Issues in the development of a survey simulation tool to explore robust estimation of models of annoyance due to aircraft noise

K. Foertsch, P. Davies

Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN 47907-2031, USA

ABSTRACT

A community survey simulation tool is described that has been developed to examine the effects of sampling populations around airports on the estimation of annoyance models. Aircraft noise exposure is predicted on a fine grid around airports for single operations with Integrated Noise Model (INM) or Noise Model Simulation (NMSim); these are used with airport operations scenarios to determine noise exposure at points on a finely resolved grid. Survey populations are defined for the areas surrounding the airports using available microdata and aggregate data from the U.S. Census; households and associated individuals are assigned to noise grid points. Demographic data can be incorporated into the simulation tool to examine the potential for non-acoustic confounding factors. Various sampling methodologies and signal-to-noise ratios are used in Monte Carlo simulations to examine how they affect parameter estimates in an annoyance model that includes number-of-events and a measure of average event sound level as predictor variables.

INTRODUCTION

A goal of current research is to predict how a community will react to changes in airport operations. Advances in aircraft design are resulting in reductions in sound level produced by aircraft, but aviation forecasts predict continuing growth in air traffic in the years to come. A central question surrounding the topic of annoyance to aircraft noise concerns a trade-off between the number-of-events and the sound level of aircraft. Most predictive models of community annoyance to aircraft noise are functions of the average A-weighted sound energy and do not explicitly contain an independent number-of-events term. Different combinations of aircraft sound levels and number-of-events can produce the same A-weighted energy average, but do they result in the same level of community annoyance?

Several researchers have developed models that are an alternative to energy equivalence-based models such as L_{dn}, L_{den}, or L_{night}. These alternative models typically contain two types of terms: measures of the sound level of events, e.g., average of individual events' PNL_{max}, L_{Amax}, *ASEL*, etc.; and number-of-events (*N*) above a certain sound level (Rylander et al. 1980), or the logarithm of the number-of-events above a certain sound level (TRACOR 1970, 1971; Connor & Patterson 1972, 1976; Rice 1977a, b; Powell 1980; Bullen & Hede 1986a, b; Vogt 2005; Le Masurier et al. 2007). The noise exposure forecast (*NEF*), Australian variant (*ANEF*), and Noise and Number Index (*NNI*) are community noise metrics that also involve number-of-events terms, see, e.g., Bradley (1996). While such measures are used in some countries, other countries have moved to using metrics based on average A-weighted sound energy. In the ANASE study in the UK (Le Masurier et al. 2007) there was an argument made for including a number-of-events term based on a comparison of the data collected in that study to data collected earlier in the Aircraft Noise Index Study (Brooker et al. 1985), though there is some disagreement as to whether the data supported that conclusion (Brooker 2008a, b).

In an analysis of older community survey data, Fields (1984) examined the ratio of the estimated coefficients for the two types of terms (logged number-of-events over sound level) for different field studies; the ratios varied considerably from study to study, ranging from -3.7 to 23.8. Some of this variation can be attributed to different ranges of variables within particular studies and the use of different sound level metrics, as well as difference in the instrument used to measure annoyance, but even taking this into consideration, there does not seem to be a great deal of consistency between the estimated models from different field studies.

Community surveys have been the tool of choice in the formation of annoyance models but are also the validation tool for models proposed from the results of laboratory-based studies. However, there are challenges in estimating parameters in models containing both number-of-events and sound level terms from survey data. The number-of-events heard and the average sound level of events are often correlated because when events are louder, more are heard. If a number-of-events term is added to an energy-based term in a model, then, unless the mix of aircraft is changed at an airport, an increase in the number-of-events will lead to an increase in average sound level. As the degree of correlation between the variables changes (as it will from study to study) then the model parameter estimates will also vary. When, for a sample population, there is partial correlation between variables used in the model, it is impossible to separate out each variable's individual contribution to the output (annoyance). A more extreme form of this problem is called multicollinearity; this is when there is a high degree of correlation between variables (columns in the data matrix used for the estimation are close to being linearly related). This is a well-known problem in the estimation of regression models, see, e.g., Belsley (1991). The result is an ill-conditioned data matrix so that estimation results are very sensitive to perturbations in the data matrix. In addition, the variance of the estimates increases as the data matrix becomes more ill-conditioned. This is relevant because the variables are measured or predicted and thus, a certain level of uncertainty will be present in the data matrix.

