11th MEETING OF THE SCIENTIFIC COMMITTEE

11 to 16 September 2023, Panama City, Panama

SC11 – JM10

Update up to June 2022 of the Chilean Jack Mackerel abundance index based on the catch per fishing set

Republic of Chile

Update up to June 2022 of the Chilean Jack Mackerel abundance index based on the catch per fishing set.

By Ignacio S. Payá C.

August 2023.

Stock Assessment Department.

Chilean Fisheries Development Institute.

Abstract

The CPUE index of the Central-South Chile purse seiner fleet is one of the most important indices of the Chilean Jack Mackerel stock assessment model. The CPUE is calculated as the catch divided by the days out of the port multiply by the vessel holding capacity. An alternative index based on catch by fishing set was estimated by Caballero et al. (2020) for the 1994-2020 period. The aim of this document was to update this index up to June 2022. The data combined the scientific observations collected by observers on board from IFOP and the fishing industry. CPUE index was estimated using a statistic model that includes a distribution of compound probability that describes the joint probability of success and a catch per fishing set. The updated index was significantly different from the nominal data and it was equal to the index estimated by Caballero et al. (2020). During the 2006-2022 period, the updated CPUE index trend was similar to the CPUE index used in the stock assessment model. During 1994-2005, these CPUE indices had different trends. The updated model explained 10.2% total deviance and the sample size was low in some of the early years. Further analyses were identified as GAM and geostatic analysis and time-space models.

Background.

CPUE abundance index of the Central-South Chile purse seiner fleet is one of the most important abundance indices of the Chilean Jack Mackerel stock assessment model. The CPUE index is based on the CPUE data by fishing trip, which is calculated as the catch divided by the product of days out of the port by the vessel holding capacity. Caballero et al. (2020), considering that fishing trips may hide resolution on fishing operations and their strategies, proposed an alternative CPUE index based on catch per set for the 1994-2020 period. This alternative CPUE index was estimated using a statistic model that includes a distribution of compound probability that describes the joint probability of success and a catch per set fishing (Figure 1). The new model predicted similar results to the standardized CPUE by trip, but some differences were also noted. During the 2006-2022 period, the trends were similar between the two CPUE indices. But during the 1994-2006 period, the CPUE index by set increased while the CPUE index by trip was stable.

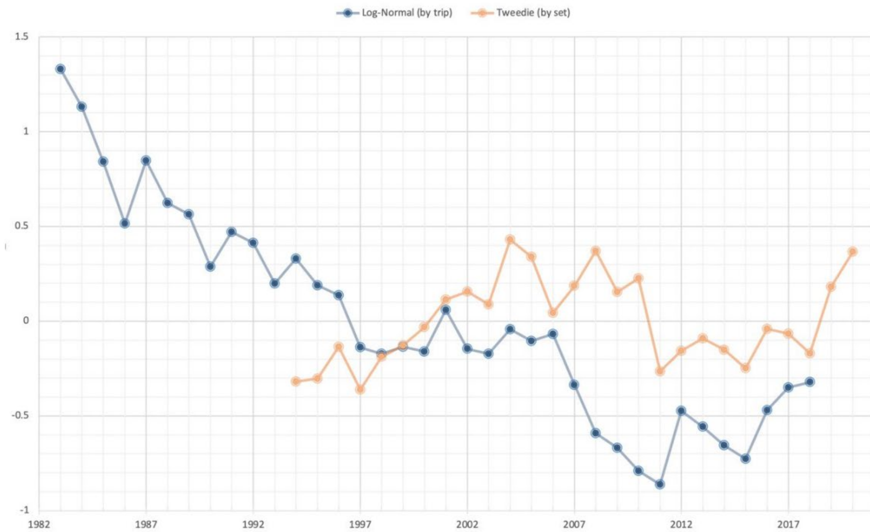

Figure 1. Standardized CPUE by the method Log-Normal (blue) and currently used by the SPRFMO for the jack mackerel stock assessment and the CPUE standardized by a compound model. Please note that the CPUE Log-Normal is based on the catch per fishing trip while the other CPUE is based on catch per fishing set. Both series were scaled to the average value of the CPUE for the 1994-2020 period (taken form Caballero et al. 2020)

Aim

Update the estimate of the abundance index based on the catch per set with data up to June 2022, and compare it with the abundance index in use in the stock assessment model.

