SCRS/2015/106 Collect. Vol. Sci. Pap. ICCAT, 72(2): 406-414 (2016)

STANDARDIZATION OF CATCH RATES FOR THE EASTERN TROPICAL ATLANTIC BIGEYE TUNA CAUGHT BY THE FRENCH PURSE SEINE DFAD FISHERY

I. Katara[1], D. Gaertner[1], A. Maufroy[1], E. Chassot[2]

SUMMARY

The drifting Fish Aggregating Device (dFAD) purse seine fishery is complex and fishing effort depends on a multitude of factors. Traditional indices of fishing effort such as searching time are meaningless for this fishery. We composed a comprehensive list of 28 candidate variables that describe the dFAD fishery and used them as predictors of fishing effort in bigeye tuna CPUE standardization of the French purse seiners operating in the Eastern Atlantic Ocean during 2007-2013. We performed variable selection using penalized maximum likelihood in GLM and GLMM frameworks, aiming to improve prediction accuracy and interpretability of the selected models. We applied the Lasso (Least Absolute Shrinkage and Selection Operator) regression models to derive the true parsimonious model, because the number of candidate independent variables is large compared to the number of observations. The penalized model selection process retained explanatory variables such as: the skipper, the vessel, the price of targeted tuna species, the density and spatial distribution of FADs and the number/type of deployed buoys. The inclusion of these predictors in CPUE standardization models provided realistic estimates of uncertainty. We propose the systematic collection of selected explanatory variables and their usage in dFAD related tuna CPUEs standardization in a mixed model framework.

RÉSUMÉ

La pêche à la senne ayant recours aux dispositifs de concentration des poissons dérivants (DCPd) est complexe et l'effort de pêche dépend d'une multitude de facteurs. Les indices traditionnels de l'effort de pêche tels que le temps de recherche sont dénués de sens pour cette pêcherie. Nous avons dressé une liste exhaustive de 28 variables potentielles qui décrivent la pêche sous DCPd et les avons utilisées comme prédicteurs de l'effort de pêche pour standardiser la CPUE du thon obèse des senneurs français opérant dans l'océan Atlantique oriental entre 2007 et 2013. Nous avons effectué la sélection de variables en utilisant la vraisemblance maximale pénalisée dans les cadres de GLM et GLMM, dans le but d'améliorer l'exactitude des prévisions et l'interprétabilité des modèles sélectionnés. Nous avons appliqué les modèles de régression de Lasso (Least Absolute Shrinkage and Selection Operator) pour obtenir le vrai modèle parcimonieux, car le nombre de variables indépendantes potentielles est élevé par rapport au nombre d'observations. Le processus de sélection du modèle pénalisé a retenu des variables explicatives telles que le capitaine, le navire, le prix des espèces de thonidés ciblées, la densité et la distribution spatiale des DCP et le nombre / type de bouées déployées. L'inclusion de ces prédicteurs dans les modèles de standardisation de la CPUE a fourni des estimations réalistes de l'incertitude. Nous proposons la collecte systématique des variables sélectionnées explicatives et leur utilisation pour standardiser les CPUE de thonidés se rapportant aux DCPd dans un cadre d'un modèle mixte.

RESUMEN

La pesquería de cerco con dispositivos de concentración de peces a la deriva (DCPd) reviste una gran complejidad y el esfuerzo pesquero depende de multitud de factores. Los índices tradicionales del esfuerzo pesquero, como tiempo de búsqueda, pierden su razón de ser en esta pesquería. Se ha creado una lista exhaustiva de 28 posibles variables que describen la pesquería con DCPd y las hemos utilizado como predictores del esfuerzo pesquero en la estandarización de la CPUE de patudo de la pesquería de cerco francesa que operó en el

[1] IRD (UMR MARBEC) Av. Jean Monnet, CS 30171, 34203, Sète, France.
[2] IRD (UMR MARBEC), SFA, BP 570 Victoria, Seychelles.

