

# Toward a standardised abundance index for Chilean jack mackerel: spatio-temporal modelling of acoustic survey data in northern Chile

Maria Grazia Pennino<sup>a</sup>, Sebastián I. Vásquez<sup>b,c</sup>, Nicolás Alegría<sup>b</sup>, Aquiles Sepúlveda<sup>b</sup>

<sup>a</sup>*Institut de Ciències del Mar, Barcelona, Spain*

<sup>b</sup>*Instituto de Investigación Pesquera, Talcahuano, Chile*

<sup>c</sup>*Programa de Doctorado en Oceanografía, Universidad de Concepción, Chile*

---

## Abstract

We developed a spatio-temporal modelling framework to estimate an alternative abundance index for Chilean jack mackerel (*Trachurus murphyi*) in northern Chile, integrating environmental covariates while accounting for spatial structure and uneven sampling coverage. Models were fitted using the R package *sdmTMB*, which implements a stochastic partial differential equation (SPDE) approach. Given the high proportion of zero observations, a delta modelling framework was adopted, combining binomial (occurrence) and Gamma (positive density) components. Alternative spatio-temporal structures were evaluated, including independent, autoregressive (AR1), and random walk formulations, and anisotropy was considered to account for directional spatial dependence. Environmental covariates (sea surface temperature, chlorophyll-a, sea level anomaly, and wind-induced turbulence) were incorporated using smooth functions after imputation of missing values and assessment of collinearity. The resulting model produced a standardised abundance index showing a marked increase from around 2015 onwards, with the highest values observed in recent years.

The model-based index was strongly correlated with the historical index over the common period (2006–2024) (Pearson = 0.93; Spearman = 0.97), indicating consistency in the main temporal trends. However, this agreement is partly driven by a shared long-term trend, and differences in magnitude were observed, particularly in recent years. Sensitivity analyses

indicated that the temporal trajectory of the index is relatively consistent across alternative spatial domains, although substantial differences in absolute magnitude were observed depending on the integration area, highlighting the influence of spatial coverage and potential extrapolation. Uncertainty was moderate for most years but increased in specific periods, and the strong increase observed in recent years should be interpreted with caution.

Overall, the proposed framework provides a flexible approach to derive abundance indices from survey data, but results should be interpreted with caution, particularly in relation to recent increases and spatial extrapolation. Ongoing work focuses on improving model robustness through enhanced validation, evaluation of extrapolation effects, and refinement of environmental and spatial components.

---

*Email address:* [grazia.pennino@ieo.csic.es](mailto:grazia.pennino@ieo.csic.es) (Maria Grazia Pennino)

## 1. Introduction

Acoustic surveys provide key information for the stock assessment of small pelagic species, but derived abundance indices can be affected by changes in spatial coverage and sampling design. In northern Chile, survey coverage has varied substantially over time, which may introduce bias in annual abundance indices and complicate the interpretation of temporal trends. In particular, changes in spatial coverage may confound variation in abundance with shifts in spatial distribution, making it difficult to disentangle ecological signals from sampling artefacts.

In addition, traditional acoustic indices are derived directly from survey observations and do not explicitly account for the ecological and environmental processes influencing species distribution. This limitation may reduce their ability to capture changes in habitat suitability and spatial dynamics, especially in systems where environmental variability plays a key role in structuring distribution patterns. Spatio-temporal modelling approaches provide a flexible framework to address these challenges by explicitly accounting for spatial autocorrelation, temporal structure, and environmental drivers.

Here, we develop a spatio-temporal modelling framework to estimate an alternative abundance index for jack mackerel (*Trachurus murphyi*) in northern Chile, integrating environmental covariates and accounting for spatial structure. The objectives of this study are to: (i) derive a model-based abundance index that accounts for spatial and environmental variability; (ii) evaluate its sensitivity to changes in spatial coverage; and (iii) compare its temporal behaviour with the historical acoustic index currently used in stock assessment.

This work represents an initial step towards developing an environmentally informed abundance index, and results should be interpreted with caution, particularly in relation to spatial coverage, model structure, and associated uncertainty.

## 2. Methodology

Spatio-temporal models were fitted using the R package `sdmTMB`, which implements a stochastic partial differential equation (SPDE) approach to represent continuous spatial fields through a Gaussian random field approximation. Given the high proportion of zero observations and the aggregated nature of acoustic data, a delta modelling framework was adopted. This approach separates the modelling process into two components: (i) a binomial model describing the probability of occurrence, and (ii) a Gamma model describing the positive density conditional on presence.

