



Akademie věd České republiky
Ústav teorie informace a automatizace, v.v.i.

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

Václav Šmídl, Radek Hofman

Application of Sequential Monte Carlo Estimation for Early Phase of Radiation Accident

No. 2322

May, 2012

ÚTIA AV ČR, P.O.Box 18, 182 08 Prague, Czech Republic
Tel: +420 266052420, Fax: +420 266052068, Url: <http://www.utia.cas.cz>,
E-mail: smidl@utia.cas.cz, hofman@utia.cas.cz

This report constitutes an unrefereed manuscript, which is intended to be submitted for publication. Any opinions and conclusions expressed in this report are those of the authors and do not necessarily represent the views of the Institute.

Abstract

The early phase of radiation accident is characterized by minimum number of measured data and high uncertainty in both atmospheric conditions and radiation situation. Our goal is to provide an accurate method of radiation situation assessment that is capable to respect the uncertainty and provide informative predictions of its evolution for the involved decision makers. We propose a state space model based on atmospheric dispersion model, numerical weather model with local corrections and random walk on the model corrections and release evolution. This model is highly nonlinear and is estimated using sequential Monte Carlo. Since the model is significantly more complex than previously considered models and its estimation with naive proposal densities become too computationally demanding. We propose to construct a proposal density using problem specific simplification followed by application of the Laplace approximation. Properties of the resulting estimation procedure are illustrated on a twin experiment.

Keywords: radiation protection, atmospheric dispersion model, importance sampling

Contents

1	Introduction	4
2	Early phase of a Radiation Accident	5
3	State Space Model and its Sequential Monte Carlo Estimation	6
3.1	Probability density elicitation	8
3.2	Release scenario	9
3.3	Measurement model	9
3.4	State space model of the scenario	10
3.4.1	Wind field model	10
3.4.2	Activity release model	11
3.4.3	Atmospheric dispersion model	11
3.5	Summary of the state space model	12
3.6	Proposal density	13
4	Results	14
4.1	Simulation scenario	14
4.2	Model calibration	15
4.3	Puff model implementation	15
4.4	Estimation results	16
4.5	Prediction results	17
5	Conclusion	19
A	Conjugate update of inverse Gamma likelihood and Gamma prior	21
B	Laplace Approximation of shifted inverse Gamma likelihood and Gamma prior	22

List of Figures

1	Twin experiment of a radiation accident.	6
2	Comparison of different distributions with equal first two moments. . .	8
3	Effective number of particles for two choices of proposal densities. . . .	16
4	Posterior distributions of the released activity Q_t in time $t = 1, \dots, 6$.	17
5	Trajectories of the first puff in each particle at time $t = 24$ based on data available at times $t = 4, 8, 12$	18
6	Contour plots of the predicted accumulated dose in 4 hours after the release.	18
7	Distribution of the accumulated dose at different locations after 4 hours ($t=24$) from the start of the release.	19

1 Introduction

We are concerned with a scenario of a hypothetical accident in a nuclear power plant followed by an atmospheric release of radionuclides. The radioactive effluent forms a plume that is moving over the terrain according to the current meteorological situation. Urgent protective measures must be introduced as soon as possible to protect the public from harmful effects of ionizing radiation. The necessary countermeasures are typically prescribed by the law and classified according to the expected radiation level into several severity categories. Determination of these expected values is thus the most important input for decision making of the crisis management authority. The most reliable source of information is direct measurement of radiation level. However, detailed measurements are typically available only several hours or days after the release when the radioactive plume has already passed. This time frame is known as the *late phase*, and many post-accidents tasks can be defined to retrospectively analyze what happen.

In this paper, we are concerned with the *early phase* of an accident, i.e. the time frame when the plume is still in the air over the monitored area. We seek a way how to determine as much information as possible from the available measurements. In this phase, the measurements are rather sparse since we can not use mobile measuring devices but only the on-line measurements form the radiation monitoring network. These sensors typically measure total radiation dose. The closest measurement points are on the ring around the power plant, covering all directions of potential release. In principle, it is possible to start to evaluate predictions of the accident consequences within minutes from detection of an anomaly. Each new measurement should be used to improve accuracy of the predictions as soon as possible to support the decision making of the crisis management team.

This task is known in the environmental literature as data assimilation, and the dominant methods in this field are interpolation (Eleveld et al., 2007; Winiarek et al., 2010), variational approach (Jeong et al., 2005; Kovalets et al., 2009), genetic algorithms (Haupt et al., 2009; Cervone et al., 2010). However equivalence of this task to statistical estimation is known (Anderson et al., 1999), and statistical techniques such as Monte Carlo Markov Chain (Senocak et al., 2008; Delle Monache et al., 2008) and sequential Monte Carlo (Johannesson et al., 2004; Hiemstra et al., 2011) are gaining popularity.

Monte Carlo methods are usually considered to be computationally expensive, especially in cases with many unknown parameters. This is typically the case with the classical particle filter (Gordon et al., 1993), where the transition probability is used as the proposal density. Many general techniques for better proposal design were published, for example (Oh and Berger, 1992; Pitt and Shephard, 1999; Cornebise et al., 2008). Any of these methods has good potential to improve performance of the sequential Monte Carlo method since the uncertainty in the transitional density is much higher than that in the observation density. Moreover, evaluation of the likelihood function for a single particle is computationally expensive which makes this application an excellent area where sophisticated adaptive proposals will have significant impact. As a first step in this direction, we propose a method based on problem-specific simplification.

The simplified posterior is then approximated using the Laplace method (Kass et al., 1990). We show that this proposal significantly increases computational efficiency of the estimation procedure.

2 Early phase of a Radiation Accident

Release of radioactive material into the atmosphere is the last possible resort of any accident in a nuclear power plant. It is an extremely rare event, however with severe consequences for potentially many people living in proximity of the power plant. Awareness of radiation security has been increased after the Chernobyl accident, and almost every country is now equipped with monitoring network of on-line connected receptors continually measuring radiation levels. At present, these measurements are monitored by state authorities and stored for later use. The purpose of the network is to serve the responsible decision makers to assess the situation in the case of radiation accident. However, as recent experience from the Fukushima-Daiichi accident suggests, the measurements alone are insufficient since the expert group of decision-makers in the early phase of the accident will not have enough time to analyze the data.

Therefore, we aim at creation of an automatic system that processes the measured data online and continually produces prediction of radiation situation e.g. 2 hours ahead. Continuous operation of such system would allow to estimate local corrections of the weather forecast in the spirit of (Sloughter et al., 2010), and thus providing on-line calibration of the weather conditions. Automated evaluation of predictions will allow the decision makers to focus on practical issues without any delay caused by evaluation of the measurements.

The challenge in the task is twofold. First, the procedure must be computationally affordable to compute on a commodity PC. Second, the predictions should be considered as trustworthy by the decision-makers. The latter challenge is particularly demanding since there are almost no data sets on which it would be possible to validate the models. Even in the case of Fukushima-Daiichi accident, no data from the close monitoring stations in the early phase of the accident are available. For this reason, the research is done on so called twin experiments, where the measured data are replaced by simulated data obtained from atmospheric dispersion models with known parameters and weather conditions. The task for estimation is to recover the parameters and weather conditions from the simulated data.

