FOUNDATIONS OF COMPUTING AND DECISION SCIENCES
Vol. 37 (2012) No. 4

DOI: 10.2478/v10209-011-0014-9

BAYESIAN-BASED METHODS FOR THE ESTIMATION OF THE UNKNOWN MODEL'S PARAMETERS IN THE CASE OF THE LOCALIZATION OF THE ATMOSPHERIC CONTAMINATION SOURCE

M. BORYSIEWICZ[1], A. WAWRZYNCZAK[1,2], P. KOPKA[1,3]

Abstract. In many areas of application it is important to estimate unknown model parameters in order to model precisely the underlying dynamics of a physical system. In this context the Bayesian approach is a powerful tool to combine observed data along with prior knowledge to gain a current (probabilistic) understanding of unknown model parameters. We have applied the methodology combining Bayesian inference with Markov chain Monte Carlo (MCMC) to the problem of the atmospheric contaminant source localization. The algorithm input data are the on-line arriving information about concentration of given substance registered by distributed sensor network. We have examined different version of the MCMC algorithms in effectiveness to estimate the probabilistic distributions of atmospheric release parameters. The results indicate the probability of a source to occur at a particular location with a particular release rate.

Keywords: Bayesian inference, stochastic reconstruction, MCMC methods

1. Introduction

In many fields of applications we face the task of having to draw conclusions from imperfect, very often fragmentary information. In those cases it becomes important to estimate some model's unknown parameters to predict more precisely the underlying dynamics of a considered physical system. In this context the Bayesian approach is a powerful tool to combine observed data along with prior knowledge to gain a current (probabilistic) understanding of unknown model parameters. In particular, it provides a very natural framework for updating the state of knowledge about a considered dynamic

[1] National Centre for Nuclear Research, Swierk - Otwock, Poland
[2] Institute of Computer Science, Siedlce University, Poland
[3] Institute of Computer Science of the Polish Academy of Sciences, Warsaw, Poland

system the more precisely, the more new data is available. For complex systems, such updating needs to be carried out via stochastic sampling of unknown model parameters.

One of the fields of application of the Bayesian approach is the problem of the location of the dangerous substance release from given concentration of the released substances in the points where the sensors are placed.

Knowledge of the temporal and spatial evolution of a contaminant released into the atmosphere, either accidentally or deliberately, is fundamental to adopt efficient strategies to protect the public health and to mitigate the harmful effects of the dispersed material. In this context there arise the questions to be answered: What was released? When and where was it released? How much material was released? Moreover, we have to answer these questions as soon as possible. In general, we are able to develop a model to predict precisely concentration fields of a pollutant. However, to create the model realistically reflecting the real situation based only on a sparse point-concentration data is not trivial. This task requires specification of set of parameters, which depends on the considered model. The stated problem is in some sense ill-posed. It should be noted, that only given a downwind concentration sparse measurement and knowledge of the wind field, the determination of the source location and/or its characteristics could be ambiguous. Non-inverting problems of this type are termed inverse problems: problems that can be solved in one direction but for some physical reason cannot be solved in the opposite direction. Such problems are widely encountered in several fields [15]. For instance the group method of data handling (GMDH) [6], [10] and it modifications seems to be successful as a method of inductive modeling and forecasting of complex processes and systems. The main idea of the GMDH is to have the algorithm construct a model of optimal complexity based only on the data. The goal is to get mathematical model to describe the processes, which will take place at object in the future. GMDH solves it, by sorting-out procedure, i.e. consequent testing of models, chosen from set of models-candidates in accordance with the given criterion. More recent developments utilize genetic algorithms or the idea of active neurons and multileveled self-organization to build models from data e.g. [2], [17].

In all inverse problems the aim is to infer the unknown state from measured consequences of that state. In the case of gas dispersion, the unknown state is the gas source distribution of strengths and locations; and the measured consequences are the gas concentrations for the associated wind conditions and measurement locations. Our aim is to find the source distribution that will generate predicted concentrations closest to those actually measured. To do this we have developed a dynamic data-driven event reconstruction model which couples data and predictive models through Bayesian inference to obtain a solution to the inverse problem.

The key idea behind statistical inversion methods is to recast the inverse problem in the form of statistical inference by means of Bayesian statistics. In the framework of Bayesian statistics all quantities included in the mathematical model are modeled as random variables with joint probability distributions. This randomness can be interpreted as parameter variability, as it is related to the uncertainty of the true values which is expressed in terms of probability distributions. The solution of the inverse problem corresponds to summarizing a probability distribution when all possible knowledge of the measurements, the model and the available prior information, has been incorporated. This distribution which is referred to as posterior distribution describes the degree of confidence about the estimated quantity conditioned on the measurements [18].

A comprehensive literature review of past works on solutions of the inverse problem for atmospheric contaminant releases can be found in [9]. A variety of approaches to solving the atmospheric dispersion inverse problem have been explored including non-linear optimization, back-trajectory, Green's function, adjoint, and Kalman filter methods [13]. However, these methods often fail due to the inherent complexities, high-dimensionality, and/or non-linearity of the underlying physical system [7]. In [7] and [8] dynamic Bayesian modeling was introduced, and the Markov chain Monte Carlo (MCMC); in [5] and [3] sampling approaches to reconstruct a contaminant source for synthetic data were presented.

