Smoothing of flight altitude data for input into Collision Risk Models

An overview of potential smoothing methods

H.M. Madden

WAARDENBURG Ecology

we consult nature.

Smoothing of flight altitude data for input into Collision Risk Models

An overview of potential smoothing methods

H.M. Madden

Status: Final report

Report nr:	25-378
Project nr:	23-0667
Date of publication:	11 December 2025
Photo credits cover page:	Great black-backed gull, Austevoll, Norway, Ruben Fijn/ Waardenburg Ecology
Project manager:	Dr A. Gyimesi
Second reader:	Dr. R.C. Fijn
Name & address client:	Rijkswaterstaat Water, Verkeer en Leefomgeving Lange Kleiweg 34, 2288 GK RIJSWIJK
Reference client:	Zaaknummer 31192825, kenmerk RWS-2024/2345
Approved for publication:	Dr. R.C. Fijn Head of the department of Bird Ecology

Please cite as: Madden, H.M. 2025. Smoothing of flight altitude data for input into Collision Risk Models. Report 25-378. Waardenburg Ecology, Culemborg.

Keywords: impact assessments, KEC, offshore, mortality estimates, seabird, Wozep

Waardenburg Ecology Varkensmarkt 9, 4101 CK Culemborg, 0345 512710
info@waardenburg.eco, www.waardenburg.eco

WAARDEN
BURG
Ecology

Preface

Offshore wind energy in the Netherlands contributes to the energy transition and realisation of the Dutch government’s climate ambitions. However, the construction and presence of offshore wind farms have ecological impacts on wildlife. The Offshore Wind Ecological Programme (Wozep) stipulates that offshore wind energy projects may not have major negative ecological effects on birds, bats or marine mammals.

Within the Framework for Assessing Ecological and Cumulative Effects (Dutch name: Kader Ecologie & Cumulatie; hereafter KEC), the cumulative effects of all existing and planned Dutch and foreign wind farms in the southern North Sea are predicted and evaluated. Given the potential negative effects of wind energy on the natural environment, the KEC includes several environmental impact assessments of projected future wind farms.

One focus of these assessments is potential bird mortality. Waardenburg Ecology has been tasked with estimating potential future bird collision-based mortality based on (scenarios of) planned wind farms in the North Sea. Since GPS altitude data are often prone to uncertainty, however, such projections may over- or underestimate the risk of mortality when included in collision risk models. In this study, we tested and evaluated three potential methods to ‘smooth’ altitude measurements.

The project was guided by Martine Graafland and Meik Verdonk (RWS).

Data statement

Data ownership: Ecowende CV. Data were used with their permission.

Summary

Various methods exist to ‘smooth’ altitude measurements derived by GPS-loggers on live birds, and thus reduce errors in the data. Smoothing is a data processing technique used to reduce random ‘noise’ or uncertainty in GPS data, helping make patterns and trends more visible. Instead of showing every fluctuation, which is likely misrepresentative of the true behaviour of the species, smoothing creates a cleaner representation of the underlying signal. This report provides an overview of three smoothing techniques, using spatial GPS-derived flight altitude data collected on live great black-backed gulls as an example.

By applying smoothing to GPS data the reliability of altitude measurements can be enhanced, thereby improving the accuracy of collision risk estimates with (offshore) wind farms for (sea)birds. Ultimately, refining these methods will contribute to more robust projections of potential bird mortality associated with offshore wind farm developments, informing both ecological impact assessments and future policy decisions.

Contents

1 Introduction

1.1 Background

Seabirds are vulnerable to collision with (and thus mortality from) offshore structures such as wind turbines (Furness et al. 2013). The Kader Ecologie & Cumulatie (KEC) programme assesses the cumulative ecological impact of offshore wind farms in the North Sea. A key component of KEC is collision risk modelling (CRM) for seabirds and migratory species. CRMs estimate the probability of birds colliding with turbine blades based on flight height distribution (not just mean altitude) and time spent within the rotor-swept zone. Collision risk estimates based on GPS-derived spatial data are often used by wind farm developers and policymakers to estimate mortality levels of new offshore wind farms (OWFs), yet little is known about whether and to what extent individuals change their behaviour to avoid colliding with turbines and/or blades.

The KEC modelling framework explicitly relies on species-specific flight height distributions for CRMs. These distributions are drawn from multiple sources, including Johnston et al. (2014) and telemetry-based studies. The knowledge base updates for KEC aims to incorporate new demographic and behavioural data and review the methodological choices for CRMs. Environmental Impact Assessments (EIA) for offshore wind farms must justify the basis for collision risk estimates, including flight height data sources. However, no single method is universally optimal; choices for the best method are often species- and site-specific, as well as how they deal with uncertainty.

Whilst various methods exist to measure flight altitude (barometric altimeter, radar, visual observations), GPS devices deployed on (sea)birds to monitor their aerial behaviour provide fine-scale, high-resolution data, as well as insight into the flight height distributions (FHD) of birds. However, vertical flight height measurements within raw spatial datasets often contain positioning inaccuracies of 10-50 m (Skone *et al.* 2001), which can distort CRMs. In reality, birds do not change altitude every few seconds; such fluctuations are artefacts rather than true behaviour of individuals. This can lead to ambiguity in flight height estimates and uncertainty when including such data in collision risk models.

1.2 Objective

Earlier KEC versions summarised flight heights using mean values and variation, however this approach is no longer sufficient since the CRM depends on the full distribution of flight heights, not just averages. For KEC 6.0 and beyond, the goal is to generate species-level flight height distributions that are less influenced by individual variation and more suitable for stochastic CRMs using, for example, stochLAB (Caneco *et al.* 2022).

Currently, CRMs incorporate flight height distributions based on the proportion of time an individual spends in each one-metre altitude bin, aggregated across all observations for that species. This binning approach allows the model to represent the full vertical profile of flight behaviour rather than relying on summary statistics. By using these proportions, CRMs can simulate collision risk more accurately under stochastic frameworks, as they capture the variability and probability of occurrence at different heights within the rotor-swept zone. However, these distributions typically do not account for different behavioural states or environmental covariates (e.g. foraging vs. commuting, over land or sea, wind speed, visibility), which can influence flight height patterns. As a result, the model assumes a single, pooled distribution for each species, potentially overlooking context-dependent variation that could affect collision risk estimates.

The aim of this report is to provide an overview of the ‘smoothing’ techniques currently in existence that can be applied to GPS-derived altitude data to reduce errors. Smoothing is not yet standardised across all species but is recommended for datasets with high variability or sparse observations. Each technique will be described, and an overview provided with the pros and cons of each in relation to large spatial datasets. Explanations of R scripts are provided in the Results section and appendices. Finally, we provide advice on which technique to apply in future stochastic collision rate calculations for the KEC and EIAs for future offshore wind farms.

2 Overview of smoothing techniques

Flight height data are frequently summarised using a single metric, typically average flight altitude and its variation. However, for various analyses capturing the entire distribution of flight altitudes is more important rather than focusing solely on mean values. This is particularly crucial when assessing collision risk: a wide variance in flight heights can mean that, even if the average altitude lies outside the collision risk zone, individuals may still spend a significant amount of time within it. Time spent in the collision risk zone is not only influenced by mean flight height but also its variability. As such, collision risk assessments should incorporate the full distribution of flight heights, rather than relying solely on fixed-effect estimates from linear models, a recommendation likely applicable to all research concerning vertical space use.

GPS altitude errors in the range of 10-50 m are common, whereas birds do not drastically change their altitude every few seconds. These errors can distort analyses and thus need to be accounted for. One way to deal with this is to incorporate ‘smoothing’ into GPS-derived flight altitude data. Smoothing reduces irregular variations and inaccurate fluctuations in altitude measurements to obtain a more realistic representation of a bird’s true flight path. Below, we provide an overview of various smoothing methods and the pros and cons of each. In the following chapters, more details about the different methods are provided.

Table 2.1 Overview of methods that can be implemented to 'smooth' GPS-derived altitude data, R packages required, computational demands, ease of use, and suitability for collision risk models.

Study	Methods	R (or other) packages used	Computational demands	Ease of use	Suitability for stochastic collision risk models (sCRM)
Ross-Smith *et al.* 2016	Bayesian state-space models based on GPS data	rjags called via MCMC in R	High (test on 3 birds, 1000 iterations took 15 mins, excluding script-writing). For 25 birds @ 5000 iterations, runtime would be ~11 hours	Requires Bayesian MCMC expertise, constructing state-space models	Outputs are full flight height distributions, with uncertainty; percentage of time in collision risk window
Péron *et al.* 2020	Hierarchical state-space models	crawl, ctmm, TMB in R C++ script	High (test on 25 birds took >4 hours, excluding script-writing)	Requires Bayesian MCMC expertise, constructing state-space models	Output could be adapted for CRM but focus is on data correction/pre-processing
Johnston *et al.* 2014	Continuous flight height distributions based on visual survey data	Cubic splines, bootstrapping, non-linear maximisation (nlm) function in R	Moderate – high (test on aggregated dataset with 50 bootstraps took 10 mins, excluding script-writing). A full analysis on 25 birds with 200 bootstraps would take ±35 hours.	More accessible to ecologists comfortable with R and generalised regression methods	Outputs are continuous distributions of flight height for many species, with confidence intervals. Useful for collision risk models but no correction for measurement error; less information on individual variability
Garde *et al.* 2023	Local polynomial regressions; Locally Estimated Scatterplot Smoothing (LOESS)	stats, stochlab in R	Low (test on 25 birds took less than 5 minutes, excluding script-writing)	More accessible to ecologists comfortable with R	Outputs are full flight height distributions, easy to implement in CRM, but no confidence intervals and possible boundary bias

2.1.1 State-space models

State-space models can be used to correct location inaccuracies in satellite tracking data (Table 2.1). A state-space model is a type of stochastic framework (a model that incorporates randomness or uncertainty) used to represent how a latent state variable (i.e. flight height) evolves over time when it is indirectly and imperfectly measured. These models distinguish between two separate components: the state process, which governs the true but unobserved changes in the variable, and the observation process, which accounts for measurement errors. When correcting for spatial error, the state variable being modelled is the physical location itself (Péron *et al.* 2020).

