Ecological Modelling 222 (2011) 1048–1059

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

# Optimization of an artificial neural network for identifying fishing set positions from VMS data: An example from the Peruvian anchovy purse seine fishery

Rocío Joo[a,b,*], Sophie Bertrand[a,b], Alexis Chaigneau[c,b], Miguel Ñiquen[b]

[a] Institut de Recherche pour le Développement (IRD), Centre de Recherche Halieutique Méditerranéenne et Tropicale, Avenue Jean Monnet, BP 171, 34203 Sète Cedex, France
[b] IMARPE, Esquina Gamarra y General Valle S/N Chucuito, Callao, Lima, Peru
[c] Institut de Recherche pour le Développement (IRD), Laboratoire d'Océanographie et de Climatologie: Expérimentation et Analyse Numérique (LOCEAN), UMR CNRS, Université P. et M. Curie, IRD, MNHN, Paris, France

ARTICLE INFO

*Article history:*
Received 8 January 2010
Received in revised form 24 August 2010
Accepted 26 August 2010
Available online 29 September 2010

*Keywords:*
Artificial neural networks
Sensitivity analysis
Fishing set locations
Vessel Monitoring System
Anchovy purse seine fishery

ABSTRACT

The spatial behavior of numerous fishing fleets is nowadays well documented thanks to satellite Vessel Monitoring Systems (VMS). Vessel positions are recorded on a frequent and regular basis which opens promising perspectives for improving fishing effort estimation and management. However, no specific information is provided on whether the vessel is fishing or not. To answer that question, existing works on VMS data usually apply simple criteria (e.g. threshold on speed). Those simple criteria generally focus in detecting true positives (a true fishing set detected as a fishing set); conversely, estimation errors are given no attention. For our case study, the Peruvian anchovy fishery, those criteria overestimate the total number of fishing sets by 182%. To overcome this problem an artificial neural network (ANN) approach is presented here. In order to set both the optimal parameterization and use "rules" for this ANN, we perform an extensive sensitivity analysis on the optimization of (1) the internal structure and training algorithm of the ANN and (2) the "rules" used for choosing both the relative size and the composition of the databases (DBs) used for training and inferring with the ANN. The "optimized" ANN greatly improves the estimates of the number and location of fishing events. For our case study, ANN reduces the total estimation error on the number of fishing sets to 1% (in average) and obtains 76% of true positives. This spatially explicit information on effort, provided with error estimation, should greatly reduce misleading interpretations of catch per unit effort and thus significantly improve the adaptive management of fisheries. While fitted on Peruvian anchovy fishery data, this type of neural network approach has wider potential and could be implemented in any fishery relying on both VMS and at-sea observer data. In order to increase the accuracy of the ANN results, we also suggest some criteria for improving sampling design by at-sea observers and VMS data.



## 1. Introduction

The ecosystem approach to fisheries management is increasingly calling for spatially explicit indicators (e.g. Pikitch et al., 2004; Babcock et al., 2005). Among those indicators, fishing effort is crucial for at least two reasons: (1) the need to control the compliance with spatially explicit management measures (such as inshore restrictions or Marine Protected Areas), and (2) the need to improve the interpretation of catch-per-unit of effort (CPUE) data in terms of fish stock abundance (see for instance problems related with spatial hyper aggregation, Rose and Kulka, 1999).

The effort deployed by fishing fleets can be spatially examined by VMS data. VMS provide high-resolution records of vessel positions on a regular time basis, ranging from few minutes to few hours depending on the fishery. However, while they are available for numerous fisheries and vessels (www.fao.org), this type of data has been barely used in fisheries science and management. One of the reasons is that VMS data do not provide explicit information on whether a vessel is fishing or not. Then, a first step for processing those data consists in estimating from position records where fishing events (set, haul, trawl, etc.) probably occurred. This is not a trivial task and several approaches have been undertaken, from very coarse ones (such as unique speed thresholds applied to different activities, Witt and Godley, 2007; and areas of high VMS poll density summed up with kernel home ranges, Harrington et al., 2007) to more refined ones (such as a combination of speed thresholds, directionality and other complementary rules; Deng et al., 2005; Mills et al., 2007). Although those methods usually detect

* Corresponding author. Tel.: +33 0676535663.
*E-mail addresses:* kaoridrummer@gmail.com (R. Joo), Sophie.Bertrand@ird.fr (S. Bertrand), alexis.chaigneau@ird.fr (A. Chaigneau), mniquen@imarpe.gob.pe (M. Ñiquen).

0304-3800/$ – see front matter 
doi:10.1016/j.ecolmodel.2010.08.039

quite properly true positives (a true fishing set detected as a fishing set), false positives and over (or under) estimation errors are rarely assessed. For the present case study, the Peruvian purse seine anchovy fishery, Bertrand et al. (2008) showed that the use of a simple speed threshold on raw VMS data leads to an overestimation of the number of fishing sets of 182%. Alternatively, a general linear regression modeling approach (GLM) identified 65% of true positives and 16% of false positives, leading to a global underestimation of the total number of fishing sets of 19%. Thus, none of these two approaches are satisfying since they strongly bias the estimation of the number of fishing sets.

To solve this problem for the Peruvian anchovy fishery, Bertrand et al. (2008) proposed an ANN approach based on a Multilayer Feed-Forward Network (MFN). This methodology was chosen because (1) ANNs do not require to know nor to assume any probability distribution function, (2) ANNs are adapted for working with large datasets linked by complex non linear relations, and (3) among ANNs, MFNs are commonly used because of their simplicity and the wide availability of software tools. This particular ANN is first trained on a subset of fishing trips for which fishing set positions are documented by at-sea observers (~1% of the total fishing trips) and then used to estimate the location of fishing sets for the remaining trips monitored by VMS only. This ANN is designed to overcome the overestimation problem as it aims at: (1) accurately assessing the total number of fishing sets and (2) maximizing the true-to-false positives ratio. Bertrand et al. (2008) applied this tool on a rather limited spatial area (7°S–10°S along the Peruvian coast) and temporal window (2000–2002), correctly identifying 83% of the real fishing sets (true positives) with a total overestimation of 0.5%.

Based on those satisfying preliminary results, and before implementing such a tool in routine for the monitoring dashboard of the Peruvian anchovy fishery, there is a critical need to check the behavior and validate an optimal parameterization of the ANN when confronted to variable situations in time and space. We are particularly concerned in answering 3 main questions:

(1) At-sea observers data are usually not available on a day-to-day basis; then, when using the ANN in real-time during the fishing season, what is the effect of estimating fishing sets using an ANN trained on an earlier period?
(2) Two types of vessels, steel and wooden hulls, are participating to the industrial anchovy reduction fishery; although wooden vessels are only recently and gradually incorporated to the VMS monitoring. What may be the effect in the ANN behavior of this change in the fleet composition in VMS data?
(3) Management rules differ between two large regions in the coast of Peru (coastal restrictions, fishing bans, total allowable catch); does the ANN need different optimizations for the two regions or can we use a single ANN along the entire Peruvian coast?

Those questions are addressed performing a global optimization of the ANN. ANN optimization ranges from trial-and-error sensitivity analysis (e.g. Dedecker et al., 2004) to more complex and efficient approaches such as genetic algorithm or simulated annealing (e.g. Mühlenbein, 1990; Sexton et al., 1999; Bernardos and Vosniakos, 2007). These latter optimizations mainly concern network architecture (e.g. number of neurons and layers, and shape of activation functions) and convergence rules for the training algorithm (Bernardos and Vosniakos, 2007). Here, we refer to the optimization of (1) the internal structure and training algorithm of the ANN and (2) the "rules" used for choosing the relative size and composition of the DBs for the ANN training and inference. To address these two aspects with comparable methods, we use a trial-and-error sensitivity analysis.

In the next section, we present some characteristics of the Peruvian fishery, the data used in this study, the ANN architecture and training algorithm; and then we describe and perform a series of sensitivity tests. From the results obtained in terms of ANN performance, we draw practical recommendations for an optimal and robust use of this tool, in the specific case of the Peruvian anchovy fishery. Finally, we discuss on how this type of neural network approach can have wider potentials and could be implemented, with adjustments on the input variables, in any fishery relying on both VMS and at-sea observer data.