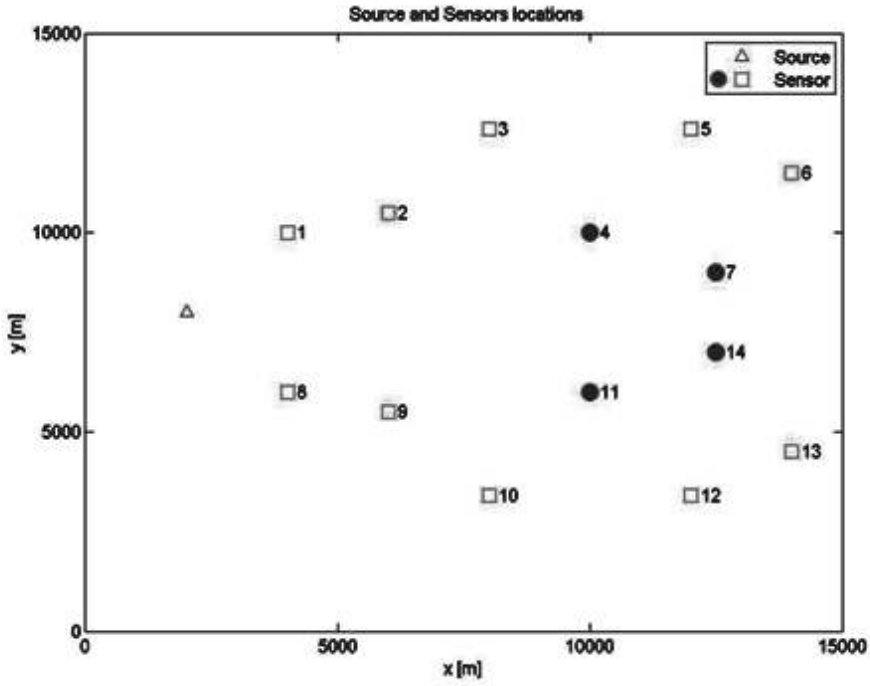

Figure 1. Distribution of the sensors and the release's source

1.1. Synthetic data

In this paper we have implemented stochastic models based on the MCMC sampling to find the contamination source location based on the concentration of given substance registered by the maximum 14 sensors distributed over 15km x 15km (Figure 1). The synthetic concentration data (Figure 2), used in testing the algorithm, were generated with use of the atmospheric dispersion Gaussian plume model [11], [16]. In this experiment the contamination source was located at x=2km, y=5km, z=50m within the domain (Figure 1). The release rate was assumed to change with time within interval q≈5000 g/s up to q≈7000 g/s which resulted in the change of the concentration measured by the sensor in subsequent time intervals (Figure 2). The wind was directed along x axis with speed 5m/s.

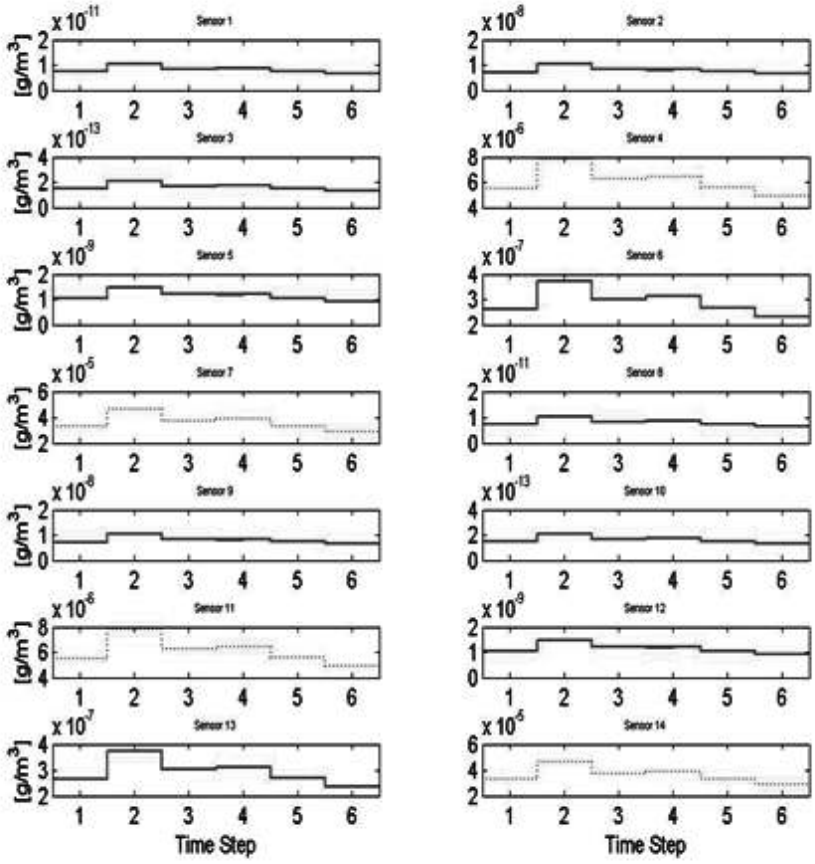

Figure 2. The synthetic concentration registered by the 14 sensor in 6 subsequent intervals (time steps)

2. Reconstruction procedure

2.1. Bayesian inference

A good introduction to Bayesian theory can be found in [3] and [1]. Bayes' theorem, as applied to an emergency release problem, can be stated as follows:

$$P(\mathrm{M}|\mathrm{D}) = \frac{\mathrm{P}(\mathrm{D}|\mathrm{M})\mathrm{P}(\mathrm{M})}{\mathrm{P}(\mathrm{D})} \tag{1}$$

M represents possible model configurations or parameters and D are observed data. For our application, Bayes' theorem describes the conditional probability P(M|D) of certain source parameters (model configuration M) given observed measurements of concentration at sensor locations (D). This conditional probability P(M|D) is also known as the posterior distribution and is related to the probability of the data conforming to a given model configuration P(D|M), and to the possible model configurations P(M), before taking into account the measurements. The probability P(D|M), for fixed D, is called the likelihood function, while P(M)-the prior distribution. P(D) is the marginal distribution of D and is called prior predictive distribution [18] it serves as a scaling factor; so the Bayes theorem can be written as follows:

$$P(M|D) \propto P(D|M)P(M) \tag{2}$$

To estimate the unknown source parameters M using (2), the posterior distribution P(M|D) must be sampled. P(D|M) quantifies the likelihood of a set of measurements D given the source parameters M.

Value of likelihood for a sample is computed by running a forward dispersion model with the given source parameters M and comparing the model predicted concentrations in the points of sensors location (within a considered domain) with actual observations D. The closer the predicted values are to the measured ones, the higher is the likelihood of the sampled source parameters (Figure 3).

As the sampling procedure we use an MCMC with the Metropolis–Hastings algorithm to obtain the posterior distribution P(M|D) of the source term parameters given the concentration measurements at sensor locations [3], [5]. This way we completely replace the Bayesian formulation with a stochastic sampling procedure to explore the model parameter space and to obtain a probability distribution for the source location.

