Validation and Optimization of Vessel Underwater Radiated Noise Prediction Using Measurements in United States Waters

Vanessa M. ZoBell[1], Dan F.B. Flynn[2], Kaitlin E. Frasier[1]

[1]Marine Physical Laboratory of the Scripps Institution of Oceanography
University of California, San Diego
[2] Energy Analysis and Sustainability,
U.S. DOT Volpe National Transportation Systems Center

Tug operating near the Channel Islands at dusk.

Corresponding Author: vmzobell@ucsd.edu
Project Period: July 1st, 2025 – October 31st, 2025

MPL Technical Memorandum #675

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (26-11-2025)	**2. REPORT TYPE** White Paper	**3. DATES COVERED (From - To)**
4. TITLE AND SUBTITLE Validation and Optimization of Vessel Underwater Radiated Noise Prediction Using Measurements in United States Waters		**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Vanessa M. ZoBell, Dan F.B. Flynn, Kaitlin E. Frasier		**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)		**8. PERFORMING ORGANIZATION REPORT** MPL TM 675
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation John A. Volpe National Transportation Systems Center 220 Binney Street Cambridge, MA 02142-1093		**10. SPONSOR/MONITOR'S ACRONYM(S)** USDOT Volpe Center **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION / AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This research aimed to improve the accuracy of predicting vessel underwater radiated noise (URN) in U.S. waters by validating and optimizing the current JOMOPANS-ECHO (J-E) source level model using a vast dataset of 25,129 usable measurements from Southern California and the Gulf of Mexico. By integrating these measurements into the USDOT Volpe Center MARINE-T decision support tool, researchers addressed substantial unpredicted variations in the original model, which was primarily trained on vessels in British Columbia. Using a derivative-free global search algorithm to iteratively tune five key model coefficients., the study achieved an average reduction in Root Mean Square Error (RMSE) of 4.5 dB across various vessel classes. While optimization significantly improved fit for large vessels like bulkers and tankers—particularly at low-frequency energy peaks—discrepancies remained for smaller or underrepresented categories like tugs, where operational modes (e.g., transiting versus active towing) drastically alter acoustic signatures. The findings underscore the necessity of expanding regional datasets and adopting multi-output modeling approaches to better capture the complex spectral characteristics of diverse global shipping fleets.

15. SUBJECT TERMS

Marine Transportation, Underwater Radiated Noise, MARINE-T, AIS, Optimization

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Dan Flynn
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unlimited	23	**19b. TELEPHONE NUMBER (include area code)** 617-494-3665

Standard Form 298 (Rev. 8-98)

Objective and Purpose

The project objective was to analyze vessel underwater radiated noise (URN) levels from existing databases based on recordings at three monitoring stations located in the Santa Barbara Channel and Gulf of Mexico, and compare those to the current JOMOPANS-ECHO (J-E) source level model, which is fit to a database of source levels from vessels operating in British Columbia, Canada (MacGillivray et al. 2021). The J-E model provides best fit predictions based on a large sample of empirical URN measurements collected at this single location over many years. The underlying dataset, as reflected in the reference paper for the J-E model and other published reports on the ECHO dataset, in some cases can still reflect substantial unpredicted variation among source levels with the same type, length and speed. The model fits less well for vessel types, lengths and speeds that are under-sampled in the input data, and/or for which additional important drivers in predicting URN are unaccounted for (e.g., loading, machinery noise). Variables that are not accounted for can include both those that affect the noise radiated by the vessel (vessel-specific design and operation characteristics) and those that affect the use of the monitoring data to accurately calculate source level (accounting for distance and oceanography between the source and the receiving recorder). By developing systems that integrate information from both Canadian and United States (U.S.) monitoring locations to inform vessel operators of their predicted URN, there is an opportunity to improve performance in both arenas. First, integration can increase the sample sizes for vessel types and operational modes that are currently not well constrained by the J-E model's core output predictors (speed, length, type). Second, it enables identification of measurement specifics (e.g., frequencies) that suggest the need for additional work to ensure comparability across regions and monitoring systems.

Using these U.S. datasets, which pair acoustic data and Automated Identification System (AIS) data, source levels were calculated in one-third-octave bands spanning the 15 Hz to 4 kHz band. The effective bandwidth of the original recordings is either 10 Hz to 10 kHz, or 10 Hz to 100 kHz, depending on the deployment Gassmann et al. 2017, ZoBell et al. 2021, Johnson et al. 2025). Recognizing that vessel URN can vary significantly due to differences in ship types, operational conditions, designs, and environmental factors, this project seeks to integrate source level estimates for vessels operating in U.S. waters for the MARINE-T (*Maritime Analytics for Research and Innovation on Noise and Energy – Tool*) project developed by the Department of Transportation's (USDOT's) Maritime Administration (MARAD). To achieve these goals, Scripps Institution of Oceanography (SIO) collaborated with the USDOT Volpe National Transportation Systems Center (Volpe Center) acousticians to interpret the source levels estimated in United States waters and identify variations among different vessel types and operating conditions. Ultimately, this work optimized the accuracy of the MARINE-T model by augmenting the existing application of the J-E source level model, incorporating information based on source levels of vessels operating in U.S. waters, which will support improved regulatory and environmental assessments.

Project Description

To advance understanding and prediction of vessel URN in U.S. waters, SIO partnered with the USDOT Volpe Center to explore and interpret SIO's existing URN database in support of Volpe's decision support tool development effort for MARAD, MARINE-T. The specific URN metric that was analyzed during the project was monopole source level (or source level), which will be used interchangeably in this report. SIO provided expertise to the Volpe group to evaluate differences between the acoustic signatures obtained from the Southern California and Gulf of Mexico monitoring stations and the J-E source level model **(Figure 1**). Integrating information from previous SIO work directly comparing ECHO model estimates with Southern California measurements (Frasier et al. 2022), and using published accounts of the underlying empirical data supporting the J-E model predictions, SIO optimized the J-E model based on their additional sampling.

Task 1 of the project focused on comparing SIO source level measurements with MARINE-T model predictions based on the J-E framework, using 2022 data to evaluate model performance by vessel type. The MARINE-T tool incorporates the J-E model for its core of URN estimation. The Volpe Center implemented the J-E model based on published work, covering the vessel types, sizes, and speeds represented in the U.S. EEZ waters (MacGillivray et al. 2021).

Building on these results, Task 2 examined how variations in the MARINE-T/J–E model coefficients influenced predictions and assessed approaches for tuning them to better match the SIO dataset. Model optimization was performed separately for each vessel class across multiple coefficients, using SIO measurements from 2016–2023 in Southern California and 2020–2023 in the Gulf of Mexico. Because Tasks 1 and 2 are interrelated, their methods and results are presented together in the following sections.

