Aircraft noise dose-response relations for national parks

Grant S. Anderson[a], Amanda S. Rapoza[b], Gregg G. Fleming[c] and Nicholas P. Miller[d]

(Received: 12 April 2011; Revised: 14 July 2011; Accepted: 14 July 2011)

An analysis of visitor survey responses and concurrent noise exposure was performed using data from ten sites in four scenic U. S. National Parks. Data collection was structured to learn the effects of air-tour aircraft noise and high-altitude jet noise on the experience of park visitors at scenic overlooks and on short hikes. The analysis utilized multilevel logistic regression and resulted in six dose-response relations: two responses (annoyance and interference with natural quiet), paired with three response dichotomizations (slightly or more, moderately or more, and very or more). Each of those six relations retained the same set of regression predictors. Individual-visitor Leq from all aircraft (averaged over the visitor duration at the site) proved to be the most reliable/accurate predictor of all noise dose metrics tested. The relation with visitor Leq was significantly strengthened by inclusion of three additional dose-related predictors: the energy-percentage due to tour helicopters for each visitor, the same due to fixed-wing tour aircraft, and the interaction of these two percentages. In addition, the relation was also strengthened by inclusion of the following context variables: Scenic overlook or short hike, natural quiet very important (or not) to that visitor, visitor group includes only adults (or not), and first-time visit at that site (or not). For a given noise exposure, visitors expressed more negative response regarding interference with natural quiet than regarding annoyance. In addition, visitor response to a given dose of air-tour noise was less severe when there were low-to-moderate levels of high-altitude jet noise present. © ***2011 Institute of Noise Control Engineering*.**

Primary subject classification: 68.3; Secondary subject classification: 75.1

1 INTRODUCTION

In the United States, the need to predict public annoyance to transportation noise was fully realized when military jet aircraft were first flown in the 1950s[1]. Subsequently, the advent of civilian air travel resulted in dozens of social surveys around the world, mostly concerning community response to aircraft noise, during the 1960s and 1970s[2].

Various attempts to generalize these surveys—both their data sources and their analysis methods—culminated when Theodore Schultz converted their various dose measures to a common noise metric and harmonized their responses[3]. The resulting "Schultz dose-response curve" relates day-night average sound level (DNL, denoted by the symbol L_{dn}) to the percent of the surveyed residential population that reported high annoyance. Schultz's analysis was later supplemented with additional data, to result in a dose-response curve that still underlies many current federal transportation-noise findings and regulations[4–10]. This supplemented dose-response relation is widely used as an estimate of community response to transportation noise in residential settings[11].

Unfortunately, a dose-response relation for people living in residential areas does not suffice for predicting response to aircraft noise within national parks. Park visitors are not at home, are not noise-exposed for months/years, and are not generally exposed to the same high levels of aircraft noise experienced near airports. Moreover, near major airports the predominant aircraft are large commercial jets, on approach or departing. In contrast,

[a] Grant Anderson Consulting, 76 Brook Trail, Concord MA 01742; email: gndrsn@comcast.net.
[b] John A. Volpe National Transportation System Center, US DOT/Research and Innovative Technology Administration, 55 Broadway, Cambridge, MA 02142; email: Amanda.Rapoza@dot.gov,
[c] John A. Volpe National Transportation System Center, US DOT/Research and Innovative Technology Administration, 55 Broadway, Cambridge, MA 02142; email: Gregg.G.Fleming@dot.gov.
[d] Harris Miller Miller & Hanson Inc., 77 South Bedford Street, Burlington, MA 01803; email: NMiller@hmmh.com.

the park-visitor environment often includes low-flying tour aircraft, either helicopters or fixed-wing propeller-driven aircraft, along with high-flying commercial jets.

Predicting the aircraft annoyance of park visitors is the focus of several research efforts in response to a 1987 Congressional mandate[12]. Those efforts were initiated by the U.S. National Park Service (NPS) in the early 1990s expressly in response to that mandate, which jointly tasks the NPS and the U.S. Federal Aviation Administration (FAA) to assess visitor response to noise. During that research, visitor responses were measured by two research teams at selected scenic national parks, using a survey questionnaire and concurrently measured sound exposures. That questionnaire, plus the data analysis and findings, were previously published in a series of reports[13–16]. From this research, there exists a database of some 2,600 visitor questionnaires with associated direct measurements of visitor-specific aircraft noise exposure. This paper includes a brief summary of the measurement portions of those reports, their measured sound levels, and their combined, homogeneous dose-response database.

In 2000 Congress further mandated FAA, in cooperation with NPS, to develop Air Tour Management Plans (ATMPs) for parks and tribal lands where air tour operations occur or are proposed[17]. The objective of these ATMPs is to develop acceptable/effective measures to mitigate or prevent the significant adverse impacts, if any, of commercial air-tour operations upon park resources, visitor experiences, and tribal lands.

An ongoing research program (development of methods to assess noise impacts from aircraft operating in and above national parks) identified additional analysis of these existing dose-response data as an important research step. That additional analysis was aimed towards (1) understanding the salient aspects or combinations of aspects of the noise exposure (sound level, length of exposure, time between exposures, number of exposure events, and/or source of the noise), (2) identifying additional site-specific or visitor-specific factors which may significantly influence visitor response, (3) developing a homogenous set of dose-response relationships that can be applied beyond the parks/sites represented in those studies, (4) constructing a prediction spreadsheet for use during ATMP development, and (5) providing guidance about future data-taking in types of park environments not yet included. This paper describes that additional analysis and its findings.

2 LIST OF SYMBOLS AND ABBREVIATIONS

This paper uses numerous abbreviations. These have been structured so the abbreviations provide substantial clues to the full description of each term or metric. Where needed, equation symbols use subscripts identical to their corresponding abbreviations.

2.1 Parks and Site Types

BRCA: Bryce Canyon National Park
GRCA: Grand Canyon National Park
HALE: Haleakala National Park
HAVO: Hawaii Volcanoes National Park
Overlook: Site at which visitors walk a short distance from their cars/buses to an overlook, mainly to view scenery
ShortHike: Site at which visitors walk along a short hike or trail—up to approximately $1^1/_2$ km (3/4 mile)

2.2 Aircraft Types

All: All aircraft types (Helos, Props, Jets)
Helos: Tour helicopters
Jets: High-altitude jet aircraft
Props: Tour fixed-wing, propeller-driven aircraft
Tours: Tour aircraft types (Helos, Props)

2.3 Ambient Types

LclAmb: Thirty seconds of non-aircraft sound level, measured just before/after (within five minutes of) each aircraft event
NatAmb or Nat: Natural ambient (wind in trees, water flow, animal sounds)—that is, all sounds except those caused by humans—therefore constituting the acoustic environment that NPS is tasked to preserve.
TradAmb or Trad: Traditional ambient (wind in trees, water flow, animal sounds, plus ground transportation, voices)—that is, all sounds except aircraft—therefore constituting the ambient for traditional FAA studies, in which aircraft noise protrudes into this non-aircraft, ambient environment.

2.4 Visitor Doses and Ambients (Computed from Measured Data)

All these variables are computed separately for each individual visitor. The time interval for each computation is that visitor's duration at the study site—arrival to start of interview.

L50, L90, L99: Exceedence percentages of the A-weighted sound-level history
LeqAll, LeqHelos, LeqJets, LeqProps, LeqTours, LeqNatAmb, LeqTradAmb (L_{eqAll} and so forth): A-weighted energy-equivalent sound level of the indicated type of aircraft or ambient
LmaxAll, LmaxHelos, LmaxJets, LmaxProps, LmaxTours: A-weighted maximum sound level of the indicated type of aircraft

NFIMaxNatAmb, NFIMaxTradAmb: Maximum (i.e., longest) noise-free time interval (minutes), determined with the indicated type of ambient

NFINumbNatAmb, NFINumbTradAmb: Number of noise-free time intervals, determined with the indicated type of ambient

NFINumbNatAmbPerHour, NFINumbTradAmbPerHour: Number of noise-free time intervals per hour, determined with the indicated type of ambient

NQ: Natural Quiet

NumbAll, NumbHelos, NumbJets, NumbProps: Number of the indicated type of aircraft

PHelos, PProps: Percentage of the indicated type of aircraft, based upon logged aircraft counts

PEnHelos, PEnProps ($P_{EnHelos}$ and so forth): Percentage of the indicated type of aircraft, based upon that source's energy contribution to LeqAll. See Eqn. (2) below for computation equation.

Plsx: Plus x dB

PTAbvNatAmbL50, PTAbvTradAmbL50: Percent time above L50 of the indicated type of ambient

PTAbvLclAmb + x: Percent time above LclAmb + x dB

PTAudHelos, PTAudJets, PTAudProps: Percent time indicated type of aircraft is audible

SELHelos, SELJets, SELProps ($L_{AEHelos}$ and so forth): A-weighted Sound Exposure Level of the indicated type of aircraft

2.5 Mediating Variables (from Questionnaire)

Numerical values of these variables depend upon whether (1) they are being mathematically regressed during analysis (called "regression" below), or (2) they are being substituted into the final regression equations during ATMP application of those equations (called "application" below).

AdultsOnly ($M_{AdultsOnly}$): During regression, equals 1 if that visitor's travel group consisted of only adults; equals 0 otherwise. During application, equals 0.01 times the percentage of visitor groups that consist of only adults.

ImpNQ_VorMore: Importance of Natural Quiet to the visitor: Very or more. During regression, equals 1 if that visitor responded "yes"; equals 0 if "no." During application, equals 0.01 times the percentage of "yes" visitors.

SiteType ($M_{SiteType}$): Equals 1 if ShortHike; equals 0 if Overlook.

SiteVisitBefore ($M_{SiteVisitBefore}$): During regression, equals 1 if that visitor has visited the site before; equals 0 otherwise. During application, equals 0.01 times the percentage of visitors who have visited the site before.

2.6 Visitor Responses (from Questionnaire)

Annoy: Annoyance

IntWithNQ: Interference with Natural Quiet

SorMore: Slightly or more

MorMore: Moderately or more

VorMore: Very or more

E: Extremely

2.7 Mathematics

σ_i: Standard uncertainty (standard error) of predictor i

ρ_{ij}: Correlation between the uncertainties of predictors i and j

C_i: Regression coefficient for the ith predictor

D_i: Dose value for the ith predictor

e: Base of natural logarithms = 2.71828…

log(x): logarithm of x to the base 10

M_i: Mediator value for the ith predictor

n: Number of predictors

N(μ, σ^2): Normal (Gaussian) probability distribution, with mean μ and standard deviation σ

Prob[x]: Probability of x

p-value: An assessment statistic for regressions

R: Response—either 0 (no) or 1 (yes)

SU: Standard Uncertainty (standard error) of regression coefficients

Var[x]: Variance uncertainty of x

x_i: Predictor variable

z: *logit* of the regression—that is, the "variable" term in the logistic equation $R = \frac{1}{(1+e^{-z})}$.