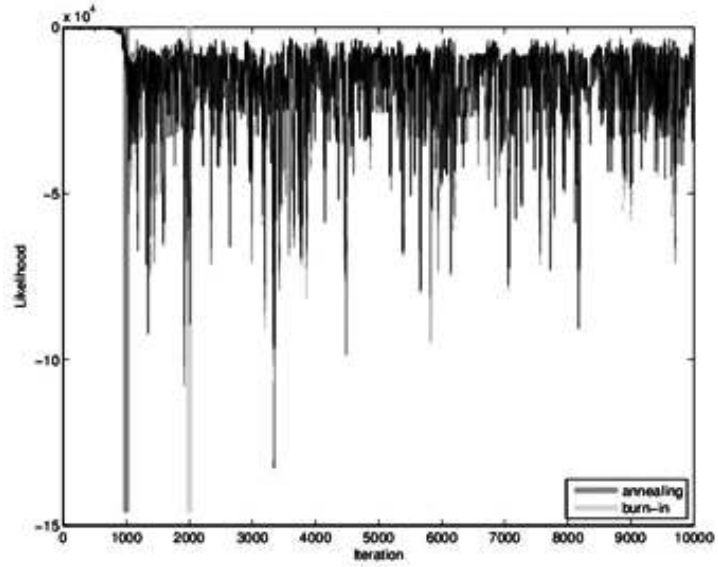

Figure 3. The values of the likelihood function versus number of iterations for one of the MCMC algorithms with marked the annealing and burn-in phases.

The Markov chains are initialized by taking samples from the prior distribution (in different ways presented further in this paper). To lower the computational cost, we limit the prior distribution to the two dimensional space fixing the vertical position constant both for the source and sensors location at 50 m.

2.2 The likelihood function

A measure indicating the quality of the current state of Markov chain is expressed in terms of a likelihood function. This function compares the predicted from model and observed data at the sensor locations as:

$$\ln[\mathrm{P(D|M)}] = \ln[\lambda(\mathrm{M})] = -\frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}\left[\log\left(\mathrm{C_i^M}\right)-\log\left(\mathrm{C_i^E}\right)\right]^2}{2\sigma_{\mathrm{rel}}^2} \tag{3}$$

where λ is the likelihood function, C_i^M are the predicted by the forward model concentrations at the sensor locations *i*, C_i^E are the sensor measurements, σ_{rel}^2 is an error parameter chosen accordingly to expected errors in the observations and predictions, *N* is the number of sensors.

After calculating value of the likelihood function (Figure 3) for the proposed state its acceptance is performed as follows:

$$\frac{\ln(\lambda_{\mathrm{prop}})}{\ln(\lambda)} \geq \mathrm{RND}(0,1) \tag{4}$$

where λ_{prop} is the likelihood value of the proposal state, λ is the previous likelihood value, and *RND(0, 1)* is a random number generated from a uniform distribution in the interval (0, 1).

It is important to note that condition (4) is more likely to be satisfied if the likelihood of the proposal is only slightly lower than the previous likelihood value. It gives a chance to choose even a little "worse" state, because the probability of acceptance depends directly on the quality of proposed state. Different likelihood functions can also be developed [14].

2.3 Posterior distribution

The posterior probability distribution (2) is computed directly from the resulting Markov chain paths defined by the algorithm described above and is estimated with

$$P(\mathrm{M|D}) = \frac{1}{n}\sum_{i=1}^{n}\delta\left(\mathrm{M_i}-\mathrm{M}\right) \tag{5}$$

which represents the probability of a particular model configuration giving results that match the observations at sensor locations. Equation (5) is a sum over the entire Markov chain of length *n* of all the sampled values M_i. Thus $\delta\left(M_i-M\right)=1$ when $M_i = M$ and *0* otherwise. If a Markov chain spends several iterations at the same location value of $P(M|D)$ increases through the summation (increasing the probability for those source parameters).

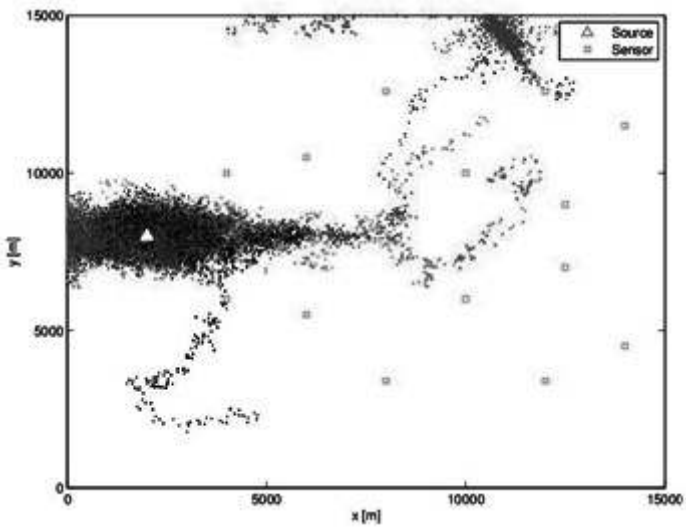

Figure 4. The traces of five Markov chains in the *x,y* space. The source location is marked by triangle and the sensors by squares. The samples came from results of Standard MCMC algorithm.

2.4 Forward dispersion model

A forward model is needed to calculate the concentration C_i^M at the points i of sensor locations for the tested set of model parameters M at each Markov chain step. As a testing forward model we selected the fast-running Gaussian plume dispersion model [11], [16].

The Gaussian plume dispersion model for uniform steady wind conditions can be written as follows:

$$\mathrm{C(x,y,z)} = \frac{\mathrm{q}}{2\pi\sigma_\mathrm{y}\sigma_\mathrm{z}\mathrm{U}} \exp\left[-\frac{1}{2}\left(\frac{\mathrm{y}}{\sigma_\mathrm{y}}\right)\right] \cdot \left\{\exp\left[-\frac{1}{2}\left(\frac{\mathrm{z-H}}{\sigma_\mathrm{z}}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{\mathrm{z+H}}{\sigma_\mathrm{z}}\right)^2\right]\right\} \qquad (6)$$

where $C(x,y,z)$ is the concentration at a particular location, U is the wind speed directed along x axis, q is the emission rate or the source strength and H is the height of the release; y and z are the distance along horizontal and vertical direction, respectively. In the equation (6) σ_y and σ_z are the standard deviation of concentration distribution in the crosswind and vertical direction. These two parameters are defined empirically for different stability conditions by [12] and [4]. In this case we restrict the diffusion to the stability class C (Pasquill type stability for rural area). Thus, in creation of the testing data we have fixed this coefficient as:

$$\sigma_\mathrm{y} = 0.22\mathrm{x} \cdot (1 + \mathrm{x} \cdot 4 \cdot 10^{-5})^{-0.5}, \qquad \sigma_z = 0.2x \,.$$

