Exploitation of particle filter for forecast error covariance structure estimation

Radek Hofman
Institute of Information Theory and Automation ASCR
Pod Vodárenskou věží 4
18208 Praha 8
hofman@utia.cas.cz

***Abstract*— The paper presents a scheme for estimation of spatio–temporal evolution of a quantity with unknown model error. Model error is estimated on basis of measured–minus–observed residuals evaluated upon measured and modeled values. Methods of Bayesian filtering are applied to the problem. The main contribution of this paper is application of general marginalized particle filter algorithm to the linear–Gaussian problem with unknown model error covariance structure. Methodology is demonstrated on the problem of modeling of spatio–temporal evolution of groundshine–dose from radionuclides deposited on terrain in long–time horizon.**

***Index Terms*— marginalized particle filter, groundshine–dose, Bayesian estimation, covariance structure estimation**

I. Introduction

The task of estimation of time evolution of a spatially distributed quantity is widely applied in different branches of "Earth sciences" such as meteorology and oceanography [11]. During the last years, there have arisen tendencies for application of an advanced data assimilation algorithms also in the field of radiation protection [15], [18], [19]. It is related to the renaissance of nuclear energy which can be observed. The application of advanced statistical methods can increase reliability of consequence predictions of possible releases from nuclear power–plants. Their reliability is in the field of radiation protection mission–critical as the problem deals with the population health.

There were developeded several models for modeling of evolution of living environment contamination for different release scenarios. The only connection with physical reality are measurements with errors (sparse both in time and in space). In our work, we attempt to make groundshine–dose model predictions more reliable in a way of adjusting them towards measurements incoming from terrain. This process is called data assimilation [11]. Its principle consists in combining of the information provided by the model with the measured data. Exploiting information on sources of uncertainty, we are able to produce improved estimate of the true situation on terrain.

If the problem is treated as linear–Gaussian, it can be successfully solved via Kalman filter (KF) [10]. The unavoidable condition for utilization of Kalman filter is knowledge of model error covariance structure but in many cases it is unknown due to the problem background. In this paper is presented methodology for application of the Kalman filter to the problems where the model error covariance structure is unknown and has to be estimated upon actual data before application of the filter. This results in marginalized particle filter described in [21].

Model error covariance is represented by a covariance matrix. As the total number of its elements is much higher the number of measurements, we can't estimate all of them. Simplified model error covariance structure parametrization based on idealized assumptions is introduced. For finding the most plausible values of these parameters, the similar approach as proposed in [3] or [14] based on modeled–minus–observed residuals is used. Instead of maximum likelihood estimates, we use marginalized particle filter for estimation of both the model error covariance parameters and groundshine–dose distribution. The marginalized particle filter is a powerful combination of the particle filter and the Kalman filter, which can be used when the underlying model contains a linear substructure which is being subject to Gaussian noise.

The performance of this methodology is demonstrated on modeling of groundshine–dose evolution in long–time horizon of several months [6]. As the problem is complex, the groudshine–dose evolution model is an idealized approximation of the true physical process. Calculations are performed on a subset of polar network around the source of pollution. The model error covariance parametrization proposed here follows the physical background of the problem.

The outline of this paper is as follows. Section II briefly discusses Bayesian filtering. Kalman filter, particle filter and marginalized particle filter are successively presented there. In Section III, the assimilation algorithm is proposed and the problem of model error covariance estimation is described. Application of the algorithm on modeling of long–term evolution of groundshine–dose from radionuclide deposition on terrain is presented in Section IV. Specific model error covariance parametrization suitable for the problem is developed there. In Section V, experimental results with simulated measurements are demonstrated and the conclusion is given.

This work is part of the grant project GAČR No. 102/07/1596, which is funded by the Grant Agency of the Czech Republic.

II. BAYESIAN FILTERING

Bayesian approach to filtering is applicable to all linear or nonlinear stochastic systems [7], [12]. Let the stochastic system be defined by discrete–time state–space transition equation (1) and observation equation (2)

$$\boldsymbol{x}_t = f(\boldsymbol{x}_{t-1}) + \boldsymbol{b}_t \tag{1}$$

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t) + \boldsymbol{\epsilon}_t \tag{2}$$

Here, t is time index, $\boldsymbol{x}_t$ is unobservable system state vector, $\boldsymbol{b}_t$ is the additive dynamic noise vector. Vector $\boldsymbol{y}_t$ is the measurement vector which provides us indirect information about the system state and $\boldsymbol{\epsilon}_t$ its noise. Both the densities of noise terms are assumed to be independent and known. Functions $f(\cdot)$ and $h(\cdot)$ are generally non–linear. State transition function $f(\cdot)$ propagates the prior state to the current one. Forward observation operator $h(\cdot)$ transforms vectors from state–space to the measurement space, thus constitutes relation of the actual measurements to the current state.

The goal is to acquire posterior density $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ where $\boldsymbol{Y}_t = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$ are available measurements. In the following text, the state process $\{\boldsymbol{x}_t\}$ is assumed to be Markovian of the first order. It means that given the present state, future states are independent of the past states:

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{x}_{t-2}, \ldots, \boldsymbol{x}_0) = p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) \tag{3}$$

Realization of the process at time t contains all information about the past, which is necessary in order to calculate the future behavior of the process.

