

Light-level geolocation in polar regions with 24-hour daylight

Simeon Lisovski

Swiss Ornithological Institute, Department of Bird Migration, Seerose 1, CH-6204 Sempach, Switzerland
Simeon.Lisovski@gmail.com

Lisovski, S. 2018. Light-level geolocation in polar regions with 24-hour daylight. *Wader Study* 125(2): xx–xx.

Solar geolocation has become one of the most frequently used tools in wader migration research. Geolocators provide location estimates based on recorded light intensities, and more specifically, on the changes in light during the twilight periods, allowing an increase in knowledge on how waders migrate across the globe. Yet, quite a number of species breed in polar regions where they experience 24-hour daylight. Given the lack of recorded twilight times on the geolocators during this period, analysis methods have been unable to resolve birds' positions when they are in constant daylight. This is especially problematic in the older geocator generations that could record light only over a narrow range of light intensities. Some newer geolocators record the full light range, allowing even small changes in light intensity during the day (under bright light conditions) to be detected. However, the common methods for estimating locations are not designed for changes in high light regimes that lack dark periods during the night. Previously, I developed and implemented a method for analysing continuous light records, which evaluates the likelihood of a measured light cycle being from a given location, leading to first estimates of breeding sites in Sanderlings *Calidris alba* and Great Knots *C. tenuirostris*. However, the final decision on the breeding site was somewhat subjective and a formal description of the method was still lacking. Here, I describe a new development for estimating high-latitude positions, which is implemented in a freely available R Package called *PolarGeolocation*.

Keywords

animal tracking
 solar geolocation
 Arctic breeding sites
 template-fit
 shorebirds
 animal movements

INTRODUCTION

Solar geolocation data loggers are simple tracking devices that record ambient light levels for the purpose of estimating locations and thereby reconstructing animal movement trajectories. Despite the disadvantages of these archival tags over GPS and other satellite-linked tracking devices, which mainly concern their low spatial and temporal precision, geolocators have been proven to be extremely useful tools. Their low weight (<0.5 g) and relatively low cost broadens the range of species that can be tagged (Bridge *et al.* 2011) as well as enabling larger sample sizes to be gathered.

In recent years, multiple analytical tools have been developed to estimate locations from light recordings. The most common approach is the 'threshold method'. This method requires the definition of each twilight event (i.e. sunrise or sunset event) in a dataset as the time point corresponding to the moment when solar irradiance reaches some arbitrary, but constant, threshold level (Hill 1994, Ekstrom 2004). Latitude is then estimated by the duration of time between consequent pairs of twilights and the longitude by the time of solar noon or midnight.

While this approach is simple, it is plagued by many well-known problems such as high error rates near an equinox, generally biased estimates, unrealistic assumptions of constant shading, and a null assumption of no movement (Lisovski *et al.* 2012). 'Template-fit methods' were therefore developed to overcome some of these limitations. The advantage of this approach is that, rather than using just a single value per transition, the rate of change of solar elevation (and therefore light intensity) over time is analysed (Musyl *et al.* 2001, Ekstrom 2007). Template-fit methods have been shown to be relatively robust in dealing with the effects of shading, and location estimates are less affected by individuals moving between consecutive twilight periods (Ekstrom 2007).

The next step to improve precision was to develop models that allow the incorporation of constraints. These might include, for instance, movement models that define the range of movement speeds, or spatial masks that restrict location estimates to land or ocean. One of the most important features of these (mainly) probabilistic models (e.g. MCMC simulations in R package *SGAT*; Sumner *et al.* 2009), Particle filters (e.g. in R package *FLightR*;

Rakhimberdiev *et al.* 2015), and Kalman filters (e.g. in R package *Trackit*; Nielsen & Sibert 2007), is that they come with a quantification of the error associated with each location estimate.

Despite these methodological improvements, some migrants (notably some of the very interesting long-distance migratory waders) are still able to escape scrutiny by flying into their Arctic breeding sites where they experience 24-hour daylight. Currently, established tools cannot provide any estimates of their whereabouts during these periods that lack sunrise and sunset times. Previously, I developed and implemented a method for analysing continuous light records, which evaluates the likelihood of a measured light cycle being from a given location, leading to first estimates of breeding sites in Sanderlings *Calidris alba* (Lisovski *et al.* 2016a) and Great Knots *C. tenuirostris* (Lisovski *et al.* 2016b). In this paper I describe 'the problem' of estimating locations in these regions in more detail and develop an alternative method that allows us to estimate location, even when the sun does not fall below the horizon.

