Estuarine, Coastal and Shelf Science 312 (2025) 109034

Contents lists available at ScienceDirect

Estuarine, Coastal and Shelf Science

journal homepage: www.elsevier.com/locate/ecss

Spatiotemporal shifts and influence of environmental parameters on estuarine-dependent fishes in Texas bays

Michelle Zapp Sluis [a,*], Masami Fujiwara [b], Fernando Martinez-Andrade [c], R.J. David Wells [a]

[a] Department of Marine Biology, College of Marine Sciences and Maritime Studies, Texas A&M University, Galveston, TX 77553, USA
[b] Department of Ecology and Conservation Biology, Texas A&M University, College Station, TX 77843, USA
[c] Coastal Fisheries Division, Texas Parks and Wildlife Department, 6300 Ocean Dr., NRC Bldg. Suite 2500, Corpus Christi, TX 78412, USA

ARTICLE INFO

Keywords:
Estuary
Fish
Gulf of Mexico
Random forest classification

ABSTRACT

Spatiotemporal shifts are occurring for estuarine-dependent species in Texas bays. To better understand what factors are causing these shifts, a random forest classification analysis was applied to the presence-absence data for seven estuarine-dependent species collected over 38 years. Five of the species showed an increase in presence and expanded their distributions northward, while the remaining two species declined in numbers and retracted their distributions to northern bays. The dominant factor influencing the presence of these species was year, followed by distance to major bay inlet and distance to major river mouth. While these factors may not be directly related to climate change, environmental fluctuations can impact year class success and alter the parameters of inlets and river flow. Studies examining multiple environmental and spatial conditions are needed to better understand the complexity of the changes in species composition that are occurring.

1. Introduction

Species composition within the major bays along the Texas coast have gradually changed over the past 30 years, with climate change being linked to the expansion of tropical and sub-tropical species into the region (Fujiwara et al., 2019; Pawluk et al., 2022; Torres Ceron et al., 2023). Conditions have become more favorable to warm-water species that are more prevalent in the estuaries during the summer months, with a decline in cold-water species that are more prevalent in the spring (Fujiwara et al., 2022). This decline was noticeable for species already within their southern limits and maximum temperature tolerance, resulting in diminished ranges over the years (Fujiwara et al., 2019). Additional trait analysis found that not only was there an increase in warm-water species, but that these species were also predominately long-lived, large, late-maturing predators (Pawluk et al., 2022).

Studies examining the variation of species composition among Texas bays determined that temperature was not the prominent, or even significant, environmental factor, but instead discerned salinity was the most influential (Fujiwara et al., 2019, 2022; Torres Ceron et al., 2023). Freshwater input into Texas bays decreases with decreasing latitude, creating a unique salinity gradient along the coast from the hyposaline Sabine Lake in the north to the hypersaline Laguna Madre in the south. While the northward expansion of species may have been limited to those that can tolerate a wide range of salinities, climate related changes may alter salinity gradients within the bays due to reduced precipitation, leading to a decline in river discharge. Therefore, as salinities increase in the northern bays, species can expand their range into systems they previously could not tolerate.

Torres Ceron et al. (2023) identified the top 20 species that are significantly contributing to the changes in species composition in at least four or more bays along the Texas coast. In this study, the top three species (gafftopsail catfish [*Bagre marinus*], ladyfish [*Elops saurus*] and gray snapper [*Lutjanus griseus*]), along with four additional commercially and recreationally important species (common snook [*Centropomus undecimalis*], Atlantic tarpon [*Megalops atlanticus*], Gulf flounder [*Paralichthys albiguttata*] and southern flounder [*Paralichthys lethostigma*]), were further analyzed to determine additional factors that may contribute to their changes in abundance and, if applicable, northern expansion. In addition to the usual environmental parameters, distance to major bay inlets and major river mouths was included in this study as these factors were not previously examined in other studies and, due to their potential to affect the salinity gradient of the bays, may be more of an influential factor than salinity alone.