The simulation tool has been developed to help determine if it is possible to sample a population around an airport in a manner that minimizes collinearity in the data matrix that is used to estimate annoyance model coefficients. The focus of the investigation reported here is on collinearity between acoustical variables, but demographic information is incorporated into the simulation tool obtained from US Census data (US Census Bureau 2000). Health data could also be incorporated in the future by using CDC data (CDC 2011).

SURVEY SIMULATION APPROACH

There are many possibilities in sample design, but the main focus of this work was to study the effect of stratification. The essence of stratification is the classification of a survey population into groups, or strata, based on available information about the population. Samples are then further selected from each of the strata. In practice, strata can be formed by using any available information about a population, both acoustical and non-acoustical. Only acoustical variables were explored as stratification variables in this work, but correlations between acoustical and non-acoustical variables in the population were examined.

The simulation tool was designed with the intent of representing possibilities at existing airports. Thus, it was required to select particular airports and associated communities on which to base the survey simulations. Here results from only one airport are given, but the approach would be the same for each airport included in a survey. Possibilities for future work include exploring whether including multiple airports in a single survey can be helpful in controlling collinearity. A description of the methodology applied in a survey simulation follows.

Acoustical environment specification

The Integrated Noise Model (INM) or Noise Model Simulation (NMSim) is used to generate acoustical output of single operations on a grid of points spaced at 0.1 nautical miles (nm). The grid of points is centred on the airport and extends outward in each cardinal direction (for the specific airport used as an example in this paper, the grid extended 10 nm outward). An "operation" is defined as one arrival or departure of a single aircraft operating at a certain flight profile on one flight path. Based on actual known airport operations, a realistic annual operations scenario is specified and used in conjunction with the stored acoustical output to generate sound level metrics at each location on the grid.

Demographic specification

U.S. Census 2000 data is used to specify the demographics of the population surrounding the airport. Both aggregate-level tabulations, reported in Summary File 3 (SF3) and 5 % public use microdata sample (PUMS) files are used in the demographic specification. The 5 % PUMS data are computerized versions of the census questionnaires, as coded and edited during census processing.

Combinatorial optimization (Huang & Williamson 2001; Ryan et al. 2009) is used to design the demographics of the population surrounding the airport; it is a method of selecting microdata records that best fit the aggregate-level tabulations for each census block given in SF3. A population is generated for the census block groups of the county containing the selected airport from microdata belonging to a larger geographical region that encompasses the airport.

Geographic specification of potential survey participants

Each grid point used in the acoustical environment specification has associated latitude and longitude. U.S. Census 2000 TIGER/Line shapefile data, or polygonal shape data, for the census block groups of the county containing the airport are collected. From an analysis of the spatial data, each of the grid points is associated with a census block group. The number of households in each census block group is compared with the number of acoustical grid points within it. Linear interpolation of the acoustical data grid is performed so that the number of acoustical grid points is larger than the number of households in each census block group. In each census block group, households (and associated individuals) are randomly assigned to acoustical grid points. Note that sound level metrics are converted to "mean square pressure or equivalent" before interpolation.

Annoyance specification

An average state of annoyance is defined for the population. In the example given in this paper, the average annoyance (A) due to aircraft noise at each grid point was specified using:

$$A = 0.069\, PNL_{max,av} + 0.22\, log_{10}\,(N) - 4.278, \qquad (1)$$

where $PNL_{max,av}$ is a logarithmic average of the Perceived Noise Level (*PNL*) of events above a threshold of 80 dB, and N is the annual number of events above that threshold. This is a model derived from a reanalysis of older aircraft noise survey data. Shown in Figure 1 is an annoyance map generated using Equation (1) for a particular flight operations scenario. Also shown on this map are the *Ldn* contours.