3 METHODS

For this data-collection and analysis effort, parks/sites were needed that allowed acquisition of the maximum amount of data in the minimum amount of field time. Additionally, for resulting dose-response relations to be most broadly applicable and also to be determined with maximum certainty, a large range of doses was needed. To satisfy both requirements, selected parks/sites needed high visitation rates, relatively frequent overflights, and aircraft overflights both close to, and distant from, park visitors.

With these selection constraints in mind, sites were selected through joint discussions between research staff and NPS park personnel, supplemented by visits to candidate parks/sites. This interactive and iterative process produced the selected parks/sites in Table 1. Included are four national scenic parks, with four short-hike, six scenic overlook, and two back-country sites divided among them.

3.1 Data Collection

Data collection proceeded at these selected parks/sites during the summers of 1992, 1997 and 1998.

Table 1—Selected parks and sites.

Park	Site	Site type	Trail length [km (miles)]	Approximate location
BRCA	Queens Garden Trail	Short hike	1.1 (0.7)	Descends toward Queens Garden from Sunrise Point, at the rim of the Canyon to the junction of Horse's Trail.
	Rainbow Point	Overlook	———	At the south end of State Road 63 looking north over the rim of Bryce Canyon.
	Fairyland	Overlook	———	Approximately 2 km (1 mi) off of State Road 63 heading east toward the rim of Bryce Canyon.
	Bryce Point	Overlook	———	Approximately 3 km (2 mi) off of State Road 63 heading east toward the rim of Bryce Canyon.
HALE	Sliding Sands Trail	Short hike	0.72 (0.45)	Descends into Haleakala crater from the top rim. The trail length walked by interviewed visitors was the portion most commonly walked before visitors turn back.
GRCA	Pima Trail	Short hike	1 (0.6)	Located on the South Rim, extending approximately 1 km (0.6 mi) west from Pima Point along the Canyon Rim.
	Lipan Point	Overlook	———	Located on the South Rim, at the end of a 0.8-km (half-mile) long paved road off East Rim Drive, State Route 64.
	Point Imperial	Overlook	———	On the North Rim, at the end of a 13-km (eight-mile) drive from State Route 67.
	Pima Point	Overlook	———	Approximately 6 km (4 mi) west of the visitor's center off of West Rim Drive.
	Hermit Basin	Back country	———	Approximately $2^1/_2$ km ($1^1/_2$ mi) along Hermit Trail and 400 m (1,300 ft) below Hermit's Rest on the South Rim.
	Havasu Creek	Back country	———	At Havasu Creek, on the Colorado River about 600 m (2,000 ft) below the South Rim.
HAVO	Wahaula Temple and the hike to it	Short hike	1.22 (0.76)	At the end of the Chain of Craters Road. At the time of the measurements, the trail led across an old lava flow to the remains of the temple (now completely destroyed).

The same methods and questionnaire were used along the Big Dune Trail at White Sands. However, those data were not included in the analysis because no air tours flew there and jets at that site were not typical of any other site. In particular, jets there (1) included many "stealth" aircraft (a great novelty to study participants), and (2) were not "high altitude," but were mostly landing/departing from adjacent Holloman Air Force Base.

Table 2—Survey years, days and visitors.

Park	Survey years	Site	Number of retained survey days	Number of retained visitors
BRCA	1997, 1998	Queens Garden Trail	9	905
		Rainbow Point	1	58
		Fairyland	2	140
		Bryce Point	1	43
HALE	1992	Sliding Sands Trail	5	217
GRCA	1992, 1998	Pima Trail	1	24
		Lipan Point	3	194
		Point Imperial	5	397
		Pima Point	2	156
		Hermit Basin	0	0
		Havasu Creek	0	0
HAVO	1992	Wahaula Temple and the hike to it	4	186
Totals	3 years		33	2320

Collection involved simultaneous measurement of sound level, observer logging of sound source, logging of visitor entry/exit times, and surveying of exiting visitor's response to aircraft sounds. The data-collection protocol ensured that visitor-specific noise doses could be calculated, based upon the time each visitor entered the study area and the time that visitor was interviewed. During data collection, aircraft activity included 5-to-30 tour Helos and/or Props per hour, and most often high-altitude jets in addition (as almost everywhere in the U.S.).

3.1.1 Visitor-response questionnaire

A questionnaire with the same basic format and core response questions was utilized for each of these data-collection efforts. This questionnaire was developed through an intensive consultative, planning and review process by NPS personnel and a number of acousticians, recreation sociologists, social survey experts and community dose-response survey experts.

The questionnaire was administered to groups of visitors (mostly single families) by an interviewer who read the questions. Each member of the group then independently completed an answer sheet that did not have the questions printed on it. Visitor answers provided the responses and mediator variables used in analysis.

All visitors, aged 16 and older, who were fluent in English were interviewed. From time to time, visitation rates exceeded the resources of the interview teams, thereby causing some visitor groups to be skipped. In addition to these visitor responses, the interviewer recorded observations such as group size, activity, and time of the interview.

Table 2 documents the survey years, along with the number of retained survey days and visitors at each site. Data from the two backcountry sites, Hermit Basin and Havasu Creek, were not retained (62 visitors) due to low sample size at Hermit Basin and the high ambient noise from Havasu Creek, which made measurement of aircraft sound levels impossible.

3.1.2 Sound levels

Simultaneous with visitor presence at the site, one-second A-weighted sound levels were continuously monitored by acoustical specialists with equipment capable of measuring very-low (below zero dB) noise levels. Measurement locations were chosen to be close enough to the on-site visitors to adequately capture their doses. Being close to visitor locations required several microphones at two of the sites, so that measured levels could be combined with visitor timings near those locations to compute each visitor's total dose.

In addition to best-available acoustical instrumentation, protecting each microphone from wind was a specially constructed double windscreen—consisting of an outer fabric windscreen of diameter 51 cm (20 inches) and a concentric, standard open-foam windscreen. Wind was continuously monitored and recorded. In addition, supplemental instrumentation collected digital audio tapes for portions of each measurement day, but these tapes were not used for the current analysis.

3.1.3 Observer logs and listening priority

Simultaneous with these acoustical measurements, observations of aircraft activity and moment-by-moment sound sources were logged on a laptop computer by trained personnel who passed an audiometric hearing test.

Those observations were structured by the following priorities: (1) tour aircraft (Helos or Props), (2) jet

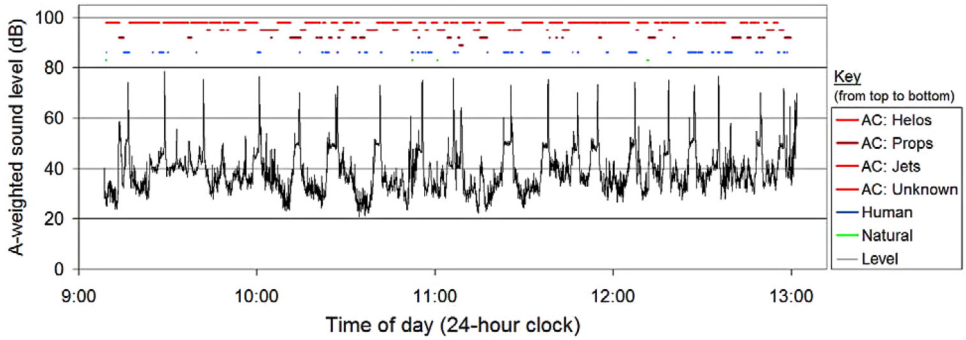

Fig. 1—Representative time history: one site, one day. Continuous line in black: one-second Leq time record at one site microphone. Short horizontal lines at the top: logged sound source from moment to moment (vertical ordering same as in key).

aircraft, (3) other human-produced sound such as autos and speech, and (4) natural ambient (natural sounds, as well as the complete absence of any audible sound). At any given time, only the highest-priority event was logged, thereby preserving no information about the presence or absence of lower-priority sound sources. After several preliminary experiments, this listening-priority method proved most reproducible and consistent across different observers.

Consistent with this priority order, the onset of each noise event was logged by pressing a time-stamp key on the computer, thereby entering the clock time into a new spreadsheet row. Then when the observer could be sure of the type of sound source, that source type was entered into the same line by a second key press. If that sound source proved incompatible with the listening priority (for example, the button should not have been pressed), the logger could rectify the initial entry in the spreadsheet before proceeding. Further comments about the source could be typed into a comments column on the same spreadsheet line.

3.2 Data Processing and Summary

The source logs and output from the acoustical instrumentation were first processed into sound-level time histories, associating each second of measured sound level with the logger's identified source. Graphs such as Fig. 1 help visualize this association. As shown along the top, each second in time is tagged with just one specific sound source. That second contributed to the "time audible" of that source type, while that second's sound levels contributed to that source type's sound-level metrics through the standard metric definitions.

With this time history and sound-source tagging, individual visitor doses (from aircraft) and ambient sound level metrics were computed from visitor arrival and departure times.

The resulting database contains one record for each retained visitor, and that visitor record contains all collected visitor parameters—doses, mediators and responses. Table 3 contains visitor percentages of selected non-numeric parameters (yes/no values and male/female). Omitted from this table are various parameters not useful to analysis because they were missing for many study sites.

Complementing this table is Fig. 2, which summarizes three numeric visitor parameters. These numeric distributions suggest a region of validity for the eventual dose-response relations. For example, almost all visit durations were less than 120 minutes, in effect defining the word "short" in "short hike."

The more extensive Fig. 3 provides several summaries of acoustical conditions experienced by the participants in the study: decibels, numbers and durations in separate frames. As shown, for the aircraft sound levels at the top of the figure, median visitor exposure to Helos exceeded exposure to Props, which exceeded exposure to Jets—but not by very much. The figure also shows the expected relative levels among Leq (lowest),

Table 3—Visitor percentages.

Characteristic	Percentage
U.S. citizen	71
Only adults in visitor group	69
View scenery and relax while in park	65
Natural Quiet very/extremely important	64
Arrived before noon	54
Male	49
Saw aircraft during site visit	47
Day hike while in park	30
Visited park before	28
Visited site before	17
Has ever taken an air tour	11

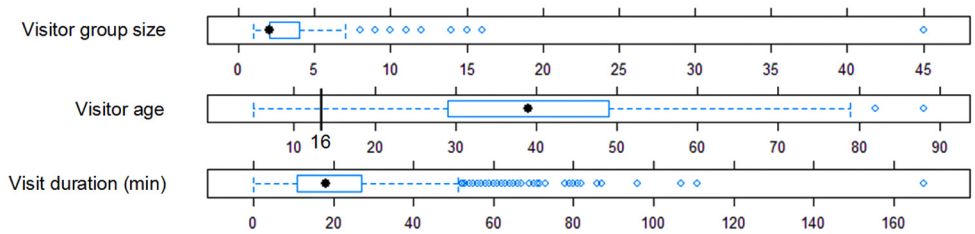

Fig. 2—Visitor characteristics. Visitors younger than 16 years old were excluded from analysis. Box-whisker plots: left-most circle (or left end of "whisker" line) = minimum, left edge of box = lower quartile, dark circle = median, right edge of box = upper quartile, right-most circle = maximum.

Lmax (middle) and SEL (highest). Also note that the Leq of ambient-sound exposure is comparable to that of aircraft.

