Visual exploration of migration patterns in gull data

Konzack, M.; Gijsbers, P.; Timmers, F.; van Loon, E.; Westenberg, M.A.; Buchin, K.

DOI
10.1177/1473871617751245

Publication date
2019

Document Version
Final published version

Published in
Information Visualization

Info Vis

Article

Information Visualization
2019, Vol. 18(1) 138–152

DOI: 10.1177/1473871617751245
journals.sagepub.com/home/ivi
SAGE

Visual exploration of migration patterns in gull data

Maximilian Konzack[1], Pieter Gijsbers[1], Ferry Timmers[1], Emiel van Loon[2], Michel A Westenberg[1] and Kevin Buchin[1]

Abstract

We present a visual analytics approach to explore and analyze movement data as collected by ecologists interested in understanding migration. Migration is an important and intriguing process in animal ecology, which may be better understood through the study of tracks for individuals in their environmental context. Our approach enables ecologists to explore the spatio-temporal characteristics of such tracks interactively. It identifies and aggregates stopovers depending on a scale at which the data is visualized. Statistics of stopover sites and links between them are shown on a zoomable geographic map which allows to interactively explore directed sequences of stopovers from an origin to a destination. In addition, the spatio-temporal properties of the trajectories are visualized by means of a density plot on a geographic map and a calendar view. To evaluate our visual analytics approach, we applied it on a data set of 75 migrating gulls that were tracked over a period of 3 years. The evaluation by an expert user confirms that our approach supports ecologists in their analysis workflow by helping to identifying interesting stopover locations, environmental conditions or (groups of) individuals with characteristic migratory behavior, and allows therefore to focus on visual data analysis.

Keywords

Visual analytics, animal behavior, data visualization, spatio-temporal visualization, geovisualization

Introduction

The study of animal movement has long been of interest. Starting in the 1990s, the availability of new technology has led to increasingly detailed and diverse types of data relating to movement.[1] Global Positioning System (GPS)-based movement tracks are currently among the most frequently collected types of data. Because of its relative accuracy and high sampling rate, the GPS technology drastically improved the ability to describe and gain new insights about animal movement. The continuing miniaturization allows at the same time to collect data for an increasing range of species and conditions.[2] Today, animal movement tracks form one of the main data sources when studying the mechanics of movements, navigational cues, or drivers of movement, constantly leading to new insights on animal physiology, behavior, and demography.[3] Through those advances, ecologists are starting to ask novel questions on the causes of movement and its consequences for individuals, populations, and ecosystems for which formal analysis techniques are not always readily available. Kays et al.[2] claim that interdisciplinary research between data scientists—computer scientists, statisticians—and ecologists will be required to develop these new tools efficiently and will eventually lead to new insights and scientific breakthroughs.

[1]TU Eindhoven, Eindhoven, The Netherlands
[2]University of Amsterdam, Amsterdam, The Netherlands

Corresponding author:
Maximilian Konzack, TU Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
Email: m.p.konzack@tue.nl

The goal of data visualization is to provide insights into data.[4] New visualization techniques and visual encodings help users to understand their data sets. Furthermore, visual analytics can be used to generate knowledge from large and often complex data sets by developing and deploying analytical and visualization techniques.[5]

To support an ecologist in her search for new knowledge, a visualization expert needs to spend time and effort to understand the relevant questions, data, and the general ecological context of a study system—knowledge which has often not be acquired a priori. An ecologist does, however, not commonly have an overview of the analysis and visualization possibilities that are technically feasible or how these would help in answering certain ecological research questions. Thus, a knowledge gap between domain experts and visualization designers exists.[4,6] Explorative visualization can help in filling this gap as it provides means for ecologists to discover new trends, to present a data set visually, to identify pertinent subsets, to compare the movement of individuals, and to locate moving entities, among other tasks. Abstracting such tasks helps to reason about similarities and differences between them, to distinguish between different goals, and to guide data abstraction.[7] Eventually, exploratory visualization provides a novel analytical means and therefore leads to new ecological insights. In addition, it may help the users to build trust in their generated knowledge base.[5]

The current practice of ecologists to investigate and visualize movement is by developing and using MATLAB or R libraries.[6,8] Those results are either analyzed and visualized statically or plotted on top of satellite maps. Few bird ecologists examined movement data in interactive visualizations with multiple coordinated views or used a Google Earth–based tool for exploring GPS data from a bird's eye perspective, both of which have been of limited value.[6] It seems that a tight integration of different spatio-temporal views of the data, with a flexible selection, would be beneficial to ecologists to focus on data analysis mechanisms rather than laborious coding.[9]

In this article, we present such a technique. It comprises a novel visual analytics approach to help explore animal migration patterns interactively. Our approach provides analytical and visual means to understand different aspects of migration through an aggregation at various spatial scales, with interlinked geographical maps, and views on spatio-temporal events. Migration is ubiquitous in ecology. It is the seasonal displacement of individuals between sites. In our approach, we identify and aggregate stopovers. A stopover is a break within a migratory trajectory. Functionally, stopovers are important for foraging, resting, or socializing with conspecifics, but stopovers can also be used diagnostically to recognize different migration strategies: along the coastline, over sea, or over land.

Past research on visualizing gull migration lack an aggregation of the trajectories[8] or impose visual clutter by drawing the results of the clustering as colored data points on a map.[9] Our visual analytics approach remedies this by employing a stopover aggregation visualization, a density map, and a calendar view (see Figure 1). The stopover aggregation and density map are plotted on top of interconnected zoomable geographic maps. The user can select stopovers and impose constraints on spatio-temporal properties of the selection.

We applied our approach to a data set of 75 migrating Lesser Black-backed Gulls (*Larus fuscus*), which we will henceforth denote as "gulls" in this article. We evaluated our tool by consulting an expert user[10] with an expertise on bird migration to assess the strengths and weaknesses of our approach. We identified ecological research questions and requirements on the visual design and mapped them to analytical tasks that the expert user completed. In this article, we use the terms moving entities and individuals interchangeably.

Our qualitative evaluation confirms that our approach helps ecologists in their analysis of migration patterns, so that they are able to identify and to isolate groups of individuals with a certain migration behavior visually rather than in non-visual computations. This speeds up and fosters their analytical workflow because our approach empowers ecologists to focus on interpreting the data and on developing new questions without being distracted by coding or by algorithmic technicalities.

Related work

To identify different homogeneous movement episodes in trajectory data, trajectories are commonly segmented, that is, cut into parts, according to characteristics of the movement. Segmentation together with classification or clustering of movement data helps to summarize and visualize large trajectory data sets. Visualization techniques on movement data support users in identifying new trends within data sets. We survey past research on visualization and clustering of movement data from various application backgrounds to give an overview of recent achievements in these fields.

