Classification of Historical Trap Effort in St. Thomas and St. John Logbooks

Caribbean Fisheries Branch
Sustainable Fisheries Division, Southeast Fisheries Science Center

Abstract

Gear-specific fishing effort is an important component in catch-rate analyses that inform stock assessments. Analysis of self-reported commercial logbook data from St. Thomas and St. John by gear group revealed distinct species proportions between landings from lobster and fish trap trips. These distinct catch compositions were used to train a classification tree to categorize historical trips where trap type was unspecified. This analysis successfully extends the time series of trap-specific, fishery-dependent logbook data by a decade, directly supporting sensitivity analyses of the SEDAR 91 Caribbean Spiny Lobster stock assessment for St. Thomas and St. John.

Introduction

Since the mid-1970s, commercial fishing license holders in the U.S. Virgin Islands have been required to submit logbook reports for each fishing trip. Over time, modifications to the species and gear types listed on these forms have complicated the gear-specific analyses required for fisheries stock assessments.

Beginning in 2011, logbook forms explicitly differentiated trap types into lobster traps, fish traps, or "both." Although these gears can capture many of the same species, lobster trap trips yield significantly higher proportions of Caribbean Spiny Lobster (*Panulirus argus*) than trips utilizing fish traps (Figure 1). Because earlier forms did not specify trap type, we developed a classification tree model to predict historical trap types based on recent catch profiles. This approach allowed us to extend fishery-dependent indices of abundance by a decade to support subsequent sensitivity runs in the SEDAR 91 assessment.

Methods

Caribbean Commercial Logbook data for St. Thomas and St. John were prepared as follows:

1. Filtered the dataset to encompass the years 2000–2024.
2. Calculated the number of gear groups reported per trip, removing trips that recorded multiple gear groups.
3. Retained trips associated exclusively with the "Trap" gear group (reported as Traps, Fish Traps, Lobster Traps, or both; Figure 2).
4. Calculated the total number of unique species, total landed pounds, and the percentage of lobster (lobster pounds / total pounds) for each trip.
5. Developed a classification tree model using the `rpart` package in R to categorize trap types based on the percentage of lobster landings and the number of unique species reported.
6. Evaluated an expanded model including additional covariates (year, month, weekday, and total pounds landed) to test for significant improvements in predictive power.

Results

Figure 2 displays the annual number of trips by reported trap type. From 2011 to 2024, when specific trap types were explicitly available, noticeable differences were observed between the fisheries (Figure 1). As outlined in Table 1, fish trap trips and combined mixed-gear trips averaged a high diversity of catch, featuring a mean of 12.6 and 11.9 species per trip, respectively. Conversely, lobster trap trips averaged just 1.4 unique species per trip.

The results of the primary classification tree are presented in Figure 3. The inclusion of additional covariates in the expanded model did not alter the results, as no additional variables were identified as significant predictors.

The final classification criteria were defined as:

- **Fish Trap Trips:** $< 27\%$ of total landings consist of Caribbean Spiny Lobster.
- **Lobster Trap Trips:** $\geq 83\%$ of total landings consist of Caribbean Spiny Lobster and < 6 unique species landed.
- **Combined Fish & Lobster Trap Trips:** $\geq 27\%$ of total landings consist of Caribbean Spiny Lobster and ≥ 6 unique species landed.

The classification tree model demonstrated strong predictive performance, achieving an overall accuracy of 90.0% (95% CI: 89.6%–90.5%) and a Cohen's Kappa of 0.83. Overall accuracy measures the percentage of all predictions that the model got exactly right out of the total sample.

$$\text{Overall Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Trips}}$$

Cohen's Kappa (κ) evaluates if the model is better than a random classifier based on frequencies of the data. A value of 0 implies the model is no better than random guessing, while a value of 1 reflects perfect agreement.

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

Table 2 provides the predicted vs. observed classification matrix along with class-specific sensitivity and precision statistics. The measurement of sensitivity addresses the model's accuracy in identifying a given category (i.e., what percentage of each actual reference group did the model correctly classify?). Precision measures the proportion of positive predictions that were correctly assigned (i.e., what percentage of each predicted group belonged to the correct reference group?).

Classification reliability was highest for Lobster Traps (95.2% sensitivity, 98.1% precision) and Fish Traps (95.7% sensitivity, 88.7% precision). The combined 'Fish Trap + Lobster Trap' category exhibited a lower sensitivity (63.9%), reflecting the intermediate, overlapping catch profiles characteristic of mixed-gear trips.