Bayesian estimation procedure consists of two recursively repeated step. The first step transits the state estimate to the next time step according to the probability density function (PDF) $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$. This step is called time update (4). In the second step called data update (5), the information provided by actual measurements $\boldsymbol{y}_t$ is included into the current estimate given by the PDF $p(\boldsymbol{x}_t|\boldsymbol{Y}_{t-1})$.

$$p(\boldsymbol{x}_t|\boldsymbol{Y}_{t-1}) = \int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) p(\boldsymbol{x}_{t-1}|\boldsymbol{Y}_{t-1}) d\boldsymbol{x}_{t-1} \tag{4}$$

$$p(\boldsymbol{x}_t|\boldsymbol{Y}_t) = \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{Y}_{t-1})}{\int p(\boldsymbol{y}_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{Y}_{t-1}) d\boldsymbol{x}_t} \tag{5}$$

The state evolution is initialized by a probability density function $p(\boldsymbol{x}_0|\boldsymbol{Y}_{-1}) = p(\boldsymbol{x}_0)$ which represents all the prior information on the problem and also our subjective judgments. This density is often called background–field or just the prior.

If both the measurement density $p(\boldsymbol{y}_t|\boldsymbol{x}_t)$ and the state transition density $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ are parametric, the problem can be solved analytically. Provided that the system is linear–Gaussian, the integrals (4, 5) lead to KF recursion.

A. Kalman filter

In the following text $N(\boldsymbol{\mu}, \boldsymbol{Q})$ is assumed to be a Gaussian PDF with mean value $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{Q}$. KF is simple implementation of the Bayesian filter and it provides the optimal Bayesian solution. Its usage is limited to the case of linear estimation with the Gaussian noise where

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = N(\boldsymbol{M}\boldsymbol{x}_{t-1}, \boldsymbol{Q}_t) \tag{6}$$

$$p(\boldsymbol{y}_t|\boldsymbol{x}_t) = N(\boldsymbol{H}\boldsymbol{x}_t, \boldsymbol{R}_t) \tag{7}$$

Matrices $\boldsymbol{M}$ and $\boldsymbol{H}$ are matrices of linear (linearized) operators $f(\cdot)$ and $h(\cdot)$, respectively. Matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ are known error covariance matrices of model error and measurement error, respectively. Under these assumptions (4, 5) lead to KF equations for time update and data update steps [10]. The equations perform recursive update of the first two moments of estimated Gaussian distribution $p(\boldsymbol{x}|\boldsymbol{Y}) = N(\hat{\boldsymbol{x}}, \boldsymbol{P})$ – the mean value $\hat{\boldsymbol{x}}$ and its covariance matrix $\boldsymbol{P}$.

B. Particle filter

In more general cases where analytical solution of integrals (4, 5) is not known, there are their numerical approximations based on sequential Monte Carlo methods also known as particle filters.

Particle filter (PF) is more general implementation of Bayesian filter which can be used to approximate the posterior density function for the state in non–linear and non–Gaussian filtering problems [7]. It is based on recursive estimation of the PDF $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ which is represented as a set of M so called particles $\boldsymbol{x}_t^{(i)}$ and its associated normalized weights $\tilde{q}_t^{(i)}$ as $\{\tilde{q}_t^{(i)}, \boldsymbol{x}_t^{(i)}\}|_{i=1\ldots M}$. The posterior PDF $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ can be approximated with their help as $\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t)$.

$$p(\boldsymbol{x}_t|\boldsymbol{Y}_t) \approx \hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t) = \sum_{i=1}^{M} \frac{1}{M} \delta(\boldsymbol{x}_t - \boldsymbol{x}_t^{(i)}), \tag{8}$$

where $\delta(\cdot)$ is the Dirac δ-function and $\boldsymbol{x}_t^{(i)}$ are samples from approximated PDF. In (8), all the weights $\tilde{q}_t^{(i)}$ are equal to $\frac{1}{M}$. Our goal is usually to estimate the mean value of a function defined on our approximated distribution $E_{\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t)}[g(\boldsymbol{x}_t)]$. The approximation $\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ satisfies condition

$$\lim_{M \to +\infty} E_{\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t)}[g(\boldsymbol{x}_t)] \overset{a.s.}{\to} E_{p(\boldsymbol{x}_t|\boldsymbol{Y}_t)}[g(\boldsymbol{x}_t)], \tag{9}$$

where $\overset{a.s.}{\to}$ means almost sure convergence and $g(\boldsymbol{x}_t)$ is arbitrary function bounded for support $\Omega = \{\boldsymbol{x}_t | p(\boldsymbol{x}_t|\boldsymbol{Y}_t) > 0\}$.

