SPRFMO
South Pacific Regional Fisheries Management Organisation

11th MEETING OF THE SCIENTIFIC COMMITTEE

11 to 16 September 2023, Panama City, Panama

SC11 – JM12

CJM abundance index estimated by spatiotemporal SPDE-based GLM and compare with other CPUE indices

Republic of Chile

Chilean jack mackerel abundance index estimated by spatiotemporal SPDE-based GLM and compare with other CPUE indices.

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 off port multiply by the vessel holding capacity. An alternative index based on catch by fishing set was estimated by GLM using a statistic model that included a distribution of compound probability that describes the joint probability of success and a catch per fishing set (Caballero et al., 2020 and Payá, 2023a). However, this model explained 10.2% of total deviance and had small sample sizes in some years, therefore, further analyses were identified. This document reports CPUE index estimates done using spatiotemporal SPDE-Based GLMMs with template model builder (sdmTMB). Two models were fitted to the data, a spatiotemporal model and a spatio-temporal GLM. The estimated abundance indices were not significantly different between the two models. In relation to the CPUE index estimated by GLM using catch per fishing set (Caballero et al., 2020 and Payá, 2023a), the indices had similar trends. In relation to the CPUE abundance index used in the stock assessment model, the indices had a similar trend for the 2006-2022 period, but not for the 1994-2005 period. For the 2006-2022 period all 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 days off port. These results are part of a work in progress and further analyses were identified.

Background.

The 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 off port and 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, which was updated up to June 2022 by Payá (2023). 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. 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 1). The alternative abundance index still requires improvements because the model explained 10.2% of total deviance, and the sample size was low in some years. Further analyses were identified as GAM, and geostatic analysis and time-space models.

Figure 1. CPUE abundance index estimated based on catch per days off port (dop; blue) and catch by set (set; pink). For comparison the indices were divided by their 2017 value (taken from Payá 2023).

The sdmTMB R package (Anderson et al. 2022) implements geostatistical spatial and spatiotemporal GLMMs using TMB for model fitting and R-INLA to set up SPDE (stochastic partial differential equation). Therefore, this R package was used in this working document to estimate alternative jack mackerel CPUE abundance indices.

Aim

To estimate Chilean jack mackerel abundance index using the spatiotemporal SPDE-based GLM with template model builder (sdmTMB).

Data

The database compiled by Caballero et al. (2020) and updated up to 2022 was used. It was composed by the daily logbooks of fishing sets associated to the industrial purse seine fleet that operated in the Chilean central-south macro zone. The records were from the IFOP scientific observers on board fishing vessels for June/1994–December/2022 period that were supplemented with INPESCA records of the fishing industry logbooks for June/1994–April/2022 period.

The fishing set location in latitude and longitude units were transformed to UTM using EPSG: 32718 – WGS 84 / UTM zone 18S (https://maps.omniscale.com/en/openstreetmap/epsg-32718) (Figure 2)

