State Space Model for Light Based Tracking of Marine Animals: Validation on Swimming and Diving Creatures

Anders Nielsen, John R. Sibert, Suzanne Kohin and Michael K. Musyl

Abstract Determining light-based geolocations of electronically tagged marine animals by using irradiance measurements in an integrated random walk movement model is presented for improving the precision of the reconstructed geographical tracks. The model avoids making any irradiance threshold assumptions, or constraining the movement of the tag between dawn and dusk, which has limited the efficacy of previous models. Unlike previous models which relied on using previously calculated raw geolocations from outside sources, the new model uses irradiance data to produce the most probable track. The model generates two estimates of geographic positions per day (at dawn and dusk). Prior to this study, the model has been successfully evaluated for tags mounted on moorings, drifters, and in simulated cases. This paper documents the application of the model to wild caught swimming and diving mako sharks and blue marlin. The reconstructed tracks via the new model are compared to tracks adjusted using sea-surface temperature and to very accurate satellite-based tracks. The model performs well for the tested tags, which is surprising as these are pop-off satellite archival tags (PSATs), from which only a small fraction of the light record around approximately dusk and dawn is stored. The integrated model was vastly superior to classical purely light-based methods in all cases.

Keywords Light based geolocation · Geolocation errors · Unscented Kalman filter · Archival tags · pop-up satellite archival tags

Introduction

Electronic data storage tags (archival or data storage tags and pop-up satellite archival tags (PSATs)) routinely used for tracking marine animal's behaviors are designed to record measurements of the animal's ambient environment, including

A. Nielsen (✉)
Pelagic Fisheries Research Program, Joint Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa, 1000 Pope Road, MSB 312, Honolulu, HI 96822, USA
e-mail: anders.nielsen@hawaii.edu

J.L. Nielsen et al. (eds.), *Tagging and Tracking of Marine Animals with Electronic Devices*, Reviews: Methods and Technologies in Fish Biology and Fisheries 9,
DOI 10.1007/978-1-4020-9640-2_18,

temperature, pressure, and irradiance. In theory, geographic positions (geolocations) can be estimated from irradiance data although sometimes the error estimates can be substantial (e.g. Musyl et al., 2001). More accurate tracks of animal's movements for terrestrial animals, and marine animals that stay at the surface for extended periods, can be obtained using radio-linked satellite tracking methods. However, many marine species, for example tuna, live almost exclusive below the surface where radio waves cannot penetrate, hence radio-linked satellite tracking is impossible. The movements of these species must be estimated using archived measurements of light, temperature, and pressure (depth). It is possible with continued development of sensors, that parameters such as geomagnetism, ocean colour, temperature-at-depth, salinity, etc. could also be used to calculate a position on the globe.

Fisheries research organizations spend large amounts of funding and effort on deployment of electronic tags. Despite the resources being devoted, the technology is not fully developed. This is especially true for the estimation of geographical postion from irradiance measurements. However, for some marine species there really is no alternative other than using data about the ambient environment to estimate where the animal has travelled. Directly observing these animals in situ is only possible for very short periods (e.g. acoustic tracking), and impractical on the high seas. Electronic tags can stay on animals for years, and potentially provide valuable information about post release survival, spawning areas, habitat preferences, feeding bouts, and migration corridors (Arnold and Dewar, 2001).

Estimating geographic positions from a time series of depth corrected irradiance measurements is a challenging task, and many approaches have been suggested (see for instance Smith and Goodman, 1986; Musyl et al., 2001; Ekstrom 2004). The challenge is partly because the irradiance data provide weak information about the position, especially concerning latitudes around equinoxes, and partly because of the highly correlated noise structure originating from outside sources, for instance local weather, light attenuation with depth, water clarity, clock errors, and resolution of light sensor. Previous approaches have ignored the high correlations and either reduced data to threshold crossing times, or fitted data to predefined template functions via least squares in order to derive location estimates. This has caused problems, as tagged marine animals have mistakenly been placed hundreds of kilometers from their actual location (e.g. Welch and Eveson, 1999; Musyl et al., 2001; Sibert et al., 2003). Close to equinoxes these algorithms frequently fail completely to estimate the latitude coordinate, and produce geolocations that differ by up to 40° of latitude from the true locations (Musyl et al., 2001).

To get an accurate representation of tracks from these very uncertain raw (i.e. not post-processed or filtered) geolocations, several post-processing schemes, some of which involve including temperature to improve location estimates, have been suggested (e.g. Sibert et al., 2003; Teo et al., 2004; Nielsen et al. 2006). These have been fairly successful, and are now widely used (e.g. Musyl et al., 2003; Wilson et al., 2005, 2007), but using the two step approach, of first estimating the raw geolocations and then using these raw geolocations as data in a separate model, is not optimal.

A recently developed statistical model showed increased geolocation accuracy compared to previous approaches (Nielsen and Sibert, 2007). The model represents a complete reworking of light-based geolocation and is unique in many ways. The model avoids making any irradiance threshold assumptions, or constraining the movement of the tag between dawn and dusk. For example, a simple model to estimate location based on sunrise and sunset times is only optimal if the subject does not move—a situation unlikely in many marine species. This model generates two estimates of geographic positions per day (at dawn and dusk). The covariance structure in the model is designed to mimic what is actually seen in light observations. A further unique—but often undervalued—feature of the model is that it is a single integrated model instead of a series of independent models (or procedures), each with its own assumptions and uncertainties, working one after another.