* Corresponding author.
E-mail address: zappsluis@tamu.edu (M. Zapp Sluis).

https://doi.org/10.1016/j.ecss.2024.109034
Received 20 February 2024; Received in revised form 30 October 2024; Accepted 13 November 2024
Available online 14 November 2024

2. Methods

Gillnet sampling data utilized in this study were collected by the Texas Parks and Wildlife Department (TPWD) as part of the Marine Resource Monitoring program (Martinez-Andrade, 2018). Collections were analyzed from January 1982 through December 2019 across the eight major bays along the Texas coast: Sabine Lake, Galveston Bay, Matagorda Bay, San Antonio Bay, Aransas Bay, Corpus Christi Bay, Upper Laguna Madre, and Lower Laguna Madre (except for Sabine Lake, where sampling began in January 1986). Gillnet sampling occurred semiannually during spring (April–June) and fall (September–November), with 45 locations sampled in each bay each season (Fig. 1). During each 10-week sampling period, at least 3 gillnet deployments were completed each week to ensure collections were obtained over the entire period.

Gillnets (182.9 m × 1.2 m) comprised of four sections with varying mesh sizes: 76 mm, 102 mm, 127 mm, and 152 mm. The net was set perpendicular to the shore, with the smallest mesh size positioned closest to the shore and the bottom part of the net secured to the seabed. Deployment of the net occurred within an hour before sunset, and retrieval took place within 4 h after sunrise on the following day. The sampling locations were chosen with a stratified random sampling approach to ensure that the entire shoreline of each bay is represented during a given season. All fish collected were identified to the most specific taxonomic level, usually species level, and counted. Water temperature, salinity, and dissolved oxygen levels were measured using a YSI© meter, while turbidity was gauged using a HACH© turbidimeter. These environmental parameters were measured at the offshore end of the net during each sampling event. More detailed information about the sampling procedure can be found in Martinez-Andrade (2018).

In addition to the four environmental variables collected by TPWD, monthly mean sea level, the North Atlantic Oscillation (NAO) Index (a climate index calculated based on high and low pressure patterns in the North Atlantic), the distance of the sampling locations to the nearest major bay inlet and distance of the sampling locations to the mouth of the nearest major river were obtained as variables to further explain presence-absence observations of selected fishes. The monthly mean sea level was obtained from NOAA Center for Operational Oceanographic Products and Services (https://tidesandcurrents.noaa.gov/). The North Atlantic Oscillation Index was obtained from NOAA's Climate Prediction Center (https://www.cpc.ncep.noaa.gov/). Distances from sampling locations to major bay inlets and river mouths were calculated in ArcGIS (version 10.7.1).

The feature variables had smooth yearly patterns as well as variations among major bays and seasons. These large-scale spatiotemporal patterns can cause spurious associations between the response variables and feature variables (e.g. Pyper and Peterman, 1998). Therefore, a generalized additive model with year as a smooth explanatory variable and major bays and seasons as factors was fitted to each feature variable. Then, the deviations from the original data and the fitted model (residuals) were calculated, with the residuals being used for the subsequent analysis.

To identify the conditions affecting the presence-absence (incidence) of selected species with environmental and habitat variables, with bays, season, and year as explanatory (feature) variables, a Random Forest classification analysis (Hastie et al., 2009) was applied with the incidence data of individual species as the response variable. The Random Forest classification reduces overfitting with a cross-validation method. It is also robust against co-linear feature variables because it samples a subset of feature variables when building a classification tree. Finally, as

