The following supplements accompany the article

Potential feeding habitat of fin whales in the western Mediterranean Sea: an environmental niche model

Jean-Noël Druon[1,*], Simone Panigada[2], Léa David[3], Alexandre Gannier[4], Pascal Mayol[5], Antonella Arcangeli[6], Ana Cañadas[7,10], Sophie Laran[8], Nathalie Di Méglio[3], Pauline Gauffier[9]

[1]Joint Research Centre of the European Commission, Maritime Affairs Unit,
Institute for the Protection and Security of the Citizen, Via Fermi, TP 051, 21027 Ispra (VA), Italy
[2]Tethys Research Institute, c/o Acquario Civico, Viale G.B. Gadio 2, 20121 Milan, Italy
[3]écoOcéan Institut, 18 rue des Hospices, 34090 Montpellier, France
[4]Groupe de REcherche sur les Cétacés (GREC), BP 715, 06633 Antibes cedex, France
[5]Souffleurs d'Ecume, Hôtel de Ville, 83470 La Celle, France
[6]Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Via Vitaliano Brancati, 48, 00144 Rome, Italy
[7]ALNITAK, C/ Nalón 16, La Berzosa, 28240 Hoyo de Manzanares, Madrid, Spain
[8]Centre de Recherche sur les Mammifères Marins, UMS 3462, Université de La Rochelle/CNRS, 17000 La Rochelle, France
[9]Circé, Cabeza de Manzaneda, 3, 11390 Pelayo, Algeciras, Spain

[10]*Present address:* ALNILAM Research and Conservation, Cándamo 116, 28240 Hoyo de Manzanares, Madrid, Spain

*Email: jean-noel.druon@jrc.ec.europa.eu

Marine Ecology Progress Series 464: 289–306 (2012)

Supplement 1: Pre-processing of satellite data for limited edge recovery

Computation of the horizontal gradient of sea-surface temperature (SST) and chlorophyll *a* (chl *a*) for deriving fronts tends to erode the habitat coverage around missing values. To reduce the occurrence of irregular SST or chl *a* fields due to cloud occlusions of small size, several iterations of median filtering were applied to the original data when sufficient information was available in the neighbourhood of 2 pixels, i.e. 9.2 km. At the last iteration, the filtered values were superimposed on the original data to avoid divergence from the original information. A Gaussian filter was then applied to decrease the error in the recovered pixel value; the error may, in fact, be substantial in areas of strong gradients. The quality of the recovered data in the vicinity of missing values was tested using daily SST and chl *a* data with good coverage. An important cloud mask from another day was first applied on this data, which was then filtered as described above, and the result was compared to the original data. The mean absolute error for SST between the recovered and original values was 0.2°C (Aulanier & Druon 2010), which is below the mean error of the observation by the satellite sensor (~0.5°C). Considering the range of chl *a* in which fin whales *Balaenoptera physalus* were observed (5th and 95th percentile values are 0.10 and 0.34 mg m^{-3}, respectively, where n = 1809), the mean error value of 0.005 mg m^{-3} for chl *a* recovery leads to relative errors of 5.0 and 1.4% for the 5th and 95th percentiles. These relative levels of error are well below those of the satellite-derived chl *a* estimates which are 33% for MODIS-Aqua and 34% for SeaWiFS (in oceanic waters deeper than 200 m; Sean Bailey pers. comm.). The median filter and Gaussian smoothing procedure allows for the recovery of ca. 8% of the SST and chl *a* data. The relative gain in coverage is much higher after the gradient calculation with +42% for the SST gradient and +38% for the chl *a* gradient, and +57% for habitat coverage. The use of the median filter and Gaussian smoothing procedure is therefore particularly relevant in habitat coverage in cases of dappled cloud occlusions.