An example of one such experiment is displayed in Figure 1. The typical measurement network is rather sparse with two circles of receptors, the first one inside the power plant area and the second in the closest inhabited villages. Note that the radiation plume is rather narrow so only a few receptors may register radiation level significantly higher than the radiation background. The task is to use these few data points and predict the total absorbed dose two hours ahead.

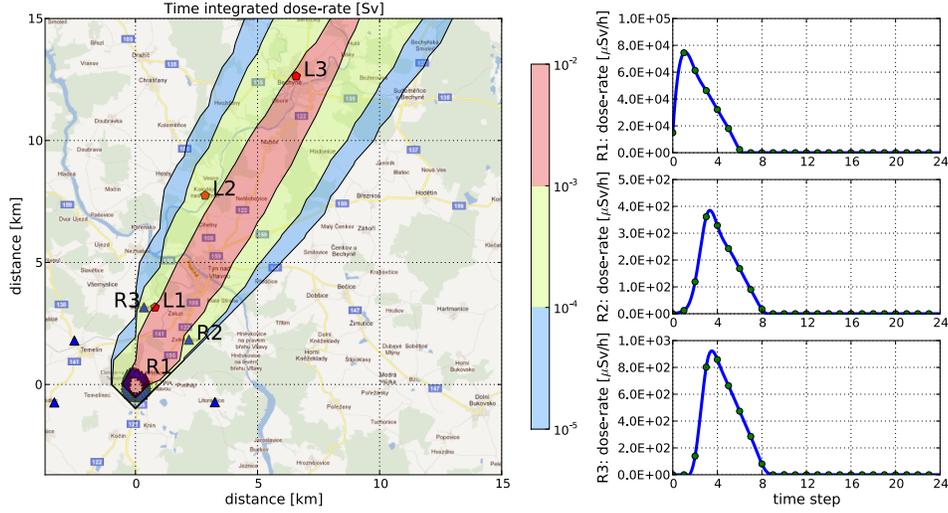


Figure 1: Twin experiment of a radiation accident. **Left:** position of nominal total radiation dose accumulated 4 hours after the release; measurements sites are denoted by triangles and points of interest are denoted by pentagons. **Right:** Simulated measurements in the three radiation dose receptors with highest observed values.

3 State Space Model and its Sequential Monte Carlo Estimation

The spatio-temporal distribution of the pollutant can be modeled by a discrete-time stochastic process:

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{v}_t), \quad (1)$$

$$\mathbf{y}_t = h(\mathbf{x}_t, \mathbf{w}_t), \quad (2)$$

where \mathbf{x}_t is the state variable, \mathbf{y}_t is the vector of observations and $f(\cdot)$ and $g(\cdot)$ are known functions transforming the state into the next time step, or observations, respectively. Both transformations are subject to disturbances \mathbf{v}_t and \mathbf{w}_t which are considered to be random samples from a known distribution.

The model (1)–(2) can thus be formalized in terms of probability density functions

$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_t), \quad \mathbf{x}_t \sim p(\mathbf{x}_t|\mathbf{x}_{t-1}). \quad (3)$$

Here, $p(\cdot|\cdot)$ denotes the conditional probability density of its argument. By *Bayesian Filtering* we mean the recursive evaluation of the filtering distribution, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, using

Bayes rule (Peterka, 1981):

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}, \quad (4)$$

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1}, \quad (5)$$

where $p(\mathbf{x}_1|\mathbf{y}_0)$ is the prior distribution, and $\mathbf{y}_{1:t} = [\mathbf{y}_1, \dots, \mathbf{y}_t]$ denotes the set of all observations. The integration in (5), and elsewhere in this paper, is over the whole support of the involved probability density functions.

Equations (4)–(5) are analytically tractable only for a limited set of models. The most notable example of an analytically tractable model is linear Gaussian (3) for which (4)–(5) are equivalent to the Kalman filter. For other models, (4)–(5) need to be evaluated approximately. Sequential Monte Carlo technique (Gordon et al., 1993; Doucet et al., 2001) is based on approximation of the posterior density by a weighted empirical density

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \approx \sum_{i=1}^n w_t^{(i)} \delta(\mathbf{x}_{1:t} - \mathbf{x}_{1:t}^{(i)}), \quad (6)$$

where $\mathbf{x}_{1:t} = [\mathbf{x}_1, \dots, \mathbf{x}_t]$ is the state trajectory, $\{\mathbf{x}_{1:t}^{(i)}\}_{i=1}^n$ are samples of the trajectory, $w_t^{(i)}$ is the weight of the i th sample, $\sum_{i=1}^n w_t^{(i)} = 1$, and $\delta(\cdot)$ denotes the Dirac δ -function.

The main appeal of Sequential Monte Carlo methods is in the fact that this approximation can be evaluated for an arbitrary model (3) given a suitable proposal function, $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$, yielding

$$w_t^{(i)} \propto \frac{p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}. \quad (7)$$

Convergence of the approximation to the exact posterior has been studied theoretically, e.g. in (Crisan and Doucet, 2002), and it was concluded that the error of the approximation grows with time. One way to remedy the situation is the use of *resampling* procedure, where the existing particles are copied or removed based on their $w_t^{(i)}$ such that the new particles have equal weights (Doucet et al., 2001).

An important task in the early phase of a radiation accident is prediction of its future development. Prediction of the state trajectory for h time steps ahead, $\mathbf{x}_{t:t+h}$, is given by:

$$p(\mathbf{x}_{t:t+h}|\mathbf{y}_{1:t}) = \sum_{i=1}^n w_t^{(i)} \prod_{\tau=1}^h p(\mathbf{x}_{t+\tau}^{(i)}|\mathbf{x}_{t+\tau-1}^{(i)}). \quad (8)$$

Thus the predictor is a combination of predictors for each Monte Carlo trajectory weighted by the latest available weights. Note that the observation data enter the prediction formula only via the weight $w_t^{(i)}$, (7). Hence, reevaluation of the predictions for a new observed data is computationally cheap.

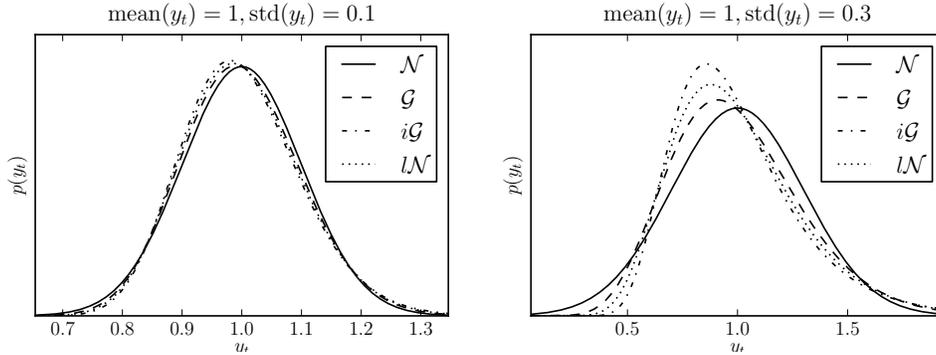


Figure 2: Comparison of different distributions with equal first two moments. All distributions have the same mean value $\mu_t = 1$, but different standard deviations $\gamma = 0.1$ (left) and $\gamma = 0.3$ (right).