Segmentation algorithms have been successfully applied to migrating geese[11,12] and gulls[9] among many other studies.[13–19] These approaches segment individual trajectories into pieces of similar movements. Buchin et al.[20] present algorithms to summarize segmentations of a larger number of trajectories in a flow

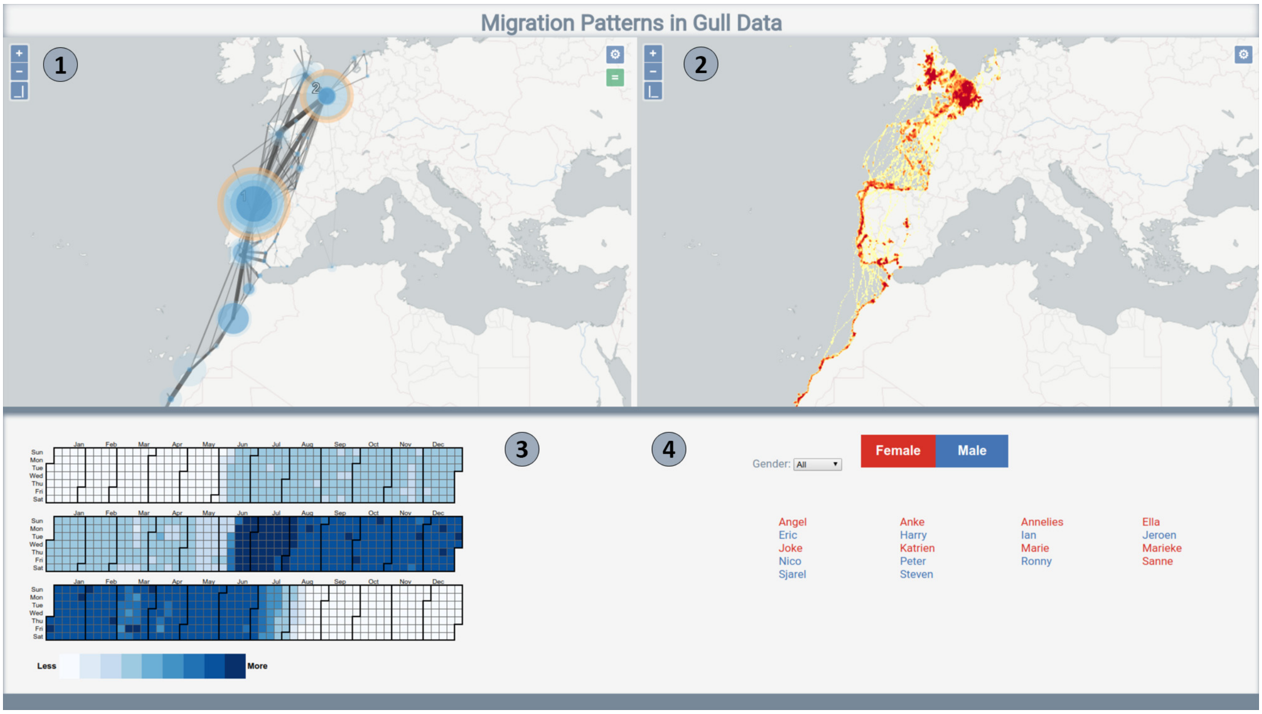


Figure 1. Overview of the visual analytics tool: the stopover aggregation visualization (1) enables the user to investigate and select stopovers. The density map (2) shows the spatial distribution of the selected moving entities. Within the calendar view (3) the temporal distribution of the stopovers of the selected entities over time is visualized. The list of gulls (4) shows the names and genders of the selected entities.

diagram, which they apply to trajectories of football players to analyze spatial formations and plays.

To aggregate movement data, Andrienko and Andrienko[21] transformed trajectories into aggregate flows between spatial regions within their visual analytics approach. This type of aggregation allows groupings of essential characteristics of the movement. In their approach, the user can control the overall level of abstraction of the visualization interactively. Andrienko and Andrienko[21] applied their approach to deer and stork data and to trajectories in an urban context, where the aggregated results are visualized as flow maps or as transition matrices.

Density visualization is a powerful visualization technique for analyzing many trajectories. One such example is the work by Willems et al.,[22] which presents a multiscale density visualization for trajectories. They demonstrate it on vessel trajectories. The visualization of the density fields is derived from the convolution of the dynamic vessel position with a kernel that takes the speed of the movement into account. The density fields are illuminated as height maps on top of the heat map. Additionally, Willems et al.[22] visualize the individual's contribution of a moving point within the overall density.

Scheepens et al.[23] use a similar approach to visualize densities of maritime trajectories, for which they apply a cascade of filtering and selection mechanisms on top of the density maps. These selection mechanisms are sensitive to a user role, either a domain expert or an operator, and to the task at hand. An operator is usually only concerned with events connected to his work task, such as surveillance of a particular port.

Densities of movement data do not have to be visualized necessarily as a raster map on top of a geographical map. Slingsby et al.[24] deploy hierarchical, interactive treemaps to explore spatio-temporal movement patterns of couriers in London.

Clustering on urban data has been studied by Lu et al.[25,26] on taxi data. Lu et al.[25] explore *Origin Destination* (OD) pairs as an interactive selection of clustered regions. The underlying summarization uses a modified DBSCAN algorithm, which is a density-based clustering algorithm. Lu et al.[26] ranked trajectories of taxi data based on their similarity of travel time. Within their visual analytics tool, they visualized the ranking as bands over time. Before ranking the trajectories, they are mapped onto the underlying street network first and then segmented.

Slingsby and van Loon[8] present an exploratory visual analysis approach for animal movement data. They applied it, as we did, on gull data, and they devised ecological and visualization requirements on the analytical process. Their visual encodings range from point plots over density maps to tile maps and, thus, cover partially our visualization techniques. The visual analytics software consists of a central view, a satellite map with an overlaid visual encoding and user interactions, and two interconnected timelines: one for a sequence of days and the other one on the times during those days. They did not, however, consider a summary and aggregation of the gulls' trajectories, which we provide and use as the main technique to interact with the user in our study.

Spretke et al.[9] developed a visual analytics approach for migrating seagulls. It supports interactive data exploration and enrichment of movement data by adding attributes dynamically from existing ones, and incorporating weather information, such as wind and temperature. They clustered the trajectories of gulls into three states: day migration, night migration, and stopover. They also segmented the trajectories based on spatio-temporal characteristics (45 min resting and less than 2 km continuous flight). Their visualization with interconnected views lacks an abstraction for visualizing clusters since Spretke et al.[9] mapped the clustering to only colors, and plotted the clusters as plain data points yielding visual clutter.

Kölzsch et al.[27] reflected on the visual design of migrating birds by exploring different visualization techniques to encode spatio-temporal characteristics of migration. They related ecological research questions to the visual design which inspired us to link our ecological research questions to the requirements of our visual design.

Problem definition and requirement analysis

Animal migration is an intriguing phenomenon in nature and has, as a consequence, always received much attention as a research topic in biology. It is increasingly being studied through visual and quantitative techniques due to the availability of tracking data in combination with relevant environmental data layers.[8,12,28,29]

In our study, we focus on the interactive analysis of animal migration tracks, and in this section, we lay out a number of important ecological research questions that can be partially answered through these means. By relating those domain research questions to analytical requirements, we aim for a holistic problem analysis[30] of migratory animal movement.

Ecological research questions on migration patterns

Even though there is a large body of knowledge about bird migration[31,32] and some common principles are generally recognized (a migrating organism would maximize its fitness behaviorally by minimizing energy consumption, time expenditure, or the risk incurred during migration),[33–35] our understanding of underlying drivers as well as the (behavioral, ecological, and physiological) mechanisms is still far from complete.

In the case of the focal species in this study (the Lesser Black-backed Gull), for instance, exact energy budgets are unknown, the comparative advantage of migration (vs overwintering in the breeding area) is unclear, and the reason for the wide spread in overwintering sites by individuals from a single colony is also unknown.[36] However, some aspects of the migration of this species have been studied, leading, for example, to the conclusion that it minimizes the energetic costs rather than time spent during migration.[28]

In this study, we explore how an interactive analysis of trajectory data can help to gain more insight in the possible role of individual differences, sex, and time-dependent conditions (such as weather patterns or ephemeral food resources) as well as the characteristics of stopover sites during migration. The ecological questions that we consider are listed in Table 1. They have been designed to cover insights into a summary of stopovers and the exploration of the relation between trajectories and the actual movement.

