Uncertainty assessment for outdoor sound propagation

O. Leroy (1), B. Gauvreau (1), F. Junker (2), E. de Rocquigny (3) and M. Bérengier (1)

(1) Laboratoire Central des Ponts et Chaussées, section Acoustique Routière et Urbaine, 44341 Bouguenais, France
(2) Electricité de France R&D, Département Analyses Mécaniques et Acoustique, 92141 Clamart, France
(3) Laboratoire Mathématiques Appliquées aux Systèmes, Ecole Centrale Paris, 92295 Châtenay-Malabry, France

PACS: 43.28.En, 43.28.Fp, 43.50.Rq

ABSTRACT

Intrinsic variability due to micrometeorological effects and/or ground effects, measurement uncertainty and model uncertainty are the main sources of spreading of the parameters influencing outdoor sound propagation. Thus spreading associated to outdoor SPL is a complex combination of deterministic, stochastic and epistemic uncertainties, and can be quantified thanks to a probabilistic process. This statistical process is presented in this paper and is called Calibration Under Uncertainty (CUU). Quantitative uncertainty assessment involves a pre-existing physical system to be studied, input data which can be measured or derived from measurements, and a sufficient amount of available (experimental and/or numerical) data with an eventual human expertise. CUU couples information from experimental and modelled data taking into account their own uncertainties (measurements errors, lack of knowledge on physical behavior, etc.) under specific assumptions. Quantify the global uncertainty on SPL, rank or apportion the contributions of influential parameters to a given output quantity of interest, compare experimental and effective parameters, and more generally understand the whole input-output structure are the main tasks of such a statistical method. CUU process has been applied to more or less complex cases using a large experimental set of data (Lannemezan 2005 (F)). An application to near ground sound propagation has been first led to understand the relative influence of ground parameters. A more complex case considering large distances and including micrometeorological effects has also been fulfilled with promising results which are presented in this paper.

INTRODUCTION

Transportation or industrial noise is known as a source of people annoyance. As noise regulations are more and more demanding, threshold values are harder and harder to be measured or calculated with the required accuracy. Moreover, using a single deterministic value is not sufficient to fully describe fluctuating acoustic levels for outdoor sound propagation. These levels mainly depend on variable micrometeorological and ground effects described by specific calculated or measured parameters. In a complete outdoor sound propagation study, these uncertain parameters constitute a set of input data for propagation models characterizing boundaries and medium characteristics. Uncertainty linked to output data is then a combination of input data uncertainty and model uncertainty.

In recent years, some methods have been developed to take into account variability in time and space. Fuzzy logic theory and multidimensional analysis have been applied to derive a classification of meteorological data regarding the effect on acoustic propagation (Zouboff et al. 1994). Geostatistical methods have been used to interpolate a set of experimental data, considering spatial coherence (Baume et al. 2006; 2009). Proper orthogonal decomposition has been applied to numerical results in order to develop compact representation of the sound field (Pettit and Wilson 2007). But a key issue is also to estimate the combined uncertainty that is to say a combination of model uncertainty, measurement uncertainty and physical variability in order to increase control of the prediction accuracy.

A probabilistic method called Calibration Under Uncertainty (CUU) has been proposed to estimate the combination of uncertainties (Leroy et al. 2008). This method is based on maximum likelihood estimation and integrates both experimental and numerical data and is a possible way to characterize and also quantify the combined uncertainty. It requires various and robust statistical tools such as metamodeling method (to replace numerical model during optimization process) or geostatistical tools. In this paper CUU is applied to data from the experimental campaign of Lannemezan 2005 (F) (Junker et al. 2006) and to a numerical propagation model (Parabolic Equation (Lihoreau et al. 2006)). The aim is to predict sound pressure level spectra and associated uncertainties for specific propagation conditions. Epistemic and stochastic uncertainties are estimated at two different distances from the source in order to show the effects of spatial variability.

DATASET

Preliminary

The dataset is constituted by numerical and experimental data. A response surface of the model (Kleijnen et al. 1997; 2005) is built in order to minimize CPU time during minimization process. This response surface also called metamodel is based on a specific design of numerical experiments. The robustness of CUU is better considering a set of data than considering a single data. Consideration of groups of data increases the rate of information and enhances the calibration process. A variance can be calculated for each estimated parameter (this variance is proportional to $n - 1$, where $n$ is the number of samples). It is then possible to describe global propagation conditions in terms of uncertainty. Experimental spectra representative of specific propagation conditions have then to be grouped.
Experimental data

Experimental campaign

Experimental data come from an experimental campaign which has been carried out in 2005 in Lannemezan (France) during three months. The experimental site has been chosen according to numerous considerations such as the quietness of the area, the ground flatness, the absence of obstacles between source and receivers, the soil type (grass land cover), etc. Experimental protocol is represented Figure 1 (Junker et al. 2006).