## 2. Materials and methods

### 2.1. Some insights into the Peruvian anchovy reduction fishery, its monitoring and management and the data used

The Peruvian anchovy fishery is characterized by the remarkable size of its production (~7 millions t.y$^{-1}$ since 1999) and its sensitivity to the intense regional climatic variability on various spatio-temporal scales (Chavez et al., 2008). Indeed, climatic scenarios such as El Niño or la Niña events directly condition the extent of the anchovy habitat, modifying its catchability and driving its population dynamics (Bertrand et al., 2004a). To cope with this strong natural variability, fishing authorities adopted an adaptive management for the industrial fishery (Chavez et al., 2008). This management is adaptive since catch limits are re-assessed every ~6 months and opening and closure periods decided on the basis of daily monitoring of the ecosystem, the fish population and the fishery. The fishing activity is monitored by the Instituto del Mar del Peru (IMARPE) through landings statistics, VMS and at-sea observer data. At-sea observer data provide, for a small sample (~1%) of fishing trips, detailed information on time spent on steaming and searching, and on the position and catch composition of the fishing sets, among others (Bertrand et al., 2004b). Accepting observers on board is not a legal obligation for fishing companies. The observer program is therefore run on a voluntary basis by the vessels and relies on "gentlemen's agreement" between IMARPE and the fishery. While landings and VMS data are daily available to fishery authorities, at-sea observer data need several weeks to be centralized and formatted. This difference in timing of availability for the different datasets means that fishing set positions need to be estimated mainly with an ANN trained on an earlier period. The management of this fishery also depends on the region. In the north-centre region of Peru (NC, from the frontier with Ecuador at ~3°S–16°S), where most of the landings take place, industrial fishing is forbidden within the first 5 nm from the coast. Total landings are limited by a total allowable catch (TAC) and effort is limited by fairly long fishing bans (at least for the period under study; Peru implemented a new individual quota system in 2009 which should lead to longer fishing seasons). In the southern region (S, from 16°S to the frontier with Chile), coastal restrictions vary from 1.5 to 3 nm from the coast. As total landings in that area are not limited by a TAC, fishing closures are much more limited in duration and are mainly implemented if the number of juveniles in the catches exceeds a given threshold. The pelagic industrial fleet is composed of two types of vessels, both of them delivering their catches to fish meal plants for reduction: a steel fleet (steel-hulled and at least 110 m$^3$ of fish-hold capacity), and a wooden fleet (wooden-hulled and with a fish-hold capacity ranging from 30 to 110 m$^3$). Both fleets are legally obliged to use VMS tracking devices since 1999. While the steel fleet was almost entirely covered with VMS by 2000, the coverage of the wooden fleet has been much more gradual.

In this study, we use the complete VMS and at-sea observer data from late 1999 to 2007 available along the Peruvian coast. VMS data provide on a ~1-h basis, precise geographical position records for the entire steel fleet and a continuously growing part

**Table 1**
Characteristics of the databases.

| | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|
| *Complete Database* | | | | | | | | |
| Fishing season duration (days; north-centre region) | 182 | 224 | 193 | 169 | 168 | 103 | 49 | 50 |
| Total fishing trips monitored by VMS | 36,866 | 33,769 | 54,263 | 36,667 | 52,857 | 55,890 | 39,158 | 38,431 |
| Median duration between 2 consecutive VMS records (h) | 1.00 | 0.83 | 0.70 | 0.65 | 0.70 | 0.71 | 1.00 | 1.00 |
| *Training Database (TD)* | | | | | | | | |
| Fishing trips monitored by both VMS and at-sea observers | 257 (0.7%) | 248 (0.7%) | 761 (1.4%) | 340 (0.9%) | 479 (0.9%) | 537 (1%) | 282 (0.7%) | 248 (0.6%) |
| Number of vessels | 42 | 45 | 68 | 53 | 70 | 90 | 36 | 46 |
| Number of position records | 6,743 | 6,841 | 22,792 | 14,377 | 16,990 | 15,576 | 7,945 | 6,613 |
| Percentage of the records corresponding to fishing events | 11.0% | 10.0% | 9.6% | 7.0% | 8.3% | 9.6% | 9.8% | 10.5% |
| Percentage of fishing trips within the north-centre region | 91.4% | 88.7% | 81.5% | 92.6% | 83.5% | 84.7% | 99.7% | 87.9% |
| Percentage of 'wooden' vessels* | 0.0% | 0.0% | 0.0% | 0.0% | 2.9% | 6.7% | 5.6% | 15.2% |
| Percentage of 'wooden' fishing trips[a] | 0.0% | 0.0% | 0.0% | 0.0% | 0.8% | 2.6% | 0.6% | 15.3% |

[a] Quantities are computed from the training database (clean vessel trajectories).

of the wooden fleet (Table 1). We pre-process raw VMS data as in Bertrand et al. (2005, 2007) to retain only trips at sea from the bulk of vessel positions. To exclude the not-fishing trips (e.g. transit between ports), we select trips having a minimum speed lower than 3 knots (a necessary but not sufficient criteria for possible fishing activity). Also, to exclude the very few trips that targeted other pelagic species than anchovy (e.g. jack mackerel or mackerel), we only retain trips lasting less than 5 days (most of anchovy-oriented fishing trips actually last ∼24 h). From the at-sea observer DB, we use the following information: vessel specifications, ports, time of departure and arrival, and fishing set geographical positions.

## 2.2. Artificial Neural Network approach

### 2.2.1. Training databases

The training DB is built by matching the trips monitored by at-sea observers with their respective VMS records. Since the training DB needs to be very accurate we discard any trip that may cause inconsistencies. First, to avoid possible wrong matches caused by vessel homonymy, we do not consider trips for which fishing set positions reported by observers differ by more than 5 nm from the closest VMS record. Second, as a fishing set lasts only 1–2 h, we discard any trip having a gap between two consecutive records higher than 2 h. General statistics on the training DBs are presented in Table 1 and a particular vessel track is shown for illustration in Fig. 1. Training DBs are initially based on a 1-year period.

### 2.2.2. The network architecture

The MFN used in this study learns how to transform input data into a desired output response through one or more hidden neural layers. Theoretical works by Cybenko (1989) and Hornik et al. (1989) suggest that one hidden layer is enough for approximating any complex nonlinear function with great accuracy. Our MFN is composed of three neural layers: the input layer, the middle "hidden" layer, and the output layer (Fig. 2). Neurons of a given layer are linked to the neurons of the next layer by activation functions. The input layer consists of five neurons, each one representing a particular variable computed from VMS observations: speed and time associated with the VMS record, absolute change of heading between the previous and the current step ($|\Delta\theta|$, a step is defined as the segment between two consecutive records), acceleration between the previous and the current step ($\Delta a_{-1}$) and acceleration between the current and the next step ($\Delta a_{+1}$). Neurons from the hidden layer compute a weighted sum of the input variables through a first activation function; then they send through a second activation function a result to the single-neuron output variable. Input and hidden layers are linked by a hyperbolic tangent activation function, characterized by a fast convergence (Bishop, 1995). Hidden and output layers are linked by a logistic activation function which returns values ranging between 0 and 1, allowing probabilistic interpretations of the results. The establishment of an output threshold allows partitioning the continuous probability into the categories "fishing set" (value of 1) and "not a fishing set" (value of 0).

### 2.2.3. The training algorithm

In the training DB, records corresponding to fishing sets are labeled "1" for "fishing set" and "0" for "not a fishing set". This Boolean variable is compared with the ANN outputs to estimate its performance and to improve training. This supervised training is an iterative procedure that uses a back-propagation and fast-convergence Levenberg–Marquardt algorithm. Its role is to adjust

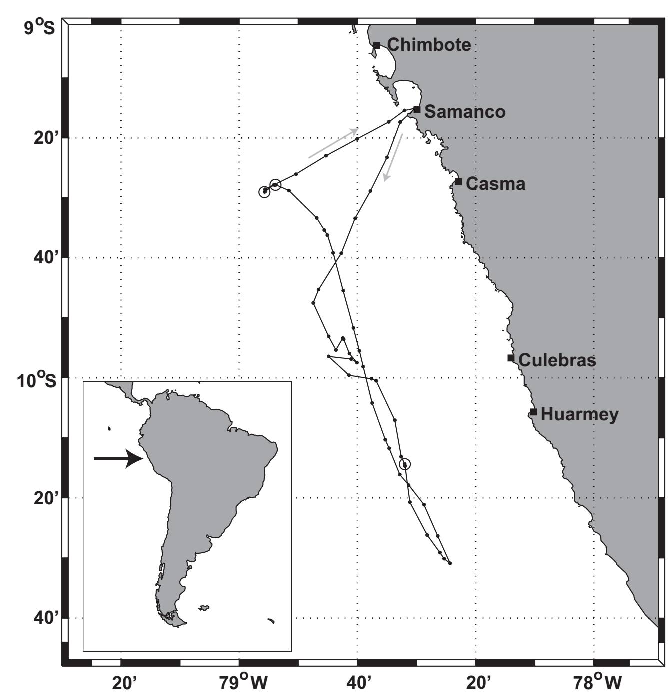


**Fig. 1.** Example of VMS positions (black dots) combined with at-sea observer information (black circles indicating fishing set positions) for a particular fishing trip held in 2008. Grey quivers show the trip direction. Inset indicates the study region of Peru.

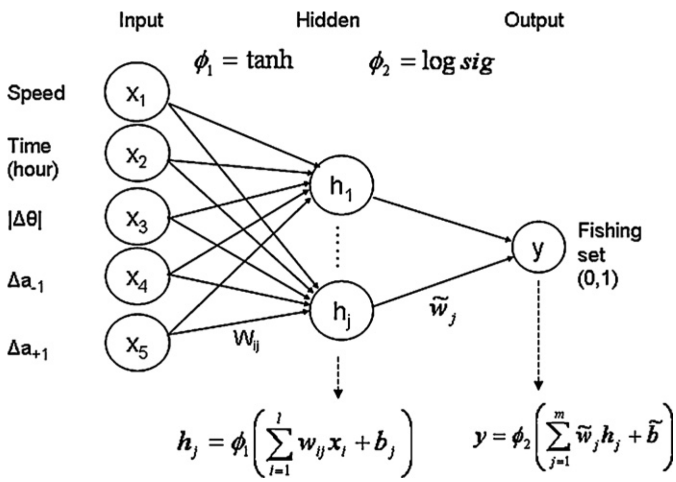


**Fig. 2.** Schematic representation of the neural network architecture. Each input neuron ($x_i$) has a corresponding weight ($w_{ij}$) associated with a hidden neuron ($h_j$) by an activation function ($\phi_1$). Each hidden neuron has a corresponding weight ($\tilde{w}_j$) associated with the output neuron ($y$) by another activation function ($\phi_2$). $b_j$ and $\tilde{b}$ are the bias (equivalent to intercept in the statistical jargon) associated with the hidden and the output layer, respectively.

the ANN weights used in the hidden layer to minimize the error function (in our case, the Mean Squared Error or MSE).

