Using GPS tracking to understand the impact of management interventions on visitor densities and bird populations

Pouwels, R., van Eupen, M., Walvoort, D. J. J., & Jochem, R.

Using GPS tracking to understand the impact of management interventions on visitor densities and bird populations

ROGIER POUWELS; rogier.pouwels@wur.nl

MICHIEL VAN EUPEN; michiel.vaneupen@wur.nl

DENNIS J.J. WALVOORT; dennis.walvoort@wur.nl

RENÉ JOCHEM; rene.jochem@wur.nl

Corresponding author:

Rogier Pouwels, Wageningen Environmental Research, WUR, P.O. Box 47, 6700 AA, Wageningen, The Netherlands

Tel.: +31 6 13297369; fax: +31 317 419 000

Summary

To manage the potential conflict between outdoor recreation and nature conservation, managers of nature areas need information to select effective interventions. For large nature areas information on visitor use is often lacking and managers often make decisions based on expert judgement. In this paper we use monitoring data gathered with GPS devices to develop a tool and derive rules of thumb managers can use to estimate the impact of management actions on visitor densities. Using a dataset of 1563 tracks from the New Forest, UK, we developed a random forest model and identified which landscape and environmental features account for the spatial variation in visitor densities. The random forest model shows that distance to car park, distance to roads and openness are the most important factors for predicting visitor densities. The model was used as a tool to assess the impact of potential management interventions on the population of Nightjar. As developing this type of tool requires a lot of data we also derived rules of thumb and a simple algorithm that managers of other nature areas can use to estimate the impact of their interventions on visitor densities. The derived rules of thumb show that changing the location of car parks in relation to tarmac roads can help managers to reduce local visitor densities by 80%. Further research in other nature areas should verify the feasibility of these rules of thumb and the simple algorithm.

Key words

Random Forest; management tool; rules of thumb; outdoor recreation; bird conservation

1 Introduction

In many nature areas the dual mandate to protect natural values and accommodate visitors is a source of potential conflicts (Reed and Merenlander 2008) because recreation can have a negative impact on biodiversity values (Larson et al. 2016). On the other hand, allowing recreation in protected areas is thought to be important to build societal support for conservation in general and local nature management in particular (Thompson 2015). Nature managers can take measures to mitigate undesired effects of recreation on nature values, but these measures might have consequences for societal support. Consequently, managers need to plan actions with care and involve stakeholders in their decision making (Sutherland et al. 2014, McCool 2016). They need adequate monitoring data on the temporal and spatial distribution of visitors to know where biodiversity values coincide with visitor use (Hadwen et al. 2007, Hammitt et al. 2015). However, such data are often lacking (Eagles 2014) as methods are time consuming and often expensive (Orsi and Geneletti 2013, Cessford and Muhar 2003). Besides information on the current situation, managers also need to know what options they have to change visitor densities and what impact their measures are likely to have on social or ecological disturbance thresholds (Sayan et al. 2013, Larson et al. 2018). They need to understand what features of the landscape and path network will determine the temporal and spatial distribution of visitors (Hammitt et al. 2015).

Visitor densities tend to be very heterogeneous in nature areas (Hammitt et al. 2015). Entrances and car parks act as gateways to an area (Beunen et al. 2008, Larson et al. 2018, Weitowitz et al. 2019). From these gateways visitors disperse using the path network (Meijles et al. 2014). Their distribution reflects the choices they make during their visit (Wolf et al. 2015). Research shows that different features influence visitor choices: specific attraction points, weather, physical features of the landscape, features of the path network, visitor preferences, the time they have available, the motives they have for visiting the area, the composition of the group and other visitors and users of the area (Arnberger and Haider 2007, Beeco and Brown 2013, Böcker et al. 2013, Hallo et al. 2012, Shoval 2010, Maldonado et al. 2011, Taczanowska et al. 2014, Torbidoni 2011, Van Marwijk et al. 2009, Schamel and Job 2017). As all these features will interact during a visit, it is difficult to identify which features account most for differences in visitor densities (Shoval et al. 2010) and which management actions will be effective in altering visitor distribution.

In recreation studies GPS devices are considered to be promising for gathering information on visitor densities and visitor behaviour (Beeco and Brown 2013). They provide accurate data on distribution, speed of movement and time spent at specific locations (D'Antonio et al. 2010, Beeco and Brown 2013). In recent years monitoring with GPS devices has also been used in combination with graph theory to evaluate the use of path structure (Taczanowska et al. 2014, 2017), in combination with recreation suitability mapping (Beeco et al. 2014), in combination with Public Participation GIS (Korpilo et al. 2017) and for spatial analyses of movement patterns (Van Marwijk and Pitt 2008, Renso et al. 2012) or developing simulation models of visitor flows (Gimblett, R. and Skov-Petersen 2008, Van Marwijk 2009, Taczanowska 2009). However, most studies using GPS devices for monitoring have focused on their utility for visual analyses and to find hotspots (Beeco et al. 2013). Few studies use monitoring information to understand what drives visitor densities in nature areas (Beeco et al. 2014). The exceptions are studies by Meijles et al. (2014), Olson et al. (2017) and Zhai et al. (2018). However, although both studies provide managers with information about which features determine visitor densities, this information might still lack relevance to managers. Managers not only need to know which features drive visitor densities, but also how visitor densities depend on these features, what the type of response curve is (Monz et al. 2013). This information would enable them to link potential management interventions, such as changing the features that drive visitor densities, to recognized values such as social and ecological thresholds.

In this study we aim to develop tools and rules of thumb that managers can use in decision-making processes with stakeholders to generate support for potential management interventions when visitor densities exceed social or ecological thresholds. For this support managers need to know how their interventions will lead to a change in visitor densities. We use monitoring data from GPS devices gathered in the New Forest (UK) to develop a random forest model (Breiman 2001) to identify which landscape and environmental features account for the spatial variation in visitor densities in the area. This model is then used as a tool to estimate visitor densities for the whole area. To illustrate its possible applications we use it to assess the impact of potential interventions on the population size of Nightjar (*Caprimulgus europaeus*), one of the protected species in the New Forest and sensitive to disturbance (Langston et al. 2007). As developing this type of tools needs much data and specialized expertise we also derived rules of thumb that managers can use to estimate the impact of management actions on visitor densities.

2 Study area

The New Forest is a large forest-heathland complex and Natura 2000 site in the United Kingdom. The area is around 57000 ha in size and was designated as a Natura 2000 site for 11 habitat types, two habitat directive species and seven bird species (JNCC 2015a, JNCC 2015b). It is a mosaic of woodland, heathlands, grasslands and mire systems and is managed by the Forestry Commission. Several hundred thousand residents live and work in small villages and medium-sized towns within the area or within a radius of 10 km. The New Forest is also a popular holiday destination all year round and is famous for its herds of horses (the New Forest pony) that roam the area. The area is easily accessible, with over 100 car parks from where visitors can use the dense network of over 2500 km walking trails (Fig. 1). An estimated 13.3 million people visit the area each year (Gallagher et al. 2007).

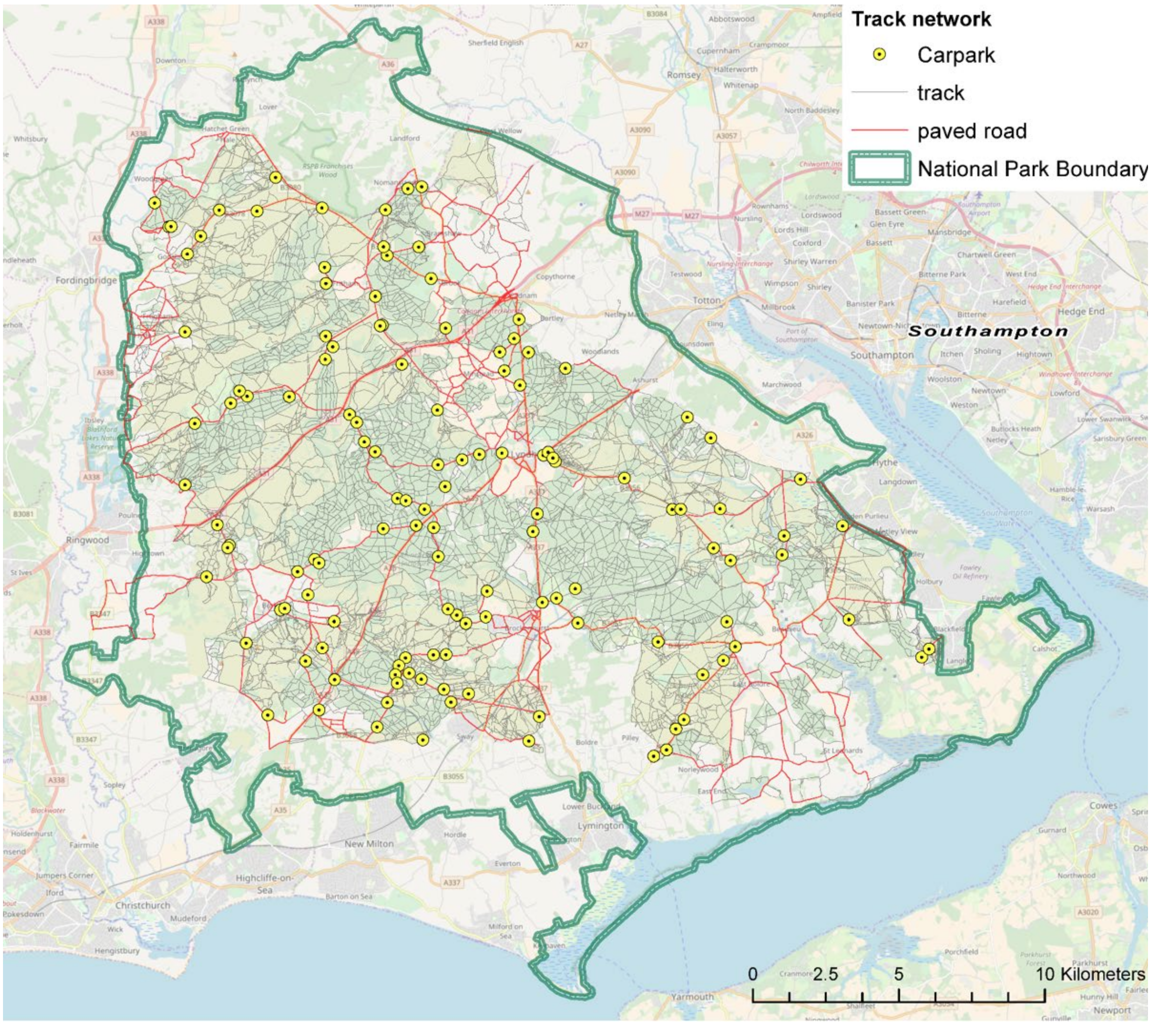

Fig. 1. The New Forest study area located west of the city of Southampton in the UK. Indicated are car parks, path network and roads.

