CCSBT-ESC/2608/38
(ESC Agenda item 9)

2026 年ミナミマグロ資源評価のための GAM22 の変動係数の計算

Calculation of the coefficient of variation of GAM22 CPUE index for the SBT assessment in 2026

伊藤智幸、ダーシー・ウェバー、アナ・パルマ
Tomoyuki ITOH[1], Darcy WEBBER[2], Ana PARMA[3]

1: Fisheries Resources Institute, Japan Fisheries Research and Education Agency

[2] Quantifish Ltd, Tauranga, New Zealand

[3] Centro Nacional Patagonico

要約

本文書は、資源評価に用いた GAM22 の年可変の CV 計算について示す。シミュレーションで求めた CV 値は極めて低い年があったため、20%以上になるように調整をした。2025 年には CV 値が高かったが、これは 2025 年の特異的現象ではなく、データの最近年であるためと示唆された。

Summary

This document describes the calculation of the year-specific CV for GAM22, which was used in the stock assessment. As the CV values derived from the simulation were very low in some years, their log-scale variances were added to a baseline variance corresponding to a log-SD of 0.20. Although the CV value was high for 2025, retrospective analyses showed a similar pattern in the final two to three years of each series.

Introduction

In the CCSBT stock assessment, "GAM22"—derived from the CPUE of the Japanese longline fishery—is used as the stock abundance index. Previously, the coefficient of variation (CV) for longline CPUE was represented by a constant log-scale SD of 0.20 across all years. Standard GAM22 production did not provide year-specific CV estimates. With the introduction of new stock assessment software based on RTMB, the incorporation of a time-varying CV was considered at OMMP16 (Anon. 2026). This document describes the simulation-based method for calculating the CV and its treatment in the assessment. It also examines the increase in CV values observed over the past two to three years.

Method for calculating the CV

The CV was estimated through simulation. In GAM22, predicted values are obtained by applying the fitted GAM components to a dummy dataset (Itoh and Takahashi 2026). The same dummy dataset and area-weighting scheme were used to calculate the CV.

One thousand coefficient sets were simulated; the approximately 146,000-row dummy dataset was processed in blocks of 10,000 rows for memory management, not resampled. For the binomial and CPUE sub-models, coefficients were simulated separately from multivariate normal distributions using the fitted coefficients and unconditional variance-covariance matrices. Predicted CPUE values were obtained by combining these components, and an annual index was calculated using the same area weighting as GAM22. The annual raw CV was calculated by dividing the standard deviation of the annual index by its mean.

The R code used is provided in Appendix A.

Adjustment of CV

The raw CV (CV_{raw}) in early years was well below the level of 20% used in past assessments. This point was considered a concern, and to avoid overweighting the early CPUE indices in the assessment, it was decided to retain the historical log-SD of 0.2 and add the raw year-specific GAM22 estimation variance in log space (Fig. 1).

$$CVassessment = sqrt(exp[0.2^2 + log(1 + CVraw^2)] - 1)$$

The raw CVs range from 1.27% to 21.96% with a mean of 5.41%; the resulting assessment CVs range from 20.24% to 30.17%, with a mean of 21.34%.

Increase of CV in recent years

When calculating the CV using catch data up to 2025, the value for 2025 was relatively high. Given that the stock abundance index also rose significantly in 2025, we investigated

whether this high CV value was unique to that year. We calculated CVs using dummy data subjected to a retrospective analysis covering the past 10 years within the GAM22 framework. The results showed that CV values tended to be elevated in the final two to three years of each series, with a trend of increasing values as the data approached the present (Fig. 2). This repeated pattern is consistent with a terminal-year effect and does not, by itself, identify a 2025-specific event.

Reference

Anon. (2026). Report of the Sixteenth Operating Model and Management Procedure Technical Group. Commission for the Conservation of Southern Bluefin Tuna, 22-26 June, Seattle, USA.

Itoh, T., and N. Takahashi (2026). Update of CPUE abundance index using GAM for southern bluefin tuna in CCSBT (GAM22) up to the 2025 data. CCSBT-OMMP/2606/06.