3.1 Probability density elicitation

As any Bayesian technique, the sequential Monte Carlo is defined for models with fully specified probability density functions. However, full distribution of measurement error of a measuring device is typically not available. More often, the error of the measuring device is characterized by two factors: (i) measurement range, and (ii) maximum error, either absolute or relative. The full distribution function has to be either obtained experimentally or assumed to belong to a chosen parametric family. In this text, we assume that the characteristics of the measurements of quantity y_t are available in the form of its statistical moments:

$$\text{mean}(y_t) = y_t^{\text{true}} = \mu_t, \quad \text{std}(y_t) = \gamma\mu_t, \quad (9)$$

for relative observation error, or $\text{std}(y_t) = \sigma_t$, for absolute error. The first equality represents the assumption of unbiased mean value of the measurements. The expected standard deviation is set from the provided measurement error which is assumed to be equal to 2 standard deviations.

The choice of the full distribution, $p(y_t)$, can be done via moment matching. The following parametric forms have the same moments as (9):

$$y_t \sim \mathcal{N}(\mu_t, (\gamma\mu_t)^2), \quad \text{or } \mathcal{N}(\mu_t, \sigma_t^2), \quad (10)$$

$$y_t \sim \mathcal{G}(\gamma^{-2}, \gamma^2\mu_t), \quad (11)$$

$$y_t \sim i\mathcal{G}(\gamma^{-2} + 2, (\gamma^{-2} + 1)\mu_t). \quad (12)$$

Here, $\mathcal{N}(\mu_t, \sigma_t^2)$ denotes Normal distribution with mean μ_t and variance σ_t^2 ; $\mathcal{G}(k, \theta)$ denotes Gamma distribution with shape parameter k and scale parameter θ ; and $i\mathcal{G}(\alpha, \beta)$ denotes the inverse Gamma distribution with shape parameter α and scale parameter β . Comparison of these distributions for $\mu_t = 1$, and various γ is displayed in Figure 2. Note that for relative error as low as 10%, all densities are very similar, their differences

become more obvious for high relative error, when the gamma distributions assign very low probabilities to the region around zero.

From these three choices, the Normal distribution (10) is typically favored since it has the greatest entropy of all possible distributions (Dowson and Wragg, 1973). However, this distribution has to be truncated if the variable is positive by the definition. On the other hand, the Gamma distribution seems a natural choice for positive variables.

3.2 Release scenario

We assume a continuous release of radioactive pollutant from a nuclear power plant at known location, $[s_{1,pp}, s_{2,pp}]$ and known altitude $s_{3,pp}$. In vector notation, $\mathbf{s} = [s_1, s_2, s_3]$ will denote three dimensional vector of space coordinates. At time t , an amount of radioactive material of activity Q_t is instantaneously released into the atmosphere forming a radioactive plume which is subject to wind and dispersion. The radiation activity of the plume is decreasing via radioactive decay.

3.3 Measurement model

The radiation monitoring network considered in this text is equipped by several dose-rate receptors and a meteorostation. The dose-rate receptors are of the Geiger-Müller type with fixed operation range, typically from several nSv/h to Sv/h. From the meteorostation we consider only the anemometer. Statistical properties of these sensors is now evaluated.

The measured value of the radiation dose is a sum of natural background radiation, y_{nb} , and the dose from the contaminant, y_Q . According to the studies of the gamma dose receptors (Thompson et al., 2000), the error of measurement is typically proportional to the measured dose with a constant of proportionality γ_y , typically in the range of 7–20%. Therefore, we assume that the measurements at i th receptor have moments

$$\text{mean}(y_{i,t}) = y_{nb,i} + y_{Q,i,t}, \quad (13)$$

$$\text{std}(y_{i,t}) = \gamma_y(y_{nb,i} + y_{Q,i,t}), \quad (14)$$

where $y_{nb,i}$ are long term averages of radiation background at the location. We choose the inverse Gamma density,

$$p_G(y_{i,t}|\mathbf{x}_{1:t}) = i\mathcal{G}(\gamma_y^{-2} + 2, (\gamma_y^{-2} + 1)(y_{nb} + y_Q)). \quad (15)$$

Technically, (15) should be truncated at support defined by the ranges of the device. However, we avoid such complication since all measurements considered in this text are well inside of the range of the device such that the truncated versions the distribution is indistinguishable from (15).

Accuracy of the anemometers is typically also available in terms of relative error on the wind speed, v_t , and constant error of wind direction, ϕ_t , i.e.

$$\begin{aligned}\text{std}(v_t) &= \gamma_v v_t(\mathbf{s}_{meteo}), \\ \text{std}(\phi_t) &= \sigma_\phi,\end{aligned}$$

where \mathbf{s}_{meteo} is the location of the meteorostation. From all choices of possible distribution (10)–(12) we select the inverse Gamma form (12). Distribution of the wind direction is considered as normal with fixed variance, (10).

3.4 State space model of the scenario

In definition of the state space model, we follow Johannesson et al. (2004) and parameterize the release by a puff model. However, we significantly extend the parametrization to consider additional phenomena.

3.4.1 Wind field model

Accurate predictions of the wind field is the most critical variable in the task of prediction of the radiation situation after the release. There are many numerical tools that compute numerical prediction of the wind field on a grid. However, these predictions may not be well calibrated for the location and may not be in agreement with the values measured by the available anemometer. Therefore, we need to perform our own local correction of the wind field. Due to sparsity of measurements, we need a really simple parametrization of the model. We have chosen the following model

$$v_t(\mathbf{s}) = \tilde{v}_t(\mathbf{s})a_t, \tag{16}$$

$$\phi_t(\mathbf{s}) = \tilde{\phi}_t(\mathbf{s}) + b_t, \tag{17}$$

where $\tilde{v}_t(\mathbf{s})$, $\tilde{\phi}_t(\mathbf{s})$ are the wind speed and wind direction predicted by the numerical model at location \mathbf{s} , respectively. In this text, we assume that the forecasted values \tilde{v} and $\tilde{\phi}$ are precomputed and do not change during estimation. Constants a_t and b_t are unknown biases of the prediction model at time t . Correction of the wind field forecast is then achieved by estimation of a_t and b_t .

The constants a_t and b_t are expected to vary in time, with moments

$$\begin{aligned}\text{mean}(a_t) &= a_{t-1}, & \text{std}(a_t) &= \gamma_a a_{t-1}, \\ \text{mean}(b_t) &= b_{t-1}, & \text{std}(b_t) &= \sigma_b.\end{aligned}$$

Matching of these moments to the distributions yields

$$\begin{aligned}p(a_t|a_{t-1}) &= \mathcal{G}(\gamma_a^{-2}, \gamma_a^2 a_{t-1}), \\ p(b_t|b_{t-1}) &= t\mathcal{N}(b_{t-1}, \sigma_b, \langle b_{t-1} - \pi, b_{t-1} + \pi \rangle).\end{aligned} \tag{18}$$

The Gamma density was chosen for a_t since its variance is proportional to its mean value. On the other hand, the variance in the direction is constant, hence we have chosen the truncated Normal on the unit circle.