THE PROBLEM

Many geolocator devices, and notably the older generations (e.g. British Antarctic Survey Mk Loggers), resolve ambient light intensities of a specific range only. For good reasons, this range is set to resolve light during the very early sunrise and very late sunset when the rate of change in light is fastest. As a result, the loggers record and save the same maximum light levels during most of the day, starting approximately from when the sun is a few degrees above the horizon during the ascent, until the sun is again a few degrees above the horizon during the descent. However, the number of 24-hour daylight days a logger records at high latitudes also depends on the sensitivity of the logger, and loggers may even record maximum light for 24 hours if the sun goes below or touches the horizon. Thus, the problem exists even below the Arctic Circle (66°33'47.2"). Consequently, if placed in the Arctic (or Antarctic) these loggers record the same value over the entire summer period when the sun stays above the horizon altogether. However, some of the more recent devices (e.g. Intigeo Loggers from Migrate Technology Ltd.) can record absolute (rather than relative) light intensity over virtually the entire light range, thus making it possible to detect changes in light even when the sun remains above the horizon (Fig. 1). Analysis of the variation of light intensity even in 24-hour daylight should even be possible with existing methods, but only if the thresholds used are set very high (e.g. 8 lux for the data shown in Fig. 1c). Theoretically, curve-methods that makes use of the rate of change over time to estimate location (Sumner *et al.* 2009) should also be able to estimate locations if the logger resolves small changes during 24-hour daylight periods. However, in practice these estimates are associated with such large error rates that they become spatially uninformative. This is because the 'shallowness' of the slopes of the changes in light around the twilight periods at high latitudes means that

