Analysis of local habitat selection and large-scale attraction/avoidance based on animal tracking data: is there a single best method?

Moritz Mercker¹*, Philipp Schwemmer², Verena Peschko², Leonie Enners² and Stefan Garthe²

¹Correspondence: info@bionum.de
²BIONUM statistical consulting office, Hamburg, Finkenwerder Norderdeich 15 A, Hamburg, Germany
Full list of author information is available at the end of the article
†Equal contributor

Abstract

Background: New wildlife telemetry and tracking technologies have become available in the last decade, leading to a large increase in the volume and resolution of animal tracking data. These technical developments have been accompanied by various statistical tools aimed at analysing the data obtained by these methods.

Methods: We used simulated habitat and tracking data to compare some of the different statistical methods frequently used to infer local resource selection and large-scale attraction/avoidance from tracking data. Notably, we compared the performances of spatial logistic regression models (SLRMs), point process models (PPMs), and integrated step selection models ((i)SSMs) and their interplays with habitat, tracking-device, and animal movement properties.

Results: We demonstrated that SLRMs were inappropriate for large-scale attraction studies and prone to bias when inferring habitat selection. In contrast, PPMs and (i)SSMs showed comparable (unbiased) performances for both habitat selection and large-scale effect studies. However, (i)SSMs had several advantages over PPMs with respect to robustness, user-friendly implementation, and computation time.

Conclusions: We recommend the use of (i)SSMs to infer habitat selection or large-scale attraction/avoidance from animal tracking data. This method has several practical advantages over PPMs and additionally extends SSMs, thus increasing its predictive capacity and allowing the derivation of mechanistic movement models.

Keywords: animal movement; autocorrelation; bio-logging; habitat selection; point process; resource selection; species distribution; telemetry; avoidance

1 Introduction

The identification of factors that influence species distribution and resource selection is an important ecological issue [1] that has traditionally been addressed using appropriate regression methods based on presence-absence or count data [2, 3, 4].

However, recent technical developments involving radio- and telemetry-based approaches (e.g., Argos-, global positioning system-, or very-high frequency-based methods [5, 6, 7, 8]) allow single-animal tracks to be recorded, thus providing a new and highly valuable alternative source of information to help answer such questions. These approaches allow the locations and movements of species to be analysed at a much finer spatio-temporal scale than previously possible.
Frequent ecological questions associated with animal tracks concern either the selection/avoidance of a certain resource/habitat or structure, or alternatively changes in behaviour related to such covariates. Both questions can again be applied at different spatial scales. On a small scale, the direct interaction of an animal with its surroundings can be investigated; e.g., by evaluating if certain habitats are used more intensively than others. On a large-scale however, certain areas or structures may assert effects on animal movement, for example as a result of direct perception (e.g., over the sea possible for $\geq 10\, km$ [9]) or spatial memory [10, 11]. Animal movement is often eventually shaped by a complex combination of all of the above-mentioned types of perceptions and memories interacting at various spatio-temporal scales [8, 12, 13, 14].

In contrast to count- or presence-absence data, tracking data only contain information about single points in time and space where the animal is present, with no information for any other locations [15]. The statistical analysis of such ‘point observation data’ is often challenging, and various approaches have previously been developed and discussed (e.g., summarized in [8]). In summary, there are two distinctly different approaches for the analysis of animal tracking data.

The first approach is given by animal movement models that rely solely on tracking data points, e.g., by evaluating step sizes and turning angles. Variations on these models mainly differ from each other in the way in which time and space dependencies, as well as latent behavioural states, are entered into the model [8, 16, 17, 18]. Indeed, one of the main strengths of these approaches is the analysis of animal behaviour, possibly changing in short-range or large-scale interaction with the environment. ‘Discrete-time hidden Markov models’ are a prominent and increasingly used example [17, 18, 19, 20, 21]. However, habitat selection cannot be directly assessed using such movement models, given that they lack quantification of habitat availability.

The second approach augments tracking points with an additional set of artificially created points (‘pseudo-absences’, ‘dummy points’, or ‘available steps’) to quantify habitat availability. The tracking points thus represent the used habitats, while dummy points are chosen optimally to measure/represent how much of each habitat type is available. This approach is frequently used in the context of spatial point observation data, and various sophisticated strategies have been used to select appropriate pseudo-absences (e.g., [22, 23]), frequently followed by an analysis using spatial regression or machine learning techniques, such as MAXENT or spatial logistic regression models (SLRMs) (e.g., [24, 25, 26, 27]).

In the context of animal tracking data, such ‘dummy point approaches’ must however be used with care: first, the artificial generation of pseudo-absence points is an ad hoc approach and thus associated with several statistical disadvantages and criticisms: e.g., the choice of location and number of these points are often not straightforward, even though the regression results may sensitively depend on these choices [26, 27, 28]. Second, strong spatial, temporal, and angular autocorrelations...
are frequent challenges presented by tracking data [29, 30]: a tracking point is most likely to appear in spatial and temporal proximity to the previous point, and turning angles often correlate with their predecessors, inducing directional persistence. The importance of taking account of autocorrelations increases with the sampling frequency relative to the velocity of the animal [8], and neglecting these issues may distinctly bias the results [31, 32]. The correct implementation of autocorrelation in the context of ecological modelling is however complex: e.g., when considering spatial autocorrelation based on mixed modelling, fixed and random spatially varying covariates may be collinear (‘spatial confounding’), potentially causing biased interference [33, 34, 35], especially if covariates and residual structures act on similar spatial scales [36].

Two new methods have become available during the last decade, providing a rational basis for the choice of dummy points on the one hand, and integrating appropriate correlation structures on the other. First, point process models (PPMs) naturally and automatically resolve many of the questions and pitfalls arising from previous techniques [8, 26, 27, 37]. In particular, PPMs allow the role and number of dummy points to be deduced purely mathematically by aiming an efficient estimation of an integral as a part of the PPM likelihood [26, 37]. Additionally, PPMs represent a generalisation of many other frequently used methods [26, 38, 39], and the PPM likelihood can be approximated using standard generalised linear modelling regression software [27, 40] (possibly including mixed and/or additive modelling). The latter ensures a flexible and individual implementation, including an appropriate integration of spatio-temporal autocorrelation [40]. The second is step-selection models (SSMs) [41], which have been developed from the viewpoint of an individual, in contrast to the population viewpoint adopted in PPMs and related models [42]. SSMs have recently been combined with simultaneous animal movement models estimating movement and resource selection parameters, leading to integrated step selection analysis (iSSM) methods [43]. The above-mentioned different viewpoints of PPM and (i)SSM approaches become noticeable e.g., when looking at autocorrelations in the data. In (i)SSM methods, these dependencies are used to create appropriate data stratification and to choose reasonable dummy point locations, whereas in PPMs, these correlations are considered to be a statistical nuisance [42, 43]. It is therefore not surprising that both methods can lead to divergent conclusions [42]. Finally, all three methods (PPMs, SSMs, and (i)SSMs) can be used not only to test hypotheses, but also to estimate utilization distributions respectively for predictions [8, 43, 44].

However, the variety of available statistical tools makes it difficult to select the most appropriate one for analysing a particular set of tracking data and addressing a specific research question. Although partial comparisons between the different approaches have been presented (e.g., [42, 43, 45]), a more comprehensive comparison of the different frequently used methods in terms of their interplays with habitat, movement and tracking-device properties has not, to the best of our knowledge, yet been presented.
In the current study, we used simulated data to systematically analyse and compare the performances of different statistical approaches (namely SLRMs, PPMs, SSMs, and (i)SSMs) with respect to local resource/habitat selection, as well as large-scale attraction/avoidance processes. We analysed and compared the statistical powers (i.e., rate of detecting an existent effect) and false-positive rates (type I error rates) in the interplay with habitat, animal movement, and tracking-device properties (Fig. 1). Based on these results, we provide practical guidance for the conditions under which each of these methods works reliably.

2 Methods

2.1 Overview
To cover the potential variety of animal tracking data, we simulated a broad range of reasonable animal tracking data, resulting from an interplay among various habitat types, animal movement properties, and tracking-device outcomes. To investigate the statistical power of different frequently used methods, different variations of SLRMs, PPMs, and (i)SSMs were applied to these simulated tracking data, the latter influenced by local habitat selection and/or large-scale attractions of varying strength. In contrast, if the rate of false-positive findings (type I error rate) was the focus, they were applied to simulated data with no underlying attraction effects.

