Proceedings of the 2008 IEEE
International Conference on Information and Automation
June 20 -23, 2008, Zhangjiajie, China

Probabilistic Inversion Techniques in Quantitative Risk Assessment for Power System Load Forecasting

Xiaofeng Wang, Chao Du, Zuoliang Cao
School of Mechanical Engineering
Tianjin University of Technology
263 Hongqinan Rd, NanKai District, Tianjin, P.R. China
wangxiaofeng@tjut.edu.cn, unicorn@vip.sina.com, zlcao@126.com

Abstract—**Expert judgment is frequently used to assess parameter values in quantitative risk assessment. Experts can however only be expected to assess observable quantities, not abstract model parameters. This means that we need a method for translating expert assessed uncertainties on model outputs into uncertainties on model parameter values. So we use Probabilistic Inversion (PI) method. The probability distribution on model parameters obtained in this way can be used in a variety of ways, but in particular in an uncertainty analysis or as a Bayes prior. In this paper probabilistic inversion problems are first defined, existing algorithms for solving such problems are also discussed and the algorithms based on iterative algorithms are introduced. Those computational algorithms have proven successful in various projects. Such techniques are indicated when we wish to quantify a model which is new and perhaps unfamiliar to the expert community. There are no measurements for estimating model parameters, and experts are typically unable to give a considered judgment. In such cases, experts are asked to quantify their uncertainty regarding variables which can be predicted by the model. Applications to power system load forecasting in NingXia province of China is discussed. This study illustrates two such techniques, Iterative Proportional Fitting (IPF) and PARmeter Fitting for Uncertain Models (PARFUM) which provide useful tools for the practicing quantities risk assessment. In addition, we also illustrate how expert judgment on predicted observable quantities in combination with probabilistic inversion may be used for model validation.**

Keywords—**Probabilistic inversion (PI), IPF, PARFUM, Probabilistic risk analysis, Load forecasting, Coincidence Factor**

I. INTRODUCTION

Probabilistic risk analysis (PRA), also called quantitative risk analysis (QRA) or probabilistic safety analysis (PSA), is currently being widely applied to many sectors, including transport, energy, chemical processing, aerospace, and military[1]. PRA is a technique that seeks to define and quantify the probability that and adverse event will occur [2]. The role of risk assessment in structuring and quantifying information and uncertainty supporting decision-making in governmental agencies and companies has grown steadily during the last decades. In particular, decision-making related to the operation of hazardous socio-technological systems has adopted risk assessment as an approach to evaluate the acceptability of operating a system, and risk as a measure to compare the performance of systems. In fact risk assessment is not procedurally similar in the different application areas, and varies with respect to the use of experience data, expert judgment, risk modeling, decision rules and criteria. Furthermore, even in an industry sector, risk assessments differ complicating risk comparison and risk communication of assessment results.

Uncertainty analysis has been a standard procedure in probabilistic risk analysis since the Reactor Safety Study. The following problem was considered in uncertainty analysis: place a joint distribution on the parameters of a model in such a way that the "push forward" distribution on model output reflects the uncertainty of the phenomena being modeled.

In the ideal case of uncertainty analysis, the model's parameters have a clear physical meaning, data on the values of these parameters is available, and joint uncertainty distributions on the parameter values can be obtained from data. Unfortunately, parameters often do not possess a transparent physical meaning, data is not available, and the uncertainty analyst must resort to experts' subjective probability judgment.

Model parameters, however, may not correspond to transparent physical measurements with which the experts are familiar, and the uncertainty analyst must resort to experts' subjective probability judgments. In this case, Probabilistic Inversion will be used to "pull back" subjective probability distributions to abstract modeling parameters.

There exist few algorithms to solve the probabilistic inversions, namely: conditional sampling [6], PREJUDICE [6], PARFUM (PARameter Fitting for Uncertain Models) [3] and IPF (Iterative Proportional Fitting) [4].

Elicitation variables and target variables:

The unknown model parameters for which we wish to find a joint distribution are called target variables. Elicitation variables are subjective physically meaningful quantities with which the experts are familiar. Elicitation variables can be measured by procedures in which experts quantify their uncertainty on the basis of their expertise.

In mathematical prose, probabilistic inversion problems arise in quantifying uncertainty in physical models with expert judgment. The uncertain parameters X to quantify do not possess a clear physical meaning and are not associated with physical measurements with which experts are familiar. We must then find the observable quantities Y functionally related with X. Extracting uncertainties of X from uncertainties of Y specified by experts is clearly an inverse problem.