In real cases we are not able to sample directly from $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ but we are able to evaluate it in discrete points (at least up to proportionality). We can draw independent samples $\boldsymbol{x}_t^{(i)}$ from a chosen known proposal distribution (importance function) $q(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ and use them for approximating of $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$. The estimated density $p(\boldsymbol{x}_t|\boldsymbol{Y}_t)$, its approximation $\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ and importance function $q(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ are related as follows

$$p(\boldsymbol{x}_t|\boldsymbol{Y}_t) = \frac{p(\boldsymbol{x}_t|\boldsymbol{Y}_t)}{q(\boldsymbol{x}_t|\boldsymbol{Y}_t)} q(\boldsymbol{x}_t|\boldsymbol{Y}_t) \approx$$

$$\approx \hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t) = \sum_{i=1}^{M} \frac{p(\boldsymbol{x}_t^{(i)}|\boldsymbol{Y}_t)}{q(\boldsymbol{x}_t^{(i)}|\boldsymbol{Y}_t)} \frac{1}{M} \delta(\boldsymbol{x}_t - \boldsymbol{x}_t^{(i)}), \tag{10}$$

where $\frac{1}{M}\sum_{i=1}^{M}\delta(\boldsymbol{x}_t-\boldsymbol{x}_t^{(i)})$ is an approximation of $q(\boldsymbol{x}_t|\boldsymbol{Y}_t)$ since $\boldsymbol{x}_t^{(i)}$ are sampled from this PDF. If we denote $q_t^{(i)} = \frac{p(\boldsymbol{x}_t^{(i)}|\boldsymbol{Y}_t)}{q(\boldsymbol{x}_t^{(i)}|\boldsymbol{Y}_t)}\frac{1}{M}$, the estimated PDF can be approximated as

$$\hat{p}(\boldsymbol{x}_t|\boldsymbol{Y}_t) = \sum_{i=1}^{M}\tilde{q}_t^{(i)}\delta(\boldsymbol{x}_t-\boldsymbol{x}_t^{(i)}), \tag{11}$$

where $\tilde{q}_t^{(i)} = q_t^{(i)}/\sum_{j=1}^{M}q_t^{(j)}$, $\sum_i^M \tilde{q}_t^{(i)} = 1$, $\tilde{q}_t^{(i)} \geq 0$ are normalized weights. This normalization will for finite M introduce a bias in the estimate. However, from the strong law of large numbers the estimate is asymptotically unbiased. This algorithm is called sampling–importance–sampling (SIS).

If we choose the posterior PDF from the previous step as proposal distribution in the current, we can via recursive evaluation of normalized weights perform Bayesian filtering. In this case will weight update result in

$$q_t^{(i)} \propto \tilde{q}_{t-1}^{(i)}p(\boldsymbol{y}_t|\boldsymbol{x}_t^{(i)}) \tag{12}$$

This algorithm also suffers from degeneracy problem, so we have to implement a resampling algorithm, more in [4]. Resampling should eliminate particles with small weights and multiply particles with large weights. After resampling all the weights are set to $\frac{1}{M}$. If we perform resampling in each step, the weights can be computed as $q_t^{(i)} = p(\boldsymbol{y}_t|\boldsymbol{x}_t^{(i)})$. This modification of SIS algorithm with resampling in each step is also known as sampling–importance–resampling (SIR).

Disadvantage of this method is that we have to be able to generate random samples from complicated distributions and this is for high dimensional problems computationally prohibitive. The computational complexity rapidly increases along with the state–space dimension.

C. Marginalized particle filter

When structure of the model (1, 2) allows analytical marginalization over a subset of states, we can reduce the computational burden. Let's consider factorization of the state vector $\boldsymbol{x}_t = \begin{bmatrix}\boldsymbol{x}_t^l & \boldsymbol{x}_t^n\end{bmatrix}^T$ where $\boldsymbol{x}_t^l$ is a subset of analytically tractable states and $\boldsymbol{x}_t^n$ is the rest. Provided that the $\boldsymbol{x}_t^l$ and $\boldsymbol{x}_t^n$ are conditionally independent, substitution of the factorization into (8) and application of the chain rule gives

$$p(\boldsymbol{x}_t^l, \boldsymbol{x}_t^n|\boldsymbol{Y}_t) = p(\boldsymbol{x}_t^l|\boldsymbol{x}_t^n, \boldsymbol{Y}_t)p(\boldsymbol{x}_t^n|\boldsymbol{Y}_t), \tag{13}$$

where $p(\boldsymbol{x}_t^l|\boldsymbol{x}_t^n, \boldsymbol{Y}_t)$ is analytically tractable and $\boldsymbol{x}_t^n$ is given by the particle filter. Assuming that $\boldsymbol{x}_0^l \sim N(\hat{\boldsymbol{x}}_0, \boldsymbol{P}_0)$ and to be governed by a linear model implies that $p(\boldsymbol{x}_t^l|\boldsymbol{x}_t^n, Y_t)$ is conditionally linear–Gaussian and can be computed via Kalman filter [22]. Substitution of (8) into (13) for $\boldsymbol{x}_t^n$ leads to

$$p(\boldsymbol{x}_t|\boldsymbol{Y}_t) \approx \sum_{i=1}^{M}\tilde{q}_t^{(i)}\delta(\boldsymbol{x}_t^n-\boldsymbol{x}_t^{n,(i)})N(\hat{\boldsymbol{x}}_t^{l,(i)}, \boldsymbol{P}_t^{(i)}) \tag{14}$$

The joint PDF is estimated as a mixture of a parametric distribution of Gaussian type and of a nonparametric one. The estimated PDF is represented by a weighted sum of Gaussians, where each particle has a Gaussian distribution attached to it. This modification of PF is called marginalized particle filter (MPF) and details on its implementation can be found in [21], [22].