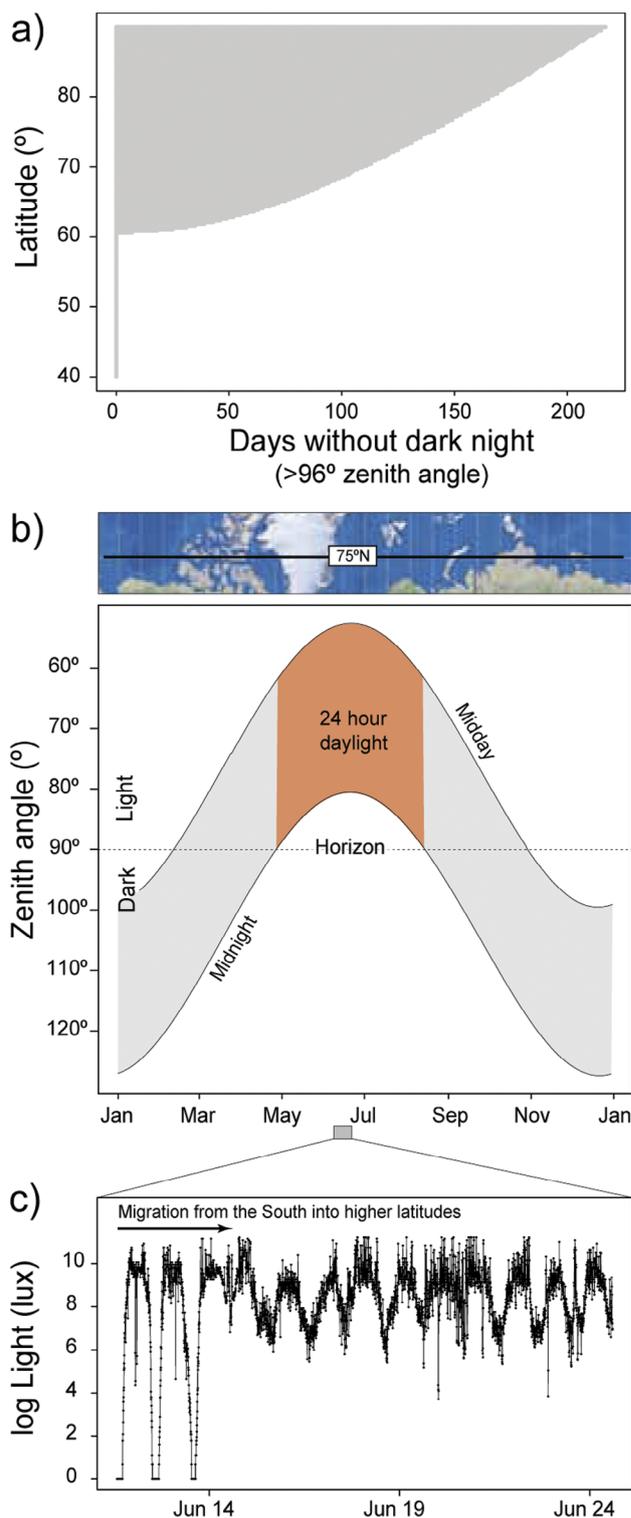


Fig. 1. (a) The number of days without clear twilight across latitudes. (b) The range of sun elevation angles at 75°N. (c) An example of light recordings before and shortly after a Sanderling crosses the Arctic Circle and reaches its breeding ground. The range of sun elevation angles indicate the annual period during which the sun does not fall below the horizon, resulting in 24-hour daylight. The geolocator data were collected on a Sanderling by a recently developed tag from Migrate Technology Ltd. that is able to resolve changes in light even in bright light conditions. Such changes can be used to get an estimate of locations using the method described in this paper.

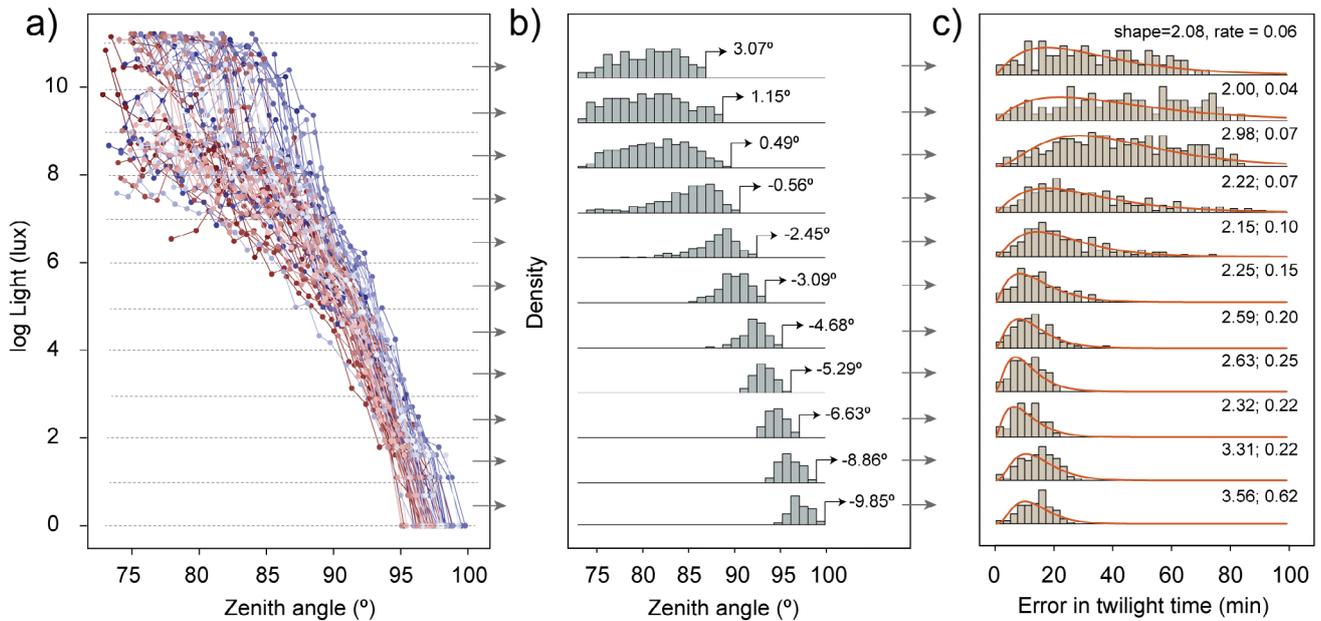


Fig. 2. (a) Light intensity recordings over zenith angles from a Great Knot at a known location (Roebuck Bay, Australia). Red lines indicate sunrise and blue lines indicate sunset. (b) The error distribution for several light intensity groups. (c) The translated error distribution in minutes with a fitted gamma density distribution model (orange lines).

any impacts of shading (from clouds, feathers, the immediate environment, etc.) result in larger errors than at lower latitudes where the natural variation in light levels is so much greater.