3.3 Analysis

In brief, analysis consisted of (1) choosing responses of interest (Sec. 3.3.1), (2) *a priori* choosing preferred doses (Sec. 3.3.2), and then (3) regressing doses against responses, in two analysis stages (Sec. 3.3.6). The first stage determined if any of the alternative doses performed significantly better than the preferred ones. The second stage then tested the significance of various dose combinations and other mediating variables, to determine the final dose-response relations.

3.3.1 Responses

Of the several possible visitor responses from the questionnaire, visitor answers to the following two questions were chosen for analysis:

1. **Annoyance.** Were you bothered or annoyed by aircraft noise during your visit to [this site—e.g., Queens Garden Trail]? Choices: Not at all, slightly, moderately, very, or extremely[18].
2. **Interference with Natural Quiet.** How much did the sound from aircraft interfere with your appreciation of the natural quiet and sounds of nature at [this site]? Choices: Not at all, slightly, moderately, very, or extremely.

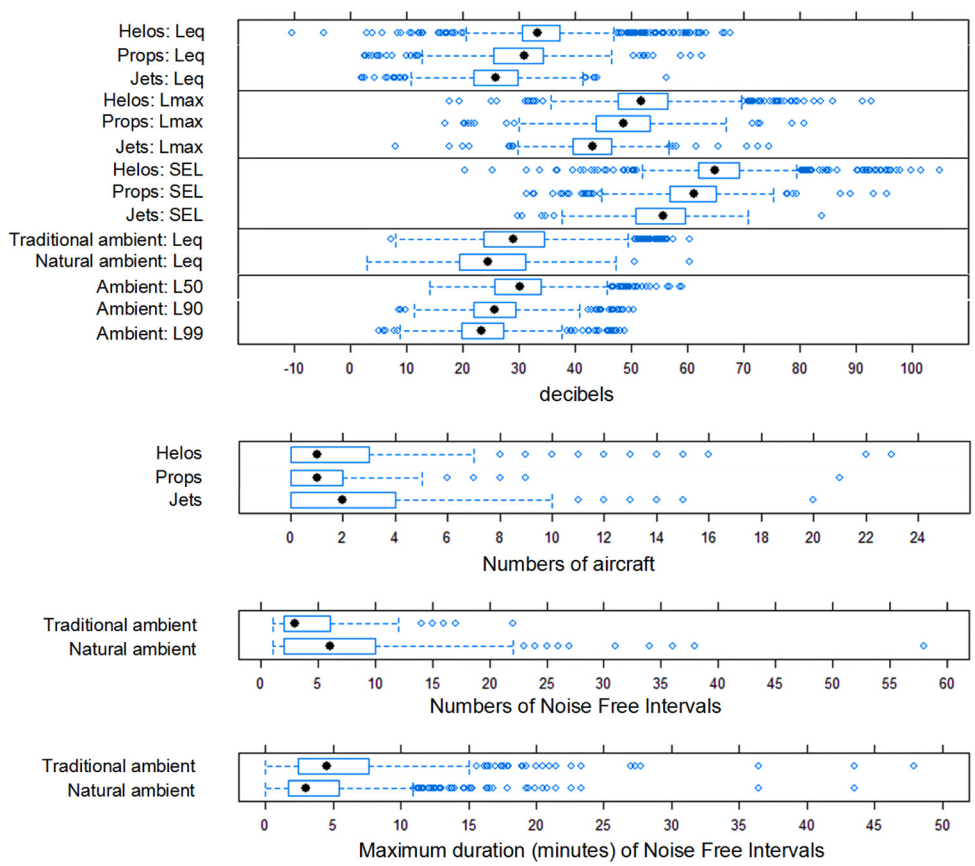

Fig. 3—Visitor doses and ambients. Box-whisker plots: see Fig. 2 caption.

Annoyance addresses the NPS mandate to improve visitor experience and is consistent with FAA noise-assessment criteria. Interference with natural quiet provides a visitor assessment of the affect of aircraft sounds on natural quiet, one of the natural park resources that NPS is mandated to preserve for posterity.

Three of the four possible dichotomizations were chosen for analysis:

- "Very or more" is analogous to the FAA's threshold for aircraft noise impacts in residential areas near airports.
- "Moderately or more" is considered by NPS to be perhaps more appropriate for a park setting.
- "Slightly or more" is of interest because its inverse ("not at all") denotes total lack of visitor response to noise.

The fifth dichotomization, using only "extremely," had insufficient data and was therefore unsuitable for regression analysis. In addition, of the 114 visitors (out of 2320, in total) who experienced no logged aircraft during their visit, 99 responded "not at all" to both response questions. The other 15 zero-dose visitors were retained in the analysis, nevertheless.

3.3.2 Doses

Against these responses/dichotomizations, not all ways to describe the measured dose could be included in the final dose-response relations, at least not without the process becoming overly burdensome. One method to narrow down the long list of doses is to regress them all separately against the responses and choose the ones that provide the tightest fit. Instead of this, however, a short list of "preferred doses" was decided *a priori* as a "starting point" for analysis:

- LeqAll: the Leq due to all aircraft
- PTAudAll: the percent time aircraft, of any type, are audible.

These two doses were chosen for the following reasons: First, aircraft Leq (with a penalty for nighttime operations) is used by FAA and many other federal/state agencies to assess and regulate environmental noise of all types. In addition, it is easily, accurately and efficiently measured/modeled by existing acoustical instrumentation and computer programs.

Furthermore, it is consistent with knowledge about human response to noise—particularly that human response increases with (1) increased aircraft sound emission (more horsepower, higher throttle, or less noise control), (2) decreased distance, (3) increased numbers of aircraft (per day or per hour), and (4) increased aircraft duration (aircraft hovers, circles around, or flies slower).

Percent time aircraft are audible is used as the main NPS metric to determine loss of natural quiet (time that human-related sound is *not* audible). Moreover, this metric is easy to measure and to understand. On the other hand, it ignores how loud the aircraft are when they are audible.

Alternative doses to these were tested (see Sec. 3.3.6), to learn if any of them are significantly better than the preferred doses (none were).

Thought was given to starting the analysis with LeqTour instead of LeqAll, since the main purpose of the study was to assess the impact of tour aircraft, especially during ATMP studies throughout the park system. However, the visitor questionnaire asked about "aircraft," not "tour aircraft." This forced the use of LeqAll as the primary metric instead of LeqTour. To compensate, aircraft type percentages were later tested, found statistically significant, and then inserted into the final regressions. This allows separate evaluation of air tours and high-altitude jets.

3.3.3 Multilevel logistic regression

Multilevel logistic regression was chosen to develop dose-response relations[19,20]. First, logistic regression properly confines the resulting curves (and also their 95% confidence regions) to lie between zero and unity (100%). Furthermore, the multilevel version of logistic regression can properly separate out the site-to-site differences that are potentially inherent in the data sampling plan.

With those differences separately computed, the regression mathematics allows proper computation of analysis uncertainties, because it (1) accounts for visitor-to-visitor variability assuming the proper underlying statistical distribution, and (2) accounts for site-to-site variability by analyzing Site as the "random" component of the multilevel regression. As a result, multilevel logistic regression avoids underestimates of prediction uncertainty, which can happen when site-to-site variabilities are not included in the regression equation.

Equation (1) results from multilevel logistic regression:

$$z = C_0 + \mathrm{N}(0,\, \sigma^2_{Site}) + \sum_{i=1}^{n} C_i x_i$$

$$R = \frac{1}{1 + e^{-z}}$$

$$= \mathrm{Prob}[\text{"yes" for the dichtomized response}]. \quad (1)$$

The first line of this equation is the regression's logit z. Except for that line's second term, it is identical to the full regression equation for linear regression. That second term is the contribution of the site-to-site variability

(quantified by its mean, normalized to zero, and standard deviation, σ_{Site}), determined by the multilevel portion of the mathematics. The second line of the equation converts the logit z to a dose-response curve: a sigmoid-shaped curve that is asymptotic to zero at the left and to unity at the right (or *vice versa* for negative values of z, depending upon the regression coefficients C).

The regression process (see Sec. 3.3.6) involves progressively testing additional predictors x_i, one at a time, for acceptance into the dose-response relation, based upon a set of predictor-acceptance goals.

3.3.4 Predictor-acceptance goals

Acceptance/rejection of predictors into multilevel regressions is not as straight forward as for their non-multilevel cousins. For example, different combinations of predictors shift uncertainty back and forth between uncertainty in the regression coefficients C_i, themselves, and the site-to-site variability σ_{Site}, which can contribute strongly to the prediction uncertainty. Moreover, regression-coefficient uncertainties, by themselves, do not show the full picture when the predictors are not statistically independent.

For these reasons, an automatic regression-testing process was not used to retain/reject predictors. Instead, various outputs of each "test" regression were examined, assessed and balanced to decide acceptance or rejection. Used in this balance were the following regression-acceptance goals, in order of importance:

- Reduction in "prediction" uncertainty (the 95% uncertainty region on the regression's graph). As graphed and discussed below, that uncertainty region is most relevant for later application to a single non-study site.
- Reduction in σ_{Site} (generally "tracks" the first goal). A small site-to-site standard deviation is essential for future applications to non-study sites.
- Statistically significant regression coefficient for the newly introduced predictor—as well as retention of statistical significance for predictors already in the regression.
- Reduction in Akaika Information Criterion (AIC) and the G-statistic. AIC sums up the regression residuals (technically, the "deviance"), so it is also important to monitor during regression. The G-statistic's p-value shows whether the reduction in total deviance is statistically significant[21].

3.3.5 Regression uncertainties

Standard uncertainties (standard errors) of the regression coefficients were automatically computed by the regression software. Of more interest, however, are "prediction" uncertainties when the regression results are applied to non-study park sites. In particular, when Eqn. (1) is used to compute response from its predictor values at a non-study site, how uncertain is the computed response?

To that end, the covariance matrix of the regression coefficients is determined during the iterative process seeking the best regression fit. That matrix incorporates not only the uncertainties of each regression coefficient in Eqn. (1), but also the correlation between every pair of coefficient uncertainties. For a linear logit, it is a complete description of the statistical distribution of those coefficients. In addition, the multilevel aspect of the software determines the site-to-site variance that was unexplained (not adjusted for) by the regression—that is, the σ^2_{Site} term in the equation.

From the outputs for each regression, prediction uncertainties were determined from 1000 simulations of Eqn. (1). Coefficient simulations centered on each coefficient's regressed value, but scattered about that value per a random draw from the coefficient covariance matrix (a multivariate normal distribution). Simulations of the site-to-site variability consisted of a random draw from that normal distribution.

In all, this process resulted in 1000 simulations, each producing a coefficient value (C_0 or C_i) and a site-to-site-variability value (σ^2_{Site}). Substitution of each simulation into Eqn. (1), along with one specific set of input values (one "input point") for the predictors, themselves, produces one simulation of R. The resulting 1000 simulations of R determine the prediction uncertainty in R at that input point. Results of this process appear in several graphs, below.

This simulation process accounts for the sampling uncertainty of the regression coefficients, plus the site-to-site variability. Accounting for the additional uncertainty of the input predictor values at a non-study site is left to the application spreadsheet, based upon user-input uncertainty estimates. That application spreadsheet was developed during the study, in response to one of the study objectives stated in the introduction above.