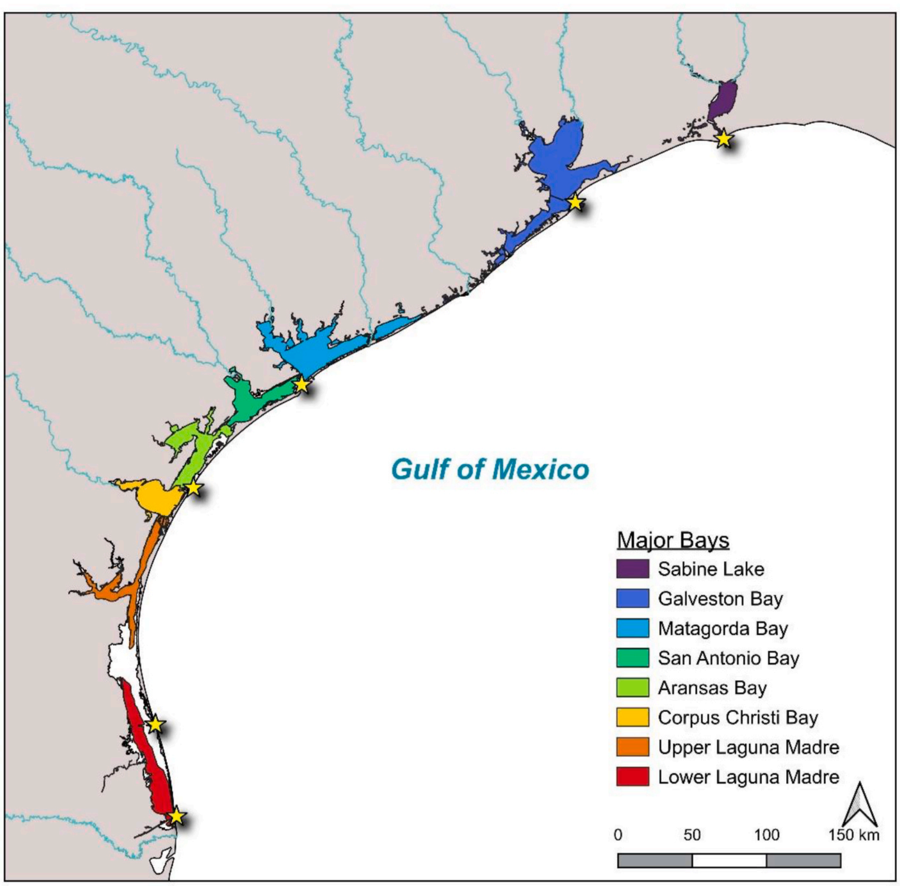

Fig. 1. Locations of major bays along the Texas coast. Major rivers are denoted by blue lines and major bay inlets are marked by yellow stars. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

this method is a non-parametric method, it is not affected by observations that substantially deviate from other observations (often called outliers) as much as parametric methods.

Each data set for each species consists of 27,126 observations (or gillnet deployments). The data were randomly divided into two subsets, with 70% for building the best-fit model (the training data) and 30% for validating the model (the testing data). The training data for gray snapper, tarpon, common snook, and Gulf flounder are unbalanced (more absences than presences), resulting in a class imbalance problem (Megahed et al., 2021). To overcome this issue, the data were randomly down-sampled such that presences were at least 25% of the absences. Then, the Synthetic Minority Over-sampling Technique (SMOTE; Chawla et al., 2002) was used to oversample the observations of presences. In our analysis, the SMOTE algorithm selected five nearest neighbors from the presence data in the feature space. It then created a new sample at a random location along the lines connecting the original sample and its chosen neighbors. This was repeated until the data were balanced. Finally, the balanced data were used for hyper-parameter selection and model fitting.

The hyper-parameter (parameters used in Random Forest Classification Algorithm) selection was done with the out-of-bag sampling prediction error rate (i.e. errors associated with unselected data in a bootstrap sampling) from the following sets: (1) number of variables randomly sampled as candidates at each split (4,6,8,10) and (2) minimum size of terminal node (4,6,8,10). Other hyper-parameter values were kept as the default values in the statistical package used (i.e. the common values). The Random Forest classification model with the best hyper-parameters based on the out-of-bag sampling prediction error rate was selected. The selected classification model was applied to the test data set to assess its performance. Performance was measured by accuracy, sensitivity, specificity, and the Area under the ROC (Receiver Operating Characteristic) Curve (AUC). Finally, the importance of feature variables was determined using the mean decrease in accuracy. In this approach, the variable that decreased the accuracy more by removing it was considered the more important variable. The Random Forest classification analysis was repeated for all species separately. The analysis was performed using the "randomForest" package (Liaw and Wiener, 2002) under R Version 4.3.1 (R Core Team, 2021). The diagnostic of the Random Forest classification was performed using the "confusionMatrix" function in the "caret" package (Kuhn, 2008) and the "auc" function in the Metrics package (Hamner and Frasco, 2018).