Questions E1 to E3 focus on the relation between attributes of the movement tracks (that can be considered as predictor variables) and migration decisions (that can be considered as response variables). However, both the predictor and response variables have not been operationalized, so that inferential testing through, for example, multiple regression is not feasible yet. Rather, an explorative analysis is required to help to define those variables.

Questions E4 to E6 deal with the uniqueness of stopover sites. For these questions, the reference (e.g. direct surroundings, other individuals) is not clearly defined, hence part of the challenge is here to discover the type and scale of reference that is meaningful.

Requirements for analysis tasks

We now map these ecological research questions to more abstracts requirements that our approach needs to support. These requirements are listed in Table 2. Moreover, they help abstracting generic tasks from spatio-temporal characteristics of migratory trajectory data.

Table 1. Overview over the ecological research questions that we explore in our visual analytics tool for migration.

	Ecological research question
E1	How do individuals' differences, sex, and temporal conditions relate to the route choice?
E2	How do individuals' differences, sex, and temporal conditions relate to the choice of the stopover site?
E3	How do individuals' differences, sex, and temporal conditions relate to the timing of stopping and commencing to migrate?
E4	What is special about the places to where migrating individuals move relative to the direct surroundings (at the same time/within the same time window)?
E5	What is special about the places to where migrating individuals move relative to the place where they come from (at the same time/within same time window)?
E6	What is special about the places to where migrating individuals move relative to other individuals (at the same time/within same time window)?

Table 2. Summary of the requirements for our visual analytics tool.

	Requirements for the visual analysis
T1	Identify spatial patterns
T2	Identify temporal patterns
T3	Identify stopovers
T4	Compare groups and individuals

In our approach, we want to *identify spatial patterns* (*T*1). This requirement covers a comprehensive visualization of all trajectories, allowing the user to understand and compare spatial patterns in migration (E1) across different scales. A grouping of gulls with similar movements provides insights into different categories of migration behaviors (E6). A sequence of stops from an origin to a destination (E5) can be expressed as such a grouping. The behavior of individuals should be distinguishable from overall group patterns to investigate how the migration strategy of an individual deviates from the group movements (E6).

We want to *identify temporal patterns* (*T*2) across several scales, ranging from day/night patterns over days to seasonal patterns. A visualization should allow to specify an episode to constrain the selection to lie within a start and end date (E4, E5, E6).

Another analytical requirement is to *identify stopovers* (*T*3). This requirement deals with a more aggregate view on the data to identify important or often used places (E2) where migrating gulls come together (E6). Stopovers can also be considered at the level of an individual to visualize its migration strategy on top of an exchangeable map, such as a geographic or a topographic map, to investigate the surroundings of a stopover (E4) with different visual cues. A visualization should furthermore provide insight into the proximity of the stopover. Statistics on a stopover help to understand the nature of the stopover (E2).

Within our analytical framework, we want to *compare groups and individuals* (*T*4). This requirement concerns grouping individuals that show similar migration strategies (E1), for example, travel mostly along a coastline, over land, or over sea. A visualization should enable not only a visual linkage between these groups (E5) but also the comparison of one or more individuals with a group (E6). Subgroups can be selected by the user individually or by characteristics of the gulls, such as gender (E1, E2, E3).

Visual analytics approach

Our visual analytics approach (see Figure 1) enables users to explore migration patterns interactively. We identify stopovers, and aggregate them in a visualization, so that the user can investigate and interact with stopovers. The tool allows the user to select a sequence of stopovers from an origin to a destination. Such a selection imposes a direction of movement for moving entities within a migration. The selected group of moving entities is rendered in the density map, the calendar view, and the list of gulls. In the density map, the spatial usage of the selected gulls is shown while the calendar displays the counts of stopovers per day of selected or all stopovers. The geographic maps for the density map and the stopover aggregation are interconnected.

Our visual analytics approach is implemented as an interactive website (http://www.win.tue.nl~kbuchin/proj/gullmigration/), so that it is widely accessible to researchers who want to explore migration patterns.

By visually exploring a data set, trust into the knowledge base of the data set can be built,[5] and it allows ecologists to focus on data analysis instead of implementing source code to isolate groups with different migration behaviors.[9]

In this section, we first discuss the algorithmic techniques which underlie our visual analytics tool and relate these to the requirements of our tool.

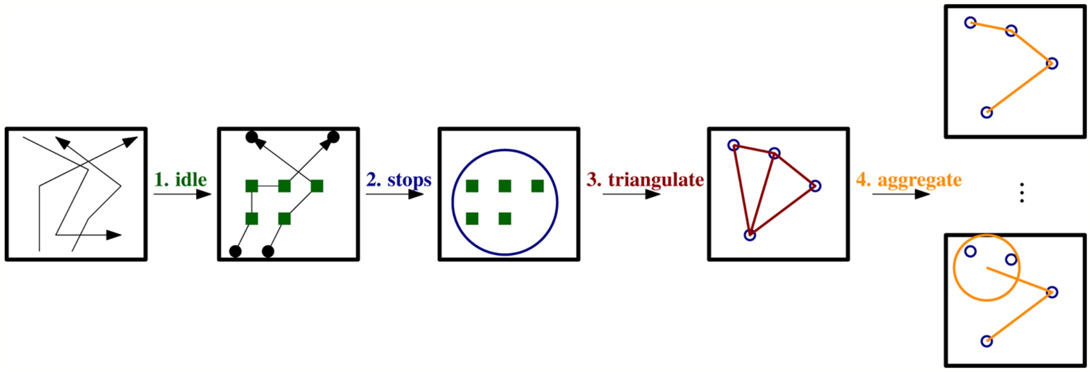


Figure 2. The computational process to aggregate stopovers consists of four phases: (1) the identification of idle points, (2) the computation of stops, (3) the triangulation of stops, and (4) the aggregation of stopovers on various scales. The first two phases identify stopovers (T3). The resulting triangulation from phase (3) is used in phase (4) to compute summaries on different spatial scales from fine to coarse (T1).

Subsequently, we link the requirements to the visualization design.

Computational methods

To allow a clustering across different spatial scales (T1), we have applied a single-linkage agglomerative clustering to aggregate stopovers. Since gulls have along their migration routes frequent stops which are heterogeneous in duration and local detours, we cannot readily apply stopover criteria used in existing algorithms[12] for segmenting the trajectories. Our algorithm supports two parameters to facilitate flexible stopover definitions. These parameters are thresholds on the speed of a point to its successor within a trajectory, defaulted to 3.5 km/h, and a maximum distance between two moving entities, defaulted to 500 m. The distance threshold determines whether two points from distinct trajectories are within the same stopover. Our chosen defaults provide sensible parameters to describe stopover criteria for gulls.

In Figure 2 we show the aggregation algorithm that we employed; it has four phases. The first two phases focus on identifying stopovers (T3), and phases (3) and (4) exploit spatio-temporal characteristics of migration (T1, T2). The first three phases are executed sequentially. After that, phase four is executed for different spatial scales. We discuss each phase in more detail in the following.

First, we classify points as *idle* if the speed of a point with respect to the previous point of the trajectory of an individual is below the given threshold. This allows us to distinguish between movement and non-movement of an individual.

Next, we compute stopovers (T3) by employing Ritter's Bounding Sphere algorithm[37] onto the idle points and thresholding the distance between two idle points. A stop is the smallest disk containing idle points (Figure 2 shows five idle points that together define a stop). Ritter's algorithm, with a running time of $O(nd)$ in general for n points in d dimensions, is exceptionally efficient with a linear running time in our case, since we compute the sphere in the plane and n points at a stop. A drawback of this algorithm is that the obtained disk is approximately 5% larger than the optimal minimum-radius circle. By identifying these stops, we are able to represent them as a visual abstraction.