Considering reconstructing a track to be one problem, where the result is a track, as opposed to a few hundred independent problems, where the result of each is a position, has potentially many benefits. It allows the model to assess the strength of the correlation between neighboring points and use that to stabilize the track estimation; all common model parameters are estimated on the basis of the most data; hypotheses about the movement can be tested within the model; the observation noise can be correctly propagated through the model to the most probable track; and finally no information is lost in the exchange of data between different sub-models.

Prior to this study, the integrated model has been evaluated for tags mounted on moorings, drifters, and in simulated cases (Nielsen and Sibert, 2007). It was concluded that the integrated model was superior to classical approaches (raw geolocations from the tag manufacturer), all model parameters were identifiable, and that uncertainties were correctly propagated from the light measurements to the most probable track.

An additional complication when analyzing movements from tags deployed on swimming and diving fish is accounting for diving behavior. Tags widely used are popoff archival tags, where the light measurements are depth-corrected on board the tag, which may not be optimal. As expected, depth-corrected light measurements tend to be noisier than light measurements taken directly at the surface. Furthermore, the popoff tags only transmit a small fraction of the recorded data, due to limitations in data transmissions capabilities and battery capacities, and this fraction is selected on board the tag by a proprietary algorithm. The combined model proved robust enough to handle these added complications.

This paper briefly presents the integrated state space model for estimating the most probable track directly from depth-corrected light measurements. The model's performance is then evaluated for real swimming and diving fish by comparing the derived tracks either to tracks reconstructed using a state space model with sea surface temperature correction on conventionally derived location estimates (Nielsen et al., 2006) or to very accurate tracks obtained using radio-linked satellite tracking methods. The integrated model was vastly superior to classical purely light-based methods in all cases, and gives a sound statistical foundation for estimating the most probable track and quantifying its precision.

Materials and Methods

Data

The data used for model evaluation are from 3 blue marlins (*Makaira nigricans*) and 3 shortfin mako sharks (*Isurus oxyrinchus*) tagged with PSATs. Marlins were tagged with model PAT2, Wildlife Computers, Inc., Redmond, WA, USA while the mako were tagged with model PAT3 tags. The mako sharks were also double-tagged with a radio-linked satellite transmitting tag (Wildlife Computers, Inc., model SPOT3) on their dorsal fin in order to obtain more accurate tracks through the Argos system using Doppler shift (see details at http://www.clsamerica.com/argos-system.html). For blue marlin and mako sharks, latitudes and longitudes were calculated from the PAT2 and PAT3 transmitted depth-corrected light data using Wildlife Computer's proprietary software (WC-GPE suite) with minimal data elimination to remove obviously uninformative points. Details of the tagging, analysis and behavior of the marlins and mako sharks will be described elsewhere.

Irradiance is only useful for geolocation if a change in position is correlated with a change in irradiance. At midnight, even a fairly big change in position (say 5°) will not change the irradiance noticeably—it will still be dark. Similarly at noon, the change in solar altitude resulting from changing position a few degrees will change the irradiance slightly, but local weather conditions, variable water clarity, and other uncontrollable features will have a greater influence on the measured irradiance. Data collected during the periods around sunrises and sunsets are most relevant parts for light-based geolocation.

An automatic procedure to select the periods around these solar events can be set up in many ways. It is likely of minor importance which selection procedure is chosen, as long as it is robust and the resulting intervals include the informative parts of data. A procedure is suggested in Nielsen and Sibert (2007). For the tag types used in this paper, the selection of light reading points was done on board the tag by a procedure known only by the tag manufacturers.

This use of light records in the new model to derive geolocations was unintended by the manufacturer. These light readings were originally stored in the PAT tags to aid the researcher in choosing appropriate light curves for estimating geolocation via a proprietary software package from Wildlife Computers (see Musyl et al., 2001 and Hill and Braun, 2001). The light data compose 12 points for PAT2 and 9 points for PAT3 approximately collected around local dawn and dusk.

Observation times around the ith solar event are denoted $\tau^{(i)} = (\tau_1^{(i)}, \ldots, \tau_{n_i}^{(i)})$, and the corresponding light observations are denoted $l^{(i)} = (l_1^{(i)}, \ldots, l_{n_i}^{(i)})$. Finally the average observation time within the ith interval is denoted $\bar{t}_i$. These average times $\bar{t}_1 \leq \cdots \leq \bar{t}_{2N}$ will be the times where geolocations are computed by the model.

Model

The following briefly describes the model. Further details about the model, and the unscented Kalman filter used to estimate model parameters and the most probable track, are available in Nielsen and Sibert (2007).

The model for the observed light measurements along an animal's track is a state space model, where the transition equation describes the movements along the sphere. A random walk model is assumed:

$$\alpha_i = \alpha_{i-1} + c_i + \eta_i, \quad i = 1, \ldots, 2N \tag{1}$$

Here α_i is a two dimensional vector containing the coordinates $(\alpha_{i,1}, \alpha_{i,2})$ in nautical miles along the sphere from a translated origin at time $\bar{t}_i$. c_i is the drift vector describing the deterministic part of the movement, η_i is the noise vector describing the random part of the movement, and N is the number of days in the track. The deterministic part of the movement is assumed to be proportional to time $c_i = (u\Delta\bar{t}_i, v\Delta\bar{t}_i)'$. The random part is assumed to be serially uncorrelated and follow a two dimensional Gaussian distribution with mean vector 0 and covariance matrix $Q_i = 2D\Delta\bar{t}_i I_{2\times 2}$. Here D is a model parameter expressing the diffusive movement component and $I_{2\times 2}$ is the two dimensional identity matrix.

