FITTING OF SEGMENTED GAUSSIAN PLUME MODEL PREDICTIONS ON MEASURED DATA

Petr Pecha and Radek Hofman
Institute of Information Theory and Automation
Pod Vodarenskou vezi 4
182 08 Prague 8, Czech Republic
E-mail: pecha@utia.cas.cz

KEYWORDS

Pollutant Spreading, Simulation Models, Uncertainty Bound

ABSTRACT

An improvement of mathematical model predictions of environmental pollution can be achieved on basis of assimilation of model simulations with real observations incoming from terrain. In this article we pay attention to development and investigation of applicability of one simple empirical method of objective analysis based on least square approach. Output background fields of resulting potentially dangerous endpoints are modified by measurements in such a way, that resulting respond surface is fitted towards measurements through the iterative adjustment of a certain selected set of model input parameters. In spite of a certain limitations this approach has occurred to be applicable for the first preprocessing of the model predictions and simulated measurements. It can support robustness of decision making and can contribute to early detection of possible fatal decision maker errors due to misinterpretation of input parameters of an accidental release scenario.

ASSESSMENT OF ACCIDENT CONSEQUENCES

Potential failures occurred in man-made processes can cause dangerous phenomena resulted in accidental releases of harmful substances into the living environment. Hazard evaluation and decision-making focused on early warning and protection of population has the highest priority. Reliable and up to date information represents basic inevitable conditions for effective management of intervention operations targeted on consequence mitigation during emergency situations. This appeared to be a basic lesson for further progress of emergency preparedness procedures, which has arisen from Chernobyl accident where lack of reliable information has shown to be the main reason of poor effectiveness of countermeasures. Decision making has to be supported by proper user-friendly simulation software tool complied with advanced theoretical methodology with access to all necessary relevant latest data. Crisis management should come out from reliable simulation of space and time of accident evolution, which should take into account all available information including physical knowledge of problem, expert judgement of input data, online measurements from terrain and others. The subject of investigation concerns evaluation of consequences of radioactivity propagation after an accidental release from nuclear facility. Transport of radioactivity is simulated by mathematical models from initial atmospheric propagation, deposition of radionuclides on the ground and spreading through food chains towards human body. In the final step a hazard estimation based on doses of irradiation is integrated into the software system HARP. Our access is mentioned in (Pecha et al. 2007).

FROM DETERMINISTIC TO PROBABILISTIC APPROACH AND DATA ASSIMILATON

Recent trends in risk assessment methodology insist in transition from deterministic procedures to probabilistic approach which enables generate more informative probabilistic answers on assessment questions. Corresponding analysis should involve uncertainties due to stochastic character of input data, insufficient description of real physical processes by parametrisation, incomplete knowledge of submodel parameters, uncertain release scenario, simplifications in computational procedure etc. Simulation of uncertainty propagation through the model brings data not only for the probabilistic assessment mentioned above (Pecha et al. 2005) but also for another main task of analysis called assimilation of model predictions with real measurements incoming from terrain. Data assimilation represents the way from model to reality and can substantially improve the model predictions.

There are several important sources of information that can enter the assimilation procedures. Basic physical knowledge is included in prior fields (resulted vectors) predicted by simulation model. Assumptions related to the random characteristics of model inputs are supported by some kind of expert judgements (Goosens 2001). Substantial benefit can result from accessibility of data incoming from terrain. Merging of all these contending resources is a principle of assimilation and had shown to be very promising in many branches of contemporary Earth sciences (e.g. Drécourt 2004). Each such resource can be known on a certain degree of details (e.g. dense or rare measurements in space and time, complete or only partial knowledge of model error covariance structure, cases with indirect observations etc). Available information determines the option of suitable assimilation technique. We are considering the assimilation techniques in broader sense (Hofman 2007) from simple interpolation (poor model predictions, but dense and precise observed data) up to advanced statistical methods when full description of error covariance structure is needed - e.g. in (Kalnay 2003).

