# Technical brief –

# update of harbour porpoise summer distribution in the Belt Sea

**University of Veterinary Medicine Hannover, Foundation**

Institute for Terrestrial and Aquatic Wildlife Research (ITAW)

## FINAL REPORT

STIFTUNG · TIERÄRZTLICHE HOCHSCHULE HANNOVER · 1778

***Authors***

Anita Gilles, Rémi Pigeault, Vincent Benoit, Nadya Ramirez-Martinez

University of Veterinary Medicine Hannover, Foundation, Institute for Terrestrial and Aquatic Wildlife Research, Büsum, Germany

Signe Sveegaard

Department of Ecoscience, Marine Mammal Research, Aarhus University, Denmark

***Under contract from Aarhus University, Danish Centre for Environment and Energy (DCE), Roskilde.***

***Contact:***

Eva Friis Møller, Professor, Deputy Head of Department, Department of Ecoscience - ECOS Roskilde, Denmark

**Büsum, 02 Dec 2025**

## Table of Contents

## INTRODUCTION

The harbour porpoise (*Phocoena phocoena*), the only resident cetacean species in the Baltic Sea, is sensitive to a wide range of anthropogenic activities. In the western Baltic Sea, the Belt seas, the Sound and the southern Kattegat the so called "Belt Sea population" has recently been listed as Endangered (HELCOM 2024; from previous Vulnerable). A delineation of the Belt Sea (summer) harbour porpoise management unit, encompassing the core home range area of this population, has been formulated based on morphological studies (Wiemann et al., 2010; Galatius et al., 2012; Lah et al., 2016), satellite telemetry and passive acoustic monitoring (Sveegaard et al. 2015). A critical step towards understanding how sensitive the harbour porpoise is to various threats in different regions of the Baltic Sea is to predict its distribution and determine which areas can be characterised as important for the species.

Predictive modelling of marine mammal habitat is a powerful tool in marine science as it integrates heterogeneity in marine ecosystems and provides important information for ecological studies, management purposes, mitigation of anthropogenic impacts (e.g. Forney et al. 2012) and for understanding the processes that influence interannual and seasonal variability in species distributions (Gilles et al. 2011, 2016, 2025; Becker et al. 2016, 2017, 2018; Pigeault et al. 2024a). Marine mammals exhibit discernible fluctuations in their distributions. However, the drivers responsible for these variations remain unclear, in part due to the dynamic and poorly understood ecological relationships between, for instance, harbour porpoises, their environment and their prey species. In the absence of a comprehensive understanding of these ecological relationships, the here applied species distribution modelling can facilitate the elucidation of these relationships for harbour porpoises in the western Baltic Sea. The objective of this study is to predict and map the long-term summer distribution of harbour porpoises by fitting a habitat-based density model to the high-quality visual survey data collected for harbour porpoises in the "Belt Sea" assessment unit (as defined in Sveegaard et al. 2015) over the last 20 years.

This spatial modelling approach, described in Gilles et al. (2016), utilises physical and biological characteristics as proxies for prey abundance. The ecological theory of species distribution modelling assumes that distribution is at least partly related to environmental variables, and that the relationship between species occurrence and environmental parameters can therefore be used to predict the distribution of the species in question (Austin 2007). Although behavioural factors such as migration, predator avoidance and social interactions influence the distribution of cetaceans, many of the distribution patterns are determined by the foraging response of top predators in a dynamic system (Redfern et al. 2006). This is particularly evident in the case of the harbour porpoise, which exhibits a marked preference for predictable hotspots characterised by high food availability. In the absence of specific prey density data at the necessary spatial resolution for habitat prediction modelling, physical and biological characteristics of the sea are utilised as proxies.

The result of the project can be used for sensitivity mapping of harbour porpoises in the context of marine spatial planning or in assessing sensitivities towards a broad range of anthropogenic activities in the Belt Sea.

# METHODS

## Spatial modelling

The modelling process used in Gilles et al. (2016), applying the workflow described in Gilles et al. (2025) for the spatial modelling of the SCANS-IV survey data, was reproduced for this study. Briefly, the modelling framework is based on Generalized Additive Models (GAMs) to analyse the survey data, with the objective of establishing the relationship between the number of observed animals (the response variable) and the environmental variables (the explanatory variable) (Hastie & Tibshirani 1990, Wood 2017). This framework allows to model non-linear relationships between cetacean habitat and the marine environment.

### ***Data preparation and processing***

A comprehensive dataset was compiled from multiple dedicated line-transect ship or mainly aerial surveys, encompassing the currently defined management area of the Belt Sea population (Sveegaard et al. 2015). The data presented herein were obtained from two sources: large-scale regional and more frequent national monitoring surveys. The first dedicated large-scale survey covering the management unit included here was SCANS-II in 2005 (Hammond et al. 2013), thereafter MiniSCANS in 2012 (Viquerat et al. 2014), SCANS-III in 2016 (Hammond et al. 2021), MiniSCANS-II in 2020 (Unger et al. 2021) and the SCANS-IV survey, which was recently conducted in 2022 (Gilles et al. 2023). The national monitoring of the western Baltic Sea waters included visual aerial survey efforts from Germany (2006-2024) and Denmark (2021-2024).

All surveys followed the standardised SCANS field protocol and used the same data collection software (Scheidat et al. 2008; Gilles et al. 2009, 2016, 2023; Hammond et al. 2013, 2021; SAMMOA 2022). Survey effort and sightings data were quality-checked and several plausibility checks were performed. Data were then aggregated into the common dataframe.

The line-transect effort data were segmented into continuous portions of effort of approximate 10 km mean length, conforming with Becker et al. (2020), Gilles et al. (2016, 2025) and Virgili et al. (2019). Harbour porpoise sighting data were assigned to each segment. The effective area searched was estimated for each segment, based on the survey-specific effective strip widths (including g(0)). The effective area searched was subsequently included as an offset in the model structure. This procedure accounts for both varying segment lengths and the different detection probabilities recorded during the surveys. The covariates were extracted at a daily resolution within a buffer of 5 km around the segment centroids (see **Table 1** for candidate covariates). A suite of environmental covariates, i.e. spatial, static and dynamic covariates, were considered. The selected habitat predictors are assumed to be proxies for the unmeasured underlying ecological processes driving species distributions rather than direct drivers.

***Table 1***. *Candidate environmental covariates used in the density surface models, shown to be important in previous modelling exercises (e.g. Gilles et al. 2016, 2025; Lacey et al. 2022; Pigeault et al. 2024b).*

