A NEW EMPIRICAL MODEL DERIVATION FOR THE ESTIMATION OF TURBULENT BOUNDARY LAYER POWER SPECTRAL DENSITY AIRCRAFT USING MACHINE LEARNING REGRESSION

Zachary Huffman *[1] and Joana Rocha †[1]
[1]Department of Mechanical and Aerospace Engineering, Carleton University, Ottawa, ON

1 Introduction

At cruise conditions, the most significant source of aircraft cabin noise are wall-pressure fluctuations induced by the turbulent boundary layer (TBL). This noise is often significant and has been associated with an elevated risk of cardiovascular disease, hearing loss, and sleep deprivation in flight crews [1, 2]. Therefore, the development of an accurate empirical model that predicts noise generated from TBLs is an important ongoing research topic. Early models, such as those from Lowson and Robertson, were primarily derived by simplifying and solving the Reynolds-averaged Navier-Stokes equations for the pressure fluctuation term and adding length and velocity scales to best match experimental data. Subsequent models were derived by applying statistical and mathematical techniques to simplify earlier models, or via modifications that addressed apparent shortcomings. Past research has had varying success. Most models are accurate only near their design Mach and Reynolds numbers [3]. Only recently was the possibility of using machine learning techniques explored by Dominique, who produced a highly accurate TBL model via an artificial neural network [4]. This paper extends Dominique's work by applying a different machine learning technique, nonlinear least squares regression (NLS).

2 Method

The model was derived in five steps, as outlined in [5]. First, an exploratory data analysis (EDA) was performed, which sought to identify data sources and possible candidate variables. In total, 14 data sources were available, consisting of wind tunnel data procured at Carleton University. 12 experimental runs were placed in the training dataset, which had an average airspeed of 10.6 m/s and Reynolds number of approximately 850,000; while the testing dataset had an average airspeed of 8.92 m/s and Reynolds number of 650,000. Additionally, a total 23 candidate variables were to be considered –most were selected based on their appearance in previous TBL noise models, but some other common dimensionless fluid dynamic parameters were also considered.

The second step was dimensional analysis, which sought to establish a priori knowledge on the final form of the model via consideration of the dimensions. Specifically, the since the PSD of pressure fluctuations in the TBL has units of [Pa^2/Hz], Equation 1 represented a highly simplified model that could theoretically predict the PSD:

$$\emptyset(f) = \frac{A[P]^2}{1/t} \tag{1}$$

where P is any candidate variable in units of Pa, $1/t$ is any candidate variable in units of Hz, and A is some coefficient to be fit to the model.

While dimensional analysis made it theoretically possible to create candidate models and iteratively add complexity (via the addition of terms), it left two unanswered problems: how to fit coefficients/exponents and how to identify the performance of each model. These problems were solved via the third and fourth steps, model development and testing, respectively. For the third step, the NLS algorithm was implemented in R to fit the model against all possible candidate variable combinations. Next, model testing was performed via a combination of its Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Mean Squared Prediction Error (MSPE). After the optimum form of any given candidate model was selected, its shortcomings were identified, a new model form was proposed, and steps three and four were repeated. This process stopped only when it became clear that added complexity failed to yield a superior model. At this point, the final selected model underwent validation, which sought to assess model accuracy against outside data. However, due to the limited availability of outside data, only freestream velocities from 7.20 m/s to 27.1 m/s and Reynolds numbers from approximately 530,000 to 6,200,000 could be considered [6].

3 Results

In total, iterative model generation procedure allowed for 186 unique model forms to be tested. The best performing model, per the statistical tests, is presented in Equation 2 below:

$$\emptyset(f) = \frac{(0.39031)M^{2.6241}q^2}{\left({U_\tau}/{\delta}\right)[(0.019972)St^{5.2179} + Re_T^{1.3968}]} \tag{2}$$

where M is the Mach number, q is the dynamic pressure, U_τ is the friction velocity, δ is the boundary layer thickness, St is the Strouhal number, and Re_T is the turbulence Reynolds number. Figures 1-3 show plots of Equation 2 against internal training data, testing data, and validation data, respectively. Additionally, note that the iterative procedure itself provided several useful insights into the nature of TBL modelling. First, the optimum candidate variables to use in the velocity/time term in the denominator has been a source of disagreement amongst researchers; however, the algorithm repeatedly suggested that either U_τ/δ or U_∞/δ^* (freestream velocity and displacement boundary layer thickness) were the optimum options. Second, the optimum Reynolds number in

* zacharyhuffman@cmail.carleton.ca
† joana.rocha@carleton.ca

the denominator has been disputed, but this paper suggests Re_T is optimum; and finally, concerning pressure in the numerator, this paper suggests q is optimal.

4 Discussion and Conclusion

Overall, Equation 2 showed a better agreement with all three plots, albeit with a tendency to underestimate PSD in the medium-to-high frequency regions. Furthermore, while less apparent in the plots, the model tended to lose accuracy against validation data if it was at higher airspeed and Reynolds number. However, given the limited number of cases covered in validation datasets, this should not be seen as an absolute conclusion. Possible causes of this discrepancy include the limited number of unique parameters in the training dataset leading to low data over-fitting issues; the noisiness of the training datasets at low and high frequencies necessitating cleansing; compilation issues with the NLS algorithm, reducing the number of available candidate models to choose from; the inability of statistical tests to filter out frequencies that were less important to human hearing; limited validation data; and the highly negative exponent on St leading to the PSD roll-off occurring at too low of a frequency.

