540 None

541 * Funding

542 SAL thanks PICT-BID 0725/2014 for financial support. JMC and CHF were supported by NSF ABI
543 1458748. The griffon vulture project was supported by ANR-07-BLAN-0201. RGJ was supported by
544 pre-doctoral grant (FPI/BES-2016-077510).

545 * Authors' contributions

546 All authors conceived the project and revised the manuscript. GP performed the analyses and
547 simulations and wrote the first draft of the manuscript. OD, RG, SAL, KS, and ES procured the data.

548 * Acknowledgements

549 A. Scharf and H. Wehner (MPI) flew the drone and assisted with data collection. M. Quetting flew the
550 ultra-light aircraft and E. Lempidakis assisted with data collection. We warmly thank everyone
551 involved in the collection and management of the raptor tracking data, and in particular C. Itty for
552 the eagle data, and the organizations *Asters* and *Vautours des Baronnie*s for the help with field
553 experiments.

554 **Supplementary material**

555 Appendix A: Supplementary figures A1, A2, A3.

556 **Literature cited**

- 557 Albertsen, C. M., K. Whoriskey, D. Yurkowski, A. Nielsen, and J. M. Flemming. 2015. Fast fitting of
558 non-Gaussian state-space models to animal movement data via Template Model Builder.
559 Ecology 96:2598–2604.
- 560 Arlettaz, R., N. Perrin, and J. Hausser. 1997. Trophic Resource Partitioning and Competition between
561 the Two Sibling Bat Species *Myotis myotis* and *Myotis blythii*. The Journal of Animal Ecology
562 66:897.
- 563 Auger-Méthé, M., C. Field, C. M. Albertsen, A. E. Derocher, M. A. Lewis, I. D. Jonsen, and J. Mills
564 Flemming. 2016. State-space models' dirty little secrets: even simple linear Gaussian models
565 can have estimation problems. Scientific reports 6:26677.
- 566 Balotari-Chiebao, F., J. E. Brommer, P. Saurola, A. Ijäs, and T. Laaksonen. 2018. Assessing space use by
567 pre-breeding white-tailed eagles in the context of wind-energy development in Finland.
568 Landscape and Urban Planning 177:251–258.
- 569 Besbeas, P., and B. J. T. Morgan. 2017. Variance estimation for integrated population models. AStA
570 Advances in Statistical Analysis 101:439–460.
- 571 Bishop, C. M., R. J. Spivey, L. A. Hawkes, N. Batbayar, B. Chua, P. B. Frappell, W. K. Milsom, T.
572 Natsagdorj, S. H. Newman, G. R. Scott, J. Y. Takekawa, M. Wikelski, and P. J. Butler. 2015. The
573 roller coaster flight strategy of bar-headed geese conserves energy during Himalayan
574 migrations. Science 347:250–254.
- 575 Borkenhagen, K., A.-M. Corman, and S. Garthe. 2018. Estimating flight heights of seabirds using
576 optical rangefinders and GPS data loggers: a methodological comparison. Marine Biology
577 165:17.

- 578 Bouten, W., E. W. Baaij, J. Shamoun-Baranes, and K. C. J. Camphuysen. 2013. A flexible GPS tracking
579 system for studying bird behaviour at multiple scales. *Journal of Ornithology* 154:571–580.
- 580 Bradshaw, C. J. A., D. W. Sims, and G. C. Hays. 2007. Measurement error causes scale-dependent
581 threshold erosion of biological signals in animal movement data. *Ecological Applications*
582 17:628–638.
- 583 Brost, B. M., M. B. Hooten, E. M. Hanks, and R. J. Small. 2015. Animal movement constraints improve
584 resource selection inference in the presence of telemetry error. *Ecology* 96:2590–2597.
- 585 Bruder, J. A. 1997. Bird hazard detection with airport surveillance radar. Pages 160–163 *Radar*
586 *Systems (RADAR 97)*. IEE.
- 587 Buderman, F. E., M. B. Hooten, J. S. Ivan, and T. M. Shenk. 2015. A functional model for characterizing
588 long-distance movement behaviour. *Methods in Ecology and Evolution* 7:264–273.
- 589 Calabrese, J. M., C. H. Fleming, and E. Gurarie. 2016. ctm: An r package for analyzing animal
590 relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution*
591 7:1124–1132.
- 592 Cleasby, I. R., E. D. Wakefield, S. Bearhop, T. W. Bodey, S. C. Votier, and K. C. Hamer. 2015. Three-
593 dimensional tracking of a wide-ranging marine predator: flight heights and vulnerability to
594 offshore wind farms. *Journal of Applied Ecology* 52:1474–1482.
- 595 Crocker, C., and G. Jackson. 2018. The use of altimeters in height measurement. [http://www.hills-](http://www.hills-database.co.uk/altim.html)
596 [database.co.uk/altim.html](http://www.hills-database.co.uk/altim.html).
- 597 Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian Computation
598 (ABC) in practice. *Trends in Ecology and Evolution* 25:410–418.
- 599 D'Eon, R. G., R. Serrouya, G. Smith, and C. O. Kochanny. 2002. GPS Radiotelemetry Error and Bias in
600 Mountainous Terrain. *Wildlife Society Bulletin* 30:430–439.
- 601 Davy, C. M., A. T. Ford, and K. C. Fraser. 2017. Aeroconservation for the Fragmented Skies.
602 *Conservation Letters* 10:773–780.
- 603 Dokter, A. M., S. Åkesson, H. Beekhuis, W. Bouten, L. Buurma, H. van Gasteren, and I. Holleman.
604 2013. Twilight ascents by common swifts, *Apus apus*, at dawn and dusk: acquisition of
605 orientation cues? *Animal Behaviour* 85:545–552.
- 606 Dunn, J. E., and P. S. Gipson. 1977. Analysis of Radio Telemetry Data in Studies of Home Range.
607 *Biometrics* 33:85–101.

