111Equation Chapter 1 Section 1
Supporting Statement
Human Response to Aviation Noise in Protected Natural Areas
OMB Control Number 2120-0744

B. Collections of Information Employing Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results. When Item 17 on the OMB Form 83-I is checked "Yes", the following documentation should be included in the Supporting Statement to the extent that it applies to the methods proposed:

1. *Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.*

Project Rationale

Data from this collection are being used to develop regression models that quantify relationships between visitor responses to aircraft noise (per the survey instruments) and in-situ noise exposure, measured simultaneously by trained professionals. These relationships will be used to predict visitor response and set noise-impact thresholds for use in air tour management and planning. As such, the data will be used to predict percentage response of all recreational visitors exposed to aircraft noise in National Park Units covered by the National Parks Air Tour Management Act (see Supporting Statement A for a full discussion of the Act's relevance to the proposed research effort).

As noted in Supporting Statement A, noise from air tour overflights is a key management issue in the National Parks. Previous research on the response of national park visitors to aircraft noise has been limited in one major respect: responses were obtained at only a limited subset of park activities and site types (front-country short hikes and overlooks). Within the current collection cycle, sufficient dose-response data have been collected for backcountry day-hikes visitors. However, this dataset continues to be limited - we have not been able to collect sufficient data for backcountry overnight and front-country cultural/historic visitors. Data collection for these site types has not yet been completed due to logistical and funding constraints. Overnight visitors remain a challenging population to sample for two reasons. First, they represent only a small portion (<1%) of total recreational visits (See Table 2). Second, the data drop-out rate, caused by our physical inability to adequately measure various metrics of aircraft noise—at particular locations, times of day, or under adverse weather conditions, or under adverse (second-by-second) ambient-noise conditions, is much higher than normal for these types of visitors. Thus far, our data drop-out rate for overnight visitors is approximately 50%, in contrast to 20% for most other types of visits.
.

Project Purpose

In more detail, the proposed survey effort continues to expand on previous work in three ways:

1. ***Low aircraft activity at previously studied site types.*** For previously studied site types (frontcountry overlooks and short hikes), it provides additional data for low aircraft activity, to (1) obtain statistical

significance of one-or-more additional (physically important) aircraft noise metrics and (2) thereby better justify future application to low-aircraft-activity time periods. These additional data will also increase the number of sites for each site type, to enable better comparisons of site types among park units—thereby more precisely determining site-to-site variability.

2. ***New site types.*** It increases the number of site types represented in the survey collection by extending survey collection to activities/site types not previously studied (frontcountry day hikes and historical/cultural sites, backcountry multi-day hikes, and camp sites)—thereby determining these site-type “offsets” from the site types in the current database. In this way it includes different visitor experiences from those previously measured, including multi-hour and multi-day visits.

3. ***Multiple survey instruments.*** It simultaneously tests multiple survey instruments in the same settings to compare methodologies. This robust comparison of methodologies will allow researchers to identify the strengths and weaknesses of each instrument and provide the means for selecting the best survey instrument to support park management policy decisions, on a park-by-park basis.

Inference Population

The goal of this research is to produce regression models of aviation noise dose-response relationships—models that are empirically valid and generalizable to all National Park visitors exposed to aircraft noise. More than 292 million people visited national parks and other units of the National Park Service during 2014. Of these, about 64 million were recreational visits to parks in the Air Tour Management Plan program. This research will concentrate on parks in the Air Tour Management Plan program, but is not limited to these parks, as nearly all park locations experience some level of aircraft overflights, whether from commercial air tours or other air traffic (e.g., general aviation high and altitude commercial jet overflights).

Potential Respondent Universe and Selection Method

The respondent universe will be English-speaking individuals 18 years of age and older who visit specific study areas in National Parks, and engage in specific activities such as viewing scenic/ natural/historical features, observing wildlife, short-duration hiking (for one to two hours), day-hiking (more than two and generally fewer than eight hours), taking limited backcountry excursions (overnight hikes/camping), or participating in ranger-led activities. On each survey day, teams of trained surveyors will be stationed at each selected point during typical visitation hours, which may be site and/or park dependent. The surveyors will recruit study participants by contacting as many visitor groups as possible based on surveyor availability.

Survey Frames

The proposed research will inform the development of Air Tour Management Plans (ATMPs). The primary survey frame for this research consists of all National Parks, with concentration on the approximately 100 parks that require ATMPs (see Appendix B-1 for map of ATMP parks). To date, we have collected data at five of the eight parks proposed in the original request. For the proposed collection cycle, we wish to conduct surveys at approximately three additional parks.

We will coordinate with the National Park Service to select parks that represent the geographic distribution and available activity types of parks as a whole, and ATMP parks in particular, to ensure that the regression models resulting from this research are generalizable to our inference population. This coordination also will help ensure the acceptability of our results to NPS park managers, who sometimes maintain that all parks/sites are unique.

A secondary sampling frame will be site-type. Researchers have identified a finite number of general site-types, or categories of sites, which can be used to represent the range of actual sites. These categories are based on a number of visitor and visit characteristics which have been shown to influence perception of, and response to, aircraft noise. These are:

- Visitor activity at the time of exposure (hiking, viewing a feature, resting/relaxing, listening to someone/something).
- Presence of other people, especially children (i.e., crowding), and/or other human-caused sounds.
- The amount of time spent within the current visit, and/or number/length of previous visits.
- The importance of natural quiet as a reason to visit a particular location.

Together, these visit and site characteristics can define a finite number of general site-types that can be used to represent the possible range of actual sites (refer to Table 1). In other words, a site where visitors are hiking in a remote (un-crowded) location for a period >24 hours would be one type of site – multi-day backcountry.

Table 1. Proto-Typical Site Characteristics

Site Type	Visitor activity	Time spent in visit	Intensity of Visitation	Spatial Character of Visitor Use
Multi-Day Backcountry	Multiple: hiking, relaxing, camping	>24 hrs	Low	Linear
Backcountry Day-Hike	Hiking	4-12 hrs	Low-Moderate	Linear
Frontcountry Day-Hike	Hiking	4-12 hrs	Moderate-High	Linear
Front country Short Hike	Hiking	<4 hrs	Moderate-High	Linear
Frontcountry Viewpoint / Destination	Viewing	<4 hrs	High	Concentrated
Frontcoutnry Historical	Viewing, Listening to Interpretive talks	< 4 hrs	Moderate	Concentrated

NPS visit statistics give us information on the potential number of respondents in each site-type strata of our sample.[i] Visit statistics are available for those who obtain backcountry overnight permits, allowing us to determine the potential universe of backcountry overnight visits. Visit statistics are also available for type of park unit (National Park, National Historic Site, etc.). Thus, we can estimate the number of historic site visits by summing the visits to parks designated as historic sites, battlefields, memorials, or monuments (see Table 2 footnote). However, visit statistics are not collected by activity or for specific sites and trails. Thus, the relative numbers of day-hike, short-hike, and viewpoint visits are not known, but we estimate that these visits make up the remainder of the total recreational visits. Table 2 summarizes the potential respondent universe, proposed sample size, and samples collected to-date by stratum. Rather than sample in proportion to the number of visitors in each stratum, we wish to sample an equal number of visitors in each stratum to develop relationships for each site-type with equal confidence.

