



**University of
Zurich**^{UZH}

Inferring migration routes of individual nocturnal bird migrants

GEO 511 Master's Thesis

Author

Mike Werfeli

14-720-635

Supervised by

Dr. Peter Ranacher

Dr. phil. II Felix Liechti (felix.liechti@vogelwarte.ch)

Faculty representative

Prof. Dr. Robert Weibel

29.09.2020

Department of Geography, University of Zurich

Abstract

Bird migration is a complex phenomenon, in which birds fly from their breeding grounds to the non-breeding locations of residence in fall and back in spring. To complete the migration, the birds need to adapt their behaviour. Therefore, the birds change the size of some organs to save weight and thus adapt their foraging behaviour. Additionally, the birds change their day and night activities to migrate at night. Bird migration is recorded by tagging the birds with tracking devices. In dependence of their body weight, the birds are equipped with a GPS or a multisensor data logger. The multisensor logger records environmental data like light intensities, air pressure and bird activity. These information are used for route inference. Especially the light intensities are used to estimate the positions a bird probably visited along a migratory route. The most probable migration route is computed by executing Bayesian inference. Migration routes are thus, inferred by executing MCMC techniques, where migration routes are proposed and tested on their probability of occurrence. The route with the highest probability is then saved as mean migration route. The error of migration route inference is $\pm 300\text{km}$, where the greatest uncertainties are found at positions estimated at solar equinox. With inclusion of additional environmental data, the position estimation is improved. The positions found on the mean migration route track are characterized by smaller standard deviations and confidence intervals, when wind and bird activity data are included to the route inference. Additionally, the inferences including wind and activity model represent the migration routes better than the migration route inferences applied so far. In addition, this thesis shows, that the wind data has an influence on the position estimation. These route inferences seem to follow the wind conditions, what results in non-breeding location positions being shifted to the West.

Key words: Bird migration, Migration route inference, Bayesian inference, geography, GIS, wind modelling

Acknowledgements

Firstly, I would like to thank Prof. Dr. Robert Weibel and Dr. Felix Liechti for offering me the opportunity to conduct this research. As well I would like to thank my supervisor Dr. Peter Ranacher, who guided me through each stage of the process and supported my research with many fruitful discussions. Also, I would like to express my great appreciation to my co - supervisor Felix Liechti, who supported my research at each stage. As well I would like to thank Dr. Simeon Lisovski for the helpful discussions and inputs at the beginning of this master thesis. Lastly, I would like to thank my friends, colleagues and family for their constant support.

Contents

Abstract	3
Acknowledgements	5
1 Introduction	9
2 Related work	13
2.1 Definitions	13
2.2 Data acquisition	14
2.3 Positioning by light	16
2.4 Migration route preprocessing	17
2.5 Route inference	19
2.6 Basic route inference (NO - No wind inference)	22
2.7 Wind support and wind drift	24
2.8 Summary and research gap	24
3 Methods	27
3.1 Work flow and procedures	27
3.2 Modelling activity	28
3.3 Wind data and wind model	30
3.3.1 Wind data base	30
3.3.2 Wind data query	31
3.3.3 Modelling wind	31
3.4 Model comparison	34
3.5 Wind and activity models	35
3.5.1 Wind model	35
3.5.2 Weak wind	35
3.5.3 Inefficient bird	36
4 Results	37
4.1 Wind model inference	37
4.2 Weak wind inference	39
4.3 Inefficient bird inference	40
4.4 Wind conditions found on the mean track	42
4.5 Non-breeding locations of residence	43
4.6 Positions estimated around the equinox	45
4.7 Entire migration route	47
5 Discussion and conclusion	49
6 Conclusion and future work	53
Appendix - A	67

1 Introduction

Bird migration is a complex phenomenon, where bird populations travel between sites visited in an annual cycle. Bird migration takes place between breeding grounds (where offspring are born and raised) and non-breeding locations of residence. In winter, the birds find more favourable conditions at non-breeding locations (Newton, 2008). Therefore, birds migrate from the breeding sites to the non-breeding locations in fall and return to the breeding grounds in spring (Fusani and Gwinner, 2004; Conklin et al., 2013; Liechti et al., 2015; van Wijk, Bauer and Schaub, 2016; van Wijk et al., 2017). Small nocturnal bird migrants are thus small birds which migrate at night from their breeding grounds to their non-breeding location. Especially for small nocturnal migrants physical and behavioural adaptations are necessary to be able to fly from breeding to non-breeding locations and back. Physical and behavioural adaptations are necessary for a safe and quick migration (Liechti, 2006; Hernández-Pliego et al., 2017; Williams et al., 2017). Before departure the birds build up muscles and fat reserves, which influence the distance the birds can fly. Whilst migration, the birds use up their fat reserves until these reserves reach a critical threshold. With the fat reserves used up, the birds stop their migration bout and refuel (Cooke, 1915; Alerstam, 1979; Pennycuik, 1998; Erni et al., 2002; Fusani and Gwinner, 2004; Liechti, 2006; Hernández-Pliego et al., 2017; Williams et al., 2017).

The bird migration is described by three main parameters, which are the safety of the bird from predators en route, the time it takes to migrate and the energy the bird has to invest to complete its migration (Alerstam and Lindström, 1990). Safety en route determines the survival of the bird. For successful migration, some birds adapt their day and night activity rhythms and perform migratory flights only at night. This adaptation brings the advantage of the migratory flight not interfering with daytime foraging habits of predators (Åkesson, 1993; Fusani and Gwinner, 2004; Liechti, 2006; Alerstam, 2011). Migration time however, describes the time the bird needs to complete the migration and therefore, consists of active flight time and the time spent at stopover sites. Birds use stopover sites to rest and refuel after long flight phases. The longer a flight phase is (e.g. over the Mediterranean) the more time the bird needs to rest and refuel (Erni et al., 2002; Alerstam, 2011; Conklin et al., 2013; Grishchenko et al., 2019). The invested energy, thus, correlates with the distance which the birds cover for migration. On migration some birds have to maximize their net energy intake but minimize migratory energy costs. On one hand, birds are able to reduce the costs of carrying heavy organs by building and rebuilding those between the migratory flights (Alerstam, 2011; Schmaljohann and Dierschke, 2005). On the other hand, birds benefit from favourable wind conditions and use these for energy optimization while migrating (Erni et al., 2002; Schmaljohann and Dierschke, 2005; Liechti, 2006; Alerstam, 2011; Conklin et al., 2013; Grishchenko et al., 2019).

These parameters described in the previous paragraph furthermore influence the choice of the migration route and the stopover sites. In order to understand the phenomenon of bird migration, route choice and stop over site selection, different tracking techniques are applied to record bird behaviour and the migration routes. Furthermore, tracking birds allows to analyse and understand the influence environmental factors have on bird migration. The most commonly used tracking techniques are GPS trackers attached to the birds, which directly record the birds' position and geolocators. The geolocators record light intensities from which the bird positions are inferred (Fusani and Gwinner, 2004; Bächler et al., 2010; van Wijk et al., 2018; Grishchenko et al., 2019). The devices are applied depending on the birds' body weight, where the device's weight should not be greater than 5% of the birds' body weight (Erni et al., 2002; Fusani and Gwinner, 2004; Alerstam, 2011; Lisovski et al., 2012). Because the geolocator is a lighter device than the GPS sensor, it

is attached to birds which are to light to carry a GPS sensor. The data recorded by the geolocators is used to reconstruct the most likely migration route the tracked bird was flying. Therefore, different algorithms have been developed, which estimate the bird position from measured light intensities. By adding simple movement models, and applying Bayesian inference, the most probable migration route is computed (Sumner et al., 2009; Bäckman et al., 2017).

European Hoopoes (*Upupa Epops*) tracked in this study are small nocturnal migrants (figure 1), which breed amongst other regions in the valleys of the southern part of the Swiss alps in the canton Vallais. In the



Figure 1: Image of a European Hoopoe (*Upupa Epops*) in flight.

(Image source: <https://www.vogelwarte.ch/de/projekte/prioritaetsarten/artenfoerderung-wiedehopf>)

winter period the birds stay in sub Saharan Africa (Sahel zone) and migrate to and from the non-breeding location in spring and autumn respectively. European Hoopoes are therefore tracked using geolocators (van Wijk, Souchay, Jenni-Eiermann, Bauer and Schaub, 2016; Bächler et al., 2010). So far, European Hoopoes are thought to migrate from Switzerland to sub Saharan Africa on different routes. The most frequently used fly ways are crossing the Mediterranean Sea via the Italian peninsula or on an axis between the Balearic Islands, Corsica and Sardinia. However, bird tracks recorded by geolocators are affected by position errors of 150 km or more (Barbaro et al., 2008; Bächler et al., 2010; Lisovski et al., 2012; Conklin et al., 2013; Liechti et al., 2015; van Wijk, Bauer and Schaub, 2016; van Wijk et al., 2018).

En route the birds encounter different environmental conditions, of which the most important is the wind. Wind conditions in the atmosphere impact the flying speed of a bird in relation to the ground (*ground speed*). With tail winds the speed of the birds in relation to the air (*air speed*) is increased and the birds are flying faster in relation to the ground (greater ground speed) than birds which are flying without tailwinds. Headwinds however, slow the ground speed of the birds. Birds are capable of compensating atmospheric conditions by adapting their flying direction in the air (*heading*) and adapting their airspeed (Liechti, 2006; Safi et al., 2013). Nevertheless, the birds are bound to the wind conditions en route. Winds along the flight-paths are thus, thought to influence and sign the migration route of the birds. However, the birds have a high site fidelity and tend to return to breeding sites where reproduction has been successful (Åkesson, 1993; Liechti, 2006; Blackburn and Cresswell, 2016; Aurbach et al., 2020). The effects of wind on bird migration routes are still not completely understood and proven. Furthermore, it is not fully understood how much a migrating bird is suffering from wind drift (the migration route being shifted in wind direction). To ana-

lyze the effects environmental conditions like wind have on migrating birds, wind data from the European Unions' Copernicus programme is included to a route simulation software (SGAT). Bird migration routes and stop over selection are crucial for understanding and developing strategies and ultimately protecting the birds and their habitats.

In order to answer these questions and contribute to a better understanding of the migration process of nocturnal bird migrant, a tool was developed, which includes real wind conditions and computes the support component it has on the migratory flight. The ultimate goal of this study is to develop a tool which can be used to demonstrate the effect, different wind conditions (velocities) have on the migratory route and thus build a model which can be used to better explain the bird migration routes.

2 Related work

Migration routes of small birds like Hoopoes can not be measured by tagging the birds with a GPS locator, because GPS locators are heavier than 5% of the bird's body weight (Erni et al., 2002; Fusani and Gwinner, 2004; Alerstam, 2011; Lisovski et al., 2012). Therefore, Hoopoes are tagged with multisensor loggers, which record different parameters of the bird's behaviour and environment. The most important measurements taken along the migration route are the light intensities. Light intensities are applied for position estimation (Hill and Braun, 2001; Lisovski et al., 2012). To compute the most probable migration route applying the measured data, a Markov Chain Monte Carlo approach is used. Thereby, for each position along the migration route the probability is computed, that the analyzed bird did visit the position.

In the following subsections, the methods applied for migration route inference are described in more detail, to make the process of bird migration more clear. Thereafter, the new contributions developed for this study, are described.

2.1 Definitions

To facilitate the explanation and discussion of bird migration and the methods applied to infer bird migration routes, a common vocabulary needs to be defined. The most important terms for bird migration describe the velocity or speed of the bird. Velocity is a vector used to describe the change of position of the bird (displacement) in a certain amount of time. The velocity vector has a magnitude and a direction. Speed describes the distance a bird was flying in a unit of time (distance travelled). Related

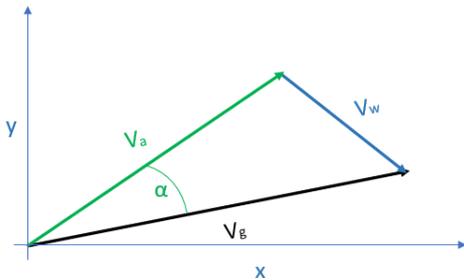


Figure 2: Sketch illustrating the different velocity vectors involved in this analysis. Figure adapted from (Liechti et al., 1994)

to speed and velocity are the terms of ground speed and air speed. Airspeed is used to describe the speed of the bird relative to the air. This speed is what the bird is actually flying in the air. Ground speed is the speed of the bird relative to earth's surface and thus, consists of the speed the bird is flying (airspeed) and the speed of the medium the bird is moving in, in this case the wind speed. Therefore, the ground speed is a sum of wind speed and the bird's airspeed. In figure 2 the introduced variables are introduced. \vec{v}_a is the vector of the airspeed, whereas \vec{v}_w is the wind velocity influencing the bird migration. \vec{v}_g is the vector of the

resulting ground speed (Liechti et al., 1994)

Two other important terms are the direction and the heading. The direction is defined as the angle between the x-axis of the coordinate system and the velocity vector of the bird and describes the orientation of the bird's displacement relative to the x-axis. The heading is the orientation of the bird in the air, necessary to fly to the desired direction. The heading might differ from the direction because the bird compensates wind effects by turning its heading towards the wind direction. Furthermore, the altitude describes the altitude above mean sea level at which the bird is flying.

Finally, the migration route is the route between the breeding site and the non-breeding location.

2.2 Data acquisition

This study analyses data recorded by *SOI-GDL3pam* logger geolocators (figure 3). These data loggers have been attached to many European Hoopoes and collected data of the birds' direct environment.



Figure 3: Image of a SOI-GDL3pam logger attached to the European Hoopoes to collect bird behaviour and environmental data (Dhanjal-Adams, 2019)

The data loggers are attached on the back of the bird similar to a backpack, to measure light intensities where as little shading effects occur as possible. The data logger is, thus, positioned in such a way that the birds' flying activity does not interfere with the device. Light intensities are recorded and stored in an interval of 2 minutes. This high temporal resolution was chosen to reduce the short term shading effect of clouds (Lisovski et al., 2012). Ambient air pressure was recorded on broader time intervals of 30 minutes. These longer time intervals were chosen mainly to save storage space and save energy and thus increase the geolocator's life span. Additionally these intervals were chosen, because the general weather conditions and

the bird behaviour do not change on faster time intervals. Same recording intervals were chosen for temperature measurements, following the same argument (Dhanjal-Adams et al., 2018; Liechti et al., 2018). The bird activity is measured by an accelerometer on the z -axis and is categorized into three classes executing PAMLR. Therefore, a dynamic activity threshold value is applied, categorizing the activities into low activity, high activity and continuous high activity. The data categorized as continuous high activity is related to bird migration, because of the Hoopoes' high flapping activity in flight. In addition to activity data, the respective recording timestamps stored on the logger are used to compute the time span in which the bird was on a migration flight.

Additionally, the flying altitude of the bird is measured with the pressure of the ambient air the bird flies in. The air pressure is not transformed to an altitude above ground, because the wind data is assimilated to air pressure levels.

Bird activity along the z - axis is recorded on a higher temporal resolution. Activity is measured with a frequency of 10 Hz on period of 3.2s every 5 minutes, where 32 values are generated. To store only one activity value, an algorithm computes the mean of these values on board and stores this value on the device (Bäckman et al., 2017; Liechti et al., 2018; Meier et al., 2018).

The birds were equipped with a geolocator at their breeding site in the valley of Valais, Switzerland. After the successful migration to their wintering site and back, the birds were recaptured and the data was read out. This data is stored in several files on an internal storage device, following a predefined data structure to facilitate the data reading and processing in R.

Wind data is retrieved from the European Centre for Medium - Range Weather Forecast (ECMWF), where assimilated and re - analysed wind data (*ERA-Interim*) are available on - line (Berrisford et al., 2011; European Centre for Medium-Range Weather Forecast, 2020). The wind data is provided on a daily basis, where four Coordinated Universal Time (UTC) time stamps and 37 different pressure levels can be selected. Wind data is aggregated and archived at 00:00, 06:00, 12:00 and 18:00 UTC on pressure levels of

1000 - 1 hPa (table 1). For this study, wind data was retrieved on atmospheric pressure levels between 1000 and 650hPa because of recorded bird behaviour. 1000hPa are found at sea level, where European Hoopoes are found to not fly on higher altitudes than 650hPa. These wind data contain information on wind

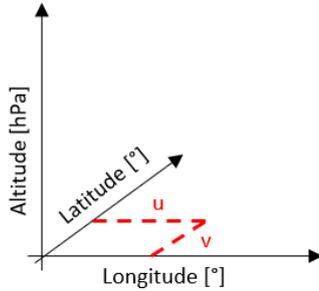


Figure 4: Illustration of the U and V wind components

speed in U and V direction (East ward and North ward wind respectively (see figure 4)). East ward winds are characterized by positive U wind speeds whereas West ward winds are characterized by negative wind speeds. Positive wind speeds in V direction signify a North ward wind whereas negative wind speeds signify South ward wind. The wind data is stored in *netCDF* files, from where it is read in to an R module. In this R module the data is rasterized and stored in a PostgreSQL data base (*DB*). The DB contains 30 fields, of which 4 contain spatial and temporal information. The other field contain U and V component for pressure levels ranging from 1000 to 650hPa.

Table 1 illustrates the pressure levels found in the ECMWF data sets, in the range in which the Hoopoe is assumed to fly (1000 - 650hPa). At these pressure levels, the wind speed measurements were aggregated. The air pressure ranges $P_{k\frac{1}{2}} - P_k$ which were aggregated to the indicated pressure level are illustrated in table 1. Finally, the assimilated pressure levels in table 1 are used in this study to retrieve the wind data at the pressure level the bird was flying. The pressure range between is assimilated to one pressure level (table 1)

Table 1: Subset of available pressure levels used in this study for wind speed retrieval and the pressure level ranges (Modellevels) which are aggregated to one pressure level. Table adapted from (Berrisford et al., 2011)

Pressure levels (hPa)	Modellevels	
	$P_{k\frac{1}{2}}$	P_k
1000	994.75	998.39
975	972.99	979.06
950	940.55	949.35
925	920.19	930.37
900	897.11	908.65
875	871.42	884.27
850	843.26	857.34
825	812.83	828.05
800	780.35	796.59
775	746.06	763.2
750	710.26	728.16
700	673.24	691.75
650	635.31	654.27

2.3 Positioning by light

Before route inference can be carried out, an initial migration path must be proposed. The measured light intensity patterns are used to compute the latitude and longitude of the bird positions. Based on the light intensity values over time, the daily positions are reconstructed executing astronomical equations. These astronomical equations determine the longitude from the midpoint of a sun rise and set event (local noon or local midnight). Latitude is computed from the day length, the time between sun rise and set events (Hill and Braun, 2001; Ekstrom, 2004; Lisovski et al., 2012; Lisovski and Hahn, 2012; Lisovski et al., 2019). Since the data logger device stores only light intensity values and the respective time stamps, it is necessary to define sun rise and sun set (figure 5). A threshold must be defined that it indicates the solar elevation angle when the first sun rays appear at the horizon. Therefore, twilight occurs between a sun elevation angle of 108° (nautical twilight) and $90^\circ - 108^\circ$ (Hill and Braun, 2001; Ekstrom, 2004; Lisovski and Hahn, 2012; Lisovski et al., 2019). For better discrimination of sun rise and sun set times and to take bird movement

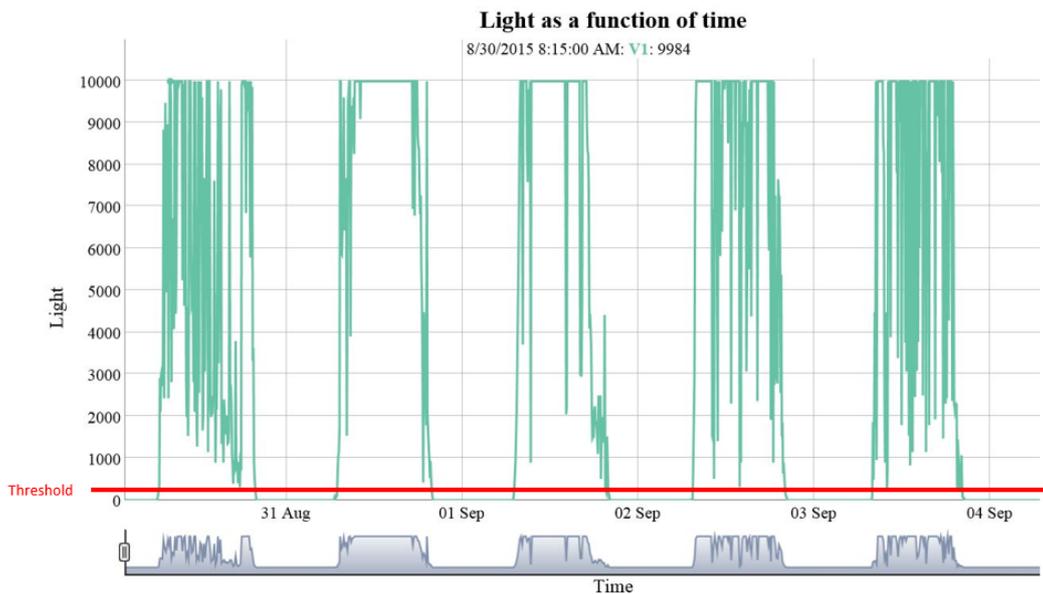


Figure 5: Illustration explaining how the threshold method is applied. The red line indicates the threshold value at which the light intensity (at arbitrary units) is defined as twilight

in to account, the threshold estimation is calibrated, which is performed on known bird locations like the breeding site. At the breeding site, sun elevation angle is more stable than en route and therefore used for calibration. The calibration (orange line on figure 6) is first applied on the light values measured at the breeding site 'in-habitat calibration' (dashed orange line on figure 6). In a second step, the twilight values are calibrated with the sun set times measured at the non-breeding locations (Lisovski et al., 2012; Lisovski and Hahn, 2012). The calibration is applied to improve the estimation of the solar zenith angle, which is used for position estimation. Calibration is especially applied to account for data errors which might occur in twilight annotation. Errors in twilight annotation result from shading caused by topography, vegetation or clouds, which reduce the light intensity measured (Lisovski et al., 2012; Lisovski and Hahn, 2012; Conklin et al., 2013; Liechti et al., 2013). The calibration can be further improved, applying the Hill-Ekstrom calibration, a template fit method based on the physical relationship between solar angle and near-surface light intensity (Hill and Braun, 2001; Ekstrom, 2004; Lisovski et al., 2012; Lisovski and Hahn, 2012; Liechti et al., 2015; Rakhimberdiev et al., 2015). These positions are used as starting point for route inference.

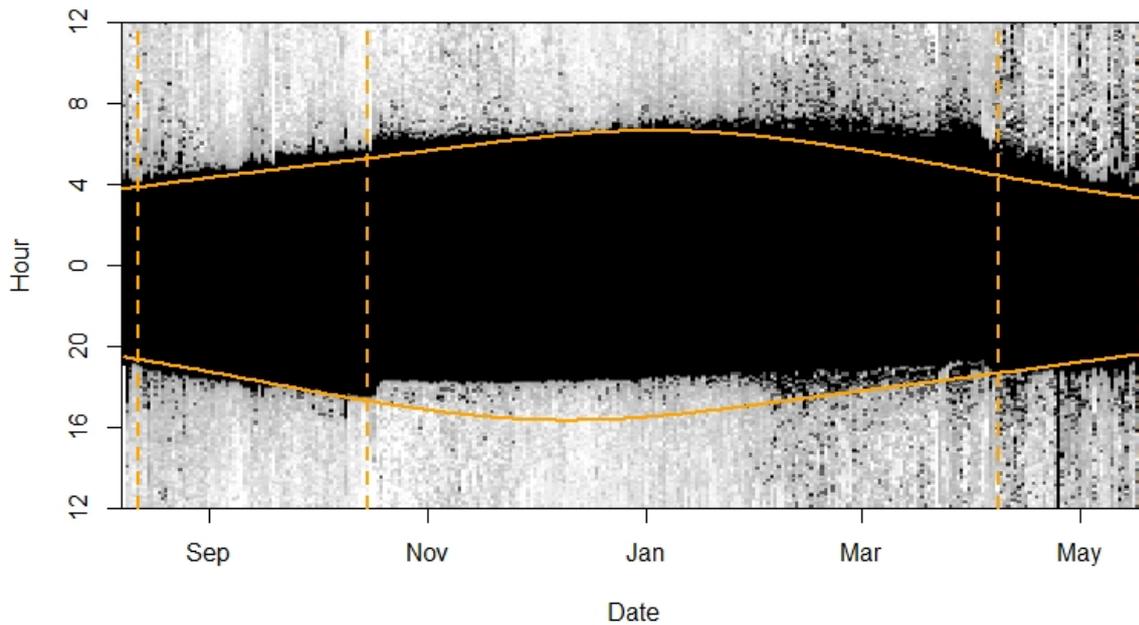


Figure 6: Light image showing day and night period for a bird equipped with the geolocator. The horizontal orange line is the calibrated twilight threshold, whereas the vertical orange lines are the start and end dates applied for calibration

2.4 Migration route preprocessing

The light positions are computed for each day along the migration route. Therefore, the light intensities are first differentiated into sun rise and sun set. Subsequently, the sun rise and sun set times are transformed into latitudes and longitudes (Sumner et al., 2009; Stutchbury et al., 2009). With each latitude and longitude pair a positions are built, which are then chronologically connected to build a migration route. This route does not account for any stop over sites, where the bird was not migrating but foraging. The first raw route, illustrates all positions recorded, while the bird was carrying the geolocator. Because of measurement errors, this route contains outliers created by false classification because of the shading of sun light which is caused by clouds or vegetation (Sumner et al., 2009; Stutchbury et al., 2009; Bächler et al., 2010; Liechti et al., 2018). A first route is then constructed by ordering the positions which where estimated with the light intensities chronologically. As a result, a migration route is generated which contains one position for each day the bird did carry the geolocator. Figure 7 shows a migration route constructed with raw positions.

