Efficient Sequential Monte Carlo Sampling for Continuous Monitoring of a Radiation Situation

Václav Šmídl, Radek Hofman

Institute of Information Theory and Automation, Prague, Czech Republic
tel: +420 26605 2420, fax: +420 26605 2068, email: smidl@utia.cas.cz

February 13, 2013

Abstract

The monitoring of a radiation situation around a nuclear power plant is a demanding task due to the high uncertainty of all involved variables and availability of measurements from a sparse monitoring network. Assessment of the situation requires an experienced specialists that may be unavailable during critical times. Our goal is to provide an automated method of instantaneous radiation situation assessment that does not underestimate its uncertainty. We propose a state space model based on the atmospheric dispersion model, local correction of a numerical weather model, and a temporal model of the released activity. This state space model is highly nonlinear and evaluation of the likelihood function requires extensive numerical calculations. Sequential Monte Carlo method is one of the few possibilities how to estimate the state recursively. Since the simple bootstrap approach yields an extremely computationally demanding algorithm, we investigate the use of existing techniques for the design of a more efficient proposal density. We propose combining the Laplace approximation and various population Monte Carlo methods. Data from an existing monitoring network were used to calibrate relevant parts of the model. Performance of the methods in a real radiation emergency situation is evaluated in a simulated experiment due to the lack of real data.

Keywords: radiation protection, atmospheric dispersion model, importance sampling, proposal density, population Monte Carlo, information geometry estimation

Contents

List of Figures

1 Introduction

We are concerned with a scenario in which the release of radionuclides into the atmosphere occurs following a hypothetical accident of a nuclear power plant facility. The radioactive effluent forms a plume that moves over the terrain according to the current meteorological situation. Urgent protective measures must be introduced as soon as possible to protect the public from the harmful effects of ionizing radiation. The necessary countermeasures are typically prescribed by law and classified according to the expected radiation level into different categories based on severity. Determining this expected value is therefore the most important factor for the decision making of the crisis management authority. The most reliable source of information is direct measurements of the radiation level. However, detailed measurements in all affected areas are typically available only several hours or days after the beginning of the release. Therefore, it may be too late to warn people in affected areas.

For this scenario, we aim to design an algorithm that processes the stream of measurements available from radiation monitoring networks (RMN) and provides a prediction of the radiation situation in real time. The algorithm combines techniques for simulation of dispersion of the pollutants in the atmosphere with statistical evaluation of their probability using data from the RMN. 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), and genetic algorithms (Haupt et al., 2009; Cervone et al., 2010). These methods typically provide only point estimates of the estimated parameters, without indication of their reliability. Reliability of these estimates is then judged from complex sensitivity studies, such as (Twenhöfel et al., 2007).

An alternative to this approach is statistical estimation which is also known in this field (Anderson et al., 1999). Indeed, the value of statistical techniques such as the 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) has been demonstrated and the techniques are gaining in popularity. An obvious advantage of these methods is their ability to evaluate the reliability of their estimates using only observed data. Moreover, the sequential Monte Carlo (SMC) approach is able to process a continuous stream of data in real time, which is necessary for a continuous monitoring system. We follow the pioneering work of Johannesson et al. (2004), where a state-space model for the task of locating a source of the release was proposed and its estimation via SMC was studied. We modify the state space model for our scenario using realistic conditions of a selected power plant. We show that estimation of the new model from data provided by a sparse RMN via the classical particle filter (Gordon et al., 1993) is computationally inefficient.

The main concern is then to find a strategy of generating efficient proposal densities. Many techniques were proposed specifically for SMC, e.g. (Pitt and Shephard, 1999), other techniques are based on adaptation of non-sequential techniques for the sequential scenario (Cornebise et al., 2008) and references therein. A specific property of the studied model is a computationally expensive evaluation of the likelihood function. Moreover, the likelihood is sharply peaked and the parameter evolution model has high variance. These properties make this application an excellent area where sophisticated adaptive proposals have significant

impact. We design an efficient proposal density by combining well known techniques (Pitt and Shephard, 1999; Cornebise et al., 2008) with lesser known concepts of information geometry (Kulhavý, 1996). We show that this proposal significantly increases computational efficiency of the SMC procedure.

2 Radiation Accident

Awareness of radiation security has increased after the Chernobyl disaster, and every nuclear power plant is now equipped with a network of radiation sensors surrounding the power plant. The sensors are connected to the central emergency warning system and continually measure radiation levels. Evaluation of the radiation situation and prediction of its evolution from these readings is very challenging for two reasons. Firstly, the spread of the pollutant in the atmosphere is a complex stochastic process with many unknowns. Secondly, the number of continually measured quantities is very limited. Each of the factors is now described in detail.

2.1 Dispersion of a Pollutant in the Atmosphere

The radioactive material is located in a very small area inside the power plant, hence the location of its release is known with sufficient accuracy. After the pollutant is released to the atmosphere, its concentration, C, depends mainly on the following phenomena: (i) advection by the wind, (ii) dispersion by turbulent processes in the atmosphere, and (iii) radioactive decay of the pollutant and its deposition. The first two processes are modeled by the following partial differential equation:

$$\frac{\partial C}{\partial \tau} + \sum_{k=1}^{3} u_k \frac{\partial C}{\partial s_k} = \sum_{k=1}^{3} K_k \frac{\partial^2 C}{\partial s_k^2}. \tag{1}$$

Here, τ is the continuous time, $\mathbf{s} = [s_1, s_2, s_3]$ are spatial coordinates in the Cartesian coordinate system, $\mathbf{u} = [u_1, u_2, u_3]$ are wind speeds in the Cartesian coordinate system, and $\mathbf{K} = [K_1, K_2, K_3]$ are diffusivity coefficients parameterizing atmospheric dispersion. However, (1) is not analytically solvable for general functional forms of u_k and K_k. Various simplified solutions arise under different assumptions (Holmes and Morawska, 2006).

In this paper, we consider the Gaussian puff approximation which is based on an instantaneous point-source release of activity. It is obtained as an analytical solution of (1) for constants u_k and K_k, and the following boundary conditions (Hanna et al., 1982): (i) the concentration is approaching zero with increasing distance for the source, (ii) the release source is infinitely small at $t = 0$, (iii) integral of the concentration over the whole space is the original released activity, Q, in becquerel Bq.

Continuous release can then be approximated as a sequence of puffs, (Zannetti, 1990), released at discrete time steps with sampling period Δt, with symbol t being reserved for the current time. A puff with activity Q_κ is released during one sampling period starting at time $\tau = \kappa\,\Delta t$ and ending at $\tau = (\kappa + 1)\Delta t$, and temporal evolution of its concentration in continuous time τ is:

$$C_\kappa(\mathbf{s},\tau) = \frac{Q_\kappa e^{-\lambda(\tau-\kappa\Delta t)}}{(2\pi)^{3/2}\sigma_1\sigma_2\sigma_3}\exp\left[-\frac{(s_1-l_{1,\kappa,\tau})^2}{2\sigma_1^2} - \frac{(s_2-l_{2,\kappa,\tau})^2}{2\sigma_2^2} - \frac{(s_3-l_{3,\kappa,\tau})^2}{2\sigma_3^2}\right]. \quad (2)$$

Here, λ denotes the decay constant of the modeled radioisotope. For the purpose of evaluation of the measurements the concentration is evaluated at discrete times t. Since t is generally used to denote the current time, the index of the puff will be often used relative to the current time, i.e. $\kappa = t - 1$. Time evolution of the location of the κth puff center is fully determined by the wind field at the previous location

$$\begin{aligned} l_{1,\kappa,t+1} &= l_{1,\kappa,t} - \Delta t\, v_t(\mathbf{l}_{\kappa,t})\sin(\phi_t(\mathbf{l}_{\kappa,t})), \\ l_{2,\kappa,t+1} &= l_{2,\kappa,t} - \Delta t\, v_t(\mathbf{l}_{\kappa,t})\cos(\phi_t(\mathbf{l}_{\kappa,t})), \\ l_{3,\kappa,t+1} &= l_{3,\kappa,t}, \end{aligned} \quad (3)$$

where $v_t(\mathbf{l}_{\kappa,t})$ and $\phi_t(\mathbf{l}_{\kappa,t})$ are the wind speed and wind direction at location $\mathbf{l}_{\kappa,t}$, respectively. Specifically, $l_{3,\kappa,t}$ is the effective release height given by the height of the release plus the plume rise. For practical reasons, the rate of dispersion is here parametrized using dispersion coefficients $\sigma_k = \sqrt{2K_k t}$, (Gifford, 1976).

Remark 1. The puff model described in this Section is only one of many approximations available. In this text, we use the term *dispersion model* in the sense of a computer program that represents the spatial concentration of a pollutant in a parametric form and is capable to compute its propagation in time given the wind field and the new released activity. Such a model may include additional physical phenomena such as dry and wet deposition of the pollutant, or terrain profile.

2.2 Continuous Monitoring System

A radiation accident has several phases that can be formally distinguished (Raskob et al., 2010): (i) *prerelease phase*: an abnormal situation happened in the power plant, however, the radiation has not leaked outside of the reactor building yet, (ii) *early phase*: the radiation material is being released from the power plant into the atmosphere and the radioactive cloud is still within the monitored area, and (iii) *late phase*: the radiation cloud has already passed. An example of a hypothetical release is displayed in Figure 1.