An important risk with learning-from-data methods is over-fitting: the ANN may excessively adjust to the training sample over iterations, losing its ability to generalize to new situations. We overcome that risk with an early-stopping criterion, wherein the training DB is randomly divided into 3 partitions: training, validation and test sets (Demuth and Beale, 1998). First the ANN is trained iteratively on the training set. Then the trained ANN is fed with the validation partition and the output is used to estimate the validation error, computing the MSE function. A better trained ANN leads to smaller errors for the validation set up until the ANN starts to over-fit to the training data; at this point, validation errors begin to increase. When a systematic increase of the validation errors through the iterations occurs, the training is stopped (Prechelt, 1998). The test set, independent from the previous processes, is used to evaluate the ANN performance. In order to limit the specific effect of one particular partition of the global training data into training, validation and test subsets, the ANN is run $n$ times ($n$ training replicates—TRs) using $n$ different partition replicates. Results of the ANN (training and inference) are given as averages of these $n$ replicates.

As any other local optimization method, it may lead to a local minimum, i.e. a training MSE much higher than the "best" minimum. To overcome that limitation, we established a maximum acceptable value for the MSE called $MSE_{max}$. In case $MSE > MSE_{max}$, the TR is discarded.

**Table 3**
Synthesis of the different parameters for which the sensitivity of the ANN is tested.

| | |
|---|---|
| Training databases | - Size (annual, biennial, triennial, 7-year period) |
| ANN architecture | - number of neurons in the hidden layer<br>- threshold of the output<br>- relative size of the partitions of the training database<br>- number of training replicates to reach robust results<br>- $MSE_{max}$<br>- size of the training database (annual, biannual, etc.) |
| Inference databases | - distribution of the fishing trips among regions (north-centre and south)<br>- distribution of the fishing trips among fleet segments ('steel' versus 'wooden')<br>- difference in the proportion of fishing events in the databases used for training and for inference<br>- difference in the total sizes (number of records) of the databases used for training and for inference<br>- absolute size of the database used for inference |

### 2.2.4. Indicators of ANN performance

There are four possible ANN results for each VMS observation (Table 2): (1) it is a fishing set location and the ANN correctly identifies it (true positive, TP), (2) it is a fishing set location but the ANN does not recognize it (false negative, FN), (3) it is not a fishing set location and the ANN wrongly identifies it as one (false positive, FP), and (4) it is not a fishing set location and the ANN correctly establishes it (true negative, TN). This last case is not of direct interest. Using these 4 definitions, if the number of observed and identified fishing sets (OF and IF, respectively) are equal then the number of false negatives and false positives are obviously equal and can be interpreted as positioning errors.

## 2.3. Sensitivity tests on the ANN behavior

We analyze the sensitivity of the ANN behavior to variations in 3 classes of parameters (Table 3): (1) composition and size of the training DBs; (2) ANN architecture; and (3) composition and size of the inference DBs.

### 2.3.1. Training databases

Several characteristics of the training DBs vary temporally (Table 1), which may result in a variable training performance of the ANN depending on the periods.

**Table 2**
Indicators used to measure ANN performance in the sensitivity analysis. The index $i$ denotes the training replicate and $n$ is the total number of replicates from which average performance indicators are computed.

| Acronym | Formulae | Description |
|---|---|---|
| OF (from observers) | $OF = \frac{\sum_{i=1}^{n} of_i}{n}$ | Observed fishing sets |
| IF (ANN results) | $IF = \frac{\sum_{i=1}^{n} if_i}{n} = TP + FP,\ \frac{\sum_{i=1}^{n} if_i}{\sum_{i=1}^{n} of_i} \times 100$ | Identified fishing sets |
| TP | $TP = \frac{\sum_{i=1}^{n} tp_i}{n},\ \%TP = \frac{\sum_{i=1}^{n} tp_i}{\sum_{i=1}^{n} of_i} \times 100$ | True positives (a real set identified by the ANN) |
| FP | $FP = \frac{\sum_{i=1}^{n} fp_i}{n},\ \%FP = \frac{\sum_{i=1}^{n} fp_i}{\sum_{i=1}^{n} of_i} \times 100$ | False positives (a real set not identified by the ANN) |
| FN | $FN = \frac{\sum_{i=1}^{n} fn_i}{n},\ \%FN = \frac{\sum_{i=1}^{n} fn_i}{\sum_{i=1}^{n} of_i} \times 100$ | False negative (a non fishing set record identified as a fishing set by the ANN) |

**Table 4**
Training performance using annual, biennial, triennial and 7-year period training databases. Parameters used for the ANN are: 0.5 for threshold, 3 for hidden neurons, 50 training replicates, 0.02 as $MSE_{max}$ value (0.03 and 0.05 for triennial and 7-year period databases, where a smaller value is not possible), and 20%, 20% and 60% as proportions for the test, validation and training partitions, respectively.

| Annual training | Training DB | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|---|
| | # Records in the training partition | 4,046 | 4,105 | 13,675 | 8,626 | 10,194 | 9,346 | 4,767 | 3,968 |
| | % IF | 100.2 | 100.7 | 97.9 | 101.7 | 99.4 | 100.1 | 100.5 | 97.2 |
| | % error in number of fishing sets | 0.2 | 0.7 | −2.1 | 1.7 | 0.6 | 0.1 | 0.5 | −2.8 |
| | % TP | 80.2 | 77.1 | 74.9 | 73.9 | 71.1 | 75.1 | 76.2 | 74.7 |
| | % FP | 20.0 | 23.6 | 23.0 | 27.8 | 28.3 | 25.0 | 24.3 | 22.5 |
| | % FN | 19.8 | 22.9 | 25.1 | 26.1 | 28.9 | 24.9 | 23.8 | 25.3 |

| Biennial training | Training DB | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 | 2006-2007 |
|---|---|---|---|---|---|---|---|---|
| | # Records in the training partition | 8,150 | 17,780 | 22,301 | 18,820 | 19,540 | 14,113 | 8,735 |
| | % IF | 98.5 | 100.6 | 100.7 | 99.5 | 101.3 | 99.3 | 100.2 |
| | % error in number of fishing sets | −1.5 | 0.6 | 0.7 | −0.5 | 1.3 | −0.7 | 0.2 |
| | % TP | 78.5 | 80.2 | 77.2 | 70.3 | 77.5 | 77.5 | 79.3 |
| | % FP | 20.0 | 20.4 | 23.5 | 29.2 | 23.8 | 21.8 | 20.9 |
| | % FN | 21.5 | 19.8 | 22.8 | 29.7 | 23.0 | 22.5 | 20.7 |

| Triennial training | Training DB | 2000–2002 | 2001–2003 | 2002–2004 | 2003–2005 | 2004–2006 | 2005–2007 |
|---|---|---|---|---|---|---|---|
| | # Records in the training partition | 13,906 | 15,844 | 19,497 | 16,900 | 14,584 | 10,849 |
| | % IF | 97.9 | 100.1 | 99.4 | 100.7 | 100.9 | 98.7 |
| | % error in number of fishing sets | −2.1 | 0.1 | −0.6 | 0.7 | 0.1 | −1.3 |
| | % TP | 75.8 | 75.2 | 75.6 | 73.2 | 76.8 | 76.3 |
| | % FP | 22.1 | 24.9 | 23.9 | 27.5 | 24.1 | 22.4 |
| | % FN | 24.2 | 24.8 | 24.4 | 26.8 | 23.2 | 23.7 |

| 7-years long training | Training DB | 2000-2006 |
|---|---|---|
| | # Records in the training partition | 54,759 |
| | % IF | 97.3 |
| | % error in number of fishing sets | −2.7 |
| | % TP | 75.1 |
| | % FP | 22.2 |
| | % FN | 3.2 |

(1) The length of the fishing season decreased significantly over the study period because of a considerable increase of the fleet capacity which meant that the TAC was used up in fewer days (Fréon et al., 2008; Aranda, 2009). The race between vessels, making them try to fish more and faster, may have impacted fishing strategies and thus the probability distribution of the ANN input variables.
(2) The average duration between consecutive VMS records varied from 1 h in 2000 to 0.5 h in 2004, increasing again to1 h in 2006 and 2007 mainly due to changes in the companies providing VMS services. The probability distribution of the input variables for the ANN tightly depends on this duration (see Section 2.2.2) and its variation may affect the ANN performance.
(3) While the VMS coverage of the wooden fleet has increased over time, observers still have few opportunities to embark on those smaller vessels and therefore this segment of the fleet remains under-sampled: while training DBs gathered at best 15% of wooden vessels in 2007, their real proportion in the fleet increased from 22% in 2000 to 50% in 2003 (Bouchon, personal communication). ANN performance may be also affected by this evolution in the fleet structure.
(4) From 2000 to 2005, there were no large differences in the percentages of trips among regions between the observers' samples and the VMS dataset. However, since 2006, the real number of southern fishing trips increased to 21–24%, a change which was not reflected by the observers' sample.
(5) Training DBs gathered from 36 (in 2006) to 90 (in 2005) fishing trips. This variability strictly depends on observers' ability to be accepted onboard. A DB with a large number of vessels and trips has wider information on fishing patterns for the ANN to learn from. ANN performance may be also sensitive to variations on both the number of records – correlated with the amount of fishing trips – and the proportion of records corresponding to fishing sets.

