pISSN 2508-1888
eISSN 2466-2461
http://dx.doi.org/10.14407/jrpr.2016.41.2.149

Journal of
Radiation Protection and Research

Paper

Received July 17, 2015 / 1st Revised May 31, 2016 / Accepted June 13, 2016

A Methodology for Estimating the Uncertainty in Model Parameters Applying the Robust Bayesian Inferences

Joo Yeon Kim, Seung Hyun Lee, and Tai Jin Park
Korean Association for Radiation Application, Seoul, Republic of Korea

ABSTRACT

Background: Any real application of Bayesian inference must acknowledge that both prior distribution and likelihood function have only been specified as more or less convenient approximations to whatever the analyzer's true belief might be. If the inferences from the Bayesian analysis are to be trusted, it is important to determine that they are robust to such variations of prior and likelihood as might also be consistent with the analyzer's stated beliefs.

Materials and Methods: The robust Bayesian inference was applied to atmospheric dispersion assessment using Gaussian plume model. The scopes of contaminations were specified as the uncertainties of distribution type and parametric variability. The probabilistic distribution of model parameters was assumed to be contaminated as the symmetric unimodal and unimodal distributions. The distribution of the sector-averaged relative concentrations was then calculated by applying the contaminated priors to the model parameters.

Results and Discussion: The sector-averaged concentrations for stability class were compared by applying the symmetric unimodal and unimodal priors, respectively, as the contaminated one based on the class of ε-contamination. Though ε was assumed as 10%, the medians reflecting the symmetric unimodal priors were nearly approximated within 10% compared with ones reflecting the plausible ones. However, the medians reflecting the unimodal priors were approximated within 20% for a few downwind distances compared with ones reflecting the plausible ones.

Conclusion: The robustness has been answered by estimating how the results of the Bayesian inferences are robust to reasonable variations of the plausible priors. From these robust inferences, it is reasonable to apply the symmetric unimodal priors for analyzing the robustness of the Bayesian inferences.

Correspondence to Joo Yeon Kim
jykim@ri.or.kr

Keywords: Atmospheric dispersion, Robust Bayesian inference, ε-contamination, Uncertainty

1. INTRODUCTION

In the event of an atmospheric release of radioactive, chemical or biological materials, a timely transport and fate estimation which predicts the current and future locations and concentrations of the material in the atmosphere or deposited on the ground is important for consequence assessment and for deployment of emergency response actions. Such predictions help analysts to make time-critical decisions regarding precautions for their own safety, protective actions such as evacuation or sheltering of the peoples, and design of efficient field measurement plans. However, all of the model parameters are not necessarily known well in certain cases particularly in emergency situations where severe time constraints are imposed. In cases like a malignant activity of terrorists, even the source location may not be known.

Since the overall uncertainty for the atmospheric dispersion is determined by a comparison of modeling predictions with environmental measurements for conditions similar to those assumed by the model, it is necessary to obtain the site-specific data such as meteorological and topographical ones for analyzing the atmospheric dispersion. Unfortunately, not enough model validation studies have been performed to allow a reliable statistical analysis of the uncertainty associated with the Gaussian plume model. In this study, the robust Bayesian inference would be introduced as a new probabilistic approach for estimating the uncertainty analysis of the Gaussian plume model for the inputs in the model. And, a validation study of the Bayesian answer would be also tried.

2. MATERIALS AND METHODS

2.1 Definition of robust Bayesian inference

From the Bayes' theorem of Equation 1, we take prior beliefs about various possible hypotheses and then modify these prior beliefs in the light of relevant data which we have collected in order to arrive at posterior beliefs.

$$(posterior) \propto (prior) \times (likelihood) \qquad (1)$$

Any real application of Bayesian methods must acknowledge that both prior distribution and likelihood function have only been specified as more or less convenient approximations to whatever the analyzer's true belief might be. If the inferences from the Bayesian analysis are to be trusted, it is important to determine that they are robust to such variations of prior and likelihood as might also be consistent with the analyzer's stated beliefs. This variation of types or parameters in the prior and likelihood is defined as "contamination" in the robust Bayesian analysis [1, 2]. It has been known that the prior tends to be more contaminated than the likelihood, because the former is quantified by the subjective knowledge of the analyzer while the latter is constructed through the objective evidence.