océano Atlántico oriental desde 2007 hasta 2013. Se ha realizado una selección de variables utilizando la máxima verosimilitud penalizada en marcos GLM y GLMM, para mejorar la precisión de la predicción y la capacidad e interpretación de los modelos seleccionados. Se aplicaron modelos de regresión Lasso (Least Absolute Shrinkage and Selection Operator) para derivar el modelo austero real, porque el número de posibles variables independientes es muy amplio en comparación con el número de observaciones. El proceso de selección del modelo penalizado mantuvo variables explicativas como el patrón, el buque, el precio de las especies de túnidos objetivo, la densidad y distribución espacial de los DCP y el número/tipo de boyas plantadas. La inclusión de estos predictores en los modelos de estandarización de la CPUE proporcionó estimaciones realistas de la incertidumbre. Se propuso una recopilación sistemática de las variables explicativas seleccionadas, así como su utilización en la estandarización de la CPUE para los túnidos relacionada con los DCPd en un marco de modelo mixto.

KEYWORDS

Bigeye tuna, CPUE, Catch/effort, Mixed models, Purse seining

1. Introduction

Indicators of abundance constitute a vital component of fish stock assessments (Hilborn and Walters, 1992). For tropical tuna, abundance indices are derived through the standardization of Catch per Unit of Effort (CPUE) from commercial data. Knowledge of catchability is necessary to acquire an unbiased indicator of abundance over time (Wilberg *et al*., 2009). In the case of the tropical tuna purse seine fisheries, catchability has drastically increased due to the increasing use of drifting Fish Aggregating Devices (dFADs) since the early 1990s *(Ariz et al*., 1992; Fonteneau *et al*., 2013) and their rapid technological development. dFADs equipped with electronic devices (i.e. GPS and echo-sounders; Lopez *et al.*, 2014; Torres-Irineo *et al*., 2014) have broken the link between searching time and effective fishing effort (Fonteneau *et al*., 1999). For juvenile yellowfin (*Thunnus albacares*), juvenile bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*) that are mainly caught by surface fisheries, there is currently no reliable dFAD-fishing CPUE index to evaluate the trend in abundance (e.g. assessment of the WCPO SKJ; IOTC, 2014)).

For these reasons the European Union has implemented a research project called "Catch, Effort, and eCOsystem impacts of FAD-fishing" (CECOFAD) to provide reliable estimates of abundance indices and accurate indicators on the impact of dFAD-fishing on juvenile bigeye tuna, juvenile yellowfin tuna, skipjack, and bycatch (Gaertner *et al*., 2015). In this framework, we attempt to develop a reliable CPUE index for bigeye tuna caught in association with dFADs by the French purse seiners operating in the Eastern Atlantic. Statistical procedures of model selection were adapted to the large number of available explanatory variables and the hierarchical structure of the data. We discuss the contribution of different explanatory variables to CPUE standardization and suggest a statistical framework to incorporate significant variables in CPUE standardization models.

2. Materials and Methods

2.1 Data

The *Institut de Récherche pour le Dévelopemen*t (IRD) has been collecting catch and effort data for the French purse seiners operating in the Atlantic Ocean since the early 1960s (Postel, 1969). Here, we used logbooks of the French purse seiners for the period 2007-2013 (Chassot *et al*., 2015) and restricted our analysis to fishing sets documented as related with floating objects, i.e. logs and dFADs. We put together a comprehensive list of candidate predictors to describe changes in fishing efficiency, provided either directly from the French purse seiner tuna association ORTHONGEL (2007-2013) or from scientific literature (**Table 1**). The final dataset consists of 3,838 logbook records**. Table 2** gives a thorough description of the candidate predictors.

2.2 CPUE Standardization

BET CPUE follows a skewed zero-inflated distribution and as such were standardized in a two stages delta-lognormal GLM (Fletcher *et al*., 2005; Lo *et al.*, 1992). At stage 1, the probability of a positive BET CPUE C_i, for i = 1,…,N, is modeled as a linear combination of x_i predictors with a_i coefficients:

$\mathbf{Pr}(\boldsymbol{C_i} > 0) = \boldsymbol{InvLogit}(\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{a_i x_i})$.