In the third phase, we take the centers of all disks, which represent stopovers of idle points, and compute a Delaunay triangulation[38] on those centers. Because a Delaunay triangulation minimizes the minimum angle within the triangulation, that is, avoiding narrow triangles, it is a suitable means to compute all possible edges between stopovers.

Finally, to aggregate the stops, we perform a single-linkage clustering by applying Kruskal's algorithm[39] on the Delaunay triangulation from fine to coarse scale (T1), where the distances between the corresponding stops—disks—serve as edge weights. We span an edge uv between the centers of the smallest enclosing balls at u and v only for the moving entities incident to this edge. We exclude in this computation those individuals who are not traveling along that edge uv. This step allows us to construct a summary of the stopovers that reduces visual clutter across multiple spatial scales.

The density map is dynamically computed on a set of gulls and allows comparisons between different groups (T4). The computation of a density map consists in our approach of three steps: first, we interpolate the data linearly for each individual, using a sample resolution of 15 min to bypass irregular sample

Figure 3. Color scale for the quintiles in the density map.

intervals. Then, we bin the counts for each individual on a grid of all possible locations. The counts of a cell are the occurrences of all moving entities with that grid cell. Eventually, we discretize the binned values to five quintiles (see Figure 3) and compute the contour lines of the density map. Such a density map allows users to perceive the distribution and the spatial extent of the trajectories (T1).

As we are interested in investigating day and night patterns of trajectories (T2), we need to classify subtrajectories as daytime or nighttime. To determine whether a data point of a trajectory, given as longitude, latitude, and a time stamp, occurs during day, night, or twilight, we used the method described in Forsythe et al.[40] It is robust across the latitude, and the computational error ranges from maximum 1 min near the equator up to 2 min near 60° north latitude.

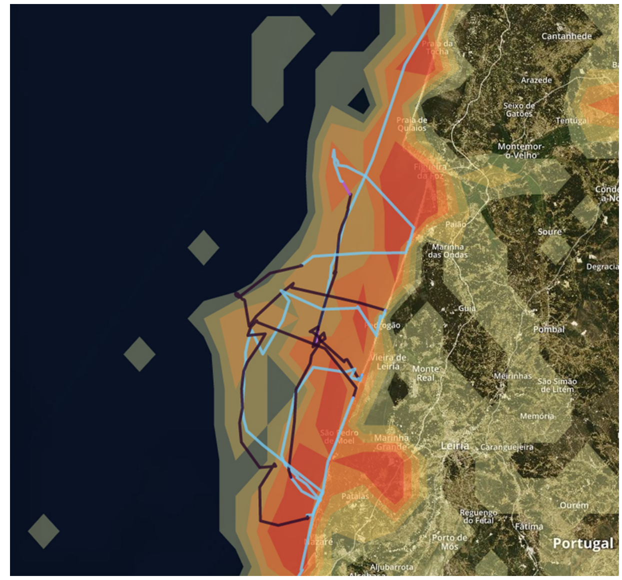


Figure 4. Activity at night of gull "Sanne." The coloring of the individual gull, Sanne, shows considerable movement at night, in black, during the migration.

Visualization techniques

The stopover aggregation (T3) provides an overview of the stopovers and the segmented trajectories, moving among the stopovers. We represent each stopover as a disk, of which the radii encode the quantiles on the number of trajectories at the stopover. This encoding allows us to visualize the spatial distribution of the stops within a stopover. The selected stopovers have an orange halo, and the number indicates the sequential ordering of the stopovers from an origin to a destination (see Figure 1). Edges are colored in a light shade of gray per default without any selection. The more moving entities of the selection travel along an edge, the stronger it will be saturated in a darker shade of gray. Our encoding enables users to have a visually salient selection.

To change the spatial layout of the stopover aggregation from coarse-grained to fine-grained (T1), we allow the user to select an aggregation level (see phase four in section 'Computational methods') by providing a slider, which is only visible if no selection has been made. The map type can be changed to a satellite view (see Figure 4) to investigate the vicinity of the stopover (T3), as it is provided by other state-of-the-art map services.

The density map provides an overview of the spatial usage of a set of moving entities or all of them (T1). We ensure that individual trajectories that deviate substantially from others are clearly visible using an appropriate kernel size. This way, the density map also supports task T4. As for the stopover aggregation, the map type can be changed to a satellite view as well (T3).

To investigate the migration strategy of an individual, we provide a trajectory visualization. The trajectory is drawn on top of the density map (Figure 4). Day, night, and twilight are color coded in light blue, black, and purple (T2), respectively. This allows a user to see whether a gull travels long or short periods on a particular day and also how many days the entire journey takes. The text labels on top of the trajectory show the stopovers of an individual and can be turned off and on. They indicate the sequence and the direction of the movement for an individual.

The calendar view is shown when a stopover, a set of gulls, or an individual gull has been selected (T4). It provides information on the distribution of stops at each day and which gulls stopped at a specific day (T2). Such temporal information can be encoded in various ways, such as a timeline or a punch card chart. A calendar has the advantage of showing multiple years at the same time. Our tool allows users to toggle between showing the distribution of the selected stopovers or of all stopovers. The number of stops are visualized in the same saturated scale of blue as the one used in the stopover aggregation. Selecting a time frame within the calendar helps to exploit temporal patterns (T2), and such a selection is visualized as a contour in the same shade of orange as the one used for the selection in the stopover aggregation (see Figure 6).

After evaluating the metadata of the used data set,[41] we focused on visualizing only names and gender of the gulls in addition to their trajectories because the ancillary data did not provide any pertinent statistics (categorical and numerical) beyond gender and name. The gulls within the selection are alphabetically sorted by the names of the gulls. Males are colored in blue and females in red. Subselections of previous selections are supported in various ways (T4).

Exploratory analysis process

Our visual analytics tool provides two ways to explore migration: analyzing migration patterns at a level of stopovers and investigating the spatio-temporal characteristics of a single moving entity. Within the more comprehensive analytical process in exploring stopovers, we support at any time the inspection of an individual.

The analysis process for exploring stopovers is the following:

1. Spot interesting stopovers;
2. Select a promising stopover;
3. Investigate the selected stopover(s);
4. Refine the selected stopover(s);
5. Inspect the individual trajectories of the result set.

Users start by attaining a general overview of the data set by zooming, panning, and inspecting the interconnected maps. By hovering on interesting stopovers (see Figure 5(a)), users gain insight on the structure and the relevance of the stopovers (step 1).

Next, users select a stopover of their interest by clicking on it (step 2). This will update all other views: the density map, the calendar, and the list of gulls.

Subsequently, users explore the nature of the selection (step 3). Within the calendar view, users can toggle between showing the counts of all stopovers or only the selected ones. By hovering a day in the calendar, a tooltip (Figure 5(b)) is shown with the names of the gulls stopping at that day. Alternatively, the inspection process on a single gull can help to gain insights here, too (step 5).