DATA ASSIMILATION (DA) USING MINIMISATION TECHNIQUE (MT)

In this article we are introducing one simple particular method based on nonlinear optimisation technique. During assimilation we assume precise measurements and thus the procedure cannot be presented as pure statistical DA. On the other hand it requires proper environmental model which is able to describe uncertainty propagation (Pecha et al. 2005). Our model is based on segmented Gaussian plume model (SGPM) approach that can account approximately for dynamics of released discharges and short-term forecast of hourly changes of meteorological conditions. For near area from the source and constant meteorological conditions can be used also simplified version of Gaussian straight-line plume model (GPM). Implemented numerical difference scheme enables simulate approximately formation of important parent-daughter pairs.

The objective multi-dimensional function F of N variables (subjected to bounds) is minimised starting at initial estimate. Commonly used Nelder-Mead direct search or Powell minimisation methods are tested here for elementary scenarios of accidental harmful discharges. Applicability bounds are examined for which satisfactory results at acceptable time of computation are achieved.

PRINCIPLES OF APPLICATION WITHIN ATMOSPHERIC DISPERSION MODELLING

Even for the simplest formulation of atmospheric dispersion and deposition in terms of Gaussian straight-line propagation the model M is nonlinear. In the following paragraphs we shall concentrate on accidental radioactivity release into atmosphere and its further deposition on terrain. Approximation in terms of source depletion scheme accounts for removal mechanisms of admixtures from the plume due to radioactive decay and dry and wet deposition on terrain (Pecha et al. 2007). Let us proceed directly to the examination of the resulting fields of radioactivity deposition of a certain nuclide on terrain. The output is assumed to be represented by vector Z having dimension equal to the number N of total calculating points in the polar grid (in our case N= 2800, what means 80 radial sections and 35 concentric radial zones up to 100 km from the source of pollution). General expression for dependency of Z on model input parameters Θ_1, Θ_2, , ..., Θ_K can be formally written as

$$Z = M(\theta_1, \theta_2, \ldots, \theta_K) \qquad (1)$$

Let there are R receptor points on terrain where the respective values are measured. Generally, the number of receptors is much lower then N and we meet the problem with rare measurements expressed by observation vector $Y \equiv (y_1, y_2, \ldots, y_R)$. Positions of sensors generally differ from the points of calculation grid. We shall use terminology from data assimilation for introduction of observation operator H, specially for its linear observation matrix H. H is R × N matrix and transforms vectors Z from model space (having length N) into corresponding vector $\dot{Z}$ in observation space (having length R) according to matrix notation $\dot{Z} = H \cdot Z$. Components $\dot{z}_r$ of vector $\dot{Z}$ represent model predictions interpolated at the positions of simulated observations r = 1, ... , R. Innovation vector $D = Y - H \cdot Z$ is defined.
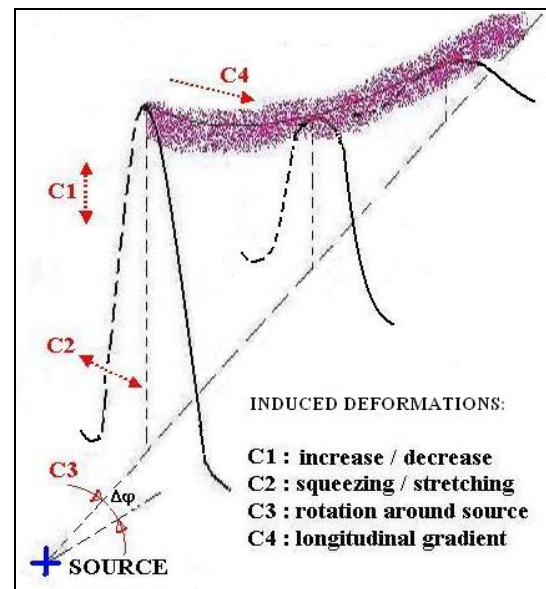