3.4.2 Activity release model

We consider a continuous release of the activity into the atmosphere. Quantity of the release is assumed to be time-varying but with slow dynamics of changes. Similarly to other quantities, we assume that the change in the released activity is proportional to its previous value, i.e. $\text{mean}(Q_t) = Q_{t-1}$, $\text{std}(Q_t) = \gamma_Q Q_{t-1}$, yielding Gamma transition density

$$p(Q_t|Q_{t-1}) = \mathcal{G}(\gamma_Q^{-2}, \gamma_Q^2 Q_{t-1}). \quad (19)$$

The constant of proportionality γ_Q governs our prior belief in variability of the release, $\gamma_Q \rightarrow 0$ models stationary release. The model is inappropriate at the beginning of the release when $Q_{t-1} = 0$. Therefore, we assume that detection of the release is performed with the following prior model:

$$p(Q_1) = \mathcal{G}(\alpha_{Q_1}, \beta_{Q_1}). \quad (20)$$

3.4.3 Atmospheric dispersion model

The continuous release is approximated by a sequence of puffs, (Zannetti, 1990), each puff representing the released material during one sampling period $\langle t, t + 1 \rangle$. The puff model is thus formed by a sequence of puffs, each with its original activity of an instantaneous release $\{Q_1, \dots, Q_t\}$. Concentration of the pollutant in the atmosphere from a single puff is:

$$C_i(\mathbf{s}, \tau) = \frac{Q_i}{\sqrt{2\pi^3/2}\sigma_1\sigma_2\sigma_3} \exp \left[-\frac{(s_1 - l_{1,i,\tau})^2}{2\sigma_1^2} - \frac{(s_2 - l_{2,i,\tau})^2}{2\sigma_2^2} - \frac{(s_3 - l_{3,i,\tau})^2}{2\sigma_3^2} \right]. \quad (21)$$

where $\mathbf{s} = [s_1, s_2, s_3]$ are spatial coordinates of a receptor and $\mathbf{l}_{i,\tau} = [l_{1,i,\tau}, l_{2,i,\tau}, l_{3,i,\tau}]$ are coordinates of the center of the i th puff. Time evolution of the puff center is fully determined by the wind field at the previous location

$$\begin{aligned} l_{1,\tau} &= l_{1,i,t} - \Delta\tau v_t(\mathbf{l}_t) \sin(\phi_t(\mathbf{l}_t)), \\ l_{2,\tau} &= l_{2,i,t} - \Delta\tau v_t(\mathbf{l}_t) \cos(\phi_t(\mathbf{l}_t)), \\ l_{3,\tau} &= l_{3,i,t}, \end{aligned} \quad (22)$$

where $v_t(\mathbf{l}_t)$ and $\phi_t(\mathbf{l}_t)$ are given by (16)–(17).

Contrary to previous works, e.g. (Johannesson et al., 2004; Haupt et al., 2009), concentration of the pollutant $C(\mathbf{s}, \tau)$ can not be measured directly but only via measurements of time integrated radiation dose rates y_t [Sv]. Contribution of the puff model to the total dose measured on the j th receptor (13) is

$$y_{Q,j,t} = \sum_{i=1}^t c_{i,j,t} Q_i, \quad (23)$$

where the coefficient $c_{i,j,t}$ is computed as (Raza et al., 2001):

$$c_{i,j,t} = \frac{\omega K E \mu_a}{\rho} \frac{1}{Q_i} \int_{t-1}^t \Phi_i(\mathbf{s}_{R_j}, \tau, E) d\tau. \quad (24)$$

Here, K is the dose conversion factor; E is the gamma energy produced by decay of the assumed radionuclide; ω is the ratio of absorbed dose in tissue to the absorbed dose in air; ρ is the air density. Fluency rate $\Phi_i(\mathbf{s}_{R_j}, E)$ from the i th puff is calculated as a three dimensional integral over the volume of the puff:

$$\Phi_i(\mathbf{s}_{R_j}, \tau, E) = \int_{\Omega} \frac{C_i(\mathbf{s}, \tau) B(E, \mu r) \exp(-\mu r)}{4\pi r^2} d\mathbf{s}, \quad (25)$$

$$B(E, \mu r) = 1 + k \mu r, \quad k = \frac{\mu - \mu_a}{\mu_a}. \quad (26)$$

Ambient activity concentration $C_i(\mathbf{s}, \tau)$ is defined by (21). B is the linear build-up factor, μ and μ_a are linear and mass attenuation coefficient, respectively; $\Omega \subset \mathbb{R}^3$ is a spatial domain of integration ($\mathbf{s} \in \Omega$); and $r = \|\mathbf{s}_{R_j} - \mathbf{s}\|$ is the distance of spatial locations \mathbf{s} and receptor \mathbf{s}_{R_j} .

Typically, the decision makers are concerned with the total committed dose up to some reference time since the release start

$$D_{j,t} = \sum_{\tau=1}^t \sum_{i=1}^t c_{i,j,t} Q_i. \quad (27)$$

The observation operator transforming activity concentration in the air to the time integrated dose rate (23)–(26) is highly nonlinear. The relation (24) was derived only for release of a single nuclide with one energy level E . In the case that the puff's inventory consists a nuclide with more energy levels of gamma radiation or of a mixture of nuclides, computation of the coefficients (24) is more complex, however, the estimation methodology works without any change.

3.5 Summary of the state space model

Summarizing results from previous Sections, the state of the considered dynamical system is composed of the wind field model, and parameterization of all puffs in the puff model, i.e.

$$\mathbf{x}_t = [a_t, b_t, Q_1, \dots, Q_t, \mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_t],$$

with measurements composed of one anemometer and m radiation dose measurements

$$\mathbf{y}_t = [v_t, \phi_t, y_{1,t}, y_{2,t}, \dots, y_{m,t}].$$

Parameter evolution model is then composed of all considered models

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = p(a_t | a_{t-1}) p(b_t | b_{t-1}) p(Q_t | Q_{t-1}) p(\mathbf{l}_t | \mathbf{l}_{t-1}, a_t, b_t), \quad (28)$$

given by (18), (19) and the observation model is

$$p(\mathbf{y}_t|\mathbf{x}_t) = p(v_t, \phi_t|a_t, b_t) \prod_{k=1}^m p(y_{i,t}|Q_{1:t}, \mathbf{l}_{1:t}).$$

3.6 Proposal density

A common choice of proposal functions $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ in (7) is the transition probability

$$q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \equiv \prod_{\tau=1}^t p(\mathbf{x}_\tau|\mathbf{x}_{\tau-1}). \quad (29)$$

The main advantage of this choice is analytical and computational simplicity due to the fact that the transitional probabilities in (7) cancel out to yield:

$$w_t^{(i)} \propto p(\mathbf{y}_t|\mathbf{x}_t). \quad (30)$$

This simplification allows to compute more particles in situations where it is cheap to draw the particles and evaluate (30). However, it is not the case in our scenario where evaluation of (7) involves computation of integrals (24).

The problem as formulated above has two advantageous structural properties that can be exploited to create efficient proposals.

1. The availability of anemometer observations v_t, ϕ_t provides almost direct observability of forecast biases a_t, b_t .
2. For the choice of inverse Gamma distribution for the likelihood (15), the Gamma transition models (18),(19) are conjugate with posterior density in the form of Gamma distribution, see Appendix A.