Movement models can be designed so that flight heights never fall below zero. This allows incorrect readings (e.g. -7 m) to provide useful information, as the model recognises a minimum error of 7 m. These erroneous data points can enhance state-space model performance, making them more suitable for vertical movement analysis. A challenge in state-space modelling is weak identifiability, whereby high sampling variance makes it hard to distinguish between measurement error and real movement. This is especially true when individuals in datasets have more variation between groups than within them. Negative flight height errors can help overcome this problem by aiding in the separation of measurement and process variances (Péron *et al.* 2020).

Various packages are available in R to fit state-space models: *crawl* (Johnson & London 2018) and *ctmm* (Continuous-Time Movement Modeling; Fleming & Calabrese JM 2023) use Kalman filtering, offering high performance, well-suited for large datasets. These methods assume Gaussian error structures and linear dynamics, which can be limiting in the presence of non-Gaussian noise (data that do not follow a normal distribution), irregular sampling intervals, or behavioural state changes. *TMB* (Template Model Builder; Kristensen *et al.* 2016) is a powerful C++-backed package that enables users to define custom state-space models, including those with non-Gaussian errors, latent behavioural states and non-linear dynamics. It offers high flexibility but is computationally intensive, especially when fitting complex models to large datasets.

Bayesian approaches (model-building incorporating prior knowledge and observed data to make inferences), such as those implemented via Monte Carlo Markov Chain (MCMC) methods in e.g. *rjags* (Plummer 2025), can be used for complex, non-linear, or truncated models. However, these are also computationally intensive, often requiring hours or days to run, and are generally more appropriate when precision and uncertainty quantification are critical, such as with very small sample sizes, complex movement environments, or conservation-critical applications. Furthermore, the focus of state-space modelling is primarily on data correction/pre-processing, not direct integration into CRM frameworks. The outputs require adaptation before they can be used in Band-derived models or StochLAB. Nevertheless, state-space modelling is appropriate when data have significant or non-uniform error, such as GPS altitude estimates with variable accuracy, or when behavioural states need to be inferred from uncertain movement data.

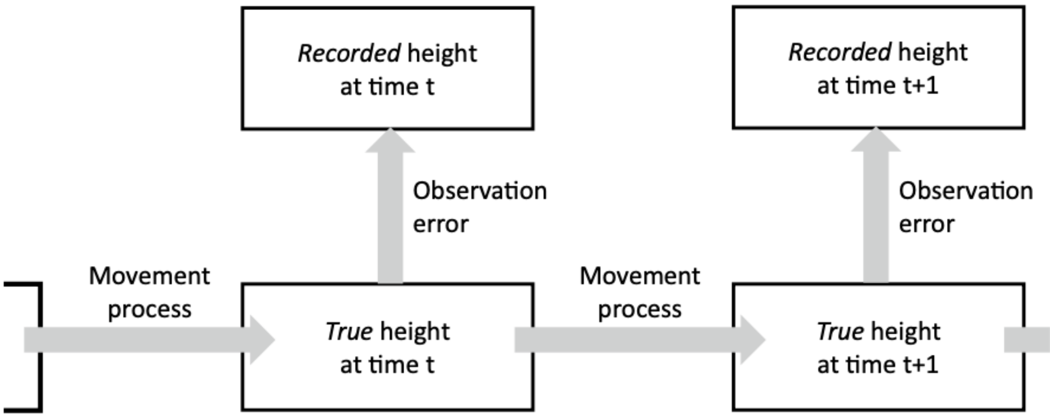


Figure 2.1 *Schematic overview of the principles of a state-space model as applied to the correction of sampling errors in flight height data. The movement (or state) process accounts for the distribution of true flight heights. The observation process introduces sampling errors of various origins and yields the recorded flight heights. It also accounts for the sampling schedule. By fitting this model to recorded flight height time series, it is possible to retrospectively compute the corrected flight height, an estimate of the true flight height (Péron et al. 2020).*

Examples of state-space models

Ross-Smith *et al.* (2016) used flight height distributions per state rather than relying on raw GPS altitudes, which may over- or underestimate flight heights, especially at low altitudes (Table 2.1). They also included covariates such as light level (day/night/twilight), habitat (marine/coastal/terrestrial), and speed (to distinguish behaviour such as flying vs floating/walking). Because flight altitude error is modelled and varies with dilution of precision (DOP), the model pulls observed inaccurate altitudes towards the latent/mean for that state (especially when observational uncertainty is higher), which acts as a smoothing effect.

The authors used a Bayesian state-space model for GPS-derived flight heights, treating each altitude measurement as an observation with error. The authors did not assume the observed altitude was true; instead they used a latent ("true") altitude of the bird, and the observed value was a noisy version of that. The magnitude of the observation error was modelled as a function of DOP. For each state, a log-normal distribution of true flight heights was assumed. Random effects were included per individual, whereby each bird had its own variation in mean altitude. Covariates influencing mean flight height for flying states were also included. The state of each fix was determined by speed and habitat, classifying the data into meaningful groups. Since some states may have contained reliable height data, by modelling the latent height separately and using the error model, noisy observations were smoothed over (Ross-Smith *et al.* 2016).

To assess potential turbine collision risks, the authors examined four vertical flight zones representing blade sweep heights from existing offshore designs. Analyses focused on flying birds in marine environments. For the lesser black-backed gull, behavioural state had a strong influence on estimated flight height. In marine habitats, 50% of flights occurred within ±13 m of sea level. Birds tended to fly higher during daylight, with altitude dropping significantly after dark (Figure 2.2). These findings highlight clear spatial and temporal patterns in flight altitude that have implications for collision risk estimates, especially under different light conditions and in contrasting habitats (Ross-Smith *et al.* 2016).

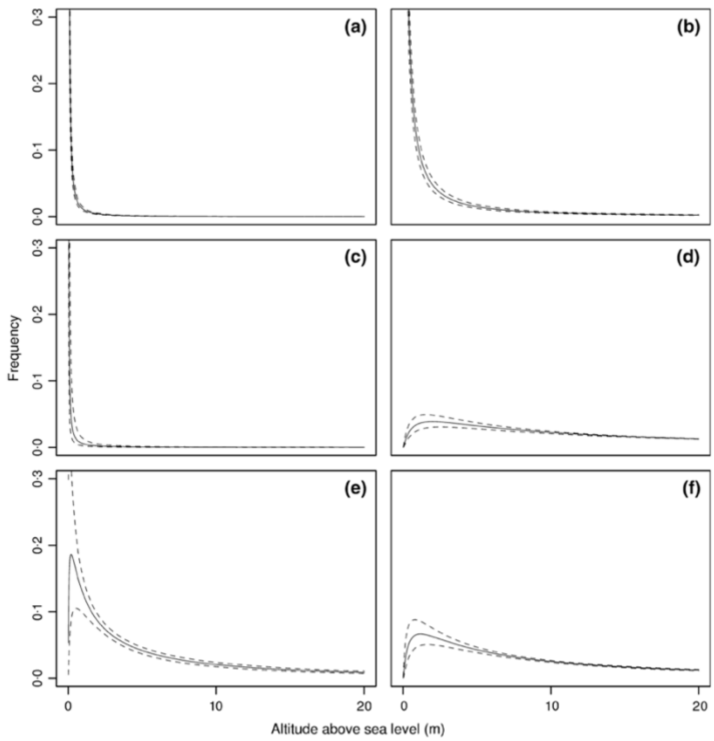


Figure 2.2 *Modelled heights for lesser black-backed gulls during the day in (a) state 1 (4 km h^{-1} 'terrestrial'); (b) state 2 (1-4 km h^{-1} 'terrestrial'), (c) state 3 (1-4 km h^{-1} 'coastal & marine'), (d) state 4 (>4 km h^{-1} 'terrestrial'), (e) state 5 (>4 km h^{-1} 'coastal'); and (f) state 6 (>4 km h^{-1} 'marine'). Solid lines – median; dashed lines – 2.5% and 97.5% credible intervals (Ross-Smith et al. 2016).*

Péron *et al.* (2020) performed field trials to quantify the magnitude and characteristics of vertical (and horizontal) errors. Based on the data collected, the authors used simulations to explore how errors and movement dynamics propagate into observed flight height distributions. They modelled true vertical movement using a stochastic process model, specifically an Ornstein-Uhlenbeck model, which allowed them to isolate and explore the effects of different error components under controlled conditions. This enabled the variance of the observation error to be estimated and sometimes linked to DOP metrics as covariates or via a log-linear model.

Improbable (e.g. negative) flight heights were treated as information that constrained the error distribution and were retained so the model could allow some probability mass to produce negative observations given the latent true height. The authors fit models via computational methods (e.g. Kalman filtering, Bayesian MCMC), and the recorded distribution of flight heights was compared to the corrected (latent estimated) distribution to assess the amount of bias, variance inflation and probabilistic time at critical heights (i.e. collision risk zones; Peron *et al.* 2020).

The authors found that bias in recorded flight height is highly sensitive to flight behaviour and error characteristics, making it difficult to correct without explicit modelling of error. Relying solely on mean flight height as a summary metric can be misleading in collision risk assessments since even if the mean altitude falls outside the rotor-swept zone, high variance means a bird may be at risk. Based on simulations and real raptor data, even if birds fly above or below the danger zone (average altitude), they could still spend a significant proportion (<20%) of time within it (Figure 2.3). Modelling the full distribution of flight heights rather than using mean values alone was recommended (Péron *et al*. 2020).