Data

The same database used by Caballero et al. (2020) was updated up to June 2022. Daily logbooks of fishing sets associated to the industrial purse seine fleet that operated in the Chilean central-south macro area in the June-1994 - June-2022 period was analysed. Database was composed by the records of the IFOP scientist observers on board fishing vessels that were supplemented with records of the fishing industry provided by INPESCA.

Statistical Model

The same statistical models fitted by Caballero et al. (2020) was applied, that is, a generalized lineal model (GLM, McCullagh & Nelder, 1989) under a Poisson-Gamma distribution composed by Smyth & Jϕrgensen (2002). For more details of the formulation where the rate of catch per set is estimated through the different levels of predictor variables (factors and covariates), revise SC8_JM06 Annex. The model includes year, month, and area as fixed effects and holding capacity of the vessel and the days out of port as covariates. The systematic part of the model had the following form:

$$g(\mu_i) = \beta_0 + \sum_{k=1}^{45} \beta_k x_{ik} \qquad \forall\ i = 1, \dots, m$$

where μ_i represents the mean catch in the i_{th} lance, β_k corresponds to the parameter that measures the linear relation of covariate x_{ik} observed to the i_h set through the linking function $g(\cdot)$, the parameter β_0 represents the intercept term in the intercept in the linear equation and m corresponds to the number of total sets on study, where the covariables are:

x_{i1}: Is the holding capacity (m^3) of the vessel that conducted the i-th set.

x_{jk}: Is the indicate variable related to the area where study the set was conducted, where x_{ik} = 1 if the i-th set was conducted in the area k and x_{ik} = 0 otherwise, for kk = 2, … ,7.

x_{ik}: Is a indicate variable related to the year in which the set was conducted, where x_{ik} = 1 if the i-th set was conducted in the corresponding year and x_{ik} = 0 otherwise, for k = 8, … ,35.

x_{ik}: Corresponds to the indicative variable related to the month in which the set was conducted, where x_{ik} = 1 if the i-th set was conducted in the corresponding month and xik = 0 otherwise for k = 36, … ,46.

x_{i47}: Are the days out of port (nº of days in decimal fraction) of the vessel where the i-th set was conducted.

The area factor was composed by seven fishing grounds (zone 4 to zone 10) described in the table 1 and figure 2.

Table 1. Fishing grounds established for the Central-South macro region.

Fishing grounds	**latitude (°S)**
Caldera	26°00’ – 28°00’
Coquimbo	28°00’ – 32°10’
San Antonio	32°10’ – 34°50’
Talcahuano	34°50’ – 38°30’
Valdivia	38°30’ – 41°00’
Chiloé	41°00’ – 43°30’
Guaitecas	43°30’ – 47°00’