3.3.6 Regression process

The regression process is diagrammed in Fig. 4.

Starting condition First, the two chosen responses, combined with their three chosen dichotomizations (Sec. 3.3.1), further combined with the two preferred doses (Sec. 3.3.2), result in twelve preferred dose-response regressions. These constitute the initial regression equations, for later testing and possible augmentation with additional predictors.

Preliminary regressions (single doses) Analysis then asked whether or not any other dose, when

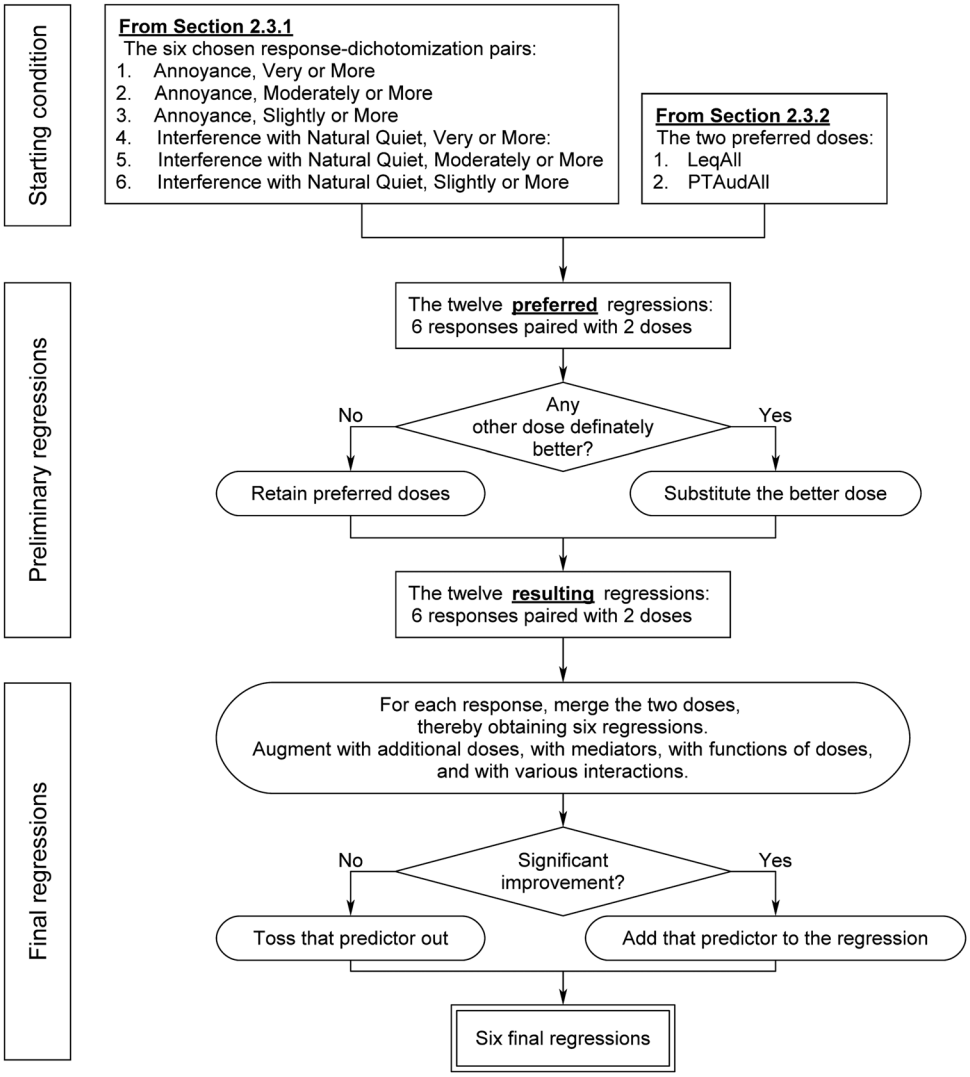

Fig. 4—Summary of the regression process.

substituted for these two preferred doses, produced a definitely "better" fit to the data, per the predictor-acceptance goals of Sec. 3.3.4. If definitely better per those goals, then the substitution was made. If not, then the substitution was not made. All these tests involved regressions with a single dose term—plus the SiteType mediator that had proven so important in prior partial analyses of these data.

Final regressions (augmented doses, plus mediators) For each response, the two doses were merged into one regression for that response, and then that merged dose was tested against the predictor-acceptance goals of Sec. 3.3.4. Then the resulting six regressions were augmented with additional doses, with mediators, with functions of doses, and with various interactions—again testing against the same predictor-acceptance goals at each step.

In brief, augmented predictors consisted of other dose metrics and other mediators, combined additively—for example, LeqAll + "one or more dose components." Also tested were various functions of various doses. Such functions were suggested *a priori* by analyst experience (two of these are retained in the final regressions). Tested in addition were various interactions between pairs of accepted doses—that is, the simple product of two doses. Each such function and each interaction produced a single predictor term x_i in Eqn. (1), thereby preserving the linearity of the logit z.

This final-regression process involved all six regressions, assessed simultaneously. Sometimes one or another of these regressions met one or another of the acceptance goals, but other regressions did not. Nevertheless, the same acceptance/regression decisions were made for all regressions, thereby retaining consistency in the format and use of the final results.

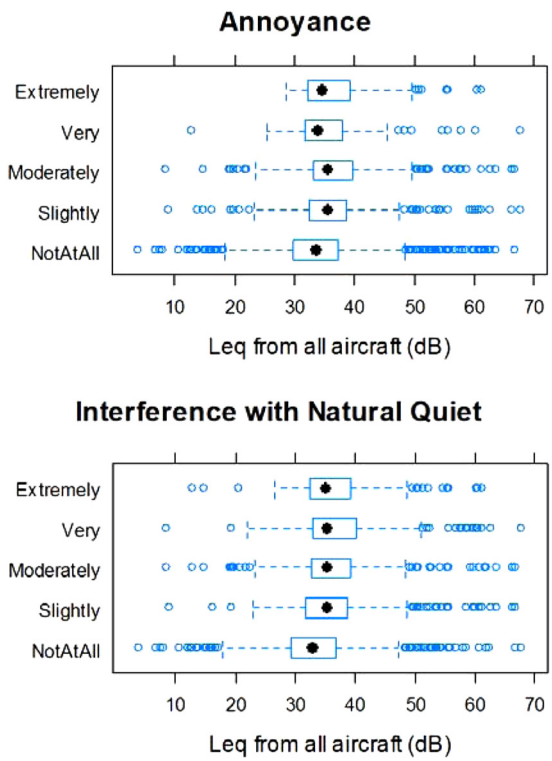

Fig. 5—Visitor responses. Box-whisker plots: see Fig. 2 caption.

4 RESULTS

4.1 Preliminary Regressions

The analyzed responses appear in Fig. 5, along with the five response choices for each. Each response is plotted against the Leq from all aircraft. From bottom to top in each frame, the box-whisker plots shift somewhat to the right. This shows an increased response as dose increases, most apparently for low aircraft sound levels.

However, the weak association and wide scatter in the figure indicate that human response is not likely based on sound level alone. Additional physical and psychological factors may have a strong influence on response, as is true near large airports. This analysis was designed to discover and quantify these additional influencing factors, specific to the park environment.

Preliminary regressions tested 34 individual, alternative doses against these responses. When these alternative doses were compared to the preferred doses:

- Many showed no descriptive power.
- None were significantly better than the preferred doses.
- Many were significantly worse than the preferred doses.

As a result, the two preferred doses were retained as the starting point for final regressions. Table 4 contains the rejected doses.

Table 4—Tested-then-rejected doses: Preliminary regressions.

Doses
LeqHelos, LeqProps, LeqJets, LeqTours
LmaxAll, LmaxHelos, LmaxJets, LmaxProps
NFIMaxTrad, NFIMaxNat
NFINumbTrad, NFINumbNat
NumbAll, NumbHelos, NumbJets, NumbProps
PTAbvAmb + 5, PTAbvAmb + 10, PTAbvAmb + 15, PTAbvAmb + 20
PTAbvAmbTradLeq, PTAbvAmbTradL50, PTAbvAmb-TradL90, PTAbvAmbTradL99
PTAudAll, PTAudHelos, PTAudJets, PTAudProps, PTAudTours
SELAll, SELHelos, SELJets, SELProps, SELTours

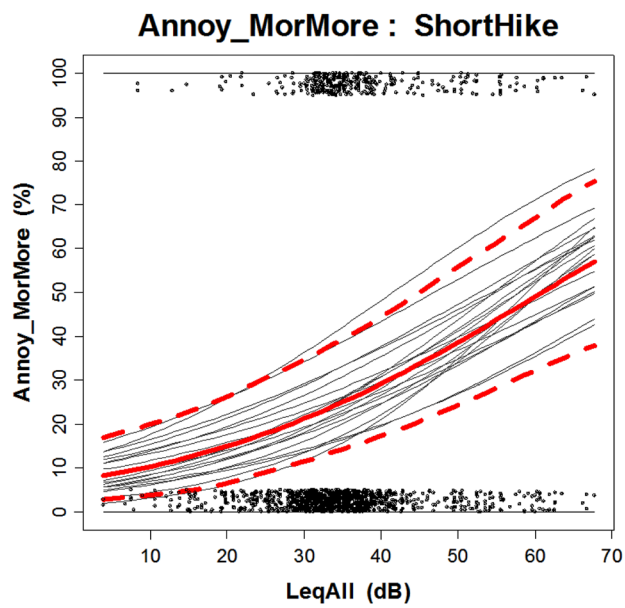

Fig. 6—Representative dose-response curve for Short Hikes: Preliminary regressions (single dose term). Underlying data points: those just below 100% are for visitors responding "yes," while those just above 0% are for visitors responding "no." To minimize data overlap, all data points are graphically "jittered" away from exactly 100% and 0%, their (percentage) values during regression. Solid red line is the regression line. Dashed red lines are the 95% confidence bounds for the regression. Faint lines are regression simulations (see text).

One of these preliminary regressions appears in Fig. 6. In that figure, the regression's output (probability of "yes" for the dichotomized response) has been converted to a percentage, for ease of use. The very light lines in the figure are 20 simulations of that regression, accounting for the sampling uncertainty of the regression coefficients as well as the additional site-to-site variability, as discussed in more detail in Sec. 3.3.5.

The dashed lines in the figure show the 95% confidence region for the regression, based upon 1000 such simulations. Because the underlying probability distribution for logistic regression is not symmetric about the regression line, the 95% confidence region here is not the "central" region containing 95% of the simulations. Instead, it is the narrowest region (vertically) within those 1000 simulations, determined separately for 50 points along the x-axis, and then smoothed.

This confidence region shows the predictive uncertainty of the regression, for each input value of LeqAll.

4.2 Final Regressions

4.2.1 The equations and their graphs

A desired outcome of final regressions was to reduce the 95% uncertainty region of the prediction. With that reduction, the confidence in applying the results to other, non-study sites will be greatly increased. To that purpose, various combinations and functions of some 34 dose and 18 mediator components were tested and one combination of 7 regression predictors was retained. Table 5 contains the rejected dose and mediator components. Some 20 other predictors were not tested, for lack of sufficient data.