The above framework produced four different data frames for final comparative analyses: two data frames for the analysis of local habitat selection (with vs. without underlying simulated attraction effect), and two data frames focussing on large-scale attraction (also with vs. without underlying simulated attraction effect). Each data frame consisted of approximately 7,000 rows nested in 400 different animal tracks (i.e., each track has been analyzed with various methods), which we considered to be an appropriate order of magnitude to detect practically relevant differences between the compared methods. In particular, each row depicted one unique method–track combination comprising information about (1) the statistical method used, (2) several parameters related to habitat, movement, and tracking-device properties (including the respective strengths of the above-mentioned two attraction effects), and (3) the outcome variable (effect) defined by the binomial response ‘significant effect detected’ vs. ‘no significant effect detected’ (where the \( \alpha = 0.05 \) significance level has been used).

In a final step, these four final data frames were further analysed using appropriate regression model selection techniques, e.g., to evaluate and compare the average statistical powers and false-positive rates of the different methods in interplay with the several predictors related to the simulated tracking and environmental data. For the sake of clarity, in the following, variable names starting with \textit{Meth}... concern the applied statistical method of interest (i.e., a specific variation of a SLRM, PPM, or (i)SSM approach), variables starting with \textit{Hab}... represent habitat properties, \textit{Mov}..., variables define properties related to animal movement, and variables starting with \textit{Dev}... concern tracking-device properties. An overview of all considered variables is given in Tab. 1. A graphical overview of the workflow applied in this study is given in Fig. 1.
2.2 Computation and software

All statistical analyses were performed using the open source software R [46]. In particular, spatial methods and spatial visualizations were mainly based on the package raster [47], and all other statistical plots were performed using ggplot2 [48]. Additional R-packages and functions are detailed below. All computations were performed using a non-parallelized code on an Intel(R) 4-Core(TM) i5–6600 with 3.30 GHz and 16 GB RAM.

2.3 Simulated data

2.3.1 Habitat data

Habitat raster data were generated based on the raster()-function in the R-package raster [47], where habitat values (given by the variable hab) were randomly assigned (uniformly distributed between 0 and 1) to 500 × 500 = 250,000 spatial 2D pixels \( \vec{x} = (x_1, x_2) \). Possible spatial autocorrelation (i.e., local spatial clustering of certain habitat types) was quantified by the continuous variable \( \text{Hab}_{auto} \in \{1, ..., 60\} \), where a value of 1 represents no autocorrelation (i.e., the hab value for each pixel does not depend on the hab values in the surroundings) and values of \( \text{Hab}_{auto} > 1 \) represent an increasing isotropic autocorrelation. In contrast, anisotropic autocorrelation was represented by the continuous variable \( \text{Hab}_{anis} \in [0, 1] \), where a value of 0.1 represents an isotropic landscape, up to strongly anisotropic landscapes with maximal values of \( \text{Hab}_{anis} = 1.0 \), the latter resulting in habitats stretched along the x-axis. Both types of autocorrelation were generated using a ‘moving window’ approach based on the R-function focal() in the raster package. Furthermore, the smoothness of the transition between different habitat types was defined by the categorical variable \( \text{Hab}_{smooth} \), which can generate sharp, medium, or smooth/blurry transitions, also realized in the context of the above-mentioned focal()-function. Finally, after generating autocorrelation and transition smoothness, in approximately 50 \% of cases, the continuous variable hab was transformed into a binomial one (based on the threshold of 0.5) where the final type of hab was represented by the variable \( \text{Hab}_{type} \) differentiating between continuous and categorical habitat data. Some examples for simulated categorical habitat data (permuted over the variables \( \text{Hab}_{auto}, \text{Hab}_{anis} \) and \( \text{Hab}_{smooth} \)) are shown in Fig. S1, and further examples including continuous habitats are presented in Fig. S2 and Fig. 1.

2.3.2 Animal movement

Simulated animal movement is strongly related to (and extends) the ‘stepping-stone’ algorithm, as presented by Avgar et al. [43]. In particular, we simulated movement on the finest available spatial scale, i.e., on a scale of pixels. For each simulation, the starting location and the location of the attraction centre were chosen randomly within the virtual study area, and a maximum of 5,000 steps per trip was simulated, unless otherwise stated. For each simulated time step \( t \) with location \( \vec{x} = (x_1, x_2) \), the probability of choosing the neighbouring pixel \( \vec{y} = \vec{x}_{t+1} \) out of the eight nearest neighbours \( N \) as the next point was given by

\[
p(x_{t+1} = \vec{y}) = \frac{F(\vec{x})}{\sum_{\vec{y} \in N} F(\vec{y})},
\]
with
\[
F(\vec{y}) = N(0, Mov_{SD}) + 
\exp \left( Mov_{\omega} \cdot \text{hab}(\vec{x}) - \mu ||\vec{y} - \vec{x}|| - Mov_{\sigma} \cdot \alpha_{att} - f(Mov_{ran}, Mov_{ran2}) \cdot \alpha_{pers} \right),
\]
where a more detailed motivation of the general structure of this stepping stone algorithm is given by Avgar et al. [43]. Here, \(N(0, M_{SD})\) represents a normally distributed random component in attraction strength, i.e., scaling the strength of random movement vs. directed/biased movement in the animal path, quantified by the movement standard deviation \(Mov_{SD} \in [0, 2.5]\). \(Mov_{\omega} \in [0, 1.0]\) quantifies the strength of resource selection, while the term \(\mu ||\vec{y} - \vec{x}||\) with \(\mu = 1.8\) penalizes larger Euclidean distances to neighbouring pixels. Furthermore, \(Mov_{\sigma} \in [0, 0.1]\) penalizes angular deviations \(\alpha_{att}\) from the direct path between \(\vec{z}\) and the attraction centre (and thus introduces directional bias towards the centre, "biased random walk" [8]), and \(f(Mov_{ran}, Mov_{ran2})\) finally penalizes angular deviations \(\alpha_{pers}\) from the direction of the foregoing movement step, thus leading to directional persistence ("correlated random walk" [8]). Notably, \(f(Mov_{ran}, Mov_{ran2}) = Mov_{ran}/(1 + Mov_{ran2} \cdot \text{hab}(\vec{y}))\) includes the variable \(Mov_{ran} \in [0, 2.5]\) for the general strength of directional persistence, but also the variable \(Mov_{ran2} \in [0, 1.0]\) antagonizing this effect if local habitat values (averaged over all neighbours) are high. The latter effect thus induced a less-directed and more-random search behaviour in appropriate habitats. To avoid boundary effects, simulations were stopped as soon as a pixel belonged to the boundary of the simulated area. However, tracks with a size of \(\leq 2000\) points were discarded. Finally, after generating the animal track, a virtual time \(t\) was added to the data. In particular, for each time point \(t\) the next time point \(t+1\) was generated by \(t + 1 = 30 + \text{pois} \left( \lambda = 30 + 30 \cdot (1 + Mov_{ran3} \cdot \text{hab}(\vec{y})) \right)\) sec. Here, \(Mov_{ran3} \in [0, 1.0]\) reflects the possibility that animals slow down their movement in appropriate habitats, the latter again represented by high values of \(\text{hab}(\vec{y})\). I.e., for \(Mov_{ran3} = 0\), time intervals were on average 1 min long and independent of local habitats (but intervals may display some temporal stochasticity due to natural variation in animal velocity). In contrast for \(Mov_{ran3} \geq 0\), the average interval length between steps increases linearly with local average \(\text{hab}\) values. Some example tracks with varying values for \(Mov_{\omega}, Mov_{\sigma}\) and \(Mov_{ran}\) are given in Fig. S2.

### 2.3.3 Tracking device and final tracking data

The simulated animal movement data at the spatial pixel scale (c.f., previous subsection) were subsequently reduced to a much coarser temporal resolution, mimicking the data collected by a tracking device. A total of 150 tracking points were selected from the approximately 5,000 raw data points. For this selection, the temporal regularity of tracking points was represented by the categorical variable \(\text{Dev}_{\text{reg}}\), where a value of 1 represents temporally equidistant tracking points (with respect to the virtual time \(t\)), and levels 2 and 3 represent increasing random irregularity, the latter realized via increasing variance of a random normal deviation from selected equidistant time points. Due to the inherent temporally discrete nature of simulated animal movement data, the nearest existing time point rather than the exact selected one was chosen. After selection of the spatio-temporal subset
('tracking data') from the raw animal movement data, as described above, spatial bias was added, quantified by the variable Dev.bias ∈ [0,3] depicting the standard deviation of a normally distributed random error separately added to each point and coordinate. These final animal tracking data were used to generate the following additional variables: total spatial extent of the trip defined by

\[ Dev\_ext = \frac{(max(x) - min(y)) + (max(x) - min(y))}{2}, \]

and total number of tracking points defined by the variable Dev.N.