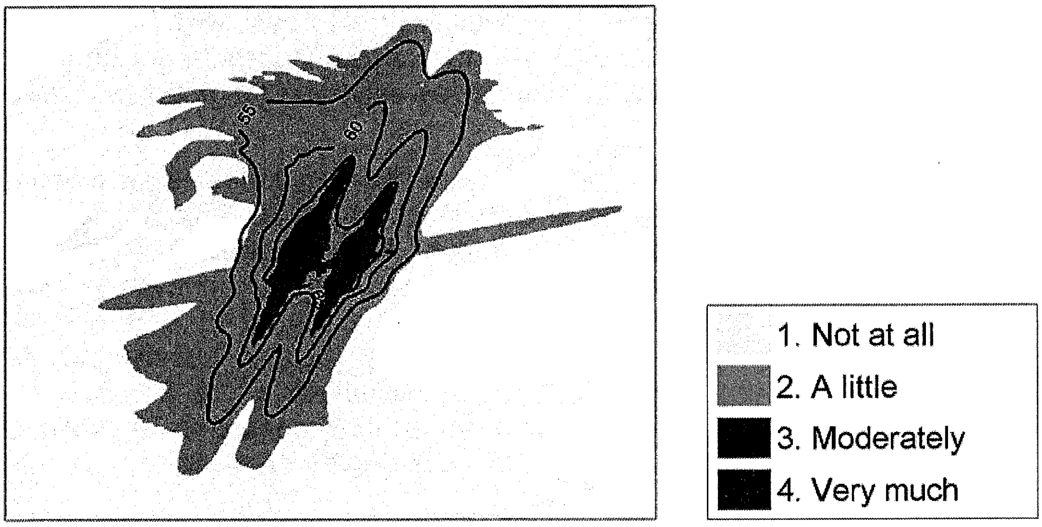

Figure 1: Annoyance contours developed using a model derived from a reanalysis of older survey data (Equation (1)). Also shown in red are the *Ldn* 55, 60, 65 and 70 dBA contours. This is based on an operations scenario that included 113,930 flights per year, random variation of 89 different flight tracks and 15 different aircraft.

SPECIFICATIONS FOR SIMULATIONS AND ANALYSIS

Scenarios for Monte Carlo simulation are created through combinations of values of different specification variables, those being the signal-to-noise ratio (R^2), or fraction of output variance attributable to the annoyance relationship in the population, the size (n) of the sample drawn from the total population for a noise survey, and the number of strata in classifying the population prior to sampling.

R^2 is controlled through varying the annoyance assigned to households from the model. For each household, annoyance values are assigned based on an assumption that annoyance scores would be distributed about the average value predicted from Equation (1), the variation being due to the many unmeasured things that affect responses at the time of filling in a questionnaire. The actual annoyance assigned to a grid point was thus: $A + na$ where na is a zero mean uniformly distributed random variable with the required variance: $var(na) = var(A)(1\text{-}R^2)/R^2$.

Choosing a single sample involves randomly selecting a household within a stratum, and then randomly selecting an adult (person at or above 18 years of age) residing in that household. Once a household is chosen, it is excluded from consideration, so that repeats are avoided. Thus, stratified random sampling without replacement, or simple random sampling without replacement (for the case of no stratification) is used to sample the population of grid points. Strata boundaries are determined so that each stratum contains a number of population households no less than two times the sample size used for a noise survey divided by the number of strata. If this is not possible for certain cases, the criterion is relaxed.

For each designed scenario, Monte Carlo simulation is used in which a noise survey is simulated for a pre-set number of trials. For each trial, or noise survey, acoustical variables of a model hypothesized to have generated the population annoyance, and the annoyance response generated as described above, are recorded for each of the sampled individuals. For the simulations used in this investigation the hypothesized model is the same as the simulated model. Non-acoustical variables describing population characteristics can also be recorded. For the simulations reported here, two socio-economic variables were recorded: the total household income and the age of the sampled individual. For each trial, two regressions were performed; these are referred to as "standard" (acoustical variables only) and "augmented" (acoustical and non-acoustical variables).