This technique offered several advantages that were not present in earlier TBL modelling techniques. First, it permitted 186 unique model forms to be tested. This is unmatched in prior literature, except by Dominique. However, the use of the neural networks in the Dominque model did not permit the same number of statistics, plots, and visual analysis to be performed/calculated on each model [4]. Second, the openness of the NLS technique allowed unanswered TBL-modelling questions to be answered, such as the optimum candidate variables to use in the numerator and denominator. Finally, as the model generation technique did not rely on TBL-specific physics, it can be generally applied to other engineering applications. Overall, the performance of the model presented and advantages demonstrate that future researchers may consider extending these techniques to novel problems and datasets.

References

[1] C. D. Zevitas et al., "Assessment of noise in the airplane cabin environment," *Journal of Exposure Science and Environmental Epidemiology*, vol. 28, p. 568–578, Mar. 2018.

[2] J. E. Robertson, “Prediction of in-flight fluctuating pressure environments including protuberance induced flow,” Wyatt Laboratories, Huntsville, AL, Tech. Report, WR 71-10, Mar. 1971.

[3] J. B. Blitterswyk and J. Rocha, “An experimental study of the wall-pressure fluctuations beneath low Reynolds number turbulent boundary layers,” *The Journal of the Acoustical Society of America*, vol. 141, pp. 1257-1268, Feb. 2017.

[4] J. Dominique et al., “Artificial Neural Networks Modelling of Wall Pressure Spectra Beneath Turbulent Boundary Layers,” *Physics of Fluids*, vol. 34, no. 3, Mar. 2022.

[5] C. Chatfield, *Problem Solving: A Statisticians Game*. Chatfield, UK: Chapman & Hall, 1988.

[6] N. Thompson and J. Rocha, “Comparison of Semi-Empirical Single Point Wall Pressure Spectrum Models with Experimental Data,” *Fluids*, vol. 6, no. 270, Jul. 2021.

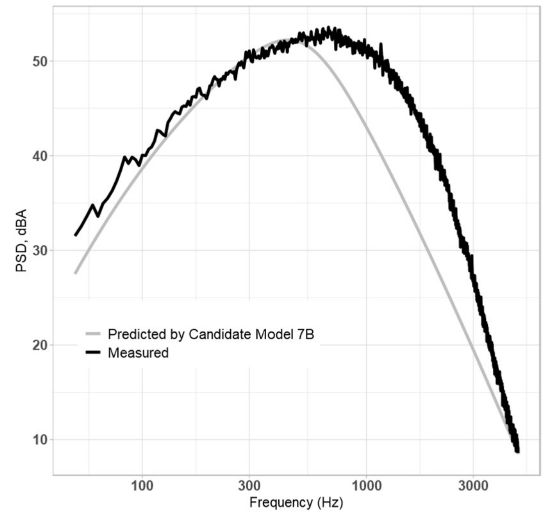

Figure 1: Plot of Equation 2 against selected training data.

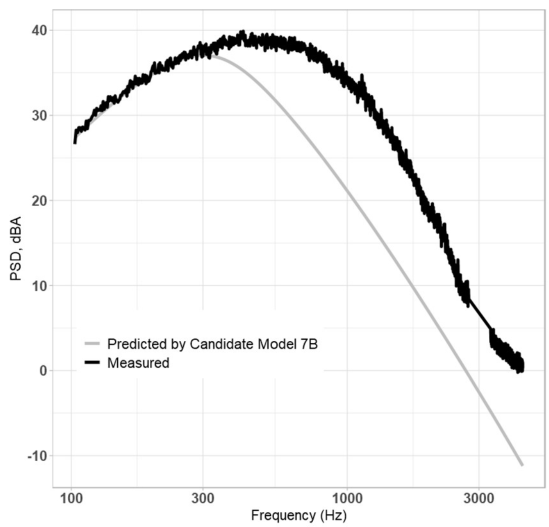

Figure 2: Plot of Equation 2 against selected testing data.

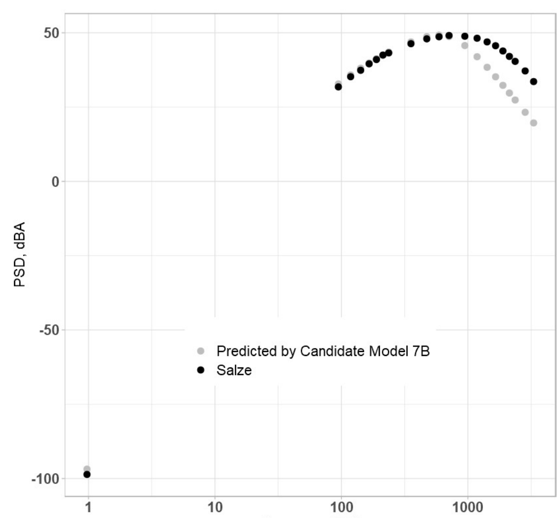

Figure 3: Plot of Equation 2 against selected validation data.