Supplement 2: Calibration of the habitat model

The calibration (i.e. estimation of the parameters) that may be seen as a tuning of the model was performed using a fraction of the dataset of fin whale *Balaenoptera physalus* geo-located sightings. The model validation is performed using the same approach as for the calibration but using an independent dataset. To do so, we defined the cost function (f_{min}) as follows:

$$f_{\mathrm{min}} = \sum D_{\mathrm{out}} - W_{\mathrm{f}} \sum D_{\mathrm{in}}$$

where D_{out} and D_{in} are the distances between a fin whale sighting and the nearest habitat boundary according to the occurrence of the sighting either outside or within the habitat boundary (see Fig. S1). W_{f} is a weighting coefficient. The cost function f_{min} allows defining the habitat size by taking the mean relative distances of fin whale sightings to the closest habitat boundary into account (Fig. S1). Note that, as the potential habitat refers to specific oceanic features in the vicinity of sightings rather than actual fin whale sightings, it was necessary to set the habitat size (using W_{f}). The hypothesis of an optimal habitat size was subsequently tested. Therefore, depending on the W_{f} value and the calibrated parameterization using f_{min}, a single sighting may be located inside or outside the potential habitat (e.g. D_{in4} and D_{out4} in Fig. S1). In other words, W_{f} was introduced to ascertain the existence of an optimum compromise between the size of the habitat and the distances of sightings to these habitats: the higher the W_{f}, the smaller the size of the habitat, leading to a higher number of sightings outside the habitat. Indeed, a habitat that would include all the field sightings would cover a large part of the western Mediterranean Sea which would have little pertinence.

Furthermore, it is expected that a proportion of fin whale sightings should be outside of the habitat because most marine species are likely to spend significant time seeking their favourable habitat and, to a lesser extent, migrating towards or away from a habitat specifically for breeding. In addition, the occurrence of clouds in the vicinity of sightings may hide potential habitat, thus erroneously increasing the distance from the habitat. To circumvent these 2 difficulties, the 90th percentile of the shortest distances of sightings to the preferred habitat is used to optimize the parameter set (i.e. 10% of the highest D_{out} values are removed for minimizing f_{min}). Note that above the 90th percentile, the increase of the percentile distance between sightings and the closest habitat (i.e. the increase of D_{out}) is about twice that of the distances below the 90th percentile whatever the W_{f} value.

In order to derive an optimal value for W_{f} for fin whales, we calibrated the model using the cost function for a wide range of W_{f} values (from $W_{\mathrm{f}} = 1$ to 8, with a step of 0.1 around the optimum value). Fig. S2A shows the relationship identified between the habitat size and mean 90th percentile distance of observations from the closest habitat with the W_{f} tested values. After normalization of both *x*- and *y*-axes, and computation of distance of each point from the origin (Fig. S2B), values of $W_{\mathrm{f}} = 3.1$ (SeaWiFS/MODIS-Terra) and 3.4 (MODIS-Aqua) were found to optimize the compromise between the habitat size and the number of fin whale sightings outside of the habitat. This optimization leads to a potential habitat that is, on average between sensors, 54% smaller than that found with a value of 1 for W_{f} (Fig. S2A). Note that in Fig. S2A the model performs slightly better using SeaWiFS/MODIS-Terra sensors than with MODIS-Aqua sensors with near optimum W_{f} values, a similar habitat size and a lower mean distance to habitat (of ca. 0.8 km). The authors suggest that this pattern is derived from the differences in the presence data used for the calibration between both pairs of satellite sensors (caused by differences in sensor lifespan, cloud cover at the time of observation and temporary sensor malfunction).

The model calibration based on the cost function f_{min} was performed by the '*fminsearch*' function of Matlab software (Matlab 2006). The function is made for the unconstrained nonlinear optimization of a scalar objective function of several variables and is based on Nelder-Mead simplex (direct search) method (Nelder & Mead 1965). In our case, the 5 model parameters (for each pair of satellite sensors) were determined by educated guesses from the previous research. The optimization is performed through the iterative evaluation of the objective function and was stopped by determining a minimum difference between the best current solution and previous solutions. We noticed that after 40 to 60 iterations the solution converged without any large changes in parameter or objective function values, thus providing an accurate calibration solution.