Given a joint distribution for: $Y = [Y_1, Y_2, Y_3, ..., Y_n]$ and the functions are $G_i : R^m \rightarrow R, i = 1,2,.....n$, find joint distributions for $X = [X_1, X_2, X_3, ..., X_m]$ such that $G(X) = [G_1(X), G_2(X), G_3(X), ..., G_n(X)] \sim Y$: where ~ means has the same distribution as. Note that this problem may be infeasible or if it is feasible it may have more than one solution. Therefore, we must have some method of selecting a preferred distribution in case of non-uniqueness and some method of selecting a best fitting distribution in case of non-existence.

II. PRA FLOW

The goal of a PRA for cyber security is to identify the potential threats and system vulnerabilities, quantify the likelihood of those threats and produce mitigation strategies based on both the risk and associated costs. The PRA process describes the sequence of risk activities which represent in Figure 1. The process can be described in the following three steps:

- To Identify of project risk variables
- To perform a Monte Carlo simulation
- To interpret results of the analysis and to make a decision regarding acceptable risk. This process may be iterated if desired by the site team.

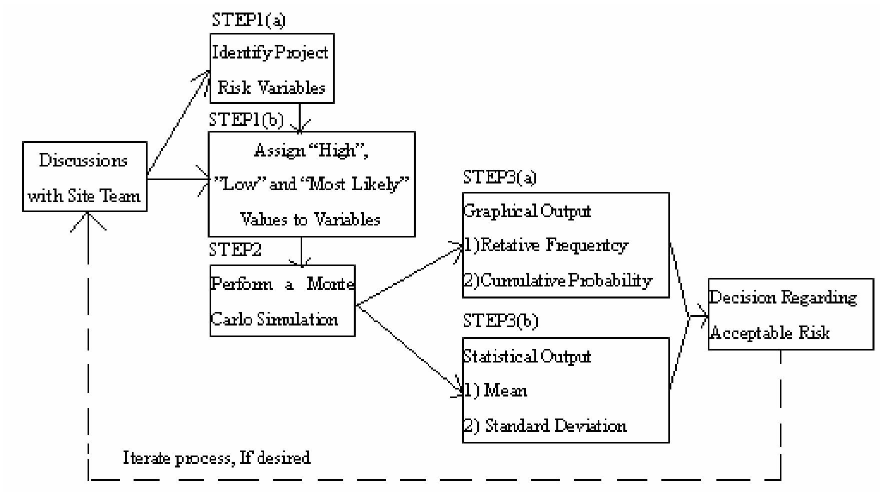

Figure 1. PRA process

The first step in the PRA process is to assess risk, measure the probability of cost and schedule overrun or under run by identifying project variables which are expected to vary greatly. These variables have especially large and volatile ranges, hence, much uncertainty. A number of methods have been used in identifying risks to an organization. Most techniques focus on enumeration of initiating events through familiarization with the system and its potential vulnerabilities. For safety related systems, FEMA, HAZOPS, and PHA have been used to identify events [5,6].

The second stage of PRA is to conduct Monte Carlo simulation on the total project budget by the use of software, including the risk variables which have been identified and "ranged" as described in the paragraph above. Monte Carlo simulation uses the selected, user-defined probability distributions of the identified project risk variables to perform random modeling; that is, given the unique distribution of each project risk variable, the simulation produces repeated variable values ("x") by simulating or "performing" many random repetitions or trials. Total cost is defined as a function of various random variables, "x." Each time a set of "x's" are randomly generated in the simulation process, a value for total cost is calculated. Once the simulation is complete, a distribution for total project cost is obtained. The precision of the approximation improves as the number of simulation trials increases. Monte Carlo simulation is thus able to replicate real-life occurrences, in its ability to "model" projected events and generate an expected value for the objective function (e.g., total cost or total schedule) under study. For example, the simulation results may indicate an 83% likelihood (relatively high probability) that a project will need to use its contingency reserve in order to avoid a funding overrun; or, results may show a 65% probability (relatively low) that the project will be complete within the budgeted schedule duration.