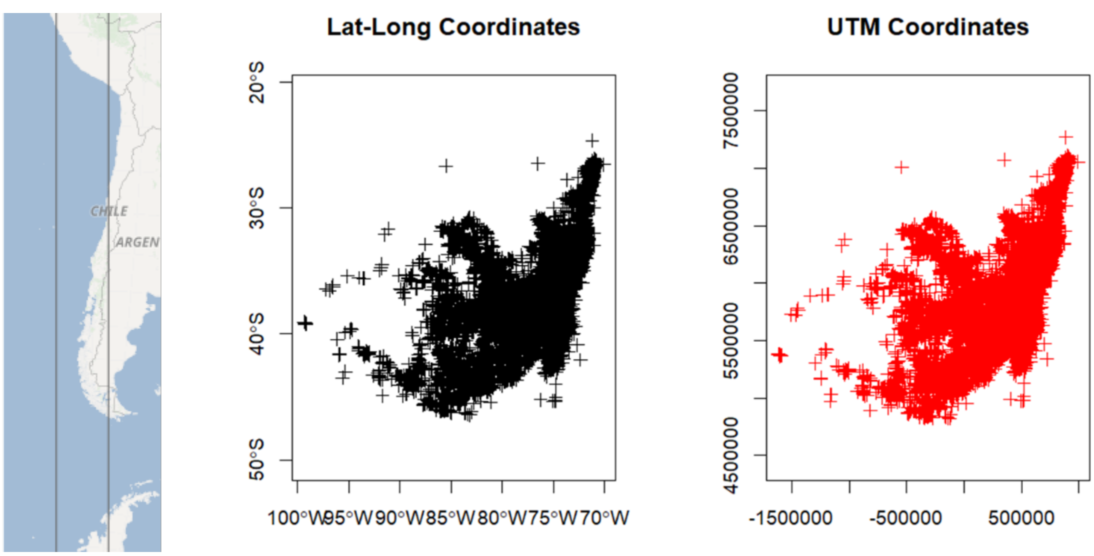

Figure 2. The left map shows the box area of the coordinate reference system according to the EPSG: 32718 – WGS 84 / UTM zone 18S. The fishing set locations in latitude-longitude units (centre plot) and UTM coordinates (right plot) for the 1994-2023 period.

The spatial grid.

The grid was defined with a cutoff = 100 (Figure 3):

mesh <- make_mesh(cjm, xy_cols = c("X", "Y"), cutoff = 100.

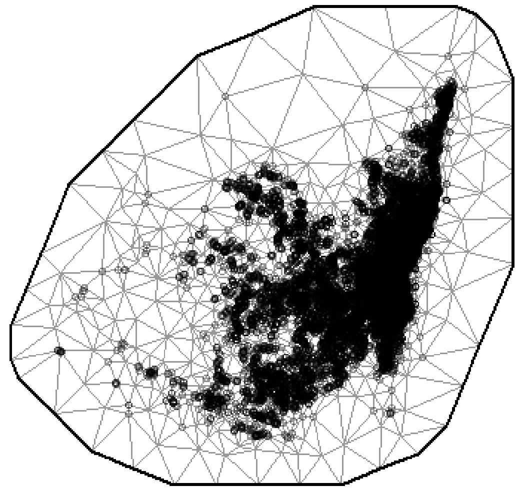

Figure 3. Spatial grid.

Statistical Models.

Two models were fitted to data:

Spatiotemporal model

The CPUE (density) was modelled as a function of the smoothness of the vessel holding capacity (hc) with spatiotemporal variability set to an auto-regressive function of order 1. A Tweedie error distribution with a log link function (Dunn, 2017) was used:

fit_spatiotemporal <- sdmTMB(density ~ s(hc, k = 5), data = cjm, mesh = mesh, **time = "year",** family = tweedie(link = "log"), spatial = "off", **spatiotemporal = "ar1",** control = sdmTMBcontrol(newton_loops = 1))

Spatio-temporal GLM.

The CPUE (density) was modelled as function of the year and month as fixed effects and the smooth of the vessel holding capacity with spatiotemporal define by the SPDE (stochastic partial differential equation). A Tweedie error distribution with a log link function was used:

m <- sdmTMB(data = cjm, formula = density ~ 0 + as.factor(year) + as.factor(month) + s(hc, k = 5), time = "year", mesh = cjm_spde, family = tweedie(link = "log"), control = sdmTMBcontrol(newton_loops = 1))

The abundance indices were also estimated by south and west areas. The south area was defined by northing < 6000, while the west area by easting <500 (Figure 4).