Methods

SIO and Volpe reviewed the existing library of URN measurements from three primary receiver locations located near shipping lanes in the U.S. These recording stations were specifically selected for their proximity to heavily trafficked shipping lanes and their potential to provide high-quality URN measurements with paired Automatic Identification System (AIS) data. From the 36,565 total initial measurements, 25,129 were usable based on URN criteria (single vessel present, no instrumentation noise, minimal background noise; **Table 1**). The measurements were cross-referenced with the 2022 Volpe dataset, resulting in a subset of 1,348 measurements across all ship types. This subset allowed for the comparison of the MARINE-T/J-E source level estimates and SIO measurements.

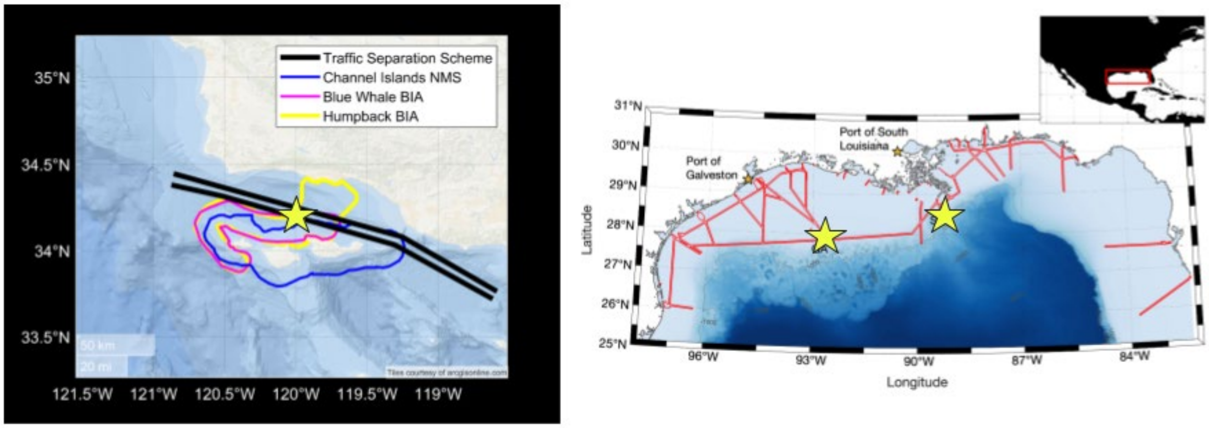


Figure 1. Primary SIO recording stations for ship recordings used in this study. Left: Site B located between the shipping lanes in the Santa Barbara Channel off of Southern California. This site records traffic associated with the port of Los Angeles/Long Beach. Right: Southern Louisiana and Galveston shipping lane listening station locations in the Gulf of Mexico/America.

Table 1. Number of transit counts in each associated region, time frame, and ship type. Vessel type counts represent usable measurements. Ship types were classified using AIS data and grouped according to the categories defined by MacGillivray et al. (2021).

Region	Southern California	Gulf of Mexico	Total
Timeframe	2016 - 2023	2020 - 2023	
Total URN Measurements	20,229	16, 336	36,565
Usable URN Measurements	13,262	11,867	25,129
Bulker	6,860	4,183	11,043
Container	2,694	694	3,388
Tanker	1,069	5,481	6,550
Tug	497	325	822
Recreational	266	2	268
Fishing	43	4	47
Naval	1	0	1
Government / Research	2	0	2
Cruise	121	160	281
Passenger	86	43	129
Dredger	2	0	2
Other	1,621	975	2,596
Unique Vessels	3,815	5,261	

Model coefficient effects and optimization

The effects of five model coefficients - K, K_{LF}, D, D_{LF}, and V_c were investigated across each vessel category. In a series of experiments, each of these model coefficients was adjusted across a range of possible values, while holding all other coefficients constant. The original set of model coefficients from the SIO/MARINE-T data comparison in Task 1 was passed into the model repeatedly, and evaluated against the measurements. The effect of the coefficient adjustment on RMSE was recorded for each case.

Model coefficients were tuned for each vessel type using an iterative process to adjust the five coefficients starting from the original values to improve fit with the SIO URN measurements. Coefficients were optimized by minimizing the root mean square error (RMSE) between predicted and observed one-third-octave source levels (in dB). Practical bounds were enforced on all parameters to maintain physical plausibility, and optimization was performed using a derivative-free global search algorithm (patternsearch). For the speed coefficient (Vc), the bounds were defined as the mean speed over ground (SOG) for each vessel type ± one standard deviation.

The patternsearch algorithm explores the parameter space by evaluating the objective function on a mesh of candidate points around the current estimate. At each iteration, it polls nearby points in a set of search directions (a generalized pattern) and moves to a new point if a lower RMSE is found. If no improvement occurs, the mesh size is reduced, refining the search locally. This process continues until convergence, yielding an approximate global minimum without relying on gradient information.

Candidate coefficient sets were iteratively evaluated, and the history of parameter values and RMSE was tracked to monitor convergence. Fitting was performed independently for each vessel type, without coefficient sharing across types, to produce type-specific coefficients that balanced fit quality and interpretability. Model improvement was assessed by comparing RMSE across vessel types and frequency bands. RMSE in relation to vessel operations (SOG) and design (length) was investigated to determine if higher RMSE was associated with certain conditions.

Results

Agreement between model predictions and measurements varies across vessel types for the 2022 dataset analysis. For the best-represented classes, clear and consistent patterns emerge. Model–data agreement is frequency dependent in all cases, with strong correspondence observed at certain frequency bands.

Bulkers and Tankers: Across the large vessel classes in the 2022 dataset, model estimates and measurements were among the closest for bulkers and tankers with a mean difference of -3.3 (RMSE 7.3) and 0.4 (RMSE = 6.7), respectively **(Table 2)**. The model appears to slightly underestimate observations below ~30Hz, possibly due to differences in actual vs. reported draft which affects the Lloyd's mirror correction at these frequencies, and/or differences in draft distributions for bulkers between datasets **(Figure 2, Figure 3)**. Main energy peaks are broader in the observed data with greater energy at ~80 Hz, versus 50 Hz in the modeled spectra. A notch at 100 Hz present in the model predictions is not observed in the data - this disagreement is noted across multiple vessel categories. A notch at ~200 Hz in the observations is not predicted by the model. The model underestimates observed amplitudes above 500 Hz for a subset of bulkers (note bimodality in bulker plot above 500Hz), while tankers and some bulker estimates are in good agreement (<5 dB difference).

Container Ships and Vehicle Carriers: The 2022 model and measurement results show close agreement at low frequencies (≤50 Hz) for these two vessel categories, suggesting that the model captures draft-related effects well **(Figure 2, Figure 3)**. Above ~60 Hz, model estimates are typically lower than measured levels by approximately 10 dB. This discrepancy may reflect differences in recording environments and the greater attenuation of higher frequencies during propagation. Additionally, the composition of operations and design encompassing the dataset used to train the J-E/MARINE-T model may have included faster ships operating at higher engine loads, which may produce more high-frequency noise than the slower, transiting cargo vessels represented in our measurements.