Equation (2) contains the full set of final regressions.

$$
\begin{aligned}
z &= C_0 + C_1\left(M_{SiteType}\right) + C_2\left(L_{eqAll}\right) \\
&\quad + C_3\left(P_{EnHelos}\right) + C_4\left(P_{EnProps}\right) \\
&\quad + C_5\left(P_{EnHelos}\right)\left(P_{EnProps}\right) \\
&\quad + C_6\left(M_{ImpNQ_VorMore}\right) + C_7\left(M_{AdultsOnly}\right) \\
&\quad + C_8\left(M_{SiteVisitBefore}\right) \\
R &= \frac{1}{1+e^{-z}} \\
&\textit{where} \\
M_{SiteType} &= \ 1 \text{for ShortHike}, 0 \text{ for Overlook} \\
L_{eqAll} &= 10\log\left(10^{L_{eqHelos}/10} + 10^{L_{eqProps}/10} \right. \\
&\quad \left. + 10^{L_{eqJets}/10}\right) \\
P_{EnHelos} &= 100\left(\frac{10^{L_{eqHelos}/10}}{10^{L_{eqAll}/10}}\right) \\
P_{EnProps} &= 100\left(\frac{10^{L_{eqProps}/10}}{10^{L_{eqAll}/10}}\right)
\end{aligned}
\tag{2}
$$

Table 5—Tested-then-rejected predictors: Final regressions.

Various combinations of these predictors with the preferred doses:LeqAll and PTAudAll	
Doses	Mediators
(LeqAll) ˆ2, (LeqAll)ˆ3	Age
LeqAll/100	Aircraft seen at this site?
LeqJets, LeqTours	First visit to this park?
LmaxAll	Gender
logNFIMaxTrad, logNFIMaxNat	Group size 1 or 2
logNFINumbTradPerHour, logNFINumbNatPerHour	Home state (or home country)
logPEnTours	ImpNQ_SorMore, ImpNQ_MorMore
logPTAbvTradAmbL90, logPTAbvTradAmbL99	Mode of transportation
logPTAudAll	Nights in this park
logPHelos, logPTours	Number of visits to this park
NFIMaxTrad, NFIMaxNat	Specific study interviewer
NFINumbTradPerHour, NFINumbNatPerHour	Trail length
PEnTours	U.S. citizen?
PEnTours/100	Visit duration
PTAbvTradAmbL99	
PTAudAll	
PHelos, PTours	
SELAll, SELTours	
Interaction: LeqAll with PEnTours	
Interaction: LeqAll/100 with PEnTours/100	
Interaction: PEnHelos with PEnProps	

As shown, both PEnHelos and PEnProps are non-linear combinations of more-basic doses, even though those two combinations each enter the regression equation linearly. When one of these aircraft types does not fly at a site, its PEn equals zero (consistent with its Leq "equaling" minus infinity). If LeqAll equals minus infinity, then there is no aircraft noise to assess.

The regression coefficients of Eqn. (2) appear in Tables 6 and 7 for Annoyance and IntWithNQ, respectively. Of the 2320 visitors in the database, 1575 contributed to these final regression results (32% dropout). Twenty-six percent of this dropout was due to lack of LeqAll. During these visitor durations, no aircraft Leq's were computable, because aircraft sound was masked by the ambient during all logged aircraft events during that visit. The other six percent of this dropout was due to missing questionnaire data. For these, five percent skipped one or both of the response questions, while one percent skipped either the SiteVisitBefore or the ImpNQ question.

A representative set of dose-response curves appears in Figs. 7 and 8 for annoyance and interference with natural quiet, respectively. These provide a visual summary of Eqn. (2) for both responses, all three dichotomizations, both site types, and the average values of the additional mediators (see Table 8).

Only one dose appears explicitly in these plots: LeqAll. And so the figure is mute about how the effect of PEnHelos and PEnProps (and their interaction) were included. To include these, they were linearly regressed against LeqAll, and the results of that regression were substituted into the dose-response equation—thereby accounting for their effects through LeqAll. In essence, for each particular value of LeqAll (position along the horizontal axis), the most likely values (through linear regression) of PEnHelos and PEnProps were used to compute the vertical position of everything plotted at that x-value.

Table 6—Accepted predictors: Annoyance.

Response Dichotomization	Predictor	Regression coefficient	SU (Standard Uncertainty)	p-value
Slightly or more	Intercept	–3.075	0.295	0.000
	LeqAll	0.0184	0.0073	0.011
	PEnHelos	0.00753	0.00205	0.000
	PEnProps	0.00444	0.00232	0.056
	Interaction: (PEnHelos)(PEnProps)	0.000103	0.000081	0.204
	SiteType	1.223	0.167	0.000
	ImpNQ_VorMore	0.360	0.119	0.002
	SiteVisitBefore	0.480	0.157	0.002
	AdultsOnly	0.349	0.134	0.009
Moderately or more	Intercept	–4.084	0.357	0.000
	LeqAll	0.0158	0.0082	0.054
	PEnHelos	0.00831	0.00260	0.001
	PEnProps	0.006690	0.00290	0.021
	Interaction: (PEnHelos)(PEnProps)	0.000106	0.000091	0.246
	SiteType	1.128	0.194	0.000
	ImpNQ_VorMore	0.763	0.148	0.000
	SiteVisitBefore	0.551	0.170	0.001
	AdultsOnly	0.186	0.157	0.237
Very or more	Intercept	–5.378	0.541	0.000
	LeqAll	0.0102	0.0115	0.379
	PEnHelos	0.00472	0.00378	0.211
	PEnProps	0.00307	0.00432	0.477
	Interaction: (PEnHelos)(PEnProps)	0.000047	0.000137	0.733
	SiteType	1.804	0.324	0.000
	ImpNQ_VorMore	0.815	0.227	0.000
	SiteVisitBefore	1.020	0.213	0.000
	AdultsOnly	0.164	0.230	0.477

Table 7—Accepted predictors: Interference with Natural Quiet.

Response Dichotomization	Predictor	Regression coefficient	SU (Standard Uncertainty)	p-value
Slightly or more	Intercept	–3.157	0.296	0.000
	LeqAll	0.0300	0.0075	0.000
	PEnHelos	0.01117	0.00199	0.000
	PEnProps	0.00627	0.00223	0.005
	Interaction: (PEnHelos)(PEnProps)	0.000145	0.000081	0.072
	SiteType	1.115	0.162	0.000
	ImpNQ_VorMore	0.340	0.118	0.004
	SiteVisitBefore	0.536	0.161	0.001
	AdultsOnly	0.287	0.130	0.027
Moderately or more	Intercept	–3.467	0.310	0.000
	LeqAll	0.0241	0.0074	0.001
	PEnHelos	0.00830	0.00218	0.000
	PEnProps	0.00440	0.00248	0.076
	Interaction: (PEnHelos)(PEnProps)	0.000101	0.000083	0.225
	SiteType	0.981	0.171	0.000
	ImpNQ_VorMore	0.329	0.124	0.008
	SiteVisitBefore	0.608	0.158	0.000
	AdultsOnly	0.398	0.140	0.004
Very or more	Intercept	–4.352	0.371	0.000
	LeqAll	0.0252	0.0084	0.003
	PEnHelos	0.00777	0.00272	0.004
	PEnProps	0.00685	0.00299	0.022
	Interaction: (PEnHelos)(PEnProps)	0.000016	0.000096	0.865
	SiteType	0.817	0.195	0.000
	ImpNQ_VorMore	0.712	0.155	0.000
	SiteVisitBefore	0.618	0.173	0.000
	AdultsOnly	0.293	0.166	0.078

4.2.2 Missing predictors

Missing from these final regressions are two sound metrics shown to be relevant in prior analyses of these data:

- **Percent Time Audible.** Addition of Percent Time Audible into a regression already containing aircraft Leq proved no better than aircraft Leq, alone. And so Percent Time Audible was not allowed into the regression along with aircraft Leq. Note that aircraft Leq accounts for aircraft sound duration in its computation, as does Percent Time Audible. Prior analyses did not examine both aircraft Leq and Percent Time Audible simultaneously in the same regression, so did not discover this joint-dose conclusion.
- **Ambient sound levels.** Inclusion of ambient sound level (in any form) did not improve the regression fit—either on its own or when combined with aircraft levels into relative sound level, as in prior analyses. The authors believe this happened for two reasons: (1) the measured sites contained large numbers of aircraft, with Lmax values significantly above ambient levels, thereby potentially minimizing the importance of ambient sound, and (2) ambient Leq correlated strongly with aircraft Leq for these data. These two conditions may not apply during future back-country measurements, so the possible inclusion of ambient sound levels should be considered during future measurements/analysis.

4.2.3 Predictor trade-offs

The final dose-response relations contain regression coefficients for both aircraft sound levels and three mediating variables. Therefore, regression-coefficient ratios directly show the trade-offs between these mediators and aircraft sound level. Those trade-offs compute to the following, in the absence of high-altitude jet aircraft (ranges are over the six dose-response relationships):

- **Visitors for whom natural quiet is very/extremely important** (call them "yes" visitors) are 10-to-50 dB more sensitive to aircraft