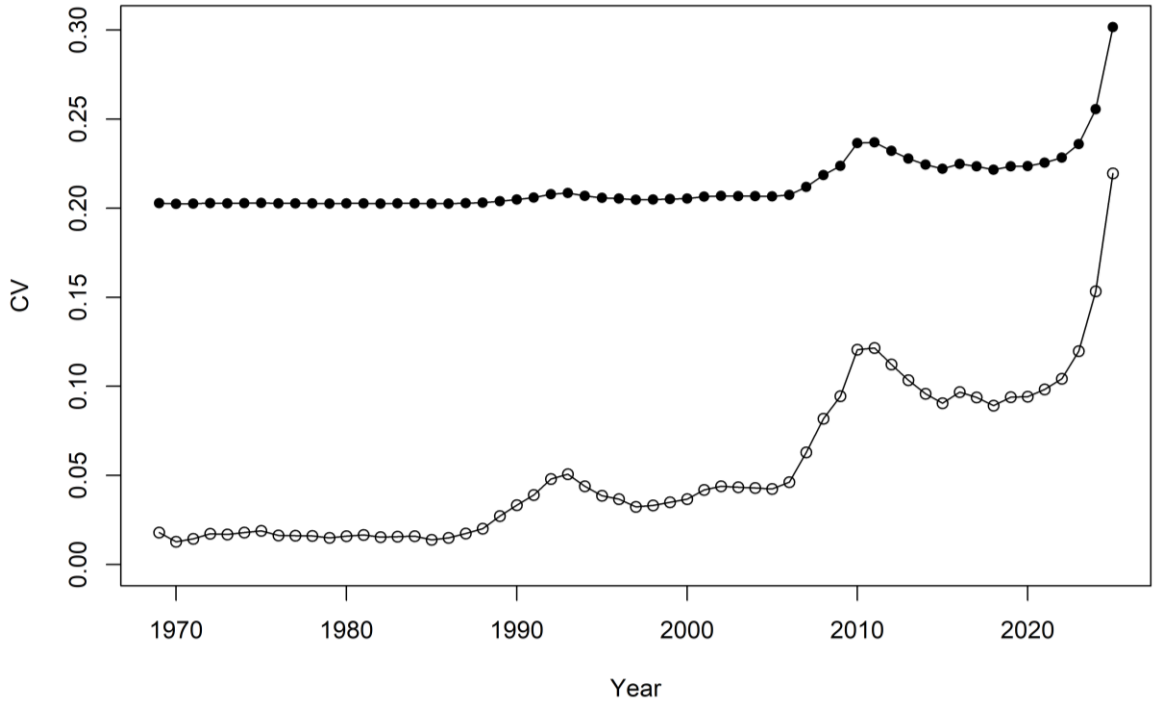


Figure 1 Coefficient of variation of GAM22.

Open and closed circles denote raw and assessment CV values, respectively.

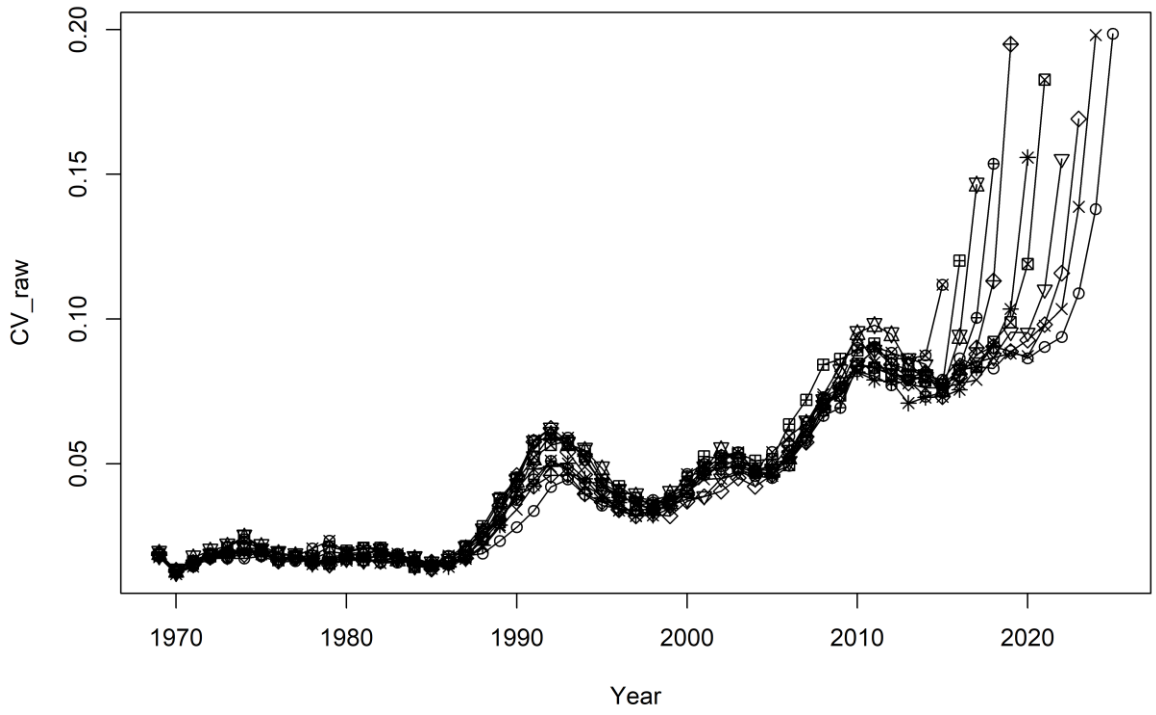


Figure 2 Coefficient of variation of GAM22 in the retrospective analysis.

