NPFC-2025-TWG CMSA11-WP08

What caused the increase in retrospective pattern in the chub mackerel stock assessment in the Northwest Pacific?

Shota Nishijima and Momoko Ichinokawa

*Fisheries Resources Institute, Japan Fisheries Research and Education Agency*

(Email: nishijima\_shota02@fra.go.jp)

**Summary**

In the 2025 provisional assessment of the Northwest Pacific stock of chub mackerel, retrospective patterns quantified by Mohn’s rho for stock biomass and recruitment over the five-year retrospective analysis increased compared to the values in the previous year’s assessment. An investigation into the cause revealed that the primary factor was that all index values for 2023 were lower than the predicted values of the model without 2023 indices, leading to a downward revision of recent stock biomass and recruitment when including the 2023 indices. The one-year shift in the reference period used to calculate Mohn’s rho also contributed to the increase in the values. In contrast, changes in the stock assessment model settings had little effect. It was also found that revisions to catch-at-age data prior to 2022 contributed to a reduction of Mohn’s rho. Retrospective pattens does not necessarily mean estimation bias and rather the update of estimates with new data is a crucial step in stock assessments. For the chub mackerel stock assessment, it is important to include the latest available abundance indices to mitigate future shifts in abundance estimates that would inevitably occur if the latest data were excluded.

# 1. Introduction

In TWG-CMSA of NPFC, the initial stock assessment of chub mackerel in the Northwestern Pacific was finalized using the state-space assessment model (SAM) in the previous year. In the 2024 stock assessment, no substantial retrospective pattern was detected, likely due to the incorporation of process errors for fish older than age 0 (S28-ProcEst(SA2024) in Fig. 1). However, in this year’s updated base case (S01-InitBase, NPFC-2025-TWG CMSA11-WP06), moderately large retrospective patterns were observed, particularly for stock biomass and recruitment (Fig. 1). Since large retrospective patterns are commonly regarded as indicators of instability and unreliability of stock assessments, it is important to investigate the underlying causes of their increase. Therefore, this study examines the effects of individual changes made since last year’s stock assessment on the retrospective pattern. Note that the last year’s base case stock assessment used the data up to the fishing year 2022 and we here compared this model of last year with updated models with data through the fishing year 2023.

# 2. Hypotheses and Their Testing

Considering that the changes made since last year may have contributed to the increase in retrospective pattern, we formulated the following six hypotheses:

1. Changes in model settings increased retrospective pattern.
2. The inclusion of 2023 index values increased retrospective pattern.
3. The inclusion of 2023 catch-at-age data increased retrospective pattern.
4. Revisions to standardized index values prior to 2022 increased retrospective pattern.
5. Revisions to catch-at-age data prior to 2022 increased retrospective pattern.
6. The one-year shift in the reference period used to calculate Mohn’s rho increased retrospective pattern.

Other changes included updates to weight-at-age and the addition of Russian fishery CPUE. However, the revisions to weight-at-age were minor (NPFC-2025-TWG CMSA11-WP03), and we judged their impact on the increased recruitment in the retrospective analysis to be negligible, so this factor was not included as a hypothesis. In addition, the changes in estimates when excluding Russian fishery CPUE were quite small (NPFC-2025-TWG CMSA11-WP06), indicating that its inclusion had little impact. Each of the six hypotheses was examined using the methods described below.

## H01: Changes in model settings increased retrospective pattern

Compared to last year’s base case (S28-ProcEst(SA2024)), this year’s updated base case (S01-InitBase) adopted a different structure for process errors in numbers-at-age and F-at-age and observation errors in catch-at-age as a result of model selection (NPFC TWG CMSA11-WP06). Other model settings remained the same as last year. We estimated parameters and conducted a retrospective analysis using a model in which the process and observation error structures were set to match that of last year’s base case (H01-Modconfig). As a result, the values of Mohn’s rho were almost identical to those of S01-InitBase (Fig. 1), leading us to reject the hypothesis that the change in model settings increased the retrospective pattern.

## H02: The inclusion of 2023 index values increased retrospective patten

To test this hypothesis, parameter estimation was first conducted using the model configuration of S01-InitBase without the 2023 index values. The expected values of the 2023 indices from this model were then used as pseudo-observations in a SAM analysis (without observation errors). A retrospective analysis using this model (H02-Idx\_2023) showed that the absolute values of Mohn’s rho decreased by 0.13–0.16 compared to S01-InitBase (Fig. 1). However, the level of retrospective pattern was still larger than that observed in last year’s base case.

