

A Unified Reef Visual Census (RVC) and REEF Citizen Science Index

Brice X. Semmens and Dan A. Greenberg

SEDAR94-AP-01

26 March 2025



This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Please cite this document as:

Semmens, Brice X., and Dan A. Greenberg. 2026. A Unified Reef Visual Census (RVC) and REEF Citizen Science Index. SEDAR94-AP-01. SEDAR, North Charleston, SC. 11 pp.

A Unified Reef Visual Census (RVC) and REEF Citizen Science Index

The Reef Environmental Education Foundation (REEF) Volunteer Fish Survey Project is a citizen science initiative that collects opportunistic fish sightings from thousands of volunteer recreational divers, resulting in a dataset with unparalleled spatial and temporal extent. This index uses the multivariate state-space modeling approach from Greenberg et al. (2024) to formally integrate the REEF citizen science data with the scientifically structured Reef Visual Census (RVC) data. This integrated approach leverages the extensive sampling of the REEF data to augment and fill gaps in the more temporally limited RVC data, allowing for information sharing and the creation of augmented RVC indices for separate management regions.

Authors: Brice X. Semmens and Dan A. Greenberg

Working Paper Number: SEDAR 94-WP-XX

Data Type: Fishery Independent

Time Series: 1993-2024

Sampling Intensity: The REEF survey is an opportunistic citizen science program where sampling intensity varies based on volunteer diver effort. The RVC survey is a scientifically rigorous survey employing a probabilistic, stratified-random sampling design. The combined survey effort for both programs in the Southeast Florida and Florida Keys regions is detailed in Tables 1 and 2 of the working paper.

Major Changes to Sampling Design/Gear Changes: The methodologies for the RVC and REEF surveys have been consistent over the time series. The significant change is the novel modeling framework that integrates these two disparate datasets to produce indices of abundance that are jointly estimated and can influence each other. This is particularly useful for when there are missing years (e.g. with changes to RVC sampling frequency in later years) to help fill in the gaps in the time-series.

Size/Age Data: The RVC program collects both fish counts and associated length-frequency measurements. A key strength of the integrated model is that it preserves the link between the final abundance index and this essential RVC length data. REEF has recently begun collecting length data via the SMILE project, but there are currently limited Hogfish measurements available from this initiative.

Data Filtering Techniques: REEF survey data were filtered prior to analysis using the following procedures:

- Dives shorter than 20 minutes or longer than 120 minutes were excluded.
- Night dives (surveys starting before 5:00 AM or after 8:00 PM) were removed.
- Only surveys from experienced divers were included (≥ 8 surveys over the period)

- Only data from well-surveyed sites were retained (≥ 5 surveys with at least 3 years of observations).
- Zero-observation records were generated to ensure a complete presence/absence record for each survey.

Standardization & Modeling

This analysis incorporates two key methodological innovations over historical approaches for the RVC index:

- **State-Space Framework:** A state-space framework is used to filter out year-to-year high-frequency sampling error from long-term structural process error.
- **Information Sharing:** An approach is employed that shares information with abundance time-series estimated from citizen science (REEF) to fill in temporal gaps in the RVC data.

A multivariate Bayesian state-space model, based on the framework of Greenberg et al. (2024), was used to generate the index. This approach assumes that the true population abundance for each survey changes from year to year according to a random walk. The model estimates two separate, latent population trajectories (state vectors), one informed by the RVC data ($x_{RVC,t}$) and the other by the REEF data ($x_{REEF,t}$).

The primary goal of this analysis is to generate augmented annual state estimates from the RVC survey, using REEF observations to improve precision and fill in data gaps, which is particularly valuable in years and locations where RVC surveys are missing. To accomplish this goal, we parameterized a model that simultaneously estimates the temporal correlation (ρ) between the year-to-year fluctuations of the two state vectors. This correlation term allows the model to "borrow information" between the two time series. The contribution of the REEF time series to the RVC state estimates is therefore up- or down-weighted based on the extent to which the two observation programs agree, as measured by ρ .

Submodel Variables

The overall model is composed of two observation models linked to a process model with two correlated state processes.