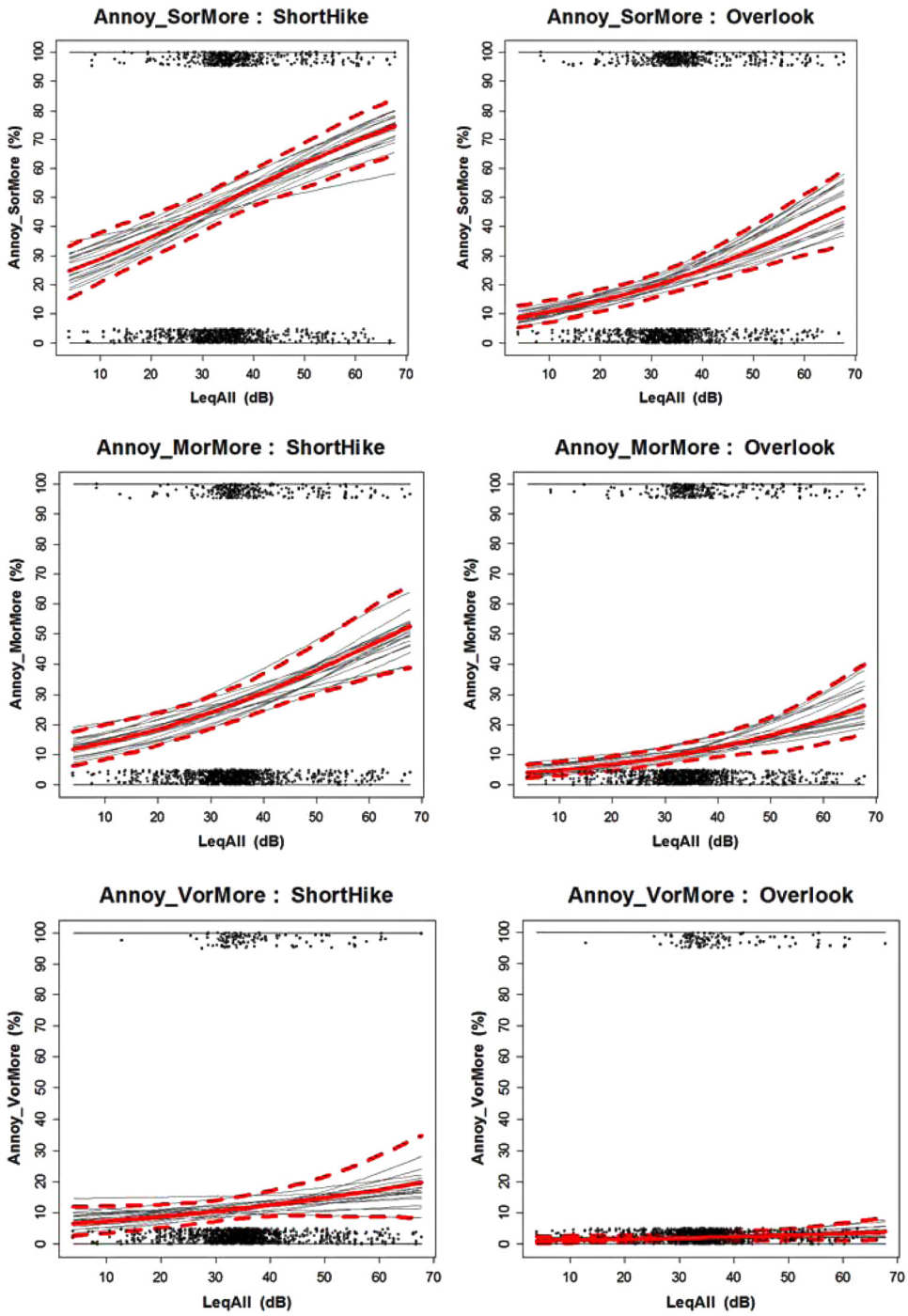

Fig. 7—Representative dose-response curves: Annoyance. Plots are computed for these average visitor percentages: ImpNQ_VorMore = 64%, SiteVisitBefore = 17%, AdultsOnly = 69%. Also see Fig 6. caption.

noise than other visitors (call them "no" visitors). For example, an aircraft Leq of 15 dB would elicit the same response from "yes" visitors as would an Leq some 10-to-50 dB higher from "no" visitors.

- **Visitors who have visited the site before** are 20-to-35 dB more sensitive to aircraft noise than other visitors.
- **Visitors in adults-only groups** are 10-to-20 dB more sensitive to aircraft noise than other visitors.
- **Visitors on Short Hikes** are 35-to-70 dB more sensitive to aircraft noise than visitors at Overlooks.
- **Response jump from SlightlyOrMore to ModeratelyOrMore** requires an aircraft-Leq increase of approximately 20-to-40 dB.

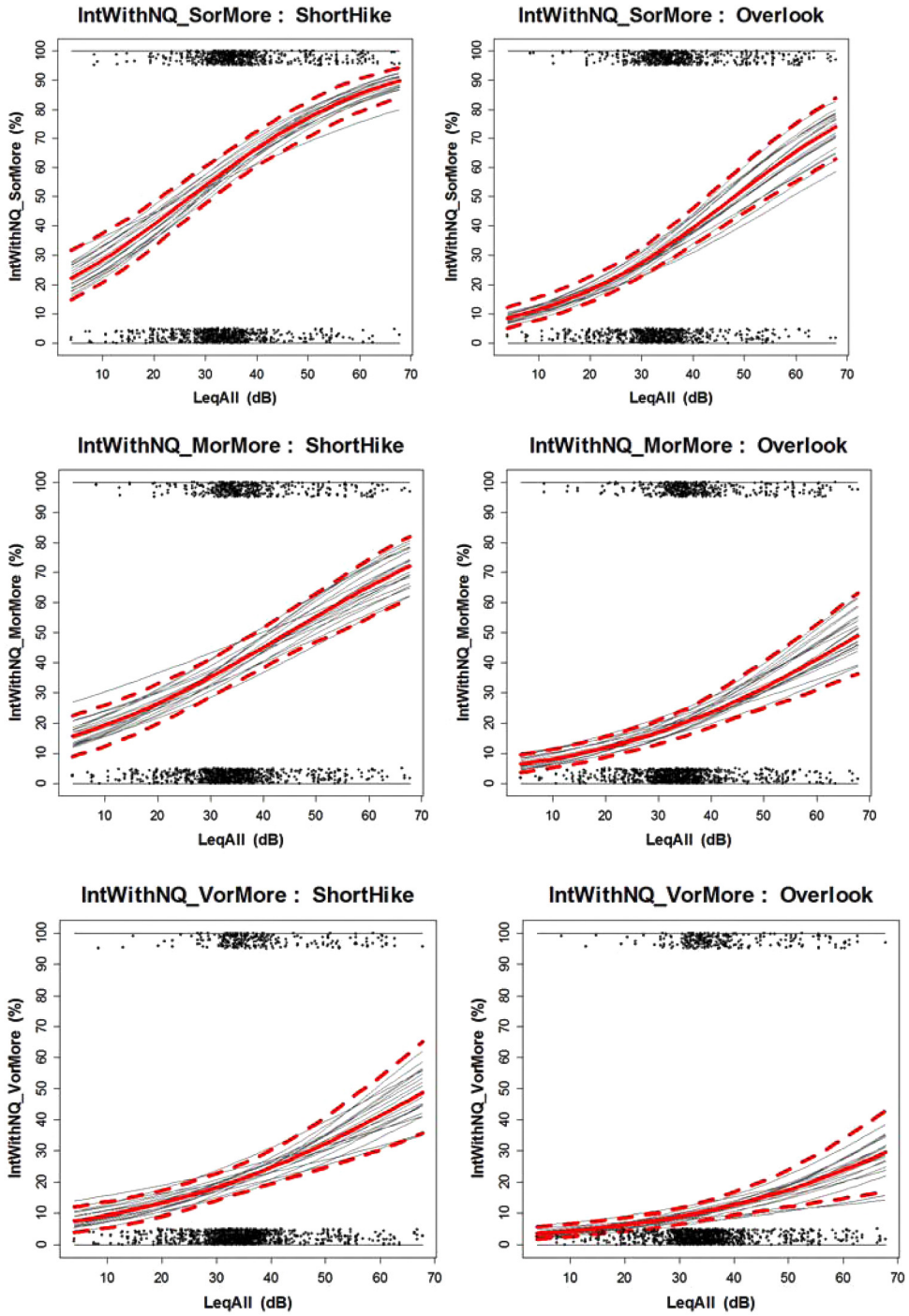

Fig. 8—Representative dose-response curves: Interference with Natural Quiet. Plots are computed for these average visitor percentages: ImpNQ_VorMore = 64%, SiteVisitBefore = 17%, AdultsOnly = 69%. Also see Fig. 6 caption.

- **Response jump from ModeratelyOrMore to VeryOrMore** requires an aircraft-Leq increase of approximately 20-to-30 dB.

4.2.4 Influence of high-altitude jets on response

Many of these dose-response relations were also plotted as contour sets, thereby showing the response dependence upon two input predictors at the same time. One of these contour sets appears to the left in Fig. 9. This set shows how response depends jointly upon LeqTours (horizontal axis) and LeqJets (vertical axis). In addition, it shows the relation between the contours and two of its corresponding dose-response curves, to the right. Contour sets for other final regressions show the same contour pattern—though with higher or lower contour values.

Table 8—Average mediator values during field study.

		Average percentages		
Site type	Site	ImpNQ_VorMore	VisitBefore	AdultsOnly
Short hikes	Queens Garden Trail	66	13	74
	Sliding Sands Trail	74	17	89
	Pima Trail	75	25	33
	Wahaula Temple	56	21	97
Overlooks	Rainbow Point	66	16	64
	Fairyland	64	9	59
	Bryce Point	53	27	70
	Lipan Point	62	19	73
	Point Imperial	68	8	70
	Pima Point	58	14	60
Short hikes	AVERAGE	68	19	73
Overlooks	AVERAGE	62	16	66
Overall	AVERAGE	64	17	69

As shown by the two green squares in the figure, tour aircraft are much more intrusive than are high-altitude jets, for the same sound level. In particular, the lower green square shows that 40 dB tour aircraft annoy 36% of visitors, compared to only 16% for 40 dB jets at the upper green square.

As shown by the two red lines, increasing tour-aircraft noise (left to right in the figure), holding jet noise constant, always results in increased annoyance. However, in the interior of the figure, where sound from both aircraft types are comparable, increasing jet noise (bottom to top in the figure) results in *decreased* annoyance. For example, the two green circles show that annoyance *decreases* from 30% to 19% as jet noise *increases* by 15dB, while holding tour noise constant. Moreover, uncertainty analysis shows this effect to be systematic (95% certainty), rather than a result of randomness.

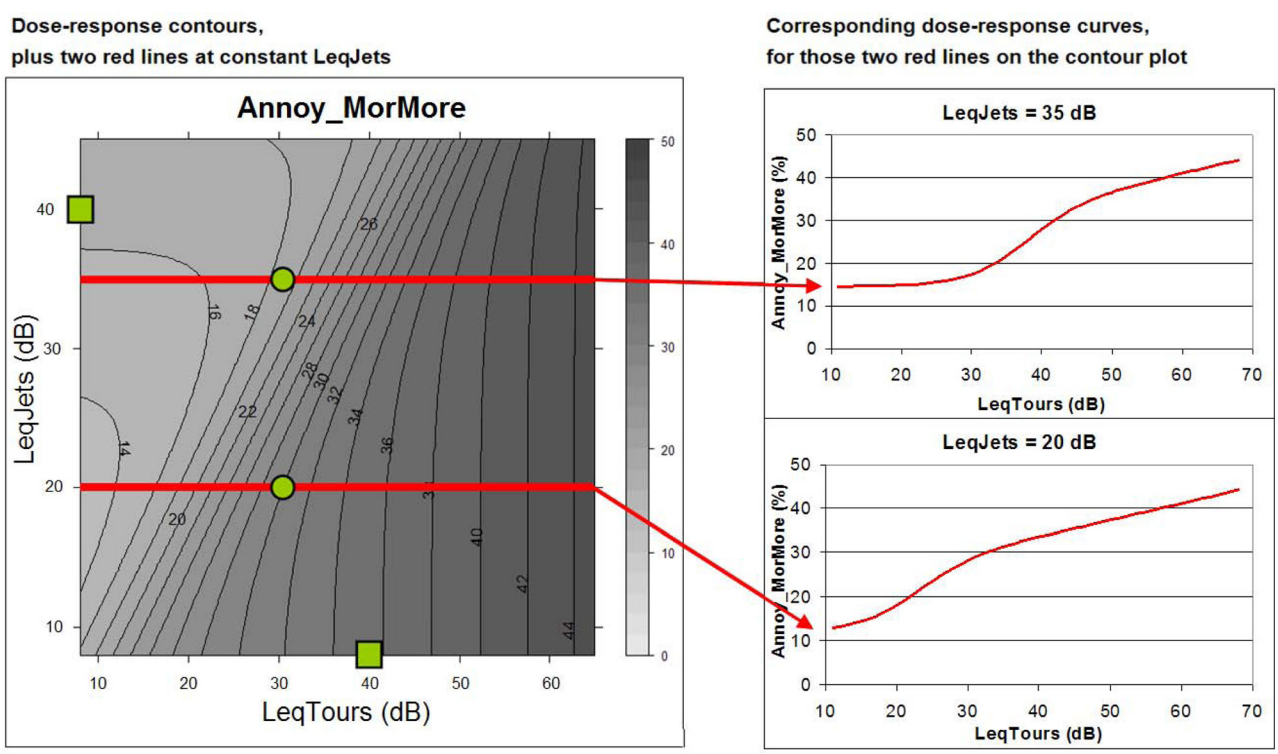

Fig. 9—Joint influence of noise from both tour aircraft and high-altitude jets: Dose-response contours compared to dose-response curves.