The initial model formulation included a spatial random field representing persistent spatial structure and a spatio-temporal random field accounting for interannual variability. Several alternative temporal structures were evaluated for the spatio-temporal component, including no temporal structure (off), independent yearly effects (iid), a first-order autoregressive process (AR1), and a random walk (RW). Model comparison was based on a combination of information criteria (AIC) and predictive performance assessed through spatial cross-validation. While the AR1 formulation provided the best fit according to AIC, alternative structures showed comparable or, in some cases, improved predictive performance, highlighting uncertainty in the choice of temporal structure.

To account for potential directional dependence in spatial correlation, anisotropy was explicitly evaluated by comparing isotropic and anisotropic model formulations. The anisotropic model was selected based on improved model fit, although this may also influence the model's extrapolation behaviour.

Environmental covariates were incorporated into the model, including sea surface temperature (SST), chlorophyll-a concentration, sea level anomaly (SLA), and wind-induced turbulence. Missing values in environmental covariates were imputed using a spatial k-nearest neighbours (kNN) approach ( $k = 5$ ), applied independently within each year to preserve spatio-temporal consistency. While this approach allows retention of all observations, it may introduce smoothing effects and additional uncertainty that are not explicitly

propagated into the model.

Covariates were transformed when necessary ( $\log(1 + x)$ ) and standardised (mean = 0, SD = 1). Potential collinearity among covariates was assessed using Pearson correlation matrices, and exploratory analyses using GAM smoothers indicated non-linear relationships with the response. Covariates were therefore modelled using smooth functions, although the flexibility of these functions may increase the risk of overfitting and extrapolation beyond the observed range of environmental conditions.

Annual abundance indices were derived by integrating model predictions over a regular spatial grid covering the study area. As this approach involves prediction in areas with limited or no sampling in some years, the resulting index may be sensitive to spatial extrapolation effects and the definition of the integration domain. To evaluate this effect, a sensitivity analysis was conducted by restricting the integration domain to a comparable area defined from survey footprints. This allowed assessment of the extent to which temporal patterns in the index are influenced by changes in spatial coverage and domain definition.

Finally, the model-derived index was compared with the historical acoustic index over the common period (2006–2024), using standardised series and correlation-based metrics. It is noted that correlation-based comparisons primarily reflect agreement in overall temporal trends and may not fully capture differences in magnitude or short-term variability between indices.

### **3. Results**

#### *3.1. Spatio-temporal model performance*

The spatio-temporal delta model reproduced the main structural features of the data, including the high proportion of zero observations and the spatial aggregation of positive densities. Model diagnostics suggested an overall reasonable fit, although residual variability remained relatively high, as expected for zero-inflated and spatially heterogeneous data.

Alternative spatio-temporal structures (off, iid, AR1 and RW) were evaluated. The

AR1 formulation provided the best model fit according to AIC, whereas the model without a spatio-temporal component (off) showed slightly better predictive performance in cross-validation. This indicates a trade-off between model fit and predictive skill. The AR1 structure was retained to represent temporal persistence, although this choice introduces uncertainty in model selection.

The inclusion of anisotropy improved model fit ( $\Delta\text{AIC} = 25.7$ ), indicating directional spatial dependence consistent with the coastal configuration of the study area. The estimated temporal autocorrelation parameter ( $\rho$ ) was low to moderate, suggesting limited persistence in spatial distribution between consecutive years.

### *3.2. Effect of environmental covariates*

The inclusion of environmental covariates improved model fit relative to both the base and linear formulations. The model incorporating smooth (non-linear) environmental effects achieved the lowest AIC (113925), compared to 114023 for the linear model and 114240 for the base model, indicating a clear improvement in fit.

Environmental effects were more pronounced in the occurrence component than in the positive density component, suggesting that environmental conditions are more strongly associated with habitat suitability than with local biomass levels. In particular, sea surface temperature (SST) and sea level anomaly (SLA) exhibited clear non-linear effects on occurrence, with SST showing an optimal intermediate range and SLA displaying a generally positive relationship. In contrast, chlorophyll concentration and wind-induced turbulence showed weaker and more variable effects.

In the positive density component, environmental effects were generally weaker, indicating that local biomass variability is likely driven by additional processes not explicitly captured in the model, such as aggregation dynamics or unobserved ecological interactions. However, these relationships should be interpreted with caution, as they are based on correlative modelling and may be influenced by spatial structure, imputation of missing values, and potential extrapolation effects.

