

Comparative performance of generalized additive models and boosted regression trees for statistical modeling of incidental catch of wahoo (*Acanthocybium solandri*) in the Mexican tuna purse-seine fishery

Raul O. Martínez-Rincón^{a,*}, Sofía Ortega-García^b, Juan G. Vaca-Rodríguez^{c,d}

^a Becario PIFI-CONACyT, Centro Interdisciplinario de Ciencias Marinas, Departamento de pesquerías, Av. Instituto Politécnico Nacional s/n Col. Playa Palo de Santa Rita Apdo. Postal 592 C.P. 23096 La Paz, Mexico

^b Becario COFAA, Centro Interdisciplinario de Ciencias Marinas, Departamento de pesquerías, Av. Instituto Politécnico Nacional s/n Col. Playa Palo de Santa Rita Apdo. Postal 592 C.P. 23096 LaPaz, Mexico

^c Programa Nacional de Aprovechamiento del Atún y de Protección de Delfines (PNAAPD), Km 107 Carretera Tijuana-Ensenada, Campus CICESE, Ensenada, Baja California, Mexico

^d Facultad de Ciencias Marinas, Universidad Autónoma de Baja California (UABC), Km 103 Carretera Tijuana-Ensenada, Campus UABC, Ensenada, Baja California, Mexico

ARTICLE INFO

Article history:

Received 23 July 2011

Received in revised form 27 February 2012

Accepted 4 March 2012

Available online 5 April 2012

Keywords:

Eastern Pacific Ocean

Chlorophyll-a

Sea-surface temperature

Floating objects

ABSTRACT

Wahoo (*Acanthocybium solandri*) is distributed in all tropical and subtropical oceans and caught incidentally by the tuna purse-seine fishery in the Eastern Pacific Ocean (EPO). Generalized additive models (GAM) and boosted regression trees (BRT) were used to analyze relationships between presence of wahoo in logbook data from the Mexican tuna purse-seine fishery with environment, geographic area and set type (unassociated, associated with dolphins or floating objects set). Model performance was evaluated using changes in deviance in the fitted models and the area under the receiver operating characteristic curve (ROC). Results indicate little difference between the performance of GAM and BRT models. Both methods were consistent with predictions of presence of wahoo with respect to the variables used. Set type was the single most important predictor of variation in presence of wahoo, with highest probability of incidental catch in sets made on floating objects. With respect to environmental factors, sea surface temperature (20–25 °C) and chlorophyll-a concentration (<2 mg m⁻³) determined the highest probability of incidental catch of wahoo. The coast of Baja California Sur, Mexico and south of the equator were predicted to have a high probability of incidental catch of wahoo.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The wahoo, *Acanthocybium solandri* (Scombridae), is a highly migratory pelagic species distributed in tropical and subtropical waters worldwide (Collette, 2002; Oxenford et al., 2003). Larval and adult wahoo occur in both oceanic and coastal waters of the Pacific Ocean between latitudes of 30°N and 25°S. Key habitats for the species include island margins, oceanic shoals, thermal fronts and current margins (Garber et al., 2001). Wahoo also tend to associate with natural flotsam or artificial fish aggregating devices (FADs) and are incidentally caught in fisheries targeting these structures (Laurans et al., 1999; Reynal et al., 1999). Spawning for this species has been observed throughout the year in equatorial waters between 14°N and 15°S (Matsumoto, 1967).

Wahoo are exploited by commercial fisheries worldwide. Most of commercial catches of wahoo is originated from the Atlantic Ocean, but in recent years, fisheries in Pacific Ocean has been

increasing (FAO, 2011). Recreational catch is largely unknown, but may significant throughout most of their geographic range and potentially exceed that of commercial catches in some regions (Zischke, 2012). In the Eastern Pacific Ocean (EPO) wahoo are incidentally caught and retained as a byproduct in the purse-seine tuna fishery, which primarily sets over unassociated schools, over dolphin-associated tuna schools, or around natural (e.g., algae, animal carcasses, logs, etc.) or artificial (e.g., FADs) floating objects. They are also exploited by recreational fisheries, with limited management arrangements in place for the species.

There are increasing concerns regarding the long-term sustainability of some species caught incidentally as bycatch in tuna seine fisheries. This is primarily due to limited information on the biology, habitat preference and distribution of these species due to their relatively low importance compared to the primary tuna species, which has hindered quantitative population assessments to guide management.

In this study we compare the performance of generalized additive models (GAM) and boosted regression trees (BRT) to determine the environmental and geographical variables that influence the distribution of wahoo.

