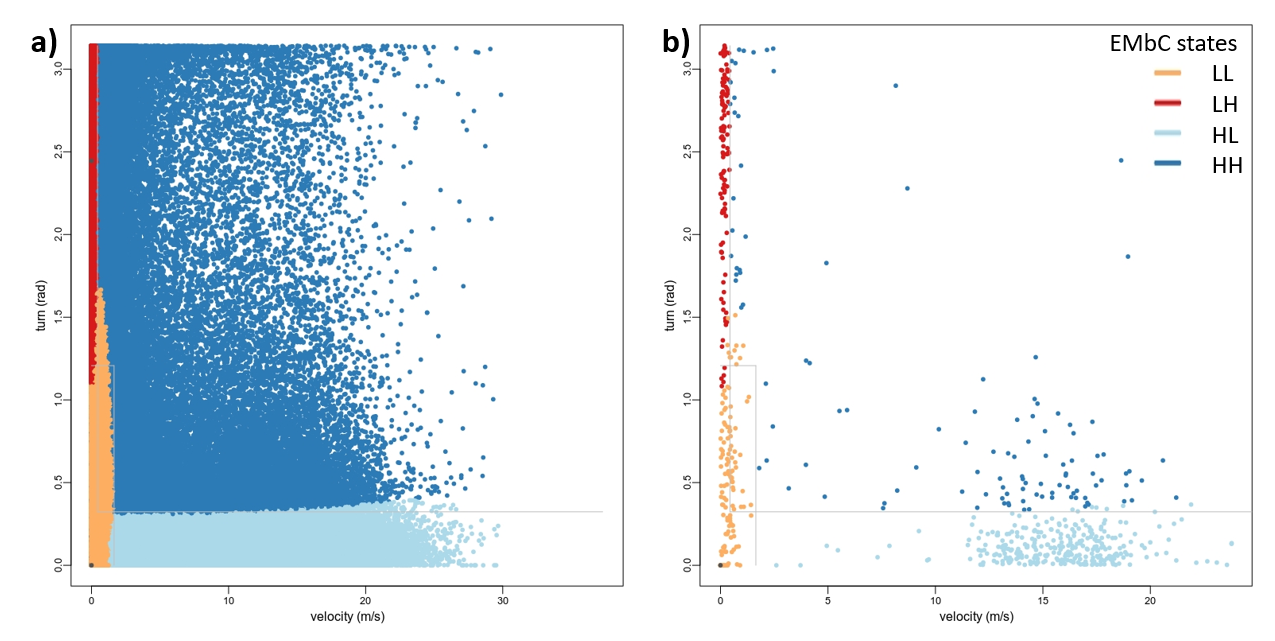
**Supplement 1. Identification of foraging locations using EMbC and performance of EMbC compared to known dive locations**

**EMbC clustering**

EMbC led to stable clustering of the four different states (Figure S1.1). Shags typically forage in dive bouts, i.e. performing a series of dives within a given area for some time, , before moving to another area to continue foraging or return to the colony (Watanuki et al. 2008, Christensen-Dalsgaard et al. 2017). We therefore expected diving to be associated with low velocities and EMbC states LL (low speed, low turning angle) or LH (low speed, high turning angle). We anticipated EMbC states HL (high speed, low turning angle) and HH (high speed, high turning angle) to be associated with arrival at/departure from a foraging site and/or commuting to/from foraging sites.

**Fig. S1.1.** EMbC Cluster semantics. Bivariate (speed / turn) scatter-plots showing the expectation maximization clustering into four clusters. Plot a) shows the data of all birds, b) of one individual bird from Hornøya as example.