### 2.4 Statistical analysis of simulated animal tracks

As noted in the Introduction, we applied and compared three main classes of statistical methods (SLRMs, PPMs, and (i)SSMs, and several variations) using simulated tracking data. To obtain optimal comparability, several measures during tracking data analysis were unified. First, the same level of dummy point numbers per tracking point (namely \( N\_dummy \in \{8, 80, 224\} \)) were used in all models. Second, central predictors were entered in all models in a similar manner: hab as a linear predictor for habitat selection studies, and attrac.dist for the evaluation of large-scale attraction measuring the Euclidean distance to the attraction centre. As an alternative to attrac.dist, the variable attrac.ang was used, calculating the cosine of the angular deviation from a straight line between the previous tracking data point and the attraction centre, thus representing the directional bias towards the centre. attrac.ang was not used for SLRMs, because a pure spatial approach does not allow the consideration of angular deviations depending on the temporal order of tracking points. Finally, all compared models were nested in simulated tracks (i.e., each simulated animal track was analysed with all models) to minimize the level of unexplained variance in the final comparative analyses.

**Spatial logistic regression models**

SLRMs were implemented using the gam() function in the R-package mgcv [49] with a binomial error distribution and logit-link function, where true tracking points served as presence points in the outcome variable, and dummy points were treated as true absences. In particular, dummy points were selected in two different manners: first, randomly chosen within the minimum convex polygon (MCP) around the tracking data (calculated by the mcp()-function from the sp-package [50]) depicting a simple and frequently used approach to estimate home ranges (e.g., Ref. [51, 52]) and leading to the model Meth_SLRM_mcp; and second, based on spatio-temporal tracking data (i.e., step size and turning angle distributions) as provided by the amt-package, leading to the model Meth_SLRM_amt. Notably, step lengths were based on a gamma distribution, whereas turning angles relied on a von Mises distribution [41, 43]. Although we used SLRMs to investigate the performance of pure spatial methods with respect to the analysis of tracking data, we used this spatio-temporal dummy point selection approach to obtain an ‘optimal set’ of spatially distributed dummy points, assuming that sophisticated pure spatial selection methods would be at least as good as this spatio-temporal selection method. Finally, to account for possible spatial autocorrelation, we optionally added a spatial 2D thin plate regression spline \( s(x_1, x_2) \) to the predictor (where the optimal number of knots was estimated based on generalized cross-validation [49]), leading to the models Meth_SLRM_amt_ac and Meth_SLRM_mcp_ac.
(Integrated) step selection models

SSM and (i)SSM analyses relied on conditional logistic regression functions \((\text{fitssf}, \text{fitissf})\), as provided by the R-package \texttt{amt} [44]. In particular, data were stratified by the time points of tracking points (‘used steps’) and temporally associated dummy points (‘available steps’) [41, 43]. Similar to \textit{Meth\_SLRM\_amt}, dummy points were generated based on step lengths and turning angles, assuming a gamma von Mises distribution [43, 44]. However, several predictor formulations are possible considering temporal, spatial, and directional correlations [44]. We therefore performed a model selection step based on the Akaike information criterion (AIC) [53] comparing 10 different models reflecting different combinations of main effects and interactions in relation to the different autocorrelation terms, namely temporal distance to the foregoing tracking point, logarithm of the spatial distance to the foregoing tracking point, and cosine of the turning angle [44]. The best models from this model selection step were termed \textit{Meth\_SSM\_1} and \textit{Meth\_iSSM\_1}, respectively. A second variant of these models was generated by augmenting the above-mentioned model selection by a second AIC-based model selection step, where the above models were compared with models where the logarithm of the spatial distance to the foregoing tracking point \((\log(d_x))\) was replaced by \(d_x\), \(\sqrt{d_x}\) and \(\log(\sqrt{d_x})\), respectively, leading to the models \textit{Meth\_SSM\_2} and \textit{Meth\_iSSM\_2}. Finally, in order to estimate the general importance of model selection within the (i)SSM framework, we fitted the (i)SSM model without any model selection procedure using \(\log(d_x)\), the cosine of the turning angle, as well as their interaction term as autocorrelation-related predictors \(\textit{Meth\_iSSM\_no\_ms}\).

Point process models

The applied PPMs mainly relied on the approach of Johnson et al. [40]. In particular, we approximated the spatio-temporal PPM likelihood by the expression

\[
\sum_{j,k} w_{jk} \left( u_{jk} \ln(\lambda(t_j, \vec{x}_{jk})) - \lambda(t_j, \vec{x}_{jk}) \right)
\]

with point process intensity \(\lambda()\) depending on time points \(t_j\) and spatial 2D points \(\vec{x}_{jk}\). Furthermore, \(u_{jk} = 1/w_{jk}\) for tracking points and \(u_{jk} = 0\) for dummy points with appropriate quadrature weights \(w_{jk}\) [27, 40], where the index \(j\) always refers to discrete time points and the index \(k\) to the discretisation in 2D space. This expression is proportional to a weighted Poisson likelihood with weights \(w_{jk}\) and observations \(u_{jk}\), such that standard generalized linear/additive modelling (GLM/GAM) software can be used [27, 40]. Dummy point generation was based on the R-package \texttt{mvQuad} [54]. In particular, for each tracking point, we created a rectangular grid of dummy points (centred around the previous tracking point and randomly rotated) including quadrature weights based on the rectangle rule. The latter outperformed several other possible quadrature weights with respect to maximal PPM statistical power, while showing a reasonable type I error rate (results not shown). However, the optimal spatial extent of the dummy point grid in PPMs is \textit{a priori} not clear, in contrast to SSM methods where dummy points represent ‘available steps’ and are thus defined by average spatial step lengths of the tracking point time series.
contrast, for each time point \( t_j \) in PPMs, the grid has to be large enough to cover sufficient area where the conditional intensity surface substantially differs from zero on the one hand, but small enough to obtain a satisfactory spatial resolution on the other. We compared two different approaches to achieve this goal. In the first model (\( \text{Meth}_{\text{PPM}_1} \)), the side length \( l \) of the dummy point grid associated with tracking point \( x_{jk} \) was based on the distance between \( x_{jk} \) and the foregoing tracking point \( x_{(j-1)k} \) by \( l = 2| x_{jk} - x_{(j-1)k} | \). The idea was that, similar to (i)SSMs, dummy point grid size was locally adapted to the spatial step length. In the second model (\( \text{Meth}_{\text{PPM}_2} \)), we determined the optimal \( l \) value empirically by fitting several models using the values \( l = 5, 10, 15, ..., 180 \), and finally choosing the \( l \)-value that led to the highest z-score with respect to the predictor of interest (e.g., habitat selection has been analysed) for further PPM analysis. I.e., we selected the model in which the most reliable attraction effect was detected.

All PPMs were fitted based on GAM-software [55] using the \texttt{bam()}-function in the \textit{mgcv} package [49]. Additive models allow the investigation and formulation of nonlinear relationships between different variables [49, 55], which we used to formulate possibly nonlinear autocorrelation-related predictors within the PPM, extending the linear approach of Johnson et al. [40]. In particular, autocorrelation has been integrated as a Markov process (similar to the spatio-temporal PPM approach of [40]), and considering a possibly nonlinear dependency of the tracking data on step duration, step length, and step heading, the same variables are suggested to integrate possible autocorrelation within the (i)SSM framework [43]. In particular, optimal models were selected based on the AIC (with AIC calculations based on the PPM-likelihood approximation and not on the standard AIC provided by the \textit{mgcv} package), comparing several linear and tensor-smooth combinations of these variables. Additionally, as proposed by Hooten et al [8], we optionally augmented the final selected models with a spatial 2D thin plate spline \( s(x_1, x_2) \) representing the large-scale home range and leading to the model \( \text{Meth}_{\text{PPM}_2, hr} \). To investigate the general importance of model selection within the PPM framework, we also fitted a model without any model selection procedure using a 3D-tensor-smooth for the above-mentioned autocorrelation-related variables (\( \text{Meth}_{\text{PPM}_2, no-ms} \)). Finally, in addition to the PPMs using a Poisson distribution, we alternatively fitted PPMs using a negative binomial (\( \text{Meth}_{\text{PPM}_2, nb} \)) and a quasi-Poisson (\( \text{Meth}_{\text{PPM}_2, qp} \)) distribution, possibly accounting for over- or underdispersion [8].