To study the effects of the first 4 aspects on ANN performance globally, we examine the training performances year by year. To address the fifth aspect (size of the training DB), we compare the training performance when using DBs covering 1-year, 2-year, 3-year and 7-year periods.

### 2.3.2. Network architecture

The values of 5 parameters of the ANN architecture are optimized through sensitivity analysis. First, the number of neurons to include in the hidden layer. Too few neurons lead to large training and generalization errors due to under-fitting and high residuals, and too many neurons also lead to large generalization errors due to over-fitting and high error variability (Sarle, 1997). The best trade-off is the lowest number of neurons that minimizes the result errors. We perform sensitivity tests on the ANN in order to find this optimal value for each training DB. Secondly, the threshold applied to the continuous output variable of the ANN determines the minimum acceptable value to consider a VMS record as a fishing set.

Thirdly, the size of each partition (training, validation and test subsets) of the global training DB: literature is not conclusive about an "ideal" size for each partition since such a size is conditioned by the particular applications (Sarle, 1995; Zhang et al., 1998; Lek and Guégan, 2000). Fourthly, the number of TRs is also needed to be optimized by sensitivity analysis since it is a critical parameter to obtain robust ANN results. Finally, the fifth parameter to be optimized is $MSE_{max}$. If small, it reduces the risk of local minima, improving the quality of the results. On the other hand, the smaller the $MSE_{max}$, the longer the time required for computations. If it is too small, the ANN may never reach to overcome the restriction.

**Table 5**
Variation in the percentage of true positives (%TP; in parenthesis %IF) for the different sizes of training databases, when parameters diverge from values considered as optimal: 0.5 for threshold, 3 for hidden neurons, 50 training replicates, 0.02 as $MSE_{max}$ value (0.03 and 0.05 for 3-year and 7-year period databases, where a smaller value is not possible), and 20%, 20% and 60% as proportions for the test, validation and training partitions, respectively.

| Variable to test | Values tested | 1-year long Databases | | 2-year long Databases | | 3-year long Databases | | 7-year long DB |
|---|---|---|---|---|---|---|---|---|
| | | 2001 | 2005 | 2000-2001 | 2005-2006 | 2000-2002 | 2004-2006 | 2000-2006 |
| Threshold | 0.6 | −2.5 (−3.7) | −3.5 (−4.5) | −3.0 (−3.6) | −3.7 (−5.1) | −3.8 (−5.3) | −5.4 (−8.9) | −6.6 (−11.3) |
| | 0.7 | −9.7 (−10.8) | −6.3 (−8.8) | −6.9 (−9.6) | −8.4 (−13.4) | −10.2 (−14.5) | −14.6 (−22.0) | −17.2 (−26.7) |
| | 0.8 | −34.7 (−40.6) | −16.5 (−20.5) | −19.3 (−26.3) | −19.0 (−26.3) | −34.4 (−41.7) | −37.2 (−48.0) | −42.7 (−58.3) |
| Number of hidden neurons | 4 | −1.9 | −1.8 | −0.7 | −0.6 | −0.5 | 1.0 | −0.3 |
| | 5 | −5.9 | −3.4 | −2.0 | −1.8 | −0.8 | 0.5 | 0.1 |
| | 6 | −6.3 | −4.8 | −3.0 | −2.2 | −1.0 | 1.6 | 0.3 |
| | 7 | −8.9 | −6.2 | −3.6 | −2.0 | −1.0 | 0.5 | 0.2 |
| | 8 | −11.4 | −7.0 | −4.4 | −3.3 | −1.0 | 1.0 | −0.1 |
| Data partition: test, validation and training partitions (%) | 10/10/80 | 7.2 | 2.4 | −0.2 | −2.7 | −0.1 | −1.0 | 0.5 |
| | 20/30/50 | −4.9 | −0.8 | −1.4 | −1.0 | −0.2 | −0.8 | 0.2 |
| | 30/30/40 | −0.9 | −2.8 | −2.7 | −2.9 | −0.7 | −0.9 | −0.2 |
| | 40/40/20 | −13.6 | −8.9 | −5.6 | −6.6 | −2.7 | −2.4 | −1.1 |
| | 50/40/10 | −24.5 | −15.1 | −10.3 | −10.6 | −6.2 | −4.9 | −3.1 |
| Number of training replicates | 20 | 1.5 | 1.7 | −1.5 | −0.8 | 0.3 | −0.2 | 0.4 |
| | 30 | 0.9 | −0.8 | −0.7 | −0.8 | 0.3 | −0.3 | 0.5 |
| | 40 | 0.4 | 0.4 | −0.4 | −0.6 | 0.3 | −0.7 | 0.2 |
| | 60 | 1.1 | −0.4 | −0.3 | −0.4 | −0.7 | 0.3 | 0.1 |
| | 70 | 0.4 | 0.6 | −0.6 | −0.4 | 0.5 | 0.4 | 0.5 |
| | 80 | 0.7 | 0.0 | −0.7 | −0.3 | 0.3 | −0.3 | 0.1 |
| $MSE_{max}$ value | 0.03 | −8.3 | −7.0 | −2.9 | −3.6 | – | – | – |
| | 0.04 | −15.3 | −8.8 | −3.5 | −2.6 | 1.3 | 0.8 | – |
| | 0.05 | −14.9 | −11.0 | −4.4 | −4.4 | −0.8 | −1.4 | – |
| | 0.06 | −16.2 | −10.8 | −4.9 | −5.9 | −1.3 | −1.3 | 0.3 |
| | 0.07 | −13.1 | −13.2 | −5.0 | −6.9 | −1.3 | −2.0 | −0.1 |

#### 2.3.3. Inference databases

For the reasons presented in the introduction and in Section 2.1, inference of fishing set positions in the considered fishery will be mainly performed using an ANN trained on an earlier period. We thus compare the inference performance of ANNs trained on DBs covering the immediately preceding 1-year, 2-year and 3-year periods. Then we study the effects of different compositions between training and inferring DBs in terms of (1) proportion of fishing trips between regions (NC versus S), (2) proportion of steel versus wooden vessels, (3) proportion of records corresponding to fishing sets, and (4) total sizes (i.e. number of records). We finally examine the pure effect of the size of the inference DB by inferring fishing set positions for the same period in two different ways:

(1) Feeding the ANN with data belonging to one fishing season and then inferring the corresponding fishing set positions.
(2) Feeding the ANN with data belonging to an entire year and inferring the corresponding fishing set positions. Then, only trips corresponding to the same fishing season in (1) are kept for comparison.

## 3. Results

### 3.1. Sensitivity analysis on the training databases

Training results are synthesized in Table 4. They are consistent with the preliminary ones in Bertrand et al. (2008). ANNs prove to be performing over the full available time series and the entire Peruvian coast. When considering training performances for DBs based on 1-year periods, the global error concerning the total number of fishing sets lies in [−2.8%; 1.7%]. It is significantly smaller than the ones of simple threshold on speed (+182%) and GLM (−19%) approaches evaluated on the 2000–2002 period (Bertrand et al., 2008). While maintaining overall error estimation low, a high percentage of true positives (%TP) is also obtained, with values lying in [71.1%; 80.2%].

With respect to the size of the DBs, we observe that ANNs trained on 1-year period DBs obtain the best TP rates (up to 80.2%). Nevertheless, using longer DBs stabilizes the ANN performance results (global error of the number of fishing sets varies by 2.8% for 3-year period DBs in contrast with the 4.5% for 1-year period DBs; %TP varies by 3.6% for 3-year period DBs in contrast with the 9.1% for 1-year period DBs). The 7-year training DB (2000–2006), has an inferior performance in detecting fishing sets (%TP: 75.1%). The fact that in general both %FN and %FP remain similar was expected and will be discussed later.

### 3.2. Sensitivity analysis on the ANN architecture

Sensitivity results to parameters of the ANN architecture are synthesized in Table 5.

#### 3.2.1. Threshold

The ANN is particularly sensitive to variations in the threshold value used to classify the output results as fishing sets or not fishing sets. A slight increase in the threshold value produces an important decrease of both %IF and %TP. As a result, criterion based only on maximizing %TP will also maximize %IF leading to an overestimation of the number of fishing sets. We thus define the best threshold value as the one that minimizes the estimation error of the total number of fishing sets, while maximizing %TP. According to this criterion, the optimum threshold value varies from 0.5 to 0.56 depending on the considered training DB.

