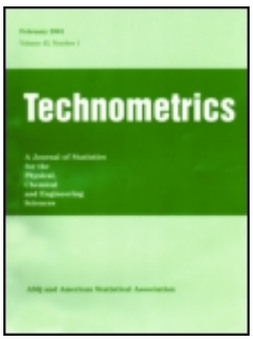

Technometrics

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: http://www.tandfonline.com/loi/utch20

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

Václav Šmídl & Radek Hofman

To cite this article: Václav Šmídl & Radek Hofman (2014) Efficient Sequential Monte Carlo Sampling for Continuous Monitoring of a Radiation Situation, Technometrics, 56:4, 514-528, DOI: 10.1080/00401706.2013.860917

To link to this article: https://doi.org/10.1080/00401706.2013.860917

Supplementary materials for this article are available online. Please go to http://www.tandfonline.com/r/TECH

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

Václav ŠMÍDL and Radek HOFMAN

Institute of Information Theory and Automation
CZ-182 08 Prague, Czech Republic
(*smidl@utia.cas.cz*; *hofman@utia.cas.cz*)

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 limited availability of measurements from a sparse monitoring network. Assessment of the situation requires experienced specialists who may be unavailable during critical times. Our goal is to provide an automated method of instant radiation situation assessment that does not underestimate its uncertainty. We propose a state-space model based on an 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. The sequential Monte Carlo method is one of the few options for estimating 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. The proposed tailor-made proposal was found to be much more computationally efficient than previously published methods. The adaptive Monte Carlo methods thus represent a compelling computational approach for the evaluation of probabilistic environmental models. The data used and a Python implementation of the methods are available as supplementary material online.

KEY WORDS: Atmospheric dispersion model; Importance sampling; Information geometry estimation; Population Monte Carlo; Proposal density; Radiation protection.

1. INTRODUCTION

We are concerned with a scenario in which the release of radionuclides into the atmosphere occurs following a hypothetical accident at a nuclear power plant facility. The radioactive effluent forms a plume that moves over the terrain according to the current meteorological situation. Protective countermeasures 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 severity categories. 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 approaches (Jeong et al. 2005; Kovalets et al. 2009), and genetic algorithms (Haupt et al. 2009; Cervone, Franzese, and Grajdeanu 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, van Troost, and Bader (2007).

An alternative approach is statistical estimation, for example, Anderson et al. (1999). Indeed, the value of statistical techniques such as the Markov chain Monte Carlo (Delle Monache et al. 2008; Senocak et al. 2008) and sequential Monte Carlo (Johannesson, Hanley, and Nitao 2004; Hiemstra, Karssenberg, and van Dijk 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, Hanley, and Nitao (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 from a selected power plant. We show that estimation of the new model from data pro-

TECHNOMETRICS, NOVEMBER 2014, VOL. 56, NO. 4
DOI: 10.1080/00401706.2013.860917
Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/tech*.

vided by a sparse RMN via the classical particle filter (Gordon, Salmond, and Smith 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, for example (Pitt and Shephard 1999), other techniques are based on adaptation of nonsequential techniques for the sequential scenario (Cornebise, Moulines, and Olsson 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. Due to flexibility of the importance sampling methodology, these techniques can be freely combined to take advantage of properties of the application specific model. Specifically, we design an efficient problem-specific proposal density by combining two well-known general-purpose approaches (Pitt and Shephard 1999; Cornuet et al. 2012) with lesser known concepts of information geometry (Kulhavý 1996). The proposed combined proposal is significantly more efficient on the studied application. Therefore, only 1000 particles are sufficient for reliable estimation of the state variables. This result demonstrates that computational overhead of modern adaptive Monte Carlo methods over simpler alternatives is much smaller than is generally perceived.

2. RADIATION ACCIDENT

Awareness of radiation security has increased after the Chernobyl disaster, and every nuclear power plant is now surrounded by a network of radiation sensors. 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. First, the spread of the pollutant in the atmosphere is a complex stochastic process with many unknowns. Second 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: (1) advection by wind, (2) dispersion by turbulent processes in the atmosphere, and (3) 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, Equation (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 article, 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 Equation (1) for constants u_k and K_k, and the following boundary conditions (Hanna, Briggs, and Hosker 1982): (1) the concentration is approaching zero with increasing distance for the source, (2) the release source is infinitely small at $t = 0$, (3) 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 step. 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

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

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, that is, $\kappa = t - 1$. Time evolution of the location of the κth puff center is fully determined by the wind field at the previous location

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

where $v_t(\mathbf{l}_{\kappa,t})$ and $\phi_t(\mathbf{l}_{\kappa,t})$ are the wind speed and the 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 parameterized using dispersion coefficients $\sigma_k = \sqrt{2K_k t}$, (Gifford 1976).