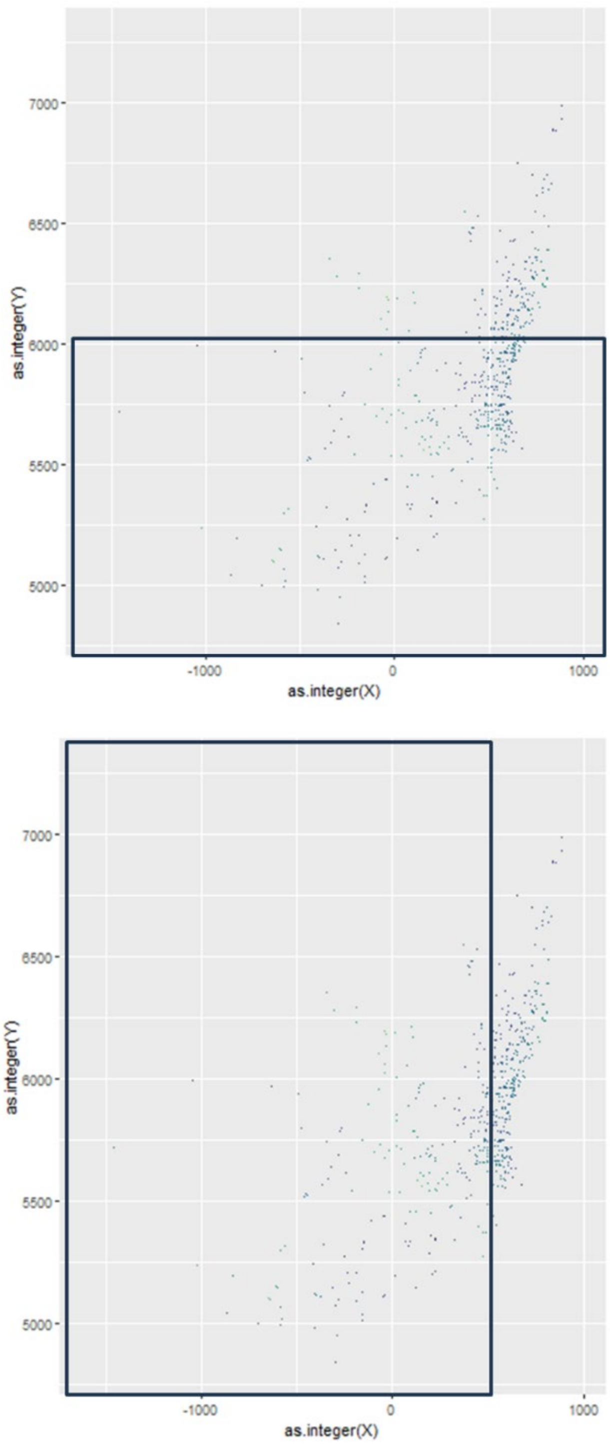

Figure 4. Boxes of the south (<6000 northern, upper plot) and west area (<500 easting, lower plot). The points represent the locations of fishing sets, X the easting and Y the northing.

Results

Spatiotemporal model

Summary of the spatiotemporal model results ae shown in table 1.

Table 1. Summary of the spatiotemporal model results.

```
Spatiotemporal model fit by ML ['sdmTMB']
Formula: density ~ s(hc, k = 5)
Mesh: mesh (isotropic covariance)
Time column: year
Data: cjm
Family: tweedie(link = 'log')

             coef.est coef.se
(Intercept)    -2.02    0.05
shc             0.49    0.18

Smooth terms:
Std. Dev.
sds(hc)      0.13

Dispersion parameter: 0.25
Tweedie p: 1.41
Spatiotemporal AR1 correlation (rho): 0.34
Matern range: 337.24
Spatiotemporal SD: 0.50
ML criterion at convergence: -13110.263
```

The gravity center had a trend to move to coastal areas in the most recent years (Figure 5). The abundance index had an increase trend in 1994-2006, a decrease trend in 2007- 2011, an increase trend in 2002-2019, and a decrease in 2020-2022 (Figure 6).