- **Process Model:** The two latent population states are assumed to evolve based on a correlated random walk process:

$$\begin{bmatrix} x_{RVC} \\ x_{REEF} \end{bmatrix}_t = \begin{bmatrix} x_{RVC} \\ x_{REEF} \end{bmatrix}_{t-1} + \begin{bmatrix} w_{RVC} \\ w_{REEF} \end{bmatrix}_t, \text{ where } \begin{bmatrix} w_{RVC} \\ w_{REEF} \end{bmatrix}_t \sim \text{MVNormal} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma \right)$$

- The covariance matrix Σ contains the process variances for each survey on the diagonal of the matrix and the covariance on the off-diagonal, which is a function of the correlation parameter ρ .

- **RVC Observation Model:** Annual pre-computed RVC index estimates (mean encounter rates and associated standard errors, as reported in Tables 1 and 2) were modeled using a **Normal** distribution. Each annual estimate was linked to the RVC-specific latent state vector ($x_{RVC,t}$), with observation uncertainty directly propagated from the survey-derived standard errors.
- **REEF Observation Model:** REEF's ordinal abundance data were modeled using an **ordinal logistic regression**. A latent continuous abundance for each survey was linked to the REEF-specific state vector ($x_{REEF,t}$) and standardized using a comprehensive set of covariates, including dive site, diver identity, habitat type, diver experience level, month, date, dive duration, depth, current, and visibility.

Abundance Indices

The final augmented annual indices of abundance for Southeast Florida and the Florida Keys are presented in figures 1 and 2, respectively, below. In addition, the

Uncertainty and Measures of Precision

Uncertainty in the annual abundance estimates is quantified using Bayesian credible intervals derived from the posterior distributions of the state-space model.

Comments on Adequacy for Assessment

A central motivation for this integrated modeling approach is that the RVC and REEF surveys have complementary strengths that, when combined, yield more informative state estimates than either survey can provide alone. The RVC is a scientifically rigorous, probabilistic survey that produces density-based encounter rate estimates (fish per unit area), but its temporal coverage is uneven, particularly in Southeast Florida, where surveys were not conducted prior to 2013 and were intermittent thereafter (Table 1). The REEF survey, by contrast, has been continuously conducted since the early 1990s across the full study region, generating a long and spatially extensive record of volunteer diver observations. By jointly modeling these two time series within a bivariate state-space framework, information from the REEF data is able to constrain RVC state estimates in years when no RVC survey was conducted, effectively bridging temporal gaps in the management-critical index. Crucially, the degree to which REEF informs the RVC state is governed by the estimated correlation parameter (ρ) between the two time series—if the surveys track population trends in agreement, more information is shared; if they diverge, each series is weighted more independently. This adaptive borrowing of strength reduces uncertainty in RVC state estimates without imposing the REEF signal where it is not supported by the data.

This augmented RVC index, where RVC state estimation is informed by REEF data to the extent the two time series covary, holds promise as an index of abundance, and is recommended for use in sensitivity analysis for the SEDAR 94 stock assessment.

1. **Resolves Data Gaps:** It resolves a critical issue by creating a single, continuous index for the Southeast Florida management unit, using the REEF data to refine annual abundance estimates and bridge temporal gaps in the RVC survey series.

2. **Improves Precision:** The integration of REEF data enhances the spatiotemporal coverage of the index, which can reduce uncertainty and provide a more complete picture of population trends
3. **Preserves Key Data Linkages:** The model augments, rather than replaces, the RVC index, preserving the vital connection to the RVC's scientifically collected length-frequency data needed for the stock assessment model.
4. **Ecologically Meaningful:** Because the RVC data are density-based (counts per unit area), the final augmented index can be interpreted as an indicator of population density, providing a more tangible metric for management.

Beyond its current role as an auxiliary data source, the REEF survey holds potential as a future stand-alone index of hogfish abundance, particularly as temporal coverage continues to expand. The joint model simultaneously estimates an independent REEF state trajectory that is standardized for a comprehensive suite of observer and environmental covariates, yielding a continuous, calibrated index of relative abundance extending back to 1993 (Florida Keys) or 1998 (Southeast Florida). The integrated model suggested a fairly high correlation between the two indices in Southeast Florida (median $\rho = 0.476$; maximum a posteriori $\rho = 0.832$), and a lower but positive correlation among indices in the Florida Keys (median $\rho = 0.249$; maximum a posteriori $\rho = 0.270$). As the REEF time series lengthens and diver effort accumulates at well-sampled sites, the precision of annual REEF state estimates is expected to improve further, and the empirical estimate of ρ will become more stable. In regions such as the Dry Tortugas where RVC surveys are conducted but REEF coverage is limited, increased REEF participation could similarly enable the kind of data integration demonstrated here. Looking ahead, a sufficiently long and well-sampled REEF time series could support its use as a primary fishery-independent index in future stock assessments. The current analysis thus serves both an immediate need (augmenting the RVC index for SEDAR 94) and a longer-term goal of developing citizen science monitoring capacity that complements, and may eventually supplement, traditional scientific surveys.