The measurement equation of the state space model describes the expected light measurements at the position α_i at times $\tau^{(i)}$. The calculation of the expected light measurements can be partitioned into three steps. (1) Position α_i is transformed into degrees of longitude and latitude by a function z as described in Sibert et al. (2003). (2) Solar altitudes θ at the position $z(\alpha_i)$ at times $\tau^{(i)}$ are calculated from standard astronomical algorithms (Meeus, 1998). (3) Expected light measurements are calculated from the solar altitudes. The function φ describing this relationship is not a known function, and hence it must be estimated within the model.

The function φ is represented in the model by a cubic spline function interpolating the points $(\widetilde{\theta}_1, \widetilde{\varphi}_1), \ldots, (\widetilde{\theta}_{n_\varphi}, \widetilde{\varphi}_{n_\varphi})$, where $\widetilde{\theta}_1 \leq \cdots \leq \widetilde{\theta}_{n_\varphi}$ are chosen equidistant over the angles in the dataset. Notice here that the selection of data intervals around solar events substantially narrows the angle interval where this function is needed. The $\widetilde{\varphi}$'s are model parameters. To stabilize the numerical optimization and to reflect our knowledge about the relation between solar altitude and irradiance the $\widetilde{\varphi}$'s are restricted to be increasing. Combining (1)–(3) leads to the following measurement equation:

$$l^{(i)} = \varphi_{\widetilde{\varphi}}\left(\theta(z(\alpha_i), \tau^{(i)})\right) + \varepsilon_i, \quad i = 1, \ldots, 2N \tag{2}$$

For easier future reference the entire non-linear mapping from position to expected light measurement is denoted Λ, which reduces the measurement equation to:

$$l^{(i)} = \Lambda\left(\alpha_i, \tau^{(i)}\right) + \varepsilon_i, \quad i = 1, \ldots, 2N \tag{3}$$

The measurement error ε_i is assumed to follow a Gaussian distribution with mean vector 0, and covariance matrix $\Sigma^{(i)}$, where

$$\Sigma^{(i)}_{j,k} = \begin{cases} \sigma_1^2 + \sigma_2^2 + \sigma_3^2 & , \quad \text{if } \ j = k \\ \sigma_1^2 + \sigma_2^2 \exp\left(-|\tau_j^{(i)} - \tau_k^{(i)}|/\rho\right), & \text{if } \ j \neq k \end{cases} \tag{4}$$

where $j, k = 1, \ldots, n_i$. Here σ_1, σ_2, σ_3, and ρ are model parameters.

The covariance structure in (4) reflects the intuition that two light measurements taken near the same solar event are more similar (correlated), than two taken at separate solar events. If visibility is low one morning all measurements will consequently be lowered. The covariance structure furthermore allows the correlation between two light measurements near the same solar event to decrease, as the time between them increases.

Now the state space model is completely defined. All parameters of this model are:

$$\vartheta = (u, v, D, \widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{n_\varphi}, \sigma_1, \sigma_2, \sigma_3, \rho) \tag{5}$$

The track is predicted and the negative log likelihood is computed via the unscented Kalman filter (Julier et al., 2000). The model parameters are all maximum likelihood estimated.

Kalman Filter Model With Sea Surface Temperature

The Kalman filter model with sea surface temperature (kfsst) was used as baseline for the marlin tracks for which no Argos locations (except pop-off points) were available. The model used light based longitude and latitude estimates determined using the tag manufacturer's software, and the sea surface temperature obtained by averaging the daily near surface temperature measurements. The predicted sea surface temperature at a given location was obtained from satellite images (Reynolds 8 day). All details of this model and corresponding software are described in Nielsen et al. (2006).

Results

In the double tagging study the Argos positions are taken to be the true positions of the mako sharks, as the uncertainties associated with Argos are negligible compared to the uncertainties in any light based estimates.

For longitude, the double tagging data showed that the geolocations were always reasonably estimated by both the raw geolocations from tag manufacturer's algorithm and by the light based state space model. On average, the estimates from the light based state space model were closer to the Argos positions than the raw geolocations. The estimated 95% confidence region covered more than 95% of the Argos positions, which makes it a conservative estimate (Figs. 1(A), 2(A), and 3(A)).

For latitude in the double tagging study, the estimated 95% confidence regions covered less than 95% of the Argos positions, but the estimates from the light based state space model were much closer to the Argos positions than the raw geolocations (Figs. 1(B), 2(B), and 3(B)).