3. Results

The variables of importance for the top three species (gafftopsail catfish, ladyfish and gray snapper) that significantly contribute to the changes in species composition along the Texas coast were year, distance to bay inlets and distance to river mouths based on mean decrease accuracy values (Fig. 2). The random forest classification model for gafftopsail catfish had an accuracy of 81% and an AUC of 0.80 (Table 1). Gafftopsail catfish were present in all bay systems, displaying an increase in catches over time, with predominance in the upper to middle bay systems (Galveston Bay through Corpus Christi Bay) beginning around 1998 (Fig. 3). The random forest classification model for ladyfish

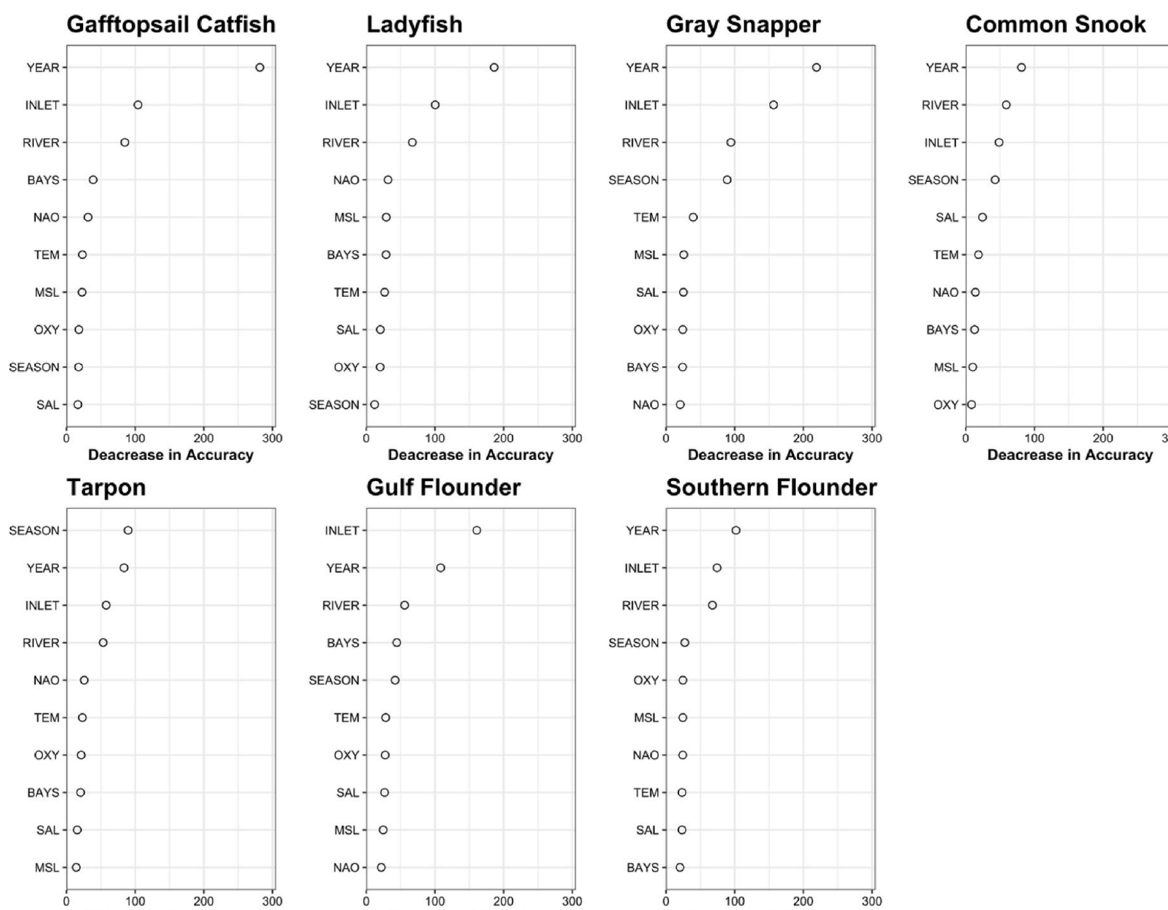

Fig. 2. Random forest variable importance based on mean decrease accuracy scores for seven species collected in Texas bays through TPWD gillnet surveys from 1982 to 2019. Variables are listed in the order of importance and include year, distance to major inlet (INLET), distance to major river mouth (RIVER), bays, and the residuals of feature variables (North Atlantic Oscillation [NAO] Index, mean sea level [MSL], temperature [TEM], salinity [SAL], dissolved oxygen [OXY] and season).

Table 1
Random forest classification model performance measured by accuracy, sensitivity, specificity and area under the ROC curve (AUC) to identify the conditions affecting incidence of seven species collected from Texas bays from 1982 to 2019 through TPWD gillnet surveys.

Model Validation	Gafftopsail Catfish	Ladyfish	Gray Snapper	Common Snook	Tarpon	Gulf Flounder	Southern Flounder
Accuracy	0.81	0.77	0.89	0.85	0.84	0.88	0.70
Sensitivity	0.83	0.83	0.90	0.86	0.84	0.89	0.91
Specificity	0.77	0.68	0.68	0.65	0.49	0.57	0.29
AUC	0.80	0.76	0.79	0.76	0.66	0.73	0.60