III. Assimilation procedure based on MPF

The unavoidable condition for application of Kalman filter is knowledge of model error represented in (1) by the noise vetor $\boldsymbol{b}_t$. We assume $\{\boldsymbol{b}_t\}$ to be the white noise process where $\boldsymbol{b}_t \sim N(\boldsymbol{0}, \boldsymbol{Q}_t)$. Matrix $\boldsymbol{Q}_t$ is corresponding covariance matrix. The value of $\boldsymbol{Q}$ should reflect total (unknown) model error, which is in each step contribution to the forecast error due to differences between the model and the true process. In KF [10], forecast error covariance matrix $\boldsymbol{P}$ evolves as

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{M}_{t|t-1}\boldsymbol{P}_{t-1|t-1}\boldsymbol{M}_{t|t-1}^T + \boldsymbol{Q}_t, \tag{15}$$

where $\boldsymbol{M}$ is matrix of linear (linearized) operator for the state transition from time $t-1$ to t. It is obvious that if $\boldsymbol{Q}$ is neglected, the predicted forecast error will be underestimated. This could cause divergence from the true state (its good estimate) because smaller model error will handicap the information provided by measurements.

We assume that the $\boldsymbol{Q}$ is unknown and attempt to estimate it in each assimilation step. As the total number of elements of $\boldsymbol{Q}$ to be estimated is much higher than the number of measurements, we can't estimate all of them. Simplified covariance model based on idealized assumptions has to be introduced.

Schematically, let the model error covariance matrix be approximated as a function $Q(\boldsymbol{\theta}) : \Re^{dim(\boldsymbol{\theta})} \rightarrow \Re^{[dim(\boldsymbol{x}),dim(\boldsymbol{x})]}$ of a parameter vector $\boldsymbol{\theta}$, where $\Re^{[m,n]}$ is a space of real matrices of dimension $m \times n$.

$$\boldsymbol{Q}_t = Q_t(\boldsymbol{\theta}_t) \tag{16}$$

Function Q has to be chosen properly in order to produce positive semi–definite symmetric matrices which can be covariance matrices.

For finding the most plausible values of $\boldsymbol{\theta}$ a similar approach as proposed in [3], [14] based on modeled–minus–observed residuals is used. Instead of maximum likelihood estimates proposed there we use MPF introduced in Section II. When the measurements are available, we can evaluate residual vector $\boldsymbol{v}_t = \boldsymbol{y}_t - \boldsymbol{H}\hat{\boldsymbol{x}}_t$ having the same dimension as the measurement vector. Covariance of $\boldsymbol{v}$ derived in [3] has the form

$$E[\boldsymbol{v}_t\boldsymbol{v}_t^T] = \boldsymbol{H}_t\boldsymbol{P}_{t|t-1}\boldsymbol{H}_t^T + \boldsymbol{R}_t = \boldsymbol{S}_t \tag{17}$$

We assume $\boldsymbol{v_t} \sim N(\boldsymbol{0}, \boldsymbol{S}_t)$. If we substitute (15) into (17) for $\boldsymbol{P}_{t|t-1}$ and use covariance parametrization (16) of $\boldsymbol{Q}_t$ we obtain

$$\boldsymbol{S}_t(\boldsymbol{\theta}) = \boldsymbol{H}_t[\boldsymbol{M}_t\boldsymbol{P}_{t-1|t-1}\boldsymbol{M}_t^T + Q_t(\boldsymbol{\theta})]\boldsymbol{H}_t + \boldsymbol{R}_t \tag{18}$$

From (15) can be seen that the parametrization of model error covariance leads to parametrization of forecast error covariance $\boldsymbol{P}$. The most plausible value of parameters are found in each time step via PF from likelihood $p(\boldsymbol{v}_t^{(i)}|\boldsymbol{\theta}_t^{(i)}) =$

$N(\mathbf{0}, \boldsymbol{S}(\boldsymbol{\theta}_t^{(i)}))$ for random parameter vectors $\boldsymbol{\theta}_t^{(1)}, \ldots, \boldsymbol{\theta}_t^{(M)}$ and corresponding residual vectors $\boldsymbol{v}_t^{(i)}$. The likelihood is the higher, the higher is the probability that difference between modeled and measured values is zero given covariance (18). These parameters are then used in (15, 16) for forecast error propagation. Incorporation of this algorithm into KF assimilation scheme results in MPF for estimation of joint PDF $p(\boldsymbol{x}_t, \boldsymbol{\theta}_t | \boldsymbol{Y}_t)$ which is the mixture of Gaussian and nonparametric distributions

$$\underbrace{p(\boldsymbol{x}_t, \boldsymbol{\theta}_t | \boldsymbol{Y}_t)}_{MPF} = \underbrace{p(\boldsymbol{x}_t | \boldsymbol{\theta}_t, \boldsymbol{Y}_t)}_{KF} \underbrace{p(\boldsymbol{\theta}_t | \boldsymbol{Y}_t)}_{PF}, \qquad (19)$$

where $\boldsymbol{x}_t$ is the state vector and $\boldsymbol{\theta}_t$ is the vector of parameters used for estimation of current model error covariance structure.