2.5 Final comparative analysis of model performance

As explained in the introduction to the Methods section, the final analysis comparing the performances of the different statistical methods was based on four different data frames: two analysing habitat and two analysing large-scale attraction. Notably, in each case, the analysis was based either on simulated tracking data with an existing underlying attraction effect (for evaluating statistical power) or with no underlying effect (for evaluating false-positive rates). In all cases, the binomial variable \textit{effect} (‘detected significant effect’ vs. ‘no significant effect detected’) was evaluated, likely depending on some of the 16 different variants of the applied statistical methods, but also possibly interplaying with the 15 different predictors related
to habitat, animal movement, and tracking-device properties (c.f., Tab. 1).

For the final comparative analyses, generalized estimating equation (GEE) models [56] with an 'exchangeable correlation structure' [2] and a binomial probability distribution were used, realized with the `geeglm()`-function in the R-package `geepack` [57]. We favoured GEEs over logistic generalized linear mixed models (GLMMs [2, 58, 59]), because residuals related to the trip IDs (used as a random intercept) strongly violated the normality assumption in GLMMs. Here, GEEs are known to be much more robust against miss-specification in correlation structures [2].

For the comparison of average values of statistical power in terms of false-positive rates between the different statistical methods, the variable `method` was used as the only predictor in GEEs. If the model performance in interplay with several habitat, movement, and tracking device-related properties was investigated instead, a very high number of possible predictors resulted from the variables presented in Tab. 1, because not only the main effects, but also the interaction terms between method- and tracking-related variables, were considered. We therefore applied an efficient model selection technique before GEE analysis, namely the 'least absolute shrinkage and selection operator' (LASSO) technique [60, 61] (based on the R-package `glmnet` [62]). This technique is known to perform reliable model selection even if predictor numbers are high [4]. Notably, we used logistic regression models within the LASSO procedure and combined them with cross-validation to select a promising set of predictors [63]. The cross-validation method is a systematic approach to adjust a certain degree of freedom (here, predictor number and choice) to maximize the generalization of the model to independent data (i.e., its predictive performance) [64]. In a second step, the selected predictors were analysed by GEEs as described above, finally allowing us to determine the circumstances under which each method performed best.

We analysed the impact of the tracking data size on model performance and computing time largely as described above (c.f. also Fig. 1). However, only the size of the simulated animal tracks was varied, such that the final tracking data size ranged between $N = 33$ and 300 tracking points; all other variables (related to habitat, movement, and tracking-device properties) were set to constant average values across all simulations. The latter aim was to minimize the unexplained variance. The final comparative analysis was again based on GEEs if statistical power or false-positive rates were investigated, while appropriate GAMs were separately applied for each method for the analysis of computation time [49, 55].

3 Results

3.1 Habitat selection

3.1.1 Average performance

We first compared the average performances of the 15 different variants of the three main statistical methods using 80 dummy points per tracking point. Corresponding GEE results are shown in Fig. 2.
| variable name   | explanation                                                                                                                                 |
|----------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Dev_reg        | temporal regularity of tracking points                                                                                                     |
| Dev_bias       | strength of spatial location error                                                                                                          |
| Dev_ext        | spatial extent of the animal trip                                                                                                            |
| Dev_step       | average spatial step size of the animal trip                                                                                                 |
| Dev_var        | variance of the spatial step size of the animal trip                                                                                         |
| Hab_auto       | strength of spatial habitat autocorrelation                                                                                                  |
| Hab_anis       | anisotropy of spatial habitat autocorrelation                                                                                                |
| Hab_smooth     | smoothness of transition between habitats                                                                                                     |
| Hab_type       | continuous or categorical habitat                                                                                                              |
| Mov_SD         | strength of randomness in animal movement                                                                                                     |
| Movω           | movement bias towards attraction centre                                                                                                |
| Mov_ran        | strength of directional persistence                                                                                                           |
| Mov_ran2       | reducing directional persistence in preferred habitats                                                                                         |
| Mov_ran3       | introducing slower movement in preferred habitats                                                                                               |
| Meth_SLRM_mcp  | spatial logistic regression model, dummy points are randomly generated inside the minimal convex polygon                                       |
| Meth_SLRM_amt  | spatial logistic regression model, sophisticated generation of dummy points based on spatio-temporal information                                 |
| Meth_SLRM_mcp_ac | as for SLRM_mcp but with spatial 2D smooth aiming to reduce spatial autocorrelation                                                          |
| Meth_SLRM_amt_ac | as for SLRM_ams but with spatial 2D smooth aiming to reduce spatial autocorrelation                                                            |
| Meth_SSM_1     | step selection model including model selection with respect to the autocorrelation terms                                                     |
| Meth_SSM_2     | as SSM_1 but with refined model selection with respect to spatial autocorrelation                                                            |
| Meth_iSSM_1    | integrated step selection model including model selection with respect to the autocorrelation terms                                           |
| Meth_iSSM_2    | as iSSM_1 but with refined model selection with respect to spatial autocorrelation                                                           |
| Meth_iSSM_no ms | as iSSM_1 but without model selection procedure                                                                                            |
| Meth_PPM_1     | spatio-temporal point process model including model selection, the spatial scale of the dummy point grid is related to local tracking data step size |
| Meth_PPM_2     | spatio-temporal point process model including model selection, the optimal spatial scale of dummy point grid is empirically determined as PPM_2 but augmented with spatial 2D smooth representing home range |
| Meth_PPM_2_hr  | as PPM_2 but using a negative binomial instead of a Poisson probability distribution                                                         |
| Meth_PPM_2_nb  | as PPM_2 but using a quasi-Poisson instead of a Poisson probability distribution                                                            |
| Meth_PPM_2_qp  | as PPM_2 but without model selection procedure                                                                                               |

Table 1: Possible predictors for statistical power and type I error rates in the context of habitat selection and large-scale attraction based on animal tracking data. Above the double line, predictors related to habitat (Hab,...), animal movement (Mov,...) and tracking-device (Dev,...) properties are shown; below the double line predictors related to the statistical method applied to simulated tracking data (Meth,...) are presented.
SLRMs without consideration of spatial autocorrelation appeared to show upward biased type I error rates, especially if dummy points were selected based on the simple MCP approach. In the model \textit{Meth\_SLRM\_mcp}, the type I error rate was 30\%, which is far above the nominal level of 5\%. The application of more sophisticated schemes for dummy point selection (namely the spatio-temporal method used in model \textit{Meth\_SLRM\_amt}) appeared to reduce this bias. However, our results suggested that the type I error rate was still above the nominal level, and could be even more pronounced if sophisticated but pure spatial methods for dummy point selection were applied. We therefore do not recommend using SLRMs without consideration of spatial autocorrelation to infer habitat selection from tracking data. This approach would only be appropriate if the tracking data were spatially uncorrelated; however, the existence of home ranges means that this will be rarely the case. PPM approaches using a quasi-Poisson distribution instead of a Poisson distribution (model \textit{Meth\_PPM\_2\_qp}) also showed distinctly upward biased type I error rates, making this approach also inappropriate for the analysis of habitat selection.

With respect to the statistical power (currently excluding the above models with biased type I error rates), (i)SSM approaches showed the highest average power, which appeared to be robust against the different variants of these models, such as different model selection procedures or SSMs vs. iSSMs. PPMs generally showed equally robust but slightly reduced power, except for the model \textit{Meth\_PPM\_1} where the dummy point grid extension was based on the local distance between subsequent tracking points. The latter approach led to distinctly reduced power and type I error rate, indicating that the empirical choice of the dummy point grid extension was critical for good PPM performance. All the remaining PPM variants (possibly including a negative binomial instead of a Poisson distribution, an additional 2D home range smooth in the predictors, or lacking the model selection procedure) showed similar magnitudes of statistical power to each other and to the (i)SSMs.

Finally, SLRM approaches including a spatial smooth to take account of spatial autocorrelation showed the lowest power of all the tested models with appropriate false-positive rates, where the sophisticated methods for dummy point selection (\textit{Meth\_SLRM\_amt\_ac}) performed better than the model with the simple MCP procedure (\textit{Meth\_SLRM\_mcp\_ac}), as expected. Notably however, the total differences in average statistical power among the three main methods were relatively small: PPM approaches differed by $< 5\%$ from (i)SSM approaches, and the difference between (i)SSM and SLRM approaches was $< 10\%$.