<table>
<tr><th>Covariate</th><th>Description</th><th>Source</th><th>Initial resolution</th></tr>
<tr><td>X</td><td>Longitude converted to ETRS89 (EPSG:3035)</td><td></td><td></td></tr>
<tr><td>Y</td><td>Latitude converted to ETRS89 (EPSG:3035)</td><td></td><td></td></tr>
<tr><td>Water depth (depth)</td><td>Mean water depth (m)</td><td>EMODnet Digital Bathymetry (DTM 2020)</td><td rowspan="2">200 m</td></tr>
<tr><td>Slope</td><td>Slope of the seabed (°) calculated with R package raster, version 3.6-26 (Hijmans 2013).</td><td>Derived from bathymetry</td></tr>
<tr><td>Mean sea surface temperature (SST and SST_8d)</td><td>Daily temperature (°C) on the survey date but also averaged with the previous 7 days.</td><td rowspan="6">Global Ocean Physics Reanalysis. E.U. Copernicus Marine Service Information (CMEMS)</td><td rowspan="6">0.083° × 0.083°</td></tr>
<tr><td>Spatial sea surface temperature deviation on a radius of one cell (SST_SDSpace)</td><td>Spatial deviation in daily temperature (°C) within the radius of one cell, on the survey date.</td></tr>
<tr><td>Temporal sea surface temperature deviation over 8 days (SST_SDTime)</td><td>Temporal deviation in daily temperature (°C) over 8 days, calculated with the standard deviation.</td></tr>
<tr><td>Difference of temperature between surface and sea floor (ΔTemperature)</td><td>Difference between SST and bottomT (°C).</td></tr>
<tr><td>Eddy kinetic energy (EKE)</td><td>Eddies calculated as the current velocity (m/s).</td></tr>
<tr><td>Mixed layer thickness (MLD)</td><td>Ocean mixed layer thickness defined by sigma theta (m).</td></tr>
<tr><td>Net primary productivity (NPPV and NPPV_8d)</td><td>Expressed as carbon in sea water (mg/m$^3$/day), calculated on the survey date but also averaged with the previous 7 days.</td><td>Global Ocean Biogeochemistry Hindcast (CMEMS)</td><td>0.25° × 0.25°</td></tr>
</table>

### ***Data analysis***

All data processing was undertaken in software R version 4.4.0 (R Core Team 2024), and modelling was conducted using R package *mgcv*, version 1.9-1 (Wood 2017).

### ***Model structure, fitting and selection***

A multi-stage modelling approach was implemented with the objective of reducing bias in the density estimates generated from the habitat-based models. Methods largely followed the one described in Gilles et al. (2025), using Generalised Additive Models (GAM) to link environmental covariates to observations and we refer to this source for detailed information. Briefly, smooth functions were fitted using restricted maximum likelihood (REML) with automatic term selection (Marra & Wood 2011). Cubic regression splines were used for all covariates, with a maximum number of knots set to 10. This model-fitting method helps to avoid overfitting of the smooth functions by including a penalization (Marra & Wood 2011). The method can reduce the estimated degrees of freedom of a covariate term towards zero if it does not contribute sufficiently to account for the variability in the data.

In addition to the spatial-autocorrelation smooth *te(X, Y)* of locational covariates, a deviation by year from this main spatial-autocorrelation smooth was tested in the model, using the factor smooth *bs = 'fs'* (Pedersen et al. 2019). This enables the spatial autocorrelation to reflect yearly survey observations when available, while otherwise conforming to the main spatial-autocorrelation smooth, *te(X, Y)*. In doing so, it avoids extrapolating into spatio-temporal frames without survey coverage. However, from visual inspection, predictions of harbour porpoise density better matched the observations without the yearly deviation in the spatial-autocorrelation smooth, and including this term reduced model parsimony; therefore, the term was removed during model selection.

Models were fitted for each possible combination of two to five uncorrelated covariates (that is covariates with a Pearson's pairwise correlation coefficient < 0.50). The five models with the best goodness of fit, based on leave-one-out cross-validation, were selected and their respective predictions were stacked (Yao et al. 2017) for further investigation, with their respective contribution to the final prediction estimated with the loo R-package (version 2.8.0, Vehtari et al. 2017).

For accurate uncertainty quantification, a pseudo-posterior approach was taken (e.g. King et al. 2000). The pseudo-posterior approach allows for seamless quantification of uncertainty for any derived quantities (e.g. abundance, CV, 95% confidence interval) from model parameters. Maximum likelihood estimates of parameters and their associated covariance matrix were extracted from fitted models (using the function rmvnorm from package *mvtnorm,* version 1.3-3, Genz & Brentz 2009) and used to generate a sample of 1,000 values from a pseudo-posterior, assuming a multivariate normal distribution for the parameters (King et al. 2000). This sample was used to carry out predictions at a daily level over the survey period. As the variability of some smooth functions was high and the upper values could reach extreme densities that are ecologically unrealistic due to the over-dispersion parameter, a threshold in densities was set to the 99.9% quantile of the initial predicted density (i.e. by sample, cell and day). Densities above this limit were removed from the samples as they could not be considered as ecologically realistic. Finally, the predicted densities were averaged over the survey period for each cell and sample, providing a pseudo-posterior distribution per cell.

For SCANS-IV, a prediction grid of 10x10 km was used to account for the extensive area over which the density was predicted (Gilles et al. 2025). Here, the spatial extent of the prediction grid was limited to the Belt Sea management unit (Sveegaard et al. 2015). Therefore, to facilitate a more nuanced prediction, a 5x5 km prediction grid was developed, one which would prove more useful in the context of a sensitivity analysis. On the temporal extent, daily prediction grids of 5x5 km were created from the first to the last summer survey date, specific for each survey year, with each grid populated using daily covariate mean values. This approach entailed predicting at a finer resolution than that of the training data. However, given the size of the study area and previous evidence that predictions at this resolution yield realistic results (see Gilles et al. 2016), this was deemed to be appropriate. The same evidence was drawn in the context of species distribution modelling for fish species (Núñez-Riboni et al. 2021).

### *Model evaluation*

The performance of the model was evaluated using several established metrics. These included the percentage of explained deviance, deviance residuals, information criteria and visual inspection of predicted and observed distributions.

QQ plots, degrees of freedom, fitted relationships, predicted species distributions and abundances were inspected for this selection of models. Goodness-of-fit and model performance diagnostics were overall consistent among selected models, and the model contributing the most to the final prediction was finally selected.

# RESULTS AND DISCUSSION

## Spatial modelling

### *Searching effort and sightings*

Survey data collected in the summer months (June-August) constituted the largest proportion of available data and, accordingly, predictions are valid for the summer. The total number of effort segments, groups and individuals sighted per segment is reported in **Table 2** for each year. The maps of spatial distribution of realised line-transect effort and sightings for the years 2005 to 2024 are presented in **Figure 1**.

We aggregated 34,346 km of on-effort survey data with 1,653 sightings of harbour porpoise groups. From this, a total of 3,560 effort segments, with a final segment mean length of 9.6 km (SD = 2.0 km), were included in the modelling.

***Table 2****. Summary of 2005-2024 survey data from visual surveys, showing number of effort segments, number of harbour porpoise groups and individuals sighted in the summer of each year, as used for model fitting.*

| Year | Total no. of effort segments | No. of effort segments with groups | % effort segments with groups | Number of groups | Number of individuals |
|---|---|---|---|---|---|
| 2005 | 316 | 68 | 21.5 | 131 | 162 |
| 2006 | 135 | 26 | 19.3 | 37 | 40 |
| 2008 | 138 | 39 | 28.3 | 57 | 81 |
| 2010 | 149 | 42 | 28.2 | 61 | 74 |
| 2011 | 154 | 26 | 16.9 | 33 | 38 |
| 2012 | 72 | 38 | 52.8 | 104 | 141 |
| 2013 | 149 | 39 | 26.1 | 55 | 78 |
| 2015 | 242 | 53 | 21.9 | 73 | 82 |
| 2016 | 301 | 91 | 30.2 | 349 | 462 |
| 2018 | 125 | 7 | 5.6 | 9 | 10 |
| 2019 | 249 | 55 | 22.1 | 79 | 102 |
| 2020 | 524 | 433 | 25.4 | 210 | 259 |
| 2021 | 202 | 70 | 34.7 | 144 | 211 |
| 2022 | 453 | 123 | 27.2 | 178 | 252 |
| 2023 | 188 | 49 | 26.1 | 65 | 74 |
| 2024 | 163 | 48 | 29.5 | 68 | 79 |
| **Total** | **3560** | **907** | **25.5** | **1653** | **2145** |

