

Data stochasticity and model parametrisation impact the performance of species distribution models: insights from a simulation study

Supplementary Information

Charlotte Lambert^{*1} and Auriane Virgili²

¹Littoral ENvironnement et Sociétés UMR 7266 CNRS-LRUniv, 2 Rue Olympe de Gouges, 17000 La Rochelle, France

²Observatoire Pelagis UAR 3462 CNRS-LRUniv, 5 allée de l'Océan, 17000 La Rochelle, France

1 Environmental conditions

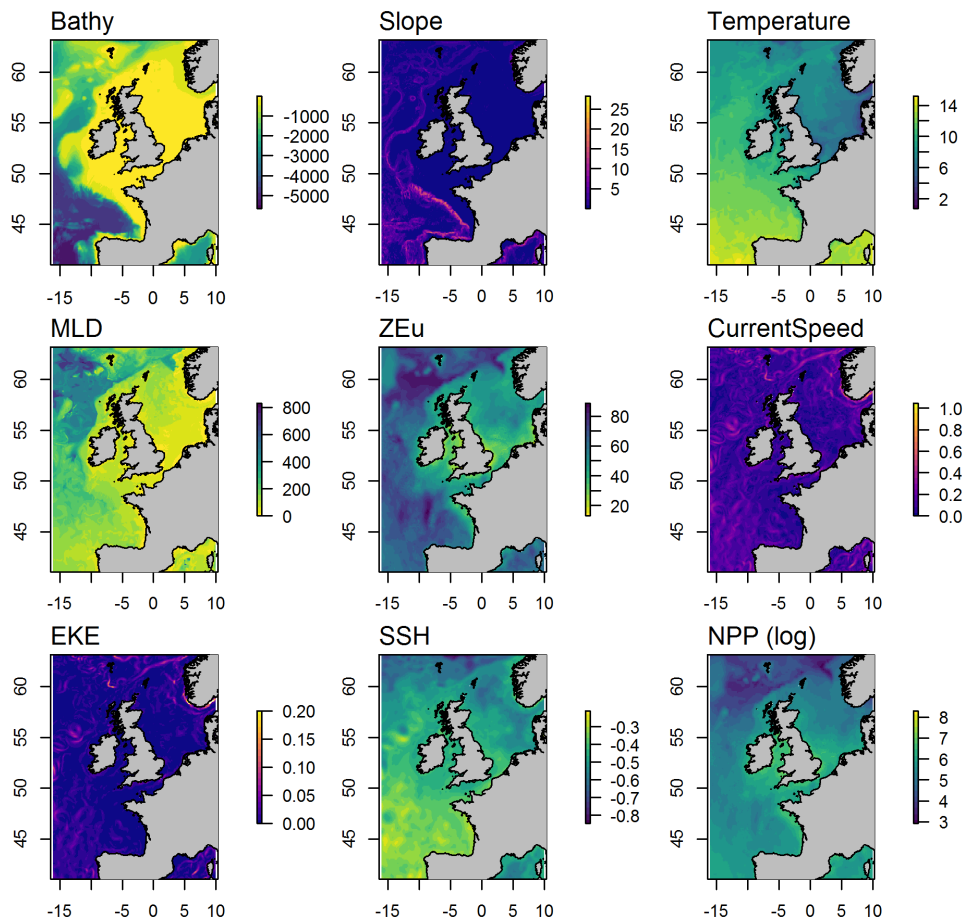


Figure S1-1. Maps of environmental conditions used to simulate species distribution and perform the habitat modelling, for the Eastern North Atlantic region. Bathy: bathymetry; MLD: mixed layer depth; ZEu: euphotic depth; EKE: eddy kinetic energy; SSH: sea surface height; NPP: net primary production. See main text for details.