These raw, unclassified and unanalyzed position contain very imprecise positions. Most prominently are the positions found to be in the middle of the Gulf of Guinea, several kilometres of the coast of Nigeria, Benin, Togo, Ghana and the Ivory Coast. Further, there are point positions in the middle of the Mediterranean, several kilometres of the coast of Italy (Sardinia), France, Spain and Algeria. In addition, there are two positions which can be classified as outliers, found in the Democratic Republic of the Congo and in Chad (figure 7).

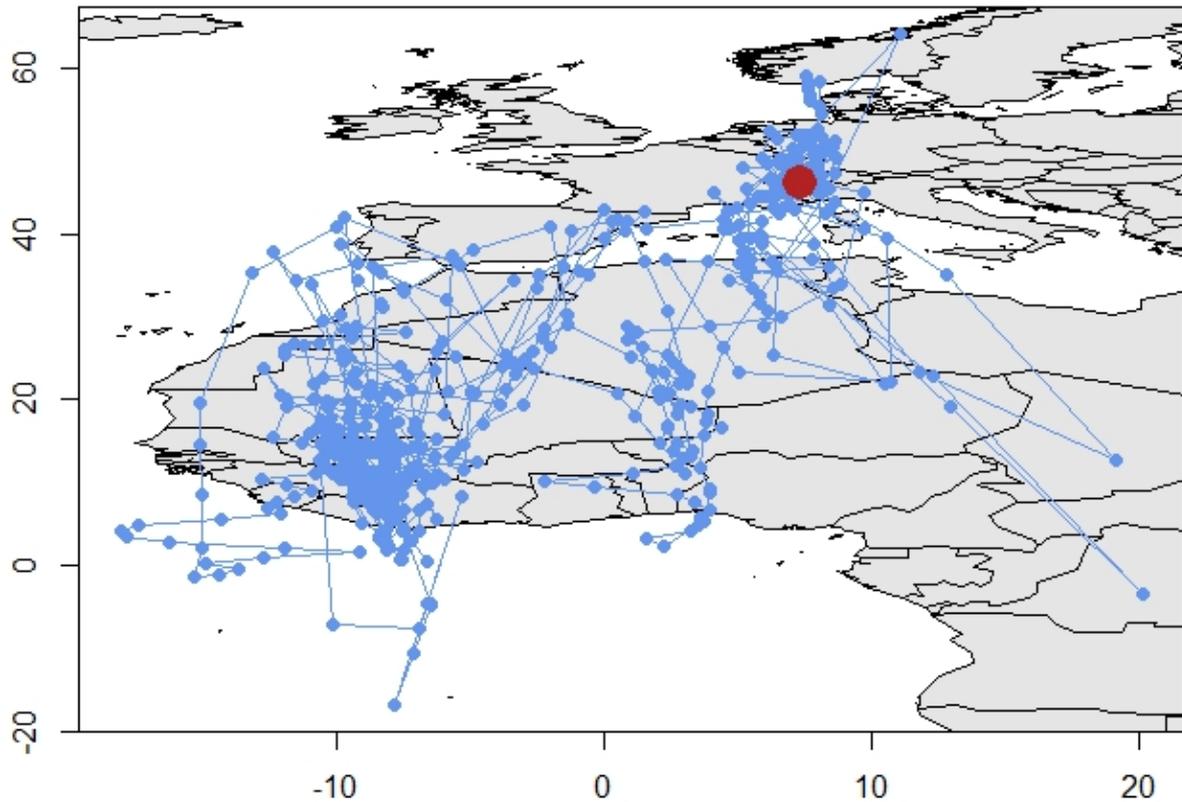


Figure 7: Route constructed given the positions estimated by the light intensities measured on the geolocator en route

After computing this first migration route it was necessary to simplify the migration route and reduce the number of point positions. Therefore, the measured positions which appear to characterize a stop over site are aggregated to one stop over position. Stop over positions are described by the mean value of the positions latitude and longitude respectively. Activity and air pressure data are used to aggregate these multiple positions to one single position. The aggregated positions are found to lie between two migration flight bouts and are thus additionally characterized by a break of 24h between two subsequent migratory flights.

Additionally, the fixed positions are introduced to the route. Fixed positions are position coordinates which are well known to be visited by the birds, like the breeding site where the birds did start and end their migration. These positions are introduced to force the simulation to head to a specific direction. This can be used for example if the non-breeding locations of a bird is known very well and the migration route computed in the inference algorithm needs to reach this specific destination. At fixed locations, the algorithm does not sample new positions, it computes the probability of the same position (Sumner et al., 2009; Lisovski et al., 2012, 2019).

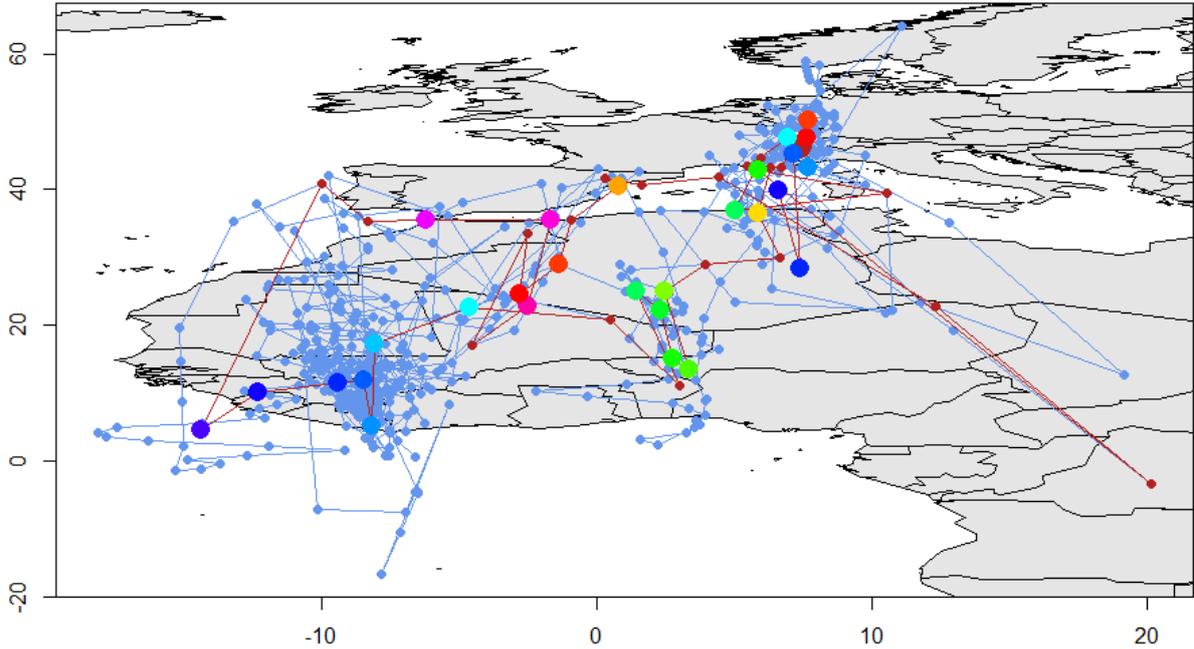


Figure 8: Aggregated positions which from a first chain which is used for route inference

Figure 8 illustrates the position aggregation step described in the previous paragraph. For visual validation the simplified route is overlaid on top of the light positions. The larger and coloured dots are positioned in the middle of point clusters and are thus representing the stop over sites. This route does neither contain any information of the wind found en route, nor any information concerning the birds' physiological capabilities. This route assigns the single positions probabilities, given the light data. These positions need to be adjusted and adapted to approximate the position the bird most probably visited.

2.5 Route inference

The first route with the aggregated positions is the starting point for route inference. In the route inference, all possible flight routes and their respective probabilities are computed. Each position along the migration route is described by a set of coordinate pairs and respective probabilities of occurrence, what is called posterior distribution.

To compute the possible migration routes, a Bayesian approach is applied to approximate the posterior distribution, wherefore three main elements are required: (1) a prior, (2) a likelihood and (3) a posterior. The *prior* distribution $p(\Theta)$ describes the knowledge of the parameters before any data is observed. The *likelihood* $p(y|\Theta)$ gives the probability of observing data y given a set of parameters Θ . The *posterior distribution* $p(\Theta|y)$ is computed, with the prior distribution and the likelihood, following the Bayes' rule, shown in formula (1) (Binz et al., 1981; Storrer, 2009; Sumner et al., 2009; Sheldon et al., 2013; McElreath, 2015).

$$p(\Theta|y) = \frac{p(y|\Theta)p(\Theta)}{\int p(y|\Theta)p(\Theta)d(\Theta)} \quad (1)$$

The posterior distribution $p(\Theta|y)$ describes the knowledge of the set of parameters after data y have been observed. Therefore, the Bayes' rule is a robust and useful mechanism to update the posterior distribution, based on observed data (Sumner et al., 2009; Demšar et al., 2013; Sheldon et al., 2013; McElreath, 2015).

Before the Bayes' rule is implemented the prior and the likelihood have to be defined. First, a **prior** is constructed, before the inference starts. Thereby, anything that is known about the bird migration route is used to construct the prior density distribution. The prior contains information about previous known general migration patterns, habitat preferences and preferred stop over sites (Hoopoes do not stop over water bodies). In this case, the prior is used to constrain the inferred positions along the migration route to be over dry land. To force the positions to be on land, a land mask was generated, where the probabilities of positions found to be on dry land are kept as they are, whereas the probabilities of positions found to be on the ocean are set to a very low number ($\sim 10^{-6}$).

After setting up the prior, the **primary location data** is set up. This is the first migration route computed with light positioning only, which is composed of a chronological ordered set of states (positions). This route (state) is then used for route inference, where Markov Chain Monte Carlo (MCMC) methods are applied.

Following the primary location data, the **likelihood** of the route is computed. The likelihood includes the wind data to the movement model, which is used to determine the probabilities of the different routes. The movement model thus, is executed to assign low probabilities to positions the bird is not able to reach under the given circumstances (Sumner et al., 2009; Sheldon et al., 2013; Lisovski et al., 2019).

Route inference by applying the Bayes' rule (formula (1)) is not solvable analytically because of the integral in the denominator. Therefore, it is necessary to apply other methods to compute the posterior distribution of the positions in the route and estimate the possible migration routes. As the migration route is described by a set of ordered positions the Markov Chain Monte Carlo methods are applied to estimate the unobserved bird positions probabilistically (Roberts and Stramer, 2001; Sheldon et al., 2009; Rakhimberdiev et al., 2015). To model bird migration, Hidden Markov Models (HMMs) are applied, which assume

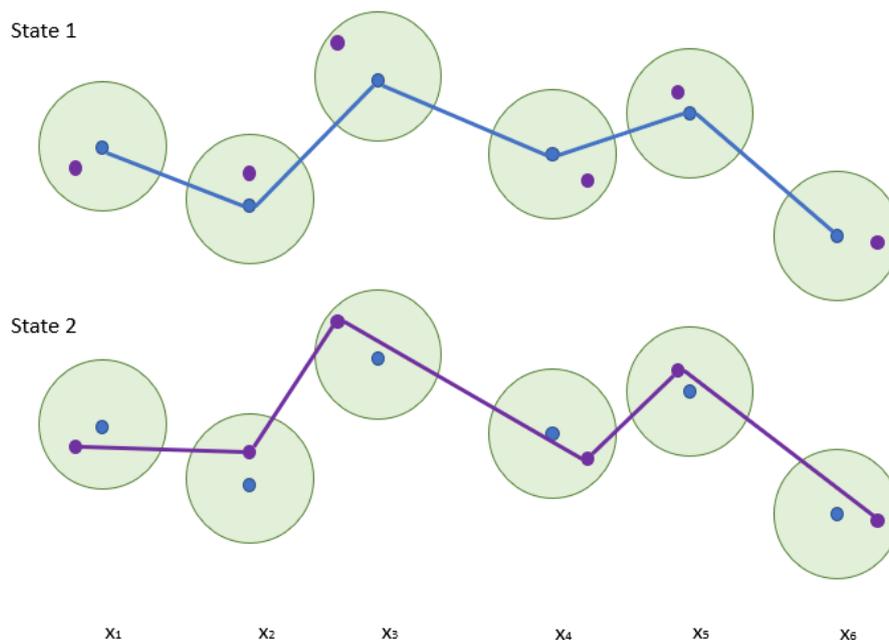


Figure 9: Schematic sketch of the MCMC method applied in this study for route inference

generative models for a sample path. Therefore, a sequence of states is randomly drawn applying a hidden experiment (Goodman, 1986; Sheldon et al., 2009; Demšar et al., 2013; Sheldon et al., 2013; Rakhimberdiev et al., 2015). The experiment is hidden as the bird positions along the migration track are not observed but are rather randomly drawn from an error shape built around each state in the Markov Chain. Each error shape represents the uncertainty of each state along the chain (green ovals around blue positions in chain 1 in figure 9). New positions (purple points in the green error shape in chain 1 in figure 9) are drawn randomly to generate a new chain (purple path in chain 2 in figure 9). The error shape is defined in such a way that all randomly drawn points can be described by two models: (1) a physical model (an observational model) and (2) a movement model (process model), respectively (Sheldon et al., 2009; Dean et al., 2013; Sheldon et al., 2013; McElreath, 2015; Rakhimberdiev et al., 2015). Therefore, different transition paths (flight ways) between different states (measured positions) are randomly generated and evaluated in regard to their probability given the observed data, where for each position in one chain the respective probability is computed given the light measurements, the birds' physiological capabilities and the wind data.

Markov Chain Monte Carlo methods are used to infer nonstandard, complex multivariate distributions. In this study an acceptance - rejection sampling algorithm was implemented and executed, the Metropolis - Hastings (M - H) algorithm. The M - H algorithm is an algorithm first implemented by Metropolis et al. in 1953 in a physical context and was later generalized by Hastings in 1970 (Roberts and Smith, 1994; Chib and Greenberg, 1995). The M - H algorithm was developed for Markov Chain Monte Carlo simulations, where the target distribution is known but the transition kernel to get to the target distribution is not. To get an approximation of the transition kernel, samples are drawn from the probability space of the posterior distribution (target density). Every time these samples are drawn, a transition kernel is constructed (iteration). The transition kernel converges to the target density for large number of iterations (Metropolis et al., 1953; Hastings, 1970; Roberts and Smith, 1994; Chib and Greenberg, 1995; Roberts and Stramer, 2001; Roberts and Rosenthal, 2001).

The Metropolis - Hastings algorithm starts at an arbitrary chain x and iterates n times. Being an acceptance - rejection sampling algorithm, the M - H algorithm is set up to be an "one block at the time" algorithm (Chib and Greenberg, 1995). Therefore, the algorithm creates a new random sample (a migratory route) and computes their probability given the available data (activity, bird velocity and wind velocity). After computing the new log probabilities, the probabilities of both the new and the old migration route are subtracted from one another. The difference is then compared to a probability value, which is drawn randomly from a uniform density distribution. If the difference is greater than the randomly drawn probability, the new position is accepted, otherwise it is rejected and the old position is kept in the new chain (Metropolis et al., 1953; Hastings, 1970; Roberts and Smith, 1994; Chib and Greenberg, 1995; Roberts and Stramer, 2001; Roberts and Rosenthal, 2001; Sumner et al., 2009; Lisovski et al., 2019).

The probability of the positions along the chain are estimated as the probability that the bird reaches a point P_x from a previous point P_{x-1} in the given flight time. Therefore, the necessary speed to travel the distance between the points P_x and P_{x-1} is computed by dividing the distance by the time the bird was flying. Subsequently, the probability that the bird was flying the necessary speed is computed (ν_n). Assuming there is no negative ground speed, the density distribution of the possible flight speeds is modelled by a gamma distribution. However, the gamma distribution is described by a shape parameter (α) and a rate parameter (β). These two parameters are computed by the wind model described before (see formulae (6) and (7)). Knowing the density distribution of the bird speeds on the analyzed segment, the probability of

v_n is computed. Thereafter, the flight speed probability is multiplied with the position probabilities and compared to the acceptance level.

After completing the iterations and converging the Markov Chain to the best possible approximation of the posterior distribution the mean route is plotted as solid line indicating the mean tracks. Furthermore, the percentile tracks based on the MCMC Chains are plotted as raster indicating the percentiles in different colors (Ruginski et al., 2016). Additionally, a location summary for each stationary position is provided, where the mean latitude and longitude, as well as their median, standard deviations, 2.5, and 97.5% percentile are listed. Latitude and longitude values and the respective 95% confidence intervals are plotted in chronological order for visual representation of the inference results. Additionally, each position's standard deviation is shown with error bars, thus illustrating the position precision.

2.6 Basic route inference (NO - No wind inference)

In the migration route inference implemented in SGAT so far the basic MCMC algorithm is executed. Therefore, the two required models are implemented. As described in section 2.5 the MCMC method requires two models, the physical model and the physiological model. In this route simulation, the physical model is applied to describe the bird positions en route estimated given the light intensities (see section 2.3). The physiological capabilities of the birds are described by the flying speed en route. In this first approach on each route segment the physiological capabilities are mathematically described the same way and are not taking

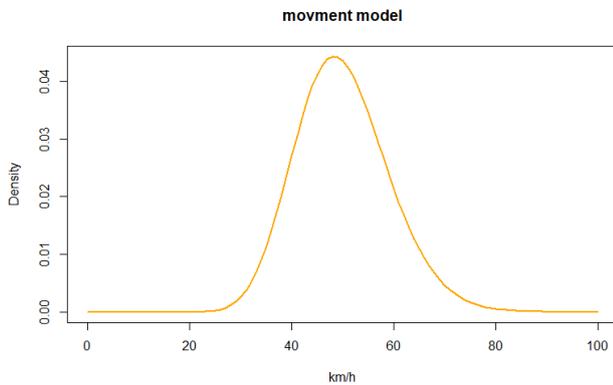


Figure 10: Density distribution given the fixed alpha and beta values applied on each route segment

flight time and distance into account. Therefore, the gamma density distribution of flying speeds is described by the same parameters on each route segment. The shape (α) and rate (β) parameters are set in such a way, that the highest flying speed probability is at a mean flying speed of $13.89 \frac{m}{s}$. The α and β parameters are therefore set to: $\alpha = 30$ and $\beta = 0.6$. This flight speed however, is thought to represent high flight speeds measured on an escape flight (Bolch and Bruderer, 1982; Liechti and Bruderer, 1995; Bruderer and Boldt, 2001). Figure 10 shows the plot of the applied movement model (Lisovski et al., 2019). On the x - axis, the speeds are plotted against the respective probabilities on the y - axis. The plot illustrates the gamma model

applied on each route segment, to estimate how probable the proposed route segment is, given the very basic gamma distribution. This gamma distribution is included to the position likelihood and is developed to reduce great distances, which the bird physiologically can not reach (Sumner et al., 2009). Currently, the route inference is computed applying the before mentioned gamma density distribution (No wind (NO)) and neglecting the new developed activity and wind model.

Given the movement model plotted in figure 10, a new route is inferred. Outliers found in the first chain are eliminated because of the low probability assigned to these positions. Figure 11 illustrates the most probable route a bird is flown given the light intensities and the basic flying speed density distribution. In this figure, the solid line represents the mean route, hence the most probable migration route. Stop over

sites are represented by coloured points and are labelled alphabetically, to indicate the chronology the bird was approaching them. The error bars indicate the uncertainty related to the latitude and longitude values of each position and are thus strongly related to the uncertainty found in the data and in the models describing the movement. Most of the outliers mentioned in section 2.4 are classified as improbable positions.

Table 3 contains arrival and departure dates at stopover sites and the corresponding latitude and longitude values. Given the data applied in this route inference, the locations of residence in Africa are found with great possibility in the South Western region of Mali and in the South Eastern region of Mauritania, where the bird arrived mid October (18.10) and staid until end of March (29.03.). At the wintering site in the Western Sahel zone the standard deviation of the inferred position latitude values is computed to be between 0.3° (12.01.) and 1.8° (23.01.), whereas the standard deviation of the longitude is computed to be between 0.95° (12.01) and 2.13° (23.03.). These positions correspond to the positions i, j and k on the map in figure 11. Stop over sites visited for a short period of time are less accurate than long visited, what is expressed in great standard deviations ($1.30^\circ - 2.31^\circ$ for longitude values and $1.09^\circ - 2.76^\circ$ for latitude values).

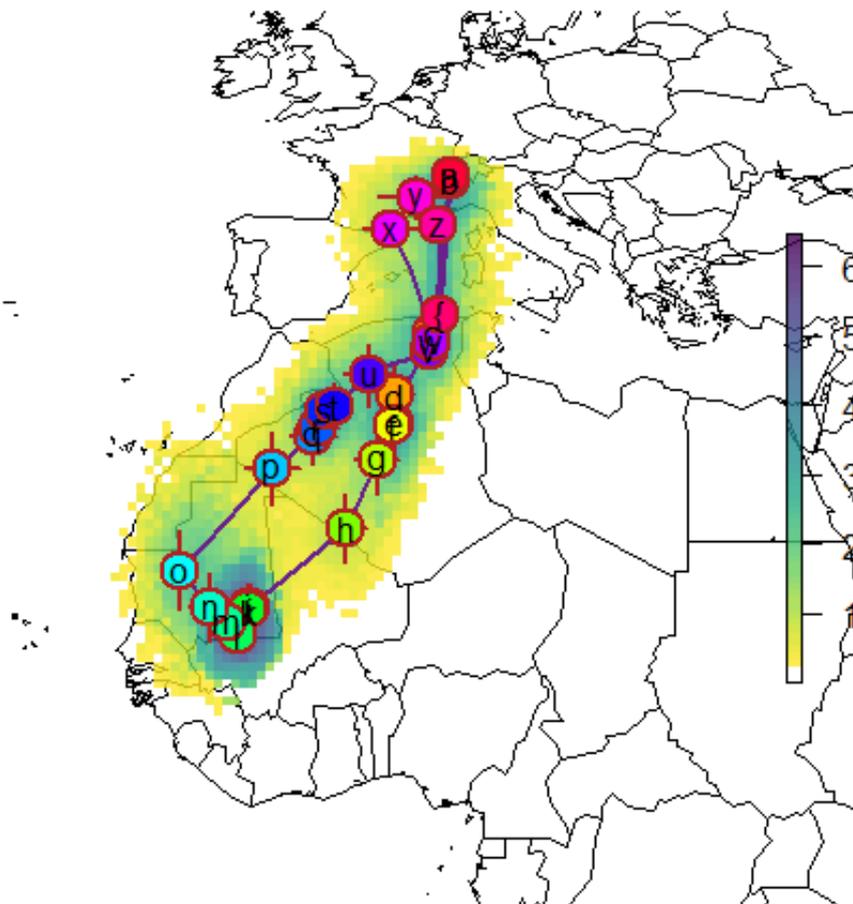


Figure 11: Result of route inference by applying MCMC method only, neglecting bird activity and wind influence

Additionally, the 95% confidence interval of the estimated latitude and longitude values is computed

to illustrate the range in which the estimated latitude and longitude values are spreading. Therefore, the 2.5 and 97.5 percentiles of the latitude and longitude values are computed for each position in the Markov Chain. The mean of all confidence intervals in the chain is 4.35° for longitude coordinates and 4.85° for latitude coordinate respectively. The mean confidence interval covers 1.85° for longitude values and 0.12° for latitude values respectively.