### *3.3. standardised abundance index*

The standardised abundance index showed relatively low and stable values during the early years of the series, followed by an increase from around 2015 onwards, with the highest values observed in recent years (Figure 1). This increasing pattern was observed in both the base and environmental models, although the environmental model produced a more moderate increase, suggesting that part of the variability captured by the base model may be explained by environmental covariates.

However, the magnitude of the increase in recent years, particularly after 2020, is associated with greater uncertainty, as reflected by wider confidence intervals. Therefore, the magnitude of the recent increase should be interpreted with caution.

The decomposition of the delta index indicates that this increase is associated with both an expansion in spatial occupancy and an increase in positive densities, although the relative contribution of each component varies over time.

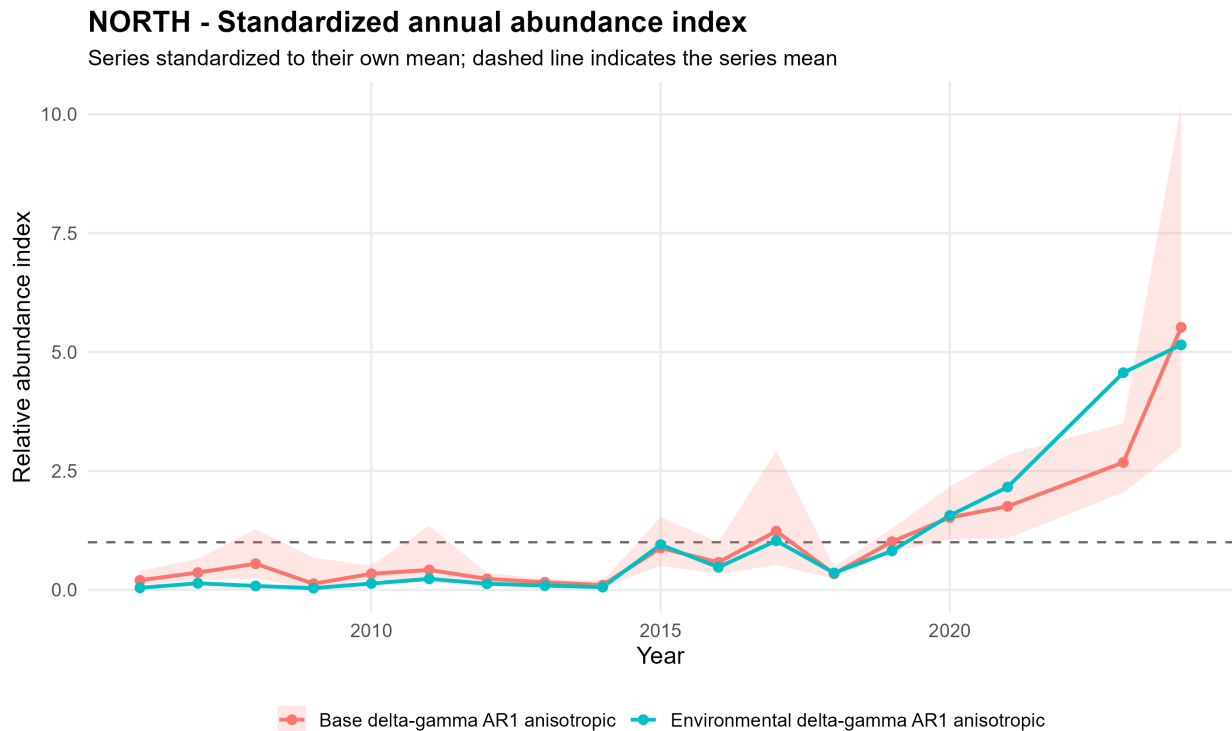


Figure 1: standardised annual abundance indices derived from the base and environmental spatio-temporal delta-gamma models. Shaded areas represent 95% confidence intervals. The strong increase observed in recent years is associated with greater uncertainty.

### 3.4. Sensitivity to spatial coverage

The comparison between the full-domain index and the index restricted to a comparable spatial area showed consistent temporal patterns but substantial differences in magnitude (Figure 2). The ratio between the restricted and full-domain indices ranged from 0.09 to 0.89, indicating that a considerable fraction of the estimated abundance may lie outside the comparable area in some years.