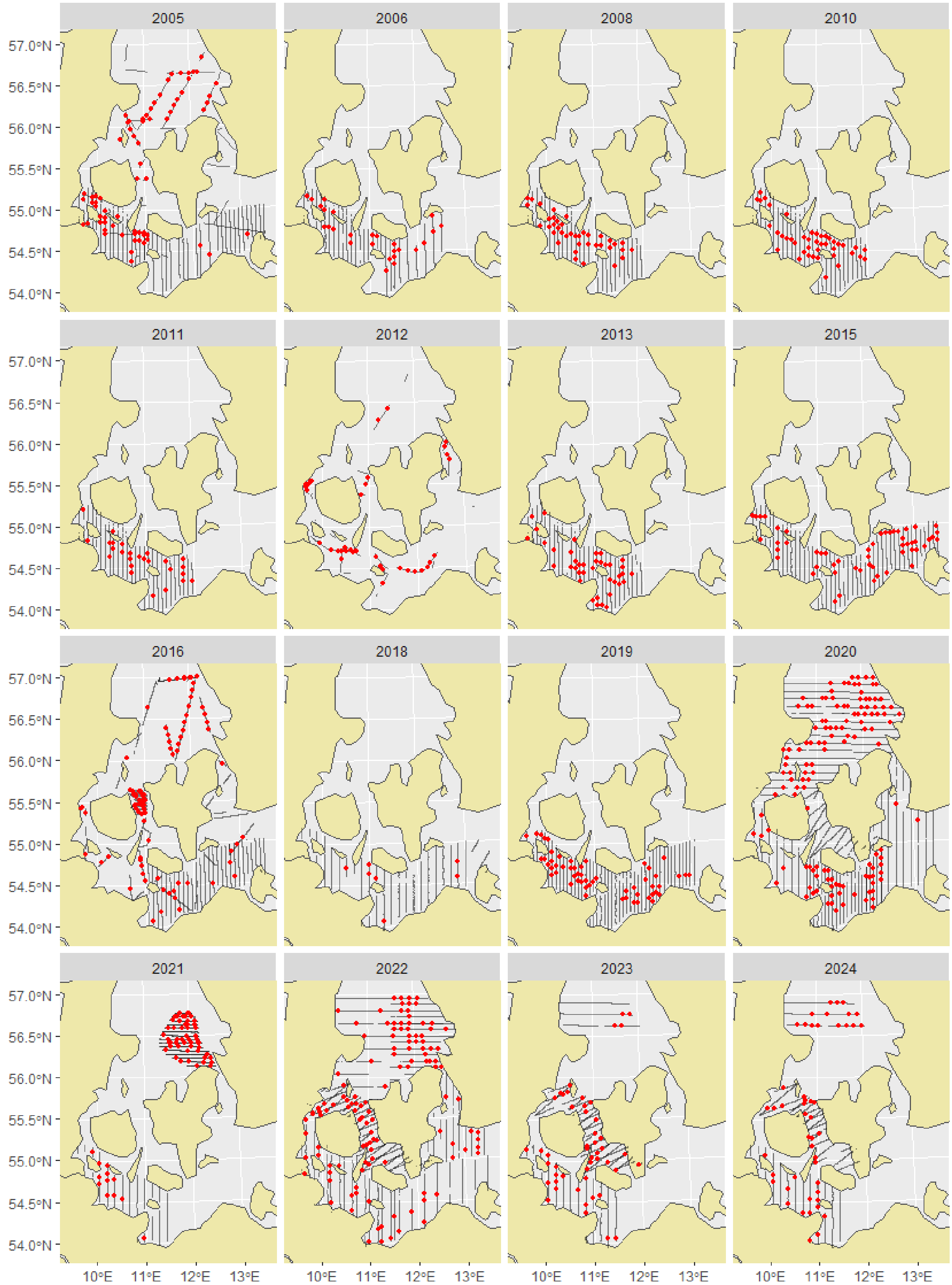


***Figure 1**. Data used for model fitting: this represents ship or aerial survey effort transects (black) and sightings of harbour porpoise groups (coloured dots) for the survey years 2005-2024. To ensure data quality and representativeness of local porpoise occurrence, data collected under poor sighting conditions were discarded (see definition in Gilles et al. 2025).*

### ***Model results and predicted density surfaces***

The best model explained 24.9% of the deviance using a negative binomial error distribution and a two-dimensional tensor product smooth of location, and further covariates as seen in **Table 3**.

***Table 3***. *Model description and diagnostics for the final selected model of harbour porpoise individuals. Covariates in the models are described in* ***Table 1***.

| Distribution | Model covariates | Estimated degrees of freedom | Model degrees of freedom | % Deviance explained |
|---|---|---|---|---|
| Negative Binomial (0.424) | X,Y | 12.9 | 25.1 | 24.9 |
| | Slope | 3.1 | | |
| | ΔTemperature | 1.8 | | |
| | EKE | 2.1 | | |
| | MLD | 1.3 | | |
| | SST_8d | 3.9 | | |

The best model also included several dynamic covariables such as sea surface temperature (SST_8d) as well as Eddy kinetic energy (EKE), difference of temperature between surface and sea floor (ΔTemperature) and Mixed layer thickness (MLD). Although the locational covariates (x,y) still explained most of the model's deviance, the model thereby also captured distinct species-environment relationships involving dynamic variables, also from the water column (3-D), that are consistent with harbour porpoise ecology and habitat use. These relationships suggest useful proxies that can effectively describe secondary production and prey aggregations.

The visualisation of the prediction generated from the GAM clearly delineates regions where the predicted porpoise density is greatest and where it levels off (towards the east). The maps showing surfaces of predicted density and associated estimated coefficient of variation (CV) are shown in **Figure 2**. The patterns of predicted density are influenced by the covariates retained in the models (see **Table 3**), the fitted smooth functions (see **Figure 3**), and spatial variation in the covariates' values in the prediction grid.

The maps of CVs provide a measure of the confidence in predicted density across the survey area. Lower CVs are generally associated with areas of higher density when predictions are interpolations *sensu* Pigeault et al. (2024b). On the other hand, high CV values may betray either extrapolations, a high between-day variance in predictions, or be associated with areas of very low density. CVs for predicted harbour porpoise density (**Figure 2**) are relatively low across most of the survey area, whereas the confidence in the predictions in areas of low density is generally poorer. The magnitude of the CV is influenced by the number of sightings as well as by how well the models fit the data.