2.7 Wind support and wind drift

In literature, the different effects wind velocity has on a migrating birds are modelled in different ways. The most common wind modelling approach is described by Liechti et al. (1994), where the relationship between bird flight heading and air speed, and wind speed and direction are summarized to a flight track with ground speed. In these models, the wind support and wind drift are included in one model, showing the resulting flight direction and ground speed resulting from the influence of the wind on the migrating bird (Alerstam, 1979; Liechti et al., 1994; Liechti, 1995, 2006). Therefore, the wind vector and the flight vector of the bird are added to a resulting vector. The vector addition is executed, by adding the bird flight vector to the wind vector. Therefore, the direction of the migration bout changes, because a not complete wind drift compensation is assumed.

2.8 Summary and research gap

The approach implemented in the SGAT model (Lisovski et al., 2019) the bird positions are inferred applying positioning by light and by modelling the movement of the bird with a gamma distribution. Bird migration route positions are estimated by applying astronomical equations, wherefore sun rise and sun set times are then used to compute latitude and longitude values (Hill and Braun, 2001; Ekstrom, 2004; Bäckman et al., 2017). Migration route inference is computed by executing a MCMC method. Therefore a physical and a movement model are required. The physical model is implemented with a gamma distribution around the mean velocity of $13.89 \frac{m}{s}$, whereas the physical model describing the positions is modelled with a error function. Within the inference the probability of a proposed chain of positions is computed and compared to a threshold probability value (Sumner et al., 2009). During the iteration steps executed, those positions which are assigned a greater probability are kept for the next iteration step. Following the law of large numbers, with this algorithm, the most probable route is inferred after a great number of iterations (Chib and Greenberg, 1995; Roberts and Rosenthal, 2001; Sumner et al., 2009; Storrer, 2009; Lisovski et al., 2019). However, the position errors caused by the uncertainties of the light positioning techniques ($\pm 300 km$) are rather great. Route inference implemented in SGAT however, does neither include wind conditions nor bird activity data.

Therefore, different questions emerge:

RQ1: How do endogenous and environmental factors affect the migration route inference?

RQ2: Do activity and wind data improve the migration route inference?

RQ3: Do wind conditions en route change the stop over positions in the route inference?

To answer the research questions different hypothesis where formulated. First, it is expected, that the activity data and the wind data on one hand improve the migration route inference. On the other hand,

including the wind data is expected to influence the migration route. The bird is hence, expected to follow the wind conditions it encounters en route. Wind conditions are thus, expected to change the stop over position and the bird is expected to drift a bit with the wind.

3 Methods

To improve the positions found along the migration route, two new models are developed and presented in this section. First, the activity model is discussed, where the flight activity data of the birds is included to better estimate the distances the bird probably did fly. Secondly, the wind data is included to the movement model, to estimate the wind support a bird did encounter on a migration bout.

3.1 Work flow and procedures

The work flow of this thesis is illustrated in figure 12. After data acquisition and bird recapture, the data is processed using two R packages (PAMLR and SGAT). With these two packages the most likely bird migration routes are inferred applying Bayesian inference. Bayesian inference is a statistical method to infer a probability distribution (*posterior distribution*) from observed data, given prior knowledge about the observed phenomenon.

To include the activity and wind data for better position estimation, two new models are generated. One model is used describe the bird behaviour and especially the distance the bird most probably did fly. To model the distances flown, it is necessary to know when a bird was on a migratory flight and how long the bird was flying (Lisovski et al., 2019). The second model is used to describe the wind conditions a bird encounters en route. The wind directions and forces are measured with uncertainties, which are used to define the probability a specific wind direction and speed occurred in a particular moment of the bird migration.

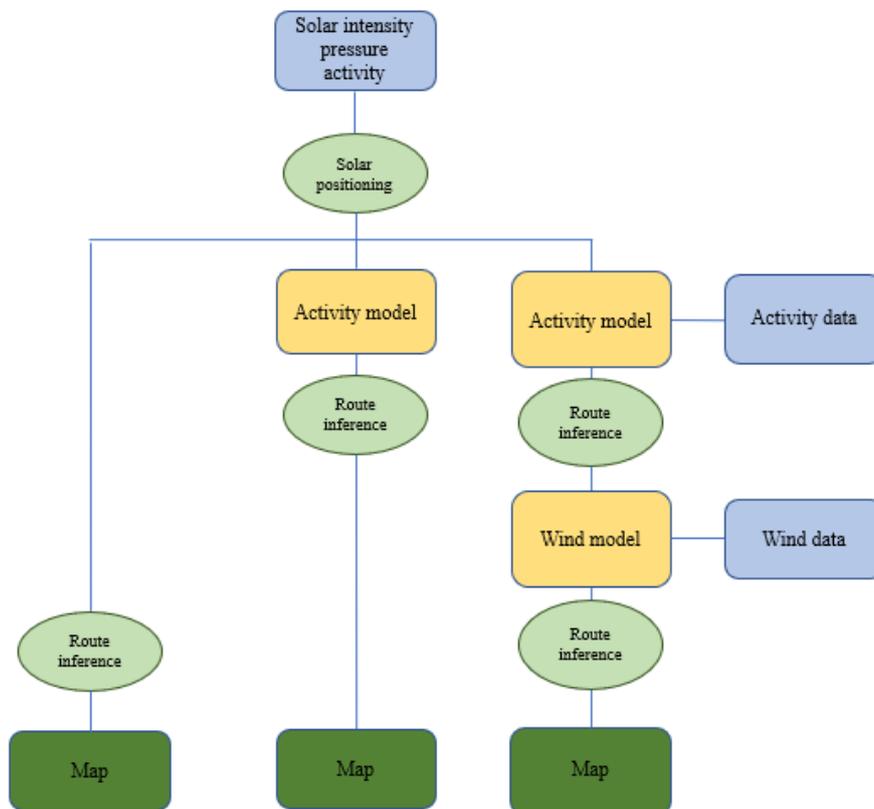


Figure 12: Illustration of the general work flow of this master thesis. Where the yellow squares represent the models applied for route inference, the blue squares represent the data needed and the green circles represent the algorithms applied. The final output is represented by green squares

These two models are used to infer the total probability of positions along a possible migration route. Including activity and wind information allows to better estimate the possible migration routes and stop over sites a bird chooses to use. Initially, the bird positions are estimated executing solar positioning in SGAT. According to the certainty this positioning method has, these positions are assigned probabilities, too. The inclusion of the two models allows to update the probability of each position without deleting any position. With these two models added to the solar positioning, an estimation of the wind influence can be made. Furthermore, the birds' flying speed in relation to the ground (*ground speed*) can be made on each migration segment along the route.

Applying MCMC technique allows to compute a desired posterior distribution by random sampling in the probability space (Raftery et al., 2007; Sumner et al., 2009). Thereafter, the possible positions along a migration track are plotted on a map, representing the likelihood of the bird flying along these positions.

After data preprocessing in PAMLR, the migration routes are inferred applying SGAT. With SGAT the migration routes are proposed and evaluated, given the light intensities, the measured activity data and wind data.

3.2 Modelling activity

The bird activity is a very important parameter for migration route inference because it conditions the distance the bird probably flew. After categorizing the activity levels, the time spans of continuous high activity are used to estimate the distance the bird was probably flying. To transform the time period in which the bird was migrating to a distance (μ_d) which the bird was possibly flying, a constant mean flying velocity must be assumed. In this study a velocity of $13.89 \frac{m}{s}$ is assumed for migration flights, which is multiplied with the migration time (formula (2)) (Bolch and Bruderer, 1982; Liechti, 1995; Bruderer and Boldt, 2001; Nilsson et al., 2014).

$$\mu_d = \text{activity} \cdot \text{mean bird speed} \quad (2)$$

Being an assumed mean value, a rather high standard deviation (σ_d^2) of $0.97 \frac{m}{s}$ found in the literature is applied to model the distances flown (formula (3)) (Nilsson et al., 2014).

$$\sigma_d^2 = \text{activity} \cdot \text{bird speed standard deviation} \quad (3)$$

The main goal of the activity model is to propose different distances the bird can fly in a given time span and at a given speed, as well as to estimate the probability the bird flew the proposed distances given the before mentioned parameters. To compute the probabilities of a phenomenon a mathematical function is selected, which best describes the probability density of the phenomenon. The distances flown given activity and mean velocity are best described with the gamma function because it does not allow any negative distances to occur. It weights distances around the mean distance higher than greater distances and has a nearly exponential decrease towards the higher end (Lukacs, 1955; Binz et al., 1981; Storrer, 2009). The gamma distribution has two parameters, the shape parameter (α) and the rate (β). These two parameters are derived after applying some algebraic transformation on the mean and variance formulae (formula (4) and (5)) (Binz et al., 1981; Storrer, 2009).

$$\mu_d = \frac{\alpha}{\beta} \quad (4)$$

$$\sigma_d^2 = \frac{\alpha}{\beta^2} \quad (5)$$

Where μ_d is the mean distance the bird was flying in a given time and σ_d^2 is the variance of the flying distance. Given the mean and the standard deviation of the bird activity, the rate parameter β is computed applying formula (6) (Lukacs, 1955; Thom, 1958; Stacy and Mihram, 1965; Binz et al., 1981; Storrer, 2009).

$$\beta_{dist} = \frac{\mu_d}{\sigma_d^2} \quad (6)$$

Where μ_d is the mean distance flown by the bird and σ_d is the standard deviation of the distance flown. The shape parameter is computed after the rate parameter, applying formula(7) (Lukacs, 1955; Thom, 1958; Stacy and Mihram, 1965; Binz et al., 1981; Storrer, 2009).

$$\alpha_{dist} = \mu_d \cdot \beta_{dist} \quad (7)$$

Where μ_d is the mean flying distance and β_{dist} is the rate parameter of the gamma distribution computed in formula (6). The resulting density function of flying distances are plotted 13.

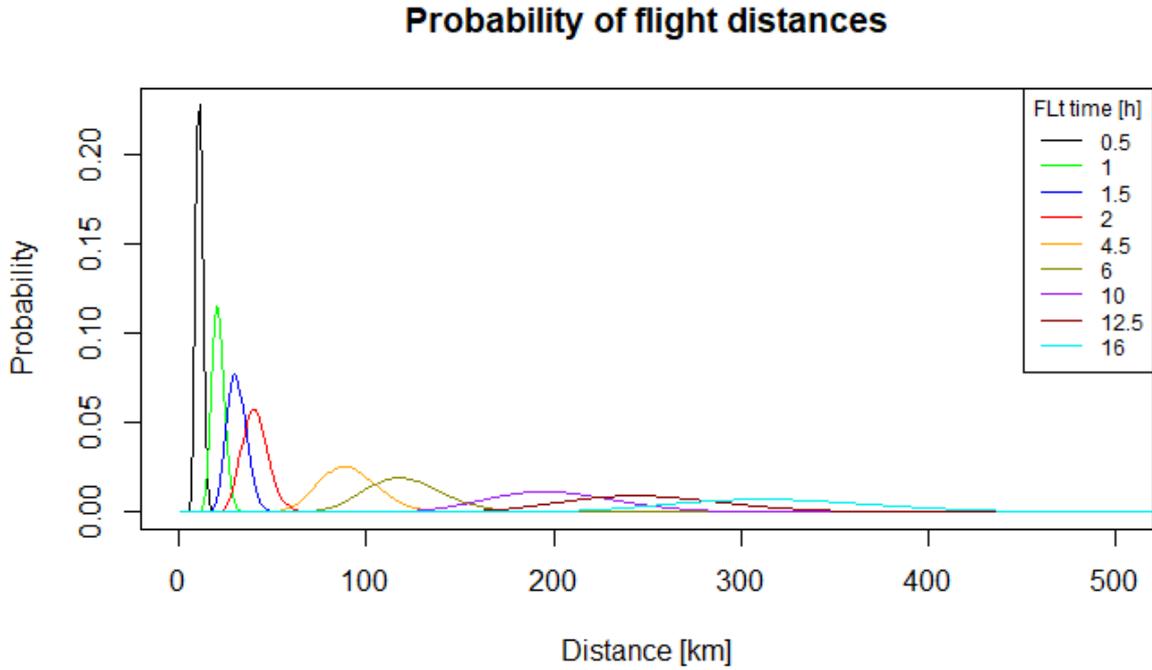


Figure 13: Flying distances at different time spans of high bird activity and their probabilities of occurrence

In figure 13, the model of different flying times and thereof resulting distances are illustrated in one plot. Figure 13 illustrates a almost exponentially decreasing probability with increasing flight distance. The mean distance flown when the bird was active for 0.5h (cyan curve in figure 13) is assigned a probability of 22.7% (0.227), whereas the mean distance covered when the bird was flying 16h (green curve in figure 13) is assigned a probability of 0.7% (0.007). This steep decline of probability is caused by the increasing uncertainties in the data, which are modelled with a large standard deviation.

3.3 Wind data and wind model

Wind data was downloaded from **ECMWF** web site, which has a size of 19.9GB and was recorded by the European Union's Copernicus programme (Berrisford et al., 2011; European Centre for Medium-Range Weather Forecast, 2020). To facilitate data retrieval, a PostgreSQL data base was set up where the wind data was stored. In figure 14 a schematic overview is shown, how the wind data management was executed and how the PostgreSQL database was set up.

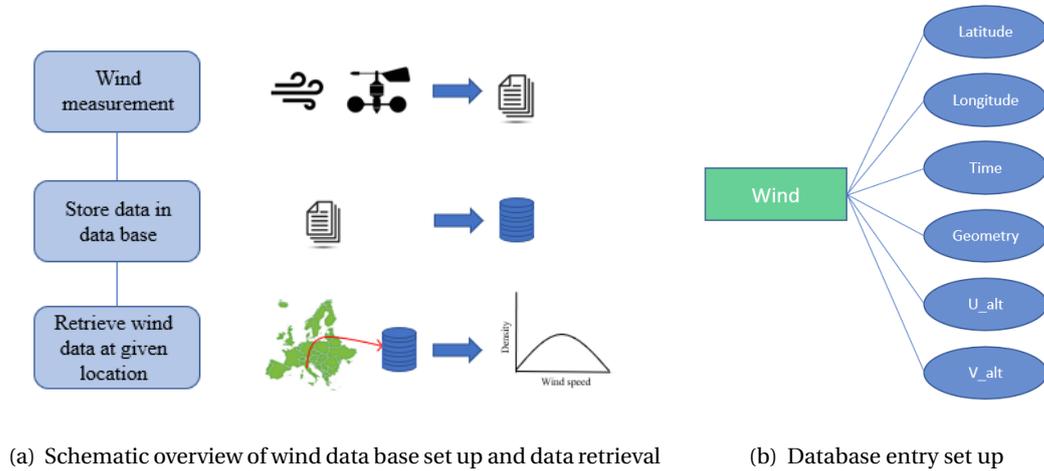


Figure 14: Schematic overview of the wind database set up and use

3.3.1 Wind data base

The data base holds 5.3 million entries (rows) which contain 30 fields (columns) in which the attributes of the observed wind conditions are stored. Each wind entry has attributes of the position off measurement, the wind speed in U and V direction at each available pressure level respectively and the time stamp of the wind data aggregation. The latitude and longitude where transformed to a geometric point object (figure 14(b)). The geometric point objects where defined to facilitate and speed up wind data queries on the data base (Zhang et al., 2009; Zhang and Yi, 2010; Fröhlich, 2018; Makris et al., 2019).

3.3.2 Wind data query

The attributes which are used to query wind data from the data base are indexed. Spatial objects, like the Geometry are indexed with a spatial index. With the spatial index, it is not necessary to execute a query as sequential scan, where the entire data base is scanned for a geometry object. Applying a spatial index therefore reduces computational costs. Scanning for example a data base with 10 000 entries without an index would

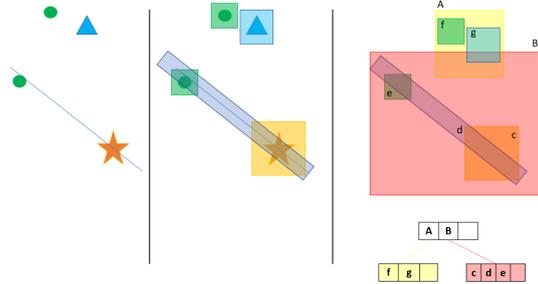


Figure 15: Sketch to illustrate the basic idea of a spatial index applied on the available wind data

cause 100 000 000 comparisons, whereas scanning the same data base with spatial index would cost approximately 20 000 comparisons (Fang et al., 2008; Simion et al., 2013; Makris et al., 2019). The spatial index speeds up spatial queries because the index first creates bounding boxes around the objects to be indexed and secondly around a cluster of objects. The geometries of these bounding boxes are indexed in a hierarchical tree structure, where the bounding box of a cluster is stored on a higher level than the bounding box of a geographic object (see figure 15). Spatial indexing therefore

structures the geographic space and facilitates the search for coordinates in a data base. Especially the 5.3 million geometry entries (geographic point positions of the wind) need to be queried as quickly as possible.

3.3.3 Modelling wind

Because the wind is a continuous phenomenon, the nine closest wind measurements to the bird position are queried from the data base. The 9 returned geographic positions closest to the bird position at a given date and time were used to compute the mean wind velocity the bird encountered during migration (figure 16(a)). The mean wind velocity was computed by executing mean weighted by distance on the wind velocities retrieved from the wind database (figure 16(b)). With this method, additionally, positional uncertainties were taken into account and smoothed. Formula 8 is executed to compute the weighted mean of the wind velocity and direction at the position of the bird at a given time. The mean weighted wind velocity is computed by summing all winds multiplied by their respective weight and dividing this sum by the sum of weights (Finch, 2009; Saghir et al., 2017).

$$\mu_w = \frac{\sum_{i=1}^n (x_i \cdot w_i)}{\sum_{i=1}^n w_i} \quad (8)$$

Where μ_w is the weighted mean wind velocity, x_i is the wind velocity and w_i is the weight applied at the i^{th} position. The weights were computed to give wind measurements closer to the bird position more weight than those further away (see figure 16(b)). Therefore, the distance between the bird position and each wind measurement position was computed. The latitude and longitude values were plugged in to the Haver-sine formula to compute the distances necessary to compute the weights (Basyir et al., 2018; Hijmans, 2018). For better computability, the weights were normalized to scale between 0 and 1 (see formula (9)). Hence, the weight is computed, dividing the difference of maximum distance and the actual distance by the difference

of minimum distance and maximum distance.

$$w_i = \frac{d_i - d_{max}}{d_{min} - d_{max}} \quad (9)$$

Where w is the weight applied to compute the mean wind speed, d is the distance between the bird position and the position of the wind measurement, d_{max} is the greatest distance and d_{min} is the smallest distance between the bird position and the wind measurement coordinate.

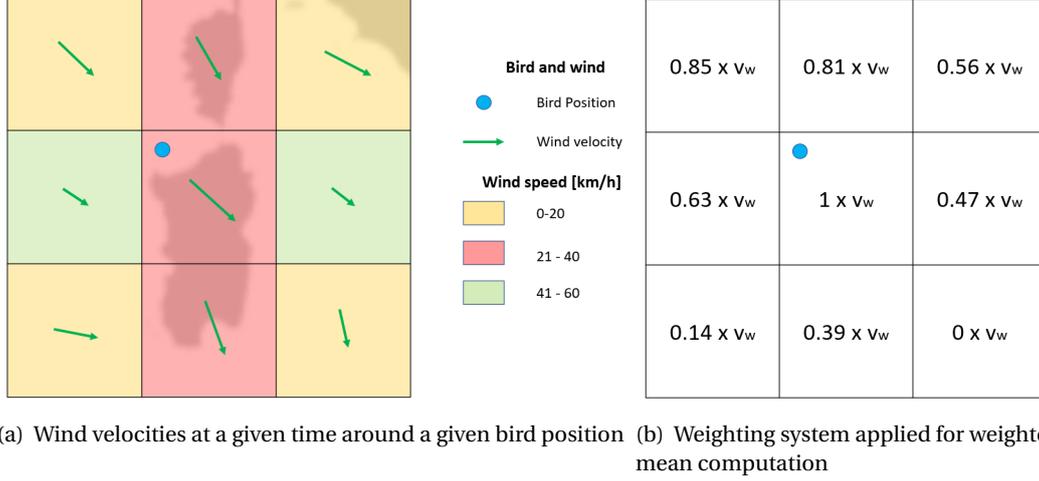


Figure 16: Wind situation recorded around the estimated bird location, returned from database query

After computing the weighted mean of the queried wind data, the wind vector needs to be partitioned into two new vectors of which one is perpendicular and the other is parallel to the bird flight vector. In figure 17 the vector partitioning is illustrated, where figure 17(b) shows a possible situation in which the bird can be found en route. Figure 2 however shows the vector partitioning, where v_b is the bird flight vector (air speed vector), v_w is the wind vector and v_g is the ground speed vector. The red vector v_c perpendicular to v_b is the compensation vector. This vector indicates how strong the bird has to compensate the wind conditions to reach its goal. The purple vector v_{wa} parallel to v_b is the wind support vector. α is the angle between the air speed vector and the wind vector and β is the angle between the compensation vector v_c and the wind vector. To compute the wind support vector formula 10 was applied. Following the law of trigonometry, the cosine of the angle α multiplied with the hypotenuse results in the adjacent leg of the analyzed triangle.

$$v_{wa} = \cos(\alpha) \cdot v_w \quad (10)$$

In this study, the partition of v_w parallel to v_b is computed by multiplying v_w with the cosine of angle α . With formula 10 the magnitude of v_{wa} is computed. The direction is the same like the air speed vector v_b . The magnitude of the compensation vector v_c is computed by multiplying v_w with the sine of β , whereas the compensation angle β is computed with the sum of the inside angles of the triangle analyzed.

After computing the central metric of the wind add vector (speed and direction) it was possible to compute the ground speed v_g of a migrating bird. Therefore, the wind add vector \vec{v}_{wa} and the bird migration vector \vec{v}_b were added. Because these vectors point to the same direction, the wind speed and the bird migration speed were summed. Thereby, the mean distance the bird probably did fly (μ_d) is computed differently

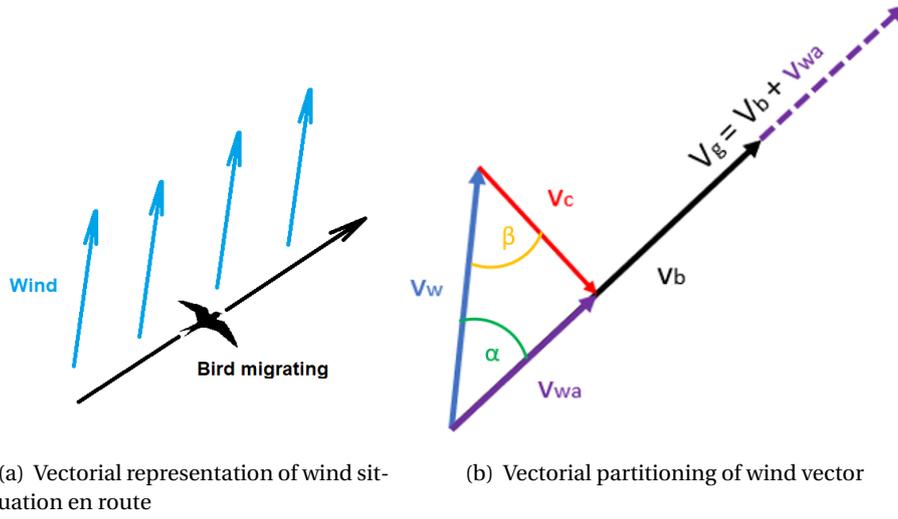


Figure 17: Vectorial partitioning of a wind vector to compute wind add parallel to the bird migration vector

(formula (11)).

$$\mu_d = \text{activity} \cdot (\text{mean bird speed} + \text{wind speed}) \quad (11)$$

Finally, the probabilities of all distances a bird most likely flew in the current wind situation are computed. Applying this new method, the density distribution of the flight speeds is a mathematical composition of two different density distributions. On one hand there is the distribution of the bird movement (activity data and physiological capabilities), which assimilates a gamma distribution. On the other hand there are the wind speeds which are assumed to follow a Weibull distribution, because it is not known how well Hoopoes can take advantage of strong winds. For wind data modelling, the symmetric Weibull distribution was used (Diniz et al., 2004; Wang et al., 2018). Because the Weibull distribution has a scale and a shape parameter which are difficult to compute numerically, the Weibull distribution is approximated with a special normal distribution (Kulkarni and Powar, 2011). This special case of the normal distribution starts at a wind speed of $0 \frac{m}{s}$ and ends at $2 \cdot \mu_w$. Given this assumption, the standard deviation of the wind speeds is computed executing formula (12).

$$sd_w = \mu_w - \frac{\mu_w \cdot 2.5}{2} \quad (12)$$

Where sd_w is the standard deviation of the wind measurement and μ_w is the mean wind speed (wind speed returned from the data base).