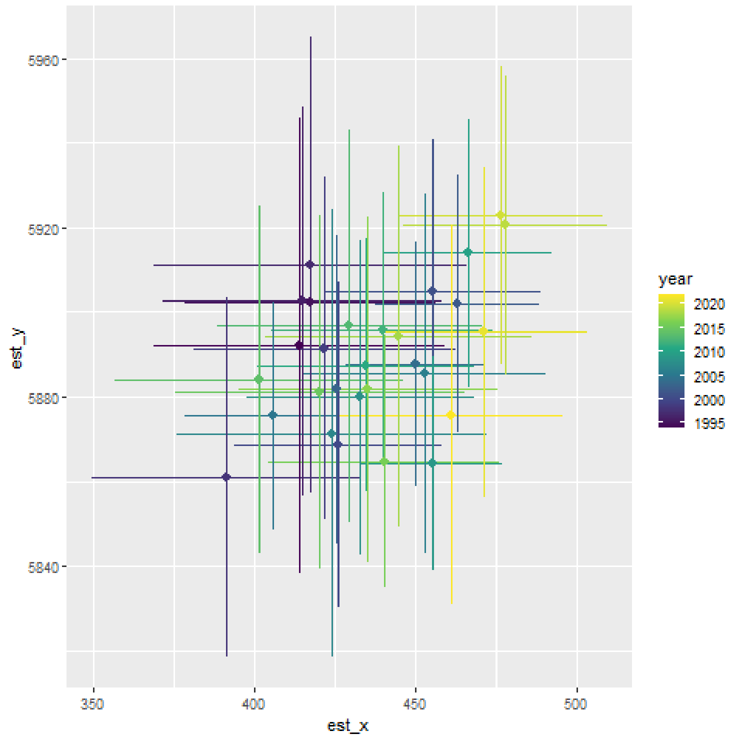

Figure 5. Gravity centres by year estimated using spatiotemporal model, the lines represent the 95% intervals.

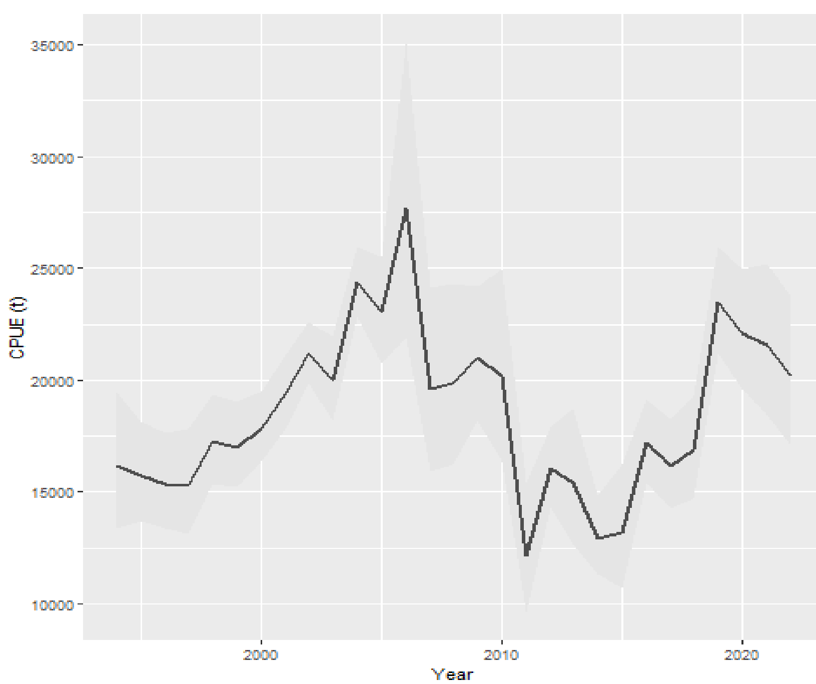

Figure 6. Abundance index estimated by spatiotemporal model. The grey band represents the 95% intervals.

Spatio-temporal GLM.

The summary of the model results is presented in Table 2, and the abundance index in Table 3.

Table 2. Parameters estimated by spatiotemporal GLM model.

	coef.est	coef.se
as.factor(year)1994	-2.10	0.22
as.factor(year)1995	-2.01	0.16
as.factor(year)1996	-1.89	0.15
as.factor(year)1997	-2.02	0.17
as.factor(year)1998	-1.93	0.15
as.factor(year)1999	-1.97	0.16
as.factor(year)2000	-1.77	0.13
as.factor(year)2001	-1.62	0.15
as.factor(year)2002	-1.76	0.12
as.factor(year)2003	-1.69	0.11
as.factor(year)2004	-1.56	0.10
as.factor(year)2005	-1.45	0.10
as.factor(year)2006	-1.38	0.17
as.factor(year)2007	-1.78	0.14
as.factor(year)2008	-1.88	0.08
as.factor(year)2009	-1.93	0.08
as.factor(year)2010	-2.02	0.09
as.factor(year)2011	-2.62	0.10
as.factor(year)2012	-2.10	0.13
as.factor(year)2013	-2.09	0.15
as.factor(year)2014	-2.33	0.15
as.factor(year)2015	-2.31	0.11
as.factor(year)2016	-1.99	0.12
as.factor(year)2017	-2.07	0.15
as.factor(year)2018	-2.12	0.15
as.factor(year)2019	-1.51	0.15
as.factor(year)2020	-1.66	0.15
as.factor(year)2021	-1.66	0.16
as.factor(year)2022	-1.52	0.20
as.factor(month)2	-0.11	0.03
as.factor(month)3	-0.02	0.03
as.factor(month)4	0.05	0.03
as.factor(month)5	0.07	0.03
as.factor(month)6	0.09	0.03
as.factor(month)7	-0.09	0.03
as.factor(month)8	-0.20	0.03
as.factor(month)9	-0.43	0.04
as.factor(month)10	-0.54	0.04
as.factor(month)11	-0.52	0.04
as.factor(month)12	-0.24	0.03
shc	0.56	0.19

```
Smooth terms:
          Std. Dev.
sds(hc)       0.16

Dispersion parameter: 0.24
```

```
Tweedie p: 1.41
Matern range: 1.90
Spatial SD: 0.15
Spatiotemporal SD: 0.41
ML criterion at convergence: -13409.269
```