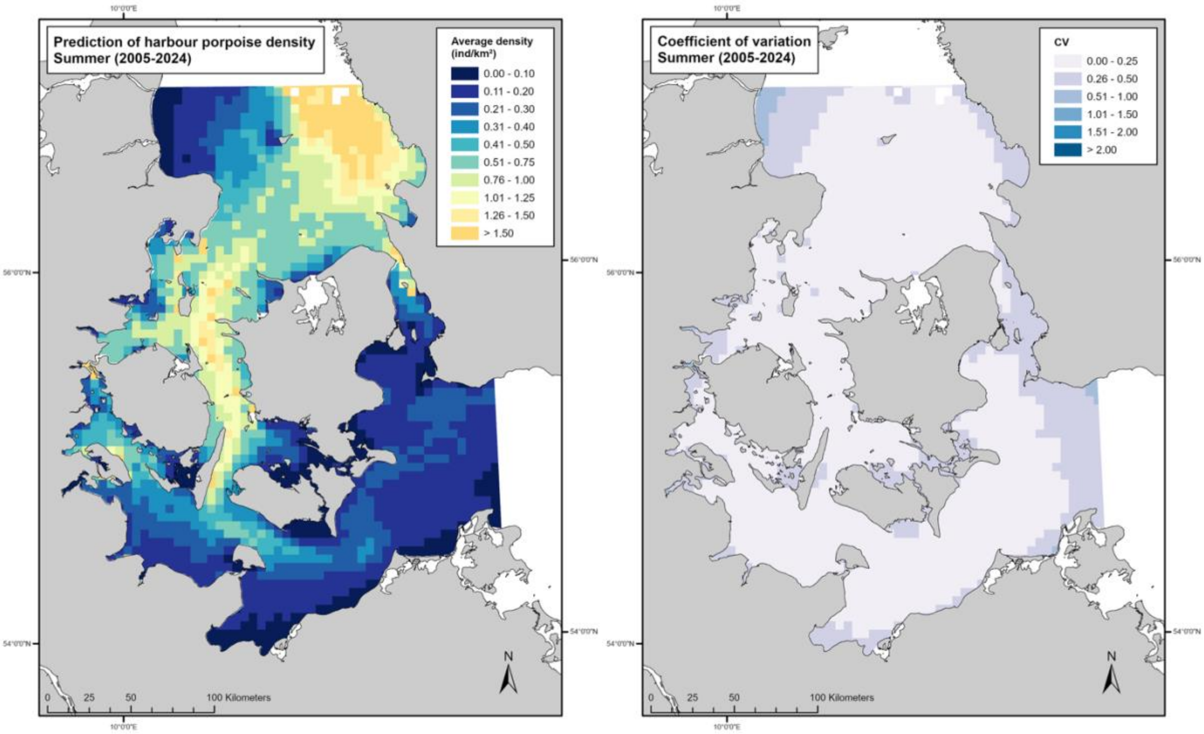


***Figure 2**. Predicted surface of estimated summer density [left] and associated coefficient of variation (CV) [right] for harbour porpoise, fitting the model to survey data collected between 2005-2024 in the management unit of the Belt Sea harbour porpoise population.*

***Model diagnostic***

The fitted smooth relationships between relative density and the spatial-autocorrelation smooths (main smooth and deviation by year), as well as each covariate selected in the final best model, are shown in **Figure 3**. The spatial autocorrelation te(X, Y) explained most of the variance in this final model, while other covariates contributed an equivalent magnitude to the model. Slope and SST_8d (i.e., a proxy for thermal fronts) also contributed to a higher magnitude, as shown by the larger amplitude of their respective partial effects on the y-axis.

The Q-Q plot (**Figure A.1.**) shows that the deviance residuals follow the theoretical quantiles of the distribution family used (negative binomial). This suggests that the chosen distribution is appropriate for the fitted model.

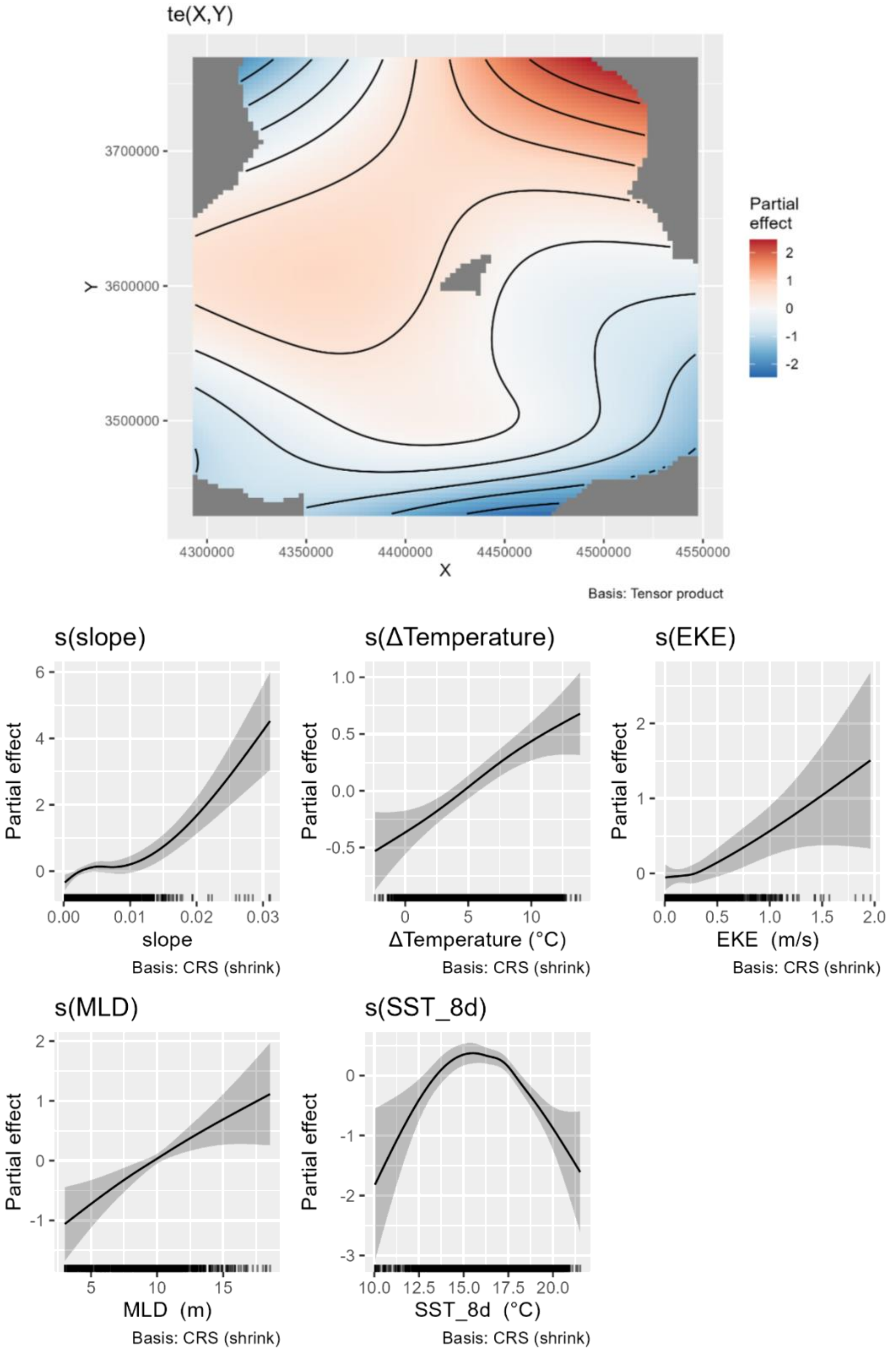


***Figure 3***. *Functional plots of environmental variables relative to harbour porpoise density as indicated by the estimated smooth functions for the selected covariates in the best model (see **Table 3** for explanation on covariates). Y-axes represent the contribution of the term (linear or spline, on the log-scale) to predict the response variable. Zero on the y-axes corresponds to no effect of the predictor variable on the estimated response variable. Scaling of y-axis varies among predictor variables to emphasize model fit. The shading reflects 2× standard error bands (i.e., 95% confidence interval); tick marks ("rug plot") above the X-axis show data values.*

In conclusion, our modelling exercise yielded a prediction of a multi-year average density surface (see **Figure 2**), driven by a substantial number of dedicated visual surveys and, consequently, data points, resulting in a powerful dataset for predicting harbour porpoise summer density distribution in the management area of the Belt Sea population. Consequently, the predicted density surface now reflects habitats that are important over time (also in earlier years) for the harbour porpoise and highlights the areas that are sensitive to anthropogenic impacts.