An attractive idea, particularly for studying robustness with respect to the prior, is to elicit a plausible prior π_0 and, realizing that any prior "close" to π_0 would also be reasonable, choose Γ to consist of all such "close" priors [1, 2]. A rich and attractive class to work with is that of ε-contamination

$$\Gamma = \{\pi = (1-\varepsilon)\pi_0 + \varepsilon q : q \in Q\}, \qquad (2)$$

where ε determines the amount of probabilistic deviation from the plausible prior π_0 that is allowed and $0<\varepsilon<1$ reflects how "close" is π to π_0. Q is a class of possible contaminations.

A natural goal of a robustness investigation is to find the variability of the posterior quantities, such as the posterior mean, variance and credible set, as π varies over Γ. If the variability of the posterior quantity is small, then one can be assured of robustness with respect to the elicitation process. If the variability is large, one does not have robustness with respective to Γ allowing for further investigation or refinement of the contaminated prior.

Unfortunately, the variability of the posterior quantity of interest will often be excessively large when Q = {all distributions} is used, because this Q contains many unreasonable distributions. Indeed, it is argued therein that more reasonable Q are the class of unimodal (U) distributions and the class of all symmetric unimodal (SU) ones. The contaminated prior taking a uniform type will be a pretty sensible prior and the class of uniform type does contain priors with effective tails substantially larger than π_0 [2].

2.2 Gaussian plume model

The Gaussian plume model has been used as one suggested in the Canadian Standard Association (CSA). This standard provides a modified Gaussian plume model to evaluate the time-integrated concentration at downwind distances from 100 m to 100 km

Corresponding author : Joo Yeon Kim, jykim@ri.or.kr
18F, Seoul-forest IT Valley, 77, Seongsuil-ro, Seongdong-gu, Seoul, 04790, Korea

 http://dx.doi.org/10.14407/jrpr.2016.41.2.149

and makes reference to the Pasquill atmospheric stability categories A to F for the purpose of calculation.

The sector-averaged relative concentrations, $(\chi_k/Q)_S$, for the centerline of the plume at ground level are obtained by setting y and z equal zero and assuming the height of the capping inversion to be much greater than H, the equation for calculating them is given by

$$\left[\frac{\chi_k(x,y,0)}{Q}\right]_S = \frac{1}{\pi u_k \Sigma_y \Sigma_z} \exp\left(-\frac{H^2}{2\Sigma_z^2}\right) f(\Sigma_z, H, h_i) \ \ (s/m^3), \tag{3}$$

2.3 Uncertainty in model parameters

Model parameters to be uncertain are assumed as dispersion coefficients and wind speed. Their distributions were assumed as log-normal ones for the prior and the likelihood by Hamby [3]. The means of the priors of two dispersion coefficients were assumed as the values calculated from the equations of the dispersion coefficients recommended in CSA model. Their standard deviations were assumed as 10% of the midpoint values determined by logarithmic interpolation between the Pasquill curves of E and F-stabilities at a given downwind distance in this study.

The uncertainties for each input are expected to be within a factor of 2 for vertical dispersion coefficient, a factor of 0.5 to 0.6 for horizontal dispersion coefficient according to Turner, if the uncertainty for the ground level centerline concentration is divided by the inputs in the model [4]. From the stated information of uncertainties of the inputs, the standard deviations of the likelihoods of horizontal and vertical dispersion coefficients were set as 50% and 200% of the midpoints of dispersion coefficients at a given downwind distance calculated in CSA model, respectively.