Root mean square error RMSE is an overall measure of how well the geolocations match the Argos positions. The RMSE was in all cases lower for the most probable track from the light based state space model than for the raw geolocations

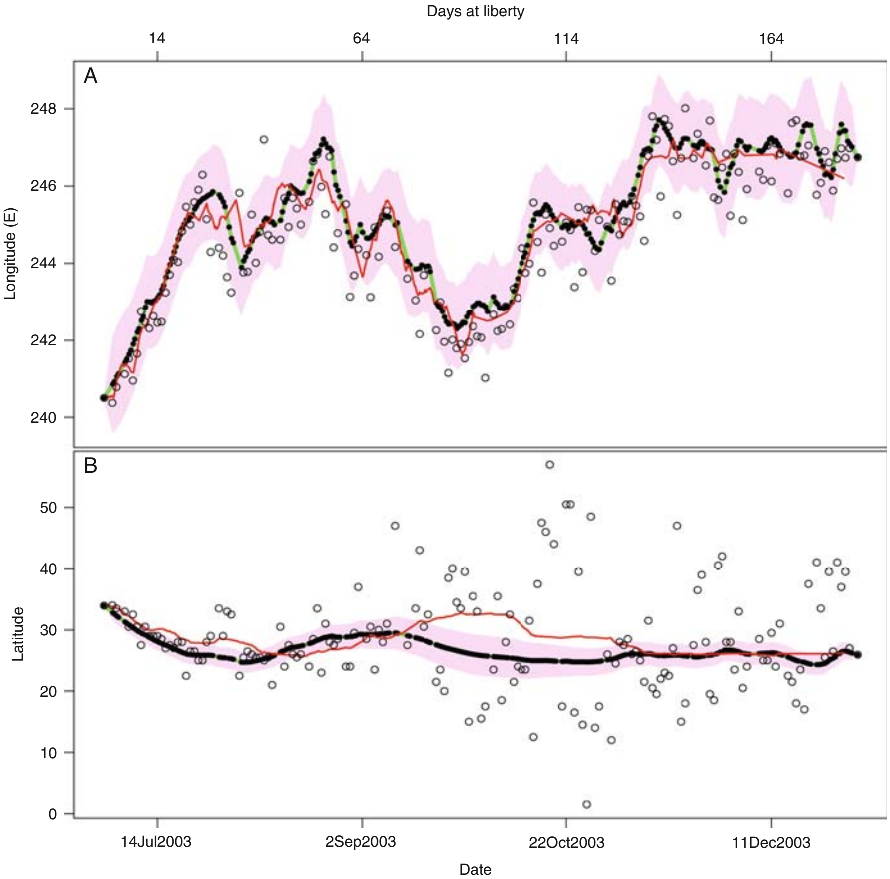

Fig. 1 Longitude (**A**) and latitude (**B**) for a mako shark (tag id 1901). *Solid dots* are estimated points for the most probable track and *shaded areas* are corresponding 95% confidence regions. *Circles* are estimated raw geolocations from the tag manufacturer. *Red lines* are Argos geolocations. Notice the different range in (**A**) and (**B**)

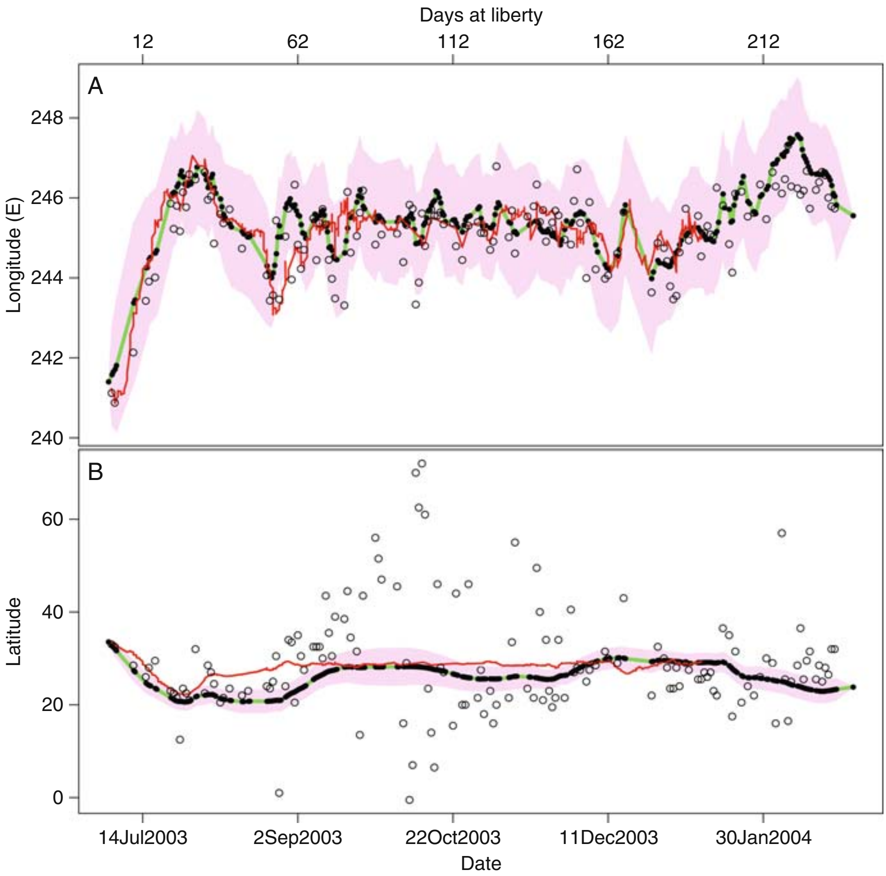

Fig. 2 Longitude (**A**) and latitude (**B**) for a mako shark (tag id 1902). Notation described in caption of Fig. 1

(Table 1). The main improvement is in the latitude coordinate, but the longitude is also improved (Figs. 1, 2, and 3).

