



Partnership for AiR Transportation
Noise and Emissions Reduction
An FAA/NASA/Transport Canada-
sponsored Center of Excellence



The Number-of-Events as a Predictor Variable in Aircraft Noise Annoyance Models

A PARTNER Project 24 Report

prepared by
Kevin Foertsch, Patricia Davies

December 2013

REPORT NO. PARTNER-COE-2013-002

The Number-of-Events as a Predictor Variable in Aircraft Noise Annoyance Models

A PARTNER Project 24 Report

Kevin Foertsch, Patricia Davies

PARTNER-COE-2013-002

November 2013

This work is funded by the US Federal Aviation Administration Office of Environment and Energy under FAA Award No. 07-C-NE-PSU, Amendment Nos. 006, 007, and 012, Award No. 07-C-NE-PU, Amendment Nos. 013 and 028, and Award No. 09-C-NE-PU, Amendment Nos. 001, 008, 009, 014, and 019. The project was managed by Mehmet Marsan of the FAA.

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, Transport Canada, the U.S. Department of Defense, or the U.S. Environmental Protection Agency

The Partnership for AiR Transportation Noise and Emissions Reduction — PARTNER — is a cooperative aviation research organization, and an FAA/NASA/Transport Canada-sponsored Center of Excellence. PARTNER fosters breakthrough technological, operational, policy, and workforce advances for the betterment of mobility, economy, national security, and the environment. The organization's operational headquarters is at the Massachusetts Institute of Technology.

**The Partnership for AiR Transportation Noise and Emissions Reduction
Massachusetts Institute of Technology, 77 Massachusetts Avenue, 33-240
Cambridge, MA 02139 USA
<http://partner.mit.edu>**

**THE NUMBER-OF-EVENTS AS A PREDICTOR
VARIABLE IN AIRCRAFT NOISE ANNOYANCE
MODELS**

Sponsored by
FAA/NASA/Transport Canada PARTNER Center of Excellence

HL 2013-1

Submitted by: Kevin Foertsch, Graduate Research Assistant
Patricia Davies, Principal Investigator

Approved by: Patricia Davies, Director
Ray W. Herrick Laboratories

MAY 2013

ACKNOWLEDGEMENTS

Opinion, findings, conclusions and recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of PARTNER Center of Excellence sponsor organizations.

This work was done under the supervision of Patricia Davies, Professor of Mechanical Engineering, Purdue University. This research would not have been possible without the help of Dr. James Fields who supplied past surveys data and spent time discussing the research with the investigators.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	xv
NOMENCLATURE	xviii
ABSTRACT	xxi
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. LITERATURE REVIEW	4
2.1 Alternative Annoyance Models	5
2.2 Annoyance Model Estimation and Sample Design	7
2.3 Holistic Annoyance Models	10
2.4 Conclusions	13
CHAPTER 3. NOISE SURVEY DATA SETS	15
3.1 Sound Level Metrics	18
3.1.1 DNL to $PNL_{max,av}$	19
3.2 L_{av} to $PNL_{max,av}$	27
3.3 Simulation of UKD-604	28
3.4 Data Visualization: Plotting Program	35
3.5 Data Set Combination	40
3.5.1 Annoyance Response Combination	40
3.6 Summary	42
CHAPTER 4. METHODOLOGIES IN MODEL DEVELOPMENT AND PARAMETER ESTIMATION	44
4.1 Individual Data Set Results and Methodologies	44
4.1.1 Relative Variable Importance	46
4.1.2 Decibel Equivalent Number Effect, \hat{k}	51
4.1.3 Standard Deviation of \hat{k} , $\hat{\sigma}_{\hat{k},c}$	52
4.1.4 Average Semipartial R^2 ($\overline{R^2}_{sp}$)	59
4.1.5 Comparison of Importance Measures	66
4.1.6 Continuous and Aggregate Data	67
4.2 Combination Data Set Results	72
4.2.1 Decibel Equivalent Number Effect, \hat{k}	73
4.2.2 Standard Deviation of \hat{k} , $\hat{\sigma}_{\hat{k},c}$	74

	Page
4.2.3 Average Semipartial R^2	75
4.2.4 Continuous and Aggregate Data	75
4.2.5 The Effect of Airport	79
4.3 Summary	85
CHAPTER 5. SURVEY SIMULATION	87
5.1 Sample Design	87
5.2 Survey Simulation Program	90
5.2.1 Single-Event Acoustical Data	91
5.2.2 Noise Metric Module	93
5.2.3 Synthetic Population Generator	101
5.2.4 Survey Sampling Module	108
5.3 Example Survey Simulation: Design	126
5.3.1 Survey Design Decisions	126
5.3.2 Survey Sampling Decisions	127
5.3.3 Survey Analysis Decisions	131
5.4 Example Survey Simulation: Results	132
5.4.1 Validation of Survey Simulation Program	132
5.4.2 Effects of Rescaling and Curtailing Annoyance Data	134
5.4.3 Results of Survey Simulations	135
5.4.4 Annoyance Model Comparisons	153
5.5 Summary	156
CHAPTER 6. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK	162
6.1 Summary and Conclusions	163
6.2 Recommendations for Future Work	167
LIST OF REFERENCES	170
APPENDICES	
APPENDIX A. ADDITIVE-LOG ANNOYANCE MODELS AND SOUND LEVEL METRICS	179
A.1 Additive-Log Annoyance Models	179
A.2 Sound Level Metrics	189
A.2.1 Definition of the Perceived Noise Level (PNL) and Related Metrics	190
A.2.2 Definition of the Sound Exposure Level (SEL_A) and Related Metrics	191
APPENDIX B. VARIANCE ESTIMATION	193
B.1 Jackknife Repeated Replication	195
B.2 Taylor Series Linearization	197
APPENDIX C. MULTICOLLINEARITY	204

	Page
APPENDIX D. STANDARDIZED STORAGE FORMAT	207
APPENDIX E. FLIGHT TRACK CREATION	218
E.1 Flight Track Maps	218
E.2 Flight Track Data	222
APPENDIX F. INSTRUCTIONS FOR CREATING A SYNTHETIC POPU- LATION	226
APPENDIX G. TRANSFORMING .UF3 TO .TXT	231
APPENDIX H. SURVEY SIMULATION PROGRAM INFORMATION . . .	232
H.1 Folder Structure and Program Files	232
H.2 Text Files	233
H.2.1 Basic specification	234
H.2.2 Variable specification	234
H.2.3 Annoyance model specification	236
H.2.4 Analyses specification	236
APPENDIX I. COMPUTER PROGRAM DESCRIPTIONS	238
I.1 Noise Metric Module	238
I.2 Aircraft Operations Scenario Analysis	244
I.3 Airport 3 Aircraft Operations Scenario Adjustment	245
I.4 Aircraft Operations Scenario Creation from INM Output	246
I.5 Aircraft Model and Flight Track Contributions	247
I.6 Variation of Noise Exposure within Census Block Groups	248
I.7 Synthetic Population Generator	249
I.8 Survey Sampling Module	254
I.9 Variance of a Ratio of Regression Coefficients	261
I.10 Multicollinearity	264
I.11 Extra Sum of Squares	266
I.12 Census Block Group Identification	267
I.13 Data Set Combination	268
I.14 UKD-604 Data Set Simulation	274
I.15 AUL-210 Data Set Simulation	276
I.16 Plotting Program	278
I.17 Flight Track Creation	281
APPENDIX J. SURVEY SIMULATION RESULTS	286

LIST OF TABLES

Table	Page
2.1 Correlation ($\rho_{L, \log_{10}(N)}$) between the sound level and $\log_{10}(N)$ variables for survey data sets, where N is the number-of-events. The sound level variable (L) and the sound level threshold (L_t) for counting number-of-events are given. $PNL_{max,av}$ is the average Maximum Perceived Noise Level. MNL is the Maximum A-weighted Sound Level (L_{Amax}) of the noisiest aircraft regularly using the airfield at least 3 times per day. L_{av} is the arithmetic average Maximum A-weighted Sound Level. Data sets with site-level data are notated with the superscript *. Data sets that resulted from the use of stratification by both sound level and number-of-events are notated with the superscript S	9
3.1 Information about noise survey data sets. $PNL_{max,av}$ is the average Maximum Perceived Noise Level of aircraft events measured in PNdB and N is the average daily number of aircraft events. $PNL_{max,av}$ is described in Appendix A.2.1.	16
3.2 Multiple linear regressions results for ten transformations of DNL to $PNL_{max,av}$. The regression model derived in each case was: $A = b_L PNL_{max,av} + b_N \log_{10}(N_T) + b_C$, where A is a 5-point verbal annoyance scale. $\hat{k} = \hat{b}_N / \hat{b}_L$ is referred as the “decibel equivalent number effect”.	25
3.3 Observed trends in the average standard deviation of L_{av} ($\overline{\sigma_{L_{av}}}$) and the average normalized standard deviation of N_{av} ($\overline{\sigma_{N_{av}}^*}$).	32
3.4 Information about the UKD-604 Output Areas. For each Output Area is given the airport, L_{av} , N_{av} , the assigned average standard deviation of L_{av} ($\overline{\sigma_{L_{av}}}$), and the assigned average normalized standard deviation of N_{av} ($\overline{\sigma_{N_{av}}^*}$). n is the sample size.	32
4.1 Basic information about the individual data sets, separated by airport.	44
4.2 Individual data set results for models of annoyance that are a linear combination of $PNL_{max,av}$ and $\log_{10}(N)$. $DEFT$ is the Design Factor and the estimated 95% confidence intervals are given for k . Standard deviation estimates of the regression coefficient estimates are given in parentheses.	47
4.3 Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the individual data sets.	62

Table	Page
4.4 Multiple linear regression results for the UKD-604 data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). The 95% CI (confidence interval) is for k . Standard deviation estimates of the regression coefficient estimates are given in parentheses. UKD-604* is the first UKD-604 data set utilizing variation in sound level and number-of-events. UKD-604** is the second UKD-604 data set with further variation applied to sound level and number-of-events. UKD-604 is the third UKD-604 data set which was created only from the reported data and does not include variation in sound level and number-of-events.	65
4.5 Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the UKD-604 data sets. UKD-604* is the first UKD-604 data set utilizing variation in sound level and number-of-events. UKD-604** is the second UKD-604 data set with further variation applied to sound level and number-of-events. UKD-604 is the third UKD-604 data set which was created only from the reported data and does not include variation in sound level and number-of-events.	65
4.6 $PNL_{max,av}$ group boundaries for the survey data sets in the aggregate data analyses.	69
4.7 $\log_{10}(N)$ boundaries for the survey data sets in the aggregate data analyses.	69
4.8 Multiple linear regression results from the aggregate data analyses. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. n_g is the number of aggregate groups. Standard deviation estimates of the regression coefficient estimates are given in parentheses. ¹ The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$	70
4.9 Multiple linear regression results from the unaggregated data analyses. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). $DEFT$ is the Design Effect for \hat{k} . The 95% CI (confidence interval) is for k . n is the sample size. Standard deviation estimates of the regression coefficient estimates are given in parentheses.	71

Table	Page
4.10 Average semipartial R^2 values $\left(\overline{R_{sp}^2}\right)$ and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ in the unaggregated data analysis.	72
4.11 Labels and annoyance scale modification descriptions for the combination data sets.	73
4.12 Multiple linear regression results from the combination data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). <i>DEFT</i> is the Design Effect for \hat{k} . The 95% CI (confidence interval) is for k . Descriptions of the combination data sets can be found in Table 4.11. The predictor variable data are the same for all combination data sets.	76
4.13 Average semipartial R^2 values $\left(\overline{R_{sp}^2}\right)$ and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the combination data sets. The predictor variable data are the same for all combination data sets.	77
4.14 $PNL_{max,av}$ and $\log_{10}(N)$ group boundaries for the combination data sets in the aggregate data analysis.	77
4.15 Multiple linear regression results from the aggregate data analyses of the combination data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. Average semipartial R^2 values $\left(\overline{R_{sp}^2}\right)$ are given. n_g is the sample size. The predictor variable data are the same for all combination data sets.	78
4.16 Changes in jackknife replicate estimates of k $\left(\hat{k}_r\right)$ as airports that cause \hat{k} to decrease the most are removed from the full data set. The estimated 95% confidence intervals for k for the data sets without the airports in the column headings are given at the bottom of the table. The replicate estimates are those from the Combination #1 data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: <i>A</i> denotes UKD-024/Heathrow (LHR), <i>B</i> denotes USA-022/Chicago (ORD), and <i>C</i> denotes USA-022/Denver (DEN).	81

Table	Page
4.17 Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the midpoint Miedema scaling combination data set (Combination #1) as airports that cause \hat{k} to decrease the most are removed from the full data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-024/Heathrow (LHR), B denotes USA-022/Chicago (ORD), and C denotes USA-022/Denver (DEN).	82
4.18 Changes in jackknife replicate estimates of k (\hat{k}_r) as airports that cause \hat{k}_r to increase the most are removed from the full data set. The estimated 95% confidence intervals for k for the data sets without the airports in the column headings are given at the bottom of the table. The replicate estimates are those from the Combination #1 data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-604/Heathrow (LHR), B denotes USA-032/Boston (BOS), and C denotes AUL-210/Perth (PER).	83
4.19 Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the midpoint Miedema scaling combination data set (Combination #1) as airports that cause \hat{k} to increase the most are removed from the full data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-604/Heathrow (LHR), B denotes USA-032/Boston (BOS), and C denotes AUL-210/Perth (PER).	84
5.1 Aircraft models used at Airport 1.	92
5.2 Aircraft models used at Airport 2.	93
5.3 Aircraft models used at Airport 3.	94
5.4 Number of household and person microdata records used and the average number of households and persons per census block group at the three airports.	107
5.5 The average, minimum, and maximum number of iterations, i , for the census block groups of the three airports.	107
5.6 Statistical hypothesis test decision matrix. H_0 is the null hypothesis: The i^{th} predictor variable p_i is not in the population mean annoyance model. The alternative hypothesis (H_A) is that the i^{th} predictor variable p_i is in the population mean annoyance model. α is the Type I error rate and β is the Type II error rate.	119

Table	Page
5.7 Strata for Airport 1. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.	136
5.8 Strata for Airport 2. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.	137
5.9 Strata for Airport 3. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.	138
5.10 Lower alternative values of \hat{k} that can be stated to be significantly different from $\hat{k} = 10$ with a 95% of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error.	139
5.11 Upper alternative values of \hat{k} that can be stated to be significantly different from $\hat{k} = 10$ with a 95% of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error.	140
5.12 Multiple linear regression results for one Monte Carlo trial around Airport 3 for three different levels of R_n^2 (0.2, 0.8, and 1.0) and thus, annoyance modifications. $\hat{\sigma}_{\hat{k},c}$ is the jackknife standard deviation estimate of \hat{k} . The 95% CI (confidence interval) is for k . The standard deviation estimates of the regression coefficient estimates are given in parentheses. ¹ The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$	140

Table	Page
<p>5.13 Multiple linear regression results for Equation (5.31) estimated using the three methods of adjusting annoyance; the data of Airport 2 were used. $\overline{\hat{b}_L}$ is the mean of the estimates of b_L, likewise for the other regression coefficient estimates; the standard deviations are given in parentheses. The 95% CI (confidence interval) is for the expected value of \hat{k} ($E[\hat{k}]$) and is calculated from $\overline{\hat{k}}$ and $\hat{\sigma}_{\hat{k}}$. $\overline{R^2}$ is the mean of the R^2's from the multiple linear regressions; the interquartile range is given in parentheses. For each trial, the predictor variable data are the same for all adjustment methods; the only variable that changes is the annoyance variable. "Normal" is the adjustment applied with Equation (3.26), "DNL cutoff" is the adjustment applied with the use of <i>DNL</i> data, and "Scale cutoff" is the adjustment applied by limiting the annoyance values at the extremities.</p>	141
<p>5.14 Multiple linear regression results for estimates of Equation (5.31). $\overline{\hat{b}_L}$ is the mean of the estimates of b_L, likewise for the other regression coefficient estimates; the standard deviations are given in parentheses. $\overline{\hat{k}}$ is the mean of the 100 estimates of \hat{k}; the range of \hat{k} is given in parentheses. The 95% CI (confidence interval) is for the expected value of \hat{k} ($E[\hat{k}]$) and is calculated from $\overline{\hat{k}}$ and $\hat{\sigma}_{\hat{k}}$. $\overline{\rho_{L, \log_{10}(N)}}$ is the mean correlation between L ($PNLT_{max,av,80}$) and $\log_{10}(N)$ ($\log_{10}(N_{80})$); the interquartile range is given in parentheses. $\overline{R^2}$ is the mean of the R^2's from the multiple linear regressions; the interquartile range is given in parentheses.</p>	144
<p>5.15 Multiple linear regression results for estimates of Equation (5.32). $\overline{\hat{b}_D}$ is the mean of the estimates of b_D, likewise for the other regression coefficient estimate; the standard deviations are given in parentheses. $\overline{R^2}$ is the mean of the R^2's from the multiple linear regressions; the interquartile range is given in parentheses. ¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.</p>	146
<p>5.16 Averages of average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ of Equation (5.31). "Full S" denotes the 6000-person sample resulting from stratified sampling of the population, "Full R" denotes the 6000-person sample resulting from simple random sampling of the population, "RSD" denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and "RR" denotes the 2000-person sample extracted from the Full S sample using simple random sampling.</p>	147

Table	Page
<p>5.17 Ranges of the scaled $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data in a typical sample. From Equation (5.26), 0.08 is the scaling applied to $PNLT_{max,av,80}$ and 1.12 is the scaling applied to $\log_{10}(N_{80})$. “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling.</p>	149
<p>5.18 Results of the retrospective power analysis estimated using the jackknife estimate of $\hat{\sigma}_{\hat{k}}$ in each Monte Carlo trial. “FT S.D. 10” is the fraction of Monte Carlo trials in which \hat{k} was concluded with 95% confidence to be significantly different from the equal-energy value ($k = 10$). $\bar{\hat{k}}_{0.80}$ is the mean upper alternative value of \hat{k} distinguishable from the equal-energy value with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, where $P = 0.80$; the interquartile range is also given in parentheses. $\bar{\hat{\sigma}}_{\hat{k}}$ is the mean jackknife standard deviation of \hat{k} over the 100 Monte Carlo trials. “FT $\bar{\hat{k}}_{0.80} \geq 12.36$” is the fraction of Monte Carlo trials in which the upper alternative value calculated was greater than the value from the prospective power calculations ($\hat{k} = 12.36$).</p>	151
<p>5.19 Results of the retrospective power analysis estimated using the Taylor estimate of $\hat{\sigma}_{\hat{k}}$ in each Monte Carlo trial. “FT S.D. 10” is the fraction of Monte Carlo trials in which \hat{k} was concluded with 95% confidence to be significantly different from the equal-energy value ($k = 10$). $\bar{\hat{k}}_{0.80}$ is the mean upper alternative value of \hat{k} distinguishable from the equal-energy value with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, where $P = 0.80$; the interquartile range is also given in parentheses. $\bar{\hat{\sigma}}_{\hat{k}}$ is the mean jackknife standard deviation of \hat{k} over the 100 Monte Carlo trials. “FT $\bar{\hat{k}}_{0.80} \geq 12.36$” is the fraction of Monte Carlo trials in which the upper alternative value calculated was greater than the value from the prospective power calculations ($\hat{k} = 12.36$).</p>	152

Table	Page
5.20 Average difference (E) between the predicted household annoyance and the annoyance generated by the population mean annoyance model. Each cell represents the average difference in predicted household annoyance, calculated by using Equation (5.36). “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling. “N&L” and “ DNL ” denote the model predictions from the mean estimates of Equations (5.32) and (5.32), respectively.	155
D.1 Column assignments for the standardized storage format.	207
H.1 Text file names, sections that they are described in, and the modules that use the text files. “Y” indicates that the module uses the text file. “N” indicates that the module does not use the text file.	233
J.1 Simulation-level statistics for \hat{b}_L	287
J.2 Simulation-level statistics for \hat{b}_N	287
J.3 Simulation-level statistics for \hat{b}_{01}	288
J.4 Simulation-level statistics for \hat{b}_D	288
J.5 Simulation-level statistics for \hat{b}_{02}	289

LIST OF FIGURES

Figure	Page	
3.1	Overflight A-weighted mean square pressure profile (blue) and decibel (referenced to 20 μ Pa) profile (black) for: (a) Boeing 737, (b) Boeing 747, (c) Boeing 757, and (d) MD80. Each blue overflight profile shape can be well-approximated by a triangle, shown in red; the height of the triangle represents the A-weighted maximum mean square pressure of a single aircraft event ($p_{A,max}^2$). The height of the green rectangle represents the A-weighted equivalent mean square pressure of a single aircraft event ($p_{A,eq}^2$). The blue, red, and green pressure profiles all contain the same amount of energy, or enclose the same amount of area. The duration of each triangular-approximated event is noted. An averaging time of 3 seconds was used to construct the profiles.	22
3.2	Noise metrics calculated using the Noise Metric Module. Each plot contains 650,000 total points from noise metric calculations for simulated operations around three airports. The red lines are linear regression models derived from the linear ranges of the relationships: (a)&(b) <i>DNL</i> 10-80, (c) <i>DNL</i> 10-70, (d)&(e) <i>SEL_{A,av}</i> 20-80, (f) <i>L_{Amax,av}</i> 0-70. The estimated linear relationships are: (a) $SEL_{A,av} = 0.99 DNL + 14.5$, (b) $L_{Amax,av} = 1.04 DNL - 0.3$, (c) $PNL_{max,av} = 1.1 DNL + 5.2$, (d) $L_{Amax,av} = 1.0 SEL_{A,av} - 15.1$, (e) $PNL_{max,av} = 1.1 SEL_{A,av} - 11.1$, (f) $PNL_{max,av} = 1.1 L_{Amax,av} + 5.5$	26
3.3	Cumulative metrics predicted from other cumulative metrics plotted against those directly calculated from individual event data using the Noise Metric Module (Chapter 5, Section 5.2.2 and Appendix I.1). Each plot contains 650,000 points from single event predictions and simulated airport operations around three US airports. The red reference line has a slope of 1. (a) <i>SEL_{A,av}</i> predicted from <i>DNL</i> . (b) <i>L_{Amax,av}</i> predicted from the predicted <i>SEL_{A,av}</i> . (c) <i>PNL_{max,av}</i> predicted from the predicted <i>L_{Amax,av}</i>	27
3.4	Flowchart of the UKD-604 person-level data set simulation process.	30
3.5	Illustration of factor levels and treatments. The continuous variables <i>L_{Amax}</i> and the number-of-events are each divided into factor levels. Combinations of factor levels form treatments.	36

Figure	Page	
3.6	Each plot shows average annoyance vs. average Maximum Perceived Noise Level ($PNL_{max,av}$). Separate lines in each plot highlight different number-of-events groups. Each line is marked by a certain shade of blue. Darker shades of blue indicate larger number-of-events groups. The progression of number-of-events groups is: (a) 0-24, 25-49, 50-99, 100-199, 200-488; (b) 0-49, 50-99, 100-199, 200-399, 400-504; (c) 0-34, 34-68, 68-102, 102-134.	39
5.1	Flight tracks for (a) Airport 1, (b) Airport 2, and (c) Airport 3. Grid points are spaced by 1 nautical mile. Red tracks are arrival flight tracks, blue tracks are departure flight tracks.	95
5.2	Recommended (a) verbal and (b) numerical annoyance scales (Fields et al., 2001).	110
5.3	Census block groups surrounding (a) Airport 1, (b) Airport 2, and (c) Airport 3. Runways are shown in black.	114
5.4	Simulated household locations (shown in red) for the populations surrounding (a) Airport 1, (b) Airport 2, and (c) Airport 3. Roads (in blue) and runways (in black) are also shown.	115
5.5	Simulated household locations (shown in red) for the population surrounding Airport 3 prior to the use of water polygon shapefiles. Roads (in blue) and runways (in black) are also shown.	116
5.6	Census block groups (orange) and water polygons (blue) in the vicinity of Airport 3. Runways are shown in black.	116
5.7	Superposition of two hypothetical distributions for \hat{k} , each with a standard deviation of 1.5. The black curve represents the distribution of \hat{k} under a null hypothesis ($k = 10$). The red curve represents the distribution of \hat{k} under an alternative hypothesis ($k = 14$). The blue regions denote the $\frac{\alpha}{2}$ probability regions. The green region denotes the Type II error region, and contains a probability of β . Here, $\alpha = 0.05$, $\beta = 0.397$, and $P = 1 - 0.397 = 0.603$	120
5.8	Strata boundaries for (a) Airport 1, (b) Airport 2, and (c) Airport 3. Combinations of $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ for households in the population are shown by the black dots. The strata boundaries are shown by the red lines. The correlation between the $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data for the entire population around the airport is 0.59, 0.71, and 0.49 for Airports 1, 2, and 3, respectively.	129
5.9	Household locations (shown in red) for representative samples of the populations around (a) Airport 1, (b) Airport 2, and (c) Airport 3. Roads (in blue) and runways (in black) are also shown.	133

Figure	Page
5.10 Annoyance model error plots for Airport 1. The error plots in (a) and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; and (e) and (f), RR estimates. The error plots in the left column ((a), (c), and (e)) are from using the N&L model estimates and in the right column ((b), (d), and (f)), <i>DNL</i> model estimates. The resolution of the plots is 0.125 nmi.	157
5.11 Annoyance model error plots for Airport 2. The error plots in (a and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; in (e) and (f), RR estimates; in (g) and (h), Full R estimates. The error plots in the left column ((a), (c), (e), and (g)) are from using the N&L model estimates and in the right column ((b), (d), (f), and (h)), <i>DNL</i> model estimates. The resolution of the plots is 0.125 nmi.	158
5.12 Annoyance model error plots for Airport 3. The error plots in (a) and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; and (e) and (f), RR estimates. The error plots in the left column ((a), (c), and (e)) are from using the N&L model estimates and in the right column ((b), (d), and (f)), <i>DNL</i> model estimates. The resolution of the plots is 0.125 nmi.	159
6.1 (a) Individual data set \hat{k} plotted against the year of the survey. (b) The \hat{k} of a combination data set, where one \hat{k} results from removing an individual data set from the combination data set, plotted against the maximum average daily number-of-events (N_{max}). There is one outlier in (b), shown in red.	165

NOMENCLATURE

$CNEL$	Community Noise Equivalent Level
CNR	Composite Noise Rating
DNL	Day-Night Average Sound Level
$DENL$	Day-Evening-Night Average Sound Level
$EPNL$	Effective Perceived Noise Level
$EPNL_{av}$	Average (energy-based) Effective Perceived Noise Level
k	Decibel equivalent number effect
L_{Amax}	Maximum A-weighted Sound Level
$L_{Amax,av}$	Average (energy-based) Maximum A-weighted Sound Level
$L_{Amax,av,70}$	Average (energy-based) Maximum A-weighted Sound Level of aircraft events with $L_{Amax} > 70$ dBA
L_{av}	Average (arithmetic) Maximum A-weighted Sound Level
L_{Aeq}	24-hr A-weighted Equivalent Sound Level
$L_{Aeq,d}$	A-weighted Equivalent Sound Level for the daytime (07:00-22:00)
$L_{Aeq,n}$	A-weighted Equivalent Sound Level for the nighttime (22:00-07:00)
$L_{Aeq,0607}$	A-weighted Equivalent Sound Level for a 1-hr early morning period (06:00-07:00)
$L_{Aeq,0719}$	A-weighted Equivalent Sound Level for a 12-hr daytime period (07:00-19:00)
$L_{Aeq,1922}$	A-weighted Equivalent Sound Level for a 3-hr evening period (19:00-22:00)
$L_{Aeq,2223}$	A-weighted Equivalent Sound Level for a 1-hr late evening period (22:00-23:00)

$L_{Aeq,2306}$	A-weighted Equivalent Sound Level for a 7-hr nighttime period (23:00-06:00)
$L_{Aeq,diff}$	Difference between L_{Aeq} at the most and least exposed side of the building
L_{eq}	24-hr Equivalent Sound Level
$L_{eq,d}$	Equivalent Sound Level for the daytime (07:00-22:00)
$L_{eq,n}$	Equivalent Sound Level for the nighttime (22:00-07:00)
N	Number of aircraft events
N_{70}	Number of aircraft events with $L_{Amax} > 70$ dBA
N_{80}	Number of aircraft events with $PNLT_{max} > 80$ PNdB
$N_{PNL_{max,av},80}$	Number of aircraft events with $PNL_{max} > 80$ PNdB
$N_{PNLT_{max,av},80}$	Number of aircraft events with $PNLT_{max} > 80$ PNdB
$N_{EPNL_{av}}$	Number of aircraft events contributing to $EPNL_{av}$
$N_{SEL_{A,av}}$	Number of aircraft events contributing to $SEL_{A,av}$
$N_{SENEL_{A,av}}$	Number of aircraft events contributing to $SENEL_{A,av}$
$N_{L_{Aeq}}$	Number of aircraft events contributing to L_{Aeq}
$N_{L_{Aeq,d}}$	Number of aircraft events contributing to $L_{Aeq,d}$
$N_{L_{Aeq,n}}$	Number of aircraft events contributing to $L_{Aeq,n}$
$N_{L_{Aeq,0607}}$	Number of aircraft events contributing to $L_{Aeq,0607}$
$N_{L_{Aeq,0719}}$	Number of aircraft events contributing to $L_{Aeq,0719}$
$N_{L_{Aeq,1922}}$	Number of aircraft events contributing to $L_{Aeq,1922}$
$N_{L_{Aeq,2223}}$	Number of aircraft events contributing to $L_{Aeq,2223}$
$N_{L_{Aeq,2306}}$	Number of aircraft events contributing to $L_{Aeq,2306}$
$N_{L_{eq}}$	Number of aircraft events contributing to L_{eq}
$N_{L_{eq,d}}$	Number of aircraft events contributing to $L_{eq,d}$
$N_{L_{eq,n}}$	Number of aircraft events contributing to $L_{eq,n}$
N_{av}	Number-of-events with L_{Amax} above 65 dBA
NEF	Noise Exposure Forecast
NNI	Noise and Number Index

PNL_{max}	Maximum Perceived Noise Level
$PNL_{max,av}$	Average (energy-based) Maximum Perceived Noise Level
$PNLT_{max,av}$	Average (energy-based) Tone-Corrected Maximum Perceived Noise Level
$PNL_{max,av,80}$	Average (energy-based) Maximum Perceived Noise Level of aircraft events with $PNL_{max} > 80$ PNdB
$PNLT_{max,av,80}$	Average (energy-based) Tone-Corrected Maximum Perceived Noise Level of aircraft events with $PNLT_{max} > 80$ PNdB
L_{10}	Sound Level exceeded 10% of the time
L_{20}	Sound Level exceeded 20% of the time
L_{30}	Sound Level exceeded 30% of the time
L_{40}	Sound Level exceeded 40% of the time
L_{50}	Sound Level exceeded 50% of the time
L_{60}	Sound Level exceeded 60% of the time
L_{70}	Sound Level exceeded 70% of the time
L_{80}	Sound Level exceeded 80% of the time
L_{90}	Sound Level exceeded 90% of the time
L_{100}	Sound Level exceeded 100% of the time
R^2	Coefficient of determination or the “R-squared” value
R_{sp}^2	Semipartial R^2
$\overline{R_{sp}^2}$	Average semipartial R^2
$\overline{\overline{R_{sp}^2}}$	Average of $\overline{R_{sp}^2}$
SEL_A	A-weighted Sound Exposure Level
$SEL_{A,av}$	Average (energy-based) A-weighted Sound Exposure Level
$SENEL_{A,av}$	Average (energy-based) A-weighted Single Event Noise Exposure Level
$WECPNL$	Weighted Equivalent Continuous Perceived Noise Level

ABSTRACT

Foertsch, Kevin K. MSE, Purdue University, May 2013. The Number-of-Events as a Predictor Variable in Aircraft Noise Annoyance Models. Major Professor: Patricia Davies, School of Mechanical Engineering.

Aircraft noise may have a number of direct adverse effects on the communities surrounding airports, including annoyance. The annoyance reactions of individuals and communities to aircraft noise are predicted with annoyance models, which are normally functions of predictor variables that describe the noise exposure. The number of aircraft events that a person is exposed to (the number-of-events), has been hypothesized as a significant contributor to annoyance. However, most models of annoyance to aircraft noise are functions only of the average sound energy of the total noise exposure. The purpose of this research is to quantify the relative effects of sound level and number-of-events in historical noise survey data sets and to develop a survey simulation tool to help in the design of future surveys so that the collected data will be sufficient to compare the performance of proposed annoyance models. The models considered here are *DNL* and those that are a function of sound level and number-of-events. Seven historical data sets were collected and analyzed individually and in combination. Multiple linear regression models were estimated using the annoyance, sound level, and number-of-events variables in the data sets. The contributions of sound level and number-of-events to the prediction of annoyance were compared. Most regression models could not be distinguished from an equal-energy annoyance model. A general-purpose tool was developed to simulate annoyance surveys around airports. Monte Carlo simulations were performed to evaluate the effectiveness of survey sampling approaches. Annoyance surveys were simulated around three airports in the United States. The use of stratification, as opposed to simple random sampling, resulted in more robust estimation of annoyance models.

CHAPTER 1. INTRODUCTION

Many predictive models of community annoyance to aircraft noise are functions only of the average sound energy of the total noise exposure. Currently in the United States, the Federal Aviation Administration (FAA) uses the Day-Night Average Sound Level (*DNL*) to predict the annoyance induced by aircraft noise in communities surrounding airports. *DNL* is the A-weighted average sound energy from aircraft events occurring over a 24-hour period, with a 10 dB penalty applied to nighttime sound levels occurring between 10:00 PM and 7:00 AM. Researchers have proposed that an additional acoustical factor, the number of aircraft events that a person is exposed to (the number-of-events), might significantly contribute to annoyance. Alternative annoyance models have been proposed; the models described in the literature are typically multiple linear regression models which contain two predictor variables: the number-of-events (e.g., the total number-of-events or the number-of-events exceeding a certain sound level), and the average sound level of the events (e.g., the average Maximum A-weighted Sound Level or the average Maximum Perceived Noise Level). These models allow the study of the relative importance of sound level and number-of-events to the prediction of annoyance.

The research presented in this thesis is separated into two main parts: the analysis of noise survey data from previous studies, and the simulation of noise surveys. The motivation for the research is to examine and quantify the relationship between physical descriptors of aircraft noise (measures of sound level and the number-of-events) and community annoyance with current survey data and to identify and demonstrate improvements to how future noise surveys are designed and conducted. An ultimate aim of the research on annoyance modeling is to improve the ability to predict changes in community annoyance resulting from changes in air traffic scenarios at airports. The objectives of the research are:

1. Identify trends in annoyance as number-of-events and sound level varies
2. Generate and compare various types of statistical models using available survey data sets to determine how much of the variance of the annoyance responses can be accounted for by the number-of-events term and the sound level term in the models.
3. Quantify the robustness of model parameter estimates that can be generated from available data; offer recommendations for gathering data to improve estimates
4. Demonstrate the effect of survey sampling techniques on model estimation through simulation
5. Offer strategies in the design of future noise surveys to determine the unique contributions of sound level and number-of-events on annoyance and to enable discrimination between the performance of various models

An outline of the work reported in this thesis is as follows. Chapter 2 contains a review of past work in the modeling of community annoyance to aircraft noise, with a description of the findings that have helped establish the direction of the research reported here. Chapter 3 contains a description of noise survey data sets that will be used in the statistical analyses described in Chapter 4. Also in Chapter 3, methods of transforming sound level and annoyance data in the data sets so that there is consistency across all data are described. This allows data set combination. Also, a survey data set visualization program is demonstrated, and a process of simulating a person-level survey data set from site-level data is described. In Chapter 4 objectives 1-3 are addressed by using multiple linear regression techniques and additional statistical analysis of the data sets. The concept of relative variable importance is described, and the relative importance of sound level and number-of-events to the prediction of annoyance is quantified by two separate measures. In addition, the effect of including or excluding an airport in a multiple-airport data set is investigated. The work described in Chapter 5 addresses objectives 4 and 5. In the chapter, a noise survey simulation program developed for the research is described and the results

of simulations of noise surveys around three airports are shown. The noise surveys were designed to facilitate the precise quantification of the relative importance of the number-of-events and sound level terms in the models. The conclusions that can be drawn from the research and suggestions for future research are given in Chapter 6.

CHAPTER 2. LITERATURE REVIEW

Since the end of World War II, annoyance and complaints about noisy or low-flying aircraft have been a public concern (Borsky, 1954). Within a few years of the start of jet air transport service in the late 1950's, an initial round of a large-scale social surveys were conducted in the vicinity of major international airports (Fidell and Pearsons, 2007). The results of some of the early surveys are reported in (Schultz, 1982, Ch. 4). From the advent of investigations on the annoyance to aircraft noise, the contribution of the number of aircraft events that a person hears (the number-of-events) to the experience of annoyance to aircraft noise has been in question.

Mathematical models are used around the world to predict community response to aircraft noise. Most models are functions only of the average sound energy of the total noise exposure. Currently in the United States, the Federal Aviation Administration (FAA) uses the Day-Night Average Sound Level (DNL) to predict the annoyance induced by aircraft noise in communities surrounding airports. DNL is the A-weighted average sound level in decibels occurring over a 24-hour period with a 10 dB penalty applied to nighttime aircraft events occurring between 10:00 PM and 7:00 AM (U.S. Environmental Protection Agency, 1974). DNL can be expressed by:

$$DNL = 10 \log_{10} \left[\frac{1}{24} \left(15 \left(10^{\frac{L_{Aeq,d}}{10}} \right) + 9 \left(10^{\frac{L_{Aeq,n}+10}{10}} \right) \right) \right], \quad (2.1)$$

where $L_{Aeq,d}$ is the A-weighted Equivalent Sound Level (L_{Aeq}) for the daytime (7:00 AM to 10:00 PM) and $L_{Aeq,n}$ is the L_{Aeq} for the nighttime (10:00 PM to 7:00 AM). At its core, DNL is an average sound energy and is calculated by taking into account all sound energy experienced by a receiver at a given location. Therefore, the number-of-events is not included in DNL as an explicit term, only implicitly in the averaging

process. In Chapter 3 and Appendix A.1 it is shown that Equation (2.1) can be expressed as a model that includes an explicit number-of-events term.

2.1 Alternative Annoyance Models

Researchers have proposed alternative annoyance models that separate the contributions of sound level and number-of-events (Bjorkman, 1991; Bullen and Hede, 1986a,b; Labiale, 1983; Le Masurier, Bates, and Taylor, 2007a; McKennell, 1963; MIL Research Ltd., 1971; Powell, 1980; Rice, 1977a,b, 1978; Rylander, Bjorkman, and Ahrlin, 1980). These annoyance models are multiple linear regression models and are functions of a sound level variable, a number-of-events variable, and an intercept. A reason for this proposition is a suspicion that people who are annoyed by aircraft noise might be reacting to more than just the total sound energy of the exposure.

Consider an example of two hypothetical aircraft operating scenarios that a person could be exposed to in a typical day. The first scenario is composed of 125 daytime events and 12 nighttime events each with a Sound Exposure Level (SEL_A) of 96 dBA per event. The second scenario is composed of 30 daytime events and 3 nighttime events each with a SEL_A of 102 dBA per event. The scenarios differ in the number-of-events and sound level per event, but $DNL = 70$ dBA in both scenarios. Although the scenarios are of equal DNL will they prompt an equal annoyance reaction? The motivation for exploring the alternative annoyance models in which sound level and number-of-events are separated is to determine the extent to which the individual descriptors affect annoyance.

If an annoyance model is a function of a measure of sound level (in decibels) and the logarithm (base 10) of the number-of-events, it is referred to as an “additive-log” model. The additive-log model can be expressed as:

$$\hat{A} = \hat{b}_L L + \hat{b}_N \log_{10} N + \hat{b}_0, \quad (2.2)$$

where \hat{A} is the annoyance variable estimate, L is the sound level variable, \hat{b}_L is the sound level coefficient estimate, N is the number-of-events variable, \hat{b}_N is the number-of-events coefficient estimate, and \hat{b}_0 is the intercept estimate. In an analysis of survey data, Fields (1984) evaluated the effect of the number-of-events in a data set using a ratio he called the decibel equivalent number effect, defined as $\hat{k} = \hat{b}_N/\hat{b}_L$. It can be seen that \hat{k} is the number of decibels which increases annoyance by the same amount as a tenfold increase in the number-of-events.

Equation (2.2) can be rewritten as:

$$\frac{\hat{A}}{\hat{b}_L} = L + \frac{\hat{b}_N}{\hat{b}_L} \log_{10}(N) + \frac{\hat{b}_0}{\hat{b}_L}. \quad (2.3)$$

In the instance that L is an average of individual aircraft events' 24-hr A-weighted Equivalent Sound Level (L_{Aeq}), then:

$$L = 10 \log_{10} \left(\frac{\overline{p_{av}^2}}{p_{ref}^2} \right), \quad (2.4)$$

where $\overline{p_{av}^2} = \frac{1}{N} \sum_{i=1}^N 10^{L_{Aeq,i}/10}$, p_{ref} is the reference pressure of 20 μPa , and N is the number-of-events. Equation (2.3) can then be written as:

$$\frac{\hat{A}}{\hat{b}_L} = 10 \log_{10} \left(\frac{\overline{p_{av}^2}}{p_{ref}^2} \right) + \frac{\hat{b}_N}{\hat{b}_L} \log_{10} N + \frac{\hat{b}_0}{\hat{b}_L}. \quad (2.5)$$

This can be thought of as an ‘‘equal-energy’’ model when $\hat{k} = 10$, because doubling the number-of-events has the same effect as doubling the energy of each event, i.e., both result in an addition of $10 \log_{10}(2)$. \hat{k} is a measure of the importance of the number-of-events relative to sound level. Although *DNL* is not expressed an additive-log model, $k = 10$ is implied in its calculation (Federal Aviation Administration, Office of Environment and Energy, 2009; Fields, 1984). When $\hat{k} = 10$, the additive-log model is an equal-energy model, like *DNL*.

Typically, the sound level variable in the additive-log model is an average sound level, calculated by averaging single-event sound levels. The average sound level and number-of-events data for a noise-exposed area is either calculated by including all aircraft that operate in the vicinity of the area or by including only the aircraft events with a single-event sound level above a threshold. The average sound level could be, e.g., an average of the Maximum Perceived Noise Levels ($PNL_{max,av}$) (Brooker, Critchley, Monkman, and Richmond, 1985; Josse, 1968; McKennell, 1963; MIL Research Ltd., 1971), an average of the Maximum A-weighted Sound Levels ($L_{Amax,av}$) (Bullen and Hede, 1986a,b; Le Masurier et al., 2007a; Rylander and Bjorkman, 1997; Rylander et al., 1972, 1974, 1980), or an average of the Effective Perceived Noise Levels ($EPNL_{av}$) (Bullen and Hede, 1986a,b). Different methods of averaging can be used, but the norm is to use the logarithmic, or energy-based average. In two laboratory studies, researchers explored many different sound level variables in the additive-log model. (Powell, 1980; Rice, 1977a,b). They and other researchers found that for the variables examined the particular choice of sound level variable was not critical, as both single-event metrics and metrics that involve averages of the single-event metrics were correlated highly with one another (Bishop, 1975; Hede and Bullen, 1982; Pearsons, Howe, Sneddon, Silvati, and Fidell, 1997; Powell, 1980; Rice, 1977a; TRACOR, Inc., 1971). In Appendix A.1, formulae for additive-log models which have been used as official noise ratings to predict the reaction of individuals to noise, as well as additive-log approximations of other noise ratings are given.

2.2 Annoyance Model Estimation and Sample Design

In past research on the use of additive-log models to predict the annoyance to aircraft noise, k was found to be less than 10 (below equal-energy) (Rice, 1977a; Rylander et al., 1972, 1974, 1980), above 10 (greater than equal-energy) (Bullen and Hede, 1986a,b; Le Masurier et al., 2007a; McKennell, 1963; Powell, 1980), and that results can depend on the acoustical data used (MIL Research Ltd., 1971) or the annoyance

variable used (Brooker et al., 1985). However, sensitivity analyses of \hat{k} have revealed that large variation in \hat{k} is possible without an accompanying large variation in the fit of the model; a good fit of the model to the data, in terms of the coefficient of determination (or R^2 value), can be realized for a large range of \hat{k} (Le Masurier et al., 2007a; MIL Research Ltd., 1971; Powell, 1980; Rice, 1977a). The insensitivity of the fit of the annoyance model to the data is a function of the correlation between the sound level and number-of-events variables for the data sets used. Typically, as aircraft events become more noisy, more people hear them, and in practice it is difficult to collect uncorrelated data (Fields, 1980). Correlation is the two-dimensional case of a general problem of interrelation between predictor variables in a regression model, known as multicollinearity, see, e.g., (Belsley, 1991a). Multicollinearity is described in Appendix C. When predictor variables are correlated, regression coefficient estimates tend to have large variability due to the poorer conditioning of the matrices used in the estimation, which makes the solution sensitive to perturbations in the data. In the case of the additive-log annoyance model, correlation also increases the variance estimate of \hat{k} because it is a function of the two regression coefficient estimates. In one of the earliest surveys of annoyance to aircraft noise, McKennell (1963) stated that the precise determination of the level/number trade-off (\hat{k}) was prevented by correlation between the acoustical variables. To address this problem, he recommended that sample design, specifically stratification, be strategically used to collect acoustical data that would facilitate model identification, and the determination of \hat{k} with more precision.

In light of the results and recommendations from McKennell (1963), future surveys made use of stratification prior to sampling a population. 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. Stratification is described in more detail in Section 5.1 of Chapter 5. Brooker et al. (1985); Josse (1968); MIL Research Ltd. (1971) stratified by measured sound level and number-of-events data. Rylander et al. (1972, 1974, 1980) stratified by

sound level and number-of-events data from predicted sound level contours and aircraft movement data. Le Masurier et al. (2007a) stratified by simulated sound level and number-of-events data calibrated with measurements. TRACOR, Inc. (1971) and Connor and Patterson (1972) stratified by sound level only, through the use of Perceived Noise Level (PNL) contours. Hede and Bullen (1982) stratified only by Noise Exposure Forecast (NEF), which is an additive-log model with a k of 10 and a nighttime event weighting. In Table 2.1 are shown the correlations between the acoustical variables realized for different surveys. For some data sets, the acoustical data are person-level so that each acoustical data observation represents the noise exposure for one person. In other data sets, the acoustical data are site-level so that all of the persons within a given geographical region are assigned the noise exposure experienced at a certain spatial point within that region, e.g., the population-weighted centroid.

Table 2.1. Correlation ($\rho_{L, \log_{10}(N)}$) between the sound level and $\log_{10}(N)$ variables for survey data sets, where N is the number-of-events. The sound level variable (L) and the sound level threshold (L_t) for counting number-of-events are given. $PNL_{max,av}$ is the average Maximum Perceived Noise Level. MNL is the Maximum A-weighted Sound Level (L_{Amax}) of the noisiest aircraft regularly using the airfield at least 3 times per day. L_{av} is the arithmetic average Maximum A-weighted Sound Level. Data sets with site-level data are notated with the superscript *. Data sets that resulted from the use of stratification by both sound level and number-of-events are notated with the superscript S .

Survey	L	L_t	$\rho_{L, \log_{10}(N)}$
McKennell (1963)	$PNL_{max,av}$	80 PNdB	0.70
Josse (1968) ^{*S}	$PNL_{max,av}$	80 dBA	0.50
MIL Research Ltd. (1971) ^S	$PNL_{max,av}$	80 PNdB	0.24
TRACOR, Inc. (1971)	$PNL_{max,av}$	None	0.24
Connor and Patterson (1972)	$PNL_{max,av}$	None	0.48
Rylander et al. (1972) ^{*S}	MNL	70 dBA	-0.08
Hede and Bullen (1982)	$PNL_{max,av}$	None	-0.49
Brooker et al. (1985) ^S	$PNL_{max,av}$	80 PNdB	-0.10
Le Masurier et al. (2007b) ^{*S}	L_{av}	65 dBA	0.53

In general, the stratification of a population by the acoustical variables can be used to reduce the correlation between the variables (Allison, 1999). Although there is not a maximum-allowable correlation, with all else being the same, reduced correlation leads directly to lower variance estimates for the regression coefficient estimates and \hat{k} . As mentioned, it is typical for some correlation to exist between sound level and number-of-events data because normally, as the average sound level of events increases, the number that people notice also increase. The combinations of either a small number of events with a large average sound level or a large number of events with a small average sound level are less likely than either a small number of events with a small sound level or a large number of events with a large average sound level. The correlation between acoustical variables that can be realized in a sample is not limited by the correlation between variables in the population. However, there may not be an adequate number of either a small number/large sound level or large number/small sound level combinations at the extremes of the acoustical variable ranges in the population to achieve a small correlation between the variables in the sample, if the full ranges of the variables are to be represented.

2.3 Holistic Annoyance Models

Many annoyance models are in the form of Equation (2.2), which are functions only of acoustical descriptors of noise. However, the variance in the annoyance variable accounted for by the acoustical descriptors is usually very low. This could be due to inherent variability in people's responses to noise, or inaccuracy in measurements of annoyance or the acoustical variables. Job (1988) thought concentration on improvements in noise measurement accuracy was unlikely to substantially increase the relationship between noise exposure and annoyance. The effects of nonacoustical variables have been investigated as predictors of annoyance. Generally, the nonacoustical factors are classified into three main groups: attitude variables, socio-demographic variables, and situational variables (Marquis-Favre, Premat, and Aubrée, 2005a). At-

titude variables can be categorized into six groups: fear of danger of a noise source, attitude toward a noise source, sensitivity to noise, activity during exposure, perception of the neighborhood, and the global perception of the environment (Nelson, 1987). Socio-demographic variables include, e.g., gender, marital status, culture, household size, education level, social status, income, age, length of residence, type of home, and type of housing occupancy (owning or renting). Situational variables include, e.g., time spent at home, exposure to noise at work, soundproofing at home, and general state of health.

In some studies, the cumulative effects of multiple nonacoustical variables on the annoyance to noise have been investigated (Bertoni, Franchini, Magnoni, Tartoni, and Vallet, 1993; Fields, 1993; Fields and Walker, 1982; Job, 1988; Miedema and Vos, 1999; Vallet, 1996). (Job, 1988), (Fields, 1993), and (Miedema and Vos, 1999) are reports of meta-analyses of survey data. In these types of studies researchers usually explored the correlations between the variables and annoyance. Findings on the effects of nonacoustical variables have been mixed across different surveys. In general, attitude variables have been found to correlate more strongly with annoyance than either socio-demographic or situational variables with the accompanying conclusion that socio-demographic variables are not strong indicators of annoyance (Fields, 1993; Miedema and Vos, 1999). Also, although attitudinal variables correlate with annoyance, it is uncertain what the causal relationships are if there are any, e.g., fear may increase annoyance or vice versa, or they may be expressions of one underlying construct.

In some studies, models of annoyance which involve nonacoustical predictor variables have been proposed which are sometimes referred to as “holistic” models. The models are either multiple linear regression models (Le Masurier et al., 2007a; Leonard and Borsky, 1973; Miedema and Vos, 1999; Raw and Griffiths, 1988; Schreckenberg and Schuemer, 2010; TRACOR, Inc., 1971), path models (Leonard and Borsky, 1973; Taylor, 1984; Yano, Sato, Bjorkman, and Rylander, 2002), or structural equations (Kroesen, Molin, and van Wee, 2008; Schreckenberg, 2011). Path models are formed from path analysis, which is an extension of multiple linear regression analysis that

accounts for causal structure between variables. When the variables in path analysis are not single variables, but mathematically-derived constructs through the analysis and combination of two or more single variables, path analysis is termed structural equation modeling, from which structural equations are formed. Causal models are capable of modeling indirect and reciprocal effects between variables (Kroesen et al., 2008) but the application of causal models in sociological research has been criticized because the predictor variables assumed to influence annoyance cannot be considered to be independent of annoyance at the questionnaire stage (Alexandre, 1999).

In a discussion of nonacoustical variables, Guski (1999) stated that at best, about one-third of the variance in annoyance reactions can be accounted for by variation in acoustical variables and another one-third can be accounted for by variation in nonacoustical variables. TRACOR, Inc. (1971) found that 78% of the variance in annoyance was accounted for by a combination of the Composite Noise Rating (*CNR*) and seven nonacoustical variables, including self-rated fear and sensitivity, versus 37% of the variance accounted for by *CNR* alone. Leonard and Borsky (1973) found that in a regression analysis, 10% of the variance in annoyance over a two-month period during June and July was accounted for by *CNR*, which was raised to 58% when three nonacoustical variables, including self-rated fear, were included. For the months of August and September, the percentage of variance accounted for increased from 14% to 45%. Raw and Griffiths (1988) found that 28% of the variance in annoyance was accounted for by the 18-hour A-weighted Sound Level exceeded 10% of the time at the facade of a house ($L_{A,10}$) which was increased to 50% when five nonacoustical variables, including self-rated sensitivity, were included in the analysis. Schrecken-berg and Schuemer (2010) found that 18% of the variance in long-term annoyance was accounted for by a 16-hour $L_{Aeq,d}$ from 6:00 AM to 10:00 PM, with 64% of the variance accounted for when twelve nonacoustical predictors were included in the model, including self-rated fear and sensitivity. In the path analysis of Leonard and Borsky (1973), *CNR* was stated to be a significant predictor of annoyance in the August and September data, but not in the June and July data. This finding was addressed in a

criticism of path analysis (Alexandre, 1999). The results of Kroesen et al. (2008) were interesting in that although fear and sensitivity were strong correlates with annoyance in the survey data, they were unable to account for additional variance over other variables that were included in the model, which might have been due to correlations between these variables and others that were included in the model.

In some studies where the effects of nonacoustical variables were considered, measures of noise exposure were found to correlate less with annoyance than some nonacoustical variables (Kroesen et al., 2008; McKennell, 1963; Taylor, 1984; TRACOR, Inc., 1971; Yano et al., 2002) though this may be due to a faulty conclusion from path analysis (Alexandre, 1999), the choice of a cumulative energy metric over one based on maximum sound levels (Taylor, 1984), or a limited range in the noise exposure variable used (Job, 1988; Kroesen et al., 2008). It is expected that acoustical variables would be correlated with annoyance due to the noise of aircraft, for in order for noise to be annoying, it must be heard. The strong correlations between attitudinal and annoyance variables could, e.g., be an indicator that the variables are measurements of one underlying construct. It could also be that the annoyance felt by survey respondents in the vicinity to airports is not primarily due to noise, but to, e.g., fear of danger. There is value in including nonacoustical variables to improve the prediction of annoyance at the person-level. However, nonacoustical variables which vary considerably from person to person and are descriptive of specific attributes of persons (such as those presented above) cannot be used for noise abatement programs, unlike social variables which are common to people of a community, such as the relationship between the operators of the noise source and the community, the history of noise exposure, and general expectation of residents (Guski, 1999).

2.4 Conclusions

Aircraft noise may have a number of direct adverse effects on the communities surrounding airports, including annoyance. Annoyance models are used to assess noise

exposure, establish criteria for acceptable noise environments, and predict the reactions of individuals and communities to noise exposure (Schultz, 1982). Most models of annoyance to aircraft noise are functions only of the average sound energy of the noise exposure, but it is possible that people react to more than sound energy. Specifically, the number-of-events has been proposed as an important contributor to annoyance. Social surveys have been used extensively to study the relationships between noise exposure and annoyance, some also taking into account the possible effects of nonacoustical factors of annoyance. Surveys are integral in the formation and validation of annoyance models. However, the data collected from historical surveys of the annoyance to aircraft noise have resulted in an assortment of proposed annoyance models, with little cross-validation (Fidell and Pearsons, 2007). Contributing factors to the lack of agreement could include inherent variability in people's responses, variation in aircraft operations scenarios, or even differences in housing construction, which affects indoor noise exposure. The tendency for the sound level and number-of-events data to be correlated in the population has also contributed to the difficulty of identifying additive-log models with precision and determining the relative effects of sound level and number-of-events on annoyance. There are data sets of many historical surveys of the annoyance to aircraft noise which would be valuable for research into the effects of different variables on annoyance (Bassarab, Sharp, and Robinette, 2009). In addition, it is important to understand the effect of sampling strategy in noise surveys on data collection and model estimation so that surveys can be designed and implemented which allow the precise identification of annoyance models.

CHAPTER 3. NOISE SURVEY DATA SETS

In this chapter the various noise survey data sets that were used in the research are described. Detailed information on each data set can be found in an updated catalog of 628 social surveys of residents' reaction to environmental noise from the years 1943 to 2008 (Bassarab et al., 2009). This is an update of the catalog by Fields (2001). Seven noise survey data sets were used in a reanalysis of survey data. These were chosen because of the availability of the data. The original survey name with the identifier used by Bassarab et al. (2009), the reference of the main report, ranges of sound level and number-of-events that people were exposed to, and the sample size (n) of the seven studies are given in Table 3.1.

USA-022, USA-032, and USA-044 were different phases of a endeavor that was sponsored by the National Aeronautical and Space Administration. The collective data set is referred to as USA-054. All of the survey datasets were adapted to a standardized storage format in which annoyance, sound level, number-of-events, sample design, and nonacoustical variables are arranged in a defined order. The standardized format allocates space for 58 annoyance variables (verbal, numerical, and numerical with verbal endpoints scales; annoyance and annoyance frequency ratings for individual activities, overall activity interference annoyance, and overall frequency of annoyance), 55 sound level variables, 37 number-of-events variables, 19 sample design variables, and 40 nonacoustical variables. The standardized format is shown in Table D.1 in Appendix D. For two historical data sets, multiple airports were included in the surveys. For these, in addition to being stored as a single noise survey data set, the data sets associated with the individual airports were separated and stored.

UKD-024, UKD-130, USA-022, USA-032, and USA-044 were acquired for this research through interaction with Dr. Jim Fields, who has derived annoyance models from the data (Fields, 1984). Fields also provided additional files that were used to

Table 3.1. Information about noise survey data sets. $PNL_{max,av}$ is the average Maximum Perceived Noise Level of aircraft events measured in PNdB and N is the average daily number of aircraft events. $PNL_{max,av}$ is described in Appendix A.2.1.

Study name/identifier	Main report	Range of $PNL_{max,av}$	Range of N	n
1967 Heathrow Aircraft Noise Study (UKD-024)	(MIL Research Ltd., 1971)	81.5 to 112.5	0 to 487.4	4655
1976 Heathrow Concorde Noise Survey (UKD-130)	(McKennell, 1977)	83.0 to 114.0	20 to 500	2618
1967 USA Four-Airport Survey (Phase 1 TRACOR Survey) (USA-022)	(TRACOR, Inc., 1971)	66.0 to 113.7	6 to 503	3537
1969 USA Three-Airport Survey (Phase 2 TRACOR Survey) (USA-032)	(TRACOR, Inc., 1971)	76.7 to 123.6	3 to 268	2902
1970 USA Small City Airports	(Connor and Patterson, 1972)	77.2 to 115.5	1.1 to 31.2	1960

Table 3.1. Continued

Study name/identifier	Main report	Range of $PNL_{max,av}$	Range of N	n
(Small City TRACOR Survey) (USA-044)				
1980 Australian Five Airport Survey (AUL-210)	(Hede and Bullen, 1982)	81.1 to 117.1	0.7 to 132.8	2720
2005 Attitudes to Noise from Aviation Sources in England (ANASE) (UKD-604)	(Le Masurier et al., 2007a)	78.3 to 95.9	0.6 to 299.1	2437

import, process, and store the data sets within Statistical Package for the Social Sciences (SPSS). The relevant files for each data set included the data set raw data text file and a SPSS command file. The command file contained commands which imported and processed the raw data. Because the command files were created in the early 1980s, the syntax had to be updated to current coding standards. He also provided many other resources documenting his analyses of the survey data as well as guides to using the resources efficiently. This was helpful to test the integrity of the data sets and to “calibrate” the analysis programs developed in the research reported in this thesis. The AUL-210 data set had been previously acquired over the Internet from the UK Data Archive by a colleague. For all data sets but UKD-604, the person-level acoustical and annoyance data were available. With UKD-604, the person-level data were not available. Instead, site-level were available from Le Masurier et al. (2007b) and were used to simulate a person-level data set. The simulation process is described in Section 3.3.

The sound level metrics and annoyance scales varied among the survey data sets which are shown in Table 3.1. In order to compare the findings of studies and combine data sets, it was necessary to establish a common set of sound level metrics and annoyance scales among the data sets. Methods used to transform sound level metrics and annoyance scales are described in the following sections. The sound level metrics referred to are well-known in the environmental noise community. For the less familiar reader, an overview of various sound level metrics and their definitions can be found in (Bennett and Pearsons, 1981; Marquis-Favre et al., 2005b; Schultz, 1982). The common sound level metric that will be used throughout this chapter, $PNL_{max,av}$, is described in Appendix A.2.1.

3.1 Sound Level Metrics

$PNL_{max,av}$ was chosen as the common sound level metric. The UKD-024, UKD-130, USA-022, USA-032, and USA-044 data sets all contain $PNL_{max,av}$ data. AUL-

210 does not contain $PNL_{max,av}$ data, but has DNL and number-of-events data. A method of estimating $PNL_{max,av}$ from DNL and number-of-events metrics was devised and is described in Section 3.1.1. A similar problem arose with the UKD-604 data where a different metric was used. The transformation of that metric to $PNL_{max,av}$ is described in Section 3.2.

3.1.1 DNL to $PNL_{max,av}$

DNL is the Day-Night Average Sound Level. Strictly speaking, DNL should be calculated for a point in space by considering all sounds to which the point in space is exposed over a 24-hour time period. DNL can be expressed by:

$$DNL = 10 \log_{10} \left[\frac{1}{24} \left(15 \left(10^{\frac{L_d}{10}} \right) + 9 \left(10^{\frac{L_n+10}{10}} \right) \right) \right], \quad (3.1)$$

where L_d is the A-weighted Equivalent Sound Level (L_{Aeq}) for the daytime (7:00 AM to 10:00 PM) and L_n is the L_{Aeq} for the nighttime (10:00 PM to 7:00 AM) (U.S. Environmental Protection Agency, 1974). However, DNL can be expressed in various formulations, which facilitates other calculations. For example, the DNL of the Australian study (AUL-210) was calculated from a formula that utilizes tabulated A-weighted Sound Exposure Level (SEL_A) for individual aircraft (Hede and Bullen, 1982, p. 36). In the formula, DNL is expressed as a sum (over all aircraft types and operation types) of the energies captured in the SEL_A 's (one SEL_A quantifies the A-weighted energy in the entire duration of an aircraft operation) multiplied by the sum of the number of daytime events and ten times number of nighttime events. Thus, the formula for DNL used in the Australian study is:

$$DNL = 10 \log_{10} \left(\sum_i \sum_j (N_{i,j} + 10N'_{i,j}) 10^{\frac{SEL_{A_{i,j}}}{10}} \right) - 49.4, \quad (3.2)$$

where $SEL_{A_{i,j}}$ is the SEL_A for aircraft of type i performing operation j , $N_{i,j}$ is the average number of such aircraft operations per day between 7:00 AM and 10:00 PM,

and $N'_{i,j}$ is the average number such operations between 10:00 PM and 7:00 AM. In the Australian study, each SEL_A was defined as the total energy in an event above the background noise, not only that within 10 dB of the peak. Thus, each $SEL_{A_{i,j}}$ quantifies the total energy in an aircraft event. Periods of background noise without aircraft events were not considered in the DNL calculation.

Equation (3.2) can be approximated by the use of an average A-weighted Sound Exposure Level ($SEL_{A,av}$):

$$DNL = SEL_{A,av} + 10\log_{10}(N_d + 10N_n) - 49.4, \quad (3.3)$$

where $SEL_{A,av}$ is a Sound Exposure Level for the average aircraft event, N_d is the average number of daytime events, and N_n is the average number of nighttime events.

In addition to DNL , the average daily number-of-events with Maximum A-weighted Sound Level (L_{Amax}) greater than or equal to 70 dBA ($N_{\geq 70}$) was reported in the Australian study. This information was integrated into Equation (3.3) by separating contributions from aircraft events with $L_{Amax} \geq 70$ dBA from those with $L_{Amax} < 70$ dBA:

$$DNL = SEL_{A,av \geq 70} + 10\log_{10}(N_{\geq 70,d} + 10N_{\geq 70,n}) + \\ SEL_{A,av < 70} + 10\log_{10}(N_{< 70,d} + 10N_{< 70,n}) - 49.4, \quad (3.4)$$

where $SEL_{A,av \geq 70}$ is the average A-weighted Sound Exposure Level for events with $L_{Amax} \geq 70$ dBA, $N_{\geq 70,d}$ is the average number of average daytime events with $L_{Amax} \geq 70$ dBA, $N_{\geq 70,n}$ is the average number of nighttime events with $L_{Amax} \geq 70$ dBA, $SEL_{A,av < 70}$ is the average A-weighted Sound Exposure Level for events with $L_{Amax} < 70$ dBA, $N_{av < 70}$ is the average number of daytime events with $L_{Amax} < 70$ dBA, and $N_{< 70,n}$ is the average number of nighttime events with $L_{Amax} < 70$ dBA. In order to use the $N_{\geq 70}$ data with Equation (3.4), $N_{\geq 70}$ was split into daytime and nighttime

components. This was done assuming 90% of the events in $N_{\geq 70}$ occurred during the day. Thus, for the AUL-210 data:

$$N_{\geq 70,d} = 0.9N_{\geq 70}, \quad (3.5)$$

$$N_{\geq 70,n} = 0.1N_{\geq 70}. \quad (3.6)$$

The next step was to calculate $SEL_{A,av < 70}$ so that by rearranging Equation (3.4), $SEL_{A,av \geq 70}$ could be calculated. This was required because both $SEL_{A,av < 70}$ and $SEL_{A,av \geq 70}$ are required to calculate an overall energy-based average SEL_A of all events. Rearranging Equation (3.4) yields:

$$\begin{aligned} SEL_{A,av \geq 70} = & DNL - 10 \log_{10}(N_{\geq 70,d} + 10N_{\geq 70,n}) - \\ & SEL_{A,av < 70} - 10 \log_{10}(N_{< 70,d} + 10N_{< 70,n}) + 49.4. \end{aligned} \quad (3.7)$$

$SEL_{A,av < 70}$ was calculated by deriving a relationship between the A-weighted equivalent mean square pressure of an average aircraft event ($p_{A,eq,av}^2$) and the A-weighted maximum mean square pressure of the average aircraft event ($p_{A,max,av}^2$). The relationship between $p_{A,eq,av}^2$ and $p_{A,max,av}^2$ was established through a visual analysis of sound level profiles for flyover recordings. Examples of sound level profiles for different flyover recordings are shown in Figure 3.1. The method of calculating $SEL_{A,av < 70}$ is given by Equations (3.8) to (3.15).

