A quantile regression approach to define optimal ecological niche (habitat suitability) of cockle populations (Cerastoderma edule)

Lehuen Amélie, Dancie Chloé, Grasso Florent, Orvain Francis\*

2023-10-09

Table of contents

# 1. Introduction

# 2. Materials and Methods

All data processing was conducted in R version 4.3.0 (2023-04-21 ucrt) except for MARS 3D pre-treatment in Matlab 2019a. Significance levels are p < .0001 with “\*\*\*\*”, p < .001 with “\*\*\*”, p < .01 with “\*\*”, p < .05 with “\*”.

## 2.1 Study area

Several historically known areas of the Seine estuary with different habitats or communities were studied, mainly mudflats and subtidal areas ([Figure 1](#fig-map)).

|  |
| --- |
| Figure 1: Maps of habitats areas defined in the study area. Dots represent the location of the biological samples |

## 2.2 Biological model

## 2.3 Datasets

### 2.3.1 Biological data

The raw data (n= 50,948) were harmonised and grouped in a single database which contain a total of 31,079 observations, and 187 sampling stations (with some variation in coordinates from year to year), with an average of 87 stations sampled in each campaign, mainly in Oct, Sep, Nov. A series of 5-year periods was chosen among the periods covered by the dataset, from 2000 to 2019 (the years before 2000 were discarded as they contained too few observations, n=216), with 1 to 2 sampling campaigns per year: 2000-2005, including the construction of ‘Port 2000’ which caused major disruptions in the estuary; 2006-2010; 2011-2015; 2016-2019. 627 different species are contained in the records.

### 2.3.2 Hydro-Morpho-Sedimentary data

The All Seine Bay MARS 3D model was calculated using a grid of 78,400 cells (700 by 112), focused on the Seine Estuary with a mesh of 5,184 cells (103 by 68), where the cells were on average 100m long, with the coordinates of Longitude -0.0361 to 0.3217 and Latitude 49.3382 to 49.5488. Each area defined a group with several cells (North Upstream Mudflat: n= 158, North Median Mudflat: n= 713, North Downstream Mudflat: n= 733, South Mudflat: n= 1998, Ilot Oiseaux: n= 23, Channel: n= 1290, Cote Fleurie: n= 133, OffShore: n= 2554, Octeville: n= 100).

The HMS model cells corresponding to the location of the stations in the biological dataset totalled 157 with the following distribution: Channel, n=6; Cote Fleurie, n=4; Ilot Oiseaux, n=1; North Downstream Mudflat, n=9; North Median Mudflat, n=39; North Upstream Mudflat, n=6; Octeville, n=13; OffShore, n=46; South Mudflat, n=33. The positions in the tidal area are: upper intertidal ([0-25%[ immersion time, n= 19), medium intertidal ([25-75%[ immersion time, n= 41), lower intertidal ([75-100%[ immersion time, n= 48), subtidal (100% immersion time, n= 93).

### 2.3.3 Calculation and selection of predictors

A correlation study was conducted on all HMS variables to select the most pertinent factors and to avoid autocorrelation between factors ([Figure 2](#fig-mars_corr)). All the environmental factors were calculated as annual means of model variables.

* **Daily maximum current speed** [m.s-1] was preferred to current speed [m.s-1] to reflect the hydrodynamics that occur on the mudflats due to tidal cycles (Corr = 0.73\*\*\*\*).
* **Inundation time** [%] is usually used to reflect the tidal position of benthic organisms.
* **Daily salinity range** (min-max of the day) was preferred to salinity to emphasise the influence of the tide (Corr = -0.59\*\*\*\*).
* **Temperature** [degC] was picked upon temperature daily range [degC] for its accessibility (Corr = -0.52\*\*\*\*) and the range being reflected by the inundation time (Corr = -0.75\*\*\*\*).
* **Mud content** [%] (particles <63µm) was preferred to the total concentration of sediment [kg.m-3] to reflect the condition of the habitat (Corr = -0.77\*\*\*\*).
* **Daily maximum bed shear stress** [N.m-2] is poorly correlated with bed shear stress [N.m-2] (Corr = -0.06 ) but the former was nevertheless preferred as evidence for the extreme local conditions that can lead to erosion of the sediment.
* **Bathymetry** [m] relative to the mean sea level, is correlated with several factors, but poorly with the daily salinity range (Corr = -0.08 ). It was kept for its wide accessibility in any estuary.
* MES mud [g.l-1] was not selected to focus on benthic erosion phenomena more than maximal turbidity zone (MTZ) processes even though it was not correlated to a lot of factors.