For the marlin tracks, the Kalman filter model with sea surface temperature (kfsst), and the light based state space model produced very similar longitude estimates, and they were mostly in close agreement with the longitude coordinate of the raw geolocations (Figs. 4, 5, and 6). The only exceptions were a few raw longitudes on tag 20546 (Fig. 5(C)), which were obviously outliers produced by the raw geolocation algorithm, as they were 5 or even 10 degrees longitude away from all other geolocations near the same time.

The latitude coordinates of the raw geolocations were highly variable and severely biased for the marlin tracks (especially Figs. 5(D) and 6(D)). For those tracks the range of the estimates was 70 degrees of latitude, but that scale of movement is completely unrealistic, and not supported by the kfsst or the light based state space model. The latitude coordinates of the light based state space model and

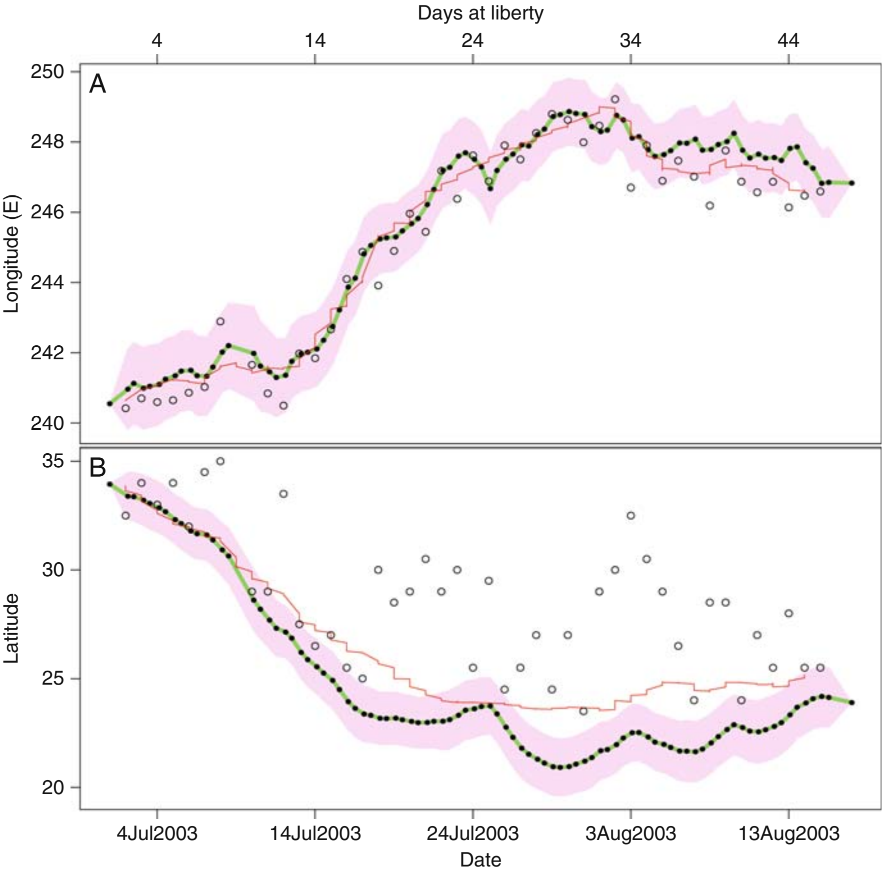

Fig. 3 Longitude (**A**) and latitude (**B**) for a mako shark (tag id 25107). Notation described in caption of Fig. 1

Table 1 Root mean square error (RMSE) in nautical miles of the raw geolocations and of the most probable track from the light based state space model for the mako shark tracks

	Tag ID		
Method	1901	1902	25107
Raw geolocation	503	716	208
Most probable track	183	187	113

of kfsst differ most in periods where SST from the tag was not (or very sparsely) available to kfsst, but are much closer in periods where kfsst had additional information from SSTs (Figs. 4(B), 5(B), and 6(B)). The confidence intervals around the estimated latitude estimates from kfsst and the light based state space model have