Finally, the standard deviation of the distance the bird probably flew is computed the same way like the mean distance. The standard deviation (σ_d^2) is computed by multiplying the bird activity with the sum of the standard deviation of the bird speed and the standard deviation of the wind speed (formula (13)).

$$\sigma_d^2 = \text{activity} \cdot (\text{standard deviation bird speed} + \text{standard deviation wind speed}) \quad (13)$$

Because the bird movement was considered to be more important than the wind conditions in which the bird was flying, the density distribution contributing to the probability of the positions en route (posterior distribution) was thought to follow a gamma distribution to not allow reverse movement. The α and β parameters are therefore computed executing formula (6). In figure 18, the different probability distributions

of different flying times are illustrated in one plot. Comparing this plot with the figure 13 different changes are noticed. First, the probabilities of flight distances are in general lower than those without wind support. Secondly the distances the bird has probably flown increase. Third, the probabilities of the probably flown distances decrease with increasing flight duration (7).

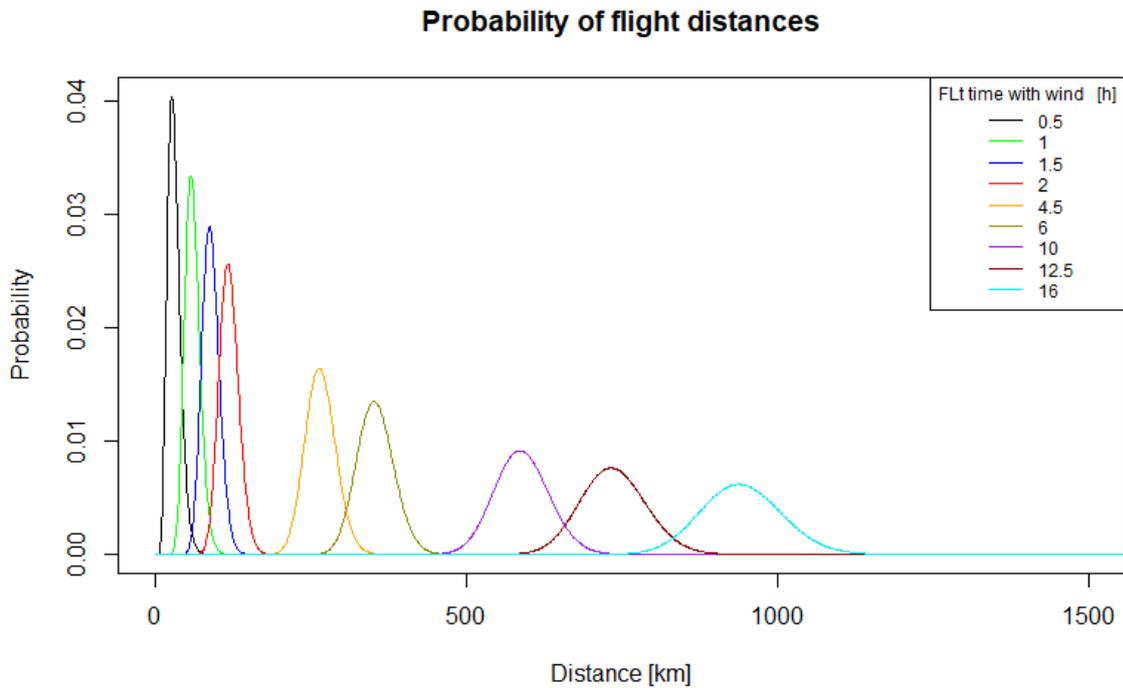


Figure 18: Flying distances at different time spans of bird activity, including wind conditions and their probabilities of occurrence

3.4 Model comparison

After route inference, the resulting migration routes and their probabilities need to be compared in order to answer the questions analyzed in this study. First, the mean latitude and longitude values of the different models are compared and their differences tested for significance. Therefore, a two sided t - test is executed to assess if the latitude and longitude values along the mean MCMC of two compared models differ significantly (\neq). This data is then compared to a threshold value, the p - value. The p - value is a statistical value indicating the threshold at which the H0 hypothesis of non significant difference of means is rejected and H1 hypothesis of significant difference of means is accepted. In this study, applying a two sided t - test a threshold value (p - value) of 0.05 is applied for significance testing. To correctly compare the computed t - statistics of the data analyzed with the p - value, the appropriate threshold value corresponding to the p - value needs to be read out from a t - test table applying the correct degrees of freedom. In this study a threshold value (CV (critical value)) of 1.96 was applied (Storrer, 2009). Applying the t - test, the hypothesis (H0) of equal means (non significant difference of means) is assumed. The alternative hypothesis (H1) assumes non equal means (significant difference of means). Hence, a t - value greater than the CV indicates a significant difference, whereas a smaller t - value indicates no significant difference of the mean values compared.

Secondly, the model results are compared in terms of how good they describe the measured data (light intensity, activity and wind data). Therefore, the marginal likelihood or marginal data density of each model result is computed and compared. The marginal likelihood is a Bayesian model selection criterion, which is widely used to compare and pair wise select Bayesian models through comparison of their posterior distributions (Chib and Jeliazkov, 2001; Raftery et al., 2007; Chan and Grant, 2015; Pajor, 2017). To compute the marginal likelihood, the sampling density is integrated with respect to the prior distribution of the parameters. Thereby the model likelihood is averaged and thus reduced to one number and the models can be better compared (Berger and Pericchi, 1996; Chib and Jeliazkov, 2001; Tipping and Faul, 2003; Robert and Wraith, 2009; Pajor, 2017). Because the computation of an integral causes some difficulties and computational costs, alternative way were found to approximate the marginal likelihood of a Bayesian model. To approximate the marginal likelihood the harmonic mean of all probabilities is computed, what has become one of the most applied methods to approximate the marginal likelihood, despite its simplicity (Gelfand and Dey, 1994; Raftery et al., 2007; Robert and Wraith, 2009; Pajor, 2017). The harmonic mean of the two models to be compared are then compared, computing the Bayes factor (formula (14)), where one model harmonic mean is divided by the second model harmonic mean (Gelfand and Dey, 1994; Berger and Pericchi, 1996; De Santis and Spezzaferri, 1997; Chib and Jeliazkov, 2001; Raftery et al., 2007; Chan and Grant, 2015; Pajor, 2017).

$$BF = \frac{\text{Harmonic mean 1}}{\text{Harmonic mean 2}} \quad (14)$$

The resulting factorial number indicates approximatively, which model describes the observed data better and can thus be assumed to describe the observed phenomenon better.

3.5 Wind and activity models

To answer the research questions the activity and wind data were modelled in different ways, to analyze the influence the wind conditions have on the migration route inference. In the following sections, the different models are introduced. The four different models are illustrated in figure 19, where the different gamma distributions are shown. The plot in figure 19 shows how the probabilities of different distances the bird likely flew change with changing wind influence.

3.5.1 Wind model

For this experiment (Wind model (WM)), the wind speed data was included to the route inference without any modification. The wind speeds returned from the database were partitioned into the two vectors v_{wa} and v_c , where the wind speed uncertainty was defined for all wind speed values to be a constant value of $3.5 \frac{m}{s}$. Hence, the wind speeds are assumed to be normally distributed. This experiment was carried out to demonstrate the effects a normal distribution of the wind speed values have on the migration route.

3.5.2 Weak wind

In this experiment (Weak wind inference (WW)) the wind speed data was included to the movement model without any modification. The wind speeds returned from the database were partitioned into the two vectors v_{wa} and v_c . The uncertainty of v_{wa} was computed applying the Weibull approximated distribution. Thereafter, the v_{wa} vector was used to compute the ground speed of the migrating bird. For this wind model, it was assumed, that the bird uses high wind speeds less efficient than low wind speeds. The gamma density

Density distributions of the distances flown

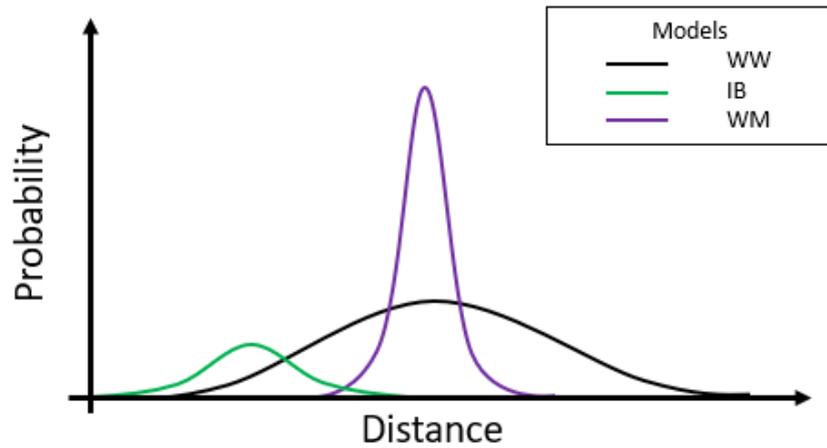


Figure 19: Density functions of the different wind and activity models applied in route inference

distribution in figure 19 shows how the probabilities are assigned to the different distances the bird probably was flying.

3.5.3 Inefficient bird

For this experiment (inefficient bird (IB)) the wind speed was halved after the partitioning of the wind vector v_w . The halved wind support v_{wa} was included to the movement model, to simulate the bird migrating with half wind use efficiency. This experiment was carried out to demonstrate the effects of a bird which is capable of using only half of the wind speeds. With half wind use efficiency, the gamma distribution in figure 19 is shifted to the left, assigning shorter distances than the weak wind inference high probabilities. In addition to that, the probabilities of the distances are lower than in WW.

4 Results

In this section the results of the study, as well as the results of the experiments carried out to construct the models are described. The experiments described, are necessary to understand the underlying algorithms and methods of the route inference and to prove that the wind conditions en route have an influence on bird migration and behaviour. Furthermore, these experiments had an impact on the technical construction of the wind model applied in the inference.

4.1 Wind model inference

Including a normal distributed wind speed probability and thus increasing the probability, the bird did follow the general wind conditions, decreased the area of probability where the bird was expected to fly. In addition to that, the areas of higher probability are smaller, too. The yellowish areas extend in such a way that they include the eastern Spanish coast line and the Balearic islands. Including the normal distributed

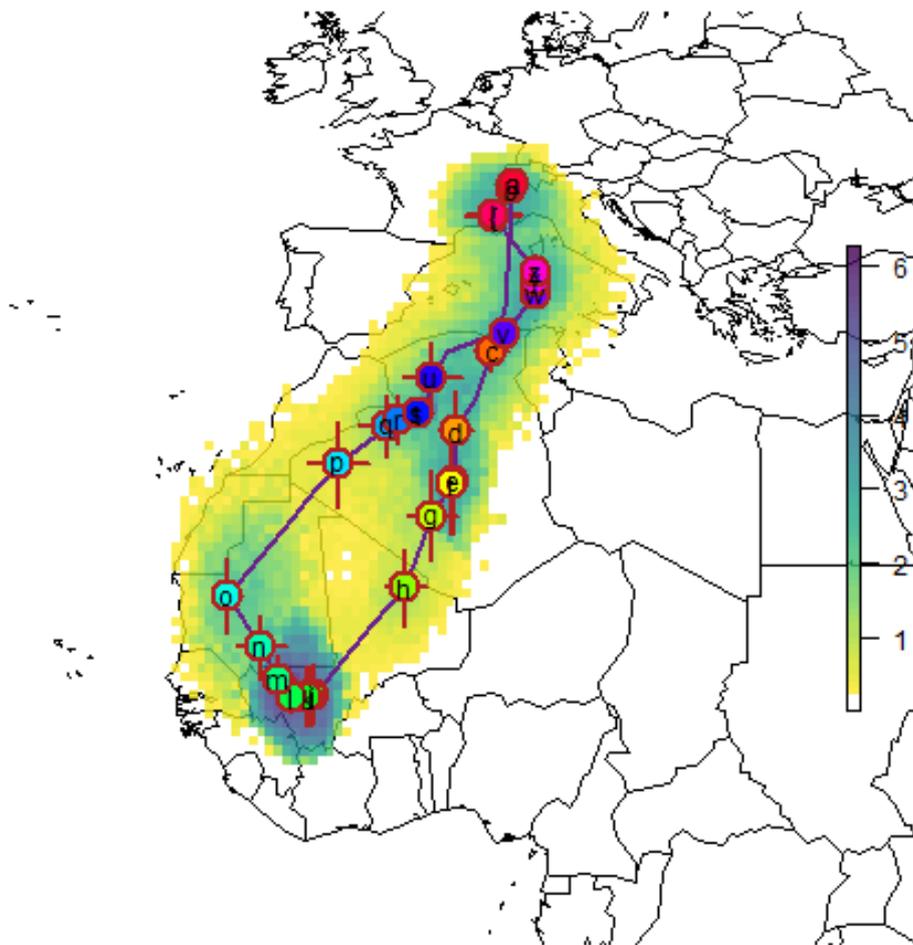


Figure 20: Result of route inference applying MCMC methods, which include the newly developed activity and wind model, where the wind model includes normal distributed wind speed data.

wind data includes the Italian peninsula, especially the Italian West coast to the area of probability. Applying the constant standard deviation of the wind speed, narrows the areas of higher probability (2 – 6%).

Especially at the non-breeding location, the which are assigned the highest probability, are smaller and less diffuse delimited. This inference includes the islands of Corsica and Sardinia as stop over sites. Additionally, this region is delimited as one separate stop over region which is assigned a probability of 3 – 5% on the migratory bouts from the non-breeding locations back to the breeding grounds. The positions at the non-breeding locations in Africa are found to be most South, compared to the other inferences. However, the region where the non-breeding sites are found is inferred with a greater confidence, what is shown with a greater area assigned the highest probability of 6%.

In table 6 all position information, as well as the arrival and departure dates of this route inference are shown. The standard deviation of the non-breeding locations increased significantly compared to NO. Especially the standard deviations of the latitude, improved in comparison to the other inferences, where the standard deviation ranges from $0.98^{\circ} - 2.39^{\circ}$. The standard deviation of the longitude values however, range from $0.31^{\circ} - 1.06^{\circ}$. These values are significantly smaller than those of SW but non significant different from WW. Over all, the standard deviation of this route inference ranges from $0.28^{\circ} - 3.33^{\circ}$ for latitude values and from $0.31^{\circ} - 3.29^{\circ}$ for longitude values. The smaller standard deviations indicate a better position precision (postion o and p in figure 20). The positions with the smallest standard deviations where inferred to be at the non-breeding location. The confidence intervals indicate a similar picture. The confidence intervals range over 6.64° and 6.48° for latitude and longitude respectively.

However, the two models (no wind NO and Wind model WM) need to be compared and finally the selection criterion determining the model best describing the observed data is computed. Therefore, the Bayes factor is computed, by dividing the marginal likelihood of WM by the marginal likelihood of NO. As described in section 3.4 the marginal likelihood is approximated by computing the harmonic mean of all probabilities of each position along the migration route. The Bayes factor of WM versus NO ($BF = \frac{WM}{NO}$) is 4.96. Additionally, the Bayes factor comparing WM with IB is computed to be 4.13 and the BF comparing WM with WW is computed to be 2.46. Table 2 shows all Bayes factors comparing the harmonic means of all models implemented.

Table 2: Look up table for all Bayes factors computed to compare and evaluate the different models applied in this study. This table is read row wise, where the result represents the division of the row value by the column value.

	ML NO	ML IB	ML WW	ML WM
ML NO	1.00	0.83	0.50	0.20
ML IB	1.20	1.00	0.60	0.24
ML WW	2.02	1.68	1.00	0.41
ML WM	4.96	4.13	2.46	1.00

4.2 Weak wind inference

Including the wind data to the migration route inference, results at first sight in increased area, the Hoopoe was possibly flying through on its migration. Besides the increased yellowish area (0.1 – 1% probability), the greenish area (2 – 4% probability) is increased in size, too. The locations of residence of Hoopoes however, do not change drastically at first sight, they are shifted slightly towards West - South Westerly direction. Therefore, the locations of residence are found to be with greatest probability (6%) in Mauritania and Senegal. Applying environmental conditions and a more precise activity model to infer the migration routes, adds different new countries to the flight range of Hoopoes. The area where Hoopoes are migrating to, in-

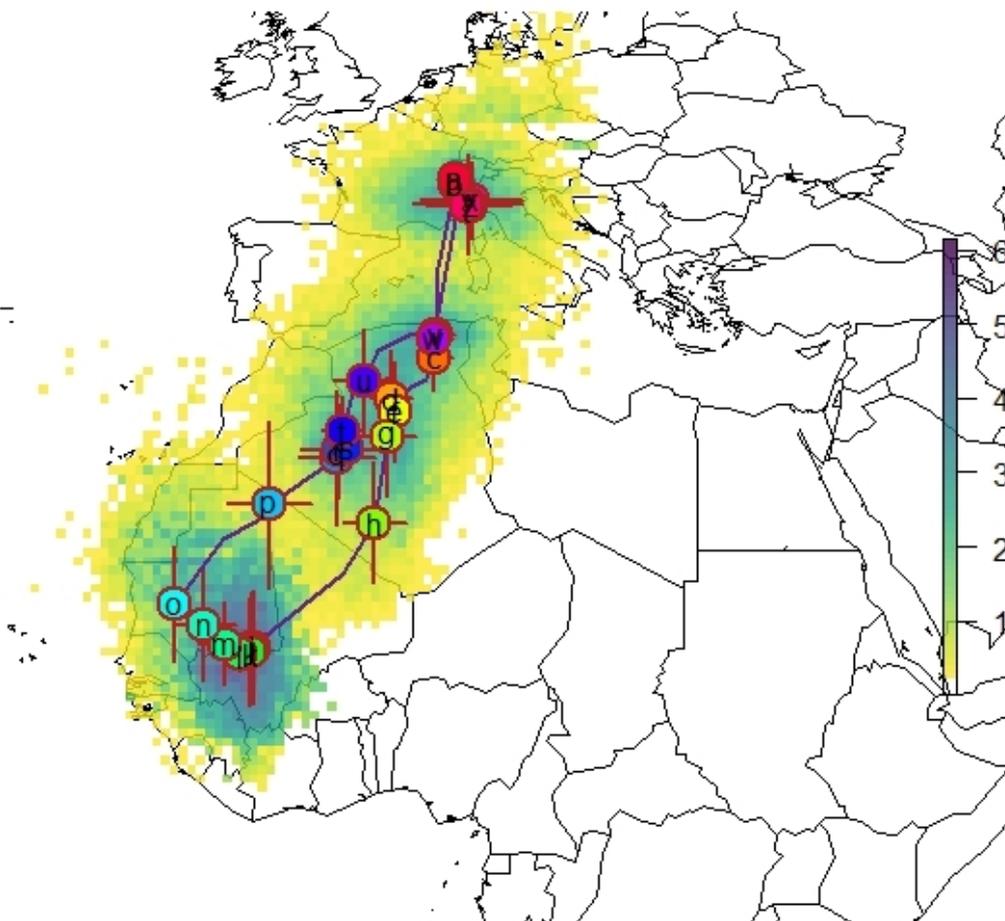


Figure 21: Bird migration route inference including Weibull distributed wind conditions

creases in size and includes countries like Gambia, Guinea, Guinea - Bissau and parts of Liberia, the Ivory Coast and Burkina Faso (see figure 21). The areas where the bird did probably cross the Mediterranean are wider spread across the ocean and now reach from Italy to Spain, including the strait of Gibraltar. The mean migration route however, remains in the region of Corsica and Sardinia. Additionally, the probability is more evenly spread across the Mediterranean than the previous inference predicted. The regions around the Balearic Islands, Sardinia and Corsica are assigned slightly higher probabilities ($\geq 2\%$). The migration route inference proposes migration routes which include the Canary Islands. The fire-brick line represents

the mean migration route, hence the route which is assigned the highest probability. Finally, the error bars indicating the spread of the coordinate latitude and longitude values are increased in size. This increase indicates a greater spread of those values and therefore a greater uncertainty, introduced by the new data and the newly computed bird flight vector.

Table 5 contains arrival and departure dates at stopover sites and the corresponding latitude and longitude values. The arrival and departure dates did not change, but the standard deviation of the locations of residence has increased significantly. Applying full wind data and bird activity data, the standard deviation of the locations of residences' latitude values range between 0.95° (12.01) and 2.51° (23.03), whereas the standard deviation of the locations of residences' longitude values ranges between 0.32° (12.01) and 1.79° (23.03). These values quantify the standard deviation of the positions i, j, and k in figure 21. However, the positions estimated en route are less accurate than those at the locations of residence. These positions' standard deviations are estimated to range from 0.89° to 4.89° for latitude values and from 0.97° to 3.89° for longitude values. The mean confidence interval of the positions a migrating bird possibly visited is estimated to include 9.56° for latitude and 6.62° for longitude values. The 95% confidence interval illustrates the range in which the position latitude and longitude values are spreading. These results suggest, the confidence interval as well as the standard deviation of the positions along the migration route estimated applying model in WW are greater than those estimated applying model in IB. The standard deviations of the latitude and longitude values are summarized to a mean value, which includes and summarizes all positions and sd values in one numerical value. The standard deviations of the latitude values however, are slightly greater than the mean longitude values. The confidence intervals represent a similar situation. The mean value of all confidence intervals of the latitude values is greater than the mean confidence intervals of the longitude values.

4.3 Inefficient bird inference

At first sight, the migration route of IB differs from NO, especially in the region of the locations of residence. The wintering site is greater compared to the wintering site estimated in the previous inference. Furthermore, the probability that the bird did reach these positions looks more concentrated than in NO. Additionally, the stop over sites across the Mediterranean did change location. The bird did not cross the ocean multiple times but stopped on Sardinia. In figure 22 the three stopover regions are better differentiated. One is found close to the breeding site, the second in Northern Africa and the third at the wintering site.

However, the area where the bird did fly through is increased, especially the greenish areas (2 – 3%) in Algeria and across the Mediterranean. The yellowish areas (0.1 – 1%) are increased and now include the South Eastern coast of Spain and the South Western coast of Italy. The mean track which is represented as solid line, is shifted to westerly direction. Additionally, the error bars around the stop over positions are greater, indicating a greater spread of possible latitude and longitude values for the respective position. Especially the errorbars of the latitude values are increased in size.