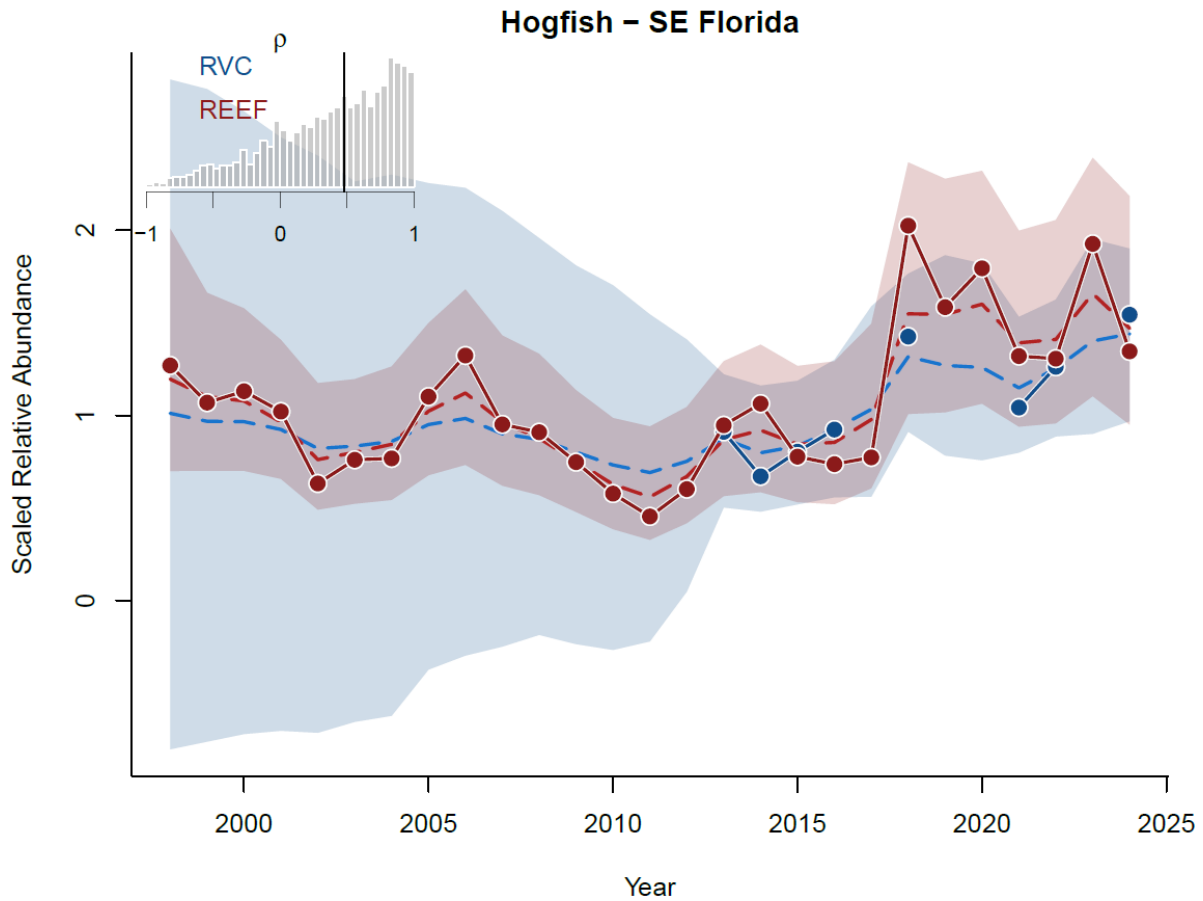


Figure 1. Indices of relative abundance (mean encounter rates - scaled to the median estimate throughout the time period) for RVC (blue) and REEF surveys for Hogfish in Southeast Florida. Solid points represent the observed annual estimate in each index, while the dashed line indicates the latent state abundance without additional year-to-year observation error in mean estimates. Shades represent the 90% equal-tail credible intervals for the state abundance of each index. The inset histogram shows the posterior distribution of the correlation parameter, ρ , indicating the temporal synchrony (ie. correlation in year-to-year abundance changes) between the indices.