#### 3.2.2. Number of neurons in the hidden layer

Variations of the number of hidden neurons have smaller effects in %TP. Since ANN with 1 or 2 hidden neurons could not converge, the minimum number of hidden neurons is fixed at 3. In general, 3 neurons in the hidden layer provide the best results, maximizing

**Table 6**
Optimal relative size of data partitions for each training database.

| Data Partition | Annual DB | | Biennial DB | | Triennial DB | | 7-year long DB |
|---|---|---|---|---|---|---|---|
| | 2001 | 2005 | 2000–2001 | 2005–2006 | 2000–2002 | 2004–2006 | 2000–2006 |
| Training | 55–60 | 60–70 | 35–70 | 45–75 | 45–80 | 35–75 | 60–80 |
| Validation | 20 | 15–25 | 10–40 | 10–30 | 10–40 | 15–35 | 15–25 |
| Test | 20–25 | 15–25 | 20–55 | 10–45 | 10–45 | 10–50 | 10–25 |

%TP. Variations of the number of neurons in the hidden layer have much less impact on the results when the size of the training DB is increased. In the case of the 7-year period training DB, the %TP varies only by ±0.3% when the number of hidden neurons varies between 3 and 8 (Table 5).

### 3.2.3. Size of the training DB partitions

In order to vary the relative sizes of data partitions – expressed as percentages of the whole DB-, we simultaneously change the sizes of the validation and test partitions – and accordingly, the size of the training partition is equal to 100% minus the other two. The size of the test partition varies from 10% to 50% whereas the validation one, from 10% to 40%. In general, variations in partition sizes significantly affect the %TP. Instead of an optimal size for each partition, an optimal region with ranges of sizes is obtained (Table 6). For practical purposes, and based on the optimal region, sizes of 60%, 20% and 20% of the original DB are respectively chosen for training, testing and validating the ANNs.

### 3.2.4. Number of training replicates

The ANN self-stabilizes when using at least 50 TRs. Nevertheless and independently of the considered training DB (1-year, 2-year, 3-year or 7-year periods), Table 5 confirms that the number of TRs does not strongly impact the ANN results (%TP varies by ±2% when the number of TRs changes).

### 3.2.5. $MSE_{max}$

In contrast, $MSE_{max}$ plays an important role on the rate of true positives. For annual DBs, an increase in the $MSE_{max}$ value strongly decreases the %TP (Table 5). This impact is gradually reduced as the size of the training DB is increased. Considering also the latent risk of local minima, the smallest $MSE_{max}$ value remains the best option. As it depends on the data structure, the minimum $MSE_{max}$ may differ between training DBs (Table 5), varying from 0.02 for 1-year and 2-year period DBs to 0.05 for the 2000–2006 DB.

In summary, the sensitivity analysis indicates that the best ANN parameter values are: (1) ranging from 0.5 to 0.56, an adaptive threshold value which minimizes the estimation error of the total number of fishing sets; (2) 3 neurons in the hidden layer; (3) DB partitions of 60%, 20% and 20% for the training, validation and test sets, respectively; (4) 50 TRs for the training procedure; and (5) an adaptive $MSE_{max}$ value ranging from 0.02 to 0.05 depending on the DB considered. An ANN with these parameter values obtains the training results presented in Table 4.

In addition, Table 7 presents a ranking of the tested parameters, based on the sensitivity of the ANN performance to variations in parameter values. We compare the variations in %TP when parameters are modified by a small amount: threshold from 0.5 to 0.6, neurons from 3 to 4, TRs from 50 to 60, data partitions from 60%, 20% and 20–50%, 30% and 20% for the training, validation and test partitions, and an increase in 0.01 from the minimum $MSE_{max}$. Ranking for annual DBs differ from the ones for longer DBs and between each other. ANNs trained on 1-year period DBs are more sensitive to $MSE_{max}$, meaning that the ANN has a great need of this parameter to regulate the error (it cannot do it by itself). In general, the network is more sensitive to the threshold of the output value, followed by the $MSE_{max}$, the number of hidden neurons, data partition sizes, and finally, the number of TRs.

## 3.3. Sensitivity to the inference parameters

We have previously detailed the actual need for inferring fishing set positions with ANNs trained on earlier DBs (e.g. introduction and section 2.1). In that sense, this section describes how inference behaves for a given time period, using ANNs trained on different time periods.

Table 8 presents annual inference results using training DBs built on the immediately preceding 1-year, 2-year and 3-year periods. In general, ANNs trained over 2-year period DBs show the best inference results (higher %TP); except for inferring over 2005, where training on 1-year period DBs is much more appropriate.

The threshold value assures a small estimation error for the total number of fishing sets, and consequently, similar values between the two types of errors – %FN and %FP – in the training stage. It is not expected to find the same behavior in performance indicators of inference. Here, depending on the inference DB, %IF can be lower or higher than %OF, producing over or under estimations of the total number of fishing sets. For the 2000–2002 period and comparing with speed threshold and GLM approaches, the ANN continued providing better inference results when using any of the nearest-in-time training DBs.

### 3.3.1. Effect of the distribution of fishing trips among regions

To test for this effect over inference performance, we selected two 1-year period DBs, 2004 (training) and 2005 (inference), which originally had similar distributions of fishing trips among regions (86.5% and 87.9% of fishing trips performed in the NC region, respectively). No significant difference was observed in ANN performance

**Table 7**
Ranking of parameters to which the ANN is more sensitive. The ranking is based in the magnitude of %TP variation when the values of the parameters are modified (threshold from 0.5 to 0.6, hidden neurons from 3 to 4, data partition from 60%, 20% and 20% to 50%, 30% 20%, training replicates from 50 to 60, or increasing minimum $MSE_{max}$ in 0.01). We present one ranking for each specific database and one general ranking (based on the synthesis of the specific ones). Each annual, biennial and triennial ranking is weighted 0.5, so annual, biennial, triennial and the 7-year long ranking have the same weight in the general ranking.

| Variable to test | Annual Databases | Biennial Databases | | Triennial Databases | | 7-year long DB | | Sensitivity ranking |
|---|---|---|---|---|---|---|---|---|
| | 2001 | 2005 | 2000–2001 | 2005–2006 | 2000–2002 | 2004–2006 | 2000–2006 | |
| Threshold | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| $MSE_{max}$ value | 1 | 1 | 2 | 2 | 2 | 3 | 2 | 2 |
| Number of hidden nodes | 4 | 3 | 4 | 4 | 4 | 2 | 3 | 3 |
| Data partition | 2 | 4 | 3 | 3 | 5 | 4 | 5 | 4 |
| Number of training experiences | 5 | 5 | 5 | 5 | 3 | 5 | 4 | 5 |

**Table 8**
Inference performance using 1-year, 2-year and 3-year period training databases. 'Δ% fishing events' indicates the difference in the proportion of records corresponding to fishing sets in both training and inference databases. Parameters used for the ANN are: 0.5 for threshold, 3 for hidden neurons, 50 training replicates, 0.02 as $MSE_{max}$ value (0.03 and 0.05 for triennial and 7-year period databases, where a smaller value is not possible), and 20%, 20% and 60% as proportions for the test, validation and training partitions, respectively. Estimation error is given in terms of total number of fishing sets.

| | | Inference databases | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| Annual training | **Training DB** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** | **2006** |
| | # Records in the training partition | 4 046 | 4 105 | 13 675 | 8 626 | 10 194 | 9 346 | 4 767 |
| | Δ% fishing events | 1.0 | 0.4 | 2.6 | −1.3 | −1.3 | −0.2 | −0.7 |
| | %IF | 101.8 | 101 | 112.4 | 93.3 | 102.5 | 96.2 | 73.9 |
| | **% error in number of fishing sets** | **1.8** | **1** | **12.4** | **−6.7** | **2.5** | **−3.8** | **−26.1** |
| | %TP | 77.3 | 80.0 | 81.1 | 72.9 | 80.9 | 77.8 | 60.0 |
| | %FP | 24.5 | 21.0 | 31.3 | 20.4 | 21.6 | 18.4 | 13.9 |
| | %FN | 22.7 | 20.0 | 18.9 | 27.1 | 19.1 | 22.2 | 40.0 |
| Biennial training | **Training DB** | | **2000–2001** | **2001–2002** | **2002–2003** | **2003–2004** | **2004–2005** | **2005–2006** |
| | # Records in the training partition | | 8 150 | 17 780 | 22 301 | 18 820 | 19 540 | 14 113 |
| | Δ% fishing events | | 0.9 | 2.7 | 0.3 | −1.9 | −0.9 | −0.8 |
| | %IF | | 107.4 | 118.4 | 102.5 | 96.7 | 98.2 | 80.9 |
| | **% error in number of fishing sets** | | **7.4** | **18.4** | **2.5** | **−3.3** | **−1.8** | **−19.1** |
| | %TP | | 83.2 | 84.2 | 77.0 | 75.9 | 77.8 | 62.1 |
| | %FP | | 24.2 | 34.2 | 25.5 | 20.8 | 20.4 | 18.8 |
| | %FN | | 16.8 | 15.8 | 23.0 | 24.1 | 22.2 | 37.9 |
| Triennial training | **Training DB** | | | **2000–2002** | **2001–2003** | **2002–2004** | **2003–2005** | **2004–2006** |
| | # Records in the training partition | | | 13 906 | 15 844 | 19 497 | 16 900 | 14 584 |
| | Δ% fishing events | | | 1.3 | 1.1 | −0.4 | −1.4 | −1.0 |
| | %IF | | | 117.2 | 100.7 | 105.3 | 98.2 | 81.7 |
| | **% error in number of fishing sets** | | | **17.2** | **0.7** | **5.3** | **−1.8** | **−18.3** |
| | %TP | | | 82.2 | 74.0 | 80.3 | 75.9 | 62.2 |
| | %FP | | | 35.0 | 26.7 | 25.0 | 22.3 | 19.5 |
| | %FN | | | 17.8 | 26.0 | 19.7 | 24.1 | 37.8 |

**Table 9**
Effects of the relative composition of training and inference databases in terms of regions (North-Centre and South) and fleet segment (steel and wooden vessels).

| | Training DB | Inference DB | Δ%TP | Δ%FP | Δ%FN |
|---|---|---|---|---|---|
| Region effect | 2004 NC | 2005 NC+S | 0.0 | 1.2 | 0.0 |
| | 2004 NC+S | 2005 NC | 1.6 | 0.3 | −1.6 |
| Fleet effect | 2006 Steel | 2007 Steel+Wooden | −0.6 | −6.2 | 0.6 |
| | 2006 Steel+Wooden | 2007 Steel | 9.3 | 3.0 | −9.3 |

when either the training or the inference DB is restricted to the NC region (Table 9).