In order to facilitate management and risk scenarios related to the most recent time, another model was fitted but only including surveys from 2020 to 2024 (i.e., miniSCANS-II and SCANS-IV as well as national monitoring surveys in Denmark and Germany). The results of this model are presented in the Appendix (**Table A. 1**; **.** ). The selected best model has a good fit to the data and explained 16.7% of deviance, yet overall fewer dynamic co-variates were selected. The predicted surface shows that harbour porpoise density, especially in the region of the Great Belt, was reduced in the most recent period. This was previously demonstrated in the modelling report for SCANS-IV (Gilles et al. 2025), however, using data exclusively from the 2022 SCANS-IV survey and predicting on a larger grid (10x10 km).

## ACKNOWLEDGMENTS

The survey data collection used in this study was primarily funded by the Danish Ministry of Environment and Gender Equality (Denmark), the German Federal Agency for Nature Conservation (BfN) and the Ministry for the Environment, Climate Action, Nature Conservation and Nuclear Safety (BMUKN, Germany) and the Swedish Agency for Marine and Water Management (Sweden). The analysis performed in this study by TiHo-ITAW subcontracted by Aarhus University, was funded as part of the project "Environmental mapping and screening of the offshore wind potential in Denmark" by the Danish Energy Agency.

# REFERENCES


Austin, M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecol. Modell. 200: 1–19. doi:10.1016/j.ecolmodel.2006.07.005

Becker, E.A., Forney, K.A., Fiedler, P.C., Barlow, J., Chivers, S.J., Edwards, C.A., Moore, A.M., Redfern, J.V. (2016). Moving towards dynamic ocean management: how well do modeled ocean products predict species distributions? Remote Sens. 8: 149. doi: 10.3390/rs8020149

Becker, E.A., Forney, K.A, Thayre, B.J., Debich, A.J., Campbell, G.S., Whitaker, K., Douglas, A.B., Gilles, A., Hoopes, R., Hildebrand, J.A. (2017). Habitat-based density models for three cetacean species off Southern California illustrate pronounced seasonal differences. Front. Mar. Sci. 4: 121. doi: 10.3389/fmars.2017.00121

Becker, E.A., Forney, K.A., Redfern, J.V, Barlow, J., Jacox, M.G., Roberts, J.J., Palacios, D.M. (2018). Predicting cetacean abundance and distribution in a changing climate. Divers. Distrib. 25: 626–643. doi:10.1111/ddi.12867

Becker, E.A., Forney, K.A., Miller, D.L., Fiedler, P.C., Barlow, J., & Moore, J.E. (2020). Habitat-based density estimates for cetaceans in the California Current Ecosystem based on 1991-2018 survey data. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-638.

EMODnet Digital Bathymetry (DTM 2020). EMODnet Bathymetry Consortium https://doi.org/10.12770/bb6a87dd-e579-4036-abe1-e649cea9881a

EMODnet Bathymetry World Coastline. EMODnet Bathymetry Consortium https://doi.org/10.12770/36821cff-56db-4e96-8426-ddfe68240c4b

Forney, K.A., Ferguson, M.C., Becker, E.A., Fiedler, P.C., Redfern, J.V, Barlow, J., Vilchis, I.L., Ballance, L.T. (2012). Habitat-based spatial models of cetacean density in the eastern Pacific Ocean. Endang. Species Res. 16: 113-133. doi:10.3354/esr00393

Galatius, A., Kinze, C.C., Teilmann, J. (2012). Population structure of harbour porpoises in the Baltic region: evidence of separation based on geometric morphometric comparisons. J. Mar. Biol. Assoc. U.K. 92: 1669–1676. doi: 10.1017/S0025315412000513

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Bornkamp, B., Maechler, M., Hothorn, T. (2021). Package 'mvtnorm'. J. Comput. Graph. Stat. 11(950-971), 155.

Gilles, A., Scheidat, M., Siebert, U. (2009). Seasonal distribution of harbour porpoises and possible interference of offshore wind farms in the German North Sea. Mar. Ecol. Prog. Ser. 383: 295-307.

Gilles, A., Adler, S., Kaschner, K., Scheidat, M., Siebert, U. (2011). Modelling harbour porpoise seasonal density as a function of the German Bight environment: implications for management. Endang. Species Res. 14: 157–169

Gilles, A., Viquerat, S., Becker, E.A., Forney, K.A., Geelhoed, S.C.V., Haelters, J., Nabe-Nielsen, J., Scheidat, M., Siebert, U., Sveegaard, S., van Beest, F.M., van Bemmelen, R., Aarts, G. (2016). Seasonal habitat-based density models for a marine top predator, the harbor porpoise, in a dynamic environment. Ecosphere 7 (6), e01367. https://doi.org/10.1002/ecs2.1367

Gilles, A., Authier, M., Ramirez-Martinez, N.C., Araújo, H., Blanchard, A., Carlström, J., Eira, C., Dorémus, G., Fernández-Maldonado, C., Geelhoed, S.C.V., Kyhn, L., Laran, S., Nachtsheim, D., Panigada, S., Pigeault, R., Sequeira, M., Sveegaard, S., Taylor, N.L., Owen, K., Saavedra, C., Vázquez-Bonales, J.A., Unger, B., Hammond, P.S. (2023). Estimates of cetacean abundance in European Atlantic waters in summer 2022 from the SCANS-IV aerial and shipboard surveys. Final report published 29 September 2023. 64 pp. https://www.tiho-hannover.de/itaw/scans-iv-survey

Gilles, A., Authier, M., Pigeault, R., Ramirez-Martinez, N.C., Benoit, V., Carlström, J., Eira, C., Geelhoed, S.C.V., Laran, S., Sequeira, M., Sveegaard, S., Taylor, N.L., Saavedra, C., Vázquez-Bonales, J.A., Hammond, P.S. (2025). Spatial models of cetacean density in European Atlantic waters based on SCANS-IV summer 2022 survey data. Final report published 14 May 2025. 31 pp plus Appendix. https://tinyurl.com/3rv246v5

Global Ocean Biogeochemistry Hindcast. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). DOI: 10.48670/moi-00019 (Accessed on 26-05-2025)

Global Ocean Physics Reanalysis. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). DOI: 10.48670/moi-00021 (Accessed on 31-03-2025)

Hammond, P.S., Macleod, K., Berggren, P., Borchers, D.L., Burt, L., Cañadas, A., Desportes, G., Donovan, G.P., Gilles, A., Gillespie, D., Gordon, J., Hiby, L., Kuklik, I., Leaper, R., Lehnert, K., Leopold, M., Lovell, P., Øien, N., Paxton, C.G.M., Vázquez, J. A. (2013). Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. Biol. Conserv. 164, 107-122. https://doi.org/10.1016/j.biocon.2013.04.010

Hammond, P.S., Lacey, C., Gilles, A., Viquerat, S., Börjesson, P., Herr, H., Macleod, K., Ridoux, V., Santos, M.B., Scheidat, M., Teilmann, J., Vingada, J., Øien, N. (2021). Estimates of cetacean abundance in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. SCANS-III project report 1, 39 pp. https://scans3.wp.st-andrews.ac.uk/files/2021/06/SCANS-III_design-based_estimates_final_report_revised_June_2021.pdf

Hastie, T.J., Tibshirani, R.J. (1990). Generalized additive models. Chapman & Hall, London.