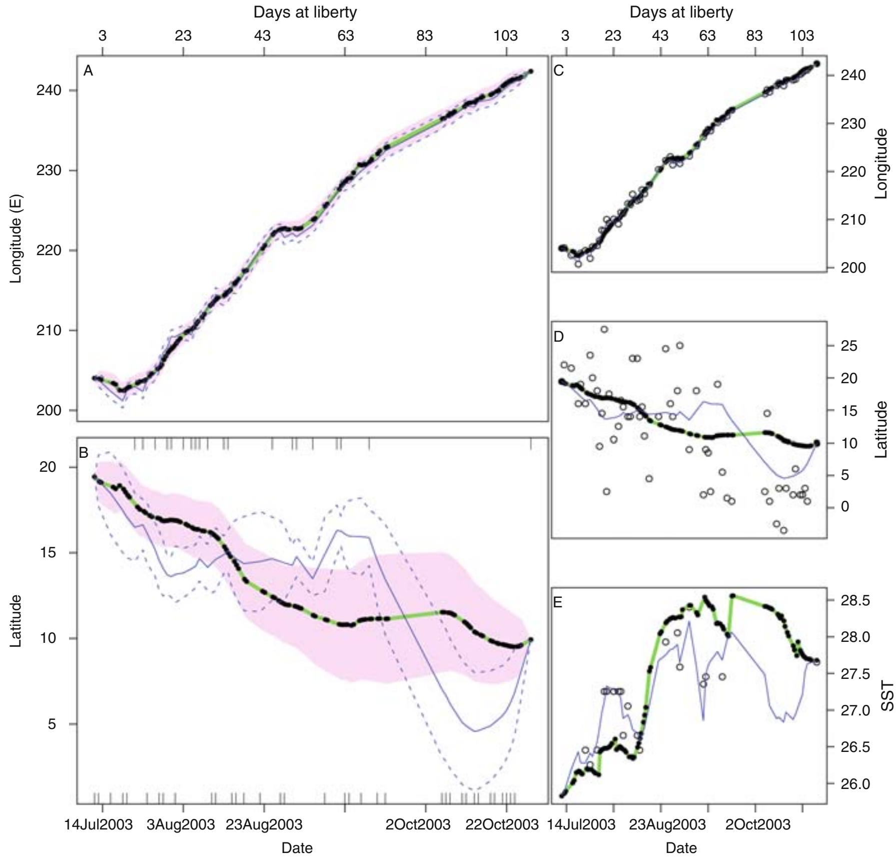

Fig. 4 Track of Marlin (tag id 20545). Solid dots are estimated points on the most probable track from the light based state space model and shaded areas are the corresponding 95% confidence regions. *Solid blue lines* are the estimated most probable track from kfsst and the *dashed lines* are the corresponding 95% confidence regions. *Circles* are estimated raw geolocations from the tag manufacturer. The five different plots show longitude (**A**), latitude (**B**), longitude and latitude scaled to include raw geolocations (**C** and **D**), and finally SST along the two estimated tracks (**E**). Inner tick marks below the latitude plot (**B**) indicates that SST was not available to kfsst at that point, whereas tick marks above indicates that SST was available

the same overall width, but in periods where SST from the tag was not available, the light based state space model has much narrower confidence intervals (Figs. 4(B), 5(B), and 6(B)).

Reconstructing SST via satellite data along the most probable track from the light based state space model showed that in two cases the most probable track from the light based state space model matched the observed SST as well as the most probable track from kfsst (Figs. 5(E) and 6(E)). In one case the most probable track from kfsst matched the observed SST more closely (Fig. 4(E)).

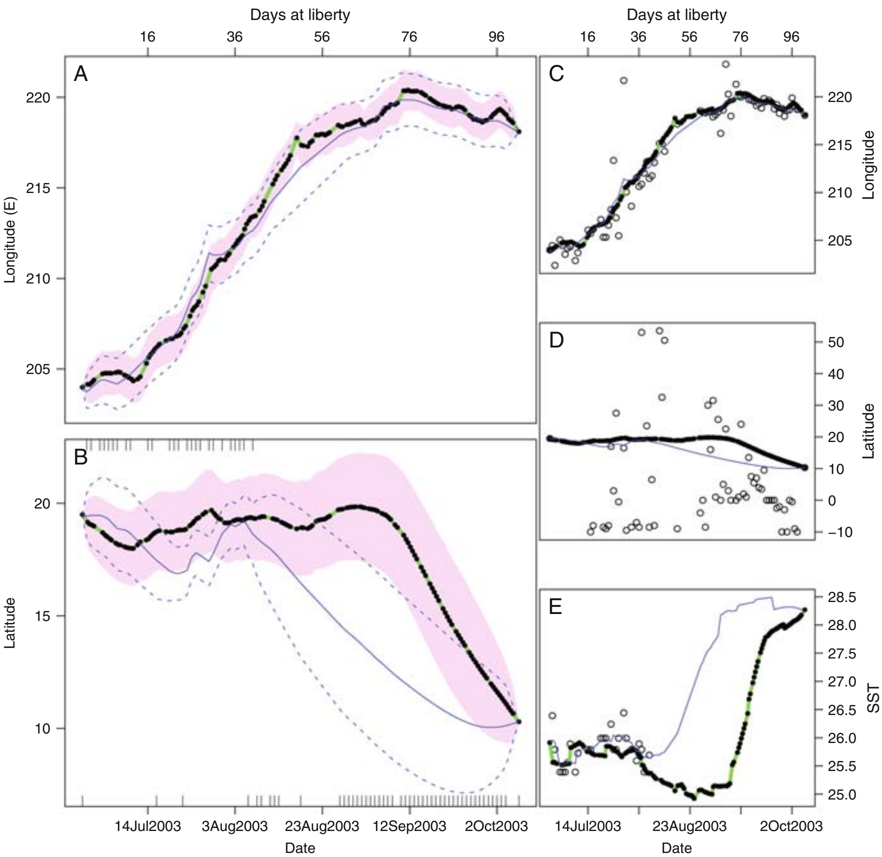

Fig. 5 Track of Marlin (tag id 20546). Notation described in caption of Fig. 4

The raw latitude geolocations from the mako sharks were generally less problematic than from the marlins. Uncertainties were high for the mako sharks, but the huge systematic biases observed in marlin (e.g. Fig. 6(D)) were not seen for the mako sharks. The reasons for differences in raw latitude biases are unknown but may be attributed to the different tag models used or the season and area in which the animals were tagged.