IV. Assimilation scenario

The algorithm described in Section III is demonstrated on assimilation scenario introduced in this section.

In case of an accidental aerial release of radioactive pollutants into the living environment, the radioactive plume is depleted during passing over the terrain. This phase is called the plume phase. Due to the deposition processes the plume leaves a radioactive trace on the ground.

After the plume phase (when the radioactive cloud exits the area of interest) post–emergency phase follows. It covers latter stages of accident consequence evolution. Post–emergency phase may extend over a prolonged period of several weeks or many years depending on the source of radiation and local conditions. It ends when environmental radiation levels resume to normal. The main exposure pathways in this phase are groundshine and ingestion. The deposited material cause irradiation and through the root system migrates to the edible parts of crops consumed by people and livestock. Among many radionuclides released during emergency situations we focus only on ^{137}Cs since it is one of the most important nuclides in long–time perspective. Its half–time of decay is long (30 years) and also analysis after the Chernobyl accident had shown that it is one of the most significant nuclides in these types of accidents having detrimental long–term effect on population health.

Our assimilation scenario covers the post–emergency phase. The source of pollution is placed into the centre of polar network. We perform our calculations on subset of this network in successive time steps $t \in \{0, 1, \ldots, t_{MAX}\}$. Groundshine–dose in ordered set of analyzed spatial points forms our state vector $\boldsymbol{x}$. We assume $\boldsymbol{x} \sim N(\hat{\boldsymbol{x}}, \boldsymbol{P})$. Let $\hat{\boldsymbol{x}}_0$ be an initial estimate of groundshine–dose and $\boldsymbol{P}_0$ its corresponding error covariance matrix. This background–field is given by probabilistic version of Atmospheric Dispersion Model (ADM) and constitutes the prior characterization of the problem. It is based on segmented Gaussian plume model and it is part of the HARP system, more in [15]. We assume sparse measurements $\boldsymbol{y}_t$ of actual gamma dose–rate to be available in each time step. These measurements are assumed to be conditionally independent with known error. Assimilation procedure consists of two iteratively repeated steps: In time update step (4) current state estimate together with its error covariance matrix are propagated forward in time. The model error is estimated and accounted for. Following data update step (5) produces so called analysis – adjusts the model prediction to be in accordance with actual measurements. Along with this two Kalman filter steps is in each time step estimated model error covariance structure.

A. The model

Groundshine dose evolution is modeled according to semi–empirical formulas from Japan model OSCAAR. This abbreviation stands for Off–Site Consequence Analysis code for Atmospheric Releases in reactor accidents. It has been developed within the research activities on probabilistic safety assessment at the Japan Atomic Research Institute [9]. Let s be a spatial coordinate. Relation between $SD(s)$ – the initial ^{137}Cs deposition $[Bq \times m^{-2}]$ and $D_g(t, s)$ – the groundshine dose–rate $[Sv \times s^{-1}]$ at time t is given by

$$D_g(t, s) = SD(s) \times SF(s) \times L \times DF_g \times R(t) \times E(t), \quad (20)$$

where $SF(s)$ is the integrated shielding factor for groundshine at location s, L is geometric factor and DF_g = 5.86E–16 $Sv \times s^{-1} / Bq \times m^{-2}$ is the integrated dose–rate conversion factor for groundshine. $R(t)$ is a unit–less factor taking into account the radioactive decay of the deposited radionuclude with half–time of decay T_y (21).

$$R(t) = \exp\left(-\frac{t}{T_y} \ln 2\right) \qquad (21)$$

Term $E(t)$ is a unit–less factor taking into account the decrease of groundshine due to environmental processes, such as radionuclide migration deeper into the soil, weathering, leaching etc. The experiments had shown that the mitigation of groundshine due to environmental removal mechanisms follows relation given by superposition of two exponentials – fast and slow component (22).

$$E(t) = d_f \exp\left(-\frac{t}{T_f} \ln 2\right) + d_s \exp\left(-\frac{t}{T_s} \ln 2\right) \qquad (22)$$

Non–negative parameters d_f and d_s are weights of decay effect for fast and slow component, respectively. It must hold true that $d_f + d_s = 1$. T_f and T_s are half–times of fast and slow decay terms for groundshine.

Ground deposition model formulas are semi–empirical, it means that some of equation parameters are determined empirically on basis of measurements and the parameter values depend on the local conditions in the area of model application – soil type etc. The dose conversion factor was calculated by the method of Kocher in which the exposed individual was assumed to stand on a smooth, infinite plane surface with uniform concentration of source of radioactivity. Some parameters used in ground exposure model are treated as random variables with a given probability distribution. The parameter distribution types were determined for ^{137}Cs from the Chernobyl disaster and they are listed in [9]. The appropriate data for other radionuclides are not available but it is assumed that the long–term influence of the most of

them is not significant. For elements with high half–time of decay and similar fixation rates as ^{137}Cs are assumed the same equations of groundshine mitigation. The uncertainty of model parameters plays role in the whole model error and has to be accounted for.