From Figure 3.1, the relationship between $p_{A,max}^2$ and $p_{A,eq}^2$ was established by equating the energies (areas underneath the curves) of a triangular aircraft event approximation and a rectangle. The energy contained within a triangular aircraft event approximation can be expressed as:

$$E = \frac{1}{2}D(p_{A,max}^2 - p_{A,eq,bg}^2), \quad (3.8)$$

where D is the duration of the event, $p_{A,max}^2$ is the A-weighted maximum mean square pressure and height of the triangle, and $p_{A,eq,bg}^2$ is the A-weighted steady-state back-

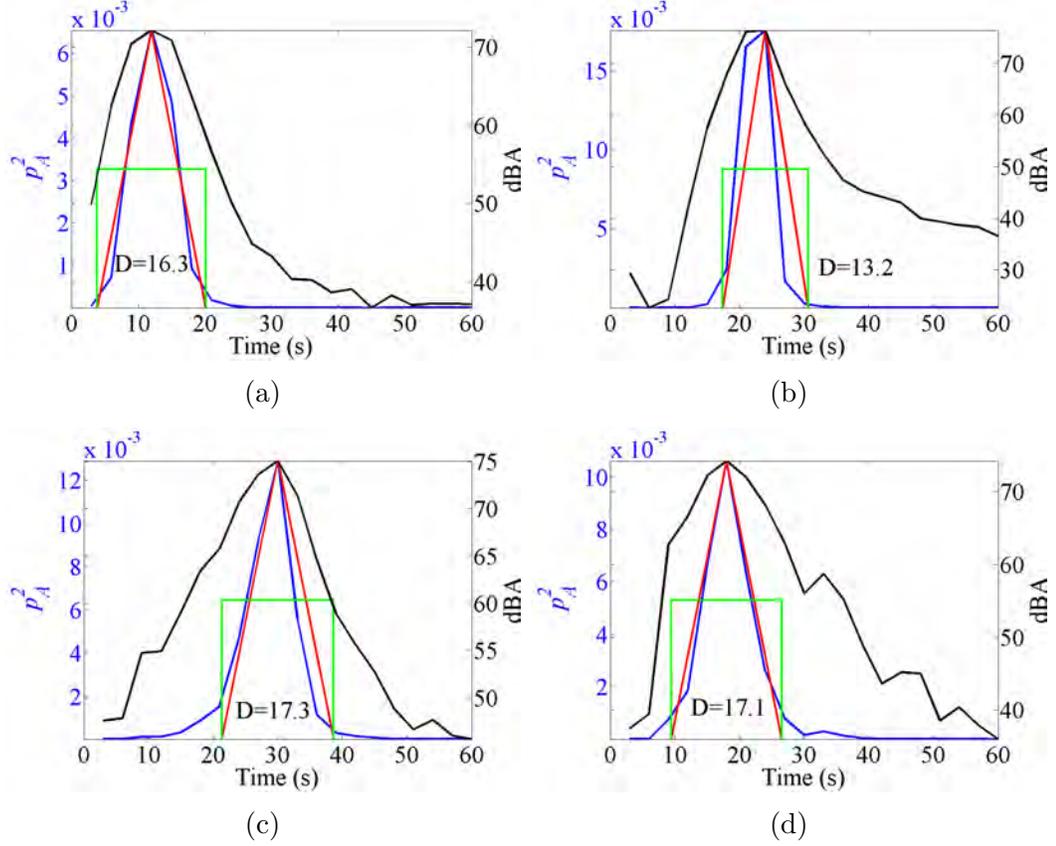


Figure 3.1. Overflight A-weighted mean square pressure profile (blue) and decibel (referenced to $20 \mu\text{Pa}$) profile (black) for: (a) Boeing 737, (b) Boeing 747, (c) Boeing 757, and (d) MD80. Each blue overflight profile shape can be well-approximated by a triangle, shown in red; the height of the triangle represents the A-weighted maximum mean square pressure of a single aircraft event ($p_{A,max}^2$). The height of the green rectangle represents the A-weighted equivalent mean square pressure of a single aircraft event ($p_{A,eq}^2$). The blue, red, and green pressure profiles all contain the same amount of energy, or enclose the same amount of area. The duration of each triangular-approximated event is noted. An averaging time of 3 seconds was used to construct the profiles.

ground mean square pressure. The energy contained within the rectangle can be expressed as:

$$E = D(p_{A,eq}^2 - p_{A,eq,bg}^2), \quad (3.9)$$

where $p_{A,eq}^2$ is the A-weighted equivalent mean square pressure of an aircraft event. Equations (3.8) and (3.9) can be equated to yield:

$$p_{A,eq}^2 = 0.5(p_{A,max}^2 + p_{A,eq,bg}^2). \quad (3.10)$$

Adopting Equation (3.10) to the average aircraft event with $L_{Amax} < 70$ dBA yields:

$$p_{A,eq,av<70}^2 = 0.5(p_{A,max,av<70}^2 + p_{A,eq,bg}^2), \quad (3.11)$$

where $p_{A,eq,av<70}^2$ is the average A-weighted equivalent mean square pressure of aircraft events with $L_{Amax} < 70$ dBA and $p_{A,max,av<70}^2$ is the average A-weighted maximum mean square pressure of aircraft events with $L_{Amax} < 70$ dBA. In AUL-210, the A-weighted Sound Exposure Level can be expressed as:

$$SEL_A = 10 \log_{10} \left(\frac{p_{A,eq}^2}{p_{ref}^2} \right). \quad (3.12)$$

Adopting Equation (3.12) to the average aircraft event with $L_{Amax} < 70$ dBA yields:

$$SEL_{A,av<70} = 10 \log_{10} \left(\frac{p_{A,eq,av<70}^2}{p_{ref}^2} \right), \quad (3.13)$$

where p_{ref} is the reference pressure of $20 \mu\text{Pa}$. Inserting Equation (3.11) into Equation (3.13) yields the following expression for the average Sound Exposure Level of the events with $L_{Amax} < 70$ dBA:

$$SEL_{A,av<70} = 10 \log_{10} \left(\frac{0.5 (p_{A,max,av<70}^2 + p_{A,eq,bg}^2)}{p_{ref}^2} D \right). \quad (3.14)$$

Thus, $SEL_{A,av<70}$ can be calculated by assuming a $p_{A,max,av<70}^2$ for the events with $L_{Amax} < 70$ dBA. $p_{A,max,av<70}^2$ can be calculated directly from an assumed average of L_{Amax} for aircraft events with $L_{Amax} < 70$ dBA ($L_{Amax,av<70}$):

$$p_{A,max,av<70}^2 = p_{ref}^2 10^{\frac{L_{Amax,av<70}}{10}}. \quad (3.15)$$

Equations (3.15), (3.5), and (3.6) can be inserted into Equation (3.7) to calculate $SEL_{A,av\geq 70}$. Then, an overall $SEL_{A,av}$ can be calculated from:

$$SEL_{A,av} = 10 \log_{10} \left(\frac{N_{<70} 10^{\frac{SEL_{A,av<70}}{10}} + N_{\geq 70} 10^{\frac{SEL_{A,av\geq 70}}{10}}}{N_T} \right), \quad (3.16)$$

where N_T is the total number-of-events. From $SEL_{A,av}$, the average A-weighted maximum mean square pressure can be calculated:

$$p_{A,max,av}^2 = 2 \left(\frac{10^{\frac{SEL_{A,av}}{10}}}{D} p_{ref}^2 \right) - p_{A,eq,bg}^2. \quad (3.17)$$

$L_{Amax,av}$ can be calculated from:

$$L_{Amax,av} = 10 \log_{10} \left(\frac{p_{A,max,av}^2}{p_{ref}^2} \right). \quad (3.18)$$

Finally, $PNL_{max,av}$ can be calculated from:

$$PNL_{max,av} = L_{Amax,av} + 13. \quad (3.19)$$

Ten different transformations of DNL to $PNL_{max,av}$ were realized by assuming ten different levels of $L_{Amax,av<70}$. This was done to explore the effect of changing the average L_{Amax} of aircraft events with $L_{Amax} < 70$ dBA ($L_{Amax,av<70}$) on multiple linear regression results. $L_{Amax,av<70}$ was ranged from 60 to 69 dBA; multiple linear regression results of the AUL-210 annoyance variable on a the ten resulting estimates of $PNL_{max,av}$ and N_T are shown in Table 3.2. From an examination of Table 3.2, as

the $L_{Amax,av}$ of the events with L_{Amax} below 70 dBA was increased, \hat{k} systematically increased, but only slightly. Thus, changing $L_{Amax,av < 70}$ did not significantly change the regression results. In light of these results, only the $PNL_{max,av}$ data resulting from $L_{Amax,av < 70} = 65$ dBA is used throughout the rest of the document.

Table 3.2. Multiple linear regressions results for ten transformations of DNL to $PNL_{max,av}$. The regression model derived in each case was: $A = b_L PNL_{max,av} + b_N \log_{10}(N_T) + b_C$, where A is a 5-point verbal annoyance scale. $\hat{k} = \hat{b}_N / \hat{b}_L$ is referred as the “decibel equivalent number effect”.

$L_{Amax,av < 70}$	\hat{b}_L	$\hat{\sigma}_{\hat{b}_L}$	\hat{b}_N	$\hat{\sigma}_{\hat{b}_N}$	\hat{b}_C	$\hat{\sigma}_{\hat{b}_C}$	\hat{k}
60	0.055	0.0045	1.017	0.059	-3.95	0.50	18.5
61	0.055	0.0045	1.017	0.059	-3.95	0.50	18.5
62	0.055	0.0045	1.016	0.059	-3.94	0.50	18.5
63	0.055	0.0045	1.016	0.059	-3.93	0.50	18.5
64	0.055	0.0045	1.016	0.059	-3.92	0.50	18.6
65	0.055	0.0045	1.015	0.059	-3.91	0.50	18.6
66	0.054	0.0045	1.015	0.059	-3.9	0.50	18.6
67	0.054	0.0044	1.014	0.059	-3.88	0.50	18.7
68	0.054	0.0044	1.013	0.059	-3.85	0.50	18.7
69	0.054	0.0044	1.011	0.059	-3.81	0.50	18.8

The quality of the transformation method was assessed by comparing the relationships between calculated metrics directly from a computer program that computes noise exposure metrics using single-event data and an aircraft operations scenario. The single-event data is simulated using the Integrated Noise Model (INM) 7.0. This program is called the Noise Metric Module and is described in Chapter 5, Section 5.2.2 and Appendix I.1. In Figure 3.2 are shown plots of the relationships between the Noise Metric Module calculations of DNL , $SEL_{A,av}$, $L_{Amax,av}$, and $PNL_{max,av}$. The relationships between metrics are approximately linear though there is a departure from linearity at the higher sound levels. The results of fitting linear models to the data, not including the higher level regions, are also shown in Figure 3.2. Regression equations are given in the figure caption.

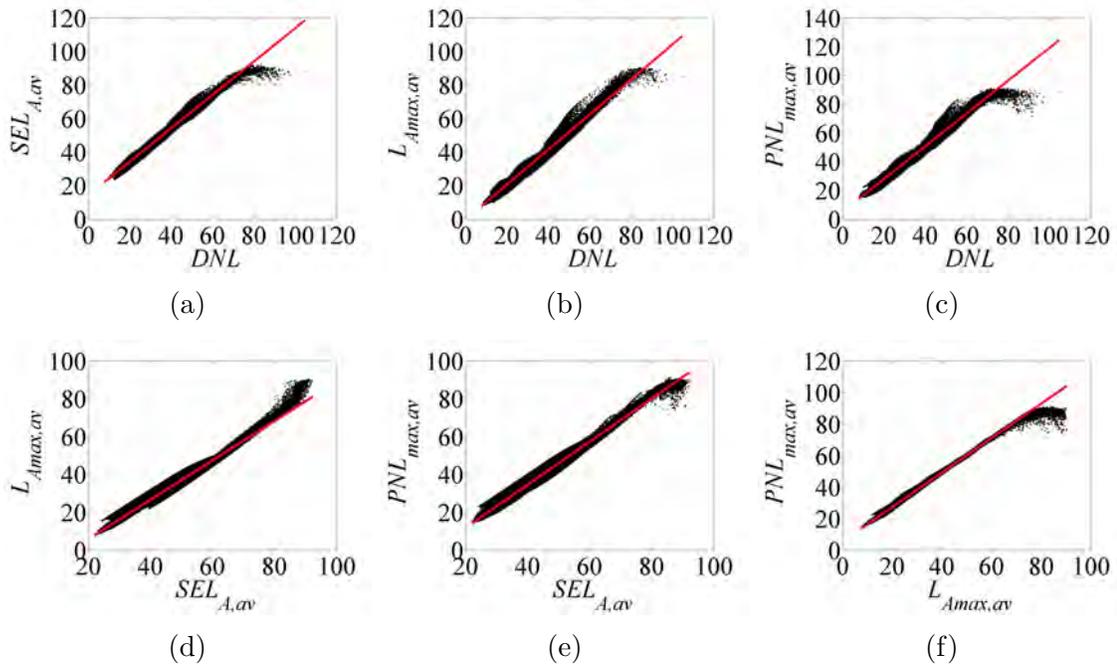


Figure 3.2. Noise metrics calculated using the Noise Metric Module. Each plot contains 650,000 total points from noise metric calculations for simulated operations around three airports. The red lines are linear regression models derived from the linear ranges of the relationships: (a)&(b) DNL 10-80, (c) DNL 10-70, (d)&(e) $SEL_{A,av}$ 20-80, (f) $L_{Amax,av}$ 0-70. The estimated linear relationships are: (a) $SEL_{A,av} = 0.99 DNL + 14.5$, (b) $L_{Amax,av} = 1.04 DNL - 0.3$, (c) $PNL_{max,av} = 1.1 DNL + 5.2$, (d) $L_{Amax,av} = 1.0 SEL_{A,av} - 15.1$, (e) $PNL_{max,av} = 1.1 SEL_{A,av} - 11.1$, (f) $PNL_{max,av} = 1.1 L_{Amax,av} + 5.5$.

The transformation method involves three linear transformations: DNL to $SEL_{A,av}$, $SEL_{A,av}$ to $L_{Amax,av}$, and $L_{Amax,av}$ to $PNL_{max,av}$. The relationships between the predicted metric values and the metric values calculated from the Noise Metric Module individual aircraft event metric data are shown in Figure 3.3. The predictions are reasonably good up to $PNL_{max,av} = 70$ PNdB with an overestimation at the higher levels.

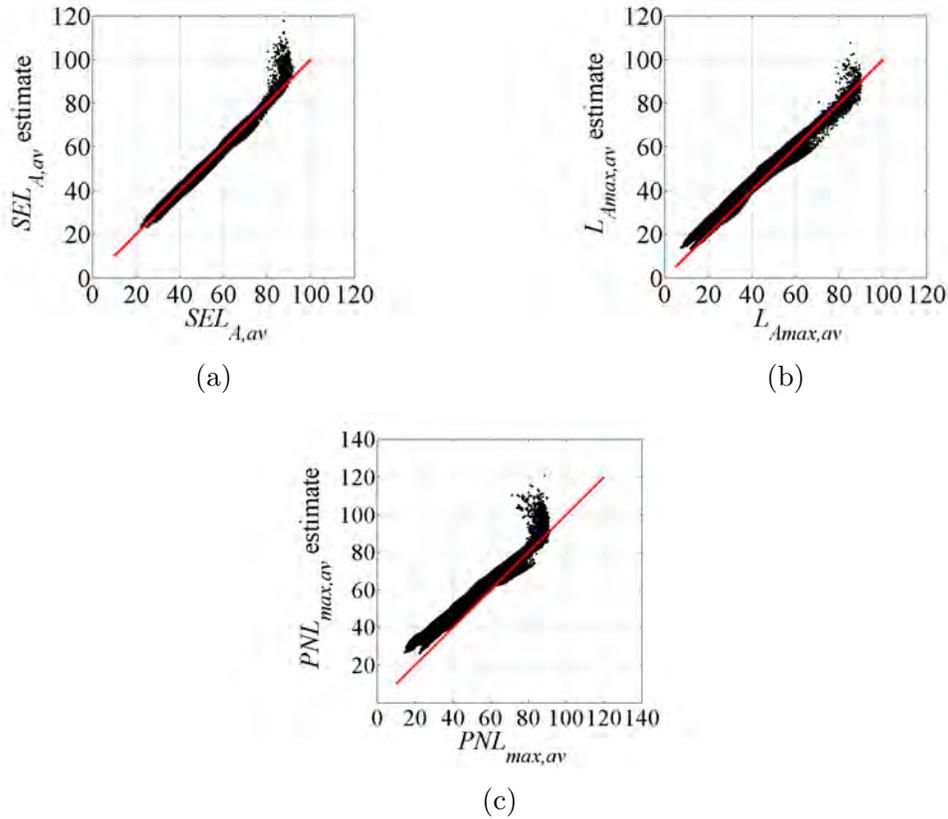


Figure 3.3. Cumulative metrics predicted from other cumulative metrics plotted against those directly calculated from individual event data using the Noise Metric Module (Chapter 5, Section 5.2.2 and Appendix I.1). Each plot contains 650,000 points from single event predictions and simulated airport operations around three US airports. The red reference line has a slope of 1. (a) $SEL_{A,av}$ predicted from DNL . (b) $L_{Amax,av}$ predicted from the predicted $SEL_{A,av}$. (c) $PNL_{max,av}$ predicted from the predicted $L_{Amax,av}$.

3.2 L_{av} to $PNL_{max,av}$

UKD-604, the data set from the 2005 survey in England, contained noise exposure data in terms of L_{av} , which is an arithmetic average of the L_{Amax} of individual aircraft events with a L_{Amax} above 65 dBA. Prior to the collection of the UKD-604 data, the accepted practice had been to report sound level metric data in logarithmic (energy-based) averages. An arithmetic average is an average of the decibel values. A logarithmic average is an energy average, and involves taking the antilog of the

decibel values prior to averaging. In Annex B of Appendix A2 of the UKD-604 report (Le Masurier et al., 2007b), there are tables for 15 geographic sites that contains logarithmic and linear averages for SEL_A for 34 aircraft models for both arrivals and departures. Descriptive statistics were calculated for the difference between the logarithmic and linear averages for arrivals and departures at 4 sites. At each site, the logarithmic average was always greater than the linear average for each aircraft model. For arrivals, the average difference between the two calculations was 0.4 dB. For departures, the average difference was 1.1 dB. For both arrivals and departures, the average difference between the logarithmic and linear averages is quite small considering that a 1 dB difference in sound level is barely noticeable to the average listener. Because the disagreement between arithmetic and logarithmic SEL_A was small, L_{av} values were not modified from reported values. L_{av} was simply increased by 13 dB to approximate $PNL_{max,av}$.

3.3 Simulation of UKD-604

Person-level acoustical data were not available for UKD-604. The only data available for UKD-604 were those reported in the study publication appendices (Le Masurier et al., 2007b). Acoustical data are reported for the population-weighted centroids of geographic areas called Census Output Areas, which were used for the 2001 Census in England . In Table 3 of Appendix A6.3 of Le Masurier et al. (2007b) are shown acoustical data in terms of the number-of-events operating in a 30-day period with a L_{Amax} above 65 dBA (N_{av}), and the arithmetic average of those sound levels (L_{av}). Although acoustical data are given only for one point in an Output Area, the full breakdown of annoyance responses on a 5-point verbal annoyance scale are given for each Output Area in Table 10 of Appendix A6.3 (Le Masurier et al., 2007b).

A person-level UKD-604 data set was simulated from the information available in the publication appendices. The simulation process was designed as an optimiza-

tion algorithm; optimization proceeded until the stopping criteria was fulfilled. The simulation process is shown visually in a flowchart in Figure 3.4 and is as follows:

1. Calculate probable ranges of L_{av} and N_{av} for each Output Area.
2. Specify an arbitrary multiple linear regression annoyance model with L_{av} , N_{av} , and an intercept as predictor variables.
3. Enter the optimization algorithm:
 - (a) For each Output Area:
 - i. Generate arrays of L_{av} (incremented by 0.2 dBA) and N_{av} (incremented by 2 events) over the probable ranges.
 - ii. Generate all possible combinations of L_{av} and N_{av} .
 - iii. Calculate the annoyance for each L_{av} and N_{av} combination using the annoyance model.
 - iv. Rescale the annoyance calculations to have the same mean and standard deviation as the UKD-604 annoyance responses for the Output Area.
 - v. Based on the range and values of the UKD-604 annoyance responses for the Output Area, map the rescaled annoyance calculations to the UKD-604 annoyance responses so that each UKD-604 annoyance response category has an associated group of rescaled annoyance calculations, and thus, L_{av} and N_{av} combinations.
 - vi. For each UKD-604 annoyance response category, and for the number of respondents with that annoyance response, randomly sample with replacement from the associated group of L_{av} and N_{av} combinations so that each respondent is assigned a L_{av}/N_{av} combination.
 - (b) Re-derive the annoyance model using the annoyance responses and the assigned L_{av}/N_{av} combinations.
 - (c) Implement the stopping rule: Halt optimization if each regression coefficient estimate of the annoyance model has not changed by 6% over 20 iterations from the values observed at the beginning of the 20 iterations.

calculations were used to study how L_{av} and N_{av} varied within census block groups at the airports using a computer program described in Appendix I.6. At each airport:

1. The arithmetic mean of $L_{Aeq,d}$ grid point values was calculated for each census block group.
2. The arithmetic average $L_{Aeq,d}$ were used to associate the census block groups with one of six $L_{Aeq,d}$ levels ranging from 40 dBA to 65 dBA in 5 dB increments. Each census block group was associated with the $L_{Aeq,d}$ level that was closest to the arithmetic average $L_{Aeq,d}$.
3. Each Output Area from UKD-604 was also associated with one of the six $L_{Aeq,d}$ values by use of $L_{Aeq,d}$ calculations performed at the population-weighted centroids, as reported in (Le Masurier et al., 2007a, Appendix A6.3, Table 3).
4. The standard deviation of L_{av} ($\sigma_{L_{av}}$) and the normalized standard deviation of N_{av} ($\sigma_{N_{av}}^*$) were calculated for each census block group. For each census block group, $\sigma_{N_{av}}^*$ was calculated by dividing the standard deviation of N_{av} ($\sigma_{N_{av}}$) by the maximum N_{av} . This was done because N_{av} counts were on a linear scale and varied greatly between census block groups.
5. For each $L_{Aeq,d}$ level, the averages of $\sigma_{L_{av}}$ ($\overline{\sigma_{L_{av}}}$) and $\sigma_{N_{av}}^*$ ($\overline{\sigma_{N_{av}}^*}$) were calculated from the data of the associated census block groups.

From a study of the results from all three airports, general trends in $\overline{\sigma_{L_{av}}}$ and $\overline{\sigma_{N_{av}}^*}$ were observed as $L_{Aeq,d}$ varied. As $L_{Aeq,d}$ increased, $\overline{\sigma_{L_{av}}}$ increased and $\overline{\sigma_{N_{av}}^*}$ decreased. In Table 3.3 are shown the established trends for the $L_{Aeq,d}$ groups. In Table 3.4 are shown the assignments of the average standard deviations to the UKD-604 Output Areas.

Table 3.3. Observed trends in the average standard deviation of Lav ($\overline{\sigma_{Lav}}$) and the average normalized standard deviation of Nav ($\overline{\sigma_{Nav}^*}$).

$L_{Aeq,d}$ (dBA) Group	$\overline{\sigma_{Lav}}$	$\overline{\sigma_{Nav}^*}$
40	0.50	17.0
45	0.50	17.0
50	0.75	15.0
55	1.00	14.0
60	1.75	13.0
65	2.50	12.0

Table 3.4. Information about the UKD-604 Output Areas. For each Output Area is given the airport, Lav , Nav , the assigned average standard deviation of Lav ($\overline{\sigma_{Lav}}$), and the assigned average normalized standard deviation of Nav ($\overline{\sigma_{Nav}^*}$). n is the sample size.

Output Area	Airport	n	Lav	$\overline{\sigma_{Lav}}$	Nav	$\overline{\sigma_{Nav}^*}$
H1H	Heathrow	61	75.1	2.50	7431	0.12
H1J	Heathrow	57	70.6	1.75	9114	0.13
H1K	Heathrow	59	72.1	1.75	6513	0.13
H1L	Heathrow	62	74.9	1.75	2641	0.13
H1M	Heathrow	59	72.5	1.75	3795	0.13
H1P	Heathrow	63	68.4	1.00	2620	0.14
H3A	Heathrow	61	69.8	0.75	906	0.15
H3B	Heathrow	63	70.2	0.75	864	0.15
H3C	Heathrow	62	68.2	0.50	171	0.17
H3D	Heathrow	58	68.0	0.75	1059	0.15
H3E	Heathrow	63	68.4	1.00	1054	0.14
H5A	Heathrow	60	69.6	0.75	1028	0.15
H5B	Heathrow	61	69.9	1.00	3210	0.14
H5C	Heathrow	61	70.3	1.75	6815	0.13

Table 3.4. Continued

Output Area	Airport	n	Lav	$\overline{\sigma_{Lav}}$	Nav	$\overline{\sigma_{Nav}^*}$
H5D	Heathrow	61	70.2	1.75	5456	0.13
H5E	Heathrow	63	69.3	0.75	627	0.15
H5F	Heathrow	61	68.5	1.00	4111	0.14
O2A	Luton	60	68.8	0.75	458	0.15
O2C	Leeds	54	69.9	0.75	1109	0.15
O2D	Manchester	60	70.7	1.00	2829	0.14
O2E	Manchester	61	68.0	1.00	2998	0.14
O2F	Manchester	62	68.8	1.00	6554	0.14
O4A	Leeds	60	70.8	1.00	1282	0.14
O4B	Manchester	62	70.4	1.75	7118	0.13
O4C	Manchester	60	73.8	1.75	8319	0.13
O4D	Birmingham	60	68.4	0.75	1555	0.15
O4E	Birmingham	59	72.1	1.00	2742	0.14
O4F	Manchester	59	75.2	2.50	8599	0.12
O4G	Southampton	62	72.3	0.75	755	0.15
O5A	London City	65	80.1	1.75	2223	0.13
O6A	Birmingham	62	81.3	0.50	4333	0.17
O6B	Birmingham	64	76.0	1.75	4333	0.13
O6C	Birmingham	59	77.6	1.75	4333	0.13
O6D	Southampton	21	66.7	0.50	926	0.17
O6E	Southampton	60	66.4	0.50	652	0.17
O6F	Southampton	60	67.6	0.50	16	0.17
R01	Heathrow	16	67.4	0.50	26	0.17
R02	Heathrow	18	67.1	0.50	40	0.17
R03	Heathrow	16	68.5	0.50	72	0.17
R04	Heathrow	16	67.9	0.50	168	0.17
R05	Heathrow	15	69.1	0.75	183	0.15

Table 3.4. Continued

Output Area	Airport	n	Lav	$\overline{\sigma_{Lav}}$	Nav	$\overline{\sigma_{Nav}^*}$
R06	Heathrow	16	69.7	0.50	130	0.17
R07	Heathrow	16	68.3	1.00	484	0.14
R08	Heathrow	16	68.4	0.75	479	0.15
R09	Heathrow	17	68.3	0.50	655	0.17
R10	Heathrow	16	69.0	0.75	789	0.15
R11	Leeds	16	68.4	0.50	253	0.17
R12	Manchester	14	66.8	0.50	70	0.17
R13	Manchester	15	67.8	0.50	34	0.17
R14	Manchester	15	69.4	0.50	97	0.17
R15	Manchester	16	67.8	0.50	45	0.17
R16	Gatwick	15	67.0	0.75	467	0.15
R17	Stansted	15	66.6	0.75	1360	0.15
R18	Manchester	17	67.9	0.50	205	0.17
R19	Manchester	15	67.6	0.50	362	0.17
R20	Manchester	16	68.1	0.50	276	0.17

The average standard deviations shown in Table 3.4 were used to calculate probable ranges for Lav and Nav . For the i^{th} Output Area:

$$r_{Lav,i} = \frac{[10\sqrt{12} \overline{\sigma_{Lav,i}}]}{10}, \quad (3.20)$$

$$r_{Nav,i} = [\sqrt{12} \overline{\sigma_{Nav,i}^*} Nav_i], \quad (3.21)$$

where $r_{Lav,i}$ is the probable range for the Lav of the i^{th} Output Area, $r_{Nav,i}$ is the probable range for the Nav of the i^{th} Output Area, $[\]$ denotes rounding to the nearest integer, $\overline{\sigma_{Lav,i}}$ is the standard deviation of Lav assigned to the i^{th} Output Area, and $\overline{\sigma_{Nav,i}^*}$ is the normalized standard deviation of Nav assigned to the i^{th} Output Area. Lav_i and Nav_i are the values of Lav and Nav simulated at the population-weighted

centroid of the i^{th} Output Area. The annoyance model converged in 120 iterations, changing significantly only over the first iteration from the arbitrary annoyance model.

3.4 Data Visualization: Plotting Program

A program based on graphical user interfaces was created to visualize the three-dimensional sound level, number-of-events, and annoyance data of survey data sets. Rather than plotting each individual observation in a data set, the program groups observations and plots the average point, where each point represents an arithmetic average sound level, an arithmetic number-of-events (logged base-10 or unlogged), and an arithmetic annoyance. Each group is customizable and is formed from an intersection of a range of number-of-events and a range of sound level. In this context of grouping data points, number-of-events and sound level are referred to as “factors” and an intersection of a range of number-of-events and a range of sound level is referred to as a “treatment”. The grouping allows a reduction of visual clutter so that general trends can be observed in data sets. In addition to plotting individual data sets, the program can simultaneously plot multiple data sets. In (Fields, 1984, p. 449-450), trends of average annoyance versus average sound level for different number-of-events groups were shown for different data sets, including UKD-024, UKD-130, and USA-054. The program allows data sets to be plotted with the number-of-events groupings that Fields used.

Two special features of the program can be used to design the boundaries of the factor groups in order to accomplish either an equal number of samples per treatment as closely as possible or a “balanced” sample allocation as closely as possible. A balanced sample allocation is realized when the sample sizes of the treatments are “proportional” and satisfy the following equation:

$$n_{ij} = \frac{n_i \cdot n_j}{n_T}, \quad (3.22)$$

where i is a factor level of the number-of-events, n_i is the sample size in factor level i , j is a factor level of sound level, n_j is the sample size in factor level j , n_{ij} is the sample size of treatment ij , and n_T is the total sample size. Equal-sample sizes is a special case of proportional sample sizes. The concept of factor levels and treatments is shown in Figure 3.5, in which a two-dimensional space spanned by the number-of-events and sound level (L_{Amax}) data is divided into treatments.

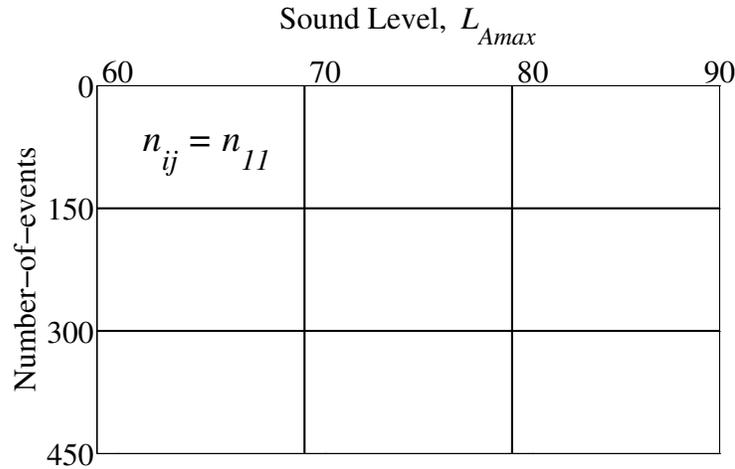


Figure 3.5. Illustration of factor levels and treatments. The continuous variables L_{Amax} and the number-of-events are each divided into factor levels. Combinations of factor levels form treatments.

In grouping the continuous data into treatments, the continuous data become categorical data. Categorical data are usually analyzed with an analysis of variance. An analysis of variance does not involve fitting a line to predictor variable data as in multiple linear regression. Analysis of variance tests if there are differences in a criterion variable (e.g., annoyance) in groups and/or treatments of a data set. Although the plotting program treatments are formed from an intersection of a range of number-of-events and a range of sound level, treatments in a general analysis of variance can be formed with the use of any available information about the sample that would categorize it into different groups. In an analysis of variance, the treatment means of the criterion variable are used to analyze differences between factor levels

and/or treatments. The analysis of variance “factor effects” two-factor model can be expressed as:

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \quad (3.23)$$

where:

$\mu_{..}$ is a constant

α_i are factor effects for factor A subject to the constraint $\sum \alpha_i = 0$

β_j are factor effects for factor B subject to the constraint $\sum \beta_j = 0$

$(\alpha\beta)_{ij}$ are interaction effects constants subject to the constraints:

$$\sum_{i=1}^a (\alpha\beta)_{ij} = 0 \quad j = 1, \dots, b$$

$$\sum_{j=1}^b (\alpha\beta)_{ij} = 0 \quad i = 1, \dots, a$$

ϵ_{ijk} are independent and zero-mean normally distributed errors

$i = 1, \dots, a; j = 1, \dots, b; k = 1, \dots, n$

a is the number of factor levels for factor A

b is the number of factor levels for factor B

n is the number of samples

The estimator of each parameter in Equation (3.23) is a function of treatment and factor level means, where each mean is a mean of criterion variable data. Kutner, Nachtsheim, Neter, and Li (2005, p. 834-842) provides further detail on the estimation of Equation (3.23) and the calculation of the factor and interaction sums of squares when the sample sizes are equal. A sum of squares represents the marginal contribution of a factor toward the prediction of a criterion variable. With equal sample sizes, the total model sum of squares can be perfectly partialled into unique components attributable to each factor and the interaction. The sums of squares sum to the total model sum of squares and are referred to as “orthogonal”.

In the case that the treatment sample sizes are not equal, the factor and interaction sums of squares as calculated with the analysis of variance equations from Kutner et al. (2005, p. 834-842) are not orthogonal. Nonorthogonality of the factors in an analysis of variance is akin to the problem of correlation between predictor variables in

a multiple linear regression analysis. analysis of variance and multiple linear regression are special cases of the “general linear model” (Cohen and Cohen, 2005). Multiple linear regression is more general in its application and can be applied when an analysis of variance cannot. Thus, in the case of unequal sample sizes, a regression formulation of Equation (3.23) is required to estimate the proper sum of squares for the factors and interaction (Cohen and Cohen, 2005, p. 4)(Kutner et al., 2005, p. 953).

In the case that the sample sizes are proportional according to Equation (3.22), the sums of squares are orthogonal with the use of a weighting procedure applied to the regression matrix (Keren and Lewis, 1993, p. 94-127). The plotting program uses this weighting procedure and the regression formulation to estimate Equation 3.23 and calculate the sum of squares for the factors and interaction. The Type III sum of squares is calculated, which is Overall and Spiegel’s method I (Overall and Spiegel, 1969). A Type III sum of squares represents the unique contribution of a factor (e.g., sound level or number-of-events) to the prediction of a criterion variable.

In the case of a balanced sample allocation, the plotting program uses statistical concepts to specify minimum sample sizes for the n_i , n_j , and n_{ij} to guide the allocation process. Two statistical methods are used: the power method and the estimation method. The power method establishes the minimum-allowable factor level sample sizes, the estimation method establishes the minimum-allowable treatment sample sizes. Both methods are described in (Kutner et al., 2005, p. 862-864).

At an early stage of the research, the utility of the balanced sample allocation approach was investigated as a means by which the unique and orthogonal sums of squares attributable to sound level and number-of-events in the prediction of annoyance could be determined. The sums of squares are important because they can be used to establish the relative importance of factors to the prediction of a criterion variable. The concept of relative importance is described in detail in Chapter 4, Section 4.1.1. However, further research (Harrell Jr., 2012) halted the pursuit of balanced sample allocation as applied to survey data. For example, the categorization of the continuous variable data into treatments assumes that the relationship between the

data and the criterion variable is flat within treatments. Also, categorization assumes there is a discontinuity in the criterion variable as treatment boundaries are crossed. Finally, the treatment boundaries are arbitrary and manipulatable. In light of these observations, models derived from categorized survey data were not investigated.

Figure 3.6 shows sample output from the plotting program for three data sets (UKD-024, USA-054, and AUL-210). Appendix I.16 contains a description of the program code and the storage location of the code.

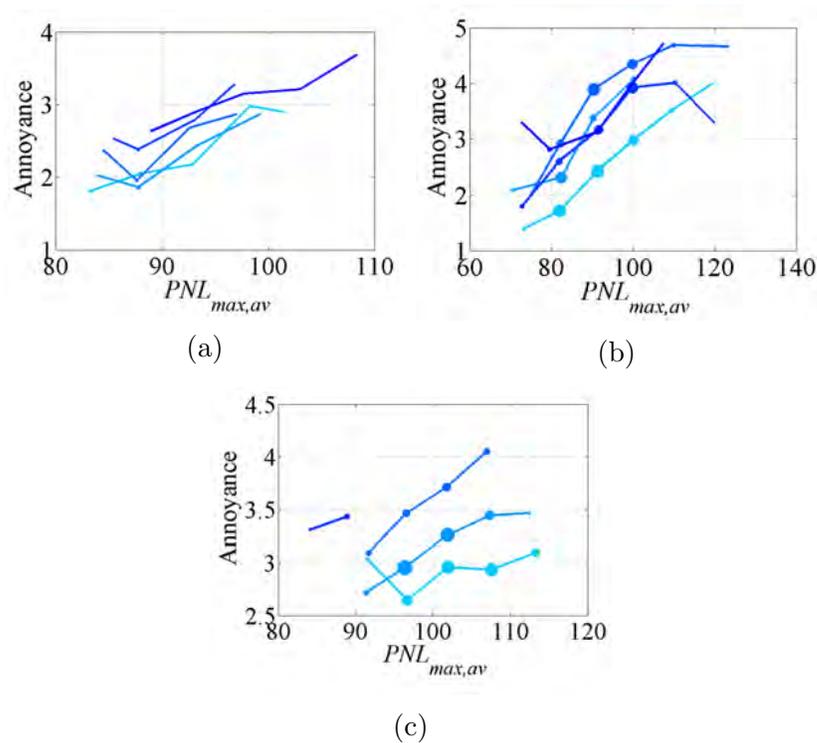


Figure 3.6. Each plot shows average annoyance vs. average Maximum Perceived Noise Level ($PNL_{max,av}$). Separate lines in each plot highlight different number-of-events groups. Each line is marked by a certain shade of blue. Darker shades of blue indicate larger number-of-events groups. The progression of number-of-events groups is: (a) 0-24, 25-49, 50-99, 100-199, 200-488; (b) 0-49, 50-99, 100-199, 200-399, 400-504; (c) 0-34, 34-68, 68-102, 102-134.

3.5 Data Set Combination

All of the individual noise survey data sets were combined and analyzed as one. Seven combination data sets were constructed. The combination data sets were created in order to study the regression relationships predicted from the data and the sensitivity of the results to combination method. The combination was performed using a computer program. Appendix I.13 contains a description of the program code and the storage location of the code.

3.5.1 Annoyance Response Combination

As described earlier, a common sound level metric, $PNL_{max,av}$, was established among the data sets. In a similar manner, a common annoyance scale had to be established among all data sets. All data sets featured a 4-point or 5-point annoyance scale that was either verbal, numerical, or numerical with verbal endpoints. Activity interference ratings, historically constructed from summing interference ratings over a collection of activities (e.g., conversation, listening to television, or reading), have often been used to predict annoyance. However, these were not used because different activities were used in constructing the ratings in separate surveys.

Midpoint Miedema Scaling

Miedema and Vos (1998) describe a process in which numerical annoyance scales are rescaled to span a continuum from 0 to 100. The rescaling is based on the following assumptions:

1. Equal intervals: Each category from a set of response alternatives occupies an equal portion of the annoyance continuum
2. Equal extremes: The extreme lower and upper category boundaries from different sets of annoyance response alternatives coincide

In this rescaling process, each annoyance category is mapped to a lower and an upper bound. The boundaries are calculated by:

$$b_i = \frac{100i}{m}, \quad (3.24)$$

where b_i is a boundary, m is the number of categories, and $i = 0, 1, \dots, m$ is the index of the boundary, starting with the lower boundary of the lowest annoyance category. By this rescaling, a single annoyance category is mapped to a lower and upper bound so that the category spans a range of the continuum from 0 to 100. This rescaling was slightly modified so that each annoyance category would be mapped to a single number. Specifically, the rescaled annoyance number was chosen to be the midpoint of the lower and upper bounds:

$$a_j = \frac{100j}{m} + \frac{100}{2m}, \quad (3.25)$$

where a_j is a rescaled annoyance response, m is the number of categories, and $j = 1, 2, \dots, m$ is the index of the annoyance category. One combination data set resulted from using the midpoint Miedema scaling.

Category Combination

Because in some data sets the annoyance scales were 4-point scales, the 5-point scales were required to be reduced to 4-point scales. This was accomplished in different ways. Different combination data sets were created by combining either 1 and 2, 2 and 3, 3 and 4, or 4 and 5. The values above the combined values were then reduced to span the range from 1 to 4. Four combination data sets resulted from using the category combination method.

Rescaling and Rounding

This method was applied to both 4-point and 5-point scales. If applied to a 4-point scale, the scale was expanded to range from 1 to 5. If applied to a 5-point scale, the scale was reduced to range from 1 to 4. The rescaling is performed as follows:

$$a_{r,j} = [\max(a_r) - \min(a_r)] \frac{a_j - \min(a)}{\max(a) - \min(a)} + \min(a_r), \quad (3.26)$$

where a_r is the rescaled annoyance scale, $a_{r,j}$ is the j^{th} rescaled annoyance response, a_j is the j^{th} unscaled annoyance response, $\max(a_r)$ is the maximum of the rescaled annoyance scale, $\min(a_r)$ is the minimum of the rescaled annoyance scale, $\max(a)$ is the maximum of the unscaled annoyance scale, and $\min(a)$ is the minimum of the unscaled annoyance scale. The reduction of a scale using this method was computationally equivalent to the combination of categories 3 and 4 with the category combination method. The expansion of a scale using this method mapped 1 to 1, 2 to 2, 3 to 4, and 4 to 5. That is, the absence of a midpoint in the scale was preserved through the rescaling process. Two combination data sets resulted from using the rescaling and rounding method.

3.6 Summary

Seven historic data sets were acquired and processed for research. The sound level data of two data sets were transformed so that a common sound level metric, $PNL_{max,av}$, could be established among all of the data sets. A process for transforming DNL to $PNL_{max,av}$ was outlined. A person-level UKD-604 data set was simulated from site-level acoustical and annoyance data reported in (Le Masurier et al., 2007b). A data set plotting program was described, examples of data set visualizations were provided. The motivation for combining data sets was presented and the various methods for establishing common annoyance scales among the data sets were described. Although major interest lies in the analysis of combination data sets, a reanalysis of the indi-

vidual data sets and comparison to prior results (Fields, 1984) is beneficial to serve as a an indicator that the data sets are equivalent to those used before and as a calibration of computer programs that were created to analyze the data sets.

CHAPTER 4. METHODOLOGIES IN MODEL DEVELOPMENT
AND PARAMETER ESTIMATION

In this chapter analyses of seven survey data sets that were introduced in the last chapter are presented. Five of the seven data sets (UKD-024, UKD-130, USA-022, USA-032, and USA-044) were used by Fields (1984). The results of the original analyses of these data sets from Fields served to guide the creation and use of data processing methodologies, so that Fields' results could be reproduced. The goal in doing this was to illustrate that the data sets and analysis methods were equivalent. This was important because a goal of the current research was to build on the results from Fields' study and analyze data sets composed of multiple individual data sets using the same techniques. Results from an analysis of individual survey data sets will be given accompanied with a description of the techniques used to generate the results. Then, results from an analysis of a combination data set will be presented.

4.1 Individual Data Set Results and Methodologies

Basic information about the individual data sets is shown in Table 4.1. The annoyance scale labels are from Table D.1 in Appendix D and n is the sample size.

Table 4.1. Basic information about the individual data sets, separated by airport.

Study/airport	Date	Annoyance	n
AUL-210/Adelaide (ADL)	6/5/80 - 7/3/80	VRB5	557
AUL-210/Melbourne (MLB)	7/24/80 - 8/7/80	VRB5	255
AUL-210/Perth (PER)	5/8/80 - 6/5/80	VRB5	467
AUL-210/Richmond (XRH)	2/20/80 - 3/26/80	VRB5	246
AUL-210/Sydney (SYD)	2/20/80 - 4/2/80	VRB5	1,195

Table 4.1. Continued

Study/airport	Date	Annoyance	n
USA-022/Chicago (ORD)	5/8/67 - 8/3/67	NUMVRB5	865
USA-022/Dallas (DFW)	5/9/67 - 7/6/67	NUMVRB5	908
USA-022/Denver (DEN)	5/24/67 - 7/10/67	NUMVRB5	990
USA-022/Los Angeles (LAX)	5/31/67 - 8/3/67	NUMVRB5	774
USA-032/Boston (BOS)	7/6/69 - 9/5/69	NUMVRB5	1,166
USA-032/New York (JFK)	5/13/69 - 9/12/69	NUMVRB5	676
USA-032/Miami (MIA)	3/17/69 - 6/21/69	NUMVRB5	1,060
	9/7/69 - 11/15/69		
USA-044/Chattanooga (CHA)	11/16/70 - 12/19/70	NUMVRB5	1,114
	1/2/71 - 1/31/71		
USA-044/Reno (RNO)	10/26/70 - 12/19/70	NUMVRB5	846
UKD-024/Heathrow (LHR)	9/8/67 - 10/2/67	VRB4	4,655
UKD-130/Heathrow (LHR)	11/21/76 - 1/15/77	VRB4	2,618
UKD-604/Birmingham (BHX)	10/5/05 - 1/6/06	VRB5	298
UKD-604/Gatwick (LGW)	11/5/05 - 2/6/06	VRB5	15
UKD-604/Heathrow (LHR)	8/5/05 - 2/6/06	VRB5	1,189
UKD-604/Leeds-Bradford (LBA)	9/5/05 - 2/6/06	VRB5	127
UKD-604/London City (LCY)	11/5/05 - 1/6/06	VRB5	65
UKD-604/Luton (LTN)	9/5/05 - 10/5/05	VRB5	60
UKD-604/Manchester (MAN)	9/5/05 - 2/6/06	VRB5	465
UKD-604/Southampton (SOU)	10/5/05 - 1/6/06	VRB5	203
UKD-604/Stansted (STN)	11/5/05 - 2/6/06	VRB5	15

Regression models were estimated in the form of:

$$\hat{A} = \hat{b}_L L + \hat{b}_N \log_{10} N + \hat{b}_0, \quad (4.1)$$

where \hat{A} is the annoyance variable estimate, L is the sound level variable, \hat{b}_L is the sound level coefficient estimate, N is the number-of-events variable, \hat{b}_N is the number-of-events coefficient estimate, and \hat{b}_0 is the intercept estimate. Models of this form are referred to as “additive-log” models.

In Table 4.2 are shown the results from the analyses of the individual data sets. $\hat{k} = \hat{b}_N/\hat{b}_L$ is referred to as the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the estimated complex (accounting for sample design) standard deviation of \hat{k} generated by using both Jackknife repeated replication and Taylor series linearization (in parentheses). Estimated 95% confidence intervals for k generated by use of the jackknife standard deviation estimate are provided. *DEFT* is the Design Factor, which is an inflation factor for the standard deviation estimate when a complex sample design is used rather than a simple random sample design. The *DEFT* was calculated using the jackknife method. R^2 is the coefficient of determination for the model. The results will be explained in the following sections.

4.1.1 Relative Variable Importance

Multiple regression analysis has two distinct applications: prediction and explanation (Courville and Thompson, 2001; Johnson and LeBreton, 2004). In the explanatory case, the interest is in the extent to which each predictor variable contributes to the prediction of the criterion variable, or the importance of each predictor variable. A goal of the current research is to determine, if possible, the relative importance of the number-of-events and sound level to the prediction of annoyance.

There are many measures of variable importance (Bring, 1994; Johnson and LeBreton, 2004; Kruskal and Majors, 1989). Researchers often examine regression coefficient estimates or zero-order correlations of predictor variables with a criterion variable to gauge importance. When predictors are uncorrelated, zero-order correlation coefficients and standardized regression coefficient estimates (beta coefficients)

Table 4.2. Individual data set results for models of annoyance that are a linear combination of $PNL_{max,av}$ and $log_{10}(N)$. $DEFT$ is the Design Factor and the estimated 95% confidence intervals are given for k . Standard deviation estimates of the regression coefficient estimates are given in parentheses.

Study/airport	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{\hat{k},c}$	$DEFT$	95% CI for k	R^2
AUL-210/ADL	0.055 (0.011)	0.729 (0.106)	-3.633 (1.280)	13.15	1.94 (2.55)	0.88	8.97 to 17.34	0.08
AUL-210/MLB	0.016 ¹ (0.032)	-0.818 ¹ (1.262)	2.675 ¹ (3.202)	-49.96	12054 (118)	109	-27318 to 27219	0.00
AUL-210/PER	0.039 (0.011)	0.686 (0.142)	-2.165 ¹ (1.142)	17.48	6.21 (3.02)	1.30	3.43 to 31.53	0.06
AUL-210/XRH	0.062 (0.020)	4.330 ¹ (3.361)	-10.430 ¹ (5.620)	69.80	81.39 (62.80)	1.34	-114.33 to 253.93	0.05
AUL-210/SYD	0.063 (0.007)	1.583 (0.227)	-5.608 (0.905)	24.97	3.94 (3.43)	1.15	16.86 to 33.09	0.08
USA-022/ORD	0.083 (0.005)	0.683 (0.211)	-5.966 (0.664)	8.24	5.42 (5.34)	2.03	-3.26 to 19.73	0.23
USA-022/DFW	0.055 (0.007)	0.045 ¹ (0.257)	-2.141 (0.476)	0.83	19.69 (16.85)	4.17	-45.74 to 47.40	0.11
USA-022/DEN	0.086 (0.005)	-0.260 (0.068)	-5.097 (0.403)	-3.02	1.50 (2.50)	2.01	-6.41 to 0.37	0.25
USA-022/LAX	0.066 (0.005)	0.404 ¹ (0.234)	-3.371 (0.737)	6.10	6.38 (6.86)	1.79	-7.67 to 19.88	0.17

Table 4.2. Continued

Study/airport	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{k,c}$	DEFT	95% CI for k	R^2
USA-032/BOS	0.056 (0.005)	0.586 (0.254)	-2.435 (0.649)	10.51	10.52 (8.27)	2.08	-10.55 to 31.56	0.11
USA-032/MIA	0.071 (0.007)	0.458 ¹ (0.244)	-4.316 (0.522)	6.43	6.31 (6.53)	1.52	-6.41 to 19.27	0.23
USA-032/JFK	0.040 (0.005)	0.562 (0.190)	-0.779 ¹ (0.523)	14.05	6.17 (6.98)	1.01	1.78 to 26.32	0.09
USA-044/CHA	0.048 (0.006)	0.495 (0.091)	-2.738 (0.498)	10.25	2.50 (3.24)	1.05	5.25 to 15.25	0.14
USA-044/RNO	0.051 (0.005)	0.221 ¹ (0.262)	-2.481 (0.648)	4.36	3.83 (4.17)	0.74	-3.28 to 12.01	0.10
UKD-024/LHR	0.069 (0.003)	0.220 (0.025)	-4.278 (0.272)	3.18	0.57 (0.77)	1.40	2.05 to 4.31	0.13
UKD-130/LHR	0.063 (0.005)	0.042 ¹ (0.087)	-3.496 (0.375)	0.66	1.32 (2.82)	0.93	-2.02 to 3.33	0.10
UKD-604/BHX	0.028 (0.011)	6.629 (0.267)	-14.319 (0.919)	234.01	808 (345)	9.68	-1629 to 2097	0.72

Table 4.2. Continued

Study/airport	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{k,c}$	DEFT	95% CI for k	R^2
UKD-604/LGW	0.437 ¹ (0.279)	2.629 (0.604)	-38.299 ¹ (22.440)	6.01	2.30 (3.80)	0.71	-1.31 to 13.34	0.63
UKD-604/LHR	0.188 (0.009)	3.100 (0.077)	-19.310 (0.732)	16.47	3.30 (2.79)	2.36	9.90 to 23.04	0.70
UKD-604/LBA	0.402 (0.055)	3.004 (0.241)	-37.366 (4.514)	7.47	18.05 (3.74)	9.14	-33.37 to 48.31	0.66
UKD-604/LCY	0.258 (0.023)	6.177 (0.198)	-34.625 (2.195)	23.92	1.52 (8.58)	0.70	17.37 to 30.48	0.95
UKD-604/LTN	0.218 ¹ (0.117)	3.199 (0.259)	-21.760 (9.726)	14.70	9.65 (20.75)	1.28	-12.09 to 41.50	0.74
UKD-604/MAN	0.140 (0.017)	2.655 (0.195)	-14.800 (1.287)	18.94	10.54 (8.93)	3.45	-2.58 to 40.47	0.47
UKD-604/SOU	-0.009 ¹ (0.024)	1.892 (0.153)	-0.903 ¹ (1.915)	-221.99	36199 (2178)	63.68	-80878 to 80434	0.44
UKD-604/STN	0.216 ¹ (0.133)	4.868 (0.352)	-23.882 ¹ (10.993)	22.56	39.10 (0.0 + 6.7i)	3.32	-101.87 to 147.00	0.95

¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.

are equivalent. The sum of the squares of either the zero-order correlations or the beta coefficients is equal to the regression model R^2 , and the relative importance of each variable can be expressed as a proportion of R^2 . However, when predictor variables are correlated, squared correlations coefficients and squared beta coefficients are no longer equivalent, do not sum to R^2 , and take on very different meanings. A zero-order correlation is a measure of linear dependence between two variables. A beta coefficient indicates how many standard deviations a criterion variable will change per standard deviation increase in a predictor variable. The magnitude and sign of the beta coefficient is affected by the inclusion of the other predictor variables.

Johnson and LeBreton (2004) give a comprehensive review of importance measures. They described three categories of importance measures:

1. Single-analysis methods: Single-analysis methods use the output from a single regression analysis. Importance is equal to a single regression metric or a combination of metrics.
2. Multiple-analysis methods: Multiple-analysis methods utilize the results from multiple stepwise regression analyses.
3. Variable transformation methods: Variable transformation methods transform the original predictor variables to a set of uncorrelated variables, regress the criterion variable on the uncorrelated predictors, and either use those results as a proxy for inferring the importance of the original predictor variables or further analyze the data to yield results that are associated directly with the untransformed predictor variables.

Two measures of variable importance will be used. The decibel equivalent number effect (\hat{k}), introduced in Table 4.2, is a single-analysis method and has historically been used to express the importance of $\log_{10}(N)$ relative to sound level. The estimation of \hat{k} is addressed in Sections 4.1.2 and 4.1.3. The second method is a multiple-analysis method and is addressed in Section 4.1.4.

4.1.2 Decibel Equivalent Number Effect, \hat{k}

Fields (1984) evaluated the effect of the number-of-events in a data set as a ratio of unstandardized partial regression coefficient estimates. He called this ratio the decibel equivalent number effect ($\hat{k} = \hat{b}_N/\hat{b}_L$), which is the number of decibels which increases annoyance by the same amount as a tenfold increase in the number-of-events. A value of $\hat{k} = 10$ indicates that the equal-energy hypothesis applies. That is, when $\hat{k} = 10$, a proportional increase or decrease in either the sound energy or number-of-events produces an equal shift in the annoyance index value. An estimate of k though is meaningless without a measure of the precision with which it can be determined, as expressed by the confidence interval. If the confidence interval includes $k = 0$, the number-of-events cannot be stated at the confidence level to be significant in the prediction of annoyance. If the confidence interval encompasses a range of values that exceeds and does not include $k = 10$, the number-of-events can be stated at the confidence level to have a greater effect on the annoyance index than sound energy. The reverse is true if the confidence interval includes a range of values that is less than and does not include $k = 10$.

Examination of the data in Table 4.2 shows that many \hat{k} cannot be determined at 95% confidence to be significant ($k > 0$) or significantly different from $k = 10$. This is due to the large standard deviations of \hat{k} . Seventeen of the twenty-five data sets do not produce a significant \hat{k} . Of the remaining eight data sets, five \hat{k} cannot be stated with 95% confidence to be significantly different from 10. Only two data sets (AUL-210/Sydney (SYD) and UKD-604/London City (LCY)) feature $k > 10$ with 95% confidence, while one data set (UKD-024/Heathrow (LHR)) features $0 < k < 10$ with 95% confidence. There is also a lack of agreement; the value of \hat{k} varies from data set to data set. There is disagreement even among three \hat{k} from the different studies at Heathrow. The \hat{k} of UKD-024 and UKD-130 cannot be stated with 95% confidence to be different from one another, but both are significantly different from UKD-604/Heathrow. In some data sets b_L cannot be stated with 95% confidence to

be different from 0. Because $\hat{k} = \hat{b}_N / \hat{b}_L$, caution should be exercised when considering \hat{k} for these data sets.

4.1.3 Standard Deviation of \hat{k} , $\hat{\sigma}_{\hat{k},c}$

Sample design affects variance calculations and must be incorporated into data analysis (Fields, 1980; Kalton, 1983; Sul Lee, Forthofer, and Lorimor, 1989). The regular variance or standard deviation formulae found in statistics texts usually relate only to unrestricted sampling (simple random sampling with replacement). Sampling “with replacement” means that after each sample is selected, it is not prevented from being sampled again. That is, the population size from which samples are selected is never decreased and samples can be selected more than once. Normally in practice, sampling is performed without replacement so that once a sample is selected, it cannot be selected again. This practice does not completely invalidate the unrestricted sampling formulae. Sampling without replacement introduces a correction factor into the unrestricted sampling formulae called the finite population correction factor, or FPC. However, if the population size is large relative to the size of the collected sample, the FPC can be ignored, and it normally is in practice. In historical surveys, more complex sample designs, such as multistage clustered sample designs (typically two-stage) using stratification and clustering (Kalton, 1983; Tomberlin, 1985) were more common than simple random sample designs. Clustering is a method of grouping a population in which the population of people is grouped into clusters (usually geographic areas), which become units for sampling. Since these sampling units are first-stage groups of the population, they are referred to as “primary sampling units”. In clustered sampling, only a *subset* of the clusters is included in the survey sample. Another method of grouping the population prior to sampling is called stratification. Stratification is the classification of a population into groups, or strata, based on available information about the population. Samples are then further selected from *each* of the strata. Stratification and clustering are described further in Chapter 5,

Section 5.1 and Appendix B. With these complex sample designs, the unrestricted formulae are not valid. The use of the unrestricted formulae with data resulting from a complex sample design may yield calculations that overestimate, or more commonly, underestimate the true variance (Fields, 1980; Kalton, 1983; Sul Lee et al., 1989). In addition, variance formulae for some statistics (such as \hat{k}) do not exist.

With complex sample designs, other methods of variance estimation can be used, e.g., the jackknife repeated replication method and the Taylor series linearization method. Both methods can be used to estimate the variance of a ratio of regression coefficient estimates for a population in which all study areas are assumed to have the same *true* partial regression coefficients. These methods are described in detail in Appendix B and an overview of the methods follows.

Jackknife Repeated Replication

With jackknife repeated replication, an estimate of a statistic of interest (here, \hat{k}) is calculated in subsamples (called replicates) of the entire sample. The jackknife variance estimate is a function of the squared differences between the replicate estimates and the mean of the estimates. Typically, a replicate is created by removing one primary sampling unit from the sample. For the individual data sets given in Table 4.2 the primary sampling unit is a geographical area in the proximity of an airport which is exposed to aircraft noise and in which people reside. The jackknife coefficient is a weighting factor that is used in the variance calculation and can be constructed to equally-weight primary sampling units or people. The standard deviations given in Table 4.2 and those reported in (Fields, 1984) were generated using equally-weighted primary sampling units. The jackknife variance estimate of \hat{k} is calculated by:

$$\hat{\sigma}^2(\hat{k}) = \sum_{r=1}^R \alpha_r \left(\hat{k}_r - \bar{\hat{k}} \right)^2, \quad (4.2)$$

where \hat{k}_r is an estimate of k in the r^{th} replicate, $\bar{\hat{k}}_r$ is the mean of the \hat{k}_r , R is the number of replicates, and α_r is the jackknife coefficient of the r^{th} replicate, defined by:

$$\alpha_r = \frac{p_i - 1}{p_i}, \quad (4.3)$$

where p_i is the number of primary sampling units in stratum i . Thus, each stratum has an associated jackknife coefficient. Further description of jackknife repeated replication is in Appendix B.1.

Taylor Series Linearization

Taylor series linearization is a method of representing a mathematical function as an infinite sum of terms based on the function's derivatives which are evaluated at a single point. Taylor series linearization can be applied to collected survey data and is also useful in planning surveys (Fields, 1984).

The process of calculating the Taylor variance estimate of \hat{k} for survey data collected with a complex sample design involves the estimation of variance components. The variance component estimates are calculated via a random effects single-factor analysis of variance on the residuals resulting from a multiple linear regression analysis, where primary sampling unit is the random factor. The residuals are the differences between the actual annoyance responses and the annoyance predictions from the estimated regression model. The variance component estimation partitions the variance of the residuals into two components: the variability that can be accounted for by primary sampling unit (the variance "between" primary sampling units) and the remaining variability not accounted for by primary sampling unit (the variance "within" primary sampling units). The between primary sampling unit variance component estimate is an estimate of the true variance of the average annoyances from the primary sampling units (σ_b^2). The within primary sampling unit variance component estimate is an estimate of the true variance of the annoyance of persons that

is not accounted for by differences in primary sampling unit (σ_w^2). The Taylor series linearization estimate of k is calculated by:

$$\hat{\sigma}^2(k) = \frac{\hat{b}_N^2}{\hat{b}_L^2} \left(\frac{\hat{\sigma}^2(b_N)}{\hat{b}_L^2} + \frac{\hat{\sigma}^2(b_L)}{\hat{b}_L^2} - \frac{2\hat{\sigma}(b_N, b_L)}{\hat{b}_N \hat{b}_L} \right), \quad (4.4)$$

where:

$$\hat{\sigma}^2(b_N) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_L^2 \sigma_{\log_{10}(N)}^2 - (\sigma_{L, \log_{10}(N)})^2} \right) (\sigma_L^2), \quad (4.5)$$

$$\hat{\sigma}^2(b_L) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_L^2 \sigma_{\log_{10}(N)}^2 - (\sigma_{L, \log_{10}(N)})^2} \right) (\sigma_{\log_{10}(N)}^2), \quad (4.6)$$

$$\hat{\sigma}(b_N, b_L) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_L^2 - \sigma_{\log_{10}(N)}^2 (\sigma_{L, \log_{10}(N)})^2} \right) (-\sigma_{L, \log_{10}(N)}), \quad (4.7)$$

where $\sigma_{\log_{10}(N)}^2$ is the variance of the $\log_{10}(N)$ data, σ_L^2 is the variance of the sound level (L) data, $\sigma_{L, \log_{10}(N)}$ is the covariance of $\log_{10}(N)$ and L , n is the number of samples, α is the number of primary sampling units, $\hat{\sigma}_b^2$ is the between primary sampling unit variance component estimate, and $\hat{\sigma}_w^2$ is the within primary sampling unit variance component estimate. Further description of Taylor series linearization and the estimation of the variance components is in Appendix B.2.

Comparison of Jackknife Repeated Replication and Taylor Series Linearization

In past work the properties of different methods of variance estimation have been discussed, see, e.g., (Fuller, 2009; Lohr, 1999; Rust, 1989; Wolter, 2003). Neither jackknife repeated replication nor Taylor series linearization is recommended as generally better than the other. Taylor series linearization is well-developed for many different common survey designs, and the resampling methods such as the jackknife can also be applied in many situations. However, the Taylor series linearization formula must be estimated for each statistic, whereas the same jackknife algorithm is applicable in all situations. If the sample is restricted in size, the Taylor estimate

of variance is often biased downward (Lohr, 1999). For nonlinear functions of the mean, in many cases the jackknife estimate of variance is larger than the Taylor estimate (Fuller, 2009) and tend to always be biased upward (Efron and Stein, 1981). With some survey data sets, confidentiality requirements might preclude the release of stratum and cluster information, but typically the jackknife replicate weights are not restricted and survey respondents are still protected from identification. In these scenarios, Taylor series linearization cannot be appropriately applied. Though both jackknife repeated replication and Taylor series linearization can be applied on collected survey data, Taylor series linearization is also often applied in the planning stage of a survey to forecast expected variances of estimators.