Tug/Tow vessels: Although underrepresented in the dataset, this category is of particular interest due to its relevance to port operations. Observed source levels were generally 15–20 dB lower than model predictions across all frequencies, and measured peak frequencies occurred near 80 Hz, which is substantially lower than the ~400 Hz peaks predicted by the model **(Figure 2, Figure 3)**. These discrepancies led to the highest mean difference across all ship types with 18.1 dB and an mean RMSE of 21.6 dB **(Table 2)**. These discrepancies likely reflect differences in operational mode: tug and tow vessels transiting in shipping lanes are often not engaged in active towing, resulting in lower engine load and reduced URN. The operational mode, size, and design characteristics of the tug/tow vessels used to train the model are unknown, which may further contribute to the mismatch.

Additional Takeaways: Smaller and more rarely represented vessels showed a variety of different patterns depending on the frequency and ship type. Passenger ship source level measurements were lower below 400 Hz than the MARINE-T model estimates. Recreational vessel measurements were overestimated by the MARINE-T model below 80 Hz and underestimated above 80 Hz for the 2022 dataset. This may be due to the wide variety of vessel types that fall into this category, and the overall small sample size. Model estimates for other minimally-represented categories fit the data well, including cruise, government/research, and fishing vessels.

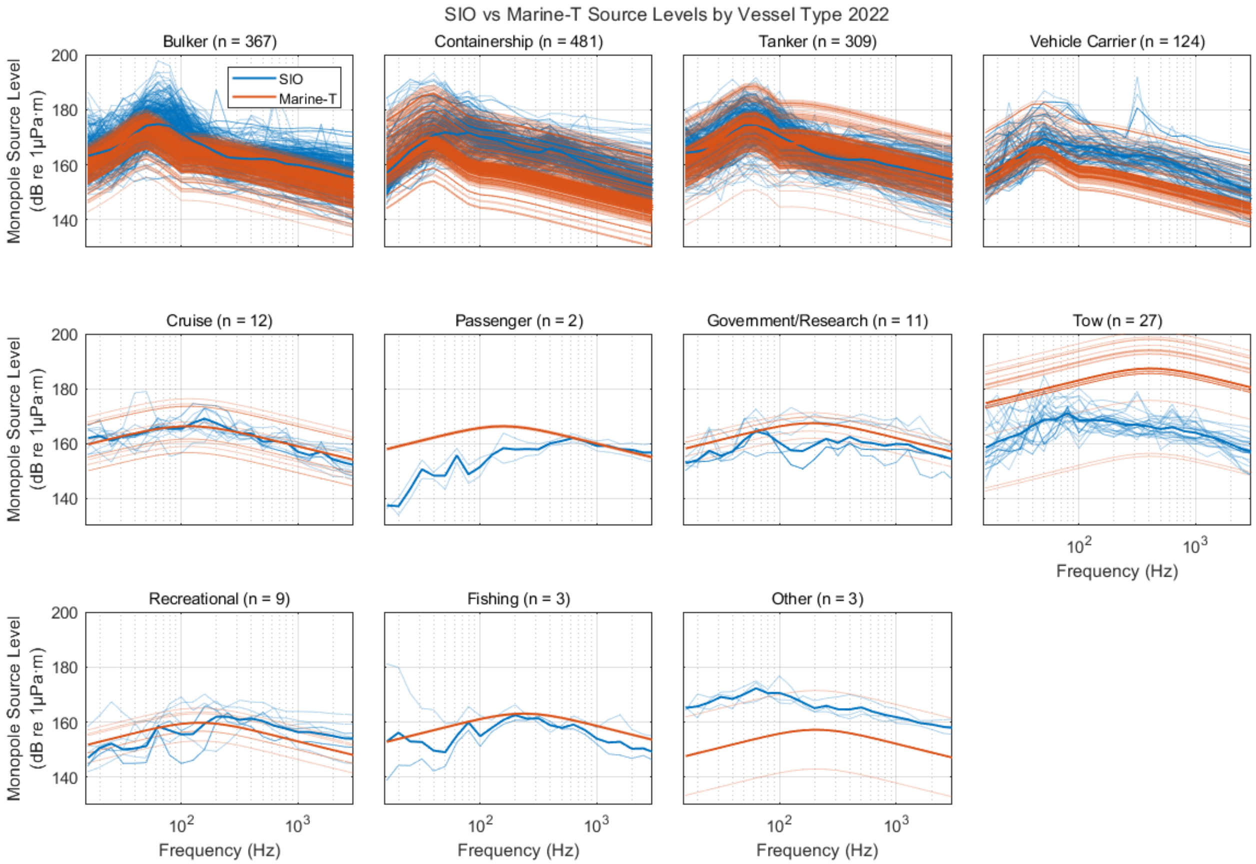


Figure 2. Comparison of SIO-measured (blue) and MARINE-T model–predicted (red) source levels for vessel passages recorded in 2022 across three sensors. Large vessel categories are shown in the top row, while smaller and less frequently observed categories are presented in the bottom two rows.