Table 4 contains arrival and departure dates of all stop over sites on the migration route as well as the corresponding latitude and longitude values. Given the additional data, the locations of residence are found to be with greatest possibility in the South Western region of Mali and the South Easter region of Mauritania. The arrival and departure times are the same like in the first inference. At the non-breeding locations in the Sahel Zone, the standard deviation of position j is estimated to be 1° for latitude and 0.31° for longitude

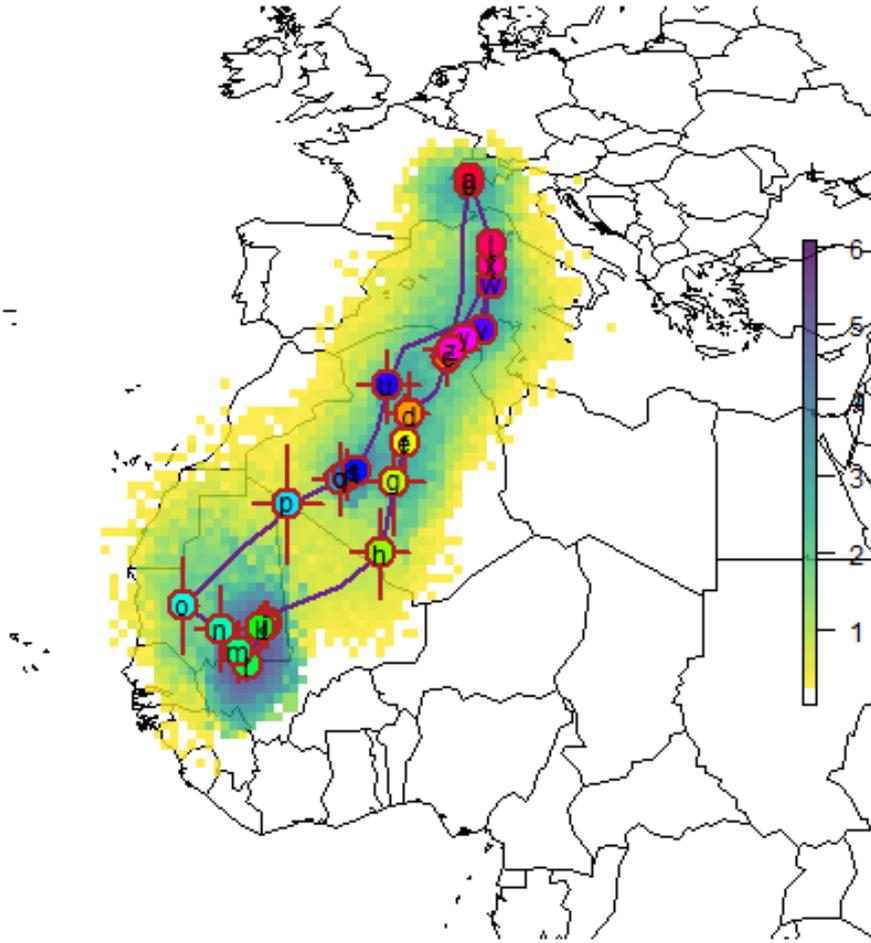


Figure 22: Result of route inference applying MCMC methods, which include the newly developed activity and wind model, where the wind model includes half wind speeds only.

values respectively. The last position in the region characterized by the greatest probability at the locations of residence (6%) is position k, where the standard deviation is a bit greater. At position k the standard deviation of the positions latitude values is 2.72° , whereas the standard deviation is 1.9° for longitude values. Stopover sites visited for a short period of time are characterized by a greater standard deviation than stopover sites visited for a long period of time. Stop over sites along the migration route are thus characterized by standard deviation from 1° to 4.12° for latitude values and from 1.3° to 3.2° for longitude values. These great standard deviation values indicate a small precision of position estimation.

Computing the 95% confidence interval of each position (see table 4) show a similar picture of the position estimation accuracy. The mean confidence interval in the mean MCMC for latitude values is 6.95° , whereas the mean confidence interval for longitude values is 5.49° . The smallest confidence interval however is 1.45° and 1.13° for latitude and longitude respectively. The greatest confidence interval is 15.31° and 11.81° for latitude and longitude values respectively. These confidence intervals are significantly greater than those of NO.

The decision criterion comparing IB with NO is 1.20 (table 2) and indicates hence, that IB is explaining the observed phenomenon 1.20 times better than NO.

4.4 Wind conditions found on the mean track

In this section, the wind conditions found on the mean track are analyzed. First the inference where the bird is assumed to only use half of the available wind speed (IB) is analyzed. Secondly, the Bayesian inference including the wind data with the full wind uncertainty (WW) is analyzed. Finally, the inference including a constant uncertainty is analyzed (WM). To analyze the wind conditions as well as the wind support a bird encountered en route, two different figures are illustrated for each Bayesian inference. First, a bar plot is generated, indicating the wind speeds found on each route segment. Secondly, a bar plot is generated to illustrate the wind support a bird encountered on each route segment.

In IB the wind conditions are almost the same like in WW. They differ most importantly because of the implemented wind use efficiency of the birds. The birds are thought to be able to use only half of the available wind speed. In addition to that, the wind speeds differ because of the different spatial positions the route positions are estimated to be. In figure 23(a) the wind speeds found on each route segment are illustrated in a bar plot. Hence, the mean wind speed found on the mean track of IB is 6.06 m/s, with a

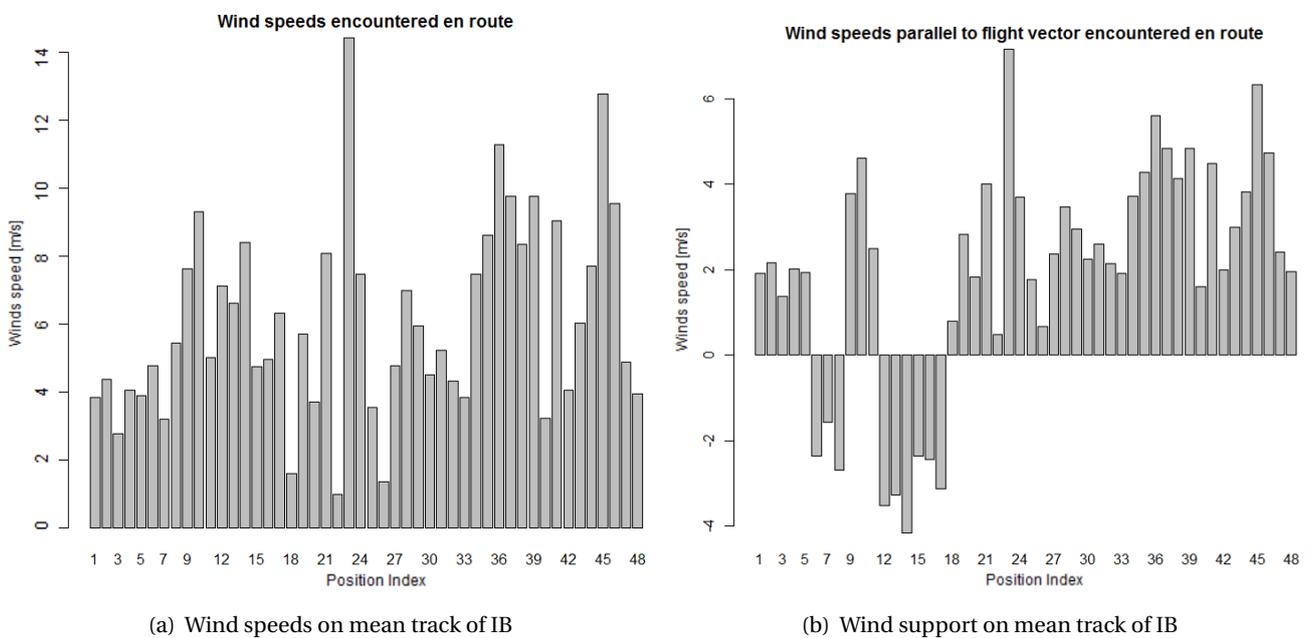


Figure 23: Wind speeds and wind support found on the mean track of IB

standard deviation of 2.86 m/s. The lowest wind speed is 0.98 m/s where the highest wind speed found is 14.44 m/s. The wind support is as well plotted in a bar plot to indicate how much wind support the bird encountered on each segment. In figure 23(b) the wind support is indicated, where the positive wind support values are tail wind speeds which increased the ground speed, whereas the negative values are reducing the ground speed. The mean wind support is 1.94 m/s with a standard deviation of 2.72 m/s. The maximum wind support is 7.15 m/s, whereas the minimum wind support is -4.17m/s. The wind conditions applied in the weak wind inference are illustrated in figure 24(a). The mean wind speed found on the mean track is 5.1 m/s, where the standard deviation is 2.47 m/s. The lowest wind speed found en route is thus 0.57 m/s and the highest wind speed is 15.4 m/s. The wind support applied in this inference is illustrated in figure 24(b). The mean wind support is 2.49 m/s with a standard deviation of 2.2 m/s. The strongest wind support is 5.99 m/s, whereas the lowest wind support is -3.57 m/s.

The wind conditions applied in the constant uncertainty inference are illustrated in figure 25(a). The

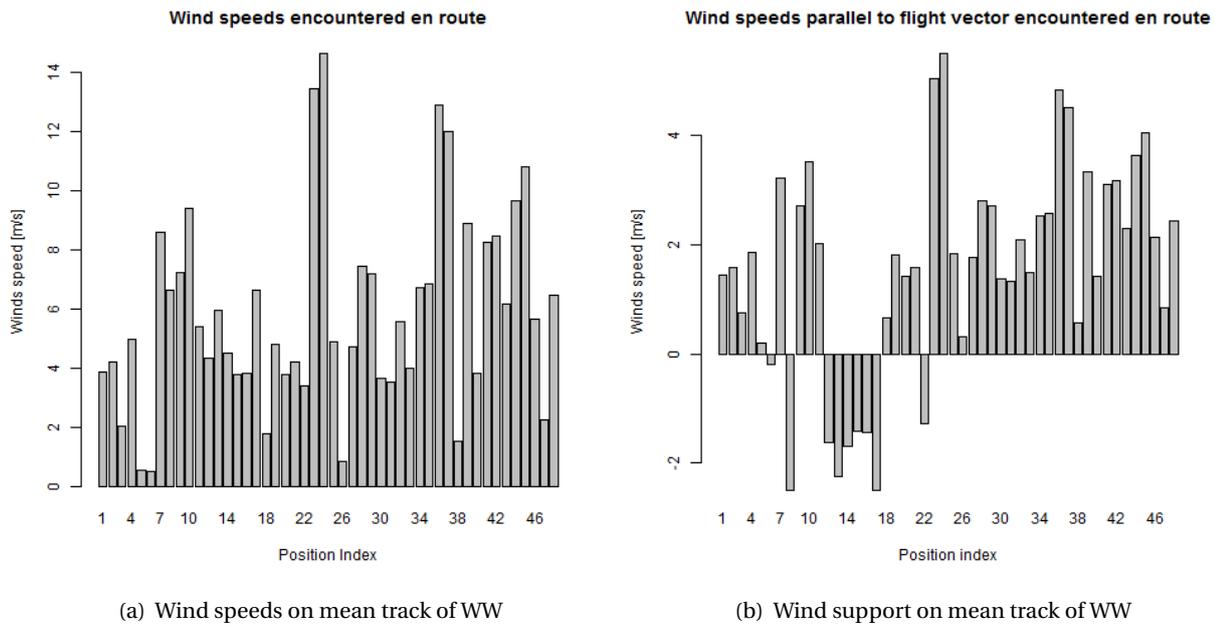


Figure 24: Wind speeds and wind support found on the mean track of WW

mean wind speed found on the mean track is 10.82 m/s, where the standard deviation is 5.91 m/s. The lowest wind speed found en route is 1.99 m/s and the highest wind speed is 26.24 m/s. The wind support applied in this inference is illustrated in figure 25(b), where the mean wind support speed is 0.64 m/s and the standard deviation is 0.85 m/s. The strongest wind support is 2.24 m/s, whereas the lowest wind support is -1.37 m/s.

4.5 Non-breeding locations of residence

Amongst the most interesting positions on the migration route are the positions at the non-breeding locations of residence, where the birds migrate to when it is winter at the breeding grounds. To demonstrate the effects the different parameters have on the migration route inference, the most central statistical values of

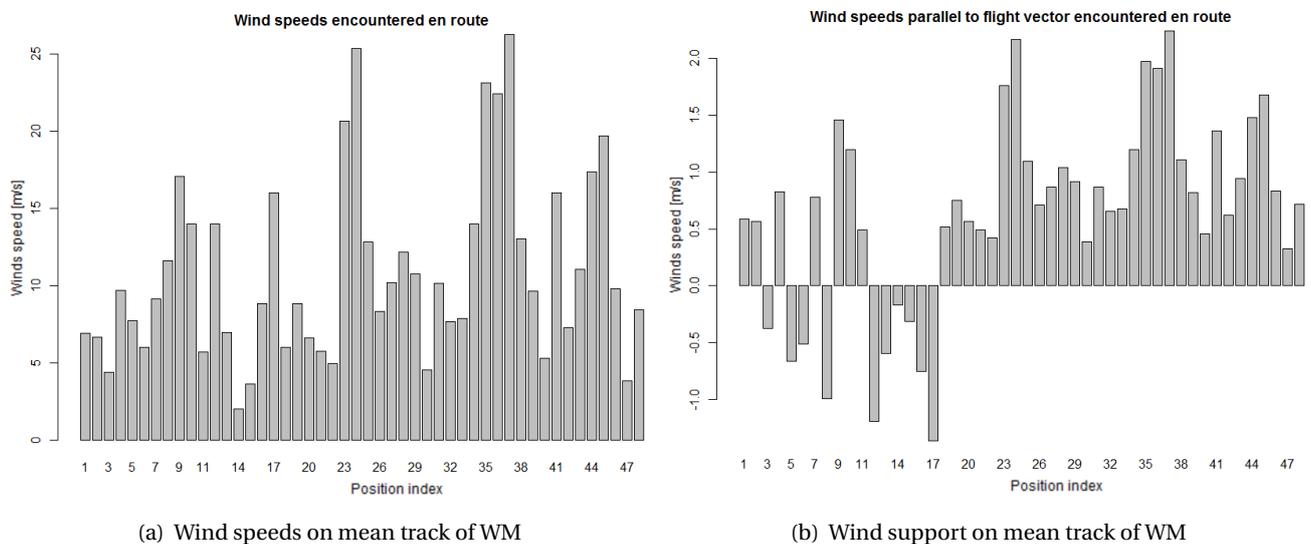


Figure 25: Wind speeds and wind support found on the mean track of WM

the positions' latitude and longitude are presented. Hence, the median, upper 95% quantile and lower 95% quantile are illustrated in different plots. Therefore, box plots are created of each position found at the non-breeding locations. Thereby, at each position, the box plots are generated for the positions' latitude and longitude values respectively. For better comparison, the box plots of each Bayesian inference are plotted in one figure. The positions found to be at the non-breeding location are the positions i , j , k , l and m . These are as well the positions characterized by the greatest difference, especially in latitude values. In the following section the positions found in the non-breeding locations illustrated in figures 11, 21 and 20 are analyzed. At the first position of the non-breeding location (figure 26), the latitude values of the Bayesian inferences are located around the same latitude, with the exception of WW and WM. WW is the most northerly position, whereas WM is the most southerly. The positions are characterized by similar uncertainties too, where NO and WM are characterized by the strongest uncertainties. Bigger uncertainties are expressed by greater whiskers and quantile ranges. In NO, the greatest uncertainty of position i is found. Most precise latitude value for position i is found in WW, where the full wind data was introduced to the inference. The longitude value however indicates different results. The models in which a influence of wind was introduced, the positions' latitude values are found within the same range. In figure 26(b) the westernmost longitude value of position i is found in NO. IB and WM are found to be more westerly in respect to WW and NO, too. The uncertainties are greatest in WW, where the full wind data is introduced to the inference. However, the WM inference shows the smallest standard deviation.

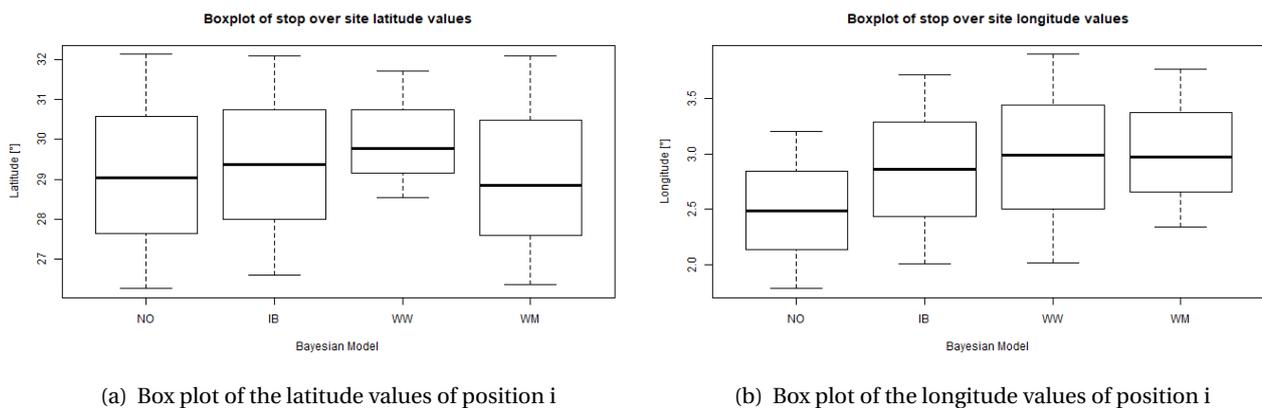


Figure 26: Box plot of the latitude and longitude values of position i found at the non-breeding locations

Figure 26 illustrates the situation at the first position found at the non-breeding location, whereas the next positions at the non-breeding locations are analyzed in the next paragraphs. These positions show different trends, especially the longitude values. In figure 27 the box plots of position j latitude and longitude value are illustrated. The most southerly position is found in WM, where the normal distributed wind speed data is introduced to the inference. However, the other four inferences did not suggest position j to be that much to the South. SW including half wind uncertainty is the most uncertain position considering latitude values and the most northerly position, what is expressed in the long whiskers found in figure 27(a). Most prominently is thereby, the long whisker indicating a great range of values deviating from the mean to the South. In addition, the upper and lower quantiles behave the same way, with a greater deviation from the mean to the South. Most precisely described are the latitude values in position j by WW, where the full wind data is introduced. In WW and WM however, the whiskers and the upper and lower quantiles indicate a

greater deviation from the mean towards the North. The longitude values then again indicate a different effect. NO is characterized by the lowest longitude values. WM thus are the most easterly positions. However, the differences are not as big as they are in latitude values. In figure 27(b) the box plots illustrate that the most uncertain position j is found in WW, where the whiskers as well as the upper and lower quantiles deviate the most from the mean value.

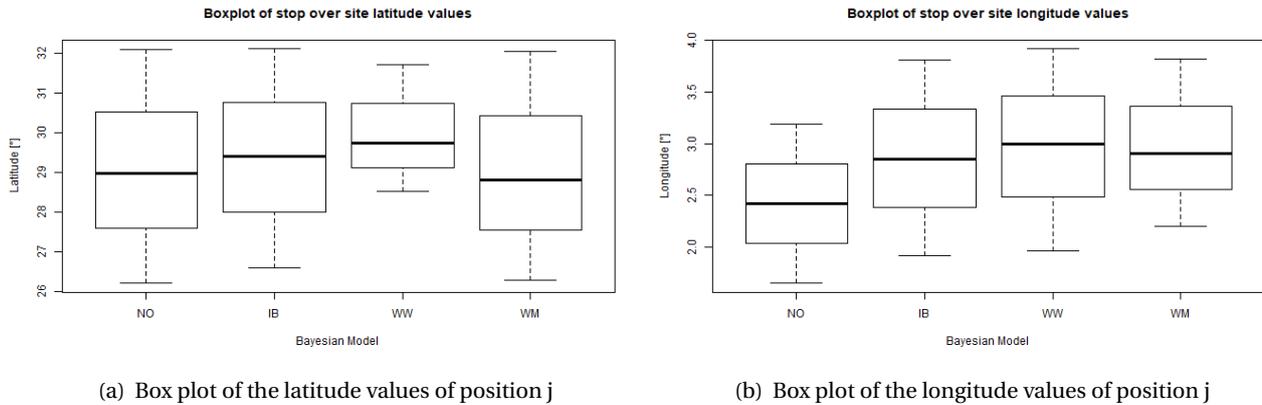


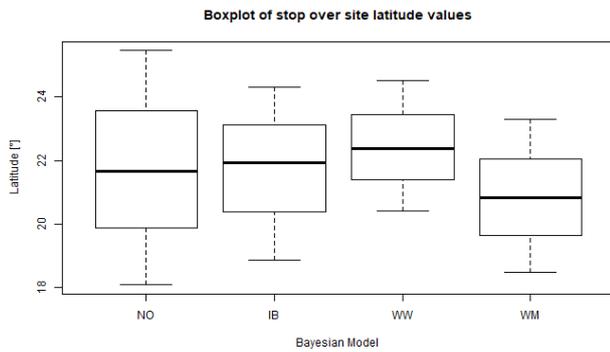
Figure 27: Box plot of the latitude and longitude values of position j found at the non-breeding locations

Positions k and l indicate the same results like position j illustrates. Latitude values are found within the same range but are significantly different. Therefore, small influences of the different model parameters are shown. The same effects are visible in the box plots of the positions' longitude values, where the four Bayesian inferences are proposing positions at the same range of longitude values, which nevertheless are significantly different. The box plots of positions k and l can be found in *Appendix - A*. The last position found in the location of is position m .

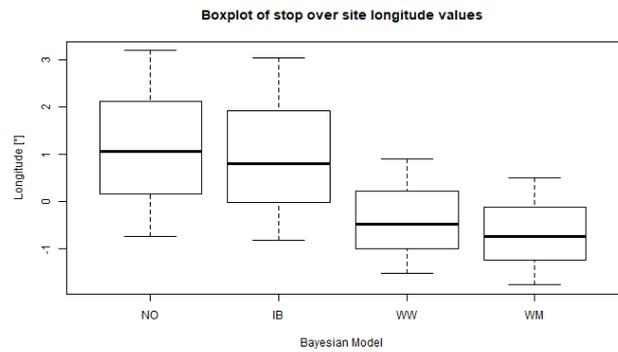
Position m is the position from where the birds start the migration in direction breeding grounds. This position results thus in interesting latitude and longitude values. In figure 28 on the next page the latitude and longitude values of position m are illustrated, indicating the respective uncertainties. On position m the uncertainties of the latitude values are still great, like in the positions before. However, in position m , NO is characterized by the greatest uncertainty of latitude values. WM is the most southerly position, followed by NO inference, where all other Bayesian inferences are found to be within the same latitude range. WW is characterized by the smallest uncertainty. The longitude values reveal a different picture than on the previous positions. Position m is found most westerly in WM, whereas WW indicates the most accurate position, being characterized by the smallest whiskers size. The other longitude values are closer, but are still significantly different. However, NO proposes the position with the greatest uncertainty, which is evenly spread around the mean, like WM. IB indicates instead a greater deviation from the mean towards East.

4.6 Positions estimated around the equinox

In this section the positions measured in the time of solar equinox are analyzed, like the positions of the non-breeding locations in the section before. Therefore, box plots of the positions analyzed are generated and compared. The positions found to be measured within the solar equinox are positions e , f and g on the migratory bouts from the breeding grounds to the non-breeding locations and the positions r , s and t on the migratory bouts from the non breeding locations back to the breeding sites respectively.