3 Methods

Our method consists of six main steps (Fig. 2). First we collected information on visitor distribution using GPS devices. In this step the monitoring data from the GPS devices was prepared for further analyses. Second, we selected explanatory variables that describe the landscape and environment of the New Forest. In the third step we performed an exploratory data analysis to better understand the relationships between the different explanatory variables and characteristics of the routes visitors had followed. In the fourth step we developed a random forest model (Breiman 2001) to estimate the importance of the variables and their interaction in explaining the spatial variation in visitor densities. In the fifth step we used this model as a tool to predict visitor density distribution for the whole area. We illustrate the possible applications of the model by using it to assess the impact of three potential management interventions on the Nightjar population. In the sixth step we derived rules of thumb based on the results of the previous steps. The steps are explained in the next six sections.

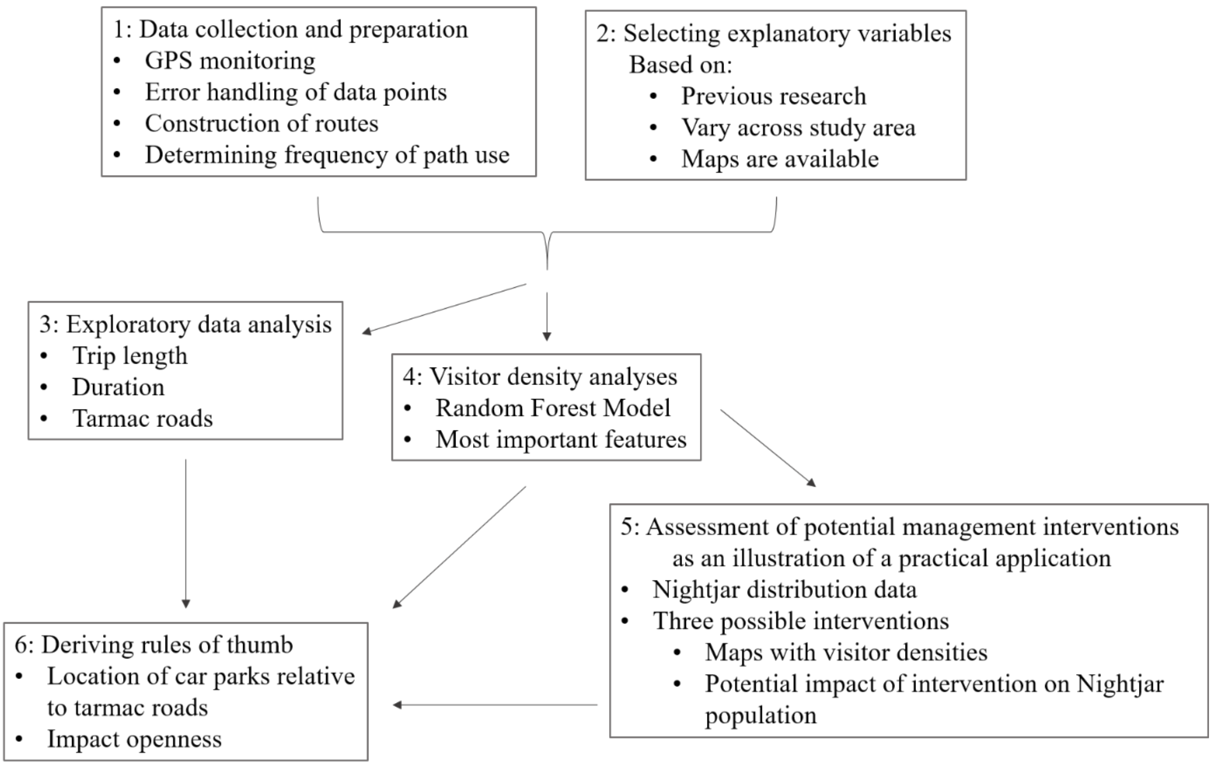

Fig. 2. Schematic overview of the method. In the first step the monitoring data was collected with GPS devices and prepared for further analyses. In the second step explanatory variables were selected that described the landscape and environment of the New Forest, followed by an exploratory data analysis in the third step. In the fourth step we developed a random forest model to estimate the importance of the explanatory variables in predicting the frequency with which visitors use specific segments of the path network. In the fifth step we used these predicted frequencies to estimate the impact of three potential management interventions on the Nightjar population. In the sixth step rules of thumb and simple algorithms were derived from the results of the previous steps.

3.1 Data collection and preparation

The monitoring data with GPS devices were collected on 80 mostly consecutive days during spring and summer in 2004 as part of the PROGRESS research project (Gallagher et al.

2007). Visitors arriving at car parks were asked to participate in the monitoring project. The GPS devices were stored in a plastic carrying case that could be clipped onto the rucksack or jacket of visitors who participated in the survey. Participants were instructed to keep the device with the built-in antenna upward and at an approximate height of 1.5m. Two models of GPS devices were used, the Garmin eTrex and Garmin eTrex Venture. The record method of the devices was set to auto to record tracks at a variable rate that creates an optimum representation of the track. On average each 51.2 meter a data point was logged. After participants returned, the data points were stored in a database using the Garmin transfer protocol. Additional information regarding the group size, number of dogs and use of a leash for the dog was added to the monitoring data by unique ID.

In total 1563 GPS tracks with 110 505 data points were collected at 41 car parks. The car parks were selected by the managers to represent differences in landscape setting and capacity and that are popular by both local visitors as well as visitors from further away. The monitoring frequency differed between car parks; some were monitored over 10 days, while others were only monitored once (Appendix A). As the number of GPS devices was limited, the proportion of visitors monitored at highly used car parks was probably lower than at less used car parks. As these proportions are unknown the dataset does not represent true visitor densities. Instead, we used the GPS tracks to construct the routes visitors were most likely to have followed (Appendix B). These routes are a set of consecutive path segments. Each path segment is delineated by two nodes. In the path network each crossing, junction and car park is represented as a node. For each segment of the path network the constructed routes were used to determine the frequency of use by visitors from a specific car park by dividing the number of constructed routes by the number of monitored visitors starting at that specific car park (see Table A.1 in Appendix A). Only car parks that had 10 or more routes in the database were taken into account, resulting in analyses based on 36 car parks (Appendix A).

3.2 Selecting explanatory variables

We selected several information sources that describe the landscape and environment of the New Forest. We applied three selection criteria: the data had to represent features that (1) were known from previous research to have an influence on visitor behaviour and visitor densities, (2) vary across the New Forest and (3) maps should be available that represent the situation around 2004 when the monitoring data was gathered. The features we selected were: car parks, path and road network, vegetation type, openness, slope and traffic noise (Table 1) and all maps were converted into a 10 x 10 m grid. This resolution was chosen to avoid information on different path segments being assigned to one cell. The information on the car parks was used to determine the distance of each point on the track to a car park, as it is known that visitor densities are higher near car parks (Meijles et al. 2014, Zhai et al. 2018). The path network was used to distinguish between different path types, as previous research suggested that visitors have different preferences for path surfacing (Beeco et al. 2014). The path network was also used to determine the distance to tarmac roads, as Henkens et al. (2006) showed that visitors avoid crossing tarmac roads and visitor densities might be lower near tarmac roads. The vegetation information was selected because vegetation types were found to determine the attractiveness of the landscape to visitors (De Vries et al. 2013), which may result in a spatial variation in visitor densities.

The openness of a landscape is considered one of the most important indicators of the visual landscape experience (Kaplan et al. 1989, Weitkamp 2011). In nature areas this openness strongly depends on the vegetation structure as perceived by visitors. Information representing landscape openness in the New Forest was not available. Instead, the Viewscape model (Jochem et al. 2016) was used to determine openness for each grid cell in the area. ViewScape calculates the visible area of the landscape from sightlines within a radius of 3 km (Jochem et al. 2016). Two openness features were used as explanatory variables: the total visible area and the variation in sightlines. Slope was selected as visitors avoid steep slopes (Beeco et al. 2014) and lower densities are expected at steeper slopes. Traffic noise was selected as visitors prefer tranquil areas and densities are expected to be lower in areas with high noise levels (Benfield et al. 2010).

Table 1. Selected explanatory variables used in the random forest model. The justification for the chosen variable is given in the third column (Source) together with the reference.

Variable	Variable determination	Source
Distance to car park	The distance to the car park was calculated as the crow flies.	Visitor densities are higher near car parks (Meijles et al. 2014, Zhai et al. 2018).
Path type	Five path types were distinguished, composed of a path network map showing nine path types: unclassified dirt tracks (72% of total length), gravel tracks (3%), tracks on lawns (6%), cycle paths (8%) and tarmac roads (11%). The map was provided by the Forestry Commission.	Visitors densities depend on path type as visitors have different preferences for types of path surface (Beeco et al. 2014) and path width (Zhai et al. 2018).
Distance to roads	Distance to the nearest tarmac road was calculated as the crow flies.	Visitors densities are expected to be lower near roads as visitors avoid crossing roads (Henkens et al. 2006).
Vegetation type	11 Groups of vegetation types were distinguished, composed of a vegetation type map containing 52 vegetation types. The vegetation map was provided by the Forestry Commission. Corine Land Cover map (EEA 2016a) was used to fill gaps in the vegetation map.	Visitor densities depend on vegetation types as they determine the attractiveness of landscapes, as perceived by visitors (De Vries et al. 2013) .
Openness:	Based on the vegetation map, two openness features were determined by the Viewscape model (Jochem et al. 2016):	Openness is an important factor for visitor preferences (Kaplan et al. 1989).
Total area	Total visible area; amount of area visible to a distance of 3 km.	
Variation	Standard deviation in the length of the 180° sightlines representing diversity in openness in the 360° view.	
Slope	The slope (in degrees) was based on the European Digital Elevation Model (EU-DEM) (EEA 2016b).	Visitors avoid steep slopes (Beeco et al. 2014).
Traffic noise	Traffic noise (in dB) was based on modelled noise levels for major traffic routes (DEFRA 2016). For the missing values a background noise of 35 dB was assumed (based on Pesonen 2000).	Visitors prefer tranquil areas and densities are expected to be lower in areas with high noise levels (Benfield et al. 2010).

3.3 Exploratory data analyses

The dataset on the derived routes was analysed for basic characteristics, such as total trip length, maximum distance from car park, average group size and presence of dogs. Information on the explanatory variables was added to explore the relationships between these variables, visitor densities and characteristics of the expected routes followed by visitors. In this step an extra GIS action was executed to add information on crossing tarmac roads to the derived routes. The path network was used to construct compartments of the New Forest that are surrounded by tarmac roads. When the derived routes occur in more than one compartment visitors are expected to have crossed at least one tarmac road or used it as part of their trip.

3.4 Visitor density analyses

A random forest model was constructed to estimate which landscape and environmental features account for spatial variation in visitor densities in the area. For the analysis, all covariates and the response were converted to a 10 x 10 m grid. We used the implementation by Wright and Ziegler (2017). Their implementation of random forests follows that of Breiman (2001) and is also suitable for large data sets. The frequency of use of a path by visitors from a specific car park was used as the response variable (‘y’ variable).