-

​**Table S1.1.** Overview of GPS deployments, logger types and GPS sampling intervals for each colony and year. Nind GPS gives the number of individual birds that were deployed with a logger successfully (data included in this study), and the number of individuals deployed with loggers in total (in brackets). Nind TDR gives the number of individuals for which TDR data were obtained. Inc. = incubation, cr = chick-rearing.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Colony & Year | Nind GPS | Sex (F/M) | Nind TDR | Tracking period | Breeding stage | Loggers used | GPS sampling interval |
| Hornøya 2011 | 31 (33) | 14/17 | 16 | 18 June - 13 July | inc & cr | IgotU + TDR | 30 sec |
| Hornøya 2012 | 25 (27) | 11/14 | 23 | 23 June - 10 July | cr | IgotU + TDR | 30 sec |
| Røst 2019 | 7 (9) | 2/5 | 0 | 07 July - 14 Sept. | inc & cr | PathTrack | > 5 min |
| Røst 2020 | 3 (7) | 0/3 | 0 | 23 June - 14 July | inc & cr | PathTrack | > 5 min |
| Sklinna 2011 | 46 (69) | 24/22 | 9 | 08 June - 16 July | inc & cr | IgotU + TDR | 1 min |
| Sklinna 2012 | 46 (49) | 24/22 | 10 | 28 June - 15 July | cr | IgotU + TDR | 1 min |
| Sklinna 2013 | 34 (49) | 16/18 | 34 | 22 june - 15 July | cr | IgotU + TDR | 1 min |
| Sklinna 2014 | 55 (79) | 25/30 | 40 | 11 June - 14 July | inc & cr | IgotU + TDR | 1 min |
| Sklinna 2015 | 29 (32) | 13/16 | 29 | 22 June - 14 July | cr | IgotU + TDR | 1 min |
| Sklinna 2016 | 35 (58) | 14/21 | 26 | 22 June - 14 July | cr | IgotU + TDR | 1 min |
| Sklinna 2017 | 58 (62) | 27/31 | 58 | 29 June - 13 July | cr | IgotU + TDR | 1 min |
| Sklinna 2018 | 40 (48) | 18/22 | 40 | 27 June - 13 July | cr | IgotU + TDR | 1 min |
| Sklinna 2019 | 51 (56) | 22/29 | 54 | 18 June - 13 July | cr | IgotU + TDR | 1 min |
| Sklinna 2020 | 37 (40) | 19/18 | 37 | 24 June - 10 July | cr | IgotU + TDR | 1 min |
| Sklinna 2020 | 10 (10) | 5/5 | 10 | 19 June - 26 June | cr | PathTrack | > 30 sec |
| Sklinna 2020 | 10 (10) | 5/5 | 10 | 24 June – 16 Aug. | cr | PathTrack | > 5 min |
| Runde 2017 | 6 (11) | 4/2 | 6 | 16 June - 23 June | cr | IgotU + TDR | 30 sec |
| Runde 2020 | 11 (11) | 5/7 | 0 | 17 June - 05 Sept. | inc & cr | PathTrack | > 5 min |
| Jarstein 2019 | 7 (8) | 4/3 | 0 | 06 June - 07 Aug. | inc & cr | PathTrack | > 5 min |
| Jarstein 2020 | 9 (10) | 4/5 | 0 | 11 June - 05 Aug. | inc & cr | PathTrack | > 5 min |

**Use of TDR-logger data to validate EMbC**

Since TDR loggers were not deployed on shags from all colonies and during all years included in the study, TDR-logger data in this paper were used exclusively to validate the use of EMbC as a method to identify foraging locations.

TDR-logger files were calibrated and summarized using the diveMove package (Luque 2007). We excluded dives that were shallower than 2 m as likely washing dives, following (Christensen-Dalsgaard et al. 2017). We used the zoc method (Luque 2007) for calibrating the surface with three sequential filters: an initial median smoothing filter with 3-s window width, followed by a 0.1 quantile filter with 3-s window width and ending with a 0.02 quantile filter with 60-s window width. The process was bound to depths between -4 and 4 m.

For each dive, depth and duration was calculated (see Table S1.2 for a summary statistic per year and colony). Dives at a speed of > 3 m/s were regarded as unlikely (Watanuki et al. 2008) and removed from the dataset. Dives were assigned to GPS positions based on date and time. We used a threshold of 30 seconds (as previously used by Christensen-Dalsgaard et al. 2017), and dives outside of this threshold were discarded, since we deemed their locations to be associated with a too large error.