- 608 Duriez, O., G. Péron, D. Gremillet, A. Sforzi, and F. Monti. 2018. Migrating ospreys use thermal uplift
609 over the open sea. *Biology Letters* 14:20180687.
- 610 Fleming, C. H., W. F. Fagan, T. Mueller, K. A. Olson, P. Leimgruber, and J. M. Calabrese. 2016.
611 Estimating where and how animals travel: An optimal framework for path reconstruction from
612 autocorrelated tracking data. *Ecology* 97:576–582.
- 613 Fleming, C. H., D. Sheldon, E. Gurarie, W. F. Fagan, S. LaPoint, and J. M. Calabrese. 2017. Kálmán
614 filters for continuous-time movement models. *Ecological Informatics* 40:8–21.
- 615 Fowler, C. M. . 2005. *The solid Earth: An introduction to global geophysics*. Cambridge University
616 Press, Cambridge.
- 617 Frair, J. L., S. E. Nielsen, E. H. Merrill, S. R. Lele, M. S. Boyce, R. H. M. Munro, G. B. Stenhouse, and H.
618 L. Beyer. 2004. Removing GPS collar bias in habitat selection studies. *Journal of Applied Ecology*
619 41:201–212.
- 620 Freitas, C., C. Lydersen, M. A. Fedak, and K. M. Kovacs. 2008. A simple new algorithm to filter marine
621 mammal Argos locations. *Marine Mammal Science* 24:315–325.
- 622 Garrett, E. S., and S. L. Zeger. 2000. Latent Class Model Diagnosis. *Biometrics* 56:1055–1067.
- 623 Gorokhovich, Y., and A. Voustianiouk. 2006. Accuracy assessment of the processed SRTM-based
624 elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain
625 characteristics. *Remote Sensing of Environment* 104:409–415.
- 626 Gould, W. R., and J. D. Nichols. 1998. Estimation of Temporal Variability of Survival in Animal
627 Populations. *Ecology* 79:2531–2538.
- 628 Gurarie, E., C. Bracis, M. Delgado, T. D. Meckley, I. Kojola, and C. M. Wagner. 2016. What is the
629 animal doing? Tools for exploring behavioral structure in animal movements. *Journal of Animal*
630 *Ecology* 85:69–84.
- 631 Hays, G. C., S. Åkesson, B. J. Godley, P. Luschi, and P. Santidrian. 2001. The implications of location
632 accuracy for the interpretation of satellite-tracking data. *Animal Behaviour* 61:1035–1040.
- 633 Januchowski, S. R., R. L. Pressey, J. VanDerWal, and A. Edwards. 2010. Characterizing errors in digital
634 elevation models and estimating the financial costs of accuracy. *International Journal of*
635 *Geographical Information Science* 24:1327–1347.
- 636 Jerde, C. L., and D. R. Visscher. 2005. GPS measurement error influences on movement model
637 parameterization. *Ecological Applications* 15:806–810.

- 638 Johnson, D. S., J. M. London, M.-A. M. A. Lea, and J. W. Durban. 2008. Continuous-time correlated
639 random walk model for animal telemetry data. *Ecology* 89:1208–1215.
- 640 Katzner, T. E., D. Brandes, T. Miller, M. Lanzone, C. Maisonneuve, J. A. Tremblay, R. Mulvihill, and G.
641 T. Merovich. 2012. Topography drives migratory flight altitude of golden eagles: Implications for
642 on-shore wind energy development. *Journal of Applied Ecology* 49:1178–1186.
- 643 Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. Bell. 2014. TMB: Automatic Differentiation and
644 Laplace Approximation. *Journal of Statistical Software* 2014:1–21.
- 645 Krone, O., and G. Treu. 2018. Movement patterns of white-tailed sea eagles near wind turbines.
646 *Journal of Wildlife Management* 82:1367–1375.
- 647 Kunz, T. H., S. A. Gauthreaux, N. I. Hristov, J. W. Horn, G. Jones, E. K. V. Kalko, R. P. Larkin, G. F.
648 McCracken, S. M. Swartz, R. B. Srygley, R. Dudley, J. K. Westbrook, and M. Wikelski. 2007.
649 Aeroecology: probing and modeling the aerosphere. *Integrative and Comparative Biology* 48:1–
650 11.
- 651 Lambertucci, S. A., E. L. C. Shepard, and R. P. Wilson. 2015. Human-wildlife conflicts in a crowded
652 airspace. *Science* 348:502–504.
- 653 Liechti, F., S. Bauer, K. L. Dhanjal-Adams, T. Emmenegger, P. Zehtindjiev, and S. Hahn. 2018.
654 Miniaturized multi-sensor loggers provide new insight into year-round flight behaviour of small
655 trans-Saharan avian migrants. *Movement Ecology* 6:19.
- 656 López-López, P. 2016. Individual-Based Tracking Systems in Ornithology: Welcome to the Era of Big
657 Data. *Ardeola* 63:103–136.
- 658 Monaldo, F. M., J. Goldhirsh, and E. J. Walsh. 1986. Altimeter height measurement error introduced
659 by the presence of variable cloud and rain attenuation. *Journal of Geophysical Research*
660 91:2345.
- 661 Murgatroyd, M., T. Photopoulou, L. G. Underhill, W. Bouten, and A. Amar. 2018. Where eagles soar:
662 Fine-resolution tracking reveals the spatiotemporal use of differential soaring modes in a large
663 raptor. *Ecology and Evolution* 8:6788–6799.
- 664 Nathan, R., W. M. W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. P. E. Smouse.
665 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings*
666 *of the National Academy of Sciences of the United States of America* 105:19052–9.
- 667 Parkinson, B. W., and J. J. J. Spilker, editors. 1996. *Global Positioning System: Theory and*

