A.3.5 MBE_{σ} for PY2

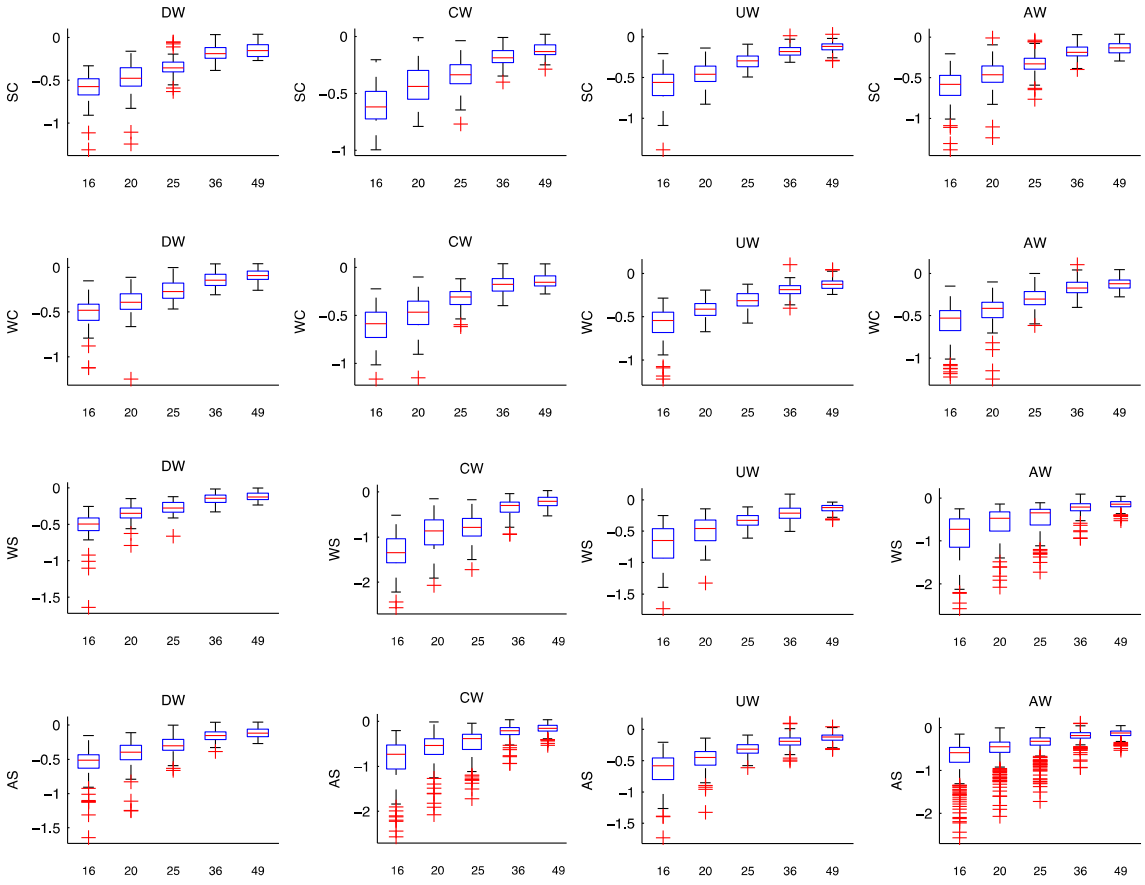

Figure A-17. Distributions of the MBE_{σ} for the PY2 model per N, given for each meteorological condition (i.e., downwind DW, crosswind (CW), upwind (UW), all wind conditions combined (AW), strongly convective (SC), weakly convection (WC), weakly stable (WS), and all stability classes combined (AS). NOTE: The vertical scale is not the same in each plot. It is varied so that the distribution for each meteorological condition can be seen.

A.3.6 MBE_σ for OKG

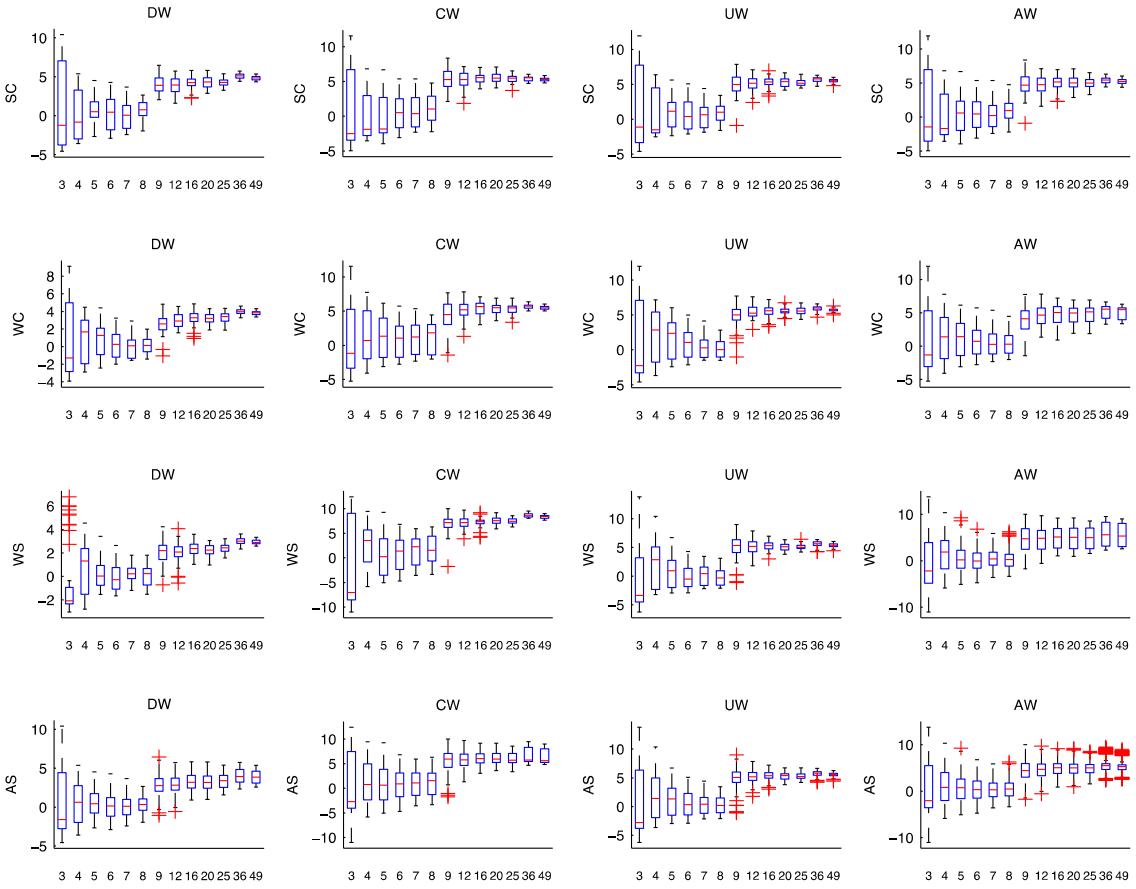

Figure A-18. Distributions of the MBE_σ for the OKG model per N, given for each meteorological condition (i.e., downwind DW, crosswind (CW), upwind (UW), all wind conditions combined (AW), strongly convective (SC), weakly convection (WC), weakly stable (WS), and all stability classes combined (AS). NOTE: The vertical scale is not the same in each plot. It is varied so that the distribution for each meteorological condition can be seen.

B Supporting Material

B.1 Starting Parameters & Bounds for Coefficients in the EAR Model

The coefficients of the EAR model (Sec. 2.5.2, Eq. 2.33):

$$P(x,y) = L_{r_0} - \beta \log_{10}\frac{r}{r_0} - \alpha_r(r - r_0) + \Delta_x(x - x_s) + \Delta_y(y - y_s) + \varepsilon,$$

are found using a constrained optimization/regression procedure (Sec. B.2). The coefficients are constrained or bounded so that the model gives SPL estimates that are physically realistic and is an attempt to deal with the overfitting issue inherent in the least-squares regression approach to solving for the EAR model coefficients (See Sec. 5.3.4 for more discussion on the topic of overfitting).

The constrained optimization procedure allows the user to specify starting parameter values and bounds for any/all of the coefficients in the model. For the 5 parameter version of the EAR model, the following constraints were used: L_{r_0} was within $\pm$ 10 dB of the level estimated at a reference distance from the source (r_0), and β was within $\pm$ 25 of 35 dB. The details and justification for the choice of the starting parameter and the bounds of each coefficient is given below. It should be noted that these parameters are specifically designed for the source/problem addressed in this dissertation. Future use of the EAR model should tailor the choice of constraints to reflect the source/problem at hand (e.g., α_r should be changed if a broadband source is considered, and β might be changed if the source is well approximated as a point source or if the problem is set in a different propagation medium).

B.1.1 Starting Parameters & Bounds for L_{r_0}

The following process was used to find the starting parameter for L_{r_0}: First, the noise monitor location closest to the reference distance from the source ($r_0 = 2$ km), was identified. Then, the SPL from that noise monitor was corrected using Eq. (2.35):

$$L_r = L_{r_0} - 20\log_{10}\frac{r}{r_0},$$

which gives the SPL relative to the SPL at the reference location (L_{r_0}) and includes the effects of spherical spreading. This is done so that the starting parameter better reflects SPLs at the reference distance, and is necessary given that the identified monitor will likely be located at some distance from the source other than r_0. For example, if the closest monitor to the reference distance is located 3 km from the source, and the SPL for this particular noise event at the monitor is 120 dB, then using Eq. (2.35) the approximate level at r_0 (where $r_0 = 2$ km) is 123.5 dB.

After the starting parameter has been specified, the next step in the procedure is to specify the bounds on L_{r_0}. In this study, an *ad hoc* analysis was done that looked at the corrected SPL (Eq. 2.35) from all CNPE output points located within 500 m of r_0. This was done for each of the 9 meteorological conditions considered in this study and the results are given in Table B-1. This table shows the estimated value of the minimum, mean, maximum, and range of estimated values of L_{r_0}, where the range is the difference between the maximum and minimum estimated value. This table also shows the actual value of L_{r_0}, which is the mean value of all SPLs located at the reference distance for each of the 9 meteorological conditions. From this analysis, bounds of $\pm$ 10 dB of the estimated level at the reference distance were chosen to encompass the maximum range in Table B-1 (i.e., 18 dB).

Table B-1. The minimum (min), mean, maximum (max), and range of estimated values of L_{r_0} for the *ad hoc* analysis performed to identify bounds on L_{r_0}. The table also shows the actual value of L_{r_0}, which is the mean value of all SPL located at the reference distance of 2 km for each of the 9 meteorological conditions. All values have units of dB.

Met Condition	Estimated L_{r_0}(min)	Estimated L_{r_0}(mean)	Estimated L_{r_0}(max)	Estimated L_{r_0}(range)	Actual L_{r_0}
SC-CW	92	95	97	5	97
SC-DW	99	100	101	2	98
SC-UW	90	91	92	2	96
WC-CW	95	100	103	7	98
WC-DW	104	105	105	1	103
WC-UW	90	92	93	2	99
WS-CW	89	101	107	18	86
WS-DW	107	107	108	1	107
WS-UW	78	80	82	3	89

B.1.2 Starting Parameters & Bounds for β

The starting parameters and bounds for β are found from an *ad hoc* analysis that examined at the geometrical decay in 1D between 1 km and 10 km for CNPE runs that included 7 different ground types (freshly fallen snow, forest floor, grassland, dirt road, asphalt road, painted concrete), 3 different wind directions relative to the propagation direction (downwind, crosswind, upwind), and 6 different stability classes (strongly convective, moderately convective, weakly convective, neutral, weakly stable, moderately stable). The value of β for each condition (126 in total) was calculated as the difference between the SPL output from the CNPE at 10 km and 1 km from the source and is shown in Figure B-1. From this analysis, the mean of 35 dB was chosen as the starting value, and bounds of $\pm$ 25 dB of 35 dB were chosen to encompass the full range of β values.

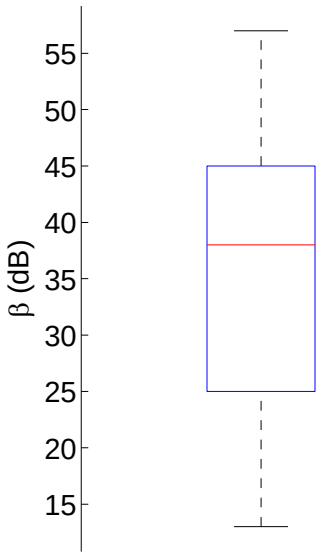

Figure B-1. Distribution of β values (dB) calculated as the difference between the SPL output from the CNPE at 10 km and 1 km for each combination of the following propagation conditions: 7 different ground types (freshly fallen snow, forest floor, grassland, dirt road, asphalt road, painted concrete), 3 different wind directions relative to the propagation direction (downwind, crosswind, upwind), and 6 different stability classes (strongly convective, moderately convective, weakly convective, neutral, weakly stable, moderately stable).

B.2 Constrained/Bounded Optimization

In this dissertation, a constrained optimization algorithm is used to solve for the coefficients in the EAR model and the power semivariogram model. Use of this type of algorithm allows one to bound the coefficients of a regression procedure. For example, recall that in order to adhere to the properties for a valid semivariogram, the power semivariogram model (Eq. 2.32) must have coefficients that are bounded (i.e., $b \geq 0$ and $0 \leq \rho < 2$), to ensure that kriging system of equations converges. In the case of the EAR model, it is of interest to bound the coefficients in the model so that the SPL estimates from the model are realistic, and is also done in attempt to keep the model from overfitting the input data points.

The constrained optimization algorithm used in this study was implemented using the *fit* function in the MATLAB Curve Fitting Toolbox.[81] This function uses a trust-region reflective algorithm,[98-101] which requires a scalar function/equation and a starting parameter, a lower bound, and upper bound for each coefficient in the equation. The trust-region reflective algorithm is a minimization optimization algorithm that finds a vector of coefficients that are a local minimum to the input scalar function $f(\mathrm{v})$. It is an iterative algorithm that starts at the point specified by the starting parameters (v in a n-dimensional parameter space). Then, the algorithm takes a step (s) within the trust region ellipsoid. If the new point $\mathrm{v} + \mathrm{s}$ better approximates the input scalar function, then the starting point is replaced with the new point and the trust region is remade. If the point does not better approximate the input scalar function, then the initial point v remains and the trust region is shrunk. In both cases, the process is iterated until the model converges.[102]

B.3 Unconstrained vs. Constrained Latin Hypercube Sampling

In this section, an exploratory analysis is presented that compares the RMSE_{SPL} distributions using 50 noise monitor geometries obtained using the constrained LHOS (cLHOS) procedure (Figure B-2 left) to 50 noise monitor geometries obtained using no constraint or unconstrained LHOS procedure (Figure B-2, right). In both cases, the analysis presented in Sec. 3.5 was run on the AC5 model with data gathered from 4 noise monitors under the 9 meteorological conditions given in Sec. 3.2.3. These graphs show a significant difference between the two sampling procedures. The cLHOS procedure results in RMSE_{SPL} values that are much lower and have less variance (dispersion) than the LHOS procedure, for all the meteorological conditions. Also, notice that the notches around the median values between the two distributions do not overlap. As explained in Sec. B.5, this means that the medians of the two distributions are significantly different at the 5% significance level, and provides evidence that the cLHOS is superior to the unconstrained LHOS sampling scheme. It can also be concluded, in general, that the choice of the noise-monitor locations is an important consideration for the problem addressed in this dissertation.