These issues are illustrated in Fig. 2, which shows a series of recorded sunrise and sunset times across sun elevation angles from a geolocator on a Sanderling at a known location in Southern Australia (Lisovski *et al.* 2016a). Each series of points in Fig. 2a represents a measured set of light values relative to sun elevation (which is known from the time and site), and the variation between the curves shows the degree of shading experienced between days. This variation in sun angle for a given light intensity threshold can be summarised (Fig. 2b, analysed in 1-lux increments) and translated into time units (actual minutes of ‘error’; Fig. 2c). Fig. 2c clearly illustrates the issue facing us, which is that errors at high light intensities are much larger than those at low light intensities. Given that high-latitude geolocators will only record quite high light intensities as shown in Figs. 1c and 4b (approx. 6–11 lux for Intigeo W65 loggers from Migrate Technology), attempts to analyse such data with existing threshold tools will result in highly inaccurate estimates of position.

A POTENTIAL SOLUTION

A potential solution for this problem is to apply a modified ‘template-fit’ model that is based on two principles: (1) there is a maximum light value for a given zenith angle (the angle between the zenith and the centre of the Sun’s disc), and (2) the measured light will follow a certain error distribution. In simple words, the changing position of the sun in the sky (caused by the Earth’s rotation),

even during 24-hour daylight, causes changes in the light intensity that can be measured by light loggers. In a perfect world, this measured light would be solely a function of the sun’s location (e.g. the zenith angle). However, when shading occurs, e.g. from clouds, the geolocator would measure a slightly reduced light intensity. Conversely, it should be impossible to measure more light than is expected for a given zenith angle (apart from lightning and anthropogenic light). By using light intensity recordings from a known location (calibration data), we can define the expected maximum light intensity over the range of experienced zenith angles (Fig. 3, left panel). In most cases this maximum light ‘template’ is zero during the night, has an increasing or decreasing slope during the twilight periods (i.e. sunrise and sunset) and reaches a maximum at a certain angle (i.e. daytime). In Fig. 3 (left panel) for instance, the maximum is reached at a zenith angle of approximately 87° (that is, 3° above the horizon).

Such a template can be used to predict the maximum light the logger could record at any given location and time (Fig. 3, orange line in middle panels). By comparing the geolocator data from a subset of measurements during stationary periods or even a single day to the expected light curves, we can identify the potential area in which the logger could have recorded these light levels (locations at which the observed measurements fall below the maximum expected light; Fig. 3).

The next step is to refine the probability of locations that are generally applicable given (1) the light measurements and (2) the expected maximum light template (e.g. the white area in Fig. 3, right panel). My proposal is to analyse the ‘errors’ from the expected light measurements – the deviation of light from the maximum light curve.