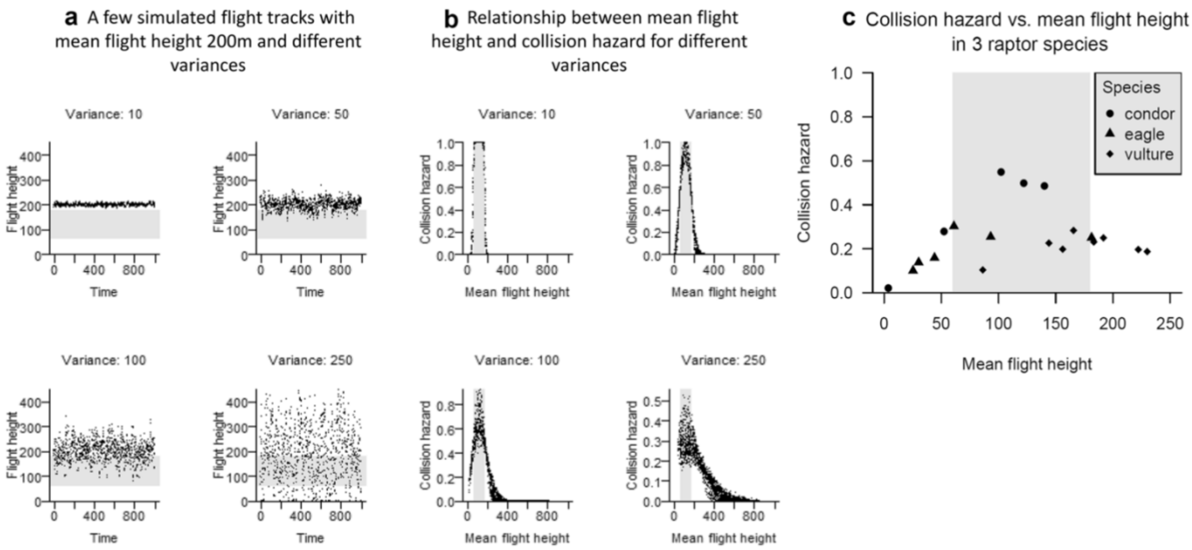


Figure 2.3 *Variance in flight height influences the percentage of time spent in the collision zone of a wind farm (grey area, 60-180 m). a) Four simulated tracks (where the true flight height is known) within the same mean flight height (200 m) but different variances (10, 50, 100 and 250 m²). b) More extensive simulations. Each point corresponds to one simulated track with a different mean flight height. c) Same as b) but using real datasets collected from three raptor species, where the corrected flight height is assumed to represent the true flight height. Each symbol stands for an individual over its entire monitoring period (Péron et al. 2020).*

The strengths of state-space modelling include realistic estimates of measurement error, disentangling how sources of error (vertical, horizontal) affect observed data, and separating true biological variation from noise. Limitations of this method include computational complexity and intensity, and the need for high-resolution data with VDOP metrics to support error modelling.

2.1.2 Multinomial/cubic spline approach

The multinomial/cubic spline approach is an alternative method that uses a two-part modelling process to create a biologically realistic flight height distribution. First, a multinomial likelihood is applied to treat the number of GPS fixes in each altitude bin as a multinomial outcome. In other words, each GPS point is assigned to one of several discrete altitude bins, and the total number of fixes is distributed among those bins. The model estimates the probability of occurrence in each bin. The model then estimates the probability of occurrence within in each bin, providing a discrete representation of flight height frequencies.

In the second step, a smooth curve – typically a cubic spline – is fitted to these probabilities to model how the likelihood of being in each altitude bin changes gradually with altitude. This smoothing ensures that adjacent bins have similar probabilities, avoiding abrupt changes that would be biologically unrealistic. The result is a continuous flight height distribution that captures the overall shape of the data while accounting for sampling uncertainty. This approach is particularly useful when data are sparse or unevenly distributed across altitude ranges, as it reduces uncertainty and produces a more interpretable curve for use in CRMs.

Various packages are available in R to fit cubic spline models: *splines* (base R) provides natural cubic splines, offering a simple and transparent approach for smoothing altitude data. These methods require manual specification of degrees of freedom and do not automatically optimise smoothing parameters, which can lead to over- or under-smoothing if not tuned carefully. *mgcv* (Wood 2017) implements generalised additive models with spline-based smoothers, including cubic regression. It automatically selects smoothing parameters and can incorporate uncertainty estimates, making it well-suited for stochastic CRM. However, *mgcv* is computationally intensive for large datasets and requires more expertise to interpret model diagnostics.

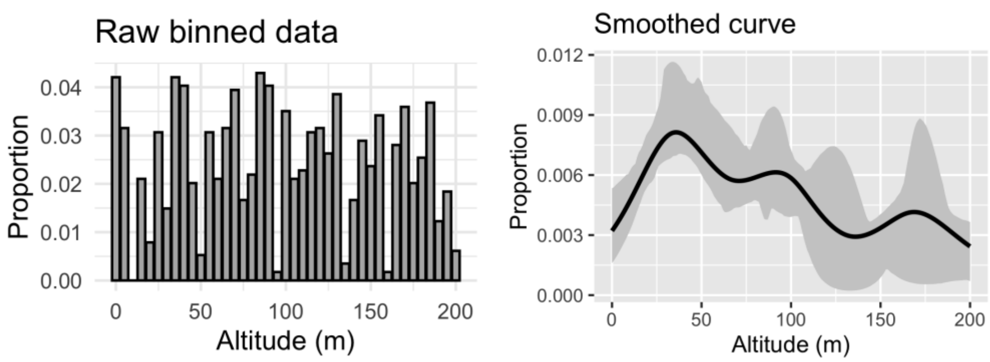


Figure 2.4 *(left) Raw binned GPS altitude data; (right) smoothed curve fitted using a cubic spline to represent continuous flight height distribution. Schematic example of the multinomial/cubic spline approach created in R package* splines*, adapted from Johnston et al. (2014).*

Overall, spline-based methods are efficient and biologically realistic for modelling FHDs, but care should be taken to balance flexibility with interpretability when applying them in stochastic CRM workflows. Spline smoothing influences the shape of FHDs, which directly affects collision risk estimates. Poorly chosen parameters can lead to biased or unrealistic results, therefore smoothed curves should be validated against the raw data.

Example of multinomial/cubic spline approach

Johnston *et al.* (2014) gathered visual survey data from 25 seabird species at 32 offshore wind farm sites in the UK and Europe to estimate flight heights at sea. Birds were classified into height bands, which varied between sites depending on the design of proposed turbines and the use of nearby structures as visual references. Flight height distributions were modelled for each species using a continuous approach i.e. rather than fitting standard distributions, a smoothing technique known as a flexible cubic spline was used on the log-transformed counts of birds at each height (Table 2.1). This process smooths over noise in the data by creating a continuous probability distribution, avoiding overfitting by using a limited number of knots to balance model flexibility with smoothness. The purpose of this is to model realistic flight behaviour, allowing more precise risk assessments and enabling uncertainty estimation by bootstrapping (resampling the data) around the smoothed flight height curve.

To fit a cubic spline to the height band data, the number of birds at each site in each flight height category were assumed to follow a multinomial distribution, whereby each bird had a certain probability of being assigned to a given height band, with all probabilities summing to one. Assuming independence between sites, the overall likelihood was calculated as the product of multinomial likelihoods across all sites. The log-likelihood function accounts for the number of birds observed in each height band at each site, integrating continuous spline estimates over the range of each band. This likelihood was maximised using the *nlm* function in R, allowing estimation of the spline coefficients. These coefficients, when used in the spline equation, produced a continuous estimate of bird flight height distributions for each species. The resulting values were standardised to produce the estimated proportion of birds flying within each 1-m height band (Johnston *et al.* 2014).

To estimate collision risk, for each bootstrap iteration the estimated proportion of flying birds within the collision risk zone was calculated. In this example, the authors determined that most seabird flight activity occurred within 20 m of the sea surface. Some species showed signs of a secondary peak in flight height at greater altitudes, as indicated by the confidence intervals. Species such as Arctic skua (*Stercorarius parasiticus*), Manx shearwater (*Puffinus puffinus*), little auk (*Alle alle*), and Atlantic puffin (*Fratercula arctica*) were heavily concentrated near the sea surface. In contrast, several gull species displayed more evenly distributed flight heights (

Figure 2.5). When turbine arrays were modelled using higher-capacity turbines, collision risk decreased by 29% when switching from 2 to 3 MW turbines, and a further 29% when switching to 5 MW turbines (Johnston *et al.* 2014). It should be noted, however, that since Johnston *et al.* used data based on visual observations, their results may be heavily skewed towards lower flying birds based on observer bias.

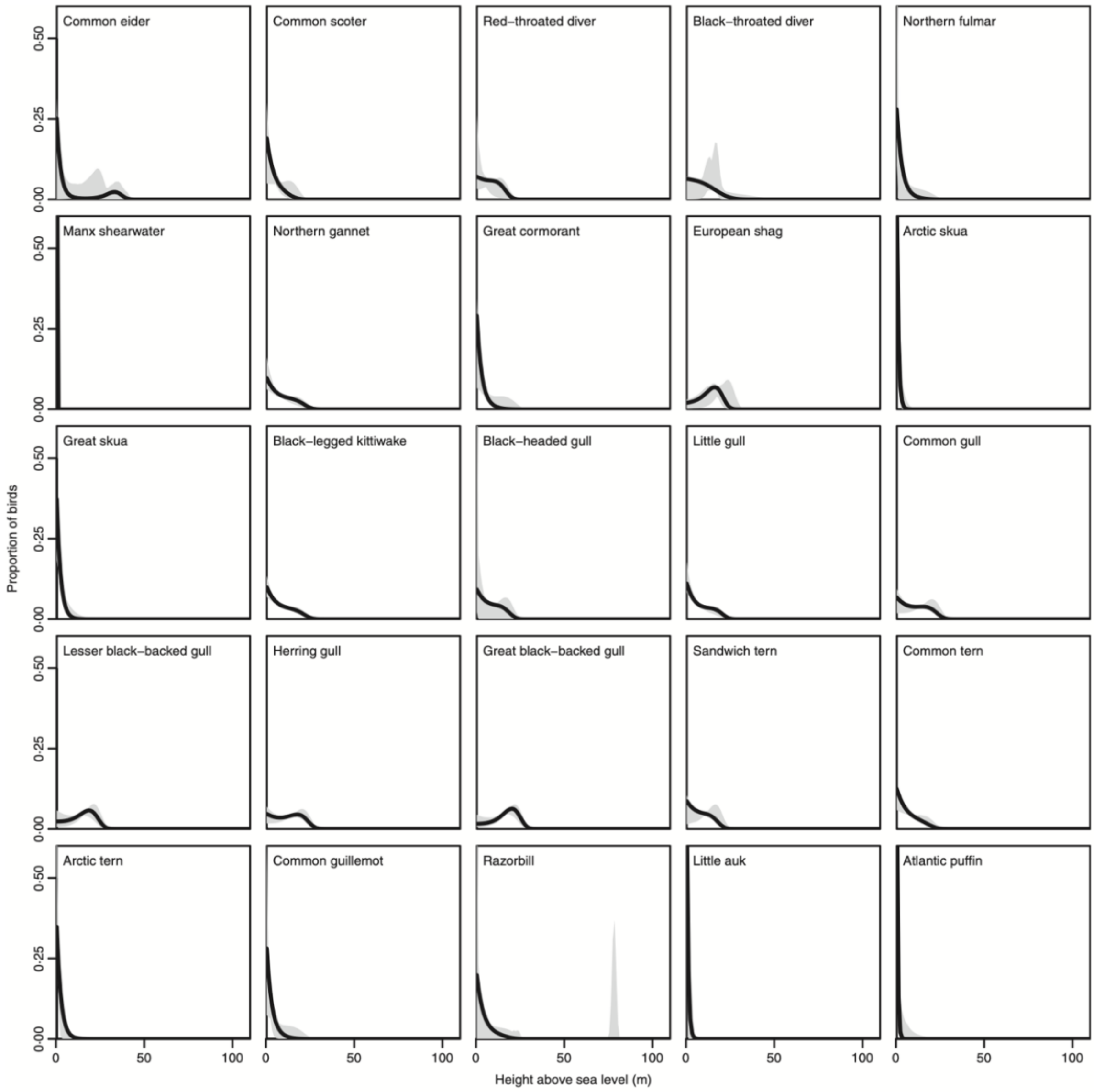


Figure 2.5 *Modelled flight height distributions (black line) and associated 95% bootstrap confidence intervals (grey area) of various seabird species. Estimates are not always in the centre of the confidence limits, because the confidence limits are nonparametric, and proportions are calculated for each bootstrap (Johnston et al. 2014).*