Radiation accidents are typically studied retrospectively from the data recorded in all phases of the accident. More and more data are being measured in the late phase which allows to deduce what has happened and what will be the final consequences. For example, aerial surveillance allows to measure radionuclide deposition on the ground with high spatial resolution which allows simple interpolation to obtain clear contour of the total dose. However, it is usually too late for the sheltering of citizens living near the power plant that are in danger already during the early phase. Since there is no time to add sensors before the early phase starts, we can use only data that are available routinely during normal operation of the plant. This is predominantly the RMN which is built around a power plant, typically in circles with poor spatial resolution, Figure 1. The amount of observed data provided by the RMN during a release is very small, e.g. less than 100 scalar values in the release displayed in Figure 1. The aim is to process this data recursively without the need of human interaction. This objective represents many challenges for modeling and computational evaluation of the results.

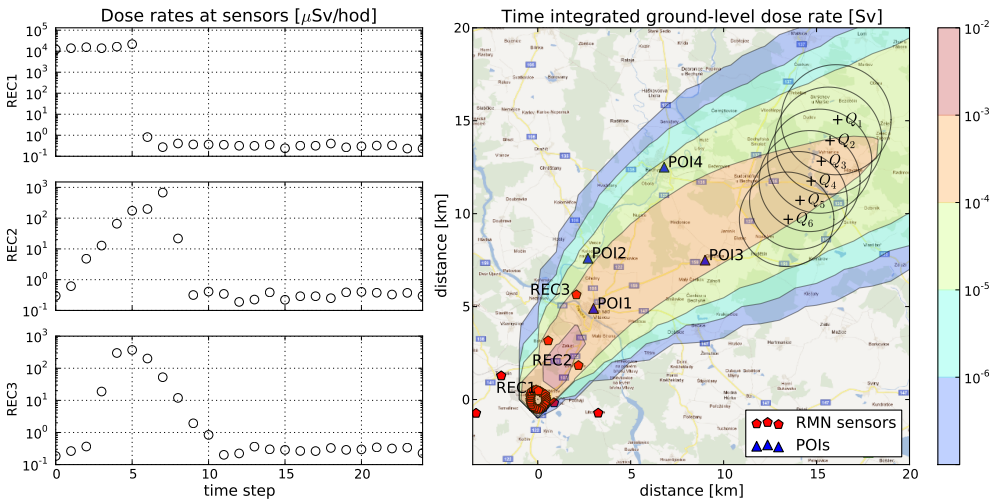

Figure 1: Simulation experiment of an accidental radiation release. **Left**: Simulated measurements of three RMN sensors with the highest observed values. **Right**: Contour plot of the total ground-level radiation dose accumulated during the first three hours after the release; locations of the radiation dose sensors are denoted by pentagons; largest populated areas (points of interest) are denoted by triangles, crosses denote centers of the puffs and circles are drawn around them with diameter $2\sigma_1$ (2).

First, meteorological parameters would only be available in the form of numerical predictions without any correction from a larger set of meteorological data. Local calibration of the weather forecast is needed. Second, the data available in the early phase are sparse in time and space. Note that the hypothetical release in Figure 1 generated only eight data points on only a few sensors of the RNM. Uncertainty about the true situation is therefore very high. If the continuous monitoring system is to be considered trustworthy by the decision makers, it must admit this uncertainty. Underestimation of the uncertainty is considered as a severe flaw of the system. Third, the procedure must be affordable to compute on an inexpensive hardware. The use of expensive equipment is not justified since the algorithm will analyze measurements of the natural background radiation most of the time.

3 State Space Model for Continuous Radiation Monitoring

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

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{v}_t), \tag{4}$$

$$\mathbf{y}_t = g(\mathbf{x}_t, \mathbf{w}_t), \tag{5}$$

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 observation, respectively. Both transformations are subject to disturbances $\mathbf{v}_t$ and $\mathbf{w}_t$ which are considered to be random samples from a known probability density.

In principle, the state variable should contain the concentration, C, which is infinite dimensional since it is a solution of the PDE (1). However, an estimation of such system from sparse measurements is difficult. Therefore, we follow Johannesson et al. (2004) and approximate the spatio-temporal distribution of the pollutant by an approximate deterministic model (in our case the puff model (2)) with unknown parameters. The state $\mathbf{x}_t$ then contains only the selected unknown input values of the dispersion model, the remaining input values will be considered to be known parameters (i.e. internal parameters of the computer program, Remark 1).

The deterministic dispersion model is treated as a mapping from the space of the state $\mathbf{x}_t$ to the expected values of the measurements, $\hat{\mathbf{y}}_t = \eta(\mathbf{x}_t)$. As an example of such mapping, the puff model from Section 2.1 will be used in this paper. Extension of the deterministic model to a full stochastic state space model thus requires to choose transition probability density (4), and the measurement probability density $\mathbf{y}_t \sim p(\mathbf{y}_t|\hat{\mathbf{y}}_t)$. Previously published approaches differ in what quantities of the dispersion model are considered to be unknown (belonging to the state variable $\mathbf{x}_t$) or known and thus internal to the deterministic dispersion model. For example, the model of Johannesson et al. (2004) was designed for localization of unknown source of radiation and the state variable was composed from the location of the source, $\mathbf{l}_{\tau,0}$, and activity of the released material Q_t. All remaining variables, including the wind field were assumed to be known. Further extension of the model for unknown dispersion coefficients is also available (Senocak et al., 2008; Delle Monache et al., 2008).

The assumptions mentioned above are only partially relevant for the purpose of the continuous monitoring system which is described in Section 2.2. The location of the source is well known. Sensitivity studies of the parameters of the puff model (Twenhöfel et al., 2007) identified the following parameters as the most significant: (i) quantity of the released material at each time step, (ii) the wind speed and wind direction. Models for these two terms has been proposed in (Johannesson et al., 2004) and (Hiemstra et al., 2011), respectively. The number of unknown parameters is much higher, however, we will analyze only these three to illustrate the challenges of the problem.

The choice of the probability density of the measurements is relatively easy since it can be deduced from the characteristics of the measuring devices. However, different types of measurements can be considered. Measurement of the concentration C, which is assumed in all previous works, is too expensive for real application. Existing RMNs are composed of sensors of the integrated dose rate, hence we have to design a new observation model for this quantity. We will now shortly discuss all components of the model and the choice of the involved probability densities.

3.1 Elicitation of Probability Density Functions

In Bayesian formalism, probability densities in (4)–(5) must be fully specified. However, exact shape of the densities in the studied application is not known and must be either obtained experimentally or assumed. For example, the manufactures specify accuracy of the measuring devices only by their relative accuracy in

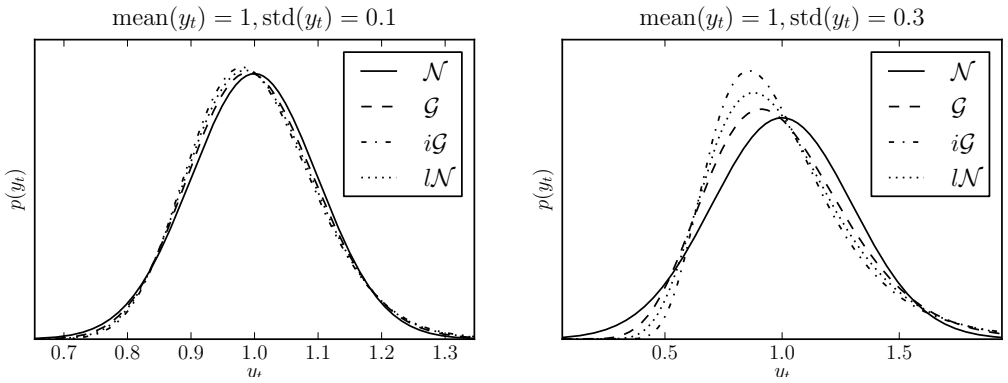

Figure 2: Comparison of different probability densities with equal first two moments. All probability densities have the same mean value $\mu_t = 1$, but different standard deviations $\gamma = 0.1$ (left) and $\gamma = 0.3$ (right).

percents of the measured value. This can be interpreted as two statistical moments of the distribution, e.g.

$$\text{mean}(y_t) = y_t^{true} = \mu_t, \quad \text{std}(y_t) = \gamma\mu_t, \tag{6}$$

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 maximum measurement error provided by the manufacturer is interpreted as being equal to 2 standard deviations.

The choice of the full density, $p(y_t)$, can be done via moment matching. The following parametric forms have the same moments as (6):

$$y_t \sim \mathcal{N}(\mu_t, (\gamma\mu_t)^2), \text{ or } \mathcal{N}(\mu_t, \sigma_t^2), \tag{7}$$

$$y_t \sim \mathcal{G}(\gamma^{-2}, \gamma^2\mu_t), \tag{8}$$

$$y_t \sim i\mathcal{G}\left(\gamma^{-2} + 2, (\gamma^{-2} + 1)\mu_t\right), \tag{9}$$

Here, $\mathcal{N}(\mu_t, \sigma_t^2)$ denotes Normal probability density with mean μ_t and variance σ_t^2; $\mathcal{G}(k, \theta)$ denotes Gamma density with shape parameter k and scale parameter θ; and $i\mathcal{G}(\alpha, \beta)$ denotes the inverse Gamma density with shape parameter α and scale parameter β. Comparison of these probability densities 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 a higher relative error, when the gamma densities assign very low probabilities to the region around zero.