* Corresponding author. Tel.: +52 612 1234658; fax: +52 612 122 53 22.
E-mail address: rmartinezr0604@ipn.mx (R.O. Martínez-Rincón).

GAM are perhaps the most widely used statistical modeling tools to analyze relationships between the distributions of the species and their environment, particularly in terrestrial and marine studies. This method is based on the use of non-parametric smoothing functions that allows flexible description of complex species responses to environment (Leathwick et al., 2006). BRT has rarely been used in ecology (Moisen et al., 2006; Elith et al., 2008). The BRT approach differs from traditional regression methods that produce the 'best' model, instead using the technique of boosting to combine large numbers of relative simple tree models and optimize predictive performance (Elith et al., 2008).

2. Data and methods

2.1. Database

2.1.1. Catch data

Observer data from Mexico's National Program for the Exploitation of Tuna and Protection of Dolphins (Programa Nacional de Aprovechamiento del Atún y Protección de los Delfines, PNAAPD) were filtered for incidental catch of wahoo (which corresponds to ~50% of fishing operations carried in the EPO from 1998 to 2007) and used for the analyses in this study. Records in this database include the date and time (start and finish) of sets, geographical position (latitude and longitude), estimated number of fish caught by species and set type, and the manner in which tunas were detected and targeted: association with floating objects, association with herds of dolphins, and as free-swimming (unassociated) schools visible at the surface.

2.1.2. Environmental data

Environmental data consisted of the mean monthly values of sea surface temperature (SST), sea surface height (SSH), chlorophyll-a concentration, and Oceanic Niño Index (ONI). Data were obtained from <http://coastwatch.pfeg.noaa.gov/erddap/index.html> and <http://www.cpc.ncep.noaa.gov/>.

2.1.3. Datasets

Database was randomly divided in two dataset: training set (70%) and test set (30%). The training set was used to build the models and the test set was used for testing model performance.

2.2. Models

Incidental catch of wahoo was the response variable for the models and was measured in units of presence/absence (1/0). The predictor variables used were: environmental variables (SST, SSH, chlorophyll-a, and ONI); spatial variables (longitude and latitude); and set type (Table 1). Models were constructed using all predictor variables and no interactions terms were used in order to better compare between models.

2.2.1. Generalized additive models

GAM is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates (Wood, 2006). All GAM were fitted in R (R Development Core Team, 2011) version 2.13.0, using mgcv package version 1.7–5 (Wood, 2011), assuming a binomial error distribution. The test dataset was used as independent data to predict the incidental catch of wahoo using the final model constructed with training dataset.

2.2.2. Boosted regression trees

All BRT models were fitted in R using the gbm and dismo packages (Ridgeway, 2010; Hijmans et al., 2011). For BRT, model fitting requires the specification of three parameters: (a) learning rate, which controls the rate at which model complexity is increased;

(b) the number of trees (even though BRT are largely resistant to model over-fitting, is still necessary to determine the optimum number of trees) and (c) the number of splits in each tree (also called the interaction depth in gbm), which controls the size of the trees. A value of one corresponds to an additive model with non interacting variables where each tree consists of a single node or decision rule. A value of two indicates that two nodes are used in each tree, which corresponds to a model with two-way interactions, and so on. Error was assumed to fit a Bernoulli distribution. For a cogent and comprehensive account of the development and application of BRT, the reader is referred to Elith et al. (2008).

Formulae developed by Friedman (2001) were implemented in the gbm library to estimate the relative influence of predictor variables. The measures are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative contribution (or influence) of each variable is scaled so that the sum adds to 100, with higher numbers indicating stronger influence on the response.

2.2.3. Relationship between response and predictor variables

The relationships between presence of wahoo and each predictor variable explained by GAM and BRT models were plotted using partial dependence plots (i.e., the effect of a variable on wahoo presence after accounting for the average effects of all other variables in the model (Friedman, 2001)). The fitted functions from partial dependence plots provide an indication of how presences of wahoo in catches depend on each predictor variable.

2.2.4. Comparative performance

Comparison of the performance of GAM and BRT models was carried out using deviance explained, percent correctly classified (PCC), sensitivity (i.e., proportion of observed positives correctly predicted), specificity (proportion of observed negatives correctly predicted) and the area under the receiver operator characteristic curve (ROC which gives a measure of the degree to which fitted values discriminate between observed values). PresenceAbsence library was used (Freeman, 2007) to get the last four values. This library provides a set of functions useful when evaluating the results of presence-absence models. Because the test data set was used for testing the performance of models, deviance explained is not available.