HELCOM (2024). HELCOM Red List II: *Phocoena phocoena* [Species information sheet]. Baltic Marine Environment Protection Commission. https://helcom.fi/wp-content/uploads/2025/11/HELCOM-Red-List-II-Phocoena-phocoena.pdf

Hijmans, R. J., Van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E.B., Shortridge, A. (2013). Raster: Geographic data analysis and modelling. R package version 3.6-26 https://CRAN.R-project.org/package=raster

King, G., Tomz, M., Wittenberg., J. (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation. Am. J. Pol. Sci. 44(2), 347–361. https://doi.org/10.2307/2669316https://doi.org/10.2307/2669316

Lacey, C., Gilles, A., Börjesson, P., Herr, H., Macleod, K., Ridoux, V., Santos, M.B., Scheidat, M., Teilmann, J., Sveegaard, S., Vingada, J., Viquerat, S., Øien, N., Hammond, P.S. (2022). Modelled density surfaces of cetaceans in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. 55 pp. Available from: https://scans3.wp.st-andrews.ac.uk/files/2022/08/SCANS-III_density_surface_modelling_report_final_20220815.pdf

Lah, L., Trense, D., Benke, H., Berggren, P., Gunnlaugsson, Þ., Lockyer, C., Öztürk, A., Öztürk, B., Pawliczka, I., Ross, A., Siebert, U., Skóra, K., Víkingsson, G., Tiedemann, R. (2016). Spatially explicit analysis of genome-wide SNPs detects subtle population structure in a mobile marine mammal, the harbor porpoise. PloS One 11, e0162792. doi: 10.1371/journal.pone.0162792

Lau-Medrano, W (2024). grec: Gradient-Based Recognition of Spatial Patterns in Environmental Data. R package version 1.6.0, https://CRAN.R-project.org/package=grec

Marra, G., Wood, S.N. (2011). Practical variable selection for generalized additive models. Comput. Stat. Data Anal. 55, 2372-2387. https://doi.org/10.1016/j.csda.2011.02.004

Núñez-Riboni, I., Akimova, A., Sell, A.F. (2021). Effect of data spatial scale on the performance of fish habitat models. Fish. Fish. *22*(5), 955-973. https://doi.org/10.1111/faf.12563

Pedersen, E.J., Miller, D.L., Simpson, G.L., Ross, N. (2019). Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7:e6876 https://doi.org/10.7717/peerj.6876

Pigeault, R., Ruser, A., Ramírez-Martínez, N.C., Geelhoed, S.C., Haelters, J., Nachtsheim, D.A., Schaffeld, T., Sveegaard, S., Siebert, U., Gilles, A. (2024a). Maritime traffic alters distribution of the harbour porpoise in the North Sea. Mar. Pollu. Bull. 208, 116925. https://doi.org/10.1016/j.marpolbul.2024.116925

Pigeault, R., Authier, M., Ramirez-Martinez, N.C., Virgili, A., Geelhoed, S.C.V., Haelters, J., Louzao, M., Saaveedra, C., Gilles, A. (2024b). Assessing the reliability of species distribution models under changing environments: A case study on cetaceans in the North-East Atlantic. Glob. Ecol. Conserv. 56, e03299. doi:10.1016/j.gecco.2024.e03299

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Found. Stat. Comput., Vienna, Austria. https://www.R-project.org/

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A.J., Qian, S.S., Read, A., Reilly, S.B., Torres, L., Werner, F. (2006). Techniques for cetacean-habitat modeling. Mar. Ecol. Prog. Ser. 310, 271–295. https://doi.org/10.3354/meps310271

SAMMOA (2022) Software dedicated for aerial survey of marine megafauna. Observatoire Pelagis LRUniv-CNRS - Code Lutin, version 2.1.2. and revised version 2.1.3. (https://www.observatoire-pelagis.cnrs.fr/tools/sammoa)

Scheidat, M., Gilles, A., Kock, K.H., Siebert, U. (2008). Harbour porpoise Phocoena phocoena abundance in the southwestern Baltic Sea. Endang. Species Res. 5:215-223 https://doi.org/10.3354/esr00161

Sveegaard, S., Galatius, A., Dietz, R., Kyhn, L., Koblitz, J.C., Amundin, M., Nabe-Nielsen, J., Sinding, M-H.S., Andersen, L.W., Teilmann, J. (2015). Defining management units for cetaceans by combining genetics, morphology, acoustics and satellite tracking. Glob. Ecol. Conserv. 3, 839–850. doi: 10.1016/j.gecco.2015.04.002

Unger, B., Nachtsheim, D., Ramírez-Martínez, N., Siebert, U., Sveegaard, S., Kyhn, L., Balle, J.D., Teilmann, J., Carlström, J., Owen, K., Gilles, A. (2021). MiniSCANS-II: Aerial survey for harbour porpoises in the western Baltic Sea, Belt Sea, the Sound and Kattegat in 2020. Joint survey by Denmark, Germany and Sweden. Final report to Danish Environmental Protection Agency, German Federal Agency for Nature Conservation and Swedish Agency for Marine and Water Management. 28 pp. URL: https://www.tihohannover.de/fileadmin/57_79_terr_aqua_Wildtierforschung/79_Buesum/downloads/Berichte/20210913_Report_ MiniSCANSII_2020_revised.pdf

Vehtari, A., Gelman, A., Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out crossvalidation and WAIC. Stat. Comput. 27, 1413-1432. https://doi.org/10.1007/s11222-016- 9696-4

Viquerat, S., Herr, H., Gilles, A., Peschko, V., Siebert, U., Sveegaard, S., Teilmann, J. (2014). Abundance of harbour porpoises (*Phocoena phocoena*) in the western Baltic, Belt Seas and Kattegat. Mar. Biol. 161(4), 745-754.

Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., Corkeron, P., Dorémus, G., David, L., Di-Mégli, N., Dunn, C., Dunn, T.E., García-Barón, I., Laran, S., Lauriano, G., Lewis, M., Louzao, M., Manocci, L., Martínez-Cedeira, J., Palka, D., Panigada, S., Pettex, E., Roberts, J.J., Ruiz, L., Saavedra, C., Santos, M.B., Van Canneyt, O., Vázquez Bonales, J.A., Monestiez, P., Ridoux, V. (2019). Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: Large-scale modelling of deep-diving cetacean habitats. Global Ecol. Biogeogr. 28(3), 300-314.

Wiemann, A., Andersen, L. W., Berggren, P., Siebert, U., Benke, H., Teilmann, J., Lockyer, C., Pawliczka, I., Skóra, K., Roos, A., Lyrholm, T., Paulus, K.B., Ketmaier, V., Tiedemann, R. (2010). Mitochondrial Control Region and microsatellite analyses on harbour porpoise

(*Phocoena phocoena*) unravel population differentiation in the Baltic Sea and adjacent waters. Conserv. Genet. *11*(1), 195-211.

Wood, S.N. (2017). Generalized additive models - an introduction with R. Chapman and Hall/CRC, 496 pp.