This variability indicates that the absolute magnitude of the index is highly sensitive to the spatial domain used for integration. Despite this, both indices displayed similar interannual trends, indicating that the temporal signal is relatively robust to changes in spatial coverage. However, these results also highlight that part of the variability in the index may reflect changes in spatial distribution and spatial extrapolation effects, rather than purely changes in abundance.

## NORTH - Sensitivity of the abundance index to spatial coverage

Environmental delta-gamma model: full integration domain vs comparable spatial area

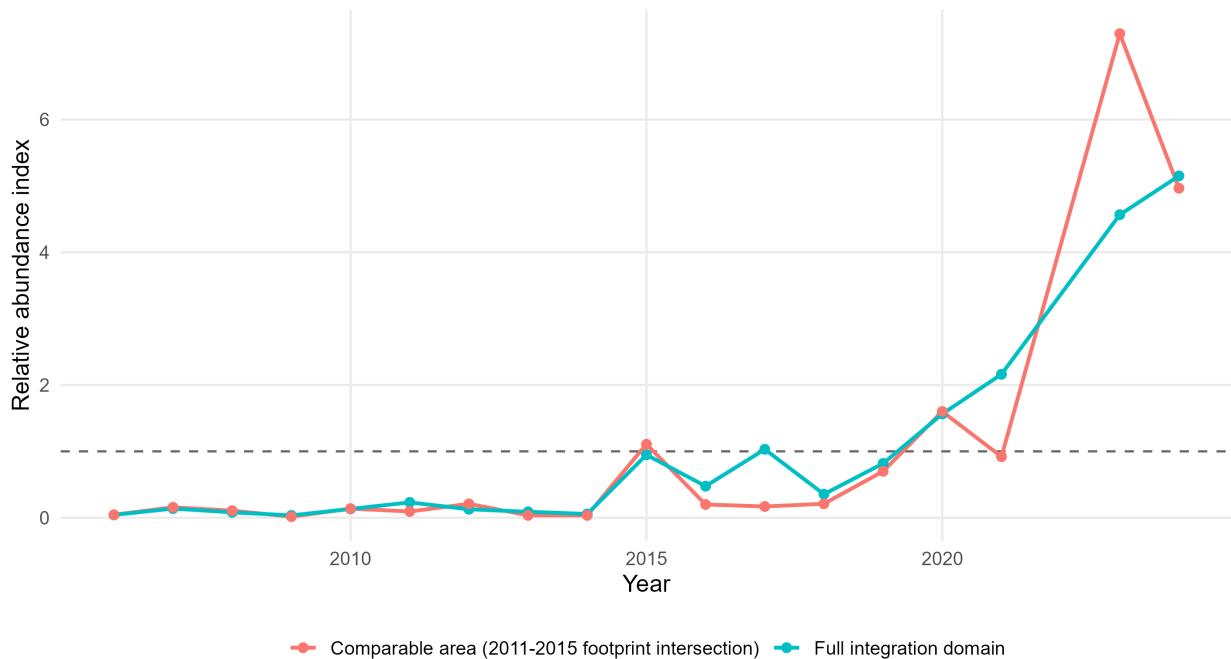


Figure 2: Sensitivity of the abundance index to spatial coverage. Comparison between indices estimated over the full integration domain and a spatially consistent comparable area (2011–2015). Differences in magnitude reflect the influence of spatial domain definition.

### 3.5. Comparison with the historical acoustic index

The model-based index showed a strong correlation with the historical acoustic index over the common period (2006–2024) (Figure 3). Both series exhibited similar long-term trends, including an increase in recent years. However, differences in magnitude were observed, particularly in the most recent years, where the model-based index showed a stronger increase than the acoustic index. This suggests that the model may be amplifying recent increases, potentially due to spatial extrapolation effects or the influence of environmental covariates.

Therefore, while the two indices are consistent in terms of overall temporal patterns, differences in magnitude and recent dynamics indicate that the model-derived index should be interpreted with caution.