However, we assume in scanning algorithm that we do not know exact behavior of the plume and consider those coefficients as unknown. Thus, the parameters $\sigma_\mathrm{y}, \sigma_\mathrm{z}$ are taken as:

$$\sigma_y = \zeta_1 \cdot x \cdot (1 + x \cdot 4 \cdot 10^{-5})^{-0.5} \quad (7a)$$

$$\sigma_z = \zeta_2 \cdot x \quad (7b)$$

where values $\zeta_1 and\ \zeta_2$ are sampled by algorithm within interval [0,0.4].
To summarize, in this paper the searched model's parameter space is

$$M = (x, y, q, \zeta_1, \zeta_2) \quad (8)$$

where x and y are spatial location of the release, q release rate and ζ_1, ζ_2 are stochastic terms in the turbulent diffusion parameterization given in (7ab).

2.5 Scanning algorithm

We assume that the information from the 14 sensors arrives subsequently in six time steps. We start to search for the source location (x,y), release rate (q) and model parameters $\zeta_1 and\ \zeta_2$ after first sensors' measurements (based on the data in time t=1,see Figure 2). Thus, scanning algorithm is run with obtaining the first measurements from the sensors (Figure 2). Based on this information we obtain the probability distributions of the searched parameters (8) starting from the randomly chosen set of parameters M (i.e. first we start from the 'flat' priori). This assumption reflects lack of knowledge about the release. The forward calculations are performed for the actual state M and likelihood function λ is calculated. Then we apply random walk procedure "moving" our Markov chain to the new position. Precisely, we change each model M parameter by the value draw from the Gaussian distribution with the variance σ_M^2 equal *200* for x and y, *100* for q and 0.02 forζ_1, ζ_2. Based on proposal state forward calculation the likelihood function λ_{prop} is again estimated. We compare this two values λ and λ_{prop} according to (4). If comparison is more favorable than the previous chain location, the proposal is accepted (Markov chain "moves" to the new location). If the comparison is "worse", new state is not immediately rejected. Bernoulli random variable (a "coin flip") is used to decide whether or not to accept the new state of chain. This random component is important because it prevents the chain from becoming trapped in a local minimum. The pseudo code of the algorithm is presented in Table 1.

In our calculation we use 10 Markov chains in each time step. The traces of five independent Markov chains in the x,y space are presented in Figure 4, the source location is marked by triangle and sensors by squares.

The number of iteration for each Markov chain $n=10000$. This number was chosen based on the numerical experiments as the number of iteration needed to reach convergence for each sampled model parameters. Statistical convergence to the posterior distribution is monitored by computing between-chain variance and within-chain variance [3]. If there are m Markov chains of length n, then we can compute between-chain variance B with

$$B = \frac{n}{m-1} \sum_{j=1}^{m} \left(\bar{M}_j - \bar{M}\right)^2 \quad (9)$$

where $\bar{M}_j$ is the average value along each Markov $\bar{M}$ and is the average of the values from all Markov chains. The within-chain variance W is

$$W = \frac{1}{m}\sum_{i=1}^{m} s_i^2 \tag{10}$$

where

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(M_{ij} - \bar{M}_i\right)^2 \tag{11}$$

The convergence parameter R is then computed as

$$R = \frac{\mathrm{var}(M)}{W} \tag{12}$$

where $\mathrm{var}(M)$ is estimate variance of M and is computed as

$$\mathrm{var}(M) = \frac{n-1}{n}W + \frac{1}{n}B \tag{13}$$

The convergence *R* value vs. the number of iteration for searched parameters presents Figure 5. One can see that the 10000 iterations satisfy the convergence condition R≈1.

Table 1. Scanning algorithm

FOR TimeStep=1:6 **FOR** j=1:ChainNumber **Draw** M ~ priori distribution **ForwardDispersion**(M) **Read** C^M **Compute** ln(λ(MFOR i=1:N $Chain^j_{TimeStep}$=$Chain^j_{TimeStep}$+M_i; M*=M_i+N(0,σ^2_M); **ForwardDispersion**(M*); **Read** C^{M*}; **Compute** ln(λ(M*)); **IF**(ln(λ(M*))/ln(λ(M))>=RND(0,1) **THEN** Mi=M*; ln(λ(M))= ln(λ(M*)); **END IF** **END FOR** **END FOR** **END FOR**	*-In different way for considered MCMC algorithms* *-Model calculation with set M (6)* *-Read values from all sensors* *-Calculating the likelihood function(3)* *-Add sample to the Markov Chain* *-Random Walk* *-Calculation with set M* (6)* *-Read values from all sensors for C^{M*}* *-Calculating the likelihood function (3)* *- condition of acceptance (4)* *-Changing Markov chain position*

One of the important aspects of stochastic procedure of calculating the posterior distribution is choosing burn-in phase. The burn-in factor represents the number of samples needed at the beginning for the Markov chain to actually reach the search state where it is sampling from the target distribution. These initial samples are discarded and not used for inference. In our calculation the burn-in was fixed at 2000 iterations. This value was chosen based on the numerical experiments as the number of iteration needed to reach the target

distribution with same approximation. Figure 6 presents the trace of x and y coordinates of source location vs. number of iteration. The burn-in is marked by vertical line.

In the algorithm we have also applied the annealing process technique which allows to escape from a bad initial location and to get "flatter" likelihood (Figure 3). It is motivated by willingness to accurately search the parameter space in the initial iterations. This procedure involves the modification of likelihood function during the initial phase of the algorithms. For the first K=1000 iterations (i = 1,…,K) the likelihood was taken as:

$$\ln[\lambda(M)]^{\frac{1}{T_i}}, \tag{14}$$

where $T_i = 1+(10-i(10/K))$. It is worth to note, that the effect of T_i ratio decreases with increasing values of iteration (up to a limit equal K).

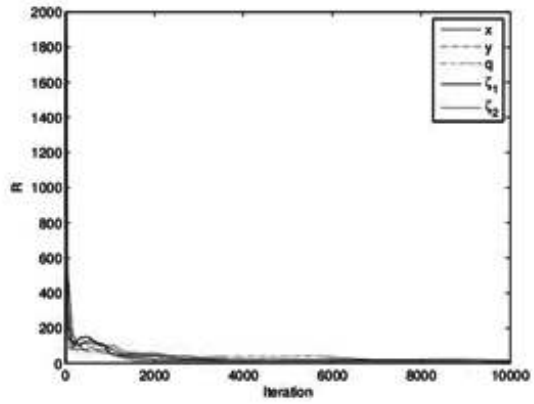

Figure 5. Convergence values R vs the number of iterations for sampled parameters.