Yao, Y., Vehtari, A., Simpson, D., Gelman, A. (2017). Using stacking to average Bayesian predictive distributions. Bayesian Anal. 13(3), 917-1007. https://doi.org/10.1214/17-BA1091

## APPENDIX

### *Model diagnostic (all data)*

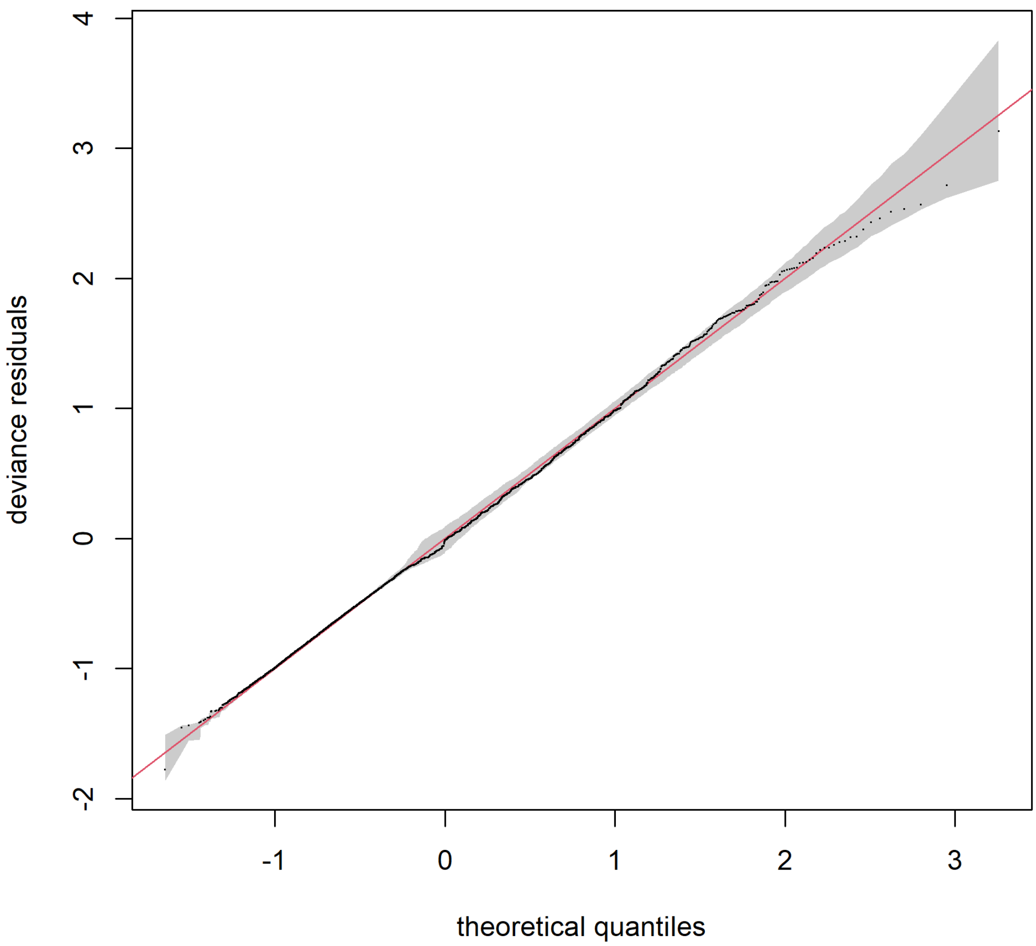


***Figure A. 1.*** *Q-Q plot for the selected final best model fitted with a negative binomial distribution (data period 2005-2024).*

### ***Model fitted to most recent data period (2020-2024)***

***Table A. 1**. Model description and diagnostics for the final selected model of harbour porpoise individuals for the period 2020-2024. Covariates in the models are described in Table 1.*

| Distribution | Model covariates | Estimated degrees of freedom | Model degrees of freedom | % Deviance explained |
|---|---|---|---|---|
| Negative Binomial (0.527) | X,Y | 14.7 | 18.5 | 16.7 |
| | Slope | 1.2 | | |
| | MLD | 2.3 | | |
| | EKE | 0.2 | | |
| | SST | 0.1 | | |

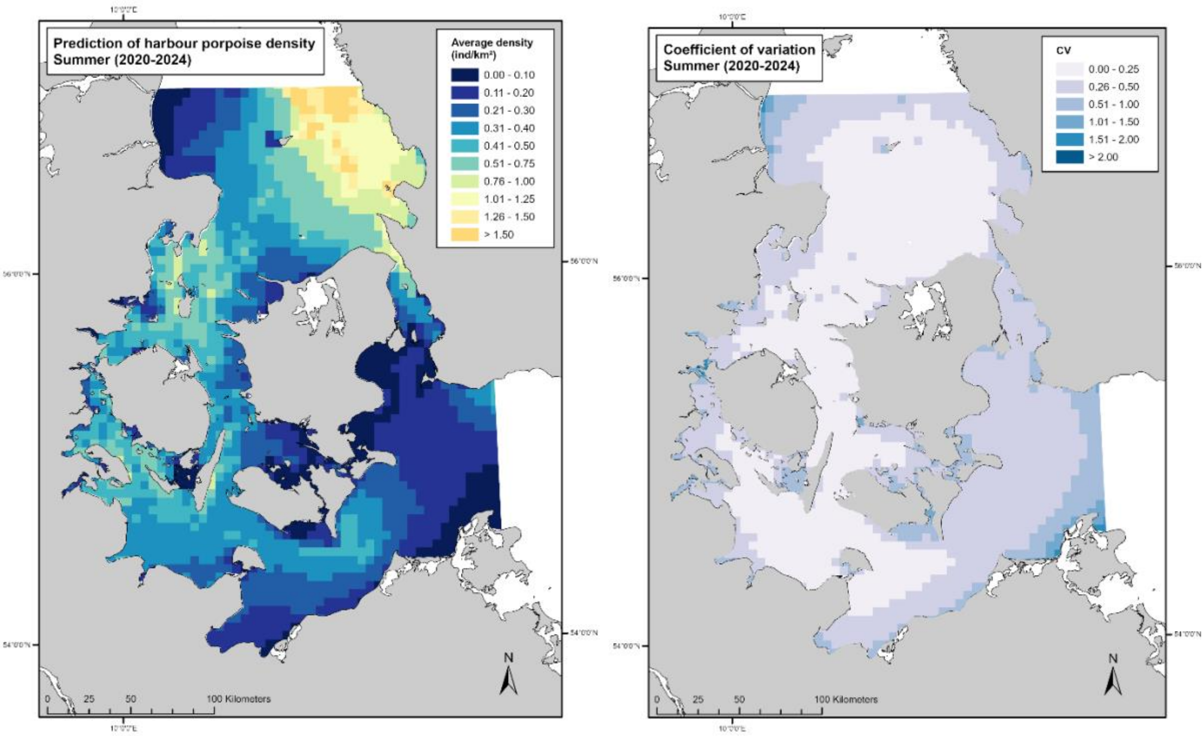


***Figure A. 2.** Predicted surface of estimated summer density [left] and associated coefficient of variation (CV) [right] for the harbour porpoise, fitting the model to survey data collected between 2020-2024 in the management unit of the Belt Sea harbour porpoise population.*