From an examination of the results shown in Table 4.2, neither jackknife nor Taylor estimates of standard deviation were consistently less than or greater than the other. They were closer in magnitude for data sets with larger sample sizes. With data sets that yielded very large \hat{k} (AUL-210/Melbourne (MLB), AUL-210/Richmond (XRH), UKD-604/Birmingham (BHX), and UKD-604/Southampton (SOU)) the jackknife standard deviation estimates of \hat{k} were also very large. In these data sets, the replicate estimates' of k (\hat{k}_r) varied greatly from the full sample estimate of \hat{k} so that squared differences between the two were large (see Equation (B.4)). The large values of \hat{k} and to a lesser degree, the small values of \hat{b}_L contributed to the also-large Taylor estimates for these data sets. For UKD-604/Stansted (STN), the Taylor estimate of standard deviation was an imaginary number. This occurred because $\hat{\sigma}_b^2$ was a negative number. This is a possibility because of the nature of the random effects single-factor analysis of variance procedure. From Equation (B.20), if the treatment mean square is smaller than the error mean square, then $\hat{\sigma}_b^2$ can be negative, which is an indication that primary sampling unit does not account for differences in annoyance responses very well.

Wolter (2003, p. 3) describes that the accuracy of variance estimators can be assessed in a number of ways, i.e., the error mean square and the quality of confidence interval estimates. Frequently, the estimator of variance with smaller error mean

square provides poorer confidence interval estimates than other variance estimators with larger error mean square. Wolter (2003, p. 306) states that since the most important purpose of a variance estimator will usually be for estimating confidence intervals for a parameter of interest, as is the case here, the confidence interval criterion is more important than the error mean square. Based on the analysis of five Monte Carlo studies, Wolter (2003, p. 316) preferred jackknife repeated replication to Taylor series linearization when the quality of confidence interval estimates is of prime importance. In light of these findings and to remain consistent with past research (Bullen and Hede, 1986b; Dempsey, Stephens, Fields, and Shepherd, 1983; Fields and Walker, 1982), the jackknife standard deviation estimate was used to estimate confidence intervals for k .

Design Effect and Design Factor

The Design Effect (*DEFF*) and the Design Factor (*DEFT*) are inflation factors of the estimated variance or standard deviation, respectively, for a complex sample design over an simple random sample design. The Design Effect is a ratio of variance estimates, the Design Factor is a ratio of standard deviation estimates. The Design Factor is the square root of the Design Effect. In the Design Effect (or Design Factor), the numerator is the variance (or standard deviation) estimate of a statistic calculated using complex sample design information (strata and primary sampling unit identifiers) with either jackknife repeated replication or Taylor series linearization and the denominator is the variance (or standard deviation) estimate of a statistic calculated assuming simple random sampling. The Design Effect and Design Factor of \hat{k} are expressed as:

$$DEFF = \frac{\hat{\sigma}_{\hat{k},c}^2}{\hat{\sigma}_{\hat{k},srs}^2}, \quad (4.8)$$

$$DEFT = \frac{\hat{\sigma}_{\hat{k},c}}{\hat{\sigma}_{\hat{k},srs}}, \quad (4.9)$$

where $\hat{\sigma}_{\hat{k},c}^2$ is the variance of \hat{k} calculated using complex sample design information (strata and primary sampling unit identifiers) and $\hat{\sigma}_{\hat{k},srs}^2$ is the variance of \hat{k} assuming a simple random sample design. The standard deviations $\hat{\sigma}_{\hat{k},c}$ and $\hat{\sigma}_{\hat{k},srs}$ are the square roots of the variances. In Appendix B there is a description of how both jackknife repeated replication and Taylor series linearization can be used to calculate variances of statistics from data resulting from either a complex sample design using stratification and clustering or a simple random sample design. Inequality between the complex sample design and simple random sample design estimates arises because homogeneity among the respondents in clusters of a complex sample design reduces statistical independence between the respondents (Kalton, 1983). The Design Effect and Design Factor have useful applications. If a complex sample design was used in a survey, the division of the survey sample size by the magnitude of the Design Effect for a statistic indicates the number of samples of a simple random sample design that would be required to yield the same variance estimate for the statistic. Also, the division of the estimated confidence interval for the statistic by the Design Factor yields the estimated confidence interval for the statistic under a simple random sample design. Commonly, the Design Factor is greater than 1, which indicates a loss of precision when using a complex sample design. Five of the data sets have a Design Factor of less than 1, which indicates the complex sample design realized greater precision that would be expected from simple random sampling.

Precision of \hat{k}

Fields (1980) suggested methodological improvements to noise surveys to reliably estimate noise index parameters and \hat{k} . An estimate of k is of little use to policymakers unless information is provided about its precision. Because even small reductions in aircraft noise level can be very expensive, it is important to precisely determine the effect of number-of-events on annoyance relative to sound level as captured by \hat{k} . Fields identified five methods to reduce the variance estimate of \hat{k} using a Taylor

series linearization formula for a simple random sample design. When a more complex sample design is used, such as a two-stage clustered sample design, the methods become:

1. Increase the sample size and the number of study areas.
2. Increase the range of the sound levels included in the sample or avoid oversampling intermediate sound levels.
3. Increase the range of the number-of-events values.
4. Decrease the correlation between sound level and number-of-events.
5. Reduce the variance of the annoyance model residuals between and within primary sampling units by using precise annoyance measures and precise noise measures.

From Equations 4.4, 4.5, 4.6, and 4.7, the relative effect of each method on the Taylor variance estimate of \hat{k} can be quantified. Methods 2, 3, and 4 deal with the values of the predictor variables in the additive-log model. Sample design techniques, such as stratification, can be used to affect the range of each variable and the relationship between the variables. However, the greatest difficulty may be in finding situations in which measures of sound level and number-of-events are not highly correlated (Fields, 1980). The general problem of correlation among predictor variables is known as multicollinearity. A review of this is given in Appendix C.

4.1.4 Average Semipartial R^2 $\left(\overline{R_{sp}^2}\right)$

The average semipartial R^2 is calculated from the results of multiple stepwise regressions. In a single stepwise regression, predictor variables are entered one-at-a-time into the regression model, and regression statistics are calculated for each one of these “steps”. For a single stepwise regression, the semipartial R^2 for the i^{th} predictor variable ($R_{sp,i}^2$) is the incremental change in R^2 that is contributed by that predictor variable. In this research, the *average* semipartial R^2 for the i^{th} predictor variable $\left(\overline{R_{sp,i}^2}\right)$ is the average of the $R_{sp,i}^2$ from $p!$ stepwise regressions, where p is

the number of predictor variables, and each stepwise regression features a different order-of-entry of the predictor variables into the regression model. The sum of all of the variables' $\overline{R_{sp}^2}$ values equals the total R^2 , and the relative importance of each predictor variable can be expressed as a proportion of R^2 . The $\overline{R_{sp}^2}$ averages a variable's direct effect (considered by itself), total effect (conditional on all predictors in the full model), and partial effect (conditional on all subsets of predictors)(Johnson and LeBreton, 2004). This measure is especially useful in data sets where there is some correlation between the predictor variables, which is inevitable in survey data. When there is correlation between predictor variables, changing the order of entry in stepwise regressions changes the R_{sp}^2 of each predictor variable. The $\overline{R_{sp}^2}$ was first proposed as an importance measure in (Lindeman, Merenda, and Gold, 1980). The process of rating the importance of predictor variables by the $\overline{R_{sp}^2}$ criteria is referred to as General Dominance Analysis (Azen, Budescu, and Reiser, 2003; Budescu, 1993) and was recommended Johnson and LeBreton (2004) as a useful measure of variable importance.

In this research, the semipartial R^2 for the i^{th} predictor variable ($R_{sp,i}^2$) was calculated by averaging the $R_{sp,i}^2$ from $(p - 1)!$ stepwise regressions in which the intercept is always entered first into the regression model. Stepwise regression models in which the intercept is not entered first were not considered because the intercept is a baseline upon which the R^2 of a model is calculated against. For example, the R^2 of a two-dimensional regression model is the fit of the model compared to a horizontal line, which is the baseline condition of $R^2 = 0$. With models in the form of Equation (4.1), there are $(3 - 1)! = 2$ possible orderings of the variables in stepwise regressions, and the $\overline{R_{sp}^2}$ of each predictor variable is calculated by averaging two R_{sp}^2 . In Table 4.3 are given the $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ along with interquartile ranges. The interquartile ranges were constructed using bootstrapping (Efron, 1979; Mooney and Duval, 1993) in which 100 resamples were realized by sampling with replacement from the original sample, each resample being the same sample size as the original. From each resample an estimate of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ were made. For

a given variable, the lower limit of the interquartile range is the 25th value and the upper limit of the interquartile range is the 75th value when the values are sorted in increasing order. The ranges are not confidence intervals and thus do not imply a probability. The ranges only show how the $\overline{R_{sp}^2}$ for each variable varies over the resamples.

Examination of the data in Table 4.3 shows that one variable was not always quantified as being more generally dominant than the others. In fact, the variables are rated as generally dominant an equal number of times, approximately; $PNL_{max,av}$ was generally dominant in thirteen data sets and $\log_{10}(N)$ was generally dominant in twelve data sets. There is also a separation in the results by study. In all data sets of the TRACOR study (USA-022, USA-032, and USA-044), $PNL_{max,av}$ was generally dominant versus those of UKD-604, in which $\log_{10}(N)$ was generally dominant. In AUL-210, there was a divide in the dominance ratings with $PNL_{max,av}$ being generally dominant in two data sets (AUL-210/XRH and AUL-210/SYD) and $\log_{10}(N)$ being generally dominant in three data sets (AUL-210/ADL, AUL-210/MLB, and AUL-210/PER).

$\log_{10}(N)$ was quantified as being more generally dominant than $PNL_{max,av}$ in all of the data sets of UKD-604. This is consistent with one main conclusions from the study: that the importance of number-of-events in predicting annoyance had increased from 1982 (Le Masurier et al., 2007a). Although the researchers concluded the rising importance of the number-of-events through the use of site-level data and an analysis of \hat{k} , the semipartial R^2 analysis is supportive of the conclusion. The $\overline{R_{sp}^2}$ results for UKD-604 are very interesting in that for all airports, $\log_{10}(N)$ was more generally dominant than $PNL_{max,av}$, though as shown in Table 3.1, the maximum number-of-events that people were recorded as being exposed to was smaller than that in some of the other studies. This could be due to the threshold of 65 dBA used in UKD-604; if the simulated L_{Amax} of an event was below 65 dBA, it was not included in the number-of-events count, though it is arguable that events below 65 dBA could have been audible to respondents. Although $\log_{10}(N)$ accounts for more

Table 4.3. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the individual data sets.

Study/airport	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
AUL-210/ADL	0.02	0.02 to 0.03	0.06	0.05 to 0.07
AUL-210/MLB	0.00	0.00 to 0.01	0.00	0.00 to 0.01
AUL-210/PER	0.02	0.01 to 0.03	0.04	0.03 to 0.05
AUL-210/XRH	0.04	0.02 to 0.06	0.01	0.00 to 0.01
AUL-210/SYD	0.06	0.05 to 0.06	0.02	0.02 to 0.03
USA-022/ORD	0.21	0.20 to 0.24	0.01	0.01 to 0.02
USA-022/DFW	0.08	0.07 to 0.10	0.03	0.02 to 0.04
USA-022/DEN	0.24	0.22 to 0.25	0.02	0.01 to 0.02
USA-022/LAX	0.17	0.15 to 0.19	0.00	0.00 to 0.01
USA-032/BOS	0.10	0.09 to 0.11	0.01	0.01 to 0.01
USA-032/MIA	0.17	0.15 to 0.18	0.06	0.05 to 0.07
USA-032/JFK	0.07	0.06 to 0.08	0.02	0.01 to 0.02
USA-044/CHA	0.09	0.08 to 0.09	0.06	0.05 to 0.06
USA-044/RNO	0.10	0.09 to 0.11	0.00	0.00 to 0.00
UKD-024/LHR	0.11	0.10 to 0.11	0.03	0.02 to 0.03
UKD-130/LHR	0.08	0.07 to 0.09	0.02	0.02 to 0.03
UKD-604/BHX	0.07	0.06 to 0.09	0.65	0.63 to 0.66
UKD-604/LGW	0.06	0.05 to 0.10	0.57	0.54 to 0.67
UKD-604/LHR	0.20	0.19 to 0.20	0.50	0.49 to 0.52
UKD-604/LBA	0.19	0.14 to 0.22	0.47	0.43 to 0.54
UKD-604/LCY	0.10	0.08 to 0.14	0.84	0.81 to 0.86
UKD-604/LTN	0.04	0.03 to 0.07	0.71	0.67 to 0.74
UKD-604/MAN	0.17	0.15 to 0.19	0.30	0.29 to 0.32
UKD-604/SOU	0.01	0.00 to 0.01	0.44	0.42 to 0.47
UKD-604/STN	0.09	0.04 to 0.17	0.86	0.78 to 0.89

of the variance in the annoyance variable than $PNL_{max,av}$ in the data sets of UKD-604, the $\overline{R_{sp}^2}$ levels for $PNL_{max,av}$ are comparable to the $\overline{R_{sp}^2}$ levels in the other data sets. The average of the $PNL_{max,av}$ $\overline{R_{sp}^2}$ values for UKD-604 is nearly identical to the average of the $PNL_{max,av}$ $\overline{R_{sp}^2}$ values of the other data sets. A threshold value was not used for event-counting in AUL-210, USA-022, USA-032, or USA-044. A threshold of 80 PNdB was used for both UKD-024 and UKD-130, which is approximately equal to 65 dBA.

The large $\overline{R_{sp}^2}$ results for $\log_{10}(N)$ observed in UKD-604 could also be directly related to the the method of simulating a person-level data set for UKD-604, as described in Chapter 3, Section 3.3. In brief, a person-level UKD-604 data set was simulated from site-level acoustical data, where each site was a Census Output Area from the 2001 Census in England. For each site, the UKD-604 documentation reported an average sound level and number-of-events count at the population-weighted centroid of the site. Although acoustical data was only available for one geographical point of a site, the full annoyance data was available. Thus, simulating a person-level data set encompassed simulating acoustical data for all persons; this was done with a converging annoyance model process.

For each site, ranges of variation for the sound level and number-of-events were defined. From these ranges, all possible combinations of sound level and number-of-events were generated. Using an arbitrary annoyance model which was a function of sound level and number-of-events, annoyance was calculated for all combinations. The annoyance calculations were rescaled to have the same mean and standard deviation as the actual UKD-604 annoyance responses. Each UKD-604 annoyance category was connected to a group of rescaled annoyance calculations, and thus, combinations of sound level and number-of-events. For each UKD-604 annoyance category, and for each sample at that annoyance level, a combination of sound level and number-of-events was randomly sampled from the group of combinations assigned to the UKD-604 annoyance category. Then, from the person-level data, the annoyance model was

re-estimated. This process was repeated until the annoyance model had sufficiently converged.

The simulation process presupposed that annoyance was related to sound level and number-of-events, so it is understandable that $\log_{10}(N)$ has large $\overline{R_{sp}^2}$ levels in the UKD-604 data sets. A possible explanation for the $\overline{R_{sp}^2}$ levels for $\log_{10}(N)$ being greater than those for $PNL_{max,av}$ is the process of generating the ranges for sound level and number-of-events for each site. Likely, the number-of-events range utilized for each site was relatively larger than that utilized for sound level so that the combinations of sound level and number-of-events assigned to persons featured greater variability in the number-of-events relative to sound level so that number-of-events, by virtue of the greater variability, accounted for the variation in annoyance to a greater degree than sound level. For this reason, a second person-level UKD-604 data set was simulated with further variability added to the sound level and number-of-events assigned to respondents in the first UKD-604 data set. The variabilities for sound level and number-of-events were randomly extracted from Gaussian distributions, which were different for each respondent. Each sound level (or number-of-events) Gaussian distribution was centered on the sound level (or number-of-events) value assigned to a respondent and featured a standard deviation that was one-half of the values shown in Table 3.3. Thus, for all respondents within a given site, the Gaussian distributions featured the same standard deviations while the means of the distributions varied with each individual. Also, a third person-level UKD-604 data set was simulated without utilizing variation in sound level and number-of-events within sites, so that for all respondents within a site, $PNL_{max,av}$ and $\log_{10}(N)$ were the same. With four of the nine airports of UKD-604 (LTN, LCY, LGW, and STN) there was only one site from which annoyance data were sampled. Thus, for these airports, the regression matrices consisted of three constants, and regressions using the data from these individual airports are not useful. However, a comparison of the regression and $\overline{R_{sp}^2}$ results from the all three UKD-604 data sets are of interest and are shown in Tables 4.4 and 4.5.

Table 4.4. Multiple linear regression results for the UKD-604 data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). The 95% CI (confidence interval) is for k . Standard deviation estimates of the regression coefficient estimates are given in parentheses. UKD-604* is the first UKD-604 data set utilizing variation in sound level and number-of-events. UKD-604** is the second UKD-604 data set with further variation applied to sound level and number-of-events. UKD-604 is the third UKD-604 data set which was created only from the reported data and does not include variation in sound level and number-of-events.

Study	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{\hat{k},c}$	95% CI for k	R^2
UKD-604*	0.073 (0.005)	3.103 (0.065)	-9.974 (0.427)	42.47	20.92 (19.67)	-5.78 to 90.72	0.56
UKD-604**	0.072 (0.005)	3.068 (0.065)	-9.790 (0.425)	42.73	20.57 (19.88)	-3.11 to 88.58	0.56
UKD-604	0.040 (0.009)	0.810 (0.046)	-1.860 (0.681)	20.27	28.30 (16.54)	-45.00 to 85.54	0.18

Table 4.5. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the UKD-604 data sets. UKD-604* is the first UKD-604 data set utilizing variation in sound level and number-of-events. UKD-604** is the second UKD-604 data set with further variation applied to sound level and number-of-events. UKD-604 is the third UKD-604 data set which was created only from the reported data and does not include variation in sound level and number-of-events.

Data set	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
UKD-604*	0.09	0.09 to 0.10	0.47	0.46 to 0.48
UKD-604**	0.09	0.09 to 0.10	0.46	0.45 to 0.47
UKD-604	0.04	0.04 to 0.05	0.14	0.14 to 0.15

As shown in Table 4.5, the $\overline{R_{sp}^2}$'s of $\log_{10}(N)$ in UKD-604* and UKD-604** are much greater than that in UKD-604. This is interesting and has implications for survey sampling. If there is a statistical model in a population surrounding an airport that can adequately predict annoyance reactions to aircraft noise and which is a func-

tion of acoustical data (as done to generate UKD-604), then it might be detrimental to model estimation if the noise exposure experienced at one spatial point is applied to many respondents of a geographical area (as done to generate UKD-604).

4.1.5 Comparison of Importance Measures

Both \hat{k} and General Dominance Analysis ($\overline{R_{sp}^2}$) are methods of expressing relative importance of variables. The estimation of \hat{k} is a single-analysis method and General Dominance Analysis is a multiple-analysis method.

Multiple-analysis methods have an appeal over single-analysis methods because they combine results from different regression analyses. In this case, Dominance Analysis averages results from two regressions. The results from both methods cannot be directly compared as “importance” is defined differently in the two methods. For \hat{k} , importance is based the change in annoyance introduced by a unit change in $\log_{10}(N)$ relative to a unit change in sound level. In Dominance Analysis, importance is based on portions of variance in annoyance accounted for by sound level and $\log_{10}(N)$. However, certain observations are helpful to differentiate the measures. The average $\overline{R_{sp}^2}$ of the predictor variables sum to the total model R^2 . The model R^2 is limited to range from 0 to 1 so that the scale of $\overline{R_{sp}^2}$ is limited in comparison to that possible with \hat{k} . Whereas a negative $\overline{R_{sp}^2}$ for a predictor variable has a definite meaning in that it is an indication that the predictor variable worsens the fit of the regression model to the data, a negative \hat{k} does not mean that an increase in the number-of-events accompanies a reduction in annoyance, because \hat{b}_L could be negative, as in UKD-604/Southampton (SOU). If the $\overline{R_{sp}^2}$ for a variable is greater than that of another variable, then it can be stated with certainty that on average, that variable accounts for more variance of the criterion variable than the other. \hat{k} is hard to interpret in that the unstandardized regression coefficient estimates depend on the scale of the predictor variable data. Simply changing the scale of a predictor variable (without changing the data) will change its unstandardized regression coefficient estimate.

The $\overline{R_{sp}^2}$, however, will not change. Finally, \hat{k} collapses the unstandardized regression coefficient estimates into one measure with the effect that the relative magnitude of the regression coefficient estimates is lost. In light of these observations, $\overline{R_{sp}^2}$ is preferable to \hat{k} as an importance measure.

Of particular interest are the results of UKD-604 compared to the other studies. From the results shown in Table 4.3, $\log_{10}(N)$ is quantified as a more generally dominant predictor variable over $PNL_{max,av}$ in every individual airport data set of UKD-604, while the opposite is observed for almost every individual airport data set of the other studies. The results of \hat{k} shown in Table 4.2 are less clear. Not only is $\log_{10}(N)$ quantified as more generally dominant, but the portions of variance in the criterion variable accounted for by $\log_{10}(N)$ are all greater than those accounted for by $PNL_{max,av}$ in the all of the other data sets. Le Masurier et al. (2007a) stated that the trend of air traffic in England has been that the number-of-events have increased significantly while the sound levels generated by individual aircraft events have been reduced, as modern aircraft with quieter turbofan engines and improved climb performance have replaced older and noisier aircraft. Thus, perhaps modern-day characteristics of air traffic result in a decrease of the importance of average sound level to annoyance and increase the importance of the number-of-events. As shown in Table 3.1, there is a 25-year gap between UKD-604 and the next-closest survey.

4.1.6 Continuous and Aggregate Data

Relationships between the predictor variable data are critical to the quality of regression results. As stated earlier, correlation between the predictor variable data can be detrimental to quantifying \hat{k} with precision. In general, it is important to consider the nature of the spread of the predictor variable data points over the space spanned by the predictor variables. As the distribution of predictor variable data points changes, the correlation structure changes. Also, the predictor variable data points determine the scope of a model; a regression model is only appropriate for the region covered

by the predictor variable data points. In addition, concentration differences in the predictor variable space affects the regression model in that areas of the predictor variable space with greater number of data points have a greater effect on the model. This last point is explored in this section.

In theory, a least squares analysis using sample weights that would balance the contributions from different predictor variable data points so that differences in concentration could be normalized is possible. However, the weighting would be arbitrary, and ordinary (unweighted) least squares analysis produces smaller standard errors than weighted least squares (Winship and Radbill, 1994).

With the historic survey data, regression analyses using aggregated data was implemented. That is, the survey data were separated into groups based on ranges of the predictor variables and averaged within groups. The result of this process was a collection of data points identified by average $PNL_{max,av}$, average $\log_{10}(N)$, and average annoyance. The averaged data were then used in regression analyses. This was done to explore how regression results changed as the contributions from predictor variable data points in different regions of the predictor variable space were normalized. For these analyses, individual airports that were part of the same noise survey were grouped together because for the multi-airport studies, some airport sample sizes were very small. Aggregating the data in these cases would have resulted in only one or two groups.

For each survey data set, twenty-five groups were formed from five $PNL_{max,av}$ groups and five $\log_{10}(N)$ groups; however some groups had zero data points within the range of $PNL_{max,av}$ and $\log_{10}(N)$ so that in effect, a smaller number of groups were realized. In Table 4.6 are given the $PNL_{max,av}$ boundaries and in Table 4.7 are given the $\log_{10}(N)$ boundaries for each of the survey data sets. In Table 4.8 are shown the multiple linear regression results using the aggregate data. In Table 4.9 are shown multiple linear regression results, excluding $\overline{R_{sp}^2}$ results, using the continuous, *unaggregated* data. In Table 4.10 are shown the average semipartial R^2 results ($\overline{R_{sp}^2}$) and interquartile ranges of $\overline{R_{sp}^2}$ for the continuous data. Variance calculations for

\hat{k} were not implemented for the aggregate data as the grouping of data nullified the geographical delineations. In addition, because the full sample was aggregated into groups based on the predictor variables, the bootstrap procedure was not used with the aggregate data sets to calculate interquartile ranges for the $\overline{R_{sp}^2}$. Since each bootstrap sample would not include one or more parts of the predictor variable space unless each bootstrap sample was equal to the original sample, interquartile ranges for the $\overline{R_{sp}^2}$ calculated using the aggregate data would be misleading.

Table 4.6. $PNL_{max,av}$ group boundaries for the survey data sets in the aggregate data analyses.

Study	$PNL_{max,av}$ boundaries					
AUL-210	81.1	88.3	95.5	102.7	109.9	117.1
USA-054	66.1	77.6	89.1	100.6	112.1	123.6
UKD-024	81.5	87.7	93.9	100.1	106.3	112.5
UKD-130	83.0	89.2	95.4	101.6	107.8	114.0
UKD-604	78.3	81.8	85.3	88.9	92.4	95.9

Table 4.7. $\log_{10}(N)$ boundaries for the survey data sets in the aggregate data analyses.

Study	$\log_{10}(N)$ boundaries					
AUL-210	-0.16	0.30	0.75	1.21	1.67	2.12
USA-054	0.00	0.54	1.08	1.62	2.16	2.70
UKD-024	0.00	0.54	1.08	1.61	2.15	2.69
UKD-130	1.30	1.58	1.86	2.14	2.42	2.70
UKD-604	-0.22	0.32	0.86	1.40	1.94	2.48

Table 4.8. Multiple linear regression results from the aggregate data analyses. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. n_g is the number of aggregate groups. Standard deviation estimates of the regression coefficient estimates are given in parentheses. ¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.

Study	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\frac{PNL_{max,av}}{R_{sp}^2}$	$\frac{\log_{10}(N)}{R_{sp}^2}$	R^2	n_g
AUL-210	0.016 (0.002)	1.087 (0.144)	-0.041 ¹ (0.132)	68.47	0.46	0.48	0.95	17
USA-054	0.028 (0.004)	0.533 (0.178)	-0.423 ¹ (0.321)	19.16	0.53	0.29	0.83	22
UKD-024	0.021 (0.002)	0.406 (0.095)	-0.047 ¹ (0.107)	19.37	0.57	0.39	0.95	18
UKD-130	0.012 ¹ (0.008)	0.603 ¹ (0.333)	-0.044 ¹ (0.138)	49.90	0.45	0.45	0.90	16
UKD-604	-0.003 ¹ (0.006)	1.362 (0.258)	-0.005 ¹ (0.208)	-415	0.27	0.53	0.79	16

Table 4.9. Multiple linear regression results from the unaggregated data analyses. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). $DEFT$ is the Design Effect for \hat{k} . The 95% CI (confidence interval) is for k . n is the sample size. Standard deviation estimates of the regression coefficient estimates are given in parentheses.

Study	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{\hat{k},c}$	$DEFT$	95% CI for k	R^2	n
AUL-210	0.054 (0.004)	1.009 (0.058)	-3.839 (0.499)	18.71	5.49 (6.36)	3.90	3.46 to 33.96	0.10	2720
USA-054	0.063 (0.002)	0.828 (0.030)	-4.215 (0.160)	13.24	4.72 (6.10)	7.41	2.36 to 24.12	0.22	8399
UKD-024	0.069 (0.003)	0.220 (0.025)	-4.278 (0.271)	3.18	0.57 (0.77)	1.40	2.05 to 4.31	0.13	4655
UKD-130	0.063 (0.005)	0.042 (0.087)	-3.496 (0.375)	0.66	1.62 (2.82)	1.13	-2.61 to 3.93	0.10	2618
UKD-604	0.073 (0.005)	3.103 (0.065)	-9.974 (0.427)	42.47	20.92 (19.67)	4.79	-5.78 to 90.72	0.56	2437

Table 4.10. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ in the unaggregated data analysis.

Study	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
AUL-210	0.03	0.02 to 0.03	0.08	0.07 to 0.08
USA-054	0.14	0.13 to 0.14	0.09	0.08 to 0.09
UKD-024	0.11	0.10 to 0.11	0.03	0.02 to 0.03
UKD-130	0.08	0.07 to 0.09	0.02	0.02 to 0.02
UKD-604	0.09	0.09 to 0.10	0.47	0.46 to 0.48

The \hat{k} resulting from use of the aggregate data, shown in Table 4.8, are greater than the \hat{k} resulting from use of the continuous data, shown in Table 4.9. This illustrates that the distribution of predictor variable data points is not balanced, or evenly represented over the predictor variable space. The aggregate R^2 values are also greater; this is likely from the smoothing effect of averaging, which lessens the “noise” in the data.

4.2 Combination Data Set Results

As mentioned in Chapter 3, Section 3.5.1, seven combination data sets were realized through various annoyance scale modifications. Three types of modifications were used to establish a common annoyance scale among all of the data sets: the midpoint Miedema scaling, category combination, and rescaling and rounding. The types of results shown in Table 4.2 for the individual data sets in are also given for the combination data sets in this section. The combination was performed using a computer program. See Appendix I.13 for a description of the program source code and the storage location of the code.

In Table 4.11 are given combination data set labels and descriptions of the modifications that were applied to annoyance scales. Prior to modifying the annoyance variables, data sets that featured the same number of annoyance categories were

grouped. This resulted in two intermediate data sets, $M1 = [\text{UKD-024}; \text{UKD-130}]$, and $M2 = [\text{USA-054}; \text{AUL-210}; \text{UKD-604}]$.

Table 4.11. Labels and annoyance scale modification descriptions for the combination data sets.

Label	Description
Combination #1	Midpoint Miedema scaling
Combination #2	Category combination: 4-point scale, $M1$: no change, $M2$: combined categories 1 and 2
Combination #3	Category combination: 4-point scale, $M1$: no change, $M2$: combined categories 2 and 3
Combination #4	Category combination: 4-point scale, $M1$: no change, $M2$: combined categories 3 and 4
Combination #5	Category combination: 4-point scale, $M1$: no change, $M2$: combined categories 4 and 5
Combination #6	Rescale and round: 4-point scale, $M1$: no change, $M2$: rescaled 1-5 to 1-4
Combination #7	Rescale and round: 5-point scale, $M1$: rescaled 1-4 to 1-5, $M2$: no change

Multiple linear regression models were again estimated in the form of Equation (4.1). In Table 4.12 are shown the results from the analyses of the combination data sets. All data sets have a sample size of 20,829, which is the sum of the individual data set sample sizes. The correlation between the $PNL_{max,av}$ and $\log_{10}(N)$ data is 0.004. In general, the results are not sensitive to combination method.

4.2.1 Decibel Equivalent Number Effect, \hat{k}

The decibel equivalent number effects predicted from the combination data sets can be stated at 95% confidence to be significant ($k > 0$). Although the predicted values of \hat{k} are all greater than the equal-energy value ($k = 10$), they cannot be stated with 95% confidence to be significantly different from the equal-energy value, as the estimated confidence intervals include $k = 10$.

4.2.2 Standard Deviation of \hat{k} , $\hat{\sigma}_{\hat{k},c}$

The complex standard deviations calculated are consistent in magnitude and are on average smaller than those predicted using the individual data sets. The jackknife estimates are always slightly greater than the Taylor estimates. As described in Appendix B, the first-stage grouping of survey respondents, or primary sampling unit, is usually used to create replicates for jackknife repeated replication. With each individual data set, the primary sampling unit was a geographically-based region surrounding an airport. For the combination data sets, the first-stage grouping of survey respondents is airport. Therefore, airport was the primary sampling unit for the combination data sets and each replicate for the jackknife was created by removing an airport from the full sample. There are twenty-five different airports for the combination data sets, and thus twenty-five replicates for the jackknife variance estimation. Although Heathrow appears three times (in UKD-024, UKD-130, and UKD-604), the three occurrences were considered separate from one another. In general, the greater the number of replicates and thus, degrees of freedom, the more precise an estimate of the variance becomes. Although twenty-five replicates is small, it is better to calculate the jackknife variance estimate than to assume there are no differences between airports.

The observed Design Factors are quite large. Neglecting the two outliers in the individual data sets (109 and 63.69 shown in Table 4.2), the average design factor was 2.4. For the combination data sets (Table 4.12), the average of the design factors is 9.5 and the estimated confidence intervals, on average, span a range of 15.9. With a simple random sample of the same size ($n = 20,829$), the estimated confidence interval for k could be expected to span a range of 1.7 ($15.9/9.5$), which is obviously preferable for quantifying the precision of \hat{k} . However, a large simple random sample of this size might not be economically feasible, or logistically possible for some airports. In theory, a geographical area consisting of *multiple* airports could be stratified by sound level and number-of-events, and a random sample could be drawn from each of

the strata, with the strata sample sizes summing to the total sample size. However, this might not realize, for each airport separately, an adequate collection of data to precisely estimate an annoyance model. Though a large simple random sample size sample can yield favorable analysis results, it may not be feasible or desirable to do so in some cases.

4.2.3 Average Semipartial R^2

As with the individual data sets, there are $(3 - 1)! = 2$ possible orderings of $PNL_{max,av}$ and $\log_{10}(N)$ in stepwise regressions for the combination data sets. In Table 4.13 are shown the $\overline{R_{sp}^2}$ and the associated interquartile ranges for $\overline{R_{sp}^2}$. As shown in Table 4.13, $PNL_{max,av}$ was always quantified as the most generally dominant predictor variable.

4.2.4 Continuous and Aggregate Data

The data were also aggregated for the combination data sets. Twenty-five groups were formed from five $PNL_{max,av}$ groups and five $\log_{10}(N)$ groups; however, one group had zero data points within the range of $PNL_{max,av}$ and $\log_{10}(N)$ so that in effect, twenty-four groups were realized. In Table 4.14 are shown the variable boundaries for all the combinations data sets. Because the combination data sets shared the same predictor variable data, only one set of boundaries was required. Within each group, the $PNL_{max,av}$, $\log_{10}(N)$, and annoyance data were arithmetically averaged. In Table 4.15 are shown the results of multiple regressions using the aggregate data. Variance calculations for \hat{k} were not implemented for the aggregate data as the grouping of data nullified the geographical delineations. In addition, the bootstrap procedure was not used to calculate interquartile ranges for the $\overline{R_{sp}^2}$ for the reasons described earlier; that since each bootstrap sample would not include one or more parts of the predictor variable space unless each bootstrap sample was equal to the original sample, interquartile ranges for the $\overline{R_{sp}^2}$ calculated using the aggregate data would be misleading.

Table 4.12. Multiple linear regression results from the combination data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. $\hat{\sigma}_{\hat{k},c}$ is the jackknife variance (Taylor variance in parentheses) of \hat{k} taking into account complex sample design information (strata and primary sampling unit identifiers). $DEFT$ is the Design Effect for \hat{k} . The 95% CI (confidence interval) is for k . Descriptions of the combination data sets can be found in Table 4.11. The predictor variable data are the same for all combination data sets.

Data set	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{\hat{k},c}$	$DEFT$	95% CI for k	R^2
Combination #1	1.091 (0.022)	14.708 (0.327)	-77.609 (2.089)	13.49	3.89 (3.03)	9.68	5.46 to 21.51	0.18
Combination #2	0.042 (0.001)	0.570 (0.013)	-2.614 (0.085)	13.44	4.57 (3.34)	10.82	4.00 to 22.88	0.17
Combination #3	0.041 (0.001)	0.544 (0.012)	-2.351 (0.079)	13.25	3.93 (2.96)	10.03	5.15 to 21.36	0.18
Combination #4	0.041 (0.001)	0.556 (0.012)	-2.233 (0.079)	13.60	3.61 (2.98)	9.05	6.14 to 21.05	0.18
Combination #5	0.042 (0.001)	0.587 (0.013)	-2.318 (0.085)	13.90	3.58 (3.30)	8.26	6.50 to 21.29	0.17
Combination #6	0.041 (0.001)	0.556 (0.012)	-2.233 (0.079)	13.60	3.61 (2.98)	9.05	6.14 to 21.05	0.18
Combination #7	0.056 (0.001)	0.763 (0.017)	-3.589 (0.109)	13.56	3.78 (2.89)	9.49	5.76 to 21.37	0.18

Table 4.13. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the combination data sets. The predictor variable data are the same for all combination data sets.

Data set	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
Combination #1	0.100	0.098 to 0.102	0.080	0.079 to 0.083
Combination #2	0.094	0.091 to 0.096	0.075	0.072 to 0.076
Combination #3	0.100	0.097 to 0.102	0.077	0.074 to 0.079
Combination #4	0.098	0.095 to 0.100	0.080	0.077 to 0.083
Combination #5	0.092	0.090 to 0.095	0.078	0.076 to 0.080
Combination #6	0.098	0.096 to 0.101	0.080	0.077 to 0.082
Combination #7	0.098	0.096 to 0.101	0.079	0.077 to 0.082

Table 4.14. $PNL_{max,av}$ and $\log_{10}(N)$ group boundaries for the combination data sets in the aggregate data analysis.

Variable	Group boundaries					
$PNL_{max,av}$	66.06	77.55	89.05	100.55	112.05	123.55
$\log_{10}(N)$	-0.22	0.36	0.95	1.53	2.12	2.70

Table 4.15. Multiple linear regression results from the aggregate data analyses of the combination data sets. The estimated regression models were in the form of Equation (4.1). $\hat{k} = \hat{b}_N/\hat{b}_L$ is the decibel equivalent number effect. Average semipartial R^2 values ($\overline{R_{sp}^2}$) are given. n_g is the sample size. The predictor variable data are the same for all combination data sets.

Study	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\frac{PNL_{max,av}}{R_{sp}^2}$	$\log_{10}\left(\frac{N}{R_{sp}^2}\right)$	R^2	n_g
Combination #1	0.541 (0.087)	10.978 (2.483)	-20.471 (8.204)	20.29	0.48	0.28	0.77	24
Combination #2	0.022 (0.003)	0.467 (0.086)	-0.603 (0.284)	20.83	0.51	0.31	0.83	24
Combination #3	0.024 (0.003)	0.434 (0.079)	-0.556 (0.260)	18.40	0.57	0.29	0.85	24
Combination #4	0.025 (0.003)	0.432 (0.078)	-0.578 (0.257)	17.15	0.60	0.27	0.87	24
Combination #5	0.026 (0.003)	0.446 (0.086)	-0.634 (0.285)	16.71	0.60	0.26	0.85	24
Combination #6	0.025 (0.003)	0.432 (0.078)	-0.578 (0.257)	17.15	0.60	0.27	0.87	24
Combination #7	0.031 (0.004)	0.591 (0.117)	-0.919 (0.386)	19.33	0.53	0.29	0.82	24

A comparison of the combination data sets results given in Table 4.12 which were estimated using the continuous data with the results given in Table 4.15 which were estimated using the aggregate data shows that the \hat{k} from the aggregate data varied more than the \hat{k} from the continuous data. This is not surprising, considering that only twenty-four points were available for the aggregate data models. As shown in Table 4.15, the \hat{k} varies with combination method. Because the predictor variable data are the same for all combination data sets, the changes in \hat{k} are due solely to the changes in average annoyance which result from varying the method of transforming and combining the annoyance scales. It is unknown how significant these changes are, however, since variance calculations for \hat{k} were not implemented.

Although estimates of the variance of \hat{k} from the aggregate data were not calculated for either the individual survey data sets or the combination data sets so that the precision of the estimates is unknown, it is interesting that the difference of \hat{k} between the aggregate data and the continuous data was less for the combination data sets than for the individual survey data sets. This could be an indication that the full collection of predictor variable data from all of the individual survey data sets is more evenly spread over the predictor variable space than the individual survey data sets alone, so that the effect of aggregating is less for the combination data sets.

4.2.5 The Effect of Airport

Of interest was the effect of including or excluding an airport on the results, especially \hat{k} and the jackknife standard deviation estimate of \hat{k} . The jackknife replicate estimates of k (\hat{k}_r) from the combination data sets were examined to find the airport that, when removed from the combination data set, resulted in the highest \hat{k} . Once identified, the airport was removed from the data set. This process was repeated two more times so that the end result was a data set that had three airports removed from it. The chosen airport for each examination was the one that had the greatest effect on \hat{k} . In order, the airports were UKD-024/Heathrow (1967), USA-022/Chicago (1967), and

USA-022/Denver (1967). The effect of including these airports in the data set was to decrease \hat{k} . Thus, when they were removed from the data set, \hat{k} increased. The same methodology was applied to identify airports that when removed from the data set, yielded the lowest \hat{k} using the remaining data. The effect of including these airports in the data set was to increase \hat{k} . In order, these airports were UKD-604/Heathrow (2006), USA-032/Boston (1969), and AUL-210/Perth (1980).

In Table 4.16 are shown jackknife replicate estimates of $k \left(\hat{k}_r \right)$ and the estimated 95% confidence intervals using the Combination #1 (midpoint Miedema scaling) data. The leftmost column contains baseline \hat{k} ; in each row is the \hat{k}_r from the full data set excluding the airport listed in the same row. Subsequent columns contain \hat{k}_r from a subset of the full data set created by excluding the airports listed in the column heading which were those that when excluded, yielded the largest \hat{k}_r ; in each row is the \hat{k}_r from the subset excluding the airport listed in the same row. The effects of excluding airports on $\overline{R_{sp}^2}$ results are shown in Table 4.17.

In Table 4.18 are shown jackknife replicate estimates of $k \left(\hat{k}_r \right)$ and the estimated 95% confidence intervals using the Combination #1 (midpoint Miedema scaling) data. The leftmost column contains baseline estimates of k ; in each row is the \hat{k}_r from the full data set excluding the airport listed in the same row. Subsequent columns contain \hat{k}_r from a subset of the full data set created by excluding the airports listed in the column heading which were those that when excluded, yielded the smallest \hat{k}_r ; in each row is the \hat{k}_r from the subset excluding the airport listed in the same row. The effects of excluding airports on $\overline{R_{sp}^2}$ results are shown in Table 4.19.

Comparing the \hat{k}_r from Table 4.16 and Table 4.18 illustrates that with the current data, results can be heavily influenced by the presence and absence of certain airports. A direct comparison of the estimated confidence intervals for k in the rightmost column of the two tables, (13.87 to 26.16) and (2.63 to 16.87), reveals that large subsets of the combination data set can be realized which yield markedly different \hat{k} and small overlap in the estimated confidence intervals. The replicate estimates \hat{k}_r change appreciably as airports are excluded from the full data set. From Table

Table 4.16. Changes in jackknife replicate estimates of k (\hat{k}_r) as airports that cause \hat{k} to decrease the most are removed from the full data set. The estimated 95% confidence intervals for k for the data sets without the airports in the column headings are given at the bottom of the table. The replicate estimates are those from the Combination #1 data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-024/Heathrow (LHR), B denotes USA-022/Chicago (ORD), and C denotes USA-022/Denver (DEN).

#	Study/airport removed	Baseline	Baseline - A	Baseline - A,B	Baseline - A,B,C
1.	AUL-210/ADL	13.05	17.02	18.85	20.99
2.	AUL-210/MLB	13.43	16.75	18.27	19.94
3.	AUL-210/PER	12.55	16.17	17.77	19.72
4.	AUL-210/XRH	13.10	16.35	17.82	19.45
5.	AUL-210/SYD	13.71	17.36	19.05	20.83
6.	USA-022/ORD (B)	14.50	18.34	–	–
7.	USA-022/DFW	13.76	17.18	18.92	20.70
8.	USA-022/DEN (C)	14.29	18.19	20.01	–
9.	USA-022/LAX	13.24	16.66	18.31	20.18
10.	USA-032/BOS	12.53	15.74	17.01	18.72
11.	USA-032/JFK	13.29	16.67	18.12	19.86
12.	USA-032/MIA	13.58	16.92	18.50	20.19
13.	USA-044/CHA	13.08	16.75	18.29	20.19
14.	USA-044/RNO	13.21	16.59	18.15	19.86
15.	UKD-024/LHR (A)	16.81	–	–	–
16.	UKD-130/LHR	14.11	17.73	19.51	21.25
17.	UKD-604/BHX	13.39	16.70	18.26	19.98
18.	UKD-604/LGW	13.50	16.83	18.36	20.04
19.	UKD-604/LHR	12.02	15.03	16.43	18.21
20.	UKD-604/LBA	13.53	16.86	18.44	20.15
21.	UKD-604/LCY	13.47	16.79	18.32	20.00
22.	UKD-604/LTN	13.49	16.81	18.35	20.04
23.	UKD-604/MAN	13.34	16.67	18.31	20.11
24.	UKD-604/SOU	13.56	16.89	18.50	20.23
25.	UKD-604/STN	13.42	16.74	18.25	19.92
	95% CI for k	4.93 to 22.04	10.20 to 23.43	11.47 to 25.20	13.87 to 26.16

Table 4.17. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the midpoint Miedema scaling combination data set (Combination #1) as airports that cause \hat{k} to decrease the most are removed from the full data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-024/Heathrow (LHR), B denotes USA-022/Chicago (ORD), and C denotes USA-022/Denver (DEN).

Data set	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
Baseline	0.100	0.098 to 0.102	0.080	0.079 to 0.083
Baseline - A	0.102	0.099 to 0.105	0.100	0.096 to 0.102
Baseline - A,B	0.093	0.089 to 0.095	0.111	0.107 to 0.114
Baseline - A,B,C	0.088	0.086 to 0.092	0.119	0.116 to 0.123

4.16, the AUL-210/ADL replicate estimate of \hat{k}_r changes from 13.05 to 20.99, which is an increase of nearly 8 dB. A change of 8 dB is approximately equivalent to a multiplication of the average sound energy by a factor of six, which is an appreciable change. In the change of \hat{k}_r from 13.05 to 20.99, $\log_{10}(N)$ is increased in importance; the average sound energy which increases annoyance by a tenfold increase in the number-of-events is reduced by a factor of six.

Samples around Heathrow are of particular interest here. Heathrow has the greatest effect on \hat{k}_r at both ends of the spectrum; UKD-024/LHR (1967) simultaneously decreases \hat{k}_r more than any other airport and UKD-604/LHR (2006) increases \hat{k}_r more than any other airport. From Table 4.2, $\hat{k} = 3.18$ in UKD-024/LHR and $\hat{k} = 16.47$ in UKD-604/LHR and furthermore the estimated confidence intervals of k do not overlap, thus it can be stated at 95% confidence that the two \hat{k} are significantly different from one another. UKD-604/LCY (2006) is the only other UKD-604 airport that features a \hat{k} that can be stated with 95% confidence to be greater than that from UKD-024/LHR (1967). However, its small sample size (65) likely dampens the effect of removing it from the full data set in comparison to UKD-604/LHR (2006) with a sample size of 1189. There are a number of possible reasons why this duality

Table 4.18. Changes in jackknife replicate estimates of k (\hat{k}_r) as airports that cause \hat{k}_r to increase the most are removed from the full data set. The estimated 95% confidence intervals for k for the data sets without the airports in the column headings are given at the bottom of the table. The replicate estimates are those from the Combination #1 data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-604/Heathrow (LHR), B denotes USA-032/Boston (BOS), and C denotes AUL-210/Perth (PER).

#	Study/airport removed	Baseline	Baseline - A	Baseline - A, B	Baseline - A, B, C
1.	AUL-210/ADL	13.05	11.42	10.18	8.83
2.	AUL-210/MLB	13.43	11.97	10.78	9.70
3.	AUL-210/PER (C)	12.55	10.96	9.75	–
4.	AUL-210/XRH	13.10	11.64	10.48	9.36
5.	AUL-210/SYD	13.71	12.19	11.01	9.88
6.	USA-022/ORD	14.50	12.91	11.44	10.28
7.	USA-022/DFW	13.76	12.27	10.97	9.88
8.	USA-022/DEN	14.29	12.85	11.64	10.52
9.	USA-022/LAX	13.24	11.60	10.20	9.07
10.	USA-032/BOS (B)	12.53	10.81	–	–
11.	USA-032/JFK	13.29	11.61	10.23	9.18
12.	USA-032/MIA	13.58	12.08	10.84	9.78
13.	USA-044/CHA	13.08	11.58	10.40	9.18
14.	USA-044/RNO	13.21	11.73	10.56	9.40
15.	UKD-024/LHR	16.81	15.02	13.59	12.65
16.	UKD-130/LHR	14.11	12.79	11.55	10.49
17.	UKD-604/BHX	13.39	11.88	10.62	9.54
18.	UKD-604/LGW	13.50	12.04	10.83	9.76
19.	UKD-604/LHR (A)	12.02	–	–	–
20.	UKD-604/LBA	13.53	12.05	10.82	9.75
21.	UKD-604/LCY	13.47	12.00	10.79	9.72
22.	UKD-604/LTN	13.49	12.01	10.80	9.73
23.	UKD-604/MAN	13.34	11.77	10.44	9.33
24.	UKD-604/SOU	13.56	12.08	10.84	9.75
25.	UKD-604/STN	13.42	11.95	10.74	9.68
	95% CI for k	4.93 to 22.04	4.20 to 19.84	3.75 to 17.87	2.63 to 16.87

Table 4.19. Average semipartial R^2 values ($\overline{R_{sp}^2}$) and interquartile range (IR) of $\overline{R_{sp}^2}$ for $PNL_{max,av}$ and $\log_{10}(N)$ of the midpoint Miedema scaling combination data set (Combination #1) as airports that cause \hat{k} to increase the most are removed from the full data set. Baseline is the full data set with no airports excluded. The removed airports are notated with letters: A denotes UKD-604/Heathrow (LHR), B denotes USA-032/Boston (BOS), and C denotes AUL-210/Perth (PER).

Data set	$PNL_{max,av}$		$\log_{10}(N)$	
	$\overline{R_{sp}^2}$	IR	$\overline{R_{sp}^2}$	IR
Baseline	0.100	0.098 to 0.102	0.080	0.079 to 0.083
Baseline - A	0.094	0.092 to 0.097	0.069	0.067 to 0.071
Baseline - A,B	0.099	0.095 to 0.101	0.062	0.060 to 0.065
Baseline - A,B,C	0.111	0.108 to 0.114	0.054	0.052 to 0.055

with Heathrow is observed. Both UKD-024/LHR (1967) and UKD-604/LHR (2006) feature large sample sizes, so the effect of removing either airport is more noticeable than that from smaller data sets. The UKD-604/LHR (2006) data set also yielded a large regression model R^2 value (0.70) (shown in Table 4.2) in addition to its large sample size, so removing it may have a large effect on the fit of the model with the remaining data. This is probably why AUL-210/SYD (1980) does not have a greater effect than UKD-604/LHR (2006); although \hat{k} for AUL-210/SYD is slightly higher and the sample sizes for both data sets are essentially equal, the R^2 value for AUL-210/SYD is only 0.08, as shown in Table 4.2.

Comparing the changes in \hat{k} and $\overline{R_{sp}^2}$ as airports are removed is also interesting. When UKD-024/LHR (1967) and USA-022/ORD (1967) are removed from the full data set, $PNL_{max,av}$ and $\log_{10}(N)$ become equally generally dominant as assessed by $\overline{R_{sp}^2}$. However, for this case, $\hat{k} = 18.34$ with an estimated 95% confidence interval which does not include the equal-energy value ($k = 10$). When only UKD-024/LHR is removed, $PNL_{max,av}$ has a greater $\overline{R_{sp}^2}$ than $\log_{10}(N)$ and thus is judged more generally dominant even though \hat{k} can be stated with 95% confidence to be greater than 10 (the equal-energy value). This seems to be a conflict of report in the im-

portance measures \hat{k} and $\overline{R_{sp}^2}$. But as mentioned earlier, the importance measures are fundamentally different in what they measure. The removal of UKD-604/HLHR (2006), USA-032/BOS (1969), and AUL-210/PER (1980) is slightly less interesting. The estimated confidence intervals for k for the full data set and the subset data sets all include $k = 10$, the equal-energy value. Also, from $\overline{R_{sp}^2}$ results, $PNL_{max,av}$ is more important than $\log_{10}(N)$ in every case, including the baseline. As airports are removed, \hat{k} decreases, and $\overline{R_{sp}^2}$ for $PNL_{max,av}$ increases.

4.3 Summary

Seven survey data sets were analyzed individually and in various combinations. Relative importance measures were introduced that can be used to infer the relative importance of one predictor variable ($PNL_{max,av}$ or $\log_{10}(N)$) compared to another in the prediction of a criterion variable (annoyance). These importance measures include the ratio of unstandardized regression coefficient estimates (referred to as the decibel equivalent number effect, or \hat{k} , in the current research) and the average semi-partial R^2 ($\overline{R_{sp}^2}$). $\overline{R_{sp}^2}$ was stated to be superior to \hat{k} as an importance measure. The importance measures were calculated for the seven data sets and for the individual airport data sets within the seven data sets. \hat{k} was not consistent across data sets and for most data sets, was accompanied by very large estimated confidence intervals. On average, $PNL_{max,av}$ was quantified as more generally dominant (using the $\overline{R_{sp}^2}$ criterion) in nearly all data sets except those of UKD-604 (2006), in which $\log_{10}(N)$ was quantified as more generally dominant. The data sets were combined according to the methods described in Chapter 3; seven combination data sets resulted. Importance measures were calculated for the combination data sets. \hat{k} was found to be significant ($k > 0$) in the combination data sets, but could not be stated with 95% confidence to be significantly different from the equal-energy value ($k = 10$). $\overline{R_{sp}^2}$ results quantified $PNL_{max,av}$ as being a more generally dominant predictor variable than $\log_{10}(N)$. In only three of the combination data sets was $\log_{10}(N)$ more generally dominant than

the intercept. The effect of including or excluding an airport in a combination data set on the jackknife replicate estimates of k (\hat{k}_r) and on the estimated confidence intervals of k was shown. The importance of determining the additive-log model regression coefficients and \hat{k} with precision was described, and was related to survey sample design. Specifically, stratification can be used to increase the ranges of the sound level and number-of-events in the sample while keeping the correlation between the variables low, which leads directly to improved regression model estimation.

CHAPTER 5. SURVEY SIMULATION

This chapter begins with a description of the importance of sample design to the collection of predictor variable data, which affects model estimation. Next, an overview of a Survey Simulation Program composed of three computer programs is described. The function of each computer program and any special techniques used in each are documented. Then, the design of a survey simulation is presented and results of simulations around three airports are given.

5.1 Sample Design

Sampling consists of selecting some part of a population to observe in order to estimate something about the whole population (Thompson, 1992). Sample design refers to the process by which one selects sampling units from a population. An excellent and concise review of sample design is given in (Kalton, 1983). Sampling differs from experimental design in that in experiments, some part of a population is deliberately perturbed to study the reaction. In sampling, the goal is to observe a population without perturbing or disturbing it. Though sample design is only one aspect of survey design (which involves many decisions such as the mode of data collection, wording of questions, number of questions, analysis methods, etc.), it can be used to affect the distribution of predictor variable data that are collected, which, as we have seen from the discussion in Chapter 4, Section 4.1.3 on the precision of \hat{k} , influences the magnitude of the variance estimate of \hat{k} . In addition, when the predictor variable data are uncorrelated, then a regression model R^2 can be uniquely decomposed into average semipartial R^2 components $\left(\overline{R_{sp}^2}\right)$ that do not change with changing order in stepwise regressions.

In particular, stratification is a powerful tool which can be used to increase the range of predictor variable data and limit the correlation between predictor variables, which accomplishes a smaller variance estimate of \hat{k} and is beneficial for policy-making (Allison, 1999; Farrar and Glauber, 1967) because the relative importance of the predictor variables can be more precisely quantified. 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. Another method of grouping a population is through clustering. With clustering, a population of people is grouped into clusters (usually geographic areas), which become first-stage units for sampling in a multistage clustered sample design. Only a *subset* of the clusters is included in the survey sample. Though stratification and clustering are different methods of grouping the population, both can be applied in surveys. In practice, stratification can be applied at each stage of a multistage clustered sample design. Stratification is especially useful when applied *to* the cluster level, or first stage of a multistage clustered sample design. That is, clusters can be classified into groups to yield much greater gains in precision than would be realized without stratification, since geographically-based clusters tend to be internally homogeneous in the predictor variables (Kalton, 1983). Stratification can be proportionate, in which the strata sample sizes are made proportional to the strata population sizes, or disproportionate. Proportionate stratification ensures that estimates of statistics made with the resulting data are no less precise than what would be realized from a simple random sample of the same size.

When a goal in the analysis of survey data is to identify with precision an explanatory regression model, and to determine relative importance of predictor variables, stratification groups formed by the predictor variable data can be very useful. With application to noise surveys and identifying additive-log annoyance models of Equation (4.1), strata can be formed by segmenting a survey population in groups based on two stratification factors, $\log_{10}(N)$ and the particular sound level measure used. Each stratification factor is divided into groups that may or may not be equally-sized

(in terms of the range of values covered in each group). One strata is formed by a combination of a range of $\log_{10}(N)$ values and sound level values.

Stratification strategies have been applied in past studies, but not consistently. McKennell (1963) stratified by measured sound level ($PNL_{max,av}$) only but after observing the high correlation of the number-of-events with annoyance and deriving an additive-log model, stated that the determination of valid sound level/number-of-events trade-off seemed primarily to be a matter of sample design, since the sound level and number-of-events data were highly correlated. MIL Research Ltd. (1971), Josse (1968), and Brooker et al. (1985) stratified by measured sound level ($PNL_{max,av}$) and number-of-events data. Le Masurier et al. (2007a) stratified by simulated sound level and number-of-events data (L_{av} and N_{av}) calibrated with measurements. In the TRACOR, Inc. (1971) surveys, stratification was not used. TRACOR, Inc. (1971) and Connor and Patterson (1972) stratified by sound level only, through the use of Perceived Noise Level (PNL) contours. Hede and Bullen (1982) stratified only by Noise Exposure Forecast (NEF), which is an additive-log model in the form of Equation (4.1) with a k of 10 and a nighttime event weighting.

Increases in computational resources and advances in noise exposure prediction capabilities have made survey simulation a viable option, where sampling and its effect on model estimation can be explored before embarking on a survey. From noise simulation programs, such as the Integrated Noise Model (INM) (He, Dinges, Hemann, Rickell, Mirsky, Roof, Boeker, Gerbi, and Senzig, 2007) and Noise Model Simulation (NMSim) (Ikelheimer and Plotkin, 2005), estimates of sound level metrics and number-of-events can be used to create strata, and identify a sample which is conducive for model identification. Stratification is helpful to account for all variables that are hypothesized to affect the criterion variable (annoyance), including demographic factors, and can help to reduce covariation between all predictor variables.

5.2 Survey Simulation Program

The Survey Simulation Program developed in this research is a collection of three computer programs. Two of the programs (the Noise Metric Module and the Synthetic Population Generator) generate acoustical and population demographic data, and the third (the Survey Sampling Module) simulates noise surveys using the outputs of the other two programs. Technical details of the Noise Metric Module are described in Appendix I.1, details of the Synthetic Population Generator in Appendix I.7, and details of the Survey Sampling Module in Appendix I.8.

The Noise Metric Module calculates the acoustical environs (number-of-events and sound level metrics) on a geographical grid of points using INM acoustical data for single aircraft events and given an aircraft operations scenario as input. The Synthetic Population Generator uses U.S. Census 2000 data (United States Census Bureau, 2002, 2003) to synthesize the demographics of the population surrounding the airport. Aggregated tabulations at the census block group level, reported in Summary File 3 (SF3) and the 5% Public Use Microdata Sample (PUMS) file, which are the recorded census responses for 5% of a state's population, are used in the demographic specification. The Survey Sampling Module is used to assign the population to grid points and to simulate noise surveys around the airport.

The Survey Sampling Module uses Monte Carlo simulation, see, e.g., (Mooney, 1997), to analyze the results of noise survey simulations. Monte Carlo simulation is useful in constructing sampling distributions of statistics not accounted for by mathematical theory. Monte Carlo techniques allow the study of the behavior of a statistic calculated from real data over a number of trials. In the Survey Sampling Module, each designed noise survey is implemented for a pre-set number of trials, i.e., a Monte Carlo simulation is performed. For each survey (or trial), regression models and related statistics are estimated. From this process, distributions of estimated statistics can be constructed. This enables examination of distributions of statistics for which analytical predictions (of distributions) are difficult, if not impossible.

Surveys were simulated around three airports: Airport 1 (small-sized, 150,000 operations per year), Airport 2 (small to medium-sized, 170,000 operations per year), and Airport 3 (large-sized, 430,000 operations per year). Note that the aircraft operations scenario used at each of the airports for the survey simulations is not reflective of actual current traffic statistics; thus the noise exposure calculated around the airports is not representative of the actual current noise exposure.

5.2.1 Single-Event Acoustical Data

The Noise Metric Module calculates the acoustical environs around airports using INM acoustical data for single aircraft events. A single event is defined as one arrival or departure of a given aircraft operating at a certain flight profile on one flight track. Prior to using the Noise Metric Module, single-event data must be created for every combination of aircraft model and flight track that is desired to be used in a simulated aircraft operations scenario at an airport. Important design choices and requirements for this program include the aircraft models to use, the flight tracks on which the aircraft are simulated, and the grid spacing of points to use in calculating the single-event data. Perhaps the most critical design choice is the specification of the flight tracks. A flight track specifies the latitude and longitude at different points in space that an aircraft uses. The heights are calculated automatically by INM from the standard profile stage numbers assigned to aircraft. Ideally, INM flight track files will be available. In the absence of pre-existing INM flight track files, the files have to be manually created. Descriptions of the flight track files and two methods that can be used to create them are contained in Appendix E.

Single-event data were calculated around the three airports. At each airport, data were calculated for the most-operated aircraft models contributing to at least 90% of the total operation count. Data were generated for the aircraft models on every specified arrival and departure flight track, operating at the the standard profile stage number 1. A grid spacing of 0.1 nautical miles was chosen as it provided a not overly-

resolved picture of the variation in noise exposure around an airport but captured changes in exposure across the landscape well. This grid spacing, though adequate for acoustical calculations, was not fine enough to assign population households to grid points. This was addressed with the Survey Sampling Module, which was used to interpolate the acoustical data matrices prior to the population household assignment. The aircraft models simulated for Airports 1, 2, and 3 are shown in Tables 5.2, 5.3, and 5.1. The flight tracks are shown in Figures 5.1b, 5.1c, and 5.1a. For Airport 1, there were 3 runways, 89 arrival flight tracks and 80 departure flight tracks. For Airport 2, there were 3 runways, 70 arrival flight tracks and 79 departure flight tracks. For Airport 3, there were 4 runways, 31 arrival flight tracks and 71 departure flight tracks.

Table 5.1. Aircraft models used at Airport 1.

INM Aircraft ID	Description
737300	Boeing 737-300/CFM56-3B-1
7373B2	Boeing 737-300/CFM56-3B-2
737700	Boeing 737-700/CFM56-7B24
747400	Boeing 747-400/PW4056
757PW	Boeing 757-200/PW2037
757RR	Boeing 757-200/RB211-535E4
767300	Boeing 767-300/PW4060
A300-622R	Airbus A300-622R/PW4158
BEC190	Beech 1900
CL601	CL601/CF34-3A
CNA560	Cessna 560 Citation V
EMB145	Embraer 145 ER/Allison AE3007
EMB170	Embraer EMB-170
FAL20	FALCON 20/CF700-2D-2
MD11GE	MD-11/CF6-80C2D1F
MD82	MD-82/JT8D-217A
SD360	Shorts 360

Table 5.2. Aircraft models used at Airport 2.

INM Aircraft ID	Description
717200	Boeing 717-200/BR 715
727EM2	FEDX 727-200/JT8D-15
737300	Boeing 737-300/CFM56-3B-1
737500	Boeing 737-500/CFM56-3C-1
737700	Boeing 737-700/CFM56-7B24
757RR	Boeing 757-200/RB211-535E4
777200	Boeing 777-200ER/GE90-90B
A300-622R	Airbus A300-622R/PW4158
A310-304	Airbus A310-304/CF6-80C2A2
A319-131	Airbus A319-131/V2522-A5
A320-211	Airbus A320-211/CFM56-5A1
BEC58P	BARON 58P/TS10-520-L
CL601	CL601/CF34-3A
CNA500	CIT 2/JT15D-4
DC1010	DC10-10/CF6-6D
DC1030	DC10-30/CF6-50C2
DC93LW	DC9-30/JT8D-9 w/ ABS Lightweight hushkit
DHC6	DASH 6/PT6A-27
DHC8	DASH 8-100/PW121
EMB14L	Embraer 145 LR / Allison AE3007A1
GASEPF	1985 1-ENG FP PROP
GIV	Gulfstream GIV-SP/TAY 611-8
GV	Gulfstream GV/BR 710
LEAR35	LEAR 36/TFE731-2
MD82	MD-82/JT8D-217A
MD83	MD-83/JT8D-219
MU3001	MU300-10/JT15D-5

5.2.2 Noise Metric Module

The Noise Metric Module calculates the acoustical characteristics of an aircraft operations scenario around an airport on a grid of points using the single-event data generated from INM. Technical details of the Noise Metric Module are described in Appendix I.1. The program can calculate DNL (from single-event L_{Aeq}), L_{Aeq} , $L_{Aeq,d}$, $L_{Aeq,n}$, $SEL_{A,av}$, $L_{Amax,av}$, $EPNL_{av}$, and $PNLT_{max,av}$. For DNL (from single-event

Table 5.3. Aircraft models used at Airport 3.