2.2.5. Spatial predictions

The test dataset was used to make the spatial predictions of GAM and BRT models for incidental catch of wahoo in EPO from 1998 to 2007.

3. Results

3.1. Generalized additive model

The final GAM for presence/absence of wahoo was constructed using the training dataset. This model explained 36% of total deviance, and the component smooth functions of the fitted model are shown in Fig. 1. The model predicts that the probability of incidental catch for wahoo is high in sets made on floating objects, when SST is in a range of 18–26 °C, at lower chlorophyll-a concentrations, and when sets are made at south of the equator. With respect to SSH, longitude and ONI, the relationship between these variables and incidental catch of wahoo was poor and no pattern was observed.

Table 1
Environmental variables used to model fish occurrence.

Variable	Description	Mean (range)
Longitude	Longitude were set was made	110.3°W (144, 77)
Latitude	Latitude were set was made	14.25°N (−14, 32)
SST	Satellite-image based estimate of SST	26.9 °C (16.01, 31.93)
SSH	Satellite-image based estimate of SSH	−0.01 m (−0.25, 0.42)
ONI	Oceanic Niño Index	0.04 (−1.60, 2.30)
Chlorophyll-a concentration	Satellite-image based estimate of chlorophyll-a concentration	0.23 mg m ^{−3} (0.03, 8.57)
Set type	Set type in three classes: unassociated, dolphin and floating object	NA

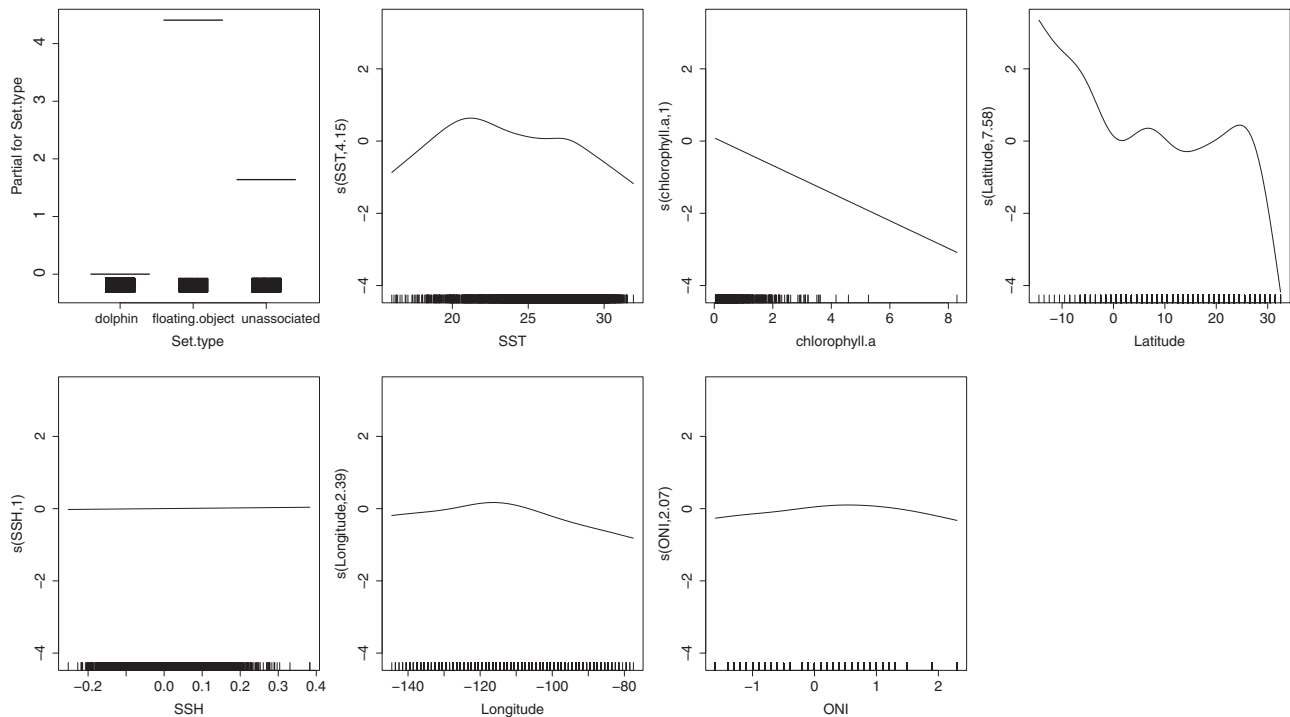


Fig. 1. Variation in wahoo presence predicted by a non interaction GAM model using environmental (SST, chlorophyll-a), spatial (latitude) and fishing characteristics (set type) as predictors. Shaded regions indicate the 5–95% confidence intervals for the predicted values.