To date, survey collection has taken place over two field seasons, spring/summer/fall 2011 and summer 2014. These collections have been concentrated at backcountry day-hikes locations. We have collected limited amounts of data at frontcountry cultural/historic and backcountry overnight locations as well. Thus, a number of additional field seasons are required to complete the proposed sampling plan, outlined further under question #2. The work is part of an ongoing effort to study additional parks and sites with different characteristics, in order to develop robust dose-response models that can inform air tour management decisions.

Table 2. Respondent universe covered by the collection and proposed sample for the universe as a whole and for each of the strata in the proposed sample

<table>
<tr><th></th><th>2014 NPS Statistics</th><th>Proposed Sample Size (including previous research)</th><th>% of Inference Population</th><th>Current Sample Size (including previous research)</th></tr>
<tr><td>National Park recreational Visits</td><td>292,800,082</td><td>19200</td><td>.006%</td><td>7317</td></tr>
<tr><td>Backcountry Overnight Visits</td><td>1,888,095</td><td>2625</td><td>0.14%</td><td>733</td></tr>
<tr><td>Backcountry Day-Hike</td><td rowspan="4">180,870,020</td><td rowspan="4">11550</td><td rowspan="4">.006%</td><td rowspan="4">3253</td></tr>
<tr><td>Frontcountry Day-Hike</td></tr>
<tr><td>Front country Short Hike</td></tr>
<tr><td>Frontcountry Viewpoint / Destination</td></tr>
<tr><td>Frontcountry Historical</td><td>110,041,967*</td><td>2625</td><td>.002%</td><td>551</td></tr>
</table>

*Derived from a summation of visits to parks units designated as International Historic Site, National Battlefield, National Battlefield Park, National Battlefield Site, National Historic Site, National Historical Park, National Memorial, National Military Park, and National Monument.

2. *Describe the procedures for the collection of information including:*
    - *Statistical methodology for stratification and sample selection,*
    - *Estimation procedure,*
    - *Degree of accuracy needed for the purpose described in the justification,*
    - *Unusual problems requiring specialized sampling procedures, and*
    - *Any use of periodic (less frequent than annual) data collection cycles to reduce burden.*

Estimation and Inference Methods

For consistency with past studies, this proposed study will use model-based estimation and inference methods—more specifically multi-level logistic regression. Our re-analysis of all relevant past data[ii] has successfully employed such regression analysis to model visitor response as a function of aircraft noise exposure and statistically significant mediator variables. Data collected in the current cycle are being used successfully to develop dose-response relationships for backcountry day-hikes. However, sufficient data to develop relationships for backcountry overnight visitors and frontcountry cultural/historic visitors have not yet been collected. Survey data from additional sites will also expand the current database to include a higher range of dose data, improving the confidence in the resulting dose-response models across the range of aircraft dose levels.

Description of the Regression Model

The newly acquired data will expand our previous analysis into several new site types (see later tables). That analysis regressed six visitor responses to aircraft noise (two response questions, combined with three dichotomizations), against a total of seven statistically significant predictors (three aircraft noise doses and four mitigating variables, including site type). That analysis will be essentially repeated with these newly acquired data.

In brief, estimation will consist of multi-level logistic regression models, guided by techniques in Gelman and Hill (2007).[iii] In more detail, the previously successful regression model was:

$$Response = \frac{100}{1 + e^{-Z}}$$

$$\begin{aligned} Z = {} & C_0 + C_1(LeqAll) \\ & + C_2(PEnHelos) + C_3(PEnProps) + C_4(PEnHelos)(PEnProps) \\ & + C_5(SiteType) + C_6(ImpNQ\_VorMore) + C_7(AdultsOnly) + C_8(SiteVisitBefore) \\ & + N\left[0, \sigma^2_{park}\right] + N\left[0, \sigma^2_{site}\right]. \end{aligned}$$

22* MERGEFORMAT ()

In this multi-level logistic regression:

1. Response (in percent) is one of the six response/dichotomization combinations investigated: the two responses listed above, dichotomized as SomewhatOrMore, ModeratelyOrMore, and VeryOrMore – all of which refer to respondents' stated level of annoyance in response to aircraft noise.
2. The logit, Z, contains nine fixed-effect terms: the logit's intercept, the basic noise metric (LeqAll[1]), two supplemental noise metrics that proved significant (PEnHelos[2] and PEnProps) plus their interaction, and four significant mitigating variables (SiteType, ImpNQ_VorMore,[3] AdultsOnly, and SiteVisitBefore).

   The coefficients of each of these contain both a deterministic and a normal random component.
3. In addition, the logit contains the two multilevel "random-effects" terms (also normal), which the multilevel software also determines from the data.

This same model, sometimes with one or two additional predictors, is planned for this current effort. In more detail:

1. *Low aircraft activity at previously studied site types.* To satisfy the first Project Purpose (see page 2), we intend to augment this regression model with one or two additional acoustic metrics that are sensitive to low aircraft activity—e.g., the percent time that tour aircraft are audible (PTAudTours).
2. *New site types.* To satisfy the second Project Purpose, we intend to augment the categorical predictor, SiteType, to include additional factors for the newly measured site types. Of main interest is the site-type "offset" between the new site types and the two already studied: frontcountry overlooks and

[1] LeqAll = equivalent sound level (all aircraft), normalized to the respondent's visit duration. This dose variable is derived from acoustic measurements taken simultaneously with the surveys, and measures sound level and duration.

[2] PEnHelos and PEnProps are measures of the total sound energy from helicopters and propeller aircraft. Both are derived from acoustic measurements taken simultaneously with the surveys.

[3] This predictor equals "yes" for visitors who consider Natural Quiet very important or more (paraphrased from the survey instrument).

short hikes. Of secondary interest is whether the other functional dependencies in the prior dose-response relations remain valid, or perhaps require generalization or modification for these new site types.

3. *Multiple survey instruments.* To satisfy the third Project Purpose, we intend to compare responses to similar questions among the three survey instruments. In that way, we will learn the comparability of the three instruments (and their respective noise-exposure methods) in measuring nominally the same response.