This method improves flexibility, especially when alternative designs are considered, and provides quantitative estimates of uncertainty. However, its reliance on categorical observer data and pooled distributions means that, when applied to a new site or under different environmental conditions, estimates should be used with caution. In particular, there is a risk of underestimating collision risk for species that fly at higher altitudes, or for variation driven by local conditions.

2.1.3 Locally estimated scatterplot smoothing (LOESS)

The simplest and by far most computationally efficient method is LOESS smoothing, a non-parametric technique that involves fitting local polynomial regressions to create a smooth line through a scatterplot to reveal trends in the data. Unlike global models that apply a single equation to an entire dataset, LOESS fits simple models (linear or quadratic) to localised subsets of the data. This allows it to adapt to complex and non-linear relationships without requiring a pre-determined mathematical form.

For each GPS point in the dataset, LOESS selects a subset of nearby data points based on a specified span (bandwidth). A span of 0.3 means a weighted least squares regression is performed on the closest 30% of bins around a point in a subset of the data. The closer the points are, the more weight they receive. The fitted local model is used to estimate the value of the smooth line at that point using the local regression. This process is repeated for each point along the x-axis to produce a smooth curve that follows the overall trend. The span parameter controls the proportion of data used for each local regression: a lower span value (0.1-0.2) captures local detail but risks overfitting (with possible noise), whereas a higher value (0.5) results in a smoother curve but runs the risk of oversimplifying the trend (Cleveland *et al.* 1992). The optimal value balances bias and variance, and it is possible to automatically select the span that best minimises prediction error (see §3.1).

LOESS is a flexible technique that can capture non-linear patterns without requiring a specific model form. It can help visualise trends in noisy time series data, smoothing GPS data to reduce measurement error. In ecological spatial studies, LOESS can be applied to smooth GPS tracking data while still preserving biologically relevant movement patterns. The results of such models can then be used as an input for collision risk modelling. Some advantages of LOESS smoothing are its non-parametric flexibility (the model does not assume a specific distributional form, making it ideal for skewed or multi-modal altitude data), its ability to handle sparse or noisy data at individual level, the ability to set smoothing strength, low computational intensity, and the ability to produce natural-looking, smooth curves that are easy to interpret.

Some limitations of LOESS include lack of optimisation for very large datasets, especially when applied separately for many individuals, boundary bias (LOESS can behave poorly at the edges of the data range, especially if those points are sparse), sensitivity to outliers or clumped data, and lack of confidence intervals. Bootstrap resampling or implementing a Bayesian approach could help address this but would significantly increase computational demands.

Example of LOESS

Garde *et al.* (2023) applied LOESS smoothing to visualise the relationship between thermal soaring behaviour and environmental variables among red-tailed tropicbirds (*Phaethon rubricauda*) based on GPS and magnetometry data. Smoothing altitude data allowed interpretation of non-linear, biologically realistic flight profiles without imposing rigid parametric assumptions i.e. the underlying constraints that are imposed on data by a statistical model (linear, quadratic).

After smoothing altitude data, the authors demonstrated that tropicbirds mostly flew within 100 m of sea level (Figure 2.6) with one individual reaching a maximum of 699 m altitude.

Peaks in flight height tended to occur near the end of foraging trips, particularly for birds returning to the colony with a tailwind (Figure 2.6).

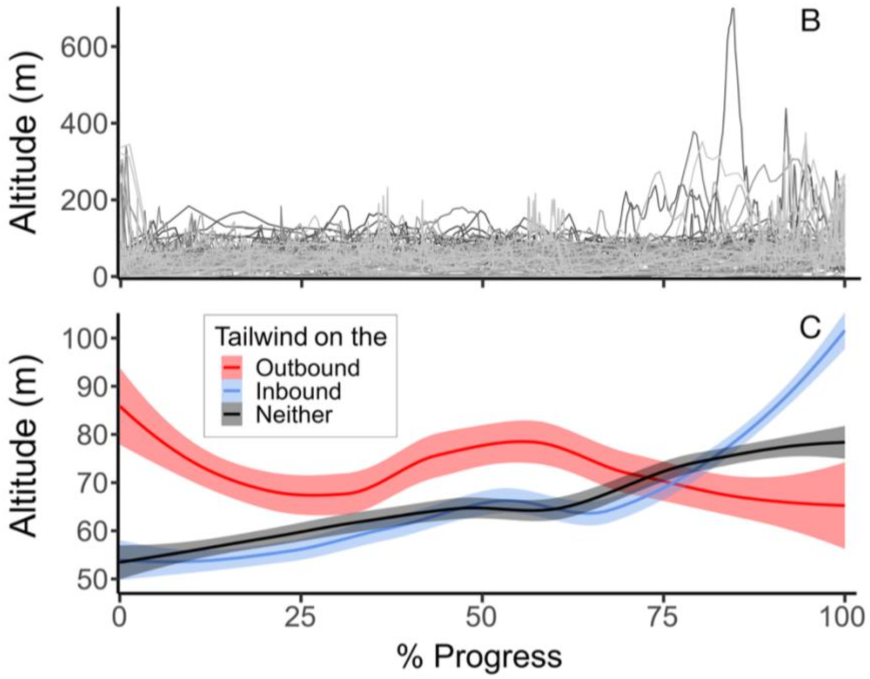


Figure 2.6 *(B) Altitude profiles of all trips. (C) Locally estimated scatterplot smoothing (loess) of altitude according to whether birds experienced a tailwind (mean TWC > 0) only in the outbound (red, n = 8) or inbound (blue, n = 21) phases of the trip, or neither outbound nor inbound (black). Garthe et al. 2023.*

In Table 2.2 we present an overview of all three smoothing techniques and the pros and cons of each for collision risk modelling. In Section 3 we present case studies for each of these methods based on GPS data from great black-backed gulls breeding in Norway.

Table 2.2 Overview of three smoothing methods tested for flight altitude data and the pros and cons of each for collision risk modelling.

Method	Pros	Cons
State-space modelling	- Explicitly accounts for observation error and latent true altitude. - Handles temporal autocorrelation and irregular sampling. - Provides mechanistic inference and uncertainty estimates. - Can incorporate covariates and random effects.	- Computationally intensive. - Requires strong statistical expertise. - Sensitive to model specification (input) errors. - Requires C++ script. - Requires high-resolution GPS data. - More complex to implement for large datasets. - Circularity in modelling if using the same variable to adjust altitude and predict risk (see Poessel et al. 2017).
Cubic spline	- Flexible for smooth, continuous curves. - Good for modelling non-linear trends. - Relatively easy to implement. - Works well for interpolation between known points.	- Assumes a global functional form (piecewise polynomial). - Can overfit if too many knots. - Does not explicitly model observation error. - Less robust for extrapolation beyond data range. - Works best with high-resolution GPS data; sparse fixes can lead to unrealistic curves.
Locally Estimated Scatterplot Smoothing (LOESS)	- Non-parametric; avoids rigid functional form assumptions. - Excellent for visualising complex, local patterns. - Adaptable to non-linear relationships. - Simple to apply and shorter computational time than state-space/cubic spline methods. - Span (bandwidth) can be automatically selected.	- No explicit error modelling. - Sensitive to span choice. - Can be computationally heavy for very large datasets. - No inclusion of states/covariates.

3 Case studies for smoothing

3.1 Methods

Below, we describe the steps taken to produce a flight height distribution based on GPS data from great black-backed gulls breeding in southern Norway.

As a case study, we examined the fight altitudes of 25 great black-backed gulls from GPS data collected between May 2024 and May 2025, originating from Norway (data ownership: Ecowende CV). The raw data (508,551 positions; Table 3.1) showed altitudes ranging from –1995 to +9950 m, including some extreme and evidently implausible values likely due to GPS inaccuracy (Figure 3.1).