3.2. Boosted regression trees

The final BRT model was constructed using the training dataset with learning rate value of 0.05, 5000 trees and tree complexity of 1. Not interaction between variables was allowed using this criterion. This model explained 47.3% of total deviance. Table 2 summarizes the relative contributions of the predictor variables in the final model.

The most significant predictor variable was set type, contributing 58.3% to the overall model. Environmental variables (SST and chlorophyll-a) had smaller, yet still important contributions to the model, ranging between 10 and 14%, suggesting important habitat preferences for wahoo. Latitude was the spatial variable that had

Table 2
Summary of the predictor variables and their relative contributions (%) to the BRT model.

Predictor variable	Relative contribution (%)
Set type	58.3
SST	14.1
Chlorophyll-a	10.7
Latitude	7.8
SSH	4.0
Longitude	2.9
ONI	2.2

the highest contribution to the model. SSH, longitude and ONI had smaller contributions to the overall model.

Fig. 2 shows the relationships between presence of wahoo and each predictor variable explained with BRT model. Set type was the single most important predictor of variation of presence of wahoo, with the highest probability of incidental catch being from floating objects, and lowest probability of incidental catch on sets associated to dolphins and unassociated schools. Environmental parameters predicted that wahoo have a highest probability of incidental catch for waters in a range of SST between 20 and 25 °C, and low chlorophyll-a concentrations (<2 mg m^{−3}). Incidental catch of wahoo with has the highest probability in sets made at latitudes south of the equator.

3.3. Predictive performance

The comparison of predictive performance between GAM and BRT indicates that the practical significance of any differences between the models is minimal. All standard accuracy measures had similar values (Table 3). For the test dataset GAM and BRT achieved 90% at the point of equal sensitivity/specificity (ROC values). The main difference observed between these models was in deviance explained, suggesting that BRT model have better predictive performance than GAM.