INM Aircraft ID	Description
737300	Boeing 737-300/CFM56-3B-1
737400	Boeing 737-400/CFM56-3C-1
737500	Boeing 737-500/CFM56-3C-1
737700	Boeing 737-700/CFM56-7B24
737800	Boeing 737-800/CFM56-7B26
747400	Boeing 747-400/PW4056
757PW	Boeing 757-200/PW2037
757RR	Boeing 757-200/RB211-535E4
767300	Boeing 767-300/PW4060
767CF6	Boeing 767-200/CF6-80A
777200	Boeing 777-200ER/GE90-90B
A319-131	Airbus A319-131/V2522-A5
A320-232	Airbus A320-232/V2527-A5
A321-232	Airbus A321-232/IAE V2530-A5
A340-211	Airbus A340-211/CFM 56-5C2
CL600	CL600/ALF502L
CLREGJ	Canadair Regional Jet
DHC8	DASH 8-100/PW121
EMB120	Embraer 120 ER/ Pratt & Whitney PW118
EMB14L	Embraer 145 LR / Allison AE3007A1
MD82	MD-82/JT8D-217A
MD83	MD-83/JT8D-219
SF340	SF340B/CT7-9B

L_{Aeq}), L_{Aeq} , $L_{Aeq,d}$, and $L_{Aeq,n}$, the program includes background noise at one or more levels which the user specifies. For $SEL_{A,av}$, $L_{Amax,av}$, $EPNL_{av}$, and $PNLT_{max,av}$, the program calculates the overall metrics and thresholded metrics. Overall metrics include the noise exposure from all operations specified in the aircraft operations scenario, no matter the sound level of single aircraft events. Thresholded metrics include the noise exposure from aircraft events with a sound level above a threshold; the user specifies the threshold(s) for each metric. For each calculated sound level metric, the number-of-events at a given grid point is the number of aircraft events that were included in the calculation of the metric. In addition, the number-of-events data are *daily* counts of events. If operations are specified in the aircraft operations

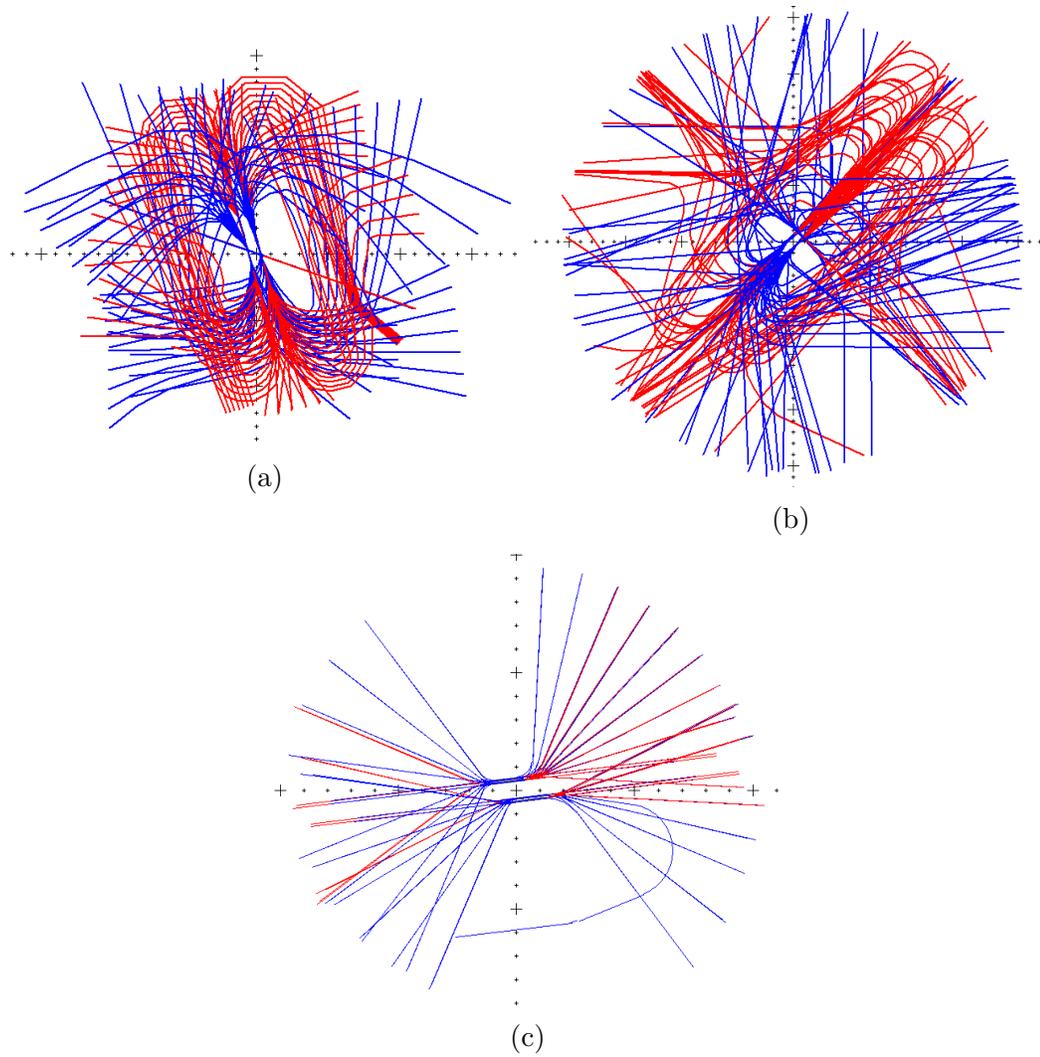


Figure 5.1. Flight tracks for (a) Airport 1, (b) Airport 2, and (c) Airport 3. Grid points are spaced by 1 nautical mile. Red tracks are arrival flight tracks, blue tracks are departure flight tracks.

scenario for which single-event data do not exist, the program can distribute the total number of data operations for which single-event data have not been generated evenly among the operations for which single-event data exist for the airport. For each metric calculated, the exposure from the single-events can be averaged by three

different methods, either logarithmically (i.e., unlogging sound levels and summing energy; then converting back to decibels):

$$\bar{S} = 10 \log_{10} \left(\frac{1}{N} \sum_{i=1}^N 10^{L_i/10} \right), \quad (5.1)$$

where \bar{S} is the average metric, N is the number-of-events, and L_i is the sound level metric in dB of the i^{th} event; arithmetically (summing sound levels directly):

$$\bar{S} = \frac{1}{N} \sum_{i=1}^N L_i; \quad (5.2)$$

or after a loudness transformation is applied to the sound levels (ISO/R-131-1959(E), 1959):

$$\bar{S} = 10 \log_2 \left(\frac{1}{N} \sum_{i=1}^N 2^{L_i - 40/10} \right) + 40. \quad (5.3)$$

Aircraft Operations Scenarios

The aircraft operations scenario is an input to the Noise Metric Module. The information contained in the aircraft operations scenario is used to access the single-event INM data. The following information is listed in a single row of the scenario:

1. Runway name
2. Operation type (arrival or departure)
3. INM flight track name
4. Average daily number of day operations
5. Day operations breakdown
6. Average daily number of night operations
7. Night operations breakdown

Essentially, each row of the scenario is reserved for a flight track. The day and night operations breakdowns are lists of aircraft on each flight track and contain the following information for each aircraft:

1. INM aircraft name
2. Standard profile stage identifier
3. Percentage of operation count

Aircraft operations scenarios can be created with the information contained in an INM ops_ft.dbf file, from information in Federal Aviation Administration (FAA) 14 CFR Part 150 studies, or from tabulated air traffic statistics from an airport. If an INM ops_ft.dbf file is available, *SCENARIO_CREATOR* (see Appendix I.4) can be used to automatically create an aircraft operation scenario in the required format. If an INM ops_ft.dbf file is not available, then the flight operations scenario must be constructed manually. A template of the aircraft operations scenario can be found on a hard drive in the possession of Dr. Patricia Davies of Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN, 47906. An aircraft operations scenario can be constructed with basic knowledge: for each runway and each aircraft that uses the runway, the average daily number of day and night arrivals and departure. Then, the 4 totals (day arrivals, night arrivals, day departures, and night departures) for each runway and aircraft can be divided by the number of arrival or departure flight tracks belonging to the runway. If an INM ops_ft.dbf is available, it is likely that the required flight track files (Appendix E) are also available. If the flight track files are not available, it is required to manually create INM flight tracks so that INM single-event data can be created. Two methods of creating flight tracks are described in Appendix E. The three methods for creating aircraft operations scenarios are described in the next three paragraphs.

INM File In the case that INM files from a FAA 14 CFR Part 150 study can be collected, the ops_ft.dbf file will likely be one of them. The ops_ft.dbf file is the Flight Operations Table that INM 7.0 creates for a “case”. It lists the number of

day, evening, and night operations for each unique combination of aircraft, operation type, and flight track. `ops_ftt` contains all of the information required to create an aircraft operations scenario. A MATLAB function, *SCENARIO_CREATOR*, was coded that creates an aircraft operations scenario based on the information contained in `ops_ftt.dbf`. A default flight operations scenarios for Airport 3 was created from an `ops_ftt.dbf` file.

Part 150 Documentation In the case that INM files from a prior FAA 14 CFR Part 150 study are not available, the Part 150 documentation should be consulted if any is available. Depending on the detail of the study documentation, it is possible that an aircraft operations scenario can be created from the reported data. An aircraft operations scenario was constructed for Airport 2 from data reported in Tables 13-15 of the Part 150 study documentation (Landrum and Brown, Inc., 2008, Tables 13-15). In Table 13 of the Part 150 documentation are given the numbers of arrivals and departures for day and night for each aircraft modeled in INM. In Table 14 of the document are given the percentages-of-use by different aircraft categories for each runway. In Tables 15 and 16 of the document are given the percentages-of-use by different aircraft categories for each arrival and departure track, respectively. This information was sufficient to specify the air traffic for each flight track. Because an INM files were not available for Airport 2, the flight track data required for the calculation of single-event data had to be manually constructed. The process of creating flight tracks for Airport 2 is also described in Appendix E.

Traffic Statistics If neither an INM `ops_ftt.dbf` nor FAA 14 CFR Part 150 study documentation are available, it remains possible to construct an aircraft operations scenario given actual traffic statistics from an airport, that is, for each runway and each aircraft that uses the runway, the average daily number of day and night arrivals and departure. For Airport 1, INM flight track files from a prior FAA 14 CFR Part 150 study were available, but it was desired to update the operations counts. An Open Records Request was filed for aircraft operations data from 2009. The acquired

traffic statistics were used to form the aircraft model list and the percentage-use of each aircraft on each flight track for Airport 1.

Nonexistent Single-Event Data Although the ideal situation is that every combination of flight track, aircraft model, and standard profile stage number specified in the aircraft operations scenario will have accompanying single-event data that has been generated, it is possible that an aircraft operations scenario can specify combinations for which single-event data do not exist. The Noise Metric Module finds every combination for which there is no single-event data and sums the total number of missing data day and night operations. The Module then allows the day and night operations counts for the operations for which there is single-event data to be increased so that the total number of operations contributing to the acoustical calculations matches the total number of operations listed in the aircraft operations scenario.

Airport 3: Specific Issues The creation of a default aircraft operations scenario for Airport 3 was complicated in that the ops_ft.dbf file specified many combinations for which there were no single-event data. The Noise Metric Module could have increased the operations counts of the combinations with single-event data, but the number of operations was so large that another approach was decided upon. The problem was two-fold: the ops_ft.dbf file specified some standard profile stage numbers with no single-event data and some aircraft models with no single-event data. The first problem was addressed by changing all of the stage numbers to 1. The second problem was addressed by pairing missing-data aircraft with aircraft that single-event data were simulated for (nonmissing-data aircraft). The pairing was performed through the use of Noise-Power-Distance (NPD) sound level data.

NPD data are a set of sound levels, expressed as a function of engine power (usually the corrected net thrust per engine), distance, and operation type (arrival or departure) (He et al., 2007). The INM NPD data are corrected for aircraft speed, atmospheric absorption, distance duration, and divergence. INM NPD data are reported

for four sound level metrics: SEL_A , $L_{A,max,s}$ (L_{Amax} calculated using a slow-scale exponential time weighting), $EPNL$, and $PNL T_{max,s}$ (Tone-Corrected Maximum PNL with slow-scale exponential time weighting). One NPD value represents a measurement of sound level taken at a certain distance from an aircraft operating at a certain thrust level. All NPD data in INM are reported at distances of 200, 400, 630, 1000, 2000, 4000, 6300, 10000, 16000, and 25000 feet.

For all missing-data aircraft and the nonmissing-data aircraft, the NPD data were extracted. Each INM aircraft has an associated “Noise Identifier” (NI) that links it to a set of NPD data. Two INM system data files were used to link aircraft and NPD data: `aircraft.dbf` and `npd.curv.dbf`. As mentioned, NPD data are reported for four sound level metrics, but not all four sound level metrics are reported for each NI. The common set of available-metric data among all NI from the missing-data and nonmissing-data aircraft was determined to be SEL_A and $EPNL$. For each available metric of each aircraft, the departure NPD data were logarithmically averaged across thrust values, so that for each metric, there was one set of NPD values that was representative of an average thrust. A logarithmic (energy-based) average is one in which decibels are unlogged prior to averaging and relogged after averaging, so that the average is an energy average. The NPD data of each missing-data aircraft was compared to the full collection of nonmissing-data aircraft through a sum of squared differences across metrics, so that for each comparison, a single number represented the departure of one aircraft’s NPD data from another. Then, for each missing-data aircraft, the nonmissing-data aircraft that departed the least was paired with the missing-data aircraft. Then, the missing-data aircraft were replaced by their paired nonmissing-data aircraft in the aircraft operations scenario. This process was performed through the use of a computer program described in Appendix I.3.

5.2.3 Synthetic Population Generator

The Synthetic Population Generator is used to simulate a population surrounding an airport using Census 2000 data (United States Census Bureau, 2002, 2003). The population is composed of a collection of persons from a state's 5% Public Use Microdata Sample (PUMS) file which best fit Summary File 3 (SF3) census block group aggregated tabulations. The simulated population is designed using combinatorial optimization (described shortly), which operates recursively, and converges on a population design that fits the tabulations in a certain number of iterations. The Census data will be described in the next two sections. Technical details of the Synthetic Population Generator are described in Appendix I.7.

Summary File 3

Summary File 3 (SF3) contains detailed population and housing data (such as place of birth, education, employment status, income, value of housing unit, year structure built) from a 1-in-6 sample of the 2000 United States Census, those that received the Census 2000 long-form questionnaire, and weighted to represent the total population.

SF3 data were chosen to be utilized for population simulation because of the amount of data contained (813 data tables), and the resolution of the data (1-in-6 sample). Alternatives to SF3 are Summary Files 1, 2, and 4. SF1 tabulations include 100% of the people, households, and housing units and report counts for census blocks. The geographic resolution of SF1 data is slightly higher than SF3 and the total count of people, households, and housing units in the aggregate tables represent the absolute numbers sampled during the census, but the number of data tabulations (286 data tables) is low relative to SF3. Particularly, there were aggregate data tables in SF3 desirable for constraining the design of a population that were not reported in SF1. SF2 and SF4 report data at the census tract level, which is less resolved than census block group. Descriptions of Census 2000 geographic entities can be found at <http://www.census.gov/geo/www/tiger/glossry2.pdf>.

The SF3 data for a state consist of 1 geographic header file and 76 data SF3 files. These files are .uf3-formatted files. The files must be transformed to simple .txt files before they can be used by the Synthetic Population Generator. This process is described in Appendix G. A single record of the census data contains all data SF3 data for a certain geographic area and spans all 76 data files and the geographic header file. The segments of each record are tied together across the data files by an identifier number known as the “Logical Record Number” (LOGRECNO).

The geographic header file lists general geographic details (e.g., land area, number of persons, number of households, etc.) for each geographical area record. The 76 data files contain counts of persons, households, and housing units for various aggregate data tables. A given aggregate data table contains count of either the number of persons, households, or housing units meeting certain characteristics in a given geographic area.

An important distinction to make is in the definitions of “household” and “housing unit” used by the Census Bureau. A “household” consists of all the people who occupy a housing unit. A “housing unit” is a space that is intended for occupancy as separate living quarters, where the occupants do not live and eat with any other persons in the structure and there is direct access from the outside or through a common hall. A house, apartment, group of rooms, or single room are examples of a housing unit.

Documentation for the SF3 data can be found at <http://www.census.gov/prod/cen2000/doc/sf3.pdf>. On pages 7-1 to 7-14 is listed the data that the geographic header file contains. In Figure 2-2, which begins on page 2-4 of the SF3 .pdf document is listed a general breakdown of the data (by aggregate data table number) contained in each of the 76 data SF3 files. On pages 7-25 to 7-534 is listed, for each aggregate data table in the 76 data SF3 files, the specific data items that it contains.

Microdata

Census PUMS (Public Use Microdata Sample) data (or microdata) contain informational records of occupied and vacant housing units in the U.S. and the people in the occupied housing units. Records corresponding to persons are computerized versions of the long-form census questionnaires, as coded and edited during census processing. A microdata file starts with a housing unit record and is followed by all of the records of the persons residing in the housing unit. This pattern is followed for all of the housing units contained in the microdata file. Each state has two microdata files which contain 1% and 5% of the recorded population. The 1% sample contains more data for each record, but is a smaller sample of housing units and persons. The 5% data file was preferred over the 1% file for simulating a population because of the greater number of persons available to design the population with. Both of the files are .txt files so no format transformation is necessary.

Records within microdata files are organized into geographic units known as Public Use Microdata Areas (PUMAs) and super-Public Use Microdata Areas (super-PUMAs). A super-PUMA is a group of PUMAs. Minimum population thresholds are set for PUMAs and super-PUMAs to maintain confidentiality of the persons. Each PUMA of a 5% file has a minimum population of 100,000. Each state is comprised of one or more super-PUMAs or PUMAs. Documentation for the microdata can be found at <http://www.census.gov/prod/cen2000/doc/pums.pdf>.

Comparison of Summary File 3 and Microdata

SF3 data contain predefined tabulations of characteristics for each record of SF3 data. Each data record contains information about a geographic entity, e.g., state, county, census tract, census block group, or census block. Each microdata data record contains either information about a housing unit or questionnaire responses of an person. It is impossible to deduce all of the questionnaire responses of a single person from SF3 data. Microdata give flexibility to an individual researcher to create any

desired aggregate tabulations from the available person/housing unit characteristics contained in the microdata. Each microdata record excludes personal identifying information and data items that could be used to identify a person.

Instructions for Creating a Synthetic Population

Basic instructions for creating a synthetic population around an airport are outlined in this section. Detailed instructions are given in Appendix F.

1. Create a project folder.
2. Collect Census 2000 Summary File 3 data from the appropriate United States Census Bureau FTP site.
3. Collect Census 2000 5% Microdata from the appropriate United States Census Bureau FTP
4. Identify a group of households and persons listed in the microdata files which will be used to simulate the population.
5. Determine geographic areas for which to simulate a population.
6. Create Synthetic Population Generator inputs.

Combinatorial Optimization

Combinatorial optimization is a recursive optimization algorithm that utilizes data constraints regarding the item to be optimized and converges on an item design that fits the constraints (Huang and Williamson, 2001; Ryan, Maoh, and Kanaroglou, 2009; Voas and Williamson, 2000, 2001). With application to population design, combinatorial optimization allows a population to be synthetically constructed with the use of two types of census data, aggregate data and microdata. The use of microdata allows population construction with persons and housing units that exist in reality. In theory, construction could proceed given only aggregate data; however, only a fictitious person, household, or housing unit could be generated from different distributions of characteristics. Ascertaining a person, household, or housing unit is

impossible given only aggregate data. In the Synthetic Population Generator, populations are designed for census block groups using combinatorial optimization. The mathematics of the combinatorial optimization algorithm are outlined in (Huang and Williamson, 2001, p. 54-55). The traditional process of combinatorial optimization applied to population design is as follows. This process is done for each census block group:

1. Randomly select a group of households (and all persons belonging to the households) from the microdata to populate the census block group.
2. Randomly choose a household from the selection to be replaced.
3. Randomly choose a proposed replacement household from the microdata.
4. Analyze the goodness-of-fit of both households to the given aggregate data constraints. Implement the replacement decision rule (explained below) to decide which household will be included in the population design.
5. Repeat until the maximum-allowable number of iterations have been completed, or the fit to the constraints is perfect, whichever occurs first.

A few modifications were made to the traditional optimization algorithm. The first modification was that the initial selection of households made in point 1 was not entirely random. Prior to selecting households from microdata, the fit of each microdata household to the Summary File 3 household tables was noted. A random selection of the best-fitting microdata households was used as the initial selection of households. The second modification was to point 3 in that the replacement household was also selected from the best-fitting households. The third modification was to point 5 in the addition of a stopping rule. For each census block group, the goodness-of-fit is analyzed every 200 iterations. Optimization is halted if the goodness-of-fit of the population to the household-level and person-level data tables has not increased in 200 iterations, which means a household has not been identified that would improve the fit.

The decision rule described in point 4 is used to determine which household and associated persons (either the original or the proposed replacement) is included in the population design. The process doesn't automatically include the household that improves the fit of the population to the aggregate data constraints. Household replacements or "retrograde swaps" leading to a moderate degradation in performance are allowed so that the algorithm might backtrack from suboptimal solutions. The probability of the retrograde swap decreases with the increase of the number of successful replacements and is determined by two parameters: the starting "temperature", a constant between 0 and 1, which is the initial probability of allowing the retrograde swap, and the number of household replacements. The two parameters are combined in a single metric called the decision ratio, which is equal to the ratio of the starting temperature to the number of replacements. As the number of household replacements increase, the probability of allowing a retrograde swap decreases.

Created Synthetic Populations

Populations were simulated around Airports 1, 2, and 3. Summary File 3 tables P145, P146, and P151 were used to constrain the population design. For Airport 1, populations were simulated for the 556 census block groups of the county the airport resides within. For Airport 2, populations were simulated for 815 census block groups in four counties surrounding the airport, since the airport is near the center of the collection of those four counties. For Airport 3, populations were simulated for 2526 census block groups within 10 nautical miles of the airport; a simulation of the population of the full county was not required, being that areas greater than eight nautical miles from the airport do not experience a significant amount of aircraft noise exposure relative to the areas near the runways (LAWA Noise Management, 2012). The number of household and person microdata records used along with the average number of households and persons per census block group at each airport are shown in Table 5.4.

Table 5.4. Number of household and person microdata records used and the average number of households and persons per census block group at the three airports.

Airport	Number of microdata records		Average number per block group	
	Household	Person	Household	Person
Airport 1	31,956	69,808	516	1,247
Airport 2	25,967	57,928	560	1,406
Airport 3	156,313	442,832	486	1,387

The starting temperature was 0.1 so that the initial probability of accepting a replacement household that worsened the fit of the population to the aggregate tables was 1 in 10. The maximum number of iterations allowable for each census block group was 4,500,000. The average, minimum, and maximum number of iterations to realize the best fit of the population to each census block group of the three airports are shown in Table 5.5. The average number of iterations used for each airport was well under the maximum-allowable number of iterations. The low number of iterations required, and at all three airports, the presence of at least one occurrence in which only the initial selection of households was required to realize the best fit of the population to the census block group is likely attributable to the use of a sample of the best-fitting microdata households to form the initial household selection.

Table 5.5. The average, minimum, and maximum number of iterations, i , for the census block groups of the three airports.

Airport	Average i	Minimum i	Maximum i
Airport 1	4,228	1	23,205
Airport 2	1,565	1	11,979
Airport 3	17,282	1	12,389

5.2.4 Survey Sampling Module

The Survey Sampling Module is used to simulate noise surveys around airports and analyze the effectiveness of different sample designs. The Module uses Monte Carlo techniques, see, for example, (Mooney, 1997) to simulate each designed noise survey a pre-set number of trials and evaluates the noise survey design from an analysis involving results from all of the trials. The Survey Sampling Module is composed of three MATLAB .m files that involve survey design, sampling, and analysis.

Survey Design Decisions

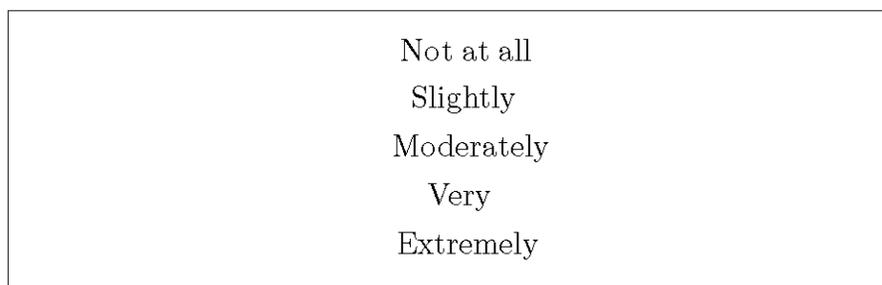
Within the Survey Sampling Module, survey design involves making decisions about different elements of a survey. The elements will be listed and then be described. The elements include:

1. Airport
2. Aircraft operations scenario
3. Population annoyance model
4. Whether or not to use a “podding” approach to assign noise exposure to respondents
5. Signal-to-noise ratio between modeled annoyance generated by the population mean annoyance model and the actual annoyance in the population
6. Maximum annoyance scale value
7. Continuous or discrete annoyance scale

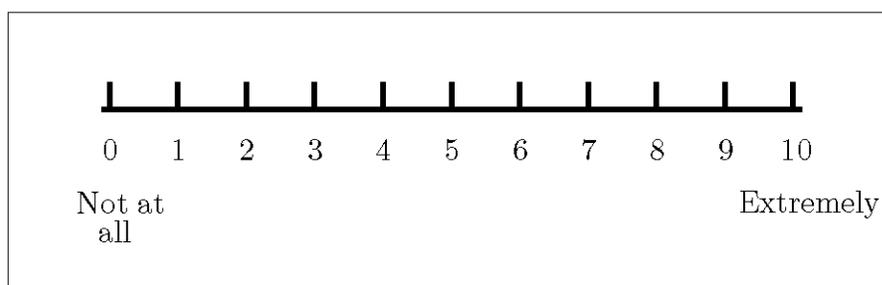
Element 1 refers to the airport around which the survey will be simulated. Element 2 refers to the aircraft operations scenario at the airport. This choice determines what acoustical data will be used for sampling and analysis purposes. Element 3 refers to the population mean annoyance model. The population mean annoyance model generates annoyance reactions at each grid point surrounding the airport, which will be assigned a household from the population created with the Synthetic Population

Generator. The population mean annoyance model annoyance values are modified with use of a signal-to-noise ratio (Element 5), which will be described shortly, to simulate variance in survey respondents' annoyance reactions. The population mean annoyance model can include acoustical and demographic variables. The demographic variables can be household-level and/or person-level. Any person-level variables are applied at the collection of a sample of respondents, since in each household there can be multiple respondents. Also applied at the collection of the sample is a person-level annoyance modification sampled from a uniform distribution which ranges from -0.5 to 0.5. This adds further variability in annoyance at the level of the person. Element 4 refers to a method of assigning acoustical data to grid points (Le Masurier et al., 2007a). Le Masurier et al. (2007a) used a "podding" approach by which all sample points within a Census Output Area were assigned the acoustical exposure simulated at the population-weighted centroid. In the Survey Sampling Module, the podding can be applied to census block groups so that all grid points within a census block group are assigned the acoustical data occurring at the household-weighted centroid. Element 6 is the maximum number of the annoyance scale used in the noise survey; 1 is the minimum value. The annoyance scale is representative of what the sampled persons in a noise survey are asked to choose their annoyance levels from in response to an annoyance question. Element 7 specifies whether the annoyance values at grid points should be discrete (rounded to the nearest integer) or continuous (not rounded). Verbal annoyance scales are discrete by nature, but numerical scales can be discrete or continuous. Recommended verbal and numerical annoyance scales are shown in Figure 5.2. An example of a discrete annoyance question is: "Thinking about the last 12 months or so, when you are here at home, how much does noise from aircraft bother, disturb, or annoy you; Extremely, Very, Moderately, Slightly or Not at all?" An example of a continuous annoyance question is: "Here is a zero to ten opinion scale for how much aircraft noise bothers, disturbs or annoys you when you are here at home. If you are not at all annoyed choose zero, if you are extremely annoyed choose ten, if you are somewhere in between choose a number between zero

and ten. Thinking about the last 12 months or so, what number from zero to ten best shows how much you are bothered, disturbed, or annoyed by aircraft noise?” (Fields et al., 2001). All of the sample individuals’ annoyance values are scaled so that each annoyance value ranges from 1 to the maximum annoyance scale number. The scaling of the annoyance scale is performed using Equation (3.26).



(a)



(b)

Figure 5.2. Recommended (a) verbal and (b) numerical annoyance scales (Fields et al., 2001).

As mentioned, the signal-to-noise ratio (Element 5) is used to simulate the presence of variability in the population at the household level. The signal-to-noise ratio (R_n^2) is used to create a Gaussian distribution centered on 0 from which annoyance modifications are sampled and applied to the grid points. The variance of the Gaussian distribution is determined from:

$$var_{est}(e) = var_{est}(y)(1 - R_n^2), \quad (5.4)$$

where $var_{est}(e)$ is the variance of the Gaussian distribution, $var_{est}(y)$ is the variance estimate of the population mean annoyance model annoyance values, and R_n^2 is the signal-to-noise ratio. Equation (5.4) is derived from the equation for the multiple linear regression R^2 . The classical definition of R^2 is:

$$R^2 = 1 - \frac{SSE}{SST}, \quad (5.5)$$

where SSE is the error sum of squares, and SST is the total sum of squares. The SSE and the model sum of squares (SSM) sum to the SST . The SSE , SSM , and SST are defined by:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (5.6)$$

$$SSM = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad (5.7)$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2, \quad (5.8)$$

where n is the sample size, y_i is the i^{th} criterion variable (annoyance) value, \hat{y}_i is the predicted value of y_i from a regression model, and \bar{y} is the average criterion variable value. SSE is the sum of the squared deviations between the actual and predicted criterion variable values (the residual errors). SSM is the sum of the squared deviations between the predicted criterion variable values and the average criterion variable value. SST is the sum of the squared deviations between the actual criterion variable values and the average criterion variable value. From SSE and SST , unbiased estimates of the variances of the residual errors and criterion variable values can be calculated by dividing each sum of squares by its degrees of freedom. These estimates are referred to as the mean square error (MSE) and the mean square total (MST):

$$MSE = \frac{SSE}{n - p}, \quad (5.9)$$

$$MST = \frac{SST}{n - 1}, \quad (5.10)$$

where p is the number of predictor variables in the regression model, $(n - p)$ is the degrees of freedom of SSE , and $(n - 1)$ is the degrees of freedom of SST . MSE is an unbiased estimate of the variance of the regression error and MST is an unbiased estimate of the variance of the criterion variable values. If both the SSE and SST are divided by $(n - 1)$ in Equation (5.5), then the equation for R^2 can be written:

$$R^2 = 1 - \frac{var_{est}(e)}{var_{est}(y)}, \quad (5.11)$$

where $var_{est}(e)$ is a biased estimate of the variance of the regression residual errors, and $var_{est}(y) = MST$ is an unbiased estimate of the criterion variable values. Equation (5.4) proceeds directly from Equation (5.11). In practice, because the population grids are so large, the downward bias of the $var_{est}(e)$ calculated from Equation (5.4) is negligible. The calculated $var_{est}(e)$ was used to construct a Gaussian distribution of errors from which the population mean annoyance model annoyance value modifications were sampled. A Gaussian distribution was used because of the assumption of the multiple linear regression model that the errors are normally distributed.

A critical function of the Survey Sampling Module is to assign households and associated people of the simulated population created with the Synthetic Population Generator to grid points. In order that each household can be assigned to a grid point, the Survey Sampling Module interpolates the acoustical data grid, if necessary, so that in each census block group, there are a number of grid points greater than or equal to the number of households. In some cases, the interpolation requirements might exceed the available computer memory. In these cases, the acoustical data matrices are interpolated as many times as possible, and some grid points are assigned multiple households. Census block group TIGER/Line shapefiles of the census block groups surrounding the airport are required for the interpolation process, so that the number of grid points in each census block group can be known. TIGER stands for Topologically Integrated Geographic Encoding and Referencing; TIGER/Line files are created by the United States Census Bureau and contain to-

pographical data of land attributes and census regions. TIGER/Line shapefiles are distributed by ESRI (Environmental Systems Research Institute) and can be found at http://arcdata.esri.com/data/tiger2000/tiger_download.cfm. The shapefiles can be downloaded at the county or state level. Because the amount of data at the state level can be overwhelming, especially in plotting applications, it is recommended that county-level data be used. Shapefiles for multiple counties can be used by the Survey Sampling Module. For each county (or state), the “Block Groups 2000” layer should be downloaded. There are three files for each download with different extensions: .dbf, .shp, and .shx. All files should be stored within the appropriate directory. The shapefiles should be stored in a folder named either after the airport or the city in which the airport resides within the “Shapefile” directory described in Appendix H.1.

The census block groups of Airports 1, 2, and 3 for which populations were simulated for in the Synthetic Population Generator are shown in Figures 5.3a, 5.3b, and 5.3c, respectively. The simulated household locations for the simulated populations surrounding Airports 1, 2, and 3 are shown in Figures 5.4a, 5.4b, and 5.4c, respectively. Road shapefiles also available from ESRI were used in the creation of the plots of Figure 5.4. The road shapefiles are contained in the “Line Features - Roads” layer.

Airport 3: Specific Issues There was a complication in the assignment of households to grid points for Airport 3. Airport 3 borders the ocean and the Census 2000 census block groups along the coastline extend a number of miles into the ocean. This initially resulted in the assignment of households to grid points in the ocean for the households in the coastline census block groups. This was remedied by using water polygon shapefiles, also available from ESRI. The water polygon shapefiles are contained in the “Water Polygons” layer. The water polygon shapefiles were used to exclude all grid points residing within the water polygons from being assigned households. Shown in Figure 5.5 are the original simulated household locations for Airport

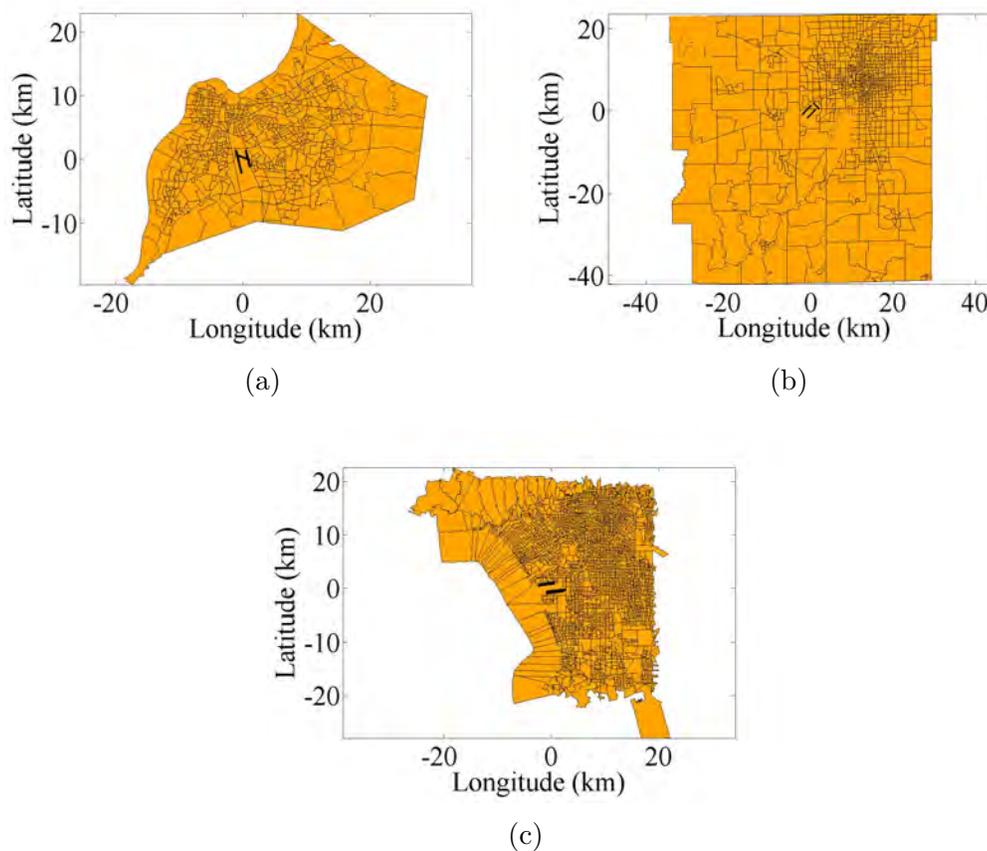


Figure 5.3. Census block groups surrounding (a) Airport 1, (b) Airport 2, and (c) Airport 3. Runways are shown in black.

3. The remedied household locations are shown in Figure 5.4c. An overlay of the census block groups and water polygons for Airport 3 are shown in Figure 5.6.

Survey Sampling Decisions

Decisions are also made regarding the sampling of the population in the Survey Sampling Module. The elements will be listed and then be described. The elements include:

1. Stratification variables (if desired), number of groups, and group boundaries
2. Sample size

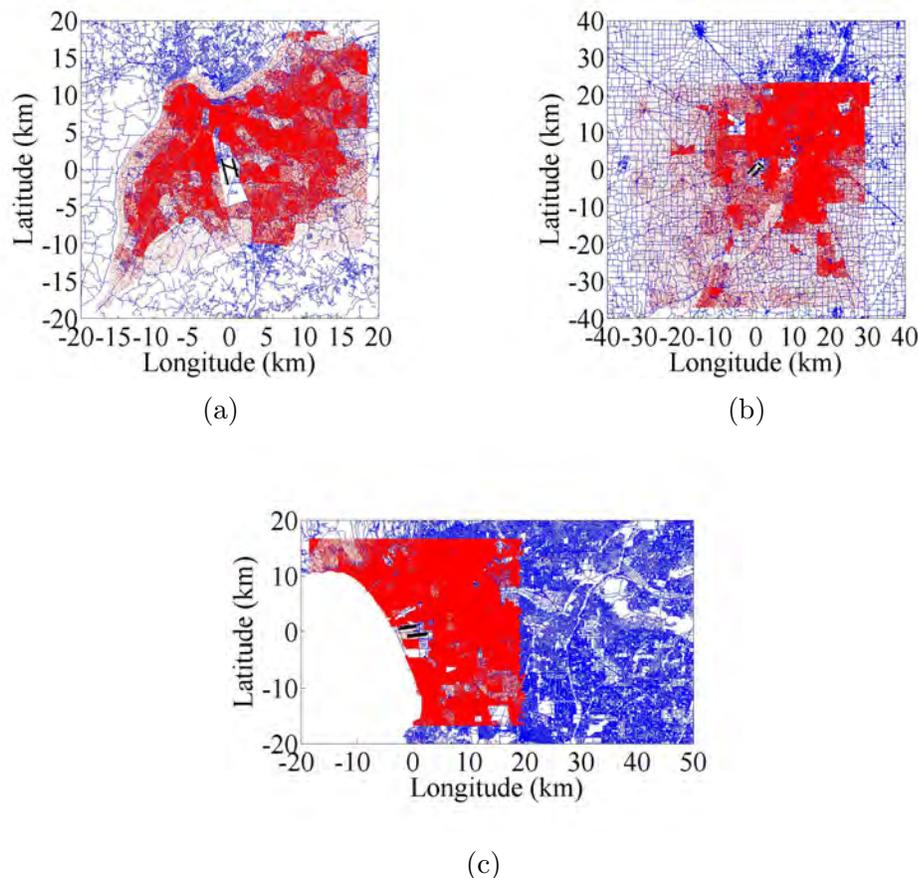


Figure 5.4. Simulated household locations (shown in red) for the populations surrounding (a) Airport 1, (b) Airport 2, and (c) Airport 3. Roads (in blue) and runways (in black) are also shown.

3. Minimum age of sampled person
4. Number of Monte Carlo trials

The sample design that the program uses is simple random sampling without replacement (SRSWOR) with optional stratification. In SRSWOR, persons are randomly sampled from the population with the requirement that once sampled, they cannot be sampled again. If stratification is desired, Element 1 in the list above determines the variables to be used to stratify the population, the number of groups for each variable, and the specific group boundaries for each variable. Stratification was described in detail in Section 5.1. In the Survey Sampling Module, acoustical

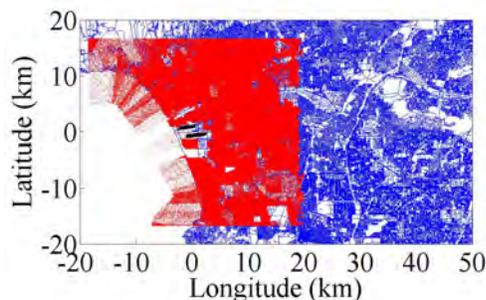


Figure 5.5. Simulated household locations (shown in red) for the population surrounding Airport 3 prior to the use of water polygon shapefiles. Roads (in blue) and runways (in black) are also shown.

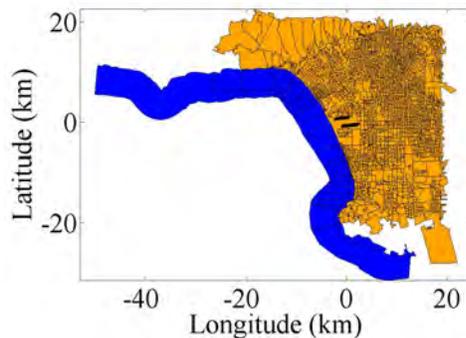


Figure 5.6. Census block groups (orange) and water polygons (blue) in the vicinity of Airport 3. Runways are shown in black.

and demographic variables can be used to stratify the population. Variables can be added by modifying the `model_term_list.txt` text file described in Appendix H.2.2.

Element 2 is the size of the sample to draw from the population surrounding the airport. Element 3 is used by the program to include only those persons above a minimum age in the sample. Element 4 is the number of Monte Carlo trials to use, or the number of times the survey design should be simulated. Each time the survey design is simulated, a new sample is collected from the population and the results are analyzed. Once all trials are completed, the survey design can be evaluated by

using the survey results from all Monte Carlo trials in aggregate analyses. The types of analyses available are described in the next section.

Survey Analysis Decisions

Decisions are also made regarding the analysis of data resulting from sampling the population in the Survey Sampling Module. Two types of analysis occur in the Survey Sampling Module. The first type is implemented at the conclusion of each Monte Carlo trial and involves only the survey data from that trial. The second type is implemented at the conclusion of all Monte Carlo trials and involves the results from all Monte Carlo trials.

For the first type of analysis, decisions are not made regarding the analyses performed on the survey data; the analyses are pre-set in the program. Decisions include choices about the multiple linear regression annoyance models to estimate from survey data, specifically in the number of annoyance models, the variables in each annoyance model, and the power that each variable is raised to. For each annoyance model that is to be estimated, the following are calculated and stored:

1. Estimates of the regression coefficients
2. Total, error, and model sum of squares (SST , SSE , and SSM)
3. Estimates of the variances of the estimates of the regression coefficients
4. The coefficient of determination or “R squared”, R^2
5. Occurrence of Type I error (for variables not in the population mean annoyance model)
6. Occurrence of Type II error (for variables in the population mean annoyance model)

In addition to decisions about the annoyance models to be estimated, the Type I error rate, α , is also specified, which is used to determine the occurrences of Type I

and Type II errors. Estimates of the regression coefficients are calculated by solving the normal equations:

$$[\mathbf{X}'\mathbf{Y}] = [\mathbf{X}'\mathbf{X}] \hat{\mathbf{b}}, \quad (5.12)$$

where \mathbf{X} is a n -by- p matrix of predictor variable data, n being the number of samples

and p being the number of predictor variables: e.g., $\mathbf{X} = \begin{bmatrix} 1, & L_1, & \log_{10}(N_1) \\ 1, & L_2, & \log_{10}(N_2) \\ \vdots & \vdots & \vdots \\ 1, & L_i, & \log_{10}(N_i) \end{bmatrix}$, where

L_i are the values of the sound level metric and N_i are the values of the number-of-events metric. \mathbf{X}' is the transpose of \mathbf{X} , and \mathbf{Y} is a n -by-1 vector of criterion variable (annoyance) values. SSE , SSM , and SST were defined earlier in Equations (5.6), (5.7), and (5.8), respectively. Estimates of the variances of the estimates of the regression coefficients are given by:

$$\hat{\sigma}^2\{\hat{\mathbf{b}}\} = MSE [\mathbf{X}'\mathbf{X}]^{-1}, \quad (5.13)$$

where the mean square error, MSE , is as defined in Equation (5.9). R^2 is calculated through Equation (5.5)

The occurrences of Type I and Type II errors are determined through the estimation of Bonferroni joint confidence intervals for the regression coefficient estimates. The confidence interval for the i^{th} regression coefficient estimate is expressed by:

$$\hat{b}_i \pm B \hat{\sigma}\{\hat{b}_i\}, \quad (5.14)$$

where \hat{b}_i is the i^{th} regression coefficient estimate, B is the Bonferroni multiple, and $\hat{\sigma}\{\hat{b}_i\}$ is the estimated standard deviation of the i^{th} regression coefficient estimate. The Bonferroni multiple is calculated by:

$$B = t\left(1 - \frac{\alpha}{2g}; n - p\right), \quad (5.15)$$

where t is a t -distribution with $n - p$ degrees of freedom, and g is the number of confidence intervals to be jointly estimated. The Bonferroni multiple is used in the confidence interval estimation so that the $1-\alpha$ family confidence intervals for the regression coefficient estimates can be estimated. The use of the normal t -multiple instead of the Bonferroni multiple (or another joint estimation multiple) to estimate the confidence intervals of more than one regression coefficient estimate simultaneously is inappropriate. The use of the Bonferroni multiple adjusts the value of α for each interval to achieve an overall family confidence at the desired α for the joint confidence region. The situations in which Type I and Type II errors occur are shown in Table 5.6.

Table 5.6. Statistical hypothesis test decision matrix. H_0 is the null hypothesis: The i^{th} predictor variable p_i is not in the population mean annoyance model. The alternative hypothesis (H_A) is that the i^{th} predictor variable p_i is in the population mean annoyance model. α is the Type I error rate and β is the Type II error rate.

Reality Conclusion	H_0 True <i>p_i is not in the model</i>	H_0 False <i>p_i is in the model</i>
H_0 Not Rejected <i>Confidence interval of p_i includes 0</i>	$1 - \alpha$ The Confidence Level: <i>The odds of correctly rejecting H_A</i>	β The Type II Error Rate: <i>The odds of incorrectly rejecting H_A</i>
H_0 Rejected <i>Confidence interval of p_i does not include 0</i>	α The Type I Error Rate: <i>The odds of incorrectly rejecting H_0</i>	$1 - \beta$ The Statistical Power: <i>The odds of correctly rejecting H_0</i>

Type I error, Type II error, and statistical power can be demonstrated through a superposition of two statistical distributions. One distribution is that of a statistic under a null hypothesis (H_0), the other is that of the statistic under the alternative hypothesis (H_A). In Figure 5.7 are shown an example of two distributions for \hat{k} . The black curve represents the distribution of \hat{k} under a null hypothesis, that $k = 10$, the equal-energy value. The red curve represents the distribution of \hat{k} under an alternative

hypothesis, that $k = 14$, in which the contribution of number-of-events exceeds that of sound energy in the prediction of annoyance. The blue (darker-shaded) regions in the tails of the null hypothesis distribution denote the $\frac{\alpha}{2}$ probability regions. If H_0 is true and an estimate of \hat{k} from survey data is less than or greater than the lower or upper $\frac{\alpha}{2}$ points, respectively, than H_0 is rejected. The green (lighter-shaded) region, which is the area under the alternative hypothesis distribution to the left of the upper $\frac{\alpha}{2}$ point, denotes the Type II error region. The Type II error, β , is the area contained within this region. If H_A is true, and an estimate of \hat{k} is less than the upper $\frac{\alpha}{2}$ point, then H_0 is not rejected. The statistical power then, is the area contained within the region of the alternative hypothesis distribution to the right of the upper $\frac{\alpha}{2}$ point.

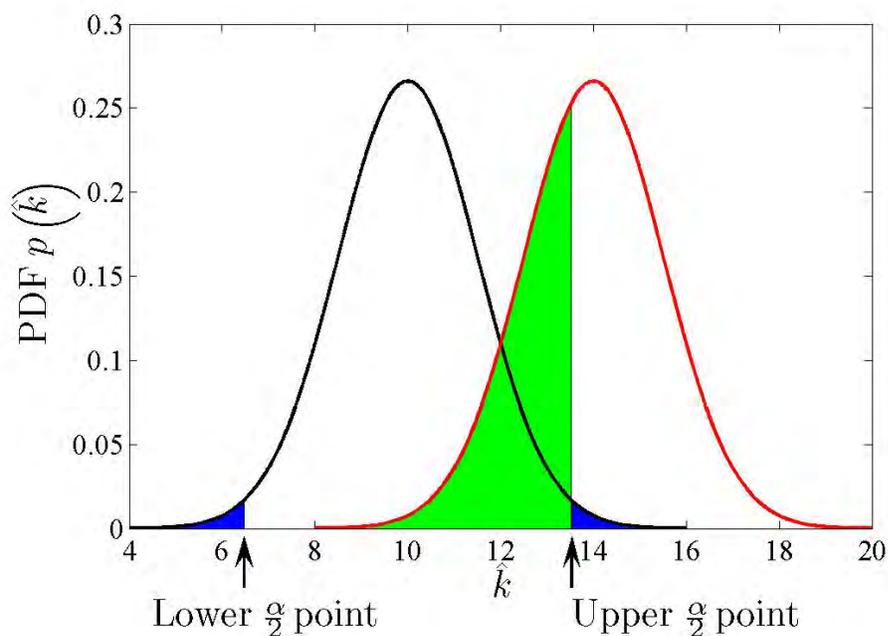


Figure 5.7. Superposition of two hypothetical distributions for \hat{k} , each with a standard deviation of 1.5. The black curve represents the distribution of \hat{k} under a null hypothesis ($k = 10$). The red curve represents the distribution of \hat{k} under an alternative hypothesis ($k = 14$). The blue regions denote the $\frac{\alpha}{2}$ probability regions. The green region denotes the Type II error region, and contains a probability of β . Here, $\alpha = 0.05$, $\beta = 0.397$, and $P = 1 - 0.397 = 0.603$.

Additional analyses for each Monte Carlo trial are also available:

1. Multicollinearity analysis (analysis Option 1 listed in Appendix H.2.4)
2. Retrospective statistical power analysis of \hat{k} (analysis Option 2 listed in Appendix H.2.4)
3. Variance analysis of \hat{k} (analysis Option 3 listed in Appendix H.2.4)
4. Extra sum of squares (analysis Option 4 listed in Appendix H.2.4)

Analysis Option 1 uses the techniques outlined in (Belsley, 1991b) to evaluate the presence of multicollinearity between the predictor variables. The general problem of multicollinearity is described in Appendix C. Analysis Options 2 and 3 deal with the decibel equivalent number effect (\hat{k}). These analyses can only be performed on annoyance models that were estimated with sound level and number-of-events data. The number-of-events data does not have to be base-10 logged. Analysis Option 2 calculates, for specified statistical powers, the minimum and maximum (alternative) values of \hat{k} that can be differentiated from a hypothetical baseline (null) value of \hat{k} using the collected survey data at one Type I error rate, α .

Analysis Option 3 uses the jackknife repeated replication and Taylor series linearization techniques outlined in Appendix B to calculate the simple random sample variance estimate of \hat{k} . Analysis Option 4 performs p regressions for annoyance models that are functions of p predictor variables. For each of the p predictor variables, a marginal regression analysis is carried out to calculate an incremental increase in the model sum of square (SSM), or equivalently, an incremental decrease in the error sum of squares (SSE), due to the inclusion of that predictor variable in the model; the incremental contribution from the variable is called the extra sum of squares of the variable. To calculate the extra sum of squares for the i^{th} predictor variable, first a regression model with all p predictor variables is estimated and the error sum of squares is calculated: $SSE(Y|X_k, k = 1, 2 \dots p)$, where the regression model is:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots b_kX_k + \epsilon. \quad (5.16)$$

Next, a regression model with all predictor variables except X_i is estimated and the error sum of squares is calculated: $SSE(Y|X_k, k = 1, 2 \dots p, k \neq i)$. The extra sum of squares for X_i is defined as:

$$SSM(X_i|Y) = SSE(Y|X_k, k = 1, 2 \dots p, k \neq i) - SSE(Y|X_k, k = 1, 2 \dots p). \quad (5.17)$$

Generally, this type of sum of squares is referred to as the Type III sum of squares, which is Overall and Spiegel's method I (Overall and Spiegel, 1969). A Type III sum of squares represents the unique contribution of a predictor variable to the prediction of a criterion variable, over and above what is contributed by all other predictor variables. The general linear test is used to test the significance of X_i to the prediction of the criterion variable. The general linear test compares the value of a test statistic (F^*) to a point on the central F -distribution, which is how F^* is distributed under the null hypothesis which is that the regression coefficient estimate of X_i is equal to zero ($\hat{b}_i = 0$). The general linear test uses the extra sum of squares to measure the significance of a predictor variable. F^* is calculated by:

$$F^* = \frac{SSE(Y|X_k, k = 1, 2 \dots p, k \neq i) - SSE(Y|X_k, k = 1, 2 \dots p)}{(n - p - 1) - (n - p)} \div \frac{SSE(Y|X_k, k = 1, 2 \dots p)}{n - p}. \quad (5.18)$$

The decision rule for the general linear test is:

$$\begin{aligned} \text{If } F^* \leq F(1 - \alpha; (n - p - 1) - (n - p), n - p) & \text{ conclude } H_0, \\ \text{If } F^* > F(1 - \alpha; (n - p - 1) - (n - p), n - p) & \text{ conclude } H_a, \end{aligned} \quad (5.19)$$

where H_0 is that $\hat{b}_i = 0$ and H_a is that $\hat{b}_i \neq 0$.

At the conclusion of all Monte Carlo trials, results across all trials are analyzed together in the second type of analysis. The following simulation-level statistics are calculated and stored for each estimated model:

1. Mean of the regression coefficient estimates

2. Mean absolute deviation of the regression coefficient estimates (from the values in the population mean annoyance model)
3. Bias of the mean regression coefficient estimates
4. Variance of the regression coefficient estimates
5. Mean absolute deviation of the regression coefficient standard deviation estimates

For the i^{th} predictor variable of an estimated annoyance model, the mean of the regression coefficient estimates is calculated by:

$$\overline{\hat{b}_i} = \frac{\sum_{t=1}^m \hat{b}_{i,t}}{m}, \quad (5.20)$$

where m is the number of Monte Carlo trials and $\hat{b}_{i,t}$ is the t^{th} estimate of the regression coefficient of the i^{th} predictor variable. The mean deviation of a regression coefficient estimate described in point 2 is calculated by using:

$$MD_{\hat{b}_i} = \frac{\sum_{t=1}^m |\hat{b}_{i,t} - b_i|}{m}, \quad (5.21)$$

where b_i is the true value of the coefficient in the population mean annoyance model. If the variable is not in the population mean annoyance model, $b_i = 0$. The bias of the mean estimate is estimated by using:

$$Bias_{\hat{b}_i} = \overline{\hat{b}_i} - b_i. \quad (5.22)$$

The variance of the regression coefficient estimates from all m Monte Carlo trials is estimated by using:

$$\hat{S}_{\hat{b}_i}^2 = \frac{\sum_{t=1}^m (\hat{b}_{i,t} - \overline{\hat{b}_i})^2}{m - 1}. \quad (5.23)$$

The mean deviation of the standard deviation estimates described in Point 5 above is calculated by using:

$$MD_{\hat{s}_{\hat{b}_i}} = \frac{\sum_{t=1}^m |\hat{s}_{\hat{b}_i,t} - \hat{S}_{\hat{b}_i}|}{m}, \quad (5.24)$$

where $\hat{s}_{\hat{b}_{i,t}}$ is the t^{th} estimate of the standard deviation of the regression coefficient estimate of the i^{th} predictor variable. $MD_{\hat{s}_{\hat{b}_i}}$ is the mean deviation of the standard deviation estimates of \hat{b}_i ($\hat{s}_{\hat{b}_i}$) from the individual Monte Carlo trials from $\hat{S}_{\hat{b}_i}$.

One additional analysis that involves the results across all Monte Carlo trials is available. The analysis is the construction of a condition number distribution, which is analysis Option 5 listed in Appendix H.2.4. The condition number is a statistic calculated in the multicollinearity analysis. The condition number, and its use as a multicollinearity diagnostic, is described in (Belsley, 1991a,b; Stewart, 1987). The condition number is equal to:

$$\kappa\{\mathbf{X}\} = \frac{\mu_{max}}{\mu_{min}}, \quad (5.25)$$

where μ_{max} is the largest singular value of \mathbf{X} and μ_{min} is the smallest singular value of \mathbf{X} . See Stewart (1973) for an introduction to matrix computations. In short, every matrix of predictor variables \mathbf{X} has an associated condition number. The condition number is an indicator of the sensitivity of the least-squares solution $\hat{\mathbf{b}}$ to perturbations in the \mathbf{X} and \mathbf{Y} data. Values close to 1 are indicative of a well-conditioned matrix. Large values are indicative of an ill-conditioned matrix, which can result in large errors in $\hat{\mathbf{b}}$. For the simulation of a sample design, a condition number distribution is formed from the condition numbers across all Monte Carlo trials.

The Survey Sampling Module can accept multiple values for certain survey sampling decisions (sample size, numbers of groups for stratification variables) so that more than one sample design can be specified and simulated. There are three analysis options that can be used to compare the results of simulations for different sample designs. The condition number distribution Option is one of these. Analysis Option 6 listed in Appendix H.2.4 merely collects the simulation-level statistics results calculated for the different sample designs. The simulation-level statistics were described earlier; they are those of the second type of analysis that involves results from all Monte Carlo trials of a survey simulation. Analysis Option 7 listed in Appendix H.2.4 is an analysis of variance in which the observed variance of a simulation-level statistic is partitioned into components attributable to the variation of sample size

and number of groups of each stratification variable. The analysis of variance results can be used to interpret the pattern of effect that each survey sample design factor has on the simulation-level statistic. An analysis of variance is performed for each simulation-level statistic specified by the user.

Analysis Option 8 listed in Appendix H.2.4 can be used to create various types of graphical output for each sample design. The different kinds of plots available are:

1. Mean of the regression coefficient estimates
2. Mean absolute deviation of the regression coefficient estimates (from the values in the population mean annoyance model)
3. Bias of the mean regression coefficient estimates
4. Standard deviation of the regression coefficient estimates
5. Mean absolute deviation of the regression coefficient standard deviation estimates
6. Mean of the regression coefficient estimates with standard deviation error bars
7. Sampling distributions of regression coefficient estimates
8. Predicted annoyance diagrams, *DNL* overlay
9. Predicted annoyance diagrams, no *DNL* overlay
10. Population annoyance model contour, *DNL* overlay
11. Population annoyance model contour, no *DNL* overlay

Plotting options 1 through 5 produce plots of the simulation-level statistics listed above. This plotting Option uses the simulation-level data collected through analysis Option 6 listed in Appendix H.2.4. Plotting Option 6 also uses the collected simulation-level data, specifically simulation-level statistics 1 and 4 to create plots of the mean regression coefficient estimates with standard deviation error bars. Plotting Option 7 creates a sampling distribution of the regression coefficient estimates for each parameter of each estimated annoyance model. Plotting options 8 and 9 create plots of the predicted annoyance from each estimated annoyance model, using the means of the regression coefficient estimates (simulation-level statistic 1), with or without overlaid *DNL* calculations, respectively. Plotting options 10 and 11 create

plots of the annoyance generated by the population mean annoyance model, with or without overlaid *DNL* calculations, respectively.

5.3 Example Survey Simulation: Design

In this section a description is given of survey simulations around Airports 1, 2, and 3 in terms of the survey design, survey sampling, and survey analysis decisions.

5.3.1 Survey Design Decisions

The same survey design decisions were used at each of the three airports. At each airport, the default aircraft operations scenario was used to specify the acoustical metrics. The population mean annoyance model was arbitrarily specified as:

$$A = 0.08 PNL T_{max,av,80} + 1.12 \log_{10}(N_{80}) - 3, \quad (5.26)$$

where $PNLT_{max,av,80}$ is the $PNLT_{max,av}$ of aircraft events with $PNLT_{max} > 80$ PNdB, and N_{80} is the number of daily aircraft events with $PNLT_{max} > 80$ PNdB. This model features $k = \frac{1.12}{0.08} = 14$. A podding approach was not used, so that the noise exposure assigned to each respondent was that of the grid point to which the respondent was assigned and not that of the population-weighted centroid of the census block group that each respondent resided within. As described in Section 5.2.4, a Gaussian distributed random variable was added to A to vary the annoyance values generated by the population mean annoyance model prior to sampling. At each airport, the variance of the Gaussian distribution was calculated by using Equation (5.4), which requires a signal-to-noise ratio (R_n^2) and a variance estimate of the population mean annoyance model annoyance values ($var_{est}(y)$). A R_n^2 of 0.2 was used at each airport to add variation to the annoyance calculations generated from Equation (5.26). Guski (1999) stated that a maximum of about one-third of the variance in person-level annoyance can be accounted for by variation in acoustical variables. From Table 4.2,

the maximum R^2 realized in the historical person-level survey data sets (excluding the simulated UKD-604 data sets) was 0.25. $R_n^2 = 0.2$ was used to simulate what might be expected in a real population. The $var_{est}(y)$ used at each airport was the variance of the nonzero annoyance values generated by the population mean annoyance model over the entire grid of points; for Airports 1, 2, and 3 the $var_{est}(y)$ was 1.84, 1.56, and 3.12, respectively. A discrete annoyance scale with five categories was used, as shown in Figure 5.2.

5.3.2 Survey Sampling Decisions

One sample design was specified at each airport. The sample design was specified by:

1. Sample size of 6,000,
2. Two stratification variables with 4 groups each: $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$, where N_{80} is the number of daily aircraft events with $PNLT_{max} > 80$ PNdB.

This sample design resulted in 16 strata formed by combinations of the $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ groups. Although the same general stratification settings were used for all airports, the boundaries of the groups for each stratification variable were separately adjusted at each airport in an attempt to divide the populations surrounding the airports into as many strata as possible with an “adequate” number of households in each stratum. An ideal number of households was defined to be $\frac{h}{s}$, where h is the number of households in the entire population surrounding the airport and s is the number of strata. An adequate number of samples was defined to be $\frac{n}{s}$, where n is the sample size. For a sample size of 6,000 and 16 strata as used in the simulations, an adequate number of households in each stratum was 375. The boundaries of the strata were adjusted with the goal of achieving at least 375 households in each stratum. However, this was not achievable at any of the airports. Adjustment of the boundaries continued until the minimum nonzero number of households in a stratum was concluded to be maximized. This was important for sampling, so that households experiencing a wide range of noise exposure conditions would be represented in each

sample drawn. As noted earlier, this improves the conditioning of the \mathbf{X} matrix in Equation (5.12). The minimum age of a sampled person was set to be 18 years old, and 100 Monte Carlo trials were performed for each survey simulation. The boundaries of the stratification variables and the number of households in each stratum are given for Airport 1 (Table 5.7), Airport 2 (Table 5.8), and Airport 3 (Table 5.9). The boundaries are also shown for Airport 1 in Figure 5.8a, Airport 2 in Figure 5.8b, and Airport 3 in Figure 5.8c. From an inspection of the plots in Figure 5.8, the entire collection of $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data at each airport do not fill the entire predictor variable space spanned by the ranges of the variables; that is, the data are not completely uncorrelated in the population. There is especially a lack of data in the combinations of a high average sound level with a low daily number-of-events. The correlation between the $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data for the entire population around the airport is 0.59, 0.71, and 0.49 for Airports 1, 2, and 3, respectively.

The sample size was settled upon through a prospective statistical power analysis. The concepts of statistical power calculations were described in Section 5.2.4. Prospective power analyses are important in planning the sample size of a survey (Cohen, 1992; Hudson, 2009; Maxwell, Kelley, and Rausch, 2008). Statistical power is the probability of rejecting the null hypothesis in favor of an alternative hypothesis when the alternative hypothesis is true. Sample size directly affects the standard deviations of statistics estimated from survey data. Confidence intervals are a function of these standard deviations; they indicate the accuracy and significance of parameter estimates.

The sample size was chosen based on the ability to distinguish a \hat{k} as significantly different from a hypothesized null value of 10 (the equal-energy value). Upper and lower alternative values of \hat{k} were calculated from different combinations of sample size and statistical power ($1 - \beta$), at one Type I error (α) rate. For each combination of sample size and power, the upper and lower alternative values of \hat{k} calculated were those which could be stated to be significantly different from the null value with a $\alpha \times 100\%$ chance of making a Type I error and a $\beta \times 100\%$ chance of making a

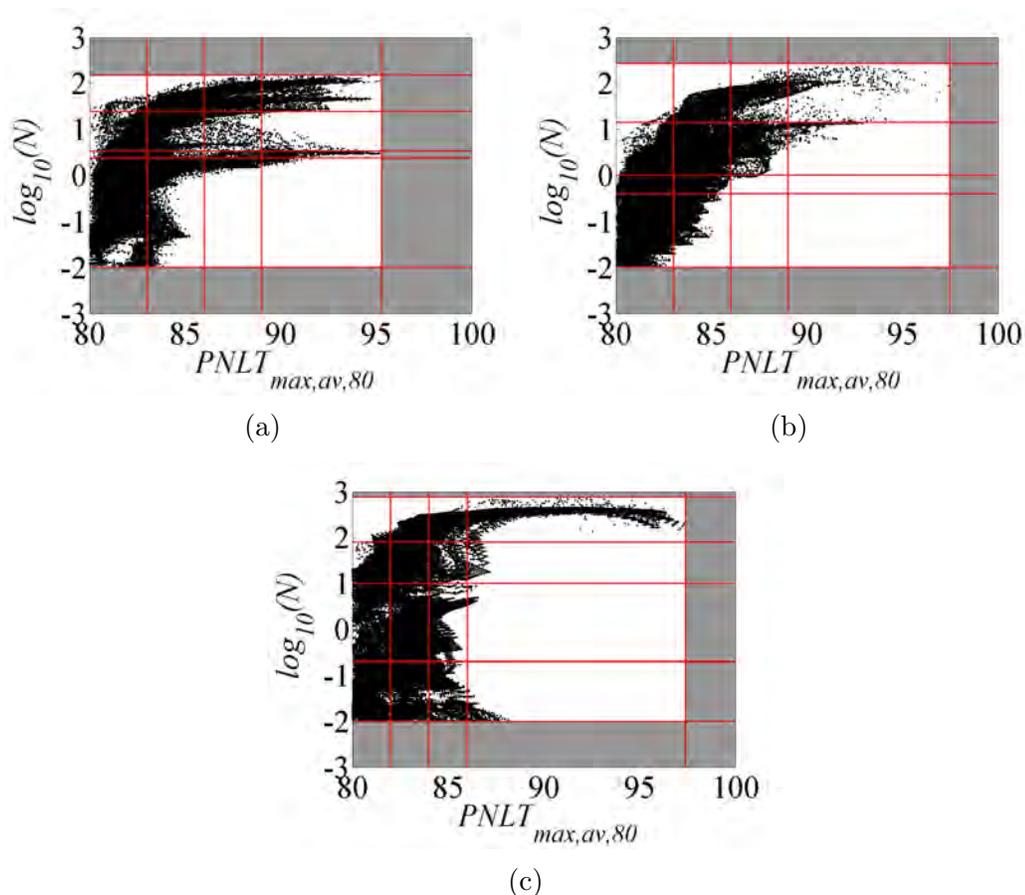


Figure 5.8. Strata boundaries for (a) Airport 1, (b) Airport 2, and (c) Airport 3. Combinations of $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ for households in the population are shown by the black dots. The strata boundaries are shown by the red lines. The correlation between the $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data for the entire population around the airport is 0.59, 0.71, and 0.49 for Airports 1, 2, and 3, respectively.