For practical reasons (data reduction to make the calculations feasible) only locations within 5 km of the car parks, as the crow flies, were taken into account. As almost 99% of the data points of the GPS tracks are within this distance from the car parks we think 5 km is a suitable value. The model we constructed consists of 500 regression trees, each of which is based on a bootstrap sample from the original data. Each bootstrap sample has the same size as the original data and was obtained by simple random sampling with replacement. This means that some records of the original data set occur more than once, and some never. Data that were not in the bootstrap sample were used for 'out-of-bag' validation. According to Breiman (2001, p.11), using out-of-bag data (about one-third of the data) removes the need for a set aside validation set. The importance of the explanatory variables used (see Table 1) was computed in three steps. First, the out-of-bag mean squared error was computed for each tree. Then, this statistic was also computed for each tree after permuting each predictor variable. Finally, the difference between the two mean squared errors was averaged over all trees (Liaw and Wiener 2002).

3.5 Assessment of potential management interventions as an illustration of a practical application

To illustrate how the data and tools could be applied to support decision making we designed three potential management interventions and estimated the visitor densities for the whole area. The visitor densities were used to assess the impact of the interventions on the Nightjar population by comparing it with the current situation. We chose management interventions that restrict visitors by temporary or permanent closures of car parks as these are one of the most commonly used methods of reducing visitor densities in sensitive parts of nature areas (Hammitt et al. 2015). The three possible interventions assessed are: 1) closure of small car parks, 2) closure of relatively isolated car parks that are located near areas with many Nightjars and 3) closure of all but 20 car parks to concentrate visitors near the border of the area or near villages, for example Lymington and Lyndhurst (see Appendix C for more details on the chosen interventions).

For this assessment we used the random forest model developed in step four together with the territories of Nightjar that were recorded during a full survey of the New Forest in 2004

(Newton 2010). The random forest model was used to predict relative visitor frequency on path segments for all car parks in the area. For each intervention the number of visitors at the car parks was used to calculate a map with the visitor densities per path per year. In order to calculate the expected impact of each intervention we used the dose–impact relationship between visitor densities and breeding pair density as described by Pouwels et al. (2017). This resulted in estimates for the total (potential) population size for the Nightjar in the New Forest area for each situation. The dose-impact relationship is based on a study in a similar landscape, the Veluwe (the Netherlands). The Veluwe is also a large forest-heathland complex and Natura 2000 site with a dense path network and high visitor numbers (See Appendix C for more details of the method used).

3.6 Deriving rules of thumb

The random forest model was used to determine the importance of the explanatory variables in explaining the spatial variation of the visitor densities in the New Forest. The relationship between the variables and visitor densities may be visualized in ‘partial dependence plots’. However, as these variables might be correlated and interact with one another, interpreting these visualizations can be complicated or even misleading (Molnar 2019). Nevertheless, managers need these relationships to estimate the impacts of interventions. Therefore, we combined the results from the data exploratory analyses for the most important variables selected by the random forest model and derived rules of thumb and simple algorithms. These rules of thumb and algorithms may be used by managers to estimate the impact of interventions on visitor densities and help them to gain support for these interventions in decision-making processes with stakeholders. We focused on interventions related to restricting visitor use by temporary or permanent closure of car parks or by changing the capacity and location of car parks.

4 Results

4.1 Exploratory data analysis

The additional information from the GPS tracks shows that 40% of the tracks represent a single visitor, 40% represent two visitors and 20% represent visitor groups with more than two people. The average number of people for each track was 2.0 visitors with a standard deviation of 1.5 visitors. Two thirds of the visitors walked their dog. Most of them were on their own and 23% had the dog off leash. This number might be biased due to the presence of the researchers at the car parks. The average trip length of visitors without dogs was 5.4 km and of visitors with dogs 3.2 km. More than half the visitors stayed within a radius of 1000 m of the car park (Fig. 3).

The data also show that 17.6% of visitors cross roads or use them during their visit. Combining this data with other variables shows that visitors without dogs in open landscapes avoid crossing roads more often than visitors in closed landscapes. The probability of crossing roads declines from 41% in the least open landscapes to 13% in the most open landscapes. For visitors with dogs the probability is less than 10% in all landscapes (Fig. 4).

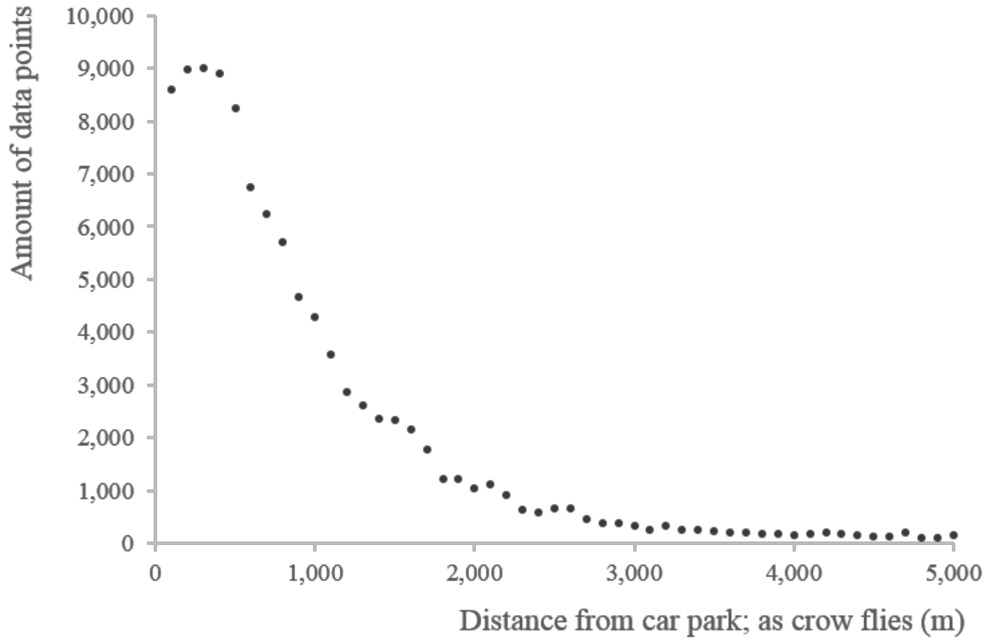

Fig. 3. Numbers of data points from GPS tracks after data handling (Appendix B) at specific distances from the car park. Just 1.6% of all data points are found at distances exceeding 5000 m and are not shown on the graph.

254

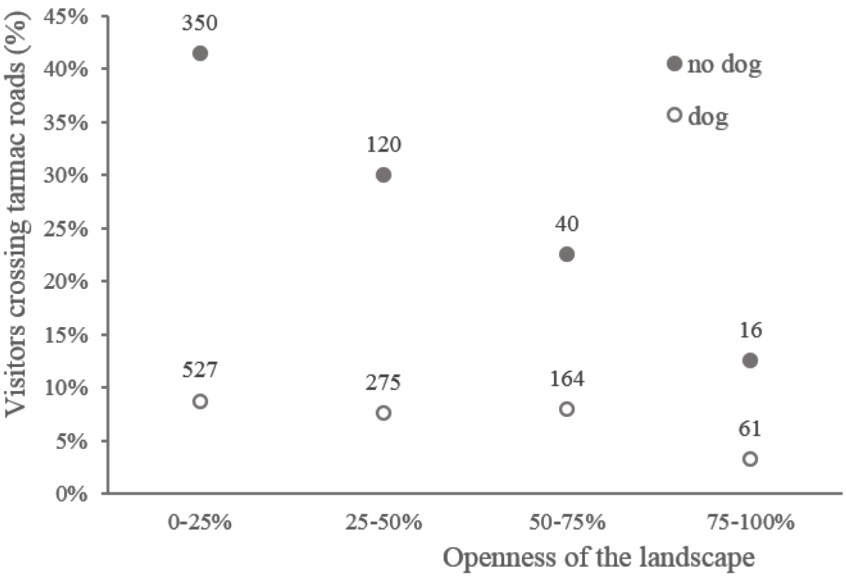

Fig. 4. Probability of visitors crossing tarmac roads in relation to average openness of the landscape along the route taken. The number of routes in each openness class is given. The average openness is defined as the percentage of area that is visible within a 3000 m radius during the entire visit.

4.2 Impact of landscape and environmental features on visitor densities

The fitted random forest model explains 74% of the variance in the data. The models show that besides distance to car park, distance to road, openness related variables, path type, slope and vegetation type are important factors for predicting visitor densities (Fig. 5). Traffic noise showed a very low importance in the first models and was removed from the final dataset, suggesting that the distance to roads is a better predictor of visitor densities than the level of traffic noise itself.

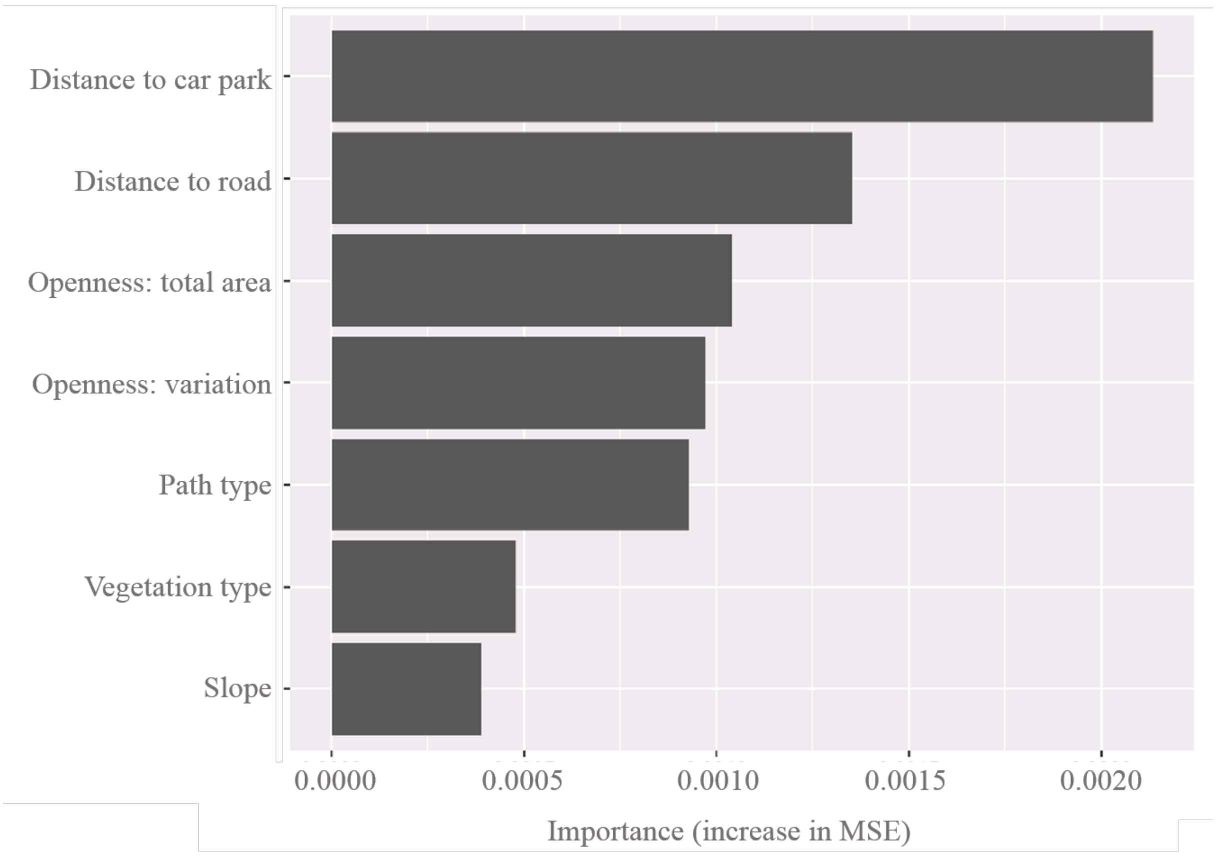

Fig. 5. Importance of variables in predicting visitor densities.