While the prior and the likelihood for wind speed were constructed by examining the data measured in real-time in the reference unit during 2004 to 2005 year. Since the data of wind speed were classified into A to G-stability by U.S. NRC [5], the data of wind speed consistent with G-stability were added to the data of F-stability for applying to the CSA model in this study. The data measured in 2004 and 2005 years were assumed as the prior and the likelihood of wind speed, respectively. The priors and likelihoods of all inputs were then transformed to normal distributions by taking natural logarithm for writing a program for calculating the sector-averaged concentrations by WinBUGS 1.4.1. Iterative simulations of 5500 were performed for calculating the concentrations with the addition of the burn-in of 500. In Table 1 and Table 2, the summaries of the priors and likelihoods are presented for dispersion coefficients and wind speed, respectively.

Table 1. The Summary of the Prior Distributions and Likelihood Functions for Dispersion Coefficients (F-stability)

Distance (m)	Input	Plausible priors	Likelihood	Contaminated priors		
				SU	U_1	U_2
1000	Σ_y	$N(3.64,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[3.55, 3.73]	U[3.55, 3.64]	U[3.64, 3.73]
	Σ_z	$N(2.69,\ 0.08^2)$	$N(prior,\ 1.11^2)$	U[2.61, 2.78]	U[2.61, 2.70]	U[2.70, 2.78]
2000	Σ_y	$N(4.29,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[4.19, 4.29]	U[4.29, 4.38]	U[4.19, 4.29]
	Σ_z	$N(3.13,\ 0.08^2)$	$N(prior,\ 1.10^2)$	U[3.05, 3.22]	U[3.05, 3.14]	U[3.14, 3.22]
3000	Σ_y	$N(4.65,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[4.56, 4.75]	U[4.56, 4.65]	U[4.65, 4.75]
	Σ_z	$N(3.38,\ 0.08^2)$	$N(prior,\ 1.10^2)$	U[3.30, 3.47]	U[3.30, 3.38]	U[3.38, 3.47]
4000	Σ_y	$N(4.90,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[4.81, 5.00]	U[4.81, 4.90]	U[4.90, 5.00]
	Σ_z	$N(3.54,\ 0.08^2)$	$N(prior,\ 1.09^2)$	U[3.47, 3.64]	U[3.47, 3.55]	U[3.55, 3.64]
5000	Σ_y	$N(5.09,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.00, 5.18]	U[5.00, 5.09]	U[5.09, 5.18]
	Σ_z	$N(3.67,\ 0.08^2)$	$N(prior,\ 1.09^2)$	U[3.59, 3.76]	U[3.59, 3.68]	U[3.68, 3.76]
6000	Σ_y	$N(5.24,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.15, 5.33]	U[5.15, 5.24]	U[5.24, 5.33]
	Σ_z	$N(3.77,\ 0.08^2)$	$N(prior,\ 1.09^2)$	U[3.69, 3.86]	U[3.69, 3.78]	U[3.78, 3.86]
7000	Σ_y	$N(5.37,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.27, 5.46]	U[5.27, 5.37]	U[5.37, 5.46]
	Σ_z	$N(3.85,\ 0.07^2)$	$N(prior,\ 1.09^2)$	U[3.78, 3.94]	U[3.78, 3.86]	U[3.86, 3.94]
8000	Σ_y	$N(5.47,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.38, 5.56]	U[5.38, 5.47]	U[5.47, 5.56]
	Σ_z	$N(3.92,\ 0.07^2)$	$N(prior,\ 1.08^2)$	U[3.84, 4.01]	U[3.84, 3.93]	U[3.93, 4.01]
9000	Σ_y	$N(5.56,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.47, 5.65]	U[5.47, 5.56]	U[5.56, 5.65]
	Σ_z	$N(3.98,\ 0.07^2)$	$N(prior,\ 1.08^2)$	U[3.90, 4.07]	U[3.90, 3.99]	U[3.99, 4.07]
10000	Σ_y	$N(5.64,\ 0.08^2)$	$N(prior,\ 0.40^2)$	U[5.55, 5.73]	U[5.55, 5.64]	U[5.64, 5.73]
	Σ_z	$N(4.03,\ 0.07^2)$	$N(prior,\ 1.08^2)$	U[3.96, 4.12]	U[3.96, 4.04]	U[4.04, 4.12]