From these three choices, the Normal density (7) is typically favored since it has the greatest entropy of all possible densities (Dowson and Wragg, 1973). However, this probability density has to be truncated if the variable is positive by the definition, the other densities are more natural choices for positive variables. Since they are almost equivalent, our choice of a particular form will be motivated primarily by analytical advantages of their subsequent estimation.

3.2 Release Scenario

In previous work, (Johannesson et al., 2004; Hiemstra et al., 2011), the model of the magnitude of the release was split into the model of the first step of the release and a random walk on its temporal evaluation. Using Gamma density, the model of the first step is:

$$p(Q_1) = \mathcal{G}(\alpha_Q, \beta_Q), \tag{10}$$

followed by random walk

$$p(Q_t|Q_{t-1}) = \mathcal{G}(\gamma_Q^{-2}, \gamma_Q^2 Q_{t-1}),\ t > 2, \tag{11}$$

where γ_Q governs the spread of the random walk.

However, random walk model of temporal evolution of the accidental release is highly unreliable due to frequent abrupt changes, as demonstrated by events in the Fukushima Daiichi accident (Katata et al., 2012), which was a sequence of rapid changes with only a few stationary periods. This would suggest the need for a hidden Markov label field switching models (10) and (11). However, this would increase the complexity of the evaluation scheme. Therefore, we propose to ignore the stationary periods and a priori assume that the puffs are temporally uncorrelated, i.e. the prior (10) holds for all t. Moreover, we aim for completely uninformative prior, i.e. for $\alpha_Q = 1$ and $\beta_Q \to 0$ for which (10) approaches the improper Jeffrey's prior on the scale parameter (Jeffreys, 1961). This model is suitable only if the measurements are informative about the released quantity.

3.3 Wind Field Corrections

We assume that the radioactive pollutant is released from a nuclear power plant of known location, $[s_{1,pp}, s_{2,pp}]$ and known altitude $s_{3,pp}$, vector notation, $\mathbf{s}_{pp} = [s_{1,pp}, s_{2,pp}, s_{3,pp}]$. From this point it is advected by the wind field (3). While it is possible to obtain numerical weather forecast from various sources, its accuracy at the power plant location is usually poor. An illustration of this fact is displayed in Figure 3 by comparison of the wind direction obtained from the numerical weather forecast and from the meteostation at the power plant.

Since accurate predictions of the wind field are the most critical variable in the task of prediction of the radiation situation in the early phase of the accident, we need to calibrate the forecast for the location of the power plant. Due to the limited amount of meteostations, we need a really simple parametrization of the calibration. We follow (Hiemstra et al., 2011) and choose model

$$v_t(\mathbf{s}) = \tilde{v}_t(\mathbf{s}) a_t, \tag{12}$$

$$\phi_t(\mathbf{s}) = \tilde{\phi}_t(\mathbf{s}) + b_t, \tag{13}$$

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. Constants a_t and b_t are unknown biases of the weather prediction model at time t. Correction of the wind field forecast is then achieved by estimation of a_t and b_t. More complex models of the local

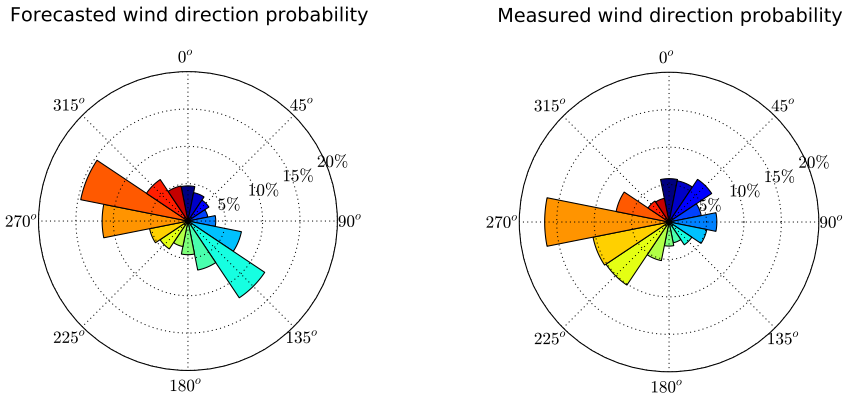

Figure 3: Comparison of histograms of the wind direction at the location of the power plant for year 2008 from the numerical model (left) and observed data (right).

wind field are available, (Monbet et al., 2007), however, their calibration requires significantly more points of measurement than we have in the area.

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 densities (7) and (8) yields

$$\begin{aligned} p(a_t|a_{t-1}) &= \mathcal{G}(\gamma_a^{-2}, \gamma_a^2 a_{t-1}), \\ p(b_t|b_{t-1}) &= t\mathcal{N}(b_{t-1}, \sigma_b, \langle b_{t-1} - \pi, b_{t-1} + \pi \rangle). \end{aligned} \tag{14}$$

The Gamma density was chosen for a_t since its variance is proportional to its mean value and allows for conjugate update. The variance in the wind direction is constant, hence we have chosen the truncated Normal with support on the unit circle.

3.4 Measurement Model

Measurement of the local concentration of the pollutant $C(\mathbf{s}, \tau)$ considered in (Johannesson et al., 2004; Haupt et al., 2009) requires a large and expensive device. More affordable Gaiger counter typically measures a time integral of radiation dose rates which is measured in sieverts per hour [Sv/h]. The radiation monitoring network considered in this text is equipped with several radiation dose sensors and a meteostation. The Geiger counters have a fixed operation range from several nSv/h to Sv/h. The meteostation is equipped with an anemometer providing measurements of the wind speed and wind direction.

3.4.1 Integrated Dose Rate of the Puff Model

In this Section, we derive a mapping $\hat{\mathbf{y}}_t = \eta(\mathbf{x}_t)$ for the puff model from Section 2.1. The purpose is to demonstrate computational cost of such operation, evaluation of the dose rate for a more complex dispersion

model would be even more costly.

If a single puff (2) of activity Q_κ is the only source of radiation, the jth sensor in the radiation monitoring network in location $\mathbf{s}_{R_j}$ would register dose $c_{j,\kappa,t}Q_\kappa$, where the coefficient $c_{j,\kappa}$ is computed as (Raza et al., 2001):

$$c_{j,\kappa,t} = \frac{K_c}{Q_\kappa} \int_{t-1}^{t} \Phi(\mathbf{s}_{R_j}, \tau, E)\ d\tau. \tag{15}$$

Here, $K_c = \omega\, K\, E\, \mu_a \rho^{-1}$ is a physical constant computed from the dose conversion factor K, the gamma energy E produced by the decay of the assumed radionuclide, the ratio of absorbed dose in tissue to the absorbed dose in air ω, the air density ρ, and the mass attenuation coefficient μ_a. Fluency rate $\Phi(\mathbf{s}_{R_j}, \tau, E)$ from the puff is calculated as the following three dimensional integral over the volume of the puff:

$$\Phi(\mathbf{s}_{R_j}, \tau, E) = \int_\Omega \frac{C_\kappa(\mathbf{s},\, \tau)\,(1 + k\,\mu r)\exp(-\mu r)}{4\pi r^2} d\mathbf{s}. \tag{16}$$

Here, ambient activity concentration $C(\mathbf{s},\, \tau)$ is defined by (2); $k = (\mu - \mu_a)\mu^{-1}$ where μ is a linear attenuation coefficient (Raza et al., 2001); $\Omega \subset \mathbb{R}^3$ 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 sensor $\mathbf{s}_{R_j}$. The full puff model, i.e. a sequence of all puffs, contributes to the jth sensor by dose

$$y_{Q,j,t} = \sum_{\kappa=1}^{K} c_{j,\kappa,t} Q_\kappa. \tag{17}$$

The observation operator transforming activity concentration in the air to the time integrated dose rate (16)–(17) is in the simplest possible form and it is still highly nonlinear. Analytical solution is not available and the dose (17) must be evaluated numerically. Many methods were proposed to increase efficiency of its evaluation, such as the "n/μ" method (Pecha and Hofman, 2011), however, it requires substantial computational resources.