^{*}Corresponding author, electronic address: charlotte.lambert@univ-lr.fr

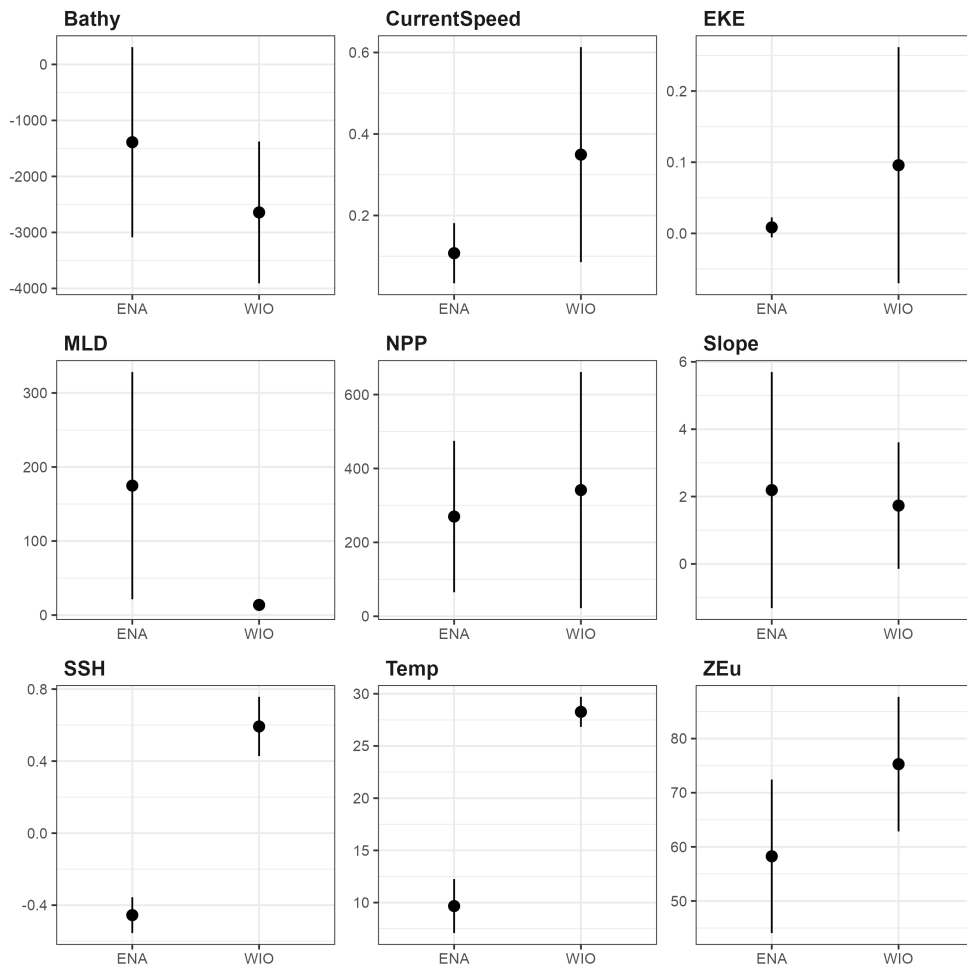


Figure S1-2. Heterogeneity of environmental conditions in the Eastern North Atlantic (ENA) and Western Indian Ocean (WIO), as expressed by mean (dot) and standard deviations (line) of every condition across each study area. Bathy: bathymetry; MLD: mixed layer depth; ZEu: euphotic depth; EKE: eddy kinetic energy; SSH: sea surface height; NPP: net primary production. See text for details.