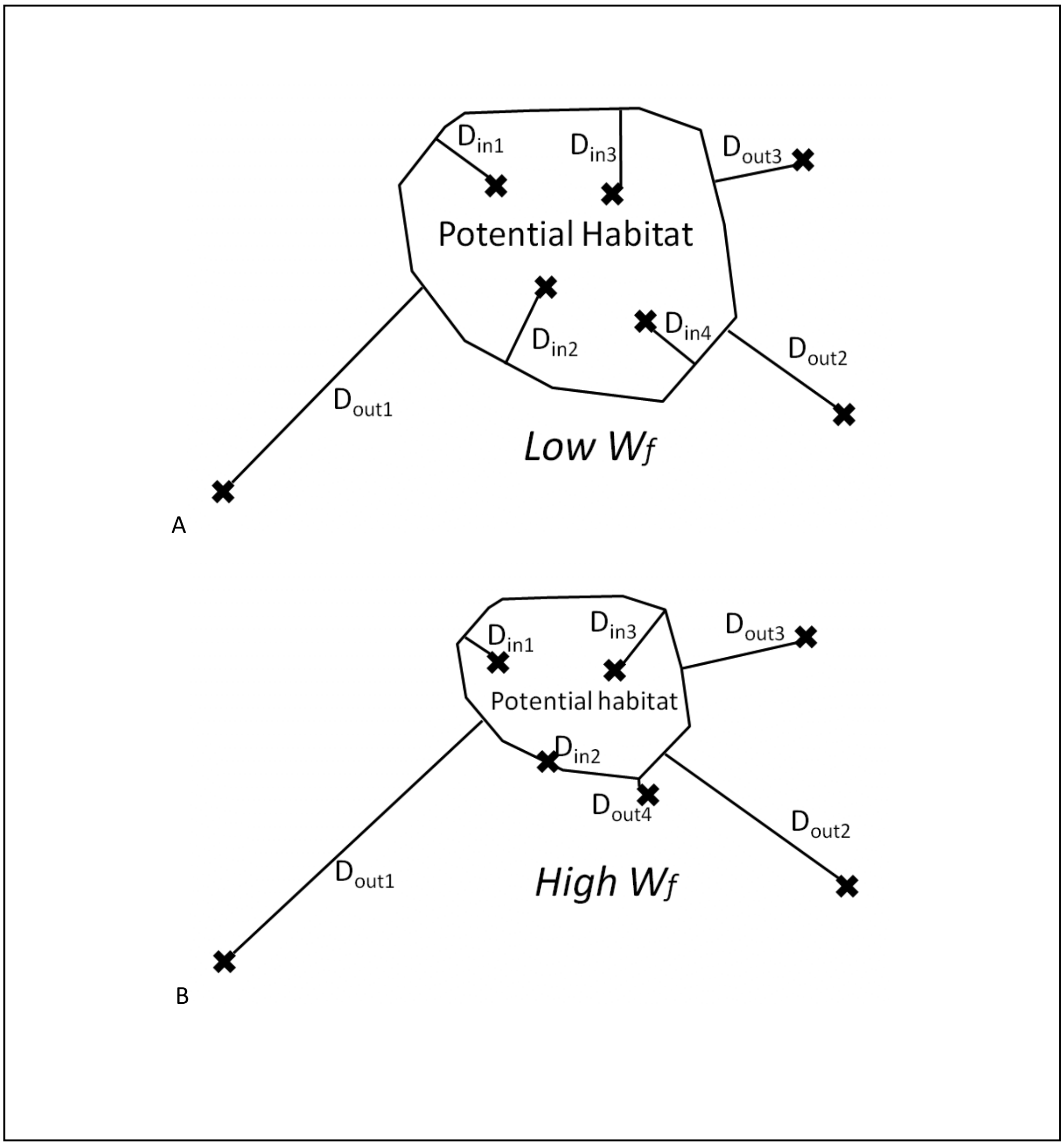

Fig. S1. Relative influence of the weight factor value W_f (A, low; B, high) on habitat size. Note that an observation may exit the potential habitat when W_f increases. D_{in}: sightings inside the habitat; D_{out}: sightings outside the habitat