In selecting the most appropriate model to fit each dataset, potential models are evaluated by first selecting the combination of dose variables, then the combination of mediator variables, that result in models with the lowest Akaike Information Criteria (AIC) values. AIC is a commonly utilized means for model selection. AIC assesses the model 'goodness of fit' using the likelihood function, while applying a penalty that increases as the number of estimated parameters increases, to discourage model overfitting .

In the first step, models including all single dose variables and combinations of dose variables are evaluated. AIC values are used to select the model that minimizes information loss (the model with the lowest AIC value). To identify the best combination of dose variables for all six dichotomizations of the *Annoy* and *Interfere* responses, the relative probabilities of all models for a given response are calculated and compared with the model with the lowest AIC value. Models with a relative probability of greater than 0.05 compared to the model with the lowest AIC value are retained as candidate models for the evaluation of mitigating variables (Step 2).

In step two, mitigating variables are added individually and in combination to each candidate model. Mitigating variables are retained if they result in models with average AIC values (across all three dichotomizations) lower than the AIC value of the model from step one. This procedure can identify the combination(s) of dose and mitigating variables that best predict visitor responses across all three dichotomizations responses simultaneously, with the objective of ensured consistency in the format of the final models to simplify implementation.

Variance Estimates and Inference Methods

Sample selection and sampling unit

This research is intended to guide decision making at National Parks for which the National Park Service and Federal Aviation Administration intend to develop Air Tour Management Plans (ATMPs), and at which there are sufficient aircraft overflights to provide a range of aircraft noise exposure against which visitor response can be regressed. Therefore, our selection of parks and sites will be informed by the ATMP program (see Appendix B-1: Map of Current ATMP Parks). In addition, as discussed above, study parks will be selected that are representative of the geographic distribution and activity types of parks nationally. Furthermore, survey collection locations within parks will be selected in coordination with National Parks Service personnel at the national and local level. To ensure that the park visitors we survey are representative of the universe of park visitors, we will take the following steps:

- English-speaking visitors over 18 will be intercepted and surveyed. The surveyors will determine English proficiency from initial conversations they have upon first contact with park visitors.
- Survey collection will take place five to seven days per week, to account for variation between weekday and weekend visitors and aircraft overflight activity.
- Survey collection will occur during typical visitation hours, which may be site and/or park dependent—to account for overflight variation during early morning, mid-morning, afternoon and perhaps sunset.

- Survey collection will be performed at a variety of site types and specific sites, to capture the range of activities in which visitors participate.

Prior empirical dose-response research in national parks (cited in the Project Rationale section, above) revealed no significant effects attributable to characteristics of sub-groups within the population (i.e., demographics). Therefore, the sampling for the proposed research will not be further stratified.

The sampling unit for this collection is the individual park visitor. Table 3 describes our anticipated sampling: first parks, then park area, then sites, then visitors (respondents).

This table represents our best estimate of the total number of surveys we will collect in this study, and our burden hour estimate (15 minutes per respondent, 4200 hours per Supporting Statement A, Question 12) is based upon this number.

Table 3. Survey collection plan, including expected number of respondents.

Site types previously measured	Park area	Site types	Target sampling sizes				
			Parks	Total sites (see note 1)	Respondents per site, per survey instrument	Respondents per survey instrument	Total Respondents
Yes	Frontcountry	Short hikes Overlooks	3	7	175	1,225	3,675
No	Frontcountry	Day hikes Historical sites	5	10	175	1,750	5,250
	Backcountry	Day hikes Multi-day hikes Campgrounds	3	15	175	2,625	7,875
TOTALS (see note 2)			11	32	-------	5,600	16,800

Note 1: Total sites split more-or-less evenly among the parks and site types.

In further explanation of this table, each selected park will contain several selected sites, depending upon the park type and its availability of desired site types. In all, we expect a total of 11 parks, some with only one site type and some with multiple site types—for an expected total of 32 sites (as tabulated).

As the table shows, we intend to sample these 32 sites separately in two site-type groups (those previously measured and those not), to conform with two of the purposes of this effort: 7 sites for the two previously studied site types combined, plus 25 sites for the 5 new site types combined.

Note that some of the new site types may have lower visitation rates than overlook and short-hike types. Nevertheless, we intend to extend survey collection, as required, to fulfill our target of 175 respondents for each site, per survey instrument.

The original application for this data collection proposed that surveys would be conducted over the course of a 20-day period (average) at each of the parks surveyed, depending upon aircraft activity and high- or low-visitation rates. Based on the data collection experiences to date (see below), we believe this continues to be an accurate schedule for low-visitation, backcountry sites. Sites with higher visitation generally require a 7-10 day data collection period. We will allow additional time for setup, breakdown, weather and other unforeseen factors that may prevent survey collection on certain days. Survey collection will be extended to fulfill our target requirements for number of observations, if necessary.

This data set will enable a robust analysis of within-park and among-park variance, for all the site types at which we administer surveys (see below for representative sample-size calculations).

Selection probabilities
Because surveyors will not know in advance how many visitors will arrive at each site, random sampling of visitors (e.g., interviewing every third visitor) will not be possible; to do so would risk not obtaining a sufficient sample. Therefore, the selection probability for each visitor to a survey site will be 100%, subject only to interviewer availability.

Our target of 175 respondents per site per survey instrument is consistent with results achieved in collections to date. In addition, our target accommodates reasonably foreseeable response exclusions due to lack of noise exposure data or errors in response (e.g., unreadable, damaged or inappropriately completed surveys). Based on both current and previous research results, we anticipate that our target number of observations is more than adequate to support the planned multilevel regression—as well as auxiliary, bivariate statistical comparisons.

Survey Collections to Date
To date, we have collected 4817 surveys in 5 national parks, as shown in Table 4, for an estimated burden of 1204 hours.

Table 4. Summary of survey data collected to date.

Park	Site	Front -country Short Hike	Front-country Historical Site	Back-country Day-Hike	Back-country Overnight Hike/Campground	
Grand Canyon	**Hermit Trail**			449	135	
Grand Canyon	**Grandview Trail**			291	126	
Grand Canyon	**Tusayan Ruins**		374			
Zion	**Taylor Creek**			453		
Zion	**West Rim Trail**			182	127	
Bryce Canyon	**Fairyland Trail**			1102		
Glacier	**Hidden Lake Trail**	335		181		
Glacier	**Sperry Trail**			540	345	
Rainbow Bridge	**Rainbow Bridge**		177			
	Total Sites	**1**	**2**	**7**	**4**	
	Total Surveys	**335**	**551**	**3198**	**733**	**4817**

Comparing Table 3 to Table 4 shows that we have thus far concentrated on data collection at backcountry locations. We have collected 80% of our target number of day-hike surveys at 11 backcountry locations, 18% of our target number of overnight-hike surveys at 4 locations, of 10% of our target number of surveys at 2 of 10 frontcountry day hikes and historical sites, and 9% of our target number of surveys at 1 of 7 frontcountry overlook and short hike sites. Data collection efforts in the current cycle will thus concentrate on the under-represented frontcountry and backcountry overnight locations.