### NORTH - Environmental model-based index compared with the historical acoustic ind

Both series are standardized to mean 1 over the common period

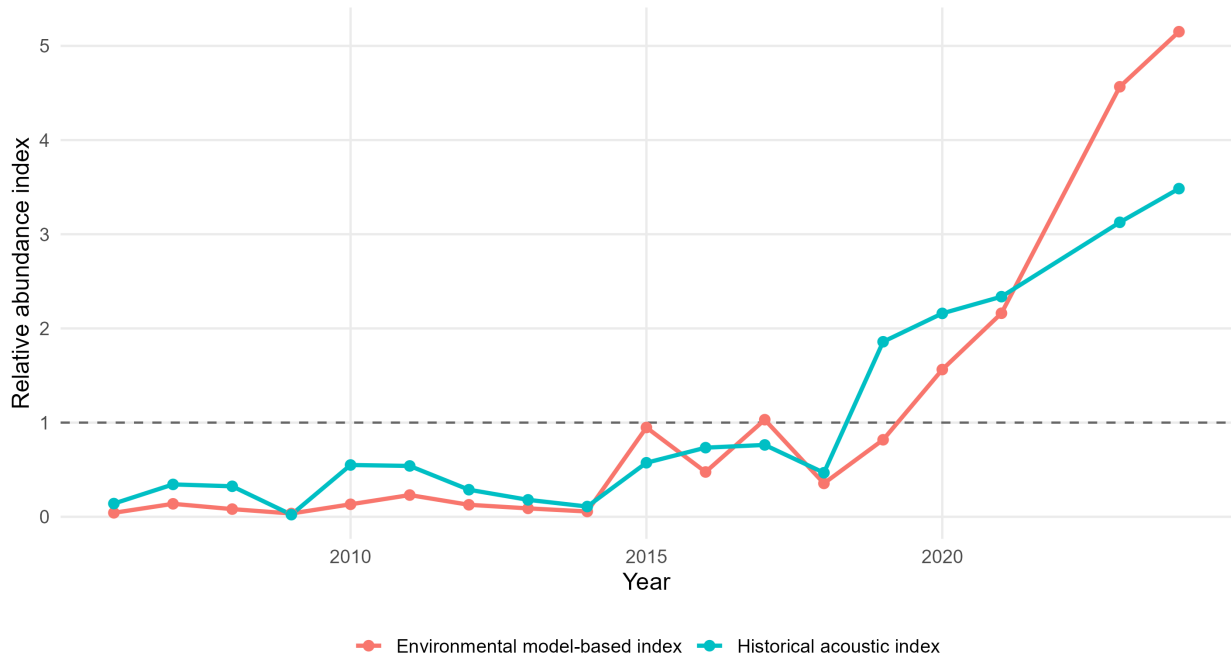


Figure 3: Comparison between the model-based index and the historical acoustic index. Both series are standardised to mean 1. Differences in recent years indicate divergence in magnitude.

## 4. Discussion

The spatio-temporal modelling approach developed in this study provides a useful framework to explore the spatial and temporal dynamics of jack mackerel in northern Chile and to derive an alternative abundance index that integrates environmental variability. The model was able to reproduce the main large-scale temporal patterns observed in the historical acoustic index, suggesting that it captures relevant features of the system.

However, several limitations need to be considered when interpreting the results. In particular, the magnitude of the index appears sensitive to the spatial domain used for integration, indicating that part of the variability may reflect changes in spatial distribution and spatial extrapolation effects rather than purely changes in abundance. This is especially relevant in recent years, where the strong increase in the index is accompanied by higher uncertainty and divergence from the acoustic index.

In addition, environmental covariates were partially imputed, which may introduce additional sources of uncertainty, including imputation error. Taken together, these results suggest that the model-derived index should be interpreted with caution, particularly regarding its absolute magnitude and recent dynamics.

Future work will focus on sensitivity analyses to systematically evaluate the effects of alternative spatial domains and mesh configurations on both the magnitude and temporal trajectory of the index.

## 5. Concluding remarks

- Spatio-temporal modelling substantially improved the representation of Chilean jack mackerel distribution dynamics by explicitly accounting for spatial structure, temporal variability, and environmental conditions.
- The model-derived abundance index reproduced the main long-term temporal patterns observed in the historical acoustic index, indicating consistency in the large-scale dynamics of the system.
- Environmental covariates were more strongly associated with occurrence patterns than with positive density levels, suggesting that environmental variability primarily influences habitat suitability and spatial occupancy.
- Sensitivity analyses demonstrated that the temporal trajectory of the index is relatively robust to changes in spatial coverage, although the absolute magnitude of the index remains highly sensitive to the integration domain and spatial extrapolation effects.
- The strong increase in the index observed in recent years should be interpreted with caution due to increased uncertainty, divergence from the historical acoustic index, and the potential influence of spatial extrapolation and model structure.