At stage 2 the mean size of positive CPUEs C_j, for j=1,…,M < N, is modeled as a linear combination of the same x_j predictors with b_i coefficients, following a lognormal distribution:

$$\boldsymbol{C_j} = \mathbf{exp}(\sum_{\boldsymbol{j=1}}^{\boldsymbol{M}} \boldsymbol{b_j x_j})$$

Our dataset comprises of less than 4,000 records, 28 continuous and categorical predictors, and more than 68 parameters to estimate (in a GLMM). For multidimentional data such as these, variable or model subset selection through stepwise selection becomes problematic. The number of possible models grows exponentially with the number of predictors and renders computation infeasible (Kuo and Mallick, 1998). Moreover, when the number of observations is not much larger than the number of predictors, ordinary least squares may result in over-fitting. Penalized maximum likelihood methods allow regression modeling when the number of model parameters is high compared to the number of observations and prevent over-fitting (Tibshirani, 1996). For model selection purposes, models were fitted with Lasso, a technique that constrains coefficient estimates (Tibshirani, 1996).

The model parameters were estimated with the glmnet[2] algorithm for elastic net models (Friedman et al., 2010, 2009; Hastie et al., 2005; Zou and Hastie, 2005). Given a linear regression with standardized predictors x_i and centered response values y_i for i=1,2,…,N and j=1,2,…,p, the glmnet algorithm estimates the regression coefficients b={b_j} to minimize:

$$\frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i=p}}^{\boldsymbol{N}} \boldsymbol{w_i}\, \boldsymbol{l}(\boldsymbol{y_i}, \boldsymbol{b_0} + \boldsymbol{b^T x_i}) + \boldsymbol{\lambda} \left[\frac{(\mathbf{1}-\boldsymbol{a})||\boldsymbol{b}||_{\mathbf{2}}^{\mathbf{2}}}{\mathbf{2}} + \boldsymbol{a}||\boldsymbol{b}||_{\mathbf{1}} \right],$$

where λ covers a range of values, l(y,η) is the negative log-likelihood contribution for observation i and a controls the elastic-net penalty (for lasso α=1). The tuning parameter λ was chosen through cross validation.

Friedman et al. (2009) suggest running the lasso to identify the set of non-zero coefficients, and then fitting an unrestricted linear model to the selected set of features. We refitted the model as a unrestricted GLM using predictors with non-zero coefficients and derived monthly estimates of the index using the least squares means (Lenth, 2013). The data reference grid was restricted to the most frequent level of each factor (Campbell, 2015; Lenth, 2013) and interactions had to be disregarded, due to computing power limitations and missing (unobserved) levels in the case of interactions.

To deal with missing/unobserved values and with the hierarchical structure of logbook data we considered mixed models. The skipper and the vessel are crossed effects. The year and grid cell interaction can be treated as a random effect to allow for the calculation of standardized indices for the unobserved (i.e. unfished) year-grid cell interaction levels from the estimated posterior mean of the distribution of the random effect, thus removing the need for imputing these values (Campbell, 2015; Maunder and Punt, 2004). The generalised linear mixed (GLMMs) Lasso models[3] are extensions of the GLM Lasso that allow for the inclusion of random effects (Groll and Tutz, 2014; Schelldorfer, 2011). To optimize the tuning parameter λ and the starting values we applied BIC-based selection and 5-fold cross validation (Groll and Groll, 2014; Schelldorfer and Meier, 2011). Models with and without random effects were compared using AIC. To derive monthly estimates of the standardized CPUE, we selected the non-zero variables of the final Lasso model, refitted the GLMM[4] and estimated the monthly indices with the least square means approach (Campbell, 2015; Lenth, 2013).

