**NSIN/NGA Global Fishing Forecast Grand Challenge
Forecast Submission Guidelines**

Participants will submit fishing effort forecasts in the form of raster data saved in ".tif" format. To be competitive, it is critically important that the forecast data are submitted in accordance with the Forecast Schedules and the Forecast Data Technical Specifications provided below. **Late or incorrectly formatted data will receive zero points toward the final score**.

Participants will forecast the geolocation and magnitude of daily and weekly fishing effort within three Regional Fisheries Management Organization (RFMO) areas:

- Joint Technical Commission of the Maritime Front (CTMFM)
- Forum Fisheries Agency (FFA)
- Southwest Indian Ocean Fisheries Commission (SWIOFC)

For each RFMO area, information necessary to complete the challenge is detailed in the "Forecast Data Technical Specifications" section at the end of this document. Although the spatial boundaries of these RFMOs are not necessary for completion of the challenge, we link them below for reference.

RFMO	Links to RFMO territorial boundaries
CTMFM	https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/fao-rfb-map-ctmfm
FFA	https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/fao-rfb-map-ffa
SWIOFC	https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/fao-rfb-map-swiofc

Forecast Schedules

In the Proposal Round **(Phase 1)**, participants have three successive attempts to forecast daily and weekly fishing effort one to three weeks into the future. For each attempt submitted on a Friday, participants will submit 21 daily and 3 weekly forecasts spanning the time period of one, two, and three weeks into the future. However, the scoring team will score only the forecasts through the third week from the first Phase 1 submission date, per the timeline below. The unscored forecasts will be used for market research.
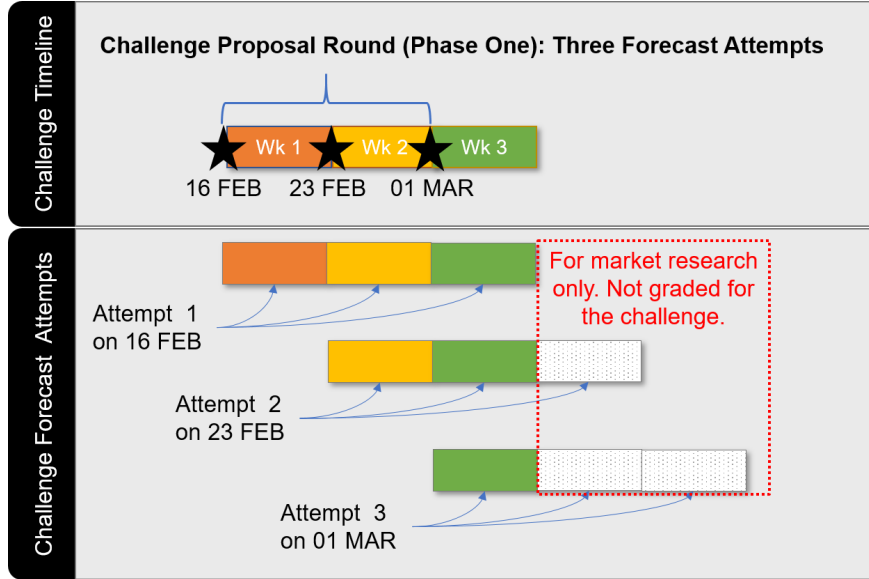
In the Pitch Round **(Phase 2)**, the finalists have five successive attempts to forecast daily and weekly fishing effort one to five weeks into the future. For each attempt submitted on a Friday, participants will submit 35 daily and 5 weekly forecasts spanning the time period of one, two, three, four, and five weeks into the future. However, the scoring team will count only the scores for the forecasts through the fifth week from the first Phase 2 submission date, per the timeline below.