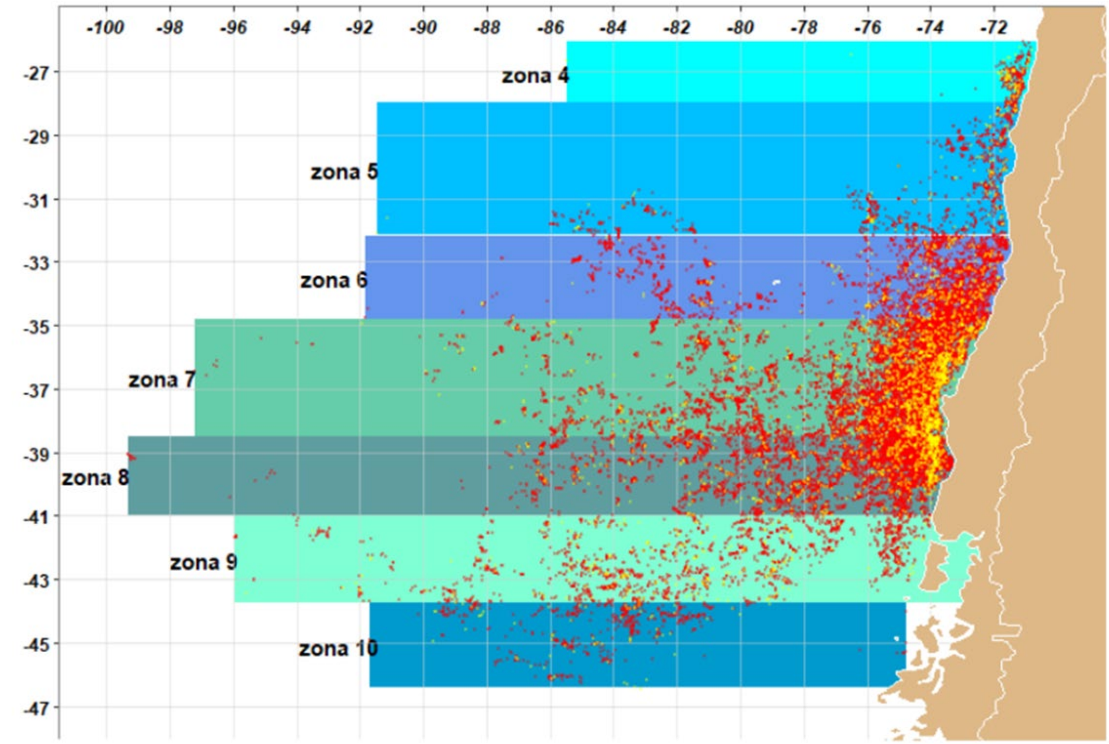

Figure 2. Spatial stratification of the central-south area in fishing subareas and identification of sets with jack mackerel sets with catch (red) and without fishing (yellow) within the period corresponding to June 1994 through July 2020 (taken from Caballero et al. 2020).

Parameters of the model were estimated through the maximum likelihood model, subject to a fixed value of the power parameter of the variance function (methodological details SC8_JM06 Annex), the significance of each of the factors was assessed through the analysis of Devianza. For the processes the software R (R Core Team, 2020) was used with statistical packages tweedie (Dunn, 2017) and statmod (Giner and Smyth, 2016).

Results

The power parameter of the variance function (p) was estimated at 1.4163, similar to the 1.484 estimated by Caballero et al. (2020), which indicates that the underlying distribution to the catch per set rates tends to maintain a compound Poisson-Gamma distribution. The fitted model allows the explication the 10.2% of the total Deviance (Table 2), similar to the 10.1% estimated by Caballero et al. (2020). Estimation of the model parameters, its standard error, t-value, and individual significance are shown in Table 3.

Table 2. Deviance analysis table of the fitted compound Poisson-Gamma model

	Df	Dev. Resid.	Df Resid.	Dev	% of dev.	cumulative % of Dev.	Pr(>Chi)	
NULL			31249	599285				
hc	1	501	31248	598784	0.1%	0.1%	6.31E-08	***
Zone	6	2082	31242	596702	0.3%	0.4%	< 2.2e-16	***
Year	28	32527	31214	564176	5.5%	5.9%	< 2.2e-16	***
Month	11	12817	31203	551358	2.3%	8.0%	< 2.2e-16	***
dop	1	13054	31202	538304	2.4%	10.2%	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3. Deviance analysis table of the fitted compound Poisson-Gamma model