There was no correlation between biological variables (biomass and density) and any of the environmental factors. Despite the high level of correlation and significance between biomass and density (Corr = 0.75\*\*\*\*), neither of the factors was fully redundant, and the two were consequently analysed in parallel.

|  |
| --- |
| Figure 2: Correlation plot of all variables extracted from MARS3D dataset corresponding to biomass and densities from biological dataset. Correlation scores are shown above the diagonal, scatter plots with a linear regression below the diagonal, while the diagonal itself shows the density plot of each variable. |

## 2.4 Models adjustments

### 2.4.1 Quantile regression

The three model types were computed with different quantiles = [0.5, 0.9, 0.95, 0.975], 0.5 being the equivalent of an OLS regression, and kept as a reference, the other values higher than 0.9 to seek for the optimum response.

### 2.4.2 Two abiotic factor sets tested

The biological responses used were density [ind.m-2] and biomass [gAFDW.m-2] (in Ash Free Dry Weight), and were compared for each model. Sets were built with two crossed abiotic factors to more accurately describe the local environment. The sets tested were:

1. Daily max current speed [m.s-1] and inundation time [%]: These variables are easily retrieved at high frequency and enable comparison with (Cozzoli et al. 2014).They are also interesting because they contain information on the position in the tidal area that could evolve with sea level rise and information on the hydrological conditions including fluctuations in the flow rate of the river linked to climate change. A significant correlation was observed between these two variables in the HMS dataset (Corr = 0.56\*\*\*\*)
2. Daily salinity range and temperature [degC]: These factors are easily measurable at high frequency (Goberville et al., 2010) (Corr = 0.02 ). They illustrate two aspects of climate change: global warming and changes in the river regime that have an impact on the salinity profile of the estuary (Lheureux et al., 2022). Temperature also has an impact on the metabolism of the fauna and hence on their activity.
3. Daily salinity range and bathymetry [m]: Both factors are accessible at high frequency, at large scale, and can be measured remotely. They provide a good geographical description of the estuary, and are strongly affected by climate change, especially by sea level rise, marine intrusion, and changes in the river regime (Corr = -0.08 ).
4. Mud content [%] and bed shear stress [N.m-2]: These variables play a determining role in building an erosion model in the HMS model (Corr = -0.10\*). In addition, the choice relies on close links between the features of the sediment and the responses of the benthic communities (Andersen et al., 2005).

### 2.4.3 Model Validation

# 3. Results

## 3.1 Description of the biological data set

Biological data for *C. edule* comprised a total of n= 543 observations. The observations were split into periods as 2000-2005 (n= 108), 2006-2010 (n= 155), 2011-2015 (n= 174), 2015-2019 (n= 106). The following treatment only focussed on the mudflats used by *C. edule* (South Mudflat (n= 218), North Median Mudflat (n= 198), North Downstream Mudflat (n= 82), North Upstream Mudflat (n= 2)). Differences in biomass [gAFDW/m²] and density [ind/m²] are detailed as a function of the period and the different areas concerned ([Figure 3](#fig-bio_struc)). The only noticeable spatial and temporal differences concerned biomass in the south mudflat and the north (median and downstream) mudflats in the period 2000-2005.

|  |
| --- |
| Figure 3: C. edule population biomass [gAFDW/m²] and density [ind/m²] in the Seine estuary, by period for each area (A) and by areas for each period (B). Dots represent the mean and standard error of each sub-group. |

## 3.2 Description of the Hydro-Morpho-Sedimentary data set

The selected predictors were observed in the same period and in the same area as the biological data ([Figure 4](#fig-mars_map)). Generally speaking, all the factors differed significantly in area and period:

* Daily max current speed [m.s-1]: the most dynamic area was the channel, with a mean of 1.05 +/- 0.21. The northern upstream and median mudflats were subject to temporal changes in the distribution of the current, which had an impact on their global mean (upstream 0.43 +/- 0.34; median 0.63 +/- 0.3). The distribution of the north median mudflat changed significantly, bimodal in the first periods, to become unimodal at the end. From a spatial point of view, the closer to the estuary mouth, the higher the mean current speed.
* Inundation time [%]: Only the north median and upstream mudflats were truly intertidal, the upstream mudflat had higher tidal locations than the median mudflat. The bimodal tidal rhythm of the north median mudflat disappeared over time, consistent with the observed current speed. At some locations, the south mudflat had low tidal values, but overall, was mainly subtidal.
* Daily salinity range: This factor varied considerably in space and over time. Offshore, salinity varied little during the day, but there was a low second modality (around 10) at all periods. Strongly influenced by the river, salinity in the channel was 15-20 during the day, but with a significant reduction as time went by. The highly dynamic variations in salinity in the three north mudflats decreased after 2005.
* Temperature [degC]: A global increase in temperature was observed in all areas over time. The north median mudflat had the highest range, due to its intertidal location.
* Mud content [%]: Offshore and the south mudflat were made of muddy fine sands, the north, downstream, and median mudflats contained more sandy muds, whereas the composition of the sediment in the north upstream and channel areas varied widely. Mud distribution was heterogeneous in all the areas.
* Daily max bed shear stress [N.m-2]: as reflected in the current speed, the channel had the highest BSS (3.02 +/- 1.35), while the BSS in the other areas was similar (1.65 +/- 0.57). Changes in BSS in the north median mudflat over time were consistent with current speed and inundation time.
* Bathymetry [m]: The depth of the channel and offshore were similar, with a mean range of 6.94-7.95. The north downstream mudflat and the south mudflat were the next deepest areas (4.16 +/- 0.95 ), the median mudflat was 1.77 +/- 3.3, and north upstream mudflat was the shallowest, -1.59 +/- 2.63 (negative bathymetry being above the mean height of sea water). The north median mudflat was bimodal before 2000, with some shallow areas and some more intertidal, but then changed, losing the upper areas and becoming unimodal.

### Map

|  |
| --- |
| Figure 4: Mars3D selected factors maps in the Seine estuary |

### Density plots

|  |
| --- |
| Figure 5: Mars3D selected factors density plot in the Seine estuary |

#### daily maximum current speed (m.s-1)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### inundation time (%)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### daily salinity range (u.s.i)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### temperature (degC)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### mud content (%)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### daily maximum bed shear stress (N.m-2)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

#### bathymetry (m)

|  |
| --- |
| Mars3D selected factors distribution in the Seine estuary, by Period for each area (A) and by areas for each Period (B) |

## 3.3 Methodology assessment

All three models were used to calculate the SDM for each set at the different quantiles chosen. A general comparison of ∆AICc scores was undertaken for all the SDMs computed ([Figure 6](#fig-aic_3mod)). The SDM scores based on density were significantly higher than those based on biomass, and the choice of the quantile had a strong impact on the score. For instance, the best scores were obtained for the biomass SDMs with the 0.5 quantile, which would not help describe the optimum ecological niche. On average, the BSpline model ∆AICc were lower than the others (RQ2bsp = 2630 +/- 1648, RQ2nli = 2723 +/- 1716, RQ2int = 2747 +/- 1699). With the same biological response and quantile (Biomass and tau=0.95), the variations between SDM ∆AICcs were quite low with respect to total variability (RQ2bsp = 1483 +/- 61, RQ2nli = 1539 +/- 62, RQ2int = 1587 +/- 55).

The predicted/observed plots ([Figure 7](#fig-rq_po_brd) shows an example from set A), completed the observations of ∆AICc, i.e. RQ2bsp > RQ2nli > RQ2int (the regression lines of each quantile were closer to the 1:1 line). However, RQ2nli performed better than RQ2bsp at some quantiles, in particular at the 95th centile, which was defined as the optimum quantile, i.e. the highest quantile that did not affect model quality, and was hence used for all subsequent analyses.