In subsequent time intervals (subsequent time steps) we investigate different version of MCMC algorithms that use (or not) the probability distributions obtained based on information from previous measurements as the priori distribution in (2) and update the marginal probability distribution with use of the newly arrived measurements. The scanning algorithm can (or not) take the advantage from the past MCMC realizations in different ways. Each type of algorithms has unique properties that have an impact on various aspects of the reconstruction of events. In this paper we examine the following MCMC algorithms:

Standard MCMC

In this algorithm, the parameter space scan in each time step t is independent form the previous ones. So, in this case we don't use information from past calculations i.e. in each time step the calculations start from scratch.

MCMC via Maximal Likelihood

This algorithm uses the results obtained in the previous time steps to run calculation with use of the new measurements. As the first location of Markov chain M_0^t it selects the set of

M parameters for which likelihood function in previous time step was the highest. So, for *t>1*:

$$M_0^t \sim \arg \max_{M \in \{M_0^{t-1}, \ldots, M_n^{t-1}\}} \ln[\lambda(M)] \tag{15}$$

With this approach, we always start with the best values of the model (previously found) and correct the result with new information from sensor.

MCMC via Rejuvenation and Extension

This algorithm as the first location of Markov chain M_0^t at the time t>1 chooses the set of parameters M selected randomly from previous realization in *t-1* with use of the uniform distribution:

$$M_0^t \sim U(M_0^{t-1}, M_1^{t-1}, \ldots, M_n^{t-1}) \text{a uniform distribution } \{1, \ldots, n\} \tag{16}$$

Applying the new knowledge (new measurements) the current chain is "extended" starting from selected position with use of the new data in the likelihood function calculation.

MCMC via Rejuvenation, Modification and Extension

This algorithm, similarly to the MCMC via Rejuvenation and Extension algorithm, as the first location of Markov chain M_0^t at the time *t>1* chooses the set of parameters *M* selected randomly from previous realization *t-1* with use of the uniform distribution as (16). However, additionally it modifies the Markov chain path obtained in previous time step *t-1*. So, if the $M_{drawned}^{t-1} = M_0^t$ is the dawned first location of Markov chain in time *t* according to (16) then the chain in time *t-1* is modified starting from this position with a new data available in time *t:*

$$(M_0^{t-1}, M_1^{t-1}, \ldots, M_{drawned}^t, \ldots, M_n^t).$$

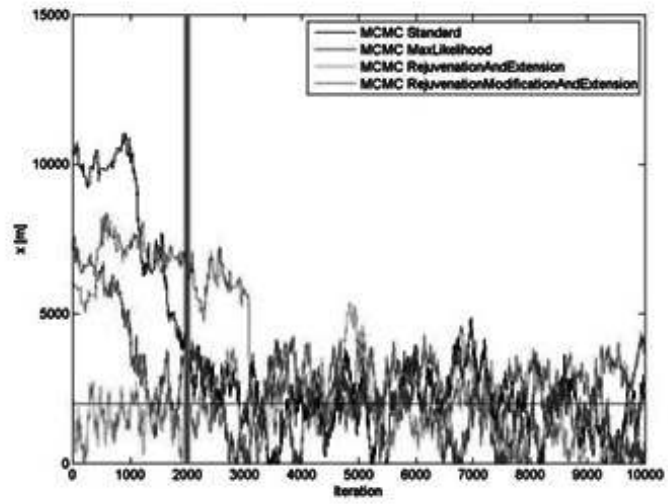

Figure 6. The trace of *x* coordinates for all considered algorithms. The target value is marked by horizontal line; the burn-in is marked by vertical line.

3. Results

Figures 7, 8, 9 and 10 presents the results of calculation with use of all four above described MCMC algorithms for $x,\ y, \zeta_1 and\ \zeta_2$ parameters. Presented marginal probability distributions were calculated based on the scanning algorithms' results from all time steps and all Markov chains.

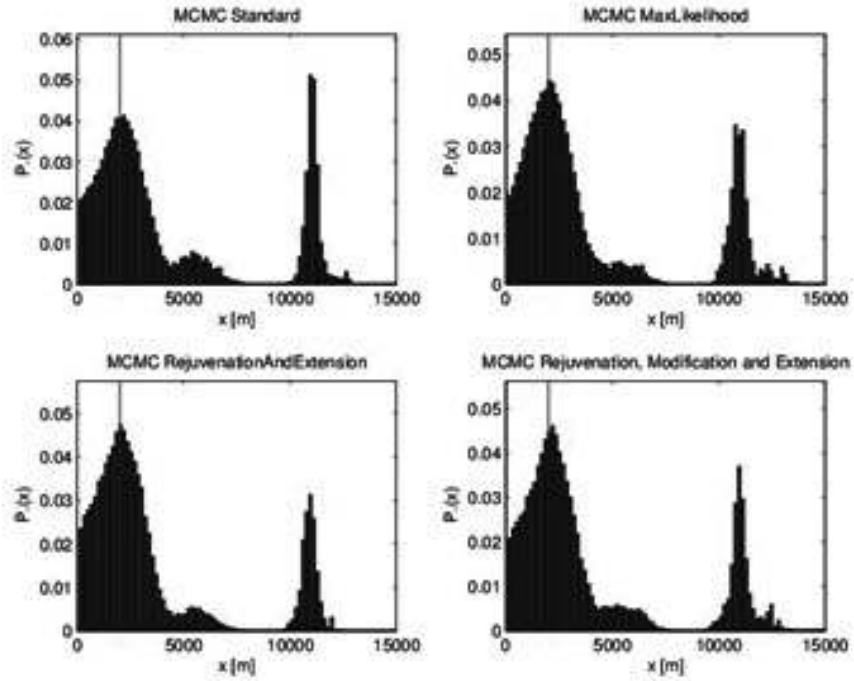

Figure 7. Posterior distribution as inferred by the Bayesian event reconstruction for all applied algorithms for *x* parameter. Posterior distributions were averaged based on the data for all time steps and all Markov chains. Vertical lines represent the target *x* value.