### 3.1.2 Dummy point numbers

To investigate the influence of dummy point numbers on model performance, we compared the statistical powers and false-positive rates among four well-performing representative models (\textit{Meth\_PPM\_2}, \textit{Meth\_SSM\_1}, \textit{Meth\_iSSM\_1} and \textit{Meth\_SLRM\_amt\_ac}) using three different levels of dummy points ($N_{dummy} \in \{8, 80, 224\}$) per tracking point (Fig. 3). For all the compared methods and dummy point numbers, type I error rates were within or below the acceptable nominal range of 5\%. With respect to the statistical power, only the PPM approach
showed a drastic decrease if only eight dummy points were used, however, there was no significant difference for 80 vs. 224. Although all the other models showed similar qualitative behaviours, SSMs, iSSMs, and SLRMs had much higher robustness with respect to \( N_{\text{Dummy}} \).

### 3.1.3 Interplay between method and tracking-data properties

To analyse the interplay between the methods and variables related to environmental, movement, and tracking-device properties (c.f., Tab. 1), we first restricted the final data frame for the power evaluation to the four above-mentioned representative methods (based on 80 dummy points per tracking point) (c.f., previous subsection). As described in more detail in subsection 2.5, we then applied a LASSO-based model selection in combination with cross-validation to the data frame to select a set of promising predictors, which were further analysed in a third step using appropriate GEEs. We were interested in the factors influencing the statistical power of the different methods in a similar way. Thus, in the first GEE, we only investigated the main effects of the variables presented in Tab. 1. In contrast, the main effects were augmented in the second GEE with interaction terms between model-related (Tab. 1 below the double line) and non-model-related (Tab. 1 above the double line) variables. In particular, the interaction terms represented conditions in which the different methods performed differently. In Tab. 2, all significant terms (to the level of \( \alpha = 0.05 \)) in the first main effect GEE (above) and the significant interaction terms in the second GEE (below) are presented.

| Parameter name | Estimate | SE  | p    |
|----------------|----------|-----|------|
| Meth\_PPM\_2  | -0.2130  | 0.0716 | 0.0029 |
| Meth\_SLRM\_amt\_ac | -0.4249  | 0.0688 | 0.0000 |
| Mov\_\omega  | 4.0999  | 0.4601 | 0.0000 |
| Hab\_auto  | -0.0236  | 0.0068 | 0.0005 |
| Dev\_bias  | -0.4458  | 0.1279 | 0.0005 |
| Hab\_smooth  | 0.2925  | 0.1417 | 0.0390 |
| Hab\_type  | -1.0178  | 0.2536 | 0.0300 |
| Mov\_ran2  | 0.8928  | 0.4333 | 0.0394 |
| Mov\_ran3  | 0.9753  | 0.3891 | 0.0122 |
| Dev\_step  | -0.1114  | 0.0417 | 0.0076 |
| Meth\_modelSLRM\_amt\_ac\_Hab\_auto  | -0.0100  | 0.0044 | 0.0249 |
| Meth\_modelPPM\_2\_Hab\_type  | -0.6554  | 0.1592 | 0.0000 |
| Meth\_modelPPM\_2\_Hab\_anis  | 0.7750  | 0.2582 | 0.0027 |
| Meth\_modelSLRM\_amt\_ac\_Hab\_anis  | 1.1949  | 0.2789 | 0.0000 |

Table 2 Significant main effects (above the double line) and interaction terms (below the double line) driving the statistical power during habitat selection studies. Results are based on GEE analyses in combination with LASSO-based model selection.

In accordance with subsection 3.1.1, the main effects in Tab. 2 revealed that the PPM and more notably the SLRM showed reduced power on average, compared with the (i)SSM (the latter served as the baseline category in the model-related predictor variable \( \text{model} \)). Furthermore, the strength of local habitat attraction (\( \text{Mov}\_\omega \)) and decrease in directional persistence and movement velocity in appropriate habitats (\( \text{Mov}\_\text{ran2} \) respectively \( \text{Mov}\_\text{ran3} \)) positively influenced the power. This was expected, given that all three processes resulted in an increased proportion of tracking points (relative to dummy points) in preferred habitats. Less expectedly, gradual continuous habitat changes appeared to have a positive influence on the...
statistical power: the average power was lower in categorical (Hab_type) than in continuous habitats, and blurry boundaries (Hab_smooth) between different habitat types resulted in a higher power than sharp transitions. Generally, this observation suggests that both categorical habitats as well as sharp boundaries provide less local information on animal–habitat interactions than continuous and spatially graded variables, affecting the power of the statistical method finally applied to these variables. Finally, increasingly strong habitat autocorrelation (Hab_auto) or location error (Dev_bias) and large average step sizes (Dev_step) decreased the power, all of which were intuitively expected.

The interaction terms in Tab. 2 revealed that each method had specific weaknesses with respect to habitat properties. (i)SSM methods coped poorly with strong habitat anisotropy, since the statistical powers of PPM and SLRM approaches decreased significantly less strongly with increasing anisotropy compared to the baseline level 'iSSM'. In contrast, the power of PPMs decreased strongly for categorical habitats, whereas the decrease for SLRMs was strongest for autocorrelated habitats.

In accordance with the above observations, if we restricted the power-related data frame during GEE analysis to continuous habitats with high anisotropy (Hab_anis > 0.5), i.e., optimal for PPM performance, PPMs had significantly higher power on average (70 %) compared with SLRMs (63 %). If we restricted the habitats to those with moderate autocorrelation (Hab_auto < 30) and high anisotropy (Hab_anis > 0.5), i.e., optimal for SLRMs, the latter showed a non-significant (by 1 %) higher statistical power than (i)SSMs. The issue of which method shows the highest statistical power may thus depend on the properties of the habitats of interest, and the optimal method could be chosen accordingly.

3.2 Long-range attraction

3.2.1 Average performance

The GEE results with respect to the average performance of the different methods in the context of inferring large-scale attraction are shown in Fig. 4. In particular, models ending in '.ang' evaluated the angular bias towards the attraction centre, whereas those ending in '.dist' considered the Euclidean distance to the centre instead. Overall, the statistical powers and false-positive rates of (i)SSMs and PPMs were similar to the analyses of local habitat selection (c.f., subsection 3.1.1). (i)SSM approaches showed rather robust behaviour with statistical power and type I error rates similar to most of the PPM approaches. Furthermore, only PPMs using a quasi-Poisson instead of a Poisson distribution (models Meth_PPM_2qp...) showed upward biased type I error rates, and PPMs without empirical determination of the optimal extent of the dummy point grid (models Meth_PPM_1...) showed strongly reduced power.

Interestingly, in PPMs, the above-mentioned angular approach for detecting long-range attraction invariably performed better than evaluating the Euclidean distance, whereas the opposite was true in (i)SSMs. Overall, (i)SSM methods using the distance-based approach showed the highest statistical power. However, similar
to analyses of habitat selection, absolute differences in power between the different appropriate PPM and (i)SSM approaches were mostly moderate (< 10 %).

In contrast to the habitat selection studies, all SLRM-based approaches performed badly. SLRMs without consideration of spatial autocorrelation showed strongly upward biased false-positive rates, whereas SLRMs incorporating these spatial correlations had drastically reduced statistical power.

### 3.2.2 Dummy point numbers

Similar to the habitat selection analyses (c.f., subsection 3.1.2), we also assessed the required number of dummy points in the context of large-scale attraction effects based on a selection of representative models. The corresponding results (Fig. 5) were highly comparable to those for the habitat selection studies (c.f., Fig. 3). We observed a robust dependency of (i)SSM performance with respect to all considered dummy point numbers. Combining PPMs with the angular approach to detect large-scale bias, a low number ($N_{\text{Dummy}} = 8$) of dummy points led to a distinct drop in statistical power compared with $N_{\text{Dummy}} \in \{80, 224\}$, even stronger than for resource selection PPMs (Fig. 3). However, this effect was not observed when evaluating the Euclidean distance. Finally, SLRMs performed approximately equally badly for all considered dummy point numbers.