Based on the range of both predictors under each set, the quantile responses of SDMs were calculated and illustrated by a surface plot ([Figure 8](#fig-3mod) shows an example from set A, interactive 3D plots below). Although the performance indicators were good, graphically, the RQ2bsp model showed overfitting which prevented both modelling and prediction. This model was thus not used for any more analysis.

|  |
| --- |
| Figure 6: ∆AICc comparison for all SDMs computed, according to the quantile, the type of model and the response. |

|  |
| --- |
| Figure 7: Example of modelled vs observed data plotted for each model based on biomass under set A. The black line represents the 1:1 ratio. |

|  |
| --- |
| Figure 8: Examples of SDM surface plots under set A: linear with interaction (A), Gaussian non-linear (B) and Cubic B-Spline linear (C). The upper panel shows the surfaces side by side, the lower panel shows the 3D plots with all processed quantiles superposed. The biological data observed with this model are represented by a contour plot, the data over the model are represented by red stars. |

## 3.4 Species Distribution Models – Optimal Ecological Niche (SDM)

### 3.4.1 Comparison of linear and nonlinear Quantile Regression

The four sets (A, B, C and D) were treated with one of the two selected models: either linear with interaction, or non-linear with a bifactorial Gaussian equation, with biomass as biological response (with predict/observed plot, SDMs with the distribution of HMS data, RQ2nli SDMs were represented in 3D graph in **?@sec-sdm\_3d**).

1. Daily max current speed [m.s-1] and inundation time [%]:
   1. RQ2int optimum was 94.82gAFDW/m², at 0 m.s-1 and 1 %; RQ2nli optimum was 233gAFDW/m² at 0 m.s-1 and 1 %.
   2. RQ2int optimum was 2129.29ind/m² at 0 m.s-1 and 1 %; RQ2nli optimum was 2195.67ind/m² at 0.43 m.s-1 and 1 %.
2. Daily salinity range and temperature [degC]:
   1. RQ2int optimum was 68.96gAFDW/m² at 0.2 u.s.i and 13.01 degC; RQ2nli optimum was 91.35gAFDW/m² at 0.2 u.s.i and 12.35 degC.
   2. RQ2int optimum was 1962.03ind/m² at 24.4 u.s.i and 13.01 degC; RQ2nli optimum was 980.01ind/m² at 0.2 u.s.i and 13.01 degC.
3. Daily salinity range and bathymetry [m]:
   1. RQ2int optimum was 172.64gAFDW/m² at 0.2u.s.i and 13.95 m; RQ2nli optimum was 147.6gAFDW/m² at 0.2 u.s.i and 5.14 m.
   2. RQ2int optimum was 3792.69ind/m² at 0.2 u.s.i and 13.95 m; RQ2nli optimum was 4160.3ind/m² at 0.2 u.s.i and 9.02 m.
4. Mud content [%] and bed shear stress [N.m-2]:
   1. RQ2int optimum was 104.58gAFDW/m² at 1 % and 0 N.m-2; RQ2nli optimum wasf 106.75gAFDW/m² at 0.35 % and 1.33 N.m-2.
   2. RQ2int optimum was 2990.53ind/m² at 1 % and 0 N.m-2; RQ2nli optimum was 1545.54ind/m² at 1 % and 0.44 N.m-2.

#### Biomass

|  |
| --- |
| Sets computed with linear model (top row, numbered 1) and non-linear with a? the? Gaussian equation (bottom row, numbered 2), the observed biological data under the model surface are represented by an isometric curve, the data over the model are represented by red stars. Each pair has its own scale to ensure visibility |

#### Density

|  |
| --- |
| Sets computed with linear model (top row, numbered 1) and non-linear with a? the? Gaussian equation (bottom row, numbered 2), the observed biological data under the model surface are represented by an isometric curve, the data over the model are represented by red stars. Each pair has its own scale to ensure visibility |

#### Biomass

|  |
| --- |
| Modelled vs observed data plot for RQ2int and RQ2nli SDM. Black line represents the 1:1 ratio |