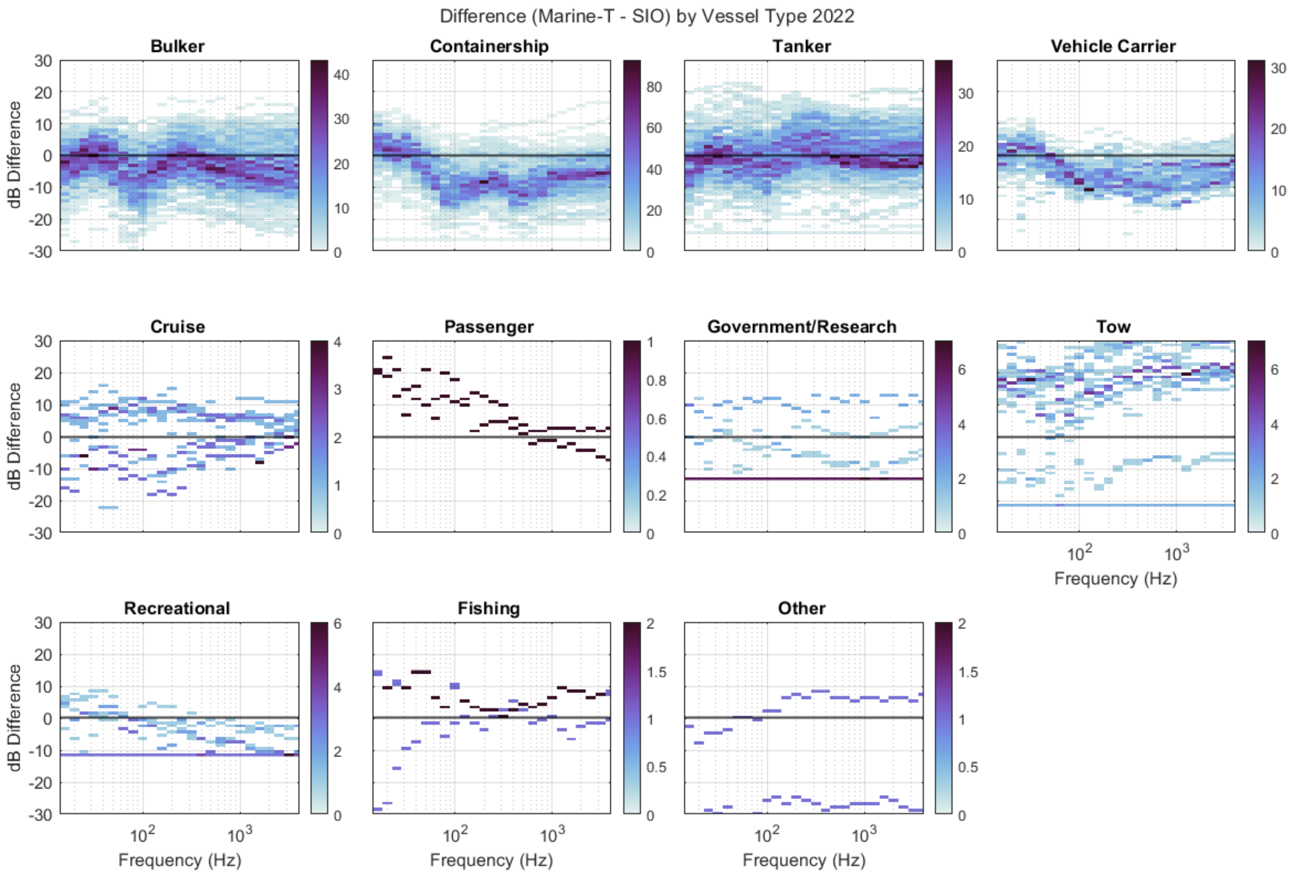


Figure 3. Decibel differences between MARINE-T model predictions and SIO measurements (MARINE-T – SIO) across frequency bands and vessel types. Model–measurement agreement varies with frequency and vessel class, with some ship types showing closer correspondence than others.

Table 2. Goodness-of-fit statistics between 2022 MARINE-T model predictions and SIO measurements by vessel type. Reported values include the mean, standard deviation, and root mean square error (RMSE) of the differences between MARINE-T model predictions and SIO observations.

Vessel Type	Mean (dB)	Standard Deviation (dB)	RMSE (dB)
Bulker	-3.3	5.0	7.3
Containership	-6.0	3.9	8.6
Tanker	0.4	5.6	6.7
Vehicle Carrier	-4.9	3.1	7.1
Cruise	0.2	7.6	8.2
Passenger	7.2	4.3	10.8
Government/Research	2.2	8.2	8.3
Tow	18.1	10.5	21.6
Recreational	-3.1	2.6	5.8
Fishing	3.3	6.6	8.5
Other	-12.3	22.1	20.2

Model Optimization

Model optimization included an increased database spanning URN measurements from transits from 2016 through 2023 in Southern California and 2020 to 2023 in the Gulf of Mexico. The J-E model was compared to measurements from all of the years represented to understand differences for each vessel type. J-E model coefficients were explored to understand their effects and reasonable adjustment ranges. The effects of these coefficients are not isolated, rather they interact and in some cases have opposing effects on the overall shape of the source level spectra. An example of the optimization of the coefficients for bulkers throughout the iterative process is shown in **Figure 4**.

K_LF was found to be the most influential model parameter in the container and bulker models, controlling the amplitude of the main spectral peak in the large vessel models. It does not affect the amplitudes at higher frequencies. K_LF is not currently used in the smaller vessel models that feature simple hyperbolic spectral shape.

D_LF is similarly only used in the more complex models for large vessels. It controls the slope on the low frequency side of the main spectral peak, and impacts the amplitude of the main spectral peak.

K controls the overall amplitude of the upper frequency portion of the spectra in the large vessel models. In the small vessel/single peak models, K controls the absolute amplitude of the spectra, with no frequency dependence.

D generally controls the salience of the notch between low and high frequency stages in the large vessel models, by adjusting the crossover point between the two stages. In the simpler small vessel models, D shifts the peak frequency and controls the slope on either side of the peak, with larger effects below the peak frequency.

Vc generally controls the frequency of the main peak in the large vessel models, with higher values corresponding to lower peak frequencies. It has marginal effects on peak amplitude. In the single peak models it affects amplitude, alongside K.

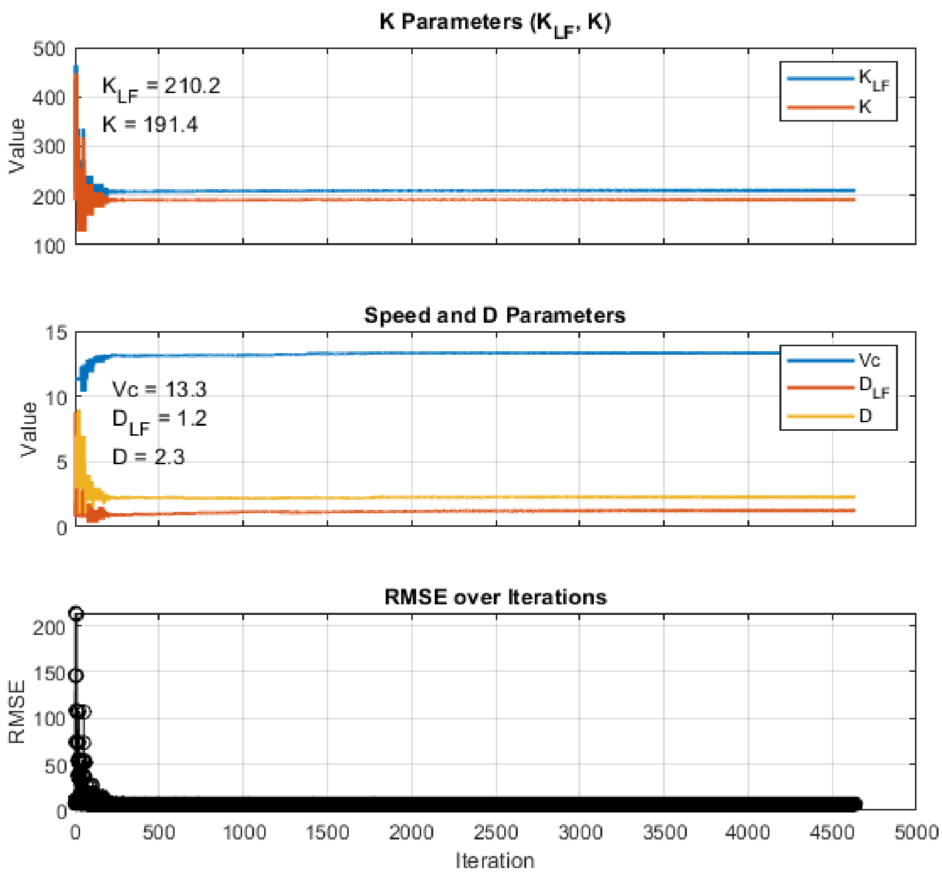


Figure 4: Bulker coefficient optimization for over 4500 iterations

Model Tuning
Through the coefficient tuning process we were able to reduce the RMSE of the model predictions by 4.5 dB on average, with changes in RMSE ranging from 0.1 dB (passenger) to 31.2 dB (dredger; **Table 3**).