3.4.2 Measuring Devices

The measured value of the radiation dose is a sum of natural background radiation, y_{nb}, and the dose from the release, y_Q (17). According to the studies of the radiation dose sensors (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 7–20%. Therefore, we assume that the measurements at jth sensor have moments

$$\begin{aligned} \text{mean}(y_{j,t}) &= y_{nb,j} + y_{Q,j,t}, & (18)\\ \text{std}(y_{j,t}) &= \gamma_y(y_{nb,j} + y_{Q,j,t}), & (19) \end{aligned}$$

where $y_{nb,j}$ is the radiation background at the jth sensor. We choose the inverse Gamma density,

$$p(y_{j,t}|\mathbf{x}_{1:t}) = i\mathcal{G}(\gamma_y^{-2} + 2,\, (\gamma_y^{-2} + 1)(y_{nb,j} + y_{Q,j,t})). \tag{20}$$

The value of natural background $y_{nb,j}$ is different for each sensor, however, it is relatively stable in time, see measurements for the whole year 2010 in Figure 4 right. Since the variance of the measurements corresponds

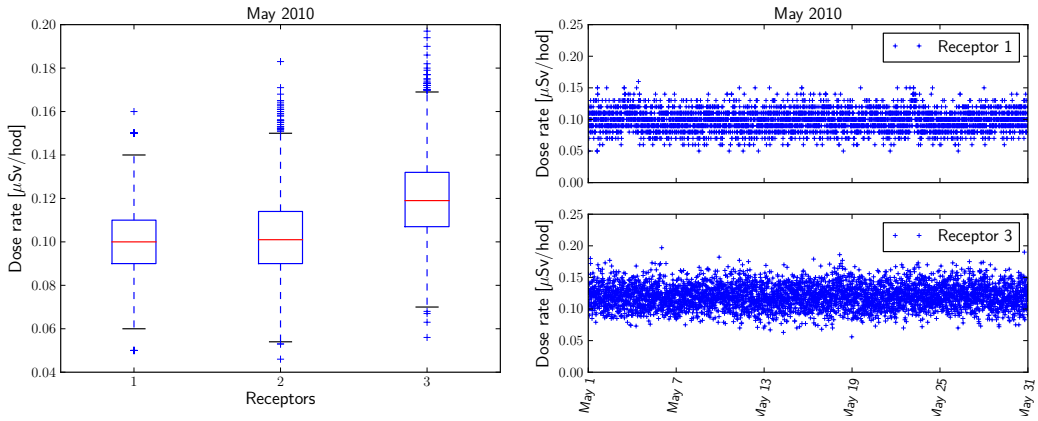

Figure 4: Natural background measured by selected sensors of the radiation network in May 2010.

well to the model (20), we consider $y_{nb,j}$ to be known and fixed at the average value on the sensor over the previous year.

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

$$\text{std}(v_t) = \gamma_v v_t(\mathbf{s}_{meteo}), \qquad \text{std}(\phi_t) = \sigma_\phi,$$

where $\mathbf{s}_{meteo}$ is the location of the meteostation. From all choices of possible densities (7)–(9) we select the inverse Gamma form (9) for v_t. Probability density of the wind direction ϕ_t is considered as Normal with fixed variance, (7).

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.

$$\mathbf{x}_t = [a_t, b_t, Q_{t-K}, \dots Q_t, \mathbf{l}_{t-K}, \mathbf{l}_2 \dots, \mathbf{l}_t],$$

where $\mathbf{l}_{t-K}$ is the location of the oldest puff that is still within the monitored area.

The vector of measurements is composed of one anemometer and J radiation dose sensors

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

Parameter evolution model is then composed of all considered models

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = p(a_t|a_{t-1})p(b_t|b_{t-1})p(Q_t), \tag{21}$$

given by (14), (10) and deterministic model (3). The observation model is

$$p(\mathbf{y}_t|\mathbf{x}_t) = p(v_t, \phi_t|a_t, b_t) \prod_{j=1}^{J} p(y_{Q,j,t}|Q_{t-K:t}, \mathbf{l}_{t-I:t}). \tag{22}$$

State space model (21)–(22) is designed for the purpose of continuous monitoring, where simplicity is extremely important due to the lack of data (see Figure 1 left, for illustration). Therefore, many elements that are in principle uncertain are considered to be known:

- Composition of the radionuclides in the release is critical for evaluation of the observations (18), since majority of the parameters in (15) is known for specific radionuclides. Different composition of the release would yield different measurements. For this purpose, many scenarios of potential power plant faults are prepared with precomputed parameters of the expected release. We assume that the specific scenario is selected from the list of prepared options in the pre-release phase of the accident by an expert or an expert system of the power plant.

- Another important parameter is the temperature of the released material which influences the altitude of the release. Once again, we assume that the temperature is available from the release scenario. In this text, we consider a common scenario that the pollutant is captured in the containment for long enough time to cool down to the temperature of the surrounding air and the height of the puff is then equal to the height of the release.

- Dispersion of the pollutant depends on many more atmospheric parameters than just the considered wind speed and direction. For example, the height of the mixing layer and dispersion coefficients. In this paper, we set these parameters to their typical values for an aggregated parameter called Pasquill's category of stability (Hanna et al., 1982) which is provided by the meteostation.

- The error of approximation of the reality by the puff model can be taken into account only via the measurement error (20). In practice, the error is spatially dependent and a detailed correction term as in (Kennedy and O'Hagan, 2001) would be appropriate.

The accuracy of the presented model is thus very dependent on the correct scenario of the release which provides all necessary input parameters for the dispersion model. Potential extensions of the state space model may thus estimate the most likely scenario from a predefined set or any of the relevant parameters of the model. Reliable estimation of these extensions is possible only with more measurements, provided e.g. by unmanned aerial vehicles (Šmídl and Hofman, 2013).

4 Sequential Monte Carlo Estimation

The model (4)–(5) is fully specified by probability density functions (21)–(22). Sequential estimation of the posterior state probability is based on recursive evaluation of the filtering density, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, using Bayes rule (Peterka, 1981):

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}, \tag{23}$$

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1}, \tag{24}$$

where $p(\mathbf{x}_1|\mathbf{y}_0)$ is the prior density, and $\mathbf{y}_{1:t} = [\mathbf{y}_1, \ldots, \mathbf{y}_t]$ denotes the set of all observations. The integration in (24), and elsewhere in this paper, is over the whole support of the involved probability density functions.

Equations (23)–(24) are analytically tractable only for a limited set of models. The most notable example of an analytically tractable model is linear Gaussian for which (23)–(24) are equivalent to the Kalman filter. For other models, (23)–(24) 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

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^{(i)} \delta(\mathbf{x}_{1:t} - \mathbf{x}_{1:t}^{(i)}), \tag{25}$$

where $\mathbf{x}_{1:t} = [\mathbf{x}_1, \ldots, \mathbf{x}_t]$ is the state trajectory, $\{\mathbf{x}_{1:t}^{(i)}\}_{i=1}^N$ are samples of the trajectory (the particles), $w_t^{(i)}$ is the weight of the ith 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 and arbitrary model (4)–(5) given a suitable proposal density, $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$, yielding

$$w_t^{(i)} \propto \frac{p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}. \tag{26}$$

An important property of (26) is the possibility of recursive evaluation. However, it often converges to a degenerate system where one particle has weight of 1 and the others zero. One way to remedy the situation is the use of the *resampling* procedure, where existing particles are copied or removed according to their $w_t^{(i)}$ so 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:

$$p(\mathbf{x}_{t:t+h}|\mathbf{y}_{1:t}) = \sum_{i=1}^{N} w_t^{(i)} \prod_{\tau=1}^{h} p(\mathbf{x}_{t+\tau}^{(i)}|\mathbf{x}_{t+\tau-1}^{(i)}). \tag{27}$$

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 $w_t^{(i)}$, (26). Hence, reevaluation of the predictions for new observed data is computationally cheap.

4.1 Choice of the Proposal Density

Proposal density is often the main factor in computational efficiency of the particle filter and was heavily studied for this purpose. The optimal proposal density is (Doucet et al., 2001):

$$\begin{aligned} q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) &= q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) q(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1}), \\ q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) &= \frac{p(x_t|x_{t-1}) p(y_t|x_t)}{\int p(x_t|x_{t-1}) p(y_t|x_t) dx_t}. \end{aligned} \tag{28}$$

However, evaluation of the integral in (28) is computationally intractable and (28) is helpful only as a theoretical concept. The goal is to approximate (28) as closely as possible, with many approaches how to achieve it. From the range of possibilities, we will focus on the following options:

- the original approximation $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) \equiv p(\mathbf{x}_t|\mathbf{x}_{t-1})$ of Gordon et al. (1993), which is often called the *bootstrap* approximation. The main advantage of this choice is simplicity of the resulting algorithm.

- local linearization of (28) via the Taylor expansion (Pitt and Shephard, 1999; Doucet et al., 2000), which is also known as the Laplace's approximation (Kass and Raftery, 1995).

- parametric representation of the proposal, $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \theta)$, and estimation of the parameter using several populations of particles. This technique is well known in the classical Monte Carlo methods, (Oh and Berger, 1992; Rubinstein and Kroese, 2004), and has been used in sequential Monte Carlo in (Cornebise et al., 2008).

Each of these approaches is well suited for models that meet their assumptions. For more complex models, such as the one in Section 3, it is advantageous to combine them for different parts of the model to improve the performance.

4.2 Problem Specific Proposal Density

The problem of generating good samples of parameters of dispersion models has already been studied in Johannesson et al. (2004), where a combination of MCMC and SMC has been proposed. This is particularly advantageous for estimating an unknown location of the source. However, this approach is not suitable for the continuous monitoring system where evaluation of the likelihood function requires a computationally expensive numerical routine, and the computational resources per each time step are fixed and limited.

On the other hand, we may use specific features of the scenario. Specifically, the radiation monitoring networks were designed to measure quantities that are essential in prediction of the consequences. Therefore, we have direct observability of the forecast biases a_t, b_t from the anemometer (with large variance though), and almost direct observability of the released activity Q_t from the first ring of sensors in the proximity of the power plant (Figure 1). These measurements are informative about the corresponding state variables, but their accuracy is insufficient for exact estimation. Note that even a small error in the wind direction will cause large error in readings of the radiation sensors.