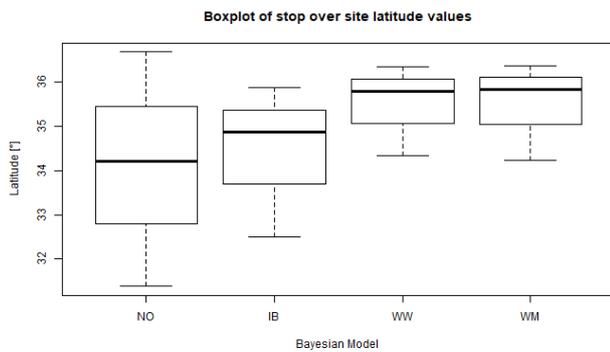


(a) Box plot of the latitude values of position m

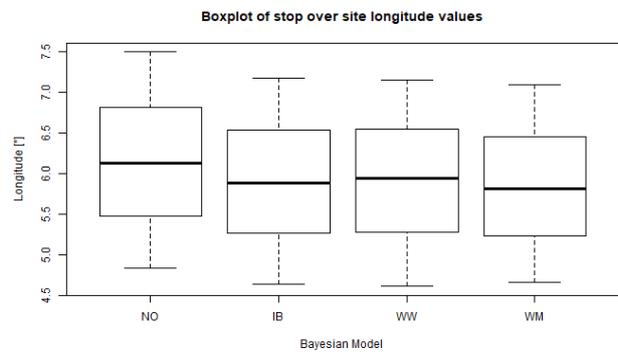


(b) Box plot of the longitude values of position m

Figure 28: Box plot of the latitude and longitude values of position m found at the locations of residence in Africa



(a) Box plot of the latitude values of position e



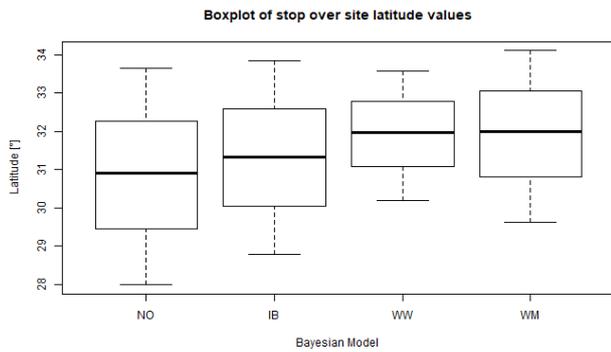
(b) Box plot of the longitude values of position e

Figure 29: Box plot of the latitude and longitude values of position e found to be measured around the solar equinox

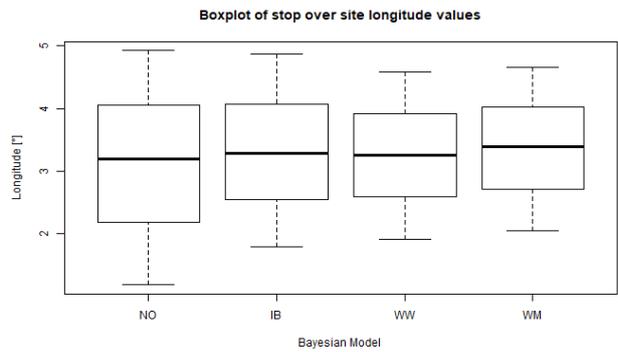
The latitude values of the position e result in a different picture than the previous positions (figure 29). The latitude values differ greatly between the different inferences. Especially NO differs from IB as well as WW differs from IB and WM. The most southerly position is inferred in NO, whereas the most northerly is found in WW. The latitude values of IB and WM are very close, where WM is the most northerly position. However, the uncertainty of the latitude values is greatest in NO. The most precise latitude values are inferred in WW. WM, WW and IB show a stronger deviation from the mean position towards the South than NO.

Another position found to be in the solar equinox is position g. To compare the positions' latitude and longitude values, box plots of each inference are plotted in figure 30. The latitude values are differing strongly, where the most extreme latitude value is found in SW, which is additionally characterized by the greatest uncertainty. NO and IB do not differ greatly, but are characterized by great uncertainties. The most precise latitude values are inferred to be in WW. The longitude values however, are more homogeneous. Like the latitude values, the most precise longitude values were inferred in WW.

The analysis of position f can be found in *Appendix - A*, because it indicates the same results, like positions e and g do. The following figure 31 illustrates the latitude and longitude values of position r which is found on the migratory bout from the non-breeding locations back to the breeding grounds. The latitude



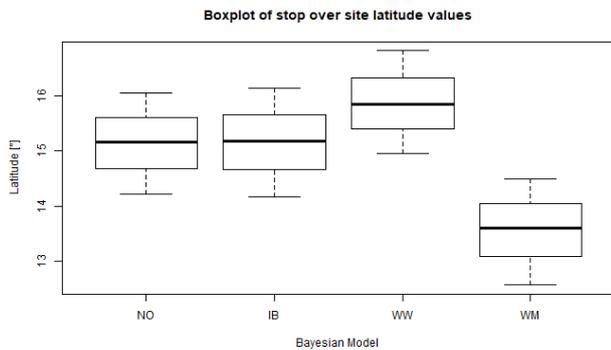
(a) Box plot of the latitude values of position g



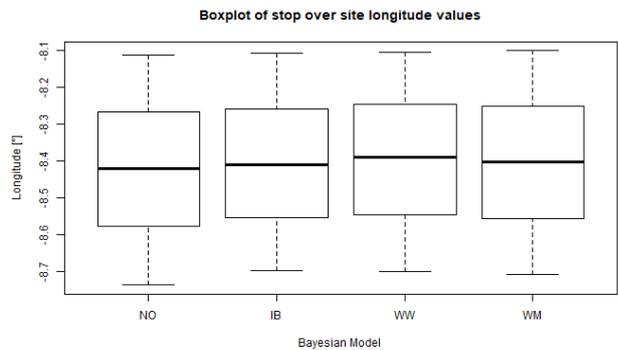
(b) Box plot of the longitude values of position g

Figure 30: Box plot of the latitude and longitude values of position g found at the solar equinox

values of the position r result in a big difference between WM and the other inferences, where WM is clearly the most southerly position. NO and IB however, almost do not differ. Furthermore, the uncertainties in



(a) Box plot of the latitude values of position r



(b) Box plot of the longitude values of position r

Figure 31: Box plot of the latitude and longitude values of position r found to be measured around the solar equinox

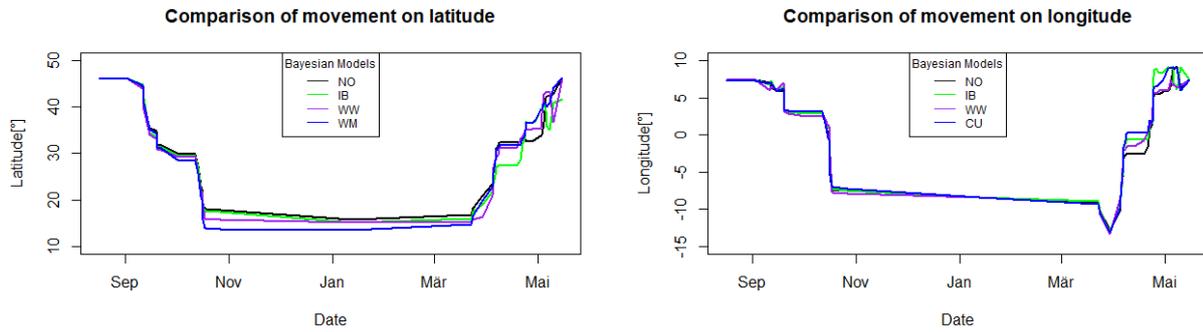
latitude values at position r seem not to differ much, where the uncertainties of IB and WM are slightly bigger. In WW, the latitude values are distributed more to the North. The longitude values however, illustrate a more homogeneous picture. The mean values do not differ significantly and the upper and lower quantiles are approximately of the same size. The whiskers too, are almost of the same size, where WW however, has a greater deviation from the mean values to the North. Additionally, WM seems to have slightly the most precise longitude values.

The remaining two positions found to be in the solar equinox can be found in *Appendix - A*. These two positions (s and t) illustrate approximately the same trends illustrated in figure 31.

4.7 Entire migration route

In this section the entire migration routes of all Bayesian inferences are compared. Therefore, the latitude and longitude values of the entire migration are plotted on one graph in chronological order. This plot illustrates the movement of the bird in both directions, latitude and longitude. For better comparison of the latitude and longitude values of the mean tracks of each inference are plotted in the same graph in figures

32(a) and 32(b). The black line illustrates the mean track of NO, the green line the mean track of IB, the orange line the mean track of WW and the blue line the mean track of WM.



(a) Latitude values of the entire migration illustrating all four Bayesian inferences

(b) Longitude values of the entire migration illustrating all four Bayesian inferences

Figure 32: Plot of the latitude and longitude values of the entire migration route representing all four Bayesian inferences

The latitude values are differing more over the entire migration route than the longitude values. Figure 32(a) illustrates the four different inference results of the latitude values recorded over the entire migration. At the start of the migration, the bird positions do not differ extremely. The first differences are found at the first stop over site, where the NO inference is estimated to be more North than the other inferences. WW is the most southerly route, whereas WM is more northerly than IB. The migration bouts on the second great migration step are close together. On the arrival at the non-breeding locations, the WM inference proposes the most southerly positions and the NO inference proposes the most northerly positions. In January, the non-breeding location is in every inference estimated to be at roughly the same latitude, with exception of WM which is the most southerly position and NO which proposes the most northerly positions. At the start of the migratory bouts from the non-breeding location to the breeding sites the NO inference has the strongest northbound trend in the first migratory bout, whereas the WW starts later with the northbound movement. The first stop over position between the northward migratory bouts are distributed across different latitude values. The second migration bout however is characterized by different flight paths resulting in different latitude values.

The longitude values are estimated to be very similar on the South bound migration bout. The greatest differences are found between WW and WM at the first stop over site, where the WM inference is estimated to arrive slightly more Westerly than the other inferences. On arrival at the non-breeding locations the WW inference is estimated to arrive slightly more to the West than the other inferences. At the non-breeding location the inferences trend to move constantly to the West, before the northbound migratory bout starts, where the IB does show the smallest West ward movement. On migration bout from the non-breeding location back to the breeding grounds, the first stop over site the positions are estimated to spread in terms of longitude values, where the WM inference is the most Easterly and NO is estimated to be the most Westerly position. The second migration bout is again very homogeneous, whereas the final bouts diverge heavily along longitude.

5 Discussion and conclusion

In this thesis an algorithm is implemented to infer the migration route of European Hoopoes. To run the route inference, the Metropolis - Hastings technique is executed to estimate the positions along the migration route. The positions were estimated with light positioning, wherefore measured light intensities are used to compute possible positions and propose a first migration route. This route is used in SGAT to infer the most likely migration routes and their respective probabilities (Sumner et al., 2009; Dean et al., 2013; Lisovski et al., 2012; Lisovski and Hahn, 2012; Dhanjal-Adams et al., 2018). This method is applied to tackle the problem of position estimation by light intensities. Solar positioning is known to result in big positional errors of several 10 - 100 km. Especially around solar equinox, where equal day length every where on the planet poses some difficulties and in areas greatly impacted by shadowing effects, result in errors of $\pm 300 km$ (Lisovski et al., 2012; Lisovski and Hahn, 2012). These great uncertainties can be reduced by introducing more measured data, like bird activity, air pressure and wind data. Therefore, the new input parameters introduced to the Bayesian inference are summarized to a mathematical model, which contributes to the mathematical description of the likelihood of a proposed position (Sumner et al., 2009). To answer the first research question: "How do endogenous and environmental factors affect the migration route inference?" three different models were developed to include different wind conditions to the route inference. Wind models which trust the mean wind speed more and thus have smaller standard deviations (WM), result in narrower areas where the bird was expected to fly. Wind models which have a greater wind speed standard deviation result however, in wider area of probability. In respect to these findings, the endogenous and environmental conditions show to have an influence on the size and position of the probability areas.

Additionally, different wind conditions and activity times result in different flight distances a bird most likely did fly. The activity model influences the flight time a bird was flying on a migratory bout, which is mostly constrained by the physiological capabilities of the bird (Pennycuick, 1998). Wind conditions did change the ground speed the bird did probably fly. Not only did favourable conditions increase the ground speed (Safi et al., 2013) and thus, the flight distances (Åkesson and Hedenström, 2000), but the side winds did change the longitudinal position of the bird. The combination of activity and wind conditions thus is not only a pushing factor, but as well a limiting factor for the flight distances and directions. These conditions affect the migration route inference in such a way, that the positions along the route are estimated to be more to the North, South, East or West than they were estimated without the inclusion of the activity and wind data. Applying different wind models, in which different wind speed standard deviations are applied result in different ground speeds and thus in different inference results. The non-breeding locations of residence are strongly affected by the wind conditions found en route. The smaller the standard deviation of the wind speeds is the greater the influence of the wind on the position estimation. On arrival at the non-breeding location, the inferences which include wind and activity data, propose more northerly positions. Additionally, the inferences which include wind and activity data show great differences. The inferences which trust the wind speed more (small standard deviation) propose the first position at the non-breeding location to be more to the North than the others. In addition to that, the influence of wind is shown on the longitude values, too. The inferences including wind and activity data, propose positions more to the West than the NO inference. At the following non-breeding locations the influence of the wind conditions are especially shown in the estimation of the latitude values of the positions. The head winds shown in figures 23(b) to 25(b) affect the route inferences in two ways. First, in inferences, where strong head winds

where assumed (IB and WW), the positions are estimated to be more to the North than in NO. Secondly, in inferences, where weak head winds were assumed (WM), the position is found more to the South than in NO. Finally, at the positions estimated around the solar equinox the inferences show similar effects. In dependence of the wind conditions found en route, the positions are found more to the North or more to the South. The longitude values are not affected to a great degree. This means, the winds included in these migration bouts, are winds which blew more in the direction of the bird movement (tail wind) or against the direction of the bird movement (head wind).

Finally, the endogenous and environmental factors have an influence on the variance of the migration route inference. The activity and wind data have proven to explain the solar position estimation better than the basic activity model (BF 1.20, 2.02 and 4.96).

The Bayes factors furthermore, are used to answer research question 2: "Do activity and wind data improve the migration route inference?". With this study the second research question is answered with yes. To answer this question, the Bayes factors of the different route inferences are computed. The BF of the route inferences which include activity and wind data are greater than 1 (table 2). A Bayes factor greater than one shows that the marginal likelihood in the numerator is greater than the marginal likelihood in the denominator. Especially, the uncertainty of the wind speed data which is included to the movement model has an influence on the BF. Smaller standard deviations of the wind speed like in WM result in greater marginal likelihoods and thus in greater BF, when compared to NO. This is because the narrower wind speed density distribution forces the inference to accept the bird movement in the wind direction. This movement is achieved by assigning the mean wind speed greater probabilities than the wind speeds deviating from the mean. Therefore, the most likely flight distances have smaller standard deviations, too. As a consequence, the area of possible positions decreases and the possible positions are assigned greater probabilities, thus the marginal likelihood increases. The WM inference is best explaining the positions estimated with light data (BF = 4.96 vs NO). WM has a greater BF value compared to IB and WW, too (table 2). To answer this question, secondly, the standard deviation and the confidence interval of the positions along the migration route are compared. Especially at the non-breeding locations, the results show an improvement of position estimation. The wind and activity data decrease the standard deviation and the confidence intervals of the latitude positions. Figures 26 to 28, show that WW proposes the smallest standard deviation and confidence intervals on latitude values, whereas WM is estimating smaller standard deviations and confidence intervals on longitude values. This is explained with the standard deviation of the wind speed data. This is explained with the standard deviation of the wind speed data. The smaller the wind speed standard deviation, the smaller the position standard deviation. In this case, the smaller wind speed supporting the bird results in a smaller standard deviation in WW, whereas WM has a constant standard deviation, which in this case is greater than in WW. As a consequence the most likely flying distance in WW is assigned a smaller standard deviation, which favours a position by assigning this position a great probability. At the positions estimated around the equinox however, this effect is more evident. The models including wind and activity data with smaller standard deviations result in smaller standard deviations of the positions.

The third research question: "Do wind conditions en route change the stop over positions in the route inference?" is answered with yes, too. The wind conditions en route change the stop over positions in the route inference. The stop over positions change in absolute values, in dependence of the wind speed and the wind speed standard deviation. Strong head winds push the inferred stop over positions further to the North, whereas strong tail winds push the inferred stop over positions further to the South. In addition to

that, the wind drift effects are considered in the migration route inference, too. The longitude values are affected by the wind speeds included to the route inference. Especially at the non-breeding locations these effects are shown, where the WM inferences is most affected by the wind conditions.

The inferred routes including different wind modelling techniques returned migration routes, which propose the crossing of the Mediterranean over the islands of Sardinia and Corsica. This migration route is one of two suggested migration routes across the Mediterranean. The two main migration routes are suggested to fly over Sardinia and Corsica or over the Balearic islands (Reichlin et al., 2009; van Wijk et al., 2018). In addition to that, the area of probability follows the Italian West coast, what is found to be a common flight behaviour of Hoopoes. (Åkesson, 1993; Alerstam, 2011; Meier et al., 2018). However, the routes inferred in this study propose ring shaped migration routes and thus contradict previous migration patterns (Reichlin et al., 2009). The route inference analyzed in this thesis, proposes two main migratory bouts on the way from the breeding grounds to the non-breeding grounds and on the way back, thus indicating the importance of stop over sites, which are used to refuel after long migratory bouts (Erni et al., 2002; Schmaljohann and Dierschke, 2005).

To answer the three research questions, this thesis expanded an existing tool for bird migration route inference (SGAT). SGAT was extended to include bird activity and wind data, to improve the route inference. To include the wind data into the route migration inference, the wind speed vector found en route is included into the bird flight speed model, to influence the most likely distances flown on a migratory bout. For inclusion of the wind speed, the wind speed was partitioned into two vector components, one parallel to the bird flight vector v_{wa} and one perpendicular to the bird flight vector v_c (Safi et al., 2013). After wind partitioning, the wind support was added to the air speed to compute the ground speed. This method of wind speed computation is different to the method proposed by Liechti et al. (1994), which suggests to add the air speed and wind speed vector to compute the ground speed vector (Liechti et al., 1994). The different ground speed computation which is applied in this thesis, assumes a bit greater wind support than the method suggested by Liechti et al. (1994). In addition to that, the air speed retrieved from literature ($13.89 \frac{m}{s}$), is considered to be a high flight speed, what enables greater flight distances. This wind model however, has the advantage of including the wind conditions and the bird activity data in a simple way. In addition it proves to be a tool which can be used flexibly, to analyze the effect of different wind and air speed conditions.

6 Conclusion and future work

Bird migration is a complex phenomenon, where many different parameters contribute to the final route. In this study the most likely migration route of a European Hoopoe was inferred, executing MCMC techniques. To infer the most probable migration route, the bird was equipped with a geolocator, which recorded amongst other parameters, light intensities, air pressure and activity. The light intensities were applied for position estimation, executing astronomical formulae. The air pressure and the activity data were applied for stop over site detection. In addition to these data, the inference was executed with the inclusion of wind data, measured by the European Union's Copernicus program. The inferences were conducted, by parametrizing the wind data differently. First, the biological capabilities of the bird were changed, introducing a reduced wind use efficiency, wherefore the bird was set to only use half of the available wind speed. Secondly, the wind data was introduced to the inference with unchanged wind speed use efficiency but a standard deviation which is increasing with increasing speeds. Finally, the wind data was modelled with a normal distribution, wherefore a constant standard deviation found in literature was applied.

The resulting migration routes were then compared in regard to different aspects. Most importantly, including wind and activity data to the bird migration route inference improves the inference results. Additionally, including wind data with small standard deviations reduces the standard deviation of the positions. Not only at the non-breeding location but as well at positions estimated around the solar equinox the positions are estimated to have smaller standard deviations compared to previous inference methods. Therefore, it can be said, that migration route inferences including wind conditions return improved results. However, the inference conducted with the constant wind standard deviation returns the best migration route inference in terms of position probability. Comparing the marginal likelihoods by computing the Bayes factor reveals, the WM inference being the best inference of those four. The half wind use efficiency inference however does not explain the observed and inferred phenomenon as good as the other three models did but does still explain the bird migration better than the original inference, which neglected bird activity and wind conditions.

With this new developed model which was included to an existing tool (SGAT) improved route migrations, including wind and activity data. Additionally, the data used for route inference explains the bird migration and thus the position estimated by light better than the original inference.

With this inference, further experiments with different wind use efficiencies or different wind support estimations can be conducted. Furthermore, the wind conditions found en route can be used to analyze first the compensation a bird must carry out to reach its desired destination. Secondly, the energy expenditure to compensate the wind conditions found en route can be estimated. Additionally, the influence of changing environmental conditions can be analyzed with this tool.

List of Figures

1	European Hoopoe	10
2	Vector components	13
3	SOI-GDL3pam geolocator	14
4	Illustration of the U and V wind components	15
5	Threshold method for light positioning	16
6	Light image showing day and night levels	17
7	First route proposal	18
8	Aggregated route proposal	19
9	Schematic sketch of the MCMC method	20
10	Basic gamma distribution	22
11	MCMC result	23
12	General work flow	27
13	Flight distances in relation to flight time	29
14	Wind database schema	30
15	Sketch to illustrate the basic idea of a spatial index applied on the available wind data	31
16	Wind situation en route	32
17	Vector geometry	33
18	Flying distances in relation to flight time	34
19	Different density functions applied	36
20	WM route inference	37
21	WW route inference	39
22	IB route inference	41
23	Wind speeds and wind support found on the mean track of IB	42
24	Wind speeds and wind support found on the mean track of WW	43
25	Wind speeds and wind support found on the mean track of WM	43
26	Box plots of position i	44
27	Box plots of position j	45
28	Box plots of position m	46
29	Box plots of position e	46
30	Box plots of position g	47
31	Box plots of position r	47
32	Entire route coordinates plot	48
33	Box plots of position k	68
34	Box plots of position l	68
35	Box plots of position s	68
36	Box plots of position t	69

List of Tables

1	Subset of available pressure levels used in this study for wind speed retrieval and the pressure level ranges (Modellevels) which are aggregated to one pressure level. Table adapted from (Berrisford et al., 2011)	15
2	Look up table for all Bayes factors computed to compare and evaluate the different models applied in this study. This table is read row wise, where the result represents the division of the row value by the column value.	38
3	Listing of the summary of the migration route	70
4	Listing of track positions including half wind speed data in to route inference	71
5	Listing of track positions after including full wind data to the route inference	72
6	Listing of track positions after including normal distributed wind data to the route inference .	73

References

- Åkesson, S. (1993), 'Coastal Migration and Wind Drift Compensation in Nocturnal Passerine Migrants', *Ornis Scandinavica* **24**(2), 87–94.
- Åkesson, S. and Hedenström, A. (2000), 'Wind selectivity of migratory flight departures in birds', *Behavioral Ecology and Sociobiology* **47**, 140–144.
- Alerstam, T. (1979), 'Wind as Selective Agent in Bird Migration', *Ornis Scaniavica* **10**(1), 76–93.
- Alerstam, T. (2011), 'Optimal bird migration revisited', *Journal of Ornithology* **152**, 5–23.
- Alerstam, T. and Lindström, A. (1990), Optimal Bird Migration: The Relative Importance of Time, Energy, and Safety, in 'Bird Migration: Physiology and Ecophysiology', Springer, Berlin Heidelberg, pp. 331–351.
- Aurbach, A., Schmid, B., Liechti, F., Chokani, N. and Abhari, R. (2020), 'Simulation of broad front bird migration across Western Europe', *Ecological Modelling* **415**(November 2019), 1–11.
- Bächler, E., Hahn, S., Schaub, M., Arlettaz, R., Jenni, L., Fox, J. W., Afanasyev, V. and Liechti, F. (2010), 'Year-round tracking of small trans-Saharan migrants using light-level geolocators', *PLoS ONE* **5**(3), 1–6.
- Bäckman, J., Andersson, A., Alerstam, T., Pedersen, L., Sjöberg, S., Thorup, K. and Tøttrup, A. P. (2017), 'Activity and migratory flights of individual free-flying songbirds throughout the annual cycle: method and first case study', *Journal of Avian Biology* **48**(2), 309–319.
- Barbaro, L., Couzi, L., Bretagnolle, V., Nezan, J. and Vetillard, F. (2008), 'Multi-scale habitat selection and foraging ecology of the eurasian hoopoe (*Upupa epops*) in pine plantations', *Biodiversity and Conservation* **17**, 1073–1087.
- Basyir, M., Nasir, M., Suryati, S. and Mellyssa, W. (2018), 'Determination of Nearest Emergency Service Office using Haversine Formula Based on Android Platform', *EMITTER International Journal of Engineering Technology* **5**(2), 270–278.
- Berger, J. O. and Pericchi, L. R. (1996), 'The Intrinsic Bayes Factor for Model Selection and Prediction', *Journal of the American Statistical Association* **91**(443), 109–122.
- Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S. and Simmons, A. (2011), 'The ERA-Interim Archive', *ERA-Interim archive* **2**, 1–27.
- Binz, J., Friedrich, M., Gebauer, P., Höhn, W., Ineichen, R., Kappus, H., Marzetta, B., Schilt, H., Schnyder, A., Thöni, W., Bächtiger, P., Ess, H., Etter, K., Florin, R., Joho, P., Plüss, A., Rosenthaler, E., Rüetschi, R., Speich, H. M. and Walter, M. E. (1981), *Formeln und Tafeln, Mathematik - Statistik - Physik*, Orell Füssli Verlag Zürich.
- Blackburn, E. and Cresswell, W. (2016), 'High winter site fidelity in a long-distance migrant: implications for wintering ecology and survival estimates', *Journal of Ornithology* **157**(1), 93–108.
- Bolch, R. and Bruderer, B. (1982), 'The Air Speed of Migrating Bird and its Relationship to the Wind', *Behavioral Ecology and Sociobiology* **11**, 19–24.