- A sound source is located far enough from the trees to avoid interactions with unwanted turbulent flows. The source height is fixed at 2m above the ground. The source is omnidirectional (12 loudspeakers) and is fed with a stationary pink noise (L_{eq} = 108dB(A)). Reference microphones were located at 10m from the source and at 2m above the ground. A total of about 50 microphones have been settled along four propagation directions (DP). The distance between the microphones is 25m long from 50m to 200m from the source. Each third octave band in the range [50Hz-4kHz] has been stored as continuous equivalent sound pressure levels with a one second integration time.
- Regarding micrometeorological devices, three dimensional ultrasonic anemometers and 10m high meteorological towers were placed respectively at 75m, 125m and 175m from the source in each propagation direction. An additional 60m high meteorological tower has been located at 200m northbound from the source and has been running during the 3 months.
- The ground impedance was monitored every four hours, on one fixed point during the whole campaign, using two microphones placed at 20m from the source (2m and 0.3m high). The effective airflow resistivity was determined by fitting the difference of the levels measured between the two microphones with a theoretical curve calculated using Delany-Bazley impedance model and an analytical propagation model (Piercy et al. 1977).

These experimental data have been filtered and validated with care before any analysis (Junker et al. 2005).

Figure 1: Experimental protocol for acoustical and micrometeorological devices. The black circled symbols indicate the devices which were let on site during the three months. Other devices were removed after three weeks of intensive measurements.

Figure 2: Experimental attenuation spectra measured at 50m and 150m relative to a microphone located at 10m from the source for homogeneous propagation conditions.

Figure 3: Histograms of measured vertical effective celerity gradients, turbulence index and effective airflow resistivity.

Numerical data

The sound propagation model used to apply the CUU is the Parabolic Equation model (PE model) (Lihoreau et al. 2006). This model enables to take into account micrometeorological phenomena and ground effects considering large distances of...
propagation. CUU requires numerous iterations and by the way numerous calculations with PE model. In order to minimize computer time during calibration process and to enable the application of CUU process with a numerical model such as PE model, interpolation between the discretization points of a calculation design is required. A solution is to create a metamodel, also called a response surface, auxiliary model, emulator, etc. Defined metamodels depend on the studied configuration; specific metamodels cannot be used in another configuration.

**Metamodel: basis**

Kleijnen et al. (Kleijnen et al. 2005) define a metamodel as an approximation of the true input/output function implicitly defined by the given simulation model. Obviously, a metamodel has to be simpler and faster than a numerical model. A metamodel treats the simulation model as a black box. It uses the input/output data without knowledge of the way the simulation model processes these inputs to get the outputs. Metamodeling is used for:

1. sensitivity analysis,
2. optimization,
3. prediction,
4. validation and verification.

There are many types of metamodels such as low-order polynomial regression models (the most popular) or exact interpolation by kriging (Den Hertog and Stehouwer 2002) (Simpson et al. 1998). The kriging method was developed in the field of geostatistics by Matheron in the 60’s and has been used in numerical design of experiments by Sacks (Sacks et al. 1989). Kriging has the advantage of being an interpolating method leading to the construction of a probabilistic model. This method will therefore depend on experimental points.

Kriging is used to build an interpolator which is a linear function of the observations taking into account the spatial structure of data. It provides a prediction as a linear combination of the observed values. The weights depend on the distance between the prediction point and the design of experiments through the chosen covariance function. In this paper, *kriging metamodel* is used. The study has been done thanks to the DACE toolbox (Lophaven et al. 2002).

**Application to Parabolic Equation model**

The predictive metamodel is based on a numerical design built from experimental parameters. The parameters are representative of the whole Lannemezan 2005 experimental campaign. A first design of experiments is built without turbulence and is based on effective airflow resistivity \( \sigma \) and on refraction parameters \( a_{log} \) and \( b_{lin} \). These last parameters are calculated from Monin-Obukhov scales and depend on stability conditions (Heimann et al. 2007, Stull 1988). Their combination allows to calculate vertical effective celerity such as \( \frac{\partial c}{\partial z} = a_{log} + b_{lin} \).

The design of experiments is a “full factorial design” such as:

- \( a_{log} = [-0.80 : 0.20 : 0.80] \text{ m.s}^{-1} \)
- \( b_{lin} = [-0.12 : 0.12 : 0.12] \text{ s}^{-1} \)
- and \( \sigma = [50 : 20 : 290] \text{ kN.s.m}^{-4} \).

Calculations taking into account turbulence being time consuming, it is necessary to reduce the numerical design. Another type of numerical design is defined: the Doehlert design (Sacks et al. 1989). Doehlert design is a rhombic network generated from a simplex and is based on a criterion of uniform distribution. The design in Table 1 is built to cover all the field of study. Calculations are made with a maximum value of turbulence. Kriging metamodels are then built for each frequency (from 100Hz to 2kHz) from the Doehlert design to reproduce the same datagrid as the numerical design without turbulence.