Discussion

The light based state space model is unique in combining all steps from depth corrected light readings to estimated most probable track in one coherent model. The model does not assume to know the relationship between solar altitude and light, but estimates this relationship within the model. It does not constrain the movement

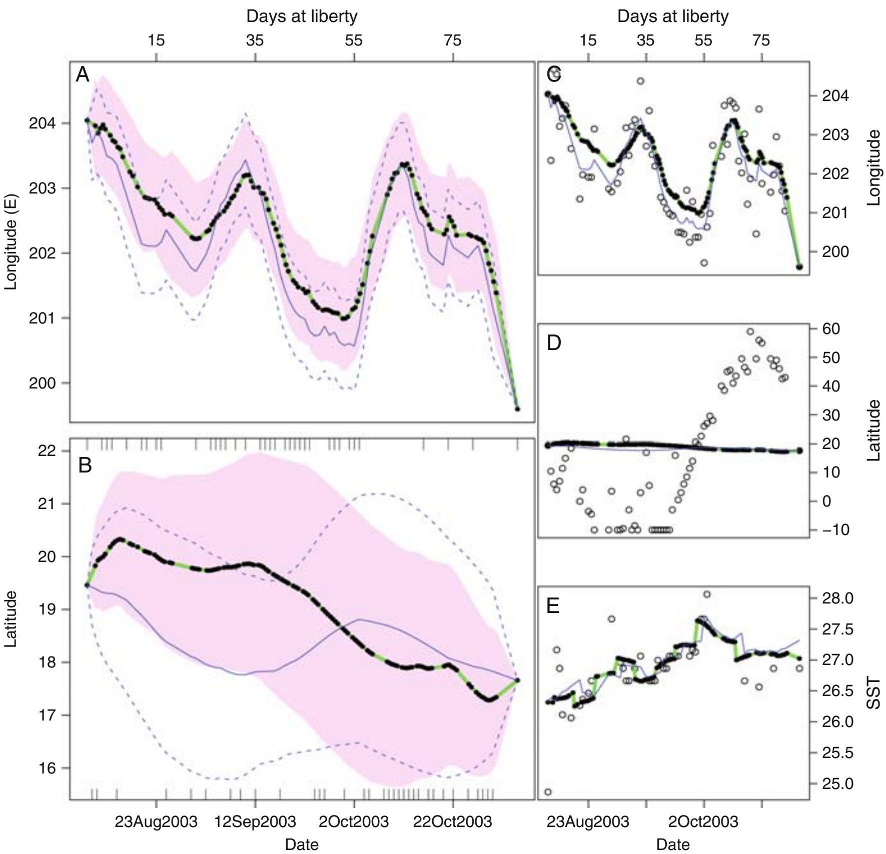

Fig. 6 Track of Marlin (tag id 20673). Notation described in caption of Fig. 4

of the tag between dawn and dusk, and it generates two estimates of geographic positions per day (at dawn and dusk). The verification procedure that this model has been subjected to is also unique. The model has been tested for tags mounted on moorings, drifters, and in simulated cases (Nielsen and Sibert, 2007). This paper validates the model for deployed tags on wild caught sharks and marlin, and as such is forced to address all additional challenges that occur with real data.

The light data returned from satellite transmitting tags used in the study contain only a small fraction of the complete data record, and this fraction is selected on-board the tag and is not necessarily the most informative data for estimating the most probable track. The tag types used here returned 12 light readings per solar event for the marlins and 9 light readings per solar event for the mako sharks. Nine points per solar event could be near the lower limit for the light based state space model to be able produce reliable geolocations, and could be the reason that the confidence regions were too narrow, but even with only 9 points per solar event, the model was able to follow the true Argos track well and with greater precision than the raw light based locations.

The fact that this model was able to use very sparse light data to get vastly improved geographic positions underscores the need for tag manufacturers to save as much of the original data as tag memory or transmission capacity allows. Further improvements would be expected with more of the irradiance data available for the model. Furthermore, the model described herein delivers much better estimation performance than the proprietary software produced by the manufacturer. Therefore, it is incumbent upon the manufacturer to allow further development of these kinds of innovations by enabling further scrutiny of the raw data collected by sensors. It is not enough to simply save the estimated geographic positions, as they are only as good as the models in use at the time the tag was constructed. Clearly as the tags evolve, so do the ways to investigate the data.

In this study, the depth corrected light data was used as returned from the tags. Depth correction is carried out in the tags separately for each day's data, but applying the same logic to depth correction as done for geolocation would suggest that it could be optimal and more stable to depth correct the entire light record with one model. Notice that this is not the same as suggesting to use the same depth correction with the same parameters for the entire record. Using one common model to depth correct the individual measurements is likely not optimal if we only have the 9 or 12 light readings per solar events returned from satellite transmitting tags, but for archival tags where the entire record is returned, it seems like a logical next step.

The algorithm used to compute the likelihood function and reconstruct the most probable track is the unscented Kalman filter (Julier et al., 2000). The unscented Kalman filter is very similar to the extended Kalman filter used in Sibert et al. (2003) and Nielsen et al. (2006), but instead of approximating the nonlinear functions, the transformed probability distributions are approximated directly. This is done by representing the distribution by a set of cleverly selected points, transforming these points by the nonlinear function, and then approximating the mean and variance of the transformed distribution, by the (possibly weighted) mean and variance of the transformed points. This approach gives a simpler implementation, not requiring derivatives of the equations in the state space model, and higher accuracy—at least corresponding to a second order Taylor approximation (Julier et al., 2000). The unscented Kalman filter represents a trade-off between accurate approximation of the model and computation time. More accurate approximations, capable of handling severe nonlinearities and probability distributions other than Gaussians, are available (e.g. particle filter (Gordon et al., 1993) and hidden Markov models (Pedersen, 2007), but these are very computer intensive. As computer power increases, these alternatives will become more appealing.