3. Results and Discussion

We developed a modeling approach to standardize CPUE from logbook data and acquired an index of abundance for tropical Atlantic juvenile BET caught by the French purse seine fishery. **Table 3** lists coefficients derived from the Lasso GLMs, binomial and lognormal. The GLMM Lasso models were more conservative than the GLMs on the selection of the fixed effects (**Table 3**), especially for the binomial model, where only the introduction of buoys equipped with a short-range radio beacon correlates to the probability of a positive bigeye catch.

[2] http://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html
[3] https://cran.r-project.org/package=glmmLasso
[4] https://cran.r-project.org/web/packages/lme4; (Bates *et al.*, 2013)

Figures **1** and **2** show the GLM and GLMM standardized monthly bigeye CPUE estimates and confidence intervals, derived with the least squares means procedure. The GLM-based time series shows higher annual bigeye CPUEs (closer to the annual data average) with narrower confidence intervals compared to the GLMM-based time series (**Figure 1**). Both standardized BET CPUE time series peak in 2010; the GLM-based time series remains stable at high CPUE levels after the peak of 2011 whereas the GLMM-based time series declines after 2011.

Our results support the use of mixed models to mitigate the effects of a spatially shifting fishery (hyperstability) and account for the correlation present in the data (repeated measurements at a vessel and/or skipper level). Moreover, variable catchability, attributed to technological developments, is modeled by incorporating relevant predictors as proposed in literature (Bishop, 2006; Bishop *et al*., 2008; Mahévas *et al*., 2011). Our results clearly show a correlation between catch rate and predictors that describe technological developments, but these relationships are often weak, possibly due to data deficiencies (e.g. short time series, low contrast, unreliable or missing data) or non-linearities (Achen, 2005). Finally, this study highlighted the role of the skipper and the vessel, also featured in Mahévas *et al*., (2011) for European fishing fleets.

The development of the FAD fishery has led to the need for new approaches to CPUE standardization if we want to derive reliable abundance indices from commercial data. To that end, we suggest the collection of data on technological advances potentially influencing fishing effort and the use of advanced statistical methods to assess the effect of technology on fishing effort.

Acknowledgments

The fisheries data analysed in this publication were collected by the Mixed Research Unit 248 MARBEC of the *Institut de Recherche pour le Développement* (IRD) through the EU-funded Data Collection Framework [DCF, Reg (EC) 1543/2000 and 199/2008]. We wish to acknowledge the contributions of ORTHONGEL, the staff of the 'Observatoire Thonier', and all the technicians and fishers who have been involved in data collection, processing, and management. We are particularly grateful to Laurent Floch (IRD) for assistance in data supply. This work is a contribution to the CECOFAD project (MARE/2012/24). CECOFAD is led by IRD (France) and involves two other Scientific partners: Instituto Español de Oceanografía (IEO, Spain), AZTI Tecnalia (AZTI, Spain), and 3 professional partners: Organisation des producteurs de thon tropical congelé et surgelé (ORTHONGEL (France), Asociación Nacional de Buques Atuneros Congeladores (ANABAC, Spain), Organización de Productores Asociados de Grandes Atuneros Congeladores (OPAGAC, Spain).

References

Achen, C.H., 2005. Let’s Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. Confl. Manag. Peace Sci. 22, 327–339. doi:10.1080/07388940500339167

Ariz, J., Delgado, A., Fonteneau, A., Gonzales Costas, F., Pallarés, P., 1992. Logs and tunas in the eastern tropical Atlantic: A review of present knowledge and uncertainties, in: Proceedings of the International Workshop on Fishing for Tunas Associated with Floating Objects, La Jolla, CA. pp. 21–65.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2013. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-4. Accessed Online Dec.