For the investigation simulations, levels of the specification variables were defined as follows. The values of R^2 chosen were 0.05, 0.20, and 0.40. Sample sizes for the simulated surveys were set at 500, 2,000, and 4,000. Four different levels of stratification were specified. Each predictor variable in the model of annoyance, $PNL_{max,av}$ and N, was divided into 2, 3, 4, or 5 groups each so that 4, 9, 16, and 25 strata, respectively, were realized. The case of no stratification was also investigated, yielding 5 total levels of stratification. Thus, through combinations of the 3 levels of R^2, 3 sample sizes, and 5 levels of stratification, a total of 45 scenarios for simulation were realized. 100 Monte Carlo trials were used for each survey simulation.

RESULTS OF FITTING MODELS TO THE ANNOYANCE "SURVEY" DATA

The parameter estimates from the 100 trials were analyzed and the biases and standard deviations for each of the parameter estimates were estimated. For the intercept and the coefficients of the acoustical variables the Type II (false negative) error rates were recorded (based on estimated 95% confidence intervals in each trial). For the socio-economic variables, which were not parameters in the simulated model, the Type I (false positive) error rates for each coefficient were calculated from the 100 augmented regression results. Also, for each trial of each simulation, the joint significance of the addition of the socio-economic variables to the standard model was assessed by calculating the marginal reduction in error sum of squares they provided. Aggregating the results over the trials for each simulation yielded an overall rate of significance (fraction of times the socio-economic variables were deemed significant contributors to the prediction of annoyance).

In the augmented models, the Type I error rates (false positives) observed for the socio-economic variables (income and age) ranged between 0.00 and 0.05, and the rates of joint significance for these variables ranged between 0.00 and 0.10. There were not significant trends for either the Type I error rates or joint significance rates with R^2, sample size, or stratification. Type II error rates (false negatives) for the acoustical variables in the standard model are shown in Figure 2. Slightly larger rates were found for the acoustical variables in the augmented regressions (not shown) but trends were very similar for both regressions.

A multicollinearity analysis was performed by using techniques outlined in Belsley (1991). The condition number of the data matrix in each trial of each simulation was calculated. The mean of the condition numbers for each 100-trial Monte Carlo simulation, along with percentile ranges, are shown in Figure 3. The mean of the 100 parameter estimates from each simulation and the corresponding standard deviations of the estimates, both normalized by their true values (see Equation (1)) are shown in Figure 4.