- Bruderer, B. and Boldt, A. (2001), 'Flight characteristics of Birds: I. radar measurements of speeds', *Ibis* **143**, 178–204.
- Chan, J. C. and Grant, A. L. (2015), 'Pitfalls of estimating the marginal likelihood using the modified harmonic mean', *Economics Letters* **131**, 29–33.
URL: <http://dx.doi.org/10.1016/j.econlet.2015.03.036>
- Chib, S. and Greenberg, E. (1995), 'Understanding the Metropolis-Hastings Algorithm', *The American Statistician Association* **49**(4), 327–335.
- Chib, S. and Jeliazkov, I. (2001), 'Marginal likelihood from the metropolis–hastings output', *Journal of the American Statistical Association* **96**(453), 270–281.
- Conklin, J. R., Battley, P. F. and Potter, M. A. (2013), 'Absolute Consistency: Individual versus Population Variation in Annual-Cycle Schedules of a Long-Distance Migrant Bird', *PLoS ONE* **8**(1), 1–9.
- Cooke, W. W. (1915), *Bird migration*, U.S. Dept. of Agriculture,, Washington, D.C. :.
URL: <https://www.biodiversitylibrary.org/bibliography/64867> <http://www.biodiversitylibrary.org/bibliography/64867>
- De Santis, F. and Spezzaferri, F. (1997), 'Alternative Bayes factors for model selection', *The Canadian Journal of Statistics* **25**(4), 503–515.
- Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R. A., Perrins, C. M. and Guilford, T. (2013), 'Behavioural mapping of a pelagic seabird: Combining multiple sensors and a hidden Markov model reveals the distribution of at-sea behaviour', *Journal of the Royal Society Interface* **10**, 1–12.
- Demšar, U., Harris, P., Brunson, C., Fotheringham, A. S. and McLoone, S. (2013), 'Principal Component Analysis on Spatial Data: An Overview', *Annals of the Association of American Geographers* **103**(1), 106–128.
- Dhanjal-Adams, K. L. (2019), 'PAMLR Manual'.
URL: <https://kiranlda.github.io/PAMLRManual/index.html>
- Dhanjal-Adams, K. L., Bauer, S., Emmenegger, T., Hahn, S., Lisovski, S. and Liechti, F. (2018), 'Spatiotemporal Group Dynamics in a Long-Distance Migratory Bird', *Current Biology* **28**(17), 2824–2830.e3.
- Diniz, S. M., Sadek, F. and Simiu, E. (2004), 'Wind speed estimation uncertainties: Effects of climatological and micrometeorological parameters', *Probabilistic Engineering Mechanics* **19**(4), 361–371.
- Ekstrom, P. A. (2004), 'An advance in geolocation by light', *Memoirs of National Institute of Polar Research. Special issue* **58**, 210–226.
- Erni, B., Liechti, F. and Bruderer, B. (2002), 'Stopover strategies in passerine bird migration: A simulation study', *Journal of Theoretical Biology* **219**, 479–493.
- European Centre for Medium-Range Weather Forecast (2020), 'Advancing global NWP through international collaboration'.
URL: <https://www.ecmwf.int/>

- Fang, Y., Friedman, M., Nair, G., Rys, M. and Schmidt, A. E. (2008), Spatial indexing in microsoft SQL server 2008, *in* 'Proceedings of the ACM SIGMOD International Conference on Management of Data', pp. 1207–1215.
- Finch, T. (2009), 'Incremental calculation of weighted mean and variance', *University of Cambridge* **1**(February), 1–8.
- Fröhlich, L. (2018), *PostgreSQL 10 Praxisbuch für Administratoren und Entwickler*, Carl Hanser Verlag München, München.
- Fusani, L. and Gwinner, E. (2004), 'Simulation of migratory flight and stopover affects night levels of melatonin in a nocturnal migrant', *Proceedings of the Royal Society B: Biological Sciences* **271**, 205–211.
- Gelfand, A. E. and Dey, D. (1994), 'Bayesian Model Choice : Asymptotics and Exact Calculations', *Royal Statistical Society* **56**, 501–514.
- Goodman, P. S. D. (1986), 'DETERMINING FISH MOVEMENTS FROM AN "ARCHIVAL" TAG : PRECISION OF GEOGRAPHICAL OF SWIMMING TEMPERATURE AND DEPTH', *NOAA Technical Memorandum NMFS National Oceanic and Atmospheric Administration (NOAA)* **5**.
- Grishchenko, M., Prins, H. H. T., Ydenberg, R. C., Schaepman, M. E., Boer, W. F. and Knegt, H. J. (2019), 'Land use change and the migration geography of Greater White-fronted geese in European Russia', *Ecosphere* **10**, 1–11.
- Hastings, W. K. (1970), 'Monte carlo sampling methods using Markov chains and their applications', *Biometrika* **57**(1), 97–109.
- Hernández-Pliego, J., Rodríguez, C., Dell'Omo, G. and Bustamante, J. (2017), 'Combined use of tri-Axial accelerometers and GPS reveals the flexible foraging strategy of a bird in relation to weather conditions', *PLoS ONE* **12**(6), 1–29.
- Hijmans, R. J. (2018), Spatial Data Analysis with R, *in* 'Spatial Data Science with R', rspatial, chapter 3, pp. 1–74.
URL: <https://rspatial.org/index.html>
- Hill, R. D. and Braun, M. J. (2001), 'Geolocation by Light Level', *Electronic Tagging and Tracking in Marine Fisheries* **1**, 315–330.
- Kulkarni, H. V. and Powar, S. K. (2011), 'A simple normal approximation for weibull distribution with application to estimation of upper prediction limit', *Journal of Probability and Statistics* **November**.
- Liechti, F. (1995), 'Modelling optimal heading and airspeed of migrating birds in relation to energy expenditure and wind influence', *Journal of Avian Biology* **26**, 330–336.
- Liechti, F. (2006), 'Birds: Blowin' by the wind?', *Journal of Ornithology* **147**, 202–211.
- Liechti, F., Bauer, S., Dhanjal-Adams, K. L., Emmenegger, T., Zehtindjiev, P. and Hahn, S. (2018), 'Miniaturized multi-sensor loggers provide new insight into year-round flight behaviour of small trans-Sahara avian migrants', *Movement Ecology* **6**(1), 1–10.

- Liechti, F. and Bruderer, B. (1995), 'Direction, speed and composition of nocturnal bird migration in the south of Israel', *Israel Journal of Zoology* **41**, 501–515.
- Liechti, F., Hendenström, A. and Alerstam, T. (1994), 'Effects of Sidewinds on Optimal Flight Speed of Birds', *Journal of Theoretical Biology* **170**, 219–225.
- Liechti, F., Scandola, C., Rubolini, D., Ambrosini, R., Korner-Nievergelt, E., Hahn, S., Lardelli, R., Romano, M., Caprioli, M., Romano, A., Sicurella, B. and Saino, N. (2015), 'Timing of migration and residence areas during the non-breeding period of barn swallows *Hirundo rustica* in relation to sex and population', *Journal of Avian Biology* **46**(3), 254–265.
- Liechti, F., Witvliet, W., Weber, R. and Bächler, E. (2013), 'First evidence of a 200-day non-stop flight in a bird', *Nature Communications* **4**, 1–7.
- Lisovski, S., Bauer, S., Briedis, M., Davidson, S. C., Dhanjal-Adams, K. L., Hallworth, M. T., Karagicheva, J., Meier, C. M., Merkel, B., Ouwehand, J., Pedersen, L., Rakhimberdiev, E., Roberto-Charron, A., Seavy, N. E., Sumner, M. D., Taylor, C. M., Wotherspoon, S. J. and Bridge, E. S. (2019), 'Light-level geolocator analyses: A user's guide', *Journal of Animal Ecology* **89**(1), 221–236.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.13036>
- Lisovski, S. and Hahn, S. (2012), 'GeoLight - processing and analysing light-based geolocator data in R', *Methods in Ecology and Evolution* **3**(6), 1055–1059.
- Lisovski, S., Hewson, C. M., Klaassen, R. H., Korner-Nievergelt, E., Kristensen, M. W. and Hahn, S. (2012), 'Geolocation by light: Accuracy and precision affected by environmental factors', *Methods in Ecology and Evolution* **3**(3), 603–612.
- Lukacs, E. (1955), 'A Characterization of the Gamma Distribution', *The Annals of Mathematical Statistics* **26**, 319–324.
- Makris, A., Tserpes, K., Spiliopoulos, G. and Anagnostopoulos, D. (2019), Performance evaluation of MongoDB and PostgreSQL for spatio-temporal data, in 'CEUR Workshop Proceedings', Vol. 2322.
- McElreath, R. (2015), *Statistical Rethinking A Bayesian Course with Examples in R and Stan*, 2nd edn, Apple Academic Press Inc.
- Meier, C. M., Karaardıç, H., Aymí, R., Peev, S. G., Bächler, E., Weber, R., Witvliet, W. and Liechti, F. (2018), 'What makes Alpine swift ascend at twilight? Novel geolocators reveal year-round flight behaviour', *Behavioral Ecology and Sociobiology* **72**(3).
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953), 'Equation of state calculations by fast computing machines', *The Journal of Chemical Physics* **21**, 1087–1092.
- Newton, I. (2008), *The Migration Ecology of Birds*, Elsevier.
- Nilsson, C., Bäckman, J. and Alerstam, T. (2014), 'Seasonal modulation of flight speed among nocturnal passerine migrants: differences between short- and long-distance migrants', *Behavioral Ecology and Sociobiology* **68**(11), 1799–1807.

- Pajor, A. (2017), 'Estimating the marginal likelihood using the arithmetic mean identity', *Bayesian Analysis* **12**(1), 261–287.
- Pennycuik, C. J. (1998), 'Computer simulation of fat and muscle burn in long-distance bird migration', *Journal of Theoretical Biology* **191**, 47–61.
- Raftery, A. E., Newton, M. A., Satagopan, J. M. and Krivitsky, P. N. (2007), Estimating the Integrated Likelihood via Posterior Simulation Using the Harmonic Mean Identity.
- Rakhimberdiev, E., Winkler, D. W., Bridge, E., Seavy, N. E., Sheldon, D., Piersma, T. and Saveliev, A. (2015), 'A hidden Markov model for reconstructing animal paths from solar geolocation loggers using templates for light intensity', *Movement Ecology* **3**(25), 1–15.
- Reichlin, T. S., Schaub, M., Menz, M. H. M., Mermod, M., Portner, P., Arlettaz, R. and Jenni, L. (2009), 'Migration patterns of Hoopoe *Upupa epops* and Wryneck *Jynx torquilla*: An analysis of European ring recoveries', *Journal of Ornithology* **150**, 393–400.
- Robert, C. P. and Wraith, D. (2009), Computational methods for Bayesian model choice, in 'AIP Conference Proceedings', Vol. 1193, pp. 251–262.
- Roberts, G. O. and Rosenthal, J. S. (2001), 'Optimal Scaling for Various Metropolis-Hastings Algorithms', *Statistical Science* **16**(4), 351–367.
- Roberts, G. O. and Smith, A. F. (1994), 'Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms', *Stochastic Processes and their Applications* **49**(2), 207–216.
- Roberts, G. O. and Stramer, O. (2001), 'On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm', *Biometrika* **88**(3), 603–621.
- Ruginski, I. T., Boone, A. P., Padilla, L. M., Liu, L., Heydari, N., Kramer, H. S., Hegarty, M., Thompson, W. B., House, D. H. and Creem-Regehr, S. H. (2016), 'Non-expert interpretations of hurricane forecast uncertainty visualizations', *Spatial Cognition and Computation* **16**(2), 154–172.
- Safi, K., Kranstauber, B., Weinzierl, R., Griffin, L., Rees, E. C., Cabot, D., Cruz, S., Proaño, C., Takekawa, J. Y., Newman, S. H., Waldenström, J., Bengtsson, D., Kays, R., Wikelski, M. and Bohrer, G. (2013), 'Flying with the wind: Scale dependency of speed and direction measurements in modelling wind support in avian flight', *Movement Ecology* **1**(4), 1–13.
- Saghir, A., Hamedani, G. G., Tazeem, S. and Khadim, A. (2017), 'Weighted Distributions: A Brief Review, Perspective and Characterizations', *International Journal of Statistics and Probability* **6**(3), 109.
- Schmaljohann, H. and Dierschke, V. (2005), 'Optimal bird migration and predation risk: A field experiment with northern wheatears *Oenanthe oenanthe*', *Journal of Animal Ecology* **74**, 131–138.
- Sheldon, D., Farnsworth, A., Irvine, J., Van Doren, B., Webb, K., Dietterich, T. G. and Kelling, S. (2013), Approximate Bayesian inference for reconstructing velocities of migrating birds from weather radar, in 'Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI 2013', pp. 1334–1340.

- Sheldon, D., Saleh Elmohamed, M. A. and Kozen, D. (2009), Collective inference on Markov models for modeling bird migration, *in* 'Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference', pp. 1–8.
- Simion, B., Ilha, D. N., Brown, A. D. and Johnson, R. (2013), The price of generality in spatial indexing, *in* 'Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, BigSpatial 2013', pp. 8–12.
- Stacy, E. and Mihram, G. (1965), 'Estimation for a Generalized Gamma Distribution', *Technometrics* **7**(3), 349–358.
- Storrer, H. H. (2009), *Einführung in die mathematische Behandlung der Naturwissenschaften I*, 3rd edn, Birkhäuser Verlag AG, Basel.
- Stutchbury, B. J. M., Tarof, S. A., Done, T., Gow, E., Kramer, P. M., Tautin, J., Fox, J. W. and Afanasyev, V. (2009), 'Tracking long-distance songbird migration by using geolocators', *Science* **323**(5916), 896.
- Sumner, M. D., Wotherspoon, S. J. and Hindell, M. A. (2009), 'Bayesian Estimation of Animal Movement from Archival and Satellite Tags', *PLoS ONE* **4**(10), 19–22.
- Thom, H. C. S. (1958), 'A Note on the Gamma Distribution', *Monthly Weather Review* **86**(4), 117–122.
- Tipping, M. and Faul, A. C. (2003), Fast Marginal Likelihood Maximization for Sparse Bayesian Models, *in* 'Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics', Key West FL, pp. 1–5.
- van Wijk, R. E., Bauer, S. and Schaub, M. (2016), 'Repeatability of individual migration routes, wintering sites, and timing in a long-distance migrant bird', *Ecology and Evolution* **6**, 8679–8685.
- van Wijk, R. E., Schaub, M. and Bauer, S. (2017), 'Dependencies in the timing of activities weaken over the annual cycle in a long-distance migratory bird', *Behavioral Ecology and Sociobiology* **71**(73), 1–8.
- van Wijk, R. E., Schaub, M., Hahn, S., Juárez-García-Pelayo, N., Schäfer, B., Viktora, L., Martín-Vivaldi, M., Zischewski, M. and Bauer, S. (2018), 'Diverse migration strategies in hoopoes (*Upupa epops*) lead to weak spatial but strong temporal connectivity', *Science of Nature* **105**(42), 1–9.
- van Wijk, R. E., Souchay, G., Jenni-Eiermann, S., Bauer, S. and Schaub, M. (2016), 'No detectable effects of lightweight geolocators on a Palaearctic-African long-distance migrant', *Journal of Ornithology* **157**(1), 255–264.
- Wang, J., Niu, T., Lu, H., Guo, Z., Yang, W. and Du, P. (2018), 'An analysis-forecast system for uncertainty modeling of wind speed: A case study of large-scale wind farms', *Applied Energy* **211** (July 2017), 492–512.
URL: <https://doi.org/10.1016/j.apenergy.2017.11.071>
- Williams, H. J., Holton, M. D., Shepard, E. L., Largey, N., Norman, B., Ryan, P. G., Duriez, O., Scantlebury, M., Quintana, F., Magowan, E. A., Marks, N. J., Alagaili, A. N., Bennett, N. C. and Wilson, R. P. (2017), 'Identification of animal movement patterns using tri-axial magnetometry', *Movement Ecology* **5**(6), 1–14.

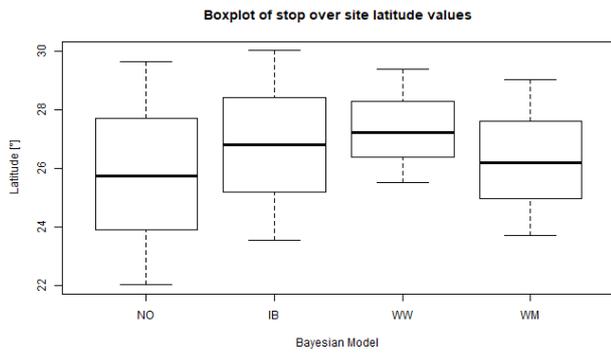
Zhang, J., Gertz, M. and Grauewnwald, L. (2009), Efficiently Managing Large Scale Raster Species Distribution Data in PostgreSQL Categories and Subject Descriptors, *in* '17th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, ACM-GIS 2009', pp. 316–325.

Zhang, L. and Yi, J. (2010), Management methods of spatial data based on PostGIS, *in* '2010 2nd Pacific-Asia Conference on Circuits, Communications and System, PACCS 2010', Vol. 1, IEEE, pp. 410–413.

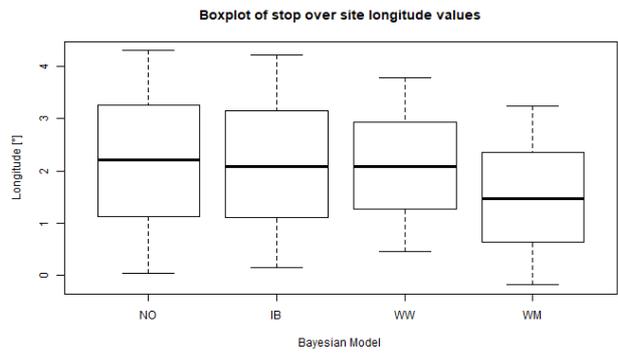
Appendix - A

In this section the tabular summaries of all migration routes are presented. Each table contains information of all point positions found in a Markov Chain, generated by the route inference algorithm. The tables are ordered as follows:

- Box plot of the latitude values of position k
- Box plot of the longitude values of position k
- Box plot of the latitude values of position l
- Box plot of the longitude values of position l
- Box plot of the latitude values of position s
- Box plot of the longitude values of position s
- Box plot of the latitude values of position t
- Box plot of the longitude values of position t
- Listing of the summary of the migration route
- Listing of track positions including half wind speed data in to route inference
- Listing of track positions including half wind speed uncertainty in to route inference
- Listing of track positions after including full wind data to the route inference

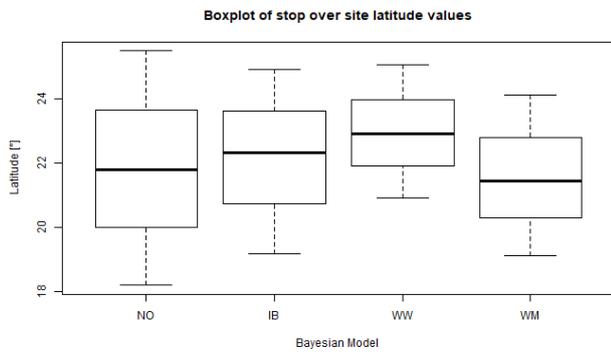


(a) Box plot of the longitude values of position k

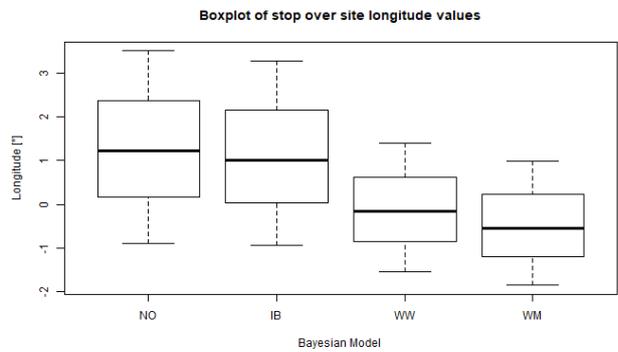


(b) Box plot of the longitude values of position k

Figure 33: Box plot of the latitude and longitude values of position k found at the locations of residence in Africa

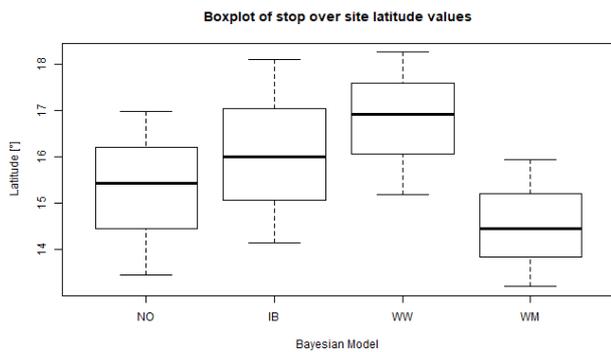


(a) Box plot of the longitude values of position l

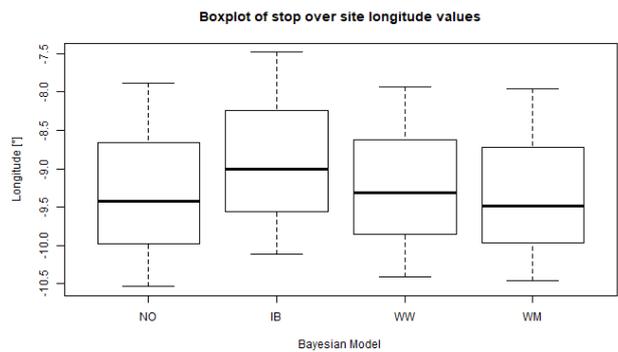


(b) Box plot of the longitude values of position l

Figure 34: Box plot of the latitude and longitude values of position l found at the locations of residence in Africa

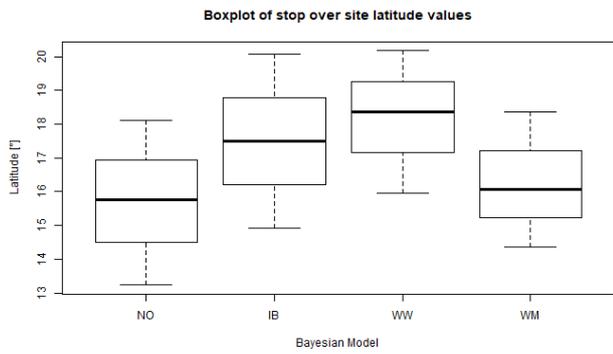


(a) Box plot of the latitude values of position s

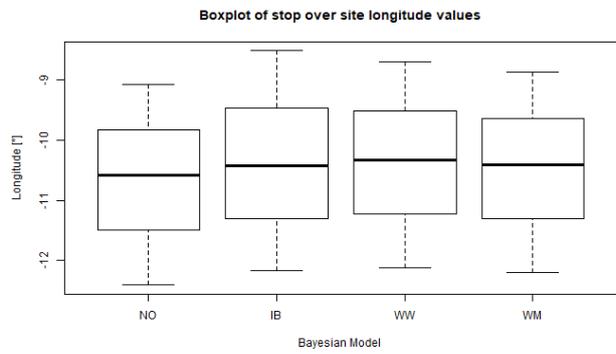


(b) Box plot of the longitude values of position s

Figure 35: Box plot of the latitude and longitude values of position s found to be measured around the solar equinox



(a) Box plot of the latitude values of position t



(b) Box plot of the longitude values of position t

Figure 36: Box plot of the latitude and longitude values of position t found to be measured around the solar equinox