Figure 7 shows the marginal probability distribution for *x* coordinate of source location within the considered domain. One can see that the Standard MCMC algorithm do not marked the target value of *x* as the value with the highest probability. At the same time all other methods hit to the target value of *x*. Additionally, algorithms MCMC via Rejuvenation and Extension and MCMC via Rejuvenation, Modification and Extension mark this value with higher probability than MCMC Max Likelihood. The same is for the ζ_1 parameter. On the other hand all methods successfully find the correct value of the *y* coordinate of the source. The reason of the high peek in the histogram is that y is the crosswind direction, thus applied model is quite sensitive to this parameter. In contrast, all methods do not find the target value of the ζ_2 parameter being responsible for the dispersion in vertical direction which do not influence the results much, as far we consider sensors and source as fixed at z=50m. We do not consider the probability of the release rate distribution, as far it was changed during creation of the synthetic data.

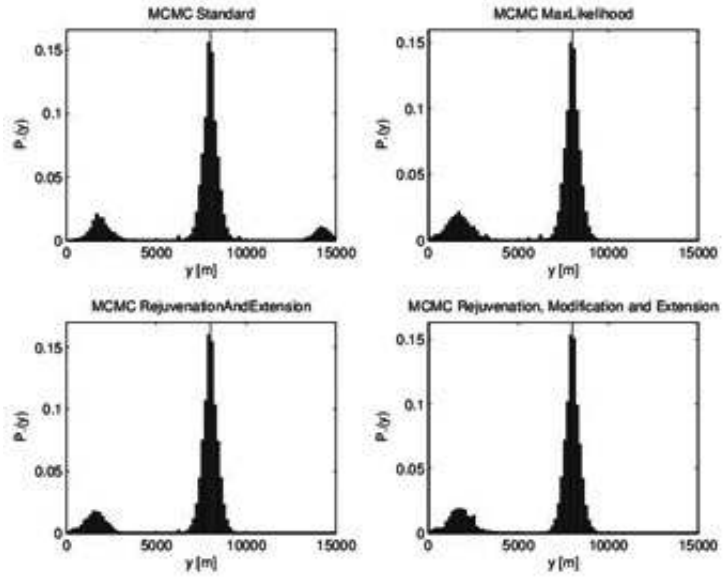

Figure 8. The same as in Figure 7 for *y* parameter

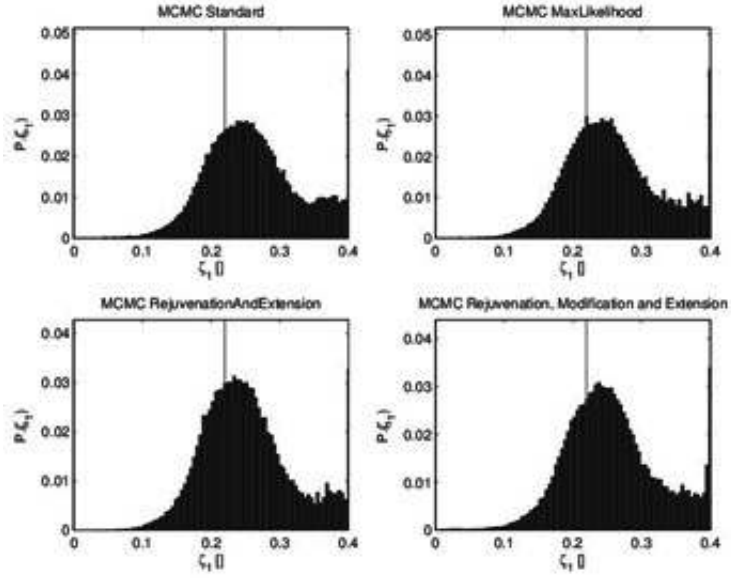

Figure 9. The same as in Figure 7 for ζ_1 parameter

It is worth to mention that three algorithms for which we obtain better results (i.e. MCMC Max Likelihood, MCMC via Rejuvenation and Extension, MCMC via Rejuvenation, Modification and Extension) use the probability distributions obtained based on information from previous measurements Markov chain in subsequent time step is sampled from the posterior distribution in previous time step. This methodology makes

those algorithms more effective in localization of the most probable value of considered parameters.

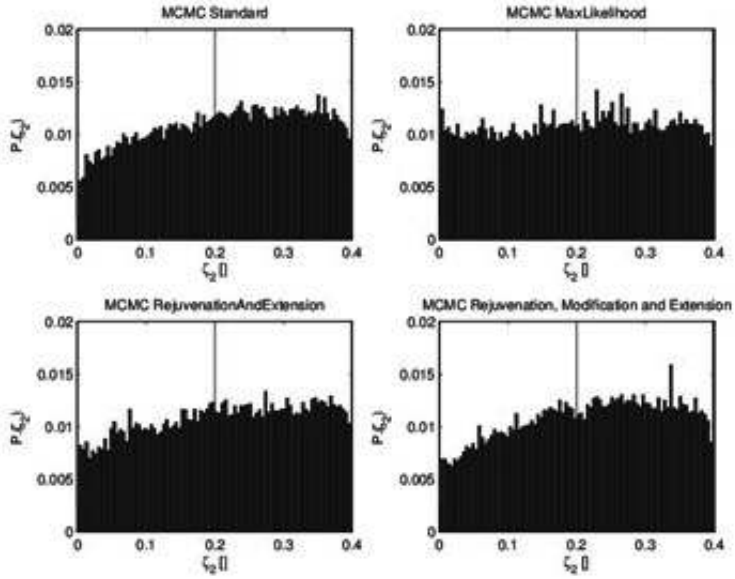

Figure 10. The same as in Figure 7 for ζ_2 parameter

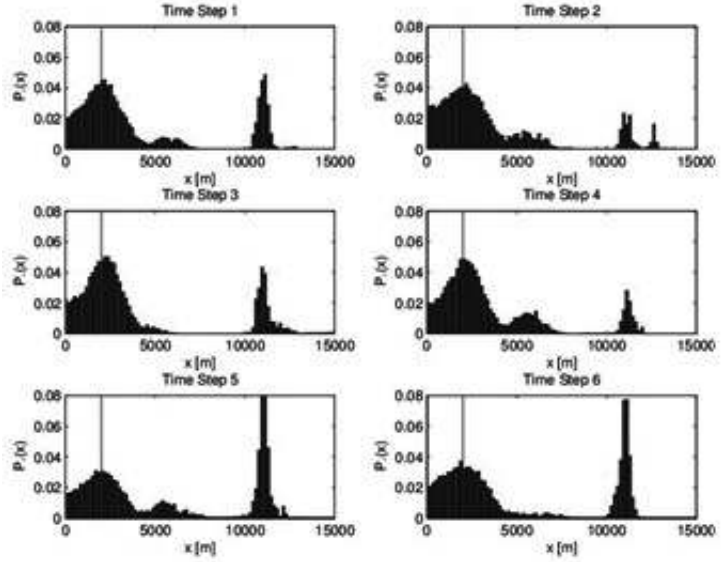

Figure 11. Posterior distribution of *x* parameter in subsequent time steps for Standard MCMC algorithm. Vertical line represents the target value of x.