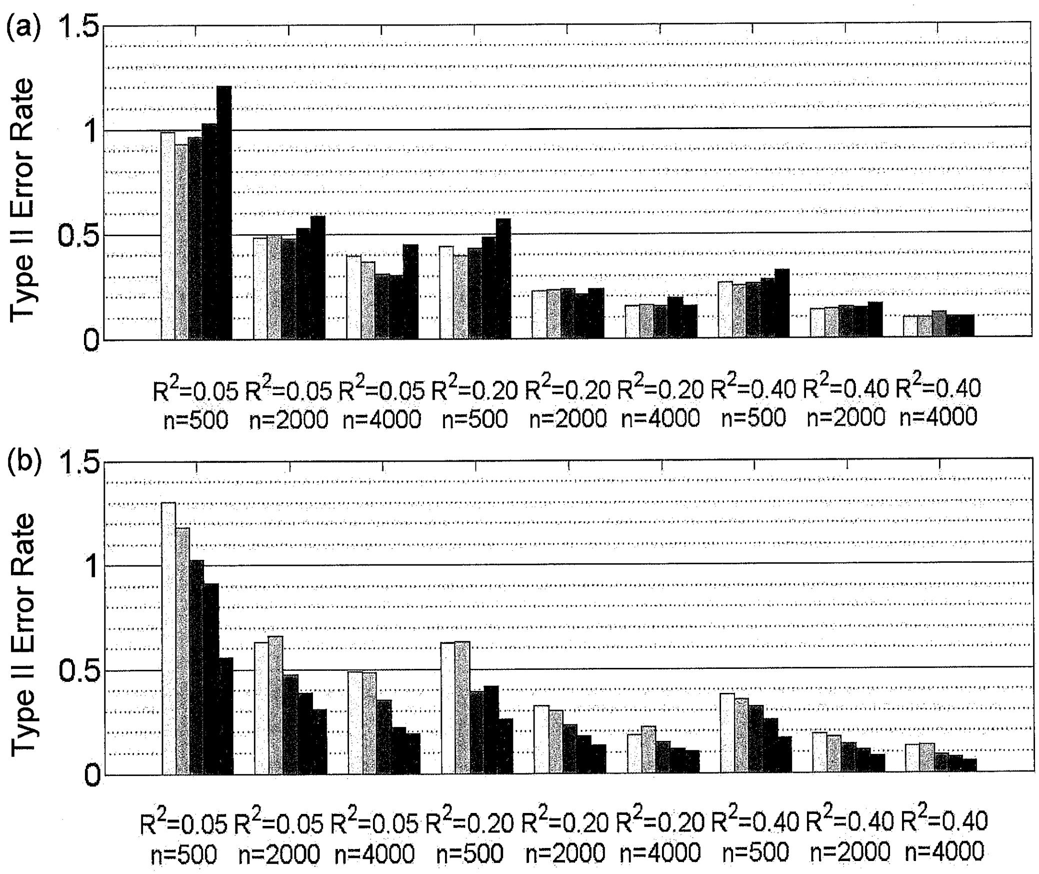

Figure 2: Type II error rates (false negatives) for the (a) number-of-events and (b) sound level terms from the standard regression fits. Shading (light to dark) indicates level of stratification (1, 4, 9, 16, 25 strata). Results from 100 trials in each simulation

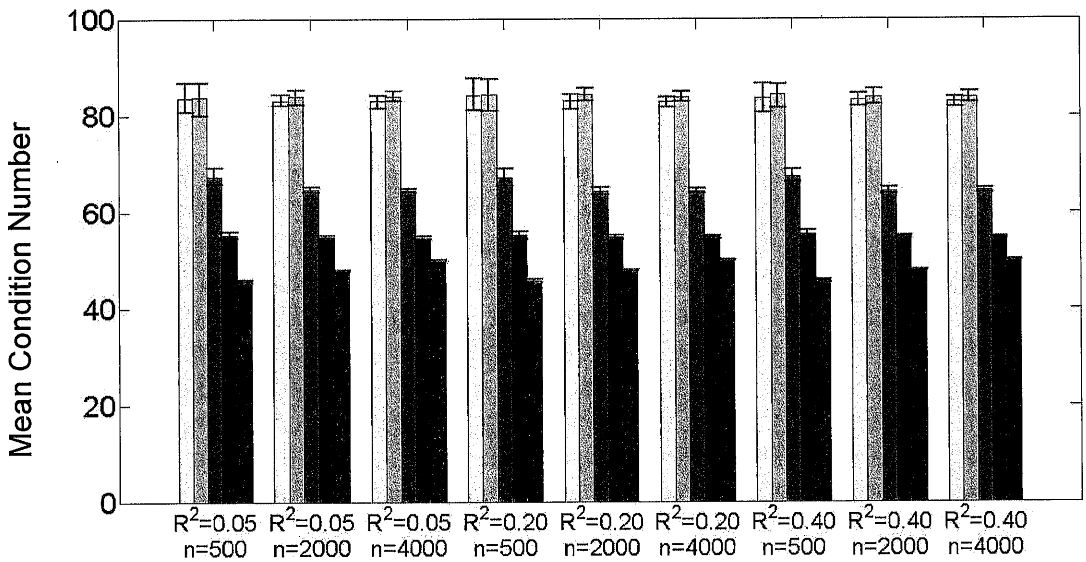

Figure 3: Mean condition number of data matrices for the 45 sets of Monte Carlo simulations organized as in Figure 2. Bars indicate the 25 % and 75 % percentiles.