- 668 Applications, vol. 1. American Institute of Aeronautics and Astronautics, Washington, D.C.
- 669 Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space
670 models of individual animal movement. *Trends in Ecology and Evolution* 23:87–94.
- 671 Péron, G., C. H. Fleming, O. Duriez, J. Fluhr, C. Itty, S. Lambertucci, K. Safi, E. L. C. Shepard, and J. M.
672 Calabrese. 2017. The energy landscape predicts flight height and wind turbine collision hazard
673 in three species of large soaring raptor. *Journal of Applied Ecology* 54:1895–1906.
- 674 Péron, G., C. A. Nicolai, and D. N. Koons. 2012. Demographic response to perturbations: The role of
675 compensatory density dependence in a North American duck under variable harvest regulations
676 and changing habitat. *Journal of Animal Ecology* 81:960–969.
- 677 Pirotta, E., E. W. J. Edwards, L. New, and P. M. Thompson. 2018. Central place foragers and moving
678 stimuli: A hidden-state model to discriminate the processes affecting movement. *Journal of*
679 *Animal Ecology*:0–3.
- 680 Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.
681 *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*:20–22.
- 682 Poessel, S. A., A. E. Duerr, J. C. Hall, M. A. Braham, and T. E. Katzner. 2018. Improving estimation of
683 flight altitude in wildlife telemetry studies. *Journal of Applied Ecology* 55:2064–2070.
- 684 Rebke, M., T. Coulson, P. H. Becker, and J. W. Vaupel. 2010. Reproductive improvement and
685 senescence in a long-lived bird. *Proceedings of the National Academy of Sciences of the United*
686 *States of America* 107:7841–6.
- 687 Rebolo-Ifrán, N., M. Graña Grilli, and S. A. Lambertucci. 2019. Drones as a Threat to Wildlife: YouTube
688 Complements Science in Providing Evidence about Their Effect. *Environmental Conservation*:1–
689 6.
- 690 Roeleke, M., T. Teige, U. Hoffmeister, F. Klingler, and C. C. Voigt. 2018. Aerial-hawking bats adjust
691 their use of space to the lunar cycle. *Movement Ecology* 6:11.
- 692 Ross-Smith, V. H., C. B. Thaxter, E. A. Masden, J. Shamoun-Baranes, N. H. K. Burton, L. J. Wright, M.
693 M. Rehfish, and A. Johnston. 2016. Modelling flight heights of lesser black-backed gulls and
694 great skuas from GPS: a Bayesian approach. *Journal of Applied Ecology* 53:1676–1685.
- 695 Sachs, G. 2005, December 16. Minimum shear wind strength required for dynamic soaring of
696 albatrosses. John Wiley & Sons, Ltd (10.1111).
- 697 Sanz Subirana, J., J. M. Juan Zornoza, and M. Hernandez-Pajares. 2013. GNSS Data Processing Volume

- 698 I: Fundamentals and Algorithms. ESA Communications, Noordwijk, NL.
- 699 Shepard, E. L. C., C. Williamson, and S. P. Windsor. 2016. Fine-scale flight strategies of gulls in urban
700 airflows indicate risk and reward in city living. *Philosophical transactions of the Royal Society of*
701 *London B, Biological sciences* 371:76–93.
- 702 Siemers, B. M., and H.-U. Schnitzler. 2004. Echolocation signals reflect niche differentiation in five
703 sympatric congeneric bat species. *Nature* 429:657–661.
- 704 Spiegelhalter, D., A. Thomas, N. Best, and R. Way. 2003. WinBUGS User Manual. *Components* 2:1–60.
- 705 Tikkanen, H., S. Rytönen, O.-P. Karlin, T. Ollila, V.-M. Pakanen, H. Tuohimaa, and M. Orell. 2018.
706 Modelling golden eagle habitat selection and flight activity in their home ranges for safer wind
707 farm planning. *Environmental Impact Assessment Review* 71:120–131.
- 708 de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and
709 observation error. *Ecological Monographs* 72:57–76.
- 710 Walter, W. D., J. W. Fischer, J. S. Humphrey, T. S. Daugherty, M. P. Milleson, E. A. Tillman, and M. L.
711 Avery. 2012. Using three-dimensional flight patterns at airfields to identify hotspots for avian–
712 aircraft collisions. *Applied Geography* 35:53–59.
- 713 Weimerskirch, H., C. Bishop, T. Jeanniard-du-Dot, A. Prudor, and G. Sachs. 2016. Frigate birds track
714 atmospheric conditions over months-long transoceanic flights. *Science* 353:74–78.
- 715 Weimerskirch, H., M. Le Corre, Y. Ropert-Coudert, A. Kato, and F. Marsac. 2005. The three-
716 dimensional flight of red-footed boobies: adaptations to foraging in a tropical environment?
717 *Proceedings. Biological sciences* 272:53–61.
- 718 Zaugg, S., G. Saporta, E. van Loon, H. Schmaljohann, and F. Liechti. 2008. Automatic identification of
719 bird targets with radar via patterns produced by wing flapping. *Journal of The Royal Society*
720 *Interface* 5:1041–1053.
- 721

722 Figure legends

723 Fig. 1: Illustration of the difference between *true* and *recorded* flight height. A: True flight height
724 above ground (h_{true}), and true elevation above ellipsoid (z_{true}). B: Adding the five sources of error, with
725 circled numbers referring to headings in Part 1. DEM stands for Digital Elevation Model. C: Two tracks
726 with the same amount of error. The bird of track 1 is flying high so all the recorded flight height data
727 remain positive despite the errors. The bird of track 2 is flying low, so some of the recorded data fall
728 below the digital elevation model.