Table 3. Abundance index estimated by spatiotemporal GLM model.

year	est	lwr	upr	cv
1994	14691	9507	22703	0.22
1995	16061	11851	21769	0.16
1996	18052	13576	24003	0.15
1997	15950	11399	22317	0.17
1998	17346	13057	23043	0.15
1999	16714	12247	22810	0.16
2000	20419	15991	26075	0.13
2001	23604	17815	31273	0.14
2002	20664	16359	26101	0.12
2003	22184	18023	27305	0.11
2004	25278	20802	30718	0.10
2005	28124	23233	34044	0.10
2006	30017	21632	41651	0.17
2007	20162	15461	26292	0.14
2008	18371	15761	21414	0.08
2009	17369	15020	20086	0.07
2010	15909	13360	18943	0.09
2011	8736	7143	10685	0.10
2012	14750	11533	18865	0.13
2013	14839	11199	19663	0.14
2014	11672	8801	15481	0.14
2015	11938	9667	14743	0.11
2016	16390	12920	20791	0.12
2017	15071	11208	20266	0.15
2018	14354	10670	19310	0.15
2019	26411	19715	35383	0.15
2020	22707	16855	30589	0.15
2021	22845	16814	31038	0.16
2022	26165	17858	38338	0.20

The residuals had a good spatiotemporal distribution without any trend in space or time (Figure 7).