The current selection of stopover(s) may be incomplete or inconclusive, so users can refine the selected stopover(s) (step 4). By adding another stopover to the selection, users define a sequence from an origin, the first selected stopover, to a destination, the most recently selected stopover. In the stopover aggregation (2) of Figure 1, we show a stopover sequence from Spain/Portugal to the Netherlands/Belgium. A deselection of stopovers is supported and restricting the selection to a specific gender is supported too, so that

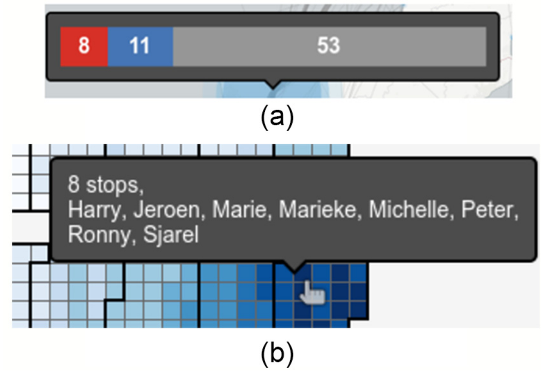


Figure 5. (a) Hovering on a stopover shows a tooltip with the amount of gulls at that stopover and their genders. (b) Within the calendar, a tooltip shows the number of stops at this day for the selection and the names of the gulls stopping at that day.

users can investigate gender-specific differences (see Figure 7). Using the slider for the aggregation level, users are able to adjust the granularity of the stopover aggregation.

Additionally, users can define a time range wherein the sequence of stopovers must lie (see Figure 6). This imposes a temporal restriction in which each individual within the selection must have at least one sequence of stopovers matching to the selection sequence, a stopover at the origin after or during the start date of the time frame, and a stopover at the destination before or during the end date of the time frame. On a single stopover selection, the origin and the destination then coincide.

It is also possible to define a subgroup manually from the current selected entities. The users apply steps 3 and 4 until they are satisfied with their findings. Eventually, they discern the individual trajectories from the result set of moving entities (step 5).

The analytical process to inspect a trajectory of a single moving entity is defined by selecting the individual first and investigating the movements of the trajectory on top of the density map. The visual encoding can be altered to exploit geographical characteristics and stopovers along the trajectory.

Evaluation

To assess the effectiveness of our visual analytics tool in terms of strengths and weaknesses, we applied it to a data set of migrating gulls over a period of 3 years and evaluated the visual analytics tool by interviewing an expert user in a 2-h session.

Our domain expert (one of the co-authors) has a background in ecology and studied bird migration in the past. He has seen a previous prototype of our

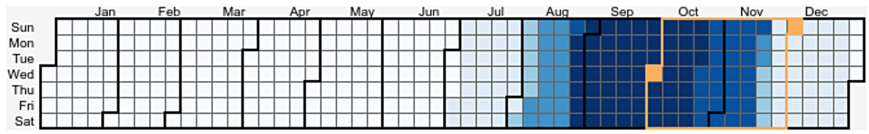


Figure 6. A selected time range from October to November within the calendar view at a stopover.

approach, but he did not experiment with the visual analytics tool nor did he study the used data set beforehand.

We conducted the evaluation with the expert user in three phases. First, we gave instructions and explanations on the visual design, the user interaction, the computation of stopovers, and the analytical features. This phase took 25 min. After that, we dealt with specific analytical questions on the data set. We developed a catalog of analytical tasks, see Table 3, covering different facets of the tool and linking ecological research questions and the requirements for our approach, see Tables 1 and 2 in section "Problem definition and requirement analysis." In the final phase, we asked to apply our tool on an area of interest and then asked questions on the usability of the tool. Those reflections will serve as a discussion of our approach.

Data set of migrating gulls

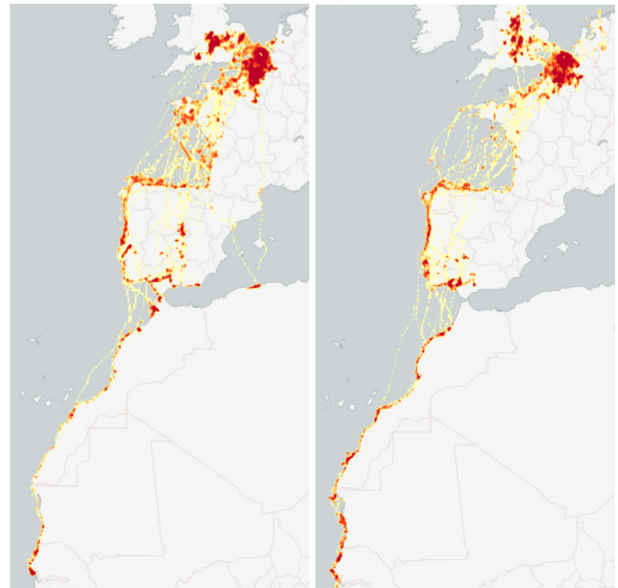
The Lesser Black-backed Gull is an interesting and challenging species to analyze, since it has a broad diet and can feed on many resources (both terrestrial and marine), can fly efficiently in many weather conditions, and can rest on both land and sea. As a consequence, migration can take place across almost any landscape, and foraging is possible almost anywhere along its migration route. This species generally adopts a fly-forage migration strategy which avoids carrying loads but switches frequently between flying and feeding instead.[28]

We applied our visual analytics tool to a comprehensive data set collected by Stienen et al.,[41] investigating our ecological research questions on the Lesser Black-backed Gull. Stienen et al.[41] collected almost 2.5 million data points from 101 gulls. This data set is unique with respect to other gull data sets, since in other data sets only 10–20 individuals were tracked over two to three seasons in a coarse resolution, whereas individuals in this data set were tracked over 3 years in a much higher resolution. The gulls have been tracked for at least 10 days, and more than half of the gulls had locations for more than 100 days. This data set contains 75 Lesser Black-backed Gulls and 26 Herring Gulls. Since the Herring Gulls stayed at their breeding site, we have focused on the Lesser Black-backed Gulls in

Table 3. Analytical tasks that the expert user completed, which have been linked to ecological research questions and the requirements of our approach.

	Analytical task	Completion time	Ecological research questions	Requirement
A1	Find a stopover with a lot of gulls.	2 min	E4	T1, T3
A2	Which gulls are migrating from England to North Africa?	2 min	E4, E5	T1, T4
A3	Find a stopover with only female gulls.	2 min	E2, E4	T4
A4	Which gulls are migrating from Brittany (North France) to England during the period beginning June 2013 until end of June 2014?	3.5 min	E3, E5	T1, T2, T3
A5	Find a gull with interesting movements during nighttime.	9 min	E1, E5	T1, T2, T3, T4
A6	Which gulls are migrating along the coastline in South France?	4 min	E6	T1, T3, T4
A7	Find a stopover with more female than male gulls.	1 min	E2, E4	T4
A8	Which gulls are migrating from the breeding spot via North Spain to England?	2 min	E4, E5	T1, T4
A9	What has been the migration strategy of gulls visiting Madrid (Spain)?	5 min	E1, E2, E5	T2, T4
A10	How would you describe the migration route of gulls at the stopover in Gibraltar?	2 min	E1, E2, E5	T2, T4
A11	Describe the migration pattern over France.	5 min	E2, E3, E4	T1, T2, T3

our study. Their breeding sites were at the Belgian and Dutch coast, and during autumn the Lesser Black-backed Gulls migrated to the South of Spain, Portugal, and North Africa. Further details about the data set and the ecological studies that it supports are provided in Stienen et al.[41]



(a) (b)

Figure 7. Density maps for (a) females and (b) males reveal that females stop more often in Brittany, the most northwestern part of France, than males and that males tend to take larger detours during migration along the sea.

Expert user evaluation

All analytical tasks (see Table 3) were completed successfully. The ecologist completed most of the tasks within 2–5 min. However, A5 consumed 9 min in total, since this task heavily relied on several inspection processes and contemplating on the meaning of "interesting movements during nighttime."

We discuss the analytical tasks sequentially from top to bottom and elaborate on the differences in solving similar tasks as well.