Table 2. The Summary of the Prior Distributions and Likelihood Functions for Wind Speed (F-stability)

Sector	Plausible priors	Likelihoods	Contaminated priors		
			SU	U_1	U_2
N	$N(-0.86,\ 0.42^2)$	$N(prior,\ 0.96^2)$	U[-1.32, -0.40]	U[-1.32, -0.86]	U[-0.96, -0.40]
NNE	$N(-1.05,\ 0.48^2)$	$N(prior,\ 0.85^2)$	U[-1.58, -0.52]	U[-1.58, -1.05]	U[-1.05, -0.52]
NE	$N(-1.00,\ 0.31^2)$	$N(prior,\ 0.76^2)$	U[-1.34, -0.66]	U[-1.34, -1.00]	U[-1.00, -0.66]
ENE	$N(-1.29,\ 0.60^2)$	$N(prior,\ 0.76^2)$	U[-1.95, -0.62]	U[-1.95, -1.29]	U[-1.29, -0.62]
E	$N(-0.51,\ 0.41^2)$	$N(prior,\ 0.87^2)$	U[-0.96, -0.06]	U[-0.96, -0.51]	U[-0.51, -0.06]
ESE	$N(-0.58,\ 0.47^2)$	$N(prior,\ 0.89^2)$	U[-1.10, -0.06]	U[-1.10, -0.58]	U[-0.58, -0.06]
SE	$N(-0.93,\ 0.55^2)$	$N(prior,\ 0.54^2)$	U[-1.53, -0.32]	U[-1.53, -0.93]	U[-0.93, -0.32]
SSE	$N(-1.11,\ 0.54^2)$	$N(prior,\ 0.64^2)$	U[-1.70, -0.52]	U[-1.70, -1.11]	U[-1.11, -0.52]
S	$N(-0.72,\ 0.68^2)$	$N(prior,\ 0.56^2)$	U[-1.47, 0.03]	U[-1.47, -0.72]	U[-0.72, 0.03]
SSW	$N(-0.78,\ 0.67^2)$	$N(prior,\ 0.56^2)$	U[-1.52, -0.03]	U[-1.52, -0.78]	U[-0.78, -0.03]
SW	$N(-0.46,\ 0.55^2)$	$N(prior,\ 0.52^2)$	U[-1.06, 0.15]	U[-1.06, -0.46]	U[-0.46, 0.15]
WSW	$N(-0.44,\ 0.48^2)$	$N(prior,\ 0.48^2)$	U[-0.96, 0.08]	U[-0.96, -0.44]	U[-0.44, 0.08]
W	$N(-0.17,\ 0.36^2)$	$N(prior,\ 0.46^2)$	U[-0.57, 0.23]	U[-0.57, -0.17]	U[-0.17, 0.23]
WNW	$N(-0.04,\ 0.34^2)$	$N(prior,\ 0.57^2)$	U[-0.42, 0.33]	U[-0.42, -0.04]	U[-0.04, 0.33]
NW	$N(-0.31,\ 0.40^2)$	$N(prior,\ 0.79^2)$	U[-0.75, 0.13]	U[-0.75, -0.31]	U[-0.31, 0.13]
NNW	$N(-0.63,\ 0.38^2)$	$N(prior,\ 0.83^2)$	U[-1.04, -0.21]	U[-1.04, -0.63]	U[-0.63, -0.21]

3. RESULTS AND DISCUSSION

3.1 Procedures for analyzing robustness

An answer for a question has been attempted by applying the robust Bayesian analysis to the Gaussian model based on the procedures of a robustness: how robust is the sector-averaged relative concentration regardless of the contamination of the prior? This question is answered through the following steps:

(1) quantify the plausible priors of the parameters of three inputs, where the plausible priors means to be the information of the priors;
(2) obtain the contaminated priors of the stated parameters;
(3) derive the classes of ε-contamination of all priors by applying the plausible and contaminated priors of the parameters;
(4) calculate the sector-averaged concentrations by applying the classes of ε-contamination of all priors; and
(5) compares the relative errors of medians of the concentrations based on the plausible and contaminated priors.