The final stage of PRA is interpretative and need to be iterated, or run again. The results of this stage may be used to generate or "feed" another round of analysis beginning with a fresh look at stage one, i.e., re-examining the triangular or range input values for the critical project risk variables which are selected by the project site team. The interpretative stage of the probabilistic method relies on graphical tools such as histograms (relative frequency polygons), or gives (cumulative frequency polygons), and may include "tornado graphs" which describe calculated sensitivities for critical variables generated from the information which is produced as a result of the Monte Carlo simulation. Visual output is quite softwaredependent, but it typically includes probability density functions and statistical parameters such as expected value (mean), standard deviation, upper and lower limits, and graphic confidence intervals for a given region (probability) expressed as a percent between 0% and 100%.

In general, a PRA is conducted in three stages: Risk identification, risk quantification, and risk evaluation and acceptance (Table 1).

TABLE I. THREE STAGES OF PRA

Stage	Question	Actions
1 Risk Identification	What can go wrong?	Identify risk source
2 Risk Quantification	What is the likelihood it would go wrong? What are the consequences?	Assess probabilities subjective, or objective Model causal relation-ships and their impacts
3 Risk Evaluation and Acceptance	What can be done? What are the options and trade-offs?	Create policy options Trade-off analysis of risk and cost/benefits of mitigation

III. THE ALGORITHM IPF AND PARFUM

IPF and PARFUM are the most recent and popular iterative algorithms applied to solve the probabilistic problem.

A. IPF algorithms

IPF is known from the literature as Iterative Proportional Fitting. The IPF algorithm projects a starting measure onto the set with fixed first margins (Ω_1) then project this onto the set with fixed second margins (Ω_2), then project this again onto the set Ω_1 and so on. Hence is we have arrived at vector p by projecting onto Ω_1, the next iteration is

$$p^{'} = p^{\Omega_2}$$

(1)

This algorithm was first used to estimate cell probabilities in a contingency table subject to certain marginal constraints. It is easy to see that if p satisfies the constraints, thus p is a fixed point of this algorithm.

The convergence of the IPF algorithm has been studied by many authors. Csiszar shows that if the IPF algorithm converges, then it converges to the I-projection of the starting probability vector on the set of probability vectors satisfying the constraints. He further showed that starting with a probability vector p, IPF converges if there is a vector r satisfying the constraints and having zeros in those cells where p is zero

$$p_{ij} = 0 \Rightarrow r_{ij} = 0, i, j = 1,..., K \tag{2}$$

For the two dimensional case the IPF algorithm is equivalent to an alternating minimization procedure studied in [7]. It is proven there that if sequences{P_n} and {Q_n} from Ω_1 and Ω_2, respectively, are obtained by alternating minimization of I($P_n|Q_n$) with respect to P_n respect Q_n, then I($P_n|Q_n$) converges to the infimum of I(P|Q) on $\Omega'_1 * \Omega_2$, where Ω'_1 is a set of all $P \in \Omega_1$ such that I(P|Q_n)<∞ for some n. Moreover, the convergence of the sequences {P_n} and {Q_n} is proven. Csiszar and Tusnady [7] proved this result for a general case where the sets Ω_1 and Ω_1 are convex sets of finite measures and the function that is minimized in alternating minimization procedure is an extended real valued function.

B. PARFUM algorithm

A variation on this is an iterative version of the PARFUM algorithm [6]. It is shown that this algorithm always converges, and that if the problem is feasible, it converges to a solution that minimizes a certain information function.

Let $Y = [Y_1, Y_2, Y_3, ..., Y_m]$ be a random vector with densities $[f_1, f_2, f_3, ..., f_m]$ and let $Gm : R^n \rightarrow R, m = 1,2,.....M$ be measurable functions. The PARFUM algorithm can be described in the following steps [8-10]:

- Choose a finite set $\chi \subset R^n$
- Define the conditional mass function Q_m of Y_m on the image $G_m(\chi)$ of χ under G_m, Where $x \in \chi$:

$$Q_m(G_m(x)) = \frac{f_m(G_m(x))}{\sum_{G_m(Z)} f_m(G_m(Z))}$$

- Define the minimally informative distribution on χ whose push-forward distribution P_m on $G_m(x)$ agree with Q_m, that is, for $x \in \chi$

$$P_m(x) = \frac{Q_m(G_m(x))}{\#\{Z \in \chi \mid G_m(Z) = G_m(x)\}}$$

where # means numbers of points.

- Find a distribution P on χ which minimizes the relative information $\sum_{m=1}^{M} I(P_m \mid P)$, where

$$I(P_m \mid P) = \sum_{X \in \chi} P_m(x) \ln(\frac{p_m(x)}{p(x)})$$

Let $S^K = \{r \in R^K \mid r_k \geq 0, \sum_{k=1}^{K} r_k = 1\}$

Proposition: Let $P_m \in S^K, m = 1,..., M$ Then

$$\min_{P\in S^K}\sum_{m=1}^{M} I(P_m \mid P) = \sum_{m=1}^{M} I(P_m \mid P^*)$$

If and only if $P^* = (1/M)\sum_{m=1}^{M} P_m$.