For all seven stock abundance indices, the 2023 index values were consistently lower than the values expected by the model excluding them (Fig. 2). In other words, the decline in the 2023 indices could not have been predicted from data up to 2022. As a result, in this year’s base case (S01-InitBase), recent estimates of stock biomass, spawning stock biomass (SSB), and recruitment were substantially revised downward, which led to the increase in the Mohn’s rho especially in the latest year (Table 2).

## H03: The inclusion of 2023 catch-at-age data increased retrospective pattern

Similar to Hypothesis 2 (H02), an analysis was first conducted excluding the 2023 catch-at-age data. The predicted values from that model were then used as pseudo-data for the 2023 catch-at-age, and parameter estimation was performed. A retrospective analysis using this model (H03-CAA\_2023) showed that the absolute values of Mohn’s rho decreased by 0.01–0.05 compared to the base case (S01-InitBase) (Fig. 1), suggesting that the 2023 catch-at-age data may have contributed slightly to the increase in retrospective pattern.

## H04: Revisions to standardized index values prior to 2022 increased retrospective pattern

All indices were standardized, and changes were also made to the data prior to 2022. To assess their impact, pseudo-data were created by replacing the data up to 2022 for six indices (excluding the Russian CPUE, which was not used in the previous assessment) with the values used in last year’s assessment. Since the indices represent trends in relative abundance and absolute values may vary due to the standardization process, the replacement was made by multiplying last year's index values by the ratio of the mean values of the indices up to 2022 between this year and last year.

A retrospective analysis was conducted using a model (H04-Idx\_to2022) in which only the indices before 2022 were replaced. As a result, the absolute values of Mohn’s rho decreased by 0.01–0.05 compared to the base case (S01-InitBase), suggesting that revisions to the pre-2022 indices may also have been a minor contributing factor to the increase in retrospective pattern.

## H05: Revisions to catch-at-age data prior to 2022 increased retrospective pattern

Since the 2024 stock assessment, discrepancies have been identified between the NPFC Annual Summary Footprint and the data used in the stock assessment (Scientific Committee, 2024). To resolve these discrepancies, the catch-at-age data prior to 2022 were revised from what was used last year (NPFC TWG CMSA11-WP04). To assess the impact of this revision, a model (H05-CAA\_to2022) was analyzed in which the catch-at-age data prior to 2022 were replaced with those used in last year’s assessment, and a retrospective analysis was conducted.

As a result, retrospective pattern as measured by Mohn’s rho increased by 0.02–0.10 compared to the current base case (S01-InitBase). This suggests that the revision of the catch-at-age data prior to 2022 contributed to a *reduction* in retrospective pattern.

## H06: The one-year shift in the reference period used to calculate Mohn’s rho increased retrospective pattern

With the addition of one more year of data, the reference period used to calculate Mohn’s rho also shifted by one year. In last year’s retrospective analysis, the model with 2017 as the terminal year underestimated stock biomass, spawning stock biomass (SSB), and recruitment, while overestimating fishing mortality (F) (pink lines at S28-ProcEst(SA2024) in Fig. 1; Table 1). However, 2017 is no longer included in this year’s Mohn’s rho calculation. In contrast, the model with 2022 as the terminal year—newly included in this year’s calculation—overestimated stock biomass, SSB, and recruitment, while underestimating F (yellow-brown lines at S01-InitBase in Fig. 1; Table 1). This indicates that the one-year shift in the reference period significantly affected the values of Mohn’s rho.

The five-year Mohn’s rho differences between the two models were 0.21 for stock biomass, 0.20 for SSB, 0.20 for recruitment, and 0.36 for F (Table 1). When the reference period was aligned to the common years (2018–2021) between the two analyses, these differences decreased to 0.11 for stock biomass, 0.04 for SSB, 0.18 for recruitment, and 0.14 for F (Table 1). These results suggest that the shift in the reference period accounted for 0.02–0.22 of the change in Mohn’s rho and was a major factor in reversing the direction of retrospective pattern for SSB and F.