To accomplish A1, the expert user inspected first different aggregation levels (T1) and zoomed at different levels within the stopover aggregation visualization. Next, he investigated the whole map by hovering several stopovers (E4, T3). Eventually, he identified the stopover in North of Spain and Portugal as a stopover with many gulls.

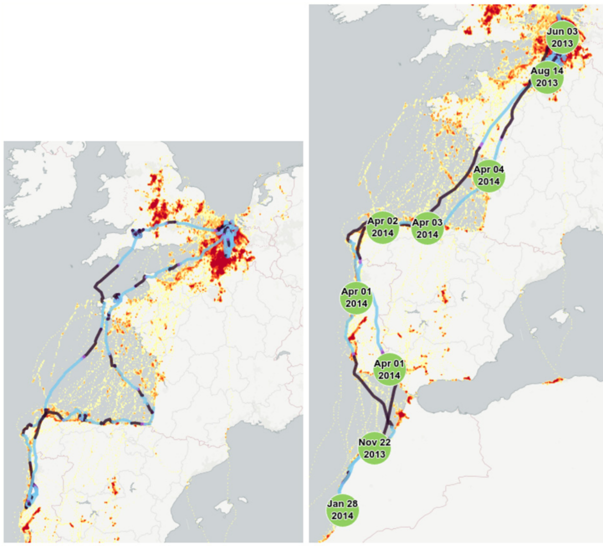
A3 and A7 are similar tasks to A1, but the expert user did not use the aggregation slider. In A3, the user browsed over a couple of stopovers (E4) and spotted three stopovers in Brittany (North France) where only female gulls stayed (E2, T4). The expert user instantly recalled the stopovers from A3 in A7 and started to hover on several stopovers (E2, E4) until he found suitable stopovers in Brittany, England, and other parts of France, where more females than males stopped. He implicitly assumed there might be stops where there are more males than females (compare Figure 7).

To solve A2, the user directly selected the singleton stopover in England and noted that there are multiple stopovers in North Africa (E4). He used the slider for the aggregation level after that until there was only one stopover in North Africa left and selected this one as a destination (E5). Subsequently, the domain expert inspected the individual trajectories of the selection using this route (T1, T4).

A4 differs from A2 by adding a temporal constraint (E3, T2) on an origin-destination selection (E5). The domain expert wanted to use the aggregation slider similarly as in A2 for Brittany, but he did not find a unique way, since, if the clustering is too coarse, there is no stopover in England, and, if it is too fine, there are many stopovers in Brittany. The expert user accepted this trade-off and selected a sequence from England to the largest stopover in Brittany (T1). Next, he restricted the time frame within the calendar appropriately (T2), enumerated the gulls Harry and Sanne, and expressed that he wanted to select a region, Brittany as such, since he thought that he might miss some birds from other stopovers in Brittany. In A8, he shared the same desire to select a region after selecting yet another origin-destination pair, over an intermediate stopover in this case. Thus, a weakness of our approach is that we do not support a selection of a region.

Within A5, finding a gull with interesting movements during nighttime (E1), the user investigated at first the complete list of gulls by hovering them (T4). He was surprised that he could not see long trajectories sometimes. After discovering that half of the gulls stayed at the breeding spot, he wanted to select all gulls from all stopovers excluding the breeding spot (E5, T4) to investigate night movements outside of the breeding spot, since he presumed that those motions at the breeding spot are probably due to human interference. He moved through the list of gulls further, while noticing that selecting a stopover (T3) might be more effective, until he found gull Angel with a suitable long trajectory (E1) (see Figure 8). Then, he defined "interesting" in this context as distinctively different patterns in travel duration, speed, or trajectory shape between a series of consecutive days and nights (E5).

He assumed implicitly that gulls can float on the sea and be moved by tidal forces, as known from the



(a) (b)

Figure 8. Overview of the movements of gulls: (a) "Angel" and (b) "Eric." Both show considerable movement at night in France, Spain, and Portugal.

literature.[8,29] So these patterns were not considered as interesting per se, only when occurring at a different rate during day and night.

We classified the locations of a trajectory based on the time stamps and the corresponding geographical position as twilight, daytime, or nighttime (see section "Computational methods") and visualized this on the trajectory. So the visualization tool does not provide direct information about the actual duration of daytime and nighttime periods for a given date. In order to compare trajectory lengths between day and night straightforwardly, the domain expert sought for periods around equinox, since daytime and nighttime are then almost equally distributed (based on this the ecologist did, for example, skip a stopover of Angel in November, since then the nighttime is longer). He switched to another individual, Eric, subsequently, and zoomed out to get an overview of the overall movements (E1). After some panning and zooming, he found interesting that Eric, as many others, traveled as much during the day as during the night. He hypothesized that there might always be light at the landmarks available to navigate.

While the trajectories of Angel and Eric (see Figure 8) do not show much difference between the lengths of day/night stretches, there are enormous differences in the lengths of the tracks between Angel and Eric. This poses whether Angel is coping with adverse wind conditions (during the segment with short distances per day/night) while Eric would have strong wind support.

In A6, the user investigated the coastline of the Atlantic ocean and selected the stopover covering this area in South France. After wondering whether he missed some gulls within the selection, he inspected the individual gulls (E6, T4). The domain expert distinguished between gulls traveling up north at the coastline, Anke, Sjarel, and Hilbran, southwards, Roxanne and Jasmin (partially), in both directions, Lea, or not all, Joke and Ian. To investigate the migration of gull Marie, a partial coastline migration, he needed to zoom further to obtain a higher resolution for the text labels of the stopovers. He wondered whether he covered all of them, and hovered and selected other stopovers south of the previous selection (T3). Eventually, he noticed that those gulls are a subset of the previously analyzed gulls.

A9 and A10 dealt with migration strategies at specific locations (E1), Madrid, and Gibraltar. To obtain a more detailed aggregation, the domain expert used the slider in A9 first. Then, he noted movements within the density map in Madrid, and selected the stopover at Madrid, enumerated Joke, Lea, and Michelle, and noticed that only females visited the stopover (E2, T4). The expert defined a migration strategy for a group or an individual by identifying the most southern stopover during their/its migration (E5). By inspecting the individual trajectories, he traced that Madrid was the most southern stopover for Joke during the migration. Lea, in contrast, visited a more southern stopover and traveled a week later than Joke. The ecologist noticed instantly that Michelle has been recorded for 2 years and pursued different migration strategies. At the stopover in Gibraltar (A10), the expert user often switched between the individual trajectories of Annelies and Ella, the gulls stopping in Gibraltar, since he was interested whether both were visiting England (E5). He summarized that Annelies is migrating in a clockwise two-way migration (E1), and Ella migrates a year later than Annelies, visits England before migrating south (anticlockwise two-way migration) (see Figure 9). Then, the ecologist used date information within the green stopover of an individual trajectory (T2) to accomplish this. He remarked that both pursued a coastline migration, and he assumed that gulls learn to shortcut during migration through experiences based on learned landmarks and resources from previous visits at the coastline. In ecology, it is commonly believed that bird species gain a better navigation capacity over the years through learning experiences of the past. Therefore, older birds are assumed to have better navigation abilities.