The advantage of this method is that it is always feasible and easily implemented. One disadvantage is that the conditional distributions Q_m might be different to those of Y_m, but this may be steered by the right choice of χ. More serious is the fact that the push forward of P need not have marginal that agree with the Y_m. This also can be influenced by steering, but is more difficult.

C. Examples

Starting from a given sample distribution, we may consider the starting point of the IPF algorithm as the uniform distribution over the sample points. IPF iteratively re-weights these sample points. IPF need not converge, but if it converges, it converges to a solution that is minimally informative with respect to the starting distribution.

We start with an initial distribution over X and generate an initial distribution over the interquantile cells, which we represent in Table 2. The marginal are shown in boldface.

TABLE II. INITIAL DISTRIBUTION OVER INTERQUANTILE CELLS

Starting distribution and marginal distributions			
0.4	0.1	0.1	0.1
0.3	0.1	0.1	0.1
0.3	0.1	0.1	0.2
1	**0.4**	**0.3**	**0.3**

1st Row projection			**2nd Column projection**		
0.169977	0.127483	0.10254	0.172973	0.12973	0.104348
0.127483	0.095612	0.076905	0.12973	0.097297	0.078261
0.101471	0.076103	0.122426	0.097297	0.072973	0.117391

3rd Row projection			**4th Column projection**		
0.170364	0.127773	0.101863	0.170433	0.127825	0.101904
0.127773	0.09583	0.076397	0.0127825	0.095868	0.076428
0.101839	0.076379	0.121782	0.101743	0.076307	0.121667

Repeating these row-column projections, we get the solution of IPF:

$$\begin{bmatrix} 0.170373 & 0.12778 & 0.101847 \\ 0.12778 & 0.095835 & 0.076386 \\ 0.101847 & 0.076386 & 0.121767 \end{bmatrix}$$

In order to compare these two algorithms, we use the same example which we represent in Table 2.

1st Row projection			**1st Column projection**		
0.133333	0.133333	0.133333	0.133333	0.1	0.075
0.1	0.1	0.1	0.133333	0.1	0.075
0.075	0.075	0.15	0.133333	0.1	0.15

1st result of PARFUM		
0.133333	0.116667	0.104167
0.11667	0.1	0.0875
0.104167	0.0875	0.15

2nd Row projection			**2nd Column projection**		
0.150588	0.131765	0.117647	0.150588	0.115068	0.091463
0.115068	0.09863	0.086301	0.131765	0.09863	0.076829
0.091463	0.076829	0.131707	0.117647	0.086301	0.131707

2nd result of PARFUM		
0.150588	0.123417	0.104555
0.123417	0.09863	0.081565
0.104555	0.081565	0.131707

Repeating it, we get the solution of PARFUM:

$$\begin{bmatrix} 0.167082 & 0.128086 & 0.104832 \\ 0.128086 & 0.095604 & 0.07631 \\ 0.104832 & 0.07631 & 0.118858 \end{bmatrix}$$

The difference of PARFUM algorithm between IPF is that, we make the row projection and column projection at the same time, and then average these two projections as our new distribution for the PARFUM algorithm. The virtue of this algorithm is that it is always converges even the problem is unfeasible which means that IPF algorithm no longer works.

D. Generalizations of IPF and PARFUM

IPF and PARFUM algorithms are both easy to apply. For easier presentation we have assumed that both margins are from S^K. This can be trivially extended. Moreover the results presented in previous subsection can be generalized to higher dimensions (M >2) [8]. Iterative algorithms can easily be adopted to satisfy joint as well as marginal constraints.

Generalization of the IPF to the continuous case has been introduced. However, the convergence of the IPF in the continuous case under certain regularity condition was proven much later. Proposition 10 and theorem 1 can be easily generalized to the continuous case. $I(p \mid q) \geq 0$. Using proposition 10 the proof of theorem goes through for continuous case.