Therefore we propose the following two stage proposal density:

1. Each state variable is estimated from the observed quantity that is the most informative. The proposal density (28) is then conditionally independent as follows:

$$q^{[1]}(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) \quad = \quad q(a_t|a_{t-1}, v_t)q(b_t|b_{t-1}, \phi_t)q(Q_t|\mathbf{y}_{1:t}, \mathbf{l}_{1:t})q(\mathbf{l}_t|\mathbf{l}_{t-1}, a_t, b_t). \tag{29}$$

2. Using the particles from the first stage, we choose a parametric form $q^{[2]}(\mathbf{x}_t|\theta)$, and estimate its parameters using information geometric approach of Kulhavý (1996).

The first stage allows to follow rapid changes in the wind field and/or the released activity. The second stage relaxes the assumption of conditional independence and provides proposal based on all measured quantities.

The second stage can be repeated several times in the sense of population Monte Carlo method (Cappé et al., 2004).

The family of population Monte Carlo (PMC) methods (Cappé et al., 2004) is based on the fact that the full set of N particles is not sampled form the same proposal density, but it is split between M populations, each of $n^{[m]}$ particles, $m = 1, \ldots, M$, $\sum_{m=1}^{M} n^{[m]} = N$. Each population is sampled from different proposal function $q(\mathbf{x}_t|\hat{\theta}^{[m]}), m = 1, \ldots, M$. We will study three variants of the PMC:

PMC: the original algorithm of Cappé et al. (2004), where the samples from the mth population are used to estimate the parameters $\hat{\theta}_{m+1}$ and are discarded afterwards. The final estimate is then based only on the samples from the Mth population. This is computationally inefficient but can be proven to converge to the optimal solution.

AMIS: a way how to increase computational efficiency has been proposed by (Cornuet et al., 2012) using the idea of deterministic mixture. Each parametric proposal generated by the previous populations is interpreted as a component of a mixture density

$$q_{AMIS}^{[m]}(\mathbf{x}_t|\theta) = \sum_{k=1}^{m} \frac{n^{[k]}}{\sum_{k=1}^{m} n^{[k]}} q(\mathbf{x}_t|\hat{\theta}^{[k]}), \tag{30}$$

and the weights (26) are reevaluated after each population. The estimates of the parameters $\hat{\theta}^{[m+1]}$ are evaluated using all re-weighted samples. This approach improves computational efficiency, however proof of convergence to the optimal value is not available.

mAMIS: a modification of the AMIS procedure that can be proven to converge to the optimal value (Marin et al., 2012). It still uses the mixture density (30) but the parameter estimates are based only on the sampled from the previous population. As a consequence, it is sufficient to recompute the weights using (30) only after the last population.

From this point of view, the proposal density (29) from the fist stage is only a smart initialization of the PMC.

4.3 First Stage: Update of Conditionally Independent Densities

For the choice of inverse Gamma density (20) for the likelihood of the measured wind speed, the Gamma transition models (14) are conjugate with posterior density in the form of Gamma density. So is the Normal likelihood of the wind direction and its Normal random walk model (14). The first two factorized densities in (29) are then

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

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

with shaping parameters

$$
\begin{aligned}
k_{a,t} &= \gamma_a^{-2} + \gamma_v^{-2} + 2, \\
\theta_{a,t} &= \left[\gamma_a^{-2} a_{t-1}^{-1} + \tilde{v}_t v_t^{-1} (\gamma_v^{-2} + 1)\right]^{-1}, \\
\mu_{b,t} &= r_{b,t} \left[\sigma_b^{-2} b_{t-1} + \sigma_\phi^{-2} (\phi_t - \tilde{\phi}_t)\right], \\
r_{b,t} &= (\sigma_b^{-2} + \sigma_\phi^{-2})^{-1}.
\end{aligned} \tag{33}
$$

Derivation of $q(Q_t|y_{1:m}, \mathbf{s}_{1:t})$ is more demanding since the likelihood (20) is not conjugate. We propose to follow the Laplace's method (Kass and Raftery, 1995), based on local linearization via the Taylor expansion. The conditional distribution on the dose Q_t is then:

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

where algorithm for evaluation of moments μ_Q, σ_Q is described in Appendix A. Since Q_t has positive support, the general results from Appendix A has to be truncated.

4.4 Second Stage: Multiple Populations

Once we have an empirical approximation of the posterior in the form of empirical density from the first stage, we may use it to initialize the population-based algorithms. We choose the parametric form to be

$$
q(\log a_t, b_t, \log Q_t|\mu_\theta, \Sigma_\theta) = \mathcal{N}(\mu_\theta, \Sigma_\theta), \tag{35}
$$

which requires for an additional Jacobian in the evaluation of the likelihood function. Estimation of the parameters $\theta = [\mu_\theta, \Sigma_\theta]$ can be done via the cross entropy (CE) minimization (Rubinstein and Kroese, 2004). The idea of CE is to choose a value of the parameter θ as the one that minimizes the Kullback-Leibler divergence between the empirical representation (25) and $q(\mathbf{x}_t|\theta)$. For parametric forms from the exponential family, the minimum can be obtained analytically. Specifically, for the Normal distribution, $q(\mathbf{x}_t|\theta) = \mathcal{N}(\mu_\theta, \Sigma_\theta)$:

$$
\hat{\mu}_\theta = \sum_i w_i \mathbf{x}_t^{(i)}, \quad \hat{\Sigma}_\theta = \sum_i w_i \mathbf{x}_t^{(i)} (\mathbf{x}_t^{(i)})' - \hat{\mu}_\theta \hat{\mu}_\theta'. \tag{36}
$$

Note however, that for extremely low number of effective samples, $n_{eff} = (\sum_i (w_t^{(i)})^2)^{-1}$, this estimate would be misleading since the covariance matrix may not be positive definite.

To derive a more robust solution, we note that the CE method is a special case of the geometric parameter estimation (Kulhavý, 1996). Specifically, it is the maximum likelihood estimate which is sensitive to the lack of data. Therefore, we propose to replace it by the Bayesian version of the geometric estimation, (Kulhavý, 1996, chapter 2):

$$
p(\theta|\mathbf{x}_t, w_t) \propto p(\theta) \exp\left(-n_{eff}\, \mathrm{KL}(p_{emp}(\mathbf{x}_t|\mathbf{y}_{1:t}),\, p(\mathbf{x}_t|\theta))\right), \tag{37}
$$

Here, $p_{emp}(\cdot)$ denotes the approximate posterior (25), and $\mathrm{KL}(\cdot,\cdot)$ is the Kullback-Leibler divergence. Substituting Normal distribution (36) into (37), we obtain a non-standard form of conjugate Bayesian update

of its parameters. Conjugate prior for the Normal likelihood is in the form of Gaussian-inverse-Wishart

$$p(\theta|\mathbf{x}_t, w_t) = \mathcal{N}(\hat{\mathbf{x}}^{[0]}, \gamma^{[0]}\Sigma_\theta)iW(\nu^{[0]}, \Lambda^{[0]}), \tag{38}$$

with statistics $\hat{\mathbf{x}}^{[0]}, \gamma^{[0]}, \nu^{[0]}, \Lambda^{[0]}$. Posterior statistics in the sense of (37) are:

$$\begin{aligned}
\gamma_t^{[m]} &= \frac{\gamma^{[0]}}{1+\gamma^{[0]}n_{eff}}, \\
\hat{\mathbf{x}}_t^{[m]} &= \hat{\mathbf{x}}^{[0]} + \gamma_t n_{eff}(\hat{\mu}_\theta - \hat{\mathbf{x}}^{[0]}), \\
\nu_t^{[m]} &= \nu^{[0]} + n_{eff}, \\
\Lambda_t^{[m]} &= \Lambda^{[0]} + \left[n_{eff}(\hat{\Sigma}_\theta + \hat{\mu}_\theta\hat{\mu}_\theta' - \hat{\mathbf{x}}_t\hat{\mathbf{x}}_t') + \frac{1}{\gamma^{[0]}}\left(\hat{\mathbf{x}}^{[0]}(\hat{\mathbf{x}}^{[0]})' - \hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)\right].
\end{aligned} \tag{39}$$

Note that (39) can approach (36) arbitrarily close for a very flat prior. However, informative prior regularizes the parameter estimates in cases with very small n_{eff}. For this application, we propose to set the prior statistics to match the moments of the unweighted samples from the first stage. This uniquely determines statistics $\hat{\mathbf{x}}^{[0]}$ and $\Lambda^{[0]}$, while statistics $\gamma^{[0]}$ and $\nu^{[0]}$ need to be chosen. The full algorithm for population Monte Carlo setup of M populations is in Algorithm 1.

5 Results

5.1 Simulation Setup

The simulated accident was a release of radionuclide ^{41}Ar with a decay half-life of 109.34 minutes. Radionuclide ^{41}Ar was chosen because it allows us to use simplified formula (16) for evaluation of the dose. Bayesian filtering is performed with a sampling period of 10 minutes, matching 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.

Observations were sampled from distributions described in Section 3.4 with mean values evaluated from the dispersion model. Integration (15) must be performed in both time and space domains. Integration in time is done in sub-steps $\Delta t/5$. Spatial integration in (16) is approximated using Gauss quadrature rules (Golub and Welsch, 1969), and the "n/μ" method (Pecha and Hofman, 2011) with $n = 15$. The same evaluation routines are used for the particles.