729

730 Fig. 2: Comparison between the standard deviation of the *recorded* flight height (y-axis) and of the
731 *corrected* flight height (x-axis), assumed to represent the *true* flight height, in three species of large
732 soaring raptors. Each point stands for one bird over its entire monitoring period. The state-space
733 model that we used to correct the flight heights, and in particular its robustness to variation in
734 sampling resolution across populations, is explained in Péron et al. (2017). The diagonal line shows
735 where the points should be if the recorded flight heights were error-free.

736

737 Fig. 3: Removing the negative *recorded* flight heights introduces biases in the distribution of the
738 remaining flight heights. Left group of panels: in simulations, where the *true* flight height is known.
739 Right group of panels: in the raptor case studies, where the *corrected* flight height is assumed to
740 represent the true flight height. In all panels, the x-axis features the variance in the true (or
741 corrected) flight height. The y-axis features the percentage bias in (a) mean flight height; (b) collision
742 risk (proportion of time spent between 60 and 180m above ground); (c) variance in flight height; and
743 (d) skewness of the distribution of flight height. A percentage bias of +10% means that the focal
744 quantity is 10% larger after we remove the negative records.

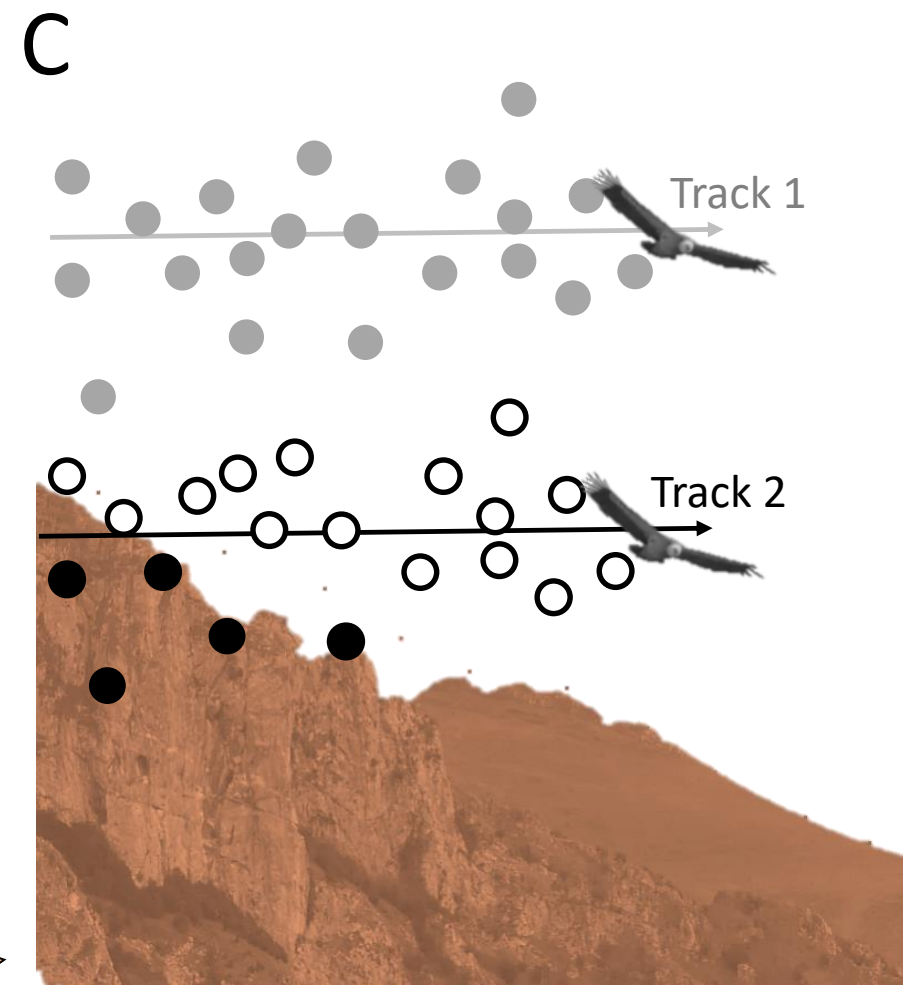
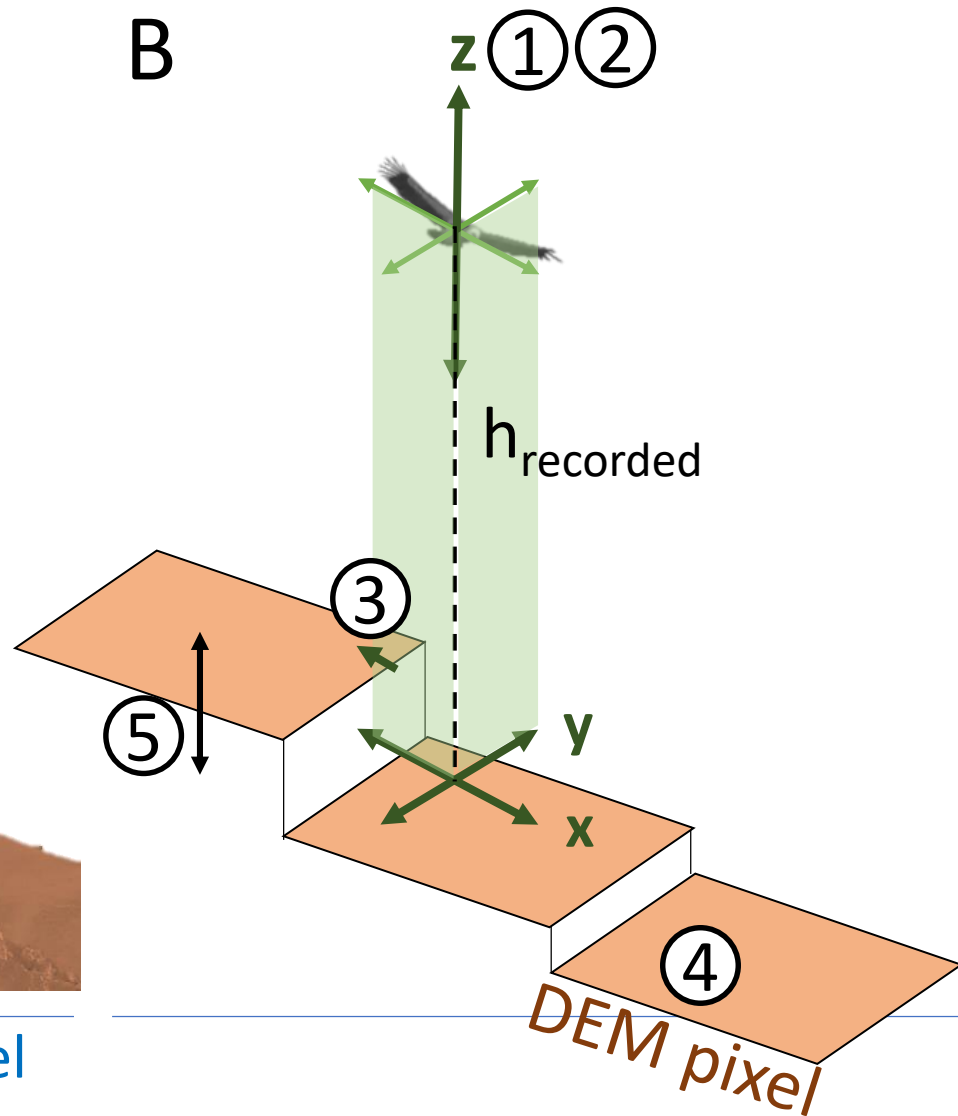
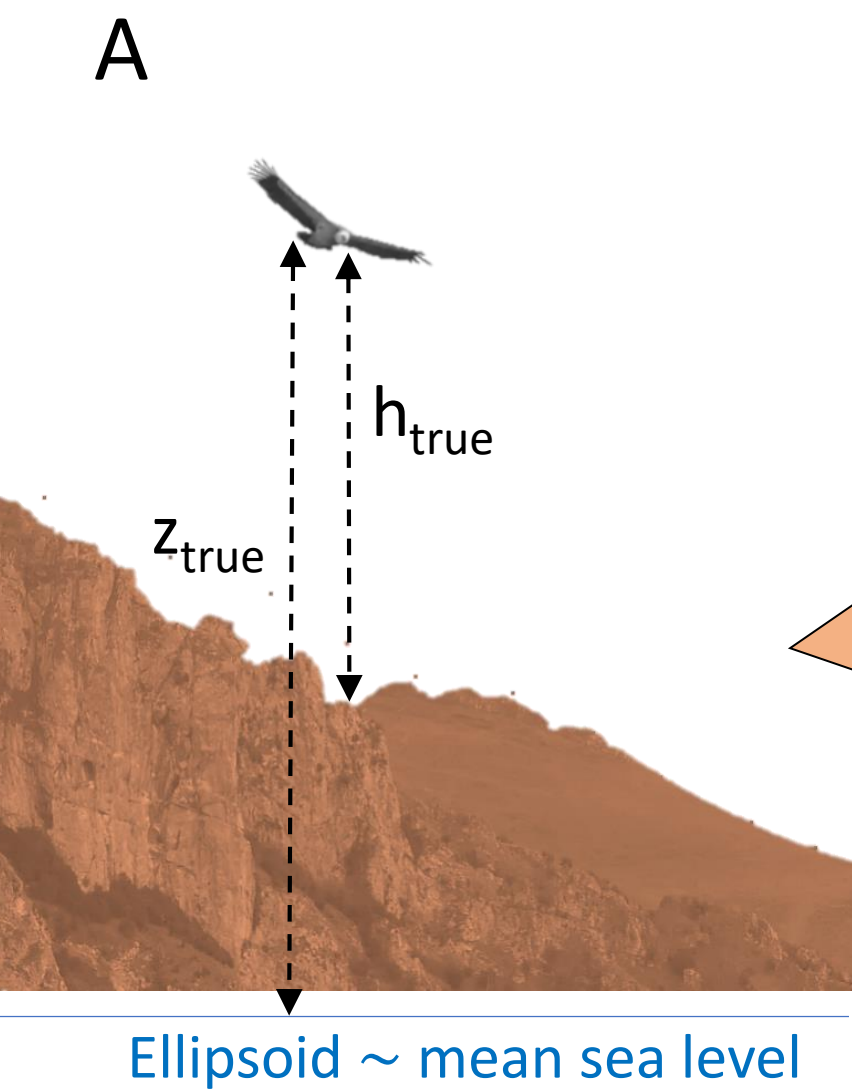
745

746 Fig. 4. The variance in flight height influences the percentage of time spent in the collision zone of a
747 wind farm (grey area, between 60 and 180 m). (a) Four simulated tracks (where the *true* flight height
748 is known) with the same mean flight height (200m) but different variances (10, 50, 100, and 250m²).
749 (b) More extensive simulations. Each point corresponds to one simulated track with a different mean
750 flight height. (c) Same as (b) but using real datasets collected from three raptor species, where the
751 *corrected* flight height is assumed to represent the true flight height. Each symbol stands for an
752 individual over its entire monitoring period.