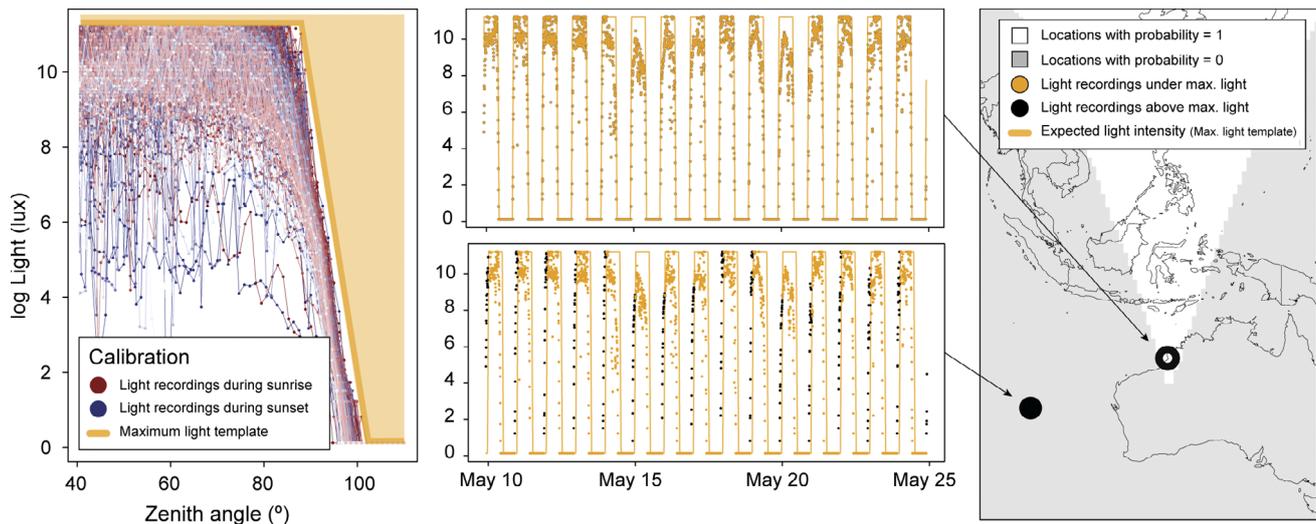


Fig. 3. The principle behind a maximum light template is to identify the range of locations that are possible for a given geolocator time series. The left panel shows light recordings from a Great Knot at a known location (Roebuck Bay, Australia; open black circle in right panel) and the fitted maximum light template (orange thick line). The two middle panels show the same geolocator recordings with the template-fit evaluation for two different locations (see arrows in right panel). The geolocator recordings are from a period in which this individual was known to be resident at the deployment site. The orange line indicates the extrapolated maximum light curve (using the maximum light template – middle panels) for the two locations. In the upper middle panel, all light recordings fall below the maximum light curve indicating that the logger could have recorded the light at this location (in fact it was exactly this location). In the lower middle panel, some measurements (black dots) exceed the maximum light curve; thus, the light could have not been recorded there. The right panel shows the result of this evaluation for all locations on the map.

These are expected to follow a certain distribution (e.g. a gamma or log-normal density distribution). Given the rather flat distribution in bright light across zenith angles (Fig. 2b), analysing the error of light measurements across certain ranges of zenith angles (e.g. 80–81°) should help refine the possibilities. Using a sequence of ranges of zenith angles allows us to fit a density distribution to each range separately and to get optimal parameters for these ranges of zenith angles. In most cases, a log-normal or gamma distribution will fit the dataset. We can then compare the recorded light intensities for each range and estimate the log-likelihood values using the defined distribution, e.g. the distribution parameters for each range of zenith angles. Finally, we can sum the log-likelihood values calculated for each zenith angle range, providing an estimate of how likely it was that the light has been recorded at the given location.

IMPLEMENTATION

To make the above described process accessible, I compiled a package called *PolarGeolocation* that is freely available from GitHub (Lisovski 2018) and can be installed directly via R (R Core Team 2014). The package contains four major functions and a tutorial that explains the workflow in more detail. A usual analysis, after loading the data and defining the twilight times (e.g. using the R Package *TwGeos*; Lisovski *et al.* 2016c), starts with the calibration. The function `getTemplateEstimate()` uses the raw light recorded at a known location for calibration, as

well as the defined twilight times, to calculate the maximum light template in addition to the parameters for defining the log-normal distribution of the groups of zenith angles (Fig. 4a). To restrict the location estimates to a sensible spatial range, a ‘mask’ is required and can be computed using the function `getMask()`. This mask also allows us to distinguish between e.g. ocean and land, and therefore assign different probabilities to these groups. In some cases, these probabilities can be set to zero to make certain locations, such as the ocean, impossible. Since spatial distortion in the most frequently used map projection (e.g. WGS84) is most pronounced in polar regions (i.e. less of the Earth’s surface falls into an e.g. 1 x 1° grid cell with increasing latitudes), the function uses an equal area projection. This particular projection allows us to specify the centre, radius and resolution of the desired spatial extent (Fig. 4c). The function `templateEstimate()` then takes the mask, the calibration, and the raw light recordings (e.g. during the breeding season of Arctic breeding waders) to provide daily estimates of the joint likelihood (measurements across a 24-hour period) for each location (i.e. grid cell), as well as the sum of measurements exceeding the maximum light template. Finally, the function `templateSummary()` provides a summary of the likelihood estimates, and can also plot a map of the spatial likelihood surface of a bird being in a given location. Thus, it estimates the most likely location of a bird, while also providing an error estimate for both longitude and latitude (Fig. 4d).