Type II error. The calculations required constructing hypothetical distributions of the null and alternative values of \hat{k} . The distributions were specified to be normal; the variance was a function of sample size so that for each sample size, a different variance was used to create distributions for the null and alternative values.

The Taylor series linearization expression of the variance of k was used to involve the effect of sample size on variance (Appendix B.2). Specifically, the prospective

power analysis used Equations (B.14), (B.26), (B.27), and (B.28) which are reprinted as Equations (5.27), (5.28), (5.29), and (5.30) below:

$$\hat{\sigma}^2(k) = \hat{\sigma}^2\left(\frac{b_N}{b_L}\right) = \frac{\hat{b}_N^2}{\hat{b}_L^2} \left(\frac{\hat{\sigma}^2(b_N)}{\hat{b}_N^2} + \frac{\hat{\sigma}^2(b_L)}{\hat{b}_L^2} - \frac{2\hat{\sigma}(b_N, b_L)}{\hat{b}_N\hat{b}_L} \right), \quad (5.27)$$

$$\hat{\sigma}^2(b_N) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_{\log_{10}(N)}^2 \sigma_L^2 - (\sigma_{L, \log_{10}(N)})^2} \right) (\sigma_L^2), \quad (5.28)$$

$$\hat{\sigma}^2(b_L) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_{\log_{10}(N)}^2 \sigma_L^2 - (\sigma_{L, \log_{10}(N)})^2} \right) (\sigma_{\log_{10}(N)}^2), \quad (5.29)$$

$$\hat{\sigma}(b_N, b_L) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_{\log_{10}(N)}^2 \sigma_L^2 - (\sigma_{L, \log_{10}(N)})^2} \right) (-\sigma_{L, \log_{10}(N)}), \quad (5.30)$$

where b_N represents the estimated coefficient of $\log_{10}(N)$, L represents a sound level metric, $\sigma_{\log_{10}(N)}^2$ is the variance of $\log_{10}(N)$, σ_L^2 is the variance of L , $\sigma_{L, \log_{10}(N)}$ is the covariance of L and $\log_{10}(N)$, and $\hat{\sigma}_e^2$ is the variance of the residuals and an estimate of the variance of the regression error terms.

From the equations, the variance is proportional to $\frac{1}{n}$, where n is the sample size. The 20,829 samples of data set Combination #7 (Table 4.11), where the sound level metric is $PNL_{max,av}$ and the annoyance scale is a 5-point discrete scale, were used to calculate $\sigma_{\log_{10}(N)}^2$, σ_L^2 , $\sigma_{L, \log_{10}(N)}$, σ_z^2 . Then various sample sizes were used in Equations (5.28), (5.29), (5.30), and (5.27) to calculate estimates of the variance of \hat{k} ; statistical power was varied over a range of values at each sample size and lower and upper alternative values of \hat{k} , which could be stated to be significantly different from the null value with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, were calculated for different combinations of sample size and statistical power. These are shown in Tables 5.10 and 5.11.

A sample size of 6,000, representative of a large survey, was chosen to be used in the survey simulations. With the assumption that collected survey data in simulations will have variances similar to those in the Combination #7 data set, then an estimate of \hat{k} lower than 7.64 or higher than 12.36 (values at the intersection of

$n = 6,000$ and $P = 0.80$) will be concluded as significantly different from 10 in at least 80% of the Monte Carlo trials. Since the chosen population mean annoyance model in this example features $k = 14$, then the population annoyance will be concluded significantly different from equal-energy in at least 90% of the Monte Carlo trials, if properties of the collected data are similar to that in the Combination #7 data set.

Two resamples of each 6,000-person sample collected. At each airport, from the full sample collected in each Monte Carlo trial of each simulation, a sample of 2,000 persons was drawn using the stratified sample design (strata assignments in Tables 5.7, 5.8, and 5.9) and a sample of 2,000 persons was drawn using simple random sampling without stratification. The resamples were performed in order to investigate if stratification could affect the simulation results, being that both were drawn from a stratified sample. The data of the resampled simulations were analyzed in the same way as that from the full-sample simulations. Thus, three simulations were realized at each of the three airports. Also, a survey simulation was performed around Airport 2 using simple random sampling without stratification. A sample of 6,000 persons was sampled from the population in each of 100 Monte Carlo trials. This simulation was performed to compare the cases of sampling the population around an airport with stratified sampling and simple random sampling. Among the three airports, the largest correlation between $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ in the population exists at Airport 2 (0.71). Thus, the effect of stratification will hypothetically be the most limited at Airport 2.

5.3.3 Survey Analysis Decisions

For each of the three simulations, two multiple linear regression annoyance models were estimated from the sample collected in each Monte Carlo trial:

$$A_1 = b_L PNL T_{max,av,80} + b_N \log_{10}(N_{80}) + b_{01}, \quad (5.31)$$

$$A_2 = b_D DNL + b_{02}. \quad (5.32)$$

Equation (5.31) is composed of the same predictor variables as the population mean annoyance model in Equation (5.26). In re-deriving the population mean annoyance model, each sample design can be ranked by its closeness to the population mean annoyance model used in the simulations. Equation (5.32) is a linear function of DNL . These regression models were also estimated using the resampled data. Of interest was to compare the accuracy of the predictions using Equations (5.31) and (5.32) with estimated coefficients and to identify the geographical areas where the models would predict different annoyance reactions to aircraft noise.

In addition to the default analysis settings carried out for each estimated annoyance model from each Monte Carlo trial, all of the additional trial analyses were performed: multicollinearity analysis, retrospective statistical power analysis of \hat{k} (for estimates of Equation (5.31) only), variance analysis of \hat{k} (for estimates of Equation (5.31) only), and the extra sum of squares calculations.

5.4 Example Survey Simulation: Results

This section contains the results from survey simulations around Airports 1, 2, and 3. The spatial locations of the households for a representative 6,000-person sample for Airports 1, 2, and 3 are shown in Figures 5.9a, 5.9b, and 5.9c, respectively.

5.4.1 Validation of Survey Simulation Program

As described in Sections 5.2.4 and 5.3.1, a Gaussian distribution of annoyance modifications was used to vary the annoyance values generated by the population mean annoyance model prior to sampling. At each airport, the variance of the Gaussian distribution was calculated by using Equation (5.4), which requires a signal-to-noise ratio (R_n^2) and a variance estimate of the population mean annoyance model annoy-

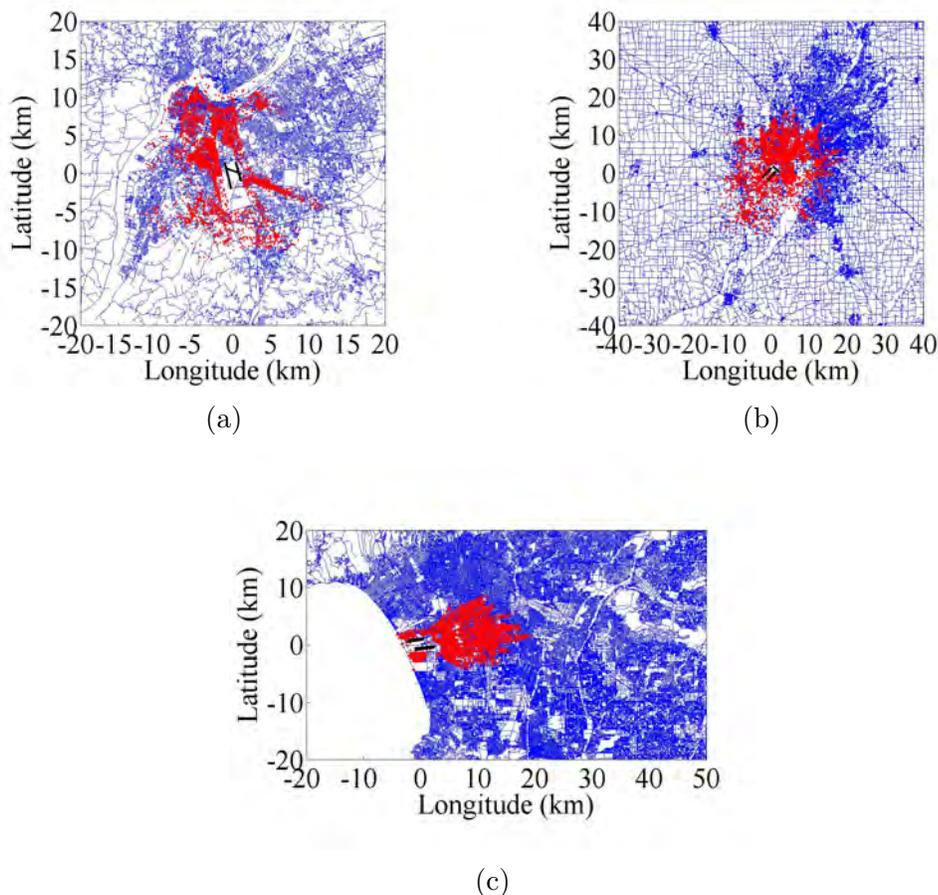


Figure 5.9. Household locations (shown in red) for representative samples of the populations around (a) Airport 1, (b) Airport 2, and (c) Airport 3. Roads (in blue) and runways (in black) are also shown.

ance values ($var_{est}(y)$). As a test of the Survey Sampling Module, the unmodified population mean annoyance model annoyance values were used in the data analyses at Airport 3. For each trial, the regression coefficient estimates for Equation (5.31) were identical to those in Equation (5.26) and the R^2 of the fit of the annoyance model to the data was 1.0. Thus, the sampling and data analyses functions of the Survey Sampling Module are correctly programmed since when there is no modification applied to the annoyance values, the Survey Sampling Module identifies the population mean annoyance model from the sample data. In Table 5.12 are shown the results from one Monte Carlo trial around Airport 3 for three different levels of R_n^2

and thus, annoyance modifications. As described in Section 5.3.1, $R_n^2 = 0.2$ is the baseline value and was used to vary the annoyance values generated from the population mean annoyance model at all three airports. Two other values were used to generate the results in Table 5.12. $R_n^2 = 0.8$ was realized by halving the Gaussian errors applied to the population mean annoyance model annoyance values. Also, the case of no errors was simulated, which is equivalent to the case of $R_n^2 = 1.0$.

5.4.2 Effects of Rescaling and Curtailing Annoyance Data

As described in Section 5.2.4, Equation (3.26) was used to rescale the modified annoyance values of the sample of respondents to range between 1 and the maximum annoyance scale number. In the survey simulations, the maximum annoyance scale number was 5. The use of Equation (3.26) to rescale the annoyance values yielded unbiased estimates for the regression coefficients, the results of which are shown in Tables 5.14 and 5.15. In addition, unbiased estimates for the regression coefficients of Equation (5.31) yields unbiased estimates of \hat{k} . As shown in Table 5.14, the 95% confidence interval for $E[\hat{k}]$ was in the vicinity of the true value ($k = 14$) in every simulation.

Two other methods of adjusting the annoyance which involve cutoffs at the extremities of the annoyance scale were investigated. Rather than purely scaling the annoyance (as with Equation (3.26)), these other methods apply cutoff limits to the annoyance scales. The first of these “cutoff” methods involves *DNL* calculations. All grid points with $DNL \leq 42$ dBA were set to the minimum annoyance value, 1. The choice of 42 dBA was made according to a polynomial approximation for the relationship between *DNL* and percent of people “highly annoyed” (*%HA*) (Janssen and Vos, 2011). The polynomial approximation is expressed by the following:

$$\begin{aligned} \%HA = & -1.395 \times 10^{-4}(DNL - 42)^3 + 4.081 \times 10^{-2}(DNL - 42)^2 \\ & + 0.342(DNL - 42). \end{aligned} \tag{5.33}$$

According to the expression, $\%HA$ is 0 at a DNL value of 42 dBA. This value was adopted at which population annoyance would be minimal. At the other end, all grid points with $DNL \geq 70$ dBA were set to the maximum annoyance value, 5. The grid points with $42 < DNL < 70$ were scaled to the range in between 1 and 5. In the second cutoff method, all annoyance values less than 1 were set equal to 1 and all annoyance values greater than 5 were set equal to 5; the annoyance is not rescaled with the second cutoff method, only limited at the extremities. When survey respondents are asked to fill in a number from 1 to 5 (Figure 5.2), responses are naturally limited. Multiple linear regression results for Equation (5.31) estimated using the annoyance data of Airport 2 are given in Table 5.13. Results are given for the three annoyance adjustment methods. All results are simulation-level and were estimated from the results of 100 Monte Carlo trials.

From Table 5.13, employing the DNL cutoff and scale cutoff methods resulted in biased regression coefficient estimates, which affected the estimates \hat{k} . The DNL cutoff method resulted in a downward bias in the estimation of \hat{k} , but did not appreciably affect the standard deviation of the 100 estimates of \hat{k} nor the estimated confidence interval for $E[\hat{k}]$. The scale cutoff method resulted in very unstable estimation of \hat{k} ; the standard deviation of the 100 estimates was very inflated and resulted in a very large confidence interval for $E[\hat{k}]$. The normal rescaling method, which does not employ a cutoff, resulted in unbiased and very stable estimation of \hat{k} .

5.4.3 Results of Survey Simulations

Multiple linear regression results for estimates of coefficients in Equation (5.31) and Equation (5.32) are given in Table 5.14 and Table 5.15, respectively. For each simulation, all statistics are simulation-level and were estimated from the results of 100 Monte Carlo trials. $\overline{\hat{b}_L}$ is the mean of the estimates of b_L , likewise for the other regression coefficient estimates. For each predictor variable, the standard deviation

Table 5.7. Strata for Airport 1. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.

Stratum	L_g	L_{min}	L_{max}	N_g	N_{min}	N_{max}	n
1	1	80.00	83.00	1	-2.00	0.38	44,234
2	1	80.00	83.00	2	0.38	0.52	2,633
3	1	80.00	83.00	3	0.52	1.39	12,714
4	1	80.00	83.00	4	1.39	2.18	335
5	2	83.00	86.00	1	-2.00	0.38	17,429
6	2	83.00	86.00	2	0.38	0.52	2,202
7	2	83.00	86.00	3	0.52	1.39	20,948
8	2	83.00	86.00	4	1.39	2.18	10,106
9	3	86.00	89.00	1	-2.00	0.38	2,790
10	3	86.00	89.00	2	0.38	0.52	469
11	3	86.00	89.00	3	0.52	1.39	2,669
12	3	86.00	89.00	4	1.39	2.18	6,294
13	4	89.00	95.29	1	-2.00	0.38	445
14	4	89.00	95.29	2	0.38	0.52	885
15	4	89.00	95.29	3	0.52	1.39	519
16	4	89.00	95.29	4	1.39	2.18	5,135

Table 5.8. Strata for Airport 2. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.

Stratum	L_g	L_{min}	L_{max}	N_g	N_{min}	N_{max}	n
1	1	80.00	83.00	1	-2.00	-0.40	76,738
2	1	80.00	83.00	2	-0.40	0.00	17,684
3	1	80.00	83.00	3	0.00	1.15	15,500
4	1	80.00	83.00	4	1.15	2.44	427
5	2	83.00	86.00	1	-2.00	-0.40	8,591
6	2	83.00	86.00	2	-0.40	0.00	8,553
7	2	83.00	86.00	3	0.00	1.15	26,088
8	2	83.00	86.00	4	1.15	2.44	6,267
9	3	86.00	89.00	1	-2.00	-0.40	0
10	3	86.00	89.00	2	-0.40	0.00	106
11	3	86.00	89.00	3	0.00	1.15	3,839
12	3	86.00	89.00	4	1.15	2.44	3,508
13	4	89.00	97.47	1	-2.00	-0.40	0
14	4	89.00	97.47	2	-0.40	0.00	0
15	4	89.00	97.47	3	0.00	1.15	511
16	4	89.00	97.47	4	1.15	2.44	887

Table 5.9. Strata for Airport 3. L_g refers to a group of $PNLT_{max,av,80}$ formed from a range, L_{min} is the minimum value of $PNLT_{max,av,80}$ for the range, and L_{max} is the maximum value of $PNLT_{max,av,80}$ for the range. N_g refers to a group of $\log_{10}(N_{80})$ formed from a range, N_{min} is the minimum value of $\log_{10}(N_{80})$ for the range, and N_{max} is the maximum value of $\log_{10}(N_{80})$ for the range. n is the number of households.

Stratum	L_g	L_{min}	L_{max}	N_g	N_{min}	N_{max}	n
1	1	80.00	82.00	1	-2.00	-0.70	35,686
2	1	80.00	82.00	2	-0.70	1.00	15,681
3	1	80.00	82.00	3	1.00	1.90	19,999
4	1	80.00	82.00	4	1.90	2.88	321
5	2	82.00	84.00	1	-2.00	-0.70	21,097
6	2	82.00	84.00	2	-0.70	1.00	35,567
7	2	82.00	84.00	3	1.00	1.90	21,518
8	2	82.00	84.00	4	1.90	2.88	14,199
9	3	84.00	86.00	1	-2.00	-0.70	6,373
10	3	84.00	86.00	2	-0.70	1.00	6,727
11	3	84.00	86.00	3	1.00	1.90	2,984
12	3	84.00	86.00	4	1.90	2.88	9,627
13	4	86.00	97.39	1	-2.00	-0.70	673
14	4	86.00	97.39	2	-0.70	1.00	259
15	4	86.00	97.39	3	1.00	1.90	321
16	4	86.00	97.39	4	1.90	2.88	16,989

Table 5.10. Lower alternative values of \hat{k} that can be stated to be significantly different from $\hat{k} = 10$ with a 95% of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error.

Sample size	$P = 0.50$	$P = 0.60$	$P = 0.70$	$P = 0.80$	$P = 0.90$	$P = 1.00$
500	4.29	3.55	2.76	1.84	0.56	-19.84
1,000	5.96	5.44	4.88	4.23	3.32	-11.1
1,500	6.70	6.28	5.82	5.29	4.55	-7.23
2,000	7.14	6.77	6.38	5.92	5.28	-4.92
2,500	7.44	7.11	6.76	6.35	5.78	-3.35
3,000	7.67	7.37	7.04	6.67	6.14	-2.19
3,500	7.84	7.56	7.26	6.91	6.43	-1.28
4,000	7.98	7.72	7.44	7.11	6.66	-0.55
4,500	8.09	7.85	7.58	7.28	6.85	0.05
5,000	8.19	7.96	7.71	7.42	7.01	0.56
5,500	8.28	8.05	7.82	7.54	7.15	1.00
6,000	8.35	8.14	7.91	7.64	7.27	1.38
6,500	8.41	8.21	7.99	7.73	7.38	1.72
7,000	8.47	8.27	8.06	7.82	7.47	2.02
7,500	8.52	8.33	8.13	7.89	7.56	2.29
8,000	8.57	8.38	8.19	7.96	7.64	2.54
8,500	8.61	8.43	8.24	8.02	7.71	2.76
9,000	8.65	8.48	8.29	8.07	7.77	2.96
9,500	8.69	8.52	8.34	8.12	7.83	3.15
10,000	8.72	8.55	8.38	8.17	7.89	3.32

Table 5.11. Upper alternative values of \hat{k} that can be stated to be significantly different from $\hat{k} = 10$ with a 95% of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error.

Sample size	$P = 0.50$	$P = 0.60$	$P = 0.70$	$P = 0.80$	$P = 0.90$	$P = 1.00$
500	15.71	16.45	17.24	18.16	19.44	124.95
1,000	14.04	14.56	15.12	15.77	16.68	91.29
1,500	13.30	13.72	14.18	14.71	15.45	76.37
2,000	12.86	13.23	13.62	14.08	14.72	67.48
2,500	12.56	12.89	13.24	13.65	14.22	61.41
3,000	12.33	12.63	12.96	13.33	13.86	56.93
3,500	12.16	12.44	12.74	13.09	13.57	53.45
4,000	12.02	12.28	12.56	12.89	13.34	50.65
4,500	11.91	12.15	12.42	12.72	13.15	48.32
5,000	11.81	12.04	12.29	12.58	12.99	46.35
5,500	11.72	11.95	12.18	12.46	12.85	44.66
6,000	11.65	11.86	12.09	12.36	12.73	43.19
6,500	11.59	11.79	12.01	12.27	12.62	41.89
7,000	11.53	11.73	11.94	12.18	12.53	40.73
7,500	11.48	11.67	11.87	12.11	12.44	39.68
8,000	11.43	11.62	11.81	12.04	12.36	38.74
8,500	11.39	11.57	11.76	11.98	12.29	37.88
9,000	11.35	11.52	11.71	11.93	12.23	37.1
9,500	11.31	11.48	11.66	11.88	12.17	36.38
10,000	11.28	11.45	11.62	11.83	12.11	35.71

Table 5.12. Multiple linear regression results for one Monte Carlo trial around Airport 3 for three different levels of R_n^2 (0.2, 0.8, and 1.0) and thus, annoyance modifications. $\hat{\sigma}_{\hat{k},c}$ is the jackknife standard deviation estimate of \hat{k} . The 95% CI (confidence interval) is for k . The standard deviation estimates of the regression coefficient estimates are given in parentheses. ¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.

R_n^2	\hat{b}_L	\hat{b}_N	\hat{b}_0	\hat{k}	$\hat{\sigma}_{\hat{k},c}$	95% CI for k	R^2
$R_n^2 = 0.2$	0.022 (0.003)	0.250 (0.004)	1.052 (0.216)	11.56	1.44	8.74 to 14.37	0.37
$R_n^2 = 0.8$	0.031 (0.002)	0.415 (0.004)	0.174 ¹ (0.185)	13.50	0.89	11.75 to 15.26	0.69
$R_n^2 = 1.0$	0.080 (0.000)	1.120 (0.000)	-3.00 (0.000)	14.00	0.00	14.00 to 14.00	1.0

Table 5.13. Multiple linear regression results for Equation (5.31) estimated using the three methods of adjusting annoyance; the data of Airport 2 were used. \widehat{b}_L is the mean of the estimates of b_L , likewise for the other regression coefficient estimates; the standard deviations are given in parentheses. The 95% CI (confidence interval) is for the expected value of $\widehat{k} \left(E \left[\widehat{k} \right] \right)$ and is calculated from \widehat{k} and $\widehat{\sigma}_k$. $\overline{R^2}$ is the mean of the R^2 's from the multiple linear regressions; the interquartile range is given in parentheses. For each trial, the predictor variable data are the same for all adjustment methods; the only variable that changes is the annoyance variable. "Normal" is the adjustment applied with Equation (3.26), "DNL cutoff" is the adjustment applied with the use of *DNL* data, and "Scale cutoff" is the adjustment applied by limiting the annoyance values at the extremities.

Adjustment Method	\widehat{b}_L	\widehat{b}_N	\widehat{b}_{01}	\widehat{k}	$\widehat{\sigma}_k$	95% CI for $E \left[\widehat{k} \right]$	$\overline{R^2}$
Normal	0.027 (0.002)	0.365 (0.017)	0.488 (0.213)	13.44	1.14	13.22 to 13.67	0.43 (0.43 to 0.44)
<i>DNL</i> cutoff	0.027 (0.023)	0.866 (0.178)	1.189 (2.363)	63.98	71.61	49.77 to 78.19	0.45 (0.43 to 0.50)
Scale cutoff	0.024 (0.003)	0.805 (0.014)	1.495 (0.264)	33.54	4.74	32.60 to 34.48	0.50 (0.49 to 0.50)

of the 100 regression coefficient estimates is given below the mean regression coefficient estimate in parentheses. The statistical significance of the i^{th} mean regression coefficient estimate \bar{b}_i was concluded through an inspection of the estimated confidence interval for the expected value of \hat{b}_i ($E[\hat{b}_i]$). The estimated confidence interval is for $E[\hat{b}_i]$ and not b_i because the 100 estimates of \hat{b}_i might be biased. The estimated confidence interval for $E[\hat{b}_i]$ is expressed as:

$$\bar{b}_i - t \left(1 - \frac{\alpha}{2}; n - 1 \right) \frac{\hat{\sigma}_{\hat{b}_i}}{\sqrt{n}} \leq E[\hat{b}_i] \leq \bar{b}_i + t \left(1 - \frac{\alpha}{2}; n - 1 \right) \frac{\hat{\sigma}_{\hat{b}_i}}{\sqrt{n}}, \quad (5.34)$$

where $t \left(1 - \frac{\alpha}{2}; n - 1 \right)$ denotes the $(\alpha/2) \times 100^{\text{th}}$ percentile of the Student's t distribution with $n - 1$ degrees of freedom, n is the number of Monte Carlo trials, α is the Type I error rate (here, $\alpha = 0.05$), and $\hat{\sigma}_{\hat{b}_i}$ is the standard deviation of the 100 estimates of b_i . Analyses of pertaining to the decibel equivalent number effect ($\hat{k} = \hat{b}_N/\hat{b}_L$) were performed only for estimates of coefficients in Equation (5.31), because Equation (5.32) does not feature separate sound level and number-of-events predictor variables, though it can be interpreted as an equal-energy ($k = 10$) model (Chapter 2, Section 2.1). \bar{k} is the mean of the \hat{k} from the 100 Monte Carlo trials. $\hat{\sigma}_{\hat{k}}$ is the standard deviation of the 100 estimates of \hat{k} and is an estimate of the true standard deviation of k . $\hat{\sigma}_{\hat{k}}$ was used to estimate a 95% confidence interval for the expected value of \hat{k} ($E[\hat{k}]$):

$$\bar{k} - t \left(1 - \frac{\alpha}{2}; n - 1 \right) \frac{\hat{\sigma}_{\hat{k}}}{\sqrt{n}} \leq E[\hat{k}] \leq \bar{k} + t \left(1 - \frac{\alpha}{2}; n - 1 \right) \frac{\hat{\sigma}_{\hat{k}}}{\sqrt{n}}. \quad (5.35)$$

$\overline{R^2}$ is the mean of the R^2 's from the multiple linear regressions. The interquartile range for $\overline{R^2}$ is given in parentheses. Also given in Table 5.14 is the mean correlation ($\overline{\rho_{L, \log_{10}(N)}}$) between L ($PNLT_{max, av, 80}$) and $\log_{10}(N)$ ($\log_{10}(N_{80})$), with the interquartile range given in parentheses. In Tables 5.14 and 5.15, "Full S" denotes the 6000-person sample resulting from stratified sampling of the population, "Full R" denotes the 6000-person sample resulting from simple random sampling of the

population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling.

Decibel Equivalent Number Effect (\hat{k}) and the Standard Deviation of \hat{k} ($\hat{\sigma}_{\hat{k}}$)

The statements in this section follow from an examination of the results in Table 5.14. The $\bar{\hat{k}}$ for all simulations are close to the true value ($k = 14$) and because the $\hat{\sigma}_{\hat{k}}$ are relatively small, the estimated confidence intervals for $E[\hat{k}]$ exclude the equal-energy value ($k = 10$) so that $E[\hat{k}]$ is concluded to be significantly different from 10 in all of the simulations. The magnitude of $\hat{\sigma}_{\hat{k}}$ decreases in progressing from the Full S simulation to the RSD simulation to the RR simulation. Larger $\hat{\sigma}_{\hat{k}}$ are expected for the RR simulation; the use of a stratified sample design (RSD simulation) ensures that predictor variable data are sampled from all strata, which has the potential to realize lower correlation between $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ over a random sample design (RR simulation). However, because the Full S sample was drawn with the stratified sample design and both of the RSD and RR samples were drawn from the Full S sample, the effects of stratifying a second time are not dramatic. A comparison of the Full S and Full R results of Airport 2 shows that the estimation results from using simple random sampling are much less stable than those from using stratified sampling; although the $\bar{\hat{k}}$ of the Full R sample of Airport 2 was close to the true value, the range of \hat{k} realized over the 100 Monte Carlo trials (46.7) is much greater than that in the Full S simulation (5.4). The correlation between $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ was also affected by stratification. The maximum correlation observed in the SS simulation (0.66) was smaller than the minimum correlation in the RS simulation (0.69); in every trial, stratification realized a lower correlation between the predictor variables. From Chapter 4, Section 4.1.3, variance estimates of \hat{k} can

Table 5.14. Multiple linear regression results for estimates of Equation (5.31). \bar{b}_L is the mean of the estimates of b_L , likewise for the other regression coefficient estimates; the standard deviations are given in parentheses. \bar{k} is the mean of the 100 estimates of \hat{k} ; the range of \hat{k} is given in parentheses. The 95% CI (confidence interval) is for the expected value of \hat{k} ($E[\hat{k}]$) and is calculated from \bar{k} and $\hat{\sigma}_{\hat{k}}$. $\overline{\rho_{L, \log_{10}(N)}}$ is the mean correlation between L ($PNLT_{max, av, 80}$) and $\log_{10}(N)$ ($\log_{10}(N_{80})$); the interquartile range is given in parentheses. $\overline{R^2}$ is the mean of the R^2 's from the multiple linear regressions; the interquartile range is given in parentheses.

Simulation	\bar{b}_L	\bar{b}_N	\bar{b}_{01}	\bar{k}	$\hat{\sigma}_{\hat{k}}$	95% CI for $E[\hat{k}]$	$\overline{\rho_{L, \log_{10}(N)}}$	$\overline{R^2}$
Airport 1, Full S	0.027 (0.002)	0.38 (0.02)	0.46 (0.22)	14.21 (12.25 to 17.52)	0.97	14.02 to 14.41	0.20 (0.20 to 0.21)	0.27 (0.26 to 0.28)
Airport 1, RSD	0.027 (0.004)	0.38 (0.02)	0.48 ¹ (0.32)	14.47 (10.21 to 18.84)	1.94	14.09 to 14.86	0.19 (0.18 to 0.20)	0.27 (0.26 to 0.28)
Airport 1, RR	0.027 (0.004)	0.38 (0.02)	0.46 ¹ (0.37)	14.43 (10.35 to 20.02)	2.21	13.99 to 14.87	0.20 (0.19 to 0.21)	0.27 (0.26 to 0.29)
Airport 2, Full S	0.027 (0.002)	0.37 (0.02)	0.49 (0.21)	13.44 (11.31 to 16.75)	1.14	13.22 to 13.67	0.65 (0.65 to 0.65)	0.43 (0.43 to 0.44)
Airport 2, Full R	0.018 (0.006)	0.25 (0.02)	1.55 (0.48)	15.62 (7.74 to 54.44)	7.51	14.13 to 17.11	0.71 (0.71 to 0.71)	0.19 (0.18 to 0.20)
Airport 2, RSD	0.027 (0.004)	0.37 (0.02)	0.53 ¹ (0.32)	13.96 (10.28 to 20.99)	2.27	13.51 to 14.41	0.62 (0.62 to 0.62)	0.42 (0.41 to 0.43)
Airport 2, RR	0.028 (0.005)	0.37 (0.02)	0.47 ¹ (0.39)	13.70 (8.56 to 21.98)	2.59	13.19 to 14.21	0.65 (0.65 to 0.66)	0.44 (0.42 to 0.45)

Table 5.14. Continued

Simulation	\widehat{b}_L	\widehat{b}_N	\widehat{b}_{01}	\widehat{k}	$\widehat{\sigma}_k$	95% CI for $E[\widehat{k}]$	$\overline{\rho_{L, \log_{10}(N)}}$	$\overline{R^2}$
Airport 3, Full S	0.021 (0.002)	0.29 (0.02)	1.02 (0.22)	13.90 (10.89 to 19.19)	1.51	13.60 to 14.20	0.18 (0.18 to 0.19)	0.42 (0.41 to 0.43)
Airport 3, RSD	0.021 (0.004)	0.29 (0.02)	1.05 (0.35)	14.52 (9.00 to 26.32)	3.14	13.90 to 15.14	0.18 (0.17 to 0.18)	0.41 (0.40 to 0.42)
Airport 3, RR	0.021 (0.004)	0.29 (0.02)	1.04 (0.37)	14.57 (9.31 to 26.63)	3.39	13.90 to 15.25	0.18 (0.17 to 0.20)	0.42 (0.41 to 0.43)

¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.

Table 5.15. Multiple linear regression results for estimates of Equation (5.32). $\overline{\hat{b}_D}$ is the mean of the estimates of b_D , likewise for the other regression coefficient estimate; the standard deviations are given in parentheses. $\overline{R^2}$ is the mean of the R^2 's from the multiple linear regressions; the interquartile range is given in parentheses. ¹The predictor variable is not significant to the prediction of annoyance at $\alpha = 0.05$.

Simulation	$\overline{\hat{b}_D}$	$\overline{\hat{b}_{02}}$	$\overline{R^2}$
Airport 1, Full S	0.051 (0.003)	0.38 (0.17)	0.19 (0.18 to 0.20)
Airport 1, RSD	0.050 (0.003)	0.39 (0.20)	0.19 (0.18 to 0.20)
Airport 1, RR	0.051 (0.003)	0.38 ¹ (0.20)	0.19 (0.18 to 0.20)
Airport 2, Full S	0.057 (0.003)	0.04 ¹ (0.18)	0.36 (0.35 to 0.37)
Airport 2, Full R	0.052 (0.003)	0.50 (0.19)	0.14 (0.14 to 0.15)
Airport 2, RSD	0.057 (0.003)	0.06 ¹ (0.20)	0.35 (0.34 to 0.37)
Airport 2, RR	0.057 (0.003)	0.02 ¹ (0.19)	0.36 (0.35 to 0.37)
Airport 3, Full S	0.058 (0.003)	-0.22 ¹ (0.19)	0.37 (0.36 to 0.38)
Airport 3, RSD	0.058 (0.003)	-0.22 ¹ (0.19)	0.36 (0.35 to 0.37)
Airport 3, RR	0.058 (0.003)	-0.22 ¹ (0.19)	0.37 (0.35 to 0.38)

be lowered by lowering the correlation between variables and increasing the ranges of each predictor variable. As described in Appendix C, a consequence of correlation between predictor variables or restricted ranges in the variables is that estimated regression coefficients tend to have large sampling variability. Stratification reduces the correlation and maintains the full ranges of the predictor variables in the samples. Also, the R^2 of the Full S estimated regression models shown in Tables 5.14 and 5.15 are larger than those of the Full R simulation; using stratification increased the variance in the annoyance accounted for by $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$.

Average Semipartial R^2 ($\overline{R_{sp}^2}$)

$\overline{R_{sp}^2}$ were estimated for every predictor variable of each of the two estimated multiple linear regression models in every Monte Carlo trial of a survey simulation. For each predictor variable, the $\overline{R_{sp}^2}$ results were averaged over the 100 Monte Carlo trials and the interquartile range was noted. The average $\overline{R_{sp}^2}$ (average of averages) results are given in this section. This average $\overline{R_{sp}^2}$ will be distinguished from $\overline{R_{sp}^2}$ through the use of two averaging bars; the average $\overline{R_{sp}^2}$ will be designated by $\overline{\overline{R_{sp}^2}}$. In Table 5.16 are shown the $\overline{\overline{R_{sp}^2}}$ for $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ of Equation (5.31) along with interquartile ranges. Because Equation (5.32) contains only one predictor variable in addition to an intercept, the $\overline{\overline{R_{sp}^2}}$ results are not shown, since they are equal to the R^2 shown in Table 5.15.

Table 5.16. Averages of average semipartial R^2 values ($\overline{\overline{R_{sp}^2}}$) and interquartile range (IR) of $\overline{\overline{R_{sp}^2}}$ for $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ of Equation (5.31). “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling.

Simulation	$PNLT_{max,av,80}$		$\log_{10}(N_{80})$	
	$\overline{\overline{R_{sp}^2}}$	IR	$\overline{\overline{R_{sp}^2}}$	IR
Airport 1, Full S	0.04	0.04 to 0.04	0.23	0.22 to 0.24
Airport 1, RSD	0.04	0.04 to 0.04	0.23	0.22 to 0.24
Airport 1, RR	0.04	0.04 to 0.05	0.23	0.22 to 0.24
Airport 2, Full S	0.13	0.13 to 0.13	0.30	0.30 to 0.31
Airport 2, Full R	0.06	0.05 to 0.06	0.13	0.13 to 0.14
Airport 2, RSD	0.12	0.11 to 0.13	0.30	0.29 to 0.31
Airport 2, RR	0.13	0.12 to 0.14	0.30	0.29 to 0.31
Airport 3, Full S	0.02	0.02 to 0.02	0.40	0.39 to 0.41
Airport 3, RSD	0.02	0.02 to 0.02	0.39	0.38 to 0.40
Airport 3, RR	0.02	0.02 to 0.02	0.40	0.39 to 0.41

In all simulations, $\log_{10}(N_{80})$ was quantified as more generally dominant than $PNLT_{max,av,80}$. As shown in Table 5.16, the differences between $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ are not slight; $\log_{10}(N_{80})$ is at least twice as important as $PNLT_{max,av,80}$ by the $\overline{R_{sp}^2}$ criterion. These results are a natural outcome of the population mean annoyance model (Equation (5.26)) used in the simulation. The mean population annoyance was generated from a sum of scaled $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data, the scaling factors being the regression coefficients in Equation (5.26). The scaled variable which varied more had the greater effect in determining the variation in the generated annoyance. The dominance of $\log_{10}(N_{80})$ over $PNLT_{max,av,80}$ proceeds from the fact that the scaled $\log_{10}(N_{80})$ data affected the variance of the generated population mean annoyance more than the scaled $PNLT_{max,av,80}$ data. The ranges of each scaled predictor variable are shown in Table 5.17; the scaled $\log_{10}(N_{80})$ data varied over a greater range than the $PNLT_{max,av,80}$ data for every simulation.

Statistical Power Analysis of \hat{k} : Prospective Versus Retrospective

Analyses of the statistical power of \hat{k} were performed only for estimates of Equation (5.31), since Equation (5.32) does not feature separate sound level and number-of-events predictor variables. As described in Section 5.3.2, a sample size of 6,000 for the Full S simulation was decided upon with the guidance of a prospective power analysis. Properties of the $PNL_{max,av}$ and $\log_{10}(N)$ data of the Combination #7 data set described in Table 4.11 in Chapter 4 were used to calculate the alternative values of \hat{k} that could be stated as significantly different from the equal-energy value ($k = 10$) with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, where α is the Type I error rate and P is the statistical power. With the assumption that the simulated survey data would have variances similar to those in the Combination #7 data set, an estimate of \hat{k} lower than 7.64 or higher than 12.36 would be concluded as significantly different from 10 in at least 80% of the Monte Carlo trials. Also, an estimate of \hat{k} equal to 14

Table 5.17. Ranges of the scaled $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data in a typical sample. From Equation (5.26), 0.08 is the scaling applied to $PNLT_{max,av,80}$ and 1.12 is the scaling applied to $\log_{10}(N_{80})$. “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling.

Simulation	Range of $PNLT_{max,av,80}$	Range of $0.08PNLT_{max,av,80}$	Range of N_{80}	Range of $\log_{10}(N_{80})$	Range of $1.12\log_{10}(N_{80})$
Airport 1, Full S	15.20	1.22	133.20	4.07	4.56
Airport 1, RSD	15.19	1.22	133.19	3.78	4.23
Airport 1, RR	15.18	1.21	130.30	3.95	4.42
Airport 2, Full S	16.74	1.34	228.30	4.33	4.85
Airport 2, Full R	12.72	1.02	224.10	4.34	4.86
Airport 2, RSD	16.74	1.34	206.63	4.01	4.49
Airport 2, RR	16.74	1.34	223.30	4.32	4.84
Airport 3, Full S	16.98	1.36	618.01	4.79	5.37
Airport 3, RSD	15.94	1.28	565.23	4.74	5.31
Airport 3, RR	16.18	1.29	445.19	4.64	5.20

(the population mean annoyance model value) would be concluded significantly different from equal-energy in at least 90% of the Monte Carlo trials.

The collected survey data of each Monte Carlo trial was used in a retrospective power analysis to calculate the upper alternative value of \hat{k} distinguishable from the equal-energy value ($k = 10$) with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, given the calculated jackknife standard deviation of $\hat{k}(\hat{\sigma}_{\hat{k}})$. An upper alternative value of \hat{k} was calculated for each Monte Carlo trial so that for each simulation, 100 upper alternative values of \hat{k} were realized. If the estimate of $\hat{\sigma}_{\hat{k}}$ in a Monte Carlo trial is larger than the value based on the data in the Combination #7 data set, than the upper alternative value of \hat{k} will be greater than the planned value of 12.36. That means that the statistical power at 12.36 will be less than 0.80. In the prospective power analysis and for a sample size of 6,000, the Taylor method calculated a planned $\hat{\sigma}_{\hat{k}}$ for \hat{k} of 0.84 from the Combination #7 data so that the hypothetical null and alternative distributions of \hat{k} were constructed with $\hat{\sigma}_{\hat{k}} = 0.84$. In the retrospective power analysis, the estimated $\hat{\sigma}_{\hat{k}}$ from the data were used to create the hypothetical null and alternative distributions. Both jackknife and Taylor estimates of $\hat{\sigma}_{\hat{k}}$ were used. In Tables 5.18 and 5.19 are shown the results of the retrospective power analysis using the jackknife and Taylor estimates of $\hat{\sigma}_{\hat{k}}$, respectively, over the 100 Monte Carlo trials. In both tables, “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling.

From an examination of the rightmost columns of Tables 5.18 and 5.19, the upper alternative value of \hat{k} at $P = 0.80$ was always greater than \hat{k} from the prospective power analysis ($\hat{k} = 12.36$), no matter which $\hat{\sigma}_{\hat{k}}$ estimates were used to create the hypothetical null and alternative distributions for \hat{k} . This means that the statistical power of distinguishing \hat{k} from the equal-energy value ($k = 10$) for all trials of all

Table 5.18. Results of the retrospective power analysis estimated using the jackknife estimate of $\hat{\sigma}_{\hat{k}}$ in each Monte Carlo trial. “FT S.D. 10” is the fraction of Monte Carlo trials in which \hat{k} was concluded with 95% confidence to be significantly different from the equal-energy value ($k = 10$). $\bar{\hat{k}}_{0.80}$ is the mean upper alternative value of \hat{k} distinguishable from the equal-energy value with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, where $P = 0.80$; the interquartile range is also given in parentheses. $\bar{\sigma}_{\hat{k}}$ is the mean jackknife standard deviation of \hat{k} over the 100 Monte Carlo trials. “FT $\bar{\hat{k}}_{0.80} \geq 12.36$ ” is the fraction of Monte Carlo trials in which the upper alternative value calculated was greater than the value from the prospective power calculations ($\hat{k} = 12.36$).

Simulation	FT S.D. 10	$\bar{\hat{k}}_{0.80}$	$\bar{\sigma}_{\hat{k}}$	FT $\bar{\hat{k}}_{0.80} \geq 12.36$
Airport 1, Full S	1.00	13.33 (13.03 to 13.57)	1.19	1.00
Airport 1, RSD	0.59	15.98 (14.92 to 16.72)	2.13	1.00
Airport 1, RR	0.54	16.01 (14.87 to 16.81)	2.15	1.00
Airport 2, Full S	0.71	14.27 (13.74 to 14.71)	1.52	1.00
Airport 2, Full R	0.00	29.17 (18.31 to 30.10)	6.84	1.00
Airport 2, RSD	0.05	17.75 (16.14 to 18.89)	2.77	1.00
Airport 2, RR	0.01	17.74 (15.85 to 18.96)	2.76	1.00
Airport 3, Full S	0.60	15.27 (14.42 to 15.94)	1.88	1.00
Airport 3, RSD	0.00	20.24 (17.17 to 21.80)	3.65	1.00
Airport 3, RR	0.00	20.39 (16.88 to 22.33)	3.71	1.00

simulations was less than that was planned in the prospective power analysis. This is not surprising, because the value of $\hat{k} = 12.36$ was estimated using the Combination #7 predictor variable data, in which the predictor variables each span a large range of values and are less correlated than the sample data collected in the survey simulations. As described in Chapter 4, Section 4.1.3, either increasing the range of the predictor variables or reducing the correlation between them reduces the Taylor estimate of $\hat{\sigma}_{\hat{k}}$. The prospective value of \hat{k} for sample size planning was calculated with the Taylor

Table 5.19. Results of the retrospective power analysis estimated using the Taylor estimate of $\hat{\sigma}_{\hat{k}}$ in each Monte Carlo trial. “FT S.D. 10” is the fraction of Monte Carlo trials in which \hat{k} was concluded with 95% confidence to be significantly different from the equal-energy value ($k = 10$). $\bar{\hat{k}}_{0.80}$ is the mean upper alternative value of \hat{k} distinguishable from the equal-energy value with a $\alpha \times 100\%$ chance of making a Type I error and a $(1 - P) \times 100\%$ chance of making a Type II error, where $P = 0.80$; the interquartile range is also given in parentheses. $\bar{\sigma}_{\hat{k}}$ is the mean jackknife standard deviation of \hat{k} over the 100 Monte Carlo trials. “FT $\bar{\hat{k}}_{0.80} \geq 12.36$ ” is the fraction of Monte Carlo trials in which the upper alternative value calculated was greater than the value from the prospective power calculations ($\hat{k} = 12.36$).

Simulation	FT S.D. 10	$\bar{\hat{k}}_{0.80}$	$\bar{\sigma}_{\hat{k}}$	FT $\bar{\hat{k}}_{0.80} \geq 12.36$
Airport 1, Full S	1.00	13.31 (13.00 to 13.53)	1.18	1.00
Airport 1, RSD	0.61	15.96 (14.91 to 16.76)	2.13	1.00
Airport 1, RR	0.55	15.99 (14.79 to 16.68)	2.14	1.00
Airport 2, Full S	0.70	14.29 (13.75 to 14.76)	1.53	1.00
Airport 2, Full R	0.00	29.17 (18.31 to 30.10)	6.84	1.00
Airport 2, RSD	0.03	17.81 (16.22 to 18.98)	2.79	1.00
Airport 2, RR	0.01	17.79 (15.89 to 18.94)	2.78	1.00
Airport 3, Full S	0.58	15.30 (14.46 to 15.96)	1.89	1.00
Airport 3, RSD	0.00	20.31 (17.17 to 21.99)	3.68	1.00
Airport 3, RR	0.00	20.44 (17.03 to 22.36)	3.73	1.00

method. A comparison of the results in the second columns of Tables 5.18 and 5.19 shows that using either the jackknife or the Taylor estimates of $\hat{\sigma}_{\hat{k}}$ resulted in \hat{k} being concluded as significantly different from the equal-energy value ($k = 10$) in over half of the trials of all of the Full S simulations. Also, a comparison of the third columns shows that the jackknife and Taylor estimates of $\hat{\sigma}_{\hat{k}}$ were similarly-sized. Thus, the fraction of Monte Carlo trials in which \hat{k} was concluded with 95% confidence to be significantly different from the equal-energy value was not sensitive to the choice of

the jackknife or Taylor estimate of $\hat{\sigma}_{\hat{k}}$. A comparison of the first-column results of the Full S simulation to the results of both the RSD and RR simulations of each airport in either table shows that using fewer samples resulted in less power to distinguish \hat{k} from the equal-energy value so that \hat{k} was usually not distinguished from the equal-energy value with 95% confidence in a majority of trials. In the Full S simulation of Airport 2, \hat{k} was concluded to be significantly different from the equal-energy value ($k = 10$) with 95% confidence in 71% of the Monte Carlo trials when the jackknife estimates of $\hat{\sigma}_{\hat{k}}$ were used to construct 95% confidence intervals for k (Table 5.18). When the Taylor estimates of $\hat{\sigma}_{\hat{k}}$ were used, \hat{k} was concluded to be significantly different from the equal-energy value with 95% confidence in 70% of the Monte Carlo trials (Table 5.19). In contrast, there were zero trials of the Full R simulation at Airport 2 in which the confidence interval of \hat{k} was narrow enough to discern \hat{k} from the equal-energy value, no matter whether jackknife or Taylor estimates of $\hat{\sigma}_{\hat{k}}$ were used. This indicates that random sampling of the population resulted in larger $\hat{\sigma}_{\hat{k}}$, and thus, wider confidence intervals for k . A comparison of the Full R and Full S simulation results in the third columns of Tables 5.18 and 5.19 shows that the average of the estimated $\hat{\sigma}_{\hat{k}}$ were much larger for the Full R simulation.

5.4.4 Annoyance Model Comparisons

The parameters of two annoyance models, Equations (5.31) and (5.32), were estimated from the sample data in each of the 100 Monte Carlo trials of the survey simulations. In Tables 5.14 and 5.15 were given the mean regression coefficient estimates averaged over the 100 Monte Carlo trials. In this section, predictions of annoyance from the estimated mean annoyance models formed from the mean regression coefficient estimates are compared to the annoyance generated by the population mean annoyance model, without the random error applied. The mean annoyance model estimate of Equation (5.31) is referred to as the “N&L model”, where N&L stands for “number

and level". The mean annoyance model estimate of Equation (5.32) is referred to as the *DNL* model. Six annoyance model comparisons were made at each airport:

1. Full S, N&L model and population mean annoyance model
2. Full S, *DNL* model and population mean annoyance model
3. RSD, N&L model and population mean annoyance model
4. RSD, *DNL* model and population mean annoyance model
5. RR, N&L model and population mean annoyance model
6. RR, *DNL* model and population mean annoyance model

In addition, the N&L and *DNL* models of the Full R simulation of Airport 2 were compared to the population mean annoyance model. For each annoyance model comparison, the root mean square (rms) error in the prediction was calculated:

$$E = \sqrt{\frac{\sum_{i=1}^h (A_{e,i} - A_{p,i})^2}{h}}, \quad (5.36)$$

where A_e is the predicted annoyance from the estimated mean annoyance model, A_p is the annoyance generated from the population mean annoyance model before random error is added, and h is the number of households. In Table 5.20, E is given for the estimated mean annoyance models at each of the three airports.

The rms error (E) between the predicted household annoyances and the mean annoyance generated by the population mean annoyance model was near one annoyance category. A numerical five-point annoyance scale such was used in the survey simulations; the numerical values could be representative of the verbal ratings: Not at all (1), Slightly (2), Moderately (3), Very (4), Extremely (5). One annoyance category could be the difference between any two adjacent verbal ratings. From an examination of the E in Table 5.20, resampling the Full S samples did not appreciably change the estimated mean regression models or the predicted annoyance. The E for the Full R simulation of Airport 2 were slightly smaller than the E for the Full S, RSD, and RR simulations of Airport 2. Although random sampling of the population resulted in less stable model estimation, the mean estimation results were comparable to those

Table 5.20. Average difference (E) between the predicted household annoyance and the annoyance generated by the population mean annoyance model. Each cell represents the average difference in predicted household annoyance, calculated by using Equation (5.36). “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling. “N&L” and “ DNL ” denote the model predictions from the mean estimates of Equations (5.32) and (5.32), respectively.

	Full S N&L	Full S DNL	RSD N&L	RSD DNL	RR N&L	RR DNL	Full R N&L	Full R DNL
Airport 1	1.24	1.26	1.24	1.26	1.24	1.26	-	-
Airport 2	0.89	0.98	0.89	0.98	0.89	0.98	0.85	0.93
Airport 3	1.18	1.22	1.18	1.22	1.18	1.22	-	-

resulting from stratified sampling. From an examination of Table 5.14, the $\bar{\hat{k}}$ of the Full R sample of Airport 2 was close to the true value, but the trial estimates of \hat{k} varied over a large range. For each airport, the difference between the predicted annoyance and mean population annoyance was similar for both the N&L model and the DNL model. However, the annoyance predictions of the N&L model were less different from the mean population annoyance than the annoyance predictions of the DNL model at all three airports. This is not surprising considering that the N&L model is composed of the same predictor variables as the population mean annoyance model used in the simulation. Plots that show the difference in annoyance predictions between two annoyance models were also created. The plots show $(A_e - A_p)$, where A_e is the predicted annoyance from an estimated mean annoyance model and A_p is the annoyance generated from the population mean annoyance model without the addition of the random variation. In Figures 5.10, 5.11, and 5.12 are shown the annoyance differences for Airports 1, 2, and 3, respectively. An inspection of the plots reveals that the estimated N&L and DNL models underpredicted annoyance close to the runways and overpredicted annoyance near the outer boundary of noise exposure.

The N&L model predicted greater annoyance than the *DNL* model at areas close to the runways, but the differences are slight. The plots show that the predictions from the mean annoyance models estimated from the resample data did not appreciably differ from the predictions from the mean annoyance models estimated from the Full S data, which is also given by the results in Table 5.20.

5.5 Summary

The importance of sample design to regression model estimation was described. A methodology of simulating noise surveys with the Survey Simulation Program (a collection of three computer programs) was described. Two of the computer programs (the Noise Metric Module and the Synthetic Population Generator) generate acoustical and population demographic data, and the third (the Survey Sampling Module) does a Monte Carlo simulation of surveys around airports using the outputs of the other two programs. Each computer program and any special techniques used by the program were described. An example of the functionality of the Survey Simulation Program was given. Acoustical and population demographic data were designed around three airports with the use of actual air traffic data and Census 2000 data. A survey was designed, implemented, and analyzed with the Survey Sampling Module at each of the airports. One sample of 6,000 persons was drawn with stratification applied to the population at each airport; two subsamples of 2,000 persons were drawn from the 6,000-person sample of each trial, one sample drawn using the survey sample design (stratification by the regression model predictor variables) and the sample other drawn using random sampling. An additional survey was simulated at Airport 2 in which a sample of 6,000 persons was drawn with random sampling of the population. Each survey was simulated 100 times around each airport. Several conclusions have been reached from an analysis of the survey simulation results. First, stratification was successful in limiting the correlation between the regression model predictor variables ($PNLT_{max,av,80}$ and $\log_{10}(N_{80})$) while simultaneously maintaining the orig-

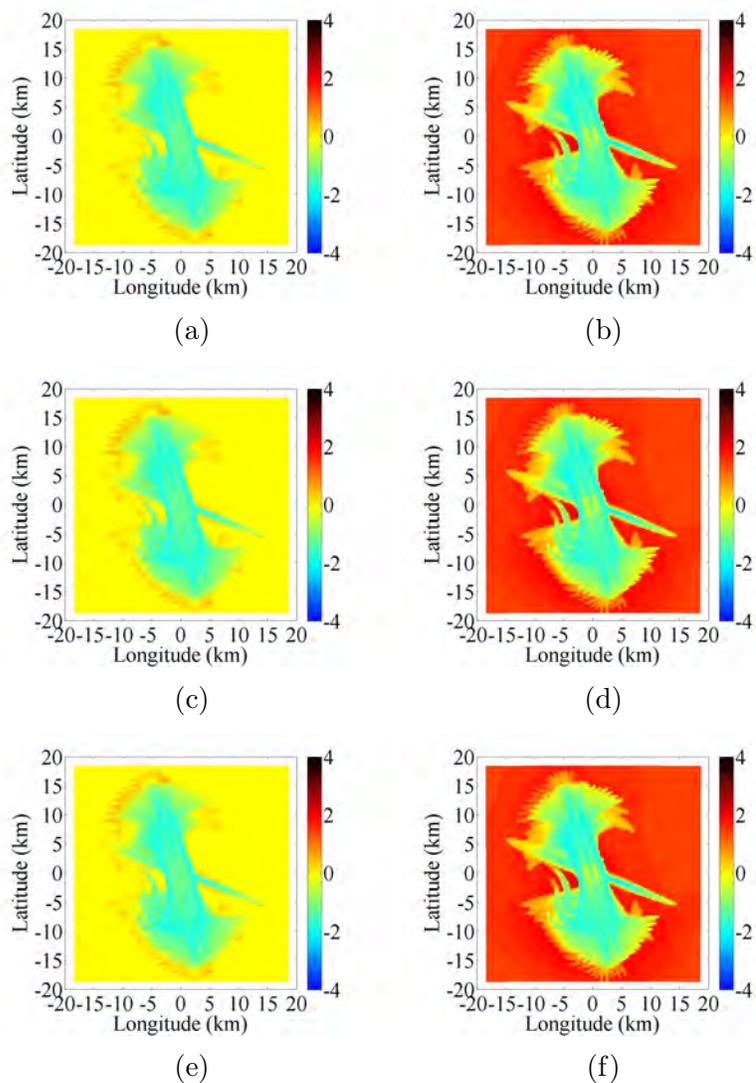


Figure 5.10. Annoyance model error plots for Airport 1. The error plots in (a) and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; and (e) and (f), RR estimates. The error plots in the left column ((a), (c), and (e)) are from using the N&L model estimates and in the right column ((b), (d), and (f)), *DNL* model estimates. The resolution of the plots is 0.125 nmi.

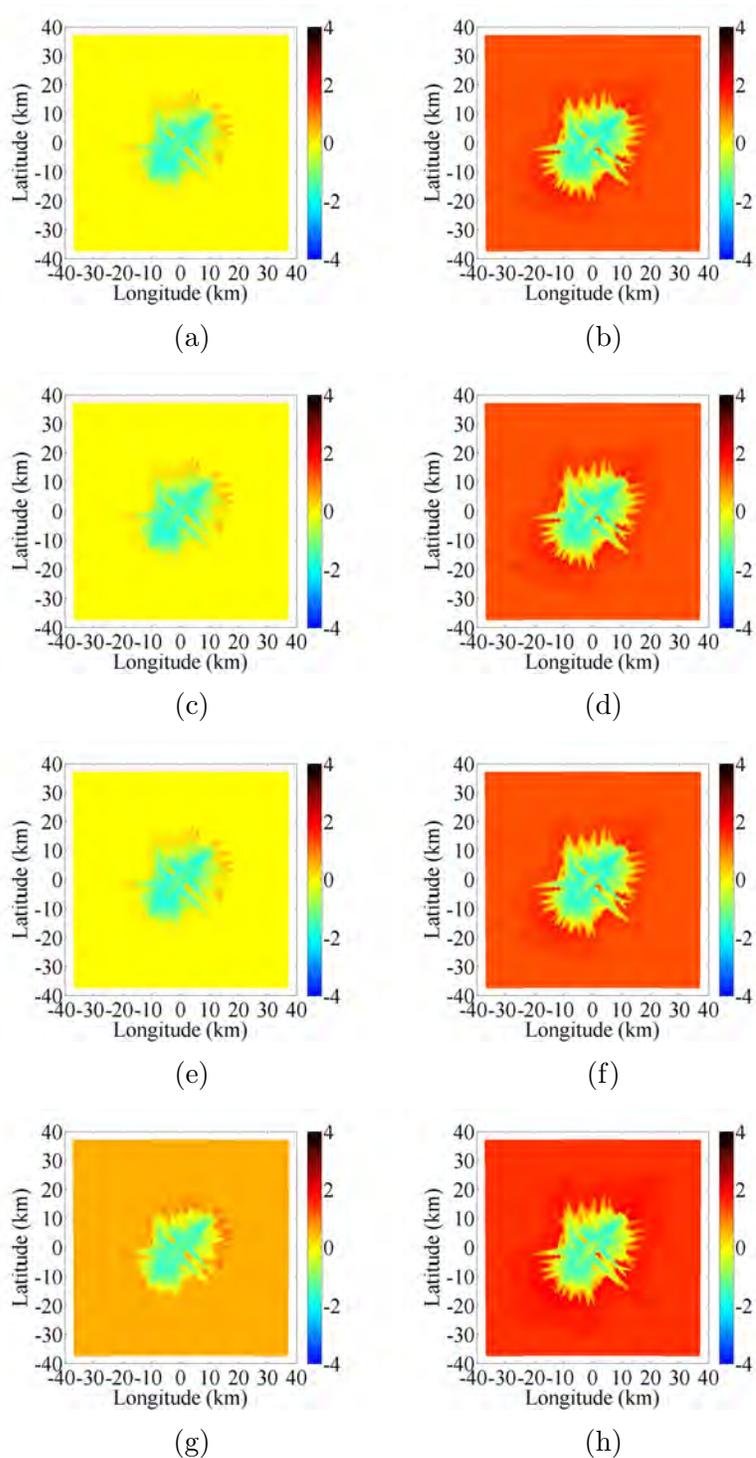


Figure 5.11. Annoyance model error plots for Airport 2. The error plots in (a and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; in (e) and (f), RR estimates; in (g) and (h), Full R estimates. The error plots in the left column ((a), (c), (e), and (g)) are from using the N&L model estimates and in the right column ((b), (d), (f), and (h)), *DNL* model estimates. The resolution of the plots is 0.125 nmi.

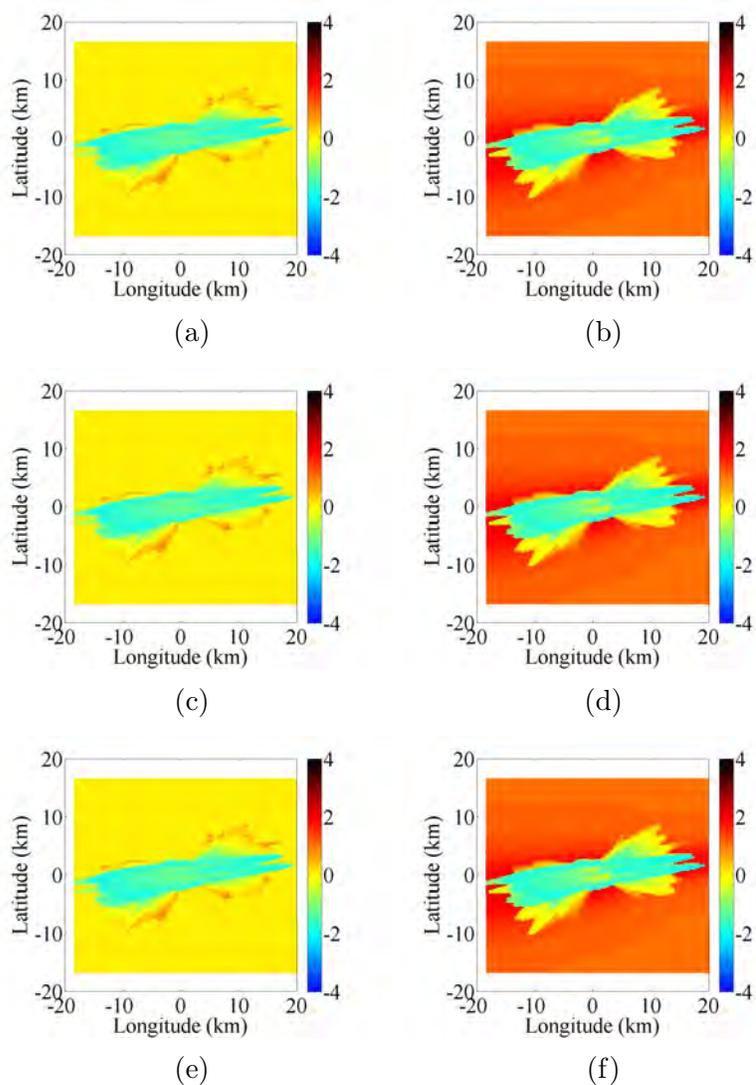


Figure 5.12. Annoyance model error plots for Airport 3. The error plots in (a) and (b) are from using the Full S sample estimates; in (c) and (d), RSD estimates; and (e) and (f), RR estimates. The error plots in the left column ((a), (c), and (e)) are from using the N&L model estimates and in the right column ((b), (d), and (f)), *DNL* model estimates. The resolution of the plots is 0.125 nmi.

inal ranges of each predictor variable in the population. As shown in Table 5.14, the correlations realized in the samples are lower than the correlations inherent in the population data, which are listed in Figure 5.8. In Chapter 4, Section 4.1.3, it was shown that variance estimates of \hat{k} can be lowered by lowering the correlation between variables and increasing the ranges of each predictor variable. Thus, stratification assisted in producing lower variance estimates of \hat{k} . The predictor variables in the population around Airport 2 were the most correlated, as can be seen in Figure 5.8b. A comparison of the results from stratified sampling and simple random sampling of the population around Airport 2 revealed that stratification had a large impact on the stability of the model estimation and resulted in less variation in \hat{k} . Second, $\log_{10}(N_{80})$ was quantified as more important to the prediction of annoyance than $PNLT_{max,av,80}$ by the \overline{R}_{sp}^2 analysis; in every simulation of every airport $\log_{10}(N_{80})$ was quantified as the more important predictor variable. This was a natural outcome of the population mean annoyance model used in the simulations which by virtue of the regression coefficients, resulted in $\log_{10}(N_{80})$ having greater weight than $PNLT_{max,av,80}$ in determining the population annoyance. Third, the retrospective power analysis revealed that there was less statistical power than what was planned for to distinguish the estimated \hat{k} in the samples from the equal-energy-value ($k = 10$). This resulted from the fact that the $PNLT_{max,av,80}$ and $\log_{10}(N_{80})$ data of the Combination #7 data set described in Chapter 4 were used to estimate a prospective $\hat{\sigma}_{\hat{k}}$; the correlation between the predictor variables of the Combination #7 data set is lower than what was realized in the simulations. Despite having less statistical power, estimates of \hat{k} were concluded to be significantly different from the equal-energy value of 10 in a majority of the 6,000-person sample stratified sampling trials. This is in contrast to results from the 6,000-person simple random sampling trials around Airport 2 in which \hat{k} was not distinguished from the equal-energy value in any of the trials. In addition, the jackknife and Taylor estimates of the standard deviation of \hat{k} ($\hat{\sigma}_{\hat{k}}$) were similarly-sized so that power results were not sensitive to which method was used to estimate $\hat{\sigma}_{\hat{k}}$. Finally, the predicted annoyance from two estimated mean annoyance

models of each simulation were compared to the population mean annoyance model predictions. At all three airports, the annoyance predictions of the estimated mean annoyance model identical in structure to the simulated population mean annoyance model were more accurate than the annoyance predictions of the other (*DNL*-based) estimated mean annoyance model. Subsampling the 6,000-person samples did not appreciably change the estimated mean annoyance models. Although model estimation was unstable in 6,000-person sample simple random sampling trials around Airport 2, the estimated mean annoyance model was comparable to that resulting from the 6,000-person sample stratified sampling.