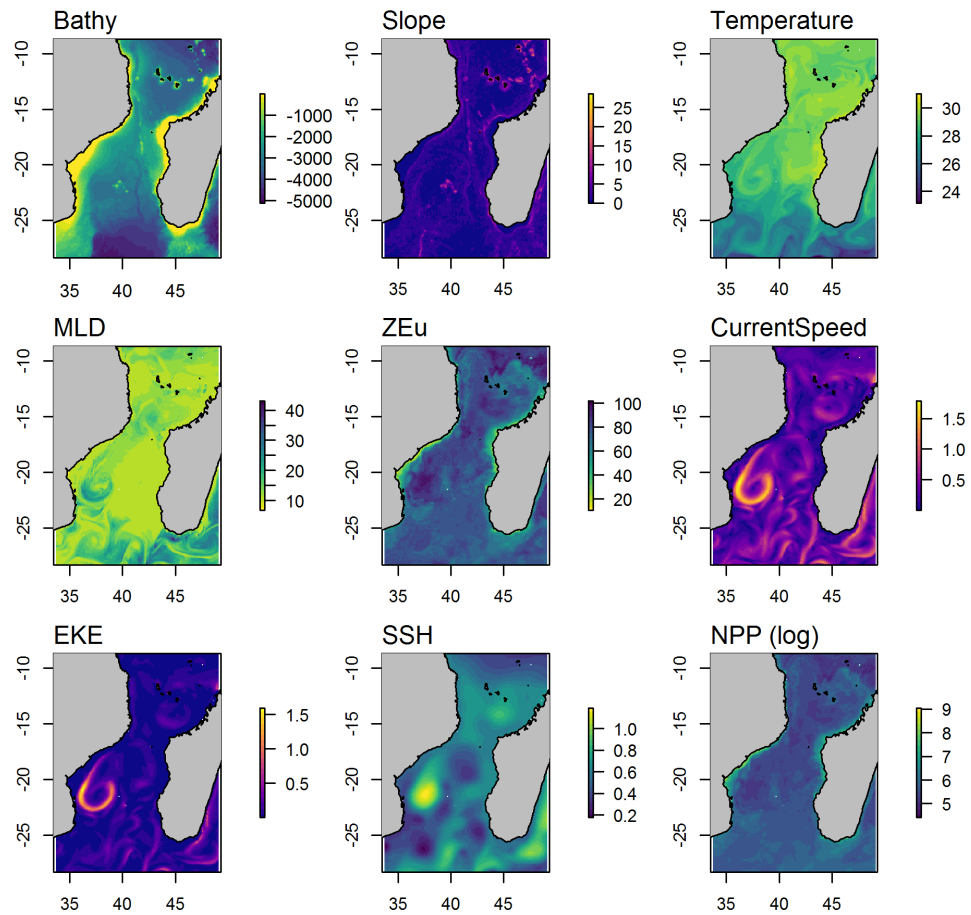


Figure S1-3. Maps of environmental conditions used to simulate species distribution and perform the habitat modelling, for the Western Indian Ocean region. Bathy: bathymetry; MLD: mixed layer depth; ZEu: euphotic depth; EKE: eddy kinetic energy; SSH: sea surface height; NPP: net primary production. See main text for details.