Bishop, J., 2006. Standardizing fishery-dependent catch and effort data in complex fisheries with technology change. Rev. Fish Biol. Fish. 16, 21–38. doi:10.1007/s11160-006-0004-9

Bishop, J., Venables, W.N., Dichmont, C.M., Sterling, D.J., 2008. Standardizing catch rates: is logbook information by itself enough? ICES J. Mar. Sci. J. Cons. 65, 255–266. doi:10.1093/icesjms/fsm179

Campbell, R.A., 2015. Constructing stock abundance indices from catch and effort data: Some nuts and bolts. Fish. Res. 161, 109 – 130. doi:http://dx.doi.org/10.1016/j.fishres.2014.07.004

Chassot, E., Floch, L., Dewals, P., Irie, D., Cauquil, P., Amande, J.M., Chavance, P., 2015. Statistics of the French purse seine fishing fleet targeting tropical tunas in the Atlantic Ocean (1991-2014). SCRS/2015/196.

Fletcher, D., MacKenzie, D., Villouta, E., 2005. Modelling skewed data with many zeros: A simple approach combining ordinary and logistic regression. Environ. Ecol. Stat. 12, 45–54. doi:10.1007/s10651-005-6817-1

Fonteneau, A., Chassot, E., Bodin, N., 2013. Global spatio-temporal patterns in tropical tuna purse seine fisheries on drifting fish aggregating devices (DFADs): Taking a historical perspective to inform current challenges. Aquat. Living Resour. 26, 37–48. doi:10.1051/alr/2013046

Fonteneau, A., Gaertner, D., Nordstrom, V., 1999. An overview of problems in the catch per unit of effort and abundance relationship for the tropical purse seine fisheries. Col Vol Sci Pap ICCAT 49, 258–278.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33, 1–22.

Friedman, J., Hastie, T., Tibshirani, R., 2009. glmnet: Lasso and elastic-net regularized generalized linear models. R Package Version 1.

Gaertner, D., Ariz, J., Bez, N., Clermidy, S., Moreno, G., Murua, H., Soto, M., 2015. Objectives and first results of the CECOFAD project. SCRS/P/2015/104.

Goujon, M., Claude, A., Le Couls, S., Mangalo, C., 2015. Premier bilan du plan de gestion des DCP mis en place par la France en Océan Atlantique. Collect Vol Sci Pap ICCAT, SCRS/2014/187 71, 573–591.

Groll, A., Groll, M.A., 2014. Package "glmmLasso."

Groll, A., Tutz, G., 2014. Variable selection for generalized linear mixed models by l 1-penalized estimation. Stat. Comput. 24, 137–154.

Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. Math. Intell. 27, 83–85.

Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty/Book and Disk. Springer Science & Business Media.

IOTC, 2014. Indian ocean skipjack tuna stock assessment 1950-2013 (Stock Synthesis). Prepared by IOTC secretariat (R.Sharma and M. Herrera). IOTC 2014 WPTT16 43 Rev 3.

Kuo, L., Mallick, B., 1998. Variable selection for regression models. Sankhyā Indian J. Stat. Ser. B 65–81.

Lenth, R.V., 2013. lsmeans: Least-squares means. R package version 1.06-05.

Lo, N.C., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on Delta-Lognornial Models. Can. J. Fish. Aquat. Sci. 49, 2515–2526.

Lopez, J., Moreno, G., Sancristobal, I., Murua, J., 2014. Evolution and current state of the technology of echo-sounder buoys used by Spanish tropical tuna purse seiners in the Atlantic, Indian and Pacific Oceans. Fish. Res. 155, 127 – 137. doi:http://dx.doi.org/10.1016/j.fishres.2014.02.033

Mahévas, S., Vermard, Y., Hutton, T., Iriondo, A., Jadaud, A., Maravelias, C.D., Punzón, A., Sacchi, J., Tidd, A., Tsitsika, E., Marchal, P., Goascoz, N., Mortreux, S., Roos, D., 2011. An investigation of human vs. technology-induced variation in catchability for a selection of European fishing fleets. ICES J. Mar. Sci. J. Cons. 68, 2252–2263. doi:10.1093/icesjms/fsr150

Maufroy, A., Chassot, E., Joo, R., Kaplan, D.M., 2015. Large-Scale Examination of Spatio-Temporal Patterns of Drifting Fish Aggregating Devices from Tropical Tuna Fisheries of the Indian and Atlantic Oceans. PLoS ONE.

Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70, 141–159.

Postel, E., 1969. Présentation des thons. Pêche maritime., in: Ressources, Pêche et Biologie Des Thonidés Tropicaux de l'Atlantique Centre, Est. FAO Doc. Techn. Pêches. p. 391.

Schelldorfer, J., 2011. An algorithm for high-dimensional generalized linear mixed models using l1-penalization. Schelldorfer, J., Meier, M.L., 2011. Package “glmmlasso.”

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Methodol. 267–288.

Torres-Irineo, E., Gaertner, D., Chassot, E., Dreyfus-León, M., 2014. Changes in fishing power and fishing strategies driven by new technologies: The case of tropical tuna purse seiners in the eastern Atlantic Ocean. Fish. Res. 155, 10–19.

Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2009. Incorporating time-varying catchability into population dynamic stock assessment models. Rev. Fish. Sci. 18, 7–24.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B Stat. Methodol. 67, 301–320.

Table 1. Datasets collected to provide additional information on changes in fishing effort.

Dataset	Description	Time Period	Format	Source
dFAD distribution	Location of dFADs	2007-2014	Monthly point maps	(Maufroy *et al.*, 2015)
Buoy purchases	Total number of buoys purchased by the French fleet, by buoy type.	2004-2014	Time series, annual	(Goujon *et al.*, 2015)
Buoy purchases	Number of buoys bought by vessel (French fleet).	2002-2014	Time series with a variable time interval	*ORTHONGEL*
Vessel characteristics	length, horsepower, capacity, vessel age, company, flag, fleet		Records per vessel	*IRD Observatoire Thonier*
Skippers	Skippers' ID per vessel and trip	2004-2014	Records per vessel and trip	*IRD Observatoire Thonier*

Table 2. Description of the predictors used in the elastic net GLMs and the Lasso GLMMs. The numbers of positive and null fishing sets are used as predictors in the lognormal models.

Variables	Short description
Year	Year at which the fishing set took place as recorded in logbooks
Month	Month at which the fishing set took place as recorded in logbooks
Time at Sea	Duration of the fishing trip, recorded in logbooks
Fishing Time	Cumulated time dedicated to fishing
Sample Area	Areas used for the estimation of species composition of the catch .
EEZ	Exclusive Economic Zone
Grid Cell	Grid cell at 1x1 degree
Skipper	ID of the skipper on each vessels per trip
Vessel	Vessel ID
Vessel age	Year of vessel service
Vessel length	In meters
Vessel power	In horsepower
Vessel Capacity	In m^3
Vessel category	Vessel category related with vessel length and capacity
YFT price	Yellowfin tuna price from the Bangkok auction
SKJ price	Skipjack tuna price from the Bangkok auction
YFT/SKJ price ratio	Price ratio
GPS buoys bought by vessel	Buoys equipped with a GPS
HF buoys bought by vessel	Buoys equipped with a high-frequency radar
HF-GPS/GPS buoys bought by vessel	Buoys equipped with a high frequency radar and a GPS
Average HF buoys per vessel	Buoys equipped with a short-range radio beacon
Average BS buoys per vessel	Buoys equipped with a beacon that communicates its position to the ship via satellite
Average BSE buoys per vessel	Buoys equipped with a GPS and an echosounder
Distance from a dFAD	Distance of the fishing set from the nearest dFAD (monthly scale)
Distance from the centre of the dFAD area	Distance of the fishing set from the centre of the dFAD area (monthly scale)
dFAD area	Total area occupied by dFADs: the sum of the areas of the polygons of the standard distance for each dFAD trajectory. Overlapping polygons were merged.
dFAD counts in buffer zone =143nm	Number of dFADs in a buffer zone around the fishing set. The buffer zone diameter equals 143nm, which is the average nearest neighbor distance between sets and dFADs occurring on the same month, over the whole time series.
dFAD counts in buffer zone = max	Number of dFADs in a buffer zone around the fishing set. The buffer zone diameter is varying on a monthly basis and is equal to the max nearest neighbor distance between the fishing sets and the dFADs for the given month.