4.3 Impact of management actions on visitor densities and current distribution of Nightjar

The random forest model shows visitor densities in the New Forest varying between 0 and 300 000 visitor groups per ha per year (Fig. 6). The current population of Nightjar in the New Forest consists of 498 breeding pairs. The potential population size, without recreation in the area, is estimated to be 805 breeding pairs, implying that current recreational use lowers the population size by 38%. All three interventions lead to an increase in population size, but only the intervention in which all but 20 car parks are closed shows a large impact on the population size (Table 2). It should be noted that these results should be seen as an indication as they are based on the dose-impact relationship from the Veluwe and that the accuracy of its use for the New Forest has not been determined.

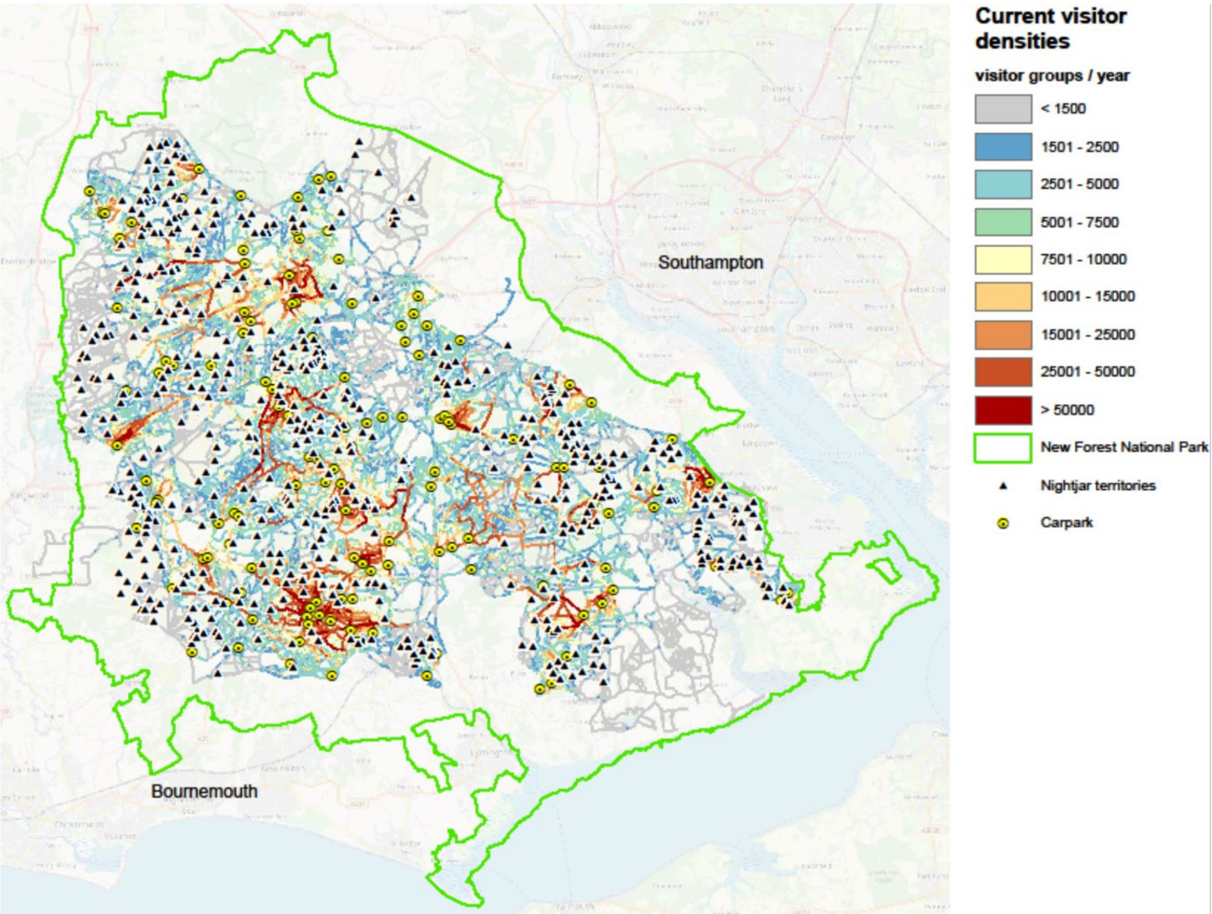

Fig. 6. Visitor density map on the New Forest based on the random forest model and the current location of car parks and Nightjar territories. The area containing the highest visitor densities is located at Wilverley in the south.

Table 2. Predicted impact of potential interventions to amend the spatial variation of visitor densities on Nightjar breeding pairs in the New Forest, UK. The figures represent the current and predicted number of Nightjar breeding pairs over four visitor density zones (column 1) for four situations: the current distribution of car parks (column 3) and three interventions (columns 4–6). Column 2 shows the predicted number of breeding pairs for a situation without recreation.

visitors (x 1000 per year per ha)	without recreation (predicted)	current (actual in 2004)	intervention 1: close small car parks (predicted)	intervention 2: close 3 car parks (predicted)	intervention 3: close all but 20 car parks (predicted)
0 – 10	805	157	177	199	578
10 – 25		156	141	144	64
25 – 50		95	90	88	35
> 50		90	92	85	28
total	805	498	500	515	705

4.4 Rules of thumb

To estimate the impact of management interventions on visitor densities we derived two rules of thumb and one simple algorithm for managers by combining the results from the data exploratory analyses for the three most important variables selected by the random forest model: distance to car park, distance to road and openness (total area). The first rule of thumb concerns the impact of tarmac roads on visitor densities: as on average 17.6% of all visitors cross roads or use them during their trip (see section 4.1) visitor densities are up to five times higher in areas on the same side of the road as the car park than on the opposite side of the road. Managers could use this rule of thumb to change visitor densities by relocating car parks. These interventions might reduce visitor densities by 80% in areas that are sensitive to disturbance without restricting visitor use completely. The presence of dogs might even be reduced by 90%. The second rule of thumb concerns the interaction between tarmac roads and openness. In woodlands the impact of tarmac roads on visitor densities is less distinct (around 78% reduction), while in open landscapes, like heathlands, the impact is larger (around 95% reduction). Combining both rules of thumb shows that managers might be able to reduce visitor densities by up to 95% by relocating a car park from one side of the road to another in open landscapes.

Results from the exploratory data analysis and the random forest model give a reliable estimate of how visitor densities decline with increasing distance from the car park. We used the frequency distribution of GPS locations, the data points (Fig. 3), to derive a simple algorithm that estimates the number of visitor groups at a specific path segment. First, we chose an algorithm that describes the sigmoid declining curve and fitted the parameters for the correlation of data points. This curve represents the probability that a visitor group is present at a specific distance (Eq. 1; Fig. 7). Next, we multiplied this by the number of visitor groups starting at a specific car park, taking into account that visitor groups will be present at a specific distance twice: when they enter the area and when they return to the car park. Finally, the number of visitor groups was divided by the number of paths segments at a specific distance class to account for a potential unevenness in path density over distance (Eq. 2). Managers can use Eq. 2 to acquire a first estimate of the number of visitors at a specific location (N_v). The parameters needed are quite easy to collect and are (1) the distance to the car park of interest, (2) the number of visitors that use the car park, and (3) the density of the path network around the car park. For locations within 5 km of more than one car park, the algorithm should be applied for each car park separately and the number of visitors per path segment should be summed.

$$f_{sdp} = 1 - \frac{d^{\alpha}}{H^{\alpha} + d^{\alpha}} \qquad \text{Equation 1}$$

f_{sdp}	fraction of data point at distance d
d	distance to car park (m)
H	parameter at which visitor presence is 50% (m); 965 in Fig. 7
α	parameter determining the rate at which visitor presence declines; 2.80 in Fig. 7

$$N_v = \left(1 - \frac{d^{2.80}}{965^{2.80} + d^{2.80}}\right) \times \frac{V_{cp} \times 2}{p_d} \qquad \text{Equation 2}$$

N_v	Predicted number of visitors present at a path segment at distance d (per day or per year)
d	distance to car park (m)
V_{cp}	number of visitors starting at a specific car park (per day or per year)
p_d	number of path (segments) at a specific distance class

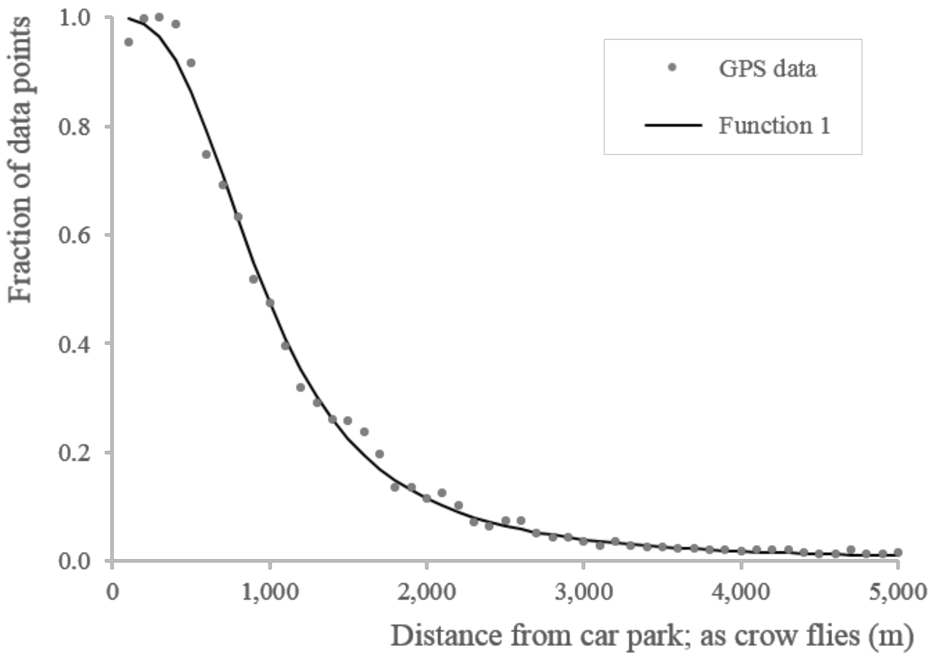

Fig. 7. Fraction of data points at a certain distance from a cark park. The parameter values of Eq. 1 are 2.80 for α and 965 m for H.