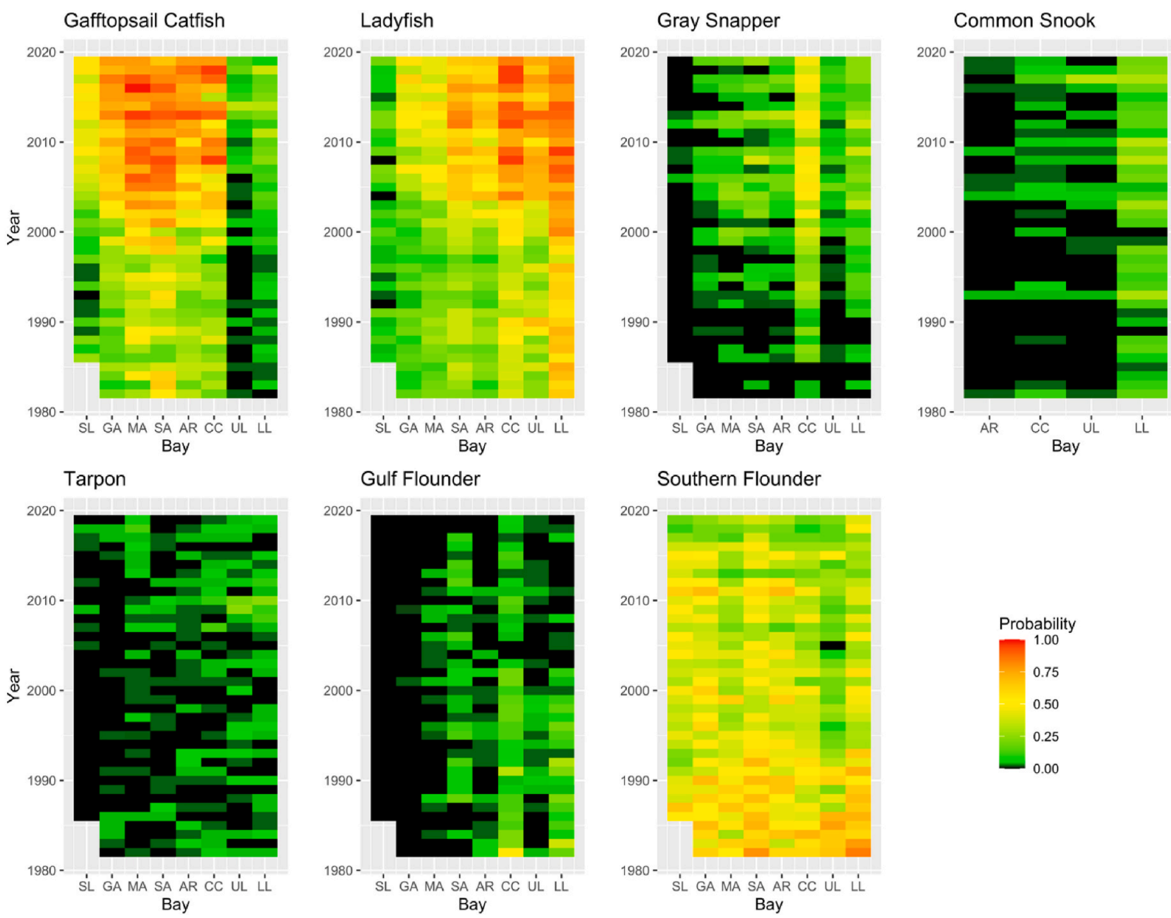

Fig. 3. Heat maps displaying the presence of seven species collected in Texas bays (SL = Sabine Lake, GA = Galveston Bay, MA = Matagorda Bay, SA = San Antonio Bay, AR = Aransas Bay, CC = Corpus Christi Bay, UL = Upper Laguna Madre, LL = Lower Laguna Madre) from 1982 to 2019 through TPWD gillnet surveys.

had an accuracy of 77% and an AUC of 0.76 (Table 1). Ladyfish were present in all bays, with higher concentrations in the southern bay systems (San Antonio Bay through the Lower Laguna Madre) and increasing presence beginning around 2004 (Fig. 3). The gray snapper random forest classification model had the highest accuracy at 89% and an AUC of 0.79 (Table 1). In addition to the three top variables of importance listed above, season was also an important variable based on mean decrease accuracies over 80 (Fig. 2). Greatest catches of gray snapper occurred in Corpus Christi Bay, with presence increasing in all bay systems over time (Fig. 3).

The variables of importance varied among the four commercially and recreationally important species (common snook, tarpon, Gulf flounder and southern flounder). For common snook, the random forest classification model had an accuracy of 85% and an AUC of 0.76 (Table 1). Year, distance to river mouths and distance to bay inlets were important variables based on mean decrease accuracies (Fig. 2). Common snook were only collected in the four lowermost bay systems (Aransas Bay, Corpus Christi Bay, Upper and Lower Laguna Madre) along the Texas coast, but