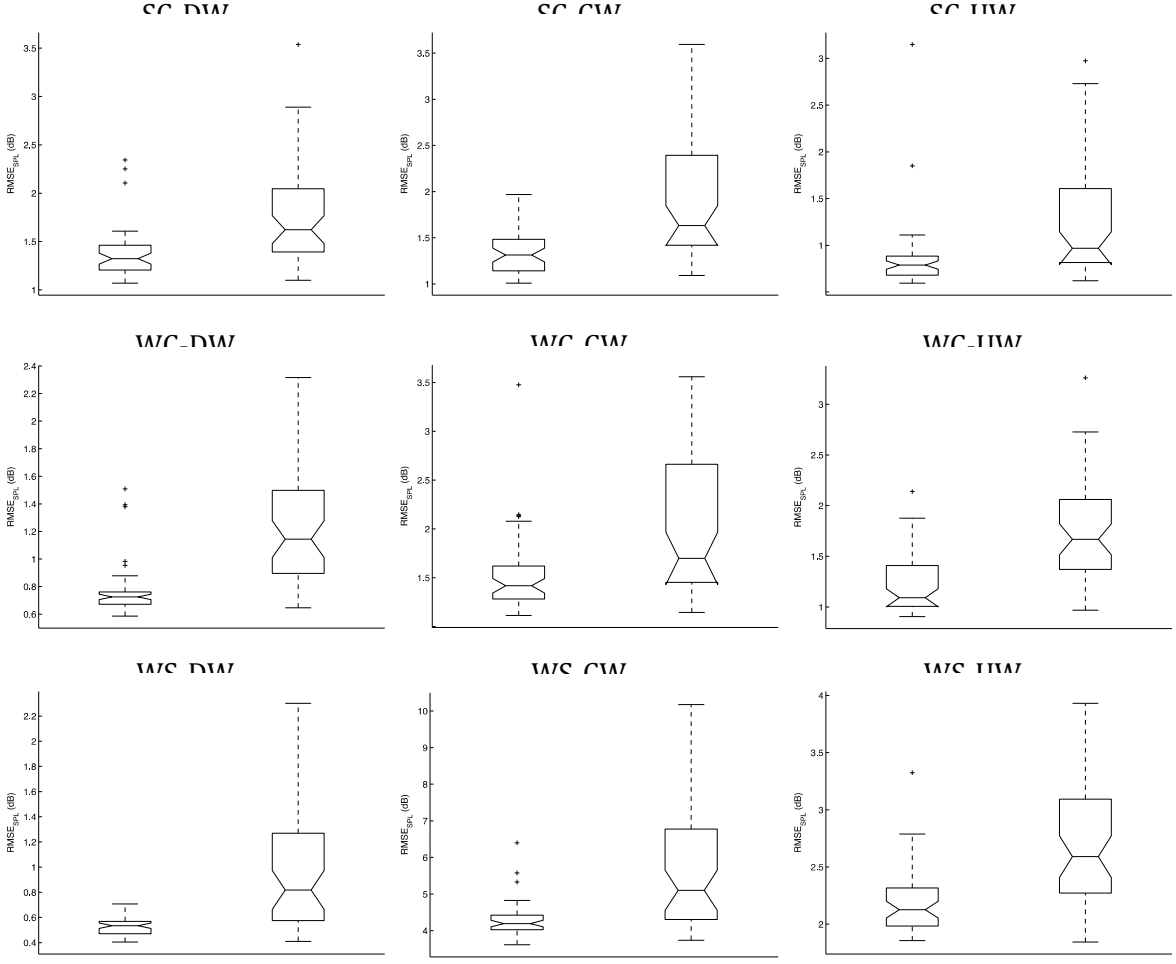

Figure B-2. Distributions of $\mathrm{RMSE_{SPL}}$ (dB) for noise monitor geometries that were obtained using the cLHOS procedure (left box plot) and using the unconstrained LHOS procedure (right plot). The $\mathrm{RMSE_{SPL}}$ values were obtained from analyses run on the AC5 model with data gathered from 4 noise monitors under the 9 meteorological conditions given in Sec. 3.2.3.

B.4 50 vs. 1000 Noise Monitor Geometries

This section addresses the question of how many noise monitor geometries are sufficient to obtain the distribution of accuracy measures. An *ad hoc* analysis was performed to compare the distributions of the OAI using 50 noise monitor geometries vs. 1000 noise monitor geometries (Figure B-3). This analysis was done using the most accurate model, as determined in Sec. 4.1, when $N = 3, 4, 9, 16,$ and 25. With the exception of a few additional outliers in the case of 1000

geometries, Figure B-3 shows that the two distributions are nearly identical. Also, notice that the notches around the median values overlap. As explained in Sec. B.5, this means that medians of the two distributions are *not* significantly different at the 5% significance level. It is therefore concluded that the use of only 50 noise monitor geometries is sufficient to capture the full distribution of results.

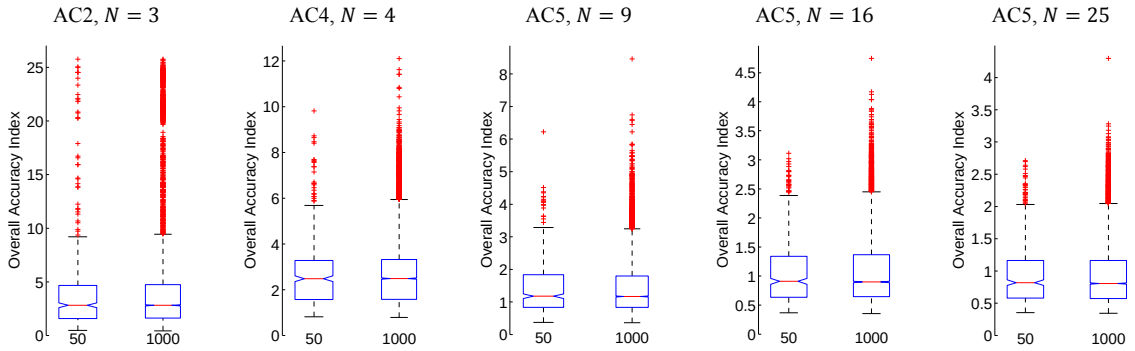

Figure B-3. Comparison of the OAI distributions when there are 50 and 1000 noise monitor geometries for the most accurate model when $N = 3, 4, 9, 16,$ and 25.

B.5 Box Plots

An example of a box plot is given in Figure B-4. The red line is the median and represents the central tendency of the data or the 50th percentile of the data. The lower blue edge of the box is the first quartile (Q1) and represents the 25th percentile of the data. The upper blue edge of the box is the third quartile (Q3) and represents the 75th percentile of the data. A red plus represents a data point that is distant from the rest of the data. The dashed black lines extending from the blue box are the whiskers. For the box plots used in this dissertation the whisker length (w) is set to 1.5 and the whiskers extend from the first and third quartile by a distance of ($w *$ IQR), where IQR is the interquartile range, which is the distance between the Q1 and Q3. If the

data are normally distributed then a whisker length of 1.5 corresponds to approximately $\pm 2.7\sigma$, and covers 99.3% of the data points.[103]

The notches around the median values, which are present on some of the box plots presented in this dissertation, are used to compare the median values of two different distributions. If the notches do not overlap, the medians of the two distributions are significantly different at the 5% significance level. The notches extend to a distance of Q2 $\pm \frac{1.57*\text{IQR}}{\sqrt{n}}$, where n is the number of data points.

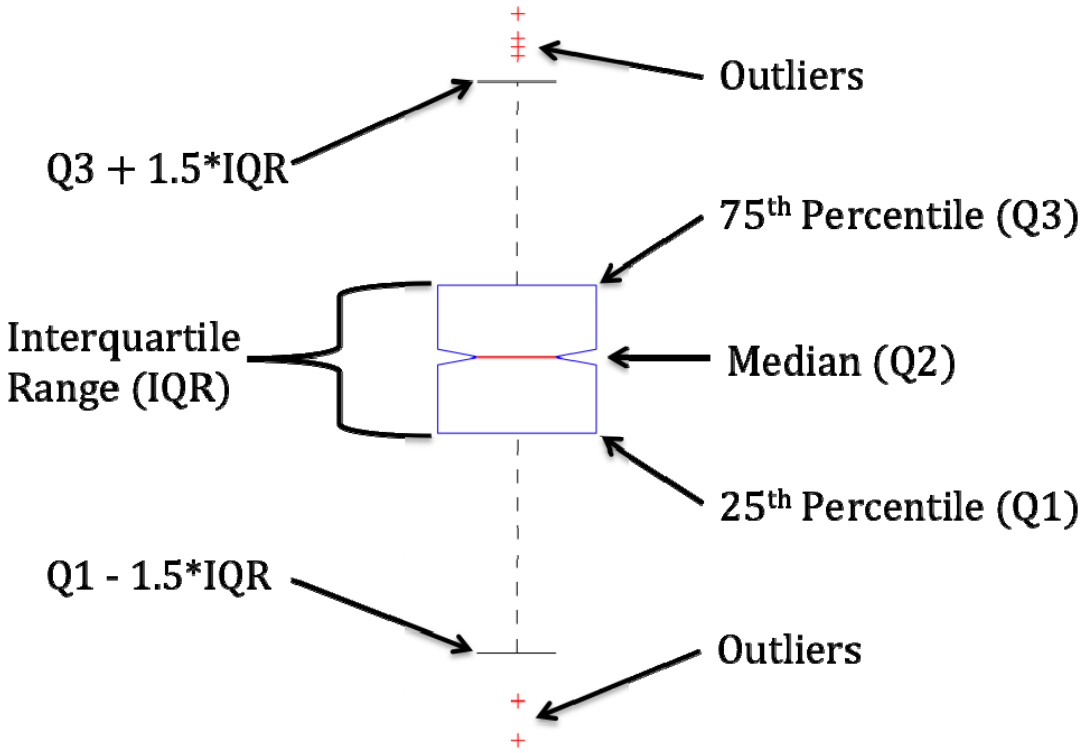

Figure B-4. Overview of the box plot, showing the first quartile (Q1, lower edge of the blue box), second quartile (Q2, median, and red line inside the blue box), third quartile (Q3, upper edge of the blue box), interquartile range (IQR, which is Q3-Q1), whiskers (dashed black lines), and points that are distant from the rest of the data (outlier, red plus).

B.6 Theoretical Semivariogram Functions

In the early stages of this research, an exploratory analysis was done to determine which theoretical semivariogram function (TSF) would be the most appropriate to use in this study. This analysis was necessary given that 1) there are many TSFs that meet the mathematical requirements as described in Sec. (2.4.2), 2) there is no historical precedent for selecting an appropriate model for the problem this dissertation addresses, and 3) in practice, the semivariogram model selection is often based on experience and the fitting procedure is sometimes done subjectively.[54]

A total of 14 common TSFs (Table B-2)[45,52] were examined in this analysis, with and without the nugget term (2.4.1, Figure 2-10). The TSFs were fit to empirical semivariogram (ESV) cloud data obtained using the kriging models described in Sec. (3.1) with the noise monitor and ground truth data described in Sec. (3.3), and the fit of the TSF to the data was assessed using Akaike's information criterion (AIC).[72] AIC is a model section procedure that takes into account the error in the fitting procedure and the number of parameters in the fitting model. In AIC models that have a more parameters are penalized, with the idea of finding the model that fits the data the best using the fewest number of parameters.

Table B-2. Theoretical semivariogram functions considered in this dissertation. The equations for each model are given in Chilès and Delfiner,[52] and Cressie.[45]

TSF	General Class	Notes
Power	Power	Unbounded, also referred to as Kolmogorov model
Linear	Power	Power model with $\rho = 1$
Stable	Stable	Bounded, 2nd best model
Exponential	Stable	Stable model with $\rho = 1$
Gaussian	Stable	Stable model with $\rho = 2$
Spherical	Spherical	Bounded
Cubic	Spherical	Used in Baume et al.[38]
Pentamodel	Spherical	
Pentaspherical	Spherical	
Rational Quadratic		
Wave of Hole-Effect		
Log		
Generalized Cauchy		
Matern		

The results from this exploratory analysis showed that in the majority of scenarios (i.e., for the kriging models, propagation conditions, and sensor densities/geometries considered), the power semivariogram model (PSM, Eq. 2.32) was the best model. This is perhaps not surprising given that this model only has two parameters and is often used to model atmospheric turbulence.[48,73,74] In addition, many of the models resulted in very similar fits, which is also not surprising given the underlying form of the models are very similar. For example, the Gaussian TSF and Exponential TSF are exactly equal to the Stable TSF:

$$\gamma(\mathrm{h}) = \begin{cases} 0, & \mathrm{h} = 0, \\ c_s\left(1 - e^{-(\mathrm{h}/c_r)^{\rho}}\right), & \mathrm{h} \neq 0 \end{cases} \tag{B.1}$$

with $\rho = 1$ and $\rho = 2$, respectively. In Eq. (B.1), $c_s \geq 0$, $c_r > 0$ and $0 \leq \rho \leq 2$ as required to obtain a valid semivariogram model (Sec. 2.4.2).

For the sake of consistency, all of the kriging models in this dissertation were fit with the PSM, which unlike most of the other models, can model both bounded and unbounded

semivariograms (Figure 2-10, Figure 2-11). Many of the other TSFs in Table B-2 did not converge to a solution, which may gave been due to the semivariogram being unbounded. Another point, and as noted in Sec. (2.4.2), objectively fitting a TSF from an ESV is still an active area of research.[67-70] It is possible that some of the other TSFs in Table B-2 may perform better than the PSM for specific noise monitor configurations and propagation conditions, and should be considered in future research efforts.

B.7 Estimating the "Actual" Value of σ

The method used to obtain the "actual" value of σ for a residual kriging model is given below. Here, "actual" is emphasized because the value is not the actual value but rather is estimated using many more data points than that obtained by using only the noise monitor locations. That is, in an asymptotic or large-sample sense this method is likely to approach the actual or true value of the variance. For example, 1024 data points are used to estimate the "actual" variance values, whereas the number of noise monitor data points used in this study ranged from $3 - 49$. The steps below outline the process of estimating the "true" variance when a RKG is used. The same formulation can be done for an ordinary kriging model (OKG) as well, except that the process starts at Step 3 and the actual SPL levels at the evaluation points are used in place of the residual data obtained in Step 2.

1. Estimate the SPL at the evaluation points using the noise monitor data and the trend model associated with the residual kriging model (e.g., an EAR model or a polynomial surface fit model).
2. Calculate the residual error between the SPL estimate obtained with the trend model and the actual SPL at the evaluation.
3. Define the "true" semivariogram function using the residuals calculated in Step 2.

4. Estimate the “actual” variance at each evaluation location using the RKG system of equations and the “true” semivariogram calculated in Step 3.

B.7.1 Estimate the SPL at the evaluation points

The first step in estimating the “true” variance is to estimate the SPL at the evaluation points with the trend model using the data at the noise monitor locations. Note: this formulation is different than the steps outline in Sec. 2.2.2 in that the SPL estimates are made at the evaluation points in this case and are made at the noise monitor locations in Sec. 2.2.2.