B. Model error covariance parametrization

The idealized model of $\boldsymbol{Q}$ chosen for this example has three parameters $\boldsymbol{\theta} = (\alpha, \beta, L)|_{\ \alpha,\beta,L\geq 0}$

$$\boldsymbol{Q}_t = \alpha_t \left[\boldsymbol{Q}_t^{(1)} + \beta_t \boldsymbol{Q}_t^{(2)}(L_t)\right] \tag{23}$$

The model error is formally partitioned into two components representing different sources of uncertainty. The partitioning has physical background. Matrix $\boldsymbol{Q}^{(1)}$ concerns the uncertainty of forecast model parameters introduced in Section IV-A. This component is found as a covariance of sample obtained via Monte–Carlo simulation with many different settings of model parameters. Component $\boldsymbol{Q}^{(2)}$, scaled with β, is structured, homogeneous and isotropic error. All other sources of uncertainty are accommodated by introduction of $\boldsymbol{Q}^{(2)}$. This component is generated by means of second order autoregressive function $\rho_L(r)$ of length–scale parameter L and Euclidean distance between two spatial locations r [5].

$$\rho_L(r) = \left(1 + \frac{r}{L}\right) \exp\left(-\frac{r}{L}\right) \tag{24}$$

The value of length–scale parameter L controls how fast the correlation between two points decreases with their growing distance. The overall covariance is scaled with α. This parametrization allows for mutual scaling of unstructured noise component $\boldsymbol{Q}^{(1)}$ given upon numerical simulation and "additional" structured noise given by $\boldsymbol{Q}^{(2)}$. MPF algorithm according to [20] modified for this case is listed in the box ALGORITHM.

In Step 1), the particles are initialized with a prior distribution. In Step 2) are evaluated residuals upon measured and modeled values for purpose of normalized weights evaluation for different covariance parameter vectors $\boldsymbol{\theta}_t^{(i)}$. For each particle, the overall covariance given by (23) has to be evaluated. During Step 3) are particles resampled – those with small weights are replaced with particles "better" in terms of likelihood. Sometimes is also in this step introduced an artificial noise to prevent particle degeneracy problem – to maintain high diversity of particles. In Step 4) is performed data and time update of KF and time update of PF. If we omit Steps 4a) and 4c) we get the standard PF. In Step 4b) is set new importance function for the next time step.

ALGORITHM

1) Initialization:
   a) $For\ i = 1, \ldots, M$ initialize $\boldsymbol{\theta}_0^{(i)} \sim p(\boldsymbol{\theta}_0)$
   b) Set $\{\boldsymbol{x}_{0|-1}^{(i)}, \boldsymbol{P}_{0|-1}^{(i)}\} = \{\hat{\boldsymbol{x}}_0, \boldsymbol{P}_0\}$
2) Normalized weights evaluation:
   $For\ i = 1, \ldots, M$ evaluate:
   a) Residuals $\boldsymbol{v}_t^{(i)} = \boldsymbol{y}_t - \boldsymbol{H}_t \hat{\boldsymbol{x}}_t^{(i)}$
   b) Model error covariance matrix parametrization:
   $$\boldsymbol{Q}_t^{(i)} = Q\left(\boldsymbol{\theta}_t^{(i)} = \{\alpha_t^{(i)}, \beta_t^{(i)}, L_t^{(i)}\}\right)$$
      i) Evaluation of $\boldsymbol{Q}_t^{(i),(1)}$ via M-C simulation with multiple groudnshine model parameters setting
      ii) Evaluation of $\boldsymbol{Q}_t^{(i),(2)}(L_t^{(i)})$ via (24)
      iii) Evaluation of overall covariance via (23)
      $$\boldsymbol{Q}_t^{(i)} = \alpha_t^{(i)} \left[\boldsymbol{Q}_t^{(i),(1)} + \beta_t^{(i)} \boldsymbol{Q}_t^{(i),(2)}(L_t^{(i)})\right]$$
   c) Residual covariance matrix $\boldsymbol{S}(\boldsymbol{\theta}_t^{(i)})$ via (18)
   d) Importance weights $q_t^{(i)} = N(\boldsymbol{0}, \boldsymbol{S}(\boldsymbol{\theta}_t^{(i)}))$
   e) Normalize weights $\tilde{q}_t^{(i)} = \frac{q_t^{(i)}}{\sum_{j=1}^{M} q_t^{(j)}}$
3) PF measurement update – resampling:
   Resample M particles with replacement
   $$Pr(\boldsymbol{\theta}_{t|t}^{(i)} = \boldsymbol{\theta}_{t|t-1}^{(j)}) = \tilde{q}_t^{(j)}$$
4) KF data/time update and PF time update
   a) KF data update:
   $$\hat{\boldsymbol{x}}_{t|t}^{(i)} = \hat{\boldsymbol{x}}_{t|t-1}^{(i)} + \boldsymbol{K}_t^{(i)}[\boldsymbol{y}_t - \boldsymbol{H}_t \hat{\boldsymbol{x}}_{t|t-1}^{(i)}]$$
   $$\boldsymbol{K}_t^{(i)} = \boldsymbol{P}_{t|t-1}^{(i)} \boldsymbol{H}^T (\boldsymbol{H}_t \boldsymbol{P}_{t|t-1}^{(i)} \boldsymbol{H}_t^T + \boldsymbol{R}_t)^{-1}$$
   $$\boldsymbol{P}_{t|t}^{(i)} = (\boldsymbol{I} - \boldsymbol{K}_t^{(i)} \boldsymbol{H}_t) \boldsymbol{P}_{t|t-1}^{(i)}$$
   b) PF time update – prediction of new particles:
   $$\boldsymbol{\theta}_{t+1}^{(i)} \sim p(\boldsymbol{\theta}_{t+1}^{(i)} | \boldsymbol{\theta}_t^{(i)})$$
   c) KF time update:
   $$\hat{\boldsymbol{x}}_{t+1|t}^{(i)} = \boldsymbol{M} \hat{\boldsymbol{x}}_{t|t}^{(i)}$$
   $$\boldsymbol{P}_{t+1|t}^{(i)} = \boldsymbol{M} \boldsymbol{P}_{t|t}^{(i)} \boldsymbol{M}^T + \boldsymbol{Q}_{t+1}(\boldsymbol{\theta}_{t+1}^{(i)})$$
5) Iterate from step 2) with $t := t + 1$