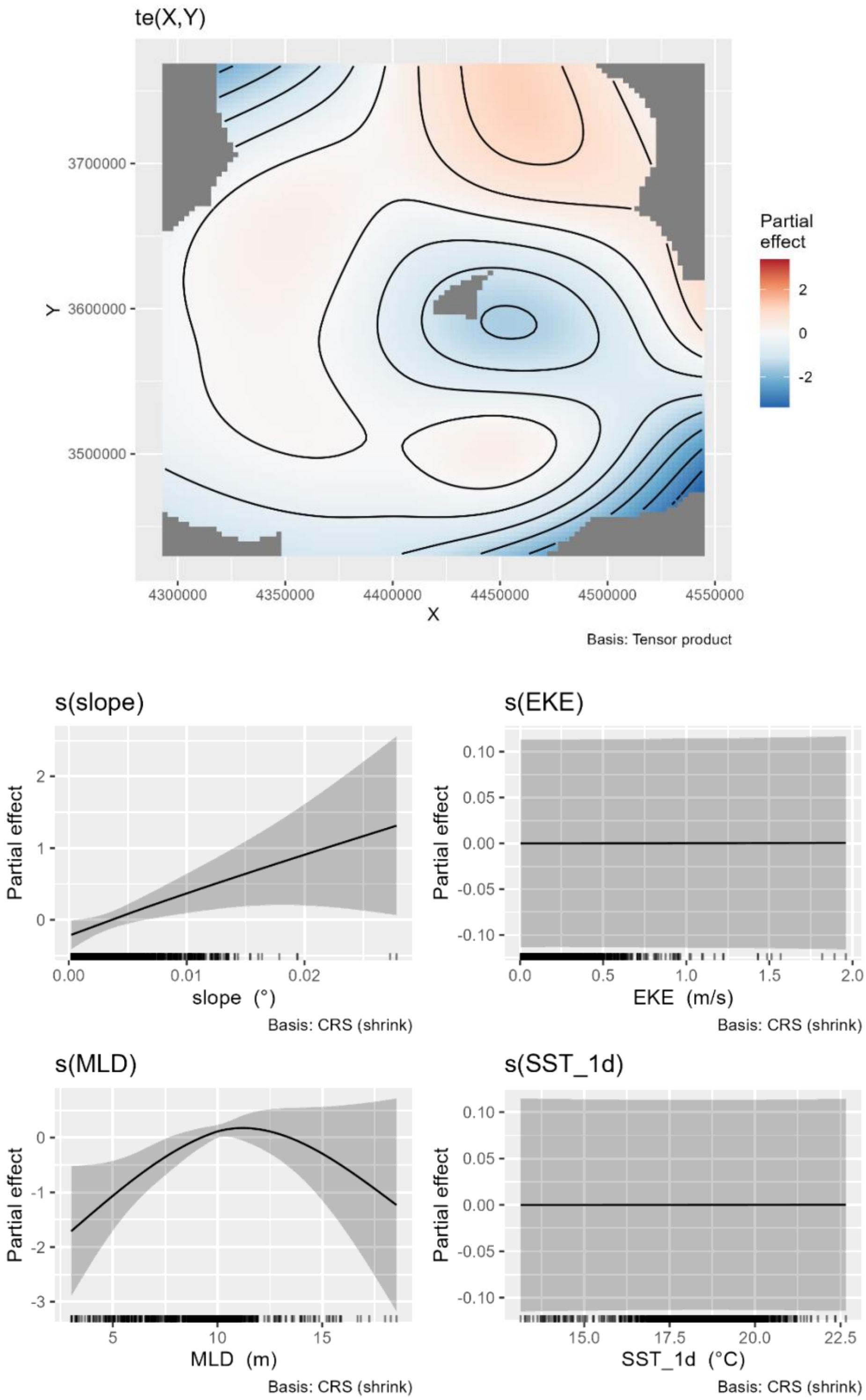


***Figure A. 3.*** *Functional plots of environmental variables relative to harbour porpoise density as indicated by the estimated smooth functions for the selected covariates in the best model for the period 2020-2024 (see* ***Table A. 1****). Y-axes represent the contribution of the term (linear or spline, on the log-scale) to predict the response variable. Zero on the y-axes corresponds to no effect of the predictor variable on the estimated response variable. Scaling of y-axis varies among predictor variables to emphasize model fit. The shading reflects 2× standard error bands (i.e., 95% confidence interval); tick marks ("rug plot") above the X-axis show data values.*

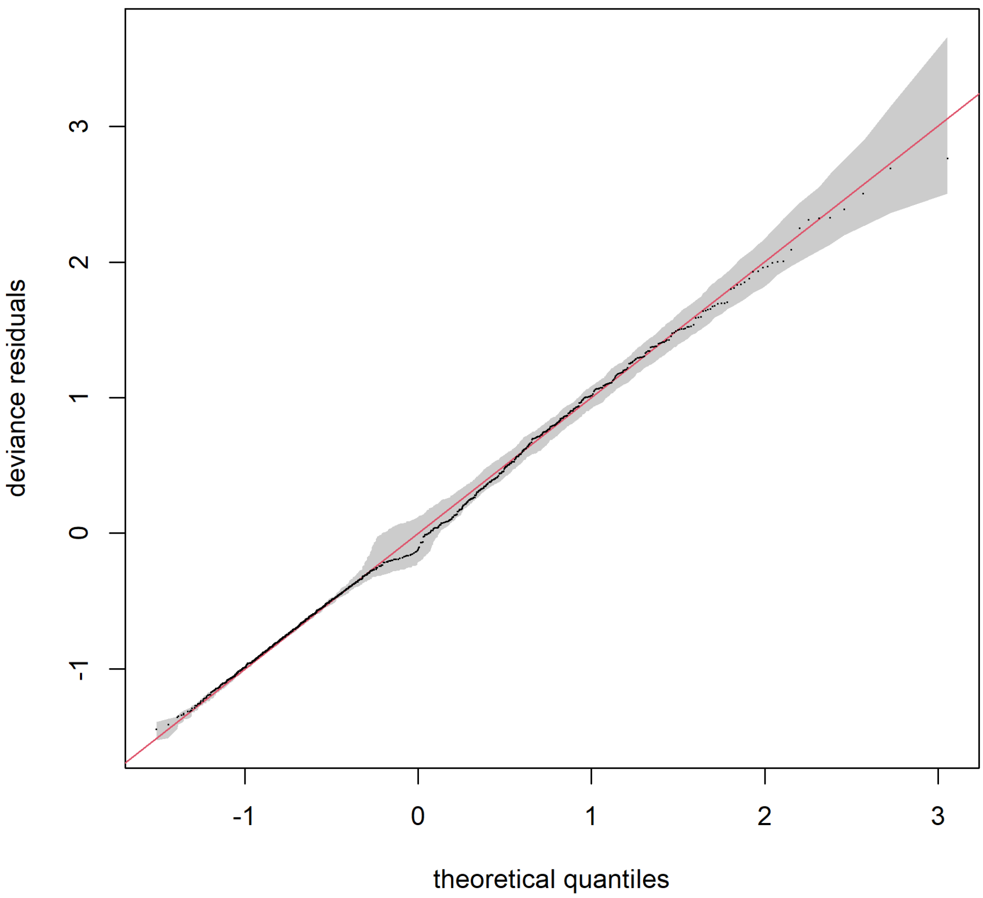


***Figure A. 4.*** *Q-Q plot for the selected final best model fitted with a negative binomial distribution for the period 2020-2024.*