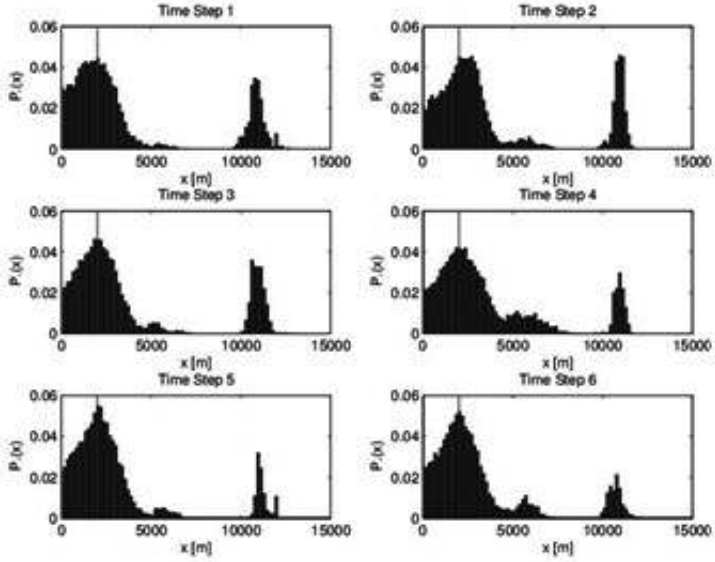

Figure 12. The same as in Figure 11 for MCMC via Rejuvenation and Extension algorithm.

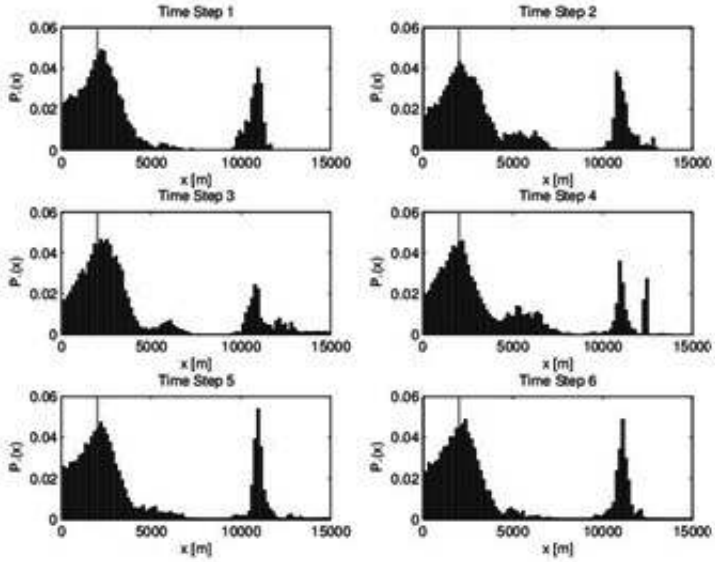

Figure 13. The same as in Figure 11 for MCMC via Rejuvenation, Modification and Extension algorithm.

Figures 12 and 13 presents how the probability distributions of x parameter are updated in subsequent time steps by MCMC via Rejuvenation and Extension (Figure 12) and MCMC via Rejuvenation, Modification and Extension (Figure 13) algorithms. One can see

that with time the probability of the target value is reached with higher probability, at the same time the false peak around x≈12000m decreases. Moreover, the MCMC via Rejuvenation, Modification and Extension algorithm seems to be more effective. On the contrary Figure 11 presents that Standard MCMC algorithm does not increase its efficiency in finding the target value with time.

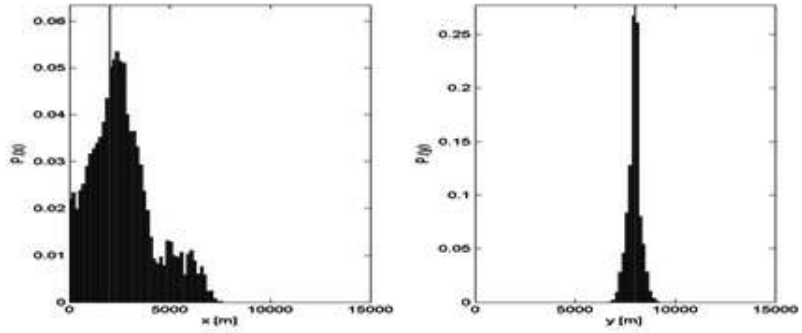

Figure 14. Posterior distribution as inferred by the Bayesian event reconstruction for MCMC via Rejuvenation and Extension algorithm for x and y parameter obtained based on 10 sensors measurements. Posterior distributions were averaged based on the data for all time steps and all Markov chains. Vertical lines represent the target values.

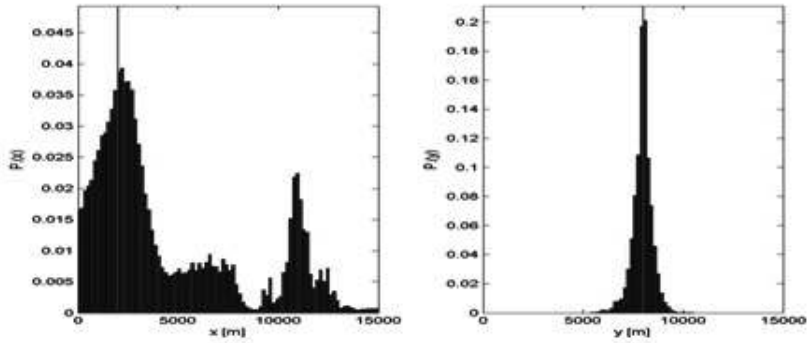

Figure 15. The same as in Figure 14 for MCMC via Rejuvenation, Modification and Extension algorithm

To confirm correctness of the proposed algorithms we have tested them assuming that among 14 sensors (Figure 1) we gather the measurements only from 10 sensors i.e. we do not have the information from four centrally located sensors designated by numbers *4, 7, 11, 14* (denoted by circles in Figure 1). This experiment allows us to see whether proposed algorithms will still be able to successfully find the target value of searched parameters.