*3.3.2. Effect of the relative composition of the fleet:*

For evaluating the effect of the relative composition of the fleet between steel and wooden vessels, we consider inferring over 2007 using the 2006 training DB. 2007 is indeed the inference DB for which we obtain the worst ANN inference performance, with an underestimation of 26.1% of the total number of fishing sets (Table 8). Moreover, 2007 is the DB for which we have by far the most important contribution of trips from the wooden fleet (15.3% in contrast with very low percentages from the previous years). Results show that when using only the steel fleet from 2006 to infer on both the steel and wooden fleets from 2007, performance is even slightly worse than when using the full 2006 training DB (Table 9). On the contrary, when using steel and wooden fleets in the 2006 training DB to infer fishing sets from the steel fleet in 2007, results significantly improve (+9.3.% for %TP). Thus, it is important that the training DB contains enough "wooden" data so the ANN can learn to correctly identify fishing sets occurred in inferences DBs containing "wooden" trips.

*3.3.3. Effect of the difference in fishing set proportions (p) between training and inference databases*

When the training $p$ is relatively high, the network is less restrictive for recognizing fishing set patterns. Therefore, it learns to "identify" a high proportion of fishing sets within the inference data. We proceed to an artificial equalization of the fishing set proportions between training and inference DBs (by removing fishing sets in the training DB to obtain the same $p$ of the inference DB). Table 10 shows that the estimation error on the total number of fishing sets is slightly minimized (Δ%IF = −2.4%). However, a similar $p$ for both DBs also leads to an important decrease in the detection of true positives (Δ%TP = −18.8%).

*3.3.4. Effect of the relative size in number of records (N) of training and inference databases*

For the considered years (Table 10), the training DB (2002) is larger than the inference DB (2003). To test if this difference in size may affect the ANN performance, we randomly removed records

**Table 10**
Variations in the performance of inference according to the relative proportion of fishing events ($p$) and the number of records ($N$) in the training and inference DBs. Experiments with equalized $p$ ($p$ = 7%) are performed removing records corresponding to fishing events in the training DB. Experiments with equalized $N$ ($N$ = 14,377) are performed removing randomly records from the training DB.

| 2002 (Training DB) → 2003 (Inference DB) | %IF | %TP | %FP | %FN |
|---|---|---|---|---|
| Training $p$ (9.6%) > Inference $p$ (7%)<br>Training $N$ > Inference $N$ (22,792) (14,377) | 112.4 | 81.1 | 31.3 | 18.9 |
| With identical $p$ ($p$ = 7%) | −2.4 | −18.8 | −15 | 18.8 |
| With identical $N$ ($N$ = 14,377) | −0.2 | −2.2 | −2.1 | 2.2 |

**Table 11**
Effect of the total size of the inference DB. Performance indicators for each fishing season, using an ANN trained on 2000. For each indicators, results of inference is given with inference performed only on the mentioned seasonal period (*S*) and inference performed on the whole year (*Y*), then results concerning the mentioned seasonal period retained.

| | %IF | | %TP | | %FP | | %FN | |
|---|---|---|---|---|---|---|---|---|
| | *S* | *Y* | *S* | *Y* | *S* | *Y* | *S* | *Y* |
| Fishing season (0.5 month): 01/01/2001 to 01/15/2001 | 69.5 | 107 | 63.9 | 93.1 | 5.6 | 13.9 | 36.1 | 6.9 |
| Fishing season (4.5 months): 03/12/2001 to 07/26/2001 | 98.6 | 100.2 | 75.1 | 75.8 | 23.5 | 24.4 | 24.9 | 24.2 |
| Fishing season (3 months): 10/09/2001 to 01/16/2002 | 121.8 | 96.4 | 80.0 | 76.4 | 41.8 | 20 | 20 | 23.6 |

from the 2002 DB to reach the same number of records than the one of the 2003 DB. The experiment results show a less accurate inference ($\Delta$%TP = −2.2%, and $\Delta$%FN = 2.1%). Rather than a balance in *N*, it seems more important to have as many observations as possible in both training and inference DBs.

### 3.3.5. *Effect of the absolute size of the inference database*

Table 11 (and Fig. 3) show – in accordance with the results of the previous test – that, in general, larger (yearly) inference DBs give better results than smaller (seasonal) inference DBs.

In summary, excepting for inference over 2007, ANNs trained on year *X* provide globally a good detection rate of fishing sets (TP) for inference over year *X* + 1 (between 72.9% and 84.2%, Table 8) while maintaining an estimation error of the total number of fishing sets quite low (from −6.7% to +12.4%). The 2007's is an instructive case. Training DBs with too few "wooden" trips will not gather enough information to accurately recognize fishing sets in another DB with a lot more "wooden" trips as the 2007's (Table 1). It is to be expected then that ignoring the "wooden" part of the inference DB would improve the inference results. The relative composition of the fleet

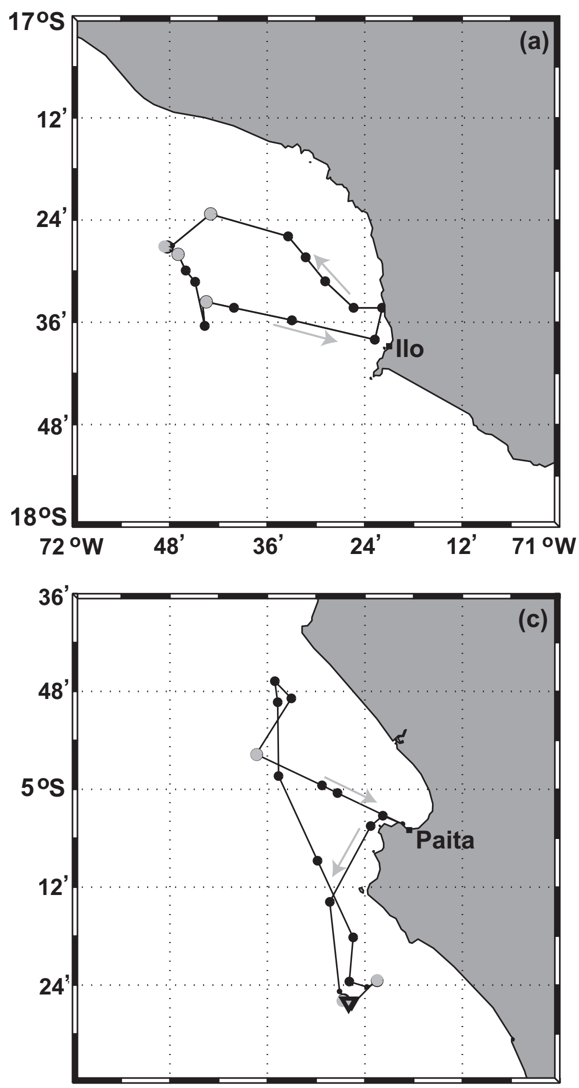


**Fig. 3.** Examples of trips (black dots for their VMS positions) with inferred fishing sets (grey circles and triangles). Inference is made over 2001's data using an ANN trained on a 2000's DB. Results for a small inference DB (a and c) and for a larger inference DB (b and d) are shown. Grey circles are true positives whereas triangles are false negatives.

between training and inference DBs is thus an important parameter for inference (Table 9). Conversely, the ANN performance is not significantly influenced by the distribution of data among regions, neither by the relative proportion of fishing sets or by the relative sizes of training and inference DBs (at least, no improvement can be obtained by balancing those parameters). Finally, the size of the inference DBs is very important for inference quality: for the same period of interest, the use of a larger inference DB improves the ANN performance (Table 11).

## 4. Summary and discussion

In this study, we tested the ability of an ANN trained on at-sea observer data to correctly infer fishing set positions in VMS data. We used an exhaustive fishery dataset from the purse seine anchovy fishery along the Peruvian coast from 2000 to 2007. Firstly, we have verified that an ANN approach provides better results than conventional criteria such as a simple threshold on speed or a GLM approach. For the studied fishery, a threshold on speed overestimates the number of fishing sets by about 182% whereas a GLM approach underestimates them by 19% (Bertrand et al., 2008). Alternatively, this ANN achieves an average of 76% of correct identification of fishing sets (%TP) and an average of 1% of estimation error for the total number of fishing sets.

Secondly, we examined the variability of the ANN behavior at training and inference, as well as its sensitivity to specific parameters. Our results show that sensitivity analysis is strongly recommended for defining the limits of confidence for the ANN. Rules of thumb or “recipes” are frequently elaborated for particular cases but they usually lack of heuristic value. In that sense, this study provides optimized parameters for tuning the ANN trained for the Peruvian purse seine industrial anchovy fishery. It also allows general recommendations for the use of ANN for inferring fishing sets from any VMS dataset.