CHAPTER 6. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

Aircraft noise can be disruptive to people in communities surrounding airports and can be a cause of annoyance. Many models of community annoyance to aircraft noise are functions only of the average sound energy of the total noise exposure, but other acoustical descriptors of noise exposure might also be important to the prediction of annoyance. One set of alternative models of community annoyance due to aircraft noise called “additive-log” models are functions of two variables: the number of events heard over a time period (the number-of-events), and the average maximum sound level of the events. In this work, the relative importance of the predictor variables as quantified by the decibel equivalent number effect and the average semipartial R^2 ($\overline{R_{sp}^2}$) in seven survey data sets consisting of 20,829 total observations around twenty-five airports has been investigated. Some adjustments were made to a couple of the data sets in order to establish a common sound level metric among the data sets, the average Maximum Perceived Noise Level ($PNL_{max,av}$) of aircraft events. In most cases, the number-of-events metric was usually the logarithm (base 10) of the number-of-events above a certain sound level. In the second part of the work, a survey simulation tool was developed to investigate how best to sample populations around airports to validate or compare various annoyance models. The utility of this computational tool was demonstrated through simulations of surveys around three airports. The focus of the simulations were to evaluate the effect of different sample designs on model estimation and to compare annoyance predictions from an additive-log model to those from a *DNL*-based model.

6.1 Summary and Conclusions

In Chapter 4, multiple linear regression analysis was used with historical survey data to generate models of annoyance in additive-log model form:

$$\hat{A} = \hat{b}_L PNL_{max,av} + \hat{b}_N \log_{10} N + \hat{b}_0, \quad (6.1)$$

where \hat{A} is the annoyance variable estimate, $PNL_{max,av}$ is the average Maximum Perceived Noise Level, \hat{b}_L is the coefficient estimate for $PNL_{max,av}$, $\log_{10}(N)$ is the logged (base 10) of the number-of-events, where N is the number-of-events, \hat{b}_N is the coefficient estimate for $\log_{10}(N)$, and \hat{b}_0 is the intercept estimate. In certain airport data sets the $PNL_{max,av}$ and $\log_{10}(N)$ data were calculated based on events above a certain sound level threshold, in others they were calculated including all aircraft operations. The decibel equivalent number effect ($\hat{k} = \hat{b}_N / \hat{b}_L$) was calculated for the data sets, which is the number of decibels which increases annoyance by the same amount as a tenfold increase in the number-of-events. The variance of \hat{k} was also calculated; in many cases, the variances were large so that k could not be determined precisely, or even quantified as being significant ($k > 0$). The magnitude of \hat{k} was not consistent among the data sets and varied over a large range.

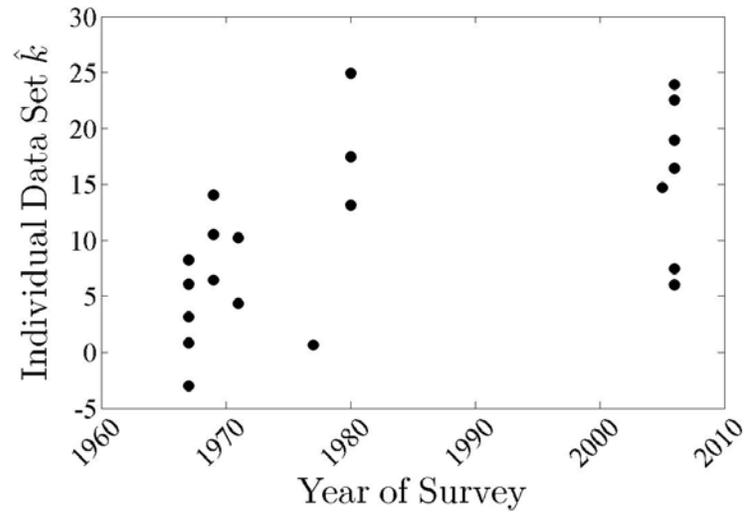
Both $PNL_{max,av}$ and $\log_{10}(N)$ were positively correlated with annoyance so that as either increased, annoyance also increased. However, $PNL_{max,av}$ and $\log_{10}(N)$ were also correlated in the airport data sets, so that a unique portion of variance of annoyance accounted for by each variable could not be determined. The $\overline{R_{sp}^2}$ calculations were used to determine average portions of the variance in the annoyance accounted for by each variable. Mostly, the $\overline{R_{sp}^2}$ calculations showed that $PNL_{max,av}$ was the more important variable in the prediction of annoyance. $PNL_{max,av}$ normally accounted for 0 to 20% of the variance in annoyance, and the contribution of $\log_{10}(N)$ was normally less than that of $PNL_{max,av}$, except in the most recent survey data.

The airport data sets were combined by the use of different strategies for establishing a common annoyance variable; seven combination data sets resulted. \hat{k} was

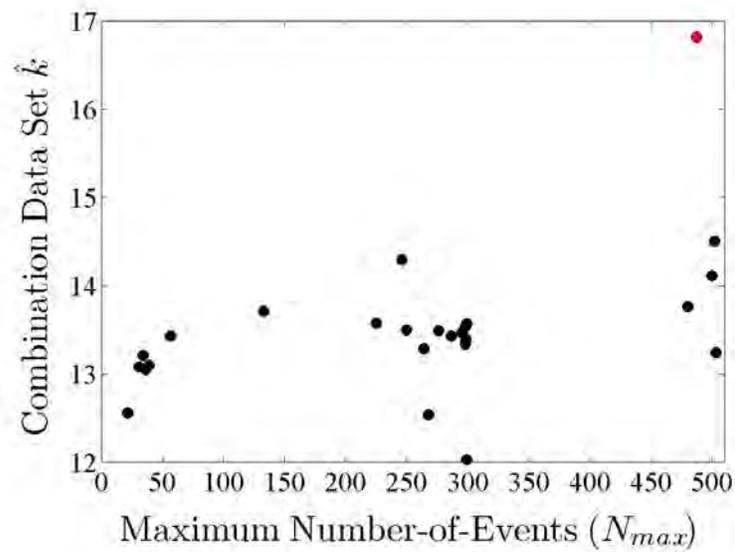
estimated in the combination data sets and was found to be consistent in magnitude across the data sets so that combination method did not appreciably affect the results. Also, \hat{k} was concluded as significant ($k > 0$) in all of the combination data sets. The estimated variance of \hat{k} was used to conclude the significance; the variance was calculated with jackknife repeated replication and involved the removal of subsamples belonging to individual airports from the full sample. The effect of removing subsamples from the combination data set on the estimation of \hat{k} was demonstrated; it was shown that a single airport could significantly influence the estimate of \hat{k} . The $\overline{R_{sp}^2}$ of $PNL_{max,av}$ was also greater than that of $\log_{10}(N)$ in the combination data sets.

In Figure 6.1a are shown the \hat{k} of the twenty-five airport data sets, which are also given in Table 4.2. There are four outlier \hat{k} not shown in the plot (-49.96 in 1980, 69.80 in 1980, 234.00 in 2006, -221.99 in 2006). There does not appear to be any correlation between survey year and \hat{k} . In Figure 6.1b are shown twenty-five \hat{k} 's of a combination data set, where one \hat{k} results from removing one airport data set from the combination data set. There does not appear to be any correlation between the maximum average daily number-of-events in a data set (N_{max}) and the \hat{k} of the combination data set without the data set. There is one outlier in Figure 6.1b, shown in red, which occurs from removing the UKD-024 data set from the combination data set. UKD-024 is the data set for the 1967 Heathrow aircraft noise study, and features the largest sample size of any data set investigated (4,655 samples).

The issues with model estimation revealed in the analyses given in Chapter 4 could be related to survey sample design. One data weakness in the historical data sets was identified as the correlation between the sound level and number-of-events, which inflates the estimated variance of \hat{k} and reduces the precision with which k can be estimated. Stratification was proposed as a sample design strategy capable of influencing the collection of predictor variable data from a sample, which might directly benefit model estimation. Stratification is the classification of a population into groups formed by the predictor variable data. Stratification was applied to the seven noise survey data sets; twenty-five strata were formed and the $PNL_{max,av}$,



(a)



(b)

Figure 6.1. (a) Individual data set \hat{k} plotted against the year of the survey. (b) The \hat{k} of a combination data set, where one \hat{k} results from removing an individual data set from the combination data set, plotted against the maximum average daily number-of-events (N_{max}). There is one outlier in (b), shown in red.

$\log_{10}(N)$, and annoyance data were averaged within strata so that the sample size was reduced to twenty-five average observations. The estimated \hat{k} of these data sets differed from the \hat{k} of the unstratified data sets, which indicated that the predictor variable data were not evenly represented over the predictor variable space.

The noise survey simulation program described in Chapter 5 was used to simulate noise surveys with and without the use of stratification, in order to evaluate the effect of sample design on model estimation. The populations were stratified by two factors: the logged (base 10) of the number of aircraft events with a Maximum Perceived Noise Level above 80 PNdB ($\log_{10}(N_{80})$) and the average Maximum Perceived Noise Level of those events ($PNL_{max,av,80}$). The intent of the survey simulations was to evaluate the effect of sample design on the ability to discover a population mean annoyance model, which generated the predictions annoyance of the survey populations. Stratification was successful in affecting the collection of the predictor variable data in ways that were beneficial to estimating the variance and confidence intervals around \hat{k} , so that k was determined with relative precision. Though the correlations between the $PNL_{max,av,80}$ and $\log_{10}(N_{80})$ data were high in the survey populations, the use of stratification realized lower correlations in the samples. The effects of stratification on model estimation were apparent when comparing the results from the stratified samples of 6,000 to those from the random samples of 6,000 at Airport 2: Using stratification, \hat{k} varied from 11.31 to 16.75 over 100 samples; without stratification, \hat{k} varied from 7.74 to 54.44. The stability of model estimation was affected by stratification. The variance of \hat{k} was estimated with two different methods: jackknife repeated replication and Taylor series linearization. The estimates resulting from the two different methods were similarly-sized. The concept of statistical power was introduced and was used in the design of the survey simulation sample size. Results from a power analysis performed after the collection of the survey data were compared to those performed before the survey data; it was found that there was less statistical power to distinguish \hat{k} from the equal-energy value ($k = 10$) in the survey data than planned for. Although there was less statistical power, \hat{k} was concluded as being significantly

different and greater than the equal-energy value in a majority of the simulations. Two regression models were estimated from the collected survey data: the first model was of identical structure as the population annoyance model, and the second model was a function of only DNL and an intercept. At all three airports, the annoyance predictions of the first model were less different from the population annoyance than the annoyance predictions of the second annoyance model. Subsampling the 6,000-person samples did not appreciably change the estimated regression models and the predicted annoyance.

In conclusion, estimated additive-log models derived from the data of the individual historical survey data sets could not be distinguished from an equal-energy type (such as DNL) in which $k = 10$. However, this could be an issue of sample design. If applied, the methodological improvements to noise surveys suggested by Fields (1980) would influence the precision of estimated noise index parameters and \hat{k} (Section 4.1.3 of Chapter 4). In particular, stratification of a population by the predictor variables in a regression model will help to realize three of Fields' five suggestions, and has shown to be more effective than simple random sampling.

6.2 Recommendations for Future Work

There are many possibilities for future research in the area of annoyance to aircraft noise. This research has introduced a collection of techniques and tools that can be expanded and built upon. Some recommendations for future work follows:

1. *Acquire more historical survey data sets.* The regression modeling and statistical techniques described in Chapter 4 can be applied to any data set with annoyance, maximum sound level, and number-of-events variables. There are many data sets that are of interest, see, e.g., Bassarab et al. (2009), but many of these are difficult to obtain, since organizations that are in possession of the data are at times reluctant to share the data. This is unfortunate and prevents the advance of the research. Recent data sets are especially desirable, since the

acoustical variables are likely estimated with greater accuracy, which affects the variance estimation of \hat{k} .

2. *Investigate other annoyance models.* This research focused primarily on the additive-log model (Equation (2.2)). However, other statistical models are also of interest and might predict annoyance more adequately than the additive-log model. These include models that use the unlogged number-of-events, piecewise regression models, and nonlinear models. In Chapter 2, Section 2.3, the concept of the holistic annoyance model was introduced. Holistic annoyance models are those which involve nonacoustical predictors variables in addition to acoustical predictors of annoyance. Researchers have investigated the effects of certain attitude, socio-demographic, and situational variables on the annoyance to noise, but there are other variables which could account for even more variance in the annoyance of individuals in communities surrounding airports. For example, housing construction type could be either an important predictor variable, or could even be used to approximate indoor sound levels from measured or simulated outdoor sound levels.
3. *Perform more survey simulations.* Use the Survey Simulation Program to simulate annoyance surveys around airports which are proposed to be included in upcoming surveys around airports in the United States. Different sample designs could be evaluated in terms of their effectiveness in gathering sufficient data to estimate statistically robust annoyance models.
4. *Evaluate the effect of using site-level data in annoyance models.* Acoustical data were simulated at the population-weighted centroids of 2001 Census geographical areas in the ANASE study (Le Masurier et al., 2007a). All survey respondents within a given geographical area were assigned the acoustical data of the centroid so that individual variation in noise exposure was not accounted for. The survey simulation tool has the capability of using person-level or site-level acoustical data in the estimation of annoyance models. Thus, the implications of using site-level acoustical data in person-level annoyance models can be thor-

oughly investigated. Related to this is the effect of uncertainty in predictor variable data measurements on model estimation. The Survey Simulation Program cannot currently model predictor variable uncertainty, but this capability should be added.

5. *Investigate various hypotheses.* Researchers recently introduced the concept of a loudness-based annoyance model developed using signal detection theory which involves a community-specific constant that intensifies or reduces the annoyance predictions (Fidell, Schultz, and Green, 1988; Fidell, Mestre, Schomer, Berry, Gjestland, Vallet, and Reid, 2011). Other researchers have suggested that reactions to aircraft and other noise sources are changing with time (Le Masurier et al., 2007a). Both of these hypotheses should be explored through further analysis of survey data and survey simulations.

LIST OF REFERENCES

LIST OF REFERENCES

- A. Alexandre. An assessment of certain causal models used in surveys on aircraft noise annoyance. *Journal of Sound and Vibration*, 44(1):119–125, 1999.
- P.D. Allison. *Multiple Regression: A Primer*. Pine Forge, 1999.
- R. Azen, D.V. Budescu, and B. Reiser. The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, 8(2):129–148, 2003.
- R. Bassarab, B. Sharp, and B. Robinette. An updated catalog of 628 social surveys of residents' reaction to environmental noise (1943-2008). Wyle Report WR 09-18, Wyle Laboratories, Inc., Arlington, Virginia, 2009. Prepared for the Federal Aviation Administration, Office of Environment and Energy.
- D.A. Belsley. Conditioning diagnostics through centering. *The American Statistician*, 38(2):73–77, 1984.
- D.A. Belsley. *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. Wiley-Interscience, 1991a.
- D.A. Belsley. A guide to using the collinearity diagnostics. *Computer Science in Economics and Management*, 4(1):33–50, 1991b.
- R.L. Bennett and K.S. Pearsons. Handbook of Aircraft Noise Metrics. National Technical Information Service N81-21871, Contract Report NASA-CR-3406, 1981. Prepared by Bolt, Beranek and Newman Inc, Canoga Park, CA.
- W.D. Berry and S. Feldman. *Multiple Regression in Practice*. Quantitative Applications in the Social Sciences. Sage Publications, Inc., 1985.
- D. Bertoni, A. Franchini, M. Magnoni, P. Tartoni, and M. Vallet. Reaction of people to urban traffic noise in Modena, Italy. In *Noise and Man 1993. Noise as a Public Health Problem. Proceedings of the 6th International Congress.*, Nice, France, June 5-9, 1993.
- D.E. Bishop. Comparisons of variability in aircraft flyover noise measurements. *Journal of the Acoustical Society of America*, 58(6):1211–1221, 1975.
- M. Bjorkman. Community noise annoyance: importance of noise levels and the number of noise events. *Journal of Sound and Vibration*, 151(3):497–503, 1991.
- P.N. Borsky. Community aspects of aircraft annoyance. National Opinion Research Center Report No. 54, 1954.
- J. Bring. How to standardize regression coefficients. *The American Statistician*, 48(3):209–213, 1994.

P. Brooker, J.B. Critchley, D.J. Monkman, and C. Richmond. United Kingdom Aircraft Noise Index Study (ANIS): Main Report. DR Report 8402, Civil Aviation Authority, January 1985.

D.V. Budescu. Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114(3):542–551, 1993.

R.B. Bullen and A.J. Hede. Reaction to aircraft noise in residential areas around Australian airports. *Journal of Sound and Vibration*, 108(2):199–225, 1986a.

R.B. Bullen and A.J. Hede. Comparison of the effectiveness of measures of aircraft noise exposure by using social survey data. *Journal of Sound and Vibration*, 108(2):227–245, 1986b.

W. Bürck, M. Grützmacher, F.J. Meister, and E.A. Müller. Fluglärm: seine messung und bewertung, seine berücksichtigung bei der siedlungsplanung, massnahmen zu seiner minderung (Aircraft Noise: Its Measurement and Evaluation, its Significance for Community Planning, and Measures for its Abatement). Technical report, German Federal Ministry of Health, Göttingen, 1965.

California Department of Aeronautics. Noise Standards, California Administrative Code, Register 70, No. 48, Chapter 9, Title 4, November 1970.

S. Chatterjee and A.S. Hadi. *Regression Analysis By Example*. Quantitative Applications in the Social Sciences. Wiley-Interscience, 2006.

J. Cohen. A power primer. *Psychological Bulletin*, 112(1):155–159, 1992.

J. Cohen and P. Cohen. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Inc, 365 Broadway, Hillsdale, New Jersey 07642, second edition, 2005.

W.K. Connor and H.P. Patterson. Community Reaction to Aircraft Noise Around Smaller City Airports. Contractor Report NASA-CR-2104, National Aeronautics and Space Administration, 1972. Prepared by TRACOR, Inc., 6500 Tracor Lane, Austin, Texas 78721.

T. Courville and B. Thompson. Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement*, 61(2):229–248, 2001.

T.K. Dempsey, D.G. Stephens, J.M. Fields, and K.P. Shepherd. Residents' Annoyance Responses to Aircraft Noise Events. Technical Report NASA-TP-2121, National Aeronautical and Space Administration, September 1983.

B. Efron. Bootstrap method: another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.

B. Efron and C. Stein. The jackknife estimate of variance. *The Annals of Statistics*, 9(3):586–596, 1981.

FAA. *Noise Standards: Aircraft Type and Airworthiness Certification, Calculation of Effective Perceived Noise Level from Measured Data*. Federal Aviation Administration (FAA), 2002. Federal Aviation Regulations Part 36, Appendix A2 to Part 36 - Section A36.4.

D.E. Farrar and R.R. Glauber. Multicollinearity in regression analysis: the problem revisited. *The Review of Economics and Statistics*, 49(1):392–107, 1967.

Federal Aviation Administration, Office of Environment and Energy. Summary report of the workshop on Aircraft Noise Impacts Research. Technical Report, Washington DC, December 10-11, 2009.

S. Fidell and K.S. Pearsons. Community response to environmental noise. In *Encyclopedia of Acoustics*. John Wiley and Sons, Inc., 2007.

S. Fidell, T. Schultz, and D.M. Green. A theoretical interpretation of the prevalence rate of noise-induced annoyance in residential populations. *Journal of the Acoustical Society of America*, 84(6):2109–2113, 1988.

S. Fidell, V. Mestre, P. Schomer, B. Berry, T. Gjestland, M. Vallet, and T. Reid. A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure. *Journal of the Acoustical Society of America*, 130(2):791–806, 2011.

J.M. Fields. Designing community surveys to provide a basis for noise policy. Technical Report NASA-TM-80110, National Aeronautical and Space Administration, June 1980.

J.M. Fields. Effect of numbers of noise events on people's reactions to noise: an analysis of existing survey data. *Journal of the Acoustical Society of America*, 75(2):447–467, 1984.

J.M. Fields. Effect of personal and situational variables on noise annoyance in residential areas. *Journal of the Acoustical Society of America*, 93(5):2753–2763, 1993.

J.M. Fields. The TNO Data Base Project Report. Contract F41624-94-D-6001, D.O. 0010, June 1997. Prepared for Ball Aerospace and Technologies Corporation.

J.M. Fields. An Updated Catalog of 521 Social Surveys of Residents' Reaction to Environmental Noise (1943-2000). Technical Report NASA/CR-2001-211257, National Aeronautical and Space Administration, 2001.

J.M. Fields and J.G. Walker. The response to railway noise in residential areas in Great Britain. *Journal of Sound and Vibration*, 85(2):177–255, 1982.

J.M. Fields, R.G. De Jong, T. Gjestland, I.H. Flindell, R.F.S. Job, S. Kurra, P. Lercher, M. Vallet, T. Yano, R. Guski, U. Felscher-suhr, and R. Schumer. Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation. *Journal of Sound and Vibration*, 242(4):641–679, 2001.

W.A. Fuller. *Sampling Statistics*. John Wiley and Sons, Inc., 2009.

W.J. Galloway. Community Noise Exposure Resulting from Aircraft Operations: Technical Review. Air Force Report AMRL-TR-73-106, Bolt Beranek and Newman Inc., November 1974.

W.J. Galloway and D.E. Bishop. Noise Exposure Forecasts: Evolution, Evaluation, Extensions and Land Use Interpretations. Technical Report No. 1862, Bolt Beranek and Newman Inc., August 1970.

W.J. Galloway and A. Pietrasanta. Land Use Planning with Respect to Aircraft Noise. Technical Report AFM 86-5, TM 5-365, NAVDOCKS P-98, Department of Defense, 1964. Also published as Bolt, Beranek and Newman Report No. 821.

R.F. Gunst. Regression analysis with multicollinear predictor variables: definition, detection, and effects. *Communication in Statistics - Theory and Methods*, 12(19): 2217–2260, 1983.

R. Guski. Personal and social variables as co-determinants of noise annoyance. *Noise and Health*, (3):45–56, 1999.

F.E. Harrell Jr. Regression modeling strategies. <http://biostat.mc.vanderbilt.edu/twiki/pub/Main/RmS/rms.pdf>, March 2012.

H. He, E. Dinges, J. Hemann, D. Rickell, L. Mirsky, C. Roof, E. Boeker, P. Gerbi, and D. Senzig. *INM 7.0 User's Guide*. U.S. Department of Transportation, Federal Aviation Administration, Office of Environment and Energy, 800 Independence avenue, S.W., Washington, DC 20591, 2007.

A.J. Hede and R.B. Bullen. Aircraft Noise in Australia: A Survey of Community Reaction. NAL Report No. 88., National Acoustic Laboratories, Canberra, Australia, 1982.

R.R. Hocking and O.J. Pendleton. The regression dilemma. *Communication in Statistics - Theory and Methods*, 12(5):497–527, 1983.

Z. Huang and P. Williamson. A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Working Paper 2001/2, Department of Geography, University of Liverpool, October 2001. Obtained from http://pcwww.liv.ac.uk/~william/microdata/workingpapers/hw_wp_2001_2.pdf.

Z. Hudson. Sample size, power, and effect size - what all researchers need to know. *Physical Therapy in Sport*, 10:43–44, 2009.

ICAO. *Annex 16 to the Convention on International Civil Aviation: Environmental Protection, Volume 1 - Aircraft Noise*. International Civil Aviation Organization (ICAO), fifth edition, August 2008.

B. Ikelheimer and K. Plotkin. *Noise Model Simulation (NMSim) User's Manual*. Wyle Acoustics Group, 2001 Jefferson Davis Highway, Suite 701, Arlington, VA 22202, 2005.

ISO/R-131-1959(E). *Expression of the physical and subjective magnitudes of sound*. The International Organisation for Standardization, Case Postale 56, CH-1211 Geneva 20, Switzerland, 1959.

S.A. Janssen and H. Vos. Dose-Response Relationship Between DNL and Aircraft Noise Annoyance: Contribution of TNO. Technical Report DOT/FAA/AEE/2011-04, Department of Geography, University of Liverpool, 2011.

R.F.S. Job. Community response to noise: a review of factors influencing the relationship between noise exposure and reaction. *Journal of the Acoustical Society of America*, 83(3):991–1001, 1988.

- J.W. Johnson and J.M. LeBreton. History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7(3):238–257, 2004.
- R. Josse. La gêne causée par le bruit autour des aéroports (The Annoyance Caused By Noise Around Airports). Final Report, DGRBT/CSTB No. 63-FR-138, Centre Scientifique et Technique du Bâtiment, Paris, March 1, 1968. p. 1-219.
- G. Kalton. *Introduction to Survey Sampling*. Quantitative Applications in the Social Sciences. Sage Publications, Inc., 1983.
- G. Keren and C. Lewis. *A Handbook for Data Analysis in the Behavioral Sciences*, volume 1. Psychological Press, 1993.
- L. Kish. *Survey Sampling*. John Wiley and Sons, Inc., 1965.
- L. Kish and M.R. Frankel. Inference from complex samples. *Journal of the Royal Statistical Society*, 36(1):1–37, 1974.
- M. Kroesen, E.J.E. Molin, and B. van Wee. Testing a theory of aircraft noise annoyance: a structural equation analysis. *Journal of the Acoustical Society of America*, 123(6):4250–4260, 2008.
- W. Kruskal and R. Majors. Concepts of relative importance in recent scientific literature. *The American Statistician*, 43(1):2–6, 1989.
- K. Kryter. The meaning and measurement of Perceived Noise Level. *Noise Control*, 6(5):12–27, 1960.
- K.D. Kryter and K.S. Pearsons. Judged noisiness of a band of random noise containing an audible pure tone. *Journal of the Acoustical Society of America*, 38:106–112, 1965.
- T.K. Kumar. Multicollinearity in regression analysis. *The Review of Economics and Statistics*, 57(3):365–366, 1975.
- M.H. Kutner, C.J. Nachtsheim, J. Neter, and W. Li. *Applied Linear Statistical Models*. Operations and Decision Sciences. McGraw Hill/Irwin, fifth edition, 2005.
- G. Labiale. Laboratory study of the influence of noise level and vehicle number on annoyance. *Journal of Sound and Vibration*, 90(3):361–371, 1983.
- Landrum and Brown, Inc. Indianapolis International Airport Noise Exposure Map Update. Part 150 Technical Report, Landrum and Brown, Inc., December 2008.
- LAWA Noise Management. California State Airport Noise Standards Quarterly Report - Los Angeles World Airports, 4Q11, Los Angeles International Airport. <http://www.lawa.org/uploadedFiles/LAX/pdf/lax4q11%20noise%20contour%20map.pdf>, October 15, 2012.
- P. Le Masurier, J. Bates, and J. Taylor. Attitudes to Noise from Aviation Sources in England (ANASE). Final Report, MVA Consultancy, 2007a. Prepared for the UK Department of Transport.
- P. Le Masurier, J. Bates, and J. Taylor. Attitudes to Noise from Aviation Sources in England (ANASE). Technical Appendices, MVA Consultancy, 2007b. Prepared for the UK Department of Transport.

S. Leonard and P.N. Borsky. A causal model for relating noise exposure, psychosocial variables and aircraft noise annoyance. In *Proceedings of the International Congress on Noise as a Public Health Problem*, Dubrovnik, Yugoslavia, May 13-18, 1973. p. 691-705.

R.H. Lindeman, P.F. Merenda, and R.Z. Gold. *Introduction to Bivariate and Multivariate Analysis*. Scott, Foresman and Company, 1980.

S.L. Lohr. *Sampling: Design and Analysis*. Duxbury Press, 1999.

C. Marquis-Favre, E. Premat, and D. Aubrée. Noise and its effects - a review on qualitative aspects of sound. Part II: noise and annoyance. *Acta Acustica United with Acustica*, 91:626-642, 2005a.

C. Marquis-Favre, E. Premat, and D. Aubrée. Noise and its effects - a review on qualitative aspects of sound. Part I: notions and acoustic ratings. *Acta Acustica United with Acustica*, 91:613-625, 2005b.

C.H. Mason and W.D. Perreault, Jr. Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, 28(3):268-280, 1991.

MathWorks, Inc. Example of *xmlread*. <http://www.mathworks.com/help/matlab/ref/xmlread.html>.

MathWorks, Inc. Combinator, combinations and permutations. <http://www.mathworks.com/matlabcentral/fileexchange/24325-combinator-combinations-and-permutations>, June 2009. Obtained from the MATLAB Central File Exchange, authored by Matt Fig.

S.E. Maxwell, K. Kelley, and J.R. Rausch. Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology*, 59:537-563, 2008.

A.C. McKennell. Aircraft Annoyance around London (Heathrow) Airport: The Social Survey. Report 55337, UK Government, London, England, 1963.

A.C. McKennell. Community Response to Concorde Flights Round London (Heathrow) Airport. Technical Report, Social and Community Planning Research, London, England, 1977.

H.M.E. Miedema and H. Vos. Exposure-response relationships for transportation noise. *Journal of the Acoustical Society of America*, 104(6):3432-3445, 1998.

H.M.E. Miedema and H. Vos. Demographic and attitudinal factors that modify annoyance from transportation noise. *Journal of the Acoustical Society of America*, 105(6):3336-3344, 1999.

MIL Research Ltd. Second survey of aircraft noise annoyance around London (Heathrow) airport. Technical Report, Office of Population Censuses and Surveys, Social Survey Division, H.M.S.O. London, England, 1971.

C.Z. Mooney. *Monte Carlo Simulation*. Quantitative Applications in the Social Sciences. Sage Publications, Inc., 1997.

- C.Z. Mooney and R.D. Duval. *Boostrapping: A Nonparametric Approach to Statistical Inference*. Quantitative Applications in the Social Sciences. Sage Publications, Inc., 1993.
- N.E. Murray and R.A. Piese. Aircraft take-off and landing noise annoyance criteria. Technical Report Report No. 22, Commonwealth Acoustic Laboratories, Sydney, Australia, May 1964.
- P.M. Nelson. *Transportation Noise Reference Book*. Butterworth, 1987.
- Y. Osada. Comparison of community reaction to traffic noise. *Journal of Sound and Vibration*, 151(3):479–486, 1991.
- J.E. Overall and D.K. Spiegel. Concerning least squares analysis of experimental data. *Psychological Bulletin*, 72(5):311–322, 1969.
- K. Pearsons, R. Howe, M. Sneddon, L. Silvati, and S. Fidell. Comparison of predictors of the annoyance of commuter, stage II, and stage III aircraft overflights as heard outdoors. Technical Report NASA/CR-97-205812, 1997.
- C.A. Powell. Annoyance due to multiple airplane noise exposure. Technical Report NASA Technical Paper 1706, 1980.
- G.J. Raw and I.D. Griffiths. Individual differences in response to road traffic noise. *Journal of Sound and Vibration*, 121(3):463–471, 1988.
- C.G. Rice. Investigation of the trade-off effects of aircraft noise and number. *Journal of Sound and Vibration*, 52(3):325–344, 1977a.
- C.G. Rice. Development of cumulative noise measure for the prediction of general annoyance in an average population. *Journal of Sound and Vibration*, 52(3):345–364, 1977b.
- C.G. Rice. Trade-off effects of aircraft noise and number of events. In *Proceedings of the 3rd International Congress on Noise as a Public Health Problem*, American Speech-Language-Hearing Association Report 10, pages 495–510, Freiburg, West Germany, September 25-29, 1978.
- K. Rust. Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics*, 1(4):381–397, 1989.
- J. Ryan, H. Maoh, and P. Kanaroglou. Population synthesis: comparing the major techniques using a small, complete population of firms. *Geographical Analysis*, 41: 181–203, 2009.
- R. Rylander and M. Bjorkman. Annoyance by aircraft noise around small airports. *Journal of Sound and Vibration*, 205(4):533–537, 1997.
- R. Rylander, S. Sorensen, and A. Kajland. Annoyance reactions from aircraft noise exposure. *Journal of Sound and Vibration*, 24(4):419–444, 1972.
- R. Rylander, S. Sorensen, and K. Berglund. Re-analysis of aircraft noise annoyance data against the db(a) peak concept. *Journal of Sound and Vibration*, 36(3):399–406, 1974.

- R. Rylander, M. Bjorkman, and U. Ahrlin. Aircraft noise annoyance contours: importance of overflight frequency and noise level. *Journal of Sound and Vibration*, 69(4):583–595, 1980.
- D. Schreckenberg and R. Schuemer. The impact of acoustical, operational and non-auditory factors on short-term annoyance due to aircraft noise. In *Inter-Noise 2010*, 2010.
- M. Kroesen D. Schreckenberg. A measurement model for general noise reaction in response to aircraft noise. *Journal of the Acoustical Society of America*, 129(1): 200–210, 2011.
- T.J. Schultz. *Community Noise Rating*. Applied Science Publishers, second edition, 1982.
- G.W. Stewart. *Introduction to Matrix Computations*. Quantitative Applications in the Social Sciences. Academic Press, 1973.
- G.W. Stewart. Collinearity and least squares regression. *Statistical Science*, 2(1): 68–84, 1987.
- E. Sul Lee, R.N. Forthofer, and R.J. Lorimor. *Analyzing Complex Survey Data*. Quantitative Applications in the Social Sciences. Sage Publications, Inc., 1989.
- S.M. Taylor. A path model of aircraft noise annoyance. *Journal of Sound and Vibration*, 96(2):243–260, 1984.
- S.K. Thompson. *Sampling*. John Wiley and Sons, Inc., 1992.
- T.J. Tomberlin. Statistical methods for efficient design of community surveys of response to noise: random coefficients regression models. Technical Report NASA-CR-177933, National Aeronautical and Space Administration, July 1985.
- TRACOR, Inc. Community Reaction to Aircraft Noise, Volume I. Technical Report Contractor Report NASA-CR-1761, National Aeronautics and Space Administration, 1971. Prepared by TRACOR, Inc., 6500 Tracor Lane, Austin, Texas 78721.
- United States Census Bureau. Census 2000 summary File 3 FTP. http://www2.census.gov/census_2000/datasets/Summary_File_3/, May 2002.
- United States Census Bureau. Census 2000 5-percent Public Use Microdata Sample (PUMS) files. <http://www.census.gov/census2000/PUMS5.html>, August-September 2003.
- U.S. Environmental Protection Agency. Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety. U.S. Environmental Protection Agency, Office of Noise Abatement and Control Technical Report 550/9-74-004, 1974.
- M. Vallet. Annoyance after changes in airport noise environment. In *Proceedings of Inter-noise '96*, pages 2329–2335, Liverpool, UK, July 30 to August 2, 1996.
- C.G. van Niekerk and J.L. Muller. Assessment of aircraft noise disturbance. *Journal of the Royal Aeronautical Society*, 73:383–396, 1969.

- D. Voas and P. Williamson. An evaluation of the combinatorial optimisation approach to the creation of sythetic microdata. *International Journal of Population Geography*, 6:349–366, 2000.
- D. Voas and P. Williamson. Evaluating goodness-of-fit measures for synthetic microdata. *Geographical and Environmental Modelling*, 5(2):177–200, 2001.
- A. Wilson. Final Report of the Committee on the Problem of Noise. Cmnd. 2056, Wilson Commitee, London, England, July 1963.
- C. Winship and L. Radbill. Sampling weights and regression analysis. *Sociological Methods and Research*, 23(2):230–257, 1994.
- K. Wolter. *Introduction to Variance Estimation*, volume 279 of *Springer Series in Statistics*. Springer, 2003.
- T. Yano, T. Sato, M. Bjorkman, and R. Rylander. Comparison of community response to road traffic noise in Japan and Sweden - Part II: path analysis. *Journal of Sound and Vibration*, 250(1):169–174, 2002.

APPENDICES

APPENDIX A. ADDITIVE-LOG ANNOYANCE MODELS AND SOUND LEVEL METRICS

In this appendix is given an overview of the additive-log model and formulae for various additive-log models that have been proposed throughout the history of the study of annoyance to aircraft noise, in addition to additive-log approximations of certain noise exposure variables. Formulae for a selection of sound level metrics are also given.

A.1 Additive-Log Annoyance Models

The additive-log annoyance model is a function of a measure of sound level (in decibels) and the logarithm (base 10) of the number-of-events. Typically, the sound level variable in the additive-log model is an average sound level, e.g., $L_{Amax,av}$, $SEL_{A,av}$, $PNL_{max,av}$, or $EPNL_{av}$. The average sound level is an average of single-event sound levels. The average sound level and number-of-events data for a noise-exposed area can be calculated by including either all aircraft that operate in the vicinity of the area or only the aircraft events with a single-event sound level above a threshold. The method of averaging can also vary. However the average sound level is calculated, the number-of-events is the number of aircraft events involved in the calculation. The additive-log model can be expressed as:

$$\hat{A} = \hat{b}_L L + \hat{b}_N \log_{10} N + \hat{b}_0, \quad (\text{A.1})$$

where \hat{A} is the annoyance variable estimate, L is the sound level variable, \hat{b}_L is the sound level coefficient estimate, N is the number-of-events variable, \hat{b}_N is the number-of-events coefficient estimate, and \hat{b}_0 is the intercept estimate. In the rest of this section are given various additive-log models that have been proposed throughout

the history of the study of annoyance to aircraft noise. Apart from the references given for each additive-log model, additional information is contained in (Bennett and Pearsons, 1981; Galloway, 1974; Schultz, 1982).

Australian Annoyance Index (AI) The Australian Annoyance Index (AI) is defined by the formula:

$$AI = PNL_{max,av} + 10\log_{10}(N), \quad (\text{A.2})$$

where $PNL_{max,av}$ is the energy-based average of the maximum values of PNL of the events, and N is the number-of-events that occur over a given hour. AI is calculated for each hour. For more information, see (Murray and Piese, 1964).

Disturbance Index (\bar{Q}) The Disturbance Index (\bar{Q}) was used in Germany primarily for rating aircraft noise and in Austria for rating traffic noise, though it was also applicable to other situations. The general formula for \bar{Q} allows a variety of noise metrics to be produced based on the selection of the sound level metric and the value of a free parameter called the “Equivalence Parameter” (Schultz, 1982). The German version of \bar{Q} can be expressed in additive-log format and is defined by the formula:

$$\bar{Q} = PNL_{max,av} + 13.3\log_{10}(N) - 47, \quad (\text{A.3})$$

where $PNL_{max,av}$ is the energy-based average of the maximum values of PNL of the events, and N is the total number-of-events affecting the location in question. For more information, see (Bürck, Grützmacher, Meister, and Müller, 1965).

24-hour A-weighted Equivalent Sound Level (L_{Aeq}) The 24-hour A-weighted Equivalent Sound Level (L_{Aeq}) is defined as the average A-weighted acoustical energy contained within a given time period of exposure to noise. L_{Aeq} forms the basis of many other noise ratings including the Day-Night Average Sound Level (DNL),

the Day-Evening-Night Average Sound Level (*DENL*), and the Community Noise Equivalent Level (*CNEL*). The standard formulation of L_{Aeq} is:

$$L_{Aeq} = 10 \log_{10} \left[\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{p_A^2(t)}{p_{ref}^2(t)} dt \right], \quad (\text{A.4})$$

where $p_A^2(t)$ is the time-varying A-weighted mean square sound pressure, p_{ref} is the reference sound pressure of 20 μPA , t_1 is the start time of integration, and t_2 is the end time of integration. L_{Aeq} can be approximated in additive-log format by the formula:

$$L_{Aeq} = \overline{SEL_A} + 10 \log_{10}(N) - 10 \log_{10}(T), \quad (\text{A.5})$$

where $\overline{SEL_A}$ is an energy-based average of aircraft events' Sound Exposure Level (SEL_A), N is the number-of-events, and T is the duration of the time period considered. $\overline{SEL_A}$ is defined in Appendix A.2.2.

Composite Noise Rating (*CNR*) The Composite Noise Rating (*CNR*) was initially designed to predict the annoyance to any noise source, though it was later adapted to military aircraft noise exposure, and then commercial aircraft. *CNR* was not originally defined formulaically, but was based on a comparison of measured noise to a collection of community noise rating curves to yield a “level rank”. The level rank was then modified with corrections based on characteristics of the noise and the community. The various complex methods of calculating *CNR* are described in (Bennett and Pearsons, 1981; Schultz, 1982). Around 1963, *CNR* was simplified as an additive-log model with *PNL* as the sound level metric and was intended for commercial aircraft noise exposure. The *CNR* for the i^{th} class of commercial aircraft using the j^{th} flight path is defined by the formula:

$$CNR_{ij} = PNL_{max,av,ij} + 10 \log_{10}(N_{d,ij} + 16.67N_{n,ij}) - 12, \quad (\text{A.6})$$

where $PNL_{max,av,ij}$ is the energy-based average of the maximum values of PNL produced by the i^{th} class of commercial aircraft operating on the j^{th} flight path, $N_{d,ij}$ is the total number of daytime operations (7:00 AM to 10:00 PM) of this type, and $N_{n,ij}$ is the total number of nighttime (10:00 PM to 7:00 AM) operations of this type affecting the location in question. The overall CNR is calculated by the formula:

$$CNR = 10 \log_{10} \sum_{i=1}^I \sum_{j=1}^J 10^{CNR_{ij}/10}, \quad (\text{A.7})$$

where I is the number of commercial aircraft classes and J is the number of flight paths. For more information, see (Galloway and Pietrasanta, 1964).

Noise Exposure Forecast (NEF) The Noise Exposure Forecast (NEF) was developed as an improvement to CNR and was intended to be applied only to commercial aircraft noise exposure. NEF uses the Effective Perceived Noise Level ($EPNL$) as the base sound level metric; $EPNL$ is a refinement of PNL , which is used in CNR . The method of calculating $EPNL$ is documented in (FAA, 2002). The NEF for the i^{th} class of commercial aircraft using the j^{th} flight path is defined by the formula:

$$NEF_{ij} = EPNL_{ij} + 10 \log_{10}(N_{d,ij} + 16.67N_{n,ij}) - 88, \quad (\text{A.8})$$

where $EPNL_{ij}$ is the $EPNL$ produced by the i^{th} class of commercial aircraft operating on the j^{th} flight path, $N_{d,ij}$ is the total number of daytime operations (7:00 AM to 10:00 PM) of this type, and $N_{n,ij}$ is the total number of nighttime (10:00 PM to 7:00 AM) operations of this type affecting the location in question. The overall NEF is calculated by the formula:

$$NEF = 10 \log_{10} \sum_{i=1}^I \sum_{j=1}^J 10^{NEF_{ij}/10}, \quad (\text{A.9})$$

where I is the number of aircraft classes and J is the number of flight paths. For more information, see (Galloway and Bishop, 1970).

Noise and Number Index (*NNI*) The Noise and Number Index (*NNI*) was developed during a survey of the impact of aircraft noise around London (Heathrow) airport. *NNI* was established from a graphical analysis, as opposed to a mathematical analysis, of the survey data. *NNI* was not established from the mathematical analysis due to the high correlation between the predictor variables. *NNI* is defined by the formula:

$$NNI = PNL_{max,av} + 15\log_{10}(N) - 80, \quad (\text{A.10})$$

where $PNL_{max,av}$ is the energy-based average of the maximum values of *PNL* of the events, and N is the number-of-events heard in a specified period. *PNL* and $PNL_{max,av}$ are defined in Appendix A.2.1. For more information, see (McKennell, 1963; Wilson, 1963).

Isopsophic Index (*R*) The Isopsophic Index (*R*) is a French additive-log model based on $PNL_{max,av}$ with a more complicated treatment of nighttime operations than other models. *R* for the daytime (6:00 AM to 10:00 PM) is defined by the formula:

$$R_d = PNL_{max,av,d} + 10\log_{10}(N_d) - 30, \quad (\text{A.11})$$

where R_d is the Isopsophic Index for the daytime, $PNL_{max,av,d}$ is the energy-based average of the maximum values of *PNL* of the daytime events, and N_d is the total number of daytime events. Noise exposure for nighttime (10:00 PM to 6:00 AM) is split into two intervals: 10:00 PM to 2:00 AM and 2:00 AM to 6:00 AM. Operations that occur in the first time period are treated as three times more significant than those that occur in the second period. *R* for the nighttime (10:00 PM to 6:00 AM) is defined by the formula:

$$R_n = PNL_{max,av,n} + 6\log_{10}(3N_{n1} + N_{n2}) - 31, \quad (\text{A.12})$$

where R_n is the Isopsophic Index for the nighttime, N_{n1} is the number of nighttime events that occur from 10:00 PM to 2:00 AM, and N_{n2} is the number of nighttime

events that occur from 2:00 AM to 6:00 AM. If $3N_{n1} + N_{n2} < 64$, then the different nighttime weightings are abandoned and the formula reduces to:

$$R_n = PNL_{max,av,n} + 10\log_{10}(N_n) - 30, \quad (\text{A.13})$$

where N_n is the number of nighttime events. For more information, see (Josse, 1968).

Weighted Equivalent Continuous Perceived Noise Level (*WECPNL*) The Weighted Equivalent Continuous Perceived Noise Level (*WECPNL*) was developed by the International Civil Aviation Organization (ICAO) as an effort to reconcile differences between the already-existing noise ratings of the member countries of ICAO. *WECPNL* is calculated in three steps. First, the Total Noise Exposure Level (*TNEL*) must be calculated:

$$TNEL = 10\log_{10} \left(\sum_{i=1}^n 10^{EPNL_i/10} \right) + 10\log \left(\frac{T_0}{t_0} \right), \quad (\text{A.14})$$

where $EPNL_i$ is the Effective Perceived Noise Level of the i^{th} event, n is the number-of-events, $T_0 = 10$ seconds, and $t_0 = 1$ second. *TNEL* is used to calculate the Equivalent Continuous Perceived Noise Level (*ECPNL*):

$$ECPNL = TNEL - 10\log \left(\frac{T}{t_0} \right), \quad (\text{A.15})$$

where T is a specified time period under consideration in seconds. *WECPNL* is calculated from the daytime and nighttime *ECPNL*, and incorporates a seasonal adjustment:

$$WECPNL = 10\log \left(\frac{5}{8} 10^{ECPNL_d/10} + \frac{3}{8} 10^{ECPNL_n/10} \right) + S, \quad (\text{A.16})$$

where $ECPNL_d$ is the *ECPNL* for the daytime (7:00 AM to 10:00 PM), $ECPNL_n$ is the *ECPNL* for the nighttime (10:00 PM to 7:00 AM), and S is an adjustment

dependent on the outdoor temperature. If the temperature is at or above 20°C less than 100 hours per month, $S = -5$. If the temperature is at or above 20°C more than 100 hours per month and at or above 25.6°C less than 100 hours per month, $S = 0$. If the temperature is at or above 25.6°C more than 100 hours per month, $S = +5$. In practice, $WECPNL$ is used in additive-log format:

$$WECPNL = \overline{ECPNL} + 10\log_{10}(N_1 + 3N_2 + 10N_3) - 40, \quad (\text{A.17})$$

where \overline{ECPNL} is the energy-based average of the $ECPNL$ produced by single aircraft events, N_1 is the number-of-events that occur from 7:00 AM to 7:00 PM, N_2 is the number-of-events that occur from 7:00 PM to 10:00 PM, and N_3 is the number-of-events that occur from 10:00 PM to 7:00 AM. In Japan and Korea, $WECPNL$ has been calculated from the average Maximum A-weighted Sound Level ($L_{A,max,av}$):

$$WECPNL = L_{A,max,av} + 10\log_{10}(N_1 + 3N_2 + 10N_3) - 27, \quad (\text{A.18})$$

where $L_{A,max,av} = \overline{ECPNL} - 13$. For more information, see (ICAO, 2008; Osada, 1991).

Noisiness Index (\overline{NI}) The Noisiness Index (\overline{NI}) was developed in South Africa from a study of aircraft noise exposure. \overline{NI} is an energy-based average sound level based on a tone-corrected A-weighted sound level for a 24-hour period. \overline{NI} is defined by the formula:

$$\overline{NI} = 10\log_{10} \sum_{i=1}^n 10^{\left(\frac{L_T + 10\log_{10}(t/t_0) + C + S}{10}\right)}, \quad (\text{A.19})$$

where t is the effective duration in seconds, t_0 is the total observation time in seconds, C is a time-of-day correction, S is a seasonal correction, and L_T is the Tone-Corrected A-weighted Sound Level calculated from one-third octave band data. The procedure for determining L_T is the same as that for determining the Tone-Corrected Perceived

Noise Level (*PNLT*). \overline{NI} has been approximated in additive-log format by the formula:

$$\overline{NI} = L_{Amax,av} + 10\log_{10}(N) - 39.4, \quad (\text{A.20})$$

where $L_{Amax,av}$ is the average Maximum A-weighted Sound Level and N is the total number-of-events. For more information, see (Schultz, 1982; van Niekerk and Muller, 1969).

Total Noise Load (*B*) The Total Noise Load (*B*) was developed in the Netherlands from a study of the response to aircraft noise around Schipol airport. *B* is an average maximum A-weighted sound level, but is not an energy-based average rating as $L_{Amax,av}$. *B* is defined by the formula:

$$B = 20\log_{10} \left(\sum_{i=1}^n w_i \times 10^{\left(\frac{L_i}{15}\right)} \right) - C, \quad (\text{A.21})$$

where L_i is the Maximum A-weighted Sound Level for the i^{th} event, w_i is the time-of-day weighting factor for the i^{th} event, n is the number-of-events, and $C = 157$ for a summation period of 1 year or $C = 106$ for a summation period of 24 hours. *B* has been approximated in additive-log format (for a 1-year period) by the formula:

$$B = \frac{4}{3}L_{Amax,av} + 20\log_{10}(N) - 157, \quad (\text{A.22})$$

where $L_{Amax,av}$ is the average Maximum A-weighted Sound Level and N is the total number-of-events. For more information, see (Galloway, 1974; Schultz, 1982).

Community Noise Equivalent Level (*CNEL*) The Community Noise Equivalent Level (*CNEL*) was developed for airports in California. *CNEL* is an energy-based average A-weighted sound level for a 24-hour period with time-of-day weightings. *CNEL* can be calculated with or without field measurements. The equation

for calculating $CNEL$ without field measurements is in additive-log format and is defined by the formula:

$$CNEL = \overline{SENEL} + 10\log_{10}(N_d + 3N_e + 10N_n) - 49.4, \quad (\text{A.23})$$

where \overline{SENEL} is the energy-based average A-weighted Single Event Noise Exposure Level ($SENEL_A$), N_d is the number-of-events that occur from 7:00 AM to 7:00 PM, N_e is the number-of-events that occur from 7:00 PM to 10:00 PM, and N_n is the number-of-events that occur from 10:00 PM to 7:00 AM. $SENEL_A$ is the California version of SEL_A . $SENEL_A$ and SEL_A are defined in Appendix A.2.2. For more information, see (California Department of Aeronautics, 1970; Schultz, 1982).

Day-Night Average Sound Level (DNL) The Day-Night Average Sound Level (DNL) was developed by the U.S. Environmental Protection Agency as a rating for all community noise. DNL is an energy-based average A-weighted sound level for a 24-hour period with a 10 dB penalty applied to nighttime aircraft events. The standard formulation of DNL is defined as an integration of energy over a 24-hour (86,400-second) period:

$$DNL = 10\log_{10} \left[\frac{1}{86,400} \left(\int_0^{54,000} 10^{L(t)/10} dt + \int_{54,000}^{86,400} 10^{L(t)+10/10} dt \right) \right], \quad (\text{A.24})$$

where $L(t)$ is the time-varying A-weighted sound level. When DNL is expressed in terms of the mean square sound pressure, Equation (A.24) becomes:

$$DNL = 10\log_{10} \left[\frac{1}{86,400} \left(\int_0^{54,000} \left(\frac{p_{A,bg}^2 + p_A^2(t)}{p_{ref}^2} \right) dt + \int_{54,000}^{86,400} 10 \left(\frac{p_{A,bg}^2 + p_A^2(t)}{p_{ref}^2} \right) dt \right) \right], \quad (\text{A.25})$$

where $p_{A,bg}^2$ is the A-weighted mean square sound pressure of the background noise which is hypothetically constant, $p_A^2(t)$ is the time-varying A-weighted mean square

sound pressure in addition to the constant background noise, and p_{ref} is the reference sound pressure of 20 μ PA. If the background noise is negligible, Equation (A.25) can be expressed:

$$DNL = 10 \log_{10} \left[\frac{1}{86,400} \left(\int_0^{54,000} \frac{p_A^2(t)}{p_{ref}^2} dt + \int_{54,000}^{86,400} 10 \frac{p_A^2(t)}{p_{ref}^2} dt \right) \right], \quad (\text{A.26})$$

Each integral is equivalent to the total energy contained in the aircraft events that occur within the specific time period. That is, each integral is equivalent to the energy-based average of the SEL_A 's of the individual events. Thus, each integral can be expressed as:

$$\int_0^{54,000} \frac{p_A^2(t)}{p_{ref}^2} dt = \sum_{i=1}^{N_d} 10^{SEL_{A,d,i}/10}, \quad (\text{A.27})$$

$$\int_{54,000}^{86,400} \frac{p_A^2(t)}{p_{ref}^2} dt = \sum_{j=1}^{N_n} 10^{SEL_{A,n,j}/10}, \quad (\text{A.28})$$

where $SEL_{A,d,i}$ is the SEL_A of the i^{th} daytime (7:00 AM to 10:00 PM) event, $SEL_{A,n,j}$ is the SEL_A of the j^{th} nighttime (10:00 PM to 7:00 AM) event, N_d is the number of daytime events, and N_n is the number of nighttime events. Each summation can be expressed as the product of the number-of-events and the energy-based average of the SEL_A 's for that specific time period:

$$\sum_{i=1}^{N_d} 10^{SEL_{A,d,i}/10} = N_d 10^{\overline{SEL_{A,d}/10}}, \quad (\text{A.29})$$

$$\sum_{j=1}^{N_n} 10^{SEL_{A,n,j}/10} = N_n 10^{\overline{SEL_{A,n}/10}}, \quad (\text{A.30})$$

where $\overline{SEL_{A,d}}$ is the energy-based average of the SEL_A 's for the daytime and $\overline{SEL_{A,n}}$ is the energy-based average of the SEL_A 's for the nighttime. Then, Equation (A.26) can be expressed as:

$$DNL = 10 \log_{10} \left[\frac{1}{86,400} \left(N_d 10^{\overline{SEL_{A,d}/10}} + 10 N_n 10^{\overline{SEL_{A,n}/10}} \right) \right]. \quad (\text{A.31})$$

Simplifying Equation (A.31) yields:

$$DNL = \overline{SEL_{A,d}} + \overline{SEL_{A,n}} + 10\log_{10}(N_d + 10N_n) + 10\log_{10}\left(\frac{1}{86,400}\right). \quad (\text{A.32})$$

Finally, DNL can be approximated in additive-log format by the formula:

$$DNL = \overline{SEL_A} + 10\log_{10}(N_d + 10N_n) - 49.4, \quad (\text{A.33})$$

where $\overline{SEL_A}$ is the energy-based average of $\overline{SEL_{A,d}}$ and $\overline{SEL_{A,n}}$. $\overline{SEL_A}$ is further described in Appendix A.2.2. For more information, see (U.S. Environmental Protection Agency, 1974).

A.2 Sound Level Metrics

Sound level metrics quantify the measurable attributes of sound. Acoustical predictors of annoyance include all of the factors that affect the audibility of sounds by people (Fidell and Pearsons, 2007). Examples of acoustical factors include the maximum or average sound pressure level, number-of-events, event rise time, event duration, impulsivity, time-of-day, source type, perceived loudness, fluctuation strength, roughness, sharpness, and tonality. In some cases, sound level metrics are directly correlated with annoyance measures and are used as annoyance models. In other cases, the sound level metrics are integrated into more complicated annoyance models, which can consist of multiple sound level metrics and/or nonacoustical variables. This section contains the definitions of a selection of sound level metrics. For more information, see (Bennett and Pearsons, 1981; Marquis-Favre et al., 2005b; Schultz, 1982).

A.2.1 Definition of the Perceived Noise Level (*PNL*) and Related Metrics

The Perceived Noise Level (*PNL*) was introduced by Karl Kryter as an improvement to Stevens' Loudness Level in the prediction of the judged noisiness of jet aircraft (Bennett and Pearsons, 1981). The development of *PNL* is given in (Kryter, 1960). The *PNL* is an instantaneous metric calculated in 0.5-second time increments. *PNL* is calculated from sound levels of an aircraft events measured in twenty-four one-third octave bands from 50 Hz to 10 kHz. The one-third octave band sound pressure levels are transformed to perceived noisiness (noys) using a process described in detail in Section A36.4.7 of (FAA, 2002). The twenty-four noy values are summed for the t^{th} time increment of the i^{th} aircraft event by:

$$N_{i,t} = n_{max} + 0.15 \left[\sum_{j=1}^{24} (n_{i,j}) - n_{max} \right], \quad (\text{A.34})$$

where $N_{i,t}$ is the total perceived noisiness for the t^{th} time increment of the i^{th} aircraft event, $n_{i,j}$ is the noy value for the j^{th} one-third octave band, and n_{max} is the maximum of the twenty-four values of n_i . The *PNL* of the i^{th} aircraft event for the t^{th} time increment is calculated by:

$$PNL_{i,t} = 40 + \frac{10 \log_{10} N_{i,t}}{\log_{10}(2)}. \quad (\text{A.35})$$

The maximum *PNL* of the i^{th} event ($PNL_{max,i}$) is the maximum of the $PNL_{i,t}$ values for the i^{th} event. From the PNL_{max} of all aircraft events, the energy-based average of the PNL_{max} ($PNL_{max,av}$) is calculated by:

$$PNL_{max,av} = 10 \log_{10} \left(\frac{1}{m} \sum_{i=1}^m 10^{(PNL_{max,i}/10)} \right), \quad (\text{A.36})$$

where m is the number of aircraft events. Several other sound level metrics are based on the *PNL*, which include the effects of pure tones and duration on the perceived noisiness of aircraft noise. The Tone-Corrected Perceived Noise Level (*PNLT*) in-

cludes the effect of pure tones on the perceived noisiness (Kryter and Pearsons, 1965). The Effective Perceived Noise Level (*EPNL*) was developed to account for duration effects on perceived noisiness, and is derived from a summation of *PNLT* values. The calculation methods of *PNL*, *PNLT*, and *EPNL* are given in (FAA, 2002).

A.2.2 Definition of the Sound Exposure Level (*SEL_A*) and Related Metrics

The Sound Exposure Level (*SEL_A*) was not developed as an official sound level metric, but was regarded as useful for the computation of other energy-based average sound level ratings, e.g., *L_{Aeq}*, *DNL*, and *CNEL* (Schultz, 1982). The *SEL_A* for an aircraft event is the A-weighted sound level that if maintained constant for one second, imparts the same A-weighted acoustic energy as a time-varying event. *SEL_A* is defined by the formula:

$$SEL_A = 10 \log_{10} \int_{-\infty}^{\infty} \frac{p_A^2(t)}{p_{ref}^2} dt / \tau_{ref}, \quad (\text{A.37})$$

where $p_A(t)$ is the time-varying A-weighted sound pressure of the event, p_{ref} is the reference sound pressure of 20 μPA , and τ_{ref} is the reference time of one second. In practice, *SEL_A* is normally integrated over the time period within 10 dB of the maximum A-weighted sound level of the event:

$$SEL_A = 10 \log_{10} \int_{t_1}^{t_2} 10^{L_A(t)/10} dt, \quad (\text{A.38})$$

where $L_A(t)$ is the time-varying A-weighted sound level of the event, and the time period of t_1 to t_2 is the time interval during which L_A is within 10 dB of the maximum level. From the *SEL_A* of all aircraft events to which a person is exposed, the energy-based average of the *SEL_A* (*SEL_{A,av}*) is calculated by:

$$SEL_{A,av} = \overline{SEL_A} = 10 \log_{10} \left(\sum_{i=1}^m 10^{(SEL_{A,i}/10)} \right), \quad (\text{A.39})$$

where m is the number of aircraft events. The Single Event Noise Exposure Level ($SENEL_A$) is a California variant of the SEL_A . For $SENEL_A$, the integration of (A.38) is limited to the period in which the sound level exceeds a specified threshold level, if this time period is shorter than t_1 to t_2 . The energy-based average of the $SENEL_A$ ($SENEL_{av}$) can be calculated from:

$$SENEL_{av} = \overline{SENEL} = 10 \log_{10} \left(\sum_{i=1}^m 10^{(SENEL_i/10)} \right). \quad (\text{A.40})$$

APPENDIX B. VARIANCE ESTIMATION

In the analysis of survey data, it is necessary to quantify the precision of statistics (measured by variance) calculated from the data in order to define ranges of confidence for the statistics. The regular variance or standard deviation formulae found in statistics texts usually relate only to unrestricted sampling (simple random sampling with replacement). Sampling “with replacement” means that after each sample is selected, it is not prevented from being sampled again. That is, the population size from which samples are selected is never decreased and samples can be selected more than once. Normally in practice, sampling is performed without replacement so that once a sample is selected, it cannot be selected again. This practice does not completely invalidate the unrestricted sampling formulae. Sampling without replacement introduces a correction factor into the unrestricted sampling formulae called the finite population correction factor, or FPC. However, if the population size is large relative to the size of the collected sample, the FPC can be ignored, and it normally is in practice. With sample designs more complex than simple random sampling such as multistage clustered sample designs with stratification and clustering, the unrestricted formulae are not valid. The use of the unrestricted formulae with data resulting from a complex sample design may yield calculations that overestimate, or more commonly, underestimate the true variance (Fields, 1980; Kalton, 1983; Sul Lee et al., 1989).

Stratification and clustering are different methods of grouping a population prior to sampling. Either or both can be applied in surveys. 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. With clustering, a population of people is grouped into clusters (usually geographic areas), which become first-stage units for sampling in a multistage clustered sample

design. Since these sampling units are first-stage groups of the population, they are referred to as “primary sampling units”. In multistage clustered sampling, only a subset of the clusters is included in the survey sample.

Clustering leads to a loss of precision compared to a simple random sample of the same size whenever the intraclass correlation coefficient is positive. The intraclass correlation coefficient is a measure of the homogeneity of surveys responses within clusters (Kish, 1965, sec. 5.4). A negative intraclass correlation coefficient indicates that the clusters are more internally heterogeneous than if they were formed randomly. Randomly-formed clusters have an intraclass correlation coefficient of 0. In practice, negative intraclass correlation coefficients occur rarely. This is because clusters are usually geographically-based, and respondents that are physically close to one another tend to have similar characteristics. It is more common for clusters to be more internally homogeneous than if they were randomly formed. The unrestricted variance formulae tend to substantially underestimate the variability (overestimate the precision) of estimates based on multistage clustered samples.

The calculation of a variance estimate is not only complicated by sample design, but also by the form of the statistic under scrutiny (Wolter, 2003). Formulae that estimate the simple random sample variance of statistics such as the mean of a variable or a regression coefficient estimate from a bivariate regression model are commonly encountered, but for other statistics such as the median, odds ratio, or a ratio of regression coefficients, there are no formulae. Certain statistics require nonstandard variance estimation techniques.

Two methods of variance estimation will be described: Jackknife repeated replication and Taylor series linearization. These methods are appropriate to use with data collected from a complex sample design in which the entire population of potential survey respondents was stratified, clusters of respondents were formed, a subset of clusters were chosen, and a group of respondents were sampled from each chosen cluster.

B.1 Jackknife Repeated Replication

Jackknife repeated replication is based on replicated sampling, and referred to as a pseudo-replication technique. Pseudo-replication is an analytical strategy applied to collected survey data. With pseudo-replication, a data set is divided into random groups (called replicates) for analysis. If stratification was used, each stratum must be represented in each replicate. Jackknife repeated replication and other methods based on replicated sampling have been applied in past noise annoyance studies (Bullen and Hede, 1986b; Dempsey et al., 1983; Fields and Walker, 1982).

Details of jackknife repeated replication can be found in (Kish and Frankel, 1974; Sul Lee et al., 1989). With the jackknife method, a statistic of interest is calculated in subsamples (called replicates) of the entire sample. In practice, replicates can be created randomly. However, a practical rule is to form the replicates in a way that each replicate has the same sample design as the parent sample. Since such detailed information on the actual sample design is rarely available, each replicate is usually formed by removing one primary sampling unit from the sample. In a multistage clustered sample design, a primary sampling unit is a first-stage grouping of the population, and is usually geographically-based. Sampling beyond the primary sampling stage is normally disregarded and is predicated on the fact that the sampling variance can be approximated adequately from the variation between primary sampling units when the primary sampling stage sampling fraction is small (Sul Lee et al., 1989). The jackknife can also be applied in the case of simple random sampling. With simple random sampling, replicates are formed by removing one sample at a time from the full sample.