A linear interpolation between the numerical designs with and without turbulence allows to access to results with intermediate values of turbulence. Examples of response surfaces at 500Hz with a maximum turbulence effect are represented Figures 4 and 5 at 50 and 150m from the source (two meters high). Responses surfaces show interferential phenomena which are characteristic of this situation.

<table>
<thead>
<tr>
<th>N</th>
<th>( \sigma ) (kN.s.m(^{-4}))</th>
<th>( a_{log} ) (m.s(^{-1}))</th>
<th>( b_{lin} ) (s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>170</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>290</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>230</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>230</td>
<td>0.2</td>
<td>0.12</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>110</td>
<td>-0.8</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>110</td>
<td>-0.2</td>
<td>-0.12</td>
</tr>
<tr>
<td>8</td>
<td>230</td>
<td>-0.8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>230</td>
<td>-0.2</td>
<td>-0.12</td>
</tr>
<tr>
<td>10</td>
<td>170</td>
<td>0.4</td>
<td>-0.12</td>
</tr>
<tr>
<td>11</td>
<td>110</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>110</td>
<td>0.2</td>
<td>0.12</td>
</tr>
<tr>
<td>13</td>
<td>170</td>
<td>-0.4</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 1: Doehlert design for PE calculations with turbulence.

Figure 4: Response surface by kriging metamodeling at 500Hz with maximum value of turbulence at 50m from the source (Mean Squared Error = 0.04 dB). 

\( a_{log} \) and \( b_{lin} \) parameters led to an overparametrization of the system. In fact a preliminary application of CUU showed that one of the parameter allowed to calibrate the model while the other one reached the boundary values of parameters which are set for the optimization process. This strong dependence between the two parameters led to the overparametrization and the metamodel finally uses normalized turbulence index, vertical effective celerity gradients and effective airflow resistivity as physical input parameters. The final kriging models for each frequencies are based on exponential models of the variogram. Finally the calculation of a third octave spectrum (from 100 to 2000Hz) at 150m from a sound source taking into account turbulence with the surrogate model lasts 0.1second (with a standard PC). For comparison a PE model calculation time with the same configuration can reach several hours.
Application to ground and medium characteristics

Measured parameters are applied as input parameters of the surrogate model. Modeled and measured spectra are subtracted to constitute residues. Covariance structure of the residues is studied thanks to variogram analysis for data located at 50 and 150 m from the source.

From a sample to another, experimental variograms present a strong variability. In order to be representative of an average and characteristic situation, mean residues are considered. The experimental variogram is calculated from these residual values. At 50 and 150 m from the sound source (Figure 6), the most suitable variogram model is a gaussian model. This model is defined by equation 2:

$$\gamma(df) = C \left( 1 - \exp \left( - \frac{df}{a} \right)^2 \right),$$

(2)

where \(a\) is the sill i.e. the distance from which two observations are distinct on average; no linear link and null covariance. At this distance, variogram value corresponds to the variance of the stochastic variable. \(C = \sigma^2 - C_0\), \(\sigma^2\) corresponds to the range i.e. the variance of the stochastic variable and \(C_0\) corresponds the nugget variance.

- At 50 m from the source, \(C_0 = 4.9 \, dB^2\), \(\sigma^2 = 4930\).
- At 150 m from the source, \(C_0 = 5.4 \, dB^2\), \(\sigma^2 = 830\).

Regarding the values of variogram parameters and the bad results of the fitting, the model obviously need to be calibrated. Residual variances have to be reduced.

**UNCERTAINTY ASSESSMENT WITH WELL KNOWN INPUTS**

**Preliminary**

A first approach consists in considering measured parameters as well-known (i.e. fixed) and to model covariance matrix of residues between measured and modeled data. This modeling is made thanks to variogram analysis.

**Variogram analysis: basis**

In classical geostatistics, the standard summary of the second-moment structure of a spatial stochastic process is its variogram (Chilès and Delfiner 1999). The variogram of a stochastic process \(U(\cdot)\) is the function

$$\gamma(f, f') = \frac{1}{2} \text{var} \left[ U(f) - U(f') \right].$$

(1)

This function enables to obtain an experimental variogram; it depends on the frequent distance between two samples and also on the chosen direction (if the field has at least two dimensions). Fitting a parametric function to an empirical variogram provides one possible way to estimate covariance parameters. Frequently in practice this is done “by eye”, without a formal criterion. Alternatively, ordinary or weighted least mean squares methods for curve fitting are sometimes used. This model enables to transform discontinuous variograms in continuous variograms valid throughout the field of study. It is essential to understand the behavior of the variogram near the origin; it is linked to the continuity of variables \(U\). Typical behaviors are:

- A parabolic behavior which characterizes a highly regular variable usually differentiable.
- A linear behavior i.e. the variable is continuous but no longer differentiable.
- A discontinuity at the origin which characterizes a nugget effect. The variable is not continuous and very irregular
- A flat curve i.e. a pure nugget effect or a white noise (total absence of spatial structure).