5 Discussion and implications for recreation management

5.1 Practical implications

In this paper we show that random forest models are suitable for modelling the complex interaction between different landscape and environmental features to explain visitor densities in nature areas. A random forest model was used as a tool to assess the impact of potential interventions on visitor densities and consequently on a population of a target species, the Nightjar, in the New Forest, UK. We focused on reallocating visitors, but interventions such as changes to path type or vegetation type could also be assessed. Although the GPS data only covered one third of all car parks, we believe that the data are representative of all car parks in the area and so the model predicts visitor densities for the whole area (Fig. 6). Random forest models based on GPS monitoring data are particularly useful in areas where managers need tools to estimate visitor densities and relate them to social or ecological thresholds. Managers could use these tools in decision-making processes with stakeholders to discuss and find support for potential interventions.

To discuss the effectiveness of interventions with stakeholders, managers need to know what measures are needed to lower visitor densities to certain desired levels. As the random forest model does not provide straightforward dose–effect relationships between a single variable and the visitor density, we derived two rules of thumb. Both use a simple algorithm to relate the location of car parks to visitor densities at specific distances from the car parks. Relocating car parks is effective as car parks act as gateways (Larson et al. 2018) and, as this present study has shown, their location accounts for much of the spatial variation of visitor densities in the area. These rules of thumb can be used by managers of nature areas who lack the resources or expertise to collect and analyse the type of data used in this study.

The distinctive downward curve of visitor densities corresponds with other distance decay curves (Yang and Diez-Roux 2012, Tratalos et al. 2013, Prins et al. 2014). The added value of our algorithm is that it is based on the data points of the GPS tracks, which is a better representation of visitor densities than a curve based on the maximum distance visitors walk (Tratalos et al. 2013, Prins et al. 2014). A distance decay curve based on the maximum distance visitors walk implies visitors walk in a straight line back and forth. Our dataset shows that not taking into account the shape of the route visitors follow will result in underestimating visitor densities by approximately 10% between 500 and 1500 m from car parks.

In the past decades other modelling approaches, like RBSim (Itami et al. 2003) and MASOOR (Jochem et al. 2008), have been developed to simulate the interaction between different landscape and environmental features and visitor flows in nature areas. Our analyses might contribute to these agent-based models as they indicate which features are important to incorporate. However, in these models the agents, the simulated visitors, navigate over a path network in pursuit of goals. During this navigation they make choices at crossings and junctions based on the local situation and the already covered route (Jochem et al. 2008). Although some of our results, like the probability to cross tarmac roads, seem useful to implement in these models the exact parameterisation is not straightforward. The probability of a visitor to cross a tarmac road during a visit is a combination of the probability to encounter a tarmac road and the probability to cross this road. Our result shows the combined effect of both probabilities while in the agent-based model only the probability to cross an encountered road should be implemented. Therefore, we propose to use the derived routes to determine which choices visitors make at crossings (i.e. Jochem et al. 2008) and determine the factors that affect visitor behaviour (i.e. Cavens et al. 2004, Taczanowska et al. 2008).

5.2 Generalization of the results

Our results show that visitor densities in the New Forest depend on the interaction between several features of the path network and landscape as well as on the accessibility of the area. That distance to car park is an important factor confirms the conclusion of Meijles et al (2014) and Zhai et al. (2018). Our finding that visitors avoid crossing tarmac roads confirms the conclusion of Henkens et al. (2006). In addition, the total visible area and the variation in 180° sightlines are important predictors of visitor densities. However, the model outcomes show that the correlation is complex and not easy to interpret. The New Forest is expected to attract visitors who prefer open areas like heathlands as well as visitors who prefer closed areas like ancient woodlands. The results of our study reflect these mixed preferences. Also, Heijman et al. (2011) showed that respondents preferred a mix of open and closed forest, making it difficult to identify a correlation between openness and visitor densities. A variable that was not found to be important by our analytical method was traffic noise. This may be due to the lack of variation provided by the data (too coarse) and the fact that a large part of the area was based on the value of background noise of 35 dB.

If our rules of thumb are to be useful, it is essential that they are applicable in a range of nature areas. Such areas should therefore have similar features to our study area. The most important features of our study area are its size (a few thousand hectares), the large path network with multiple car parks, the fact that it is a cultural landscape, common in western Europe, and the presence of just a few specific attractions. The steep distance decline curve has also been found in several other studies, suggesting that it is a generic description of the correlation between visitor densities and distance to car park (Yang and Diez-Roux 2012, Tratalos et al. 2013, Prins et al. 2014). One way of testing the validity of the algorithm for use in other areas is to compare the average trip length of visitors in the New Forest found in this study with other studies in similar areas. Such a comparison shows that the average trip length is in the same order of magnitude. Meijles et al. (2014) reported 4.8 km in a mixed forest and heathland area in the Netherlands, Taczanowska et al. (2008) reported 5.2 km in an urban forest park in Austria and Zhai et al. (2018) reported 3.4 and 3.8 km in two urban forest parks in China. Shorter lengths were reported by Sharp et al. (2008): 2.2 km for dog walkers and 2.4 km for walkers in the Dorset heaths (UK) and 2.5 km for dog walkers and 2.6 km for walkers in the Thames basin heaths (UK). In small nature areas the results might be less useful as the average trip length and maximum distance visitors penetrate into the area might be lower; Hornigold et al. (2016) uses 400 m as a typical distance covered by visitors entering nature areas in the UK.

5.3 Limitations of the used methodology

For our study we used an existing data set. The data was collected as part of the PROGRESS research project to show the managers of the New Forest how visitors behaved in the area during their visit (Gallagher et al. 2007). Although reusing this data is efficient, it may introduce a bias because the data was not tailor-made for our study. Two design choices are of interest in this respect: the selection of car parks and the survey period. Managers selected car parks that are popular by both local visitors as well as visitors from further away and represent differences in landscape setting and capacity. This might lead towards a bias of car parks that are busy. We expect that it is unlikely that poorly used car parks will greatly affect the results. Poorly used car parks will have a small impact on the overall visitor densities in the area. Nevertheless, we may not exclude that a focus on busy car parks has introduced some bias and for future research it is preferred to select survey locations randomly. The survey period was during spring and summer 2004. During this period the proportion of visitor from further away is higher in the New Forest (Sharp et al. 2008). For our study this difference should be included as the conflict between outdoor recreation and nature conservation is most prominent during spring and summer; the breeding season of bird species. However, managers should be cautious to use the tools for conflict situations during autumn and winter. Visitors might behave differently during these periods.

5.4 Dealing with GPS data

Due to the large numbers of tracks and car parks where visitors have been monitored in the area we consider the dataset to be a good reflection of visitor behaviour and visitor densities in the New Forest. Using GPS devices for monitoring purposes always has limitations due to the accuracy of the locations stored by the GPS device. Especially in woodlands, data points may lie some distance from the path network (Piedallu and Gégout 2005). Lack of accuracy can lead to errors in the dataset and we found that error handling is a time consuming part of the research (Meijles et al. 2014). Communication errors or breakdowns between the GPS device and satellites, usually for short periods, meant that some parts of the routes taken by visitors were missing. We used the travelling salesman algorithm (Appendix B) to fill these gaps, but as the algorithm always chooses the shortest distance over the path network, some of the selected paths may not actually have been used. A relatively small part of the routes followed (15%) were constructed by the algorithm and we are confident that most of the paths were selected correctly as the visual check in step four of the data preparation did not show any unexpected results. Nevertheless, further research is needed to determine the accuracy of this algorithm in selecting path segments to complete the routes followed by visitors, based on the data points collected by GPS devices.

Appendix A. Dataset of GPS tracks

The original dataset was collected as part of the PROGRESS research project in the New Forest during spring and summer of 2004 at 41 car parks (Gallagher et al. 2007). Both models of the GPS devices used, the eTrex and eTrex Venture, were manufactured by Garmin and have 12 receiver channels. The nominal position accuracy is 15 m for the eTrex and 5 m for eTrex Venture. However, Rodríguez-Pérez et al. (2007) showed a decrease in accuracy in areas with a forest canopy for comparable device models. The positional accuracy is affected by stem density due to the lowering of the signal to noise ratio and the signal interception caused by electromagnetic waves penetrating through stems and canopies. At each car park, the GPS devices were turned on before data collection to ensure that the current almanac was stored and an accurate position was acquired. At the time of data collection, no selective availability was in operation. The devices have a storage capability for 2048 data points and were set to the 'Auto' record method for recording the tracks. This method records the tracks at a variable rate to create an optimum representation of the track. After participants returned, their device was connected to a laptop. A lightweight application, using the Garmin transfer protocol, read the data points into a database.

Table A.1 shows the number of days the car parks were monitored and the number of tracks collected from each car park. Table A.2 shows monitoring was conducted less frequently on Sundays. Sharp et al. (2008) showed that residents in the New Forest tend to use different car parks than visitors living outside the area. Combining the dataset with information from Sharp et al. (2008) indicates that visitors with dogs are mainly local residents (Fig. A.1). The dataset used contains 14 columns of information (Table A.3).

Table A.1. Number of tracks gathered at each car park.

Car park	Times monitored	Total tracks
Acres Down	4	7
Anderwood	2	8
Andrews Mare	4	24
Ashley Walk	6	31
Beaulieu Heath	5	19
Blackwater	5	38
Blackwell Common	7	41
Bolderwood	7	75
Burbush Hill	6	25
Burley	5	20
Busketts Lawn	7	30
Cadnam Cricket	3	6
Clay Hill	4	29
Crockford	4	31
Deerleap	12	109
Dibden Inclosure	6	114
Fritham	4	22
Godshill Cricket	6	57
Hincheslea Moor	3	11
Kings Hat	5	30
Linford Bottom	6	61
Longslade Bottom	5	42
Longslade Heath	4	34
Millyford Bridge	4	19
Mogshade	1	2
Moonhills	9	77
Ober Corner	2	9
Pig Bush	7	56
Pipers Wait	5	32
Queens	14	119
Shatterford	8	41
Smugglers Road	5	44
Standing Hat	4	16
Turf Hill	5	57
Vereley	4	20
Whitefield Moor	6	41
Wilverley Plain	7	66
Woods Corner	2	19
Wooton Bridge	3	11
Yew Tree Bottom	4	24
Yew Tree Heath	8	46
Total	218	1563

Table A.2. Number of tracks gathered on each day of the week.