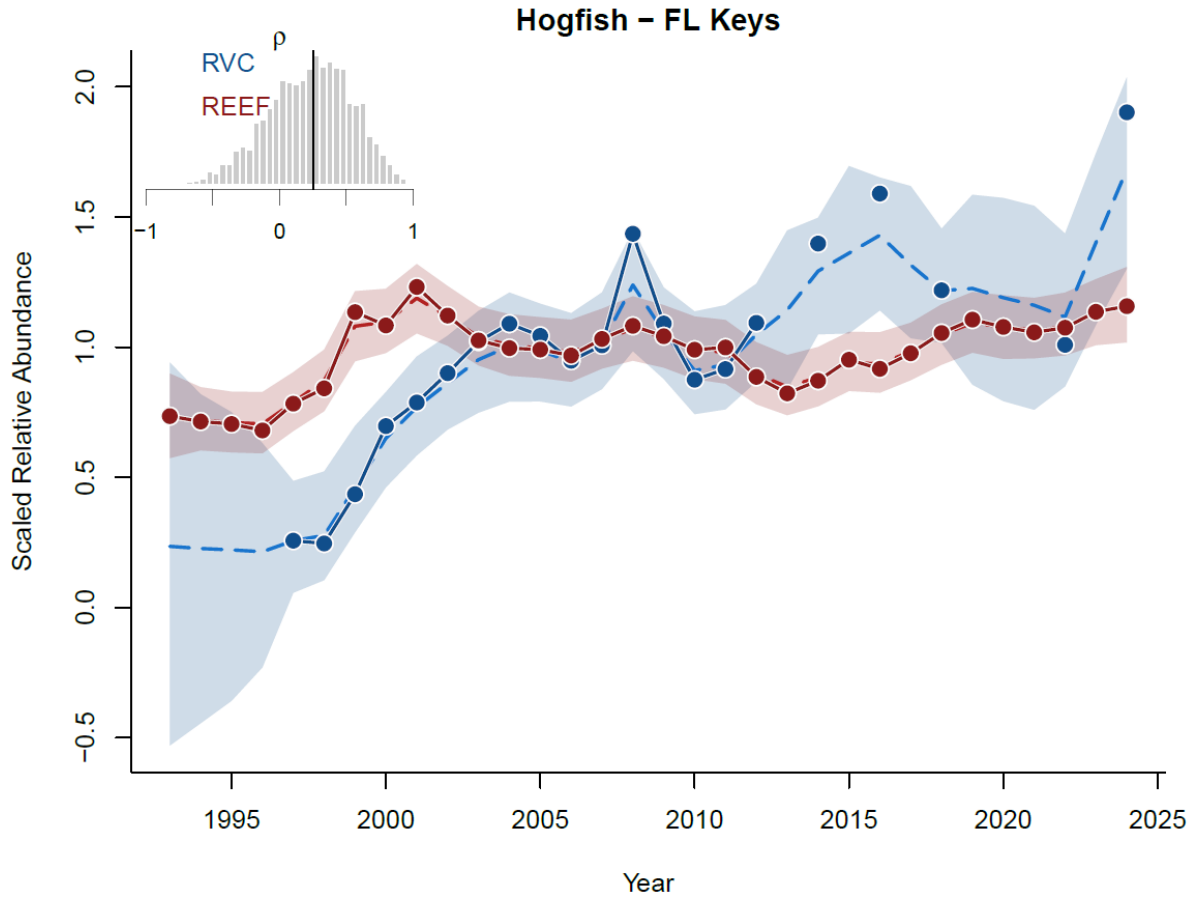


Figure 2. Indices of relative abundance (mean encounter rates - scaled to the median throughout the time period) for RVC (blue) and REEF surveys for Hogfish in the Florida Keys. Solid points represent the observed annual estimate in each index, while the dashed line indicates the latent state abundance without additional year-to-year observation error in mean estimates. Shades represent the 90% equal-tail credible intervals for the state abundance of each index. The inset histogram shows the posterior distribution of the correlation parameter, ρ , indicating the temporal synchrony (ie. correlation in year-to-year abundance changes) between the indices.