We propose a ranking of the parameters whose variation may have greatest impact on ANN behavior, in terms of variation of %TP. The two most important parameters are internal to the ANN: (1) the threshold applied to the output value to determine if the result corresponds to a fishing set or not and (2) the $MSE_{max}$. They control the results by assuring that the numbers of identified and observed fishing sets are as similar as possible, and that the training error is minimized. These regulations are crucial for obtaining good inference performance. The third important parameter is the number of hidden neurons. ANNs with low number of hidden neurons – more parsimonious models with less (theoretical) risk of over-fitting – provide better results. Data partition sizes between training, validation and test are ranked fourth. A small change in the proportions of the partitions does not significantly affect the results. However, one of the 1-year period DBs (2001) was shown to be more sensitive to changes in the partition sizes. It is evident that variations of this parameter have stronger effects on smaller DBs. Therefore, this is not an effect of the partition size but an effect of the size of the DB itself. The number of TRs is ranked at the fifth position, since it has a very weak impact on the ANN behavior. A minimum number of TRs is needed to stabilize the results, but once this number is reached, there is no significant improvement by adding replicates. Thus in general, and unless the number of TRs is very small, ANN results are consistent and are not strongly influenced by particular partitions.

It is important to underline that, aside from the threshold parameter, the ANN becomes less sensitive to variations in the parameters as the size of the training DB increases (larger periods). Thus, it is important to regulate the ANN parameters but also to use a training DB large enough to ensure robustness in the parameters and consistency in the results. For our particular ANN we estimate that a training DB should contain a minimum of 150 fishing trips (or 4000 position records).

Inference results show that the distribution of fishing activity among regions does not have significant effect on the ANN performance. In contrast, differences in the relative composition of the fleet between training and inference DBs do have an important effect on inference performance. In order to improve the quality of inference, it is therefore necessary to develop a sampling design which ensures a representative sample of the real fleet composition in the at-sea observer data. Since 2009, the Peruvian government has implemented a new individual vessel quota system for the anchovy fishery (Aranda, 2009). The goal is to motivate ship owners to decommission less productive boats, reducing the size of the fishing fleet and increasing its economic efficiency. Changes in the composition of the fleet and in fishing strategies could affect the identification of fishing sets by the calibrated ANN. Theoretically, with sampling adjustments, a dynamic training DB should withstand this effect. If not, changes in the parameters along with new sensitivity analysis should be considered.

Although the most important indicator of ANN performance is the percentage of correct classifications (%TP), it is also essential to measure the error, to interpret it and to look for its causes. In this work, we define two principal types of errors: (1) percentage of real fishing sets not detected (false negative, %FN) and (2) percentage of non-fishing set records detected as fishing sets (false positive, %FP). Although the specific causes of these errors are unknown, we provide some possible explanations. Sometimes there are fishing-like operations which do not end in a fishing set. It can be due to “bad luck” (the fish school escaped at the last moment) or a change of mind of the captain (motivated for instance by seeing neighboring vessels fishing better at short distances or by a radio call from the company indicating to move towards a recently discovered “hotspot”). Furthermore, when two or more fishing sets occur very close in time, the ANN may recognize them as a unique long fishing set. Conversely, if a fishing operation has a long duration, the ANN might identify multiple fishing sets. Nonetheless, fishing event mapping does not show any systematic geographical pattern in the errors, providing “evidence” of their stochastic nature.

Several studies in ecology sciences have compared ANN to other techniques of classical statistics such as multiple regression, discriminant analysis and other classification and clustering methods. However, few of them have made sure that the data meet the assumptions of probability density functions, linearity, homocedasticity and independence of variables, when required (Lek et al., 1996; Thorrold et al., 1998; Chen and Ware, 1999; Engelhard et al., 2003; Hanson et al., 2004). In our case study, data do not follow any known multivariate distribution, so no parametric techniques are evaluated as alternatives. Since our priority is to obtain good inference results rather than a precise model, more complex models like Bayesian dynamic models are not considered. Still, incorporating effects of changes in fishing and in fish dynamics to our model could improve forecasting results. In that matter, recurrent neural networks (Jain and Mao, 1996) or hybrid networks with time series components should be considered. Genetic algorithms have also proven to be successful in determining the ANN architecture, and they could be thought of as alternatives or complements to the sensitivity analysis in future studies (Zhang et al., 1998; Maier and Dandy, 2000; Goethals et al., 2007).

Although theoretical and technological advances can improve the quality of the results, the key element remains the data. Two important facts should be considered: (1) the sample must be large enough in order to be representative; and (2) biased data generate biased results. Therefore, sampling criteria are essential for achieving accuracy. Firstly, a larger number of fishing trips should be sampled by at-sea observers to overcome risks of too small or unbalanced training DBs: the current sample collected by IMARPE, while being a remarkable effort compared to international peers, represents approximately 2% of the total number of fishing trips, (in practice only 1.4% suitable for the training DB). Secondly, if complete randomized sampling is economically impossible, at least some rules need to be set: as previously discussed, the fleet sample covered must be representative of the global composition of the fleet ("steel" and "wooden" fleet); and observers should have a minimum and maximum of vessels to cover (and of trips per vessel), so that the variability of the fishing characteristics is effectively represented in the training sample. Furthermore, additional information collected by observers could improve ANN performance: (1) position and duration of searching for schools (in Peru, this information is collected by the observers but not used yet), so the ANN could learn to distinguish between searching and actual fishing; and (2) exact location of departure and arrival. This latter would be useful since a vessel may go directly to the factory and the factory's position can be then considered as a port and added to the algorithm for trip recognition and thus improve data processing.

It is also indispensable to guarantee high-quality satellite information. An increase in the frequency of VMS recording provides more information on the fishing activity and thus more data for the ANN to learn from. In this case study, trips including record gaps of more than 2 h were not considered, because otherwise fishing sets could be "lost" between records. With higher frequencies (one record per half hour, or even less), the quality of the VMS data would greatly increase. Also, if there are multiple VMS providers, they should clearly state the accuracy of each position, with a common labeling system, so we are sure to work only with the most accurate ones (minimum spatial accuracy of 100 m in our case). On the other hand, availability of the vessels' official registration code in the VMS data could greatly improve the quality of the data, since we would be able to link DBs by registration code rather than just by names of the vessels.

This new information on fishing sets will directly contribute to IMARPE recommendations for adaptive management of the Peruvian anchovy fishery, which operates at weekly and daily scales. To support routine use of this tool, a friendly user interface has been developed to obtain fishing set estimation in short timescales. Interface users must be aware of the DB size requirements for accurate prediction. As far as training DB are concerned, they must contain at least 150 fishing trips. For inference, we showed that larger DBs lead to better inference. Consequently, a set of at least 100 trips should be available before proceeding to inference. For research interests, where time is less pressing, we strongly recommend strict inference, based on simultaneous training and inference DBs.

Artificial neural networks have been widely used in fisheries sciences for detection, classification, distribution, clustering, and discrimination of fish (Suryanarayana et al., 2008). For the particular goal of identifying fishing sets, we consider that the ANN presented in this study can be used for different fisheries if the input variables respond to the characteristics of the fishery. This particular ANN, which uses time of the day, speed, acceleration and change of heading as input variables, is applicable for pelagic purse seine fisheries. However, change of heading is important for other fisheries targeting species that are patchily distributed; and time is also relevant if the time of fishing is not uniformly distributed. Speed is also a variable to consider in fishing set recognition for longlines and mid-water trawls. Bottom-trawl fisheries should also include bathymetry, which is linked to species and fishing effort distribution (Fall et al., 2006). If other variables, not correlated with those discussed, provide basic information on the fishing strategies, they should also be incorporated as input variables. For practical interests, parameters have to be fixed in order to guarantee a calibrated tool (neural network) which assures good and reliable results for everyday use. Short sensitivity analyses, using the criteria presented here, are strongly recommended, especially if there are changes in the input layer. In general, sensitivity analyses should be performed from time to time in case any external or internal factor alters fishing strategies.

Finally, this tool provides new opportunities for estimating fishing effort: it provides accurate, exhaustive and spatially explicit information on fishing sets at low cost. It becomes possible to estimate spatially explicit CPUE in quasi-real time. This is critical for improving the use of CPUE data as an index of fish abundance and then avoiding the misleading interpretations of CPUE that may occur, for instance when fish and fisheries hyperaggregate (Petitgas, 1994; Rose and Kulka, 1999; Babcock et al., 2005).

## Acknowledgements

The authors would like to thank Matthieu Lengaigne for priceless discussions on artificial neural networks and sensitivity results. We are also grateful to Erich Díaz and Manuel Ochoa for valuable conversations about vessel trajectories. We warmly thank Jérémie Habasque for his help in the development of the user interface for fishing set estimation. We sincerely thank Charlotte Boyd as well as Johanna Holmgren for revising the English and for helping us make the article more readable.

## References

Aranda, M., 2009. Developments on fisheries management in Peru: the new individual vessel quota system for the anchoveta fishery. Fish. Res. 96 (2009), 308–312, doi:10.1016/j.fishres.2008.11.004.