Survey Response Rates to Date
Survey response rates are provided in Table 5. The "overall response rates" are calculated as the total number of respondents who agreed to take the surveys, as a percentage of the total number of park visitors "intercepted" by our survey teams, and invited to participate. The "completed survey response rates" reflect the same pool of respondents who agreed to participate, but remove those respondents who did not end up completing the surveys. The average response rates across all survey instruments and

sites are shown in the bottom row of Table 5. The weighted averages reflect each study site's contribution to the total sample collected to date. In many cases, our response rates have been higher than the expected response rate of 70%, primarily at backcountry locations. This may be due to a number of factors unique to these locations: 1) visitation is low (sometimes as few as 20 visitors per day), allowing surveyors to approach and engage each visitor, 2) visitors are generally not rushed or in a hurry, and 3) visitors are quite willing to help improve their National Parks.

Table 5. Survey Response Rates to Date

Park	Site	Percentage of Total Sample	Overall Survey Response Rates	Completed Survey Response Rates
Grand Canyon	Hermit Trail	12.1%	93.4%	92.8%
Grand Canyon	Grandview Trail	8.7%	91.9%	91.8%
Grand Canyon	Tusayan Ruins	7.8%	68.0%	67.8%
Zion	West Rim Trail	6.4%	94.1%	93.8%
Zion	Taylor Creek Trail	9.4%	88.4%	87.9%
Bryce Canyon	Fairyland Trail	22.9%	89.4%	88.2%
Glacier	Hidden Lake Trail	10.7%	87.2%	86.9%
Glacier	Sperry Trail	18.4%	84.5%	84.1%
Rainbow Bridge	Rainbow Bridge*	3.7%	45.9%	43.4%
	Unweighted Average Response Rate		82.5%	81.9%
	Weighted Average Response Rate		85.9%	85.3%

*****Note: The extreme weather conditions (temperatures 90-110 ºF) at Rainbow Bridge resulted in lower-than-average response rates.

Sample Size Calculations in Support of Table 1

Preface concerning the three survey instruments

The three survey instruments are the products of three separate scientific groups that have been gathered together by the Volpe Center into one combined project.

To calculate sample size, we provide details here for only the first survey instrument (referred to as "The human response to aviation noise - visitor survey, version 1" in the Collection Instruments section of Supporting Statement A). This instrument employs more alternative responses and more predictors in its dose-response mathematics than do the other two instruments. Its analyses will, therefore, most likely require more data points than the other two. It is for that purpose that we chose it for these sample-size calculations. Our target sample sizes, given below, should therefore provide more-than-adequate samples for the other two survey instruments, as well.

Sample partitioning among parks, sites and visitors

Sample size calculations hinge upon the results of the previous analysis. Since that work is still ongoing and not yet reported, we include extra detail here.

In the prior analysis the regression model of Eq.2 was fit within the "R" Statistics System, using the function lmer() in the package lme4.[iv] Below is the "R" output for one of those ongoing regressions—the response/dichotomization of "Annoy_ModeratelyOrMore", which is of particular interest for ATMP application:

Generalized linear mixed model fit by the Laplace approximation
Formula: "Annoy_MorMore ~ (1|Park) + (1|Site) + 1 + SiteType + LeqAll + PEnHelos + PEnProps + SiteVisitBefore + AdultsOnly + I(PEnHelos * PEnProps) + ImpNQ_VorMore"

AIC	BIC	Log Likelihood	deviance
1514	157 3	-745.8	1492

Random effects:

Groups	**Name**	**Variance**	**Standard Deviation**
Site	Intercept	0.0152672	0.123560
Park	Intercept	0.0067854	0.082374
Number of obs: 1572, groups: Site, 9; Park, 4			

Fixed effects:

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	-4.0838577	0.3667824	- 11.134	< 2e-16 ***
SiteTypeShortHike	1.1275413	0.2063434	5.464	4.64e-08 ***
LeqAll	0.0158053	0.0083092	1.902	0.05715
PEnHelos	0.0083059	0.0026081	3.185	0.00145 **
PEnProps	0.0066900	0.0029056	2.302	0.02131*
SiteVisitBeforeYes	0.5513902	0.1699233	3.245	0.00117 **
AdultsOnlyYes	0.1855097	0.1569553	1.182	0.23723
I(PEnHelos * PEnProps)	0.0001063	0.0000916	1.161	0.24567
ImpNQ_VorMoreYes	0.7627218	0.1481054	5.150	2.61e-07 ***

Significance codes: 0 ‘***’ , 0.001 ‘**’, 0.01 ‘*’

Correlation of Fixed Effects:

	(Intr)	Site Type Short Hike	LeqAll	Pen Helos	Pen Props	Site Visit Before Yes	Adults Only Yes	I(PEnHelos * PEnProps)
SiteType ShortHike	-0.194							
LeqAll	-0.614	-0.104						
PEnHelos	-0.198	-0.133	-0.363					
PEnProps	-0.238	-0.027	-0.243	0.727				
SiteVisitBefore Yes	-0.141	0.001	0.048	-0.002	-0.030			
AdultsOnlyYes	-0.297	-0.105	0.003	0.002	-0.018	0.057		
I(PEnHelos * PEnProps)	0.026	0.059	-0.015	-0.180	-0.318	0.051	-0.002	
ImpNQ_ VorMoreYes	-0.317	0.030	0.020	-0.001	-0.008	-0.012	-0.031	0.001

Note the following: The AdultsOnly predictor was accepted in this regression for consistency with the other regressions, in which it was more significant. The interaction term was accepted because it made its two additive terms more physically (acoustically) realistic.

In this regression, specific park and specific site are the two random-effect (multilevel) parameters. By using multilevel methods with these two parameters, we learn the prediction uncertainty of our results for future application to single parks, single sites—one at a time—without underestimating that uncertainty.

To partition proposed additional data points intelligently by park and site, we need to determine, from this output, the relative uncertainty variances for park, site and visitor (respondent).