B.7.2 Calculate the residual error at the evaluation points

The next step is to calculate the residual error at the evaluation points. In order to understand why this is done, recall that the goal of RKG and the geo-acoustic models is to break the estimate into a deterministic component and a stochastic component. In this sense, everything that is not modeled with the deterministic component is the error, which is modeled as the stochastic component. This can be done (or at least approximated to a high degree of certainty) if the actual SPL levels at every point in the ROI are known, which is the case in this study. This means that the residuals–the difference between the estimated values with the deterministic model and the actual values–contain the stochastic component that the RKG model is trying to model, but in the context of the problem addressed in this dissertation, with only a few data points (i.e., a small number of noise monitors).

B.7.3 Define the "true" semivariogram

The next step in the process is to use the residuals from Step 2 to calculate "true" semivariogram. The process of estimating the empirical semivariogram cloud and fitting a theoretical semivariogram model to the data is described in Sec. 2.2, and not replicated here. However, it is worth restating that it is very difficult to model the semivariogram when there are very few data points, and in practice there are often regions in the ESV that have little or no coverage. The lack of spatial coverage (in terms of the lag distance) is illustrated below (Figure B-5). Figure B-5 (left) shows the empirical semivariogram cloud resulting from the 9 noise monitor data points using the example data points given in Sec. 2.2, and Figure B-5 (right) shows the empirical semivariogram cloud resulting from the 1024 data points from Step 2 above.

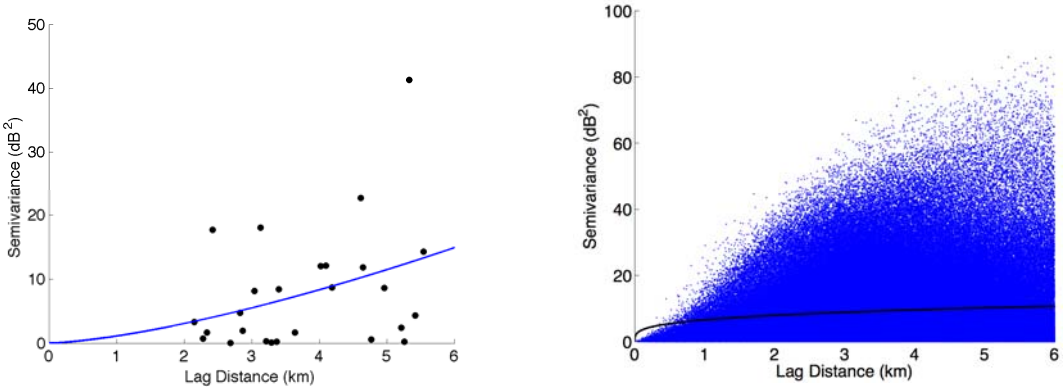

Figure B-5. The theoretical semivariogram fit (line) to the ESV cloud data (circles). The left figure is showing the semivariogram estimated from the data from 9 noise monitor locations. The right figure is showing the "true" semivariogram estimated from the data from the 1024 evaluation points.

B.7.4 Estimate the "true" variance

The variance at each evaluation point is calculated using the RKG steps outlined in Sec. 2.2, with exception that the semivariogram fit from Step 3 (Figure B-5 right) in used in place of the semivariogram that would normally be used (i.e., the semivariogram given in Figure B-5 left).

B.8 Fitting Semivariogram with Sparse Datasets

On one hand, estimating the variance from data is an attractive feature of the kriging procedure; on the other hand and especially in the case of sparse datasets, estimating the variance from data is difficult to calculate, interpret, and model.[66,104,105] The problem with sparse datasets in the context of geostatistics is that they do not contain enough information to infer the spatial variability.[67]

B.8.1 Sparse Datasets

For spatial interpolation, the term "sparse" dataset is somewhat subjective. Some authors refer to sparse datasets when there are less than 30 data points, whereas others refer to sparse as less than 100 data points.[67] For this dissertation, the number of noise monitors (N) considered (data points) range from $3 - 49$. With exception of $N \geq 36$, the number of data points fall within both definitions of sparse datasets. There have also been many suggestions in the literature regarding how many data points are needed to obtain a reliable semivariogram. Journal and Huigjbruits suggest that it is necessary to have at least 30 lag distance points and Chilès and Delfiner suggest that at least 50 are needed.[72] Using Eq. (2.31), this means that at least 8-10 data points are needed to obtain a reliable semivariogram.

B.8.2 Lag Distance Cutoff

As mention in Sec. 4.1, there is a sharp jump in the median OAI for the ordinary kriging model (OKG) as N goes from 8 to 9 (Figure 4-1). The reason for this jump is due to a change in the number of lag distance points that are used to fit the theoretical semivariogram that occurs for $N \geq 9$ as discussed in Chapter 2 (Sec. 2.2.5). When $N < 9$ all the lag distances are used to fit the theoretical semivariogram model given that there are not many lag points (recall Eq. 2.31). Whereas, for $N \geq 9$ only lag distances that are half the maximum distance of the ROI are used, as suggested by Chilès and Delfiner.[52] This is done to cut down on computation time (which is important for large datasets) and because it is the lag distance points closest to the origin (i.e., the noise monitor locations closest to the estimation point location) that receive the most weight in the kriging system of equations.

In order to see how the lag distance cutoff affects model accuracy, the evaluation procedure (Sec. 3.5) was re-run with the AC2, AC5, and the OKG models using all the lag distance points for $N \geq 9$. Figure B-6 shows the results, which compare the median OAI when all the lag distance points are used to fit the theoretical semivariogram (blue markers) to the case when the lag points that are less than half the maximum lag distance in the ROI are used (black markers). In the case of OKG, the mean OAI is significantly improved for $N \geq 9$ when all the lag points are used in fitting the theoretical semivariogram. Basically, the lag points that lie beyond the half the maximum distance in the ROI change the shape of the theoretical semivariogram that is used in the kriging procedure (Figure B-7). In Figure B-7, the fit of the semivariogram that uses all the points is closer to the true semivariogram model near the origin than the fit that only uses a portion of the lag points. Though, it should be noted that neither of the fits comes close to the true semivariogram model.

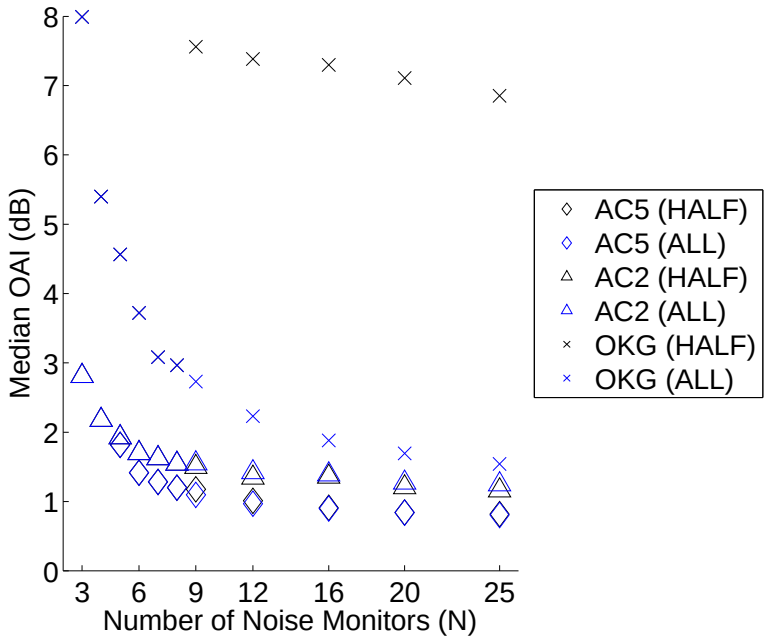

Figure B-6. Median OAI (dB) as a function of the number of noise monitors (N) across all meteorological conditions for the AC2, AC5, and OKG models. The blue markers show the results when all the lag distance points are used to fit the theoretical semivariogram, and the black markers show the results when only lag points that are less than half the maximum lag distance in the ROI are used to fit the theoretical semivariogram.

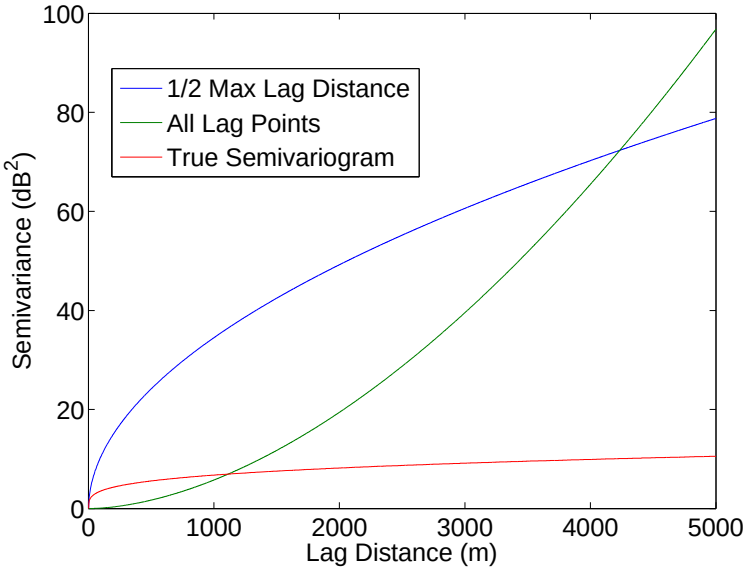

Figure B-7. Theoretical semivariogram fit using all the lag points that are less than half the maximum lag distance in the ROI (blue), using all the lag points (green), and using the procedure outlined in Sec. B.7 to obtain the "true" semivariogram (red).

The reason why the overall accuracy is improved with the change in semivariogram procedure is because the variance estimates made at each estimation point is dependent upon the distance between the estimation point and the noise monitor locations. The distance between these points is used to select the semivariogram value from the theoretical semivariogram function, which in turn is used in the kriging system of equations. As shown in Table B-3, more weight is given to noise monitor locations that are closest to the estimation point. In the context of Figure B-7, this essentially means that values of the semivariogram beyond a certain point (e.g., h $>$ 2 km) will have little influence on the SPL and variance estimates. In fact, examination of the difference between the two fits between 0 and 2 km reveals that between 0 and 1 km the semi-variance is overestimated, whereas between 1 and 2 km the semi-variance is underestimated, and thus, on average we would expect that the OAI would be better in this case in comparison to the case of blue line fit, which consistently over estimates the semi-variance.

Table B-3. The weights (w_i) obtained from solving the kriging system of equations, and the distance between the estimation point (x_0) and the noise monitor location (x_i).

Noise Monitor	Distance (km) $\|\mathbf{x}_i - \mathbf{x}_0\|$	Weight (w_i)
$\mathbf{x}_1$	1.0	0.6
$\mathbf{x}_2$	2.0	0.3
$\mathbf{x}_3$	3.6	0.0
$\mathbf{x}_4$	5.5	0.0
$\mathbf{x}_5$	6.1	0.1
$\mathbf{x}_6$	6.2	0.0
$\mathbf{x}_7$	7.6	0.0
$\mathbf{x}_8$	7.9	0.0
$\mathbf{x}_9$	9.7	0.0

C Outdoor Sound

The purpose of this appendix is to discuss the physical mechanisms that influence the way sound propagates outdoors, which in turn affects received sound pressure levels (SPLs). Particular attention is drawn to the variability of these mechanisms and how they drive the variance in received SPLs. Examples of the mechanisms that affect outdoor sound propagation are illustrated over the frequency and distance ranges relevant to this research. The information presented in this appendix support the understanding the Crank Nicholson Parabolic Equation (CNPE) simulations in Sec. 3.2, and the environmental acoustic-propagation-based regression (EAR) models in Sec. 2.5.2.

The discussion progresses from basic to more complex concepts and largely mirrors the presentation given in Salomons.[25] First, geometric spreading, atmospheric absorption, and ground surface interactions are discussed. Then the atmospheric boundary layer (ABL) is introduced, followed by a discussion of refraction from mean vertical profiles and scattering from turbulent fluctuations within the ABL.

There are numerous articles and books written on the subject of outdoor sound propagation. The reader interested in a good overview of outdoor sound propagation can refer to Chapter 15 in the Handbook of Environmental Fluid Dynamics;[22] and those interested in a more comprehensive review can refer to Embleton,[106] Attenborough et al.,[107] and Salomons.[25] Those interested in basic meteorology can refer to Stull;[88] those interested in the atmospheric turbulence can refer to Wyngaard;[108] and those interested in the atmospheric boundary layer can refer to Garratt.[109]

In many introductory acoustic textbooks, the concept of sound propagation often begins with the discussion of plane waves and other basic concepts such as the pressure amplitude, frequency, phase, wave number, sound speed, etc. In this appendix the discussion of sound propagation starts with spherical wave propagation, and it is assumed that the reader is familiar with basic acoustic concepts. The reader interested in the fundamentals of acoustics can consult Kinsler et al.;[110] and the reader interested in a more rigorous treatment of the fundamentals can consult Pierce.[111]

C.1 Point Source in Idealized Media

This section discusses the way sound propagates outdoors and starts with the assumption that the medium is a homogeneous, lossless free space and discusses the effects of geometric spreading from a point source. Next, the homogeneous medium is specified to be air, and losses due to atmospheric absorption of sound are discussed. Lastly, the free space medium is bounded to include a flat homogeneous porous ground, and the effects of different ground surface types are discussed.

C.1.1 Geometric Spreading

The source of interest for this dissertation is a 1.25 lbs. (0.57 kg) C4 explosion. It is modeled as a point source because the size of the source is small compared to the wavelengths of the sound emitted by the source. The time-independent point source solution to the acoustic wave equation in a homogeneous lossless free space medium has the following form:

$$p(r) = p_0 r_0 \frac{e^{ikr}}{r}, \tag{C.1}$$

which describes spherical wave propagation. Here p is the complex pressure, p_0 is the complex pressure amplitude at the reference distance $r_0 = 1$ m from the source, k is the acoustic wave number, and r is the distance from the source to the receiver.[22] From this equation, it can be seen that the sound pressure amplitude decays as $1/r$. The decay or attenuation of the sound pressure in all radial directions from the source is known as geometric attenuation,[25] or spherical spreading law.[111]

In practice, the sound pressure amplitude is given in terms of the sound pressure level (SPL):

$$L_P = 10 \log_{10} \frac{|p|^2}{{p_{\text{ref}}}^2}. \tag{C.2}$$

Here, L_P is the SPL in decibels (dB), and p_{ref} is the reference pressure, which is 20 μPa in air. The SPL is often given in terms of the decay relative to the source sound power level:

$$L_P = L_W - 10 \log_{10} 4\pi \left(\frac{r}{r_0}\right)^2, \tag{C.3}$$

where L_W is the sound power level,[25] or more generally in terms of any reference sound pressure level:

$$L_r = L_{r_0} - 20 \log_{10} \frac{r}{r_0}. \tag{C.4}$$

Here, L_r is the SPL at a distance r from the source, and L_{r_0} is a reference SPL at a distance r_0 from the source.[2] The $1/r$ decay in sound pressure amplitude from the effects of spherical spreading, amounts to an approximate loss in SPL of 6 dB per doubling of distance and a 20 dB loss per decade, which is easily visualized by setting $r_0 = 1$ m, $r = 2$ m, and $r_0 = 1$ m, $r = 10$ m in Eq. (C.4). An example of spherical spreading loss is given in Figure C-1. An example of

spherical spreading over the range of $1 - 10$ km, where $L_{r_0} = 100$ dB and $r_0 = 1$km for a range of distances from the source.