Approved for public release, NGA-U-2023-02867
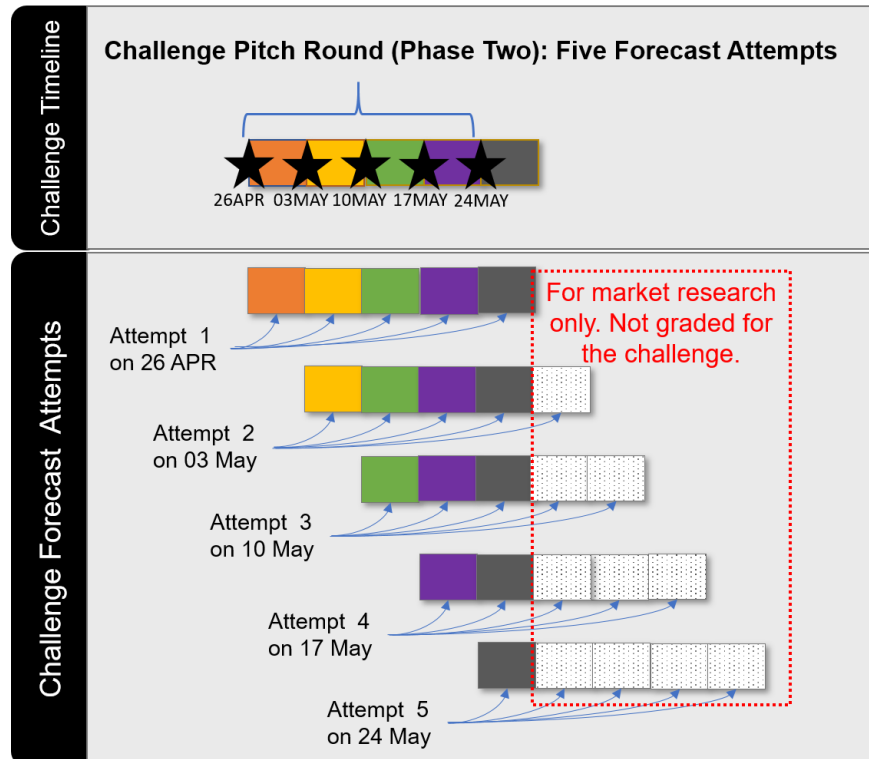
Participants will submit each forecast attempt on the Friday preceding the designated forecast week by 23:59 **Eastern Time**.

For consistency with the Global Fishing Watch Apparent Fishing Effort data product, each forecast week starts on Sunday 00:00 and ends on the following Sunday 00:00 **Zulu Time**.

Phase 1 Forecast Schedule

S	M	T	W	T	F	S
				Feb 1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	Mar 1	2
3	4	5	6	7	8	9

Black outlines denote forecast submission dates. Colored dates correspond with that of the Phase 1 forecast schedule graphic above.

Phase 2 Forecast Schedule

S	M	T	W	T	F	S
				Apr 25	26	27
28	29	30	May 1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	Jun 1

Black outlines denote forecast submission dates. The colored dates correspond with that of the Phase 2 forecast schedule graphic above.

Forecast Data Technical Specifications

Participants will submit numerical raster data (64-bit floating point) for each of the three RFMO raster extents outlined in the table below. The raster values are the forecasted total time fishing effort for all vessels for the specified forecast days or weeks, i.e., total time **per day or per week**, per 10th degree (~10 square kilometer) spatial resolution raster grid cell for the CTMFM, SWIOFC, and FFA RFMO areas.

The table below provides the parameters for the RFMO area raster coverages that **participants must match <u>exactly</u>** in their forecast submissions.[1]

RFMO	NROWS	NCOLS	NCELL	X min	X max	Y min	Y max	Cell Size	Coordinate Reference System
CTMFM	73	86	6278	-59.35	-50.75	-39.65	-32.35	10th degree	
SWIOFC	551	501	276051	29.95	80.05	-45.05	10.05	10th degree	EPSG: 4326 – WGS 84 - Geographic
FFA	764	3601	2751164	-180.05	180.05	-58.45	17.95	10th degree	

We generated the raster parameters in the table above by expanding the original Global Fishing Watch (GFW) fishing effort raster data to the <u>bounding boxes</u> of the respective RFMO territorial boundaries linked in the first page of this document.[2] The scoring team masks the forecasted and the GFW data to the actual RFMO territorial boundaries (irregular polygons) before scoring.