increasing presence was observed throughout the years (Fig. 3). Tarpon random forest classification models had an accuracy of 84% and an AUC of 0.66 (Table 1). Variables of importance for tarpon slightly differed with season prevailing, followed by year, distance to bay inlets and distance to river mouths (Fig. 2). Tarpon were collected in all bay systems, with the first presence in Sabine Lake detected in 2005, and also showed increase in catches over time (Fig. 3).

The two Paralichthyids examined in this study had differing results for their random forest classification models. The random forest classification model for Gulf flounder had an accuracy of 88% and an AUC of 0.73 (Table 1). Distance to bay inlet, year and distance to river were important variables (Fig. 2). The presence of Gulf flounder occurs in the mid to lower bay systems of the Texas coast. Dominant catches occurred in Corpus Christi Bay and the Lower Laguna Madre, but declines began to occur in the Lower Laguna Madre after the year 2000 (Fig. 3). The random forest classification model for southern flounder had an accuracy of 70% and an AUC of 0.60 (Table 1). Year, distance to bay inlets and distance to river mouths were important variables affecting the incidence of southern flounder based on mean decrease accuracy values (Fig. 2). The presence of southern flounder was greatest from 1982 to

1990, before beginning to decrease, with lower capture rates occurring in the lower bay systems (Corpus Christi, Upper Laguna Madre and Lower Laguna Madre; Fig. 3).