The D_LF and K_LF coefficients were increased for all large vessel classes optimized (bulker, container, tanker), with the largest increase in tankers (213.1 and 1.5 dB, respectively). These increases improved the agreement between the data and model for the low frequency peak portions of the spectra (<100Hz).

The K coefficient was increased by the optimization procedure in all cases, with the largest increases occurring for the smaller and sparser vessel classes. These changes, ranging from +0.1 (cruise) to + 21.3 dB (dredger), shifted the peak frequencies lower for these classes to improve model fit.

Adjustments to the D coefficient varied considerably, as this parameter was tuned to control the crossover between peaks and high frequency slopes for large vessels, and overall amplitude for smaller vessels. Adjustments ranged from -2.8 dB (dredger) to +4.9 dB (government/research). Negative adjustments to this coefficient relative to the original value tend to increase amplitudes, while positive adjustments decrease it.

Vc the reference speed for each vessel type, was also adjusted in both positive and negative directions. These adjustments may point to differences in vessel behavior in some cases (e.g. lower speeds of large vessels), as well differences in vessel composition (e.g. transit rather than towing behavior of tug/tows, or different types of recreational vessels). Adjustments ranged from -4.7 knots (recreational) to +14.6 knots (government/research).

Table 3: Optimized model coefficients per vessel type based on SIO measurements and mean RMSE ± standard deviation of optimized model and SIO measurements. Optimized coefficients are shown on the top row and original parameters are shown beneath in parentheses.

Ship type	K_LF (dB) Optimized (original)	K (dB)	Vc (knots)	D_LF (dB)	D (dB)	RMSE (dB)
Bulker	210.2 (208)	191.4 (191)	13.3 (13.9)	1.2 (0.8)	2.3 (3)	6.0 ± 3.5 (6.2 ± 3.5)
Containership	208.3 (208)	191.5 (191)	16.8 (18.0)	1.2 (0.8)	4.1 (3)	5.5 ± 3.7 (5.8 ± 3.8)
Tanker	213.1 (208)	191.9 (191)	14.3 (12.4)	1.5 (0.8)	2.1 (3)	6.1 ± 3.2 (6.4 ± 3.3)
Cruise	–	191.1 (191)	18.1 (17.1)	–	3.9 (4)	7.8 ± 4.4 (8.0 ± 4.2)
Passenger	–	196.1 (191)	9.6 (9.7)	–	5.1 (3)	10.6 ± 8.2 (10.7 ± 8.7)
Government / Research	–	211.2 (191)	22.6 (8.0)	–	7.9 (3)	4.1 ± 0.0 (8.1 ± 0.1)
Tug / Tow	–	203.5 (191)	8.0 (3.7)	–	1.7 (3)	9.1 ± 4.9 (10.3 ± 5.1)
Naval	–	198.3 (191)	9.9 (11.1)	–	2.4 (3)	3.8 ± 0.0 (11.4 ± 0.0)
Recreational	–	192.2 (191)	5.9 (10.6)	–	3.9 (3)	12.1 ± 7.1 (16.4 ± 10.2)
Fishing	–	197.7 (191)	6.6 (6.4)	–	3.4 (3)	8.0 ± 3.9 (9.7 ± 4.0)
Dredger	–	212.3 (191)	6.4 (9.5)	–	0.2 (3)	7.4 ± 1.9 (38.6 ± 7.6)
Mean Adjustment	2.5 ±2.4	6.9 ± 7.9	1.0 ± 5.1	0.5 ± 0.2	0.3 ± 2.0	-4.6 ± 2.1

Large Vessel Categories

The performance of the original J-E model and the optimized model was first evaluated for large vessel classes, including bulkers, tankers, and container ships. Bulker and tanker transits in the SIO database may include smaller vessels than those used to fit the original J-E model, resulting in higher RMSE for smaller vessels. Elevated RMSE was seen for slow-speed (< 10 knots) bulker transits and high-speed (>15 knots) tanker transits, likely reflecting differences between the SIO measurements and SOG ranges used for the original J-E model. Although coefficient optimization improved model fit at these “extreme” values, RMSE dependence on SOG and vessel length persists, such as in the case of container ships traveling <10 knots. The model specification, which relies on logarithmic length and speed related adjustments, may inherently tend to overestimate reductions at low values. Overall, model fit was improved by the optimization process, particularly below 100 Hz and above 300 Hz. Agreement between 100 and 300 Hz could be further tuned if appropriate by modifying the model to allow variation in the crossover frequency.