These properties motivate the following choice of the proposal density,

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) = q(a_t|a_{t-1}, v_t)q(b_t|b_{t-1}, \phi_t) \\ q(Q_t|Q_{t-1}, y_{1:m,t}, \mathbf{s}_{1:t})q(\mathbf{s}_t|\mathbf{s}_{t-1}, a_t, b_t),$$

where the first two factorized densities are:

$$q(a_t|a_{t-1}, v_t) \propto p(v_t|a_t)p(a_t|a_{t-1}) = \mathcal{G}(k_a, \theta_a), \quad (31)$$

$$q(b_t|b_{t-1}, \phi_t) \propto p(\phi_t|b_t)p(b_t|b_{t-1}) = \mathcal{N}(\mu_b, r_b), \quad (32)$$

with shaping parameters

$$k_{a,t} = \gamma_a^{-2} + \gamma_v^{-2} + 2, \\ \theta_{a,t} = [\gamma_a^{-2}a_{t-1}^{-1} + \tilde{v}_tv_t^{-1}(\gamma_v^{-2} + 1)]^{-1}, \\ \mu_{b,t} = r_{b,t} [\sigma_b^{-2}b_{t-1} + \sigma_\phi^{-2}(\phi_t - \tilde{\phi}_t)], \\ r_{b,t} = (\sigma_b^{-2} + \sigma_\phi^{-2})^{-1}. \quad (33)$$

Algorithm 1 Sequential Monte Carlo estimation for early phase radiation assessment

Initialization: sample state variable \mathbf{x}_t from prior densities, $p(\mathbf{x}_t)$.

At each time t do:

1. Collect measurements \mathbf{y}_t ,
 2. For each particle, do
 - (a) Update shaping parameters $k_{a,t}, \theta_{a,t}, \mu_{b,t}, r_{b,t}$ using (33) and sample new values of $a_t^{(i)}, b_t^{(i)}$ from (31) and (32).
 - (b) Compute new locations of all puff centers $\mathbf{l}_t^{(i)}$ using (16),(17) and (22).
 - (c) Evaluate radiation dose coefficients $c_{i,t}$ for each receptor.
 - (d) Update shaping parameters $\mu_{Q,t}, \sigma_{Q,t}$ of proposal density (34) and sample new values $Q_t^{(i)}$.
 - (e) Evaluate weights $w_t^{(i)}$ using (7) and normalize them.
 3. Resample the particles,
 4. Evaluate prediction of the radiation situation for specified time horizon using (8).
-

Density (31) is obtained by conjugate update of Gamma densities, Appendix A. Density (32) is derived using standard conjugate update of Gaussian densities. Derivation of $q(Q_t|Q_{t-1}, y_{1:m}, \mathbf{s}_{1:t})$ is more demanding since the likelihood is not conjugate. Therefore, we propose to use the Laplace approximation (Kass et al., 1990)

$$q(Q_t|Q_{t-1}, y_{1:m,t}, \mathbf{s}_{1:t}) = t\mathcal{N}(\mu_Q, \sigma_Q^2, \langle 0, \infty \rangle), \quad (34)$$

where moments μ_Q and σ_Q are obtained numerically, see Appendix B for derivation.

The final algorithm is described in Algorithm 1.

4 Results

4.1 Simulation scenario

The simulated accident was a 1 hour long release of radionuclide ^{41}Ar with half-life of decay 109.34 minutes. Radionuclide ^{41}Ar was chosen for two reasons: (i) ^{41}Ar is a noble gas which has no deposition and consequently no groundshine, hence we need to calculate only the gamma dose rate from cloud shine; (ii) according to the tables of radioactive isotopes (Browne et al., 1986), radionuclide ^{41}Ar emits gamma radiation at energy level 1293.57keV with branching ratio 99.1% and thus can be treated as a mono-energetic nuclide.

Bayesian filtering is performed in time steps $t = 1, \dots, 18$, with sampling period of 10 minutes. This sampling period was chosen to match the sampling period of the radiation monitoring network which provides measurements of time integrated dose rate in 10-minute intervals. The same period was assumed for the anemometer. The simulated release started at time $t = 1$ with release activity $Q_{1:6} = [1, 5, 4, 3, 2, 1] \times 1e16 Bq$.

4.2 Model calibration

Documentation provided by manufactures of the gamma dose receptors was used to establish their accuracy in terms of parameters introduced in Section 3.3. In our case, the parameters were $\gamma_y = 0.2$ for the dose monitoring stations, $\gamma_a = 0.1$ and $\sigma_\phi = 3$ for the anemometer.

Parameters of the transition model from Section 3.4.1 can be estimated from historical data. Since the observations of the wind field as well as the meteo forecast data are recorded, we can choose a fixed length window of historical data and estimate parameters γ_a, σ_b . For the presented example, we estimated $\gamma_a = 0.6$ and $\sigma_b = 10$ from a continuous windows of historical data of 1000 samples. We assume that this estimation procedure can be run recursively.

The numerical weather predictions were provided by the MEDARD system (Eben et al., 2005) with grid resolution 9km. The assumed category of Pasquill’s atmospheric stability was D.

On the other hand, we have no historical data from which we can estimate parameters of the transition model for the dose, γ_Q and α_{Q1}, β_{Q1} . These were chosen as $\gamma_Q = 1$ for the evolution model and $\alpha_{Q1} = 1, \beta_{Q1} = 0$ for the prior. Under this choice, the prior is equivalent to the non-informative Jeffrey’s prior for a scale parameter.

4.3 Puff model implementation

Implementation details of the Gaussian puff model and the observation operator transforming activity have were carefully examined to achieve most accurate results with minimum computational cost.

Activity concentration in air given by the dispersion model is evaluated using a step-wise integration scheme and it is based on a difference scheme described in detail in (Pecha et al., 2007). Integration (24) must be performed in both time and space domains. Integration in time is done in sub-steps τ that are equidistantly distributed every 120 seconds. Spatial integration in (25) is approximated using Gauss quadrature rules (Golub and Welsch, 1969). Due to symmetry in the physics behind the problem it is beneficial to perform the integration in spherical coordinates. Since we assume that all the receptors are placed 1m above the ground, the integration domain Ω around the receptor in (25) can be approximated by a hemisphere. Its radius is determined by “ n/μ ” method (Pecha and Hofman, 2011), where the integration limit considers only those sources of irradiation (volumes with non-negligible concentration of activity) located up to distance n/μ , $n \in \mathbb{R}$ from the receptor, where μ is a linear attenuation

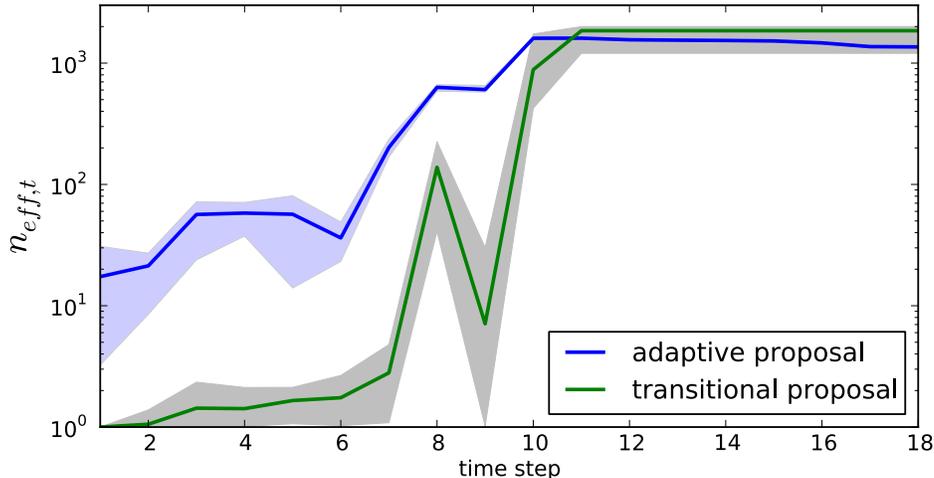


Figure 3: Effective number of particles for two choices of proposal densities. Full lines denote mean value over 10 Monte Carlo runs for each strategy. Minimum and maximum value of the Monte Carlo study are used as borders of the shaded areas.

coefficient [m^{-1}]. The calculations were performed with $n = 5$. To avoid integration over void parts of Ω , i.e. those parts where the activity concentration is negligible or zero, the integration is performed only over a subset of Ω having a non-empty intersection with the box containing significant activity assigned to i th puff. The size of the box is determined by dispersion coefficients σ_i in (21). This rather complicated choice of integration domain increases the effectively of Gauss quadrature method and makes results more accurate with given number of abscissas used in the method. The code was implemented on a multi-core architecture.