These three variances derive from the regression's logit in Eq.2, repeated here:

$$\begin{aligned} Z &= C_0 + C_1(LeqAll) \\ &+ C_2(PEnHelos) + C_3(PEnProps) + C_4(PEnHelos)(PEnProps) \\ &+ C_5(SiteType) + C_6(ImpNQ\_VorMore) + C_7(AdultsOnly) + C_8(SiteVisitBefore) \\ &+ \mathrm{N}\left[0, \sigma^2_{park}\right] + \mathrm{N}\left[0, \sigma^2_{site}\right]. \end{aligned}$$

33* MERGEFORMAT ()

Mathematically, the variance of Z is straightforwardly computed with an adaption of equation E.3 from the *ISO Guide:* [v]

$$\begin{aligned} Var[Z] &= Var[\text{visitors}] + Var[\text{parks}] + Var[sites] \\ &= Var[\text{first 9 terms of } Z] + \sigma^2_{park} + \sigma^2_{site} \\ &= \sum_{i=1}^{9}\left(\frac{\partial Z}{\partial C_i}\right)^2 Var[C_i] + 2\sum_{i=1}^{9-1}\sum_{j=i+1}^{9}\left(\frac{\partial Z}{\partial C_i}\right)\left(\frac{\partial Z}{\partial C_j}\right)\sqrt{Var[C_i]\,Var[C_j]}\,\rho_{ij} + \sigma^2_{park} + \sigma^2_{site}, \end{aligned}$$

4

4* MERGEFORMAT ()

where C_i are the regression coefficients within Z, Var[C_i] are the uncertainty variances of those coefficients, and ρ_{ij} is the coefficient correlation matrix. The first line in this equation is valid because multilevel regression considers variances at the three levels (park, site, and visitor) to be independent.

Park and site uncertainty variances are directly tabulated in the input: *Var*[parks]= 0.00679, *Var*[sites]= 0.0153. These variances derive from the standard mathematics of multilevel regression, when park and site are entered as "random effects"—that is, "levels"—in the regression. Note that these are the "unexplained" variances, after the regression controls for aircraft noise levels and all significant mitigating variables. Before control, these park and site variances were much larger.

Determining *Var*[visitors] requires evaluation of the two summation terms in Eq.4. As is apparent from the equation, this visitor variance depends upon the coefficient variances and their mutual correlations—as well as the particular values of the predictors (if they appear as a result of a partial derivative). Because the logit is "linear" in the coefficients, each partial derivative equals that coefficient's predictor value—e.g., $\partial Z/\partial C_1 = LeqAll,$ and so forth. And the partial derivative of C_0 equals unity.

As with any regression, the residual variance is a minimum at the predictor centroid. At that centroid, all the predictor values in Eq.4 equal zero, so only the variance of the logit's intercept, C_0, remains.

To most easily obtain that minimum variance, we recompute the regression after centering all the predictors—including the categorical ones, which we center by converting them to 0's and 1's —then averaging over all data points, and then centering on that average. Note that we can do this even for the SiteType categorical variable, since it has only two factors. This results in the following "R" output (without the now-unnecessary correlation matrix):

Generalized linear mixed model fit by the Laplace approximation

Formula: "Annoy_MorMore ~ (1|Park) + (1|Site) + 1 + SiteType + LeqAll.c + PEnHelos.c + PEnProps.c + SiteVisitBefore.c + AdultsOnly.c + I(PEnHelos.c * PEnProps.c) + ImpNQ_VorMore.c"

AIC	BIC	Log Likelihood	Deviance
1514	1573	-745.8	1492

Random effects:

Groups	**Name**	**Variance**	**Standard Deviation**
Site	Intercept	0.0152672	0.123560
Park	Intercept	0.0067854	0.082374
Number of obs: 1572, groups: Site, 9; Park, 4			

Fixed effects:

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	-2.080e+00	1.799e-01	-11.567	< 2e-16 ***
SiteTypeShortHike	1.128e+00	2.063e-01	5.466	4.61e-08 ***
LeqAll.c	1.581e-02	8.313e-03	1.901	0.057244 .
PEnHelos.c	1.124e-02	3.057e-03	3.677	0.000236 ***
PEnProps.c	1.319e-02	4.691e-03	2.811	0.004935 **
SiteVisitBeforeYes.c	5.514e-01	1.700e-01	3.244	0.001179 **
AdultsOnlyYes.c	1.855e-01	1.569e-01	1.182	0.237157
I(PEnHelos.c * PEnProps.c)	1.294e-04	9.129e-05	1.418	0.156214
ImpNQ_VorMoreYes.c	7.627e-01	1.481e-01	5.151	2.59e-07 ***

Significance codes: 0 '***' , 0.001 '**', 0.01 '*'

Therefore the relevant (controlled) error variances at the data centroid are:

- Park: 0.0068
- Site: 0.0153
- Visitor: 0.0324 (square of 0.1799).

Next we multiply the centroid-visitor variance by a factor of 9, to convert it to an approximate "off-centroid" value—obtaining 0.13. This factor of 9 is based upon regression plots and their 95% prediction uncertainty bounds from the ongoing analysis. Then the relevant (controlled) error variances of the data centroid are:

- Park: 0.0068

- Site: 0.0153
- Visitor: 0.32.

Then ignoring incremental costs for the moment, the most efficient way to reduce error variance for future measurements is to sample in proportion to these variances:

- 2.25 sites/park
- 20 visitors/site.

Adjustment due to expected good-data rate. Our past study resulted in a good-data rate of 60%. Most of the data drop-out was caused by our physical inability to adequately measure various metrics of aircraft noise—at particular times of day, or under adverse weather conditions, or under adverse (second-by-second) ambient-noise conditions. A much smaller part of the drop-out was due to incomplete questionnaires (for visitors who did consent to the interview).

This data drop-out occurred automatically in our regression calculations, whenever a visitor was lacking one of the regression predictors.

Since we anticipate the same good-data rate for future measurements, we adjust our target "respondents per site" upwards by the reciprocal of 0.6—that is, by 1.7. This results in:

- 2.25 sites/park
- 35 visitors/site.

Expected response rates. Based on the response rate achieved on other recent aviation-noise studies in Hawaii Volcanoes National Park, and Haleakala National Park, we expect an overall response rate of 70% or greater.[vi,vii] We see no reason why this will not be true, as well, for the present survey. Note that this is the acceptance rate of visitors when they are approached to take the survey. As such it is independent of the good-data rate, which applies to those actually-taken surveys.

Incremental cost adjustments. Incremental sampling costs modify these desired ratios. Our cost multipliers are: 2 for sites/backcountry park, and 5 for respondents/site—yielding:

- 2.25 sites/frontcountry park, and 5 sites/backcountry park
- 175 visitors/site.

These cost multipliers are best estimates prior to knowing the actual parks and sites that will be included. They will vary significantly from park to park and from site to site. In general, however, the extra visitor cost increment is always much less than the site and park increments. Moreover, that visitor increment is approximately the same for front and back country.