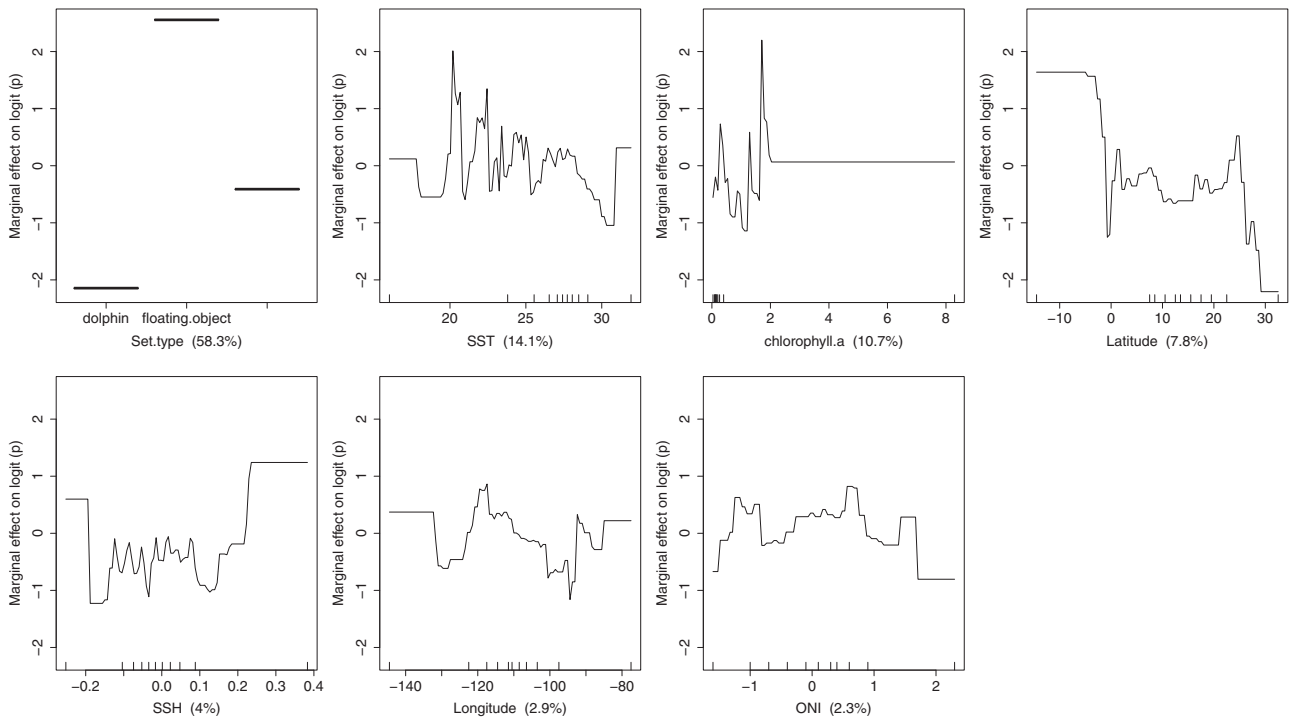


Fig. 2. Partial dependence plots for the six most influential variables in the model for wahoo. For explanation of variables and their units see Table 1. Y axes are on the logit scale and are centered to have zero mean over the data distribution. X axes represent the six most important predictor variables and their relative contribution. Rug plots inside the bottom of the X axes show the distribution of sites across that variable, in deciles.

3.4. Spatial distribution of incidental catches of wahoo

Fig. 3 shows the spatial prediction of the probability of incidental catch of wahoo as predicted by GAM and BRT models. Both models had very similar results in the spatial predictions. For both models the highest probability of incidental catch of wahoo was predicted to occur in two regions: (1) off eastern Baja California Sur, Mexico (21–26°N) and (2) both oceanic and coastal waters south of the equator (Fig. 3). Both areas had similar environmental and set data, with moderate to high presence of floating objects, SST ranging from 17 to 28 °C and low mean values of chlorophyll-a concentrations (0.21–0.35 mg m⁻²).

4. Discussion

The results indicated that there is a high level of predictability in the relationship between presence of wahoo as incidental catch in EPO with set type, environment, and spatial predictors. For both BRT and GAM, set type was the strongest predictor, with a much higher probability of incidental catch of wahoo in sets made on floating objects. Interestingly, less than 3% of sets made by Mexican purse-seine vessels were on floating objects, in contrast to the ~25% made by the international fleet on this set type (IATTC, 2010).

It is not surprising that there is a higher probability of incidental catch of wahoo in sets on floating objects, as previous research has suggested an affiliation for wahoo to occur around floating objects (either natural or artificial) (Bailey et al., 1996; Romanov, 2002; Oxenford et al., 2003; Maunder and Harley, 2006; IATTC, 2010). While the advantages of this affiliation are unknown, Dagorn and Fréon (1999) suggest that floating objects may act as a meeting point for several species, thus providing social advantages such as enhancing schooling behavior.

Sea surface temperature had the highest contribution to the BRT model. We observed a higher probability of incidental catches of wahoo in waters with SST between 20 and 25 °C in both models. A similar range of temperatures have been described as preferred by the species from electronic tagging studies. Using archival tags, Sepulveda et al. (2011) found that wahoo off Baja California Sur spend the majority of time within water temperatures between 23 and 26 °C. Theisen (2007) observed a preferred SST range of 20–25 °C using pop-up satellite archival tags on wahoo off the Atlantic US coast.

Chlorophyll-a concentration was the second environmental variable in importance to BRT model. Brill and Lutcavage (2001) suggest that chlorophyll-a concentration is an indirect surrogate measure of forage abundance for large pelagic fishes. In contrast, BRT and GAM models predicted high probability of incidental catch of wahoo when values of chlorophyll-a were lower than 2 mg m⁻³. Similar results were reported by Su et al. (2008), where they found that the highest relative densities of blue marlin (*Makaira nigricans*) in Pacific Ocean occur when chlorophyll-a is relatively low (<0.2 mg m⁻³). Sartimbul et al. (2010) found a direct correlation between chlorophyll-a and CPUE of *Sardinella lemuru* after applying a 3-month moving average, due to the time that is needed to transfer chlorophyll-a to *S. lemuru* via the food chain. We posit that the high probability of incidental catch of wahoo in environments with low primary productivity may be explained due to a similar lag in the food chain.

Table 3
Standard accuracy measures for both models.