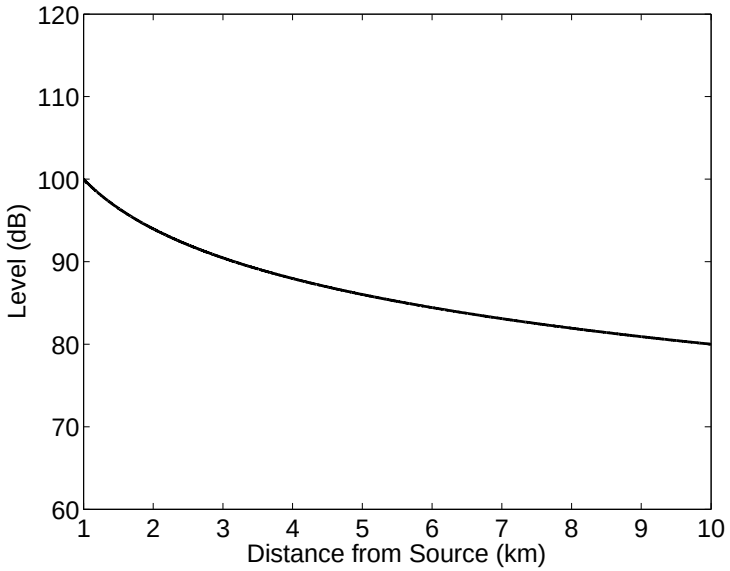

Figure C-1. An example of spherical spreading over the range of $1 - 10$ km, where $L_{r_0} = 100$ dB and $r_0 = 1$km.

C.1.2 Atmospheric Absorption

If the homogenous free space medium discussed above is specified to be air, then sound waves will lose some energy, beyond geometric spreading, to atmospheric absorption. Atmospheric absorption of sound energy is due to thermal conduction and viscosity, and molecular relaxation, which are dependent upon frequency, temperature, humidity, and atmospheric pressure.[25,111] The dominant absorption mechanism in air for low frequencies (i.e., $25 - 100$ Hz) is due to the molecular relaxation of diatomic nitrogen molecules.[111] As sound waves propagate they compress and expand the air they are propagating through. In the compression stage they excite the vibrational and rotational degrees of freedom in diatomic nitrogen and oxygen molecules, and in the expansion stage the molecules relax to their original

state. During this process, some energy is converted to heat resulting in a loss of sound energy and a phase delay.[25]

Functionally, losses due to atmospheric absorption result in an exponential decay of sound pressure amplitude or a linear decay of SPL with distance from the source. Equation (C.3) and Eq. (C.4) are updated to include the effects of atmospheric absorption:[25]

$$L_P = L_W - 10\log_{10} 4\pi \left(\frac{r}{r_0}\right)^2 - \alpha r, \tag{C.5}$$

$$L_r = L_{r_0} - 20\log_{10} \frac{r}{r_0} - \alpha(r - r_0). \tag{C.6}$$

In both equations, α is the atmospheric absorption coefficient with units of dB per unit length.

The value of α is dependent upon frequency. For the frequencies of interest to this dissertation (i.e., $25 - 100$ Hz), as shown in source spectrum given by the solid line in Figure C-2, the values of α range from 0.05 to 0.53 dB km^{-1} for 25 Hz and 100 Hz respectively (Figure C-3).

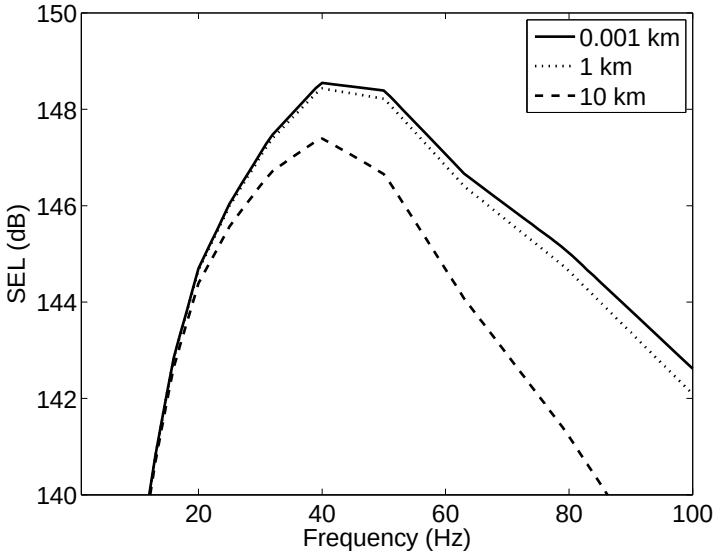

Figure C-2. Source spectrum at 1 m, 1 km, and 10 km from the source, highlighting the effects of atmospheric absorption alone (i.e., without the effect of geometric spreading) for a temperature of 20℃, relative humidity of 20%, and atmospheric pressure of 1 atm.

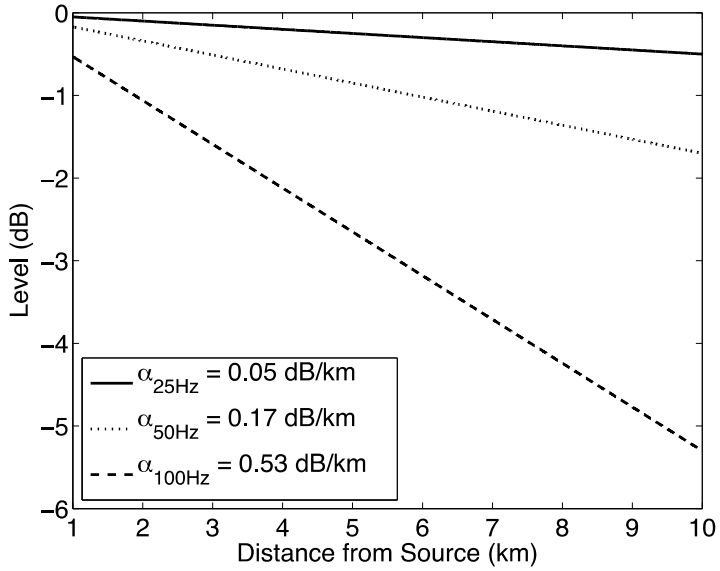

Figure C-3. Change in level (dB) due to atmospheric absorption alone (i.e., without the effect of geometric spreading) as a function of distance (km) from the source at $25, 50,$ and 100 Hz for a temperature of $20℃$, relative humidity of 20%, and atmospheric pressure of 1 atm.

The values of α shown in these figures were calculated using International Standard ISO 9613:1993(E)[112] with values of the temperature, relative humidity (RH), and atmospheric pressure that are used in the CNPE simulations presented in Chapter 3. Figure C-3 shows an example of the reduction of source SPL between $1 - 10$ km due to the effect of atmospheric absorption alone (i.e., without the effect of geometric spreading or any other phenomenon). Atmospheric absorption has little effect on very low frequency propagation. Figure C-4 shows an example of the combined effects of spherical spreading and atmospheric absorption applied to the source spectrum used in the CNPE simulations and given in Figure C-2.

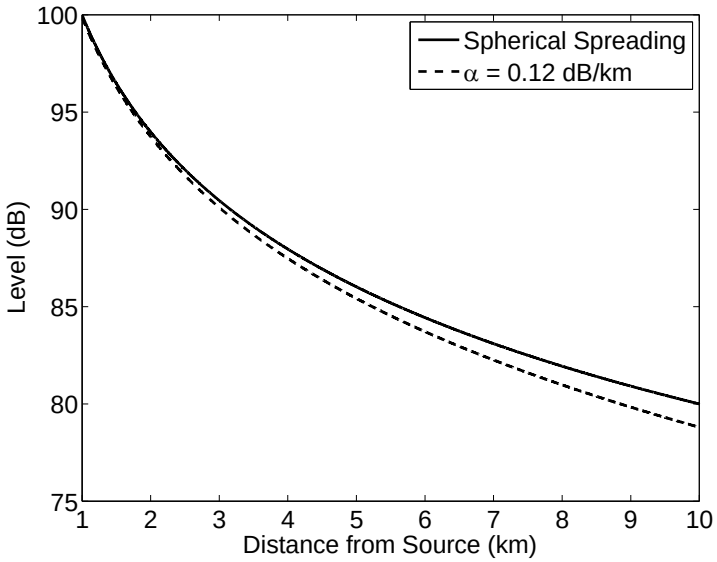

Figure C-4. An example of spherical spreading and atmospheric absorption of sound over the range of interest ($1 - 10$ km), where $L_{r_0} = 100$ dB, $r_0 = 1$ km, and $\alpha = 0.12$ dB km^{-1}. The value of atmospheric absorption was calculated for 40 Hz, 20℃, 20% RH, and atmospheric pressure of 1 atm.

The value of α also depends on temperature, RH, and atmospheric pressure. Figure C-5 shows the dependence of atmospheric absorption on temperature over a range of -20 to 40℃ and RH range of $5 - 100\%$ at 50 Hz. Here, values of α are highest when RH is low. In general, the values of α are also highest when the temperature is low. The exception to this statement, which is not shown in Figure C-5, is a local maximum of α at a RH between $0 - 5\%$ and at a high temperature.

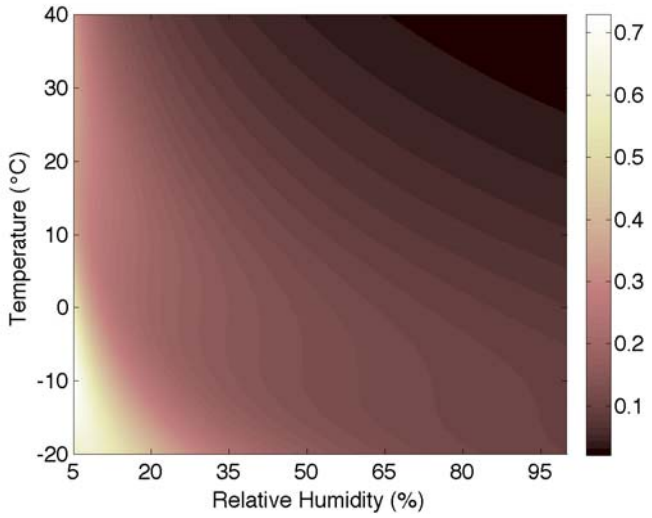

Figure C-5. Values of the atmospheric absorption coefficient for sound waves at 50 Hz and an atmospheric pressure of 1 atm over the temperature range of -20 to 40°C and relative humidity range of $5-100\%$. The color scale has dimensional units of dB km^{-1}.

Given that the RH for most areas within the continental United States rarely go below 5%,[113] RHs below 5% are ignored. Table C-1 shows the range of α values as a function of the frequencies, temperatures, and RHs values. At a frequency of 50 Hz, the values of α range between 0.021 and 0.729 dB km^{-1}, and it is concluded that the attenuation due to atmospheric absorption is minor (for the frequencies of interest to this dissertation) relative to geometric spreading and the other mechanisms that affect outdoor sound propagation (i.e., ground interactions, atmospheric refraction, and turbulent scattering.).

Table C-1. Distribution of α values at 25, 50, and 100 Hz and an atmospheric pressure of 1 atm over the temperature range of -20 to 40°C and relative humidity range of $5-100\%$.

	25 Hz	50 Hz	100 Hz
Minimum	0.005	0.021	0.086
Quartile 1	0.017	0.065	0.237
Median	0.031	0.105	0.299
Quartile 3	0.051	0.155	0.428
Maximum	0.284	0.729	1.731

C.1.3 Ground Surface

As sound waves propagate they interact with any object, or differing media encountered. An important interface for this research is the ground, given that the source and receiver(s) are located near the ground. It is common practice to model the ground as a porous medium. A porous ground means that the layer of the media above the ground is able to penetrate and fill the pores in the ground. The porous ground is treated as an effective fluid with a complex density and sound speed.

For spherical wave propagation over a flat homogeneous porous ground, Eq. (C.1) is modified to include interaction with the ground:

$$p(r) = p_0 r_0 \left(\frac{e^{ikr}}{r} + Q \frac{e^{ikr_i}}{r_i} \right) \tag{C.7}$$

where Q is the spherical wave reflection coefficient, and r_i is the distance from an image source to the receiver. The virtual image source is a convenient way to account for the interaction with a flat boundary. In a rz coordinate system the image source is located at the same r-coordinate ($r = 0$) and the negative of the z-coordinate of the real source. The spherical wave reflection coefficient Q is similar to the plane wave reflection coefficient in that $Q = 1$ is characteristic of a infinitely hard ground,[106] but is different because it includes a boundary loss factor to account for the spherical wave interaction with a complex impedance surface.[22,107] In addition to the boundary loss factor, Q is a function of frequency, distance from the source, angle of incidence, and ground impedance.[22,25,107]

In practice, it is common to specify the ground impedance in terms of flow resistivity,[87,114] though there are other models (e.g., Attenborough's 4 parameter model).[115] The flow resistivity is a measure of how easily the air moves through the pores in the ground. For this dissertation, the Komatsu model[87] is used, which is very similar to the Delany-Bazley model,[114]

except it includes different coefficients that improve the fit to measured data at lower frequencies. Values of flow resistivity for outdoor ground types typically range between 2 and 2×10^5 kPa s m^{-2}. A table of flow resistivities for various ground surfaces is given in Table C-2.[25,107,116]

Table C-2. Flow resistivities (kPa s m^{-2}) for various ground surfaces.

Ground Surface	Flow Resistivity (kPa s m^{-2})
Freshly Fallen Snow	2
Forest Floor	20
Grassland	200
Dirt Road	2×10^3
Asphalt Road	2×10^4
Painted Concrete	2×10^5

An example of the effect of ground type alone (i.e., without the effect of geometric spreading and atmospheric absorption) for a range of sources and distances is given in Figure C-6. A flow resistivity $\geq 2 \times 10^3$ kPa s m^{-2} results in an increase in SPL, which is expected given that the receiver height is close to a hard ground. On the other hand, a flow resistivity $\leq$ 200 kPa s m^{-2} results in a loss in the SPL as a function of distance from the source.

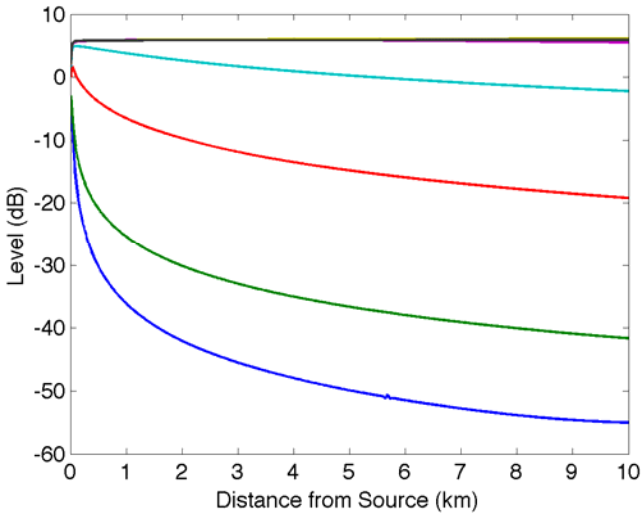

Figure C-6. Change in level (dB) due to ground effect only (i.e., without the effect of geometric spreading and atmospheric absorption) as a function of distance (km) from the C4 explosion source for a flow resistivity (kPa s m^{-2}) of 0.2 (blue line), 2 (green line), 20 (red line), 200 (cyan line), 2 x 10^3 (magenta line), 2 x 10^4 (yellow line), 2 x 10^5 (gray line).