**Table S1.2.** Descriptive statistics (mean ± SD) of shag foraging behaviour at the different colonies. Data are presented pooled for sexes and for all available dives. M/F gives the sex ratio of individuals (Males/Females). TDR data from Røst and Jarstein in 2018 were collected as part of separate pilot studies, with TDRs being attached to leg rings of the birds and logging dive data for 4-31 days during the breeding season (birds did not carry GPS loggers, so dive positions are unknown).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Location** | **Nind** | **M/F** | **N dives** | **Max depth (m)** | **Mean depth (m)** | **Mean duration (s)** |
| Hornøya 2011 | 18 | 8/10 | 6101 | 34.1 | 11.1 ± 5.4 | 37.5 ± 12.7 |
| Hornøya 2012 | 23 | 13/10 | 8357 | 31.5 | 12.0 ± 5.4 | 40.1 ± 11.9 |
| Røst 2018 | 3 | 2/1 | 15882 | 30.1 | 8.4 ± 4.0 | 35.6 ± 12.5 |
| Sklinna 2011 | 10 | 6/4 | 3864 | 56.0 | 12.0 ± 7.7 | 40.0 ± 16.8 |
| Sklinna 2012 | 10 | 4/6 | 5801 | 41.5 | 7.8 ± 5.3 | 31.5 ± 13.3 |
| Sklinna 2013 | 33 | 18/15 | 19969 | 63.3 | 10.9 ± 8.1 | 35.9 ± 17.3 |
| Sklinna 2014 | 39 | 23/16 | 16755 | 63.5 | 10.6 ± 6.9 | 36.7 ± 16.3 |
| Sklinna 2015 | 28 | 15/13 | 16610 | 50.8 | 10.6 ± 7.0 | 36.3 ± 16.2 |
| Sklinna 2016 | 24 | 17/7 | 9478 | 62.6 | 12.3 ± 8.5 | 38.8 ± 18.2 |
| Sklinna 2017 | 57 | 31/26 | 21050 | 48.6 | 9.6 ± 6.4 | 33.0 ± 14.7 |
| Sklinna 2018 | 39 | 21/18 | 22399 | 57.4 | 10.6 ± 7.3 | 35.5 ± 16.7 |
| Sklinna 2019 | 52 | 29/23 | 20157 | 50.0 | 10.6 ± 6.6 | 35.3 ± 14.9 |
| Sklinna 2020 | 58 | 28/30 | 67975 | 59.0 | 10.0 ± 8.4 | 33.5 ± 19.7 |
| Runde 2017 | 9 | 3/6 | 2675 | 61.5 | 14.6 ± 7.9 | 41.7 ± 18.6 |
| Jarstein 2018 | 3 | 3/0 | 4723 | 43.8 | 13.8 ± 10.7 | 41.7 ± 24.7 |

**Comparison between confirmed dive locations and assumed foraging locations based on EMbC clustering**

We assessed the spatial overlap between confirmed diving locations (from the matching of GPS and dive records, as described above) and EMbC-derived assumed foraging locations (i.e. GPS-time-stamps with the EMbC states LL and LH) for those colonies and individual birds, for which both GPS and TDR loggers had been successfully simultaneously deployed. Tracked individuals for which only GPS but no TDR data were collected were excluded from this part of the analyses. Data were analysed separately per year and colony to account for potential variation in foraging locations between years. In the case of Sklinna in 2020 we also analysed the data separately per logger-type / setting (either IgotU and PathTrack loggers with a GPS sampling rate of > 30 sec, or PathTrack loggers with a GPS sampling rate of > 5 min).

Graphically, for most locations and years, there was a high spatial overlap between the confirmed dive locations and the assumed foraging locations based on EMbC clustering (see examples for each colony: Figures S1.2-S1.5). We quantified the spatial overlap by creating a buffer of 200 m around each GPS location, for both known dive locations and assumed foraging locations based on the EMbC clustering. We subsequently calculated the spatial overlap between these buffer areas. This analyses was performed in the R-package sf (Pebesma 2018). We chose a radius of 200 m for the buffer as a compromise between spatial accuracy of the GPS positions (in the range of few metres), compared to actual diving locations (due to the time-lag of up to 30 sec easily 100 m or more) as well as the accuracy of environmental variables used in the analyses (≥ 450 m).

The buffer area around the assumed foraging locations based on the EMbC clustering was somewhat greater than and largely overlapped with the buffer area around known dive locations. Between 81.3 and 98.6% of the buffer area around known dive locations was contained within the buffer area around assumed foraging locations (Table S1.3). As the buffer area around known dive locations was smaller, this buffer only made up between 45.8 and 78.1% of the buffer area around assumed foraging locations based on the EMbC clustering. Notably, the overlap for the data obtained from the PathTrack loggers ≥ 5 min intervals (Sklinna 2020) was in the lower range (87.1 and 59.2%, respectively). We suppose this is caused by a lower proportion of dives matched to GPS positions (due to the longer GPS-intervals, while working with the same 30 sec threshold to match dives) rather than the EMbC algorithm failing.

With the vast majority (> 81%) of known dive locations being spatially close or identical to locations that the EMbC algorithm identified as LL and LH, we concluded that EMbC states LL and LH could be used as indicators of foraging locations for our dataset. By doing so, we might characterise some locations as foraging locations that in fact were not foraging locations (i.e. a form of overestimation). However, these locations should represent a rather small part of the total dataset, and even if not used for foraging purposes, they may still represent important habitat to the shags (possibly an area they are commuting through).