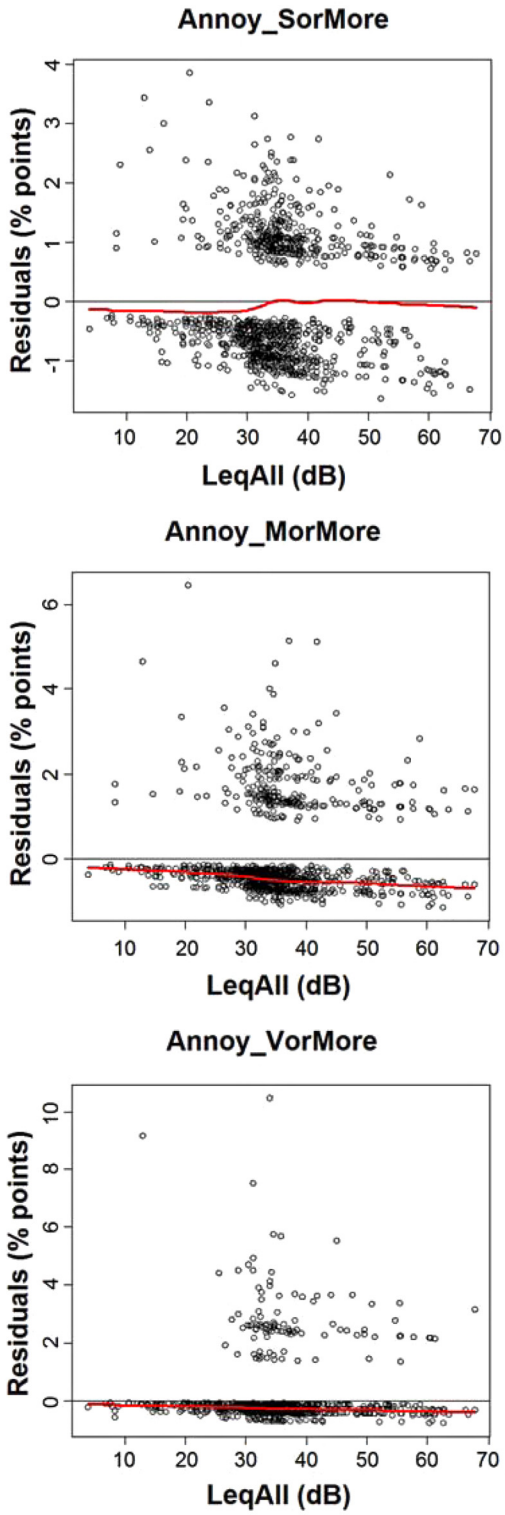

Fig. 10—Representative regression residuals. For logistic regression, residual points always form two data "clouds"—one for all visitor "yes" responses, the other for all visitor "no" responses.

As these green circles show, visitor response to a given dose of air-tour noise is less severe when there are low-to-moderate levels of high-altitude jet noise present, as well. This counter-intuitive behavior might be partly due to masking of tour noise by the less-intrusive jet noise. But it cannot be due solely to that because (1) physiological masking requires simultaneous sounds and (2) the aircraft types are often heard at different moments during the visitor's time at the site. The cause of this "apparent masking" is unknown, so future research will continue to investigate both the presence and cause of this phenomenon. The authors suspect that the presence of high-altitude jet noise during a visit partially desensitizes visitors to tour-aircraft noise during that visit.

4.3 Generalization of Results to Other Park Sites

The dose-response relations determined in this study comprise the most up-to-date quantitative assessment of air-tour aircraft noise in U.S. national parks. These multi-park, multi-site relations encompass a good range of doses (A-weighted Leq ranging from 0 to nearly 70 dB) for both tour helicopters (Helos) and fixed-wing propeller-driven aircraft (Props).

Representative regression residuals appear in Fig. 10. As seen, the locally-weighted, least-squares smoother through these residuals is quite flat and close to zero. In addition, the residuals themselves are quite small in magnitude—especially in the top two frames of the figure. In conjunction with this figure, Table 9 documents residual sizes for all six final regressions.

In addition, prediction uncertainties (see Figs. 7 and 8) are sufficiently small to provide usefully precise assessment relations. For these reasons, the authors believe that the results of this study can be generalized to other scenic-park sites—both front-country short hikes and overlooks—with the obvious caveat that some such sites might have unique characteristics not represented here.

Several park/site types await additional dose-response data and analysis: (1) multi-day hikes and campgrounds in park backcountry, (2) frontcountry day hikes, and (3) historic and/or cultural sites, for which speech interference may be important. Future plans include such additional data acquisition and analysis.

5 SUMMARY OF KEY FINDINGS

Section 3 presents and discusses the results of this study, including the six dose-response relations that

Table 9—Size of the residuals.

Response	Dichotomization	Residual size (percentage points above, or below, the predicted response)		
		minimum	average	maximum
Annoyance	Slightly or more	–1.6	–0.006	3.9
	Moderately or more	–1.1	–0.008	6.4
	Very or more	–0.8	–0.005	10.4
Interference with Natural Quiet	Slightly or more	–2.4	–0.002	3.6
	Moderately or more	–1.6	–0.004	4.5
	Very or more	–1.1	–0.003	6.7

summarize the park-visitor data. In summary of that section:

- **Dose-response relations.** Six dose-response relations derive from the study's multi-year, multi-site data. These relations appear above in Eqn. (2) with regression coefficients from Tables 6 and 7—and are graphed in Figs. 7 and 8. Per these graphed relations, percentage of annoyed visitors ranges between 0% and 70%, while percentage of visitors expressing interference with natural quiet ranges between 5% and 90% —dependent upon aircraft sound levels and the visitor's perceived response severity (some, moderate, very/extremely). These multi-park, multi-site relations can provide a basis for dose-response in park settings, compared to residential settings. They comprise the most comprehensive quantitative analysis of visitor annoyance and interference responses to measured air-tour aircraft noise in U.S. national parks. And they encompass a good range of doses (A-weighted Leq ranging from 0 to nearly 70 dB) for both tour helicopters (Helos) and fixed-wing propeller-driven aircraft (Props).
- **Significant acoustic variables.** The following acoustic variables are significant predictors of park-visitor response: Total aircraft Leq during the visit (Helos, Props, Jets), energy-percentage of that total due to helicopters (Helos), and energy-percentage of that total due to fixed-wing propeller aircraft (Props).
- **Significant non-acoustic variables.** In addition, (1) visitors on short hikes (less than 2-hour duration) are 35-to-70 dB more sensitive to aircraft noise than visitors at scenic overlooks, (2) visitors for whom natural quiet is very/extremely important are 10-to-50 dB more sensitive, (3) visitors who have visited the site before are 20-to-35 dB more sensitive, and (4) visitors in adults-only groups are 10-to-20 dB more sensitive.
- **Variables not significant.** All other variables (acoustic and non-acoustic) are statistically insignificant to the regressions. In particular, percent time audible and the level of ambient sound do not significantly affect response for visitors on short hikes or at scenic overlooks.
- **Tour aircraft vs high-altitude jets.** For the same sound levels, tour aircraft are much more intrusive than are high-altitude jets—for example, tour aircraft annoy 36% of visitors (moderately or more), compared to only 16% for high-altitude jets. In addition, visitor response to a given dose of air-tour noise is less severe when there are low-to-moderate levels of high-altitude jet noise present, as well.
- **Generalization to other parks and sites.** From these dose-response relations, prediction uncertainties (which include site-to-site variability) are sufficiently small to provide usefully precise assessment for other (non-study) short hikes and overlooks, as well—with the obvious caveat that some such sites might have unique characteristics not represented here. Generalization to other site types, especially in back country, awaits further data and analysis.

6 ACKNOWLEDGMENTS

First, the authors wish to thank the National Park Service and the Federal Aviation Administration for funding, organizing and overseeing this measurement/analysis multi-year study.

Many people contributed significantly to this overall effort. Measurements during 1992 were accomplished by a team of acousticians at Harris Miller Miller & Hanson Inc.—mainly Nicholas Miller, Christopher Menge, Richard Horonjeff, and Gonzalo Sanchez—plus two survey sociologists at Hagler Bailly Research Services, Inc.: Robert Baumgartner and Cary McDonald. 1997/1998 measurements were accomplished by

the Acoustics Group at the Volpe National Transportation System Center—mainly Gregg Fleming, Christopher Roof, Amanda Rapoza, Cynthia Lee and David Read—plus survey sociologists at Chilton Research Services: Joel Webster and Peggy Liebman.

Government oversight and general guidance from the National Park Service came from Wes Henry, Rick Ernenwein, Mike Ebersole, Kurt Fristrup, Frank Turina, Shan Burson and Vicki McCusker—plus superintendants and staff at the visited national parks. Similar oversight and guidance from the Federal Aviation Administration came from Barry Brayer, Peter Ciesla, Keith Lusk, Raquel Girvin, Hua 'Bill' He, Rebecca Cointin, Jake Plante and Julie Draper.

Additional guidance came from James Fields, Kristin Lewis, Steve Lawson, Britton Mace and Robert Manning.

7 APPENDIX: REGRESSION UNCERTAINTIES

Standard uncertainties (standard errors) of the regression coefficients appear in Tables 6 and 7, along with their p-values. Some of these p-values are higher than normally accepted into regression equations (higher than 0.05, for instance).

When regression predictors (the variables, not the coefficients) are correlated, the uncertainty of their coefficients does not show the full "uncertainty picture." Prediction uncertainty—during use of the dose-response curves (Fig. 7, for example)—depends also upon the correlation among regression coefficients (the coefficients, not the variables). The full uncertainty picture lies therefore within the covariance matrix of regression coefficients (Table 10, for example).