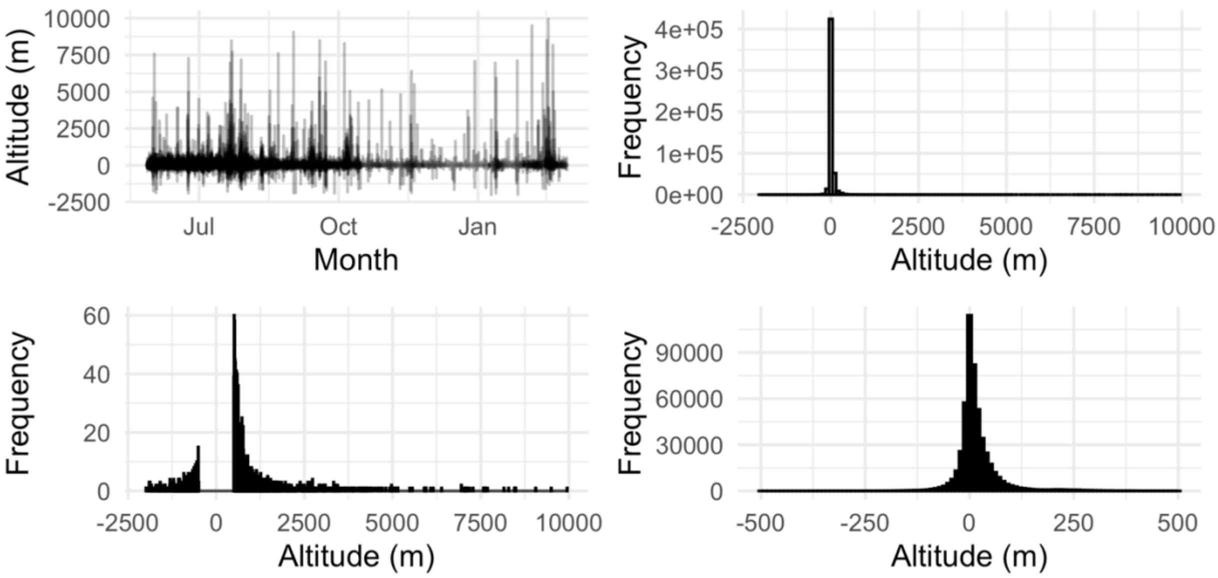


Figure 3.1 Plots showing GPS-derived flight altitudes of 25 great black-backed gulls over a one-year period (May 2024 – 2025).

One method to deal with such uncertainty is to exclude extreme and unrealistic flight heights, then filter the raw data based on dilution of precision (DOP) values. Lower DOP values (1-2) represent higher accuracy, whereas values >5 represent poor accuracy (Lato *et al.* 2022). However, not all GPS manufacturers provide vertical (VDOP) values, the most common being horizontal dilution of precision (HDOP) and/or number of satellites. Errors are more acute with high DOP values, over rough terrain, in poor or variable weather, and in areas with many obstacles that may cloud the GPS signal (Péron *et al.* 2020), whereas a higher number and more even spread of satellites produces more reliable GPS points (Table 3.1).

Table 3.1: Summary of HDOP category, satellite category and number of GPS fixes based on a raw dataset from 25 great black-backed gulls over a one-year period (May 2024 – 2025).

HDOP category	No. of fixes	Satellite category	No. of fixes
Excellent (≤1)	168516	Excellent (≥10)	76379
Good (1-2)	163028	Good (7-9)	233539
Moderate (2-5)	1296	Moderate (4-6)	192510
Poor (>5)	134	Poor (<4)	6123

3.2 Case studies

<u>State-space model (based on methods described by Ross-Smith *et al.* 2016)*</u>

<u>* A test using the methods described by Péron et al. (2021) is provided in Appendix 1</u>

We began by subsetting the data to three birds and filtering the data to retain only flying points (speed >5 km/h) and altitudes between 0-300 m. We sorted the data by individual and time, with birds being indexed from 1 to nbird. For each bird, a matrix of altitudes (alt[t, b]) and DOP classes (dop[t, b]) was created. The number of observations per bird (nobs[b]) were recorded. All matrices were padded to match the maximum number of observations. A Bayesian model was defined in JAGS syntax, consisting of three main components: 1) observation model with altitude measurements (alt[t, b]) assumed to be noisy observations of the true (latent) altitude (true_alt[t, b]), with observation error dependent on DOP class. The log of the true altitude was modelled as normally distributed around a bird-specific mean (mu.alt[b]), with individual-level variation. This ensured altitudes remained positive and allowed for log-normal vertical distribution. Each bird had a unique random intercept, composed of a global mean and an individual-specific random effect. DOP classes had their own estimated observation errors (sig.obs[1], sig.obs[2]). A population-level random effect standard deviation (sig.re) governed individual variability.

We compiled and sampled the model using the R package *rjags* (Plummer 2025), using a burn-in of 2,000 iterations to allow the sampler to reach the posterior distribution. We then used posterior sampling (5,000 iterations per chain) to yield estimates for all parameters. We used the R package *coda* (Plummer et al. 2006) to inspect the MCMC output, which included trace plots, posterior densities and Gelman diagnostics to assess convergence. Posterior means and credible intervals were extracted for parameters e.g. mu.alt, sig.re.

We exponentiated posterior draws of log_true_alt to obtain smoothed estimates of flight heights. These can be used to construct probability distributions of flight heights per bird for integration into collision risk models, particularly those requiring altitudinal percentiles (e.g. time spent in 1-m bins). We then plotted the raw vs. smoothed flight height distributions of three birds. In order to perform the same calculations on 25 birds and create an appropriate input file for collision risk modelling, 5,000 iterations would be required per sampler. This would significantly increase the computational intensity and run time. For this test (with 2,000 iterations on three birds) the run time was approximately 15 minutes; for 5,000 iterations on 25 birds the run time would be ± 11 hours.

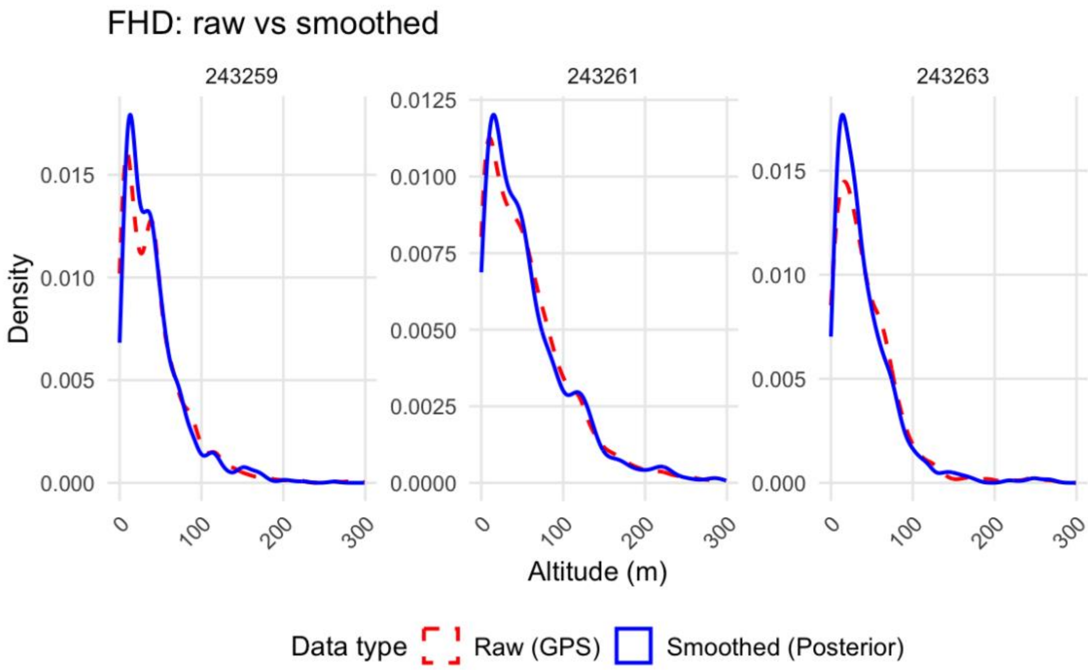


Figure 3.2 *Flight height distributions (FHD) of three great black-backed gulls (logger IDs at top of each plot) derived from GPS loggers between May 2024 and May 2025. Blue line = smoothed FHDs; red line = raw FHDs. Based on methods described by Ross-Smith et al. 2016.*

<u>Cubic spline approach (based on methods described by Johnston *et al.* 2014)</u>

We began by filtering the data to retain only flying points (speed >5 km/h) and altitudes between 0-300 m. We binned altitudes into 10-m intervals, converted GPS points to spatial objects and assigned them to grid cells (5 km x 5 km). Each grid cell became a site (site_id) for modelling purposes.

We summarised the number of GPS points that fell into each height bin per site and computed the midpoint of each bin (mid_h). We then created a height grid (h_grid) from 0-300 m and set up knot locations for spline modelling at specific heights (0, 50, 100, 150, and so on) to control the smoothness/flexibility of the spline fit. We defined a log-density spline function (logN_h) with an intercept, linear and quadratic terms, and cubic spline terms. We normalised the spline output to get a probability distribution over height (compute_p_h()). We calculated the likelihood of observing the height counts at each site given the model using logLik_all_sites(), using the binned counts and integrating over each bin.

We used optim() to find maximum likelihood estimates of the model parameters (beta_hat). We used bootstrap sampling to repeat the model fitting process 50 times (to reduce run time, rather than 200 bootstraps recommended by the authors). In each iteration, the site IDs were resampled with replacement site-level bootstraps. After optimising the likelihood for each bootstrap replicate, we stored the resulting beta parameters and applied 95% confidence intervals at each height in h_grid. We then created a plot showing the aggregated flight height distribution of all birds across all sites combined. We visualised the smooth spline curve for flight height density with shaded 95% CI and risk zone boundaries at 20 and 120m. Based on the output, 40.8% of bird flights fell within the rotor-swept zone, however area-weighted exposure was 100% suggesting birds tended to fly where the rotor area is densest (near the hub). In order to perform the same calculations per bird and create an appropriate input file for collision risk modelling, the data would need to be subset by device ID and flight height distributions computed separately for each individual. This would significantly increase the computational intensity and run time. For this test (with 50 bootstraps) the run time was approximately 10 minutes; for 200 bootstraps on 25 birds (non-aggregated data) the run time would be ± 35 hours.