Different symbols show different terminal years of the data. Note that the CV values are raw and have not been combined with the baseline assessment uncertainty.

Appendix A R code used to calculate the coefficient of variation of GAM22

```
# input data and parameters
res1        #result of GAM for the binary submodel
res2        #result of GAM for the positive-catch submodel
dumA     #dummy data for GAM prediction
B = 1000
baseline_log_sd = 0.20
lognormal_bias_correction = TRUE
unconditional = TRUE
seed = 102
row_block_size = 10000
sim_block_size = 100

              set.seed(seed)

              yrs <- sort(unique(dumA$year))
              year_id <- match(dumA$year, yrs)
              Areas <- dumA$weiM5
              n_year <- length(yrs)
              n <- nrow(dumA)



              ## Variance-covariance matrices for the fitted GAM coefficients
              V_bin <- vcov(res1, unconditional = unconditional)
              V_pos <- vcov(res2, unconditional = unconditional)

              ## Simulate coefficients from both fitted GAMs
              beta_bin <- MASS::mvrnorm(
                 n = B,
                 mu = coef(res1),
                 Sigma = V_bin,
                 tol = 1e-6
              )

              beta_pos <- MASS::mvrnorm(
                 n = B,
                 mu = coef(res2),
                 Sigma = V_pos,
                 tol = 1e-6
              )

              beta_bin <- matrix(beta_bin, nrow = B)
              beta_pos <- matrix(beta_pos, nrow = B)

              ## Residual variance from the positive/lognormal model.
              ## This assumes res2 was fitted to log(positive catch).
              #sigma2 <- summary(res2)$scale
                    sigma2 <- res2$sig2

              ## Storage for simulated annual index sums
              index_sum <- matrix(
                 0,
```

```
  nrow = B,
  ncol = n_year
)
colnames(index_sum) <- yrs

## Annual area-weight denominators
area_sum <- rep(
  0,
  length.out = n_year
  )
names(area_sum) <- yrs

row_starts <- seq.int(1, n, by = row_block_size)
sim_starts <- seq.int(1, B, by = sim_block_size)

for (row_start in row_starts) {
  row_end <- min(row_start + row_block_size - 1, n)
  rows <- row_start:row_end
  nd <- dumA[rows, , drop = FALSE]
  yr_block <- year_id[rows]
     X_areas <- Areas[rows]
  ## Accumulate annual area-weight denominators once per row block
  for (j in unique(yr_block)) {
    rows_j <- yr_block == j
    area_sum[j] <- area_sum[j] + sum(X_areas[rows_j])
  }

  ## Linear predictor matrices for this block of dummy data
  X_bin <- predict(res1, newdata = nd, type = "lpmatrix")
  X_pos <- predict(res2, newdata = nd, type = "lpmatrix")

  for (sim_start in sim_starts) {
    sim_end <- min(sim_start + sim_block_size - 1, B)
    sims <- sim_start:sim_end

    ## Simulated linear predictors
    eta_bin <- X_bin %*% t(beta_bin[sims, , drop = FALSE])
    eta_pos <- X_pos %*% t(beta_pos[sims, , drop = FALSE])

    ## Binomial probability of positive catch
    p <- plogis(eta_bin)

    ## Positive catch prediction
    if (lognormal_bias_correction) {
      mu_pos <- exp(eta_pos + 0.5 * sigma2)
    } else {
      mu_pos <- exp(eta_pos)
    }

    ## Delta-GAM prediction
    pred <- p * mu_pos

     # Accumulate area-weighted numerators by year
    for (j in unique(yr_block)) {
      rows_j <- yr_block == j
      index_sum[sims, j] <- index_sum[sims, j] +
        as.vector(crossprod(X_areas[rows_j], pred[rows_j, , drop = FALSE]))
```

```
      }
    }
}

## Divide once by the total annual area weight
index_sim <- sweep(index_sum, 2, area_sum, "/")
## Raw model-uncertainty CV by year
cv_raw <- apply(index_sim, 2, sd, na.rm = TRUE) /
  apply(index_sim, 2, mean, na.rm = TRUE)

## Combine raw estimation CV with the baseline assessment log-SD
sigma_raw2 <- log1p(cv_raw^2)
sigma_assessment2 <- baseline_log_sd^2 + sigma_raw2
cv_assessment <- sqrt(expm1(sigma_assessment2))
out <- data.frame(
  year           = yrs,
  index_mean = as.numeric(apply(index_sim, 2, mean, na.rm = TRUE)),
  cv_assessment = as.numeric(cv_assessment),
  cv_raw         = as.numeric(cv_raw)
)
```