C.2 Atmospheric Boundary Layer

Thus far, it has been assumed that the medium is homogeneous to demonstrate some of the basic mechanisms that affect outdoor sound propagation. In reality, this is not a valid assumption. The atmosphere is constantly changing and moving on various spatial scales (e.g., local, regional, global) and temporal scales (e.g., near-instantaneous, diurnal, seasonal). The part of the atmosphere where most of the motion occurs is the atmospheric boundary layer (ABL). In order to deal with the heterogeneous nature of the ABL, it is common practice to break the properties of the medium into average and fluctuating parts. The mean ABL properties are usually described in terms of vertical temperature and wind velocity profiles, which collectively affect the vertical sound speed profiles and result in the refraction of sound waves. Turbulent fluctuations in the ABL, on the other hand, result in a scattering of sound waves. This section

gives an overview of the ABL, and discusses the effects of the ABL on sound propagation (i.e., atmospheric refraction and turbulent scattering).

C.2.1 ABL Overview

The ABL is the layer of air near the Earth's surface that is influenced by the Earth's surface. That is, the layer that is affected by the heating and cooling of the Earth's surface and the friction forces of moving air over the Earth's surface.[109] The ABL is located within the troposphere, and the upper boundary of the ABL typically varies between a few hundred meters and a few kilometers. The ABL is unique from the rest of the atmosphere, in that it experiences a diurnal (daily) cycle of convective heating and cooling of the Earth's surface.[88] These cycles of heating and cooling are largely the cause of the weather we experience, the propagation conditions that affect sound propagation, and in the terms of the focus of this dissertation, the variance in received SPLs.

Though the movement of air within the ABL is dynamic and stochastic, most of the motions can be described in terms of the buoyancy forces on parcels or volumes of air. For example, if a parcel of air is warmer than its surrounding environment, the parcel will experience a positive buoyancy force and accelerate upwards. If the parcel of air is cooler than its surrounding environment, the parcel will experience a negative buoyancy force and sink.

The net buoyancy force and wind shear present in the ABL over a given region and time period define the stability of the atmosphere. The stability of the atmosphere controls the formation of the ABL, and gives an indication of whether turbulence will develop. The ABL in a given region is "unstable" if it experiences a net upward motion. Unstable ABL conditions are typical of clear sunny days when there is strong surface heating from solar radiation. During these conditions, parcels ("thermals") of warm air will accelerate upwards. The warm air then mixes

with the cooler air, resulting in large fluctuations in the temperature and wind velocity. These fluctuations are collectively known as turbulence. The ABL is "stable" if it experiences a net downward motion. Stable conditions typically occur during clear nights when the ground is cooler than the layer of air above the ground. During these conditions, turbulence is weak or non-existent. Lastly, the ABL is "neutral" if the net buoyance force is zero. These conditions are typical over overcast skies and when there is little or no heating and cooling at the surface.[88]

While the effect of convective heating and cooling of the ground is a good starting place to describe the motion of the ABL, there are other mechanisms that effect motion in the ABL. For example, large weather systems, such as regions of high and low pressure will modulate the ABL. Low pressure centers near the ground create an upward motion of air near the ground, which may lift the top layer of the ABL. High pressure centers, on the other hand, create downward motions that push and thin the ABL.[88] Other examples of large scale motions include orographic lifting, frontal wedging, and surface conduction.[117,118]

C.2.1.1 Obukhov Length & Similarity Relations

A useful way to characterize the stability of the ABL is with the Obukhov length:

$$L_o = \frac{-u_*^3\, T_0}{k_v g \langle w'T' \rangle_s}, \tag{C.8}$$

where u_* is the friction velocity – which represents the wind shear forcing on the near-ground atmosphere, T_0 is the temperature at the ground, k_v is the von Kármán constant, g is gravitational acceleration, w' is the fluctuating vertical component of the wind velocity, T' is the fluctuating component of the temperature, and $\langle \quad \rangle_s$ indicates an ensemble average at the surface. The Obukhov length (L_o) is useful in that it indicates the height where the production of turbulence transitions between wind shear and buoyance predominance.[90] The inverse of the Obukhov length

is also useful indicator. It gives the stability of the ABL. The ABL is unstable when $L_o^{-1} < 0\ \mathrm{m}^{-1}$, stable when $L_o^{-1} > 0\ \mathrm{m}^{-1}$, and neutral when $L_o^{-1} \approx 0\ \mathrm{m}^{-1}$.[25]

Assuming the ground is flat, similarity relations [88,89,119] can be used to give a realistic representation of the mean vertical temperature and wind velocity profiles for each stability class, where the mean is typically over a 10 minute time period. Similarity relations are dimensionless quantities that describe the temperature and wind velocity as a function of the Obukhov length and height above the ground. Figure C-7 shows an example of mean temperature and wind velocity profiles for unstable, neutral, and stable conditions. The temperature profiles used in this dissertation and given Figure C-7 (left) are calculated using Eq. (10) in Wilson 2003: [120]

$$T(z) = T_0 - \Gamma\left(z - z_{T_0}\right) + \frac{P_t T_*}{k_v}\left[\ln\frac{z}{z_{T_0}} - \psi_t\left(\frac{z}{L_0}\right) + \psi_t\left(\frac{z_{T_0}}{L_0}\right)\right],$$

where T_0 = 293.15 K at the reference height ($z_{T_0} = 0.01$ m), $\Gamma = 0.0098\ \mathrm{K\ m^{-1}}$, z is the height above the ground (m), $P_t = 0.095$ is the Prandtl number, $T_* = \frac{u_*^2 T_0}{L_0 \kappa g}$ comes from Salomons Eqs. (N.10 and Eq. N.11),[25] $k_v = 0.41$, and ψ_t comes from Salomons Eqs. (N.22 for convective conditions, N.26 and N.27 for stable conditions). The following parameters are used for each stability class: for strongly convective (SC, $L_o^{-1} = -0.3\ \mathrm{m^{-1}}$), weakly convective (WC, $L_o^{-1} = -0.0125\ \mathrm{m^{-1}}$) and weakly stable (WS, $L_o^{-1} = 0.0125\ \mathrm{m^{-1}}$). The wind speed profiles used in this dissertation and given Figure C-7 (right) are calculated using Eq. (N.16) in Salomons:

$$u(z) = \frac{u_*}{k_v}\left[\ln\frac{z}{z_0} - \psi_w\right],$$

where $u_* = 0.25\ \mathrm{m\ s^{-1}}$, $z_0 = 0.1$ m is the surface roughness length, and ψ_w comes from Salomons Eqs. (N.20 for convective conditions, N.26 and N.27 for stable conditions). In this calculation the wind speed at 10 m above the ground is 1.8 m $\mathrm{s^{-1}}$ for SC conditions, 2.6 m $\mathrm{s^{-1}}$ for WC conditions, and 3.2 m $\mathrm{s^{-1}}$ for WS conditions.

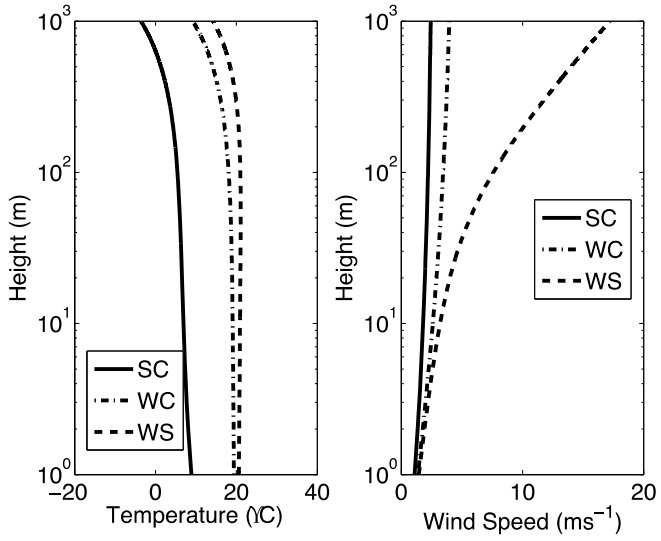

Figure C-7. The mean vertical temperature and wind velocity profiles for strongly convective (solid), weakly convective (dash-dotted), and weakly stable (dashed) conditions.

C.2.2 Atmospheric Refraction

As sound waves propagate in the ABL, they will change direction (i.e., bend or refract), due to changes in the sound speed. For sound propagation outdoors this effect is called atmospheric refraction. The effects of atmospheric refraction on received SPLs far from the source are often large when the source and receiver are close to the ground.[25]

In an atmosphere that has no wind, the sound speed is primarily a function of temperature. The sound speed is also a dependent upon humidity, but the contribution of humidity to the overall sound speed is minimal in comparison to temperature. Within a temperature range of -20 to 40 °C the sound speed has an approximately linear relationship with temperature.[25] The adiabatic sound speed for dry air is:[111]

$$c = (331 + 0.6T_C)(1 + 0.16h) \tag{C.8}$$

where c is the sound speed (ms^{-1}), T_C is the temperature in Celsius, and h are the fraction of moles that are water.[111]

Propagation of sound in a moving medium (e.g., propagation in the atmosphere that has wind), is more complex, but can be handled in the same matter as a non-moving medium if the effective sound speed is used:

$$c_{\text{eff}} = c + u\cos\theta. \tag{C.9}$$

Here c_{eff} is the effective sound speed, c is the sound speed from Eq. (C.8), u is the horizontal component of the wind speed, and θ is the angle between the wind and propagation direction.[90] Using Equation (C.9), and the mean temperature and wind velocity profiles from the Wilson[120] and Holtslag model[89*****] given above, the effective sound speed profiles can be calculated. Figure C-8 shows the effective sound speed profiles in downwind, crosswind, and upwind directions for strongly convective (Figure C-8 left), weakly convective (Figure C-8 middle), and weakly stable (Figure C-8 right) conditions. The effective sound speed increases in the downwind direction and decreases in the upwind direction as the height increases, resulting in downward and upward refraction conditions, respectively.

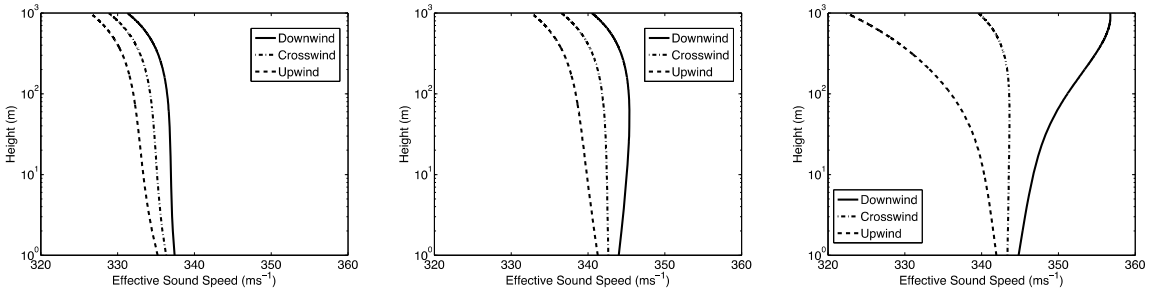

Figure C-8. An example of effective sound speed profiles for SC (left), WC (middle), and WS (right) conditions in the downwind (solid line), crosswind (dotted-dashed line), and upwind (dashed line) propagation directions.

In most outdoor sound propagation models, the ABL is broken up into horizontal layers of homogenous effective sound speeds. As sound waves propagate from one layer to another the sound waves will refract depending upon the difference between the effective sound speeds of the

*****The Holtslag model is the model given in Salomons.

two layers. Qualitatively this can demonstrated using ray theory (Figure C-9), though as noted by Embleton, the application of ray theory is only valid at very high frequencies.[106]

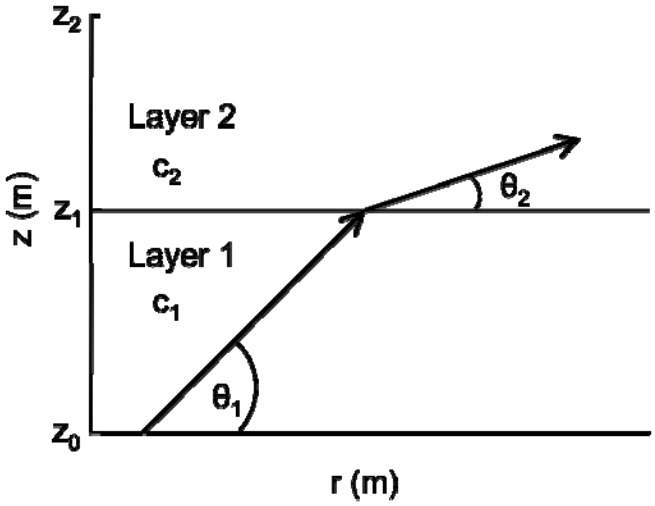

Figure C-9. Refraction of a sound ray propagating through Layer 1 into Layer 2. Layer 1 is defined as the region between z_0 and z_1, and Layer 2 is defined as the region between z_1 and z_2. c_1 and c_2 are the effective sound speeds of the two layers, and θ_1 and θ_2 are the propagation directions. The horizontal axis is the radial distance and the vertical axis is the height above the ground.

In this example, the sound speed in Layer 1 is less than the sound speed in Layer 2 and the sound ray refracts toward the lower sound speed according to Snell's law.[25] When $c_1 < c_2$ and the sound ray refracts downward ($\theta_1 > \theta_2$), which is known as downward refraction. When $c_1 > c_2$ the sound ray refracts upward ($\theta_1 < \theta_2$), which is known as upward refraction. For propagation from a point source the same ray theory principles apply, though the result is a deformation of the spherical waves.[25]

The effects of atmospheric refraction result in a redirection or focusing of sound energy, but the net effects are dependent upon the ground surface as well. For example, downward refraction near the ground can result in a ducting effect and "quiet heights", which depend on ground impedance.[121] Also, the combined effects of downward refraction and a very hard ground

can result in very little attenuation at large distances from the source, and in upward refraction conditions can result in a shadow zone of very low SPLs.[22]

An example of the effect of the effective sound speed profiles in Figure C-8 and grass-covered ground applied to the C4 explosion (Figure C-2 solid line) is given in Figure C-10. There is a large difference in the received SPLs, at a height of 1.5 meters above the ground, in the downwind and upwind directions at distances of $1 - 10$ km from the source.

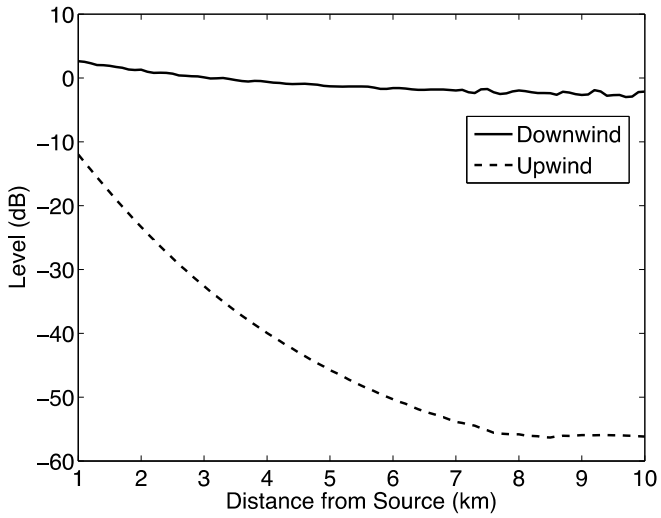

Figure C-10. Change in level (dB) due to refraction and ground effect only (i.e., without the effect of geometric spreading and atmospheric absorption) as a function of distance (km) from the C4 explosion source for a grass covered ground (flow resistivity $= 200$ kPa s m^{-2}). The source height is 3 m and the receiver height is 1.5 m above the ground.