Coefficients	Estimate	Sd. Error	t value	Pr(>|t|)	
(Intercept)	4.63E+00	1.22E-01	37.933	< 2e-16	***
hc	1.80E-04	1.92E-05	9.364	< 2e-16	***
Zone5	2.42E-01	6.40E-02	3.772	0.000162	***
Zone6	9.98E-02	5.92E-02	1.687	0.091618	.
Zone7	-1.97E-02	5.87E-02	-0.337	0.736454	
Zone8	3.85E-02	5.87E-02	0.657	0.511224	
Zone9	1.90E-01	6.03E-02	3.158	0.001592	**
Zone10	2.79E-01	6.93E-02	4.02	5.84E-05	***
year1995	6.69E-03	1.21E-01	0.055	0.955843	
year1996	2.31E-01	1.15E-01	2.005	0.045001	*
year1997	-7.96E-02	1.15E-01	-0.694	0.487891	
year1998	1.61E-01	1.10E-01	1.469	0.141823	
year1999	2.29E-01	1.10E-01	2.092	0.036478	*
year2000	3.36E-01	1.09E-01	3.083	0.002051	**
year2001	4.72E-01	1.06E-01	4.438	9.13E-06	***
year2002	5.11E-01	1.06E-01	4.803	1.57E-06	***
year2003	4.54E-01	1.07E-01	4.241	2.23E-05	***
year2004	7.21E-01	1.06E-01	6.808	1.01E-11	***
year2005	6.57E-01	1.06E-01	6.189	6.11E-10	***

year2006	4.35E-01	1.14E-01	3.817	0.000135	***
year2007	5.30E-01	1.10E-01	4.823	1.42E-06	***
year2008	6.85E-01	1.08E-01	6.335	2.40E-10	***
year2009	5.10E-01	1.09E-01	4.698	2.64E-06	***
year2010	5.70E-01	1.12E-01	5.086	3.68E-07	***
year2011	5.39E-02	1.14E-01	0.473	0.636459	
year2012	1.83E-01	1.09E-01	1.684	0.092224	.
year2013	2.55E-01	1.11E-01	2.309	0.020931	*
year2014	1.89E-01	1.16E-01	1.634	0.102268	
year2015	8.49E-02	1.11E-01	0.762	0.446074	
year2016	3.20E-01	1.13E-01	2.834	0.004594	**
year2017	2.87E-01	1.14E-01	2.512	0.01202	*
year2018	1.73E-01	1.14E-01	1.52	0.128433	
year2019	5.28E-01	1.09E-01	4.84	1.30E-06	***
year2020	6.13E-01	1.09E-01	5.648	1.64E-08	***
year2021	6.80E-01	1.10E-01	6.175	6.71E-10	***
year2022	6.21E-01	1.10E-01	5.628	1.84E-08	***
Month2	-1.04E-01	2.52E-02	-4.12	3.79E-05	***
Month3	-2.07E-02	2.50E-02	-0.828	0.407668	
Month4	5.65E-02	2.41E-02	2.345	0.019054	*
Month5	1.21E-01	2.44E-02	4.976	6.54E-07	***
Month6	2.15E-01	2.54E-02	8.463	< 2e-16	***
Month7	2.42E-01	2.57E-02	9.406	< 2e-16	***
Month8	1.32E-01	2.65E-02	4.962	7.01E-07	***
Month9	-3.40E-01	3.70E-02	-9.18	< 2e-16	***
Month10	-3.82E-01	4.00E-02	-9.543	< 2e-16	***
Month11	-2.91E-01	4.20E-02	-6.918	4.65E-12	***
Month12	-1.70E-01	2.84E-02	-6	2.00E-09	***
dop	-8.91E-02	3.28E-03	-27.153	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The updated abundance index was significantly different to the nominal data (Figure 3) and was equal to the index estimated by Caballero et al., 2020 (Figure 4).