The estimate of the jackknife variance of a statistic also has an associated standard error. The greater the number of replicates (a function of the number of primary sampling units), the smaller the standard error of the calculation and the more precisely the variance of the statistic can be estimated. A rule of thumb is to use around thirty replicates.

In the case of removing primary sampling units from strata to create replicates, there must be at least two primary sampling units in each stratum so that each stratum is represented in each replicate. The maximum possible number of replicates is equal to the number of primary sampling units. To reduce computation, as little as one replicate per stratum can be created, but the greater the number of replicates, the greater the degrees of freedom, and the smaller the variance estimate. The r^{th} replicate is created by removing all respondents from the j^{th} primary sampling unit in stratum i_d (the “donor” stratum). Each replicate has an associated jackknife coefficient:

$$\alpha_r = \frac{p_i - 1}{p_i}, \quad (\text{B.1})$$

where α_r is the jackknife coefficient of the r^{th} replicate and p_i is the number of primary sampling units in the i^{th} stratum. In this case, all replicates within the i^{th} stratum are applied the same jackknife coefficient. Equation (B.1) is standard with the use of the jackknife and equally-weights primary sampling units in a stratum. Another option is to equally-weight persons within a stratum, since primary sampling units can differ in the number of persons. To equally-weight persons, the jackknife coefficient becomes:

$$\alpha_r = \frac{n_r}{n_T}, \quad (\text{B.2})$$

where n_r is the number of samples in the r^{th} replicate and n_T is the total number of samples. In this case, the jackknife coefficient changes for each replicate. In either case of constructing the jackknife coefficient, the sampling weights of the respondents in the donor stratum must be re-weighted to maintain the presence of the stratum in the sample. For respondents in all strata other than the donor stratum, the replicate weights are equal to the original sampling weights. For respondents in the removed primary sampling unit, which is not contained in the replicate, the replicate weight

equals 0. That is, the replicate weight for the k^{th} member of the j^{th} primary sampling unit in the i^{th} stratum is:

$$w_{ijk}^r = \begin{cases} w_{ijk} & \text{if } i \neq i_d \\ \frac{w_{ijk}}{\alpha_r} & \text{if } i = i_d \text{ and primary sampling unit } j \text{ is included in replicate } r \\ 0 & \text{if } i = i_d \text{ and primary sampling unit } j \text{ is not included in replicate } r \end{cases} \quad (\text{B.3})$$

where w_{ijk}^r is the replicate weight and w_{ijk} is the original sampling weight. Let Θ represent the population parameter that the variance is to be estimated for. Then $\hat{\Theta}$ is the estimate from the full sample and $\hat{\Theta}_r$ is the estimate from the r^{th} replicate. The jackknife variance estimate of $\hat{\Theta}$ is calculated by:

$$\hat{\sigma}^2(\hat{\Theta}) = \sum_{r=1}^R \alpha_r \left(\hat{\Theta}_r - \bar{\hat{\Theta}} \right)^2, \quad (\text{B.4})$$

where R is the number of replicates and $\bar{\hat{\Theta}}$ is the mean of the replicates estimates' of Θ .

B.2 Taylor Series Linearization

Taylor series linearization is a method of representing a mathematical function as an infinite sum of terms based on the function's derivatives which are evaluated at a single point. Taylor series linearization is also referred to as the delta method (Kalton, 1983, p. 44) and the propagation of variance (Kish, 1965, p. 583). Taylor series linearization can be applied to collected survey data, as jackknife repeated replication.

The Taylor series of a real or complex function $f(x)$ about a point a is the power series:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^n(a)}{n!} (x - a)^n, \quad (\text{B.5})$$

where $n!$ is the factorial of n and $f^n(a)$ is the n^{th} derivative of f evaluated at the point a . The second-order Taylor series is:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots \quad (\text{B.6})$$

For a function of two random variables $f(X, Y)$ the first-order Taylor series about $\Theta = (E[X], E[Y])$ is:

$$f(X, Y) = f(\Theta) + \frac{f'_x}{1!}(\Theta)(X - \Theta_x) + \frac{f'_y}{1!}(\Theta)(Y - \Theta_y) + O(n^{-1}), \quad (\text{B.7})$$

where $O(n^{-1})$ contains the second-order and higher terms. Taking the expected value of (B.7) results in:

$$\begin{aligned} Ef(X, Y) &= f(E[X], E[Y]) \\ &= f(\Theta) + f'_x(\Theta)(E[X] - \Theta_x) + f'_y(\Theta)(E[Y] - \Theta_y) + O(n^{-1}) \\ &= f(\Theta) + O(n^{-1}), \end{aligned} \quad (\text{B.8})$$

because $E[X] = \Theta_x$ and $E[Y] = \Theta_y$. The variance of a function of random variables can be expressed as:

$$\sigma^2[f(X, Y)] = E [(f(X, Y) - E[f(X, Y)])^2]. \quad (\text{B.9})$$

Substituting Equation (B.7) and Equation (B.8) into Equation (B.9) results in:

$$\begin{aligned} \sigma^2[f(X, Y)] &= E \left[\left(f'_x(\Theta)(E[X] - \Theta_x) + f'_y(\Theta)(E[Y] - \Theta_y) \right)^2 \right] + O(n^{-1}) \\ &\approx f'_x(\Theta)^2 \sigma^2(X) + 2f'_x(\Theta)f'_y(\Theta)\sigma(X, Y) + f'_y(\Theta)^2 \sigma^2(Y), \end{aligned} \quad (\text{B.10})$$

where $\sigma^2(X) = E [(E[X] - \Theta_x)^2]$ is the variance of the random variable X , $\sigma^2(Y) = E [(E[Y] - \Theta_y)^2]$ is the variance of the random variable Y , and

$\sigma(X, Y) = [(E[X] - \Theta_x)(E[Y] - \Theta_y)]$ is the covariance between the random variables. When $f(X, Y) = \frac{X}{Y}$ is a ratio of random variables, $f'_x = Y^{-1}$ and $f'_y = -\frac{X}{Y^2}$. Then,

$$\begin{aligned}\sigma^2\left(\frac{X}{Y}\right) &= \frac{1}{E^2[Y]}\sigma^2(X) - \frac{2E[X]}{E^3[Y]}\sigma(X, Y) + \frac{E^2[X]}{E^4[Y]}\sigma^2(Y) \\ &= \frac{E^2[X]}{E^2[Y]}\left(\frac{\sigma^2(X)}{E^2[X]} + \frac{\sigma^2(Y)}{E^2[Y]} - \frac{2\sigma(X, Y)}{E[X]E[Y]}\right).\end{aligned}\tag{B.11}$$

Equation (B.11) can be used to estimate the variance of a ratio of regression coefficients. The general multiple linear regression model can be expressed as:

$$z_i = \beta_x x_i + \beta_y y_i + \beta_0 + \epsilon_i,\tag{B.12}$$

where:

z_i is a known criterion variable for all i ,

x_i and y_i are known predictor variables for all i ,

β_x , β_y , and β_0 are regression coefficients,

ϵ_i are independent and zero-mean normally distributed errors.

The general multiple linear regression function estimate can be expressed as:

$$\hat{z}_i = b_x x_i + b_y y_i + b_0,\tag{B.13}$$

where \hat{z}_i is the predicted criterion variable, and b_x , b_y , and b_0 are the estimates of the β 's. Then, Equation (B.11) can be expressed as:

$$\begin{aligned}\sigma^2\left(\frac{\beta_x}{\beta_y}\right) &= \frac{1}{b_y^2}\sigma^2(\beta_x) - \frac{2b_x}{b_y^3}\sigma(\beta_x, \beta_y) + \frac{b_x^2}{b_y^4}\sigma^2(\beta_y) \\ &= \frac{b_x^2}{b_y^2}\left(\frac{\sigma^2(\beta_x)}{b_x^2} + \frac{\sigma^2(\beta_y)}{b_y^2} - \frac{2\sigma(\beta_x, \beta_y)}{b_x b_y}\right),\end{aligned}\tag{B.14}$$

where $b_x = E[\beta_x]$ and $b_y = E[\beta_y]$.

Because the β 's of a regression model are unknown, $\sigma^2(\beta_x)$, $\sigma^2(\beta_y)$, and $\sigma(\beta_x, \beta_y)$ are unknown and must be estimated. Taylor series linearization can be applied in simple random sampling and also in multistage clustered sampling. The method of calculating estimates of the variances and covariance of the regression coefficients will differ depending on the sampling method. With simple random sampling, the variances and covariance estimates can be found in the regression coefficient variance-covariance matrix calculated via multiple linear regression techniques, which is standard output in many computational packages. With multistage clustered sampling in which the population was grouped into clusters prior to sampling, the process of calculating estimates of $\sigma^2(\beta_x)$, $\sigma^2(\beta_y)$, and $\sigma(\beta_x, \beta_y)$ is slightly more complicated and requires variance component estimation via a random effects single-factor analysis of variance.

Variance component estimation for the case of multistage clustered sampling involves partitioning the variance of the residuals resulting from a multiple linear regression analysis into the variability that can be accounted for by primary sampling unit (the variance “between” primary sampling units) and the remaining variability not accounted for by primary sampling unit (the variance “within” primary sampling units). The residuals are the differences between the actual annoyance responses and the annoyance predictions from the estimated regression model. The between primary sampling unit variance component estimate is an estimate of the true variance of the average annoyances from the primary sampling units (σ_b^2). The within primary sampling unit variance component estimate is an estimate of the true variance of the annoyance of persons that is not accounted for by differences in primary sampling unit (σ_w^2).

Estimates of σ_w^2 and σ_b^2 are calculated from a random effects single-factor analysis of variance on the residuals. Standard output in analysis of variance is the treatment sum of squares (*SSTR*) and the error sum of squares (*SSE*). The treatment sum of squares is a measure of the extent of differences between factor level means, where each factor level is a primary sampling unit and each mean is a mean of residuals. The

more factor level means differ, the larger $SSTR$ becomes. The error sum of squares is a measure of the random variation of the observations (residuals) around the factor level means. The more variation among the observations within each factor level, the larger SSE becomes. $SSTR$ is expressed by:

$$SSTR = \sum_{i=1}^{\alpha} n_i (\bar{z}_i - \bar{z}_{..})^2, \quad (\text{B.15})$$

where α is the number of factor levels, n_i is the number of observations in the i^{th} factor level, \bar{z}_i is the mean of the observations in the i^{th} factor level, and $\bar{z}_{..}$ is the overall mean of the observations. SSE is expressed by:

$$SSE = \sum_{i=1}^{\alpha} \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2, \quad (\text{B.16})$$

where z_{ij} is the j^{th} observation in the i^{th} factor level. Both $SSTR$ and SSE can be divided by their associated degrees of freedom to calculate the treatment mean square ($MSTR$) and error mean square (MSE). Mean squares are average squared deviations, and are estimates of variances. $MSTR$ and MSE are expressed by:

$$MSTR = \frac{SSTR}{\alpha - 1}, \quad (\text{B.17})$$

$$MSE = \frac{SSE}{n_T - \alpha}, \quad (\text{B.18})$$

where n_T is the total number of observations. $\hat{\sigma}_w^2$ and $\hat{\sigma}_b^2$ are functions of the mean squares:

$$\hat{\sigma}_w^2 = MSE, \quad (\text{B.19})$$

$$\hat{\sigma}_b^2 = \frac{MSTR - MSE}{n'}, \quad (\text{B.20})$$

where:

$$n' = \frac{1}{\alpha - 1} \left[\sum_{i=1}^{\alpha} n_i - \frac{\sum_{i=1}^{\alpha} n_i^2}{\sum_{i=1}^{\alpha} n_i} \right]. \quad (\text{B.21})$$

Then, estimates of the variances and covariance of the regression coefficients can be calculated using the following equations:

$$\hat{\sigma}^2(\beta_x) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (\sigma_y^2), \quad (\text{B.22})$$

$$\hat{\sigma}^2(\beta_y) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (\sigma_x^2), \quad (\text{B.23})$$

$$\hat{\sigma}(\beta_x, \beta_y) = \left(\frac{\hat{\sigma}_w^2}{n} + \frac{\hat{\sigma}_b^2}{\alpha} \right) \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (-\sigma_{xy}), \quad (\text{B.24})$$

where x and y represent predictor variable data, σ_x^2 is the variance of x , σ_y^2 is the variance of y , σ_{xy} is the covariance of x and y , n is the number of samples, and α is the number of primary sampling units. The estimate of Equation (B.14) can then be expressed as:

$$\hat{\sigma}^2 \left(\frac{\beta_x}{\beta_y} \right) = \frac{b_x^2}{b_y^2} \left(\frac{\hat{\sigma}^2(\beta_x)}{b_x^2} + \frac{\hat{\sigma}^2(\beta_y)}{b_y^2} - \frac{2\hat{\sigma}(\beta_x, \beta_y)}{b_x b_y} \right). \quad (\text{B.25})$$

Taylor series linearization is also useful in planning surveys (Fields, 1984). In the *planning*, not *analysis*, of a survey with a complex sample design involving clustering, Equations (B.15) to (B.21) are not used, and estimates of σ_w^2 and σ_b^2 ($\hat{\sigma}_w^2$ and $\hat{\sigma}_b^2$) are required for Equations (B.22), (B.23) and (B.24). The variance component estimation results from similar surveys can be used to guide the choice of estimates. In the event that simple random sampling without the use of clustering is planned for the sample design, the Taylor formulae becomes less complicated, as there are no primary sampling units in the sample design, and $\hat{\sigma}_w^2$ and $\hat{\sigma}_b^2$ are not required. An example of Taylor series linearization in planning a survey with a simple random sample design can be found in Chapter 5, Section 5.3.2. In this case, Equations (B.22), (B.23) and (B.24) reduce to:

$$\hat{\sigma}^2(\beta_x) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (\sigma_y^2), \quad (\text{B.26})$$

$$\hat{\sigma}^2(\beta_y) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (\sigma_x^2), \quad (\text{B.27})$$

$$\hat{\sigma}(\beta_x, \beta_y) = \frac{\hat{\sigma}_e^2}{n} \left(\frac{1}{\sigma_x^2 \sigma_y^2 - (\sigma_{xy})^2} \right) (-\sigma_{xy}), \quad (\text{B.28})$$

where $\hat{\sigma}_e^2$ is the variance of the residuals and an estimate of the variance of the regression error terms.

APPENDIX C. MULTICOLLINEARITY

Multicollinearity is the problem of interrelation, or covariation, of multiple predictor variables in a data set. Multicollinearity is not equivalent to the simple correlation between two variables. A multicollinear relation may involve more than two variables, while a correlation concerns two variables. The magnitude of a correlation indicates the magnitude of a linear dependence between two variables. In the same way, the degree of the extent of multicollinearity indicates the degree of a linear dependency among two or more variables. Correlation is two-dimensional multicollinearity. In the bivariate case, high correlation is indicative of multicollinearity. The problem of multicollinearity is common to non-experimental or observational studies in which the values of variables are not under direct control by the researcher; i.e., they cannot be assigned to survey participants, only observed. The problem of multicollinearity is well-documented (Belsley, 1991a; Farrar and Glauber, 1967; Gunst, 1983; Hocking and Pendleton, 1983; Kumar, 1975; Mason and Perreault, Jr., 1991; Stewart, 1987).

In practice, it is not common for perfect multicollinearity to exist among variables, unless by error. For example, either multiple inclusions of one variable or including a variable that is a simple sum of other variables would both lead to perfect multicollinearity. Although perfect multicollinearity is not likely, imperfect multicollinearity is not. In observational studies, multicollinearity should not be conceived as something that either “exists” or “does not exist”. Rather, multicollinearity exists in degrees (Berry and Feldman, 1985). The consequences of multicollinearity are:

1. Estimated regression coefficients tend to have large sampling variability, i.e., inflated variances.
2. The regression/model sum of squares cannot be uniquely decomposed into separate and non-overlapping contributions ascribed to the predictor variables as reflecting their effect in reducing the total variation in the criterion variable.

3. The interpretation of a regression coefficient as a marginal effect is not applicable.
4. The least-squares solutions have increased sensitivity to changes in the underlying data.

Multicollinearity is a problem of deficient data. It is always conceptually possible to increase the value of one variable in an estimated regression equation while holding the other constant. However, there may be no information about the result of such a manipulation in the estimation data (Chatterjee and Hadi, 2006). Likewise, the classical interpretation of the regression coefficient of the effect of one variable while holding the other variables constant is not applicable. Multicollinearity does not reduce the goodness-of-fit of a model to data but introduces problems in determining the separate effects of the predictor variables. If the goal of a study is to derive a model that will account for a high proportion of the variance in a criterion variable, then multicollinearity is not an issue. Multicollinearity becomes a problem when the goal is to determine the separate effects of predictor variables on a criterion variable. If multicollinearity is a problem, there are strategies to address it:

1. Increase the sample size.
2. Increase the ranges of the predictor variables.

Although these strategies will reduce the variance estimates of the regression coefficient estimates, they address the effect (increased variances), not the cause (multicollinearity). Addressing the cause involves collecting noncollinear data.

Other strategies that are not recommended are:

1. Drop a variable involved in a multicollinear relation: One of the simplest strategies to reduce multicollinearity is to drop one of the collinear variables. The strategy sidesteps the collinearity problem and introduces a new complication. Unless the true coefficient of the dropped variable in the model is zero, the model will be misspecified, resulting in biased estimates of some coefficients (Mason and

Perreault, Jr., 1991). The problem of multicollinearity is one of insufficient data, not too much data.

2. Mean-centering: Mean-centering variables involves subtracting from each variable its mean value. Mean-centering as applied to multicollinearity is addressed in detail in multiple references (Belsley, 1984, 1991a). Mean-centering typically masks the role of the constant term in any collinear relations. Although diagnostics that analyze the degree of multicollinearity in the data are seemingly improved by mean-centering, the diagnostics are concerned with the mean-centered variables, not the variables themselves. This is an issue of “structural interpretability” (Belsley, 1991a).

APPENDIX D. STANDARDIZED STORAGE FORMAT

In this appendix are given the column assignments for the standardized data set storage format. All historical data sets described in Chapter 3 were adopted to the format shown in Table D.1. A variable name with the asterisk superscript (*) indicates that the variable is included in archived data sets created by the Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (Dutch Organization for Applied Scientific Research) (Fields, 1997)

Table D.1. Column assignments for the standardized storage format.

Category	Variable	Name
Annoyance	Verbal 3	VRB3
Annoyance	Verbal 4	VRB4
Annoyance	Verbal 5	VRB5
Annoyance	Verbal 6	VRB6
Annoyance	Numerical 3	NUM3
Annoyance	Numerical 4	NUM4
Annoyance	Numerical 5	NUM5
Annoyance	Numerical 6	NUM6
Annoyance	Numerical 7	NUM7
Annoyance	Numerical 8	NUM8
Annoyance	Numerical 9	NUM9
Annoyance	Numerical 10	NUM10
Annoyance	Numerical 11	NUM11
Annoyance	Numerical 12	NUM12
Annoyance	Numerical, verbal endpoints 3	NUMVRB3
Annoyance	Numerical, verbal endpoints 4	NUMVRB4

Table D.1. Continued

Category	Variable	Name
Annoyance	Numerical, verbal endpoints 5	NUMVRB5
Annoyance	Numerical, verbal endpoints 6	NUMVRB6
Annoyance	Numerical, verbal endpoints 7	NUMVRB7
Annoyance	Numerical, verbal endpoints 8	NUMVRB8
Annoyance	Numerical, verbal endpoints 9	NUMVRB9
Annoyance	Numerical, verbal endpoints 10	NUMVRB10
Annoyance	Numerical, verbal endpoints 11	NUMVRB11
Annoyance	Numerical, verbal endpoints 12	NUMVRB12
Annoyance	Activity interference 1	ACTINT1
Annoyance	Activity interference 2	ACTINT2
Annoyance	Frequency 3	FREQ3
Annoyance	Frequency 4	FREQ4
Annoyance	Frequency 5	FREQ5
Annoyance	Frequency 6	FREQ6
Annoyance	Hearing the noise source	HEAR*
Annoyance	Open question about negative reactions to noise source	OPEN*
Annoyance	Disturbance frequency, startle	NSTRTL*
Annoyance	Disturbance frequency, vibra- tion	NVB*
Annoyance	Disturbance frequency, con- versation	NCONV*
Annoyance	Disturbance frequency, tele- phone talking	NPHONE*

Table D.1. Continued

Category	Variable	Name
Annoyance	Disturbance frequency, TV or radio listening	NTVR*
Annoyance	Disturbance frequency, sleep disturbance	NSLEEP*
Annoyance	Disturbance annoyance extent, startle	STARTL*
Annoyance	Disturbance annoyance extent, vibration	VIB*
Annoyance	Disturbance annoyance extent, conversation	CONV*
Annoyance	Disturbance annoyance extent, telephone talking	PHONE*
Annoyance	Disturbance annoyance extent, TV or radio listening	TVR*
Annoyance	Disturbance annoyance extent, waking up at night (or general waking up)	SLEEP1*
Annoyance	Disturbance annoyance extent, not being able to get to sleep at the beginning of the night	SLEEP2*
Annoyance	Disturbance annoyance extent, waking up in the morning	SLEEP3*
Annoyance	Public action taken against the source	ACTION*

Table D.1. Continued

Category	Variable	Name
Annoyance	Reacting by closing windows or reducing ventilation be- cause of noise	SHUT*
Annoyance	Extra 1	AEXTRA1
Annoyance	Extra 2	AEXTRA2
Annoyance	Extra 3	AEXTRA3
Annoyance	Extra 4	AEXTRA4
Annoyance	Extra 5	AEXTRA5
Annoyance	Extra 6	AEXTRA6
Annoyance	Extra 7	AEXTRA7
Annoyance	Extra 8	AEXTRA8
Annoyance	Extra 9	AEXTRA9
Annoyance	Extra 10	AEXTRA10
Annoyance	”Jump to column” column for additional data	AJUMP
Noise Metric - Maximum Level	$L_{Amax,av,70}$	LAMAXAV_LG70
Noise Metric - Maximum Level	$PNL_{max,av,80}$	PNLMAXAV_LG80
Noise Metric - Maximum Level	$PNLT_{max,av,80}$	PNLTMAXAV_LG80
Noise Metric - Energy Dose	$EPNL_{av}$	EPNLAV_LG
Noise Metric - Energy Dose	$SEL_{A,av}$	SELAV_LG
Noise Metric - Energy Dose	$SENEL_{A,av}$	SENELAV_LG
Noise Metric - Cumulative	$CNEL$	CNEL
Noise Metric - Cumulative	L_{eq}	LEQ24HR
Noise Metric - Cumulative	$L_{eq,d}$	LEQD

Table D.1. Continued

Category	Variable	Name
Noise Metric - Cumulative	$L_{eq,n}$	LEQN
Noise Metric - Cumulative	L_{Aeq}	LAEQ24
Noise Metric - Cumulative	$L_{Aeq,d}$	LAEQD
Noise Metric - Cumulative	$L_{Aeq,n}$	LAEQN
Noise Metric - Cumulative	$L_{Aeq,0607}$	L0607*
Noise Metric - Cumulative	$L_{Aeq,0719}$	L0719*
Noise Metric - Cumulative	$L_{Aeq,1922}$	L1922*
Noise Metric - Cumulative	$L_{Aeq,2223}$	L2223*
Noise Metric - Cumulative	$L_{Aeq,2306}$	L2306*
Noise Metric - Cumulative	$L_{Aeq,diff}$	DIFF*
Noise Metric - Cumulative	CNR	CNR
Noise Metric - Cumulative	DNL	DNL
Noise Metric - Cumulative	$DENL$	DENL
Noise Metric - Cumulative	NEF	NEF
Noise Metric - Cumulative	NNI	NNI
Noise Metric - Cumulative	$WECPNL$	WECPNL
Noise Metric - Statistical	L_{10}	L10
Noise Metric - Statistical	L_{20}	L20
Noise Metric - Statistical	L_{30}	L30
Noise Metric - Statistical	L_{40}	L40
Noise Metric - Statistical	L_{50}	L50
Noise Metric - Statistical	L_{60}	L60
Noise Metric - Statistical	L_{70}	L70
Noise Metric - Statistical	L_{80}	L80
Noise Metric - Statistical	L_{90}	L90
Noise Metric - Statistical	L_{100}	L100
Noise Metric	Extra 1	LEXTRA1

Table D.1. Continued

Category	Variable	Name
Noise Metric	Extra 2	LEXTRA2
Noise Metric	Extra 3	LEXTRA3
Noise Metric	Extra 4	LEXTRA4
Noise Metric	Extra 5	LEXTRA5
Noise Metric	Extra 6	LEXTRA6
Noise Metric	Extra 7	LEXTRA7
Noise Metric	Extra 8	LEXTRA8
Noise Metric	Extra 9	LEXTRA9
Noise Metric	Extra 10	LEXTRA10
Noise Metric	Extra 11	LEXTRA11
Noise Metric	Extra 12	LEXTRA12
Noise Metric	Extra 13	LEXTRA13
Noise Metric	Extra 14	LEXTRA14
Noise Metric	Extra 15	LEXTRA15
Noise Metric	Extra 16	LEXTRA16
Noise Metric	Extra 17	LEXTRA17
Noise Metric	Extra 18	LEXTRA18
Noise Metric	Extra 19	LEXTRA19
Noise Metric	Extra 20	LEXTRA20
Noise Metric	"Jump to column" column for additional data	LJUMP
Number Metric	N_{70}	NALAMAX70
Number Metric	$N_{PNL_{max,av,80}}$	NAPNLMAX80
Number Metric	$N_{PNLT_{max,av,80}}$	NAPNLTMAX80
Number Metric	$N_{EPNL_{av}}$	NEPNLAV
Number Metric	$N_{SEL_{A,av}}$	NSELAV
Number Metric	$N_{SENEL_{A,av}}$	SENELAV

Table D.1. Continued

Category	Variable	Name
Number Metric	N_{Leq}	NLEQ24HR
Number Metric	$N_{Leq,d}$	NLEQD
Number Metric	$N_{Leq,n}$	NLEQN
Number Metric	N_{LAeq}	NLAEQ24
Number Metric	$N_{LAeq,d}$	NLAEQD
Number Metric	$N_{LAeq,n}$	NLAEQN
Number Metric	$N_{LAeq,0607}$	N0607*
Number Metric	$N_{LAeq,0719}$	N0719*
Number Metric	$N_{LAeq,1922}$	N1922*
Number Metric	$N_{LAeq,2223}$	N2223*
Number Metric	$N_{LAeq,2306}$	N2306*
Number Metric	Extra 1	NEXTRA1
Number Metric	Extra 2	NEXTRA2
Number Metric	Extra 3	NEXTRA3
Number Metric	Extra 4	NEXTRA4
Number Metric	Extra 5	NEXTRA5
Number Metric	Extra 6	NEXTRA6
Number Metric	Extra 7	NEXTRA7
Number Metric	Extra 8	NEXTRA8
Number Metric	Extra 9	NEXTRA9
Number Metric	Extra 10	NEXTRA10
Number Metric	Extra 11	NEXTRA11
Number Metric	Extra 12	NEXTRA12
Number Metric	Extra 13	NEXTRA13
Number Metric	Extra 14	NEXTRA14
Number Metric	Extra 15	NEXTRA15
Number Metric	Extra 16	NEXTRA16

Table D.1. Continued

Category	Variable	Name
Number Metric	Extra 17	NEXTRA17
Number Metric	Extra 18	NEXTRA18
Number Metric	Extra 19	NEXTRA19
Number Metric	Extra 20	NEXTRA20
Number Metric	”Jump to column” column for additional data	NJUMP
Sample Design	Study identification number	CATALOG*
Sample Design	Country	COUNTRY*
Sample Design	City	CITY
Sample Design	Airport	AIRPORT
Sample Design	Cluster #	CLUSTER
Sample Design	Stratum #	STRATUM
Sample Design	Primary sampling unit #	PSU*
Sample Design	Phase of study if interviews were repeated in the same area	PHASE*
Sample Design	Respondent ID	PERSON*
Sample Design	Source	SOURCE*
Sample Design	Distance to source	DIST*
Sample Design	Number of people represented by point or sample weight	WEIGHT*
Sample Design	Longitude of study area	LONG
Sample Design	Longitude: actual or derived?	LONGMTHD
Sample Design	Latitude of study area	LAT
Sample Design	Latitude: actual or derived?	LATMTHD
Sample Design	Elevation of study area	ELEV
Sample Design	Elevation: actual or derived?	ELEVMTD

Table D.1. Continued

Category	Variable	Name
Sample Design	"Jump to column" column for additional data	SDJUMP
Nonacoustical variables	Age	AGE*
Nonacoustical variables	Sex	SEX*
Nonacoustical variables	Race	RACE
Nonacoustical variables	Marital status	MARSTAT
Nonacoustical variables	Education 1 , ISCD	EDUCLEV*
Nonacoustical variables	Education 2, number years	EDUCLEN*
Nonacoustical variables	Occupation 1, work status	OCCSTAT
Nonacoustical variables	Occupation 2, job ID#	OCCJOB
Nonacoustical variables	Occupation 3, hours per day	OCCHRS
Nonacoustical variables	Occupation 4, shift work?	OCCSHIFT
Nonacoustical variables	Occupation 5, noise environment at work	OCCNOISE
Nonacoustical variables	Time at home	TIMEHOME
Nonacoustical variables	Household type (construction type)	STRUCT
Nonacoustical variables	Window insulation condition	INSULAT*
Nonacoustical variables	Bedroom orientation	BEDOR
Nonacoustical variables	Kitchen orientation	KITCHOR
Nonacoustical variables	Living room/den orientation	DENOR
Nonacoustical variables	% daytime open windows	POPEND
Nonacoustical variables	% nighttime open windows	POPENN
Nonacoustical variables	Fear associated with noise source	FEAR*
Nonacoustical variables	Sensitivity to noise generally	SENSI*

Table D.1. *Continued*

Category	Variable	Name
Nonacoustical variables	User of the transportation source that is the source of the noise	USER*
Nonacoustical variables	Employee of the noise source (or the noise source industry) lives in the house	EMPLOY*
Nonacoustical variables	Social status (Scaled from 0-100)	SOCIAL*
Nonacoustical variables	Income	INCOME*
Nonacoustical variables	Number of residents in household	NHOUSE*
Nonacoustical variables	Home ownership	OWNER*
Nonacoustical variables	Type of dwelling (single, attached, multistory)	DWELL*
Nonacoustical variables	Pulse rate	PULSER
Nonacoustical variables	Blood pressure	BLOODP
Nonacoustical variables	Knowledge of enquiry	KNOW
Nonacoustical variables	Extra 1	NVEXTRA1
Nonacoustical variables	Extra 2	NVEXTRA2
Nonacoustical variables	Extra 3	NVEXTRA3
Nonacoustical variables	Extra 4	NVEXTRA4
Nonacoustical variables	Extra 5	NVEXTRA5
Nonacoustical variables	Extra 6	NVEXTRA6
Nonacoustical variables	Extra 7	NVEXTRA7
Nonacoustical variables	Extra 8	NVEXTRA8
Nonacoustical variables	Extra 9	NVEXTRA9
Nonacoustical variables	Extra 10	NVEXTRA10

Table D.1. Continued

Category	Variable	Name
Nonacoustical variables	"Jump to column" column for additional data	NVJUMP

APPENDIX E. FLIGHT TRACK CREATION

In this appendix two methods for manually creating flight tracks for the calculation of single-event acoustical data in INM 7.0 are described. In some cases, it might be possible to collect INM flight track files from prior FAA Part 150 studies. If INM flight track data can be collected, and the flight tracks are representative of current traffic, there is no need to manually construct flight tracks.

The INM flight track files are:

1. *track.dbf*: contains the names of the flight tracks
2. *trk_segs.dbf*: contains the grid points for all flight tracks

The format of each flight track file can be found in Appendix E of (He et al., 2007). In the case that INM flight track files cannot be collected, or flight tracks have changed, it will be necessary to manually construct flight tracks. Flight tracks can be constructed either by the use of flight track maps or by the use of actual flight tracks data.

E.1 Flight Track Maps

This is the first of two methods to manually create flight tracks for an airport. Adobe Photoshop and MATLAB are both used to create the flight tracks. The instructions given for Photoshop apply to version 12.0.1. Photoshop is used to create image files that a function coded in MATLAB can process. The MATLAB function calculates grid points for the flight tracks in the image files. After calculation, the grid points and flight track name for each flight track must be manually input into the INM flight track files. General instructions along with an actual example in italics will be given. The example provided is for Indianapolis International Airport flight tracks.

1. Obtain flight track maps: The flight track maps should show one or more flight tracks from an aerial view. The maps should also show geographical information such as roads, train tracks, bodies of water, etc. MATLAB will be used to establish a reference grid for the flight track maps; thus, flight track maps that show different geographical areas must be handled separately. The following instructions apply for a group of flight track maps that show the same geographical area.

Flight track maps were found in the Part 150 study, Exhibits 4-11. Exhibits 4-7 were addressed as one group; Exhibits 8-11 were a second group.

2. Create .pdf files of the flight track maps.
3. Open a flight track map in Photoshop: This will open a .pdf import utility window. Change the *Crop To:* setting to *Media Box* which imports the full document. The resolution can be specified in pixel/in or pixels/cm. The *Constrain Proportions* box should be checked. The program will offer a default recommendation for the resolution. Choosing a higher resolution (smaller pixel size) will allow finer tunings in the Photoshop grid creation process.

The recommended and chosen resolution setting was 300 pixels/in. In retrospect, a higher resolution/smaller pixel size would have been desirable. This will be described further in point 4(c)ii.

4. Establish a common coordinate grid among INM and the flight track maps: Grids will be created in both INM and Photoshop; the grids will be equivalent. The grids will be aligned in order to connect grid points in INM to points on the flight track maps. Within INM, grids can be specified in either nautical miles (nmi) or kilometers (km). 1 nautical mile is equivalent to 1852 meters. There is no requirement for the unit; choose whichever is preferable. However, if a distance on the flight track map scale is given in either nmi or km, then the INM grid spacing should be specified likewise.
 - (a) Create a grid in INM. The spacing of the INM grid should be chosen to be resolute enough that misalignments between the INM and Photoshop grids

can be noticed, but not overly-resolute to visually complicate the process.

A grid spacing of 0.5 nmi was applied in INM.

- (b) Import TIGER/Line files for the geographic area into INM: The TIGER/Line files will assist in aligning the INM and Photoshop grids. TIGER/Line files for INM can be found at <http://www.census.gov/geo/www/tiger/tgrcd108/tgr108cd.html>. Instructions for importing TIGER/Line files can be found in Section 3.5.5 of (He et al., 2007). The TIGER/Line files described here are different from those described in Chapter 5, Section 5.2.4.
- (c) Create a grid in Photoshop: It is necessary to know how a distance measurable on the flight track map relates to the units of the INM grid (either nmi or km). This requires the ruler tool in Photoshop. The measurement unit of the ruler should be set to pixels in *Preferences*. Use the ruler to find the length of the flight track map scale in pixels. Calculate the number of pixels per flight track map scale unit. Find from this the of number of pixels per INM grid unit. Use the number of INM grid units that the grid lines are separated by to calculate the required Photoshop grid spacing, z .

305 pixels was equal to the flight track map scale length of 6000 feet; this was equal to approximately 154.43 pixels per 0.5 nmi (the INM grid spacing), or roughly, 154 pixels per 0.5 nmi.

- i. Select *File*, then *New*: This creates a new project.
- ii. Specify the size of the new project template: Change the specification units to pixels. The resolution of the template should be the same as the imported flight track map; the size of the template should be square and should be set equal to z .

A new template of 154x154 pixels was created since only integer-valued dimensions are allowed; a smaller pixel size from using a higher resolution would have helped with precise sizing of the template.

- iii. Apply an inside-border to the template: Select the template (press **Ctrl+A**). Select *Edit*, then *Stroke*, set the color to be black, the border

- to be applied to the inside boundary, and the width to be two pixels; thus, each grid line in Photoshop will be four pixels wide. Define the template as a pattern (select *Edit*, then *Define Pattern*).
- iv. Create a new layer and fill it with the pattern: In the project containing the imported flight map image, create a new layer (select *Layer*, then *New*, then *Layer*) and fill it with the grid pattern (select *Edit*, then *Fill*, then choose the grid pattern that was just created). This layer will be the “grid” layer.
- (d) Align the Photoshop grid with the INM grid: Make the grid layer the active layer by clicking on it in the list of layers. Click the *Move* tool (or press **V** on the keyboard); use the arrow keys to move the grid layer one pixel at a time (press and hold the **Ctrl** key to increase the shift to ten pixels at a time). The grids can be aligned from a comparison of the placement of each grid relative to the physical map elements (roads, train tracks, bodies of water, etc.).
- (e) Mark the origin of the INM grid in Photoshop: Once the grids are aligned, the origin (0,0) point in INM can be located on the Photoshop grid. In Photoshop, create a new layer, and create a single-pixel mark in the new “origin” layer with the pencil tool at the appropriate grid location; the color of the mark should be $RGB=[0,254,0]$. Because the width of each grid line is four pixels, the intersection of height and width gridlines is a 4x4 collection of points. The location of the origin can be specified as the upper-right point of the central four points. This will require zooming in on the Photoshop grid and flight track map image.
- (f) Calculate coordinate matrices with a MATLAB function: In Photoshop, turn “off” the flight track map image layer and the grid layer. Only the origin layer should be active. Save the image as a .png file. Use the MATLAB function *COORDINATE_MATRIX_CREATOR* to calculate coordinate matrices based

on the position of the origin and the number of pixels between grid points. See Appendix I.17 for a description of the function. Two coordinate matrices will be calculated, x and y , which contain x -coordinates and y -coordinates, respectively, of the grid.

- (g) Outline flight tracks in Photoshop and calculate flight track coordinates with a MATLAB function: Once the x and y coordinate matrices are created, coordinates of flight tracks can be calculated. Turn “off” the grid layer and the origin layer. Change the color of the pencil to RGB=[254,0,0]. The following instructions apply to each flight track desired to be created in INM:
 - i. Outline the flight track: In Photoshop, create a new “flight track” layer. Mark over a flight track with a number of dots in the flight track layer. An appropriate number of dots should be used to faithfully capture the shape of the flight track. Turn “off” the flight track map layer. The flight track layer should be the only active layer. Save the layer to an appropriately-named .png file. Rename the layer. Turn “off” the current flight track layer.
 - ii. Calculate the INM grid points for the flight track: Use the MATLAB function *FLIGHTTRACK_POINT_GENERATOR* to calculate grid points for each flight track with use of the x and y coordinate matrices. The calculated grid point coordinates must be manually input into the INM file *trk_segs.dbf*. The track name must also be input into the INM file *track.dbf*. See Appendix I.17 for a description of the function.

E.2 Flight Track Data

This is the second of two methods to manually create flight tracks for an airport. Flight track data is downloaded from the Internet. MATLAB is used to collect and organize the data into a format that INM can import and process as radar flight tracks. Flight tracks are created within INM using the radar flight track information.

1. Acquire and group flight track data from the Internet: Obtain coordinate data for real-life aircraft operations from FlightAware (<http://flightaware.com>). FlightAware stores coordinate data for actual aircraft operations. Coordinate data for each operation is contained in a Google Earth .kml file, which is in .xml format. A number of .kml files sufficient to represent the range of approach and departure directions for each runway at a given airport should be downloaded and stored. Group the files by city, operation type (arrival or departure), runway, and direction of origin/destination. All files of the same group should be stored in a hard drive folder with a folder path that is formed from the combination of values of the grouping variables. City is the top level, followed by operation type, runway, and direction of origin/destination. An example folder path is *R:\airport\A\5L\Northeast*, where *airport* is the name of the airport, *A* stands for arrivals, *5L* is the name of a runway, and *Northeast* is the direction of origin. Google Earth can be downloaded from <http://www.google.com/earth/download/ge/agree.html>.

2. Organize the flight track data into the format required by INM: Use the MATLAB function

CSV_CREATOR to format and aggregate the .kml files of one group into one .csv file. See Appendix I.17 for a description of the function. This must be performed for each group of flight tracks. INM can only display the radar track data of one .csv file at a time. Separate .csv files are created so as not to over-clutter the INM display with flight track data when a .csv file is imported. For each .csv file:

(a) Import the .csv file into INM: INM features a utility called the Radar Processor that can be used to import the formatted .csv text file. From Section 3.5.6 in (He et al., 2007), the radar flight track .csv text format is as follows:

```
Code, FltId, AcType, OpType, ApFrom, ApTo, RwyId
, Time, X or Lat, Y or Long, Alt, Spd
...etc...
```

```
Code, FltId, AcType, OpType, ApFrom, ApTo, RwyId
, Time, X or Lat, Y or Long, Alt, Spd
, Time, X or Lat, Y or Long, Alt, Spd
...etc...
```

where, the fields are:

Code	Beacon/numeric code or 0000
FltId	Flight identifier
AcType	FAA aircraft type ID
OpType	Operation type: A = approach, D = depart, V = overflight, ? = unknown
ApFrom	From-airport or ???
ApTo	To-airport or ???
RwyId	Runway end or ???
Time	Time of radar return (decimal hours)
X or Lat	X value (nmi, km) or latitude (decimal deg.)
Y or Long	Y value (nmi, km) or longitude (decimal deg.)
Alt	Aircraft altitude MSL (ft, m)
Spd	Aircraft group speed (kt, km/hr) or 0

Fields must be separated by commas. A comma must be in the first column of a track-point line. Blank lines and comments lines can be anywhere. There can be any number of tracks and points per track. A line starting with the number-sign character (#) is ignored. Where noted, ??? is an acceptable field value. You can use this method to document your file or to comment-out data sections.

- (b) Create new points-type (P-tracks) in INM: Enable the display of radar tracks and view the tracks in the Input Graphics window. Follow the directions in Section 7.5.1 of (He et al., 2007) to create flight tracks. Multiple flight tracks can be created for each group of radar tracks. If multiple flight tracks are

desired, it is recommended to create dispersed tracks (Section 7.5.7 in (He et al., 2007)). This ensures that the flight tracks for one radar flight track group are evenly spaced. The dispersion process is used only to initially shape flight tracks. After creating dispersed flight tracks, it is important to make the dispersed flight tracks independent from the main flight tracks. Doing this enables the moving of individual points of the dispersed flight tracks if desired. The dispersed flight tracks are made independent of the main flight track through the manipulation of `track.dbf` and `trk_segs.dbf`. For dispersed tracks, INM labels the tracks with the same `TRK_ID1,C,8` name given to the main track, but assigns a nonzero `TRK_ID2,C,1` number to the tracks. Each dispersed track must be given a unique `TRK_ID1,C,8` name and a `TRK_ID2,C,1` number equal to zero in both `track.dbf` and `trk_segs.dbf`. `TRK_ID1,C,8` and `TRK_ID2,C,1` are field names in the `.dbf` files.

APPENDIX F. INSTRUCTIONS FOR CREATING A SYNTHETIC POPULATION

Detailed instructions for creating a synthetic population around an airport are given in this appendix. Basic instructions were given in Chapter 5, Section 5.2.3.

1. On the hard drive and inside the folder that stores the Synthetic Population Generator code, create a new “project” folder. This folder will store the output for a simulated population. The name of the folder should include the name of the airport around which a population will be simulated or the name of the city in which the airport resides (Appendix H.1).
2. Collect Census 2000 Summary File 3 data.
 - (a) Census 2000 Summary File 3 data can be found at http://www2.census.gov/census_2000/datasets/Summary_File_3/. Navigate to this FTP site.
 - (b) Find and click on the name of the state in which the airport resides. This will navigate you to a state-specific FTP site with a geographic header file and 76 data files.
 - (c) Save six files to a location on your hard drive. To save a file right-click it and choose “Save link as...”. Save the following files (“aa” represents a state abbreviation):
 - i. aa00009_uf3.zip
 - ii. aa00010_uf3.zip
 - iii. aa00011_uf3.zip
 - iv. aa00012_uf3.zip
 - v. aa00015_uf3.zip
 - vi. aageo_uf3.zip
 - (d) Use a zip tool (e.g., WinZip) to unzip each of the downloaded files.

- (e) Transform the .uf3 files to .txt (Appendix G).
 - (f) Rename the state abbreviation “aa” to “sf” of each .txt file.
 - (g) Move the .txt files to the project folder.
3. Collect Census 2000 5% Microdata.
- (a) Census 2000 Microdata can be found at <http://www.census.gov/census2000/PUMS5.html>. Navigate to this website.
 - (b) In the Table with heading “States Now Available FTP (Release date)” will be listed state names, each hyperlinked to a state-specific FTP site. Click on the name of the state in which the airport resides.
 - (c) At the state-specific FTP site, there will be a file named “PUMS5-##.txt”, where ## is the the state number of the state.
- Note: Other states’ microdata can also be used to simulate the population. The Synthetic Population Generator is capable of using the microdata from multiple 5% files.*
4. Identify a group of households and persons listed in the microdata files which will be used to simulate the population: The group of microdata households and persons which will be used in the Synthetic Population Generator are specified by super-PUMA numbers. The recommendation is to use Census 2000 super-PUMA/PUMA maps to find geographic areas and their super-PUMA numbers.
- (a) Census 2000 super-PUMA/PUMA maps for each state can be found at <http://www.census.gov/geo/www/maps/puma5pct.htm>. Navigate to this website.
 - (b) For each state from which a 5% microdata file is to be used to simulate the population, the super-PUMA number of each super-PUMA to use must be noted. The full collection of super-PUMA numbers will be integrated into an input for the Synthetic Population Generator.
- Note: The 5% microdata sample for a state is grouped into PUMAs and super-PUMAs. Collections of households and persons for the Synthetic Population Generator to use are specified by super-PUMA because if a state is*

very large it might not be desirable to use the entire state's 5% microdata file or if an airport is near a state line, it might be desirable to use subsets of the 5% files of the adjacent state(s).

5. Determine geographic areas for which to simulate a population.

Note: The Synthetic Population Generator simulates the population for census block groups but is flexible on how geographic areas are specified as input to the program. The expected format of this input is described in point 6c below.

- (a) Summary File 3 reports tabulations for the following geographical levels:
 - i. State
 - ii. County
 - iii. Place
 - iv. Tract
 - v. Block group
- (b) The recommendation is to specify collections of census block groups at the county level. This requires knowledge of county locations, names, and FIPS (Federal Information Processing Standards) codes.
 - i. Google Earth can be used to find county delineations and names. Google Earth can be downloaded from <http://www.google.com/earth/download/ge/agree.html>.
 - ii. FIPS codes can be found at <http://www.census.gov/geo/www/ansi/countylookup.html>.

Note: Though the county level is recommended, it isn't a requirement. All areas could be identified at the block group level if desired. Also, the areas can differ in geographical level.

6. Create Synthetic Population Generator inputs.

- (a) Input 1: Project folder name. This is the name of the folder that houses the downloaded Census 2000 data.

(b) Input 2: Choice of Summary File 3 data tables to constrain the population design for each area. The input is a t -by-1 MATLAB cell, where t is the number of data tables. In total, seven data tables can be used. Three of the tables contain counts of persons, three contain counts of households, and one contains counts of housing units. Each of the t data tables can be one of the following:

- i. P32 (Person-level data table) Travel time to work by means of transportation to work for workers 16 years and over who did not work at home
- ii. H34 (Housing unit-level data table) Year structure built
- iii. P34 (Person-level data table) Time leaving home to go to work for workers 16 years and over
- iv. H38 (Household-level data table) Tenure by year householder moved into unit
- v. P145 (Person-level data table) Separate population into categories by race, gender and age
- vi. P146 (Household-level data table) Households by age of householder by household type (including living alone) by presence of own children under 18 years
- vii. P151 (Household-level data table) Household income in 1999 (by race of householder)

Note: P145 and P146 must be involved in the population design. At minimum these two tables must be used. The specific choice of tables is dependent on what factors are considered important to control for.

(c) Input 3: Constraints identifying geographic areas for which to simulate a population. The input is a m -by-1 MATLAB cell, where m is the number of geographic areas. Each row contains a 1-by- p cell, where $p/2$ is the number of specification pairs. A specification pair is composed of two parts: a number identifying one of the five levels of geography and a particular numerical

value of that geography level. Each entry must be enclosed by single quotes. A hypothetical example follows:

$$\begin{aligned} \text{constraints} = \{ & \\ & \{ '3', '51', '4', '01', '6', '111111', '7', '1' \}; \\ & \{ '3', '51', '4', '01', '6', '111111', '7', '2' \} \\ & \}; \end{aligned}$$

In the example, there are two rows and thus two geographical area specified. There are eight rows and thus four specification pairs. The first column of each specification pair is a number that identifies a level of geography: '3' = State, '4' = County, '5' = Place, '6' = Tract, '7' = Block Group. The two geographic areas specified are block groups since both include the '7' identifier. The first constraint refers to: State = 51, County = 01, Tract = 111111, Block Group = 1.

- (d) Input 4: Names of the geographic areas identified by the constraints. The input is a m -by-1 MATLAB cell, where m is the number of geographic areas. In each row of the cell should be a descriptive name without spaces. Each entry must be enclosed by single quotes.
- (e) Input 5: Names of the 5% microdata files. The input is a f -by-1 MATLAB cell, where f is the number of files. Each name should be the full name (including the .txt extension) of a microdata file.
- (f) Input 6: Identities of super-PUMAs to use from the 5% microdata files. The input is a s -by-1 MATLAB cell, where s is the number of super-PUMAs. Each super-PUMA identifier should be enclosed by single quotes to preserve the presence of trailing zeros. The geographic locations and identifiers of super-PUMAs can be found via the PUMA maps described in point 4a.

APPENDIX G. TRANSFORMING .UF3 TO .TXT

A DOS command window will be used to navigate to the folder containing the .uf3 files and transform the .uf3 file extensions to .txt file extensions:

1. Navigate to the storage folder
 - (a) Type **cd**.
 - (b) Press the space bar.
 - (c) Press the **Tab** key until the top-level folder appears.
 - (d) Type a backslash (\) and press the **Tab** key until the next folder in the sequence appears. Repeat this process until the folder containing the files is shown.
 - (e) Press **Enter**
2. Use the rename command (**ren**) to rename the files as follows:
ren *.uf3 *.txt

APPENDIX H. SURVEY SIMULATION PROGRAM INFORMATION

In this appendix, the folder structure and input text file structures and contents of the Survey Simulation Program are described.

H.1 Folder Structure and Program Files

In order for the computer programs of the Survey Simulation Program to operate without error a specific hard drive folder structure must be used. The computer programs are stored in the required convention on a hard drive in the possession of Dr. Patricia Davies of Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN, 47906. The survey simulation computer programs are stored within the *SURVEY SIMULATION PROGRAM* folder and are described in Appendix I.

The Noise Metric Module program files are stored in a folder named *Noise Metric Module*. The Synthetic Population Generator program files are stored in a folder named *Synthetic Population Generator*. The Survey Sampling Module program files are stored in a folder named *Survey Sampling Module*. In addition to the folders for the computer program code, two other folders store information which the programs use. One folder is named *Shapefiles* and stores Census 2000 TIGER/Line files. These files are described in Chapter 5, Section 5.2.4. The other folder is named *Text* and stores input text files used by the Noise Metric Module and the Survey Sampling Module. The use of input text files allows the expansion of program functionality without editing program code. That is, information can be added to the text files which can be utilized by the programs. The text files are described in Appendix H.2.

For each airport around which survey simulations are desired, subfolders named after either the airport or the city in which the airport resides should be manually created in the *Shapefiles*, and *Synthetic Population Generator* folders. There is no

need to do this within the *Noise Metric Module*, *Survey Sampling Module*, or *Text* folders. As described in Chapter 5, Section 5.2.4, the county-level or state-level shapefiles for the area surrounding the airport should be stored in the subfolder within the *Shapefiles* folder.

H.2 Text Files

Input text files used by the Noise Metric Module and Survey Sampling Module are described in this section. Each entry in a text file must be enclosed by double quotes, even if it is a blank entry. Each newline in a text file is a new row. Columns are established by the use of double quotes. For entries in text files that contain text or names: each name can contain, but not begin with, commas, numbers, and spaces. The text files, the sections that they are described in, and the modules that use them are shown in Table H.1.

Table H.1. Text file names, sections that they are described in, and the modules that use the text files. “Y” indicates that the module uses the text file. “N” indicates that the module does not use the text file.

Text file name	Section	Noise Metric Module	Survey Sampling Module
airport_list.txt	H.2.1	Y	Y
analysis_options.txt	H.2.4	N	Y
annoyance_model_list.txt	H.2.3	N	Y
annoyance_model_terms.txt	H.2.3	N	Y
annoyance_model_terms_forms.txt	H.2.3	N	Y
annoyance_model_terms_values.txt	H.2.3	N	Y
anova_options.txt	H.2.4	N	Y
data_compilation_options.txt	H.2.4	N	Y
model_term_list.txt	H.2.2	N	Y
plotting_options.txt	H.2.4	N	Y
sampling_design_list.txt	H.2.1	N	Y

H.2.1 Basic specification

Two text files are described in this section: `airport_list.txt` and `sampling_design_list.txt`. `airport_list.txt` is a single-column text file that contains airport names. Name can be composed of multiple words. `sampling_design_list.txt` is a single-column text file that contains sample designs. As of this writing, only one sample design (random sampling, without replacement) is listed in the text file. The intent for the existence of this text file is that the scope of the Survey Sampling Module could be expanded in the future by inserting other sample designs and adding code to the Survey Sampling Module.

H.2.2 Variable specification

One six-column text file is described in this section. `model_term_list.txt` contains information about variables, both acoustical and nonacoustical, that the Survey Sampling Module can use in annoyance models to generate population annoyance, as stratification variables, and in annoyance models derived from survey data. Each variable is reserved a row of the text file. Acoustical variables include sound level variables (cumulative energy-based and average single-event) and number-of-events variables calculated from the Noise Metric Module. Nonacoustical variables include household-level and person-level demographic variables extracted from Census 2000 data by use of the Synthetic Population Generator. The first column contains the variable names. There are no requirements for naming variables. Columns two through four are reserved for information about acoustical variables. For nonacoustical variables, cells in these columns are left blank. Cells of the second column contain one of three words (“Logarithmic”, “Arithmetic”, or “Loudness”) which identifies the averaging method the Noise Metric Module used to generate the acoustical variable (Chapter 5, Section 5.2.2). Cells of the third column contain metric names of acoustical variables, as established by the Noise Metric Module. These include:

1. `DNL_from_LAeq`,

2. LAeq,
3. LAeqday,
4. LAeqnight,
5. SEL,
6. LAm_{ax},
7. EPNL,
8. PNLTM.

DNL_from.LAeq, LAeq, LAeqday, and LAeqnight are equivalent to DNL , L_{Aeq} , $L_{Aeq,d}$, and $L_{Aeq,n}$, respectively. SEL, LAm_{ax}, EPNL, and PNLTM are equivalent to $SEL_{A,av}$, $L_{Amax,av}$, $EPNL_{av}$, and $PNLT_{max,av}$, respectively. Cells of the fourth column contain either a background sound level or a sound level threshold. If the cell of the third column contains either DNL_from.LAeq, LAeq, LAeqday, or LAeqnight, the fourth column contains a background sound level. If the cell of the third column contains either SEL, LAm_{ax}, EPNL, or PNLTM, the fourth column contains a threshold sound level. Cells of the fifth column contain a number:

- 1, if a maximum average sound level variable ($L_{Amax,av}$ or $PNLT_{max,av}$),
- 2, if an energy-dose sound level variable ($SEL_{A,av}$ or $EPNL_{av}$),
- 3, if a cumulative energy-based average sound level variable (DNL , L_{Aeq} , $L_{Aeq,d}$, and $L_{Aeq,n}$),
- 4, if a statistical sound level variable (for possible future expansion of the Noise Metric Module),
- 10, if a number-of-events variable,
- 11, if a logged (base 10) number-of-events variable,
- 20, if a household-level nonacoustical variable,
- 21, if a person-level nonacoustical variable,
- 30, for a constant.

Cells of the sixth column are reserved for nonacoustical variables. For acoustical variables, cells in this column are left blank. If the fifth column of a given row contains the number 20 or 21, the cell in the sixth column will contain a reference to a specific household-level or person-level data item, respectively, of a simulated population around an airport. The reference identifies a column location in a variable output by the Synthetic Population Generator (see point 8 in Appendix I.7).

H.2.3 Annoyance model specification

Information about annoyance linear regression models that generate the population annoyance in the Survey Sampling Module (Chapter 5, Section 5.2.4) is contained in four different text files. `annoyance_model_list.txt` is a single-column text file that contains names of annoyance models. `annoyance_model_terms.txt` is a multiple-column text file that contains, for each annoyance model listed in `annoyance_model_list.txt`, the names of variables from `model_term_list.txt` (Appendix H.2.2) with one variable per column. The maximum number of variables for each annoyance model is thirteen. The names of the variables must match those from `model_term_list.txt`. `annoyance_model_terms_values.txt` contains the numerical coefficient values of the variables listed in `annoyance_model_terms.txt`.

`annoyance_model_terms_forms.txt` contains the power that each variable listed in `annoyance_model_terms.txt` is raised to. Every variable must have an associated power, even if it is one.

H.2.4 Analyses specification

Four text files that contain information about various analysis Options in the use of the Survey Sampling Module are described in this section. `analysis_options.txt` is the top-level analysis text file; the other three (`anova_options.txt`, `data_compilation_options.txt`, and `plotting_options.txt`) are sub-level text files for the

Options in `analysis_options.txt`. The analysis Options are described in Chapter 5, Section 5.2.4. The top-level analysis Options are:

1. Collinearity analysis,
2. Statistical power analysis of \hat{k} ,
3. Variance analysis of \hat{k} ,
4. Extra sum of squares,
5. Condition number distributions,
6. Data compilation,
7. ANOVA,
8. Plotting.

APPENDIX I. COMPUTER PROGRAM DESCRIPTIONS

In this appendix are given descriptions of various computer programs that were created throughout the research. A brief overview of each program is given accompanied by a list of the functions that make up each program. All functions were coded in MATLAB; file names of codes are capitalized and in every case, a file name is also the name of a main function. Subfunctions are sometimes also reserved a file in which case the names are also capitalized. In the instance that a subfunction is stored within a file, it will be named in lower case. The inputs and outputs of a main function are described when they are required to be provided by the user. The computer programs are stored on a hard drive in the possession of Dr. Patricia Davies of Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN, 47906. The location of each file on the drive is provided.

I.1 Noise Metric Module

The Noise Metric Module calculates the acoustical environs (sound level metrics and daily number-of-events), given an aircraft operations scenario as input, on a geographical grid of points using INM acoustical data for single aircraft events. The Noise Metric Module is described in detail in Chapter 5, Section 5.2.2. There is one main function and four subfunctions:

1. *NOISE_METRIC_MODULE*: This is the main function of the program. This function processes the input if provided input, or asks the user for input if not provided input. This function directs control to the other subfunctions.

Inputs: There is one optional input. The input is a text file that contains responses to questions that the function would normally prompt the user to answer, without the text file as input. If *NOISE_METRIC_MODULE* is run

without an input, then the function will prompt the user for responses to a number of questions, and then create a text file with the recorded responses. The input text file should be stored in the following relative folder path: ...*SURVEY SIMULATION PROGRAM\Noise Metric Module\Input Files*. An example of the text file is:

```

Number data sources:  2
Data source pathname:  F:\PURDUE\SDF-LAEQ-S1\
Data source pathname:  F:\PURDUE\Louisville-S1-Arr-Dep\
Metrics:  1 2 3 4 5 6 7 8
Averaging:  1 1 1 1 1 1 1 1
Background noise:  30 35 40 45 50
SEL thresholds:  50 55 60 65 70 75 80 85 90 95 100
LAmx thresholds:  50 55 60 65 70 75 80 85 90 95 100
EPNL thresholds:  50 55 60 65 70 75 80 85 90 95 100
PNLTM thresholds:  50 55 60 65 70 75 80 85 90 95 100
Data source number for metric:  1
Data source number for metric:  2
Airport:  Louisville
Scenario:  SDF_default_2009
Halt choice:  1
Operation increase choice:  1

```

A description of the text file items follows:

- **Number data sources:** The Noise Metric Module uses INM acoustical data for single aircraft events to calculate acoustical metrics. The first instance that a certain collection of INM acoustical data is used, the program directs MATLAB to import and store the data for easier access in the future. After being imported and stored as a series of MATLAB .mat files, the data become a “data source” that the program can use in the future. **number data sources** is the number of different storage locations of imported single-event INM data that the function is being directed to use for the calculation of acoustical metrics. Thus, an input text file cannot be used if the single-event data that are required have not been imported and stored. When the function is run, it asks the user to direct it to where single-event INM data are stored. If the data have not been imported and stored, the program performs this operation, and a text file can be used to reference the imported data in the future.
- **Data source pathname:** This item is listed n times in the input text file, where n is the number of imported single-event INM acoustical data sources to be used. A “pathname” is a hard drive folder path of a storage location of imported single-event INM acoustical data. In the example, more than one pathname is specified because in one collection of single-event data, there are data for some, but not all, of the metrics that are desired to be calculated.
- **Metrics:** This item specifies the numbers of the metrics that are to be calculated. Each of eight metrics has an assigned number:

 - 1 = DNL from single-event L_{Aeq} ,
 - 2 = L_{Aeq} ,
 - 3 = $L_{Aeq,d}$,
 - 4 = $L_{Aeq,n}$,
 - 5 = $SEL_{A,av}$,
 - 6 = $L_{Amax,av}$,

$$7 = EPNL_{av},$$

$$8 = PNLT_{max,av}.$$

Metrics 1-4 are cumulative energy-based average metrics, metrics 5-8 are average single-event metrics.

- **Averaging:** This item specifies what averaging method to use when calculating the overall metrics from the single-event INM acoustical data. An averaging method is specified for each metric to be calculated. Each of three averaging methods has an assigned number:
 - 1 = Logarithmic (Equation (5.1)),
 - 2 = Arithmetic (Equation (5.2)),
 - 3 = Loudness (Equation (5.3)).
- **Background noise:** This item specifies background noise levels to use in calculating the cumulative energy-based average metrics (DNL , L_{Aeq} , $L_{Aeq,d}$, or $L_{Aeq,n}$). If neither of these metrics are specified, then the word **NONE** should be inserted instead.
- **SEL thresholds:** This item specifies the threshold levels to use in calculating $SEL_{A,av}$. For each threshold level, $SEL_{A,av}$ is calculating involving single events with a SEL_A greater than the threshold.
- **LMax thresholds:** This item specifies the threshold levels to use in calculating $L_{Amax,av}$. For each threshold level, $L_{Amax,av}$ is calculating involving single events with a L_{Amax} greater than the threshold.
- **EPNL thresholds:** This item specifies the threshold levels to use in calculating $EPNL_{av}$. For each threshold level, $EPNL_{av}$ is calculating involving single events with a $EPNL$ greater than the threshold.
- **PNLTM thresholds:** This item specifies the threshold levels to use in calculating $PNLT_{max,av}$. For each threshold level, $PNLT_{max,av}$ is calculating involving single events with a $PNLT_{max}$ greater than the threshold.
- **Data source number for metric:** This item is listed m times in the input text file, where m is the number of metrics to be calculated, as listed in the

Metrics item. The ordering of the data source numbers corresponds to the order of the metrics. Each metric must be connected to a data source, so that the program knows where to access the single-event data for the metric. This item must be listed m times, even if there is only one data source. The data sources are numbered by the order that they are listed in the input text file.

- **Airport:** This item specifies the name of the airport around which the noise exposure is to be calculated. Possible names to use are those that appear in `airport_list.txt`, described in Appendix H.2.1.
- **Scenario:** This item specifies the aircraft operations scenario that the noise exposure will be calculated for. The aircraft operations scenario must exist as a MATLAB `.mat` file in the following relative folder path: `...\Noise Metric Module\Airport\Operations scenarios`, where **Airport** is the name of the airport, as listed in `airport_list.txt`.
- **Halt choice:** This item specifies a command for the function if it encounters flight track/aircraft/standard profile stage number combinations specified in the aircraft operations scenario for which there is no single-event INM acoustical data. The command can be 0, which directs the function to halt calculation, or 1, which directs the function to continue calculating the noise exposure with the rest of the combinations with single-event INM acoustical data.
- **Operation increase choice:** This item specifies a second command for the function if it encounters flight track/aircraft/standard profile stage number combinations specified in the aircraft operations scenario for which there is no single-event INM acoustical data. In the case that the aircraft operations scenario specifies combinations without single-event INM acoustical data and **Halt choice** is set 1, a command of 1 directs the function to distribute the total number of missing-data arrival/departure operations among those arrival/departure combinations for which single-event INM acoustical

data exists, in order to involve in calculations the total number of operations specified in the aircraft operations scenario. A command of 0 directs the function to not perform this redistribution.