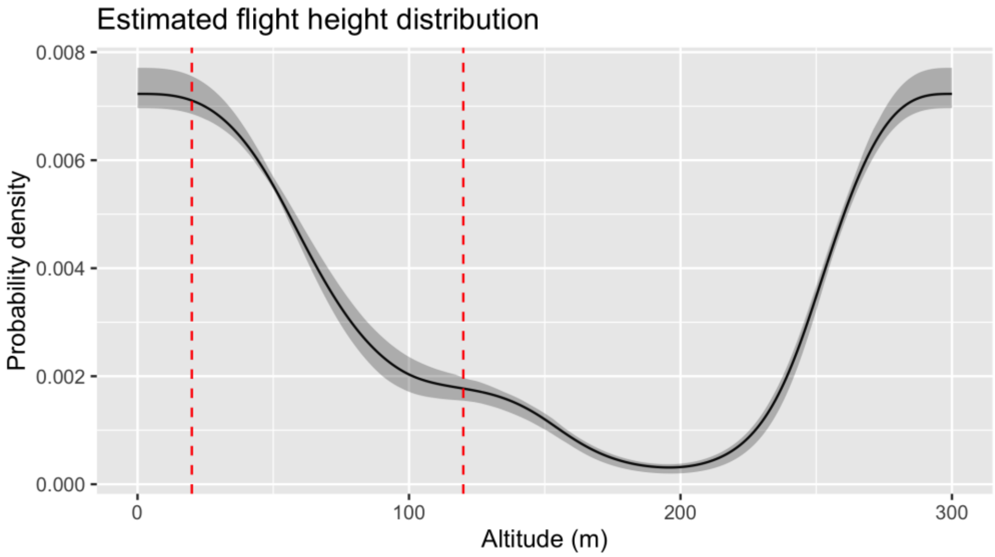


Figure 3.3 *Modelled flight height distributions (black line) and associated 95% bootstrap confidence intervals (grey area) of 25 great black-backed gulls tracked between May 2024 and May 2025 (aggregated dataset). Based on methods described by Johnston et al. 2014.*

Locally Estimated Scatterplot Smoothing (based on methods described by Garde *et al.* 2023)

In the absence of VDOP metrics, we filtered the data to remove low-quality GPS fixes based on HDOP (<2.5) and satellite count (=>5; Lato *et al.* 2022). We also filtered the data to include only GPS positions from outside the breeding season (August 2024 – April 2025), since the flight heights of breeding birds may differ to that of non-breeding birds (Schaub et al. 2023). For the purpose of this test, we retained GPS positions with speed >5 km/h and altitudes between 0 and 300 m above sea level (Shamoun-Baranes *et al.* 2011). This resulted in a dataset containing 20,744 GPS fixes (compared to the original raw dataset containing 508,551 positions). For each bird, the number of GPS fixes in each 1-m band was counted to provide an empirical distribution of flight heights. We generated a complete grid of all birds x all altitude bins, inserting zero counts when data were absent to ensure uniform input for smoothing.

We applied LOESS smoothing in the R package *stats* (R Core Team 2024) to each bird's altitude-bin frequency distribution to reduce positioning inaccuracies, capture underlying flight patterns, and avoid abrupt jumps in the data. To ensure optimal smoothing, we defined a set of candidate LOESS spans ranging from 0.1 to 0.9, in increments of 0.1. This automated selection method selected the span that best minimised prediction error, thus balancing under- and over-smoothing. For each bird, we grouped and sorted data by altitude bin, filtered to include only bins with non-zero counts when selecting the optimal span to avoid fitting noise.

For each candidate span, we fit a LOESS model and computed the mean squared error (MSE) between observed and predicted counts. MSE represents the average squared difference between predicted and actual values: low MSE means predictions are close to the true values (better fit); high MSE means predictions deviate more from the true values (poorer fit). The span with the lowest cross-validated mean MSE was selected. Using the selected best span, we used the predicted values as the smoothed counts. The smoothed values were normalised so that for each bird the sum of smoothed bin heights equalled 1, producing a valid flight height distribution representing the proportional use of altitude bins.

Finally, to assess the effect of smoothing, we compared the smoothed flight height distributions (FHD) to the raw, unsmoothed FHDs (metrics included MSE, absolute difference and correlation) across 501 altitude bins, and plotted the results per bird using the package *ggplot2* (Wickham 2016) to provide a visual comparison. We then plotted the output of the LOESS-smoothed FHDs of three birds with that from a state-space model to compare whether the two methods aligned or differed significantly from each other.

Output of LOESS smoothing

The overall MSE between smoothed and raw data was 9.21×10^{-6}, indicating extremely close agreement across the dataset. The mean absolute difference (the average of all the absolute differences in the dataset) across all birds was 0.43, and median correlation was 0.90. Total absolute difference ranged from 0.22 (birds 243279 and 243269), suggesting the model predictions were close to the reference values overall, to 1.08 (bird 243270), whereby deviations between the observed and reference values were larger (Table 3.2).

Correlation coefficients ranged from 0.59 (bird 243270), suggesting that smoothing significantly altered the pattern, to 0.97 (bird 243276), indicating that smoothing preserved the overall pattern well (Table 3.2). Four birds had a correlation coefficient of 0.95 or higher, suggesting high similarity between the smoothed and raw FHDs (Table 3.2). Bird 243270 exhibited the highest discrepancy, demonstrating that smoothing significantly altered the original flight height distribution for this individual.

Table 3.2 Overview of tagged birds, total absolute difference (sum of absolute differences between paired values across all rows) and correlation coefficients (strength and direction of the linear relationship) based on LOESS smoothing of GPS data from 24 great black-backed gulls between August 2024 and April 2025 (non-breeding season).

Bird iD	Total absolute difference	Correlation coefficient
243259	0.52	0.84
243260	0.47	0.82
243261	0.30	0.90
243262	0.40	0.89
243263	0.72	0.83
243264	0.50	0.84
243265	0.38	0.91
243266	0.41	0.86
243267	0.29	0.95
243268	0.41	0.90
243269	0.22	0.96
243270	1.08	0.59
243271	0.46	0.85
243272	0.55	0.86
243273	0.35	0.92
243274	0.37	0.93
243275	0.32	0.92
243276	0.25	0.97
243277	0.34	0.91
243279	0.22	0.96
243280	0.44	0.80
243281	0.33	0.94
243282	0.47	0.87
243283	0.53	0.86

The plots below (Figures 3.2 – 3.5) provide a visual representation of the smoothed FHDs (blue line), providing a continuous and more stable estimate of vertical space use compared to the original, raw data without any smoothing (red line). The y-axis shows the normalised frequency, and the x-axis represents flight altitude (filtered to 0-300 m).

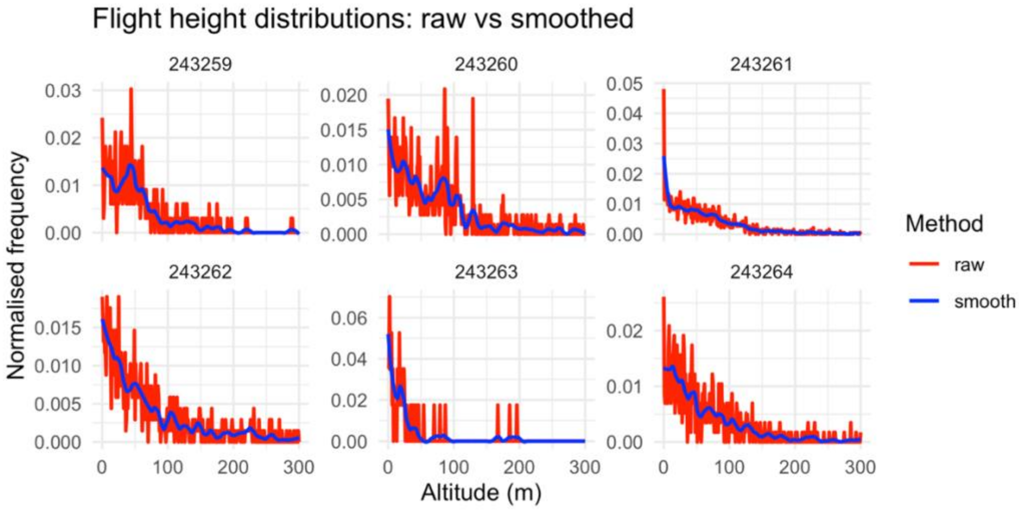


Figure 3.4 *Flight height distributions (FHD) of six great black-backed gulls (logger IDs at top of each plot) derived from GPS loggers between August 2024 and April 2025 (non-breeding season). Blue line = smoothed FHDs; red line = raw FHDs.*

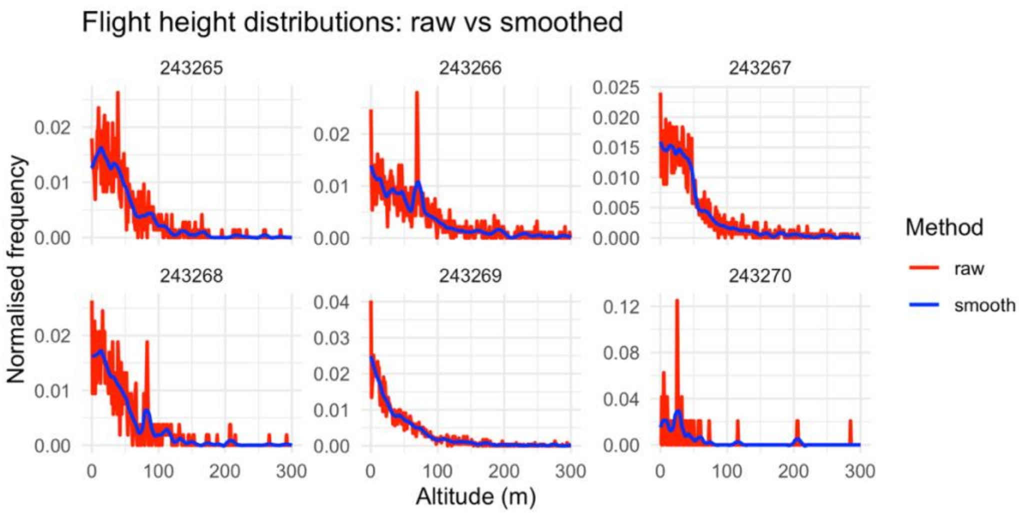


Figure 3.5 *Flight height distributions (FHD) of six great black-backed gulls (logger IDs at top of each plot) derived from GPS loggers between August 2024 and April 2025 (non-breeding season). Blue line = smoothed FHDs; red line = raw FHDs.*