# 3. Conclusions

In the current stock assessment, the main factors contributing to the increase in retrospective pattens represented by 5-years Mohn’s rho were: (1) the substantial downward revision of estimated stock biomass due to uniformly low abundance indices in 2023, and (2) the one-year shift in the reference period used to calculate Mohn’s rho. These two factors are interrelated. Specifically, the shift in the reference period resulted in the inclusion of the 2022 estimates in the calculation of Mohn’s rho. Due to the downward revision of stock biomass, the relative bias in 2022 became more pronounced (Table 2), thereby increasing retrospective patterns.

While retrospective analysis is a classical diagnostic tool that have been used to evaluate model robustness (Carvalho et al., 2021; Kell et al., 2021), a few recent studies argues that the presence of retrospective pattern does not necessarily indicate model misspecification or bias in the estimates (Breivik et al., 2023; Cadrin, 2025). In fact, in an extreme case where a model entirely ignores new data, retrospective pattens cannot occur at all. Cadrin (2025) refers to some simulation studies showing that the magnitude of the retrospective pattern was positive (negative), but the bias (compared to the true value) was negative (positive).

In all seven abundance indices used, the observed values for 2023 were uniformly lower than the values expected by the model without 2023 data (Fig. 2), providing strong evidence that the downward revision of stock biomass was an inevitable outcome when incorporating these indices. Although the 2023 index values were not included in last year’s base-case assessment, all but the Chinese CPUE were already available at the time (NPFC-2024-TWG CMSA09-Final Report). It has been shown that if these index values had been used explicitly at that time, the degree of downward revision accompanying this year’s data update would have been smaller (NPFC-2025-TWG CMSA11-WP06). In addition, sensitivity analysis on the assumptions made when using the most recent year’s indices shows that robust estimates can be obtained regardless of these assumptions (NPFC-2025-TWG CMSA11-WP07). Generally, the use of abundance index ahead of catch data will be effective for increasing the robustness and forecasting skill of fish population dynamics (Le Pape et al., 2020; Nishijima et al., 2023). Therefore, to improve both the accuracy and robustness of stock biomass estimates, it is recommended that the most recent abundance indices (i.e., 2024 in this year’s case) be included in the chub mackerel stock assessment.

 While the 2023 abundance indices contributed to an increase in retrospective pattern, the revisions to catch-at-age data prior to 2022 appear to have reduced the bias. This is evident from the fact that substituting the catch-at-age data with those used in the previous stock assessment resulted in greater retrospective pattern (Fig. 1). After the previous assessment was completed, discrepancies were identified between the sum product of catch-at-age and weight-at-age data submitted by member countries and the official statistics from NPFC’s Annual Footprint. To resolve these inconsistencies, revisions were made to the catch-at-age data. This kind of data verification process and the use of reliable data are considered to contribute significantly to the robustness of stock assessments.

 Other hypotheses were found to have a small influence on retrospective pattens (Fig. 1). The catch-at-age data for the 2023 fishing year led to downward revisions in the estimates of stock biomass and related quantities, which slightly increased retrospective patterns. The updates to abundance indices prior to 2022 also contributed marginally to the increase in retrospective pattern. In contrast, changing the specification of the process error for number at age had little impact on the magnitude of retrospective bias. Thus, the observed increase in retrospective pattern was not attributable to changes in model configuration. The SAM framework yields relatively robust estimates regardless of model configuration (NPFC-2024-TWG CMSA09-WP04). By estimating process errors in population numbers and observation errors in catch-at-age, the model can flexibly fit to the abundance indices. All of the available indices similarly show a recent decline of the stock through 2023, suggesting that there is no conflict among indices and the downward revision of stock estimates would be inevitable and valid.

# 4. REFERENCES

Breivik, O. N., Aldrin, M., Fuglebakk, E., & Nielsen, A. (2023). Detecting significant retrospective patterns in state space fish stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *80*(9). https://doi.org/10.1139/cjfas-2022-0250

Cadrin, S. X. (2025). Misinterpreting retrospective patterns in fishery stock assessment. *ICES Journal of Marine Science*, *82*(2). https://doi.org/10.1093/icesjms/fsaf014

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzel, C. R., Doering, K., Johnson, K. F., & Methot, R. D. (2021). A cookbook for using model diagnostics in integrated stock assessments. *Fisheries Research*, *240*, 105959-105959. https://doi.org/10.1016/J.FISHRES.2021.105959

Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., & Fu, D. (2021). Validation of stock assessment methods: is it me or my model talking? *ICES Journal of Marine Science*, *78*(6), 2244-2255. https://doi.org/10.1093/ICESJMS/FSAB104