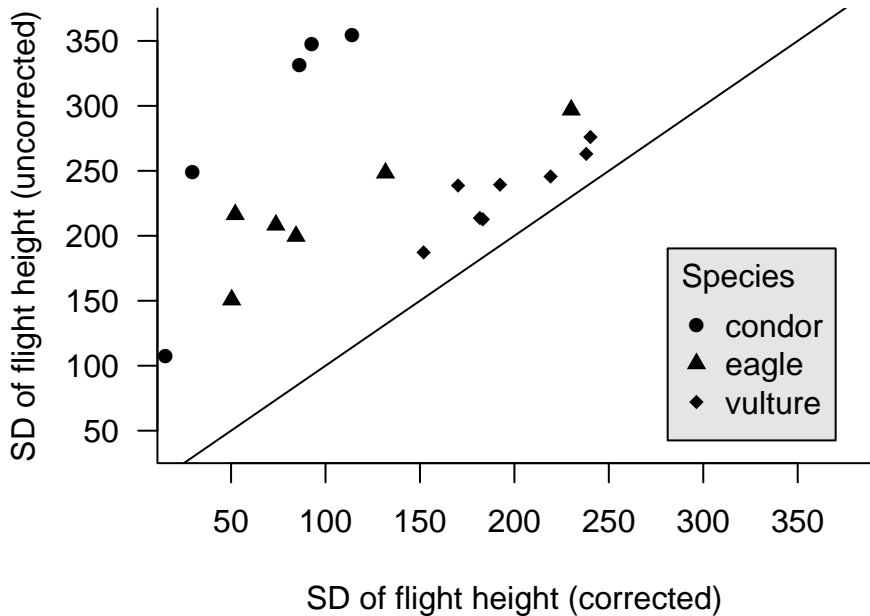
753

754 Fig. 5: Schematic overview of the principles of a state-space model as applied to the correction of
755 sampling errors in flight height data. The movement (or state) process accounts for the distribution
756 of *true* flight heights. The observation process introduces sampling errors of various origins (Part 1)
757 and yields the *recorded* flight heights. It also accounts for the sampling schedule. By fitting this model
758 to recorded flight height time series, we can retrospectively compute the *corrected* flight height, an
759 estimate of the true flight height.