4. Discussion

Year was the predominant variable affecting presence/absence of the seven fish species examined within this study, excluding tarpon and Gulf flounder in which year was the second dominant variable. This suggests that presence/absence of these species are gradually increasing or decreasing over time. Previous studies linked changes in species composition in Texas bays to an increase in warm-water species and a decrease in cold-water species (Fujiwara et al., 2019, 2022). However, temperature was not a principal factor in this study or in previous studies (Fujiwara et al., 2019; Torres Ceron et al., 2023). Recruitment variability in estuaries can be affected by several factors, including environmental (temperature, salinity, and NAO), hydrodynamic features (tidal stream transport and river flow), and density-dependent processes (prey availability; Martinho et al., 2012). These changes in species composition may be due to a combination of favorable environmental and hydrodynamic conditions for warm-water species, as well as changes in bycatch and fishing pressures over the years that can alter predator/prey interactions. Therefore, year may be representing the significance of year class success as conditions improve or decline for the selected species.

Previous research examining changes in species composition along the Texas coast determined that salinity was the most influential factor (Fujiwara et al., 2019, 2022; Torres Ceron et al., 2023). However, in this study the addition of distance to major bay inlet and distance to major river mouth resulted in these variables being the next most influential factors, respectively after year, in determining incidence for the included species. Both of these variables can affect the salinity gradients of the estuaries, with species typically remaining near bay inlets to be adjacent to areas of increased salinity (Sackett et al., 2008) and known to move away from river mouths to be closer to inlets during periods of high freshwater inflow and vice versa. Assessment of fish assemblages in estuaries along the Florida coast also determined that distance to inlets was a major feature that influenced species compositions, due to several estuarine species utilizing oceanic reproduction and relying on tidal currents for recruitment success (Gilmore, 1988; Kupschus and Tremain, 2001), resulting in greater juvenile abundance near bay inlets.

Common snook was the only species examined where distance to river mouth was more influential than distance to bay inlet. In a Florida lagoon, common snook were observed using inlets during the spawning season, then moving away and occupying the river during non-spawning months. During times of high freshwater inflow, spawning could be disrupted if the lower estuary and inlet reached salinities lower than 24 psu (Stevens et al., 2023). Common snook in Texas only occupy the four southernmost bays, where there are smaller freshwater tributaries for them to utilize, but these bays are not influenced by larger river systems that can create an influx of high freshwater inflow to disrupt spawning. Thus, as drought and reduced river flows occur, common snook may be able to expand their distribution into northern bays where lower salinity thresholds may no longer be an issue. Likewise, decrease freshwater inflow can cause population declines and for distributions to retreat from southern bays for species where lower salinities are essential. Froeschke et al. (2013) observed a higher probability of southern flounder occurring near tidal inlets with capture rates increasing with decreasing salinity, suggesting that decreased freshwater inflow can greatly impact their distribution. Both Paralichthyids examined here displayed population declines overall, with greater declines in the southernmost bays and shift to a higher presence in the northern bays of their limit.

Season was another important variable for species within the study that utilize estuaries as nursery grounds. Moreover, season was the primary influential variable for tarpon. Peak residency time of juvenile tarpon within estuaries has been observed to occur from July to December (Stein et al., 2016). Stephens et al. (2024) confirmed higher abundance in the fall gillnet surveys compared to the spring, concluding spring declines may be a result of thermal stress and the potential for tarpon to move to deeper portions of the bay, to seek refuge, that were not sampled during the surveys.

Spatiotemporal shifts in species compositions are occurring among bays along the Texas coast, with several factors affecting these distributions. The incidence of fish species examined in this study were predominately affected by year, distance to major bay inlet and distance to major river mouth. While these variables are not directly linked to climate change or temperature, the conditions around these variables can be altered by climate shifts (i.e. decrease river flow, salinity increase at inlets, etc.). Therefore, multiple environmental, hydrodynamic, spatial and density-dependent conditions should be analyzed collectively to obtain a better understanding of the features impacting species compositions.

CRediT authorship contribution statement

Michelle Zapp Sluis: Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Masami Fujiwara:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization, Investigation. **Fernando Martinez-Andrade:** Writing – review & editing, Data curation, Resources. **R.J. David Wells:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded by TPWD's Coastal Fisheries State Wildlife Grant (CA-0000985). We would like to thank the numerous staff members of TPWD for collecting data over the years.

Data availability

Data will be made available on request.