Sensors with numbers *4,7,11,14* (painted circles in Figure 1) we can consider as the quite important because they record the highest concentration which variability in

successive time steps changes the most (sensors with dotted lines in Figure 2). Figures 14 and 15 presents posterior probability distribution of x and y parameter for MCMC via Rejuvenation and Extension and MCMC via Rejuvenation, Modification and Extension algorithms obtained based on the calculations with assumed new domain (based on 10 sensors data). All other parameters were the same as in the previous set up.

We can see that reduction of information applied in the calculation of the likelihood function (3) comparing the model predicted concentrations and measured one do not caused decrease in efficiency of the proposed algorithms. The tested algorithms (MCMC via Rejuvenation and Extension and MCMC via Rejuvenation, Modification and Extension) successfully pointed out the target values of the x and y parameters as the values with the highest probabilities. This confirms the proper functioning of the proposed parameters' scanning algorithms.

4. Conclusion

We have presented a methodology to reconstruct a source causing an area contamination, basing on a set of measurements. The method combines Bayesian inference with Markov chain Monte Carlo sampling and produces posterior probability distributions of the parameters describing the unknown source. We developed dynamic data-driven event reconstruction model, which couples data and pollutant dispersion simulations through Bayesian inference. The approach successfully provide the solution to the stated inverse problem i.e. having the downwind concentration measurement and knowledge of the wind field algorithm found the most probable location of the source.

We have examine usefulness of different version of the MCMC algorithms i.e. Standard MCMC, MCMC via Maximal Likelihood, MCMC via Rejuvenation and Extension, MCMC via Rejuvenation, Modification and Extension and its modification in effectiveness to estimate the probabilistic distributions of searched parameters. We have shown the advantage of the algorithms that in different ways use the source location parameters probability distributions obtained basing on available measurements to update the marginal probability distribution of considered parameters with use of the received new information. As the most effective we pointed the MCMC via Rejuvenation, Modification and Extension algorithm. We have verified the proposed algorithms assuming two sensors' setups and showed their efficiency in cases when smaller amount of measurements was available.

The probabilistic aspect of the solution optimally combines a probable answer with the uncertainties of the available data. Among several possible solutions, the Bayesian source reconstruction is solely able to find values of the model parameters that are more consistent with the data available and its uncertainties.

The stochastic approach used in this paper is completely general and can be used in other fields where the parameters of the model bet fitted to the observable data should be found.

Acknowledgements

This work was supported by the Welcome Programme of the Foundation for Polish Science operated within the European Union Innovative Economy Operational Programme 2007-2013 and by the EU and MSHE grant nr POIG.02.03.00-00-013/09.

References

[1] Bernardo, J. M. & Smith, A. F. M., *Bayesian Theory*, Wiley, 1994.
[2] Fujimoto, K., Nakabayashi S., Applying GMDH algorithm to extract rules from examples, *Systems Analysis Modelling Simulation*, **43**, 10, 2003. 1311–1319.
[3] Gelman, A., J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*, Chapman & Hall/CRC, 2003.
[4] Gifford, F. A. Jr. Atmospheric dispersion calculation using generalized Gaussian Plum model, *Nuclear Safety*, 1960, 2(2):56-59, 67–68.
[5] Gilks, W., S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, 1996, 486.
[6] Ivakhnenko, A.G., Group method of data Handling – A Rival of the Method of Stochastic Approximation, *Soviel Automatic Control*, **13**, 43–71, 1966.
[7] Johannesson, G. et al., Sequential Monte-Carlo based framework for dynamic data-driven event reconstruction for atmospheric release., *Proc. of the Joint Statistical Meeting, Minneapolis*, MN, American Statistical Association and Cosponsors, 2005, 73–80.
[8] Johannesson, G., W. Hanley, and J. Nitao, *Dynamic Bayesian models via Monte Carlo – An introduction with examples*, Lawrence Livermore National Laboratory Tech. Rep., 2004, 53.
[9] Keats, A., E. Yee, and F.-S. Lien, Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ*., **41**, 2007, 465–479.
[10] Madala H.R., Ivakhnenko A.G., *Inductive Learning Algorithms for Complex Systems Modeling*, CRC Press, 1994.
[11] Panofsky, H. A., Dutton, J. A., *Atmospheric Turbulence*. John Wiley, 1984.
[12] Pasquill, F. The estimate of the dispersion of windborne material, *Meteorol Mag*.,**90**, 1063, 1984, 33–49.
[13] Pudykiewicz, J. A., Application of adjoint tracer transport equations for evaluating source parameters. *Atmos. Environ*., **32**, 1998, 3039–3050.
[14] Senocak I., N. W. Hengartner, M. B. Short, W. B. Daniel, Stochastic Event Reconstruction of Atmospheric Contaminant Dispersion Using Bayesian Inference, *Atmos. Environ*., **42(33)** , 2008, 7718–7727.
[15] Thomson, L. C., Hirst, B., Gibson, G., Gillespie, S., Jonathan, P., Skeldon, K. D., Padgett, M. J., An improved algorithm for locating a gas source using inverse methods. *Atmospheric Environment,* **41**, 2007, 1128–1134.
[16] Turner D. Bruce, *Workbook of Atmospheric Dispersion Estimates*, Lewis Publishers, USA,1994.
[17] Vicenç Puiga,Marcin Witczak, Fatiha Nejjari, Joseba Quevedo, Józef Korbicz, A GMDH neural network-based approach to passive robust fault detection using a constraint satisfaction backward test, *Engineering Applications of Artificial Intelligence*, **20**, Issue 7, 2007, 886–897.
[18] Watzenig,D., Bayesian inference for inverse problems – statistical inversion. *Elektrotechnik and Informationstechnik* ,**124/7/8**, 2007, 240–247.