Figure 1: "Manipulations" with Resulting Straight-line Gaussian Shape

Number K of input parameters is rather high (several tenth) and then for practical purposes only S of them are treated as random. Rest of them are assumed to be less important from viewpoint of uncertainty propagation through the model and we assign them their best estimated values. Equation (1) is then simplified to the form $Z = M(\theta_1, \theta_2, \ldots, \theta_S, \theta^b_{S+1}, \ldots, \theta^b_K)$. In other words a certain number S of selected problem-dependent optimisation parameters Θ_1, Θ_2, , ..., Θ_S are considered to be uncertain and subjected to fluctuations within some range. The function F is constructed as a sum of squares in the measurement points between the values of model predictions and values observed in terrain:

$$F(\theta_1, \ldots, \theta_S) = \sum_{r=1}^{r=R} (y_r - \dot{z}_r(\theta_1, \ldots, \theta_S))^2 \qquad (2)$$

Minimisation algorithm searches a minimum of scalar function F of S parameters starting at an initial "best estimate". In brief glance, the test points $[\theta_1, \theta_2, \ldots, \theta_S]$ of the objective function F are arranged as a S-dimensional simplex and the algorithm tries to replace iteratively individual points with aim to shrink the simplex towards the best points. Model predictions can be interpreted as Gaussian surface (or superposition of partial Gaussian extents) over the terrain. Our objective is to take into account both model predictions and available measurements incoming from the terrain and to improve spatial distribution of deposited radioactivity. We can imagine the iterative process of minimisation of function F such consecutive adjustment of the resulting respond surface, always according to the new evaluation of the parameters $[\theta_1, \theta_2, \ldots, \theta_S]$. Thus, the predicted respond surface of results is gradually "deformed by permissible manipulations" directly driven by changes of problem-dependent optimisation parameters θ_s. MT algorithm controls the procedure until the best fit of modified surface with observation values is reached. Important feature of the method insists in preservation of physical knowledge, because the new set of parameters $[\theta_1, \theta_2, \ldots, \theta_S]$ evaluated by minimisation algorithm al-

ways re-enters the entire nonlinear environmental model M according to Equation (1).

PRACTICAL IMPLEMENTATION AND RESULTS

Investigation of applicability of minimisation assimilation technique was tested on so called "twin experiment". Lack of real observations is simulated by artificial generation of measurements. Moreover, if we use for this generation the same environmental model (e.g. for a fix one set of disturbed input parameters) we can examine the problem convergence issues. In application part of the paper the results of two simulation experiments TWIN1 and TWIN2 are illustrated. TWIN1 relates to release of nuclide [131]I. Its further straight-line propagation and deposition on terrain is simulated according to simple scheme of straight-line Gaussian plume model. TWIN2 experiment deals with the problem of evolution of ^{137}Cs deposition on terrain during the plume phase. Minimisation search is applied with more complicated but more realistic segmented model SGPM.

MT applied to simple Gaussian straight-line model

Accidental one-hour release of radionuclide ^{131}I with total radioactivity 1.28 E+11 Bq discharged into atmosphere from nuclear facility is analysed. Release height is 100 m, propagation continues under constant meteorological conditions (straight-line propagation in direction North-East, mean plume velocity 1.6 m.s^{-1}, Pasquill category D of atmospheric stability, no rain). Atmospheric dispersion coefficients are calculated according to KFK-Jülich semi-empirical formulas.

In the first step all input parameters are assumed to be represented by their best estimate values denoted by θ_i^b and then the corresponding output vector Z^b presents deterministic solution of deposited activity of selected nuclide on terrain. At the same time Z^b represents initial estimate for starting of minimization iterative search. In the second step we shall further reduce the number of parameters S from equation (2) to four parameters. Corresponding four uncertainties c_1, c_2, c_3, c_4 are introduced into the model according to scheme $\theta_i = c_i \cdot \theta_i^b$ or $\theta_i = \theta_i^b + c_i \cdot f(\theta_i^b)$. Specifically, their meaning, usage in the environmental code and real choice is given in Table 1.

The function $F(\theta_1, \theta_2, ..., \theta_S)$ from (2) now has form $F(c_1, c_2, c_3, c_4)$ and minimisation algorithm handles with 4-dimensional simplex. For purposes of construction of function F we have used slight modification of probabilistic version of existing environmental model HARP (Pecha et al. 2007) where original random inputs c_1, c_2, c_3, c_4 now play more general role of uncertainties characterized only by their range of possible fluctuations (see column 4 in Table 1). Minimisation algorithm uses this constraints such lower and upper bounds for permissible manipulations with values of variables c_1, c_2, c_3, c_4 (see arrows in Figure 1). During TWIN experiments the observation vector $Y \equiv (y_1, y_2, ..., y_R)$ is simulated artificially, the simplest way is utilization of the same environmental model M.