#### Density

|  |
| --- |
| Modelled vs observed data plot for RQ2int and RQ2nli SDM. Black line represents the 1:1 ratio |

#### Biomass

|  |
| --- |
| Sets computed with linear with interaction (top, numbered 1) and nonlinear with gaussian equation (bottom numbered 2), the contour plot represents the observed HMS data |

#### Density

|  |
| --- |
| Sets computed with linear with interaction (top, numbered 1) and nonlinear with gaussian equation (bottom numbered 2), the contour plot represents the observed HMS data |

### 3.4.2 Non-linear quantile regression with bifactorial Gaussian equation models

Overall, the non-linear model performed better than the linear model, and the density-based models were generally less relevant than models based on biomass. Thus, the RQ2nli for biomass was the only model geographically applied and analysed. Each SDM for the RQ2nli model was applied on the HMS model web over the estuary, for each period ([Section 3.4.2.1](#sec-map)). The suitability index is given per period and area ([Section 3.4.2.2](#sec-suitability)) and density plots were drawn for each SDM, focusing on mudflats to facilitate comparison ([Section 3.4.2.3](#sec-densplot)).

1. Daily max current speed [m.s-1] and inundation time [%]:The channel and north mudflats were the least favourable areas, the south mudflats and offshore were more appropriate, but few locations were really optimum. These locations were in the north of the offshore area, corresponding to Le Havre beach, and Cote Fleurie, which were not known as optimal. The suitability index, which ranged from 0.1 to 0.5 and was generally stable, confirmed that the most suitable areas were the Cote Fleurie followed by the subtidal Octeville and offshore, the two latter being very similar. The suitability of the north median and upstream mudflats improved after 2005, when they became better than the channel. The north upstream mudflat had averages with large confidence interval. Regarding the density plots, the south and north downstream mudflats were subtidal, the south having a lower current speed. The maximum occurrence current speed in this model was 50-60gAFDW/m² for the south, and 35-50gAFDW/m² for the north. Both varied little over time, and the observed quantiles were consistent only with the south mudflat model during the periods 2006-2010 and 2011-2015, the others being either higher or lower. The north median and upstream mudflat had a wider range of immersion times, which was reflected in the results of the model. The model distribution also revealed the bimodal aspect of the current in the north median mudflat.
2. Daily salinity range and temperature [°C]: The salinity part of the model had a noticeable effect on the result of the model, with a clear reduction in biomass in the river and its ETM area. The closer the estuary entrance to the sea, the higher the model, the offshore area having a clear advantage that increased from 2005 on. The suitability index ranged from 0 to 0.8, and with a general increase over time. The three north mudflats underwent a significant increase during the first three periods. Octeville and offshore were dissimilar with respect to this SDM, together with Cote Fleurie, Octeville was the most suitable. The same increase was visible for the four mudflats on the density plots, all of which shifted roughly from 20 to 60gAFDW/m². The profile of all the mudflats mainly appeared to be influenced by the salinity range, the variation in salinity in the period 2000-2005 was clearly the main driver of the changes in profile. There was a very noticeable change in the north downstream and median mudflats from being highly influenced by inputs of fresh water to a more marine environment. The quantiles of observed data were generally lower than the models, but displayed the same temporal trends.
3. Daily salinity range and bathymetry [m]: The optima for this model were clearly located on Cote Fleurie, the south mudflat and the shore of Octeville, results linked to the bathymetry. However, Cote Fleurie is known to host a meagre cockle population. Offshore, a spot was high in the period 1996-1999 caused by dumping material dredged from the channel that was subsequently progressively smoothed. Except in Cote Fleurie, suitability was less than 0.5, the south mudflat being the second best. Apart from the north median and downstream mudflats which increased over the first three periods, the indexes for each area remained steady. The north downstream mudflat became more favourable for cockles over time, linked to the smaller salinity range, and to the bathymetry progressively becoming bimodal. The north median mudflat showed a widening of the results, perhaps due to the reduced salinity range or to the reduction in the very high upper tidal limit in favour of medium intertidal areas.
4. Mud content [%] and bed shear stress [Pa]: This model result was very patchy at the scale of the estuary due to the equally patchy distribution of mud. This identified a channel with high biomass potential, which did not agree with expert knowledge. There was also an area with high biomass around the borders of the ETM area offshore which decreased over time, also due to the reduced mud content. Suitability ranged from 0.25 to 0.75, with a complex pattern. The suitability of the channel, offshore and of the north downstream mudflat decreased after 2005, while the suitability of Octeville and the north upstream and median mudflats improved. The heterogeneity of the areas was also visible in the density plots, as each area had wide distribution with some local maxima that were difficult to link to physical descriptors.