Standard accuracy vales	GAM		BRT	
	Training set	Test set	Training set	Test set
Deviance explained (%)	36.00	–	47.34	–
PCC	0.963	0.960	0.965	0.961
Sensitivity	0.164	0.171	0.182	0.187
Specificity	0.995	0.993	0.996	0.994
ROC	0.881	0.902	0.897	0.906

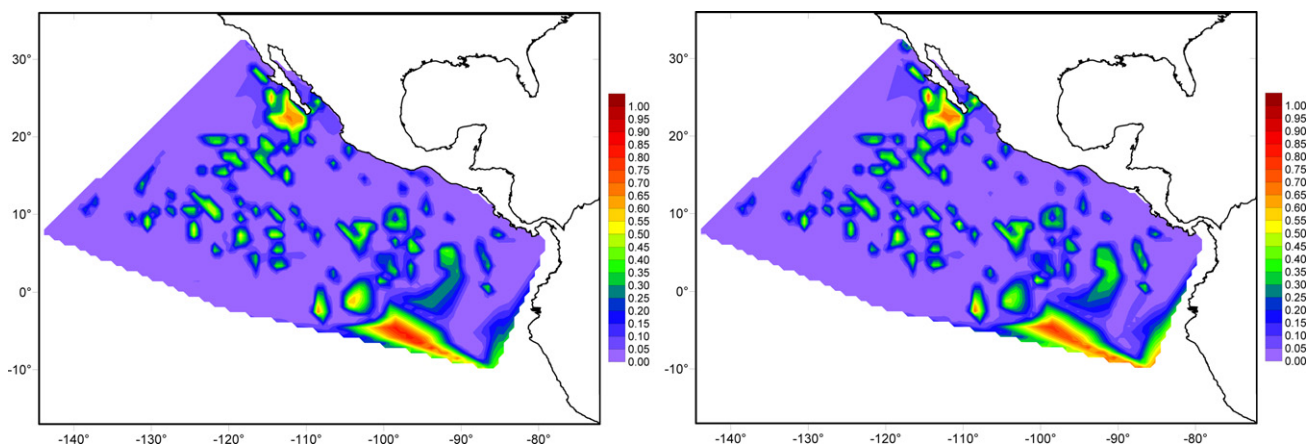


Fig. 3. Probability of incidental catch of wahoo across the fishery area as predicted from GAM (left) and BRT (right) models.

The BRT model predicted a higher probability of incidental catch of wahoo areas with positive SSH (over 20 cm). Sea surface height has been used as a proxy for upwelling and eddies, with negative values of SSH relating to oceanic features such as upwelling, presence of gyres and a shallow mixed layer (Forsbergh, 1969; Bakun, 1996; Domokos et al., 2007). If negative values of SSH relate to higher productivity, these areas should see higher pelagic fish abundance. While investigating this was beyond the scope of this study, higher incidental catch of wahoo at positive SSH may represent a similar lag in the food chain at these low productivity areas.

In this study we found large-scale spatial variation in probability of incidental catch of wahoo in the EPO. We observed that there are two regions with higher probability of presence of the species. In the north, the region located off the east-coast of Baja California Sur, Mexico (21–26°N), was predicted to have higher probability of incidental catch of wahoo. Sepulveda et al. (2011), mention that wahoo play an important role in recreational fisheries in this region, supporting private and charter vessels as well as commercial passenger fishing vessels operating out of San Diego, California. They also suggest that species has some degree of seasonal fidelity for sea mounts in the region, as revealed through archival tagging. Furthermore, we suggest that this region is important for wahoo because it has moderate quantities of floating objects – dominated by seaweed (Solana-Sansores, 2001a; Martínez-Rincón et al., 2009) – and favorable environmental conditions.

At south of the equator, a higher probability of incidental catch of wahoo was predicted in both coastal and oceanic waters. Important incidental catches of sharks and pelagic fishes also occur in this region, particularly from purse-seine sets made on floating objects (Solana-Sansores, 2001b). This region of EPO is strongly influenced by the Peruvian Coastal Upwelling (PCU) and the south equatorial current (SEC) (Pennington et al., 2006). The PCU region has been relatively well studied, primarily because it supports the largest-tonnage fishery in the world, yielding 6–12 million metric tons of anchoveta (*Engraulis ringens*) annually (FAO, 1993). This region is defined by low sea surface temperatures and high rates of primary production, most strongly developed from 4 to 16°S and to 100 km offshore (Pennington et al., 2006). The SEC has cooled water through influx of the Peru Current with chlorophyll-*a* and productivity levels in this region relatively lower when compared with equatorial upwelling and the PCU.

The prediction of incidental catch of wahoo using presence/absence data was satisfactory in our analyses, with almost 50% variation explained by the predictor variables used in BRT models. We found that GAM and BRT models had similar predictive performance and the accuracy of predictions made by models were

almost identical. Previous studies (Leathwick et al., 2006; Moisen et al., 2006; Elith et al., 2008) have found that BRT had superior predictive performance than GAM models. This is likely due to the BRT models used in these studies were set to explore more complicated interactions in predictor variables. However when both models were set without interactions between predictor variables, as is the case with this study, both models have similar results.