Table 1. Introduction of Uncertainties for Four Important Input Model Parameters

parameter	Unit	uncertainty inside code	uncertainty bounds
θ_1: Source release rate	[Bq.s^{-1}]	$Q = c_1 \cdot Q^b$	$c_1 \in \langle 0.1; 2.9 \rangle$
θ_2: horizont. dispersion	[m]	$\sigma_y(x) = c_2 * \sigma_y(x)^b$	$c_2 \in \langle 0.1; 3.1 \rangle$
θ_3: Wind direction	[rad]	$\varphi = \varphi^b + \Delta\varphi$, $\Delta\varphi = c_3 * 2\pi/80$	$c_3 \in \langle -5.0; 5.0 \rangle$
θ_4: Dry depo velocity	[m.s^{-1}]	$v_g = c_4 * v_g^b$	$c_4 \in \langle 0.1; 4.0 \rangle$

Deterministic best estimate distribution Z^b generated on the polar calculation grid in original wind direction S_{orig} (North-East) is drawn in figure 2 as TRACE I . It corresponds to the best estimate values $\{ c_1, c_2, c_3, c_4 \}_{best} \equiv \{ 1.0, 1.0, 0.0, 1.0 \}$. Selected positions of observations are labelled by red circles. For simulation of measurements in this red points we have selected a certain fixed quartet $\{ c_1, c_2, c_3, c_4 \}^{obs} \equiv \{ 1.73, 1.51, +4.00, 1.98 \}$. Artificially simulated measurements were generated using vector $Z^{obs} = M (\{ c_1, c_2, c_3, c_4 \}^{obs})$. Then the values are transformed into observation positions according to $\dot{Z}^{obs} = H \cdot Z^{obs}$. Final simulated observation vector is obtained by assignment $Y \equiv \dot{Z}^{obs}$.
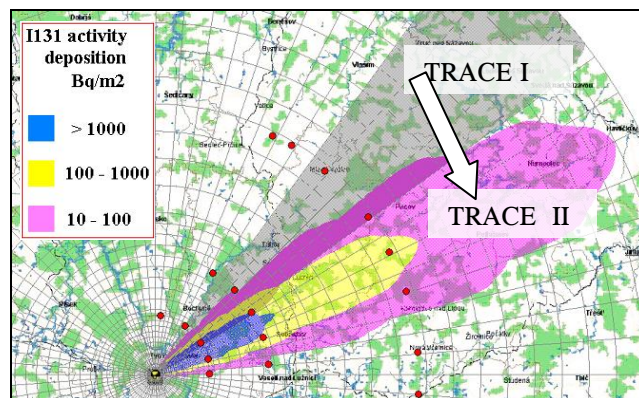

Figure 2: [131]I Deposition Levels [Bq.m^{-2}] Related to the End of Plume Progression. TWIN I experiment using Gaussian straight-line model. TRACE I and TRACE II are initial best estimate and resulting assimilation with simulated measurements (at red circles)

Minimisation algorithm in successive iterations j brings newly generated quartets $\{ c_1, c_2, c_3, c_4 \}^j$ closer and closer to the $\{ c_1, c_2, c_3, c_4 \}^{obs}$. Fast convergence of assimilated model predictions towards simulated observations has been found. 220 iterations are calculated during about 6 minutes and the following values has been found: $\{ c_1, c_2, c_3, c_4 \}^{j=220} = \{ 1.731, 1.514, +4.003, 1.982 \}$. It demonstrates very good consent with "simulated" observations generated by $\{ c_1, c_2, c_3, c_4 \}^{obs}$. The results are illustrated in Figure 2 as TRACE II isolines.