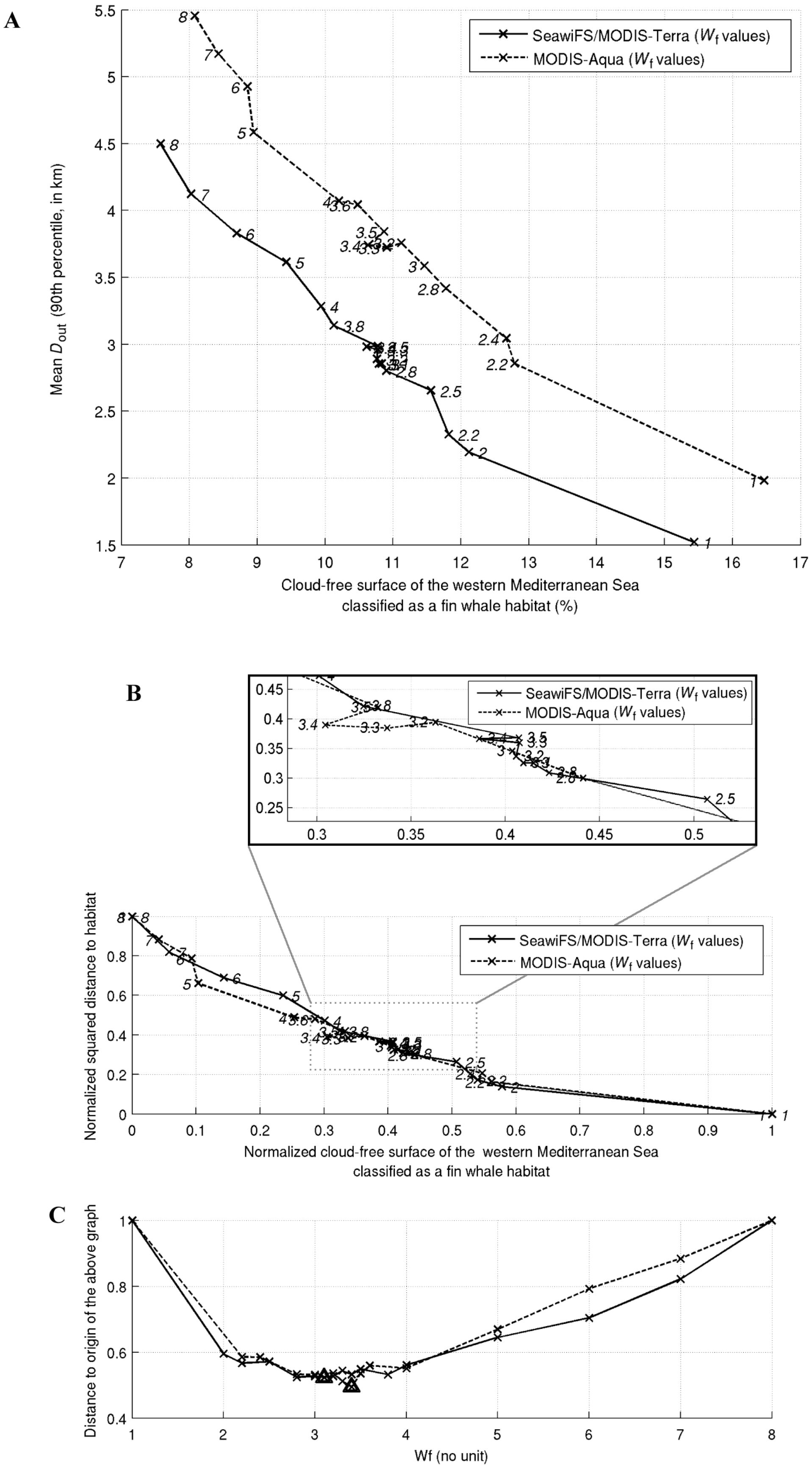

Fig. S2. Calibration of the fin whale habitat model. (A) Mean relative distances between the sightings outside the habitat and the closest potential habitat (90th percentile of D_{out}) and relative mean sea surface identified as favourable habitat (% of the western Mediterranean Sea) for an optimized parameterization using W_{f} values from 1 to 8. (B) Same as (A) with normalized coordinates. (C) Distances from the origin of the points of (B)