5. Conclusions

Our analyses indicate that while there is strong association between the incidental catch of wahoo and set type (mainly those sets made on floating object), environmental variables such as SST (20–25 °C) and chlorophyll-*a* concentrations (<2 mg m⁻³) also contribute strongly to the presence of wahoo in purse-seine catches. This information may be useful as a baseline for monitoring incidental catch of wahoo in the EPO, however consideration of inter-annual variation within environment and long-term means would be required. We consider GAM and BRT models are good statistical techniques for modeling non-target species in fisheries, because their ability to model non-linear effects between response and predictor variables.

Acknowledgments

We are very grateful to Mr. Mitchell Zischke (The University of Queensland) and Dr. Shane Griffiths (CSIRO Division of Marine Research) for comments and assistance with English that greatly improved the manuscript. Also we like to acknowledge two anonymous reviewers for providing constructive feedback. First author thank to Consejo Nacional de Ciencia y Tecnología and Programa Institucional de Formacion de Investigadores for the scholarships given.

References

- Bailey, K., Williams, P.G., Itano, D., 1996. Bycatch and discards in Western Pacific tuna fisheries: a review of SPC data holdings and literature. South Pac. Comm. Tech. Rep. 34. Noumea, New Caledonia, 171 pp.
- Bakun, A., 1996. Patterns in the Ocean. Ocean Processes and Marine Population Dynamics. California Sea Grant College System, La Jolla, 323 pp.
- Brill, R.W., Lutcavage, M.E., 2001. Understanding environmental influences on movements and depth distributions of tunas and billfishes can significantly improve population assessments. In: Sedberry, G. (Ed.), Islands in the Stream: Oceanography and Fisheries of the Charleston Bump, vol. 25. American Fisheries Society Symposium, Bethesda, MD, pp. 179–198.
- Collette, B.B., 2002. Scombridae. In: Carpenter, K.E. (Ed.), The Living Marine Resources of the Western Central Atlantic. Volume 2: Bony Fishes Part 2 (Opistognathidae to Molidae), Sea Turtles and Marine Mammals. FAO Species