### 3.2.3 Interplay between method and tracking data properties

Similar to the resource selection analysis (c.f., subsection 3.1.3), the interplay between statistical methods and tracking data properties was evaluated by combining a LASSO-based approach with final GEE evaluation, separately evaluated in a model using only main effect predictors on one hand, and in a second model including various interaction terms on the other. However, variables related to habitat properties were omitted, given that the focus was on the detection of a large-scale attraction effect possibly interplaying with movement and tracking device properties. Despite the model-related terms, which were in good accordance with subsection 3.2.1 (the model $\text{Meth}_{\text{PPM}_2 \text{ang}}$ served as the baseline level), the variable $\text{Mov}_\sigma$ was selected as the only additional significant predictor in the main effect GEE (c.f., Tab. 3), demonstrating that the statistical power with respect to large-scale attraction did not sensitively depend on animal movement or tracking-device properties, but only on the strength of the attraction itself. Furthermore, in the GEE including interaction terms, no significant interaction occurred, suggesting that this robustness held for all considered methods equally.

|                          | Estimate | SE    | p     |
|--------------------------|----------|-------|-------|
| $\text{Meth}_{\text{SSM}_1 \text{dist}}$ | 0.7074   | 0.1331| 0.0000|
| $\text{Meth}_{\text{PPM}_2 \text{ang}}$ | 0.4982   | 0.1176| 0.0000|
| $\text{Meth}_{\text{SLRM}_{\text{amt ac}}}$ | -3.4786  | 0.3029| 0.0000|
| $\text{Meth}_{\text{SSM}_1 \text{dist}}$ | 0.5869   | 0.1250| 0.0000|
| $\text{Mov}_\sigma$     | 63.1623  | 5.5495| 0.0000|

Table 3. Significant main effects driving the statistical power during large-scale attraction studies. Results are based on GEE analyses in combination with LASSO-based model selection.
3.3 Tracking data size and computation time

All the above analyses were performed based on relatively small tracking data sizes of $N \approx 150$ tracking points. However, both the performance of the different methods and their computation times may depend critically on $N$. We thus investigated the influences of tracking data size $N \in [33, 300]$ on model performance and computation time, restricting the analysis again to the four representative and well-performing models Meth\_PPM,2, Meth\_SSM,1,Meth\_iSSM,1 and Meth\_SLRM\_amt_ac. GEE analysis using the model variable and $N$ as predictors revealed that there were no significant differences in false-positive rates between the PPM (baseline level) and the other models, and no significant dependency on $N$. In contrast, and in accordance with the results presented above, GEE results revealed that iSSMs and SSMs deviated positively (but non-significantly) from the PPM in terms of statistical power, whereas the SLRM showed a significant reduction in power. Additionally, as expected, the positive dependency of statistical power on $N$ was highly significant. Adding the interaction term "model : $N$" to the model did not result in any significant interactions, suggesting that the $N$-dependent increase in power was of the same order of magnitude in all models (c.f., Tab. 4).

|          | Estimate | SE  | p    |
|----------|----------|-----|------|
| Meth\_SSM,1     | 0.2354   | 0.2475 | 0.3416 |
| Meth\_SSM,1     | -0.0329  | 0.3651 | 0.9282 |
| Meth\_SLRM\_amt_ac | -0.5161  | 0.3932 | 0.1894 |
| $N$              | 0.0184   | 0.0043 | 0.0000 |
| Meth\_iSSM,1:N  | -0.0007  | 0.0014 | 0.6319 |
| Meth\_SSM,1:N   | 0.0024   | 0.0034 | 0.4711 |
| Meth\_SLRM\_amt_ac:N | -0.0003  | 0.0027 | 0.9098 |

Table 4 GEE results analyzing the statistical power depending on the different methods in interplay with the tracking data size $N$. Here, PPM\_2 served as the baseline level in the variable specifying the statistical method.

With respect to the computation time (referring to non-parallelized computations on a desktop computer), GAMs separately applied for each of the four above-selected methods revealed that the time increased approximately linearly with the number of tracking points for all models (c.f., Fig 6). Here, iSSMs and SSMs were condensed because of the strong similarity in results. PPMs showed distinctly longer computation times than the other methods.

4 Discussion

Our simulation study demonstrated that the spatio-temporal (i)SSM and PPM approaches performed better than the pure spatial SLRMs, with respect to bias, statistical power, and robustness. In particular, regarding the analyses of large-scale attraction effects, SLRMs appear to be strongly biased and thus inappropriate. This result is not unexpected, given that the large-scale attraction process takes place on the spatial scale of the animal track, whereas contrasting dummy points are selected on a small scale. Thus, if considered as a purely spatial problem, dummy points show similar distances to the attraction centre on average, compared with the tracking points, and contrasting both therefore does not lead to reliable results. Relative differences between tracking and dummy points would thus only lead to meaningful results if the data were evaluated in a time-dependent manner, i.e., by
locally grouping tracking points with dummy points (the latter centred around the foregoing tracking point) by their associated time point. In the habitat selection case, SLRMs only perform acceptably if spatial autocorrelation is appropriately considered and sophisticated dummy point selection methods are used. But even here, SLRMs coped particularly poorly with strong habitat autocorrelation, which could be caused by spatial confounding, given that autocorrelation in SLRMs is not introduced on the scale of single steps but as a 2D smooth at the scale of the trip. This could match the scale of habitat autocorrelation, finally leading to spatial confounding [36]. In summary, we therefore strongly recommend the use of (i)SSM or PPM approaches (e.g., as presented in [40, 43, 44]) to avoid biased results, especially if the tracking data are spatio-temporally correlated.

With respect to the performances of PPM and (i)SSM approaches, there was only a small difference in their statistical powers (if a few points with respect to PPM implementation were taken into account; c.f., below), regardless of whether local habitat selection or large-scale attraction/avoidance effects are investigated. Notably, (i)SSMs generally had slightly higher power than PPMs, but if the habitats of interest are strongly anisotropic, the power of PPMs during habitat selection studies can exceed that of (i)SSMs. However, to maximize the power of PPMs and (i)SSMs, several factors need to be considered. Compared with the other methods, PPMs appeared to require a higher minimum number of dummy points in order to maximize their potential statistical power. Hence, sufficient dummy points ($\geq 80$) per tracking point should be used. This is probably because dummy points in PPMs are used to approximate a spatio-temporal integral [40, 27], which process may be strongly biased for low dummy point numbers, especially if the integral is complex. Additionally, the PPM power depends critically on an optimal spatial extension of the dummy point grid, as determined empirically in this study. Furthermore, if large-scale attraction/avoidance is evaluated, (i)SSMs should include distance to the structure of interest as a predictor, whereas PPMs should use the angular bias to this structure.

Despite these small differences in statistical power between PPMs and (i)SSMs, we recommend in general using (i)SSMs over PPMs for the following reasons: (1) the statistical power of (i)SSMs is more robust with respect to dummy point numbers compared with PPMs; (2) (i)SSMs do not need the initial empirical determination of critical measures (the latter becomes increasingly complex if more than one predictor is considered); (3) computation times for (i)SSMs are much shorter than for PPMs, which is especially important in the context of large data sets; and finally (4) (i)SSM model implementation is user-friendly and well-documented because of the provided R-package amt, in contrast to spatio-temporal PPMs for which there is currently (to the best of our knowledge) no available R-package.

The above mentioned difference in computation time between PPMs and (i)SSMs was probably on the one hand because the initial empirical procedure to determine the optimal spatial extent of the dummy point grid required several time-consuming model fits. On the other hand, our PPM implementation may have been
programmed in a less time-efficient manner compared with the SSM and (i)SSM codes provided within the well-established R-package. Further advances in non-empirical methods for determining optimal PPM dummy point grid extension, as well as a more time-efficient code for PPMs could therefore strongly reduce the described differences between PPMs and the other approaches. Indeed, time-efficient methods for estimating PPM integrals have recently been discussed by Hooten et al. [8]. Interestingly, overall, not pure spatial SLRMs but (i)SSMs showed the lowest computational costs, possibly because of the computationally costly estimation of the spatial smooth (including a cross-validation procedure [49]) in SLRMs.

Despite the general rather practical motivated recommendation using (i)SSMs over PPMs, there could be situations where PPMs could be nevertheless preferred: first, as mentioned above, if habitats are strongly anisotropic (e.g., associated with thin borders, such as coast lines) PPMs may show a higher statistical power. Second, if the focus of the study is on population behaviour rather than on the analysis of single animals, PPMs may lead to more reliable population-level estimates compared to (i)SSMs due to the Eulerian approach adopted in PPMs (in contrast to the Lagrangian viewpoint in (i)SSMs) [42].

Finally, all comparisons performed within this study only demonstrated virtual differences between SSM and (i)SSM approaches, and we were therefore unable to make any clear recommendations. However, Avgar et al. noted additional advantages of iSSMs over SSMs [43], including their predictive capacity (e.g., for landscapes different from the landscape used for the model fit), and the ability to derive and parametrize a mechanistic movement model. We therefore finally recommend the use of iSSMs.