Table 3. Variable selection using elastic net GLMs with a = 1 and GLMM Lasso. A two stage approach was followed, modeling the probability of BET presence in a set (Pr(C_s>0)) and the positive BET CPUE (lognormal) in two different models. Coefficients are listed for continuous predictors while factor variables with one or more non- zero coefficients are denoted as (f). The standard deviation is given for random effects (grey cells).

	GLM Lasso		GLMM Lasso	
	Pr(C_s>0)	lognormal	Pr(C_s>0)	lognormal
Year		(f)	(f)	(f)
Month				
Time at Sea				
Fishing Time				
Positive sets		0.08		
Null sets		0.12		
Sample Area	(f)	(f)		(f)
EEZ	(f)	(f)		(f)
Grid Cell	(f)	(f)		
Skipper	(f)	(f)	0.65	0.001
Vessel		(f)	0.49	0.1
Vessel age				
Vessel length				
Vessel Horsepower				
Vessel Capacity				
Vessel category				
YFT price	3e-05			-1e-14
SKJ price	9.4^e-05			-4e-16
YFT/SKJ price ratio				
GPS buoys bought by vessel				
HF buoys bought by vessel				
HF-GPS/GPS buoys bought by vessel				
Average number of HF buoys per vessel	0.06		-0.08	9e-15
Average number of BS buoys per vessel				
Average number of BSE buoys per vessel				-8e-16
Distance from a FAD	-0.015			-5e-02
Distance from the centre of the FAD area		4e-03		
FAD counts in buffer zone = 2.39dd		1.3e-04		
FAD counts in buffer zone = max				
FAD area		-3e-08		
Year*month	(f)	(f)		
Year*Cell	(f)	(f)	1.76	0.5
year*vessel age*category	(f)	(f)		