**Table S1.3.** Calculated buffer areas and overlap between buffer areas around known dive locations and around likely foraging locations based on the EMbC algorithm, separately by colony and year. The percentage of overlap was calculated both for the buffer area around known dive locations and likely foraging locations based on the EMbC algorithm.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Colony and Year** | **Buffer area - known dives (km2)** | **Buffer area (km2) – EmbC based foraging locations** | **Overlap (km2)** | **Overlap in % of area around known dive locations** | **Overlap in % of area around EMbC-based foraging locations** |
| Hornøya 2011 | 19.7 | 33.4 | 18.3 | 92.8 | 54.8 |
| Hornøya 2012 | 16.1 | 25.9 | 15.7 | 97.7 | 60.7 |
| Sklinna 2011 | 20.1 | 24.7 | 16.3 | 81.3 | 66.2 |
| Sklinna 2012 | 17.9 | 36.0 | 16.5 | 92.2 | 45.8 |
| Sklinna 2013 | 75.3 | 92.9 | 71.6 | 95.2 | 77.1 |
| Sklinna 2014 | 62.3 | 78.3 | 60.8 | 97.6 | 77.7 |
| Sklinna 2015 | 75.8 | 94.4 | 68.5 | 90.4 | 72.6 |
| Sklinna 2016 | 54.0 | 69.5 | 52.4 | 97.0 | 75.4 |
| Sklinna 2017 | 67.9 | 93.6 | 66.1 | 97.3 | 70.6 |
| Sklinna 2018 | 79.4 | 98.3 | 76.8 | 96.8 | 78.1 |
| Sklinna 2019 | 58.8 | 81.2 | 58.0 | 98.6 | 71.5 |
| Sklinna 2020GPS1min | 71.9 | 89.5 | 66.2 | 92.2 | 74.0 |
| Sklinna 2020GPS5min | 37.6 | 55.3 | 32.7 | 87.1 | 59.2 |
| Runde 2017 | 7.8 | 10.7 | 7.3 | 93.3 | 68.6 |

Sklinna 2020GPS1min: Deployments with IgotU-loggers with a ≥1 min GPS-interval and PathTrack loggers with a ≥30 sec GPS-interval. Sklinna 2020GPS5min: Deployments with PathTrack loggers with a ≥5 min GPS-interval. For all other years and colonies, IgotU loggers with a ≥1 min GPS-interval were used.

**Fig. S1.2.** GPS locations of shags breeding at Hornøya in 2011, and associated EMbC states for all GPS locations (top) and only locations associated with diving (bottom). Land is shown in white. Grey lines and blue shading depict bathymetry. The white star marks the breeding colony. Red = EMbc state 1, Yellow = EMbC state 2, light blue = EMbC state 3 and dark blue = EMbC state 4.

Graphical user interface, chart, map

Description automatically generatedGraphical user interface, map

Description automatically generated

**Fig. S1.3.** GPS locations of shags breeding at Runde in 2017, and associated EMbC states for all GPS locations (top) and only locations associated with diving (bottom). Land is shown in white. Grey lines and blue shading depict bathymetry. The white star marks the breeding colony. Red = EMbc state 1, Yellow = EMbC state 2, light blue = EMbC state 3 and dark blue = EMbC state 4.

Map

Description automatically generated

Map

Description automatically generated

**Chart, map, scatter chart

Description automatically generatedFig. S1.4.** GPS locations of shags breeding at Sklinna in 2020 – here only IgotU loggers and PathTrack loggers with GPS-intervals of 30 seconds, and associated EMbC states for all GPS locations (top) and only locations associated with diving (bottom). Land is shown in white. Grey lines and blue shading depict bathymetry. The white star marks the breeding colony. Red = EMbc state 1, Yellow = EMbC state 2, light blue = EMbC state 3 and dark blue = EMbC state 4.

A map of the world

Description automatically generated with medium confidence**Fig. S1.5.** GPS locations of shags breeding at Sklinna in 2020 – here only PathTrack loggers with GPS-intervals of 5 minutes, and associated EMbC states for all GPS locations (top) and only locations associated with diving (bottom). Land is shown in white. Grey lines and blue shading depict bathymetry. The white star marks the breeding colony. Red = EMbc state 1, Yellow = EMbC state 2, light blue = EMbC state 3 and dark blue = EMbC state 4.Map, scatter chart

Description automatically generatedMap

Description automatically generated

**References**

Christensen-Dalsgaard, S., J. Mattisson, T. Bekkby, H. Gundersen, R. May, E. Rinde, and S.-H. Lorentsen. 2017. Habitat selection of foraging chick-rearing European shags in contrasting marine environments. Marine Biology **164**:196.

Luque, S. P. 2007. Diving Behaviour Analysis in R. R News **7**:8-14.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal **10**:439-446.

Watanuki, Y., F. Daunt, A. Takahashi, M. Newell, S. Wanless, K. Sato, and N. Miyazaki. 2008. Microhabitat use and prey capture of a bottom-feeding top predator, the European shag, shown by camera loggers. Marine Ecology Progress Series **356**:283-293.