5.2 Model Calibration

Documentation provided by manufactures of the radiation dose sensors was used to establish their accuracy in terms of parameters introduced in Section 3.4. In our case, the parameters were $\gamma_y = 0.2$ for the dose monitoring stations, $\gamma_a = 0.1$ and $\sigma_\phi = 5$ degrees for the anemometer.

Parameters of the transition model from Section 3.3 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 $\gamma_a = 0.2$

Algorithm 1 Population Monte Carlo estimation for the continuous monitoring system

Initialization: sample state variable $\mathbf{x}_t$ from prior densities, $p(\mathbf{x}_t)$. Select the number of populations M and the number of particles in them, e.g. $n^{[m]} = \frac{N}{M}$.

At each time t do:

1. Collect measurements $\mathbf{y}_t$,
2. For each particle, $i = 1, \dots n^{[1]}$ 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 (12),(13) and (3).
   (c) Evaluate radiation dose coefficients $c_{i,t}$ for each sensor using (15).
   (d) Update shaping parameters $\mu_{Q,t}, \sigma_{Q,t}$ of proposal density (34) and sample new values $Q_t^{(i)}$.
3. Evaluate statistics $\hat{\mu}_t, \hat{\Sigma}_t$ (36) for $n^{[1]}$ particles and set $\hat{\mathbf{x}}^{[0]} = \hat{\mu}_t, \Lambda^{[0]} = \nu^{[0]}\hat{\Sigma}_t$.
4. For each population $m = 1, \dots, M-1$ do:
   (a) Evaluate weights $w_t^{(i)}$ (26). **PMC & mAMIS**: only for the last population using proposal (35). **AMIS**: for all previous samples using (30). Compute n_{eff} .
   (b) Evaluate parameters $\hat{\mathbf{x}}_t^{[m]}, \gamma_t^{[m]}, \nu_t^{[m]}, \Lambda_t^{[m]}$ using weights from step 4(a) via (39) and (36), and recompute $\hat{\mu}_\theta^{[m]}, \hat{\Sigma}_\theta^{[m]} = \frac{1}{\nu_t^{[m]}}\Lambda_t^{[m]}$.
   (c) Sample particles in the $n^{[m+1]}$ population from (35) with $\hat{\mu}_\theta^{[m]}, \hat{\Sigma}_\theta^{[m]}$.
5. Evaluate weights $w_t^{(i)}$ (26). **PMC:** for i from the last population. **AMIS & mAMIS**: for all samples using (30).

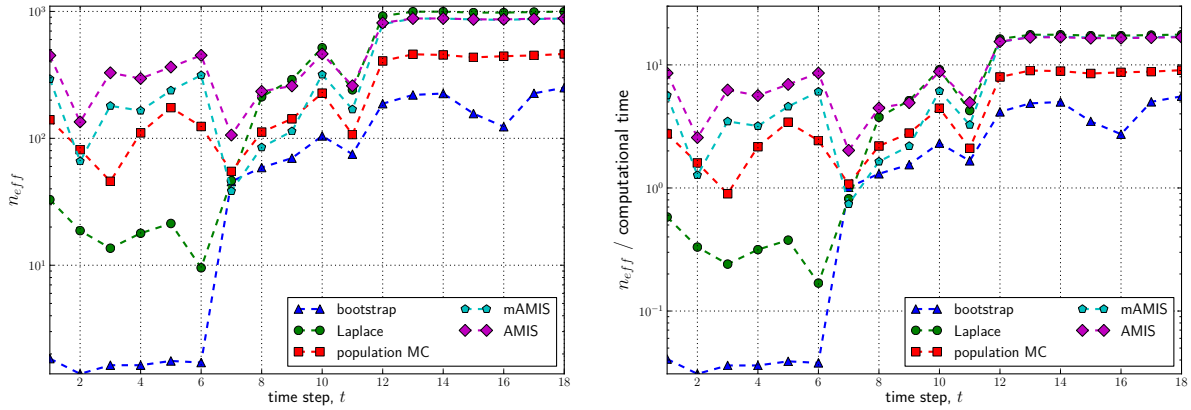

Figure 5: Release with constant release rate, known start and duration. **Left**: Average number of effective particles over 100 Monte Carlo trials. **Right**: Average number of effective particles per one second of execution time over 100 Monte Carlo trials.

and $\sigma_b = 15$ from a continuous windows of historical data of 1000 samples. This estimation procedure can be repeated for new observations. The numerical weather predictions were provided by the MEDARD system (Eben et al., 2005) with grid resolution 9km. The category of Pasquill's atmospheric stability was D.

5.3 Idealistic Scenario with Known Release Time and Duration

As noted in Section 3.2, validity of the random walk transition model (11) for the released activity is questionable. To verify if the data are sufficiently informative to justify the use of the temporally uncorrelated prior (10), we simulated a release with constant release rate Q_t from time $t = 1$, to $t = 6$, with $Q_{1:6} = [1, 1, 1, 1, 1, 1] \times 10^{16} Bq$. This scenario allows for comparison with previous approaches that were using the bootstrap proposal, e.g. (Hiemstra et al., 2011).

Results of estimation of the SMC with $N = 1000$ particles for various proposal densities are displayed in Figure 5 in the form of boxplots of Q_t together with the corresponding number of effective particles. The following proposal densities are compared: **bootstrap**, particles are drawn from the transitional density (21); **Laplace**, where all N particles are drawn from the first stage, i.e. $N = n^{[1]}$; **PMC**, with 100 particles in the first stage, 4 populations of 100 particles to adapt the parametric form, and finally 500 particles from the last population; **AMIS**, with 100 particles in the first stage and 9 adaptive populations of 100 particles; and **mAMIS**, with the same setup as AMIS. The bootstrap filter was run with the transitional density (11). All other filters used the temporally uncorrelated prior (10) with uninformative parameters, $\alpha_Q = 1$, $\beta_Q = 0$. The use of the transition proposal in the population-based methods yields results overlapping with those with the temporally uncorrelated prior (not shown in Figure 5 for clarity).

All methods have nearly identical execution time, since the most expensive operation is evaluation of the moments of the likelihood (20) via (15)–(17) which is done N times for all methods. All methods were

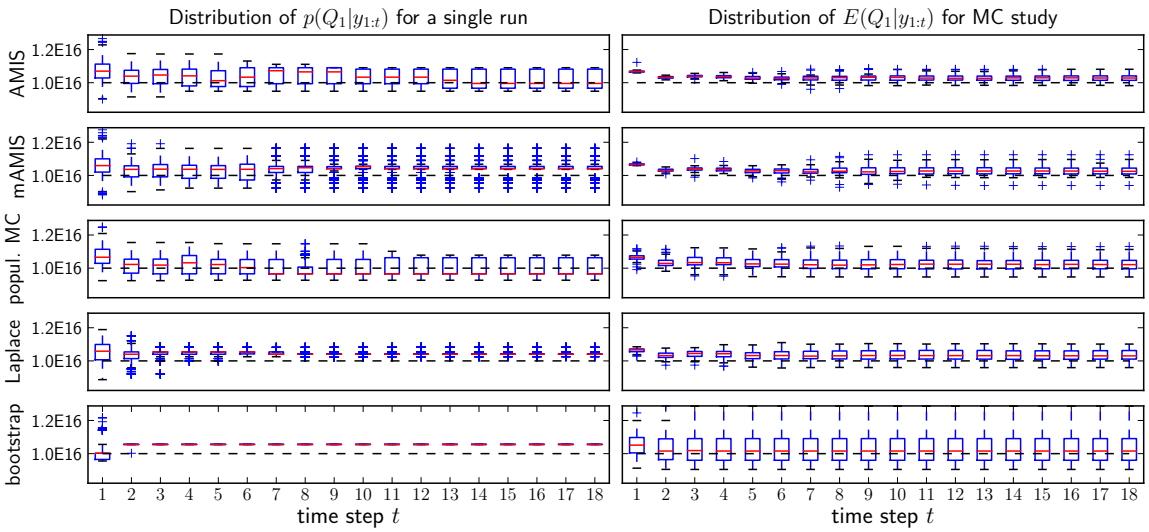

Figure 6: Release with constant release rate, known start and duration. **Left**: Estimates of the released activity of the first puff, Q_1 at times $t = 1, \ldots, 18$, from a typical run of the SMC estimation. **Right**: Distribution of the expected value of Q_1 at times $t = 1, \ldots, 18$ for 100 Monte Carlo trial runs of the SMC estimation.

implemented in the C language and their execution time is measured. The number of efficient samples per second of execution time is displayed in Figure 5 right.

Note that the most demanding part of estimation is for $t = 1, \ldots, 6$ in which the new puffs are being released. In this part, the bootstrap filter often degenerates to $n_{eff} = 1$, in spite of the fact that the initial density (10) had its mean value at the simulated value 10^{16} Bq and low dispersion of 10%. Even for the bootstrap proposal, the estimated parameters are close to the true value, Figure 6 right, hence the accuracy of the point estimate may be sufficient for some tasks. However, the lack of reliable uncertainty bounds makes the method similar to the optimization methods mentioned in the introduction. The Laplace method yields a more efficient proposal, however, the estimated n_{eff} can be also very low, as in time $t = 6$, Fig 5, left. This is not surprising, since the Laplace's approximation is not recommended to use if the number of observations is low (Kass and Raftery, 1995). After $t = 6$, the puffs are no longer generated and then $Q_t = 0, \forall t > 6$. In times $t = 7, \ldots 11$, the puffs in the air are still close to the radiation sensors and thus provide additional information about wind speed and direction. After $t = 11$ the radiation monitoring network registers only measurements of the natural radiation background and the proposals (31)–(32) for the wind speed and direction becomes optimal (28).