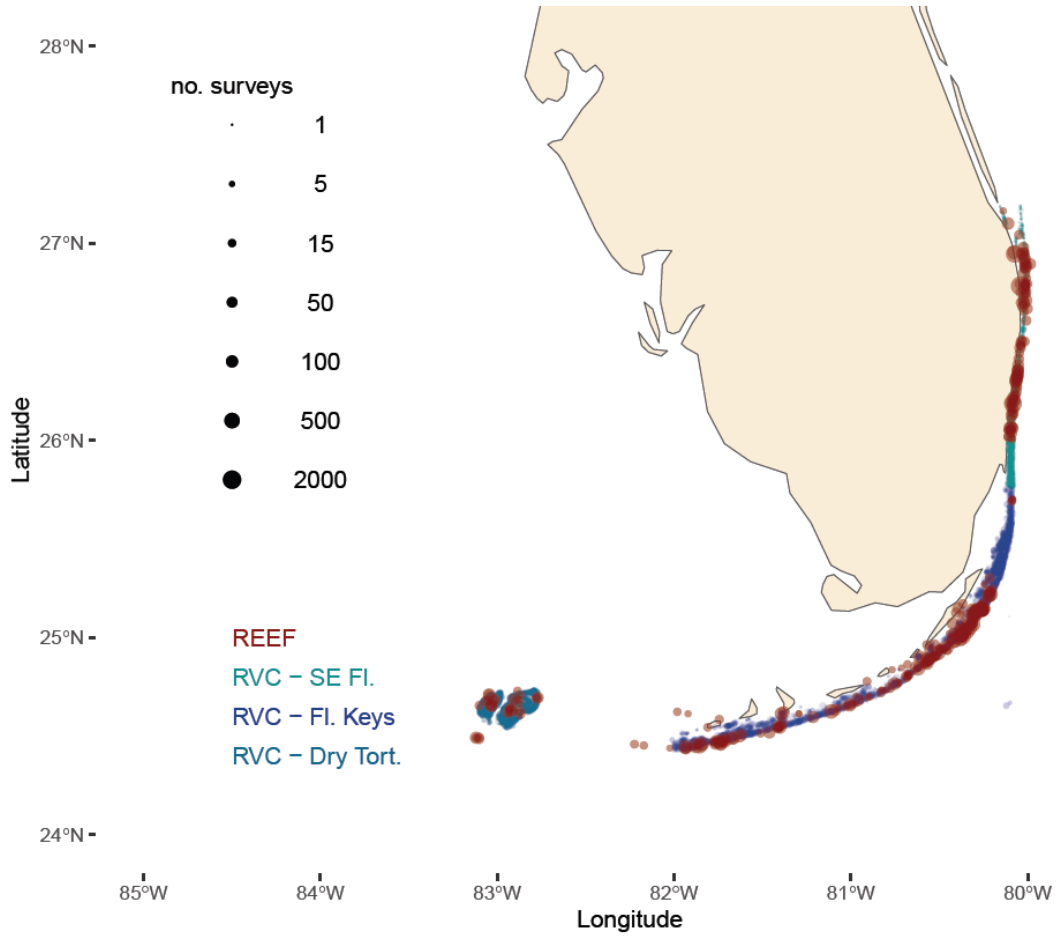


Figure 3. Sampling overlap of REEF dive sites (red) and RVC primary sample units (blue shades) surveyed in each regional survey: Southeast Florida, Florida Keys, and Dry Tortugas. Point sizes are scaled by the number of surveys throughout the time period.

Table 1. Southeast Florida RVC index including independently estimated (indicated by '(i)') annual mean and standard error (SE) encounter rates for Hogfish, as well as estimated RVC state index (mean and standard error) augmented (indicated by '(a)') by its correlation with the REEF Hogfish abundance index.