The scopes of contaminations were specified as the uncertainties of distribution type and parametric variability. The distribution was assumed to be contaminated as uniform one, and the symmetric unimodal and unimodal distributions were then applied for the robust Bayesian analysis. Since uniform distribution is defined as a vague prior one considering objectivity, which means a non-informative prior, in the Bayesian analysis, it is likely to be the most suitable type as the contaminated priors. The classes of ε -contamination of the symmetric unimodal and unimodal distributions were given by Equations 4 and 5 [6].

$$\Gamma_{su} = \{\pi = (1-\varepsilon)\pi_0 + \varepsilon q : q \ is \ U[\mu-k, \mu+k] \ for \ k>0\}, \tag{4}$$

$$\Gamma_{u} = \{\pi = (1-\varepsilon)\pi_0 + \varepsilon q : q \ is \ U[\mu, \mu+k] \ or \ U[\mu-k, \mu], k>0\}, \tag{5}$$

where, μ is the mean of plausible priors and k is the lower or upper range consistent with the standard deviations of the contaminated priors. By assuming ε as 10%, k was determined to take a more contamination of 10 % than the standard deviations of the plausible ones in this study. From this setting of k, it has been achieved that the classes of Eq. 4 and 5 do contain priors with effective tails substantially larger than π_0. For easy calculation, unimodal contamination of Eq. 5 is divided as two ranges. That is, U_1 and U_2 contaminations are assumed that their unimodal ones are equal to $U[\mu-k,\ \mu]$ and $U[\mu,\ \mu+k]$.

3.2 Robustness

The distribution of the sector-averaged relative concentrations was derived based on the symmetric unim-

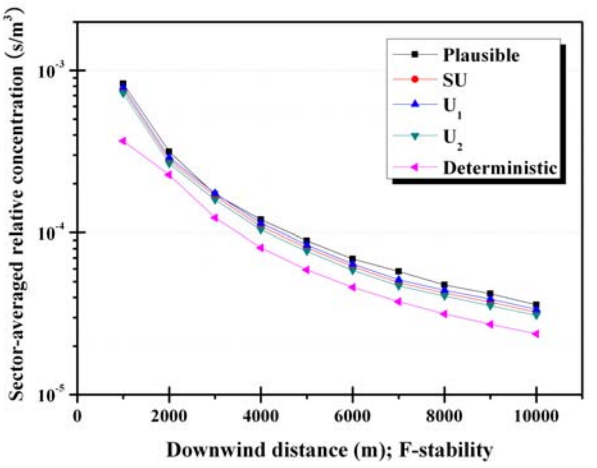

Fig. 1. The comparison of the statistical percentiles of the sector-averaged relative concentration considering the contaminated priors.

Table 3. The Relative Errors of the Medians Under F-stability Class (%)

Distance (m)	SU	U_1	U_2
1000	8.13	4.88	12.75
2000	11.36	7.61	15.74
3000	2.73	1.04	7.54
4000	9.41	5.50	13.74
5000	9.18	5.60	13.61
6000	10.16	7.00	14.95
7000	14.62	11.07	18.46
8000	10.98	7.16	14.64
9000	11.15	7.22	15.54
10000	9.74	6.18	13.86

odal and unimodal contaminated priors defined as uniform distribution. The mean of the symmetric unimodal distribution was fixed as the mean, μ, of the plausible prior for protecting a variability of its centering part. The value of lower and upper regions of the unimodal prior was set as [μ-k, μ] or [μ, μ+k] and the unimodal prior was biased to the left or right by centering μ as compared with the symmetric unimodal one. [μ-k, μ] and [μ, μ+k] contaminations are assumed as U_1 and U_2 contaminations, respectively.