These behaviors enable to explain and understand the spatial structure of the residue \(U\).
In order to avoid mathematical singularities and disease, the dimensionality of the cost function is reduced. Functional Principal Components Analysis (PCA) is used (Ramsay and Silverman 1997). Calculations are made on coefficients of functional PCA, explaining at least 95% of the variability of experimental spectra.

**CUU applied to single spectra**

**Estimation of parameters**

Estimated parameters after the calibration process are summarized in Table 2. These values are representative of the spatial variability and also of uncertainty on input parameters. Significant differences between estimated parameters at 50m and 150m appear. Estimated parameters can also be compared to measured parameters (Figure 3) usually applied as input parameters (ref previous section).

![Figure 7: Experimental variograms (circles) and pure nugget effect variograms (dotted lines) of residues at 50m and 150m from the source.](image)

**Table 2: Measured and estimated physical parameters at 50 and 150m considering independent spectra (mean and standard deviation).**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>50m</th>
<th>150m</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{dc}{dz}$</td>
<td>$0 \pm 0.01$</td>
<td>$-0.08 \pm 0.12$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$137 \pm 10$</td>
<td>$91 \pm 10$</td>
</tr>
<tr>
<td>$\mu_{norm}$</td>
<td>$0.37 \pm 0.2$</td>
<td>$0 \pm 10^{-4}$</td>
</tr>
</tbody>
</table>

**CUU applied to a set of spectra**

**Estimation of parameters**

When CUU is applied to a set of spectra, estimated parameters are representative of average propagation conditions; these parameters are summarized in Table 3. The use of an average parameter and of well-known parameters as inputs of the numerical model gives access to the epistemic uncertainty. Well known parameters such as temperature and relative humidity vary in time and this variability allows to explain a part of the epistemic uncertainty. A ratio between simultaneous and measured spectra is then calculated. Finally 30% (at 50m) and 25% (at 150m) of the spectra variability is explained thanks to estimated and well-known parameters.

Moreover estimated parameters do not report a strong spatial variability:

- homogeneous propagation conditions are well estimated at 50m and 150m.
- effective airflow resistivity values are similar. The difference between estimated parameters correspond to uncertainties of measurements (about 10kN.s.m$^{-2}$).
- normalized turbulence indices are similar regarding estimated standard deviations.

For vertical effective celerity gradients and turbulence indices, estimated standard deviations ($\sigma$) are larger at 150m than at 50m; it traduces a larger estimation uncertainty and a stronger effect of micrometeorological phenomena at 150m.

**Estimation uncertainty and combined uncertainty**

Estimation variances can be calculated generating a set of parameters thanks to Monte-Carlo runs and applying this set of parameters in the numerical model. Monte-Carlo runs are generated from a mean parameter and a standard deviation of this parameter; these values are respectively $\bar{\gamma}$ and $\sigma$. Estimation uncertainty allows to explain 38% (at 50m) and 34% (at 150m) of the spectra variability. Combined uncertainty is calculated from residual covariance matrices and from estimation
The covariance matrices of residues are estimated during the calibration process. This combined uncertainty is shown Figures 8 and 9 in comparison with experimental spectra.

**Figure 8:** Estimation and combined uncertainties associated to experimental spectra at 50m.

**Figure 9:** Estimation and combined uncertainties associated to experimental spectra at 150m.

For these specific propagation conditions, CUU process allows to determine a combined uncertainty representative of:

- physical variability,
- measurement uncertainties,
- modeling uncertainties,
- estimation uncertainty (calculated independently).

The major part of these uncertainties are stochastic and cannot be explained by the variance of the input parameters values. Moreover a skew is observed at 50m from the source at some frequencies. This is clearly an illustration of the limits of the model. A confident interval (95%) is calculated for each frequency which questions on the reliability of the parameters and/or on the efficiency of the numerical model.

The last approach considers one set of average parameters for a given group of spectra. Estimated parameters are representative of average conditions of propagation and allow to explain a part of combined uncertainty: epistemic uncertainty and uncertainty of estimation are clearly separated but unfortunately do not explain the major part of uncertainty. A part of his unexplained uncertainty is rejected into residual uncertainties. This study traduces the limitations of numerical models in specific propagation conditions to reproduce the variability (model seems to be more robust at 150m from the source) and is a first approach to improve its predictive content by using experimental data. Such an application will be applied in further works to a site with a complex topography and with a road traffic noise.

**REFERENCES**