The file names for the forecast data submissions must follow the following naming convention:

- Use only upper-case letter characters, numeric digits, and underscores in the file name. All of the data descriptors are separated by underscores in the file name.
- File name begins with a unique upper-case 5-letter code assigned to each participant: e.g., FISHR.
- The RFMO acronym follows the 5-letter participant code, i.e., CTMFM, SWIOFC, or FFA.
- Date range of the included forecast, designated as yyyymmdd_yyyymmdd, follows the RFMO acronym.
- Date that the forecast is submitted, designated as yyyymmdd, follows the forecast date range.
- End file name with .tif extension.

Example of a correctly named daily file: **FISHR_ FFA_20240211_20240212_20240209.tif**, i.e., Sunday, 11 Feb 2024, 00:00 Zulu time to Monday, 12 Feb 2024, 00:00 Zulu Time. File submitted for the 9 Feb 2024 forecast practice attempt.

Example of a correctly named weekly file: **FISHR_ CTMFM_20240211_20240218_20240209.tif** i.e., Sunday, 11 Feb 2024, 00:00 Zulu time to Sunday, 18 Feb 2024, 00:00 Zulu Time. File submitted for the 9 Feb 2024 forecast practice attempt.

---

[1] Regional Fisheries Management Organization (RFMO); number of rows (NROWS); number of columns (NCOLS); number of cells (NCELL), minimum x coordinate (Xmin); maximum x coordinate (Xmax); minimum y coordinate (Ymin), maximum y coordinate (Ymax).

[2] Using R: extend {raster}: raster_e = extend(raster, RFMO_BBOX, value=0, snap="out"). The bounding box in Appendix A refers to cell edge, not the cell centroid.

Scoring Method

The scoring team will compare the participant forecasts against the corresponding Global Fishing Watch (GFW) AIS-based Apparent Fishing Effort Data product downloaded directly from the GFW map: https://globalfishingwatch.org/map.

The scoring team will compare the forecast and GFW data using mean absolute scaled error (MASE) and Cohen's Kappa.

MASE will be used to assess forecast performance on the non-zero subset of the GFW Apparent Fishing Effort coverages. This assessment will measure how closely the forecast approximates the magnitude of "true" fishing effort in locations where fishing effort was actually observed:

$$\text{MASE} = \text{mean} \left(\frac{|e_j|}{\frac{1}{T-1} \sum_{t=2}^{T} |Y_t - Y_{t-1}|} \right) = \frac{\frac{1}{J} \sum_j |e_j|}{\frac{1}{T-1} \sum_{t=2}^{T} |Y_t - Y_{t-1}|}$$

where the numerator e_j is the forecast error for a given period (with J, the number of forecasts), defined as the actual value (Y_j) minus the forecast value (F_j) for that period: $e_j = Y_j - F_j$, and the denominator is the mean absolute error of the one-step "naive forecast method" on the training set (here defined as $t = 1..T$), which uses the actual value from the prior period as the forecast: $F_t = Y_{t-1}$.

The submitted forecast and the GFW data will then be converted to binary coverages, and Cohen's Kappa will be used to measure how closely the forecasted footprint of fishing effort matches that measured by GFW:

$$\kappa \equiv \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e},$$

where p_o is the relative observed agreement among raters, and p_e is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category. If the raters are in complete agreement then $\kappa = 1$. If there is no agreement among the raters other than what would be expected by chance (as given by p_e), $\kappa = 0$. It is possible for the statistic to be negative, which can occur by chance if there is no relationship between the ratings of the two raters, or it may reflect a real tendency of the raters to give differing ratings.

Each challenge participant will receive one overall MASE score and one overall Cohen's Kappa score, each the average of all individual forecast scores. The scoring team will count late or incorrectly formatted submissions as scores of zero in the final point scores.