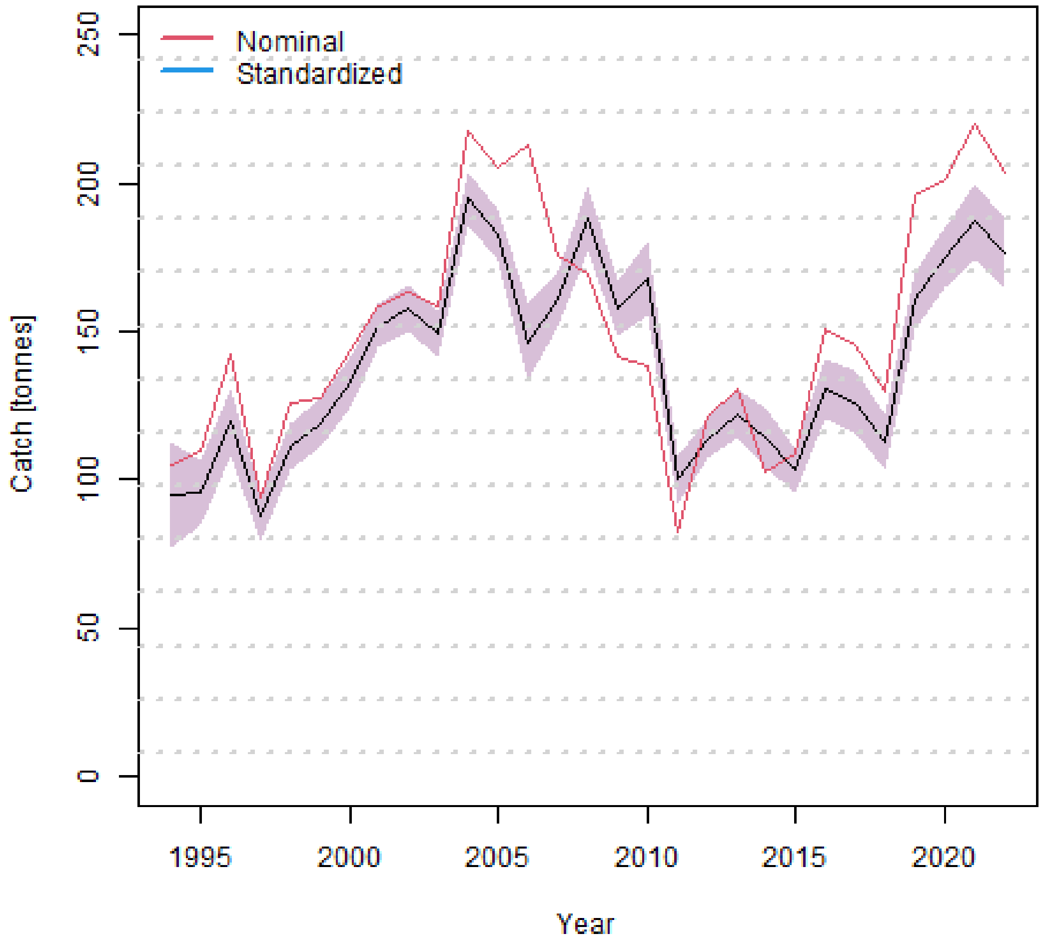

Figure 3. Chilean jack mackerel nominal catch (red line) and standardized abundance index estimated for the series 1994 to 2022, with 95% confidence interval obtained from the compound Poisson-Gamma model (pink area).