In case of feasible problems we generally advocate to use IPF algorithm as it converges to the I-projection of starting distribution onto the set of feasible distribution. However, in case of infeasibility for high dimensional problems, IPF must be terminated at some point, and the resulting distribution may be concentrated on very few samples and may distribute the lack of fit very unevenly, with some variables being wildly wrong. It is not clear when the algorithm should be terminated and which distribution should be taken as a result. The iterative PARFUM algorithm offers some advantages with respect to

IFP for infeasible cases, and the practitioner is likely to feel more comfortable with PARFUM.

IV. EXPERT JUDGMENT ON POWER SYSTEM LOAD FORECASTING IN NINGXIA PROVINCE

Power load forecasting is sorted as long period forecasting, metaphase forecasting, short period forecasting and shorter period forecasting. Long period forecasting is often used to assess the power system load of ten years after or more ten years, and metaphase forecasting is to assess the load of around five years. It is important for the national power programming sectors to assess the power load of metaphase of long period forecasting, and it is helpful to decide where the new generate or change electricity units are to install and expand the power equipments.

NingXia State Power is located in the northeast of the North-West State Power and is the important sector. Its main power load concentrates on the coal, metallurgy high-consume-energy industry and so on. We can use the Large Consumer (LC) analysis method to forecast the power load.

Table III shows the traditional forecasting method on YongLi 110KV power station project. Without more information, we only simply surmise the maximum power load of each large consumer. The "Coincidence Factor" is necessary and important here for the large consumers cannot reach their maximum power load at the same time. By the National Calibration, the Coincidence Factor for the project is given at 0.85. The Coincidence Factor is the actual total load peak values divided the total load peak values of each large consumer, which value is less than 1.

In order to improve the current forecasting method, PRA method is introduced. The experts in the field of power system are asked to given the distributions of each large consumer's power load instead of point estimation. By using Monte Carlo simulation, the total power load is given. Since the experts improve the distribution on total power load, by using PI technology, large consumers' distributions are improved and more precisely "Coincidence Factor" is calculated.

TABLE III. THE TRADITIONAL FORECASTING METHOD: LC METHOD

YONGLI 110KV POWER STATION PROJECT, UNIT(KW)					
Power Load	**2006**	**2007**	**2008**	**After 2008**	**Production/year**
Ⅰ. Colliery					
Shicaocun Colliery	2000	3000	22000	22000	600 ten thousand tons
Hongliujing Colliery	1500	1500	3000	29000	800 ten thousand tons
Maiduoshan Colliery	1500	1500	3000	29000	800 ten thousand tons
Subtotal	5000	6000	28000	80000	
Coincidence Factor	0.85	0.85	0.85	0.85	
Total	4250	5100	23800	68000	
Ⅱ.Resident	1000	1060	1123.6	1200	
Total	5250	6160	24923.6	69200	

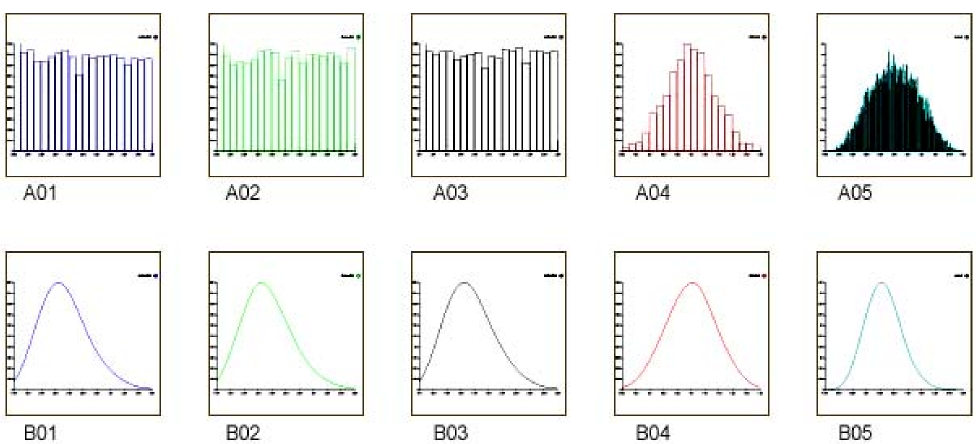

Figure 2. EJ on YongLi 110kV power station project

The variables A01, A02…A05 and B01, B02…B05 are indicated to:

A01	EJ on Shicaocun Colliery	Uniform distribution
A02	EJ on Hongliujing Colliery	Uniform distribution
A03	EJ on Maiduoshan Colliery	Uniform distribution
A04	EJ on Resident	Normal distribution
A05	Simulation result on the sum of A01, A02, A03 and A04	