Figure 4 illustrates the reconstruction of the historical time series (2000–2010) after applying these classification rules to trips with unknown trap types.

Conclusion

Accounting for trap type is critical due to differences in species composition and catch rates between trips that use fish and lobster traps. While further refinements to the model remain possible, this classification tree approach offers a robust, data-driven method to categorize historical effort. Ultimately, this allowed the SEFSC to extend the trap-specific time series by 10 years, maximizing the utility of historical logbook data for indices of abundance and associated sensitivity runs of the SEDAR 91 stock assessment.

Tables

Table 1: Summary statistics of St. Thomas and St. John logbook data by reported trap type (2011–2024).

Reported Gear Type	Total Trips (N)	Mean Species	Median Species
Fish Trap	8,719	12.6	12
Fish & Lobster Trap	2,976	11.9	12
Lobster Trap	5,518	1.4	1

Table 2: Predicted vs. observed classification matrix, predictive accuracy, and class-specific statistics for the classification tree model. The sensitivity values measure the model's accuracy in identifying a given reference category, meanwhile the precision values measure the proportion of positive predictions that were correctly assigned. For example, 95.7% of actual fish trap trips were accurately predicted as fish trap trips; however, only 88.7% of predicted fish trap trips were actually fish trap trips (i.e., 11.3% of predicted fish trap trips were actually either combined lobster and trap or lobster trap trips).

Predicted Gear Type	Reference: Fish Trap	Reference: Fish & Lobster Trap	Reference: Lobster	*Precision*
Predicted: Fish Trap	8,343	1,001	58	*88.7%*
Predicted: Fish & Lobster Trap	350	1,901	209	*77.3%*
Predicted: Lobster Trap	26	74	5,251	*98.1%*
Sensitivity	*95.7%*	*63.9%*	*95.2%*	

Figures

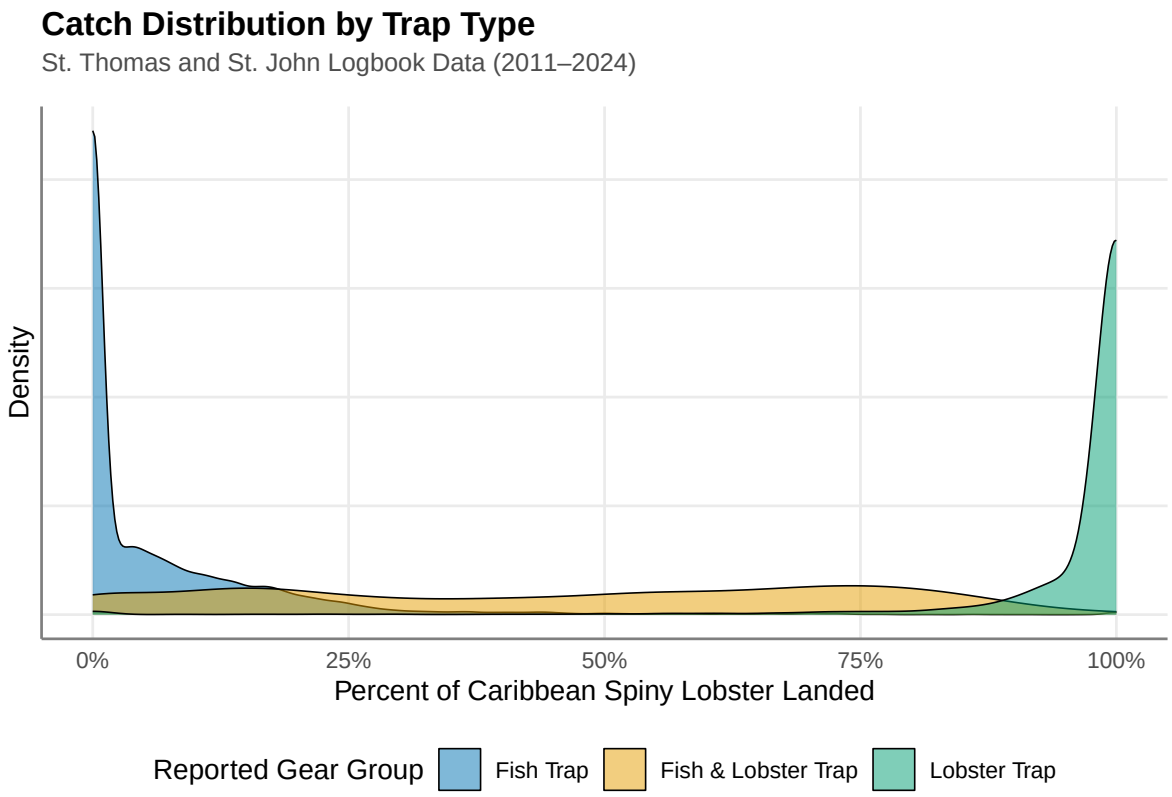

Figure 1: Catch distribution by trap type for trips from 2011–2024. Relative frequency curves are scaled to showcase distribution shapes.