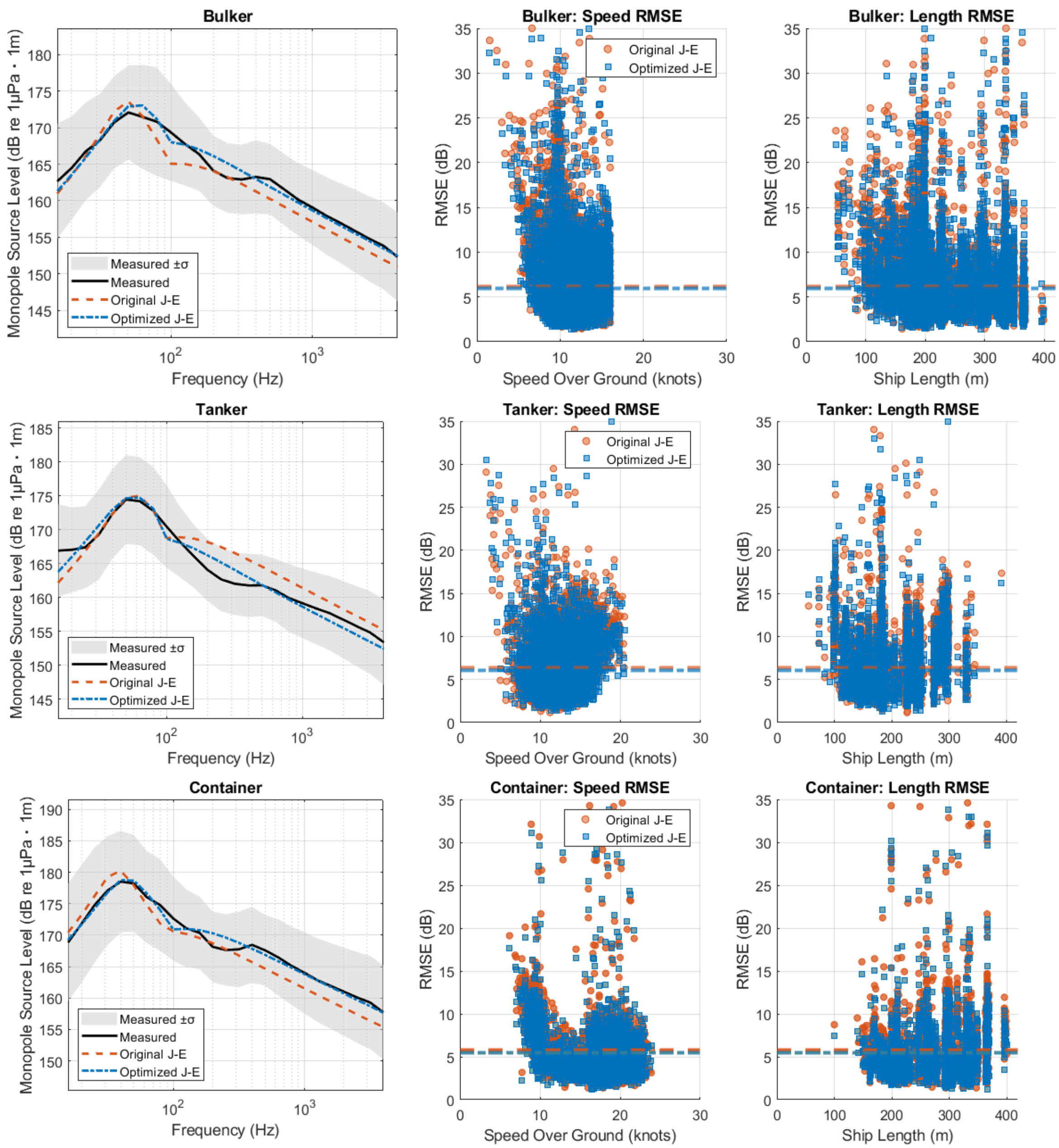


Figure 5. Measurement (black, solid), original J-E (red, dashed), and optimized J-E (blue, dotted) for large vessel classes (bulker, container, tanker). The mean RMSE in relation to SOG (knots) and ship length (m) are shown for the original J-E model (red) and the optimized J-E model (blue).

Smaller Vessel Categories with Good Agreement

Models for three categories of smaller vessels - cruise, passenger and fishing vessels - required minimal adjustment to align with the observed transits from the SIO database. Among these three, cruise ships were best fit by the single curve model. Passenger and fishing vessel spectra had somewhat more complex and varied (passenger) shapes that were more difficult to approximate with a single curve. Observations in these categories were limited, and may not adequately represent the typical features of each class.

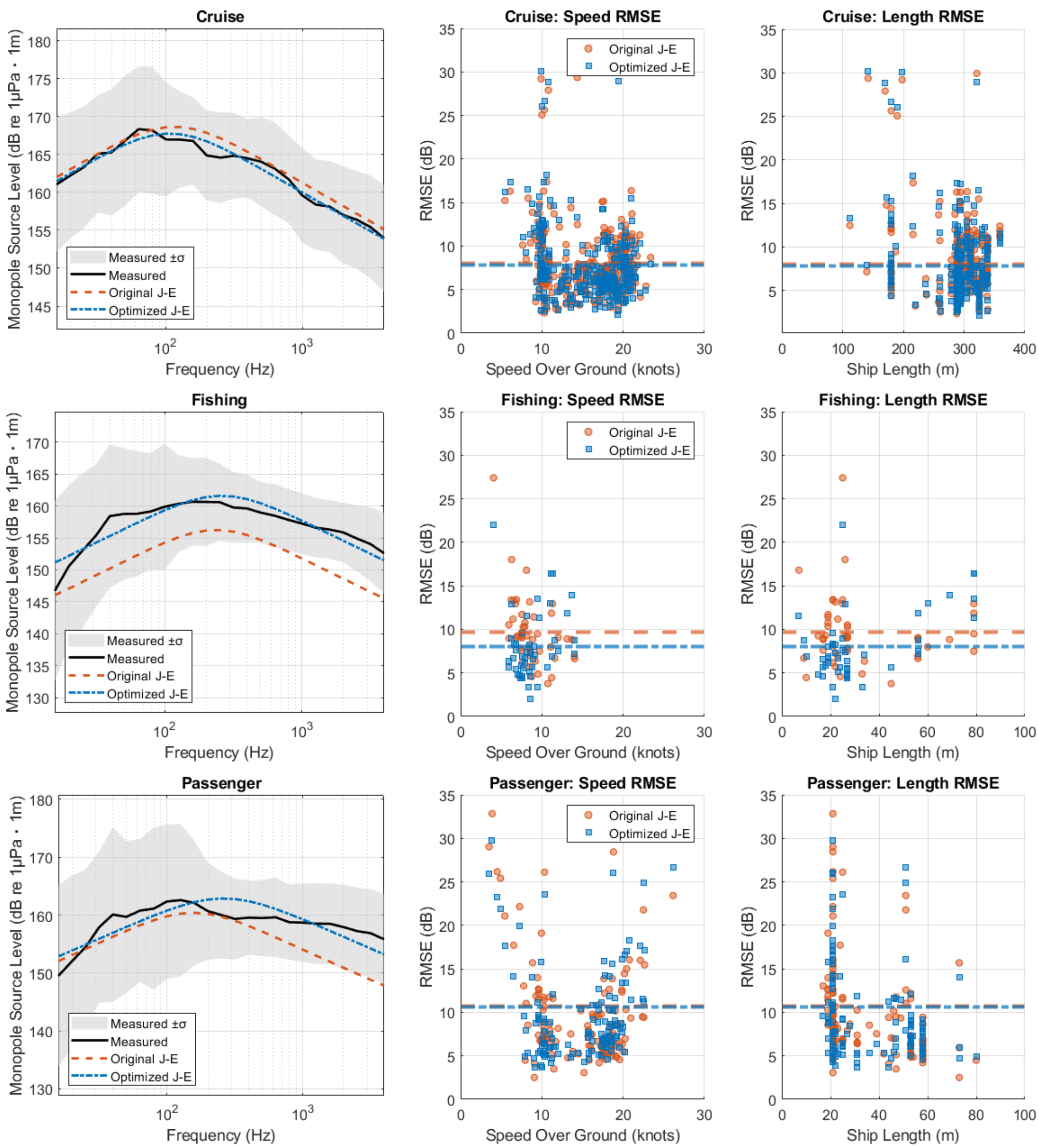


Figure 6. Measurement (black, solid), original J-E (red, dashed), and optimized J-E (blue, dotted) for small vessel classes with relatively good agreement (cruise, fishing, passenger). The mean RMSE in relation to SOG (knots) and ship length (m) are shown for the original J-E model (red) and the optimized J-E model (blue).