2 Data processing

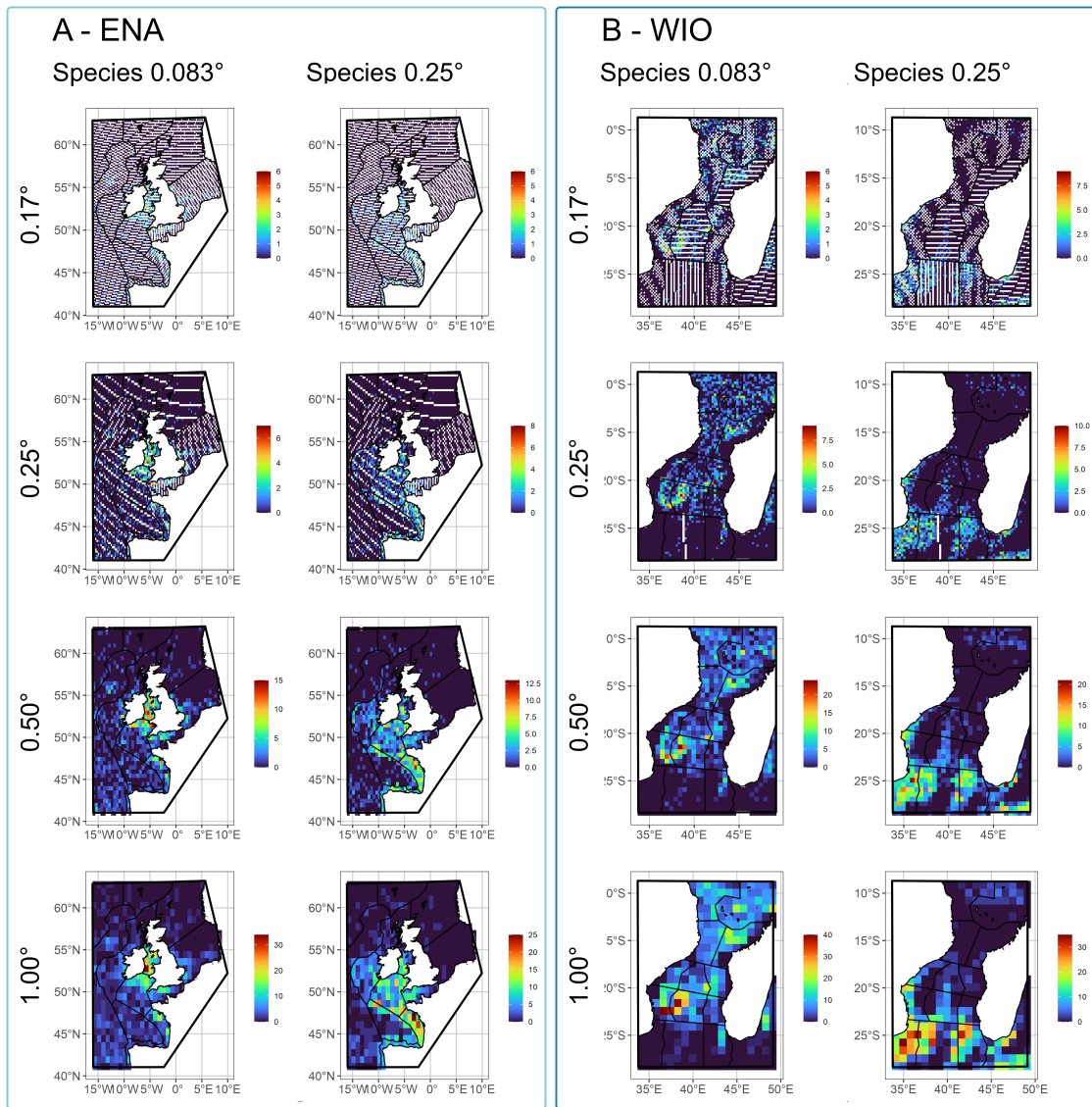


Figure S1-4. Observation maps rasterized to the 0.17, 0.25, 0.50 and 1.0° resolutions for both species in the (A) ENA and (B) WIO regions. Color ramps are for the number of individuals sighted per cell.

3 Modelling

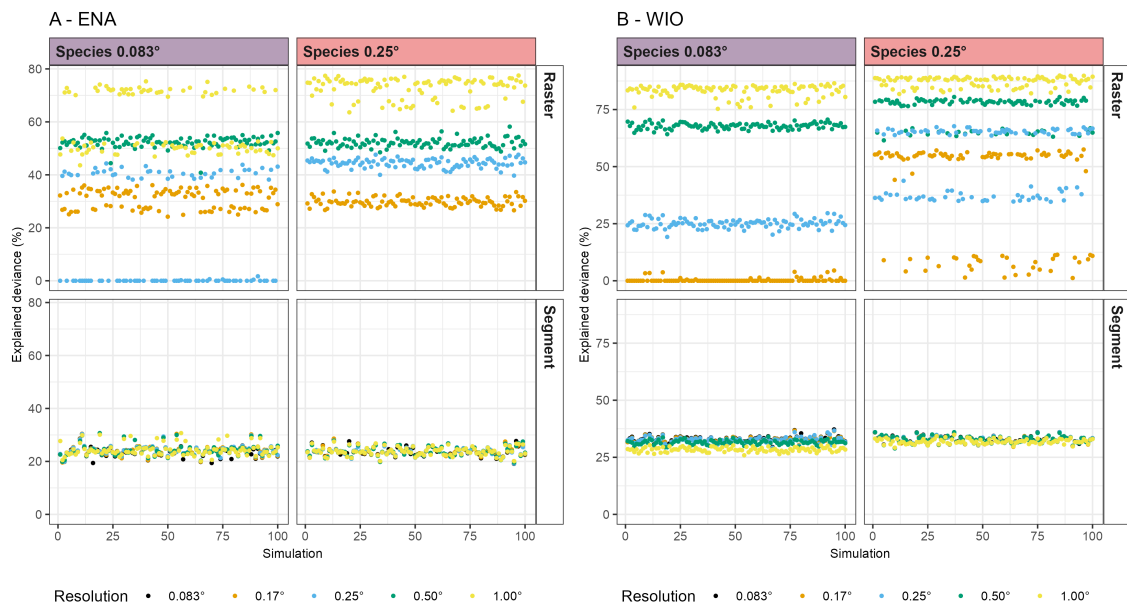


Figure S1-5. Explained deviances (%) of selected best models by sampling type, sampling resolution and species, in the ENA (A) and WIO (B). Each point is for a simulation.