Note the cost aspect of this calculation boosts the number of sites beyond what is needed for optimum variance reduction. This site-number boost has a very important positive advantage, further justifying the boost. When these results are applied during ATMP studies, the desired prediction from the regression model will be for (1) year-long averages of a huge number of visitors, but (2) only for one site at one park (the site of that particular ATMP study). Therefore, site-average predictions from the regression are of no use. Instead, we must reduce unexplained park and site variance to an absolute minimum, to allow satisfactory one-park, one-site predictions. And boosting the number of parks and sites, based upon costs, therefore also makes sense for increased prediction certainty.

Further, NPS park managers sometimes maintain that parks/sites are "unique." An increased sampling of parks/sites will help ensure that individual park managers recognize a park/site similar to theirs (in their judgment)—thereby helping them to accept application of the analysis results to their "unique" situation.

Resulting target sampling ratios. Table 6 contains the resulting target sampling ratios.

Table 6. Target sampling ratios (minimum variance, adjusted for incremental costs and good-data rate)

Site types previously measured	Park area	Corresponding site types	Target sampling ratios	
			Sites per Park	Respondents per Site
Yes	Frontcountry	Overlooks Short hikes	2.5	175
No	Frontcountry	Day hikes Historical sites,		
	Backcountry	Day hikes Multi-day hikes Camp sites	5	175

The target sampling ratios in Table 6 inform all the sample-size computations that follow.

Site types previously measured (first Project Purpose)
For site types previously measured (overlooks and short hikes, in frontcountry), we intend to augment the previous regressions with one or two additional acoustic metrics that are sensitive to low aircraft activity, such as percent time audible of aircraft (PTAud). We attempted to do this during the ongoing analysis, but were not successful in achieving our uncertainty/power goals.

For example, here is the "R" output for the most promising attempt:

Generalized linear mixed model fit by the Laplace approximation

Formula: "IntWithNQ_MorMore ~ (1|Park) + (1|Site) + 1 + SiteType + LeqAll + PEnHelos + PEnProps + PTAudTours + SiteVisitBefore + AdultsOnly + ImpNQ_VorMore + I(PEnHelos * PEnProps) + log10(PTAudTours + 0.001)"

AIC	BIC	Log Likelihood	Deviance
1488	1554	-730.9	1462

Random effects:

Groups	**Name**	**Variance**	**Standard Deviation**
Site	Intercept	0.132891	0.132891
Park	Intercept	0.00883	0.093968
Number of obs: 1208, groups: Site, 8; Park, 4			

Fixed effects:

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	-2.574e+00	5.842e-01	-4.406	1.05e-05 ***
SiteTypeShortHike	5.830e-01	2.168e-01	2.689	0.007161 **
LeqAll	3.177e-02	8.817e-03	3.604	0.000314 ***
PEnHelos	9.758e-03	2.530e-03	3.857	0.000115 ***
PEnProps	4.280e-03	2.899e-03	1.476	0.139812

PTAudTours	7.145e-03	5.725e-03	1.248	0.212076
SiteVisitBeforeYes	5.635e-01	1.732e-01	3.253	0.001141 **
AdultsOnlyYes	4.742e-01	1.561e-01	3.037	0.002389 **
I(PEnHelos * PEnProps)	9.732e-05	9.224e-05	1.055	0.291401
ImpNQ_VorMoreYes	2.500e-01	1.344e-01	1.860	0.062882 .
log10(PTAudTours + 0.001)	-7.551e-01	4.425e-01	-1.706	0.087967 .

Significance codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.'

In this regression, log10(PTAudTours + 0.001) is the new predictor, which has a p-value of only 0.088 in this regression. Corresponding to that is a z-value of –1.7.

To improve this regression performance, we need to scale that z-value from –1.7 to –2.8 —to obtain 95% certainty with 80% power. The required z-value ratio is 1.65 and its square equals 2.7. In turn, this squared value times the number of data points currently in the regression (1208 from the "R" output) indicates we need a total of 3260 good data points to achieve our goal. This, minus the 1208 already obtained, says we need approximately 2000 additional good points.

However, this estimate is overly pessimistic for the following reason. When we obtain additional data points for these two previously studied site types, we plan to concentrate our efforts on times of day with low aircraft activity (early morning and late afternoon), the direct opposite of the focus for all previous data collection. And for this reason, the newly acquired data points will have low values of PTAudTours, thereby contributing far more strongly to that predictor's regression coefficient.

From our acoustical experience with noise metrics sensitive only to low-aircraft activity (such as PTAudTours), it is our conclusion that we can safely reduce the required 2000 data points to one-third that value—that is, approximately 700 additional good data points (surveyed visitors). Remember that we will add the newly collected data to existing data for future ATMP analysis. Then we must multiply this by 1.7 to account for our expected good-point rate of 60%, to yield 1200 additional respondents.

Applying this sampling target (1200 additional respondents) to the sampling ratios in Table 6, we obtain the target sample sizes in Table 7.

Table 7. Target sample sizes: Previously measured site types – single instrument

Park area	Site types	Target sampling sizes			
		Parks	Total sites (see note 1)	Respondents per site	Total respondents (see note 2)
Frontcountry	Short hikes Overlooks	3	7	175	1,225

Note 1: Total sites split more-or-less evenly among the parks and site types.
Note 2: For all three survey instruments combined, total respondents is triple this value.

This table comprises the top portion of Table 3, above.

Site types not previously measured (second Project Purpose)
For each site type not previously measured, we plan to create a new factor within the categorical predictor SiteType, and desire to measure that factor's regression coefficient with adequate certainty and power.

For scaling, here is the relevant prior output (copied from above):

Generalized linear mixed model fit by the Laplace approximation
Formula: "Annoy_MorMore ~ (1|Park) + (1|Site) + 1 + SiteType + LeqAll + PEnHelos + PEnProps + SiteVisitBefore + AdultsOnly + I(PEnHelos * PEnProps) + ImpNQ_VorMore"

AIC	BIC	Log Likelihood	Deviance
1514	1573	-745.8	1492

Random effects:

Groups	**Name**	**Variance**	**Standard Deviation**
Site	Intercept	0.0152672	0.123560
Park	Intercept	0.0067854	0.082374
Number of obs: 1572, groups: Site, 9; Park, 4			

Fixed effects:

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	-4.0838577	0.3667824	-11.134	< 2e-16 ***
SiteTypeShortHike	1.1275413	0.2063434	5.464	4.64e-08 ***
LeqAll	0.0158053	0.0083092	1.902	0.05715 .
PEnHelos	0.0083059	0.0026081	3.185	0.00145 **
PEnProps	0.0066900	0.0029056	2.302	0.02131 *
SiteVisitBeforeYes	0.5513902	0.1699233	3.245	0.00117 **
AdultsOnlyYes	0.1855097	0.1569553	1.182	0.23723
I(PEnHelos * PEnProps)	0.0001063	0.0000916	1.161	0.24567
ImpNQ_VorMoreYes	0.7627218	0.1481054	5.150	2.61e-07 ***

Significance codes: 0 ‘***’, 0.001 ‘**’, 0.01 ‘*’, 0.05 ‘.’