5 Conclusions

To conclude, our work provides an extensive simulation study comparing the statistical powers and false-positive rates of different statistical methods frequently used to infer local habitat selection or large-scale attraction/avoidance from animal tracking data. We compared several variants of spatio-temporal PPMs, (i)SSMs, and pure spatial SLRMs, and evaluated their performances in interplay with a broad range of simulated habitat, movement, and tracking-device properties. Our results suggest that SLRMs are inappropriate for detecting large-scale attraction/avoidance effects, and only produce unbiased habitat selection inference if spatial autocorrelation and dummy point choice are considered with care. In contrast, PPMs and (i)SSMs can produce robust and unbiased results for both habitat selection and large-scale attraction studies, with only slight differences in their average statistical powers. For habitat selection studies, (i)SSMs can cope better with categorical habitats than PPMs, whereas PPMs have higher statistical power than (i)SSMs in strongly anisotropic habitats. When large-scale attraction/avoidance is investigated using PPMs, the angular deviation from the straight connection to the structure of interest should be evaluated, whereas the Euclidean distance to this structure performs better in (i)SSMs. In general, we recommend the use of (i)SSMs over PPMs,
mainly because of their more user-friendly implementation, greater robustness (e.g., with respect to dummy point numbers), and faster computation times. In particular, (i)SSMs (the recent extension of SSMs) provide the additional advantage of an increased predictive capacity in combination with the derivation of a parametrized mechanistic movement model.

**List of abbreviations**

A list of abbreviations used in this study is given in Tabl 5.

| Abbreviation | Explanation |
|--------------|-------------|
| SLRM         | 'spatial logistic regression model'; pure spatial approach for analysing tracking data |
| PPM          | 'point process model'; spatio-temporal approach for analysing tracking data, developed from an Eulerian point of view |
| SMM          | 'step selection model' (also called 'step selection function /SSF' or 'step selection analysis / SSA'); spatio-temporal approach for analysing tracking data, developed from a Lagrangian point of view |
| (i)SMM       | 'integrated step selection model'; integrating the SSM approach with a mechanistic movement model |
| MCP          | 'minimal convex polygon'; minimal convex polygon containing all tracking points, for estimation of the minimum home range |
| GLM          | 'generalized linear model'; extension of linear regression models for non-normally distributed data |
| GAM          | 'generalized additive model'; extension of GLM to describe non-linear ('additive') dependencies |
| GEE          | generalized estimation equation'; estimation method for parameters of a GLM with possible unknown correlation between outcomes |
| GLMM         | 'generalized linear mixed model'; extension of GLM to describe outcome correlation with random effects |
| LASSO        | 'least absolute shrinkage and selection operator '; technique for model selection that can especially cope with a high number of possible predictors |

**Table 5** List of abbreviations used in this study

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**

All data are available within this publication and the supporting material.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

Support for the statistical framework was obtained from the Federal Ministry for Economic Affairs and Energy according to a decision of the German Bundestag (project HELBIRD, 0325751).
Authors’ contributions

MM developed the statistical methods, programmed the R-code, and performed most of the statistical tests and graphical analyses. PS, VP, and LE discussed and developed the required features of the analysis, provided ecological input and research questions, and participated in study design and result validation. PS and SG conceived the study and participated in its coordination. All authors helped to draft the manuscript and read and approved the final manuscript.

Acknowledgements

Not applicable.

Figure titles and legends

**Figure 1:** Sketch of the approach applied to compare the performances of different statistical methods in interplay with simulated habitat, animal movement, and tracking-device properties. In particular, local habitat selection and large-scale attraction were evaluated separately, and each was further split into the evaluation of statistical power (using movement data with underlying attraction effect) and false-positive rate (using movement data without underlying attraction effects).

**Figure 2:** GEE results showing the comparative performances of different models inferring habitat selection. The average statistical power is shown in green and the false-positive rate in red. Eighty dummy points per tracking point were used in all cases. Bars represent 95% confidence intervals, points correspond to regression coefficient estimates.

**Figure 3:** GEE results showing four representative models inferring habitat selection with a focus on the dependency of model performance on dummy point number (the latter represented as circles vs. triangles vs. squares). The average statistical power is shown in green and the false-positive rate in red. Bars represent 95% confidence intervals, symbols correspond to regression coefficient estimates.

**Figure 4:** GEE results comparing the performances of different models inferring long-range attraction. The average statistical power is shown in green and the false-positive rate in red. Eighty dummy points per tracking point were used in all cases. Bars represent 95% confidence intervals, symbols correspond to regression coefficient estimates. Circles indicate that long-range attraction was measured by the angular bias towards the attraction centre, and triangles indicate use of the Euclidean distance.

**Figure 5:** GEE results showing four representative models inferring long-range attraction with a focus on the dependency of model performance on dummy point number (the latter represented as circles vs. triangles vs. squares). The average statistical power is shown in green and the false-positive rate in red. Bars represent 95% confidence intervals, symbols correspond to regression coefficient estimates.

**Figure 6** Non-parallel computation times for dummy point generation and model fit for three representative models, in relation to the number of tracking points.
Shaded areas correspond to 95% confidence bands.

**Figure S1**: Examples of simulated categorical habitat data permutated over the variables $Hab_{auto}$ (strength of spatial autocorrelation), $Hab_{anis}$ (strength of autocorrelation anisotropy), and $Hab_{smooth}$ (blurry vs. sharp transition between habitat boundaries).

**Figure S2**: Examples of simulated continuous habitat data permutated over the variables $Mov_{\omega}$ (habitat selection strength), $Mov_{\sigma}$ (strength of bias towards the attraction centre), and $Mov_{ran}$ (strength of directional persistence). Blue point represents location of the attraction centre.
Author details

1 BIONUM statistical consulting office, Hamburg, Finkenwerder Norderdeich 15 A, Hamburg, Germany. 2 Research and Technology Centre (FTZ), Hafentörn 1, 25761 Büsum, Germany.