4.4 Estimation results

The filtering posterior was evaluated using Algorithm 1 with 2000 particles and two choices of proposal selection: (i) transition probability (29), and (ii) proposal obtained by the approximate conjugate statistics on a, b and Laplace proposal on Q_t , Section 3.6. In large sample limit, these two proposals should produce the same results, however, for limited number of particles, the efficiency of the proposal density varies significantly. This is demonstrated by average number of effective particles

$$n_{eff,t} = \left(\sum_{i=1}^n (w_t^{(i)})^2 \right)^{-1}.$$

In Fig. 3 is comparison for both strategies in terms of $n_{eff,t}$. Note that especially at the beginning of the study, the tailored proposal of Section 3.6 is more than ten times more efficient, therefore, it will be used for all results in this text.

Efficiency of the Laplace approximation for the released quantity Q_t is very good, see Fig. 4, where its posterior estimates obtained just after the first vector of measurements that is influenced by Q_t are displayed. The histograms cover at most one order of the Bq scale with significant mass concentrated around the true value.

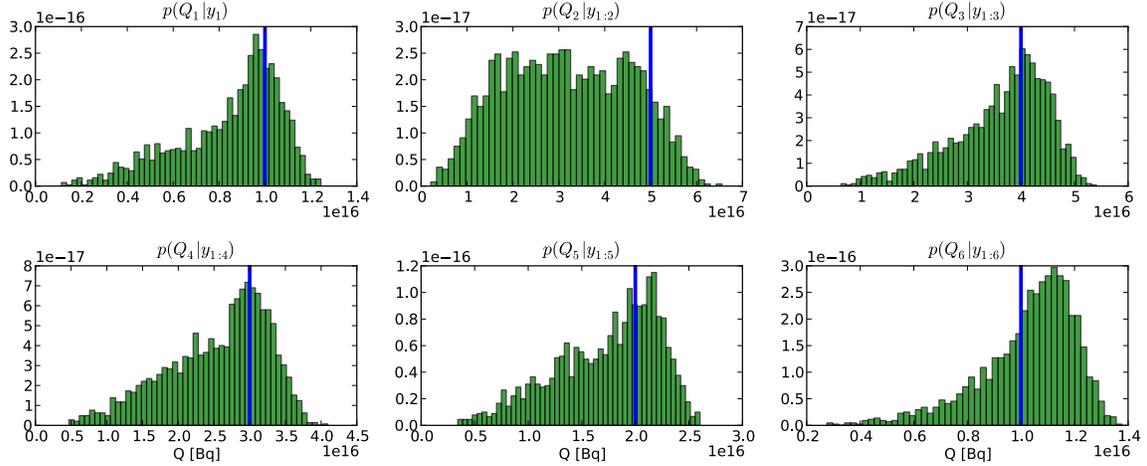


Figure 4: Posterior distributions of the released activity Q_t in time $t = 1, \dots, 6$. True values are denoted by vertical blue lines at $[1, 5, 4, 3, 2, 1] \times 1e16$ Bq, respectively.

4.5 Prediction results

One important task in the early phase of the radiation accident is the ability to extend the results of filtering to short term predictions. This can be easily achieved in sequential Monte Carlo using equation (8). Note that the future trajectory depends on the wind speed and direction that can change rapidly and is therefore highly uncertain. Any observation is then very valuable since it considerably reduces the uncertainty about reality. This is illustrated in Fig. 5 via predicted trajectories of the puff centers after 4 hours from the start of the release using different data. The most uncertain situation is at the time of the release, $t = 0$, when no data are available. With incoming data in time, the uncertainty is greatly reduced, the narrow part of the trajectories ends at point for which the observed data are available, while the ever widening part is the prediction.

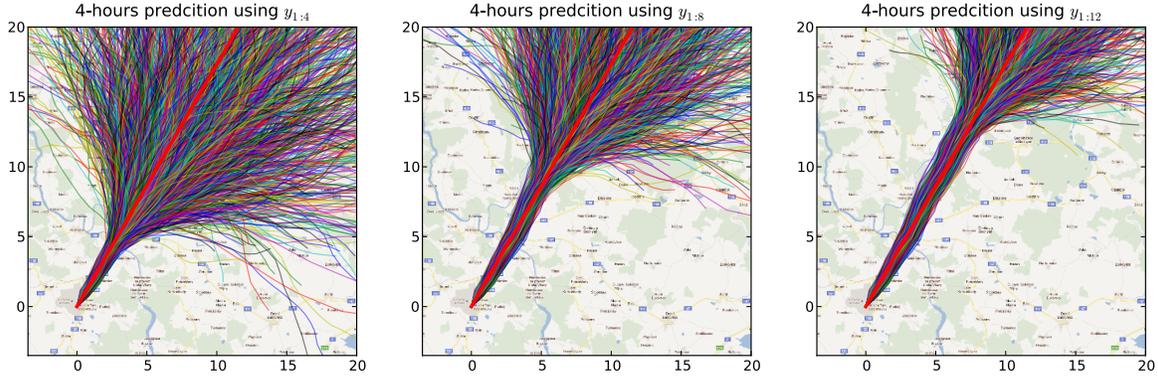


Figure 5: Trajectories of the first puff in each particle at time $t = 24$ based on data available at times $t = 4, 8, 12$, respectively from left to right.

The posterior densities (6) encode a lot of information that needs to be presented to the decision makers. One of the most valued supporting material from the human point of view is the map of isodose of the total committed dose (Bartzis et al., 2000). These are provided by many software tools. However, in the probabilistic formulation, the total absorbed dose has its own distribution. If we want to display the common contour map, we need to specify probability at which we want to draw the contour. In Fig. 6, contour plot based on the expected value of the cumulative dose is displayed.

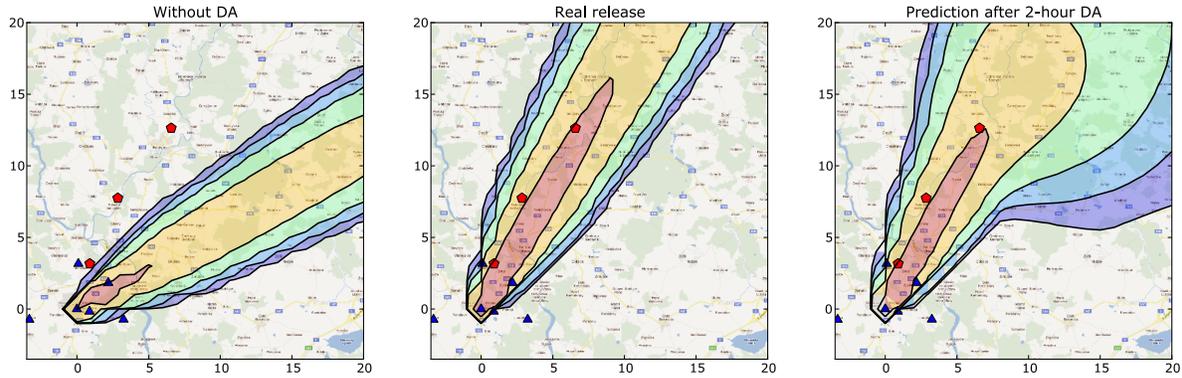


Figure 6: Contour plots of the predicted accumulated dose in 4 hours after the release. **Left:** prediction based on the uncorrected numerical weather forecast; **Middle:** “true” value simulated by the twin experiment; **Right:** expected value of the predicted density using data available up to 2-hours after the release.