#### 3.4.2.1 Application maps

#### Biomass

|  |
| --- |
| SDM models applied on the Seine estuary over the five periods. |

#### Density

|  |
| --- |
| SDM models applied on the Seine estuary over the five periods. |

#### 3.4.2.2 Suitability index

#### Biomass

|  |
| --- |
| Suitability index per period and per area for all SDM models with a 95% confidence interval on means. |

#### Density

|  |
| --- |
| Suitability index per period and per area for all SDM models with a 95% confidence interval on means. |

#### 3.4.2.3 Application density plots

#### Biomass

|  |
| --- |
| Density plot of the two predictors and the SDM model result in each area and in each period. The dots are the 95th quantile of the observed data. |

#### Density

|  |
| --- |
| Density plot of the two predictors and the SDM model result in each area and in each period. The dots are the 95th quantile of the observed data. |

#### 3.4.2.4 Application density plot details

The four models result on the estuary regarding Areas and Periods, with the two calculation modes, for both responses.

##### 3.4.2.4.1 Biomass

#### RQ2nli\_11112

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11134

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11137

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11156

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

##### 3.4.2.4.2 Density

#### RQ2nli\_11212

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11234

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11237

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

#### RQ2nli\_11256

|  |
| --- |
| Density plot of the two predictors and the SDM model result on each area and period. The dots were the 95th quantile of the observed data. |