References

Chawla, N., Bowyer, K., Hall, L., Kegelmeyer, W., 2002. SMOTE: Synthetic minority oversampling technique. J. Artif. Intell. Res. 16, 321–357.

Froeschke, B.F., Tissot, P., Stunz, G.W., Froeschke, J.T., 2013. Spatiotemporal predictive models for juvenile southern flounder in Texas estuaries. N. Am. J. Fish. Manag. 33, 817–828.

Fujiwara, M., Martinez-Andrade, F., Wells, R.J.D., Fisher, M., Pawluk, M., Livernois, M. C., 2019. Climate -related factors cause changes in the diversity of fish and invertebrates in subtropical coast of the Gulf of Mexico. Commun. Biol. 2, 403. https://doi.org/10.1038/s42003-019-0650-9.

Fujiwara, M., Simpson, A., Torres Ceron, M., Martinez-Andrade, F., 2022. Life-history traits and temporal patterns in the incidence of coastal fishes experiencing tropicalization. Ecosphere 13, e4188. https://doi.org/10.1002/ecs2.4188.

Gilmore Jr., R.G., 1988. Subtropical seagrass fish communities: population dynamics, species guilds and microhabitat associations in the Indian River Lagoon, Florida. PhD Dissertation, Florida Institute of Technology, Melbourne, FL, U.S.A.

Hamner, B., Frasco, M., 2018. Metrics: evaluation Metrics for machine learning. R package version 0.1.4.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Random forests. In: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer New York, New York, pp. 587–604.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Software 28, 1–26. https://doi.org/10.18637/jss.v028.i05.

Kupschus, S., Tremain, D., 2001. Associations between fish assemblages and environmental factors in nearshore habitats of a subtropical estuary. J. Fish. Biol. 58, 1383–1403.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2, 18–22.
Martinez-Andrade, F., 2018. Marine Resource Monitoring Operations Manual. Texas Parks & Wildlife Department. Coastal Fisheries Division, Austin, Texas, p. 137.
Martinho, F., Cabral, H.N., Azeiteiro, U.M., Pardal, M.A., 2012. Estuarine nurseries for marine fish: connecting recruitment variability with sustainable fisheries management. Management of Environmental Quality 23, 414–433.
Megahed, F.M., Chen, Y.J., Megahed, A., Ong, Y., Altman, N., Krzywinski, M., 2021. The class imbalance problem. Natural Methods 18, 1270–1272.
Pawluk, M., Fujiwara, M., Martinez-Andrade, F., 2022. Climate change linked to functional homogenization of a subtropical estuarine system. Ecol. Evol. 12, e8783. https://doi.org/10.1002/ece3.8783.
Pyper, B.J., Peterman, R.M., 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. Can. J. Fish. Aquat. Sci. 55, 2710-2710.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Sackett, D.K., Able, K.W., Grothues, T.M., 2008. Habitat dynamics of summer flounder *Paralichthys dentatus* within a shallow USA estuary, based on multiple approaches using acoustic telemetry. Mar. Ecol. Prog. Ser. 364, 199–212.
Stein, W., Shenker, J., O'Connell, M.T., 2016. A Contribution to the life history of tarpon (*Megalops atlanticus*) in the northern Gulf of Mexico. SE. Nat. 15, 496–512.
Stephens, S.A., Dance, M.A., Zapp Sluis, M., Kline, R.J., Streich, M., Stunz, G.W., Adams, A.J., Wells, R.J.D., Rooker, J.R., 2024. Habitat use and movement of Atlantic tarpon (*Megalops atlanticus*) in the northern Gulf of Mexico. PLoS ONE (in press).
Stevens, P., Webb, S., Ault, E., Ellis, R.D., 2023. Appendix 8C-4: acoustic telemetry to determine response of large-bodied fishes to freshwater inflows in the St. Lucie Estuary. South Florida Water Management District. Report No. F4492-19-22-F).
Torres Ceron, M., Fujiwara, M., Martinez-Andrade, F., 2023. Changes in species composition of fish in the bays of the northwestern Gulf of Mexico. Front. Mar. Sci. 10, 1274771. https://doi.org/10.3389/fmars.2023.1274771.