B01	Modified distribution on Shicaocun Colliery after PI	
B02	Modified distribution on Hongliujing Colliery after PI	
B03	Modified distribution on Maiduoshan Colliery after PI	
B04	Modified distribution on Resident after PI	
B05	Modified distribution on A05 by EJ	

Variable	5% quantile	50% quantile	95% quantile	Mean	Median
Sum of Load	57471	65699	75692	65976	65699

The Coincidence Factor is:

$$\frac{75692}{22000+29000+29000+1200}=93.22\%$$

TABLE IV. PRA METHOD

EJ ON YONGLI 110KV POWER STATION PROJECT, UNIT(KW)					
Power Load	**2006**	**2007**	**2008**	**After 2008**	**Production/year**
Ⅰ. Colliery					
Shicaocun Colliery	2000	3000	22000	22000	600 ten thousand tons
Hongliujing Colliery	1500	1500	3000	29000	800 ten thousand tons
Maiduoshan Colliery	1500	1500	3000	29000	800 ten thousand tons
Subtotal	5000	6000	28000	80000	
Coincidence Factor	0.95	0.95	0.95	0.95	
Total	4750	5700	26600	76000	
Ⅱ.Resident	1000	1060	1123.6	1200	
Total	5750	6760	27723.6	77200	

From the figures and tables above, the expert has a transparent physical judgment on the distribution of variable A04, it's a normal distribution. While for the other variables of A01, A02 and A03, the expert cannot tell more but its blur interval. In practice, we use uniform distribution as their original distribution.

By using Monte Carlo simulation, we get the variable A05, the sum of A01, A02, A03 and A04. Fortunately, the expert has a more professional judgment on its distribution and hence we modify it and get variable B05.

By using PI technique, we get the modified variables of A01, A02 and A03: B01, B02 and B03. New important information is shown to us and the expert: the variables of A01, A02 and A03 are familiar to Gamma distribution.

V. CONCLUSION

In risk assessment, we must frequently employ expert judgment to assess model parameters which are not directly observable and about which experimental evidence is lacking. The model under consideration may be new and even unknown to the experts. In such situations it is impossible to query experts directly about parameters in the model. We can ask experts about quantities predicted by the model with which they have some experience and some feeling. We should pull the experts' uncertainty distributions on the query variables back onto the parameters of the model. Iterative sample re-weighing methods are available to solve such problem, as illustrated in the power system load forecasting in NingXia Province. IPF and PARFUM are easy to implement and have a theoretical foundation. They provide useful tools for the practicing quantities risk assessment.

ACKNOWLEDGMENT

This paper contains the results of research on the Science and Technology Key project of Tianjin Science and Technology Development Plan (04310791R) supported by Tianjin Municipal Science and Technology Commission of the People's Republic of China, also the research involves the part of YongLi 110kV power station project in NingXia State Power of the People's Republic of China as well.

REFERENCES

[1] Tim Bedford and Roger Cooke, Probabilistic risk analysis – Foundations and Methods, Cambridge University Press,May 2001.

[2] Haimers,Y., Risk Modeling , Assessment, and Management, John Wiley and Sons, 1998.

[3] Csiszar, I. and Tusnady,G. , "Information geometry and alternating minimization procedures", Statistic Decisions ,Vol.1, 1994, pp. 205-237.

[4] R¨uschendorf L., "Convergence of the iterative proportional fitting procedure", The Annals of Statistics, 23(4) ,1995, pp. 1160–1174.

[5] Leveson,N.G., Software: System safety and computers, Addision Wesley Publishing Co., 1995.

[6] C. Du, D. Kurowicka, and R.M. Cooke, "Techniques for generic probabilistic inversion", Computational Statistics & Data Analysis, vol.50 (2006) ,pp.1164 – 1187.

[7] Csiszar I. and Tusnady G., "Information geometry and alternating minimization procedures", Statistics & Decisions, 2002(1), pp.205–237.

[8] Cooke R.M. "Parameter fitting for uncertain models: modelling uncertainty in small models", Reliability Engineering and System Safety, Vol.44,1994, pp. 89–102 .

[9] B.C.P. Kraan, Probabilistic Inversion in Uncertainty Analysis and related topics. PhD thesis, Delft University of Technology, 2001.

[10] B.C.P. Kraan and R.M. Cooke, "Post-processing techniques for the joint cec/usrc uncertainty analysis of accident consequence codes", Journal of Statistical Computation and Simulation, Vol. 71, 1997, pp. 253-268.