Day	Times monitored	Number of tracks
Monday	11	210
Tuesday	12	262
Wednesday	12	272
Thursday	13	269
Friday	12	223
Saturday	12	227
Sunday	8	100
Total	80	1563

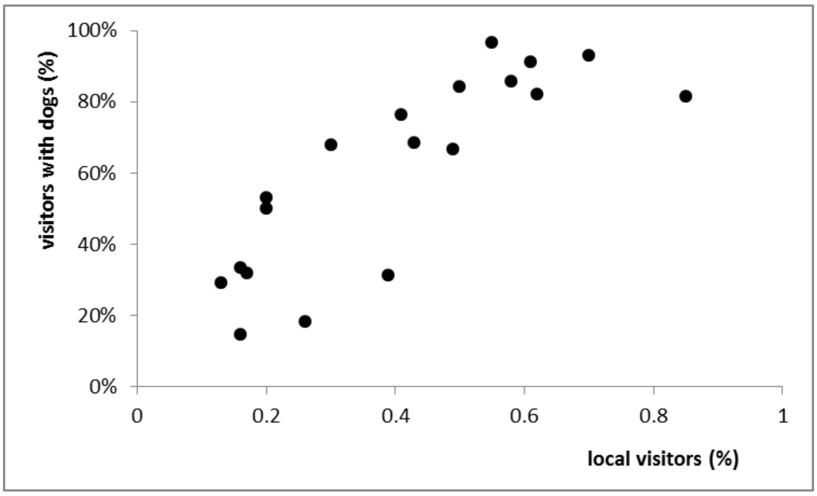

Fig. A.1. Relationship between percentage of local visitors (residents; from Sharp et al. 2008) and percentage of visitors with dogs (based on the GPS tracks) for 19 car parks that are available in both datasets.

483

Table A.3. Explanation of headers in the added file (dataset.xlsx) containing data from the GPS tracks. The first 11 rows originate from the original dataset and the final three rows were added during the preparation of the dataset.

Header	Further information
ID	unique ID for each data point
Track_ID	Unique ID for each track
Easting	Easting coordinate
Northing	Northing coordinate
Date_Time	time of storing data point
Date	Date of survey
Car_Park_Name	Name of car park where GPS device was handed out
Car_Park_Code	Code of car park where GPS device was handed out
N_People	Number of people in the visitor group
N_Dogs	Number of dogs along the visitor group
Dogs_on_Leash	"Y" means dogs were on leash and "N" means dogs were off leash
Internal_Track_ID	Unique ID for each data point starting at 1 for each track
After_Preperation	Data point taken into account after preparation step
After_GIS	Data point taken into account after GIS snapping procedure

Appendix B. Data handling of GPS tracks

The data points of each GPS track were used to construct the expected route of a visitor or visitor group in four steps. The four steps are illustrated for one track in Fig. B.1. This track was chosen as it illustrates all the potential problems we encountered in constructing routes from data points. The first step in the preparation of the dataset was the removal of outliers and data points that are considered redundant for further analyses. Outliers are data points that are located at large distances from the rest of the data points on a specific track. We found two types of outliers: outliers caused by researchers switching the GPS device on and off before arriving at a car park without resetting the device, and outliers due to errors in the communication between the GPS device and satellites (Piedallu and Gégout 2005). A visual check revealed that for some tracks two consecutive data points were outliers. To select these consecutive errors we calculated the average distance to the three previous data points and to the three following data points. We used the rule that one of the average distances had to exceed 500 m and the other at least 250 m to be considered an outlier. The dataset also contains clusters of data points at the start of a visit and at the end of a visit, due to the handling time between researchers and visitors, and at locations where visitors probably had a short stop. These clusters of data points contain many data points that may be considered redundant for determining the route followed. To decrease preparation time data points within 5 m of one another were reduced to one data point for further analysis (Fig. B.1). The removal of outliers and redundant points resulted in a 5% reduction in the number of data points.

In the second step, data points were assigned to the path network using the snapping method from the ArcGIS Toolbox (http://pro.arcgis.com/en/pro-app/tool-reference/editing/snap.htm). We used the snapping rule to assign data points to the nearest path within a distance of 50 m.

Data points that are further away from the path network were excluded for further analyses (Fig. B.1). This preparation step resulted in a 1% reduction of the data points.

The third step was the construction of the routes. Many tracks missed data points for small parts of the route followed. To fill these gaps a travelling salesman route algorithm was used in QGIS Desktop (v2.14.12) with GRASS (v7.2.0) (https://grass.osgeo.org/grass70/manuals/v.net.path.html). This algorithm constructs routes based on the order of data points. The shortest route between different data points on a path network are linked to one route. Information from the track logs was used for the order of the data points (Fig. B.1). For 10 tracks no routes could be constructed as they contained too few data points. At this stage of the analysis the resulting dataset contained 1553 routes.

Finally, in the fourth step a visual check of the constructed routes was conducted using QGIS. During the check small segments, or ‘dangling nodes’, of the routes were deleted (Fig. B.1). These segments originated from snapping a data point to the nearest path. At crossings this sometimes resulted in allocating the GPS data point to a path the visitor most likely would have crossed instead of followed. Only segments of paths were deleted when the snapped point was within 100 m, as the crow flies, of the main route a visitor had most likely followed. The set of 1553 routes was used to derive rules of thumb. For the random forest model only car parks with 10 or more routes in the database were taken into account, resulting in frequency maps for 36 car parks based on 1521 expected routes.

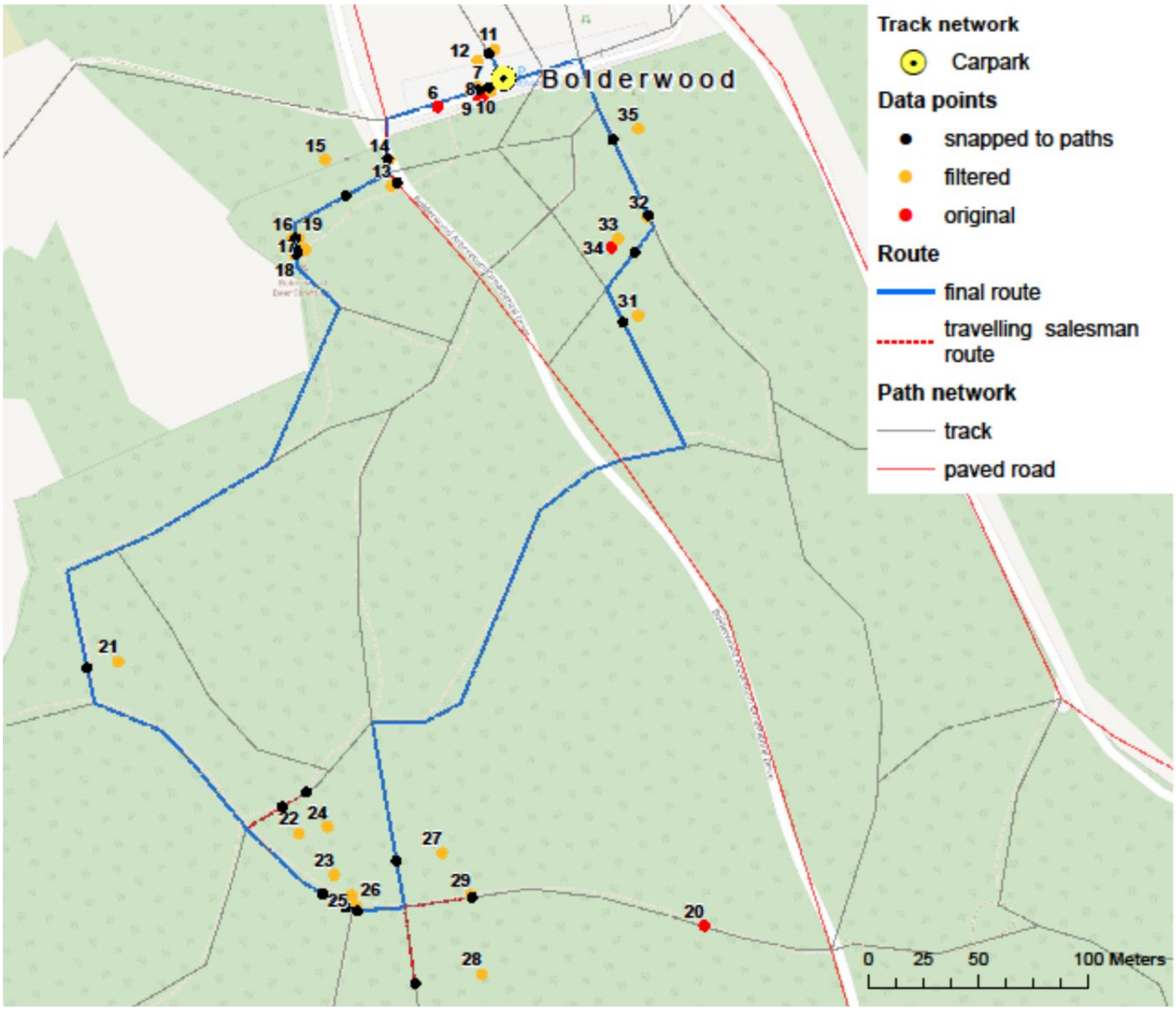

Fig. B.1. Overview of the preparation of the GPS tracks into visitor routes for track ID 1057, containing 35 data points. The red dots show the outliers and redundant data points. Numbers 1–5 are not shown as these are located 20 km from the New Forest in an urban area; the researcher probably forgot to reset the GPS device before the start of the study. Data point 20 illustrates an outlier resulting from errors in communication between the GPS device and satellites. Data point 30 is an even more extreme outlier at several hundred metres off the route and not visible at this scale. Data points 6, 10 and 34 illustrate redundant data points. For this track all selected data points (orange) are within 50 m of the path network (black dots represent the snapped data point). The travelling salesman route algorithm was used to derive the final route (the combination of the blue and dotted red lines). For this track the algorithm was needed to connect the route between data points 19 and 21 and between 29 and 31. Based on a visual check the red dotted lines towards data point 22 and 24, 28 and 29 were deleted resulting in a final route for this visitor group.

Appendix C. Description of three potential management interventions used to assess the impact of possible management actions on the Nightjar population

The model was used to assess three potential management interventions that alter the capacity and use levels of the car parks. The capacities were based on the input maps of the car parks and the use levels were based on the local knowledge of the site managers. The current capacities and use levels were altered by the researches to simulate the three interventions and illustrate the potential of the model. The total number of visitors to the New Forest (13.3 million; Gallagher et al. 2007) were distributed over the car parks based on the combination of their capacity and use level.

The three potential management interventions are:

1. Closing small car parks: All car parks with a capacity of less than 20 cars were considered closed. This resulted in the closure of 45 car parks and a redistribution of less than 10% of all visitors over the other car parks. Visitors that were expected to start from these 45 car parks in the current situation were redistributed in proportion to the number of visitors starting at the other car parks. Closing down small car parks may be expected to result in larger areas that are disturbance free.
2. Focus on suitable areas: Three relatively isolated car parks located near areas with many Nightjars were considered closed. Visitors from these three car parks were redistributed to five surrounding car parks in proportion to the capacity of these car parks. The three car parks are Andrews Mare, Yew Tree Heath and Moonhills. It was expected that this scenario would have the highest impact per redistributed visitor as the measures focus on areas that are suitable for Nightjar.