Outputs: There are eight outputs for each metric/averaging method combination specified. Each output is or contains a matrix of values, where each point of the matrix represents a geographical location on a grid surrounding an airport. The outputs are as follows:

- **lat:** This variable is a matrix of latitudes for each grid point.
- **long:** This variable is a matrix of longitudes for each grid point.
- **lat_converted_to_nmi:** This variable is the matrix of latitudes, converted to nautical miles, with reference to the southwestern-most point on the grid.
- **long_converted_to_nmi:** This variable is the matrix of longitudes, converted to nautical miles, with reference to the southwestern-most point on the grid.
- **metric:** This variable stores a matrix of sound level metric values, calculated without reference to a threshold or without involving background noise.
- **number:** This variable stores a matrix of numbers. At each point of the matrix is a number that represents the total sum of the operations specified in the aircraft operations scenario. Each number is a daily number-of-events.
- **metric_levels:** For cumulative energy-based average metrics, this variable stores, for each background noise level specified, the calculated exposure at each grid point. The background noise level is constant across grid points. For average single-event metrics, this variable stores, for each threshold level, the average exposure at each grid point involving only the single events with a sound level that exceeds the threshold.
- **number_levels:** For cumulative energy-based average metrics, this variable stores, for each background noise level specified, a matrix that is equivalent to **number**. For average single-event metrics, this variable stores, for each

threshold level, the number-of-events at each grid point with a metric value above the sound level threshold and that contributed to calculating the average metric value. Each number is a daily number-of-events.

2. *DNL_from_LAeq*: This function calculates the cumulative energy-based average metric *DNL*.
3. *LAeq_variants*: This function calculates the cumulative energy-based average metrics L_{Aeq} , $L_{Aeq,d}$, and $L_{Aeq,n}$.
4. *metric_calculate*: This function calculates the average single-event metrics $SEL_{A,av}$, $L_{Amax,av}$, $EPNL_{av}$, and $PNLT_{max,av}$.
5. *single_event_data_import*: This function imports and stores single-event INM acoustical data.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module

I.2 Aircraft Operations Scenario Analysis

The program described in this section will search for the existence of single-event INM acoustical data for flight track/aircraft/standard profile stage number combinations specified in an aircraft operations scenario. There is one main function:

1. *SCENARIO_FLAG*: This function will search, in locations specified by the user, for the existence of single-event INM acoustical data for flight track/aircraft/standard profile stage number combinations specified in an aircraft operations scenario. The function creates a notated scenario in which all flight track/aircraft/standard profile stage number combinations with single-event data are marked with a value of 1 and the missing-data combinations are marked with a value of 0. The function creates a list of the aircraft for which data is missing.

Inputs:

- **scenario**: This variable is an aircraft operations scenario in the format as that used by *NOISE_METRIC_MODULE*.

Outputs:

- **scenario_flagged**: This variable is in the form **scenario** in which all flight track/aircraft/standard profile stage number combinations with single-event data are marked with a value of 1 and the missing-data combinations are marked with a value of 0.
- **aircraft_missing**: This variable is a list of all of the aircraft models listed in the aircraft operations scenario with missing single-event INM acoustical data.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module\Various\Aircraft Operations Scenario Analysis

I.3 Airport 3 Aircraft Operations Scenario Adjustment

The codes described in this section find and replace aircraft without corresponding single-event INM acoustical data that are specified in the aircraft operations scenario for Airport 3 created from the INM ops_ftl.dbf file. The methodologies that the codes use are described in detail in Chapter 5, Section 5.2.2. There is one main function and one subfunction:

1. *SCENARIO_AIRCRAFT_REPLACE*: This is the main function of the program.

Inputs:

- **scenario**: This variable is an aircraft operations scenario in the format as that used by *NOISE_METRIC_MODULE*.

Outputs:

- **scenario_modified**: This variable is a modified aircraft operation scenario in that the aircraft models with missing single-event INM acoustical data are replaced by aircraft models with single-event data.

2. *npd_curve_analysis*: This function matches aircraft models with missing single-event INM acoustical data with aircraft models with single-event data, based on comparisons of the Noise-Power-Distance data for the aircraft. See Chapter 5, Section 5.2.2 for a description of the methodology.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module\Various\LAX Aircraft Operations Scenario Adjustment

I.4 Aircraft Operations Scenario Creation from INM Output

The program described in this section creates an aircraft operations scenario for use within the Noise Metric Module given an INM ops_ft.dbf file as input. There is one main function:

1. *SCENARIO_CREATOR*: This function organizes the information contained within an INM ops_ft.dbf file to create an aircraft operations scenario that can be used by other functions.

Inputs:

- **ops_ft**: This variable is an INM ops_ft.dbf file that has been saved as a Microsoft Excel .xls file. The name of the file should be enclosed by single quotes.

Outputs:

- **scenario**: This variable is an aircraft operations scenario which other functions can use. The variable is a MATLAB cell; the format of the variable is described in Chapter 5, Section 5.2.2.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module\Various\Aircraft Operations Scenario Creation from INM Output

I.5 Aircraft Model and Flight Track Contributions

The program described in this section analyzes an INM ops_ft.dbf file and ranks the contributions of both aircraft and flight tracks to the overall air traffic. The significant contributions (aircraft and flight tracks that contribute to at least 90% of the air traffic) are given. There is one main function:

1. *AIRCRAFT_FLIGHTTRACK_CONTRIBUTION*: This function analyzes the information contained within an INM ops_ft.dbf file. The function ranks the contributions of aircraft and flight tracks to the total number of operations and the number of night operations, for both arrivals and departures. The function ranks the full collection of aircraft and flight tracks and creates lists of the 90% contributors.

Inputs:

- *ops_ft*: This variable is an INM ops_ft.dbf file that has been saved as a Microsoft Excel .xls file. The name of the file should be enclosed by single quotes.

Outputs:

- *contributions*: This variable contains a list of four different rankings of aircraft and four different rankings of flight tracks. Each ranking is created using either total or night operations counts for either arrivals or departures.
- *significant_contributions*: For each of the rankings, this variable contains a list of the “significant” contributors, those aircraft or flight tracks that contribute to at least 90% of the operations count.
- *significant_contributions_unique_listing*: This variable is a list of every aircraft and flight track that appear in the rankings.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module\Various\Aircraft Model and Flight Track Contributions

I.6 Variation of Noise Exposure within Census Block Groups

The program described in this section analyzes the variation of acoustical data at grid points within census block groups which results from an aircraft operation scenario around an airport. There is one main function:

1. *CENSUS_BLOCK_GROUP_VARIATION*: This function categorizes the census block groups around an airport into different $L_{Aeq,d}$ groups based on the arithmetic average of the $L_{Aeq,d}$ at the grid points with each census block group. Within each $L_{Aeq,d}$ category, the function analyzes the variation of the acoustical data of another sound level metric within and across census block groups.

Inputs:

- **city**: This variable is the name of a city or airport (enclosed by single quotes) which identifies a top-level storage folder of data output from *NOISE_METRIC_MODULE*.
- **scenario**: This variable is the name of an aircraft operations scenario (enclosed by single quotes) which identifies the next level (after **city**) of a storage folder of data output from *NOISE_METRIC_MODULE*. *NOISE_METRIC_MODULE* is described in Appendix I.1.
- **averaging**: This variable is the name of an averaging method (enclosed by single quotes) which identifies the next level (after **scenario**) of a storage folder of data output from *NOISE_METRIC_MODULE*. Possible choices are ‘Logarithmic’, ‘Arithmetic’, or ‘Loudness’.
- **metric_name**: This variable is the name of an output file (enclosed by single quotes) from *NOISE_METRIC_MODULE* which is named after an average single-event sound level metric. Possible choices are ‘EPNL’, ‘LAm_{ax}’, ‘PNLTM’, or ‘SEL’.
- **threshold**: This variable is a number of a sound level threshold of **metric_name**. This variable should not be enclosed by single quotes.

Outputs:

- **census_block_groups:** This is a variable that is output by the function *CENSUS_BLOCK_GROUP_COUNT*. It is a matrix that gives the census block group identifier of every grid point at which acoustical data was calculated within *NOISE_METRIC_MODULE*.
CENSUS_BLOCK_GROUP_COUNT is described in Appendix I.12.
- **standard_deviations:** This variable contains the standard deviation of the sound level data and the standard deviation of the number-of-events data normalized to the maximum value, for each census block group. It also contains frequency tables which list the number of grid points in ranges in each of the sound level and number-of-events.
- **LAeqday_contour_average:** For each of five groups of $L_{Aeq,d}$, this variable gives the number of census block groups in the $L_{Aeq,d}$ category and the averages of the standard deviations of the sound level and number-of-events data.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Noise Metric Module\Various\Variation of Noise Exposure within Census Block Groups

I.7 Synthetic Population Generator

The Synthetic Population Generator is used to simulate a population surrounding an airport using Census 2000 data. The Synthetic Population Generator is described in detail in Chapter 5, Section 5.2.3. There is one main function and eight subfunctions:

1. *A_SYNTHETIC_POPULATION_GENERATOR:* This is the main function of the program. It calls seven of the eight subfunctions. There are outputs from the subfunctions that are saved to hard drive. These will be described for each subfunction.

Inputs:

- **folder:** This variable is the name of a folder (enclosed by single quotes) stored in the directory that contains the program functions. Data will be stored in and read from this folder.
- **choice:** This variable is a t -by-1 MATLAB cell that contains names (enclosed by single quotes) of census data tables. The census data tables will be used to constrain the population generation process. The choices are ‘P32’, ‘H34’, ‘P34’, ‘H38’, ‘P145’, ‘P146’, and ‘P151’. See Chapter 5, Section 5.2.3 for a description of the data tables. At minimum, ‘P145’ and ‘P146’ must be specified.
- **area_specification:** This variable is a n -by-2 MATLAB cell which specifies geographic areas surrounding an airport that a population will be simulated for, where n is the number of geographic areas. The first column of the variable is reserved for numerical identifiers of geographical types; the second column is reserved for numerical identifiers of specific geographical areas of the types specified in the first column. All numbers should be enclosed by single quotes. The geographical type identifiers are:
 - ‘3’ = State
 - ‘4’ = County
 - ‘5’ = Place
 - ‘6’ = Tract
 - ‘7’ = Block Group

Each row of the second column contains a numerical identifier which identifies a particular state, county, place, tract, or block group. The recommendation is to specify areas by county (see point 5b in Chapter 5, Section 5.2.3). Although geographic areas can be specified by state, county, place, or tract, the population design process is carried out for census block groups within the geographic areas. The types of areas specified by this variable do not have to be the same. For example, a full county can be specified by

a county identifier and a subset of block groups in an adjacent county can be specified by block group identifier.

- **area_names**: This variable is a n -by-1 cell that contains names of the areas identified by `area_specification`. The names should be enclosed by single quotes and should not have any spaces. Separate words in one name can be separated by underscores.
- **microdata_files**: This variable is a m -by-1 cell that contains names of microdata .txt files, where m is the number of files. This variable identifies .txt files to use in designing the population surrounding the airport. Each name should end with the .txt extension and should be enclosed by single quotes.
- **microdata_superPUMAs**: This variable is a s -by-1 cell that contains microdata super-PUMA names, where s is the number of super-PUMAs. The super-PUMA names identify a subset of the microdata to use in designing the population surrounding the airport. Not every microdata entry in file specified by `microdata_files` will be used to design the population. Only those entries that are contained within one of the super-PUMAs specified by this variable will be used. If a state's full microdata is to be used in designing a population, then every super-PUMA in the state must be specified by this variable.

Outputs: There are no outputs for

A_SYNTHETIC_POPULATION_GENERATOR.

2. *B_SUMMARY_FILE_IMPORT*: This subfunction imports and stores information contained in a Summary File 3 geographic header file for a state.

Outputs:

- **Summary_File_Import**: This is a MATLAB .mat file that stores imported Summary File 3 information. Contained in the file is a full list of the block groups reported in the Summary File 3 information and a reduced list of the block groups that are in the areas specified in `area_specification`.

3. *C_TABLE_CREATE*: For each census block group for which a population will be designed, this subfunction imports and organizes Summary File 3 data into tables (each table specified in *choice*) which will constrain the population design.

Outputs:

- *Aggregate_Data*: This is a MATLAB .mat file that stores the Summary File 3 data tables for all census block groups for which a population will be designed.
- *household_count_tracker*: This is a MATLAB .mat file that stores a variable by the same name which contains counts of the number of households per census block group for which a population will be designed.

4. *D_MICRODATA_IMPORT*: This subfunction imports and stores microdata.

Outputs:

- *Microdata*: This is a MATLAB .mat file that stores the imported microdata. Housing unit records and person records are stored separately.

5. *microdata_entry_count*: This subfunction calculates the number of housing unit records and person records in a microdata file. It is called by

D_MICRODATA_IMPORT.

6. *E_SUPERPUMA*: This subfunction organizes the imported microdata into groups based on super-PUMA for every super-PUMA of microdata that is to be used in the population design.

Outputs:

- *Microdata_SuperPUMA_#*: This is a MATLAB .mat file that stores the imported microdata housing unit and person for a given super-PUMA, where # is the number of the super-PUMA.
- *Microdata_SA*: This is a MATLAB .mat file that stores the imported housing unit and person microdata for all super-PUMAs.

7. *F_MICRODATA_REFERENCE*: This subfunction references each housing unit or person microdata record from the collection of selected super-PUMAs to a cell of every household, housing unit, or person constraint table.

Outputs:

- *Referenced_Microdata_SuperPUMA_#*: This is a MATLAB .mat file that stores the referenced housing unit and person microdata for a given super-PUMA, where # is the number of the super-PUMA.
 - *Microdata_household_ref*: This is a MATLAB .mat file that stores a list of the referenced microdata housing units for all super-PUMAs.
8. *G_OPTIMIZE*: This subfunction optimizes a population for every specified census block group using combinatorial optimization. See Chapter 5, Section 5.2.3 for details of combinatorial optimization.

Outputs: Because *G_OPTIMIZE* uses an optimization algorithm, and the number of geographic areas for which populations will be designed can be very large, the program's runtime can be very long. Rather than storing the results for every area in one variable and outputting it at the program's finish, the program saves results to hard drive in increments. The name of the output files saved to hard drive are appended with a time stamp. There are two output files per time stamp:

- *sample_frame_cell*: This is a MATLAB .mat file that stores, for each census block group, the selection of housing units and persons that were determined to best-fit the constraint tables listed in *choice*.
 - *tracker*: This is a MATLAB .mat file that stores, for each census block group, a record of the optimization process. The optimization record specifies observations of the optimization statistic for all iterations.
9. *FILE_COMBINER*: This subfunction will combine the time-stamped *sample_frame_cell* and *tracker* files that were output by the program.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Synthetic Population Generator

I.8 Survey Sampling Module

The Survey Sampling Module is used to simulate noise surveys around airports and analyze the effectiveness of different sample designs. The Survey Sampling Module is described in detail in Chapter 5, Section 5.2.4. There is one main function and three subfunctions:

1. *A_SURVEY_SAMPLING_MODULE*: This is the main function of the program. This function prompts the user to select the analyses to perform on collected survey data and to select a subset of the three subfunctions to run. There are outputs from the subfunctions that are saved to hard drive. These will be described for each subfunction.

Inputs: There are no inputs for *A_SURVEY_SAMPLING_MODULE*.

Outputs:

Stored within the ...*Survey Sampling Module\Analysis options* folder:

- **analysis_choice:** This is a MATLAB .mat file that stores a a -by-1 variable, where a is the number of analyses to perform on the collected survey data. The analysis choices are listed in Appendix H.2.4.
- **power:** This is a MATLAB .mat file that stores information relevant to a retrospective statistical power analysis.
- **extrasum:** This is a MATLAB .mat file that stores information relevant to an extra sum of squares analysis.
- **variance_method:** This is a MATLAB .mat file that stores information relevant to an estimation of the variance of a ratio of regression coefficients.
- **data_compilation_choices:** This is a MATLAB .mat file that stores information about what simulation-level statistics should be tabulated across sample designs.

- `plotting_choices`: This is a MATLAB `.mat` file that stores information about the types of plots to create.
- `anova_choices`: This is a MATLAB `.mat` file that stores information relevant to one or more analysis of variance analyses of simulation-level statistics.

2. `B_DESIGN`: This subfunction is used to design a survey.

Outputs:

Stored within the main storage folder:

- `user_choices_general_state`: This is a MATLAB `.mat` file that stores the survey design decisions. Stored within the file are the names of the airport and aircraft operations scenario, the variables and coefficients of the population annoyance model, the signal-to-noise ratio between the actual population annoyance and the annoyance generated by the population annoyance model, the number of annoyance scale categories, the choice of a discrete or continuous annoyance scale, and a choice of whether or not to use a podding approach.
- `user_choices_sampling_state_#`: This is a MATLAB `.mat` file that stores the sampling decisions for each sample design, where `#` is the number of the sample design that will be simulated. Stored within the file are the stratification variables, numbers of groups for each stratification variable, and sample size.
- `user_choices_model_state`: This is a MATLAB `.mat` file that stores survey analysis decisions in regard to the multiple linear regression annoyance models to derive from the survey data. The variables for each model and the power that each variable should be raised to are stored within the file.
- `derived_model_variables`: This is a MATLAB `.mat` file that stores, for each annoyance model to derive, predictor variable data for every household or individual that is a potential survey sample.

- **foldername:** This is a MATLAB .mat file that stores the main storage folder name within which output from the program is stored.
- **path_names:** This is a MATLAB .mat file that stores the folder names of every directory relevant to the Survey Sampling Module.
- **sampling_choices:** This is a MATLAB .mat file that stores the number of Monte Carlo trials to use for each survey simulation, the Type I error rate to use in statistical tests, and the minimum age allowable for a potential sample.
- **random_error:** This is a MATLAB .mat file that stores a matrix of errors to be applied to the annoyance generated by the population annoyance model. The matrix of errors is calculated using the chosen signal-to-noise ratio between the actual population annoyance and the annoyance generated by the population annoyance model (see Section 5.2.4 of Chapter 5).

Stored within the ...*Strata*\ folder one level beyond the main storage folder:

- **stratification_level_tracker:** This is a MATLAB .mat file that stores the order of stratification settings for all simulated sample designs. A stratification setting includes a specification of the number of strata for each stratification variable.
- **stratification_variable_boundaries_#:** This is a MATLAB .mat file that stores information about the strata boundaries for each sample design, where # is the number of the sample design.
- **stratification_variable_positions:** This is a MATLAB .mat file that stores information about the types of stratification variables in use, whether acoustical, household-level nonacoustical, or person-level nonacoustical.

- `stratification_variables_general_state`: This is a MATLAB `.mat` file that stores stratification variable data for every household or individual that is a potential survey sample.
3. *C_SAMPLE*: This subfunction is used to simulate surveys. The subfunction saves collected survey data and analysis results regarding annoyance model derivations to hard drive in increments. The name of the file saved to hard drive is appended with a time stamp. The data is stored within the *...Sampling* folder one level beyond the main storage folder. The individual time-stamped files must be combined prior to using *D_ANALYSIS*.

Outputs:

- `monte_carlo_results_sampling_state_#`: For each time stamp, this is a MATLAB `.mat` file that stores, for each Monte Carlo trial, the collected survey data and derived annoyance model results, where `#` is the number of the sample design.
 - `simulation_analyses_sampling_state_#`: This is a MATLAB `.mat` file that stores the simulation-level analyses (involving all Monte Carlo trials) of the derived annoyance model results, where `#` is the number of the sample design.
4. *D_ANALYSIS*: This subfunction is used to analyze survey data. Each analysis Option available in `analysis_choice` is reserved a separate folder for output.

Outputs:

Stored within the *...Analysis Results\ Variance of ratio of reg coeff* folder:

- `variance_results_sampling_state_#`: This is a MATLAB `.mat` file that stores, for each Monte Carlo trial, results from a variance estimation of a ratio of regression coefficients calculated using *REG_COEFF_RATIO_VARIANCE*, where `#` is the number of the sample design. *REG_COEFF_RATIO_VARIANCE* is described in Appendix I.9.

Stored within the *...Analysis Results\Collinearity* folder:

- *monte_carlo_results_collinearity_sampling_state_#*: This is a MATLAB .mat file that stores, for each Monte Carlo trial, results from a multicollinearity analysis calculated using *A_COLLINEARITY_INPUTS*, *b_collinearity_diagnostic*, and *c_collinearity_harmful*, where # is the number of the sample design. *A_COLLINEARITY_INPUTS* is described in Appendix I.10

Stored within the *...Analysis Results\Extra sum of squares* folder:

- *marginal_results_sampling_state_#*: This is a MATLAB .mat file that stores, for each Monte Carlo trial, results from an analysis of the extra sum of squares attributable to predictor variables in derived annoyance models calculated using *EXTRA_SUM_SQUARES*, where # is the number of the sample design. *EXTRA_SUM_SQUARES* is described in Appendix I.11.

Stored within the *...Analysis Results\Statistical power* folder:

- *power_results_sampling_state_#*: This is a MATLAB .mat file that stores, for each Monte Carlo trial, results from a retrospective statistical power analysis, where # is the number of the sample design.

Stored within the *...Analysis Results\Collinearity\Condition number* folder:

- *condition_number_results*: This is a MATLAB .mat file that stores, for each simulated sample design and for each derived annoyance model, the condition numbers of the annoyance model regression matrices across all Monte Carlo trials, the mean condition number, and the interquartile range of condition numbers.

Stored within the *...Analysis Results\Data compilation* folder:

- *model_#_mean_regcoeff*: This is a MATLAB .mat file that stores, for each simulated sample design, the mean regression coefficients considering all Monte Carlo trials of a certain derived annoyance model, where # is the num-

ber of the derived annoyance model. The data is calculated using Equation (5.20).

- `model_#_mean_deviation_regcoeff`: This is a MATLAB `.mat` file that stores, for each simulated sample design, the mean absolute deviations of the regression coefficients from the values in the population annoyance model, considering all Monte Carlo trials of a certain derived annoyance model, where `#` is the number of the derived annoyance model. The results are calculated from Equation (5.21).
- `model_#_bias`: This is a MATLAB `.mat` file that stores, for each simulated sample design, the biases of the mean regression coefficients of a certain derived annoyance model, where `#` is the number of the derived annoyance model. The results are calculated from Equation (5.22).
- `model_#_standard_deviation`: This is a MATLAB `.mat` file that stores, for each simulated sample design, the standard deviations of the regression coefficients considering all Monte Carlo trials of a certain derived annoyance model, where `#` is the number of the derived annoyance model. The results are calculated from Equation (5.23).
- `model_#_mean_deviation_regcoeff_standard_deviation`: This is a MATLAB `.mat` file that stores, for each simulated sample design, the mean absolute deviations of the standard deviations calculated from the full collection of regression coefficients from the individual Monte Carlo trial standard deviations from the covariance matrix, where `#` is the number of the derived annoyance model. The results are calculated from Equation (5.24).

Stored within the *...Analysis Results\ANOVA* folder:

- `model_#_TYPE_variable_#_anova`: This is a MATLAB `.mat` file that stores the results from an analysis of variance of a simulation-level statistic for a certain predictor variable of a derived annoyance model. The simulation-level statistic is of type `TYPE`, where `TYPE` can be:
 - `mean_deviation_regcoeff` (Equation (5.21))

- `bias` (Equation (5.22))
- `standard_deviation` (Equation (5.23))
- `mean_deviation_regcoeff_standard_deviation` (Equation (5.24))

A file is reserved for every predictor variable of every derived annoyance model.

Stored within the *...Analysis Results\Plots* folder:

- `model_#_mean_regcoeff_plot`: This is a MATLAB figure (.fig file) which contains a plot of the data in `model_#_mean_regcoeff`.
- `model_#_mean_deviation_regcoeff_plot`: This is a MATLAB figure (.fig file) which contains a plot of the data in `model_#_mean_deviation_regcoeff`.
- `model_#_bias_plot`: This is a MATLAB figure (.fig file) which contains a plot of the data in `model_#_bias`.
- `model_#_standard_deviation_plot`: This is a MATLAB figure (.fig file) which contains a plot of the data in `model_#_standard_deviation`.
- `model_#_mean_deviation_regcoeff_standard_deviation_plot`: This is a MATLAB figure (.fig file) which contains a plot of the data in `model_#_mean_deviation_regcoeff_standard_deviation`.
- `model_#_mean_regcoeff_standard_deviation_plot`: This is a MATLAB figure (.fig file) which contains a plot of the mean regression coefficients with error bars. Data for the plot are from `model_#_mean_regcoeff` and `model_#_standard_deviation`.
- `model_#_sampling_state_#_variable_#_sampling_distribution_plot`: This is a MATLAB figure (.fig file) which contains a plot of the sampling distribution of the regression coefficients across all Monte Carlo trial for a certain predictor variable of a derived annoyance model.
- `population_annoyance_model_contour`: This is a MATLAB figure (.fig file) which contains a plot of contours of the annoyance generated by the population annoyance model.

- `population_annoyance_model_contour_DNL_overlay`: This is a MATLAB figure (.fig file) which contains a plot of contours of the annoyance generated by the population annoyance model with *DNL* contours overlaid.
- `model_#_sampling_state_#_annoyance_diagram`: This is a MATLAB figure (.fig file) which contains a plot of contours of the annoyance from a derived annoyance model for a given sample design.
- `model_#_sampling_state_#_annoyance_diagram_DNL_overlay`: This is a MATLAB figure (.fig file) which contains a plot of contours of the annoyance from a derived annoyance model for a given sample design with *DNL* contours overlaid.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Survey Sampling Module

I.9 Variance of a Ratio of Regression Coefficients

The program described in this section estimates the variance of a ratio of multiple linear regression coefficients from a model predicted from survey data. The program can estimate the variance using two methods: the jackknife repeated replication technique and Taylor series linearization. The variance estimation methods are described in detail in Appendix B and Chapter 4, Section 4.1.3. Using either method, the program estimates the variance assuming a simple random sample design. If strata and primary sampling unit identifiers are provided for a complex sample design with stratification and clustering, complex variance estimates can be calculated. If a data set is composed of multiple airports, the complex variance is estimated with airport as the cluster identifier. There is one function:

1. *REG_COEFF_RATIO_VARIANCE*: This is the only function of the program. It estimates the simple random sample variance of a ratio of regression coefficients and will also estimate the complex variance if information about the complex sample design (strata and cluster identifiers) are provided.

Inputs: The function can accept 5, 7, or 8 inputs. If 5 inputs is provided, only the simple random sample variance will be estimated. The complex variance can be estimated using airport as a cluster identifier (7 inputs) or strata/primary sampling unit information (8 inputs).

For 5 inputs, the inputs are as follows:

- `criterion_variable`: This variable is a n -by-1 criterion variable vector, where n is the number of samples.
- `regression_matrix`: This variable is a n -by- p matrix of predictor variables, where p is the number of predictor variables.
- `method`: This variable is the name (enclosed by single quotes) of the variance estimation method. The possible choices are 'jackknife' for jackknife repeated replication and 'taylor' for Taylor series linearization.
- `column_numerator`: This variable is a column number of `regression_matrix`, the regression coefficient estimate of which will be the numerator variable in the ratio of regression coefficients.
- `column_denominator`: This variable is a column number of `regression_matrix`, the regression coefficient estimate of which will be the denominator variable in the ratio of regression coefficients.

For 7 inputs, the inputs are as follows:

- `criterion_variable`
- `regression_matrix`
- `method`
- `column_numerator`
- `column_denominator`
- `airport_id`: This variable is a n -by-1 vector that identifies the airport of each observation, where n is the number of samples.
- `weight_method`: This variable is the name (enclosed by single quotes) of the weighting method for calculating the jackknife coefficient of each jack-

knife replicate. The possible choices are ‘people’ (Equation (B.2)) or ‘psu’ (Equation (B.1)). See Appendix B.1 for details on the jackknife coefficient.

For 8 inputs, the inputs are as follows:

- `criterion_variable`
- `regression_matrix`
- `method`
- `column_numerator`
- `column_denominator`
- `airport_id`
- `weight_method`
- `data_structure`: This variable is a n -by-2 matrix with stratum identifiers in the first column and primary sampling unit identifiers in the second column.

Outputs: The number of outputs is dependent on the number of inputs. For 5 inputs (simple random sample analysis only) there are 2 outputs. For 7 or 8 inputs (complex analysis in addition to simple random sample analysis) there are 5 outputs.

For 5 inputs, the outputs are as follows:

- `out_srs`: This is a MATLAB cell of information, the exact contents of which are dependent on the variance estimation method used. Common to both methods, the cell includes the regression coefficient estimates, the ratio of regression coefficient estimates, the estimated simple random sample variance and standard deviation of the ratio, the degrees of freedom, and the estimated 95% confidence interval for the ratio of regression coefficient estimates. For jackknife repeated replication, the cell includes the replicate matrix. For Taylor series linearization, the cell includes the variances and covariance estimates of the regression coefficients from the ordinary least squares covariance matrix.

- `stdev_ratio_srs`: This is the calculated standard deviation of the ratio of regression coefficient estimates, also reported in `out_srs`.

For 7 or 8 inputs, the outputs are as follows:

- `out_srs`:
- `stdev_ratio_srs`:
- `out_complex`: This is a MATLAB cell of information, the exact contents of which are dependent on the variance estimation method used. Common to both methods, the cell includes the regression coefficient estimates, the ratio of regression coefficient estimates, the estimated complex variance and standard deviation of the ratio, the degrees of freedom, and the estimated 95% confidence interval for the ratio of regression coefficient estimates. For jackknife repeated replication, the cell includes the replicate matrix. For Taylor series linearization, the cell includes the output from the one-way random effects analysis of variance of the residuals, an estimate of the within primary sampling unit variance component, an estimate of the the between primary sampling unit variance component, and the estimated variances and covariance of the regression coefficients from Equations (B.22), (B.23), and (B.24).

*Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\
MULTIPLE USE*

I.10 Multicollinearity

The program described in this section utilizes techniques described in (Belsley, 1991a,b) to calculate diagnostics which measure the magnitude of multicollinearity between predictor variables in a statistical model. Multicollinearity is described in detail in Appendix C. For a brief description of the techniques, see (Belsley, 1991b). There is one main function and two subfunctions:

1. *A_COLLINEARITY_INPUTS*: This function processes the inputs and calls the subfunctions *b_collinearity_diagnostic* and *c_collinearity_harmful*.

Inputs:

- **criterion_variable**: This variable is a n -by-1 criterion variable vector, where n is the number of samples.
- **regression_matrix**: This variable is a n -by- p matrix of predictor variables, where p is the number of predictor variables.
- **labels**: This variable is a 1-by- p MATLAB cell that contains labels for the predictor variables in **regression_matrix**.
- **e_scale_columns**: This variable is a 1-by- m vector that indicates positions of predictor variables in **regression_matrix** that are logged (base 10) transformations of “structurally interpretable” variables. See (Belsley, 1991a, p. 183) for a discussion of structural interpretability.

Outputs:

- **results**: This variable stores intermediate and final products from the calculation of the multicollinearity diagnostic. The main output contained in the variable is a signal-to-noise analysis which gives a conclusion of each predictor variable’s involvement in a collinear relation, at different levels of stringency of the statistical hypothesis test used.
2. *b_collinearity_diagnostic*: This subfunction analyzes **regression_matrix** and determines which predictor variables are involved in “near dependencies”, where a near dependency is the existence of a near linear relationship among predictor variables of **regression_matrix** and which is indicated by the magnitude of a scaled condition index of **regression_matrix**.
 3. *c_collinearity_harmful*: This subfunction analyzes the signal-to-noise ratio of each predictor variable and performs a statistical hypothesis test to conclude each predictor variable as experiencing non-harmful collinearity, harmful collinearity, a problem of short data, or no problem. The statistical hypothesis test is implemented at different levels of stringency.

*Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\
MULTIPLE USE*

I.11 Extra Sum of Squares

The program described in this section calculates the extra sum of squares for predictor variables in a multiple linear regression model. The methodology is described in detail in Chapter 5, Section 5.2.4. There is one main function:

1. *EXTRA_SUM_SQUARES*: This function calculates the extra sum of squares for each predictor variable in a multiple linear regression model and tests the variable significance with the general linear test. It calculates the average semi-partial R^2 ($\overline{R_{sp}^2}$) for each predictor variable and uses bootstrapping to define an interquartile range of variation for each variable's $\overline{R_{sp}^2}$. The interquartile range spans from the 25th percentile to the 75th percentile. See Chapter 5, Section 5.2.4 for details on the extra sum of squares and Chapter 4, Section 4.1.4 for details on the $\overline{R_{sp}^2}$ and bootstrapping.

Inputs:

- **criterion_variable**: This variable is a n -by-1 criterion variable vector, where n is the number of samples.
- **regression_matrix**: This variable is a n -by- p matrix of predictor variables, where p is the number of predictor variables.
- **labels**: This variable is a 1-by- p MATLAB cell that contains labels for the predictor variables in **regression_matrix**.
- **bootstrap_samples**: This variable is a number that specifies the number of bootstrap samples to draw from the full data set in the calculation of the interquartile ranges of the $\overline{R_{sp}^2}$.

Outputs:

- **marginal_results**: This variable stores, for each predictor variable, the calculated extra sum of squares, statistical hypothesis test results of the variable's significance, and the $\overline{R_{sp}^2}$ and the interquartile range.

*Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\
MULTIPLE USE*

I.12 Census Block Group Identification

The program described in this section determines the census block group identifier of every grid point in a matrix. There is one main function:

1. *CENSUS_BLOCK_GROUP_COUNT*:

Inputs:

- **S**: This is a n -by-1 structure that was output from the MATLAB function *shaperead*, where n is the number of census block groups, and *shaperead* was used to read the shapefile data of census block groups in geographical areas surrounding an airport. The shapefile data is described in Chapter 5, Section 5.2.4.
- **lat**: This variable is a matrix of latitudes for each grid point.
- **long**: This variable is a matrix of longitudes for each grid point.

Outputs:

- **census_block_groups**: This variable is a matrix that gives the census block group identifier of every grid point at which acoustical data was calculated within *NOISE_METRIC_MODULE*. *NOISE_METRIC_MODULE* is described in Appendix I.1.

*Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\
MULTIPLE USE*

I.13 Data Set Combination

The program described in this section combines survey data sets and performs a multiple linear regression analysis on each combined data set, using sound level and number-of-events data from the survey data sets. Details of the combination process are described in Chapter 3, Section 3.5. One combination data set is realized by adjusting the annoyance variable of each individual data set involved in order to achieve a common annoyance scale among the data sets. The annoyance variables are adjusted by one of the techniques described in Chapter 3, Section 3.5.1. The number of combination data sets created is dependent on the number of unique annoyance scales among the individual data sets. There is one main function and seven subfunctions:

1. *COMBINATION*: This is the main function of the program. This function processes the input if provided input, or asks the user for input if not provided input. The data sets to be combined are prepared prior to combination in that criterion variable values nonadjacent to any other values can be deleted (which can indicate that the value is a missing or non-recorded value) and each survey data set and each airport within survey data sets is assigned a unique primary sampling unit identifier for complex variance estimation of \hat{k} . The Miedema scaling and combination is performed. If all data set criterion variable scales have the same number of categories, they are combined in this function. This function directs control to the other subfunctions.

Inputs: There is one optional input. The input is a text file that contains responses to questions that the function would normally prompt the user to answer, without the text file as input. If *COMBINATION* is run without an input, then the function will prompt the user for responses to a number of questions, and then create a text file with the recorded responses. The text file should be stored in the following relative folder path: `...\SURVEY DATA\Data Set Combination\Input Files`. An example of the text file is:

Rescaled minimum: 0
Rescaled maximum: 10
Variance method flag(s): 1 2 2
Variance weighting method flag(s): 0 1 2
Number of data sets: 5
Number of independent variables: 2
Data set name: TACORG2_STNDRD.mat
Dependent variable column: 17
Independent variable column(s): 95 133
Independent variable form(s): 1 4
Individual data set rescaling, number of categories:
Data set name: UKD024_STNDRD.mat
Dependent variable column: 2
Independent variable column(s): 61 117
Independent variable form(s): 1 4
Individual data set rescaling, number of categories:
Data set name: UKD130_STNDRD.mat
Dependent variable column: 2
Independent variable column(s): 61 117
Independent variable form(s): 1 4
Individual data set rescaling, number of categories:
Data set name: AUL210_STNDRD.mat
Dependent variable column: 3
Independent variable column(s): 111 133
Independent variable form(s): 1 4
Individual data set rescaling, number of categories:
Data set name: UKD604_STNDRD.mat
Dependent variable column: 3
Independent variable column(s): 96 134

Independent variable form(s): 1 4
 Individual data set rescaling, number of categories:
 Sound level variable position: 1
 Individual data set analysis indicator: 0
 Combination indicator: 1

A description of the text file items follows:

- **Rescaled minimum:** For each combination data set created and/or individual data set, two multiple linear regression analyses are performed: one with the criterion variable which spans a range of values identical to that of one or more of the original data sets, and one with a scaled criterion variable which is adjusted to span a range specified by the user. **Rescaled minimum** specifies the minimum of the rescaled criterion variable. The rescaling is useful so that all data sets, no matter the individual ranges which they span, can be normalized to span the same range so that direct comparisons can be made of regression results.
- **Rescaled maximum:** This item specifies the maximum of the rescaled criterion variable.
- **Variance method flag(s):** This item specifies the variance estimation methods to use to estimate the complex variance of a ratio of regression coefficients, with the sound level regression coefficient estimate in the denominator of the ratio, as in \hat{k} , the decibel equivalent number effect. The Data Set Combination Program is not limited to using only sound level and number-of-events predictor variables. However, a ratio of regression coefficients will only be calculated if there is a sound level predictor variable present. In this way the ratio indicates the number of decibels which have an effect equivalent to a unit increase in another variable. The choices for this item are 1 (Taylor series linearization) or 2 (jackknife repeated replication). The complex variance of the ratio is calculated with the program described in Appendix I.9. In the example provided, the jackknife method

is listed twice. This relates to the next item in the text file, **Variance weighting method flag(s)**.

- **Variance weighting method flag(s)**: In the variance estimation program, two possibilities exist for the jackknife method: either primary sampling units can be equally weighted (Equation (B.1)) or persons can be equally weighted (Equation (B.2)) in the calculation of the jackknife coefficient. See Appendix B.1 for details on the jackknife coefficient. This item specifies, for each variance method flag, the weighting method to use. The choices for this item are 0 (neither method, for the Taylor method only), 1 (equal weight to persons), or 2 (equal weight to primary sampling units).
- **Number of data sets**: This item specifies the number of data sets that are listed in the input text file. Items **Data set name**, **Dependent variable column**, **Independent variable column(s)**, **Independent variable form(s)**, and **Individual data set rescaling**, **number of categories** should be listed for each data set to be specified. Prior to listing the data sets, the item **Number of independent variables** is listed.
- **Number of independent variables**: This item specifies the number of predictor variables to be involved in the multiple linear regression analyses. An intercept is automatically added within the program. Whereas the criterion variables of the individual data sets that are to be combined can be different in the number of categories of each, the predictor variable data should be consistent across data sets. For example, $PNL_{max,av}$ data should not be combined with $L_{Amax,av}$ data.
- **Data set name**: This item specifies the name of a data set MATLAB .mat file which has a data set saved within it and is stored in a directory that can be accessed by the program. The data set should be stored in the following relative folder path: `... \SURVEY DATA \Data Sets`. The data set should contain all data necessary for a multiple linear regression analysis.

Although not required for this program, the data sets should be adopted to the standardized storage format defined in Appendix D.

- **Dependent variable column:** This item specifies the column number of the criterion variable in the data set specified by **Data set name**. Each criterion variable should be composed of a select number of integers, where each integer represents a level of disturbance with definite meaning. An example of a criterion variable that should not be specified is an activity-interference annoyance scale in which for each observation, the annoyance score is a sum of ratings of the interference caused by noise to a number of activities. With this type of scale, the annoyance score can span a large range and each score does not indicate a particular level of disturbance.
- **Independent variable column(s):** This item specifies the column number(s) of the predictor variables in the data set specified by **Data set name**.
- **Independent variable form(s):** This item specifies, for each predictor variable, how the predictor variable data should be modified prior to the regression analyses. The possibilities are 1 (the variable data is raised to a power of 1, which leaves it unmodified), 2 (the variable data is raised to a power of 2), 3 (the variable data is raised to a power of 3), or 4 (the variable data is base-10 logged).
- **Individual data set rescaling, number of categories:** This item specifies an integer that the criterion variable in the data set specified by **Data set name** should be adjusted to span to (from 1) for a multiple linear regression analysis. In the combination process, criterion variables are adjusted for individual data sets, but regression results for the individual data sets are not calculated. The specification of this item allows one to explore how reducing or expanding the range of a criterion variable for one data set affects the regression results for that data set. This scaling is different than that specified by the **rescaled minimum** and **rescaled maximum** items. This

item does not have to be specified and can be left blank after the prompt in the text file.

- **Sound level variable position:** This item specifies the position of the sound level variable in the ordering of predictor variables specified at the **Independent variable column(s)** item. This item is not repeated for each data set. Thus, the predictor variables should be listed in the same order for all data sets.
- **Individual data set analysis indicator:** This item specifies whether or not the raw data of the individual data sets should be used in their own multiple linear regression analyses. The choices are 0 (no) or 1 (yes).
- **Combination indicator:** This item specifies whether or not the individual data sets should be combined. The choices are 0 (no) or 1 (yes).

Outputs:

- **results:** This variable reports results for the combination and/or individual data sets.
 - **results_rescaled:** This variable reports results for the combination and/or individual data sets in which the criterion variables have been scaled to span **rescaled minimum** and **rescaled maximum**.
2. *scenario_2_or_3:* This subfunction is used in the combination data set analysis if the individual data sets differ in the number of criterion variable categories. In this subfunction, the criterion variable data of each data set are scaled to achieve a common annoyance scale. All possibilities of establishing a common annoyance scale are implemented within the subfunction.
 3. *combination_generator:* This subfunction determines the number of possibilities for combining criterion variables of different data sets.
 4. *regression_call:* This subfunction calls the subfunction *regression* and scales the criterion variable with **rescaled minimum** and **rescaled maximum** for regression with the the rescaled criterion variable.

5. *regression*: This subfunction performs a regression analysis of the data. It calls the *REG_COEFF_RATIO_VARIANCE* function described in Appendix I.9 to estimate the variance of a ratio of regression coefficients and the *EXTRA_SUM_SQUARES* function described in Appendix I.11 to calculate the extra sum of squares of predictor variables.
6. *plotter*: This subfunction is used to visualize the variation in jackknife repeated replication replicate estimates of a ratio of regression coefficients. It plots the replicate estimates of the ratio and overlays a normal distribution with a mean value equivalent to the mean of the replicate estimates of the ratio and a variance that is equivalent to the jackknife variance estimate of the ratio.
7. *individual_dataset_rescale*: This subfunction is used to adjust the criterion variable of individual data sets with the techniques that are used on the combination data sets.
8. *combinator*: This subfunction is used to perform permutations. It assists in calculating the number of category combination strategies that can be applied to a criterion variable (see Chapter 3, Section 3.5.1). The function was authored by Matt Fig (MathWorks, Inc., 2009). Other subfunctions included with and used by *combinator* include *perms_rep*, *perms_no_rep*, *perms_loop*, *combs_rep*, *combs_no_rep*, and *cumsum2*.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY DATA\Data Set Combination

I.14 UKD-604 Data Set Simulation

The program described in this section simulates a person-level data set from site-level data reported for the UKD-604 noise survey (Le Masurier et al., 2007a). Details of the simulation process are described in Chapter 3, Section 3.3. There is one main function and one subfunction:

1. *UKD604_DATASET_CREATOR*: This function simulates a person-level data set.

Inputs: There are no function inputs. However, the function reads Microsoft .xlsx data files that are stored in the same directory as the program functions.

- **Extracted Appendix A6.3 tables.xlsx**: This is a Microsoft .xlsx file that contains *Lav*, *Nav*, and annoyance data data from Appendix A6.3 of the UKD-604 documentation (Le Masurier et al., 2007a).
- **Site standard deviations.xlsx**: This is a Microsoft .xlsx file that contains standard deviations of *Lav* (σ_{Lav}) and the normalized standard deviation of *Nav* (σ_{Nav}^*) for the variation of site-level *Lav* and *Nav* data, as part of the simulation process. The standard deviations were calculated with the use of *CENSUS_BLOCK_GROUP_VARIATION*, which is described in Appendix I.6. The methodology used to calculate the standard deviations of each site is described in Chapter 3, Section 3.3.

Outputs: There are no function outputs. However, the function writes an output file to hard drive in the program directory.

- **ANASE_results**: This is a MATLAB .mat file that stores the person-level data set, diagnostic output from the population design algorithm, regression results of the final person-level and site-level annoyance models, sensitivity analyses of the fits of the final person-level and site-level annoyance models to the data as the decibel equivalent number effect (\hat{k}) of each is varied by changing the regression coefficients, and variance estimates of \hat{k} from *REG_COEFF_RATIO_VARIANCE*. *REG_COEFF_RATIO_VARIANCE* is described in Appendix I.9.
- **annoyance_compare_plot**: This is a MATLAB .fig file that stores a plot of mean annoyance data, as described by Le Masurier et al. (2007a, p. 7.2, Section 7.2.4). The plot displays the actual UKD-604 mean annoyance data versus the mean annoyance data as predicted by the annoyance model reported in the UKD-604 documentation (Le Masurier et al., 2007a, p. 9.14,

Model 9.3). Model 9.3 was used with the averages of the simulated L_{av} and N_{av} data within each site to generate mean annoyance predictions. Note that Model 9.3 of (Le Masurier et al., 2007a) was derived using the site-level L_{av} and N_{av} data, where each site-level observation was that estimated at the population-weighted centroid of the site.

- `annoyance_compare_plot_simulated`: This is a MATLAB .fig file that stores a plot of mean annoyance data. The plot displays the actual UKD-604 mean annoyance data versus the mean annoyance data as predicted by an annoyance model derived from the actual UKD-604 mean annoyance data and the averages of the simulated L_{av} and N_{av} data within each site.
- `mean_annoyance_plot`: This is a MATLAB .fig file that stores a three-dimensional plot of the actual UKD-604 mean annoyance data versus the averages of the simulated L_{av} and N_{av} data within each site.

2. `structure_adapt`: This subfunction separates the person-level data set designed by `UKD604_DATASET_CREATOR` into airport-specific data sets and saves all data sets to hard drive (full and airport-specific). The data sets are formatted according to the standardized storage format defined in Appendix D.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY DATA\Data Sets\UKD604

I.15 AUL-210 Data Set Simulation

The program described in this section was used to transform DNL to $PNL_{max,av}$ in the AUL-210 noise survey data set (Hede and Bullen, 1982). Two different, but mathematically-equivalent formulations were used; slightly different results were realized due to the differences in computation coupled with rounding error. Regression results reported in Chapters 3 and 4 were derived using the $PNL_{max,av}$ data which resulted from the second method of transformation, which is documented in Chapter 3, Section 3.1.1. The program also splits the full data set into five airport-specific

data sets. Details of the transformation process are described in Chapter 3, Section 3.1.1. There is one main function:

1. *AUL210_DATASET_CREATOR*: This function transforms *DNL* to $PNL_{max,av}$ using two methods of transformation.

Inputs:

- **below_70_storage_level**: This variable specifies the $L_{Amax,av}$ level of aircraft events below 70 dBA from which the $PNL_{max,av}$ to be stored is calculated. Within the function, ten different transformations of *DNL* to $PNL_{max,av}$ are computed by assuming ten different levels of the $L_{Amax,av}$ for aircraft events with $L_{Amax} < 70$ dBA. The $L_{Amax,av}$ levels range from 60 to 69 dBA in 1 dBA steps. However, the $PNL_{max,av}$ data from only one transformation are stored. This variable determines which $PNL_{max,av}$ data should be stored. The regression results reported in this document were derived using $PNL_{max,av}$ data calculated with an assumption of an $L_{Amax,av} = 65$ dBA for events with $L_{Amax} < 70$ dBA.

Outputs: There is one function output. In addition, the function saves data sets to hard drive for the full survey sample and airport-specific data sets for the five airports of the AUL-210 study. The data sets are formatted according to the standardized storage format defined in Appendix D.

- **regression_results**: This variable stores results from multiple linear regressions of the annoyance variable on $PNL_{max,av}$ and $\log_{10}(N)$. The results from ten regressions are reported, where each regression used $PNL_{max,av}$ data resulting from a different $L_{Amax,av}$ for aircraft events with $L_{Amax} < 70$ dBA.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY DATA\Data Sets\AUL210

I.16 Plotting Program

The codes described in this section can be used to visualize the sound level, number-of-events, and annoyance data of survey data sets. The plotting program is described in Chapter 3, Section 3.4. The program is based on a series of Graphical User Interfaces (GUIs). Four of the subfunctions have associated GUIs. There is one main function and eight subfunctions:

1. *A_PLOTS_INITITAL*: This is the main function of the program. This function defines the color scheme to use in the plots and calls *b_plots*.

Inputs: There is one optional input. The input is a text file that contains plotting choices that would be selected using a Graphical User Interface if the text file was not provided. The input file should be stored in the following relative folder path: ...*SURVEY DATA*\Plotting Program*Input Files*. An example of the text file is:

```
Number of data sets: 2
Data set name: UKD024_STNDRD.mat
Dependent variable column: 2
Independent variable column(s): 61 117
Data set name: UKD130_STNDRD.mat
Dependent variable column: 2
Independent variable column(s): 61 117
```

A description of the text file items follows:

- **Number of data sets**: This item specifies the number of data sets that are listed in the input text file. Items **Data set name**, **Dependent variable column**, and **Independent variable column(s)** should be listed for each data set to be specified.
- **Data set name**: This item specifies the name of a data set MATLAB .mat file which has a data set saved within it and is stored in a directory that

can be accessed by the program. The data set should be stored in the following relative folder path: `...\SURVEY DATA\Data Sets`. The data set should contain all data necessary for a multiple linear regression analysis. Although not required for this program, the data sets should be adopted to the standardized storage format defined in Appendix D.

- **Dependent variable column:** This item specifies the column number of the criterion variable in the data set specified by `Data set name`. The criterion variable should be numerical.
- **Independent variable column(s):** This item specifies the column number(s) of the predictor variables in the data set specified by `Data set name`. The predictor variables should be numerical.

Outputs: There are no function outputs. However, the plotting program writes output files to hard drive.

- **PREFIX_BALANCED:** This is a MATLAB `.mat` file that stores information about the design of treatment boundaries for a *balanced* sample allocation attempt for a certain data set. The specified treatment boundaries, the optimal sample allocation, the actual sample allocation, and output from *plots_statistical_power* including the designed minimum factor level and treatment sample sizes are stored. **PREFIX** is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in the following relative folder path: `...\SURVEY DATA>Noise Plotting Program\Results\Balanced`.
- **PREFIX_EQUAL:** This is a MATLAB `.mat` file that stores information about the design of treatment boundaries for an *equal* sample size allocation attempt for a certain data set. The specified treatment boundaries, the optimal sample allocation, the actual sample allocation, and output from *plots_statistical_power* including the designed minimum factor level and treatment sample sizes are stored. **PREFIX** is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in

the following relative folder path: ...*SURVEY DATA*\Noise Plotting Program*Results*\Equal.

- *PREFIX*_sample_dist: This is a MATLAB .mat file that stores a matrix that lists the number of data points contained in each treatment. *PREFIX* is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in the following relative folder path: ...*SURVEY DATA*\Noise Plotting Program*Results*.
 - *PREFIX*_plot_mat: This is a MATLAB .mat file that stores the raw data contributing to each treatment in separate matrices. *PREFIX* is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in the following relative folder path: ...*SURVEY DATA*\Noise Plotting Program*Results*.
 - *PREFIX*_reg_results: This is a MATLAB .mat file that stores regression results for a balanced sample allocation attempt from the function *balanced_anova*. *PREFIX* is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in the following relative folder path: ...*SURVEY DATA*\Noise Plotting Program*Results*.
 - *PREFIX*_3D_lineplot: This is a MATLAB .fig file that stores the plot of one or more data sets. The plot is not saved automatically, but is saved through the manipulation of the GUI created by *plots_change_plot_GUI_1*, which displays the plot. *PREFIX* is the prefix of the file name and is constructed from a combination of plotting choices. The file is stored in the following relative folder path: ...*SURVEY DATA*\Noise Plotting Program*Results*.
2. *b_plots*: This subfunction calls *plots_setup_GUI*, which opens a GUI, and processes the user input entered into the GUI. It calls *d_plots_create_matrices* and *e_plots_plotting*.
 3. *plots_setup_GUI*: This subfunction creates the main program GUI and stores the user input choices. It calls *plots_change_plot_GUI_2* in the case that custom factor level boundaries are desired to be specified.

4. *plots_change_plot_GUI_1*: This subfunction creates a GUI in which the plots are displayed and modifies the plots according to user input. It calls *plots_change_plot_GUI_2* in the case that the factor level boundaries are desired to be changed.
5. *plots_change_plot_GUI_2*: This subfunction creates a GUI that allows the user to specify factor level boundaries.
6. *plots_statistical_power*: This subfunction creates a GUI that allows the user to design minimum factor level and treatment sample sizes in either instance that a balanced or equal sample allocation is desired. The subfunction utilizes the power and estimation methods described in Chapter 3, Section 3.4.
7. *balanced_anova*: This subfunction is called by *d_plots_create_matrices* in the case of a balanced sample allocation attempt. It implements the regression matrix weighting procedure described in (Keren and Lewis, 1993, p. 94-127) and calculates the Type III sums of squares for the factors.
8. *d_plots_create_matrices*: This subfunction separates the data set data into treatments based on the factor level boundary choices. It calls *balanced_anova* in the case of a balanced sample allocation attempt.
9. *e_plots_plotting*: This subfunction plots the data.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY DATA\Plotting Program

I.17 Flight Track Creation

The codes described in this section can be used to create flight tracks in INM. Two different methods of creating flight tracks were described in Appendix E. Flight tracks can be created from flight track maps (Appendix E.1) or from flight track data available on the Internet (Appendix E.2). The codes for each of these methods will be described:

Flight Track Maps: There are two main functions:

1. *COORDINATE_MATRIX_CREATOR*: This function creates x and y coordinate matrices for a flight track map. Appendix E.1, point 1 to point 4f, should be reviewed prior to using this function.

Inputs:

- **city**: This variable is the name (enclosed by single quotes) of a city which specifies the name of a folder stored in the directory of *COORDINATE_MATRIX_CREATOR* and *FLIGHTTRACK_POINT_GENERATOR* from which an image file will be read and in which the output will be saved. The folder should be named after the city that the airport resides within, prior to running the function.
- **image**: This variable is the file name (enclosed by single quotes, with extension) of an image. The image should be a white Photoshop layer with the origin point of the INM grid marked with a RGB=[0,254,0] single-pixel dot.
- **pixel_separation**: This variable is the number of pixels of a flight track map in Adobe Photoshop which is equivalent to the distance between INM grid points.
- **distance_separation**: This variable is a distance (in either nautical miles or kilometers) that separates INM grid points.

Outputs: There are no function outputs. However, the function saves x and y coordinate matrices to hard drive in the folder identified by **city**.

2. *FLIGHTTRACK_POINT_GENERATOR*: This function extracts x and y spatial locations of flight track points from the coordinate matrices created by *COORDINATE_MATRIX_CREATOR*. Appendix E.1, point 4g to point 4(g)ii, should be reviewed prior to using this function.

Inputs:

- **city:** This variable is the name (enclosed by single quotes) of a city which specifies the name of a folder stored in the directory of *COORDINATE_MATRIX_CREATOR* and *FLIGHTTRACK_POINT_GENERATOR* from which an image file will be read.
- **image:** This variable is the file name (enclosed by single quotes, with extension) of an image. The image should be a white Photoshop layer with flight track points marked with RGB=[254,0,0] single-pixel dots.
- **x_coord:** This variable is a matrix of *x*-coordinates created by *COORDINATE_MATRIX_CREATOR*.
- **y_coord:** This variable is a matrix of *y*-coordinates created by *COORDINATE_MATRIX_CREATOR*.

Outputs:

- **points:** This variable is a *n*-by-2 matrix that contains extracted *x*-coordinates in the first column and *y*-coordinates in the second column.
- **x:** This variable is the first column of **points**, the values of which are limited to four decimal points. This is to facilitate input into the *trk_segs.dbf* INM file.
- **y:** This variable is the second column of **points**, the values of which are limited to four decimal points. This is to facilitate input into the *trk_segs.dbf* INM file.

Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY SIMULATION PROGRAM\Flight Track Creation\Flight Track Maps

Flight Track Data: There is one main function and five subfunctions:

1. *CSV_CREATOR:* This function reads the .kml files from a hard drive folder identified by the inputs to the function and changes the extensions of the files to .xml (.kml files are essentially .xml files). The Google Earth files should be

categorized by airport or city, operation type (arrival or departure), runway, and direction of origin (for arrivals) or destination (for departures). See point 1 in Appendix E.2 for an example of the correct folder structure. After calling the subfunctions, this function writes the data from Google Earth .kml files into a .csv text file that INM can import as a radar track file.

Inputs:

- **city**: This variable is the name (enclosed by single quotes) of a city which specifies the name of a folder which stores .kml files. This folder should be stored in the directory of *CSV_CREATOR*.
- **optype**: This variable is the name (enclosed by single quotes) of a folder stored in a **city** folder. **optype** refers to the operation type (arrival or departure). The names of the arrival and departure folders are to be specified by the user prior to running the function.
- **runway**: This variable is a runway name (enclosed by single quotes) which specifies the name of a folder stored within a **optype** folder. The folder name is to be specified by the user prior to running the function.
- **direction**: This variable is the name (enclosed by single quotes) of a folder stored within a **runway** folder. **direction** refers to the direction of origin (for arrivals) or destination (for departures). The folder name is to be specified by the user prior to running the function.

Outputs: There are no function outputs. However, the function writes and saves a .csv file to hard drive in the folder identified by the function inputs.

2. *xml_to_struct*: This subfunction calls *parseXML*. The next four subfunctions were obtained from the Internet. They were given by The MathWorks, Inc. as an example of an application of the function *xmlread* (MathWorks, Inc.). The purpose of the subfunctions are to extract the information from a .xml file and store it within a MATLAB structure.
3. *parseXML*: This subfunction runs *xmlread* which returns an XML Document Object Model node.

4. *parseChildNodes*: This subfunction operates recursively for each child node of the XML Document Object Model node.
5. *makeStructFromNode*: This subfunction extracts information from a child node and organizes it into a structure.
6. *parseAttributes*: This subfunction creates a structure of the child node attributes.

*Location: PURDUE-Kevin Foertsch\FINAL CODES AND DATA SETS\SURVEY
SIMULATION PROGRAM\Flight Track Creation\Flight Track Data*

APPENDIX J. SURVEY SIMULATION RESULTS

In this appendix are given tables of simulation-level statistics for survey simulations around Airport 1, 2, and 3. Survey simulations were described in Chapter 5. The simulation-level statistics are described in Chapter 5, Section 5.2.4. For each simulation, two multiple linear regression annoyance models were estimated from the sample collected in each Monte Carlo trial:

$$A_1 = b_L PNL_{max,av,80} + b_N \log_{10}(N_{80}) + b_{01}, \quad (\text{J.1})$$

$$A_2 = b_D DNL + b_{02}, \quad (\text{J.2})$$

where $PNLT_{max,av,80}$ is the $PNLT_{max,av}$ of aircraft events with $PNLT_{max} > 80$ PNdB, N_{80} is the number of daily aircraft events with $PNLT_{max} > 80$ PNdB, and DNL is the Day-Night Average Sound Level. In Tables J.1, J.2, J.3, J.4, and J.5 are shown simulation-level statistics for \hat{b}_L , \hat{b}_N , \hat{b}_{01} , \hat{b}_D , and \hat{b}_{02} , respectively. In the tables, “Full S” denotes the 6000-person sample resulting from stratified sampling of the population, “Full R” denotes the 6000-person sample resulting from simple random sampling of the population, “RSD” denotes the 2000-person sample extracted from the Full S sample using the stratified sample design, and “RR” denotes the 2000-person sample extracted from the Full S sample using simple random sampling. For the i^{th} predictor variable, $\overline{\hat{b}_i}$ is the mean of the regression coefficient estimates, calculated by Equation (5.20); $Bias_{\hat{b}_i}$ is the bias of $\overline{\hat{b}_i}$ from the true value, calculated by Equation (5.22); $S_{\hat{b}_i}$ is the standard deviation of the regression coefficient estimates, calculated by taking the square root of Equation (5.23); $MD_{\hat{b}_i}$ is the mean deviation of the standard deviation estimates of \hat{b}_i ($\hat{s}_{\hat{b}_i}$) from the individual Monte Carlo trials from $S_{\hat{b}_i}$, calculated by Equation (5.24). Also given are Type I and Type II error rates

(ER). The Type I error rate, for predictor variables not in the population model, gives the fraction of Monte Carlo trials in which the estimated confidence interval for \hat{b}_i did not include 0. The Type II error rate, for predictor variables in the population model, gives the fraction of Monte Carlo trials in which the estimated confidence interval for \hat{b}_i included 0.

Table J.1. Simulation-level statistics for \hat{b}_L .

Simulation	$\overline{\hat{b}_L}$	$\widehat{Bias}_{\hat{b}_L}$	$S_{\hat{b}_L}$	$MD_{\hat{b}_L}$	Type I ER	Type II ER
Airport 1, Full S	0.027	-0.053	0.002	0.000	-	0.00
Airport 1, RSD	0.027	-0.053	0.004	0.000	-	0.00
Airport 1, RR	0.027	-0.053	0.004	0.000	-	0.00
Airport 2, Full S	0.027	-0.053	0.002	0.000	-	0.00
Airport 2, Full R	0.018	-0.062	0.006	0.000	-	0.18
Airport 2, RSD	0.027	-0.053	0.004	0.001	-	0.00
Airport 2, RR	0.028	-0.052	0.005	0.000	-	0.00
Airport 3, Full S	0.021	-0.059	0.002	0.000	-	0.00
Airport 3, RSD	0.021	-0.059	0.004	0.001	-	0.01
Airport 3, RR	0.021	-0.059	0.004	0.001	-	0.00

Table J.2. Simulation-level statistics for \hat{b}_N .

Simulation	$\overline{\hat{b}_N}$	$\widehat{Bias}_{\hat{b}_N}$	$S_{\hat{b}_N}$	$MD_{\hat{b}_N}$	Type I ER	Type II ER
Airport 1, Full S	0.384	-0.736	0.020	0.011	-	0.00
Airport 1, RSD	0.383	-0.737	0.023	0.007	-	0.00
Airport 1, RR	0.383	-0.737	0.025	0.009	-	0.00
Airport 2, Full S	0.365	-0.755	0.017	0.008	-	0.00
Airport 2, Full R	0.253	-0.867	0.018	0.008	-	0.00
Airport 2, RSD	0.366	-0.754	0.020	0.006	-	0.00
Airport 2, RR	0.367	-0.753	0.020	0.005	-	0.00
Airport 3, Full S	0.291	-0.829	0.016	0.011	-	0.00
Airport 3, RSD	0.290	-0.830	0.016	0.008	-	0.00
Airport 3, RR	0.291	-0.829	0.017	0.009	-	0.00

Table J.3. Simulation-level statistics for \hat{b}_{01} .

Simulation	$\overline{\hat{b}_{01}}$	$\hat{Bias}_{\hat{b}_{01}}$	$S_{\hat{b}_{01}}$	$MD_{\hat{s}_{\hat{b}_{01}}}$	Type I ER	Type II ER
Airport 1, Full S	0.464	3.464	0.222	0.050	-	0.39
Airport 1, RSD	0.484	3.484	0.324	0.028	-	0.77
Airport 1, RR	0.464	3.464	0.373	0.076	-	0.76
Airport 2, Full S	0.488	3.488	0.213	0.009	-	0.55
Airport 2, Full R	1.549	4.549	0.477	0.022	-	0.16
Airport 2, RSD	0.535	3.535	0.320	0.055	-	0.87
Airport 2, RR	0.468	3.468	0.388	0.011	-	0.87
Airport 3, Full S	1.021	4.021	0.224	0.008	-	0.03
Airport 3, RSD	1.052	4.052	0.348	0.049	-	0.38
Airport 3, RR	1.045	4.045	0.372	0.027	-	0.43

Table J.4. Simulation-level statistics for \hat{b}_D .

Simulation	$\overline{\hat{b}_D}$	$\hat{Bias}_{\hat{b}_D}$	$S_{\hat{b}_D}$	$MD_{\hat{s}_{\hat{b}_D}}$	Type I ER	Type II ER
Airport 1, Full S	0.051	0.051	0.003	0.001	1.00	-
Airport 1, RSD	0.050	0.050	0.003	0.001	1.00	-
Airport 1, RR	0.051	0.051	0.003	0.001	1.00	-
Airport 2, Full S	0.057	0.057	0.003	0.002	1.00	-
Airport 2, Full R	0.052	0.052	0.003	0.002	1.00	-
Airport 2, RSD	0.057	0.057	0.003	0.002	1.00	-
Airport 2, RR	0.057	0.057	0.003	0.001	1.00	-
Airport 3, Full S	0.058	0.058	0.003	0.002	1.00	-
Airport 3, RSD	0.058	0.058	0.003	0.001	1.00	-
Airport 3, RR	0.058	0.058	0.003	0.002	1.00	-

Table J.5. Simulation-level statistics for \hat{b}_{02} .

Simulation	$\overline{\hat{b}_{02}}$	$\hat{Bias}_{\hat{b}_{02}}$	$S_{\hat{b}_{02}}$	$MD_{\hat{b}_{02}}$	Type I ER	Type II ER
Airport 1, Full S	0.381	3.381	0.167	0.095	-	0.07
Airport 1, RSD	0.392	3.392	0.196	0.071	-	0.29
Airport 1, RR	0.381	3.381	0.200	0.076	-	0.37
Airport 2, Full S	0.036	3.036	0.176	0.125	-	0.49
Airport 2, Full R	0.501	3.501	0.189	0.111	-	0.02
Airport 2, RSD	0.056	3.056	0.201	0.111	-	0.69
Airport 2, RR	0.022	3.022	0.193	0.103	-	0.67
Airport 3, Full S	-0.219	2.781	0.194	0.140	-	0.23
Airport 3, RSD	-0.217	2.783	0.192	0.097	-	0.41
Airport 3, RR	-0.217	2.783	0.190	0.096	-	0.45