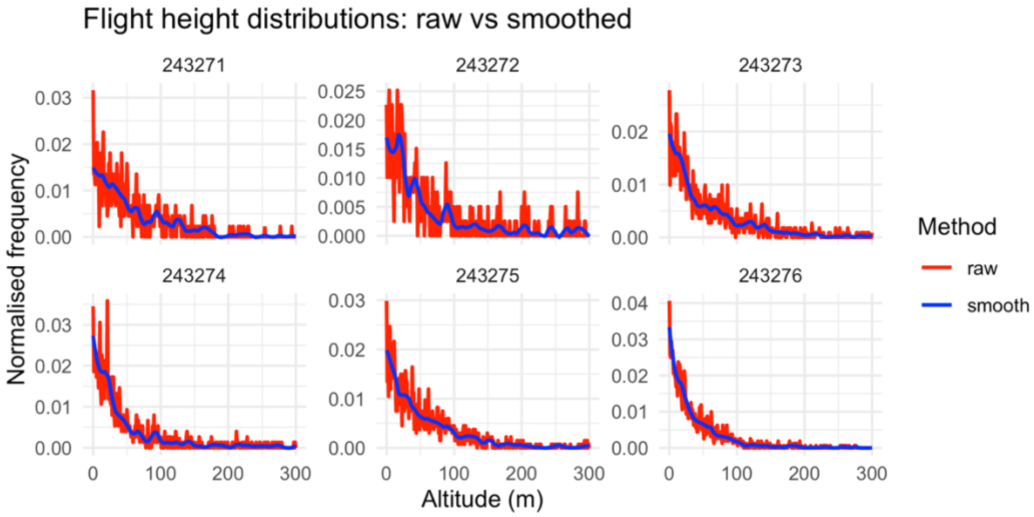


Figure 3.6 *Flight height distributions (FHD) of six great black-backed gulls (logger IDs at top of each plot) derived from GPS loggers between August 2024 and April 2025 (non-breeding season). Blue line = smoothed FHDs; red line = raw FHDs.*

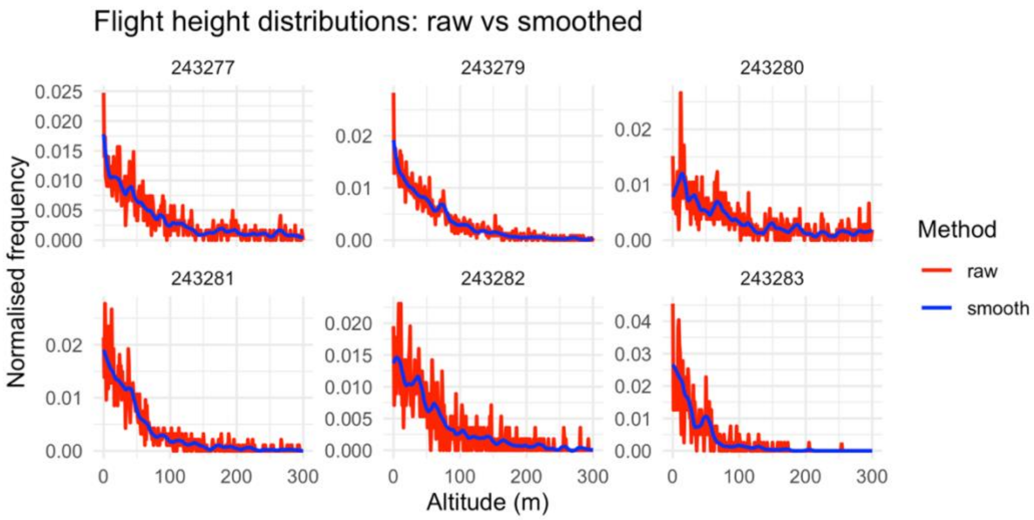


Figure 3.7 *Flight height distributions (FHD) of six great black-backed gulls (logger IDs at top of each plot) derived from GPS loggers between August 2024 and April 2025 (non-breeding season). Blue line = smoothed FHDs; red line = raw FHDs.*

After exporting and converting the outputs of the LOESS-smoothed and state space model data to a format appropriate for use in stochastic collision risk modelling, we calculated the proportion of time spent in 1-metre altitude bins by individual birds (0 – 300 m). We visualised these by comparing both outputs in the same plot for three birds (243259, 243261 and 243263). As can be seen from Figure 3.8, the proportion of time spent in each altitude bin based on the outputs of the state space and LOESS-smoothed model are similar and follow an identical trend, although the 'line' from the state space model is smoother. How sensitive the CRM is to such small differences of proportion in time at different altitude bands, and whether these produce significantly different risk estimates, was not tested in this study.

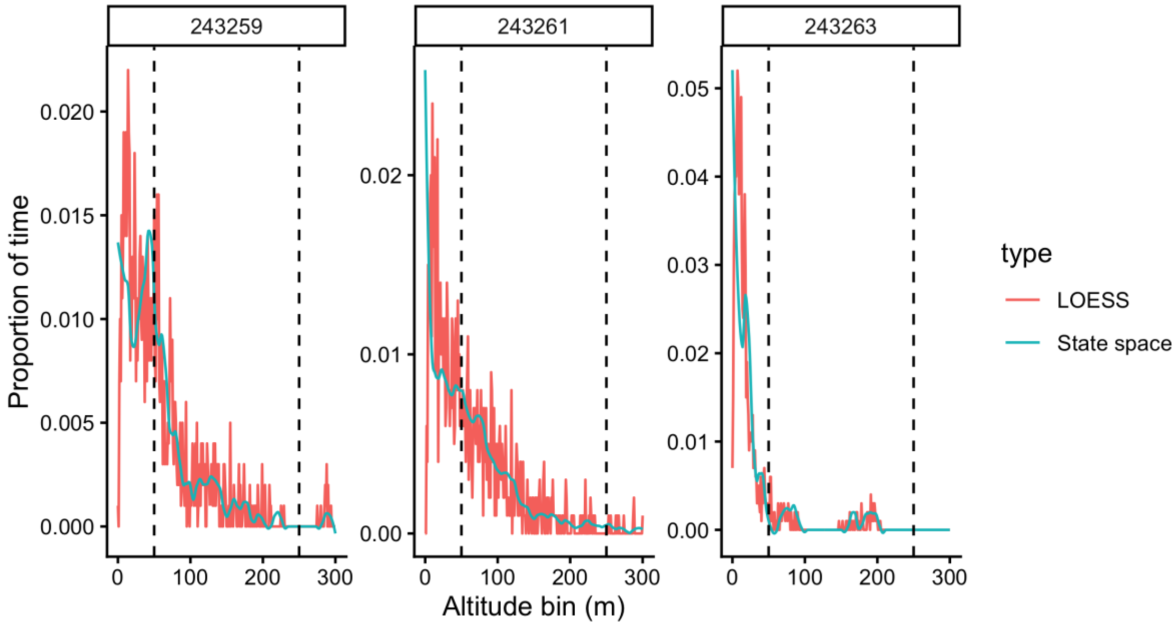


Figure 3.8 *Plots showing proportion of time spent in 1-m altitude bins by three great black-backed gulls based on GPS data collected from August 2024 to April 2025 (non-breeding season), with altitude data smoothed using LOESS (red line) and state space modelling (blue line). Dotted lines represent the hypothetical lower and upper bounds of a rotor-swept zone.*

Based on the above, and after experimenting with three different methods (Table 2.2), we concluded that LOESS is the most suitable method in existence to smooth flight height data for input into current stochastic CRMs. This approach enables data-adaptive smoothing that is tailored to the structure of each bird's data; the need to reduce noise is balanced with the ability to preserve biologically meaningful variation in vertical space use.

4 Conclusion and discussion

Stochastic collision risk models often rely on high-resolution vertical distribution data to estimate birds' potential exposure to rotor-swept zones. Smoothing helps distribute the influence of sparse or irregular bins across nearby altitudes, better reflecting true altitude use patterns. Smoothed flight height distributions improve interpolation between altitude bins and reduce erratic inputs, leading to more reliable collision risk estimates. Smoothing is therefore a valid method to reduce inaccuracies in flight height distributions while maintaining distributional structure and is a useful tool for collision risk modelling, especially if spatial sampling across altitudes/birds is uneven.

In this report, we explored several methods for smoothing GPS-derived flight altitude data to improve the accuracy of collision risk models for seabirds. After comparing three options, we selected LOESS smoothing as the most practical and effective approach for the current needs of KEC 5.0. Future iterations of the KEC may want to incorporate location-specific state-dependent and environmental covariates into flight height distributions, allowing collision risk models to reflect context-specific behaviour and improve the accuracy of collision risk estimates. However, this would introduce additional complexity to model implementation, including data requirements, model structure and increased computational effort. The decision must therefore weigh the benefits of greater accuracy and ecological realism against the challenges of data availability, parameter uncertainty, and the additional time and resources required to run and interpret more complex models. For many species, high-resolution spatial data are sparse or absent, meaning that often only density data (e.g. from aerial surveys) are available. This makes it virtually impossible to include states that are required for cubic splines or state-space models.

Of the three smoothing methods tested in this report, LOESS is the most straightforward to use, computationally efficient, and produces reliable results that are well-suited for input into current collision risk models. The choice of the smoothing parameter strongly influences the output: a large span can oversimplify patterns whereas a smaller span can introduce inaccuracies. The use of cross-validation in our test to select the best smoothing span for each bird added objectivity and transparency to the method. LOESS assumes a relatively smooth transition between altitude bins. This makes it an excellent choice for most applications, especially when working with large datasets or when ease of use is a priority. However, LOESS can smooth out sharp peaks in flight height distributions which may be relevant for collision risk (i.e. sudden increase in altitude at rotor-swept heights), which could lead to underestimating risk in certain altitude bins. It may also overweight regions with dense observations (e.g. near sea level) and does not extrapolate well beyond the observed data range. If collision risk models require predictions at altitudes with sparse data, LOESS may produce unrealistic or undefined values. Similarly, for datasets with very few observations or highly inaccurate data, LOESS can produce unstable results because local regressions rely on sufficient neighbourhood points.

Compared to LOESS, state-space models offer the most biological realism by explicitly modelling behavioural states and observation error and incorporating covariates, but they require complex data and are computationally demanding, adding significant implementation challenges. Cubic splines provide smooth interpolation without strong assumptions and are easier to implement than state-space modelling, yet they can overfit if knots are poorly chosen and do not account for ecological context. If more detailed data become available (e.g. time of day, weather, or proximity to structures), or if a more complex level of modelling is required, advanced techniques like state-space models or cubic spline approaches may be preferable. These methods can provide richer insights but are more computationally intensive and require a greater level of expertise. They also rely on location-specific high-resolution GPS data, which may not always be available for all species, and especially not for all geographical regions tested in the KEC.