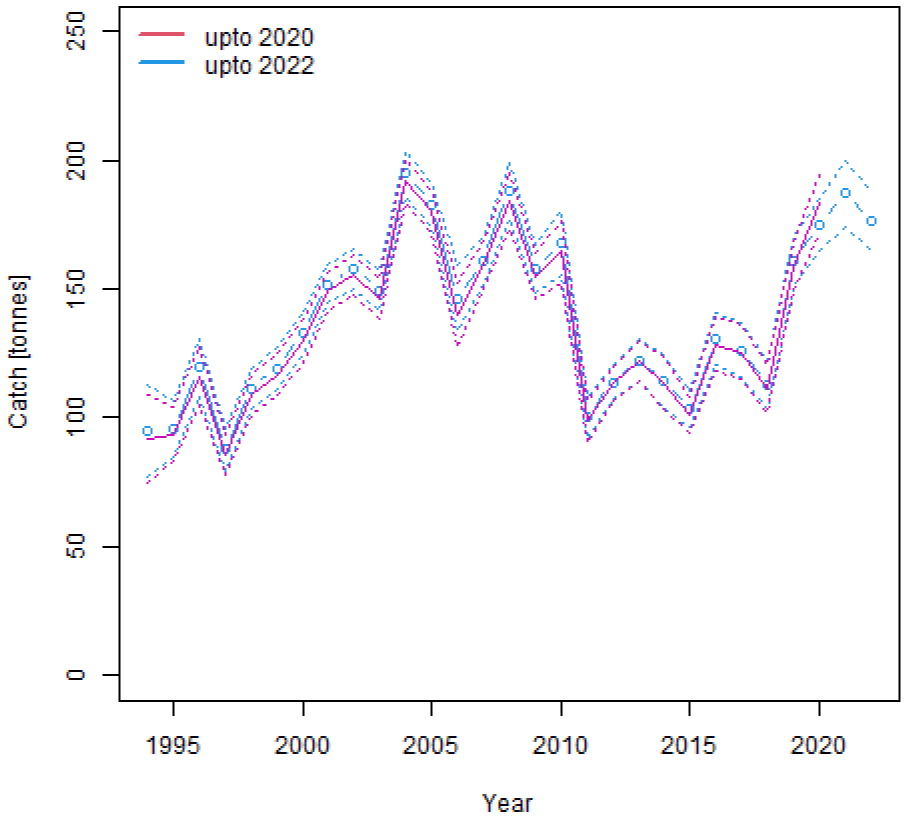

Figure 4. Abundance index based on catch per set (magenta colour) estimated by Caballero et al. (2020) y update in this working document (blue colour). The estimates and the 95% confident intervals are shown.

Discussion

In relation to the CPUE abundance index used the stock assessment model, which used the data by fishing trips (Payá 2022), the alternative abundance index based on catch by fishing set was similar for the 2006-2022 period, but not for early years of the series (Figure 5). For the 2006-2022 period both indices showed a "V" type trend with the minimal figure at year 2011, but the rate of decrease before this year and the rate of increase after this year was greater in the abundance index based on fishing trips.

Figure 5. CPUE abundance index estimated based on catch per days out of port (blue) and catch by set (pink). For comparison the indices were divided by their value in 2017

The updated estimations were equal to the ones of Cabellero et al. (2020) because the data was updated from the same database and the statistical model was fitted by updating the same R script. As a consequence, the updated results have the same limitations as low (10.2%) deviance explained by the model and small sample size for some years. The sampled catch in the database by set was a small fraction of the whole catch of the central-south Chile purse seiners, especially in the first years of the time series (Figure 6). Therefore, the estimations before year 2000 were more uncertain.

As mentioned by Caballero et al. (2020), the alternative abundance index still requires improvements. For example, the model indicates a poor log-linear relationship between the covariate holding capacity with the CPUE. This could be explained by the use of generalized additive models (GAM) as an alternative.