Accuracy of the estimates provided by all methods is visualized in Figure 6 via boxplots of estimation results for Q_1 for all studied methods. The best results are for the AMIS method, since the variance within a single run is preserved, and the method yields consistent results in the Monte Carlo study. This suggests that the AMIS method does not have a systematic bias, or it is negligible in this application.

Another important conclusion from this experiment is that the ring of sensors near to the power plant is sufficiently informative about the release activity and it is sufficient to consider uninformative and uncorrelated prior (10). This model can be reliably estimated by the proposed adaptive proposals, where both components are important: the Laplace's approximation is necessary to create the initial estimate, and the population strategies allows further improvement even in situations with very low n_{eff} as demonstrated at $t = 6$.

5.4 Continuous Monitoring

Test of continuous operation of the SMC was performed in a simulation of a sudden release of the pollutant. The simulated release started at $t = 12$ and ended at $t = 18$ with $Q_{12:17} = [1, 5, 4, 3, 2, 1] \times 10^{16}\ Bq$. The initial samples represent the period of normal operation, which is supposed to be arbitrarily long. Results of continuous estimation are displayed in Figure 7 for four selected time steps, $t = 12, 15, 18, 21$, corresponding to the exact moment of the release, and $30, 60, 90$ minutes after that. At each time step, a window of 12 puffs is estimated. During normal operation, the estimated activity fluctuates around $Q_t \approx 1e10$, because ground-level dose from such a puff is much lower than the radiation background, as demonstrated on the corresponding contour plot, Figure 7. Inaccurate estimation of the dose in this regime is thus insignificant for the main purpose of the system.

The abrupt release of activity is immediately recognized, and its activity correctly tracked. When the release is over, the estimated released activity returns immediately to the values that were estimated before the release. This confirms informativeness of data from the first ring of the RMN.

The posterior densities (25) encode useful information which needs to be presented to the decision makers. One of the most valuable supporting material from the human point of view are the contour plots of the total committed dose (Bartzis et al., 2000), which is a superposition of contours from Figure 7. However, in the probabilistic formulation, the total committed dose has its own posterior density, allowing to draw contours of its quantiles, or a histogram of the quantity of interest in the points of interest, see Figure 8. An important conclusion is that the radiological quantities simulated from the true parameters are well within the highest posterior density regions and the procedure does not underestimate their uncertainty.

6 Conclusion

We studied issues related to the design of a fully automated system of continuous monitoring of a radiation situation that would provide guidance for human decision makers during critical situations. Based on previous work, we propose a stochastic model of the accident and a measurement model that suits the existing

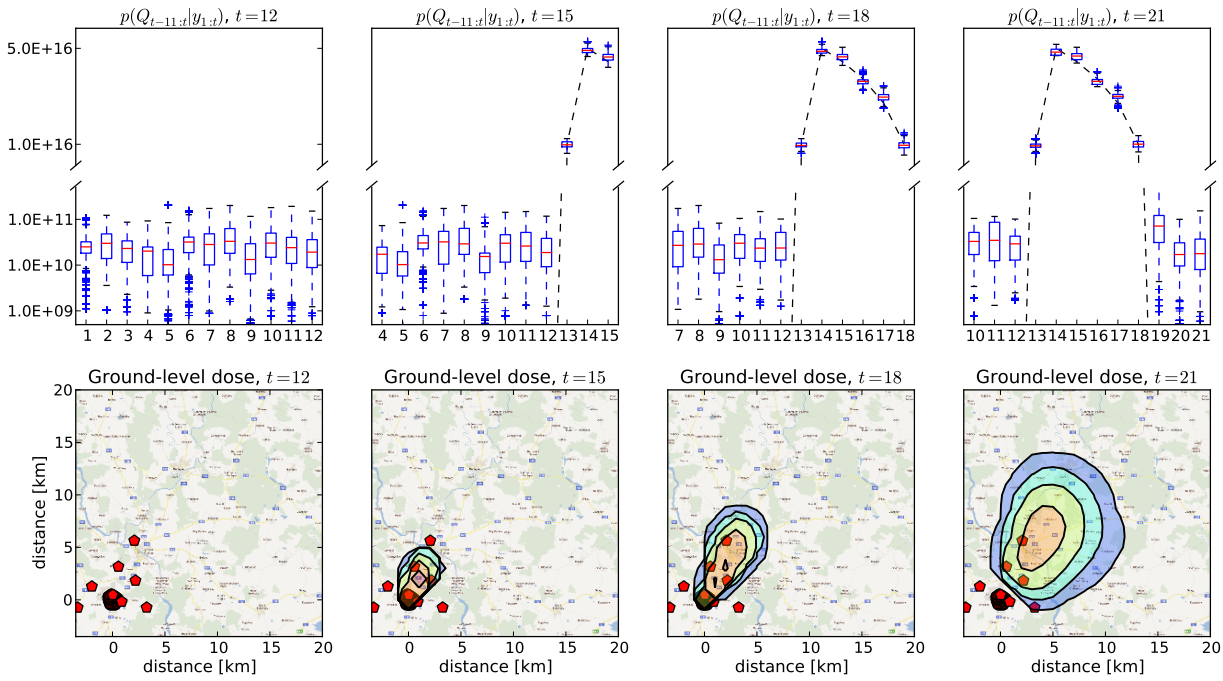

Figure 7: Continuous monitoring of a sudden release of activity. **Top row**: Boxplot of estimated activity of the last twelve (hypothetically) released puffs, and its comparison to the simulated values (dashed line). **Bottom row**: Contour plots of the mean value of the ground level gamma dose at each point on the map. Clear space corresponds to the background level, solid lines in the contours correspond to levels $[1e-6, 1e-5, 1e-4, 1e-3, 1e-2]$ Sv/h, respectively.

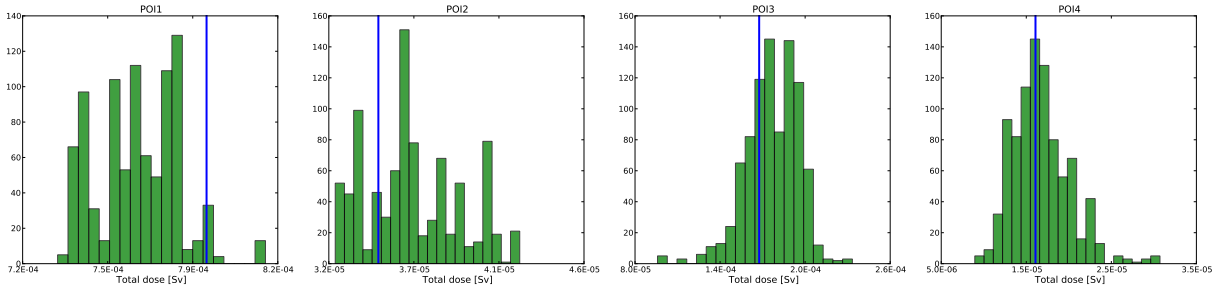

Figure 8: Histograms of the total committed dose from a radiation accident in selected points of interest (locations of the points of interest (POI) are displayed in Figure 1). Vertical line denotes the simulated value.

radiation monitoring networks. The model is highly non-linear and its evaluation requires a computationally demanding numerical procedure. The main task was to sequentially estimate the state of the model including its uncertainty bounds.

Bayesian methods and Monte Carlo techniques in particular have been shown to be effective tools for estimation of such models. They have already been used for similar tasks using the bootstrap approach. We have shown that the bootstrap filter is extremely inefficient in the considered setup and may significantly underestimate the uncertainty of the estimate. We propose a two stage strategy of design of a new proposal density based on combination of ideas of local linearization, population Monte Carlo and information geometry parameter estimation. The new proposal density significantly improved efficiency of the filter that allows to evaluate much more informative estimates in the same computational time. The new proposal also allowed to use a non-informative prior on the released activity, which simplifies tracking of abrupt changes of the release rate.

With the proposed scheme of adaptation of the proposal density, the estimation can be run with only 1000 particles which is affordable to compute in real time on an inexpensive hardware. Moreover, efficient proposal density allows to estimate more parameters of interest, e.g. the dispersion coefficients or Pasquill's category of atmospheric stability, in the future.