The final analytical task A11 of describing the migration pattern over France was driven by the user's interest in whether a stopover is used in a spring or

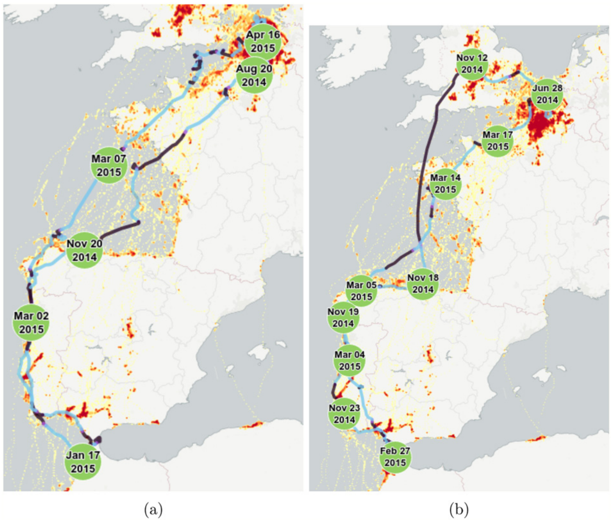


Figure 9. Trajectory visualization for (a) gull Annelies, which is migrating in a clockwise two-way migration, and (b) gull Ella, which visits England before migrating south (anticlockwise).

autumn migration. He reflected first on the meaning of a migration pattern and defined it as how many stops do the gulls make (E2) and at what time do they stop (E3, T3). To accomplish this task, he selected the largest stopover in Brittany and analyzed the temporal distribution of the selection (T2) within the calendar by toggling between showing the distribution of all stopovers or at the selected one. The domain expert concluded after investigating some other stopovers (E4) that the majority of stopovers in France is used during spring migration, such as the large stopover in Brittany, and little during autumn migration. He found this fact interesting because he assumed that the gulls would use them in both directions.

Reflections

After the task-oriented questions, we asked the expert user to apply the visual analytics tool on an area in which he is interested. He chose the stopover in England and wanted to know whether the visits to England were during, after, or before the breeding season (E2). He zoomed into the area around England and analyzed the movements within the density map (E1). Using the aggregation slider, he obtained the finest resolution of the clustering. Next, by toggling the temporal distribution between all stopovers and at the selected stopover, it became clear that the gulls visited England after breeding (E2), see Figure 10. The ecologist was surprised that the gulls visited England mainly in 2014. He hypothesized that this fact might be weather related and depended on wind conditions as well (E5). Using a temporal restriction from mid-July to September, he found that the gulls Harry, Jules, and Sanne were regular visitors of England. He was surprised that these gulls visited England that early (E3). The ecologist was also interested in investigating the difference among the entities (E6) between two time range selections. He used the tooltips within the calendar heavily to perform this comparison.

After that, we asked the ecologist to elaborate his reasoning behind the selection mechanisms. Regarding origin-destination selections, he would not select more than three stopovers at the same time, and he would like to select regions only at a very aggregate level. In particular, a gridded map to select areas of interest, such as cities or agricultural regions, would be beneficial to the expert user. This indicates that a hierarchical clustering, as we employed it, is crucial to aggregate stopovers dynamically on different spatial scales.

The time restriction within the calendar has been used by the expert user to isolate yearly cycles, and it was easy for him to identity pre-, post-, and peribreeding visits in this way.

Then, we asked him to outline his workflow after finding something of interest in our tool and to contemplate on the purpose of our approach. The expert user saw our approach as "a visual data querying tool" to generate subsets of the data set. He would select individuals showing a certain behavior and look at their space-time usage after finding something of interest. Subsequently, he would continue his research by computing some metrics, correlations, and statistics on the selected individuals in R to test a hypothesis on these groups. This confirms that our approach helps ecologists to explore and identify migration patterns visually and interactively before they proceed with non-visual analytical tasks.

By exploiting spatio-temporal relations between stopover sites, showing characteristics on a particular stopover, and visualizing spatio-temporal properties of an individual's trajectory in our tool, we enhance such analytical tasks for ecologists from manually extraction through custom prototyping—which would consume several hours—to a visual user interaction that takes a couple of minutes. Furthermore, an ecologist can identify individuals with a certain migration strategy and also infer migration routes of individuals[36] at a selected stopover site. Hence, our approach enables ecologists to visualize movement data sets with many movement tracks as well as individuals in the trajectory aggregation and thus speeds up the inspection process of discovering interesting stopover sites, individuals, or tracks drastically.

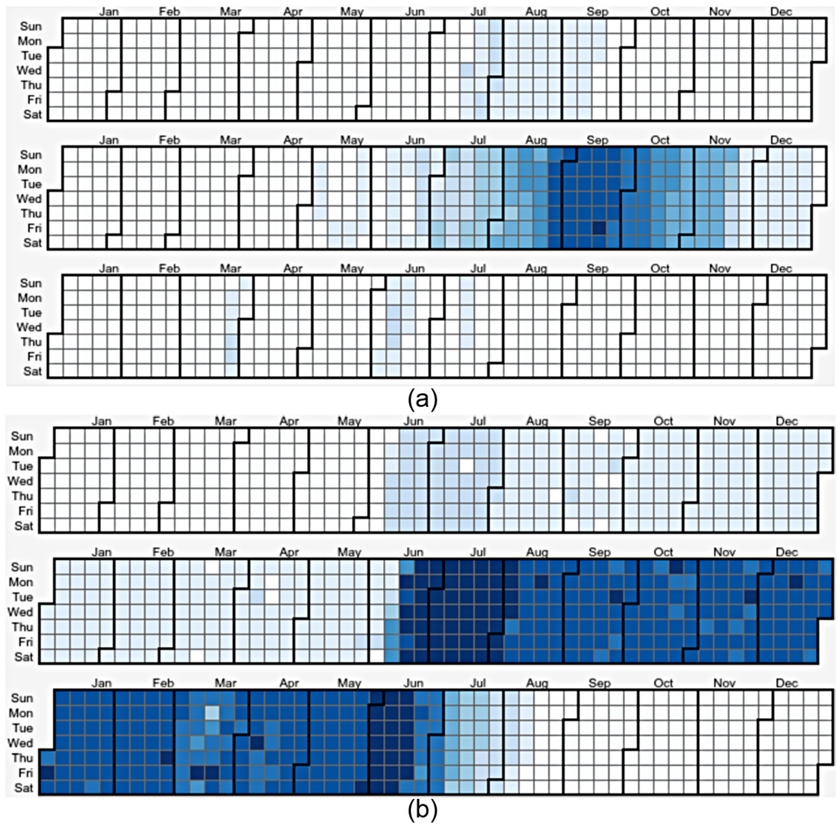


Figure 10. Stopovers with the calendar for all gulls visiting England (a) at the stopover in England and (b) at all stopovers.

Conclusion

This study presents a novel approach to visually explore migratory trajectory data. We computed an aggregation of stopovers from those trajectories and the movements between them in different spatial scales and visualized it interactively together with a density map and a calendar view. To investigate the tracks interactively, we enable the user to select stopovers and add restrictions on spatio-temporal properties of the selection. By applying our approach to a data set of 75 migrating Lesser Black-backed Gulls and by evaluating it with an expert user, we validated our visual analytics tool.

Our findings show that this exploratory visual analytics tool supports ecologists to investigate research questions on migration interactively. It especially enhances the identification of (groups of) individuals with a characteristic spatio-temporal migration behavior, but also allows discovering stopover sites with environmentally conditions which stand out.

As we used a single criterion on all stopover sites and individuals within our clustering, we think it is worthwhile to investigate varying rules per individual and per region. This would help to enhance the aggregation of stopovers to be more flexible.

As part of future work, we plan to integrate environmental data, such as information on land use, weather conditions (primarily wind), sea currents, and daylight, to facilitate the spatial exploration in the context of relevant variables.

Acknowledgements

We would like to thank our reviewers for their feedback.

Funding

M.K., M.A.W., and K.B. are supported by the Netherlands Organisation for Scientific Research (NWO) under grant no. 612.001.207.