- Identification Guide for Fishery Purposes and American Society of Ichthyologists and Herpetologists Special Publication No. 5, pp. 1701–1722.
- Dagorn, L., Fréon, P., 1999. Tropical tuna associated with floating objects: a simulation study of the meeting point hypothesis. *Can. J. Fish. Aquat. Sci.* 56 (6), 984–993.
- Domokos, R., Seki, M.P., Polovina, J.J., Hawn, D.R., 2007. Oceanographic investigation of the American Samoa albacore (*Thunnus alalunga*) habitat and longline fishing grounds. *Fish. Oceanogr.* 16 (6), 555–572.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77 (4), 802–813.
- FAO, 1993. Fishery statistics. Commodities [1991], vol. 73. Food and Agricultural Organization of the United Nations, Rome.
- FAO, 2011. FISHSTAT Plus Universal Software for Fishery Statistical Time Series. Version 2.3. FAO Fisheries Department, Fishery Information, Data and Statistics Unit.
- Forsbergh, E.D., 1969. On the climatology, oceanography and fisheries of the Panama bight. *IATTC* 14 (2), 49–385.
- Freeman, E., 2007. PresenceAbsence: An R Package for Presence–Absence Model Evaluation. USDA Forest Service, Rocky Mountain Research Station, 507 25th street, Ogden, UT, USA.
- Friedman, J.H., 2001. Greedy function approximation: the gradient boosting machine. *Ann. Stat.* 29, 1189–1232.
- Friedman, J.H., Meulman, J.J., 2003. Multiple additive regression trees with application in epidemiology. *Stat. Med.* 22, 1365–1381.
- Garber, A.F., Stuck, K.C., Franks, J.S., Garber, N.M., Geter, D.R., 2001. Characterization of the mitochondrial DNA control region of the wahoo, *Acanthocybium solandri*, from the northcentral Gulf of Mexico and Bimini, Bahamas. *GCFI* 52, 610–621.
- Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2011. Dismo: Species Distribution Modeling. R Package Version 0.6–10., <http://CRAN.R-project.org/package=dismo>.
- IATTC, 2010. Annual Report of the Inter-American Tropical Tuna Commission, 2008. Inter-American Tropical Tuna Commission, La Jolla, CA, 100 pp.
- Laurans, M., Taquet, M., Reynal, L., Lagin, A., 1999. Comparaison des captures des grands poissons pélagiques en Martinique avec ou sans DCP. In: Le Gall, J.-Y., Cayre, P., Taquet, M. (Eds.), *Peche thoniere et dispositifs de concentration de poissons*, Colloque Caraïbe-Martinique. Troishets, Martinique, October, vol. 15–19, pp. 421–434.
- Leathwick, J.R., Elith, J., Francis, M.P., Hastie, T., Taylor, P., 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Mar. Ecol. Prog. Ser.* 321, 267–281.
- Martínez-Rincón, R.O., Ortega-García, S., Vaca-Rodríguez, J.C., 2009. Incidental catch of dolphinfish (*Coryphaena* spp.) reported by the Mexican tuna purse seiners in the eastern Pacific Ocean. *Fish. Res.* 96, 296–302.
- Matsumoto, W.M., 1967. Morphology and distribution of larval wahoo *Acanthocybium solandri* (Cuvier) in the central pacific ocean. *Fish. Bull.* 66 (2), 299–322.
- Maunder, M., Harley, S.J., 2006. Evaluating tuna management in the Eastern Pacific Ocean. *Bull. Mar. Sci.* 78 (3), 593–606.
- Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E., Edwards, T.C., 2006. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecol. Model.* 199 (2), 176–187.
- Oxenford, H.A., Murray, P.A., Luckhurst, B.E., 2003. The biology of wahoo (*Acanthocybium solandri*) in the western central Atlantic. *Gulf Caribbean Res.* 15, 33–49.
- Pennington, J.T., Mahoney, K.L., Kuwahara, V.S., Kolber, D.D., Calienes, R., Chavez, F.P., 2006. Primary production in the eastern tropical Pacific: a review. *Prog. Oceanogr.* 69, 285–317.
- R Development Core Team, 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.R-project.org/>.
- Reynal, L., van Buurt, G., Taquet, M., 1999. Perspectives de développement des DCP ancres dans les Petites Antilles. L'exemple de trois îles: Guadeloupe, Martinique et Curaçao. In: Le Gall, J.-Y., Cayre, P., Taquet, M. (Eds.), *Peche thoniere et dispositifs de concentration de poissons*, Colloque Caraïbe-Martinique. Trois-Îlets, Martinique, October vol. 15–19, pp. 36–54.
- Ridgegway, G., 2010. gbm: Generalized Boosted Regression Models. R Package Version 1.6–3.1., <http://CRAN.R-project.org/package=gbm>.
- Romanov, E.V., 2002. Bycatch in the tuna purse-seine fisheries of western Indian Ocean. *Fish. Bull.* 100 (1), 90–105.
- Sartimbul, A., Nakata, H., Rohadi, E., Yusuf, B., Kadarisman, H.P., 2010. Variations in chlorophyll-a concentration and the impact on *Sardinella lemuru* catches in Bali Strait, Indonesia. *Prog. Oceanogr.* 87, 168–174.
- Sepulveda, C.A., Aalbers, S.A., Ortega-García, S., Wegner, N.C., Bernal, D., 2011. Depth distribution and temperature preferences of wahoo (*Acanthocybium solandri*) off Baja California Sur, Mexico. *Mar. Biol.* 158 (4), 917–926.
- Solana-Sansores, R., 2001a. Los objetos flotantes del pacífico oriental: formas, distribución espacial y cambios temporales. *Cienc. Mar.* 27 (3), 423–443.
- Solana-Sansores, R., 2001b. Patrones espaciales de las especies epipelágicas capturadas incidentalmente en la pesca del atún sobre objetos flotantes, en el océano pacífico oriental. *Cienc. Mar.* 27 (3), 445–479.
- Su, N.J., Sun, C.L., Punt, A.E., Yeh, S.Z., 2008. Environmental and spatial effects on the distribution of blue marlin (*Makaira nigricans*) as inferred from data for longline fisheries in the Pacific Ocean. *Fish. Oceanogr.* 17 (6), 432–445.
- Theisen, T.C., 2008. Population genetic structure, movement patterns, and environmental preferences of the pelagic wahoo, *Acanthocybium solandri*. Dissertation. Florida Atlantic University.
- Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.
- Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc.* 73 (1), 3–36.
- Zischke, M., 2012. A review of the biology, stock structure, fisheries and status of wahoo (*Acanthocybium solandri*), with reference to the Pacific Ocean. *Fish. Res.* 119–120, 13–22.