C.2.3 Atmospheric Turbulence

In general, the effect of upward refraction will result in a significant decrease in SPL. In a turbulent ABL, however, the loss in SPL due to upward refraction is less because some of the sound energy is scattered back towards the ground (Figure C-11). In the atmosphere, turbulence is always present near the ground, even on a calm day.[106] The size of the turbulent eddies in the ABL ranges from 1 mm to the height of the ABL. The two mechanisms responsible for

turbulence in the ABL are thermal convection from differential heating and mechanical flow from wind shear near the ground and other objects.[25] As sound waves propagate through the atmosphere they encounter a spectrum of different sized turbulent eddies that have temperature and wind velocities that differ and fluctuate around the mean temperature and wind velocities in that region.[106] The near-instantaneous (i.e., a few seconds) variations in temperature and wind velocities are typically on the order of $\pm 5°C$ and $\pm 1/3$ the mean wind velocity.[106] The effect of these variations, collectively called atmospheric turbulence, either refracts or scatters sounds waves altering the phase and amplitude, and causes fluctuations in received SPLs in time.

Turbulence describes a large class of complex, irregular, and unpredictable motions.[122] Turbulence is stochastic and therefore dealt with in a statistical manner. Turbulence is diffusive. Energy in turbulent flows tends to mix fluids. Turbulence causes the formation of many different sized eddies, which are coherent swirling patterns of velocity, vorticity, and pressure. Turbulence is also dissipative. Energy in large eddies cascade or break down into smaller eddies, which in turn, break down into smaller eddies until the eddy sizes reach approximately 1 mm. At a size of 1 mm, the turbulence then dissipates through viscosity and thermal conduction.[122,123]

In order to highlight the effects of turbulence, the propagation simulation involving refraction and ground effect (Figure C-10), is redone with the inclusion of turbulence (Figure C-11) using a phase screen method.[83] As can be seen by comparing the two figures, there is not much difference in received SPLs in the downwind direction and downward refraction condition. However, for the upwind direction and upward refraction condition the effects of turbulence reduce the amount of SPL loss by approximately 10 dB at distances greater than 5 km (Figure C-12). This effect is expected and noted in the literature.[25,124]

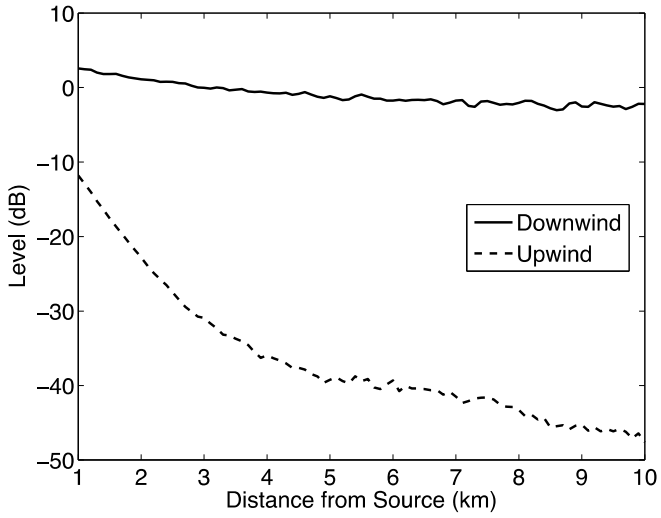

Figure C-11. Change in level (dB) due to turbulence, refraction and ground effect only (i.e., without the effect of geometric spreading and atmospheric absorption) as a function of distance (km) from the C4 explosion source for a grass covered ground (flow resistivity = 200 kPa s m^{-2}). The source height is 3 m and the receiver height is 1.5 m above the ground.

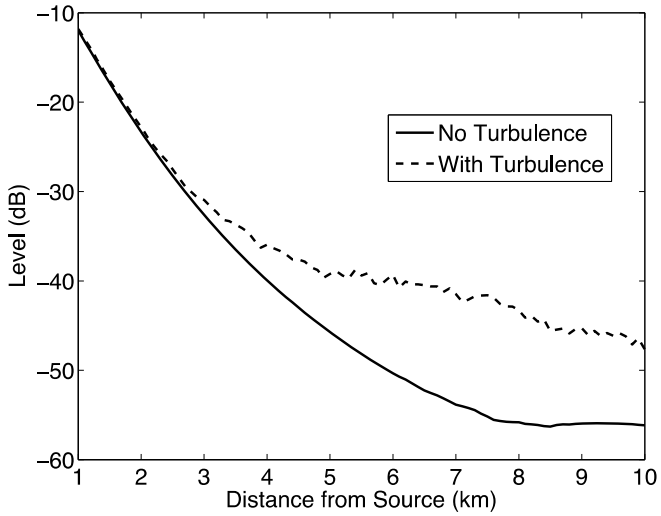

Figure C-12. Comparison of SPLs with and without the effect of atmospheric turbulence for the upwind directions in Figure C-10 and Figure C-11. The change in level (dB) as a function of distance (km) from the C4 explosion source for a grass covered ground (flow resistivity = 200 kPa s m^{-2}). The source height is 3 m and the receiver height is 1.5 m above the ground.

D Deterministic Interpolation Models

This appendix describes some common deterministic 2D spatial interpolation models, and presents results from an exploratory analysis (Sec. D.3). Some of these models have been used in noise monitoring and assessment applications, and include inverse square distance methods (also referred to as geometric spreading), minimum curvature methods (also referred to as thin plate smoothing or biharmonic splines),[42] and geographic information system (GIS) based methods,[34,35] which presume tessellations, Voronoi polygons, and Delaunay triangles to interpolate the data. The deterministic interpolation models described in this appendix are listed in Table D-1. The first four models fall into a class of linear combination models (Sec. D.1) and the next four models fall into a class of polyharmonic splines (Sec. D.2).

Table D-1. A list of deterministic interpolation models presented in this appendix.

Model	Abbreviation	Class
Nearest Neighbor	NRN	Linear Combination
Natural Neighbor*	NTN	Linear Combination
Delaunay Triangulation*	DT	Linear Combination
Inverse Distance Weighted	IDW	Linear Combination
Distance Weighted Interpolation	DWI	Polyharmonic Spline
Thin Plate Splines	TPS	Polyharmonic Spline
Polynomial Surface Fits	PSF	Polyharmonic Spline
Local Weighted Regression	LWR	Polyharmonic Spline

* These models do not perform extrapolation

D.1 Linear Combination Models

The linear combination models examined in this appendix include Nearest Neighbor (NRN), Delaunay Triangulation (DT), Natural Neighbor (NTN), and Inverse Distance Weighted (IDW) models. They each take the following general form:

$$P(\boldsymbol{x}) = \sum_{i=1}^{N} w_i \, P(\boldsymbol{x}_i), \tag{D.1}$$

where $P(\boldsymbol{x})$ is the estimated SPL at location $\boldsymbol{x}$, w_i are the weights associated with the measured data, $P(\boldsymbol{x}_i)$ are the measured SPLs at the noise monitor locations $\boldsymbol{x_i}$, and N is the number of noise monitors.[37] Given that the noise monitor and estimation heights are fixed to 1.5 m above the ground for this dissertation, $\boldsymbol{x}$ and $\boldsymbol{x_i}$ are given in terms of Cartesian coordinates (x, y). The differences between the models in this class are the methods used to determine w_i.

D.1.1 Nearest Neighbor

Interpolation procedures that involve Voronoi diagrams or Delaunay Triangulations are also known as polygonal or triangular methods.[94] One of the simplest polygonal models is the NRN model. The estimation point is assigned the value of the nearest measured noise monitor data point. The region of influence for each measured data point is a Voronoi polygon (Figure D.2) and can be described mathematically:

$$V(\boldsymbol{x}_i) = \left\{\boldsymbol{x} \mid \| \boldsymbol{x} - \boldsymbol{x}_i \|_2 \leq \| \boldsymbol{x} - \boldsymbol{x}_j \|_2 \text{ for } j \neq i \;\; i, j \in I_n\right\}, \tag{D.2}$$

where $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subset \mathbb{R}^2$ is the set of all noise monitor locations, $\| \boldsymbol{x} - \boldsymbol{x}_i \|_2$ is the Euclidean distance, $I_n = \{1, \ldots, n\}$ is the index of noise monitors, and $\mathcal{V} = \{V(\boldsymbol{x}_1), \ldots, V(\boldsymbol{x}_n)\}$ is the set of all Voronoi polygons.[94] SPL estimates with NRN interpolation is given by Eq. D.1, where the weights are given by:

$$w_i = \begin{cases} 1, & if\ \boldsymbol{x} \in V(\boldsymbol{x}_i) \\ 0, & \text{otherwise.} \end{cases} \quad \text{(D.3)}$$

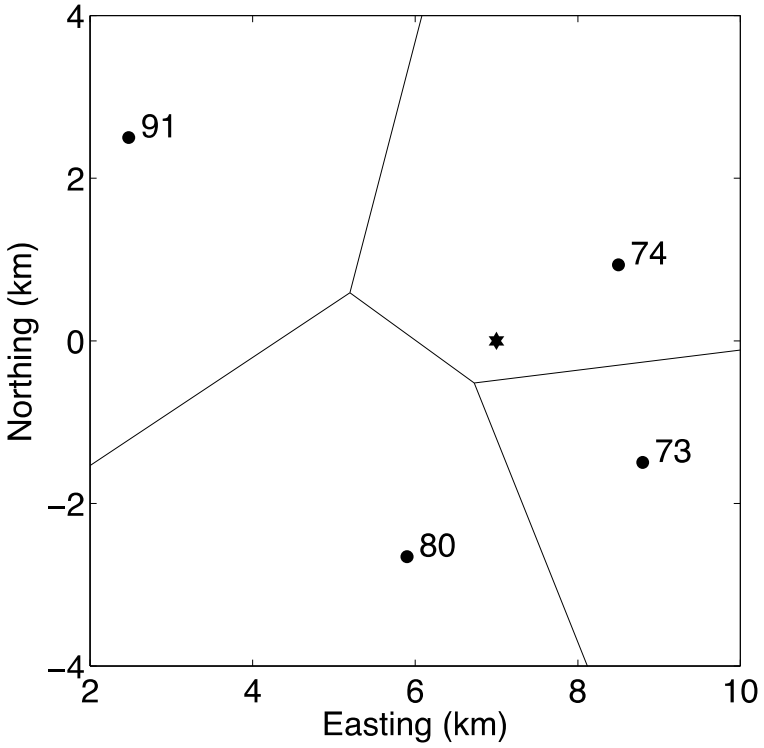

Figure D-1. Voronoi polygons drawn from noise monitor locations (circles). For NRN interpolation, the estimation point (star) is assigned the SPL of 74 dB in this example because that is the level at the near noise monitor.

D.1.2 Natural Neighbor

The NTN interpolation model is very similar to NRN model in that the model is also derived from Voronoi polygons. Here, an additional set of Voronoi polygons is derived from a set of points that include the estimation point and noise monitor locations (Figure D-2). The polygon that includes the estimation point is designated $V(\boldsymbol{x}_0)$. The polygon $V(\boldsymbol{x}_0)$ is overlaid on the original set of Voronoi polygons ($\mathcal{V}$), and the weights for NTN interpolation model are defined by

area that $V(\boldsymbol{x}_0)$ overlaps with $V(\boldsymbol{x}_i)$,[125] which can be seen in Figure D-3. Mathematically this can be written as

$$w_i = (A_{V_i \cap V_0})/A_{V_0} \tag{D.4}$$

where A_{V_i} and A_{V_0} are the areas of their respective Voronoi polygons, and ∩ indicates the intersection of the sets V_i and V_0.

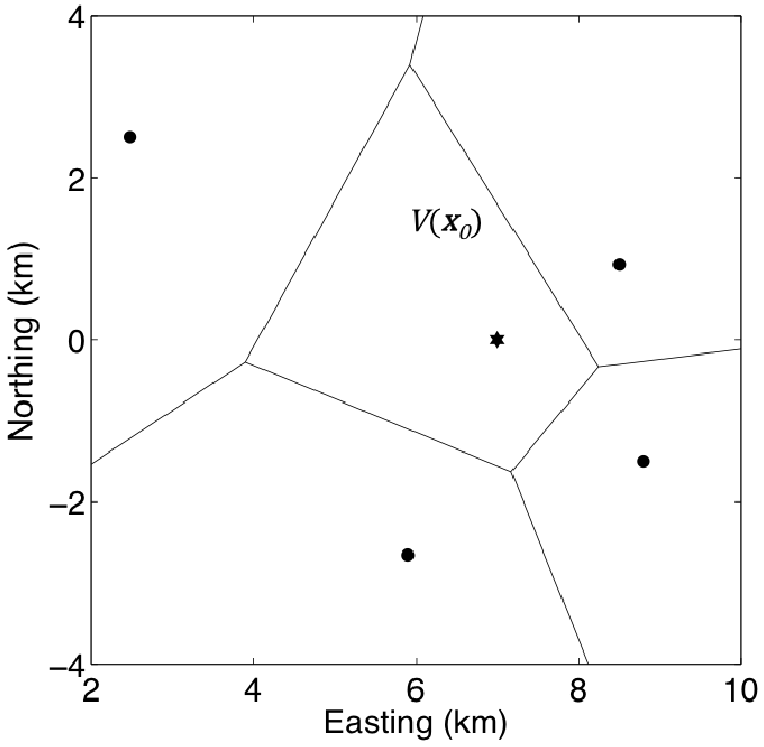

Figure D-2. Voronoi polygons derived from a set of points that include the estimation point (star) and noise monitor locations (circles). The polygon that includes the estimation point is labeled $V(x_0)$.

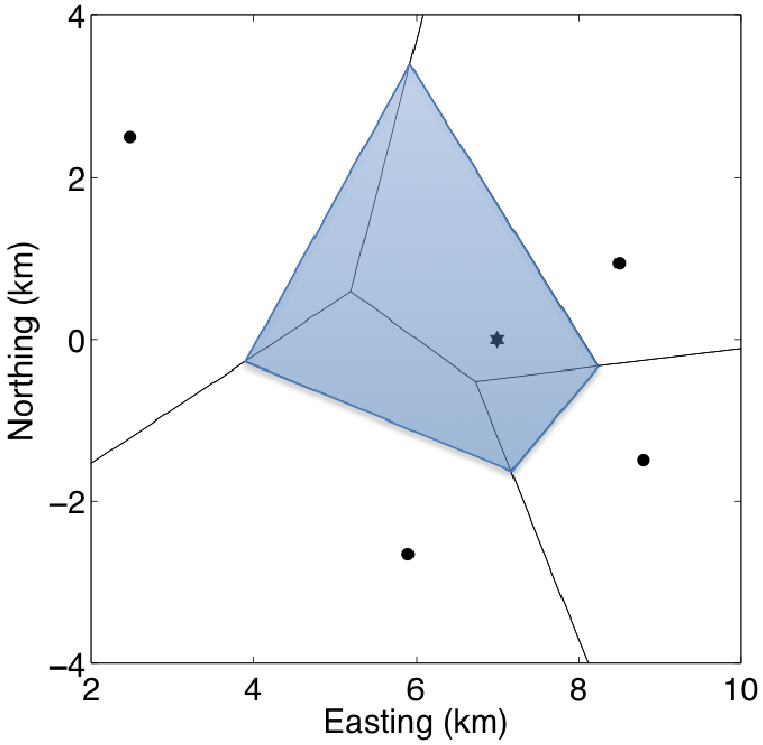

Figure D-3. Overlay of two sets of Voronoi polygons. One set of polygons is derived from the set of noise monitor locations (Figure D.2), and the other is derived from the set of noise monitor locations and estimation point (Figure D.3). The shaded region shows polygon $V(x_0)$.