However, for cautious decisions it may be more appropriate to draw contour map of the worst case scenario. Such an estimate may be overly conservative.

Therefore, we propose to provide also histograms of distributions of the cumulative dose at selected point of interest such as those displayed in Fig. 7. The locations of

these points of interest were selected such that their distance from the release site is growing, see Fig. 1. Note that while the predicted total dose at the closest location does not change after $t = 4$, Fig. 7 top row, the prediction of the same quantity at distant locations is improving with every data record, Fig. 7, middle and bottom row.

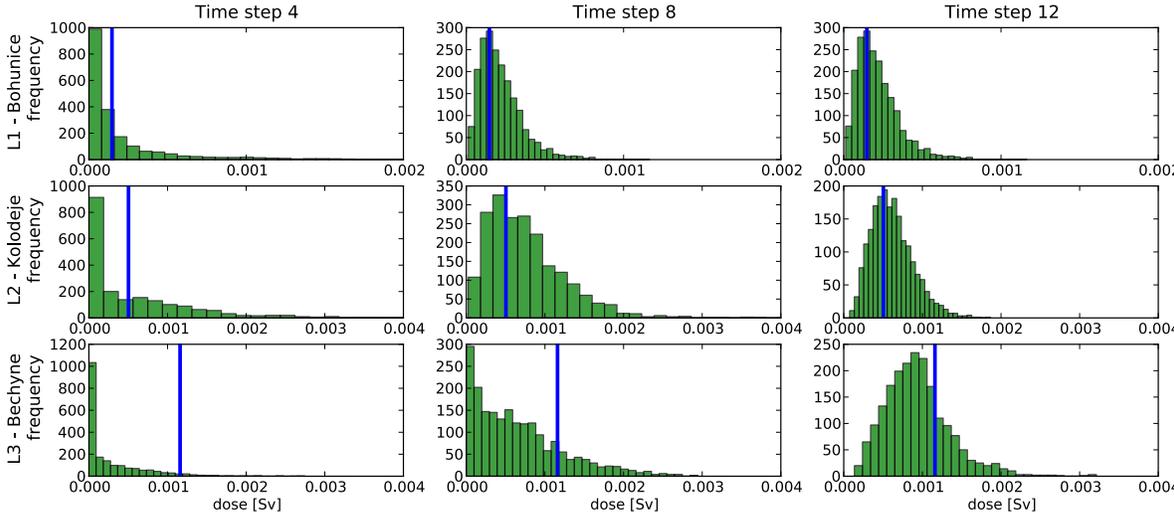


Figure 7: Distribution of the accumulated dose at different locations after 4 hours ($t = 24$) from the start of the release. Each row represents one location from Fig. 1, L1–L3 from the top. Each column corresponds to the available information based on measurements available at times $t = 4, 8, 12$, respectively.

5 Conclusion

Every accidental release of a radioactive material is burdened with uncertainty in the release dynamics and meteorological conditions. The most demanding conditions for data assimilation are in the early phase of radiation accident when only on-line measurements from the radiation monitoring network are available. Bayesian methods and Monte Carlo techniques in particular has been shown to be effective tools for data assimilation and prediction of consequences of the accident under idealized conditions. The idealization was typically either in considering important parameters of the release to be known, and/or assuming direct observability of the concentration of activity, and computational time not being an issue. However, in operational setting, many more parameters needs to be considered as unknown which leads to infeasible computational cost.

In this paper, we introduced parametric models for both the release dynamics and the meteorological forecast. Parameters of these models are estimated together with spatial distribution of the released material. Realistic model of available dose measurements from a typical radiation monitoring network was created. Computational cost

of the implied scheme is significantly reduced by introduction of application tailored proposal densities. The proposal density is computed from a simplified model using Laplace approximation. The particle filter then acts only as a corrector of the Laplace approximation which requires much less computational effort.

The performed twin experiment confirms that a particle filter with moderate size of 2000 particles is capable of computing realistic estimation of radiation situation as well as its short term forecast. Careful analysis of the puff model and its software implementation ensures that the resulting assimilation and prediction algorithms can be evaluated in real time on a contemporary personal computer. This is very promising result for potential application of this methodology for automatic operational use.

Acknowledgement

Support of grant VG20102013018 is gratefully acknowledged.

A Conjugate update of inverse Gamma likelihood and Gamma prior

Consider the inverse Gamma likelihood $p(y_t|a) = i\mathcal{G}(\alpha, (\alpha - 1)a)$ and Gamma prior $p(a) = \mathcal{G}(k, \theta)$. The Bayes rule yields

$$\begin{aligned}
 p(a|y_t) &\propto p(y_t|a)p(a) \\
 &\propto \frac{((\alpha - 1)a)^\alpha}{\Gamma(\alpha)} y_t^{-\alpha-1} \exp\left(-\frac{(\alpha - 1)a}{y_t}\right) \frac{a^{k-1}}{\theta^k \Gamma(k)} \exp\left(-\frac{a}{\theta}\right) \\
 &\propto a^{(\alpha+k-1)} \exp\left(-a\left((\alpha - 1)y_t^{-1} + \theta^{-1}\right)\right) \\
 &= \mathcal{G}\left(\alpha + k, \left[(\alpha - 1)y_t^{-1} + \theta^{-1}\right]^{-1}\right).
 \end{aligned} \tag{35}$$

The normalization factor follows from the standard result of the Gamma density.

B Laplace Approximation of shifted inverse Gamma likelihood and Gamma prior

Consider likelihood in the form of a product of inverse Gamma distributions $p(y_t|a) = \prod_{j=1}^m i\mathcal{G}(\alpha_j, \beta_j a + m_j)$ and Gamma prior $p(a) = \mathcal{G}(k, \theta)$. The Bayes rule yields

$$p(a|y_t) \propto p(a)p(y_t|a) \propto \frac{a^{k-1}}{\theta^k \Gamma(k)} \exp\left(-\frac{a}{\theta}\right) \times \prod_{j=1}^m \frac{(\beta_j a + m_j)^{\alpha_j}}{\Gamma(\alpha_j)} y_t^{-\alpha_j-1} \exp\left(-\frac{\beta_j a + m_j}{y_t}\right) \quad (36)$$