# 4. Supplementary data

## 4.1 Session information

|  |
| --- |
| Expand for R Session Info |
| ─ Session info ───────────────────────────────────────────────────────────────  setting value  version R version 4.3.0 (2023-04-21 ucrt)  os Windows 10 x64 (build 19045)  system x86\_64, mingw32  ui RTerm  language EN  collate French\_France.utf8  ctype French\_France.utf8  tz Europe/Paris  date 2023-10-09  pandoc 3.1.1 @ C:/Program Files/RStudio/resources/app/bin/quarto/bin/tools/ (via rmarkdown)  quarto 1.3.361 @ C:\\Users\\LEHUEN~1\\AppData\\Local\\Programs\\Quarto\\bin\\quarto.exe  ─ Packages ───────────────────────────────────────────────────────────────────  package \* version date (UTC) lib source  beepr \* 1.3 2018-06-04 [1] CRAN (R 4.3.0)  boot \* 1.3-28.1 2022-11-22 [1] CRAN (R 4.3.0)  colorspace \* 2.1-0 2023-01-23 [1] CRAN (R 4.3.0)  conflicted \* 1.2.0 2023-02-01 [1] CRAN (R 4.3.0)  data.table \* 1.14.8 2023-02-17 [1] CRAN (R 4.3.0)  dplyr \* 1.1.2 2023-04-20 [1] CRAN (R 4.3.0)  figpatch \* 0.2 2022-05-03 [1] CRAN (R 4.3.0)  forcats \* 1.0.0 2023-01-29 [1] CRAN (R 4.3.0)  furrr \* 0.3.1 2022-08-15 [1] CRAN (R 4.3.0)  future \* 1.32.0 2023-03-07 [1] CRAN (R 4.3.0)  GGally \* 2.1.2 2021-06-21 [1] CRAN (R 4.3.0)  gginnards \* 0.1.2 2023-05-24 [1] CRAN (R 4.3.0)  ggplot2 \* 3.4.2 2023-04-03 [1] CRAN (R 4.3.0)  ggpubr \* 0.6.0 2023-02-10 [1] CRAN (R 4.3.0)  ggridges \* 0.5.4 2022-09-26 [1] CRAN (R 4.3.0)  ggsci \* 3.0.0 2023-03-08 [1] CRAN (R 4.3.0)  grafify \* 3.2.0 2023-04-29 [1] CRAN (R 4.3.0)  gt \* 0.9.0 2023-03-31 [1] CRAN (R 4.3.0)  Hmisc \* 5.1-0 2023-05-08 [1] CRAN (R 4.3.0)  htmlwidgets \* 1.6.2 2023-03-17 [1] CRAN (R 4.3.0)  introdataviz \* 0.0.0.9003 2023-06-15 [1] Github (psyteachr/introdataviz@0519c98)  kableExtra \* 1.3.4 2021-02-20 [1] CRAN (R 4.3.0)  knitr \* 1.43 2023-05-25 [1] CRAN (R 4.3.0)  lubridate \* 1.9.2 2023-02-10 [1] CRAN (R 4.3.0)  openxlsx \* 4.2.5.2 2023-02-06 [1] CRAN (R 4.3.0)  patchwork \* 1.1.2 2022-08-19 [1] CRAN (R 4.3.0)  plot3D \* 1.4 2021-05-22 [1] CRAN (R 4.3.0)  plotly \* 4.10.2 2023-06-03 [1] CRAN (R 4.3.0)  purrr \* 1.0.1 2023-01-10 [1] CRAN (R 4.3.0)  quantreg \* 5.95 2023-04-08 [1] CRAN (R 4.3.0)  r2resize \* 1.6 2023-04-07 [1] CRAN (R 4.3.0)  RColorBrewer \* 1.1-3 2022-04-03 [1] CRAN (R 4.3.0)  readr \* 2.1.4 2023-02-10 [1] CRAN (R 4.3.0)  readxl \* 1.4.2 2023-02-09 [1] CRAN (R 4.3.0)  reshape2 \* 1.4.4 2020-04-09 [1] CRAN (R 4.3.0)  rlist \* 0.4.6.2 2021-09-03 [1] CRAN (R 4.3.0)  rstatix \* 0.7.2 2023-02-01 [1] CRAN (R 4.3.0)  scales \* 1.2.1 2022-08-20 [1] CRAN (R 4.3.0)  sessioninfo \* 1.2.2 2021-12-06 [1] CRAN (R 4.3.0)  sf \* 1.0-13 2023-05-24 [1] CRAN (R 4.3.0)  sfheaders \* 0.4.2 2023-03-03 [1] CRAN (R 4.3.0)  SparseM \* 1.81 2021-02-18 [1] CRAN (R 4.3.0)  stringr \* 1.5.0 2022-12-02 [1] CRAN (R 4.3.0)  tibble \* 3.2.1 2023-03-20 [1] CRAN (R 4.3.0)  tidyr \* 1.3.0 2023-01-24 [1] CRAN (R 4.3.0)  tidyverse \* 2.0.0 2023-02-22 [1] CRAN (R 4.3.0)  tmap \* 3.3-3 2022-03-02 [1] CRAN (R 4.3.0)  tmaptools \* 3.1-1 2021-01-19 [1] CRAN (R 4.3.0)  wesanderson \* 0.3.6 2018-04-20 [1] CRAN (R 4.3.0)   [1] C:/Users/lehuen201/AppData/Local/Programs/R/R-4.3.0/library  ────────────────────────────────────────────────────────────────────────────── |

The Scientific colour map batlow (Crameri 2018) is used in this study to prevent visual distortion of the data and exclusion of readers with colour­vision deficiencies (Crameri, Shephard, and Heron 2020).

# References

Cozzoli, Francesco, Menno Eelkema, Tjeerd J. Bouma, Tom Ysebaert, Vincent Escaravage, and Peter M. J. Herman. 2014. “A Mixed Modeling Approach to Predict the Effect of Environmental Modification on Species Distributions.” *PLoS ONE* 9 (2): e89131. <https://doi.org/10.1371/journal.pone.0089131>.

Crameri, Fabio, Grace E. Shephard, and Philip J. Heron. 2020. “The Misuse of Colour in Science Communication.” *Nature Communications* 11 (1): 5444. <https://doi.org/10.1038/s41467-020-19160-7>.