The statistical percentiles of the sector-averaged relative concentrations were calculated by applying the symmetric unimodal and unimodal priors by stability class and downwind distance. The results for F-class were summarized as the examples in Figure 1, and the medians were also then compared with those of concentrations reflecting the plausible priors in Table 3. The results of the concentrations reflecting the unimodal prior taking the region of [μ-k, μ] or [μ, μ+k] was biased to the right or left as compared with that of concentrations reflecting the symmetric unimodal one. It is noted that the parametric values sampled from the unimodal prior are smaller or larger than ones sampled from the symmetric unimodal one due to the lower or upper range of its distribution set as μ-k or μ+k from μ without symmetry. The distribution of concentrations based on the unimodal prior would be, therefore, biased to the right or left as compared with that of concentrations based on the symmetric unimodal one.

The relative errors of the medians using the unimodal prior of [μ-k, μ] were varied about 1 to 11.1 % for F-class as compared with those of ones using the plausible one, though ε is assumed as 10 % in this study. It was then expected that there was more robust in the results of the concentrations based on the unimodal prior of [μ-k, μ] than the symmetric unimodal one because of decrease of their error ranges. However, the relative errors of the medians using the unimodal prior of [μ, μ+k] were varied about 7.5 to 18.5 % for F-class as compared with those of ones using the plausible one, though ε is assumed as 10 % in this study. It was then expected that the results of the concentrations based on the unimodal prior of [μ, μ+k] was less robust than the symmetric unimodal one and unimodal one of [μ-k, μ].

4. CONCLUSION

An answer for the uncertain question has been attempted by applying the robust Bayesian analysis to the Gaussian model based on the procedures of the robustness: how robust is the sector-averaged relative concentration regardless of the contamination of the prior? The robustness was analyzed for estimating how the results of the Bayesian inferences are robust to reasonable variations of the plausible priors. The sector-averaged concentrations for stability class were compared by applying the symmetric unimodal and unimodal priors, respectively, as the contaminated one based on the class of ε-contamination. Though ε was assumed as 10 %, the medians reflecting the symmetric unimodal priors were nearly approximated within 10 % compared with ones reflecting the plausible ones. However, the medians reflecting the unimodal priors were approximated within 20 % for a few downwind distances compared with ones reflecting the plausible ones. From these robust inferences, it is reasonable to apply the symmetric unimodal priors for analyzing the robustness of the Bayesian inferences.

ACKNOWLEDGEMENTS

This work was supported by the Nuclear Power Core Technology Development Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20131520100770).

REFERENCES

1. Moreno E, Cano JA. Robust Bayesian analysis with ε-contamination partially known. J. Roy. Stat. Soc. B Stat. Meth. 1991;53(1):143-155.
2. Sivaganesan S, Berger JO. Ranges of posterior measures for priors with unimodal contaminations. Ann. Stat. 1989;17(2):868-889.
3. Hamby DM, The Gaussian atmospheric transport model and its sensitivity to the joint frequency distribution and parametric variability. Health Phys. 2002 Jan;82(1):64-73.
4. Turner DB. Workbook of atmospheric dispersion estimates: An introduction to dispersion modeling. 2nd Ed. New York NY. CRC Press. 1994;2.13–2.15.
5. U.S. Nuclear Regulatory Commission. Meteorological monitoring programs for nuclear power plants. Regulatory Guide 1.23. 2007;6-8.
6. Berger J, Berkiner LM. Robust Bayes and empirical Bayes analysis with ε-contaminated priors. Ann. Stat. 1986 Jun;14(2):461-486.