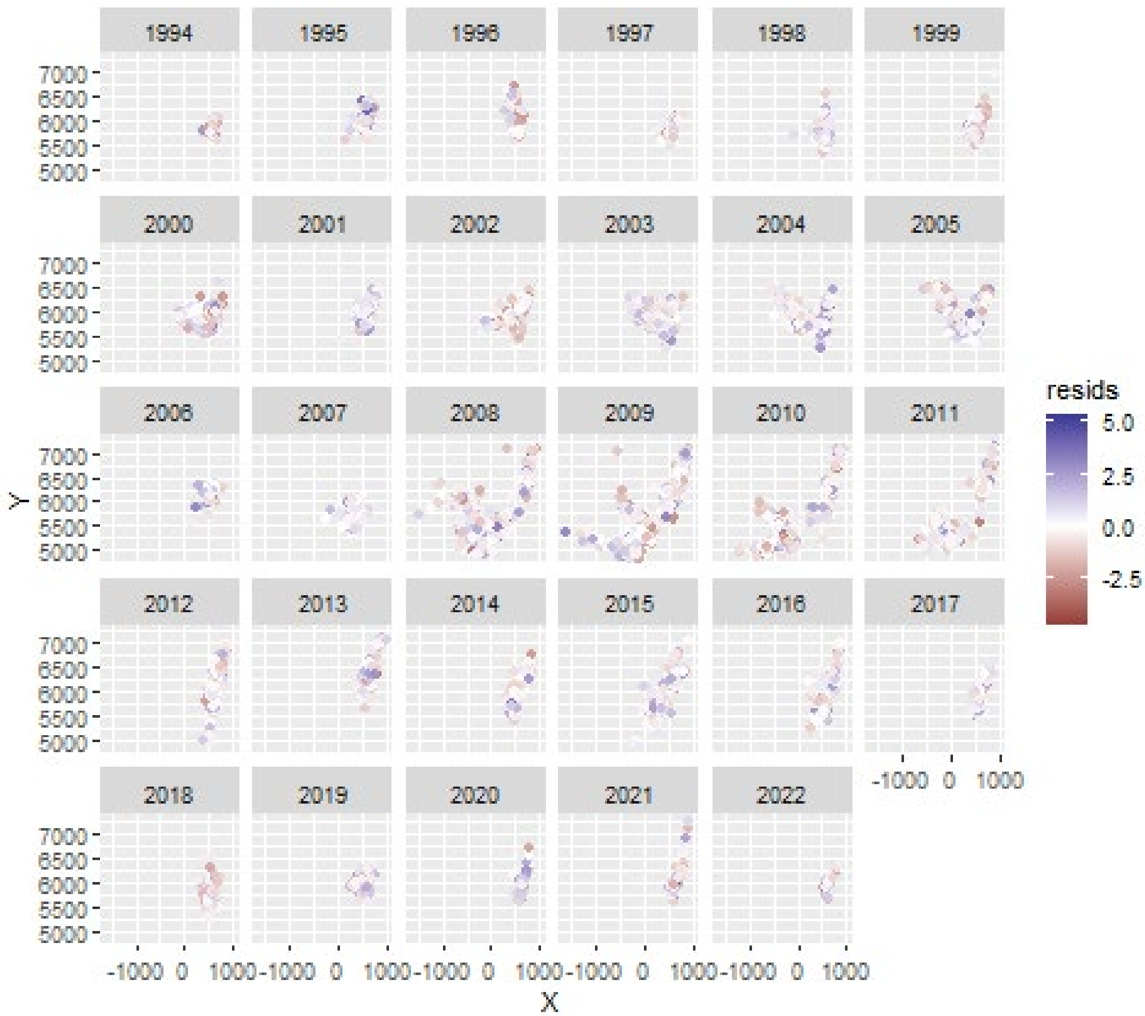

Figure 7. Residuals of the spatiotemporal GLM model by year (X = Easting, Y = Northing).

The abundance index estimated by spatiotemporal GLM had an increase trend in 1994-2006, a decrease trend in 2007- 2011, and an increase trend in 2002-2022 (Figure 8).

Figure 8. Abundance index estimated by spatiotemporal GLM model. The grey band represents the 95% intervals.

The abundance index in the south area had the same trend that the abundance index for the all area, but at a lower level that in most of the years overlapped the confident intervals of the abundance index for the whole area (Figure 9).

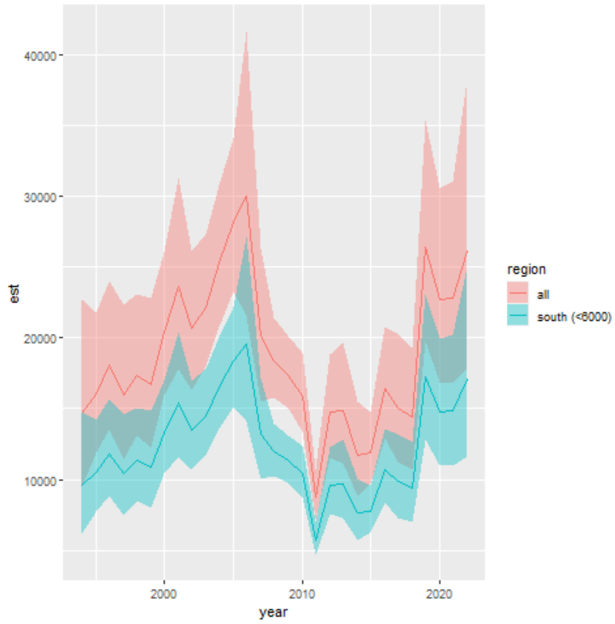

Figure 9. Abundance index for the all and south area

The abundance index in the western area had a similar trend that the abundance index for the whole area, but at a much lower level (Figure 10).