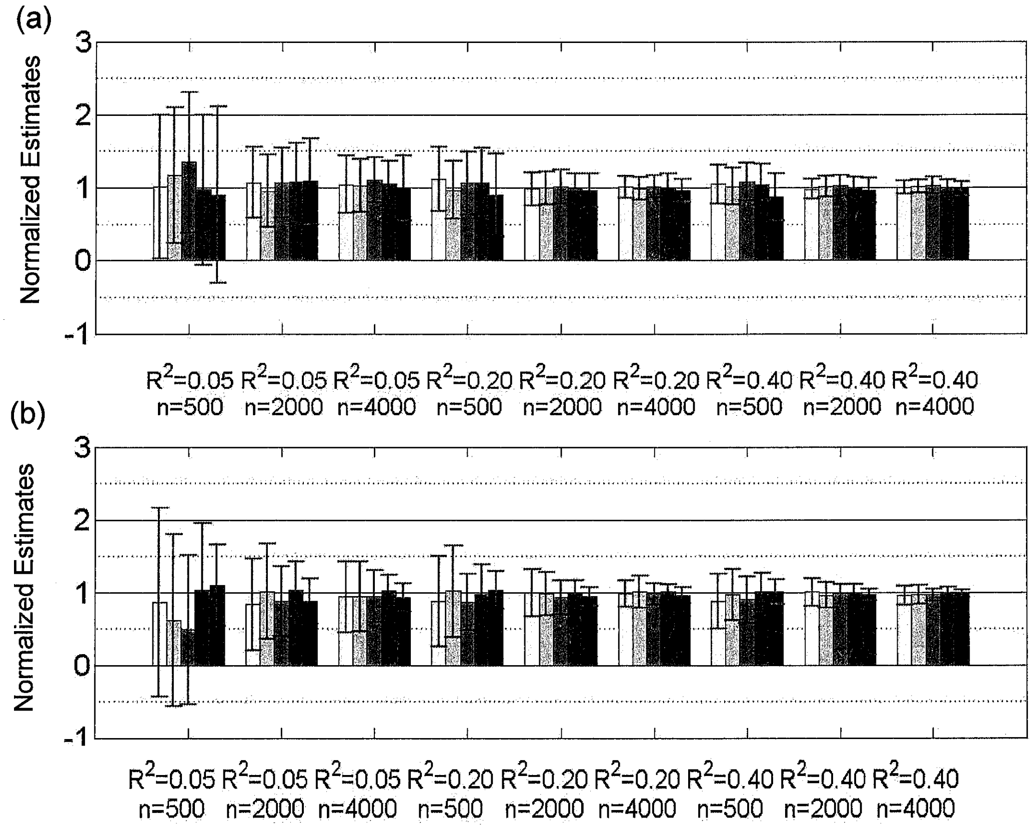

Figure 4: Parameter estimates and the standard deviation of the estimates from the 100 trials of each simulation, both normalized by the true values of the parameters: (a) the number-of-events coefficient (b) the sound level coefficient

Not surprisingly, simulations with smaller R^2 and smaller sample sizes resulted in higher standard deviation estimates. For the intercept and sound level term, increasing stratification resulted in a decrease in the standard deviation of the estimates and fewer Type II errors. For the number-of-events term, increased levels of stratification yielded little or no benefit and sometimes degraded the results. The conditioning of the data matrix improved with stratification.

SUMMARY

A methodology to simulate noise annoyance surveys around airports was described. It includes quantification of noise exposure by utilizing sound prediction software like INM or NMSim together with aircraft operations data from the airports. Socio-economic data is included in the simulation based on data from the U.S. Census Bureau. As an illustration, the simulation tool was used to conduct a series of Monte Carlo simulations to study the effects of stratification on the estimation of parameters in a linear regression annoyance model with sound level and number-of-events terms. Stratification was found to be beneficial (reduced variance) in the estimation of the coefficient of the sound level term, even at highest signal-to-noise ratio (R^2) and with the largest population sample (n). However, stratification was not generally found to be beneficial for the estimation of the coefficient of the number-of-events term (log_{10} (N)). This is surprising and warrants further investigation.