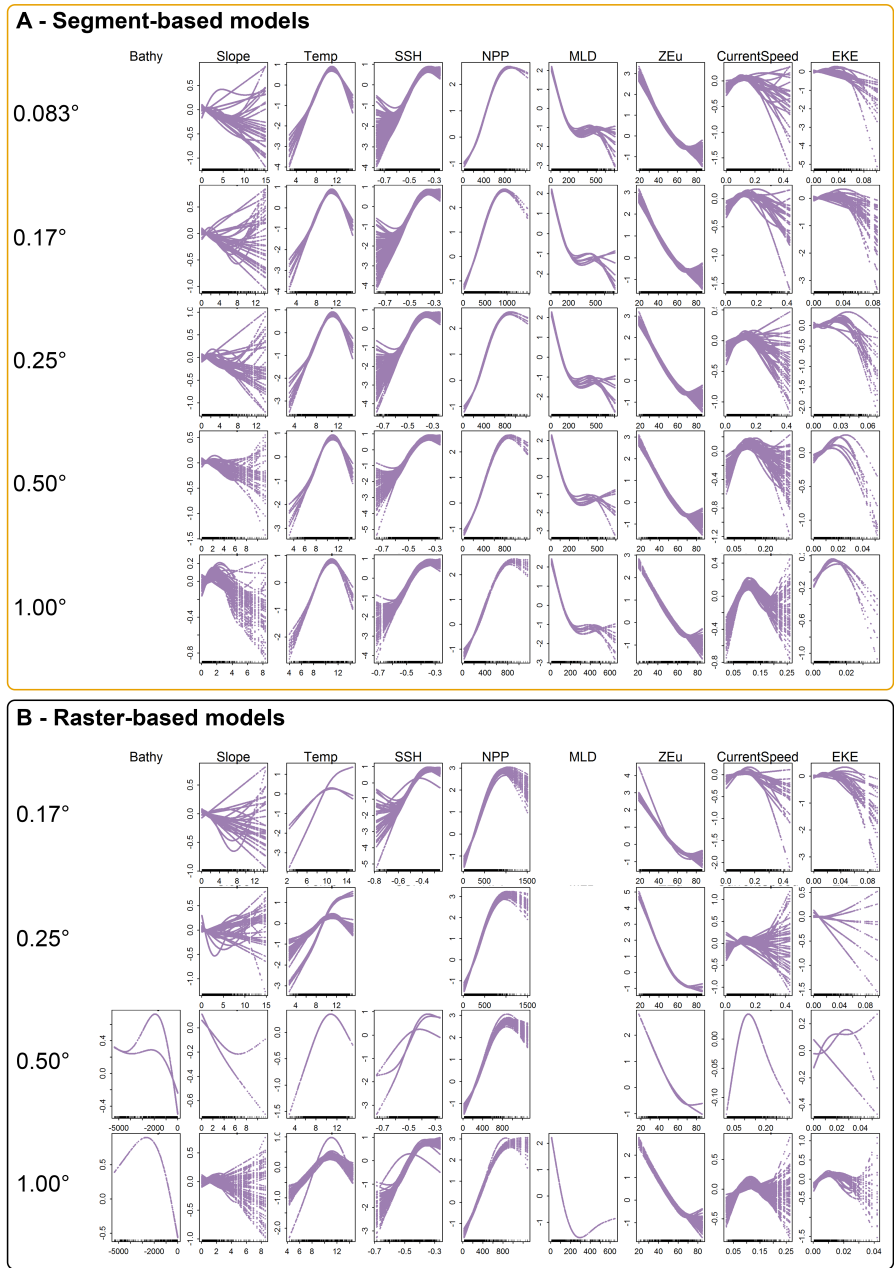


Figure S1-6. Relationships between the linear predictor ($\log(\text{individuals})$, y axis) and the environmental variables (x axis) for selected best models for the 0.083° species in the ENA, for segment-based models (above) and raster-based models (below). Each curve is for a simulation; the number of curves in each panel equals the number of times the variable was selected in the best models.