Babcock, E.A., Pikitch, E.K., McAllister, M.K., Apostolaki, P., Santora, C., 2005. A perspective on the use of spatialized indicators for ecosystem-based fishery management through spatial zoning. ICES J. Mar. Sci. 62, 469–476, doi:10.1016/j.icesjms.2005.01.010.

Bernardos, P.G., Vosniakos, G.-C., 2007. Optimizing feedforward artificial neural network architecture. Eng. App. Artif. Intell. 20, 365–382.

Bertrand, A., Segura, M., Gutiérrez, M., Vásquez, L., 2004a. From small-scale habitat loopholes to decadal cycles: a habitat-based hypothesis explaining fluctuation in pelagic fish populations off Peru. Fish Fish. 5, 296–316.

Bertrand, S., Díaz, E., Ñiquen, M., 2004b. Interactions between fish and fisher's spatial distribution and behavior: the anchovy (*Engraulis ringens*) fishery of Peru. ICES J. Mar. Sci. 61, 1127–1136.

Bertrand, S., Burgos, J.M., Gerlotto, F., Atiquipa, J., 2005. Lévy trajectories of Peruvian purse seiners as an indicator of the spatial distribution of anchovy (*Engraulis ringens*). ICES J. Mar. Sci. 62 (3), 477–482, doi:10.1016/j.icesjms.2004.12.002.

Bertrand, S., Bertrand, A., Guevara-Carrasco, R., Gerlotto, F., 2007. Scale-invariant movements of fishermen: the same foraging strategy as natural predators. Ecol. Appl. 17 (2), 331–337, doi:10.1890/06-0303.

Bertrand, S., Diaz, E., Lengaigne, M., 2008. Patterns in the spatial distribution of Peruvian anchovy (*Engraulis ringens*) revealed by spatially explicit data. Prog. Oceanogr. 79 (2008), 379–389, doi:10.1016/j.pocean.2008.10.009.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY.

Chavez, F., Bertrand, A., Guevara-Carrasco, R., Soler, P., Csirke, J., 2008. The northern Humboldt Current System: Brief history, present status and a view towards the future. Progr. Oceanogr. 79, 95–1051.

Chen, D.G., Ware, D.M., 1999. A neural network model for forecasting fish stock recruitment. Can. J. Fish. Aquat. Sci. 56 (12), 2385–2396, doi:10.1139/cjfas-56-12-2385.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. MCSS 2, 303–314, doi:10.1007/BF02134016.

Dedecker, A.P., Goethals, P.L.M., Gabriels, W., De Puw, N., 2004. Optimization of Artificial Neural Network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium). Ecol. Modell. 174, 161–173.

Demuth, H., Beale, M., 1998. Neural Network Toolbox for Use with Matlab. The MathWorks, Inc., Natick, MA.

Deng, R., Dichmont, C., Milton, D., Haywood, M., Vance, D., Hall, N., Die, D., 2005. Can vessel monitoring system data also be used to study trawling intensity and population depletion? The example of Australia's northern prawn fishery. Can. J. Fish. Aquat. Sci. 62 (3), 611–622, doi:10.1139/F04-219.

Engelhard, G.H., Dieckmann, U., Godø, O.R., 2003. Age at maturation predicted from routine scale measurements in Norwegian spring-spawning herring (Clupea harengus) using discriminant and neural network analyses. ICES J. Mar. Sci. 60 (2), 304–313, doi:10.1016/S1054-3139(03)00017-1.

Fall, M., Samba, A., Laloë, F., 2006. Fishing tactics and strategies in coastal demersal trawling fisheries in Senegal. 307-L 316. Aquat. Living Resour. 19, doi:10.1051/alr:2007001.

Fréon, P., Bouchon, M., Mullon, C., García, C., Ñiquen, M., 2008. Interdecadal variability of anchoveta abundance and overcapacity of the fishery in Peru. Prog. Oceanogr. 79 (2008), 401–412, doi:10.1016/j.pocean.2008.10.011.

Goethals, P., Dedecker, A., Gabriels, W., Lek, S., De Pauw, N., 2007. Applications of artificial neural networks predicting macroinvertebrates in freshwaters. Aquat. Ecol. 41, 491–508, doi:10.1007/s10452-007-9093-3.
Hanson, P.J., Koenig, C.C., Zdanowicz, V.S., 2004. Elemental composition of otoliths used to trace estuarine habitats of juvenile gag Mycteroperca microlepis along the west coast of Florida. Mar. Ecol. Prog. Ser. 267, 253–265, doi:10.3354/meps267253.
Harrington, J.J., Semmens, J.M., Haddon, M., 2007. Spatial distribution of commercial dredge fishing effort: application to survey design and the spatial management of a patchily distributed benthic bivalve species. Mar. Freshwat. Res. 58, 756–764, doi:10.1071/MF06101.
Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366.
Jain, A., Mao, J., 1996. Artificial neural networks: a tutorial. Computer 29 (3), 31–44, doi:10.1109/2.485891.
Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modeling nonlinear relationships in ecology. Ecol. Modell. 90, 39–52, doi:10.1016/0304-3800(95)00142-5.
Lek, S., Guégan, J.F., 2000. Artificial neuronal networks. Application to ecology and evolution. In: Environmental Sciences. Springer-Verlag, Berlin.
Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. Environ. Model. Software 15 (2000), 101–124, doi:10.1016/S1364-8152(99)00007-9.
Mills, C.M., Townsend, S.E., Jennings, S., Eastwood, P.D., Houghton, C.A., 2007. Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. ICES J. Mar. Sci. 64, 248–255, doi:10.1093/icesjms/fsl026.
Mühlenbein, H., 1990. Limitations of multi-layer perceptron networks—steps towards genetic neural networks. Parallel Comput. 14, 249–260.
Petitgas, P., 1994. Spatial strategies of fish populations. Statistic committee C.M. 1994/d:14, ICES.
Pikitch, E.K., Santora, C., Babcock, E.A., et al., 2004. Ecosystem-based fishery management. Science 305, 346–347.
Prechelt, L., 1998. Early stopping—but when? In: Orr, G.B., Müller, K.-R. (Eds.), Neural Networks: Tricks of the Trade. Springer-Verlag, London, UK, pp. 55–69 (Series: Lecture notes in computer science. No. 1524).
Rose, G.A., Kulka, D.W., 1999. Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. Can. J. Fish. Aquat. Sci. 56 (S1), 118–127, doi:10.1139/cjfas-56-S1-118.
Sarle, W., 1995. Stopped training and other remedies for overfitting. In: Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics, Pittsburgh, PA, 21–24 June 1995, Interface Foundation, pp. 352–360.
Sarle, W., 1997. Neural network FAQ, periodic posting to the Usenet newsgroup comp.ai.neural-nets (online). Available from http://www.faqs.org/faqs/ai-faq/neural-nets/ (accessed 4 February 2008; updated 17 May 2002).
Sexton, R.S., Dorsey, R.E., Johnson, J.D., 1999. Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. Eur. J. Op. Res. 114, 589–601.
Suryanarayana, I., Braibanti, A., Rao, R.S., Ramam, V.A., Sudarsan, D., Rao, G.N., 2008. Neural networks in fisheries resources. Fish. Res. 92, 115–139, doi:10.1016/j.fishres.2008.01.012.
Thorrold, S.R., Jones, C.M., Swart, P.K., Targett, T.E., 1998. Accurate classification of juvenile weakfish Cynoscion regalis to estuarine nursery areas based on chemical signatures in otoliths. Mar. Ecol. Prog. Ser. 173, 253–265, doi:10.3354/meps173253.
Witt, M.J., Godley, B.J., 2007. A step towards seascape scale conservation: Using vessel monitoring system (VMS) to map fishing activity [online]. PLoS ONE 2 (10), e1111, doi:10.1371/journal.pone.0001111.
Zhang, G., Patuwo, E., Hu, M., 1998. Forecasting with artificial neural networks: the state of the art. Int. J. Forecast. 14, 35–62, doi:10.1016/S0169-2070(97)00044-7.

**Rocío Joo** is a statistician and a doctoral student at the French Research Institute for Development (IRD), working with the Peruvian Institute for the Sea (IMARPE). She is the main author of the article and she has made the analysis for the calibration of the neural network described in it. She is also the main developer of the user interface for obtaining fishing set estimations.

**Sophie Bertrand** is a fisheries ecologist at the French Research Institute for Development (IRD), working with IMARPE since 2001. Her research is based on the statistical analysis and modelling of spatially explicit data from exploited marine ecosystems (Vessel Monitoring System data for vessels, GPS tracking for birds and mammals, underwater acoustics for fish). She has been working with the Peruvian Vessel Monitoring System data since 2002, developing from them a series of metrics, a random walk modelling approach and the beta version of the neural network which behaviour is explored here.

**Alexis Chaigneau** is a physical oceanographer at the French Research Institute for Development (IRD). His research is focused on the circulation and mesoscale activity of the Humboldt Current System based on various satellite datasets and high-resolution model simulations. He has been working with IMARPE since 2006 and he contributed to the article mainly as an advanced programmer.

**Miguel Ñiquen** is a fisheries ecologist at the Peruvian Institute for the Sea (IMARPE). He is the head of the IMARPE's department in charge of monitoring and managing the Peruvian industrial fisheries, among which, the Peruvian anchovy purse seine fishery. He is the main daily user of the tool described in the paper and as such has provided its detailed book of specifications.