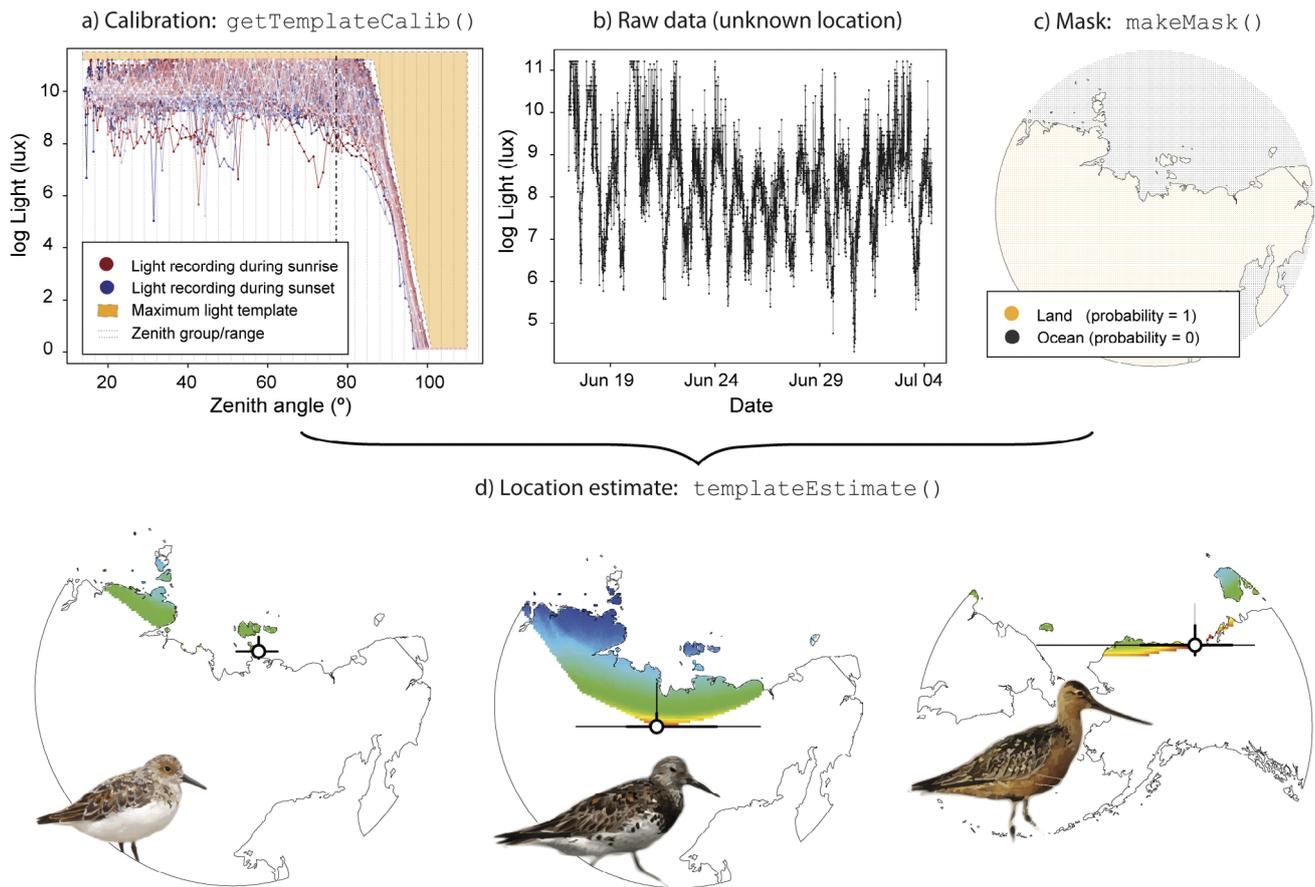


Fig. 4. The four steps to estimate locations in polar regions. **(a)** Calibration is the first crucial step to derive the maximum light template (orange) and the error distribution of light recordings across different zenith angles. **(b)** The raw data for which a location should be analysed and for which the bird has been stationary need to be defined. **(c)** A mask is provided to define the spatial extent and potential areas that are impossible, e.g. ocean. **(d)** All three sources of information (a, b, c) will be used to estimate the relative probability of locations at which the bird might have been during this period of time. The relative likelihood surface (colour scale from lowest (blue) to highest (red) likelihood; this differs among individuals), the most likely location (circle) and a measure of uncertainty (95% thin line, 99% thick line) is shown for three example datasets of a Sanderling, a Great Knot and a Bar-tailed Godwit (illustrations modified with allowance from authors: ADJ82 licensed under CC BY 4.0, Ken Gosbell, U.S. Fish & Wildlife Service – Steve Maslowski).

DISCUSSION

To date, estimating the locations of Arctic breeding waders using geolocators has been problematic due to a lack of clear sunrise and sunset periods. Here, I propose a novel method of estimating locations experiencing 24-hour sunlight. The method I describe makes use of recent tag developments that allow us to record a large spectrum of light intensities. Using the same theory as geolocation by light, we can now get sensible estimates of locations within polar regions. Based on my own experience, the method results in reasonable and useful estimates for species breeding as far north as 76°N (e.g. Sanderling and Ruddy Turnstone *Arenaria interpres* on the New Siberian Islands). While the method can be applied to geolocator data in general, resulting in good estimates of longer staging/stopover sites, it is specifically developed to identify stationary locations in high latitudes. Other methods like *GeoLight* (Lisovski & Hahn 2012), *SGAT* (Wotherspoon *et al.* 2013) and *FLightR* (Rakhimberdiev

et al. 2017) are more efficient and better suited to estimate migratory tracks under ‘normal’ conditions (e.g. at locations with dark night-time).

As with all other methods in geolocation, calibration is crucial and the results depend on the quality of the calibration dataset. The proposed method relies on a good representation of maximum light values across a range of zenith angles, e.g. light recordings measured during optimal conditions. This can be achieved either by exposing the geolocator to natural light during a period of perfect weather, or by having a long period of post-deployment and/or pre-recapture measurements at a known location. The latter is preferred, since it provides the second important information: the error distribution of the light given different weather conditions and behaviour of the individual. One might argue that, at least for the maximum light curve, a standardized curve for a certain logger type could be sufficient. However, personal experience shows that each logger is slightly (and sometimes significantly)