D.1.3 Delaunay Triangulations

Assuming that the noise monitor locations x_i are not collinear, and the number of noise monitors is finite and greater than two (i.e., $N \geq 3$), Delaunay Triangles can be drawn from Voronoi polygons. Vertices of the triangles are defined by noise monitor locations, and the edges are formed by drawing a line between vertices that share a common Voronoi polygon (Figure D-4). This construction results in a set of Delaunay Triangles $\mathcal{T} = \{T_1, \ldots T_m\}$, where m is the number of triangles. In the example given in Figure D-4, $m = 2$.

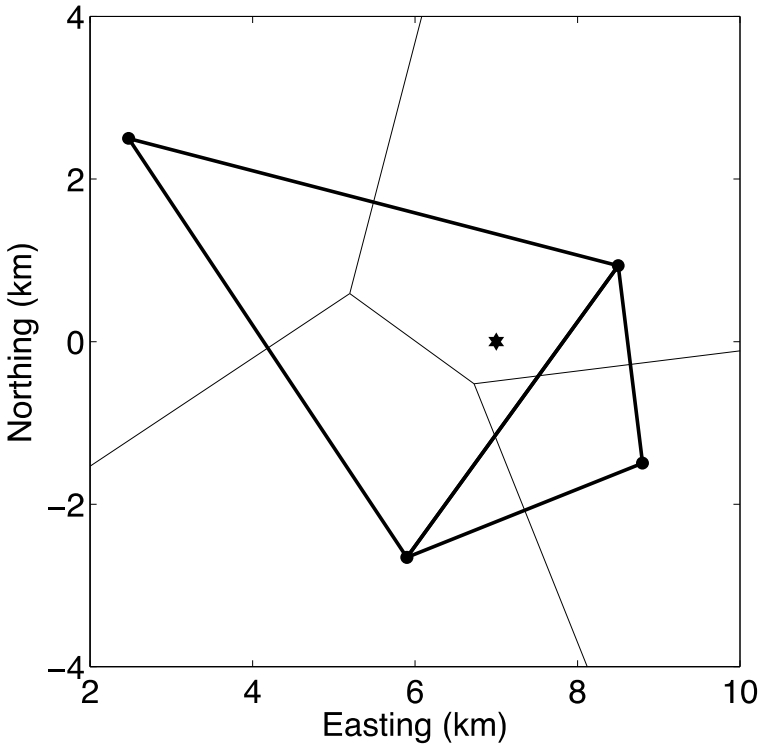

Figure D-4. Two Delaunay triangles (bold thick lines), drawn from Voronoi polygons (thin lines). The noise monitor locations (circles) form the vertices of the triangles. The estimation point (star) is located with the large triangle.

For DT linear (DTL) interpolation, the triangle containing the estimation point, and the noise monitor locations that make up the vertices of that triangle are used to choose the weights needed to solve Eq. (D.1).[125] The weights for noise monitors that define the Delaunay Triangle may be found using matrix determinants as given in de Smith et al.,[125] or can be defined by the ratio of triangle areas:

$$w_1 = A_{T_{023}}/A_{T_{123}}$$
$$w_2 = A_{T_{013}}/A_{T_{123}} \qquad \text{(D.5)}$$
$$w_3 = A_{T_{012}}/A_{T_{123}}$$

where $A_{T_{abc}}$ is the area of the triangle defined by its vertices abc.[126] Figure D-5 gives an illustration of the four triangles needed to solve Eq. D.5, where the estimation point is labeled 0 and the noise monitor locations are labeled 1-3.

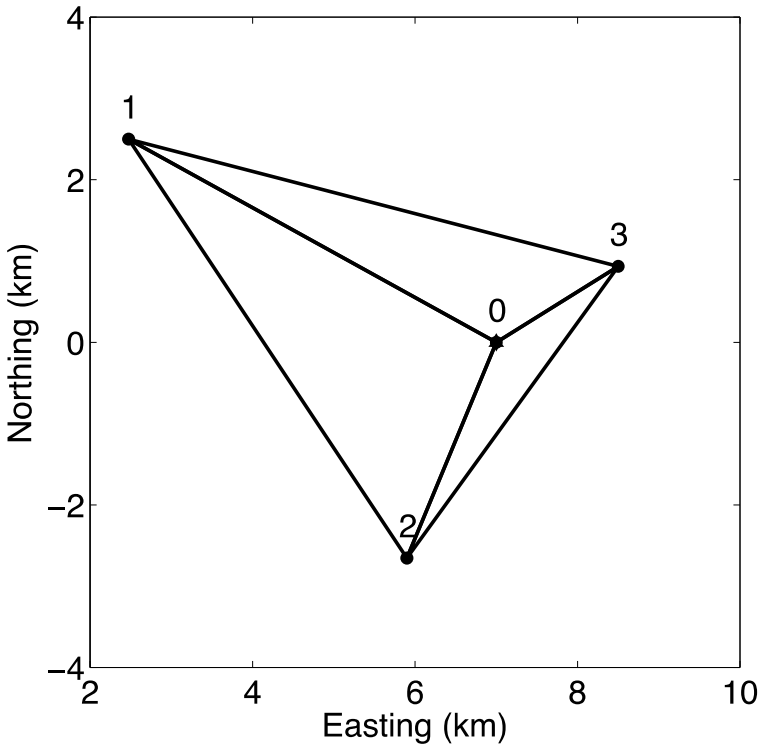

Figure D-5. Illustration of the four triangles: $T_{012}, T_{013}, T_{023}, T_{123}$. The noise monitor locations (circles) are labeled $1-3$, and the estimation point (star) is labeled 0.

Another variant of the DT model is the DT cubic spline (DTC) model. DTC uses the same Delaunay Triangles as the linear method, except a piece-wise cubic function is fit to the triangle vertices. This model uses the Clough-Toucher procedure and subdivides the triangle into 3 additional triangular regions using the centroid of the main triangle. Separate bicubic patches are fit to each of the three sub-triangles, resulting in a continuous spline-like surface.[125]

D.1.4 Inverse Distance Weighted

The IDW interpolation model, which is also known as Shepard's algorithm,[127] weights the noise monitor data according to the distance between the estimation and noise monitor locations. The noise monitors closer to the estimation points are given more weight and the weights are constrained to sum to 1. In practice the distance may also be squared or cubed to further emphasis points that are closer to the point of interest. The weights needed for Eq. (D.1) are:

$$w_i = \left(\frac{1}{d_i^{\zeta}}\right) / \sum_{j=1}^{n} \frac{1}{d_j^{\zeta}}, \tag{D.6}$$

where d_i^{ζ} is the distance between the estimation point and noise monitor location $\boldsymbol{x}_i$.[126] In practice ζ is typically set to 1, 2, and 3.

In order to demonstrate the effect of ζ a hypothetical example is given. Consider a situation in which there are 5 noise monitors located exactly $1, 2, 3, 4$ and 5 km away from an estimation point. The weight that each data point (i.e., the SPL at the noise monitor) is given is listed in Table D-2 for $\zeta = 1, 2, 3, 4$ and 5. Here it can be seen that as the value of ζ increases the number of points that significantly contribute to the estimate decreases very rapidly. With $\zeta > 3$, there is hardly any contribution from the points beyond 1 km from the estimation point.

Table D-2. Weights given to each noise monitor data point for the hypothetical case of 5 noise monitors located $1, 2, 3, 4$ and 5 km away from an estimation point for $\zeta = 1, 2, 3, 4$ and 5.

ζ	1km	2km	3km	4km	5km
1	0.438	0.219	0.146	0.109	0.088
2	0.683	0.171	0.076	0.043	0.027
3	0.843	0.105	0.031	0.013	0.007
4	0.926	0.058	0.011	0.004	0.001
5	0.965	0.030	0.004	0.001	0.000

D.2 Polyharmonic spline regression models

The polyharmonic spline regression models examined in this appendix include Distance Weighted interpolation (DWI), Thin Plate Splines (TPS), Polynomial Surface Fits (SPF), and Local Weighted Regression (LWR) models with the following general form:

$$P(\boldsymbol{x}) = \sum_{i=1}^{n} w_i \, \varphi(r) + \boldsymbol{\beta}^{\mathrm{T}} \begin{bmatrix} \mathbf{1} \\ \mathbf{x} \end{bmatrix}, \tag{D.7}$$

where $P(\boldsymbol{x})$ is the estimated SPL at location $\boldsymbol{x}$, w_i are the weights associated with radial basis function φ, $r = \|\boldsymbol{x} - \boldsymbol{x}_i\|_2$ is the Euclidean distance between the estimation points $\boldsymbol{x}$ and noise monitor locations $\boldsymbol{x}_i$, $\boldsymbol{\beta}$ are the weights associated with the polynomial $\begin{bmatrix} \mathbf{1} \\ \mathbf{x} \end{bmatrix}$.[128,129] Here the weights w_i and $\boldsymbol{\beta}$ are found using linear regression.

D.2.1 Distance Weighted

One of the simplest polyharmonic splines is one whose basis function is the distance between the estimation point and the noise monitor locations. The DWI model is essentially a piecewise linear function that has the following form:[128]

$$P(\boldsymbol{x}) = \sum_{i=1}^{n} w_i \, \| \, \boldsymbol{x} - \boldsymbol{x}_i \, \|_2. \tag{D.8}$$

In the notation of the polyharmonic splines, the basis function $\varphi(r) = r$, where r is Euclidean distance, $\boldsymbol{\beta}$ is set to zero, and the coefficients w_i are found by solving the following system of linear equations:

$$P(\boldsymbol{x}_\alpha) = \sum_{\beta=1}^{n} w_\beta \; \| \, \boldsymbol{x}_\alpha - \boldsymbol{x}_\beta \, \|_2 \; \forall \, \alpha = 1, \ldots, n \tag{D.9}$$

D.2.2 Thin Plate Splines

The TPS model is essentially the 2D analog to the 1D cubic spline model.[130] The term TPS refers to the bending of a thin plate that is fixed at various locations. Here the radial basis function in Eq. (D.7) is $\varphi(r) = r^2 \ln r$, $\boldsymbol{\beta}$ is set to zero, and the coefficients w_i are found by solving a similar set of linear equations as given in Eq. (D.9). The estimate of the SPL at estimation points $\boldsymbol{x}$ is given by the following equation:[128]

$$P(\boldsymbol{x}) = \sum_{i=1}^{n} w_i \, r^2 \ln r \,. \tag{D.10}$$

The TSP model, given by Eq. (D.10), estimates the SPL at noise monitor locations exactly (i.e., the estimate is equal to the actual SPL at noise monitor locations). In this study, the TPS model is implemented using the *tpaps* function in MATLAB.[81]

D.2.3 Polynomial Surface Fits

The 2D polynomial surface fits within the context of the polyharmonic spline equation given in Eq. (D.7), where w_i is set equal to zero:

$$P(\boldsymbol{x}) = \boldsymbol{\beta}^{\mathrm{T}} \begin{bmatrix} \mathbf{1} \\ \mathbf{x} \end{bmatrix}, \tag{D.11}$$

where $\boldsymbol{\beta}^{\mathrm{T}} = [\beta_0 \, \beta_1 \ldots \beta_k]$ is a matrix of coefficients, $\mathbf{x}$ is 2D polynomial in the position coordinates (x, y), and both $\boldsymbol{\beta}$ and $\mathbf{x}$ depend on the order of the polynomial. Five polynomial surface fits are examined in this appendix ranging from first order in *x* and *y* having 3 coefficients:

$$P(x,y) = \beta_0 + \beta_1 x + \beta_2 y \quad \text{(D.12)}$$

to fifth order polynomials in *x* and *y* having 21 coefficients, which include cross terms

$$P(x,y) = \beta_0 + \beta_1 x + \beta_2 y + \cdots + \beta_{18} x^4 y + \beta_{19} x^5 + \beta_{20} x y^4 + \beta_{21} y^5. \quad \text{(D.13)}$$

In practice, the number noise monitor (N) data points limit the use of higher order polynomials. The number of data points must be equal or preferably greater than the number of coefficients in the regression equation in order to obtain a solution. Also, there is an increased risk of over-fitting the data as the order of the polynomials and coefficients are increased. The coefficients in the polynomial equations are found by least-squares minimization of the residuals. Here the matrix form of the linear model is given as:

$$P = \mathrm{X}\boldsymbol{\beta} + \varepsilon \quad \text{(D.14)}$$

where $\boldsymbol{P}$ is vector of SPLs, $\boldsymbol{\beta}$ is a vector of coefficients, **X** is a matrix of locations, and $\boldsymbol{\varepsilon}$ is a vector of errors. Estimates of $\boldsymbol{\beta}$ are found using least-squares regression:

$$\boldsymbol{\beta} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{P} \quad \text{(D.15)}$$

where $\boldsymbol{X}^T$ is the transpose of the matrix X.[131] In this study, the PSF models are implemented using the *fit* function in the MATLAB Curve Fitting Toolbox.[81]

D.2.4 Local Weighted Regression

Local weighted regression (LWR) models are very similar to the polynomial surface fits except only a user specified percentage of the total number of noise monitors are used in the fitting procedure. This dissertation considers using 25%, 50%, and 75% of the noise monitor data by limiting the polynomial surface fits to a span or percentage of the monitors closest to the

estimation point. The weights of the data points are determined by the following tri-cube expression:

$$\lambda_i = (1 - \left|\frac{\boldsymbol{x} - \boldsymbol{x}_i}{d(\boldsymbol{x})}\right|^3)^3 \quad \text{(D.16)}$$

where λ_i are the weights for each data point in the span $\boldsymbol{x}_i, \boldsymbol{x}$ is the estimation point, and $d(\boldsymbol{x})$ is the distance between $\boldsymbol{x}$ and the farthest $\boldsymbol{x}_i$. Then a weighted regression procedure is used to solve for the polynomial coefficients ($\boldsymbol{\beta}$). Eq. (D.14) is modified to include the weighting:

$$\boldsymbol{\beta} = (\boldsymbol{X}^T \boldsymbol{\Lambda} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{\Lambda} \boldsymbol{P} \quad \text{(D.17)}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with the diagonal elements given by the weights ($\boldsymbol{\lambda}$).[132]

D.3 Analysis of Deterministic Models

In this section results from an exploratory analysis of the deterministic models presented in this appendix are given. Only the models that perform extrapolation are considered. The models are assessed in terms of the mean $RMSE_{SPL}$ (as described in Sec. 3.5) across 50 noise monitor geometries using the LHOS[†††††] procedure. For this analysis the most accurate models are shown for cases in which there are 4, 9, 16, 25, 36, and 49 noise monitors. The take away from this analysis is that the TPS model is by far the most accurate deterministic interpolation model of the models presented in this appendix (Table D-3). In general, as the number of noise monitors increases the accuracy increases (i.e., the $RMSE_{SPL}$ decreases).

[†††††] This analysis was done with a different set of noise monitor data/geometries than the cLHOS datasets that were used to analyze spatial interpolation models analyzed in Chapter 4.

Table D-3. Model accuracy of various deterministic spatial interpolation models as a function of the number of noise monitors.