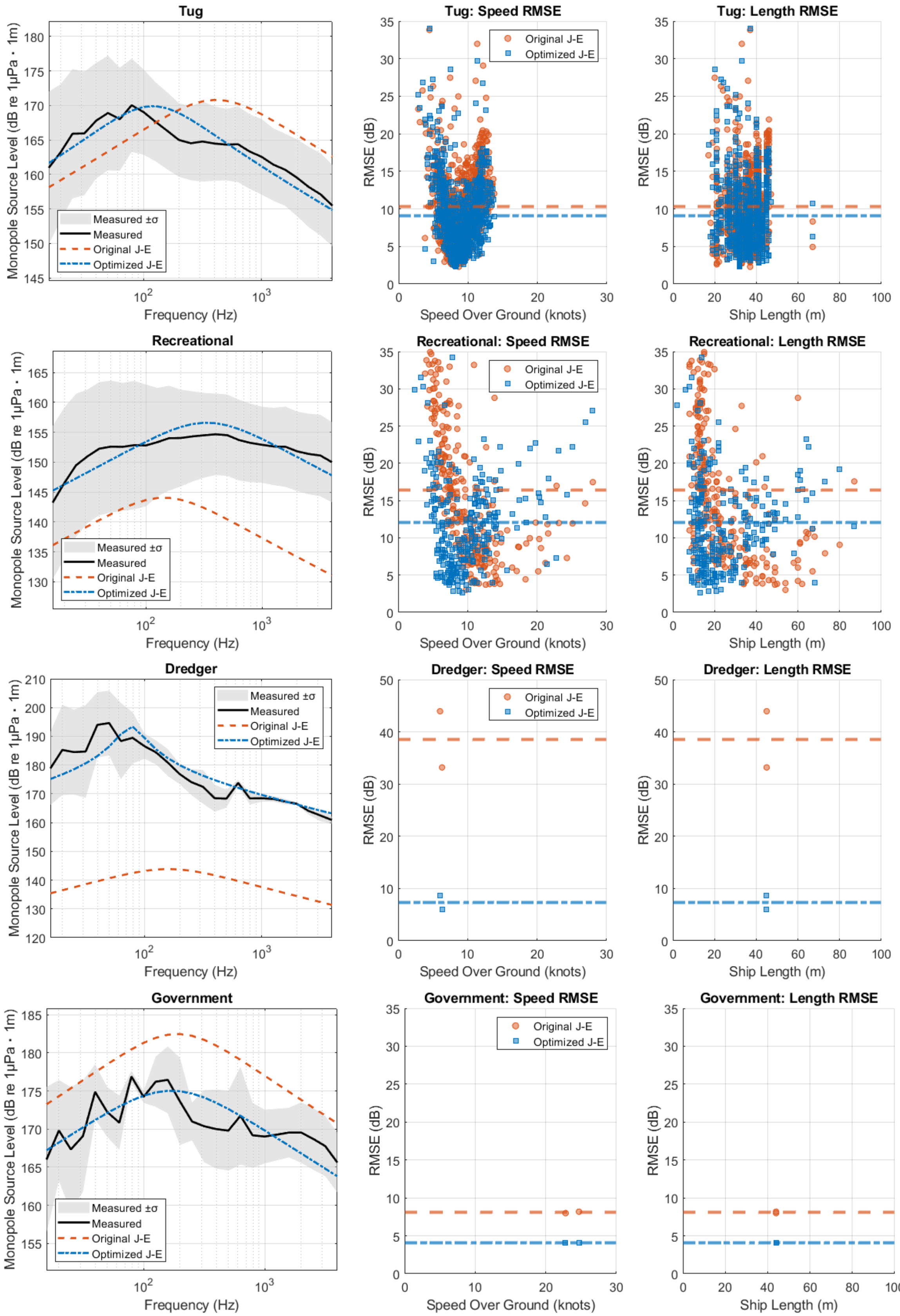
Smaller Vessel Categories with Poor Agreement

Multiple small vessel categories had poor agreement reaching over 30 dB mean RMSE between the original J-E model and the SIO measurements. The poor agreement spanned across frequency, with the J-E model predicting lower overall source level for recreational, dredger, and naval vessels and higher overall source levels for government vessels.

Tug vessels exhibited a clear mismatch in spectral shape compared to the J-E model. While the model predicted a spectral peak near 400 Hz, the measured spectra peaked around 90 Hz. As a result, the J-E model tended to underestimate source levels below ~200 Hz and overestimate them above this frequency. The mean RMSE between modeled and measured spectra reached a minimum at a SOG of approximately 9 knots, with higher RMSE values observed at speeds both above and below this point. The optimized J-E model provides a closer fit to the shape of the measured spectra and shifts the predicted peak toward the peak observed in the measurements. However, the measured spectral shapes exhibit greater variability than can be captured by a single model curve.

For recreational vessels, source level estimates from the J-E model were approximately 10–20 dB lower than the mean measured source levels. The poorest agreement occurred at relatively slower speeds (below 9 knots) and for smaller vessels (less than 20 meters in length). Optimization shifted the predicted spectra upward, resulting in improved alignment with the measured data.

It is notable that some of these vessel types (naval, government, and dredger) were represented by small sample sizes. Dredger and government vessels each had two transits represented and naval vessels were represented by a single transit. The differences between the J-E model predictions and the measured source levels were substantial, with the model underestimating levels by approximately 10-40 dB. For the dredger type in particular, the J-E model predicted source levels between 130 and 145 dB, whereas the measured values ranged from 160 to 195 dB. The reason for this discrepancy remains unclear, understanding the types of dredgers included in the J-E mode training dataset would help determine whether our vessel falls within the range for that category or if its characteristics differ in a way that limits the accuracy of the model. Model tuning brought the optimized J-E spectral predictions into closer agreement with the measurements. However, increasing the sample size for these less common vessel types may help determine whether the observed discrepancies are consistent across similar vessels or are specific to individual cases. Such additional data would improve confidence in the optimized model performance and help guide refinements needed to better capture their acoustic characteristics.


Tug
Monopole Source Level (dB re 1μPa · 1m)
Frequency (Hz)
Measured ±σ
Measured
Original J-E
Optimized J-E
Tug: Speed RMSE
RMSE (dB)
Speed Over Ground (knots)
Original J-E
Optimized J-E
Tug: Length RMSE
RMSE (dB)
Ship Length (m)
Recreational
Monopole Source Level (dB re 1μPa · 1m)
Frequency (Hz)
Measured ±σ
Measured
Original J-E
Optimized J-E
Recreational: Speed RMSE
RMSE (dB)
Speed Over Ground (knots)
Original J-E
Optimized J-E
Recreational: Length RMSE
RMSE (dB)
Ship Length (m)
Dredger
Monopole Source Level (dB re 1μPa · 1m)
Frequency (Hz)
Measured ±σ
Measured
Original J-E
Optimized J-E
Dredger: Speed RMSE
RMSE (dB)
Speed Over Ground (knots)
Original J-E
Optimized J-E
Dredger: Length RMSE
RMSE (dB)
Ship Length (m)
Government
Monopole Source Level (dB re 1μPa · 1m)
Frequency (Hz)
Measured ±σ
Measured
Original J-E
Optimized J-E
Government: Speed RMSE
RMSE (dB)
Speed Over Ground (knots)
Original J-E
Optimized J-E
Government: Length RMSE
RMSE (dB)
Ship Length (m)