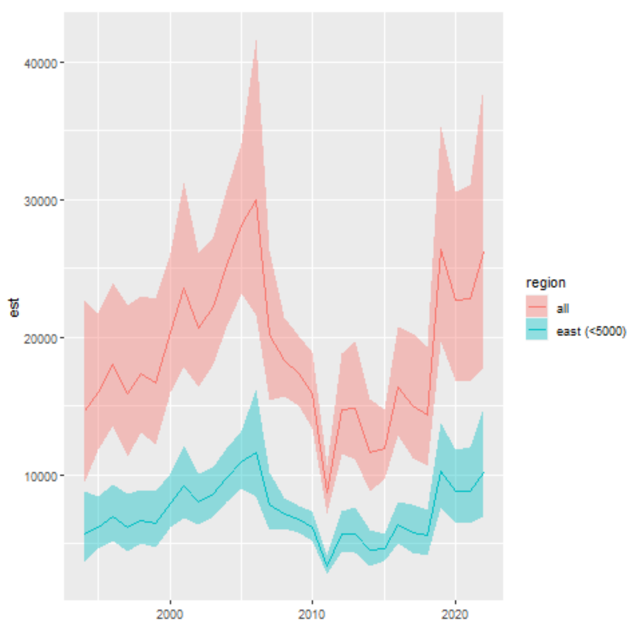

Figure 10. Abundance index for the all and the west area (easting <500).

Discussion

The abundance indices estimated by the spatiotemporal model and by the spatiotemporal GLM were not significantly different (Figure 11). However, the index estimated by the spatiotemporal GLM had an increasing trend in the last 2 years, while the other index had a decrease trend. The uncertainty was greater in index estimated by spatiotemporal GLM.

In relation with the CPUE index based on catch per fishing set estimated by Caballero et al. (2020) and updated by Payá (2023), the indices based on fishing set and estimated using sdmTMB had similar trends (Figure 12). In relation to the CPUE abundance index used in the stock assessment model, which used the data by fishing trips and updated up to June 2023 (Payá 2023b), the alternative abundance indices based on catch by fishing set was similar for the 2006-2022 period, but not for early years of the series. For the 2006-2022 period the four 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.

The results presented here are part of a work in progress. Further analyses should be done including alternative grids with different cutoff, analyse anisotropy in the Matern Function, alternative space-time GLM or GLMM models, alternative prediction grids, and indices for others areas of interest (north, west, etc.).

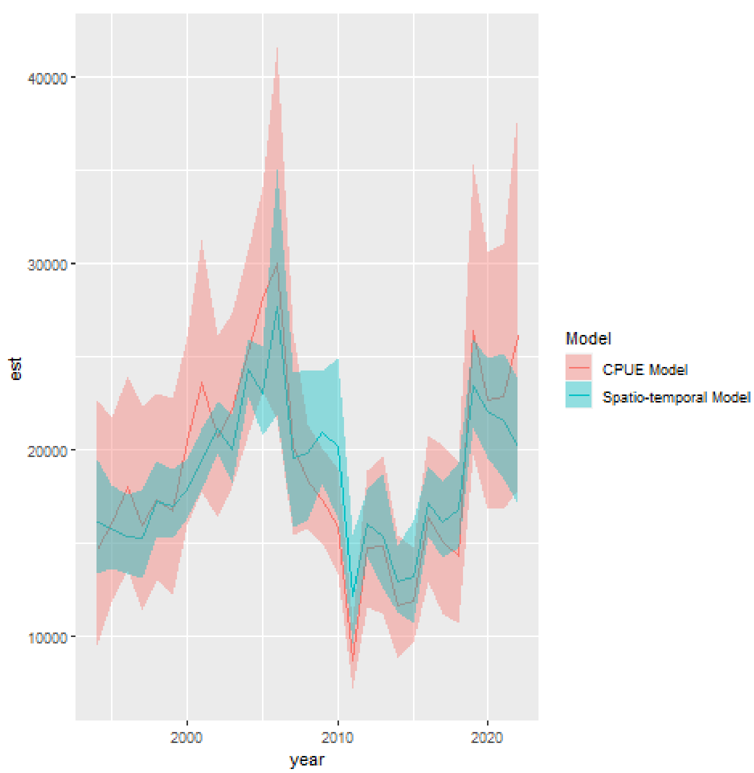

Figure 11. Abundance index estimated by the CPUE spatiotemporal GLM model (CPUE model) and by spatiotemporal model