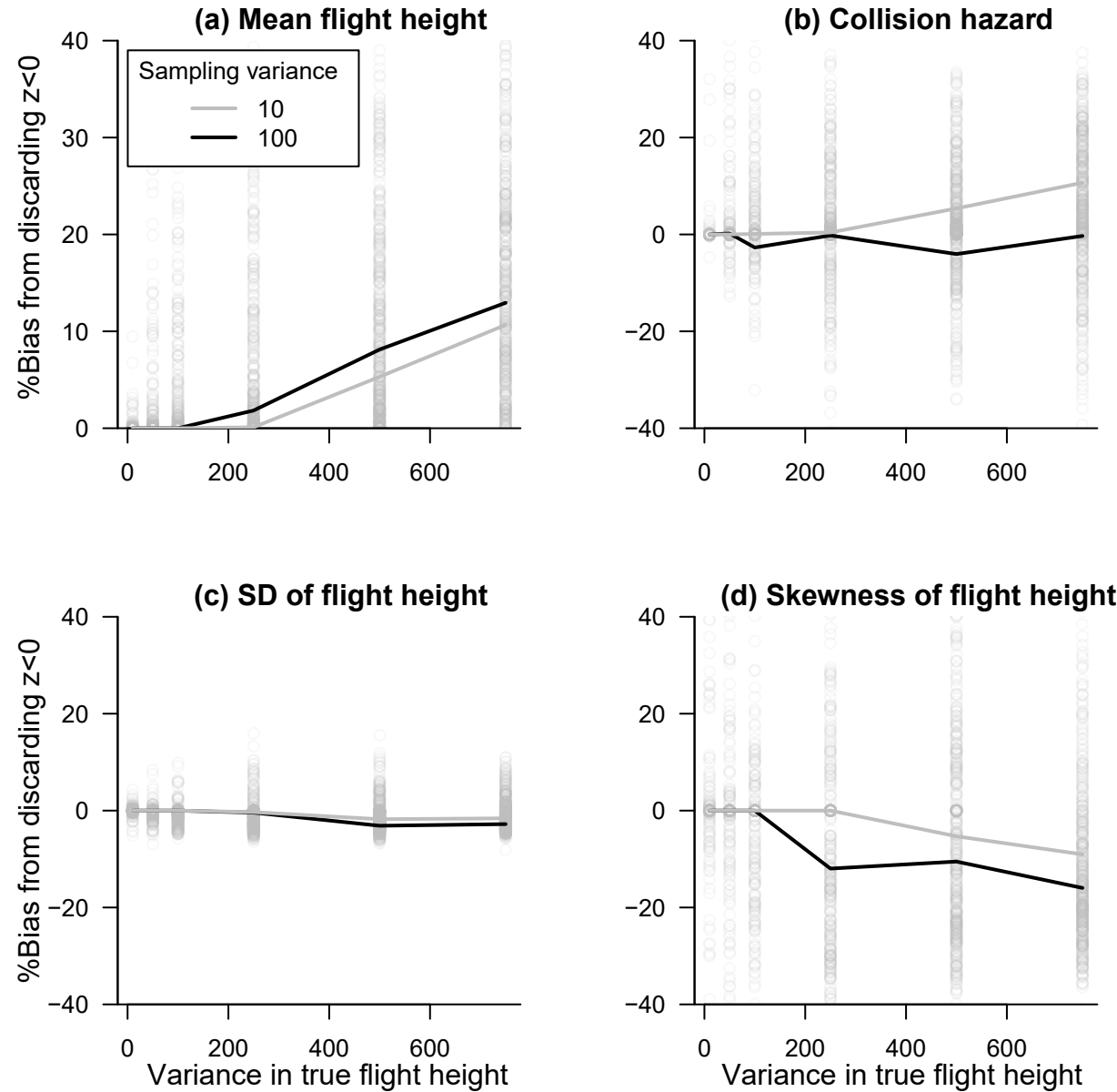
760



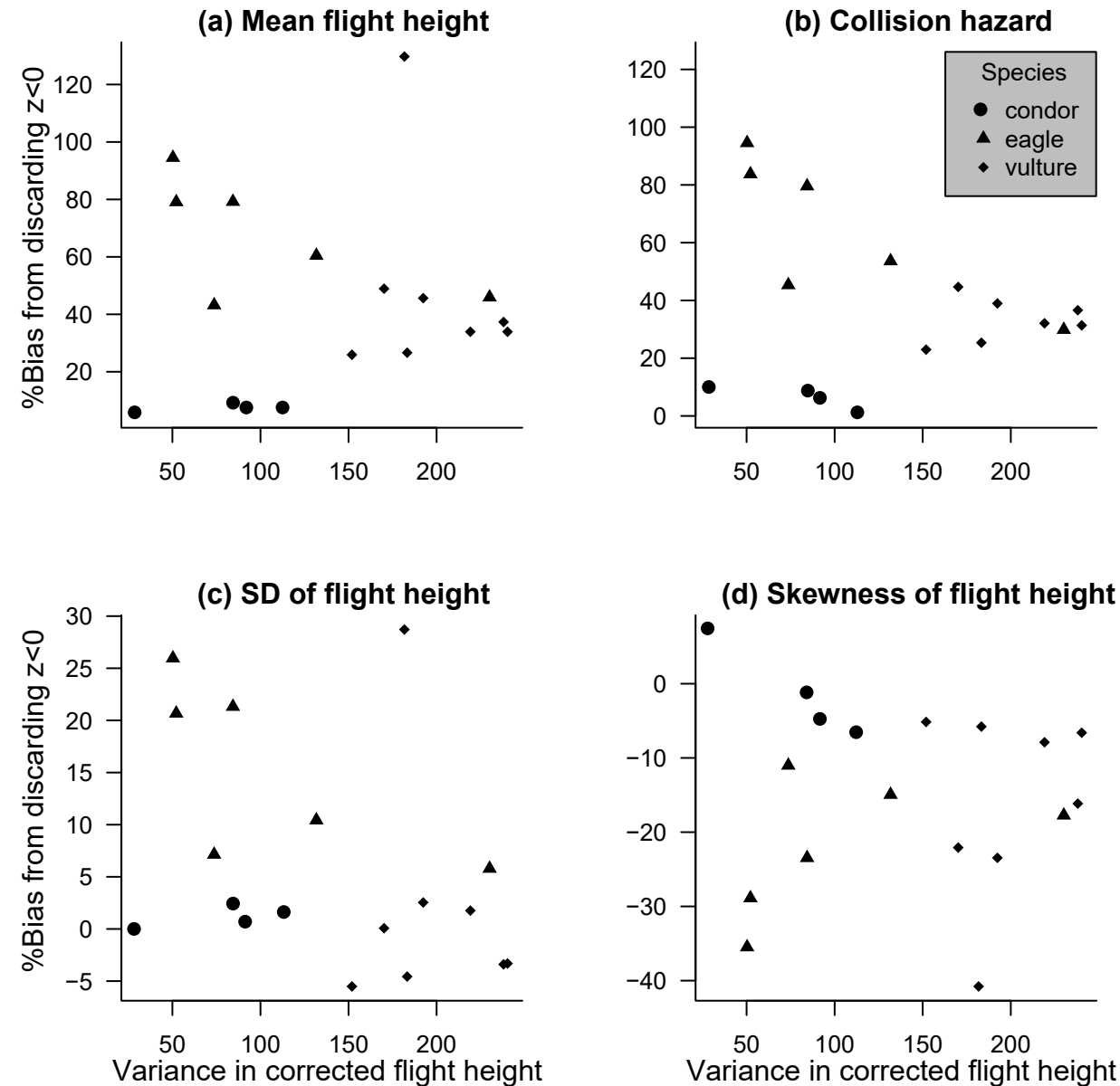
Bias in the observed variance of flight height in 3 raptor species