V. Experimental Results and Conclusion

For experimental demonstration of the algorithm, an artificial scenario with local rain during the fifth hour of the plume phase was chosen. The rain increases depletion of the plume due the wet deposition. The area of interest is subset of polar network comprising of $N = 91$ analyzed points.

The measurements were simulated from the measurement equation (2) via linear forward observation operator $\boldsymbol{H}$ where the true initial deposition $\boldsymbol{x}_0$ was assumed to be two times higher than the prior estimate $\hat{\boldsymbol{x}}_0$ obtained from ADM. The background–field (initial distribution in time $t = 0$) was $N(\hat{\boldsymbol{x}}_0, \boldsymbol{P}_0)$ where forecast error covariance $\boldsymbol{P}_0$ was calculated according to

$$\boldsymbol{P}_0 = 2\bar{\boldsymbol{P}}_0 \circ \boldsymbol{\Omega}, \tag{25}$$

where $\boldsymbol{\Omega}$ is covariance matrix generated from (24) and the $\circ$ stands for element–wise matrix product (Schur product) [14]. This was done because the background–field error covariance matrix $\bar{\boldsymbol{P}}_0$ was modeled as sample covariance from multiple calls of ADM where the rain intensity was treated as a random variable. This accommodated the uncertainty in rain intensity into $\bar{\boldsymbol{P}}_0$ and provided us a valuable physical knowledge but this process also introduced strong covariances between states. In (25), these covariances were reduced, so the

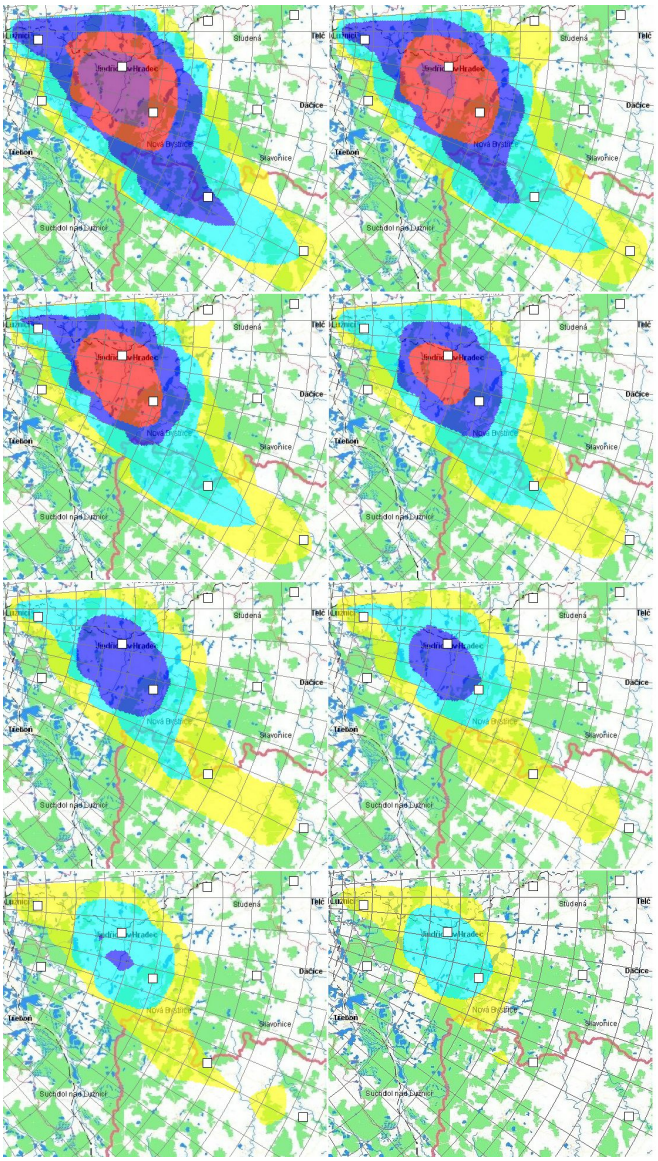

Fig. 1. Visualization of assimilated predictions for time interval 8 months with time step 1 month. Measurement locations are indicated with white squares.