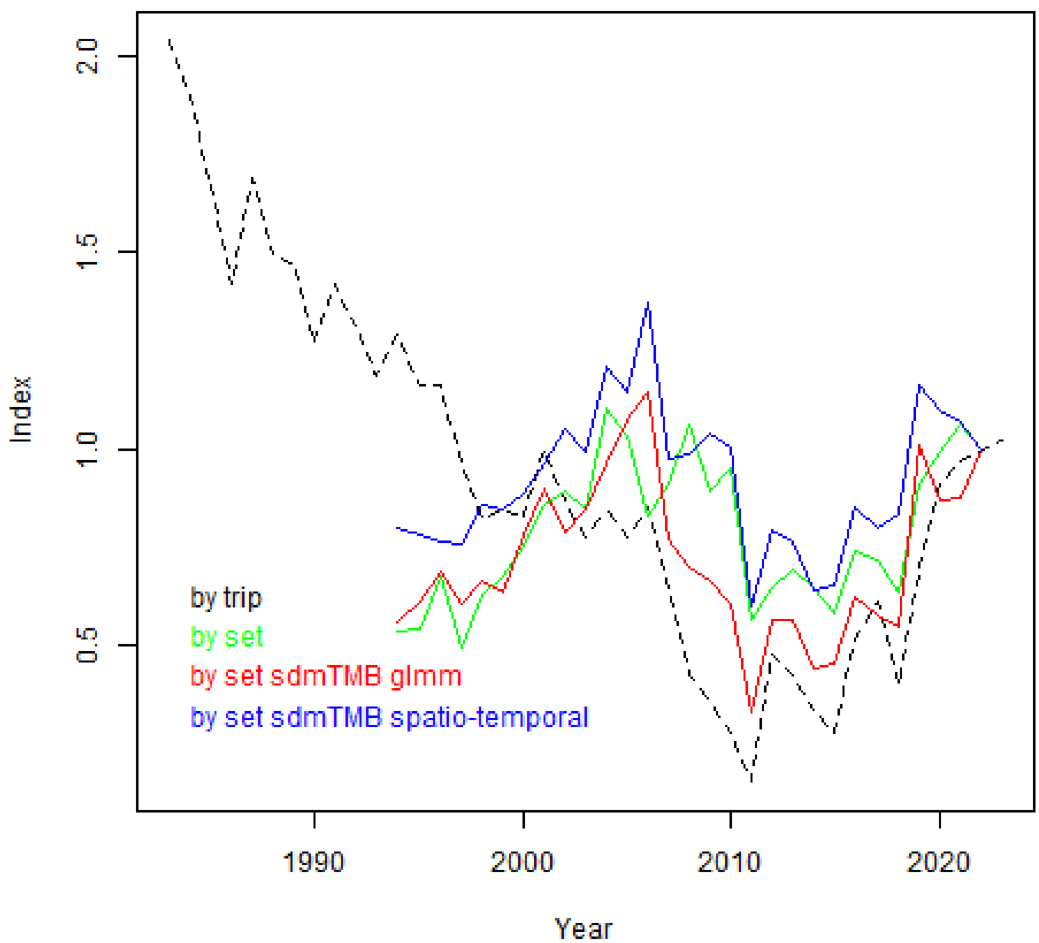

Figure 12. Abundance indices estimated by different models. CPUE index used in the stock assessment model updated up to June 2023 (trip), CPUE index based on GLM of catch per fishing set (set), CPUE estimated by the spatiotemporal model (by set sdmTMB spatio-temporal) and CPUE index estimated by the spatiotemporal GLM (by set sdmTMB glm). Indices were divided by their values in 2022 year.

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.

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

Payá I. (2023a). Update up to June 2022 of the Chilean Jack Mackerel abundance index based on the catch per fishing set. Working document for SC11_JM. Instituto de Fomento Pesquero.

Payá I. (2023b). Update of the Chilean Jack Mackerel CPUE abundance index based on catch by fishing trip in the south-central Chile. SC11_JM. Instituto de Fomento Pesquero.

Acknowledgments

I thank Aquiles Sepulveda, INPESCA director, for providing the database of catch by fishing set recorded by the INPESCA. I am also thankful to Marlene Ramirez and Erik Gaete, IFOP scientists, for providing the IFOP database and combining it with INPESCA database.