Original best estimate deposition on terrain (and at the same time initial guess entering MT) is labelled as TRACE I. Deposition after 220 iterations is calculated as $Z^{j=220} = M (\{ c_1, c_2, c_3, c_4 \}^{j=220})$ and its isolines illustrates TRACE II. The assimilated respond surface TRACE II is at the same time practically identical with

$\mathbf{Z^{obs}}$ generated according to M ({ c_1, c_2, c_3, c_4}^{obs}) originally used for artificial simulations of measurements. The shapes of TRACE I and TRACE II reflect imposed changes in values of c_1^{best} to c_1^{obs} (higher nuclide discharge), c_2^{best} to c_2^{obs}(higher peripheral dispersion), c_3^{best} to c_3^{obs}(twist by 18°), c_4^{best} to c_4^{obs}(more intensive dry deposition causing steeper longitudinal gradient).

Direct search algorithm connected with Gaussian straight-line propagation model has proved fast convergence provided that the measurements are well positioned. Its applicability depends on validity and limitations of model itself (more e.g. in (Irwing 2004)). However, the TWIN 1 results support an idea of MT application for preliminary fleeting estimation in near distances and during constant meteorological conditions.

MT with more realistic SGPM environmental model

TWIN2 scenario is formulated in connection with segmented Gaussian plume scheme (model SGPM marked as M^{SGPM}), which is much more complicated then straight-line spreading (our approach described in (Pecha et al. 2007)). The model synchronizes segmentation of release dynamics with hourly meteorological forecasts. The first two consecutive release segments of ^{137}Cs discharge (each with 1 hour duration) with released amount 2.0 E+17 Bq and 1.0 E+17 Bq has character of severe LOCA accident with partial fuel cladding rupture and fuel melting. Short-term meteorological forecast for the next 48 hours is provided by the Czech meteorological service. Then, for each hour since the release initiation there are available predictions of wind direction and speed, category of atmospheric stability according to Pasquill classification and rain precipitation. Omitting other details, the TWIN II scenario covers period of the first 3 hours from the release start and we are declaring the following plan:

i) Each of the two segments is modelled up to third hour after the release start taking into account short-term hourly meteorological forecast. The situation just after 3 hours is given by superposition of both segments in their successive meteorological hourly phases (5 phases in total). Resulting best estimate fields are calculated in analogy with Equation (1) according to scheme $\mathbf{Z^b_{3hour}} = M^{SGPM}$ ({ c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53} }$_{best}$) and is illustrated in Figure 3a as TRACE I.

ii) Number of uncertainties is increased from four to five as c_1, c_2, c_3, c_4, c_5. c_5 stands for fluctuation of mean wind velocity. If we suppose wind direction and velocity fluctuations to be independent between hourly phases, then c_3 and c_5 split to 6 independent uncertainties c_{31}, c_{32}, c_{33} (for wind direction predicted for hours 1, 2 ,3) and c_{51}, c_{52}, c_{53} (for wind velocity predicted for hours 1, 2 ,3).

iii) We have simulated artificially fictive "observation surface" according to $\mathbf{Z^{obs}_{3hour}} = M^{SGPM}$ ({ c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53}}^{obs}). Vector of simulated measurements at observation positions (see black filled squares in Figure 3b) are calculated by help of linear observation operator as $\mathbf{Y_{3hour}} \equiv H \mathbf{Z^{obs}_{3hour}}$. Their incoming is supposed at one stroke just at hour 3 after the accident start. Let us state beforehand that assimilated TRACE II from Fig. 3b nearly corresponds with the "observation surface".

iv) The main goal is to accomplish assimilation of the model predictions $\mathbf{Z^b_{3hour}}$ in compliance with measurements $\mathbf{Y_{3hour}}$ in analogy with equation (2) using BCPOL procedure of minimisation.
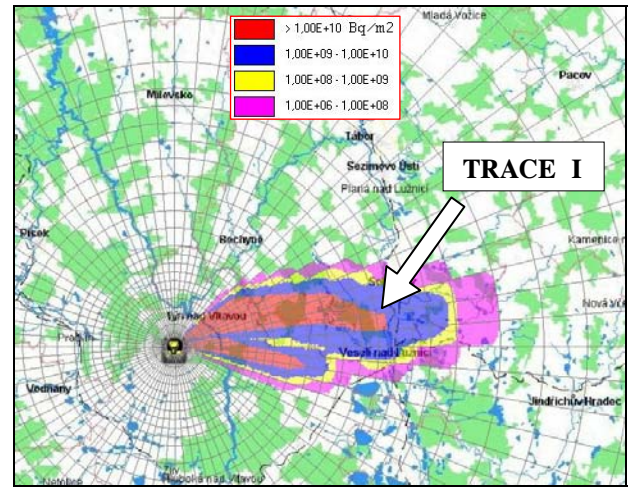