The recommended methodology for KEC 6.0 should prioritise generating species-level flight height distributions that are robust, reproducible across species/datasets, ecologically meaningful, and suitable for stochastic collision risk modelling. While incorporating behavioural states and environmental covariates would enhance realism, the added complexity and data requirements for implementation in the assessments of future offshore wind farms within the KEC framework mean this is currently not feasible. As a result, LOESS smoothing offers the most robust, user-friendly solution for current collision risk modelling scenarios. For specific (one-off) projects demanding greater detail or customisation (situations where the standard LOESS smoothing approach is insufficient because the project or species requires more nuanced modelling), alternative methods are available, and new smoothing techniques may be developed in the future that can be implemented to further refine the accuracy of collision risk estimates.

References

Caneco, B., Humphries, G., Cook, A. & Masden, E., 2022. Estimating bird collisions at offshore windfarms with stochLAB. https://hidef-aerial-surveying.github.io/stochLAB/.

Cleveland, W.S., Grosse, E. & Shyu, W.M., 1992. Local regression models. Chapter 8 of Statistical Models in S eds J.M. Chambers & T.J. Hastie, Wadsworth & Brooks/Cole.

Fleming, C.H. & Calabrese, J.M., 2023. ctmm: Continuous-Time Movement Modeling. R package version 1.2.0, https://CRAN.R-project.org/package=ctmm.

Furness, R.W., Wade, H.M. & Masden, E.A., 2013. Assessing vulnerability of marine bird populations to offshore wind farms. Journal of Environmental Management (119): 56-66.

Garde, B., Fell, A., Krishnan, K., Jones, C.G., Gunner, R., Tatayah, V., Cole, N.C., Lempidakis, E. & Shepard, E.L., 2023. Thermal soaring in tropicbirds suggests that diverse seabirds may use this strategy to reduce flight costs. Marine Ecology Progress Series 723:171-183.

Johnson, D.S. & London, J.M., 2018. crawl: an R package for fitting continuous-time correlated random walk models to animal movement data. Zenodo. https://doi.org/10.5281/zenodo.596464.

Johnston, A., Cook, A.S., Wright, L.J., Humphreys, E.M. & Burton, N.H., 2014. Modelling flight heights of marine birds to more accurately assess collision risk with offshore wind turbines. Journal of Applied Ecology 51(1): 31-41.

Kristensen, K., Nielsen, A. Berg, C.W., Skaug, H. & Bell, B.M., 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70(5): 1-21. doi:10.18637/jss.v070.i05.

Lato, K.A., Stepanuk, J.E., Heywood, E.I., Conners, M.G. and Thorne, L.H., 2022. Assessing the accuracy of altitude estimates in avian biologging devices. PLoS One 17(10): p.e0276098.

Péron, G., Calabrese, J.M., Duriez, O., Fleming, C.H., García-Jiménez, R., Johnston, A., Lambertucci, S.A., Safi, K. & Shepard, E.L., 2020. The challenges of estimating the distribution of flight heights from telemetry or altimetry data. Animal Biotelemetry 8(1): 5.

Plummer, M., Best, N., Cowles, K. & Vines, K., 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC, R News, vol 6, 7-11.

Plummer, M., 2025. rjags: Bayesian Graphical Models using MCMC. R package version 4-17, https://CRAN.R-project.org/package=rjags.

Poessel, S.A., A.E. Duerr, J.C. Hall, M.A. Braham & T.E. Katzner, 2018. Improving estimation of flight altitude in wildlife telemetry studies. Journal of Applied Ecology, 55(4): 2064-2070.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ross-Smith, V.H., Thaxter, C.B., Masden, E.A., Shamoun-Baranes, J., Burton, N.H., Wright, L.J., Rehfisch, M.M. & Johnston, A., 2016. Modelling flight heights of lesser black-backed gulls and great skuas from GPS: A Bayesian approach. Journal of Applied Ecology 53(6): 1676-1685.

Schaub, T., Millon, A., De Zutter, C., Buij, R., Chadœuf, J., Lee, S., Mionnet, A. and Klaassen, R.H.G., 2023. How to improve the accuracy of height data from bird tracking devices? An assessment of high-frequency GPS tracking and barometric altimetry in field conditions. Animal Biotelemetry 11(1): 31.

Shamoun-Baranes, J., Bouten, W. Camphuysen C.J. & Baaij, E., 2011. Riding the tide: intriguing observations of gulls resting at sea during breeding. Ibis 153: 411-415.

Skone, S., El-Gizawy, M. and Shrestha, S., 2001, June. Limitations in GPS positioning accuracies and receiver tracking performance during solar maximum. In *Proceedings of the international symposium on kinematic systems in geodesy, geomatics, and navigation. Banff, Alberta, June* (pp. 5-8).

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wood, S.N., 2003. Thin-plate regression splines. Journal of the Royal Statistical Society (B) 65(1):95-114.

Appendices

Appendix 1 TMB model demo/test (based on R script provided by Péron *et al.* 2020)

We began by fitting a stochastic process model, specifically an Ornstein-Uhlenbeck (OU) model, to GPS-derived altitude data from tracked great black-backed gulls. The goal of this demonstration was to reconstruct latent altitude trajectories, accounting for temporal irregularity in sampling and observational noise. This was done using the R package *TMB* (Kristensen *et al.* 2016), which enables fitting hierarchical state-space models efficiently using automatic differentiation.

The function fit_ou_model was first defined to fit an OU model for a single device (bird). We started by cleaning the input data: Altitude (m) was clipped to between -5 and 500 meters to reduce the influence of outliers or sensor errors, particularly those near zero or implausibly high. In the absence of vertical dilution of precision (VDOP) values, the horizontal dilution of precision (HDOP) values were capped at 5, a common threshold indicating poor GPS accuracy. Time differences between observations, dt, were cleaned to replace non-positive values with 1 second and capped at 7200 seconds (2 hours), since extreme intervals can distort model estimation. We also filtered out birds with median interval durations of 3600 s (1 hour).

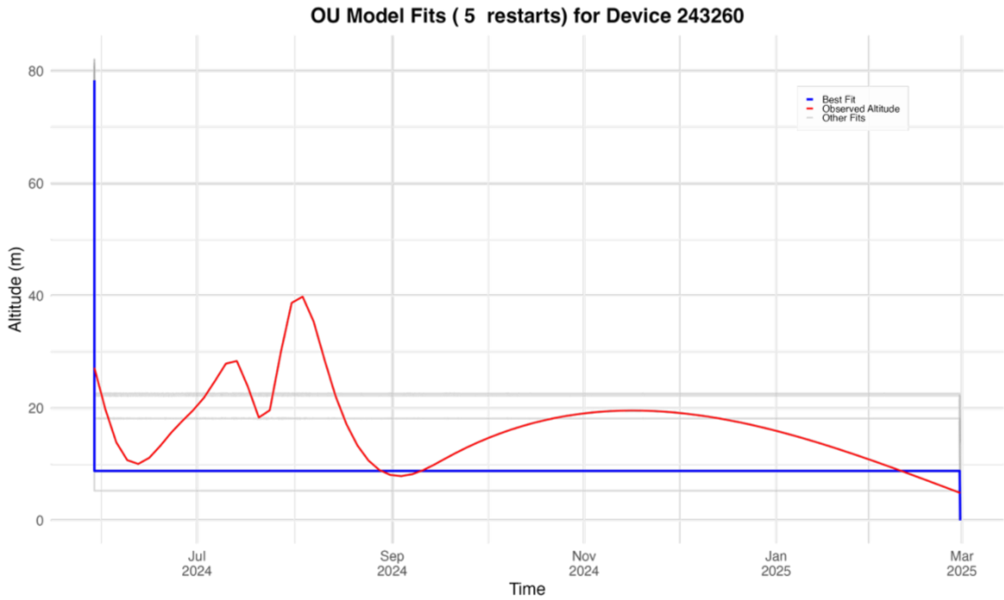
The cleaned data were passed into a list named Data, which formed the input for the TMB model. This included the clipped altitude values, the cleaned time intervals, a dummy covariate vector X (which may be replaced or extended later), the capped HDOP values, a binary flag for whether observations occur during the "non-day" (nd), and a model flag lag_model which remained fixed at 0 for this test, indicating no time-lag specific switching behaviour.

The parameters were then initialised. The parameter ltau is the log of the mean-reverting timescale (tau), and was initialised to -1. This corresponds to a timescale of exp(-1), or around 0.37 in natural units, meaning relatively fast reversion toward the mean altitude. The variance in tau was set to zero initially, implying no time-varying tau. Similarly, lsigma was initialised at 1, meaning the base volatility was exp(1), or around 2.7, giving the process room to fluctuate. The variance in sigma was set to zero. The long-term mean, mu, was initialized as the empirical mean of the clipped altitudes, assuming this provides a decent prior guess for central altitude. Its variance was also fixed at zero.

The parameter lsigma_nd was log-transformed and initialised to log(5), representing the standard deviation of the state process during the night (or any unobserved period). lsigObs, the log of the observational error, was initially set to 0, or exp(0) = 1. The parameter beta_dop was initialised at 0.2, meaning that the log observation variance would scale linearly with capped HDOP, representing a model where more uncertainty was attributed to higher DOP values. The degrees-of-freedom for the t-distribution, df_3, was initialised at 1, potentially allowing for heavier tails in the observation model. Finally, the transformed latent state vector Z_transf was initialised around 3, with small noise added to break symmetry and help with identifiability.

The model was built using MakeADFun from the TMB package, specifying the DLL (compiled C++ model) and setting Z_transf as a random effect. After building the model, the initial likelihood was checked to ensure it was finite. If not, the fitting for that device was skipped. Otherwise, the optimiser was run using the BFGS method with up to 2000 iterations and a strict relative tolerance. If optimisation was successful and produced valid parameter estimates, the fitted latent trajectory was extracted using the TMB report function.

For plotting, the fitted latent altitude trajectory (Z) was plotted over time, using the true UTC timestamps rather than accumulated time differences. Observed clipped altitude values were shown as red points, and the latent trajectory was shown as a blue line. A 95 percent confidence interval was included if standard errors were available. Plots were produced using the R package ggplot2 (Wickham 2016). Based on inspection of the plot below, the script would need refining as it did not produce the desired output.



Flight height distribution (FHD) of one great black-backed gulls (logger ID 243260) derived from GPS data collected between May 2024 and May 2025. Blue line = smoothed FHD; red line = raw FHD. Based on methods described by Péron et al. 2020.