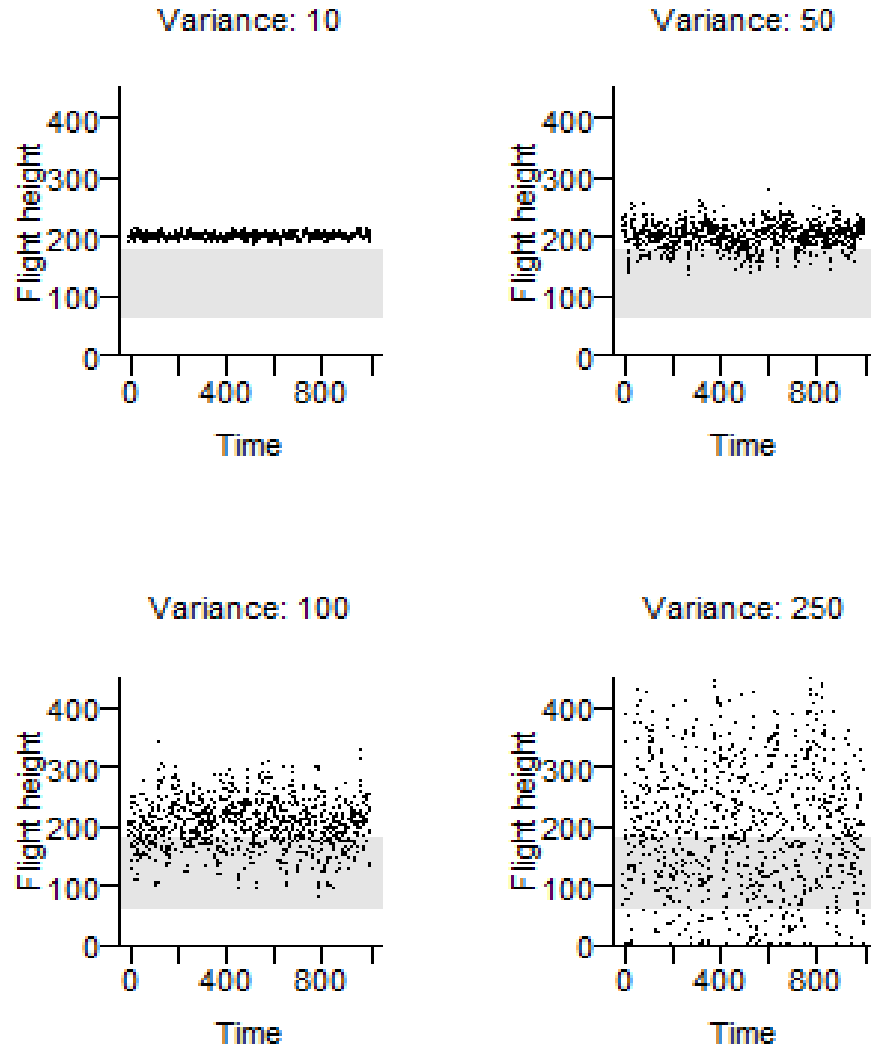
Simulations



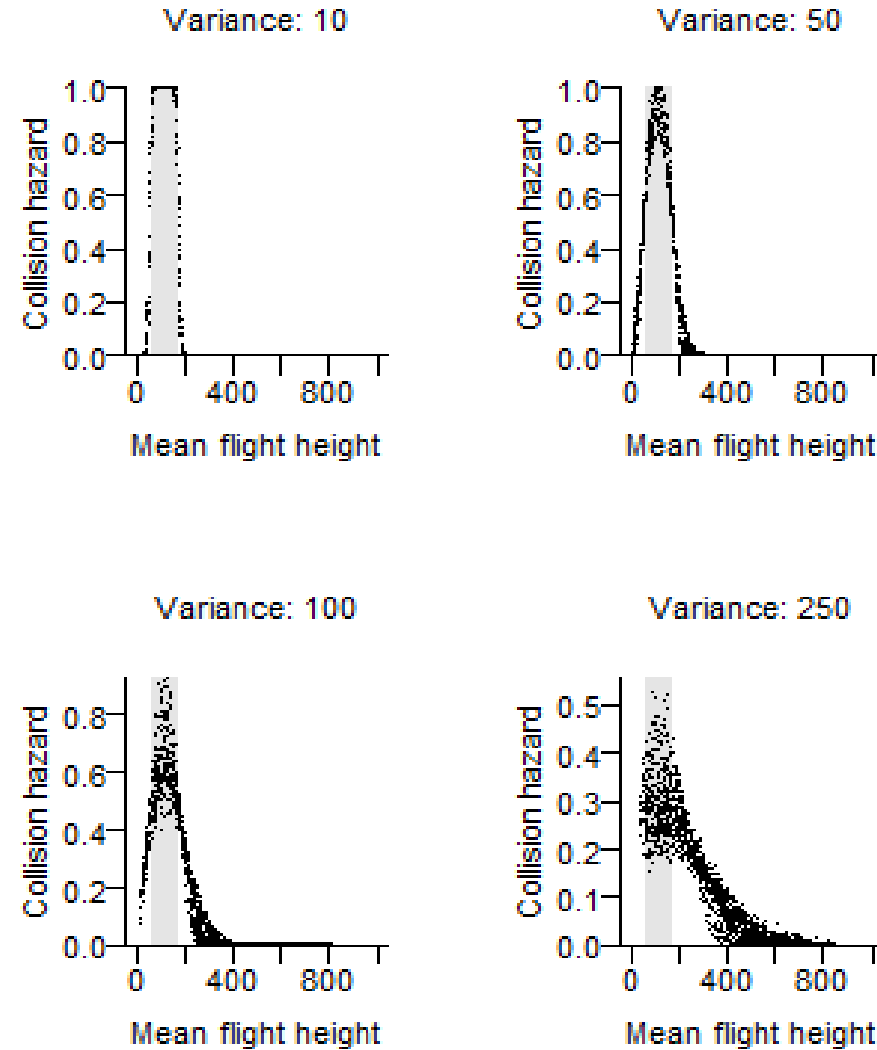
Raptor case studies



(a) A few simulated flight tracks with mean flight height 200m and different variances



(b) Relationship between mean flight height and collision hazard for different variances



(c) Collision hazard vs. mean flight height in 3 raptor species