3. Concentrate visitors in a small part of the area: All but 20 car parks were considered closed. All visitors were distributed over these 20 car parks evenly. The total number of visitors that start their trip from these car parks corresponds to the two car parks that are used most in the current situation, Bolderwood and Wilverley Inclosure. This most extreme intervention was expected to concentrate visitors in a small part of the area, resulting in large undisturbed areas and an increase in population size of Nightjar.

First, the random forest model, based on monitoring data from 36 car parks, was applied to all the car parks in the area. This gave the frequencies with which visitors would be present at certain locations. For each scenario these frequencies were multiplied by the number of visitors starting from a specific car park. The results for all the car parks were summed to derive the estimated visitor density on the path network in the New Forest.

Second, we determined the potential population of Nightjar for the situation without recreation. The Forestry Commission provided a map with the breeding pairs of Nightjar in the New Forest based on the 2004 survey, the same year as the GPS dataset (see also Newton 2010). We assumed this distribution reflects the habitat suitability for Nightjar, but should be corrected for the impact of the disturbance of visitors. In areas with high visitor numbers, the number of breeding pairs is expected to be much higher when visitors are absent. We used the dose–impact relationship of Pouwels et al. (2017) to correct the current distribution and estimate the potential population in the area for a situation without recreation by multiplying each breeding pair by the inverse of the index in Fig. C.1. We used the maximum visitor groups per ha per year within a radius of 500 m as the disturbance level (x-axis in Fig. C.1). This radius is based on research by Murison (2002) and Lowe et al. (2014).

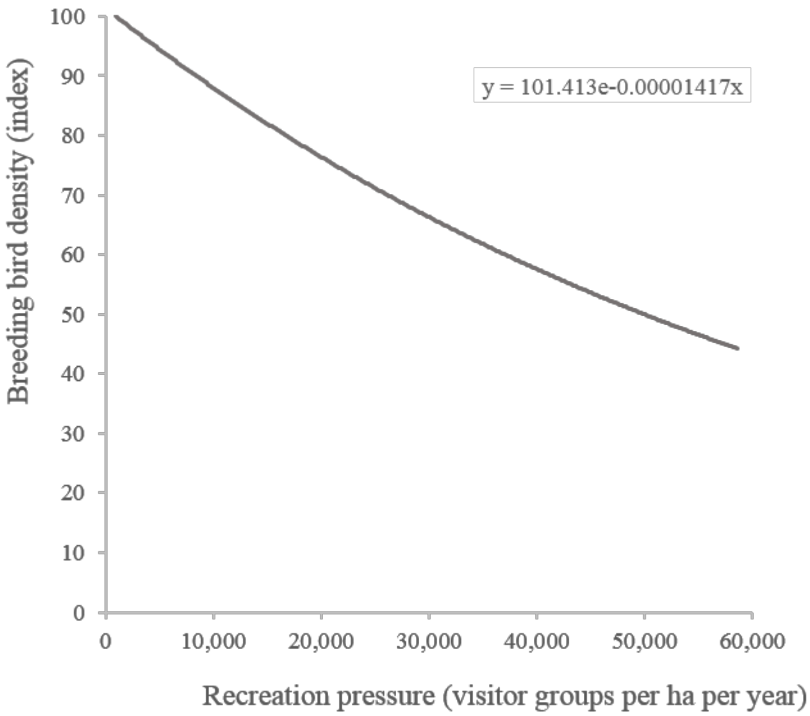

Fig. C.1. Dose–impact relationship between recreation pressure, in visitor groups per ha per year, and the breeding bird density index for Nightjar. The figure is taken from (Pouwels et al. 2017).

Finally, for the three interventions the maximum number of visitor groups per ha per year was determined within a radius of 500 m for each breeding pair. Using Eq. C.1, the corrected number of breeding pairs for a specific intervention was determined (Cpd_i). The Cpd_i is summed to predict the population size. For the current situation this resulted in the number of breeding pairs from the survey itself as $Ibpd_i$ equals $Ibpd_{current}$. $Ipbd_i$ is the index of the breeding pair density based on the recreation pressure for a specific intervention and $Ipbd_{current}$ for the current situation. As some values within the 500 m buffers are very high, we cut off the impact of visitors at the impact of 100 000 visitors, resulting in a minimum index of 25% of breeding bird densities.

$$Cbp_i = \frac{1}{Ibpd_{current}} \times Ibpd_i \qquad \text{Eq. C.1}$$

Cbp_i is the corrected number of breeding pairs for a specific intervention

$Ipbd_{current}$ is the Index of the breeding pair density based on the recreation pressure in the current situation

$Ipbd_i$ is the Index of the breeding pair density based on the recreation pressure for a specific intervention

Literature

1. Arnberger, A., & Haider, W. (2007). Would you displace? It depends! A multivariate visual approach to intended displacement from an urban forest trail. *Journal of Leisure Research, 39*(2), 345.
2. Beeco, J. A., & Brown, G. (2013). Integrating space, spatial tools, and spatial analysis into the human dimensions of parks and outdoor recreation. *Applied Geography, 38*(0), 76-85. doi: http://dx.doi.org/10.1016/j.apgeog.2012.11.013
3. Beeco, J. A., Hallo, J. C., & Brownlee, M. T. J. (2014). GPS Visitor Tracking and Recreation Suitability Mapping: Tools for understanding and managing visitor use. *Landscape and Urban Planning, 127*, 136-145. doi: http://dx.doi.org/10.1016/j.landurbplan.2014.04.002
4. Beeco, J. A., Hallo, J. C., English, W. R., & Giumetti, G. W. (2013). The importance of spatial nested data in understanding the relationship between visitor use and landscape impacts. *Applied Geography, 45*, 147-157. doi: http://dx.doi.org/10.1016/j.apgeog.2013.09.001
5. Benfield, J. A., Bell, P. A., Troup, L. J., & Soderstrom, N. C. (2010). Aesthetic and affective effects of vocal and traffic noise on natural landscape assessment. *Journal of Environmental Psychology, 30*(1), 103-111.
6. Beunen, R., Regnerus, H. D., & Jaarsma, C. F. (2008). Gateways as a means of visitor management in national parks and protected areas. *Tourism Management, 29*(1), 138-145.
7. Böcker, L., Dijst, M., & Prillwitz, J. (2013). Impact of everyday weather on individual daily travel behaviours in perspective: a literature review. *Transport reviews, 33*(1), 71-91.
8. Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.
9. Cavens, D., Gloor, C., Nagel, K., Lange, E. and Schmid, W.A. 2004 *A framework for integrating visual quality modelling within an agent-based hiking simulation for the Swiss Alps*. Paper presented at the Second International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas. Finnish Forest Research Institute, Rovaniemi, Finland.
10. Cessford, G., & Muhar, A. (2003). Monitoring options for visitor numbers in national parks and natural areas. *Journal for Nature Conservation, 11*(4), 240-250.
11. D'Antonio, A., Monz, C., Lawson, S., Newman, P., Pettebone, D., & Courtemanch, A. (2010). GPS-based measurements of backcountry visitors in parks and protected areas: Examples of methods and applications from three case studies. *Journal of Park and Recreation Administration, 28*(3).
12. De Vries, S., Buijs, A. E., Langers, F., Farjon, H., van Hinsberg, A., & Sijtsma, F. J. (2013). Measuring the attractiveness of Dutch landscapes: Identifying national hotspots of highly valued places using Google Maps. *Applied Geography, 45*, 220-229.
13. DEFRA. (2016). *Road Noise - Lden - England. Department for Environment, Food and Rural Affairs (DEFRA)*. Retrieved March 1, 2017. Latest version available at: https://environment.data.gov.uk/DefraDataDownload/?mapService=DEFRA/RoadNoiseLdenRound2&Mode=spatial
14. Eagles, P. F. (2014). Research priorities in park tourism. *Journal of Sustainable Tourism, 22*(4), 528-549.
15. EEA. (2016a). *Corine Land Cover (CLC) 2006, Version 18 (V18). European Environment Agency (EEA) under the framework of the Copernicus programme - copernicus@eea.europa.eu.* Retrieved March 1, 2017. Available at: https://land.copernicus.eu/pan-european/corine-land-cover

16. EEA. (2016b). *Slope derived from EU-DEM version 1.0. European Digital Elevation Model (EU-DEM), version 1.0. European Environment Agency (EEA) under the framework of the Copernicus programme - copernicus@eea.europa.eu.* Retrieved March 1, 2017. Available at: https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1-0-and-derived-products/slope
17. Gallagher, K., Graham, M., & Colas, S. (2007). *PROGRESS Project Handbook.* Progress Project, Lyndhurst.
18. Gimblett, R., & Skov-Petersen, H. (2008). *Monitoring, simulation and management of visitor landscapes*. Tucson: The University of Arizona Press.
19. Hadwen, W. L., Hill, W., & Pickering, C. M. (2007). Icons under threat: Why monitoring visitors and their ecological impacts in protected areas matters. *Ecological Management & Restoration, 8*(3), 177-181.
20. Hallo, J. C., Beeco, J. A., Goetcheus, C., McGee, J., McGehee, N. G., & Norman, W. C. (2012). GPS as a Method for Assessing Spatial and Temporal Use Distributions of Nature-Based Tourists. *Journal of Travel Research, 51*(5), 591-606. doi: 10.1177/0047287511431325
21. Hammitt, W. E., Cole, D. N., & Monz, C. A. (2015). *Wildland recreation: ecology and management*: John Wiley & Sons.
22. Henkens, R., Jochem, R., Pouwels, R., & Visschedijk, P. A. M. (2006, 13-17 September 2006). *Development of a zoning instrument for visitor management in protected areas.* Paper presented at the Third International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas, Rapperswil, Switzerland.
23. Heyman, E., Gunnarsson, B., Stenseke, M., Henningsson, S., & Tim, G. (2011). Openness as a key-variable for analysis of management trade-offs in urban woodlands. *Urban Forestry & Urban Greening, 10*(4), 281-293.
24. Hornigold, K., Lake, I., & Dolman, P. (2016). Recreational Use of the Countryside: No Evidence that High Nature Value Enhances a Key Ecosystem Service. *Plos One, 11*(11), e0165043. doi: 10.1371/journal.pone.0165043
25. Itami, R., Raulings, R., GMacLaren, G., Hirst, K. G., R., Zanon, D., & Chladek, P. (2003). RBSim 2: simulating the complex interactions between human movement and outdoor recreation environment. *Journal for Nature Conservation, 11*, 278-286.
26. JNCC (2015a). *Natura 2000 - Standard Data Form. Special Areas of Conservation under the EC Habitats Directive*. Retrived from. https://jncc.gov.uk/jncc-assets/SAC-N2K/UK0012557.pdf
27. JNCC (2015b). *Natura 2000 - Standard Data Form. Special Protection Areas under the EC Birds Directive*. Retrived from. https://jncc.gov.uk/jncc-assets/SPA-N2K/UK9011031.pdf
28. Jochem, R., Van Marwijk, R., Pouwels, R., & Pitt, D. G. (2008). MASOOR: modeling the transaction of people and environment on dense trail networks in natural resource settings. In R. Gimblett & H. Skov-Petersen (Eds.), *Monitoring, simulation and management of visitor landscapes* (pp. 269-293). Tucson: The University of Arizona Press.
29. Jochem, R., Meeuwsen, H., & van der Sluis, T. (2016). Imagining the future landscape with ViewScape. Paper presented at Landscape in imagination and the virtual future, Newcastle, United Kingdom.
30. Kaplan, R., Kaplan, S., & Brown, T. (1989). Environmental preference: A comparison of four domains of predictors. *Environment and behavior, 21*(5), 509-530.
31. Korpilo, S., Virtanen, T., Saukkonen, T., & Lehvävirta, S. (2018). More than A to B: Understanding and managing visitor spatial behaviour in urban forests using public participation GIS. *Journal of Environmental Management, 207*, 124-133.