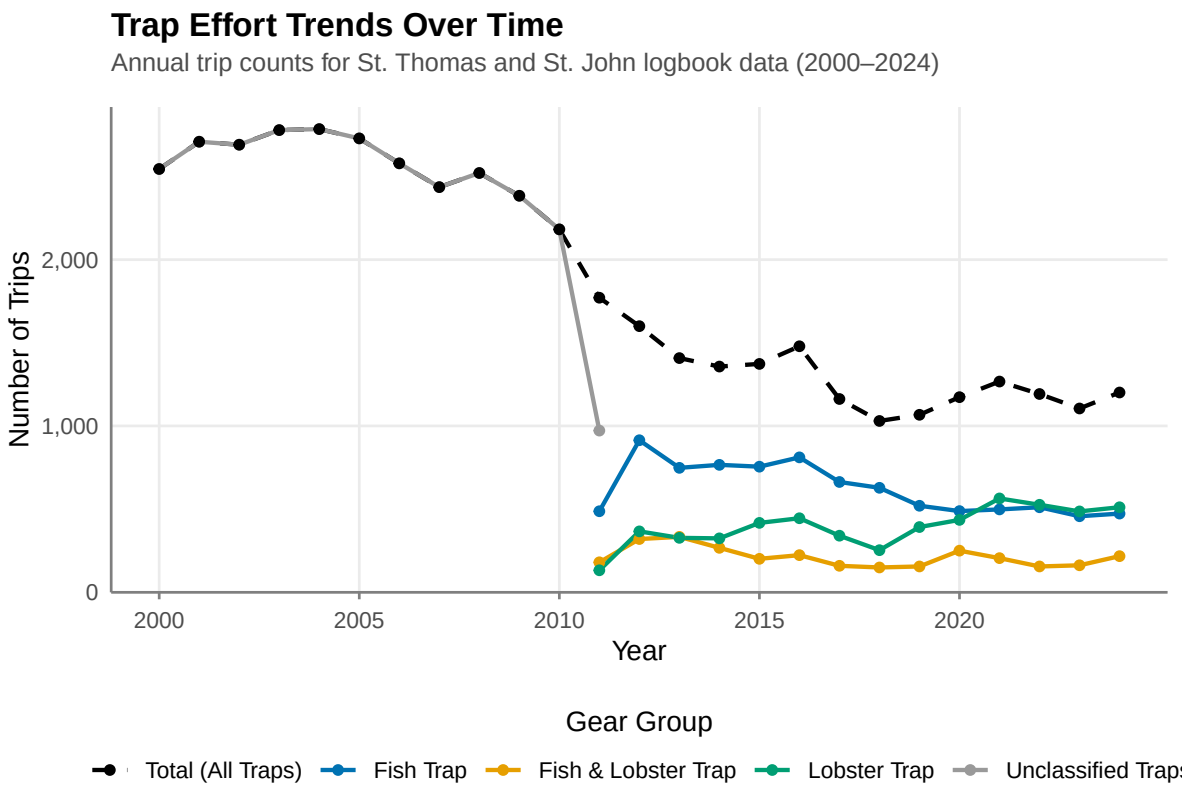

Figure 2: Annual count of logbook trips by reported trap type from 2000–2024 prior to historical back-casting.