| Year | Mean (i) | SE (i) | Mean (a) | SE (a) |
|------|----------|----------|----------|----------|
| 1998 | NA | NA | 0.139219 | 0.173073 |
| 1999 | NA | NA | 0.131781 | 0.166192 |
| 2000 | NA | NA | 0.130554 | 0.157729 |
| 2001 | NA | NA | 0.125361 | 0.153114 |
| 2002 | NA | NA | 0.113088 | 0.152272 |
| 2003 | NA | NA | 0.112866 | 0.145752 |
| 2004 | NA | NA | 0.116512 | 0.139531 |
| 2005 | NA | NA | 0.128764 | 0.127902 |
| 2006 | NA | NA | 0.135124 | 0.118798 |
| 2007 | NA | NA | 0.126414 | 0.111199 |
| 2008 | NA | NA | 0.120061 | 0.100263 |
| 2009 | NA | NA | 0.110651 | 0.093857 |
| 2010 | NA | NA | 0.10069 | 0.088313 |
| 2011 | NA | NA | 0.094916 | 0.082136 |
| 2012 | NA | NA | 0.103761 | 0.061896 |
| 2013 | 0.12846 | 0.011925 | 0.120578 | 0.033186 |
| 2014 | 0.09048 | 0.010605 | 0.11223 | 0.030617 |
| 2015 | 0.11135 | 0.01441 | 0.116332 | 0.03115 |
| 2016 | 0.12829 | 0.015169 | 0.128163 | 0.032451 |
| 2017 | NA | NA | 0.145616 | 0.046347 |
| 2018 | 0.20493 | 0.022113 | 0.183676 | 0.037692 |
| 2019 | NA | NA | 0.179351 | 0.05004 |
| 2020 | NA | NA | 0.176303 | 0.049626 |
| 2021 | 0.13904 | 0.019176 | 0.160212 | 0.034616 |

| | | | | |
|------|---------|----------|----------|----------|
| 2022 | 0.17641 | 0.021794 | 0.174266 | 0.033344 |
| 2023 | NA | NA | 0.195366 | 0.046777 |
| 2024 | 0.22299 | 0.025871 | 0.200346 | 0.040843 |

Table 2. Florida Keys RVC index including independently estimated (indicated by '(i)') annual mean and standard error (SE) encounter rates for Hogfish, as well as estimated RVC state index (mean and standard error) augmented (indicated by '(a)') by its correlation with the REEF Hogfish abundance index.

| Year | Mean (i) | SE (i) | Mean (a) | SE (a) |
|------|----------|---------|----------|----------|
| 1993 | NA | NA | 0.209141 | 0.416969 |
| 1994 | NA | NA | 0.199182 | 0.367231 |
| 1995 | NA | NA | 0.193294 | 0.313665 |
| 1996 | NA | NA | 0.193692 | 0.246784 |
| 1997 | 0.22647 | 0.03399 | 0.239419 | 0.122164 |
| 1998 | 0.213 | 0.03933 | 0.267952 | 0.116956 |
| 1999 | 0.38003 | 0.04163 | 0.434991 | 0.113622 |
| 2000 | 0.62213 | 0.05214 | 0.595497 | 0.10377 |
| 2001 | 0.69294 | 0.04809 | 0.711578 | 0.105775 |
| 2002 | 0.79409 | 0.05928 | 0.791998 | 0.100095 |
| 2003 | 0.91142 | 0.07224 | 0.871618 | 0.108932 |
| 2004 | 0.98869 | 0.1028 | 0.918975 | 0.118213 |
| 2005 | 0.92271 | 0.0647 | 0.904371 | 0.104343 |
| 2006 | 0.8285 | 0.06052 | 0.868614 | 0.101439 |
| 2007 | 0.87933 | 0.05476 | 0.93134 | 0.104427 |
| 2008 | 1.31566 | 0.07455 | 1.128586 | 0.1337 |
| 2009 | 0.96238 | 0.0478 | 0.971706 | 0.099271 |
| 2010 | 0.76086 | 0.05192 | 0.846571 | 0.112617 |
| 2011 | 0.80217 | 0.04677 | 0.868743 | 0.110251 |
| 2012 | 0.96829 | 0.05607 | 0.964844 | 0.107511 |
| 2013 | NA | NA | 1.048214 | 0.173238 |
| 2014 | 1.24541 | 0.07087 | 1.180467 | 0.127877 |
| 2015 | NA | NA | 1.255539 | 0.182008 |
| 2016 | 1.43887 | 0.08346 | 1.303468 | 0.142641 |

| | | | | |
|------|---------|---------|----------|----------|
| 2017 | NA | NA | 1.212375 | 0.16466 |
| 2018 | 1.06493 | 0.06299 | 1.124226 | 0.123396 |
| 2019 | NA | NA | 1.124683 | 0.206659 |
| 2020 | NA | NA | 1.091999 | 0.223735 |
| 2021 | NA | NA | 1.061373 | 0.216009 |
| 2022 | 0.82336 | 0.09177 | 1.035133 | 0.165911 |
| 2023 | NA | NA | 1.290336 | 0.184802 |
| 2024 | 1.90704 | 0.17469 | 1.543822 | 0.205892 |