References

1. Rutz C and Hays GC. New frontiers in biologging science. *Biol Lett* 2009; 5(3): 289–292.

2. Kays R, Crofoot MC, Jetz W, et al. Terrestrial animal tracking as an eye on life and planet. *Science* 2015; 348(6240): aaa2478.
3. Dingle H. *Migration: the biology of life on the move*. New York: Oxford University Press, 2014.
4. Van Wijk JJ. Bridging the gaps. *IEEE Comput Graph* 2006; 26(6): 6–9.
5. Sacha D, Senaratne H, Kwon BC, et al. The role of uncertainty, awareness, and trust in visual analytics. *IEEE T Vis Comput Gr* 2016; 22(1): 240–249.
6. Slingsby A and Dykes J. Experiences in involving analysts in visualisation design. In: *Proceedings of the 2012 BELIV workshop: beyond time and errors-novel evaluation methods for visualization*, Seattle, WA, 14–15 October 2012. New York: ACM.
7. Munzner T. *Visualization analysis and design*. Boca Raton, FL: CRC Press, 2014.
8. Slingsby A and van Loon E. Exploratory visual analysis for animal movement ecology. *Comput Graph Forum* 2016; 35: 471–480.
9. Spretke D, Bak P, Janetzko H, et al. Exploration through enrichment: a visual analytics approach for animal movement. In: *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*, Chicago, IL, 1–4 November 2011, pp. 421–424. New York: ACM.
10. Tory M and Moller T. Evaluating visualizations: do expert reviews work? *IEEE Comput Graph* 2005; 25(5): 8–11.
11. Alewijnse S, Buchin K, Buchin M, et al. A framework for trajectory segmentation by stable criteria. In: *Proceedings of the 22nd ACM SIGSPATIAL international conference on advances in geographic information systems*, Dallas, TX, 4–7 November 2014, pp. 351–360. New York: ACM.
12. Buchin M, Kruckenberg H and Kölzsch A. Segmenting trajectories by movement states. In: Yeh AGO, Shi W, Leung Y, et al. (eds) *Advances in spatial data handling*. New York: Springer, 2013, pp. 15–25.
13. Beyer HL, Morales JM, Murray D, et al. The effectiveness of Bayesian state-space models for estimating behavioural states from movement paths. *Methods Ecol Evol* 2013; 4(5): 433–441.
14. Gurarie E, Bracis C, Delgado M, et al. What is the animal doing? Tools for exploring behavioural structure in animal movements. *J Anim Ecol* 2016; 85(1): 69–84.
15. Lavielle M. Detection of multiple changes in a sequence of dependent variables. *Stoch Proc Appl* 1999; 83(1): 79–102.
16. Le Corre M, Dussault C and Côté SD. Detecting changes in the annual movements of terrestrial migratory species: using the first-passage time to document the spring migration of caribou. *Mov Ecol* 2014; 2(1): 19.
17. Madon B and Hingrat Y. Deciphering behavioral changes in animal movement with a “multiple change point algorithm-classification tree” framework. *Front Ecol Evol* 2014; 2: 30.
18. Thiebault A and Tremblay Y. Splitting animal trajectories into fine-scale behaviorally consistent movement units: breaking points relate to external stimuli in a foraging seabird. *Behav Ecol Sociobiol* 2013; 67(6): 1013–1026.
19. Zhang J, O’Reilly KM, Perry GL, et al. Extending the functionality of behavioural change-point analysis with k-means clustering: a case study with the little penguin (*Eudyptula minor*). *PLoS ONE* 2015; 10(4): e0122811.
20. Buchin K, Buchin M, Gudmundsson J, et al. Compact flow diagrams for state sequences. In: *International symposium on experimental algorithms*, St. Petersburg, 5–8 June 2016, pp. 89–104. New York: ACM.
21. Andrienko N and Andrienko G. Spatial generalization and aggregation of massive movement data. *IEEE T Vis Comput Gr* 2011; 17(2): 205–219.
22. Willems N, Van De, Wetering H and Van Wijk JJ. Visualization of vessel movements. *Comput Graph Forum* 2009; 28: 959–966.
23. Scheepens R, Willems N, van de, Wetering H, et al. Composite density maps for multivariate trajectories. *IEEE T Vis Comput Gr* 2011; 17(12): 2518–2527.
24. Slingsby A, Dykes J and Wood J. Using treemaps for variable selection in spatio-temporal visualisation. *Inform Visual* 2008; 7(3–4): 210–224.
25. Lu M, Wang Z, Liang J, et al. OD-wheel: visual design to explore OD patterns of a central region. In: *2015 IEEE Pacific visualization symposium (PacificVis)*, Hangzhou, China, 14–17 April 2015, pp. 87–91. New York: IEEE.
26. Lu M, Wang Z and Yuan X. Trajrank: exploring travel behaviour on a route by trajectory ranking. In: *2015 IEEE Pacific visualization symposium (PacificVis)*, Hangzhou, China, 14–17 April 2015, pp. 311–318. New York: IEEE.
27. Kölzsch A, Slingsby A, Wood J, et al. *Visualisation design for representing bird migration tracks in time and space*, 2013. Workshop on Visualisation in Environmental Sciences (EnvirVis), Jul 2013, Leipzig, Germany: 2013. http://openaccess.city.ac.uk/2384/
28. Klaassen RH, Ens BJ, Shamoun-Baranes J, et al. Migration strategy of a flight generalist, the lesser black-backed gull *Larus fuscus*. *Behav Ecol* 2012; 23: 58–68.
29. Shamoun-Baranes J, Bouten W, Camphuysen CJ, et al. Riding the tide: intriguing observations of gulls resting at sea during breeding. *Ibis* 2011; 153(2): 411–415.
30. Brehmer M and Munzner T. A multi-level typology of abstract visualization tasks. *IEEE T Vis Comput Gr* 2013; 19(12): 2376–2385.
31. Berthold P. *Bird migration: a general survey* (Ornithology series). Oxford: Oxford University Press, 2001.
32. Newton I. *The migration ecology of birds*. London: Elsevier, 2008.
33. Alerstam T and Lindström Å. Optimal bird migration: the relative importance of time, energy, and safety. In: Gwinner E (ed.) *Bird migration*. Berlin: Springer, 1990, pp. 331–351.
34. Alerstam T. Optimal bird migration revisited. *J Ornithol* 2011; 152(1): 5–23.
35. Hedenstrom A. Migration by soaring or flapping flight in birds: the relative importance of energy cost and speed. *Philos T R Soc B* 1993; 342(1302): 353–361.

36. Shamoun-Baranes J, Burant JB, Loon EE, et al. Short distance migrants travel as far as long distance migrants in lesser black-backed gulls *Larus fuscus*. *J Avian Biol* 2017; 48(1): 49–57.
37. Ritter J. An efficient bounding sphere. In: Ritter J (ed.) *Graphics gems*. San Diego, CA: Academic Press Professional, 1990, pp. 301–303.
38. de Berg M, Cheong O, van Kreveld M, et al. *Computational geometry: algorithms and applications*. Berlin: Springer, 2008.
39. Kruskal JB. On the shortest spanning subtree of a graph and the traveling salesman problem. *P Am Math Soc* 1956; 7(1): 48–50.
40. Forsythe WC, Rykiel EJ, Stahl RS, et al. A model comparison for daylength as a function of latitude and day of year. *Ecol Model* 1995; 80(1): 87–95.
41. Stienen EW, Desmet P, Aelterman B, et al. GPS tracking data of lesser black-backed gulls and herring gulls breeding at the southern north sea coast. *ZooKeys* 2016; 555: 115–124.