Le Pape, O., Vermard, Y., Guitton, J., Brown, E. J., van de Wolfshaar, K., Lipcius, R. N., Støttrup, J. G., & Rose, K. A. (2020). The use and performance of survey-based pre-recruit abundance indices for possible inclusion in stock assessments of coastal-dependent species. *ICES Journal of Marine Science*, *77*(5), 1953-1965. <https://doi.org/10.1093/ICESJMS/FSAA051>

ManabeA., GaziK. M., Oshima K. (2025) The data description for the base case stock assessment of chub mackerel *Scomber japonicus* in the northwestern Pacific Ocean for 2025 assessment, NPFC-2025-TWG CMSA11-WP03.

Nakayama, S., Nishijima, S., Ichinokawa, M., Manabe, A., Oshima, K. and Rice, J. (2025) Candidate base case scenarios for the 2025 chub mackerel stock assessment in the Northwest Pacific Ocean, NPFC-2025-TWG CMSA11-WP06.

Nishijima, S., Suzuki, S., Fukuta, R., & Okada, M. (2023). Modeling pulse dynamics of juvenile fish enables the short-term forecasting of population dynamics in Japanese pufferfish: a latent variable approach and hindcasting. *Canadian Journal of Fisheries and Aquatic Sciences*, *80*(6), 1017-1031.

Nishijima, S., Ichinokawa, M., Manabe, A., Oshima, K. and Rice, J. (2024) Sensitivity of the chub mackerel stock assessment in 2024 to observation and model uncertainty in the Northwest Pacific Ocean, NPFC-2024-TWG CMSA09-WP04.

North Pacific Fisheries Commission (2024) 9th Meeting of the Technical Working Group on Chub Mackerel Stock Assessment. NPFC-2024-TWG CMSA09-Final Report.

Scientific Committee. 2024. 9th Meeting Report. NPFC-2024-SC09-Final Report. 405 pp. (Available at www.npfc.int)

# Table 1

Relative retrospective bias each year, the five-year average (2018-2022 for S01-InitBase and 2017-2021 for S28-ProcEst(SA2024)), and the four-year average from 2017-2021 in both models for biomass, SSB, recruitment, and F.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Biomass | SSB | Recruitment | F |
| Year | S01 | S28 | S01 | S28 | S01 | S28 | S01 | S28 |
| 2022 | 0.434 | - | 0.458 | - | 0.281 | - | −0.410 | - |
| 2021 | 0.317 | 0.087 | 0.114 | 0.032 | 0.277 | 0.031 | −0.208 | 0.012 |
| 2020 | 0.396 | 0.282 | 0.148 | 0.095 | 0.438 | 0.233 | −0.294 | −0.185 |
| 2019 | 0.321 | 0.274 | 0.059 | 0.047 | 0.239 | 0.116 | −0.202 | −0.124 |
| 2018 | 0.158 | 0.109 | −0.117 | −0.143 | 0.340 | 0.199 | 0.035 | 0.179 |
| 2017 | −0.108 | −0.157 | −0.319 | −0.365 | −0.095 | −0.014 | 0.445 | 0.837 |
| Ave2018–2022 or Ave2017–2021\* | 0.325 | 0.119 | 0.132 | −0.067 | 0.315 | 0.113 | −0.216 | 0.144 |
| Ave2018–2021 | 0.298 | 0.188 | 0.051 | 0.008 | 0.323 | 0.145 | −0.167 | −0.029 |

\* The five-year averages of 2018–2022 and 2017–2021 were used for calculating Mohn’s rho of S01-InitBase and S28-ProcEst (the base case in the last year’s stock assessment), respectively. This means that the values of 2017 in S01 are not included.

# Figure 1

****Retrospective patterns in each model (column) for total biomass (upper), SSB (second row), recruitment (third row), and mean F. Mohn’s rho is shown at the upper-right corner. Different colors represent differences in the terminal year of data used in the retrospective analysis. S01-InitBase is the simple update from the base case of previous assessment (S28-ProcEst) with small modifications of model configurations and additional data of 2023. The model configurations from H01 to H05 are described in the main text.

# Figure 2


The predicted index values (blue lines) in the model without indices in 2023 (H02-Idx\_2023) and the actual indices in 2023 (red dots). Circles represent indices through 2022.