background–field became more conservative. Initialization of particles in the very first step was following: $\alpha_1 \sim Gamma(1,1)$, $\alpha_2 \sim N(10^2, 10^4)$ and $L \sim N(10^3, 10^6)$. The prediction was evaluated for the first eighth months of the post–emergency phase. Measurements were assumed to be available each month. At each time step were simulated 10 irregularly spaced measurements. For clarity, all the measurements in this example are during computation located in the same positions, so the observation operator $\boldsymbol{H}_t = \boldsymbol{H}$ is constant. Multinomial resampling described in [4] was used as a resampling algorithm in MPF. Measurement error was set according to expert judgment based on the fact that the small measured values have higher relative error than high values due to the measurement methodology. In each step, first two moments of groundshine–dose distribution approximating the truth were predicted and updated.

In the Fig. 1 we can see assimilation results. The results are in compliance with our expectations for this special scenario. Model predictions were successfully adjusted in accordance with the measurements correcting the speed of dose mitigation. More general conclusion could not be done until additional tests with various settings are performed. Applicability of the method on higher dimension is still open. Only after cautious and wide verifications the method can be included into the assimilation sub–system.

References

[1] R. Daley. *Atmospheric data analysis*. Cambridge Univ. Press, Cambridge, 1991.

[2] D. P. Dee. A simple scheme for tuning forecast error covariance parameters. In *ECMWF Workshop on Variational Assimilation*, 1993.

[3] D. P. Dee. On-line estimation of error covariance parameters for atmospheric data assimilation. *Monthly Weather Review*, 123, 1995.

[4] R. Douc, O. Cappé, and E. Moulines. Comparison of resampling schemes for particle filtering. In *Proceedings of the 4th International Symposium*. Image and Signal Processing and Analysis, 2005.

[5] G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, 125:723–757, 1999.

[6] F. Gering, W. Weiss, E. Wirth, R. Stapel, P. Jacob, H. Mller, and G. Prhl. Assessment of evaluation of the radiological situation in the late phase of a nuclear accident. *Rad. Prot. Dosim*, 109, 2004.

[7] A. J. Haug. A tutorial on Bayesian estimation and tracking techniques applicable to nonlinear and non-Gaussian processes. Technical report, MITRE Corporation, 2005.

[8] R. Hofman. Assimilation scenario for long-term deposition of ^{137}Cs. In *8th International PhD Workshop on Systems and Control a Young Generation Viewpoint, Balatonfred, Hungary*, 2007.

[9] T. Homma and T. Matsunaga. *OSCAAR Model - Description and Evaluation of Model Performance*, 2006.

[10] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, 82, 1960.

[11] E. Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge Univ. Press, Cambridge, 2003.

[12] M. Karny et al. *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, London, 2005.

[13] A. C. Lorenc. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112:1177–1194, 1986.

[14] Herschel L. Mitchell and P. L. Houtekamer. An adaptive ensemble Kalman filter. *AMS Monthly Weathr Review*, 128:416–433, 1999.

[15] P. Pecha and R. Hofman. Integration of data assimilation subsystem into environmental model of harmful substances propagation. In *Harmo11 - 11th Internal Conf. Cambridge*, 2007.

[16] P. Pecha, R. Hofman, and P. Kuča. Assimilation techniques in consequence assessment of accidental radioactivity releases. ECORAD 2008, Bergen, Norway, 2008.

[17] P. Pecha, R. Hofman, and E. Pechová. Training simulator for analysis of environmental consequences of accidental radioactivity releases. In *6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia*, 2007.

[18] C. Rojas-Palma. Data assimilation for off site nuclear emergency management. Technical report, SCK-CEN, DAONEM final report, RODOS(RA5)-RE(04)-01, 2005.

[19] C. Rojas-Palma et. al. Data assimilation in the decision support system RODOS. *Rad. Prot. Dosim*, 104, 2003.

[20] T. B. Schön, F. Gustaffson, and Nordlund P-J. The marginalized particle filter - analysis, applications and generalizations. *IEEE Transaction on Signal Proceedings*, 53, 2005.

[21] T. B. Schön, R. Karlsson, and F. Gustaffson. Marginalized particle filter for mixed linear/nonlinear state-space models. *ESAIM: PROCEEDINGS*, 19:53–64, 2007.

[22] V. Šmídl and A. Quinn. Variational Bayesian filtering. *IEEE Transactions on Signal Processing*, 2008. to be published.