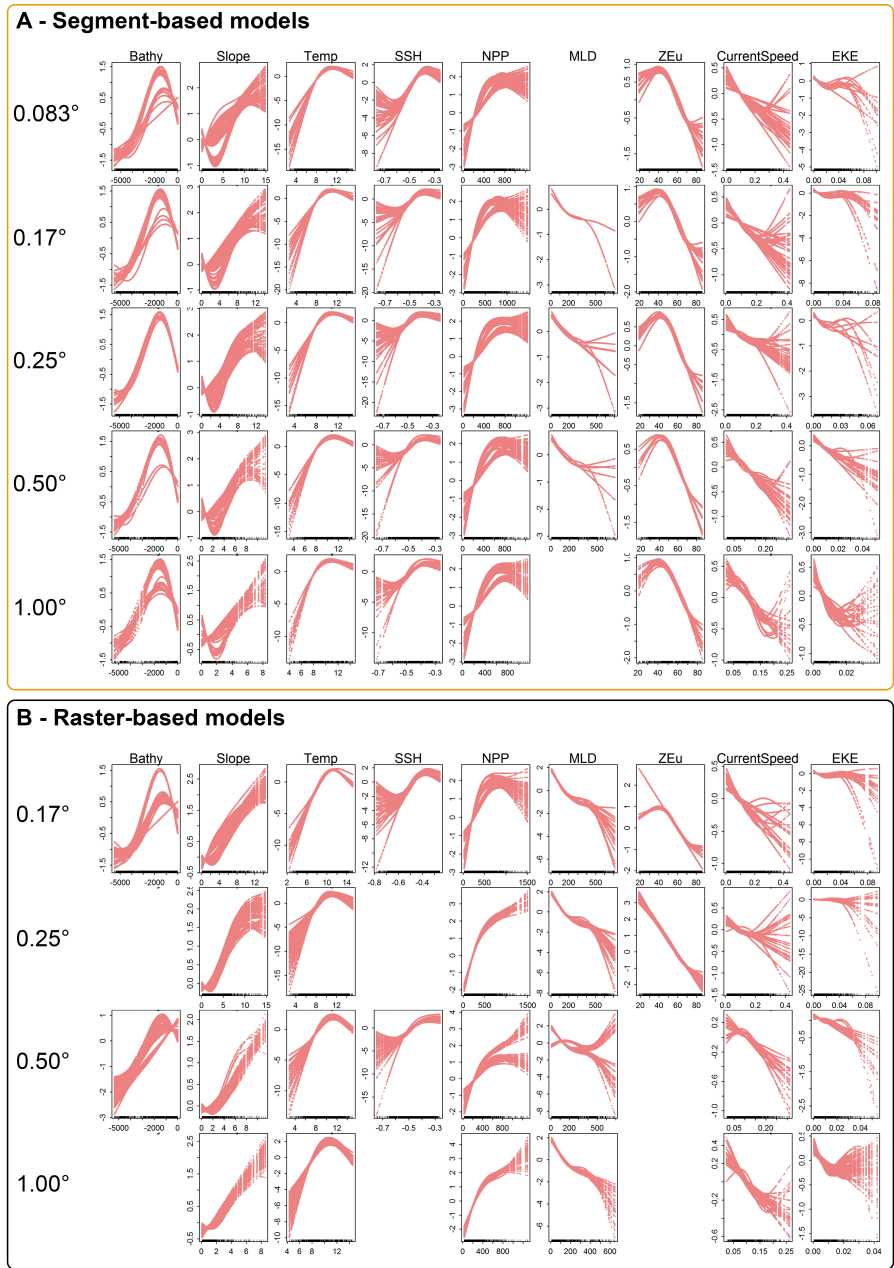


Figure S1-7. Relationships between the linear predictor ($\log(\text{individuals})$, y axis) and the environmental variables (x axis) for selected best models for the 0.25° species in the ENA, for segment-based models (above) and raster-based models (below). Each curve is for a simulation; the number of curves in each panel equals the number of times the variable was selected in the best models.