Model	N=4	N=9	N=16	N=25	N=36	N=49
TPS	3.8	1.6	1.1	0.8	0.7	0.7
PY2	NA	3.6	1.2	0.9	0.8	0.8
PY3	NA	NA	1.4	1.1	1.0	0.9
IDW4	5.4	3.4	2.4	2.0	1.6	1.4
IDW5	5.5	3.5	2.5	2.0	1.6	1.4
IDW3	5.3	3.5	2.6	2.1	1.7	1.5
NRN	6.7	4.2	3.2	2.5	2.1	1.9
IDW2	5.5	4.2	3.4	3.0	2.7	2.5
PY1	3.9	3.1	2.9	2.8	2.8	2.8
IDW1	6.7	6.1	5.8	5.7	5.6	5.6

E Supplemental Material for Kriging Derivations

This appendix shows the steps that were omitted in the derivations presented in Chapter 2. The steps omitted in the derivation of the simple kriging (SKG, Sec. E.1), the derivation of the variance in terms of semivariance (Sec. E.2), the method of Lagrangian Multipliers (Sec. E.3), and the derivation of ordinary kriging (OKG) variance (Sec. E.4) are given.

E.1 Derivation of SKG Variance

Starting with Eq. (E.1):

$$\sigma^2(\mathbf{x_0}) = \mathrm{Var}[Z^*(\mathbf{x_0}) - Z(\mathbf{x_0})] = \mathrm{E}\left[\left(Z^*(\mathbf{x_0}) - Z(\mathbf{x_0})\right)^2\right], \tag{E.1}$$

by the definition of variance.[64] Next, (Eq. E.2) is expanded:

$$\mathrm{E}\left[\left(Z^*(\mathbf{x_0}) - Z(\mathbf{x_0})\right)^2\right] = \mathrm{E}\left[\left(Z^*(\mathbf{x_0})\right)^2\right] + \mathrm{E}\left[\left(Z(\mathbf{x_0})\right)^2\right] - 2\mathrm{E}[Z^*(\mathbf{x_0})\, Z(\mathbf{x_0})], \tag{E.2}$$

and the expression for $Z^*(\mathbf{x_0})$ in Eq. (2.7) is inserted into Eq. (2.11). In order to make the derivation more manageable, the right hand side (RHS) of Eq. (2.11) is broken up into three parts (A, B, and C):

$$\mathrm{E}\left[\left(Z^*(\mathbf{x_0}) - Z(\mathbf{x_0})\right)^2\right] = \mathrm{A} + \mathrm{B} - \mathrm{C}, \text{where} \tag{E.3}$$

$$\mathrm{A} = \mathrm{E}\left[\left(Z^*(\mathbf{x_0})\right)^2\right] = \mathrm{E}\left[\left(\sum_{i=1}^{N} w_i \left(Z(\mathbf{x}_i) - m(\mathbf{x}_i)\right) + m(\mathbf{x})\right)^2\right],$$

$$\mathrm{B} = \mathrm{E}\left[\left(Z(\mathbf{x_0})\right)^2\right], \text{and}$$

$$\mathrm{C} = -2\mathrm{E}[Z^*(\mathbf{x_0})\, Z(\mathbf{x_0})].$$

Given that $m(\mathbf{x}_i) = m(\mathbf{x_0})$, the position vector $\mathbf{x}$ is removed from the rest of this derivation. The RHS of A, is expanded and the expectation is moved into the summation operator:

$$A = \mathrm{E}\left[m^2 + 2m\sum_{i=1}^{N} w_i \left(Z(\mathbf{x}_i) - m\right) + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j (Z(\mathbf{x}_i) - m)\left(Z(\mathbf{x}_j) - m\right)\right],$$

$$\mathrm{A} = \left(m^2 + 2m\sum_{i=1}^{N} w_i \left(E[Z(\mathbf{x}_i)] - m\right) + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \mathrm{E}\left[(Z(\mathbf{x}_i) - m)\left(Z(\mathbf{x}_j) - m\right)\right]\right).$$

The second term in this equation goes to zero given that $E[Z(\mathbf{x}_i)] = m$, and the expression can be simplified given that the expectation in the third term is the definition of covariance (Eq. 2.6), which yields:

$$A = m^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right]. \tag{E.4}$$

Eq. (B) is modified using the definition of covariance (Eq. 2.5) and recognizing that:

$$\begin{gathered}\mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] = \mathrm{E}\big[\big(Z(\mathbf{x}_0) - m(\mathbf{x}_0)\big)\big(Z(\mathbf{x}_0) - m(\mathbf{x}_0)\big)\big], \\ \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] = \mathrm{E}[Z(\mathbf{x}_0)^2 - 2m\, Z(\mathbf{x}_0) + m^2], \\ \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] = \mathrm{E}[Z(\mathbf{x}_0)^2] - 2m^2 + m^2, \\ \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] = \mathrm{E}[Z(\mathbf{x}_0)^2] - m^2,\end{gathered} \tag{E.5}$$

which yields,

$$\mathrm{B} = \mathrm{E}\left[\big(Z(\mathbf{x_0})\big)^2\right] = \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] + m^2. \tag{E.6}$$

Eq. (C) is expanded in a similar manner as Eq. (A), and simplified using the expression for covariance (Eq. 2.18), which yields:

$$\mathrm{C} = -2\mathrm{E}\left[\sum_{i=1}^{N} w_i\, (Z(\mathbf{x}_i) - m + m) Z(\mathbf{x_0})\right],$$

$$\mathrm{C} = -2\mathrm{E}\left[\sum_{i=1}^{N} w_i\ Z(\mathbf{x}_i)\, Z(\mathbf{x_0})\right],$$

$$\mathrm{C} = -2\sum_{i=1}^{N} w_i\, \mathrm{E}[Z(\mathbf{x}_i)\, Z(\mathbf{x_0})],$$

$$\mathrm{C} = -2\sum_{i=1}^{N} w_i\, (\mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)] + m^2),$$

$$\mathrm{C} = -2\, m^2 - 2\sum_{i=1}^{N} w_i\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)]. \tag{E.7}$$

Substituting, Eqs. (2.17, 2.19, and 2.20) into Eq. (2.15) gives the variance of the estimator at location $\mathbf{x}_0$ in terms of m, w_i, and Cov[ ]:

$$\sigma^2(\mathbf{x_0}) = m^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] + m^2$$
$$- 2\, m^2 - 2\sum_{i=1}^{N} w_i\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)], \text{ which simplifies to}$$
$$\sigma^2(\mathbf{x_0}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] \qquad \text{(E.8)}$$
$$- 2\sum_{i=1}^{N} w_i\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)].$$

As mentioned earlier (Sec. 2.3), kriging models are optimal estimators because they are unbiased and have minimal variance. For the SKG model, the estimation variance is minimal when the partial derivative of variance (Eq. 2.21) with respect to the weights is zero:

$$\frac{\partial \sigma^2}{\partial w_i} = \sum_{j=1}^{N} w_j\, \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] - 2\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)] = 0,$$

$$\frac{\partial \sigma^2}{\partial w_j} = \sum_{i=1}^{N} w_i\, \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] = 0.$$

Given that $w_i = w_j$, these equations can be combined to give the SKG system of equations:

$$2\sum_{j=1}^{N} w_j\, \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] - 2\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)] = 0, \text{ which simplifies to}$$
$$\sum_{j=1}^{N} w_j\, \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] = \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)] \text{ for } i = 1, \ldots, N. \qquad \text{(E.9)}$$

E.2 Derivation of Variance in terms of Semivariance

Using the relationship between covariance and semivariance (Eq. 2.20) and Eq. (2.11), the variance can be given in terms of the semivariance:

$$\sigma^2(\mathbf{x_0}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] - 2\sum_{i=1}^{N} w_i\, \mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)],$$

where from Eq. (2.20), the following substitutions can be made:

$$\mathrm{Cov}\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] = \mathrm{C}(0) - \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right],$$

$$\mathrm{Cov}[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] = \mathrm{C}(0) - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)],$$

$$\mathrm{Cov}[Z(\mathbf{x}_i), Z(\mathbf{x}_0)] = \mathrm{C}(0) - \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)],$$

which gives:

$$\sigma^2(\mathbf{x_0}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j \left(\mathrm{C}(0) - \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right]\right) + \mathrm{C}(0) - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)]$$

$$- 2\sum_{i=1}^{N} w_i\, (\mathrm{C}(0) - \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)]), \text{and expands to:}$$

$$\sigma^2(\mathbf{x_0}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \mathrm{C}(0) - \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \mathrm{C}(0) - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)]$$

$$- 2\sum_{i=1}^{N} w_i\, \mathrm{C}(0) + 2\sum_{i=1}^{N} w_i\, \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)], \text{which simplifies to}$$

$$\sigma^2(\mathbf{x_0}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \mathrm{C}(0) - 2\sum_{i=1}^{N} w_i\, \mathrm{C}(0) + \mathrm{C}(0) - \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right]$$

$$- \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] + 2\sum_{i=1}^{N} w_i\, \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)],$$

$$\sigma^2(\mathbf{x_0}) = \mathrm{C}(0)(1 - 2 + 1) - \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)]$$

$$+ 2\sum_{i=1}^{N} w_i\, \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)],$$

$$\sigma^2(\mathbf{x_0}) = -\sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \gamma\big[Z(\mathbf{x}_i), Z\big(\mathbf{x}_j\big)\big] - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] + 2\sum_{i=1}^{N} w_i\, \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)], \text{and}$$

$$\sigma^2(\mathbf{x_0}) = -\sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \gamma\big[Z(\mathbf{x}_i), Z\big(\mathbf{x}_j\big)\big] - \gamma[Z(\mathbf{x}_0), Z(\mathbf{x}_0)] + 2\sum_{i=1}^{N} w_i\, \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)]. \quad \text{(E.10)}$$

E.3 Method of Lagrangian Multipliers

The method of Lagrangian multipliers is important for situations where an equation needs to be minimized according to a given constraint. The method is given as follows:

$$\frac{\partial}{\partial x_\alpha}\left(f(x) - \sum_{\beta=1}^{m} \lambda_\beta\, g_\beta(x)\right) = 0, \qquad \text{for } \alpha = 1, \dots, n \quad \text{(E.11)}$$

$$g_\beta(x) = c_\beta, \qquad \text{for } \beta = 1, \dots, m \quad \text{(E.12)}$$

where $f(x)$ is the equation to be minimized, x_α are the parameters or coordinates of $f(x)$, $g_\beta(x) = c_\beta$ are the constraint equations, and λ_β are the Lagrangian multipliers associated with each constraint equation.

E.4 Derivation of OKG Variance

For the problem at hand, σ^2 (Eq. 2.21), which is a function of w_i and w_j, is minimized according to the constraints that the weights sum to 1 using the method of Lagrangian Multipliers (Sec. E.3). This results in four equations:

$$\frac{\partial}{\partial w_i}\left(\sigma^2(w_i, w_j) - \lambda_1 g_1(w_i) - \lambda_2 g_2(w_j)\right) = 0$$

$$\frac{\partial}{\partial w_j}\left(\sigma^2(w_i, w_j) - \lambda_1 g_1(w_i) - \lambda_2 g_2(w_j)\right) = 0$$

$$g_1(w_i) = \sum_{i=1}^{N} w_i = 1$$

$$g_2(w_j) = \sum_{j=1}^{N} w_j = 1.$$

Applying the partial derivatives, the first equation reduces to:

$$\left(-\sum_{j=1}^{N} w_j\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + 2\gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)]\right) - \lambda_1 + 0 = 0,$$

$$\sum_{j=1}^{N} w_j\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \lambda_1 = 2\gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)], \tag{E.13}$$

and the second equation becomes:

$$\left(-\sum_{i=1}^{N} w_i\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right]\right) + 0 - \lambda_2 = 0,$$

$$\sum_{i=1}^{N} w_i\, \gamma\left[Z(\mathbf{x}_i), Z(\mathbf{x}_j)\right] + \lambda_2 = 0. \tag{E.14}$$

It is then noted that $w_i = w_j$, and therefore $\lambda_1 = \lambda_2$, which allows Eqs. (2.33 and 2.34) to be combined:

$$2\sum_{j=1}^{N} w_j\,\gamma\left[Z(\mathbf{x}_i), Z\left(\mathbf{x}_j\right)\right] + 2\lambda = 2\gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)]$$

which leads to the system of OKG equations:

$$\begin{cases} \sum_{j=1}^{N} w_j\,\gamma\left[Z(\mathbf{x}_i), Z\left(\mathbf{x}_j\right)\right] + \lambda = \gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_0)], \quad \text{for } i = 1, \ldots \mathrm{N}, \\ \sum_{j=1}^{N} w_j = 1. \end{cases} \tag{E.15}$$

Edward T. Nykaza (VITA)

Experience

- 2011 – present, Principal Investigator, "Improved Military Noise Monitoring System."
- 2011 – present, Principal Investigator, "Characterization of the Instantaneous Sound Field Between Sensor Nodes via Passive Sensing of Blast Signatures."
- 2008 – present, Principal Investigator, "Community Attitudes Toward Blast Noise."
- 2006 - L.R. Shaffer Research and Development Achievement Award for Advances in Sleep Disturbance from Blast Noise Research.
- 2004 – present, Research Acoustical Engineer, ERDC-CERL, Champaign, IL.
- 2003, M.S. (Acoustics), Pennsylvania State University, University Park, PA, USA.
- 2001, B.A. (L.A.S.), University of Illinois, Champaign-Urbana, IL, USA.

Peer-Reviewed Publications

Nykaza, E. T., Hodgdon, K. K., Gaugler, T., Krecker, P. & Luz, G. A (**2013**). "On the relationship between blast noise complaints and community annoyance," J. Acoust. Soc. Am. **133**, 2690-2698.

Valente, D., Nykaza, E. T., Swift, H. S. & Luz, G. (**2013**). "Evolution of Metrics Used to Assess Community Response to Blast Noise," Military Operations Research **18**, 39-59.

Wilson, D. K., Valente, D., Nykaza, E. T. & Pettit, C. L. (**2013**). "Information-criterion based selection of models for community noise annoyance," J. Acoust. Soc. Am. **133**, EL195-EL201.

Wilson, D. K. *et al.* (**2012**). "Sound Outdoors and Noise Pollution," in *Handbook of Environmental Fluid Dynamics, Volume Two: Systems, Pollution, Modeling, and Measurements* (editor H.J. Fernando, Taylor & Francis Group).

Valente, D. *et al.* (**2012**). "Blast noise characteristics as a function of distance for temperate and desert climates," J. Acoust. Soc. Am. **132**, 216-227.

Cvengros, R. M., Valente, D., Nykaza, E. T. & Vipperman, J. S. (**2012**). "Blast noise classification with common sound level meter metrics," J. Acoust. Soc. Am. **132**, 822-831.

Nykaza, E. T., Pater, L. L., Melton, R. H. & Luz, G. A. (**2009**). "Minimizing sleep disturbance from blast noise producing training activities for residents living near a military installation," J. Acoust. Soc. Am. **125**, 175-184.

Nykaza, E. T., Pater, L. L. & Luz, G. A. (**2008**). "Improved procedure for correlating blast noise events with complaint logs at U.S. Army installations," Noise Control Engineering Journal **56**, 451-459.

Luz, G., Nykaza, E., Stewart, C. & Pater, L. (**2008**). "Use of actimeters to determine awakenings by sounds of large guns," Noise Control Engineering Journal **56**, 211-217.