In this regression, SiteType has two factors: ShortHike and Overlook (the reference factor). Because this predictor has only two factors, the difference between ShortHike and Overlook is contained entirely within the output’s regression coefficient for the non-reference factor, ShortHike. The z value for that coefficient is 5.46 (highly certain) while its “effect size” is 1.13. Therefore, the 1572 data points in this regression were able to determine a site-type “offset” between Overlook and ShortHike of 1.13, with very high certainty.

First we scale down the number of data points to just achieve our goal: 95% certainty, 80% power. The corresponding z-value goal is 2.8, which divides into 5.46 to obtain 1.95, which squares to 3.8. We therefore would have achieved our goal with only (1572)(1/3.8) = 414 data points, totaled over the 2 site types—for that effect size. That comes to 210 data points per site type.

Next we need to scale this value of 210, to account for the expected effect size of newly measured site types. The only bench mark we have to these unmeasured effect sizes is our current effect size, and we adopt this as our goal here. Therefore we need 210 good data points per site type. We multiply this by 5 for the new site types, to get 1100 additional good data points. Then multiplying by 1.7 yields a requirement of 2000 additional respondents. Finally, we multiply this by 2, to account for all the approximations in its computation—thereby yielding a sampling target of 4000 additional respondents.

We then wish to apply this sampling target (4000 additional respondents) to the sampling ratios in Table 2, above. First we divide 4000 by 175 respondents/site, to obtain 23 new sites. Then we split these 2:3 between front and backcountry, in proportion to the number of new site types in each—yielding 10 and 15 sites each, respectively. Then we use the two sites/park values in Table 8, to determine the required number of parks: 5 frontcountry and 3 backcountry. Our final sample-size targets appear in that table.

Table 8. Target sample sizes: New site types –single survey instrument

Park area	Site types	Target sampling sizes			
		Parks	Total sites (see note 1)	Respondents per site	Total respondents
Frontcountry	Day hikes Historical sites	5	10	175	1,750
Backcountry	Day hikes Multi-day hikes Campgrounds	3	15	175	2,625
				TOTAL (see note 2)	4,375

Note 1: Total sites split more-or-less evenly among the parks and site types.
Note 2: For all three survey instruments combined, total respondents is triple this value.

Survey-instrument comparison (third Project Purpose)

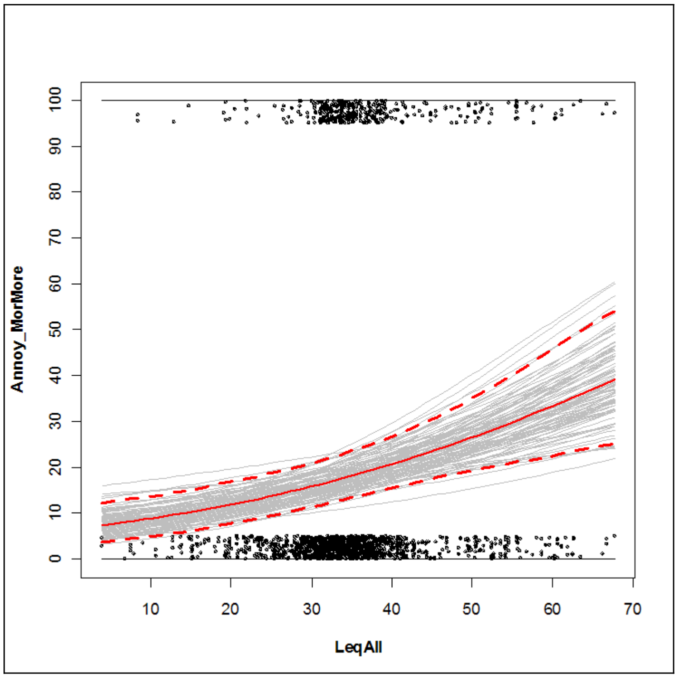
The third project purpose is to compare results from the three survey instruments, to help determine their relative strengths and weaknesses. The quantitative measure of strength/weakness depends upon how well their results compare for the overlap survey questions.

To make that comparison, we intend to regress all three sets of responses in the manner we have described above, separately by instrument. Then we will plot the three regressions against the most important acoustic metric, along with their uncertainty bounds, to look for overlap.

Figure 1 shows such a plot from the ongoing analysis, for just one survey instrument. The main acoustic dose is plotted horizontally, while the response is plotted vertically. The small points near 0% and 100% are the underlying data, one point per visitor respondent. The solid curve is the resulting dose-response regression line.

The gray lines are a one-tenth sample of 1000 simulations of that regression (sampled from the covariance matrix of the regression coefficients, augmented by the error variances of the multilevel fixed effects). In the figure, the dashed lines encompass 95% of those 1000 simulations and hence are 95% certainty bounds on the regression.

FIGURE 1. A sample prior regression, along with its
95% confidence bounds.



Non-overlap of these particular certainty bounds is not our relevant measure, however. First we must expand them from 50% to 80% power. And second we must contract them by the square root of two, because the two sets of certainty bounds are independent from each other. Only then will their overlap or non-overlap measure sameness or not, with 95% certainty and 80% power. This we intend to do, separately for each pairing of results from the three survey instruments.

As is obvious, overlap might differ for different regions of LeqAll—that is, overlap might be a function of LeqAll. *A priori* we expect overlap towards the low and high ends of the LeqAll data, where data points are sparse.

Our particular question is whether we find overlap around the data centroid and therefore along the full LeqAll axis. Such overlap would show that we failed to prove a significant difference in results from the three survey instruments. In turn, that would indicate we could use a simpler instrument (actually simpler in its required acoustic-metric determination) as substitute for a more complex instrument. This will help us in our listing of trade-offs between the instruments when determining aircraft impact for the ATMPs.

Because of the complexity of the planned overlap analysis, we have not computed required sample sizes for this determination. However, we note with satisfaction the relative narrowness of our existing confidence bounds. Moreover, we will be tripling the total number of data points (from 2500 to 2500+5600) after our additional measurements—thereby narrowing these bounds further by perhaps a factor of 1.7. Will this be enough data? That depends entirely upon the “effect size” due to a change in survey instruments, and we have no estimate at all about that size.