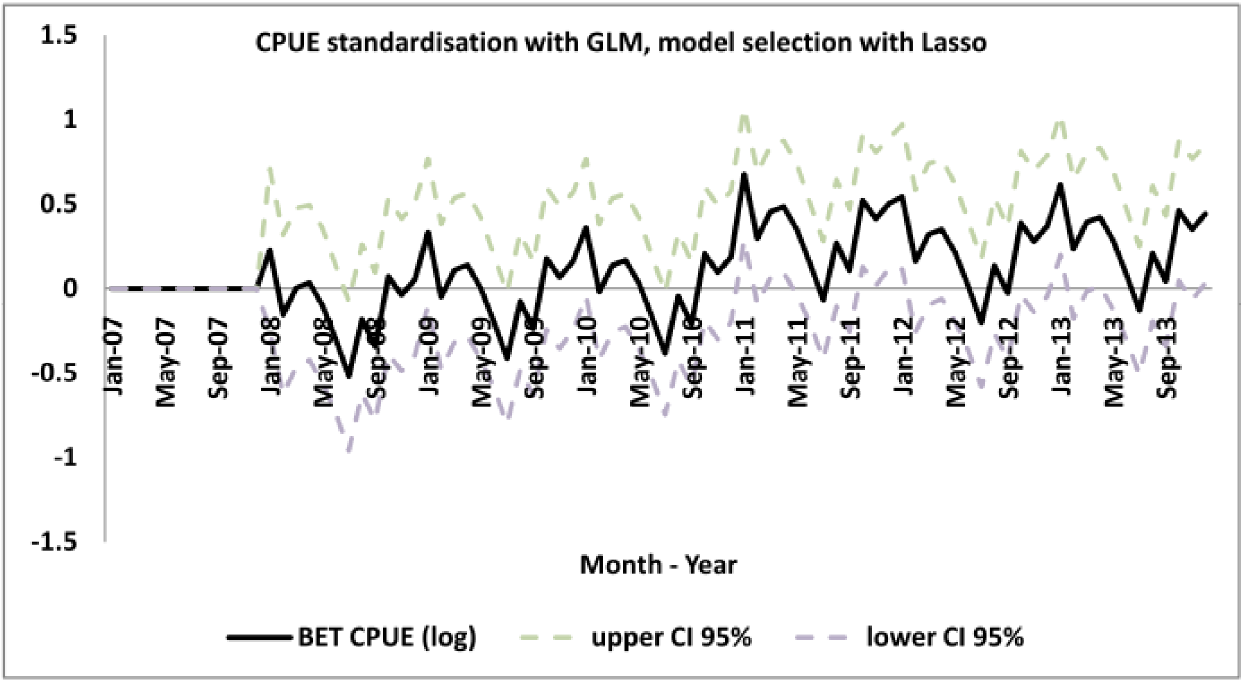

Figure 1. Fitted values (product) of the two (binomial and lognormal) GLMs derived with lsmeans.

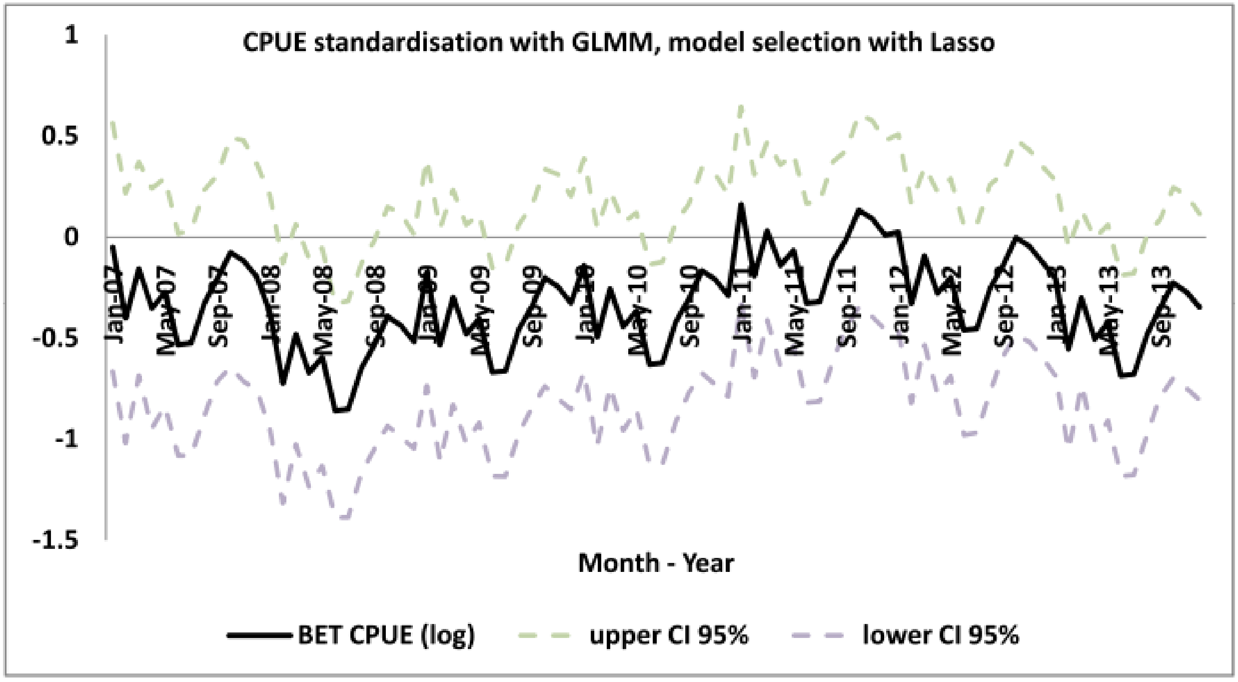

Figure 2. Fitted values (product) of the two (binomial and lognormal) GLMMs derived with lsmeans.