Figure 3a : Nominal Deposition of ^{137}Cs (just 3 Hours after the Release Start)

Deposition of ^{137}Cs on terrain after 728 iterations is calculated as $\mathbf{Z^{j=728}_{3hour}} = M^{SGPM}$ ({ c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53}}$_{j=728}$) and its isolines illustrates in Figure 3b a trail on terrain marked as TRACE II. The results represent a new distribution just at third hour after the release start, which is improved by observations. Minimisation algorithm is initiated by the best estimate solution (TRACE I) and gradually approaches to the simulated observations. In short numerical summary, TWIN2 experiment required to prepare in advance sets of parameters { c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53}} for:

best estimate: { ….}^{b} = 1.0, 1.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0 [*]

"measurements": { ….}^{obs} = 7.0, 2.0, -4.0, -5.0, -6.0, 2.55, -0.5, -0.6, -0.7 [**]

Here are examples for a particular iteration j :

{ ….}$^{j=728}$ = 7.18, 2.49, -3.94, -5.80, -6.34, 2.49, 0.21, -0.28, -0.59 (***)
{ ….}$^{j=1201}$ = 7.25, 2.03, -4.14, -5.80, -6.39, 2.59, 0.27, -0.36, -0.58

[*] *TRACE I in Figure 3a,* [**] *close to TRACE II in Figure 3b,* [***] *TRACE II in Figure 3b*

Meaning of the parameters c_1 to c_4 is the same as described in Table 1. c_5 stands for uncertainty of the mean velocity of the plume. Further splitting to c_{5i}, i=1,2,3 holds true for independent fluctuations of the mean velocity $\bar{u}_i$ forecasted for hours i. Uncertain $\bar{u}_i$ is then expressed according to $\bar{u}_i = \bar{u}_i^{best}$ (1+0.35* c_{5i}). c_{5i} bounds are <-1; +1>. More detailed recommendations for uncertainty bounds arising from expert judgement can be found e.g. in (Goossens at al. 2001).

TWIN II experiment took into consideration 9 optimisation parameters with constructive idea to discriminate according to their global or local effect (introduced into the

wind vector). The computation procedure is time consuming, but satisfactory convergence can be achieved. The question of real field measurements (e.g. Eleveld 2004) is so far opened.
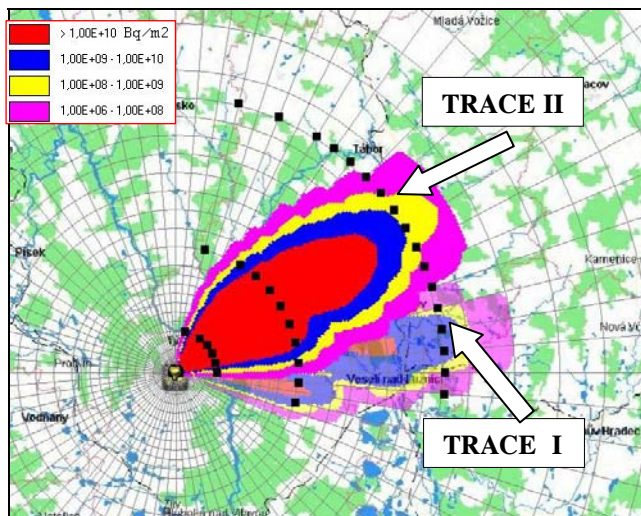

Figure 3b: Assimilation of Predicted Deposition of ^{137}Cs with Measurements just 3 Hours after the Release Start (artificially simulated measurements in black squares)

CONCLUSION

Advantage of utilisation of SGPM output fields as a fitting surface insists in preservation of physical knowledge of the model. Presented experience related to applicability of minimisation techniques indicates that number of selected optimisation parameters c_i should not be too high in order to avoid the poor convergence or even taking the wrong way (more sophisticated algorithms have to be tested). At this stage we recommend to consider five optimisation parameters included in the TWIN II experiment (where wind velocity vector is global, it means no further splitting of c_3 to further c_{3i} and c_5 to c_{5i}) and link the 6th parameter c_6 representing uncertainty in precipitation intensity forecast.