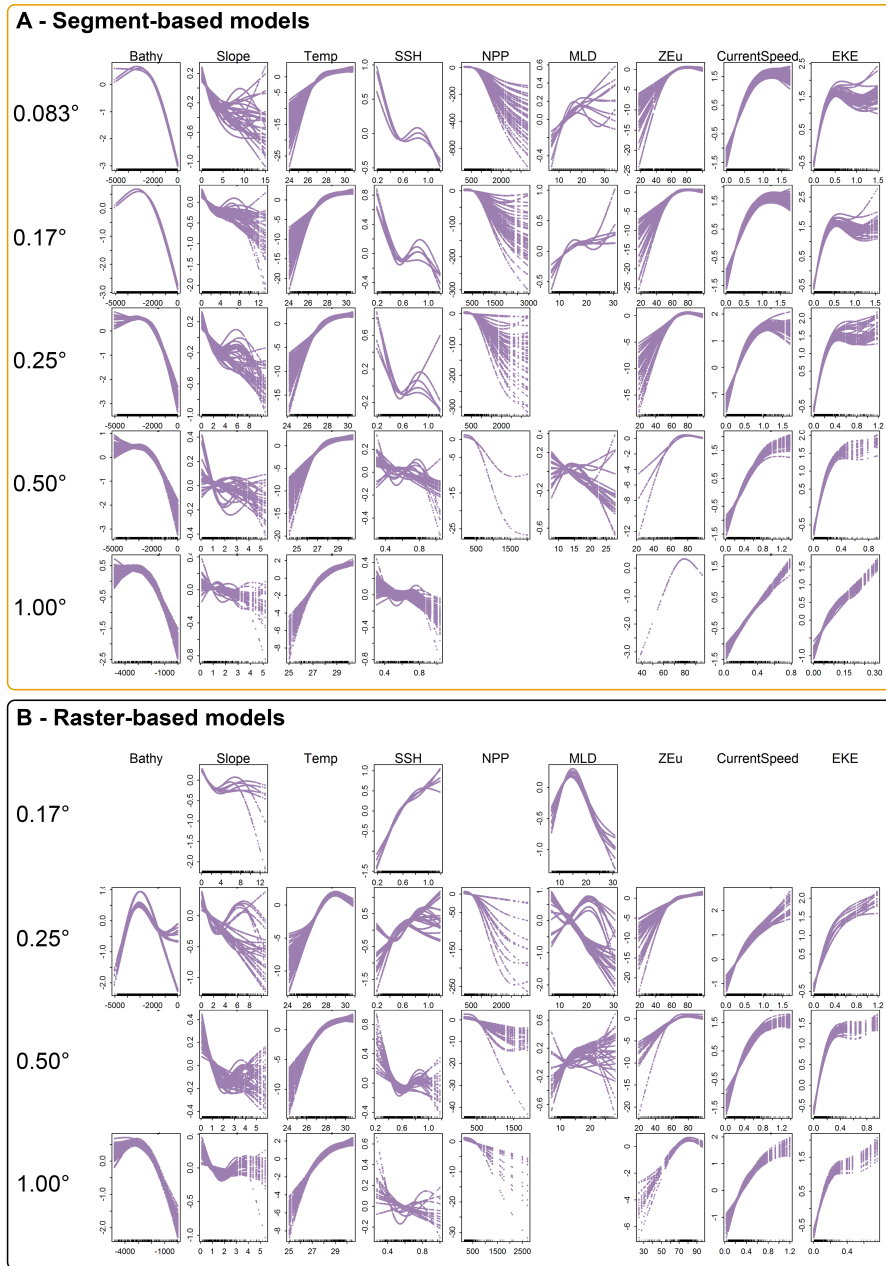


Figure S1-8. Relationships between the linear predictor ($\log(\text{individuals})$, y axis) and the environmental variables (x axis) for selected best models for the 0.083° species in the WIO, for segment-based models (above) and raster-based models (below). Each curve is for a simulation; the number of curves in each panel equals the number of times the variable was selected in the best models.