References

1. Elith, J., Leathwick, J.R.: Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40, 677–697 (2012)
2. Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M.: Mixed Effect Models and Extensions in Ecology with R. Springer, New York, USA (2009)
3. Zuur, A.F.: A Beginner’s Guide to Generalized Additive Models with R. Highland Statistics Ltd., New York, USA (2012)
4. Korner-Nievergelt, F., Roth, T., von Felten, S., Guelat, J., Almasi, B., Korner-Nievergelt, P.: Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan. Elsevier, London, UK (2015)
5. Austin, D., McMillan, J., Bowen, W.: A three-stage algorithm for filtering erroneous argos satellite locations. Marine Mammal Science 19(2), 371–383 (2003)
6. Hebblewhite, M., Haydon, T.D.: Distinguishing technology from biology: a critical review of the use of gps telemetry data in ecology. Philosophical transactions R. Soc. B 365, 2303–2312 (2010)
7. White, G.C., Garrott, R.A.: Analysis of Wildlife Radio-tracking Data. Elsevier (London), Boca Raton, Florida, USA (2012)
8. Hooten, M.B., Johnson, D.S., McClintock, B.T., Morales, J.M.: Animal Movement - Statistical Models for Telemetry Data. CRC Press, Boca Raton, Florida, USA (2017)
9. Mendel, B., Schwemmer, P., Peschko, V., Müller, S., Schwemmer, H., Mercker, M., Garthe, S.: Operational offshore wind farms and associated ship traffic cause profound changes in distribution patterns of loons (gavia spp.). Journal of environmental management 231, 428–438 (2019)
10. Collett, M., Chitika, L., Collett, T.S.: Spatial memory in insect navigation. Current biology : CB 23, 789–800 (2013). doi:10.1016/j.cub.2013.07.020
11. Poulter, S., Hartley, T., Lever, C.: The neurobiology of mammalian navigation. Current biology : CB 28, 1023–1042 (2018). doi:10.1016/j.cub.2018.05.050
12. Phillips, J.B., Schmidt-Koenig, K., Muheim, R.: True navigation: sensory bases of gradient maps. Brown MF, Cook R. G., editors. Animal Spatial Cognition: Comparative, Neural & Computational Approaches: Comparative Cognition Society (2006)
13. Cheung, A., Zhang, S., Stricker, C., Srinivasan, M.V.: Animal navigation: the difficulty of moving in a straight line. Biological cybernetics 97(1), 47–61 (2007)
14. et al, S.A.: Animal Movement Across Scales. Oxford university press, Oxford, UK (2014)
15. Aarts, G., Fieberg, J., Matthiopolous, J.: Comparative interpretation of count, presence-absence and point methods for species distribution models. Methods in Ecology and Evolution 3, 177–187 (2012)
16. Guranie, E., Ovaskainen, O.: Characteristic spatial and temporal scales unify models of animal movement. The American naturalist 178, 113–123 (2011). doi:10.1086/660285
17. McClintock, B.T., King, R., Thomas, L., Matthiopolous, J., McConnell, B.J., Morales, J.M.: A general discrete-time modeling framework for animal movement using multistate random walks 82, 335–349 (2012). doi:10.1890/11-0326.1
18. McClintock, B.T., Johnson, D.S., Hooten, M.B., Ver Hoef, J.M., Morales, J.M.: When to be discrete: the importance of time formulation in understanding animal movement. Movement ecology 2, 21 (2014). doi:10.1186/s40462-014-0021-6
19. Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E., Fryxell, J.M.: Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology (2004)
20. Jonsen, I.D., Flemming, J.M., Myers, R.A.: Robust state–space modeling of animal movement data. Ecology 86(11), 2874–2880 (2005)
21. Langrock, R., King, R., Matthiopolous, J., Thomas, L., Fortin, D., Morales, J.M.: Flexible and practical modeling of animal telemetry data: hidden markov models and extensions. Ecology 93, 2336–2342 (2012). doi:10.1890/11-2241.1
22. Pearce, J.L., Boyce, M.S.: Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43, 405–412 (2006)
23. Barbet-Massin, B., Jiguet, F., Albert, C.H., Thuiller, W.: Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3, 327–338 (2012)
24. S.J., P., Dudik, M.: Modeling of species distributions with maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175 (2008)
25. Elith, J., Phillips, S.J., Hastie, T., Dudik, M., Chee, Y.E., Yates, C.J.: A statistical explanation of maxent for ecologists. Diversity and Distributions 17, 43–57 (2011)
26. Warton, D.I., Shepherd, L.C.: Poisson point process models solve the “pseudo-absence problem” for presence-only data. The Annals of Applied Statistics 4(3), 1383–1402 (2010)
27. Renner, I.W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S.J., Popovic, G., Warton, D.I.: Poin process models for presence-only analysis. Methods on Ecology and Evolution 6, 366–379 (2015)
28. VanDerWal, J., Shoo, L.P., Graham, C., Williams, S.E.: Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know. Ecological Modelling (2009)
29. Fieberg, J., Matthiopolous, J., Hebblewhite, M., Boyce, M.S., Frair, J.L.: Correlation and studies of habitat selection: problem, red herring or opportunity? Philosophical transactions of the Royal Society B 365, 2233–2244 (2010)
30. Fleming, C.H., Fagan, W.F., Mueller, T., Olson, K.A., Leimgruber, P., Calabrese, J.M.: Estimating where and how animals travel: An optimal framework for path reconstruction from autocorrelated tracking data. Ecology 97(3), 576–82 (2016)
837 31. Hoeting, J.A.: The importance of accounting for spatial and temporal correlation in analyses of ecological data. Ecological Applications 19(3), 574–577 (2009)
838 32. F. Dormann, C., McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., Davies, R., Hirzel, A., 839 40. J. Nets, W., Daniel Kissling, W., et al.: Methods to account for spatial autocorrelation in the analysis of species 840 distributional data: a review. Ecography 30(5), 609–628 (2007)
841 33. Hodges, J.S., Reich, B.J.: Adding spatially correlated errors can mess up the fixed effect you love. The American Statistician 64, 325–334 (2011)
842 34. Hanks, E.M., Schliep, E.M., Hooten, M.B., Hoeting, J.A.: Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification: Spatial confounding in geostatistical models. Environmetrics 26(4), 10–1002331 (2015)
843 35. Hefley, H.B.E.M.R.R.E. T.J., Walsh, D.P.: Dynamic spatio-temporal models for spatial data. Spatial statistics 20, 206–220 (2017)
844 36. Paciorek, C.J.: The importance of scale for spatial-confounding bias and precision of spatial regression estimators. Statist. Sci. 25(1), 107–125 (2010)
845 37. Warton, D., Aarts, G.: Advancing our thinking in presence-only and used-available analysis. Journal of Animal Ecology 82, 1125–1134 (2013)
846 38. Aarts, G., Fieberg, J., Matthiopoulos, J.: Comparative interpretation of count, presence-absence and point methods for species distribution models. Methods in Ecology and Evolution 3, 177–187 (2012)
847 39. Johnson, D.S., Hooten, M.B., Kuhn, C.E.: Estimating animal resource selection from telemetry data. Biometrics 64, 968–976 (2008)
848 40. Johnson, D.S., Hooten, M.B., Kuhn, C.E.: Estimating animal resource selection from telemetry data using point process models. Journal of Animal Ecology 82, 1155–1164 (2013)
849 41. Thurfjell, H., Ciuti, S., Boyce, M.S.: Applications of step-selection functions in ecology and conservation 2, 4 (2014). doi:10.1186/2051-3933-2-4
850 42. Michelot, T., Blackwell, P.G., Matthiopoulos, J.: Linking resource selection and step selection models for habitat preferences in animals. Ecology 100, 02452 (2019). doi:10.1002/ecy.2452
851 43. Avgar, T., Potts, J.R., Lewis, M.A., Boyce, M.S.: Integrated step selection analysis: bridging the gap between resource selection and animal movement Methods in Ecology and Evolution 7, 619–630 (2016). doi:10.1111/2041-210x.12528
852 44. Avgar, T., Fieberg, J., Avgar, T.: Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. Ecology and evolution 9, 880–890 (2019). doi:10.1002/ece3.4823
853 45. Duchesne, T., Fortin, D., Rivest, L.-P.: Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. PloS one 10, 0122947 (2015). doi:10.1371/journal.pone.0122947
854 46. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 URL http://www.R-project.org/ (2016)
855 47. Hijmans, R.J., van Etten, J.: Raster: Geographic analysis and modeling with raster data. r package version 2.0-12. http://CRAN.R-project.org/package=raster (2012)
856 48. Wickham, H.: Ggplot2: Elegant Graphics for Data Analysis. Springer, New York, USA (2009)
857 49. Wood, S.: Generalized Additive Models: An Introduction with R. CRC press, Boca Raton, Florida, USA (2017)
858 50. Pebesma, R.S.B. E.J.: Classes and methods for spatial data in r. R News (2), 1–11 (2006)
859 51. Worton, B.J.: A convex hull-based estimator of home-range size. Biometrics, 1206–1215 (1995)
860 52. Burgman, M.A., Fox, J.C.: Bja: Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. In: Animal Conservation Forum, vol. 6, pp. 19–28 (2003). Cambridge University Press
861 53. Akaike, H.: Information theory and an extension of the maximum likelihood principle. International Symposium on Information Theory Second Edition, 267–281 (1973)
862 54. Weiser, C.: Methods for multivariate quadrature (r package version 1.0-6). URL http://CRAN.R-project.org/package=mvQuad (2016)
863 55. Hastie, T., Tibshirani, R.J.: Generalized Additive Models. Chapman and Hall, London, UK (1990)
864 56. Hardin, J., Hilbe, J.: Generalized Estimating Equations. Chapman and Hall/CRC, London, UK (2003)
865 57. Halekoh U, J.Y. Hojsgaard S: The r package geepack for generalized estimating equations. Journal of Statistical Software 2(2), 1–11 (2006)
866 58. Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S.: Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol 24(3), 127–135 (2009). doi:10.1016/j.tree.2008.10.009
867 59. Zuur, A.F., Saveliev, A.A., Ieno, E.N.: Zero Inflated Models and Generalized Linear Mixed Models Withh R. Highland Statistics Ltd., New York, USA (2012)
868 60. Tibshirani, R.: Regression shrinkage and selection via the lasso. J Roy Statist Soc B 58, 267–288 (1996)
869 61. Tibshirani, R.: The lasso method for variable selection in the cox model. Stat Med 16(4), 385–395 (1997)
870 62. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. Journal of statistical software 33, 1–22 (2010)
871 63. Usai, M.G., Goddard, M.E., Hayes, B.J.: Lasso with cross-validation for genomic selection. Genet Res (Camb). 91(6), 427–36 (2009)
872 64. Browne, M.W.: Cross-validation methods. Journal of mathematical psychology 44(1), 108–132 (2000)