different, and that differences can sometimes be attributed to the amount of transparent coating on the light sensor. Thus, it is always good practice and will pay off to have extended calibration periods when the tag is on the bird.

One important assumption of the method relates to the stationary behaviour of the bird during the period for which the location estimate takes place. The method relies on light intensity recordings over a couple of days to estimate a sensible location probability distribution. Using single days may result in highly inaccurate estimates. However, the data can be split into a series of known stationary periods and location estimates generated for each period separately. If unsure whether the individual was stationary during the breeding season, location estimates could be made for different periods, or the first and last couple of days could be excluded. Again, keep in mind that the method is most reliable if the estimate is based on a long time series. The period needed to get reliable results is, however, dependent on the quality of the recordings and if there is little shading a couple of days might be enough. Importantly for waders, if the logger is deployed on the leg of the bird, repeated periods of shading will occur during egg incubation and chick brooding as the light sensor is regularly shaded by the bird repeatedly getting up and sitting down. The shading during that period will be very different from during the calibration period and therefore light measurements during incubation and brooding should be excluded from analyses. The result shown in Fig. 4 are all based on data recorded over at least four days; in most cases, this seems to be a sufficient time period to assure reliable results.

In general, the method performs well on different wader species breeding in the Arctic and while the error is often larger than for location estimates in lower latitudes, the method provides important new information that would otherwise be impossible to collect. Implementation of this method promises to greatly improve our knowledge of where geolocator-tagged waders go in the Arctic.

ACKNOWLEDGEMENTS

I thank the Victorian and the Australasian Wader Study Group (VWSG, AWSG), especially Clive Minton, Ken Gosbell, Chris Hassel and Marcel Klaassen for continuing cooperation and sharing of data and ideas. The VWSG, AWSG and Phil Battley gave permission to use data from a Sanderling, a Great Knot and a Bar-tailed Godwit, respectively. I appreciate fruitful discussions on geolocation concepts with Eldar Rakhimberdiev, who also provided helpful comments on this manuscript, as did Ron Porter, Phil Battley and Kiran Dhanjal-Adams.