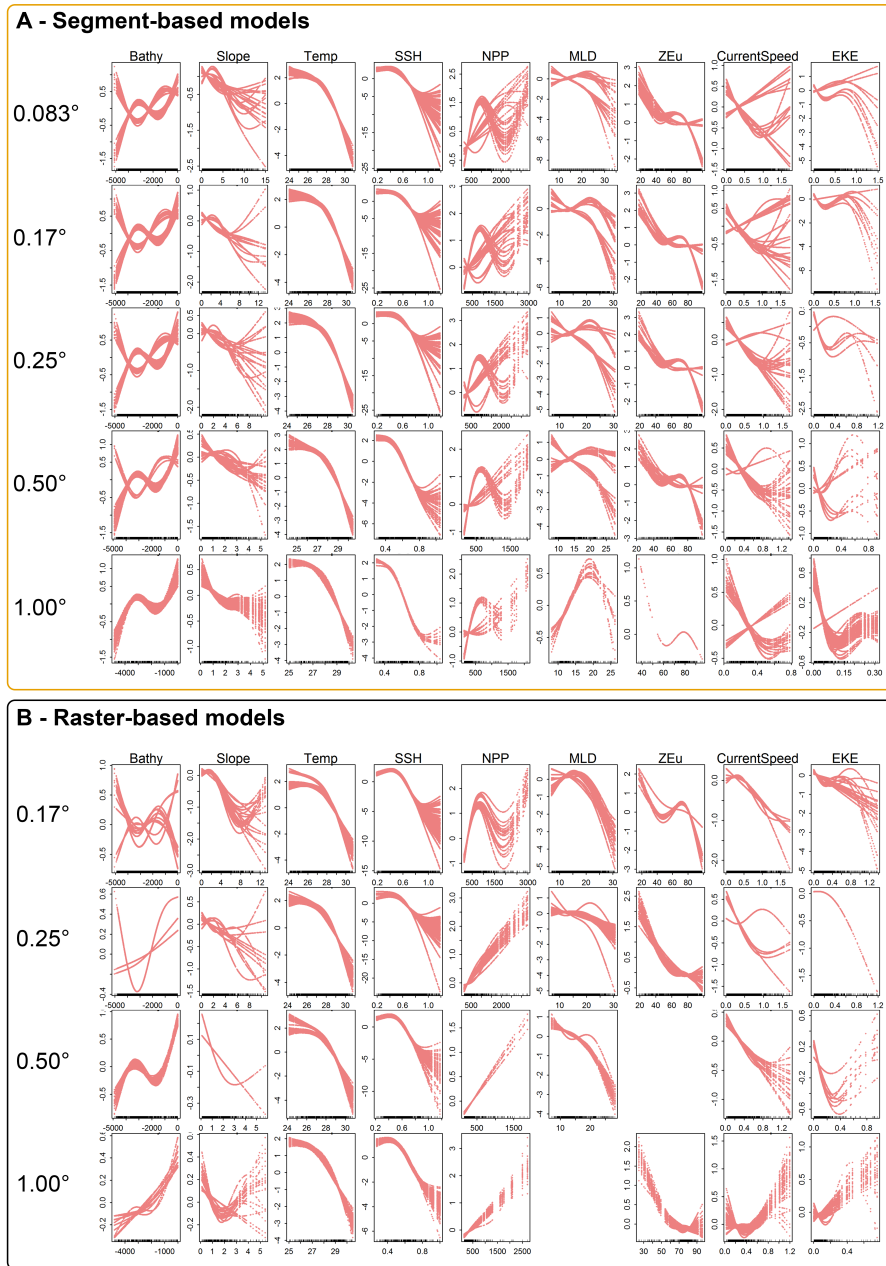


Figure S1-9. Relationships between the linear predictor ($\log(\text{individuals})$, y axis) and the environmental variables (x axis) for selected best models for the 0.25° species in the WIO, for segment-based models (above) and raster-based models (below). Each curve is for a simulation; the number of curves in each panel equals the number of times the variable was selected in the best models.