Nevertheless, even without a computation here of required sample size, we are optimistic that we will be able to say either (1) “yes” the instruments are comparable or (2) “no” they are not—at least over some particular portion of the dose (LeqAll) range. Moreover, we are prepared to make an administrative decision about future use of these survey instruments—even if the answer to our question is “maybe”—based upon their non-mathematical strengths/weaknesses.

One additional point: We understand that we might obtain a definitive answer to this project purpose, even if we reduce the number of respondents for the second and third instrument. However, since these instruments will be administered simultaneous with the first instrument (and require the same time burden per respondent), we see no reason not to collect the same full set of data for each instrument. That is what we plan. That also satisfies the objectives of all three scientific groups on our combined research team. No one group is thereby short changed.

Unusual Sampling Procedures

Certain site types, such as remote backcountry locations, typically receive fewer visitors than other site types, such as popular scenic overlook areas. These remote locations may require more intensive survey collection (i.e., longer collection periods or multiple survey efforts at a particular location) to achieve the desired number of observations.

Data Collection Cycles

The survey protocol will include a screening question to eliminate park visitors who have previously participated in the survey from the collection. Thus, for any given park visitor, participation in the survey will be a one-time, non-recurring, event. The screening question will read as follows:

> "Do you recall having taken a survey in a National Park before?" If the visitor answers "yes," the interviewer will then ask, "Do you remember the topic of the survey?"

Visitors who remember having previously taken a survey on aviation noise will not be included in the data collection.

3. *Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.*

Response Rate and Non-response Issues

Data will be collected on-site using the attached survey instruments and trained interviewers. The presence of the interviewers should lower the incidence of item non-response in the surveys. In addition, there are observable visitor characteristics that would allow for a meaningful non-response bias analysis. These characteristics are:

- Group size
- Presence of children in the group
- Type of activity (i.e., day hiking, backpacking).

These factors have been noted for all groups in the current database, and will be noted for all future-data groups, as well—including those who choose not to participate in the survey. In turn, these factors will be used to test for non-response bias within the survey data.

Brief Summary

As discussed in Supporting Statement A, all three of the survey instruments proposed for this research effort are based on survey instruments that have been previously approved by OMB.[4] As a result of these previous studies, there is a database of approximately 2500 visitor responses with associated direct measurements of aircraft noise exposure (all associated with frontcountry locations and the first survey instrument, however). This prior research did not yield results that could definitively be generalized and

[4] OMB Nos. 1024-0088, 2120-0610, 0701-0143, 1024-0224

applied to the entire universe of visitors, parks, and site types—in particular to the site types newly added here. A number of salient aspects of the relationship between noise exposure and visitor response were discovered, but the limited focus on short hikes and overlooks has limited the generalizability of these relationships.

Under the current cycle, we have collected information over the course of 2 summer seasons in 5 different parks. The 4800 surveys collected represent approximately 80% of original goals for backcountry day-hikes, 18% of original goals for backcountry overnights, and 10% of original goals for the frontcountry site types. Using this data, dose-response relationships are being developed for backcountry hikes. Relationships have not yet been developed for backcountry overnight visitors and frontcountry cultural/historic visitors as sufficient data have not yet been collected. Data for these site types have not yet been collected due to logistical and funding constraints and the desire to evaluate the initial backcountry dataset. It is anticipated that an additional 2-3 summer seasons in at least 3 additional parks will be required to provide sufficient data for these site-types.

It is important to continue the research to achieve the data acquisition levels originally proposed and to develop an understanding of whether the context of a visitor's park experience mediates his or her response to aircraft noise. This additional research will help to:

1) further provide an understanding of 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) Identify additional site-specific or visitor-specific factors which may significantly influence the visitor response.

Data Sampling Regime and Reliability

Because the proposed survey collection plan includes multiple locations within each park, we are confident of achieving collections that come closer to being representative of the true range of park visitors and activities than surveys conducted at single locations. This work is part of an ongoing effort to study as many parks and sites as possible, and therefore the studies outlined herein will be incorporated into the body of existing data on response to aviation noise exposure in National Parks, further enhancing generalization.

4. *Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.*

The survey instrument designs proposed for this research are based on tested, proven instruments used in previous studies of aviation noise impacts in National Parks. Prior to the initial collections in this cycle, pre-testing was conducted in National Parks and similar settings (e.g., National Historic Sites). Each pre-test comprised fewer than 10 respondents. Pre-testing assessed question wording, question order, response-scale design and other aspects of the survey instruments. Any aspects of the instruments that are found to be unclear or confusing were refined.

5. *Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.*

Data collection and analysis will led by: U.S. Department of Transportation, John A. Volpe National

Transportation Systems Center

Amanda Rapoza	(617) 494-6339
Cynthia Lee	(617) 494-6340
Kristin Lewis, PhD	(617) 494-2130
Joshua Hassol, PhD	(617) 494-3722
Erika Sudderth, PhD	(617) 494-3128

With consultation on survey design from:

Steve Lawson, PhD, Resource Systems Group	(802) 295-4999
Britton Mace, PhD, Southern Utah University	(435) 865-8569
Robert Manning, PhD, University of Vermont	(802) 656-3096
Peter Newman, PhD, Colorado State University	(970) 491-2839

References

[i] Ziesler, P. and B. Street. 2015. Statistical abstract: 2014. Natural Resource Data Series NPS/NRSS/EQD/NRDS – 2015/775. National Parks Service, Fort Collins, Colorado.

[ii] Anderson, G.S. Rapoza, A.S., Fleming G.G. and Miller, N.P, (2011). Aircraft noise dose-response relationships for National Parks *Noise Control Engineering Journal*. 59(5), 519-540.

[iii] Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press.

[iv] R Development Core Team (2008). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. Available from: http://www.R-project.org.

[v] International Organization for Standardization, ISO (1995), *Guide to the Expression of Uncertainty in Measurement*, Equation E.3 (with slight change of nomenclature).

[vi] Lawson, S. R., Hockett, K., Kiser, B., Reigner, N., Ingram, A., Howard, J., et al. (2007). *Social Science Research to Inform Soundscape Management in Hawaii Volcanoes National Park: Final Report.* Virginia Polytechnic Institute and State University, Department of Forestry, College of Natural Resources, Virgina. Available from: http://www.faa.gov/about/office_org/headquarters_offices/arc/programs/air_tour_management_plan/park_specific_plans/Hawaii_Volcanoes.cfm

[vii] Lawson, S. R., Hockett, K., Kiser, B., Reigner, N., Ingram, A., Howard, J., et al. (2007). *Social Science Research to Inform Soundscape Management in Haleakala National Park: Final Report.* Virginia Polytechnic Institute and State University, Department of Forestry, College of Natural Resources, Virgina. Available from: http://www.faa.gov/about/office_org/headquarters_offices/arc/programs/air_tour_management_plan/park_specific_plans/Haleakala.cfm