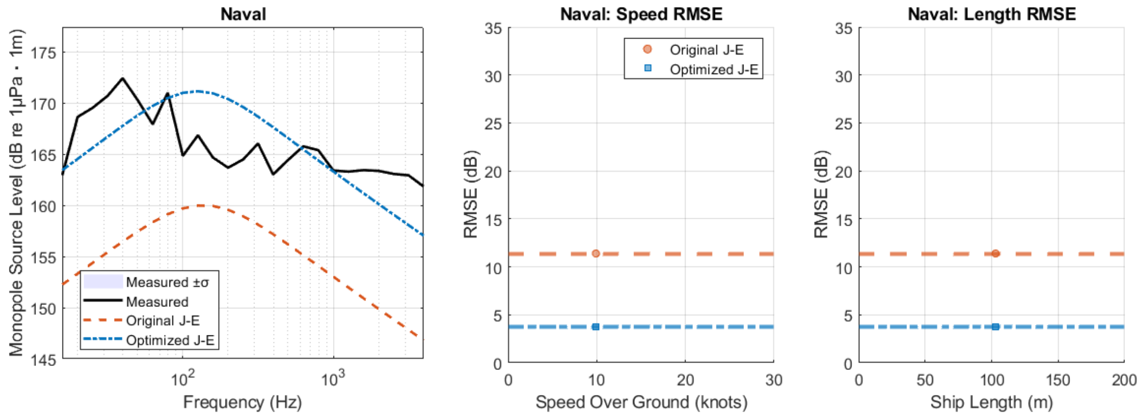


Figure 7. Measurement (black, solid), original J-E (red, dashed), and optimized J-E (blue, dotted) for small vessel classes with relatively poor agreement (recreational, tug, dredger, naval). The mean RMSE in relation to SOG (knots) and ship length (m) are shown for the original J-E model (red) and the optimized J-E model (blue).

Discussion

Estimation of vessel URN is critical for assessing noise levels, formulating management plans, and predicting the effectiveness of potential mitigation strategies. However, vessel source levels are highly variable and influenced by many different factors, including a range of operational conditions that are not documented or available for use in model predictions. The J-E model provides useful estimates of vessel source levels using a practical framework. This study demonstrates that the model can be tuned, within reasonable limits, to improve agreement with observations from new regions and to refine predictions on underrepresented vessel types, lengths, and speeds.

After optimization, model predictions showed notable improvements, though several opportunities remain for refinement. A dual- or multi-output modeling approach may refine performance, especially for vessel types within complex spectral shape. Additionally, incorporating additional years of data at the three monitoring stations may expand the diversity of ship types represented. Including more monitoring sites across more regions would also enhance coverage of rarer vessel classes and ensure that the model is not unintentionally optimized for any particular region. Retraining the original J-E model with a broader range of operational conditions and vessel designs could further improve its predictive capability. Optimization has proven to improve model performance, but the limits of the original training datasets is a remaining constraint.

Vessel URN estimates are typically derived from passive acoustic recording stations adjacent to or associated with traffic from specific ports or fairways. In principle, vessel URN measured at close range and corrected for vessel draft and acoustic environment should be largely independent of recording location. However, in practice various factors contribute to differences.

For example, the composition of AIS-based vessel categories may vary by location, with predominantly smaller or larger vessels visiting certain ports. Additionally, typical vessel speeds may vary by recording location due to proximity to destination, land, sea state, and local regulations. Model tuning may be beneficial to improve fit on new, and underrepresented scenarios. Caution is warranted, however, as some measurement differences may arise from sensor characteristics, oceanographic conditions, acoustic environment, and local ambient noise. Overfitting may occur if models are tuned to data from highly similar recording conditions. Sparse vessel classes are particularly susceptible to over-fitting, when speed and length variability is limited.

The Volpe Center researchers have recognized that within some vessel categories, specific activities such as active towing (tugs), trawling (fishing vessels) or dredging (dredgers) as opposed to simply transiting, can strongly influence vessel URN. The model-data comparison in this study further supports this observation. Inferring operational conditions either directly from AIS or indirectly based on vessel movement, location, and proximity to other vessels could improve URN estimates for these classes.

One issue identified, but not addressed in this study, was the crossover frequency in the dual-output model for large vessel classes, specifically those with the D_LF and K_LF coefficients. None of the adjustments in this study altered the position of the crossover frequency, which remained at 100 Hz. In the SIO dataset, no notch is observed at 100 Hz, but there is some evidence of a similar feature near 250 Hz. The persistence and cause of this feature should be examined more closely. The PIANO model upon initial inspection may more closely align with the 250 Hz notch in the SIO measurements (Lloyd et al. 2024). Further investigation of SIO measurements with the PIANO model may be helpful. In general the models utilizing multi-curves seemed to more effectively represent the observed source level spectra in the SIO dataset. Model fits for the tug passenger, government, naval, and dredger vessel categories could all likely be improved by including a multi-output option.

Conclusion

This study demonstrates that integrating measurements with an existing source level model provides opportunities to improve the accuracy and applicability of vessel noise predictions used in management and regulatory tools such as MARINE-T. While the J-E model offers a practical foundation, its performance varies across vessel type, size, and operational conditions. Optimization showed improvement when tuning coefficients to better reflect the diversity of vessels operating in U.S. waters. Nonetheless, model refinements remain constrained by the scope of the original training data and by real-world variability in vessel behavior, environmental conditions, and measurement characteristics across monitoring sites not incorporated in the model framework. Expanding datasets across regions and years, incorporating a wider range of ship sizes and operating modes, and adopting multi-output approaches, particularly for complex vessel classes, may further strengthen predictive capability. Additional considerations, such as operational state and investigation of spectral features like the crossover frequency, are also likely to enhance model robustness. Overall, the results underscore both the value and the limitations of tuning an empirically based model and highlight the importance of continued data integration across monitoring systems to support reliable URN estimation for environmental assessment and mitigation planning.

References

Frasier, K.E., ZoBell, V.M. MacGillivray A. O., Dolman, J.N. Ainsworth, L., Zhao J. (2022). Evaluation of ECHO Vessel Noise Correlation Models with a Novel Dataset Collected in the Santa Barbara Channel. Marine Physical Laboratory Technical Memorandum 658. Jan 2022.

Gassmann, M., Wiggins, S. M., & Hildebrand, J. A. (2017). Deep-water measurements of container ship radiated noise signatures and directionality. The journal of the acoustical society of America, 142(3), 1563-1574.

Johnson, K. H., ZoBell, V. M., Hodge, L. E., Soldevilla, M. S., Hildebrand, J. A., & Frasier, K. E. (2025). Characterization and modeling source levels of commercial vessels in the Gulf of Mexico. The Journal of the Acoustical Society of America, 158(3), 2250-2268.

Lloyd, T., Daniel, J. M., Bosschers, J., & Schuster, M. G. (2024). 'PIANO': a Physics-based Semi-empirical Source Level Model for Fleet-scale Ship URN Prediction. Technische Universität Hamburg. Universitätsbibliothek.

MacGillivray, A., De Jong, C. "A reference spectrum model for estimating source levels of marine shipping based on automated identification system data." Journal of Marine Science and Engineering 9.4 (2021): 369.

ZoBell, V. M., Frasier, K. E., Morten, J. A., Hastings, S. P., Peavey Reeves, L. E., Wiggins, S. M., & Hildebrand, J. A. (2021). Underwater noise mitigation in the Santa Barbara Channel through incentive-based vessel speed reduction. *Scientific reports*, *11*(1), 18391.