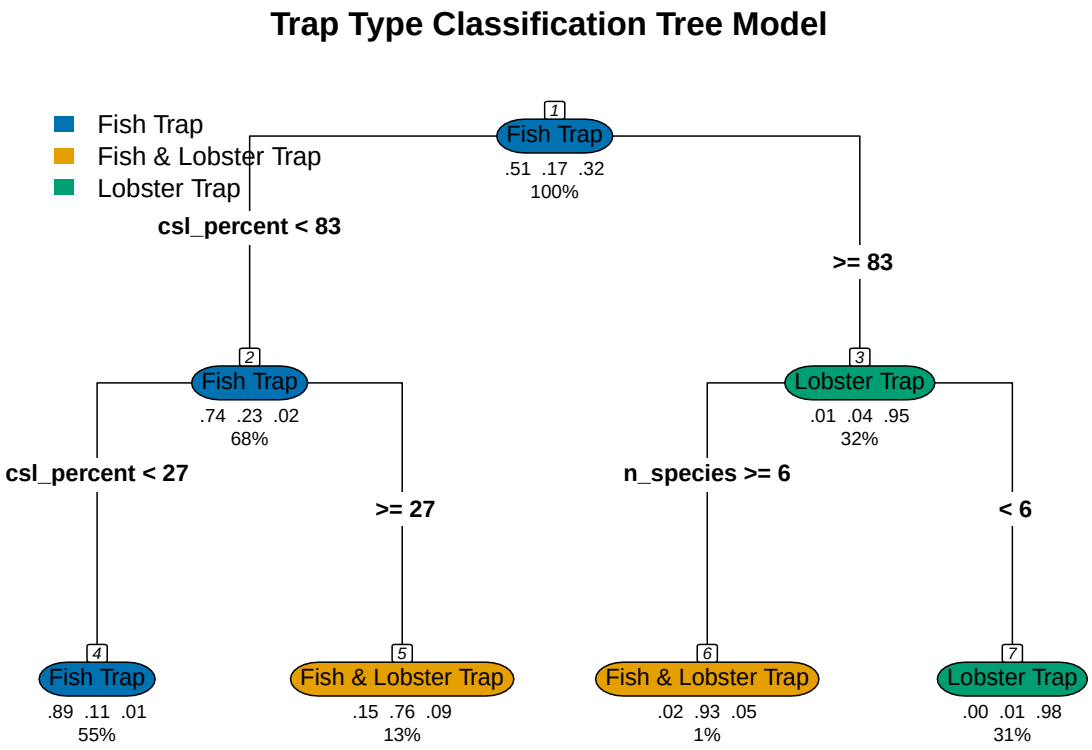

Figure 3: Classification tree structure generated by the `rpart` model mapping split criteria for unclassified historical traps.

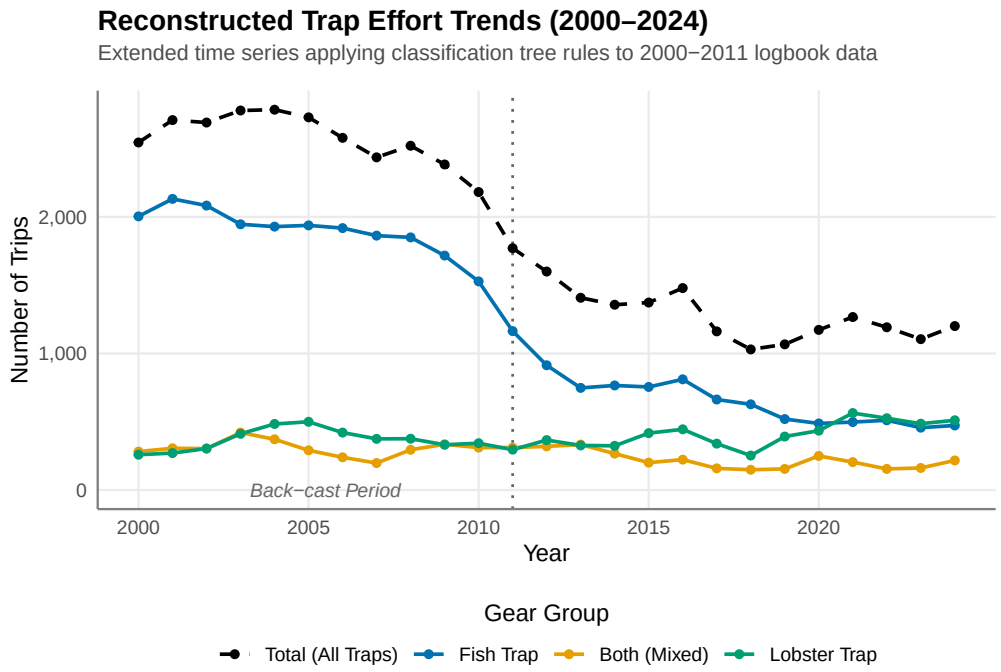

Figure 4: Reconstructed annual count of logbook trips by trap type (2000–2024) highlighting the 10-year model back-cast extension. "CSL" is Caribbean Spiny Lobster.