REFERENCES

- Bridge, E.S., K. Thorup, M.S. Bowlin, P.B. Chilson, R.H. Diehl, R.W. Fleron, P. Hartl, R. Kays, J.F. Kelly, W.D. Robinson & M. Wikelski. 2011. Technology on the move: recent and forthcoming innovations for tracking migratory birds. *Bioscience* 61: 689–698.
- Ekstrom, P.A. 2004. An advance in geolocation by light. *Memoirs of the National Institute of Polar Research, Special Issue* 58: 210–226.
- Ekstrom, P. 2007. Error measures for template-fit geolocation based on light. *Deep Sea Research Part II: Topical Studies in Oceanography* 54: 392–403.
- Hill, R.D. 1994. Theory of geolocation by light levels. In: *Elephant seals: population ecology, behaviour and physiology* (B.J. Le Boeuf & R.M. Laws, Eds.). University of California Press, Berkeley, CA, USA.
- Lisovski, S. 2018. R Package *PolarGeolocation*: Light-level geolocation analysis tool for data recorded in high latitudes. Accessed at: <https://github.com/SLisovski/PolarGeolocation>
- Lisovski, S. & S. Hahn. 2012. *GeoLight*—processing and analysing light-based geolocator data in R. *Methods in Ecology & Evolution* 3: 1055–1059.
- Lisovski, S., C.M. Hewson, R.H.G. Klaassen, F. Korner-Nievergelt, M.W. Kristensen & S. Hahn. 2012. Geolocation by light: accuracy and precision affected by environmental factors. *Methods in Ecology & Evolution* 3: 603–612.
- Lisovski, S., K. Gosbell, M. Christie, B.J. Hoyer, M. Klaassen, I.D. Stewart, A.J. Taysom & C. Minton. 2016a. Movement patterns of Sanderling (*Calidris alba*) in the East Asian–Australasian Flyway and a comparison of methods for identification of crucial areas for conservation. *Emu* 116: 168–177.
- Lisovski, S., K. Gosbell, C. Hassell & C. Minton. 2016b. Tracking the full annual-cycle of the Great Knot *Calidris tenuirostris*, a long-distance migratory shorebird of the East Asian–Australasian Flyway. *Wader Study* 123: 177–189.
- Lisovski, S., S.J. Wotherspoon & M.D. Sumner. 2016c. R Package *TwGeos*: Solar/satellite geolocation for animal tracking. Accessed at: <https://github.com/SLisovski/TwGeos>
- Musyl, M.K., R.W. Brill, D.S. Curran, J.S. Gunn, J.R. Hartog, R.D. Hill, D.W. Welch, J.P. Eveson, C.H. Boggs & R.E. Brainard. 2001. Ability of archival tags to provide estimates of geographical position based on light intensity. Pp. 343–367 in: *Electronic Tagging and Tracking in Marine Fisheries*. Kluwer, Dordrecht, The Netherlands.
- Nielsen, A. & J.R. Sibert. 2007. State-space model for light-based tracking of marine animals. *Canadian Journal of Fisheries & Aquatic Sciences* 64: 1055–1068.
- R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Accessed at: <http://www.R-project.org/>
- Rakhimberdiev, E., D.W. Winkler, E. Bridge, N.E. Seavy, D. Sheldon, T. Piersma & A. Saveliev. 2015. A hidden Markov model for reconstructing animal paths from solar geolocation loggers using templates for light intensity. *Movement Ecology* 3: 1–15.
- Rakhimberdiev, E., A. Saveliev, T. Piersma & J. Karagicheva. 2017. *FLightR*: an R package for reconstructing animal paths from solar geolocation loggers. *Methods in Ecology & Evolution* 8: 1482–1487.
- Sumner, M.D., S.J. Wotherspoon & M.A. Hindell. 2009. Bayesian estimation of animal movement from archival and satellite tags. *PLoS ONE* 410: e7324.
- Wotherspoon, S.J., M.D. Sumner & S. Lisovski. 2013. R Package *SGAT*: Solar/satellite geolocation for animal tracking. Accessed at: <https://github.com/SWotherspoon/SGAT>.