Equation (3) shows how all relevant terms combine into the prediction uncertainty of the regression's logit (and thereby the 95% uncertainty region of the curve itself)[22]:

$$\mathrm{Var}[z] = \sum_{i=0}^{n} C^2{}_i\, \sigma^2{}_i + \; 2\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} C_i C_j\; \sigma_{ij}\,. \quad (3)$$

The first term in this equation involves the coefficient uncertainty variances, σ_i^2, which are the diagonal terms of the full covariance matrix (as in Table 10). The second term involves the (covariance) off-diagonal terms σ_{ij}, which each appear twice in this symmetric matrix. In addition to those variance-covariance terms, uncertainty also depends upon the values, C_i, of all the predictor variables in the regression—for example, LeqAll. As a result of this C_i dependence, uncertainty is low near the data centroid and larger away from that centroid.

Example: Just the dose terms of Annoy_MorMore, evaluated at the data centroid

Centroid values:

- For the intersect: $C_0 = 1$ (as always)
- For LeqAll: $C_1 = 35.0$
- For PEnHelos: $C_2 = 52.6$
- For PEnProps: $C_3 = 22.8$
- For the interaction between PEnHelos and PEnProps: $C_4 = (C_2)(\,C_3) = 1199.3$.

With these centroid values, combined with the relevant terms from Table 10, Eqn. (3) evaluates to the following:

$$\begin{aligned}
\mathrm{Var}[z] &= \sum_{i=0}^{4} C_i^2\,\sigma_i^2 + 2\sum_{i=0}^{3}\sum_{j=i+1}^{4} C_i C_j\,\sigma_{ij} \\
&= [(1)^2(1.28)(10^{-1}) + (35.0)^2(6.70)(10^{-5}) + (52.6)^2(6.77)(10^{-6}) + (22.8)^2(8.41)(10^{-6}) \\
&\quad + (1199.3)^2(8.35)(10^{-9})] + 2[(1)(35.0)(-1.79)(10^{-3}) + (1)(52.6)(-1.90)(10^{-4}) + (1)(22.8)(-2.58)(10^{-4}) \\
&\quad + (1)(1199.3)(6.66)(10^{-7}) + (35.0)(52.6)(-7.81)(10^{-6}) + (35.0)(22.8)(-5.84)(10^{-6}) + (35.0)(1199.3) \\
&\quad \times(-1.02)(10^{-8}) + (52.6)(22.8)(5.48)(10^{-6}) + (52.6)(1199.3)(-4.22)(10^{-8}) \\
&\quad + (22.8)(1199.3)(-8.41)(10^{-8})] \\
&= [\text{Coefficient} - \text{uncertainty}] + [\text{Covariance adjustment}] \\
&= [0.2452] + [-0.1912] = 0.0540. \qquad (4)
\end{aligned}$$

And so the corresponding standard uncertainties are:

$$\begin{aligned}
\mathrm{SU}_{\text{coefficient-uncertainty}} &= \sqrt{0.2452} = 0.495 \\
\mathrm{SU}_{\text{Total}} &= \sqrt{0.0540} = 0.232. \qquad (5)
\end{aligned}$$

Note that the covariance adjustment has cut the total uncertainty by more than half—that is, from 0.495 to 0.232. In all, the 95% confidence bounds in graphs such as Figs. 7 and 8 take these covariances into account, thereby providing sure guidance (although

Table 10—Example covariance matrix of regression coefficients. Upper table shows covariance symbols; lower table contains their values for the example in the text (Annoyed, Moderately or More).

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
V_1	$(\sigma_1)^2$	$\sigma_{1,2}$	$\sigma_{1,3}$	$\sigma_{1,4}$	$\sigma_{1,5}$	$\sigma_{1,6}$	$\sigma_{1,7}$	$\sigma_{1,8}$	$\sigma_{1,9}$
V_2		$(\sigma_2)^2$	$\sigma_{2,3}$	$\sigma_{2,4}$	$\sigma_{2,5}$	$\sigma_{2,6}$	$\sigma_{2,7}$	$\sigma_{2,8}$	$\sigma_{2,9}$
V_3			$(\sigma_3)^2$	$\sigma_{3,4}$	$\sigma_{3,5}$	$\sigma_{3,6}$	$\sigma_{3,7}$	$\sigma_{3,8}$	$\sigma_{3,9}$
V_4				$(\sigma_4)^2$	$\sigma_{4,5}$	$\sigma_{4,6}$	$\sigma_{4,7}$	$\sigma_{4,8}$	$\sigma_{4,9}$
V_5					$(\sigma_5)^2$	$\sigma_{5,6}$	$\sigma_{5,7}$	$\sigma_{5,8}$	$\sigma_{5,9}$
V_6						$(\sigma_6)^2$	$\sigma_{6,7}$	$\sigma_{6,8}$	$\sigma_{6,9}$
V_7							$(\sigma_7)^2$	$\sigma_{7,8}$	$\sigma_{7,9}$
V_8								$(\sigma_8)^2$	$\sigma_{8,9}$
V_9									$(\sigma_9)^2$

Lower-left cells: Same values as upper-right cells, for this symmetric matrix

	Intercept	LeqAll	PEnHelos	(PEnHelos) **x** (PEnProps) PEnProps	Interaction	ImpNQ_ VorMore	SiteVisit BeforeYes	Adults OnlyYes	SiteType ShortHike
Intercept	1.28e-01	−1.79e-03	−1.90e-04	−2.58e-04	6.66e-07	−1.74e-02	−8.52e-03	−1.66e-02	−1.29e-02
LeqAll		6.70e-05	−7.81e-06	−5.84e-06	−1.02e-08	2.91e-05	6.47e-05	−2.62e-06	−1.64e-04
PEnHelos			6.77e-06	5.48e-06	−4.22e-08	−2.48e-07	−1.04e-06	−4.41e-07	−6.48e-05
PEnProps				8.41e-06	−8.41e-08	−3.35e-06	−1.47e-05	−9.15e-06	−6.01e-06
(PEnHelos) **x** (PEnProps)					8.35e-09	8.05e-09	7.99e-07	−9.54e-09	1.16e-06
ImpNQ_VorMore						2.19e-02	−3.05e-04	−7.06e-04	8.88e-04
SiteVisitBeforeYes							2.89e-02	1.52e-03	−8.67e-05
AdultsOnlyYes								2.46e-02	−3.30e-03
SiteTypeShortHike									3.78e-02

Lower-left cells: Same values as upper-right cells, for this symmetric matrix

not as quantitative) about the value of accepting additional predictors into the regression. Such graphs were used throughout the analysis to determine acceptance/rejection of regression predictors.

In particular, Tables 6 and 7 show especially high p-values for the interaction term (PEnHelos times PEnProps) in all-but-one final regression. In spite of this, that interaction was accepted into the regressions because it reduced the prediction uncertainty substantially, nearly halving the uncertainty-region width on the associated graphs. The negative covariances overcompensated for the interaction's poor p-value when determining prediction uncertainty.

8 REFERENCES

1. H. E. Von Gierke and C. S. Harris, *Annoyance response to military flight operations and the development of standard criteria for community annoyance—Environmental annoyance: characterization, measurement, and control*, Elsevier Science Publishers B.V., (1987).
2. Sanford Fidell, "An historical perspective on predicting the annoyance of noise exposure", *NoiseCon90*, (1990).
3. Theodore J. Schultz, "Synthesis of social surveys on noise annoyance", *J. Acoust. Soc. Am.*, **64**(2), 377–405, (1978).
4. Sanford Fidell, David S. Barber and Theodore J. Schultz, "Updating a dosage-effect relationship for the prevalence of annoyance due to general transportation noise", *J. Acoust. Soc. Am.*, **89**(1), 221–233, (1991).
5. Sanford Fidell, "The Schultz curve 25 years later: A research perspective", *J. Acoust. Soc. Am.*, **114**(6), 3007–3015, (2003).
6. "Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety", U.S. Environmental Protection Agency, (1974).
7. Airport Noise Compatibility Planning: Final Rule, 14 CFR Part 150, U.S. Federal Aviation Administration, (2004).
8. Environmental Criteria and Standards: Noise Abatement and Control, 24 CFR Part 51B, U.S. Department of Housing and Urban Affairs, (1979).
9. High Speed Ground Transportation Noise and Vibration Impact Assessment, U.S. Federal Railroad Administration, (2005).
10. Transit Noise and Vibration Impact Assessment, Report FTA-VA-90-1003-06, U.S. Federal Transit Administration, (2006).
11. Federal Agency Review of Selected Airport Noise Analysis Issues, Federal Interagency Committee on Noise, (1992).
12. National Parks Overflight Act: Public Law 100-91, (1987).
13. Grant S. Anderson, Richard D. Horonjeff, Christopher W. Menge, Nicholas P. Miller, William E. Robert, Christopher Rossano and Gonzalo Sanchez, "Dose-response Relationships Derived from Data Collected at Grand Canyon, Haleakala and Hawaii Volcanoes National Parks", Report No. 290940.14, Harris Miller Miller & Hanson Inc., (1993).
14. Gregg G. Fleming, Christopher J. Roof, Amanda S. Rapoza, David R. Read, Joel C. Webster, Peggy C. Liebman, Paul J. Valihura, Kristin C. Lewis, Paul D. Schomer, Jake A. Plante and Julie A. Draper, "Development of Noise Dose/Visitor Response Relationships for the National Parks Overflight Rule: Bryce Canyon National Park Study", Report DOT-VNTSC-FAA-98-6, U.S. Department of Transportation, Research and Special Programs Administration, John A. Volpe National Transportation Systems Center, (1998).
15. Amanda S. Rapoza, Gregg G. Fleming, Cynthia S.Y. Lee and Christopher J. Roof, "Study of Visitor Response to Air Tour and Other Aircraft Noise in National Parks", Report DTS-34-FA65-LR1, U.S. Department of Transportation, Research and Special Programs Administration, John A. Volpe National Transportation Systems Center, (2005).
16. Nicholas P. Miller, Grant S. Anderson, Richard D. Horonjeff, Richard H. Thompson, Robert M. Baumgartner and Pam Rathbun, "Mediating the Effects of Military Aircraft Overflights on Recreational Users of Parks: Final Report", Report No. 294470.04, Harris Miller Miller & Hanson Inc., (1999).
17. National Parks Air Tour Management Act: Public Law 106-181, (2000)
18. This 5-point response scale was adopted before international standardization to a 10-point annoyance scale, International Organization for Standardization, Acoustics – Assessment of Noise Annoyance by Means of Social and Socio-acoustic Surveys, ISO/TS 15666, (2003). Then the 5-point response scale was retained throughout this multi-decade study, for consistency. Note that the cited ISO document explicitly excludes surveys about noise annoyance in recreational areas.
19. Andrew Gelman and Jennifer Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, New York, (2007).
20. R Development Core Team, "R: A Language and Environment for Statistical Computing", ISBN 3-900051-07-0, R Foundation for Statistical Computing, (2008).
21. A very technical note: Both the AIC and G-statistic criteria require the same number of data points "before" and "after" the inclusion of the new predictor. However, introduction of a new predictor into the equation sometimes results in a loss of data points (those without a value for that new predictor). Therefore, to enforce the same-number requirement, data points lost "after" inclusion were purposely tossed away during the "before" regression, as well—but just for this acceptance testing.
22. "Guide to the Expression of Uncertainty in Measurement", International Organization for Standardization, (1995) —Equation 13, with change in notation and where partial derivatives all compute to C_i.