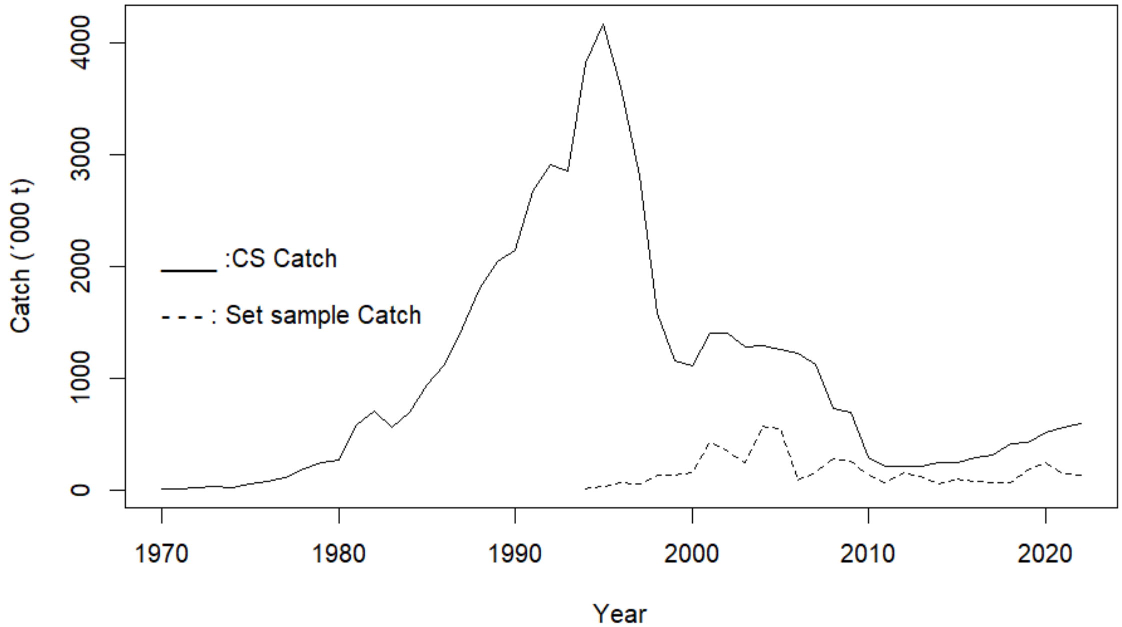

Figure 6. Whole catch of the central-south purse seiners fleet by year and the catch sampled in the catch by set database.

Another important point is that spatial correlation between sets is not formally included in the statistical model used, and so, other analytical approximations with geostatic analysis and time-space models, as spatial and spatiotemporal SPDE-Based GLMMs with template model builder (Anderson et al. 2022), should be analysed as well.

References

Anderson, S.C., E.J. Ward, P.A. English, L.A.K. Barnett. 2022. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv 2022.03.24.485545; doi: https://doi.org/10.1101/2022.03.24.485545.

Caballero L., J.C. Saavedra-Nievas, A. Barraza, C. Montenegro, A. Sepúlveda, C. González, S. Vargas, A. Aranis, G. Bohm y, Juan-Carlos Quiroz. (2020). CPUE Analysis of the jack mackerel purse-seine fishery of the center-south zone of Chile, June 1994 through July 2020. 15

p. SC8-JM06. SPFRMO 2020. https://www.sprfmo.int/assets/Meetings/SC/8th-SC-2020/SC8-JM06-CPUE-Analysis-CJM-purse-seine-fishery-center-south-Chile-1994-2020.pdf

Dunn, P. K. (2017). Tweedie: Evaluation of Tweedie exponential family models. R package version 2.3.

Giner, G, and Smyth, GK (2016). statmod: probability calculations for the inverse Gaussian distribution. R Journal 8(1), 339-351. (pinvgauss, qinvgauss, dinvguass and rinvgauss functions).

McCullagh, P., Nelder, John A. (1989); Generalized Linear Models, Series: Monographs on Statistics & Applied Probability, Chapman & Hall/CRC.

Payá I. (2022). High fish concentrations close to the coast in recent years could affect the Jack Mackerel CPUE abundance index in south-central Chile: proposal of a correction procedure. SC10_JM05. https://www.sprfmo.int/assets/Meetings/SC/10th-SC-2022/SC10-JM05-CPUE-abundance-index-in-south-central-Chile_update-and-proposed-correction-CL.pdf

Smyth, Gordon K., Jφrgensen, B. (2002). Fitting Tweedie's compound poisson model to insurance claims data: dispersion modelling. Astin Bulletin, Vol. 32, No. 1, pp. 143- 157.

Acknowledgments

I thank Aquiles Sepulveda, INPESCA director, for providing the database of catch by set recorded by the INPESCA scientific observes. I am also thank, Marlene Ramizez, IFOP data manager, for combining IFOP database with INPESCA database, and Juan Carlos Quiroz for providing an initial R script for fitting the Tweedie model.