For the marlin tracks, the width of latitude confidence intervals for the tracks reconstructed via kfsst were about the same as for the track reconstructed via the light based state space model. This is interesting, as kfsst uses auxiliary data (SST) that is not used in the light based state-space model. This means that either SST is useless for these tracks, or the light based state space model makes better use of the light data than the two step approach of first calculating raw geolocations and then running the raw geolocations through a state space model. Although SST was not available for all tracking days, comparing the latitude coordinates of the raw geolocations to the kfsst reconstructed latitudes show that in at least two of the three cases SST must be influential, as the raw geolocations are very problematic.

How informative SST is relative to the light readings will generally depend on time of year and geographic location of the study area. The optimal model should therefore use light as in the state space model presented here, but also use SST as in kfsst. In kfsst and similar models based on conventionally derived location estimates and SST, the SST measurements tends to dominate the latitude reconstruction, because the conventionally derived latitude estimates, are very uncertain. Including SST in the model described here would make the reconstructed tracks more robust to dubious SST observations, as the light based latitude estimates would be less uncertain, and hence carry more weight. Extending the state space model to also use SST is straightforward, but the focus of this study is to evaluate how to best estimate geolocations from light data only.

To advance these methods further it would be extremely helpful if a collection of validation data was made publicly available. Much can be learned from simulation studies and from cross validating animal tracks with auxiliary data, but the real test for these methods is when they are applied to real data from deployed tags. Both data from mooring studies, and double tagged animals are useful. Having such data available would also be helpful for future scientists who are considering deploying electronic data storage tags. They would be able to try their hand at analyzing real data, possibly similar to the data they were hoping to collect, and they would get an impression about the effort involved and the precision obtainable from these methods before actually spending the funds collecting the data.

Acknowledgments The work described in this paper was sponsored by the University of Hawai'i Pelagic Fisheries Research Program under cooperative agreement No. NA17RJ1230 from the National Oceanic and Atmospheric Administration.

References

Arnold G, Dewar H (2001) Electronic tags in marine fisheries research: a 30-year perspective. In: Sibert JR and Nielsen J (eds) Electronic Tagging and Tracking in Marine Fisheries Kluwer Academic Publishers, The Netherlands, pp. 7–64.

Ekstrom PA (2004) An advance in geolocation by light. Mem Natl Inst Polar Res, 58:210–226.

Gordon NJ, Salmond DJ, Smith AFM (1993) Novel approach to non-linear/non-Gaussian Bayesian state estimation. IEE-Proceedings-F, 140:107–113.

Hill RD, Braun MJ (2001) Geolocation by Light Level - The Next Step: Latitude. In: Sibert JR and Nielsen J (eds) Electronic Tagging and Tracking in Marine Fisheries Kluwer Academic Publishers, The Netherlands, pp. 315–330.

Julier SJ, Uhlmann JK, Durrant-Whyte HF (2000) A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Trans Autom Control 45:477–482.

Meeus J (1998) Astronomical Algorithms. 2nd ed. Willman-Bell, Richmond, Virginia, USA.

Musyl MK, Brill RW, et al. (2001) Ability of archival tags to provide estimates of geographical position based on light intensity. In: Sibert JR and Nielsen J (eds) Electronic Tagging and Tracking in Marine Fisheries Kluwer Academic Publishers, The Netherlands, pp. 343–368.

Musyl MK, Brill RW, et al. (2003) Vertical movements of bigeye tuna (*Thunnus obesus*) associated with islands, buoys, and seamounts near the main Hawaiian Islands from archival tagging data. Fish Oceanogr, 12:152–169.

Nielsen A, Bigelow KA, et al. (2006) Improving light-based geolocation by including sea surface temperature. Fish Oceanogr, 15:314–325.

Nielsen A, Sibert JR (2007) State-space model for light-based tracking of marine animals. Can J Fish Aquat Sci, 64:1055–1068.

Pedersen MW (2007) Hidden Markov Models for Geolocation of Fish. Master thesis, Technical University of Denmark, Available at: http://www2.imm.dtu.dk/pubdb/p.php?5183.

Sibert JR, Musyl MK, et al. (2003) Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. Fish Oceanogr, 12:141–152.

Smith P, Goodman D (1986) Determining fish movement from an "archival" tag: precision of geographical positions made from a time series of swimming temperature and depth. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFC- 60 13pp.

Teo SLH, Boustany A, et al. (2004) Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. Mar Ecol Prog Ser, 283:81–98.

Welch DW, Eveson JP (1999) An assessment of light-based geoposition estimates from archival tags. Can J Fish Aquat Sci 56:1317–1327.

Wilson SG, Lutcavage ME, et al. (2005) Movements of bluefin tuna (*Thunnus thynnus*) in the northwestern Atlantic Ocean recorded by pop-up satellite archival tags. Mar Biol 146:409–423.

Wilson SG, Stewart BS, et al. (2007) Accuracy and precision of archival tag data: a multiple-tagging study conducted on a whale shark (*Rhincodon typus*) in the Indian Ocean. Fish Oceanog, 16:547–554.