Presented minimisation technique fits the model simulation results on a certain specific situation. Any resulting effect (e.g. peripheral plume dispersion) usually depends on many other input random parameters. Thus, in no case the presented fitting technique should not be confused with parameter calibration. The problem of handling of real measurements still remains opened, the first considerations for the Czech territory are discussed in (Kuca et al. 2008). Presented approach can play a specific role among empirical assimilation techniques, especially as fast and efficient software tool for analysis of possible discrepancies between the model predictions and observations incoming from terrain. The method is incorporated into assimilation subsystem the HARP code (Hofman et al. 2007).

Realistic prediction of evolution of radiation situation during emergency gives decision makers necessary time on judgement and introduction of efficient urgent countermeasures on population protection. Reliable model predictions for the next hours in medium distances should account both for implementation of spatial meteorological forecast and development of new numerical techniques for time update of the

trajectory models (e.g. how to propagate model for the next hours starting from assimilated results TRACE II in Figure 3b). Interventions introduced on the basis of non-assimilated TRACE I could lead to fatal consequences on population health resulting from ill-anticipated impacted areas.

REFERENCES

Drécourt, J-P. 2004. "Data assimilation in hydrological modeling". Environmental&Resources, Techn. Univ. of Denmark, ISBN 87-89220-84-6..

Eleveld, H. and C. J. W. Twenhofel. 2004. "Improving the Reliability of the Prognosis of Atmospheric Dispersion Using Data Assimilation". *In Proc. Int. Conf.. on Harmonisation within Atmospheric Dispersion Modelling* (Garmisch-Partenkirchen, Germany), Vol. 1,49-53.

Goossens, L.H.J., B.C.P Kraan, R.M. Cooke, J.A.Jones, J. Ehrhardt, F. Fischer and I. Hasemann. 2001. "Uncertainty from the Atmospheric Dispersion and Deposition Module". Research report EUR 18822 EN.

Irwing, J.S. and S. R. Hanna. 2004. "Characterizing Uncertainty in Plume Dispersion Models". I*n Proc. Int. Conf.. on Harmonisation within Atmospheric Dispersion Modelling* (Garmisch-Partenkirchen, Germany), Vol.1, 287-292.

Hofman, R. and P. Pecha. 2007. **"**Integration of data assimilation subsystem into environmental model of harmful substances propagation". In Proc. *11-th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling* (Cambridge, UK, July 2-5, 2007), 111-115.

Kalnay, E. 2003. *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, ISBN 0-521-79179-0.

Kuca, P., R. Hofman and P. Pecha. 2008. **"**Assimilation techniques in consequence assessment of accidental radioactivity releases – the way for increase of reliability of predictions". In. *Proc. of ECORAD 2008 Int. Conf. on Radioecology & Environ.Radioactivity* (Bergen, Norway, 15-20 June 2008), 138-141, ISBN 978-82-90362-25-1.

Pecha, P., R. Hofman and E. Pechova. 2007. **"**Training simulator for analysis of environmental consequences of accidental radioactivity releases". In *Proc. 6th EUROSIM Congress on Modelling and Simulation* (Ljubljana, Slovenia, September 9-13, 2007), 18 pp. on conf. CD – ISBN 978-3-901608-32-2.

Pecha, P. and E. Pechova. 2005. "Modeling of random activity concentration fields in air … ". *In Proc. 10-th Int. Conf.. on Harmonisation within Atmospheric Dispersion Modelling* (Sissi, Greece), paper No. H11-069.

Rojas Palma, C. (ed.). 2005. "Data Assimilation for Off site Nuclear Emergency Management". Research Report RODOS(RA5)-RE(04)-01.