Note that contrary to the non-shifted case in (35), the resulting density is not in a standard form and the normalization can not be obtained analytically. We propose to approximate (36) using Laplace approximation Kass et al. (1990). Specifically, we approximate (36) by a probability density at its maximum value:

$$\hat{a} = \arg \max_a (\log(p(a|y_t))). \quad (37)$$

Taking the first derivative of logarithm of (36)

$$\begin{aligned} \frac{d \log(p(a|y_t))}{da} &= \frac{d}{da} \left(\sum_{j=1}^m \left[\alpha_j \log(\beta_j a + m_j) - \frac{\beta_j a + m_j}{y_t} \right] + (k-1) \log a - \frac{a}{\theta} \right) \\ &= \sum_{j=1}^m \left[\frac{\alpha_j \beta_j}{\beta_j a + m_j} - \frac{\beta_j}{y_t} \right] + \frac{k-1}{a} - \frac{1}{\theta}, \end{aligned} \quad (38)$$

we note that it is a sum of rational functions in a and thus it is strictly decreasing. Therefore, it has only a single intersection with zero at \hat{a} , which can be efficiently found using numerical methods such as the Newton Raphson method. Once the maximum value is established, we compute the second derivative of (38) at point \hat{a} and set it equal to inverse covariance matrix of the Normal distribution:

$$\frac{1}{2} \sigma_a^{-2} = \sum_{j=1}^m \left[\frac{\alpha_j \beta_j^2}{(\beta_j \hat{a} + m_j)^2} \right] + \frac{k-1}{\hat{a}^2},$$

to form the approximate posterior

$$p(a|y_t) \approx \mathcal{N}(\hat{a}, \sigma_a^2).$$

References

- H. Eleveld, Y. Kok, C. Twenhofel, Data assimilation, sensitivity and uncertainty analyses in the Dutch nuclear emergency management system: a pilot study, *International Journal of Emergency Management* 4 (3) (2007) 551–563.
- V. Winiarek, J. Vira, M. Bocquet, M. Sofiev, O. Saunier, Towards the operational estimation of a radiological plume using data assimilation after a radiological accidental atmospheric release, *Atmospheric Environment* .
- H. Jeong, E. Kim, K. Suh, W. Hwang, M. Han, H. Lee, Determination of the source rate released into the environment from a nuclear power plant, *Radiation protection dosimetry* 113 (3) (2005) 308.
- I. Kovalets, V. Tsiouri, S. Andronopoulos, J. Bartzis, Improvement of source and wind field input of atmospheric dispersion model by assimilation of concentration measurements: Method and applications in idealized settings, *Applied Mathematical Modelling* 33 (8) (2009) 3511–3521.
- S. Haupt, A. Beyer-Lout, K. Long, G. Young, Assimilating concentration observations for transport and dispersion modeling in a meandering wind field, *Atmospheric Environment* 43 (6) (2009) 1329–1338, ISSN 1352-2310.
- G. Cervone, P. Franzese, A. Grajdeanu, Characterization of atmospheric contaminant sources using adaptive evolutionary algorithms, *Atmospheric Environment* 44 (31) (2010) 3787–3796.
- J. Anderson, S. Anderson, I. Metron, V. Reston, A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review* 127 (1999) 12.
- I. Senocak, N. Hengartner, M. Short, W. Daniel, Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference, *Atmospheric Environment* 42 (33) (2008) 7718–7727.
- L. Delle Monache, J. Lundquist, B. Kosovic, G. Johannesson, K. Dyer, R. Aines, F. Chow, R. Belles, W. Hanley, S. Larsen, et al., Bayesian inference and Markov Chain Monte Carlo sampling to reconstruct a contaminant source on a continental scale, *Journal of Applied Meteorology and Climatology* 47 (10) (2008) 2600–2613.
- G. Johannesson, B. Hanley, J. Nitao, *Dynamic Bayesian Models via Monte Carlo—An Introduction with Examples*, Tech. Rep., Lawrence Livermore National Laboratory, 2004.
- P. Hiemstra, D. Karssenbergh, A. van Dijk, Assimilation of observations of radiation level into an atmospheric transport model: A case study with the particle filter and the ETEX tracer dataset, *Atmospheric Environment* (2011) 6149–6157.

- N. Gordon, D. Salmond, A. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, vol. 140, IEE, 107–113, 1993.
- M. Oh, J. Berger, Adaptive importance sampling in Monte Carlo integration, *Journal of Statistical Computation and Simulation* 41 (3) (1992) 143–168, ISSN 0094-9655.
- M. Pitt, N. Shephard, Filtering via simulation: Auxiliary particle filters, *J. of the American Statistical Association* 94 (446) (1999) 590–599.
- J. Cornebise, E. Moulines, J. Olsson, Adaptive methods for sequential importance sampling with application to state space models, *Statistics and Computing* 18 (4) (2008) 461–480.
- R. Kass, L. Tierney, J. Kadane, The validity of posterior expansions based on Laplace’s method, *Bayesian and Likelihood methods in Statistics and Econometrics* 7 (1990) 473–488.
- J. Sloughter, T. Gneiting, A. Raftery, Probabilistic wind speed forecasting using ensembles and Bayesian model averaging, *Journal of the American Statistical Association* 105 (489) (2010) 25–35.
- V. Peterka, Bayesian Approach to System Identification, in: P. Eykhoff (Ed.), *Trends and Progress in System identification*, Pergamon Press, Oxford, 239–304, 1981.
- A. Doucet, N. de Freitas, N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*, Springer, 2001.
- D. Crisan, A. Doucet, A survey of convergence results on particle filtering methods for practitioners, *Signal Processing, IEEE Transactions on* 50 (3) (2002) 736–746, ISSN 1053-587X.
- D. Dowson, A. Wragg, Maximum-entropy distributions having prescribed first and second moments, *Information Theory, IEEE Transactions on* 19 (5) (1973) 689 – 693, ISSN 0018-9448, doi:10.1109/TIT.1973.1055060.
- I. Thompson, C. Andersen, L. Bøtter-Jensen, E. Funck, S. Neumaier, J. Sáez-Vergara, An international intercomparison of national network systems used to provide early warning of a nuclear accident having transboundary implications, *Radiation protection dosimetry* 92 (1-3) (2000) 89.
- P. Zannetti, *Air pollution modeling*, Van Nostrand Reinhold, 1990.
- S. Raza, R. Avila, J. Cervantes, A 3-D Lagrangian (Monte Carlo) method for direct plume gamma dose rate calculations, *Journal of Nuclear Science and Technology* 38 (4) (2001) 254–260.
- E. Browne, R. Firestone, V. Shirley, *Table of radioactive isotopes*, John Wiley and Sons Inc., New York, NY, 1986.

- K. Eben, P. Juruš, J. Resler, M. Belda, B. Kruger, Medard - An Environmental Modelling Project for the Territory of the Czech Republic, *ERCIM News* (61) (2005) 18–19.
- P. Pecha, R. Hofman, E. Pechova, Training simulator for analysis of environmental consequences of accidental radioactivity releases, in: 6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia, 2007.
- G. Golub, J. Welsch, Calculation of Gauss quadrature rules, *Math. Comp.* (23) (1969) 221–230.
- P. Pecha, R. Hofman, Construction of observational operator for cloudshine dose from radioactive cloud drifting over the terrain, in: Proc. of 14th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 2011.
- J. Bartzis, J. Ehrhardt, S. French, J. Lochard, M. Morrey, K. Papamichail, K. Sinkko, A. Sohler, RODOS: decision support for nuclear emergencies, *Recent Developments and Applications in Decision Making*, Kluwer Academic Publishers, Dordrecht, The Netherlands (2000) 379–394.