32. Langston, R. H. W., Liley, D., Murison, G., Woodfield, E., & Clarke, R. T. (2007). What effects do walkers and dogs have on the distribution and productivity of breeding European Nightjar Caprimulgus europaeus? *Ibis, 149*(s1), 27-36.
33. Larson, C. L., Reed, S. E., Merenlender, A. M., & Crooks, K. R. (2016). Effects of Recreation on Animals Revealed as Widespread through a Global Systematic Review. *Plos One, 11*(12), e0167259.
34. Larson, C. L., Reed, S. E., Merenlender, A. M., & Crooks, K. R. (2018). Accessibility drives species exposure to recreation in a fragmented urban reserve network. *Landscape and Urban Planning, 175*, 62-71.
35. Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*(3), 18-22.
36. Lowe, A., Rogers, A., & Durrant, K. (2014). Effect of human disturbance on long-term habitat use and breeding success of the European Nightjar, Caprimulgus europaeus. *Avian Conservation and Ecology, 9*(2).
37. Maldonado, A., Wachowicz, M., & Vaźquez-Hoehne, A. (2011). Movement surface: a multilevel approach for predicting visitor movement in nature areas. *Environment and Planning B: Planning and Design, 38*(5), 864-878.
38. McCool, S. F. (2016). Tourism in Protected Areas: Frameworks for Working Through the Challenges in an Era of Change, Complexity and Uncertainty *Reframing Sustainable Tourism* (pp. 101-117): Springer.
39. Meijles, E. W., de Bakker, M., Groote, P. D., & Barske, R. (2014). Analysing hiker movement patterns using GPS data: Implications for park management. *Computers, Environment and Urban Systems, 47*, 44-57. doi: http://dx.doi.org/10.1016/j.compenvurbsys.2013.07.005
40. Molnar, C. (2018). Interpretable machine learning: A guide for making black box models explainable; https://christophm.github.io/interpretable-ml-book/. *Christoph Molnar, Leanpub*.
41. Monz, C. A., Pickering, C. M., & Hadwen, W. L. (2013). Recent advances in recreation ecology and the implications of different relationships between recreation use and ecological impacts. *Frontiers in Ecology and the Environment, 11*(8), 441-446. doi: 10.1890/120358
42. Murison, G. (2002). *The Impact of Human Disturbance on the Breeding Success of the Nightjar Caprimulgus Europaeus on Heathlands in South Dorset, England*: English Nature.
43. Newton, A. (2010). *Biodiversity in the New Forest*: Pisces Publications.
44. Olson, L. E., Squires, J. R., Roberts, E. K., Miller, A. D., Ivan, J. S., & Hebblewhite, M. (2017). Modeling large-scale winter recreation terrain selection with implications for recreation management and wildlife. *Applied Geography, 86,* 66-91.
45. Orsi, F., & Geneletti, D. (2013). Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *Journal for Nature Conservation, 21*(5), 359-368.
46. Pesonen, K. (2000). *On noise assessment and noise control engineering problems caused by seasonal variations of noise emission and excess attenuation.* Paper presented at the INTER-NOISE and NOISE-CON Congress and Conference Proceedings.
47. Piedallu, C., & Gégout, J.-C. (2005). Effects of forest environment and survey protocol on GPS accuracy. *Photogrammetric Engineering & Remote Sensing, 71*(9), 1071-1078.
48. Pouwels, R., Sierdsema, H., Foppen, R. P. B., Henkens, R. J. H. G., Opdam, P. F. M., & van Eupen, M. (2017). Harmonizing outdoor recreation and bird conservation targets in protected areas: Applying available monitoring data to facilitate collaborative management at the regional scale. *Journal of Environmental Management, 198*, 248-255.

49. Prins, R. G., Pierik, F., Etman, A., Sterkenburg, R. P., Kamphuis, C., & Van Lenthe, F. (2014). How many walking and cycling trips made by elderly are beyond commonly used buffer sizes: results from a GPS study. *Health & place, 27*, 127-133.
50. Reed, S. E., & Merenlender, A. M. (2008). Quiet, Nonconsumptive Recreation Reduces Protected Area Effectiveness. *Conservation Letters, 1*(3), 146-154.
51. Renso, C., Baglioni, M., Macedo, J. A. F., Trasarti, R., & Wachowicz, M. (2012). How you move reveals who you are: understanding human behavior by analyzing trajectory data. *Knowledge and Information Systems, 37*(2), 331-362. doi: 10.1007/s10115-012-0511-z
52. Rodríguez-Pérez, J. R., Alvarez, M. F., & Sanz-Ablanedo, E. (2007). Assessment of low-cost GPS receiver accuracy and precision in forest environments. *Journal of Surveying Engineering, 133*(4), 159-167.
53. Sayan, S., Krymkowski, D. H., Manning, R. E., Valliere, W. A., & Rovelstad, E. L. (2013). Cultural influence on crowding norms in outdoor recreation: a comparative analysis of visitors to national parks in Turkey and the United States. *Environmental Management, 52*(2), 493-502.
54. Schamel, J., & Job, H. (2017). National Parks and demographic change–Modelling the effects of ageing hikers on mountain landscape intra-area accessibility. *Landscape and Urban Planning, 163*, 32-43.
55. Sharp, J., Lowen, J., & Liley, D. (2008). Recreational pressure on the New Forest National Park, with particular reference to the New Forest SPA. *New Forest National Park Authority/Footprint Ecology*.
56. Shoval, N. (2010). Monitoring and Managing Visitors Flows in Destinations using Aggregative GPS Data. *Information and Communication Technologies in Tourism 2010*, 171-183.
57. Shoval, N., Auslander, G., Cohen-Shalom, K., Isaacson, M., Landau, R., & Heinik, J. (2010). What can we learn about the mobility of the elderly in the GPS era? *Journal of Transport Geography, 18*(5), 603-612.
58. Sutherland, W. J., Gardner, T. A., Haider, L. J., & Dicks, L. V. (2014). How can local and traditional knowledge be effectively incorporated into international assessments? *Oryx, 48*(1), 1.
59. Taczanowska, K.(2009). *Modelling the Spatial Distribution of Visitors in Recreational Areas*. Universität für Bodenkultur, Wien. Retrived from: https://zidapps.boku.ac.at/abstracts/oe_list.php?paID=3&paSID=7284&paSF=-1&paCF=0&paLIST=0&language_id=DE
60. Taczanowska, K., Arnberger, A., & Muhar, A. (2008). Exploring spatial behavior of individual visitors as a basis for agent-based simulation. In R. Gimblett & H. Skov-Petersen (Eds.), *Monitoring, simulation and management of visitor landscapes* (pp. 159-174). Tucson: The University of Arizona Press.
61. Taczanowska, K., Bielański, M., Gonzalez, L.-M., Garcia-Massó, X., & Toca-Herrera, J. L. (2017). Analyzing spatial behavior of backcountry skiers in mountain protected areas combining GPS tracking and graph theory. *Symmetry, 9*(12), 317.
62. Taczanowska, K., González, L.-M., Garcia-Massó, X., Muhar, A., Brandenburg, C., & Toca-Herrera, J.-L. (2014). Evaluating the structure and use of hiking trails in recreational areas using a mixed GPS tracking and graph theory approach. *Applied Geography, 55*, 184-192. doi: http://dx.doi.org/10.1016/j.apgeog.2014.09.011

63. Thompson, B. (2015). Recreational Trails Reduce the Density of Ground-Dwelling Birds in Protected Areas. *Environmental Management, 55*(5), 1181-1190. doi: 10.1007/s00267-015-0458-4
64. Torbidoni, E. I. F. (2011). Managing for recreational experience opportunities: The case of hikers in protected areas in Catalonia, Spain. *Environmental Management, 47*(3), 482-496.
65. Tratalos, J. A., Sugden, R., Bateman, I. J., Gill, J. A., Jones, A. P., Showler, D. A., . . . Watkinson, A. R. (2013). The conflict between conservation and recreation when visitors dislike crowding: A theoretical and empirical analysis of the spatial distribution of recreational beach users. *Environmental and Resource Economics, 55*(3), 447-465.
66. van Marwijk, R. (2009). *These routes are made for walking. Understanding the transactions between nature, recreational behaviour and environmental meanings in Dwingelderveld National Park, the Netherlands.* (PhD), Wageningen University, Wageningen.
67. van Marwijk, R., Elands, B., & Jochem, R. (2009). *Simulation of outdoor recreation: the importance of empirical data.* Paper presented at ISSRM 2009, 15th International Symposium on Society and Resource Management, Meet old and new worlds in Research, Planning, and Management. Vienna, Austria
68. van Marwijk, R., & Pitt, D. G. (2008). *Where Dutch recreationists walk: Path design, physical features and walker usage.* Paper presented at the fourth international conference on monitoring and management of visitor flows in recreational and protected areas. Management for Protection and Sustainable Development, Montecatini Terme, Italy.
69. Weitkamp, G. (2011). Mapping landscape openness with isovists. *Research in Urbanism Series, 2*(1), 205-223.
70. Weitowitz, D. C., Panter, C., Hoskin, R., & Liley, D. (2019). Parking provision at nature conservation sites and its implications for visitor use. *Landscape and Urban Planning, 190,* 103597. doi: https://doi.org/10.1016/j.landurbplan.2019.103597
71. Wolf, I. D., Wohlfart, T., Brown, G., & Bartolomé Lasa, A. (2015). The use of public participation GIS (PPGIS) for park visitor management: A case study of mountain biking. *Tourism Management, 51*, 112-130. doi: http://dx.doi.org/10.1016/j.tourman.2015.05.003
72. Wright, M. N. & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software 77*(1)*,* 1-17.
73. Yang, Y., & Diez-Roux, A. V. (2012). Walking distance by trip purpose and population subgroups. *American journal of preventive medicine, 43*(1), 11-19.
74. Zhai, Y., Baran, P. K., & Wu, C. (2018). Can trail spatial attributes predict trail use level in urban forest park? An examination integrating GPS data and space syntax theory. *Urban Forestry & Urban Greening, 29*, 171-182.