Table 3: Listing of the summary of the migration route

Time	Longitude					Latitude				
	Mean	sd	Median	2.5%	97.5%	Mean	sd	Median	2.5%	97.5%
16.08.2015 18:43	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23
02.09.2015 23:27	7.40	0.41	7.42	6.65	8.12	46.07	0.27	46.03	45.64	46.64
11.09.2015 05:06	6.96	1.83	6.93	3.60	10.52	44.31	1.01	44.20	42.61	46.59
11.09.2015 17:54	6.47	2.31	6.45	1.95	10.81	40.36	1.25	40.37	37.98	42.76
15.09.2015 11:15	5.88	1.29	5.94	3.29	8.38	35.54	1.09	35.79	32.91	36.92
19.09.2015 04:56	5.89	1.36	5.85	3.45	8.70	34.79	1.54	34.81	31.51	37.61
19.09.2015 17:58	3.26	1.38	3.26	0.55	5.90	31.99	1.62	31.98	28.41	35.16
21.09.2015 23:42	3.18	1.33	3.18	0.60	5.71	31.81	1.63	31.81	28.26	34.84
02.10.2015 11:32	2.99	0.92	2.99	1.04	4.81	29.92	1.61	29.78	27.28	33.67
11.10.2015 23:34	2.97	0.96	3.00	0.93	4.83	29.88	1.62	29.74	27.28	33.67
13.10.2015 23:35	2.07	1.74	2.08	-1.15	5.48	27.39	1.97	27.22	23.85	31.54
15.10.2015 11:42	-0.12	1.52	-0.15	-2.94	2.93	23.02	2.11	22.90	18.93	27.18
16.10.2015 05:36	-0.39	1.30	-0.49	-2.58	2.30	22.50	2.10	22.36	18.47	26.66
16.10.2015 17:33	-5.37	1.30	-5.33	-8.13	-2.90	19.27	1.79	19.28	15.91	22.87
17.10.2015 12:04	-7.26	0.92	-7.26	-9.04	-5.53	18.12	1.41	17.96	15.62	21.31
18.10.2015 12:19	-7.37	0.86	-7.39	-9.00	-5.80	18.04	1.38	17.88	15.54	21.13
27.10.2015 00:17	-7.53	0.73	-7.56	-8.95	-6.19	17.89	1.35	17.80	15.42	20.77
12.01.2016 12:42	-8.40	0.30	-8.39	-9.01	-7.82	15.86	0.95	15.85	14.07	17.77
22.03.2016 12:44	-9.22	1.31	-9.31	-11.51	-6.55	16.78	1.64	16.93	13.45	19.60
23.03.2016 13:00	-10.41	1.80	-10.33	-13.93	-7.06	18.21	2.13	18.35	13.55	21.99
29.03.2016 12:39	-12.65	1.14	-12.63	-14.91	-10.44	20.95	2.45	21.05	16.09	25.43
04.04.2016 06:16	-10.31	1.86	-10.26	-13.87	-6.78	23.36	2.76	23.33	18.23	28.68
04.04.2016 18:58	-8.18	2.31	-8.14	-12.59	-3.94	26.09	2.76	25.96	21.25	32.20
05.04.2016 12:29	-5.64	2.22	-5.66	-9.68	-1.30	28.59	2.50	28.44	24.08	33.63
06.04.2016 12:04	-3.15	1.99	-3.20	-6.82	0.96	31.02	1.71	30.97	27.90	34.54
07.04.2016 12:09	-2.77	1.68	-2.81	-5.92	0.38	31.75	1.32	31.67	29.42	34.46
08.04.2016 06:06	-2.58	0.73	-2.56	-4.14	-1.31	32.39	0.54	32.41	31.30	33.39
08.04.2016 18:38	-2.55	0.49	-2.49	-3.82	-1.64	32.39	0.40	32.39	31.55	33.19
12.04.2016 12:12	-2.55	0.48	-2.49	-3.82	-1.65	32.39	0.40	32.39	31.53	33.19
18.04.2016 05:10	-2.55	0.48	-2.49	-3.80	-1.65	32.39	0.40	32.39	31.53	33.20
20.04.2016 18:43	-1.73	1.45	-1.59	-4.65	0.69	32.52	1.32	32.54	30.12	34.85
21.04.2016 05:01	0.67	2.01	0.63	-3.08	4.46	32.97	2.19	32.98	28.85	37.17
21.04.2016 23:57	1.16	2.05	1.17	-2.85	5.06	32.76	2.11	32.97	28.51	36.14
22.04.2016 18:48	2.05	1.99	2.13	-1.96	5.62	33.04	2.48	32.94	28.46	37.84
23.04.2016 05:01	3.45	2.02	3.33	-0.27	7.44	32.71	2.20	32.82	28.50	36.81
23.04.2016 18:34	5.06	1.15	4.98	2.88	7.36	32.55	1.14	32.58	30.44	34.76
24.04.2016 04:36	5.45	0.64	5.59	3.94	6.30	32.59	0.74	32.36	31.47	34.00
24.04.2016 16:38	5.45	0.64	5.60	3.95	6.29	32.59	0.74	32.38	31.49	34.01
25.04.2016 23:42	5.46	0.64	5.61	3.99	6.36	32.61	0.75	32.40	31.50	34.03
27.04.2016 04:21	5.81	0.69	5.84	4.43	7.02	33.22	0.78	33.09	31.89	34.81
29.04.2016 23:34	5.94	0.74	5.89	4.59	7.43	33.78	0.81	33.70	32.40	35.54
02.05.2016 18:44	5.99	0.78	5.94	4.55	7.59	33.96	0.82	33.90	32.54	35.67
03.05.2016 04:26	6.89	1.23	6.89	4.43	9.20	36.62	0.88	36.57	34.99	38.41
04.05.2016 04:06	9.00	0.25	8.96	8.66	9.57	41.96	0.34	41.92	41.51	42.76
05.05.2016 04:06	8.83	1.98	8.71	5.43	12.66	42.25	1.49	42.25	39.44	44.94
05.05.2016 23:24	9.15	0.30	9.16	8.66	9.62	42.25	0.38	42.27	41.55	42.95
07.05.2016 18:43	6.19	0.28	6.21	5.69	6.62	43.03	0.03	43.02	43.00	43.12
09.05.2016 23:09	6.38	0.48	6.32	5.68	7.43	36.97	0.03	36.98	36.88	37.00
12.05.2016 18:43	7.36	0.49	7.34	6.39	8.33	36.98	0.02	36.99	36.94	37.00
16.05.2016 23:17	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23

Table 4: Listing of track positions including half wind speed data in to route inference

Time	Longitude					Latitude				
	Mean	sd	Median	2.5%	97.5%	Mean	sd	Median	2.5%	97.5%
16.08.2015 18:43	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23
02.09.2015 23:27	7.37	0.50	7.38	6.44	8.31	46.03	0.39	45.99	45.37	46.81
11.09.2015 05:06	6.74	2.10	6.57	3.08	10.93	44.81	1.47	44.62	42.26	47.97
11.09.2015 17:54	7.24	2.85	7.35	1.68	12.57	41.86	2.31	41.64	37.98	46.91
15.09.2015 11:15	5.91	1.30	5.89	3.39	8.46	34.47	1.85	34.87	30.15	36.89
19.09.2015 04:56	6.51	1.67	6.34	3.75	10.26	33.71	2.36	33.64	28.95	38.46
19.09.2015 17:58	3.30	1.63	3.28	0.31	6.46	31.38	2.62	31.32	26.25	36.34
21.09.2015 23:42	3.09	1.51	3.06	0.26	6.00	31.28	2.60	31.25	26.05	36.00
02.10.2015 11:32	2.84	0.88	2.86	1.16	4.57	29.56	3.07	29.38	23.83	34.79
11.10.2015 23:34	2.83	0.95	2.85	0.98	4.77	29.53	3.10	29.41	23.79	34.78
13.10.2015 23:35	2.13	2.13	2.09	-1.80	6.36	26.90	3.28	26.82	20.28	33.20
15.10.2015 11:42	1.06	2.19	1.01	-2.91	5.56	22.20	2.91	22.33	16.01	27.48
16.10.2015 05:36	0.96	2.05	0.79	-2.44	5.31	21.78	2.85	21.93	15.79	26.69
16.10.2015 17:33	-4.84	2.08	-4.75	-8.92	-0.72	18.96	2.20	18.94	14.73	23.11
17.10.2015 12:04	-7.08	0.94	-7.06	-8.98	-5.41	17.72	1.52	17.81	14.51	20.28
18.10.2015 12:19	-7.22	0.85	-7.15	-9.10	-5.70	17.63	1.47	17.75	14.57	20.00
27.10.2015 00:17	-7.44	0.68	-7.43	-8.86	-6.23	17.48	1.45	17.59	14.35	19.88
12.01.2016 12:42	-8.41	0.31	-8.41	-8.98	-7.81	15.16	1.00	15.18	13.15	17.09
22.03.2016 12:44	-8.87	1.44	-9.01	-11.22	-5.96	15.93	1.95	16.00	12.31	20.21
23.03.2016 13:00	-10.40	1.90	-10.42	-13.94	-6.61	17.53	2.72	17.49	12.34	22.68
29.03.2016 12:39	-13.01	1.13	-13.01	-15.21	-10.78	19.06	3.08	19.04	13.10	24.99
04.04.2016 06:16	-9.35	2.58	-9.54	-13.71	-3.75	21.78	3.67	21.72	14.88	28.61
04.04.2016 18:58	-8.59	2.92	-8.40	-14.53	-3.31	23.55	4.12	23.79	15.74	31.05
05.04.2016 12:29	-5.50	2.53	-5.54	-10.38	-0.85	25.38	3.84	25.46	17.49	32.51
06.04.2016 12:04	-1.68	2.08	-1.78	-5.49	2.66	27.04	2.36	27.06	22.30	31.69
07.04.2016 12:09	-1.31	1.63	-1.35	-4.36	1.99	27.34	1.61	27.33	24.16	30.32
08.04.2016 06:06	-0.70	0.92	-0.73	-2.37	1.05	27.40	0.98	27.43	25.22	29.34
08.04.2016 18:38	-0.60	0.83	-0.69	-1.90	0.90	27.40	0.90	27.48	25.50	29.01
12.04.2016 12:12	-0.59	0.83	-0.68	-1.87	0.89	27.41	0.90	27.50	25.51	29.04
18.04.2016 05:10	-0.59	0.83	-0.68	-1.90	0.89	27.41	0.90	27.52	25.50	29.04
20.04.2016 18:43	-0.46	1.97	-0.42	-4.29	3.13	28.70	1.96	28.76	24.92	32.43
21.04.2016 05:01	1.59	2.38	1.44	-2.44	6.84	31.86	2.47	31.91	26.93	36.69
21.04.2016 23:57	1.46	2.30	1.57	-3.27	5.64	32.97	2.31	33.04	28.48	37.41
22.04.2016 18:48	1.70	2.50	1.75	-3.21	6.31	34.77	2.73	34.68	29.79	40.13
23.04.2016 05:01	3.88	2.93	3.82	-1.92	9.54	35.96	2.67	35.77	31.25	41.65
23.04.2016 18:34	7.75	1.65	7.68	4.66	11.10	36.99	1.68	36.87	33.98	40.36
24.04.2016 04:36	8.83	0.91	8.73	7.38	10.87	36.55	0.98	36.55	34.71	38.18
24.04.2016 16:38	8.83	0.90	8.75	7.36	10.85	36.55	0.97	36.55	34.73	38.12
25.04.2016 23:42	8.30	0.76	8.32	6.81	9.72	36.43	0.51	36.59	35.16	36.98
27.04.2016 04:21	8.45	0.99	8.42	6.58	10.43	37.20	0.66	37.23	35.86	38.38
29.04.2016 23:34	9.04	0.31	9.01	8.47	9.60	39.60	0.49	39.48	39.02	40.73
02.05.2016 18:44	8.94	0.76	8.93	7.51	10.37	39.59	0.78	39.56	38.18	41.21
03.05.2016 23:24	9.04	0.34	9.04	8.27	9.61	40.94	1.14	40.76	39.10	42.79
05.05.2016 04:06	8.82	3.20	8.77	3.18	14.99	39.26	2.38	39.18	34.88	44.13
07.05.2016 04:06	7.01	1.19	7.05	4.65	9.25	35.82	0.92	36.05	33.66	36.96
09.05.2016 23:09	6.15	2.04	6.15	2.29	9.87	35.05	1.38	35.28	31.68	36.89
12.05.2016 18:43	9.03	0.34	9.02	8.30	9.61	40.86	1.04	40.73	39.11	42.66
15.05.2016 03:51	9.06	0.33	9.06	8.32	9.61	41.70	0.90	42.02	39.75	42.92
17.05.2016 19:03	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23

Table 5: Listing of track positions after including full wind data to the route inference

Time	Longitude					Latitude				
	Mean	sd	Median	2.5%	97.5%	Mean	sd	Median	2.5%	97.5%
16.08.2015 18:43	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23
02.09.2015 23:27	7.35	0.45	7.35	6.54	8.18	46.11	0.33	46.07	45.58	46.77
11.09.2015 05:06	6.01	3.35	5.79	-0.13	12.78	44.01	2.59	43.90	39.21	49.32
11.09.2015 17:54	6.47	3.58	6.66	-0.45	12.93	39.33	3.24	39.42	32.90	45.93
15.09.2015 11:15	6.12	1.39	6.13	3.54	8.87	33.93	2.49	34.21	28.56	39.16
19.09.2015 04:56	6.96	1.83	6.90	3.84	10.81	33.19	2.94	33.25	27.55	38.80
19.09.2015 17:58	3.09	1.96	3.19	-0.83	6.66	30.93	2.87	30.90	25.09	36.40
21.09.2015 23:42	2.84	1.68	2.92	-0.44	5.91	30.70	2.86	30.78	24.71	35.99
02.10.2015 11:32	2.49	0.74	2.49	1.08	3.92	29.17	2.86	29.03	23.50	35.23
11.10.2015 23:34	2.45	0.79	2.42	0.88	3.96	29.16	2.86	28.97	23.45	35.19
13.10.2015 23:35	2.16	2.24	2.22	-2.12	6.38	25.83	3.81	25.76	18.32	33.52
15.10.2015 11:42	1.15	2.24	1.23	-3.02	5.79	21.88	3.81	21.78	14.63	29.18
16.10.2015 05:36	1.13	2.12	1.06	-2.56	5.33	21.68	3.82	21.66	14.51	29.25
16.10.2015 17:33	-2.85	2.78	-2.99	-7.83	3.40	19.16	3.53	19.20	12.44	25.94
17.10.2015 12:04	-7.59	1.02	-7.61	-9.56	-5.56	15.99	1.98	16.11	11.85	19.40
18.10.2015 12:19	-7.78	0.81	-7.83	-9.29	-6.16	15.86	1.89	16.06	11.94	18.98
27.10.2015 00:17	-7.84	0.71	-7.86	-9.18	-6.41	15.81	1.85	16.01	12.01	18.78
12.01.2016 12:42	-8.42	0.32	-8.42	-9.05	-7.81	15.14	0.95	15.16	13.27	16.95
22.03.2016 12:44	-9.30	1.39	-9.43	-11.64	-6.33	15.31	1.89	15.44	11.46	18.55
23.03.2016 13:00	-10.68	1.73	-10.59	-14.23	-7.57	15.68	2.59	15.75	10.72	20.48
29.03.2016 12:39	-13.34	1.05	-13.36	-15.37	-11.27	16.46	3.32	16.53	10.08	22.91
04.04.2016 06:16	-8.74	3.30	-9.01	-14.40	-1.81	21.03	6.00	21.11	8.57	32.38
04.04.2016 18:58	-8.13	3.69	-7.88	-15.81	-1.63	26.87	6.90	26.78	13.50	40.36
05.04.2016 12:29	-4.47	2.58	-4.53	-9.25	0.47	28.56	4.89	28.65	18.77	38.87
06.04.2016 12:04	-2.12	2.45	-2.24	-6.46	3.00	29.26	3.89	29.24	21.37	37.34
07.04.2016 12:09	-1.96	2.46	-2.08	-6.52	3.19	29.59	3.52	29.64	22.57	36.02
08.04.2016 06:06	-1.62	1.29	-1.66	-4.04	0.86	31.21	2.84	31.34	25.90	35.63
08.04.2016 18:38	-1.56	1.22	-1.50	-3.72	0.74	31.20	2.77	31.37	25.93	35.32
12.04.2016 12:12	-1.55	1.22	-1.50	-3.67	0.79	31.20	2.76	31.37	25.83	35.30
18.04.2016 05:10	-0.76	1.62	-0.79	-3.73	2.37	31.12	2.65	31.25	26.06	35.60
20.04.2016 18:43	0.45	2.76	0.49	-5.29	5.67	33.13	3.21	33.18	27.05	39.76
21.04.2016 18:43	1.12	2.14	1.18	-3.13	5.31	33.64	2.72	33.70	28.31	39.73
22.04.2016 18:48	1.63	2.33	1.70	-2.91	5.83	34.53	2.66	34.55	29.23	39.78
23.04.2016 05:01	3.60	2.57	3.49	-1.08	8.70	35.06	2.22	35.01	30.81	39.65
23.04.2016 18:34	5.13	1.39	5.06	2.59	7.96	35.10	1.36	35.09	32.51	37.75
24.04.2016 04:36	5.62	0.56	5.66	4.48	6.55	35.00	0.89	34.84	33.71	36.58
24.04.2016 16:38	5.62	0.56	5.67	4.51	6.54	35.01	0.89	34.86	33.73	36.60
25.04.2016 23:42	5.61	0.56	5.65	4.45	6.58	35.01	0.90	34.84	33.69	36.62
27.04.2016 04:21	5.96	0.97	5.91	4.18	7.81	35.25	1.00	35.25	33.34	37.25
29.04.2016 23:34	5.93	1.06	5.93	3.95	8.06	35.23	1.00	35.31	33.20	36.87
02.05.2016 18:44	5.91	1.46	5.93	2.95	8.70	35.41	1.29	35.46	32.84	37.87
03.05.2016 04:26	6.40	3.50	6.35	-0.17	13.76	38.56	2.87	38.81	32.14	43.68
04.05.2016 04:06	7.45	3.21	7.47	1.04	13.28	42.69	3.79	43.86	34.94	48.25
05.05.2016 04:06	6.68	3.78	6.97	-0.26	13.51	43.10	2.94	43.26	37.67	48.46
05.05.2016 23:24	6.70	3.82	8.31	-0.32	12.97	43.23	2.57	43.52	39.18	47.98
07.05.2016 18:43	6.31	0.29	6.35	5.80	6.73	43.10	0.10	43.07	43.00	43.38
09.05.2016 23:09	6.65	0.31	6.66	5.98	7.23	36.70	0.19	36.73	36.24	36.98
15.05.2016 03:51	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23

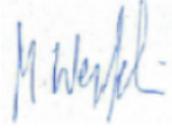
Table 6: Listing of track positions after including normal distributed wind data to the route inference

Time	Longitude					Latitude				
	Mean	sd	Median	2.50%	97.50%	Mean2	sd2	Median2	2.50%2	97.50%2
16.08.2015 18:43	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23
02.09.2015 23:27	7.31	0.43	7.29	6.58	8.10	46.07	0.31	46.01	45.61	46.72
11.09.2015 05:06	6.88	2.03	6.80	3.28	10.75	44.65	1.15	44.53	42.81	47.17
11.09.2015 17:54	7.05	2.37	7.20	1.84	11.21	40.89	1.37	40.90	38.19	43.50
15.09.2015 11:15	5.88	1.26	5.89	3.51	8.39	35.33	1.22	35.60	32.63	36.93
19.09.2015 04:56	6.25	1.46	6.16	3.71	9.23	34.28	1.97	34.06	30.77	38.59
19.09.2015 17:58	3.36	1.34	3.41	0.70	5.93	31.32	2.55	30.79	27.29	36.22
21.09.2015 23:42	3.24	1.28	3.30	0.72	5.61	31.14	2.58	30.52	27.11	35.98
02.10.2015 11:32	3.10	0.78	3.06	1.70	4.57	28.47	3.32	27.26	23.86	35.33
11.10.2015 23:34	3.07	0.84	3.04	1.50	4.73	28.40	3.33	27.20	23.77	35.25
13.10.2015 23:35	1.62	1.76	1.61	-1.82	5.01	25.65	2.88	25.08	21.26	31.85
15.10.2015 11:42	-0.34	1.51	-0.32	-3.16	2.50	20.78	2.51	20.48	16.79	26.75
16.10.2015 05:36	-0.62	1.21	-0.66	-2.81	1.75	20.11	2.43	19.70	16.12	25.79
16.10.2015 17:33	-5.22	1.29	-5.14	-8.07	-2.92	15.70	2.22	15.26	12.20	20.81
17.10.2015 12:04	-6.81	0.79	-6.86	-8.22	-5.20	13.92	1.91	13.66	11.12	18.10
18.10.2015 12:19	-6.95	0.75	-6.99	-8.24	-5.48	13.82	1.89	13.62	11.04	18.01
27.10.2015 00:17	-7.22	0.68	-7.25	-8.49	-5.94	13.71	1.90	13.51	10.77	17.81
12.01.2016 12:42	-8.40	0.31	-8.40	-9.01	-7.80	13.45	0.98	13.46	11.54	15.36
22.03.2016 12:44	-9.28	1.35	-9.43	-11.43	-6.44	14.71	1.42	14.69	11.95	17.44
23.03.2016 13:00	-10.64	1.78	-10.61	-14.01	-7.35	16.70	2.07	16.67	12.68	20.65
29.03.2016 12:39	-13.04	1.06	-13.01	-15.16	-10.97	19.98	2.39	19.97	15.27	24.58
04.04.2016 06:16	-9.98	1.84	-9.95	-13.59	-6.52	22.76	2.96	22.82	16.99	28.65
04.04.2016 18:58	-8.23	2.29	-8.10	-13.00	-4.26	25.68	3.18	25.79	19.46	31.69
05.04.2016 12:29	-5.17	2.25	-5.12	-9.64	-1.01	28.50	2.71	28.35	23.43	33.78
06.04.2016 12:04	-1.64	1.95	-1.69	-5.16	2.30	30.74	1.96	30.76	27.09	34.56
07.04.2016 12:09	-0.78	1.62	-0.88	-3.85	2.61	31.27	1.52	31.21	28.54	34.27
08.04.2016 06:06	0.10	1.18	0.39	-2.08	1.99	31.61	0.70	31.60	30.24	32.95
08.04.2016 18:38	0.26	1.11	0.62	-1.48	1.73	31.64	0.53	31.68	30.55	32.70
12.04.2016 12:12	0.27	1.11	0.64	-1.48	1.73	31.64	0.52	31.66	30.56	32.64
18.04.2016 05:10	0.27	1.11	0.64	-1.53	1.78	31.64	0.52	31.67	30.56	32.64
20.04.2016 18:43	0.44	1.90	0.49	-3.47	3.82	31.98	1.62	32.09	28.78	34.66
21.04.2016 05:01	1.97	2.33	1.92	-2.56	6.56	33.15	2.56	33.35	27.64	37.67
21.04.2016 23:57	1.57	2.09	1.60	-2.79	5.55	33.53	2.15	33.94	28.70	36.55
22.04.2016 18:48	1.83	2.16	1.97	-2.50	5.83	35.16	2.44	35.24	30.43	39.83
23.04.2016 05:01	3.65	2.20	3.55	-0.40	8.16	35.93	2.25	35.82	31.82	40.65
23.04.2016 18:34	6.11	1.43	6.07	3.49	8.89	36.73	1.37	36.70	34.28	39.28
24.04.2016 04:36	6.43	0.66	6.31	5.34	7.76	36.61	0.67	36.67	35.22	37.91
24.04.2016 16:38	6.44	0.66	6.32	5.37	7.75	36.61	0.67	36.67	35.25	37.89
25.04.2016 23:42	6.90	0.68	6.83	5.77	8.38	36.61	0.32	36.70	35.86	36.99
27.04.2016 04:21	7.41	0.79	7.33	6.05	9.00	37.31	0.40	37.33	36.53	38.04
29.04.2016 23:34	9.01	0.26	8.96	8.66	9.56	39.35	0.28	39.28	39.01	40.07
02.05.2016 18:44	8.99	0.56	8.98	7.98	10.06	39.32	0.45	39.31	38.49	40.24
03.05.2016 23:24	8.98	0.36	8.98	8.22	9.59	40.79	1.03	40.69	39.10	42.61
05.05.2016 04:06	9.10	3.29	9.02	3.16	15.22	40.57	2.85	40.53	35.48	45.56
07.05.2016 04:06	9.01	0.36	9.00	8.24	9.62	40.10	0.59	40.15	39.08	40.97
09.05.2016 23:09	9.03	0.33	9.02	8.29	9.60	40.77	1.09	40.64	39.11	42.69
12.05.2016 18:43	6.03	1.99	5.97	2.60	10.71	44.26	0.80	44.17	43.02	45.99
15.05.2016 03:51	6.07	1.96	6.01	2.70	10.68	44.42	0.81	44.33	43.06	46.19
17.05.2016 19:03	7.37	0.00	7.37	7.37	7.37	46.23	0.00	46.23	46.23	46.23

Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged.

Signature:



Date: Zurich, 29.09.2020