A Laplace Approximation of Shifted Inverse Gamma Likelihood and Gamma Prior

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

$$\begin{aligned} p(Q|y_t) &\propto p(Q)p(y_t|Q) \\ &\propto \beta_Q^{\alpha_Q} \frac{Q^{\alpha_Q - 1}}{\Gamma(\alpha)} \exp\left(-\beta_Q Q\right) \prod_{j=1}^{m} \frac{(\beta_j Q + m_j)^{\alpha_j}}{\Gamma(\alpha_j)} y_t^{-\alpha_j - 1} \exp\left(-\frac{\beta_j Q + m_j}{y_t}\right). \end{aligned} \tag{40}$$

which is not of any standard form and the normalization can not be obtained analytically. Following the Laplace's approximation, (Kass and Raftery, 1995), use local linearization via the Taylor expansion at the point of maxima of (40):

$$\hat{Q} = \arg\max_Q(\log(p(Q|y_t))). \tag{41}$$

Taking the first derivative of logarithm of (40)

$$\begin{aligned} \frac{d\log(p(Q|y_t))}{dQ} &= \frac{d}{dQ}\left(\sum_{j=1}^{m}\left[\alpha_j \log(\beta_j Q + m_j) - \frac{\beta_j Q + m_j}{y_t}\right] + (\alpha_Q - 1)\log Q - \beta_Q Q\right) \\ &= \sum_{j=1}^{m}\left[\frac{\alpha_j \beta_j}{\beta_j Q + m_j} - \frac{\beta_j}{y_t}\right] + \frac{\alpha_Q - 1}{Q} - \beta_Q, \end{aligned} \tag{42}$$

we note that it is a sum of rational functions in Q and thus it is strictly decreasing. Therefore, it has only a single intersection with zero at $\hat{Q}$, 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 (42) at point $\hat{Q}$ and set it equal to inverse covariance matrix of the Normal density:

$$\frac{1}{2}\sigma_Q^{-2} = \sum_{j=1}^{m}\left[\frac{\alpha_j\beta_j^2}{(\beta_j\hat{Q}+m_j)^2}\right] + \frac{\alpha_Q - 1}{\hat{Q}^2}, \tag{43}$$

to form the approximate posterior with its mean at the point of maxima, $\mu_Q = \hat{Q}$,

$$p(Q|y_t) \approx \mathcal{N}(\mu_Q, \sigma_Q^2). \tag{44}$$

References

Anderson, J., Anderson, S., Metron, I., and Reston, V. (1999), "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Review*, 127, 12.

Bartzis, J., Ehrhardt, J., French, S., Lochard, J., Morrey, M., Papamichail, K., Sinkko, K., and Sohier, A. (2000), "RODOS: decision support for nuclear emergencies," *Recent Developments and Applications in Decision Making, Kluwer Academic Publishers, Dordrecht, The Netherlands*, 379–394.

Cappé, O., Guillin, A., Marin, J., and Robert, C. (2004), "Population monte carlo," *Journal of Computational and Graphical Statistics*, 13, 907–929.

Cervone, G., Franzese, P., and Grajdeanu, A. (2010), "Characterization of atmospheric contaminant sources using adaptive evolutionary algorithms," *Atmospheric Environment*, 44, 3787–3796.

Cornebise, J., Moulines, E., and Olsson, J. (2008), "Adaptive methods for sequential importance sampling with application to state space models," *Statistics and Computing*, 18, 461–480.

Cornuet, J. M., Marin, J. M., Mira, A., and Robert, C. P. (2012), "Adaptive Multiple Importance Sampling," *Scandinavian Journal of Statistics*, 39, 798–812.

Delle Monache, L., Lundquist, J., Kosovic, B., Johannesson, G., Dyer, K., Aines, R., Chow, F., Belles, R., Hanley, W., Larsen, S., et al. (2008), "Bayesian inference and Markov Chain Monte Carlo sampling to reconstruct a contaminant source on a continental scale," *Journal of Applied Meteorology and Climatology*, 47, 2600–2613.

Doucet, A., de Freitas, N., and Gordon, N. (eds.) (2001), *Sequential Monte Carlo Methods in Practice*, Springer.

Doucet, A., Godsill, S., and Andrieu, C. (2000), "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and computing*, 10, 197–208.

Dowson, D. and Wragg, A. (1973), "Maximum-entropy distributions having prescribed first and second moments," *Information Theory, IEEE Transactions on*, 19, 689 – 693.

Eben, K., Juruš, P., Resler, J., Belda, M., and Kruger, B. (2005), "Medard - An Environmental Modelling Project for the Territory of the Czech Republic," *ERCIM News*, 18–19.

Eleveld, H., Kok, Y., and Twenhofel, C. (2007), "Data assimilation, sensitivity and uncertainty analyses in the Dutch nuclear emergency management system: a pilot study," *International Journal of Emergency Management*, 4, 551–563.

Gifford, F. (1976), "Turbulent diffusion-typing schemes: A review," *Nuclear Safety*, 17, 68–86.

Golub, G. and Welsch, J. (1969), "Calculation of Gauss quadrature rules," *Math. Comp.*, 221–230.

Gordon, N., Salmond, D., and Smith, A. (1993), "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEE, vol. 140, pp. 107–113.

Hanna, S., Briggs, G., and Hosker Jr, R. (1982), "Handbook on atmospheric diffusion," Tech. rep., National Oceanic and Atmospheric Administration, Oak Ridge, TN (USA). Atmospheric Turbulence and Diffusion Lab.

Haupt, S., Beyer-Lout, A., Long, K., and Young, G. (2009), "Assimilating concentration observations for transport and dispersion modeling in a meandering wind field," *Atmospheric Environment*, 43, 1329–1338.

Hiemstra, P., Karssenberg, D., and van Dijk, A. (2011), "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*, 6149–6157.

Holmes, N. and Morawska, L. (2006), "A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available," *Atmospheric Environment*, 40, 5902–5928.

Jeffreys, H. (1961), *Theory of Probability*, Oxford University Press, 3rd ed.

Jeong, H., Kim, E., Suh, K., Hwang, W., Han, M., and Lee, H. (2005), "Determination of the source rate released into the environment from a nuclear power plant," *Radiation protection dosimetry*, 113, 308.

Johannesson, G., Hanley, B., and Nitao, J. (2004), "Dynamic Bayesian Models via Monte Carlo–An Introduction with Examples," Tech. rep., Lawrence Livermore National Laboratory.

Kass, R. E. and Raftery, A. E. (1995), "Bayes factors," *J. of American Statistical Association*, 90, 773–795.

Katata, G., Ota, M., Terada, H., Chino, M., and Nagai, H. (2012), "Atmospheric discharge and dispersion of radionuclides during the Fukushima Dai-ichi Nuclear Power Plant accident. Part I: Source term estimation and local-scale atmospheric dispersion in early phase of the accident," *Journal of Environmental Radioactivity*, 109, 103–113.

Kennedy, M. and O'Hagan, A. (2001), "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 425–464.

Kovalets, I., Tsiouri, V., Andronopoulos, S., and Bartzis, J. (2009), "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, 3511–3521.

Kulhavý, R. (1996), *Recursive nonlinear estimation: a geometric approach*, Springer.

Marin, J., Pudlo, P., and Sedki, M. (2012), "Consistency of the Adaptive Multiple Importance Sampling," *arXiv preprint arXiv:1211.2548*.

Monbet, V., Ailliot, P., and Prevosto, M. (2007), "Survey of stochastic models for wind and sea state time series," *Probabilistic engineering mechanics*, 22, 113–126.

Oh, M. and Berger, J. (1992), "Adaptive importance sampling in Monte Carlo integration," *Journal of Statistical Computation and Simulation*, 41, 143–168.

Pecha, P. and Hofman, R. (2011), "Construction of observational operator for cloudshine dose from radioactive cloud drifting over the terrain," in *Proc. of 14th International Conference on Hamonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*.

Peterka, V. (1981), "Bayesian Approach to System Identification," in *Trends and Progress in System identification*, ed. Eykhoff, P., Oxford: Pergamon Press, pp. 239–304.

Pitt, M. and Shephard, N. (1999), "Filtering via simulation: Auxiliary particle filters," *J. of the American Statistical Association*, 94, 590–599.

Raskob, W., Hugon, M., Lahtinen, J., Aage, H., Ammann, M., Dyve, J., Hoe, S., Rojas-Palma, C., and Wirth, E. (2010), "Guidance on monitoring and data assimilation," *Radioprotection*, 45, 161–169.

Raza, S., Avila, R., and Cervantes, J. (2001), "A 3-D Lagrangian (Monte Carlo) method for direct plume gamma dose rate calculations," *Journal of Nuclear Science and Technology*, 38, 254–260.

Rubinstein, R. and Kroese, D. (2004), *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning*, Springer Verlag.

Senocak, I., Hengartner, N., Short, M., and Daniel, W. (2008), "Stochastic event reconstruction of atmospheric contaminant dispersion using Bayesian inference," *Atmospheric Environment*, 42, 7718–7727.

Thompson, I., Andersen, C., Bøtter-Jensen, L., Funck, E., Neumaier, S., and Sáez-Vergara, J. (2000), "An international intercomparison of national network systems used to provide early warning of a nuclear accident having transboundary implications," *Radiation protection dosimetry*, 92, 89.

Twenhöfel, C., van Troost, M., and Bader, S. (2007), "Uncertainty analysis and parameter optimization in early phase nuclear emergency management," .

Šmídl, V. and Hofman, R. (2013), "Tracking of atmospheric release of pollution using unmanned aerial vehicles," *Atmospheric Environment*, 67, 425 – 436.

Winiarek, V., Vira, J., Bocquet, M., Sofiev, M., and Saunier, O. (2010), "Towards the operational estimation of a radiological plume using data assimilation after a radiological accidental atmospheric release," *Atmospheric Environment*.

Zannetti, P. (1990), *Air pollution modeling*, Van Nostrand Reinhold.