ACKNOWLEDGEMENTS

The authors also wish to thank the FAA/NASA/TC PARTNER Center of Excellence for their financial support. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the FAA, NASA or Transport Canada.

REFERENCES

Belsley DA (1991). Conditioning diagnostics: collinearity and weak data in regression. New York: John Wiley & Sons.

Berry WD, Feldman S (1994). Multiple regression in practice. (Series Quantitative Applications in the Social Sciences). Sage.

Bradley JR (1996). NEF validation study: (2) Review of aircraft noise and its effects. National Research of Canada Contract Report, A-1505.5.

Brooker P (2008a). ANASE: Unreliable – owing to design-induced biases. Jan/Feb, 26-31. https://dspace.lib.cranfield.ac.uk/bitstream/1826/2242/1/ANASE-Unreliable-2008.pdf.

Brooker P (2008b). ANASE: measuring aircraft noise annoyance very unreliable. Significance 5(1): 18-24.

Brooker P, Critchley JB, Monkman DJ et al. (1985). United Kingdom Aircraft Noise Index Study (ANIS): Main Report DR Report 8402, for Civil Aviation Authority (CAA) on behalf of the Dept. of Transport, CAA, London. http://www.caa.co.uk/docs/33/ERCD %208402.PDF.

Bullen RB, Hede AJ (1986a). Reaction to aircraft noise in residential areas around Australian airports. J Sound Vibr 108: 199-225.

Bullen RB, Hede AJ (1986b). Comparison of the effectiveness of measures of aircraft noise exposure by using social survey data. J Sound Vibr 108: 227-245.

CDC (2011). United States Centers for Disease Control and Prevention, Data Access Tools Website. http://www.cdc.gov/nchs/data_access/data_tools.htm.

Connor WK, Patterson HP (1976). Analysis of the effect of aircraft operations on community annoyance. National Aeronautics and Space Administration Technical Report NASA-CR-2741.

FAA (2010). FAA Aerospace Forecast Fiscal Years 2011-2013. U.S. Department of Transportation, Federal Aviation Administratiion, Aviation Policy and Plans, HQ-111316.

Fields JM (1984). The effect of numbers of noise events on people's reactions to noise - an analysis of existing survey data. J Acoust Soc Am 75: 447-467.

Huang Z, Williamson P (2001). A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Population Microdata Unit, Department of Geography, University of Liverpool, Working Paper 2001/2.

Le Masurier P, Bates J, Taylor J et al. (2007) Attitudes to noise from aviation sources in England (ANASE). U.K. Department of Transportation Final Report, C34351.

Powell CA (1980). Annoyance due to multiple airplane noise exposure. National Aeronautics and Space Administration Technical Report NASA-TP-1706.

Rice CG (1977a). Investigation of the trade-off effects of aircraft noise and number. J Sound Vibr 52: 325-344.

Rice CG (1977b). Development of cumulative noise measure for the prediction of general annoyance in an average population. J Sound Vibr 52: 345-364.

Ryan J, Maoh H, Kanaroglou P (2009). Population synthesis: comparing the major techniques using a small, complete population of firms. Geographical Anal 41: 181–203.

Rylander R, Bjorkman M, Ahrlin U et al. (1980). Aircraft noise annoyance contours: importance of overflight frequency and noise level. J Sound Vibr 69: 583-595.

TRACOR Inc. (1970). Community reaction to airport and aircraft noise, Vol. II. National Aeronautics and Space Administration Technical Report NASA-CR-111316.

TRACOR Inc. (1971). Community reaction to airport and aircraft noise, Vol. I. National Aeronautics and Space Administration Technical Report NASA-CR-1761.

US Census Bureau (2000). Main Data Website, 2000 Census Data Portal and 2010 Census Data Portal. http://www.census.gov/main/, http://www.census.gov/main/www/cen2000.html, http://2010.census.gov/2010census/data/, 2011.

Vogt J (2005). The relative impact of aircraft noise and number in a full-factorial laboratory design. J Sound Vibr 282: 1085-1100.