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): (1) *prerelease phase*: an abnormal situation happened in the power plant, however, the radiation has not leaked outside of the reactor building yet, (2) *early phase*: the radiation material is being released from

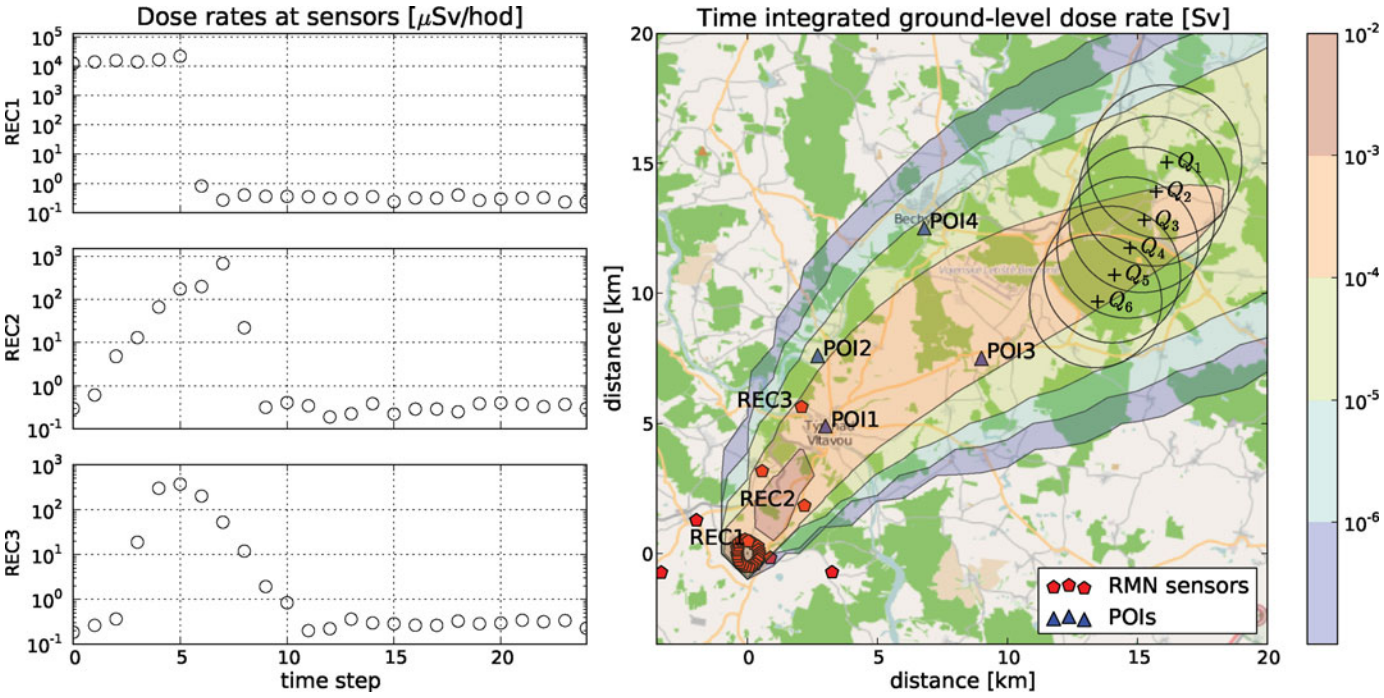

Figure 1. Simulation experiment of an accidental radiation release. Left: simulated measurements of three RMN sensors with the highest observed values. Right: contour plot of the total ground-level radiation dose accumulated during the first 3 hours after the release; locations of the radiation dose sensors are denoted by pentagons; largest populated areas (points of interest, POI) are denoted by triangles, and crosses denote centers of the puffs and circles are drawn around them with diameter $2\sigma_1$ (Equation (2)). The illustrative background map is from the *http://www.openstreetmaps.org*, © OpenStreetMap contributors.

the power plant into the atmosphere and the radioactive cloud is still within the monitored area, and (3) *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 enables a deduction of what has happened and what will be the final consequences. For example, aerial surveillance provides measurements of radionuclide deposition on the ground with high spatial resolution, and simple interpolation can provide a clear contour of the total dose. However, this analysis usually comes too late to be of use to citizens living near the power plant. Since there is no time to add sensors before the early phase starts, we must limit ourselves to 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 (right). The amount of observed data provided by the RMN during a release is very small, for example, less than 100 scalar values in the release displayed in Figure 1. Our aim is to process these data recursively without the need of human interaction. This objective represents many challenges for modeling and computational evaluation of the results.

First, meteorological parameters would only be available in the form of numerical predictions, since more accurate weather estimates using, for example, satellite data are available with unacceptable delay. 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 at most eight data points above the radiation background level on any sensor of the RNM. Moreover, more than half of the sensors did not register any increased radiation levels. 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 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 routinely analyze measurements of the natural background radiation.

3. STATE-SPACE MODEL FOR CONTINUOUS RADIATION MONITORING

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

$$\mathbf{x}_t \sim p(\mathbf{x}_t|\mathbf{x}_{t-1}) \tag{4}$$

$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_{t-1}), \tag{5}$$

where $\mathbf{x}_t$ is the state variable, $\mathbf{y}_t$ is the vector of observations. Probability distributions in Equations (4) and (5) are assumed to be known.

In principle, the state variable should contain the concentration, C, which is infinite dimensional since it is a solution of the PDE (1). However, estimation of such system from sparse measurements is difficult. Therefore, we follow Johannesson, Hanley, and Nitao (2004) and approximate the spatio-temporal distribution of the pollutant by a 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, described at the end of Section 2.1).

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 article. Extension of the deterministic model to a full stochastic state-space model thus requires choosing transition probability density (Equation (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, Hanley, and Nitao (2004) was designed for the localization of an unknown source of radiation. Therefore, 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 (Delle Monache et al. 2008; Senocak et al. 2008).

The assumptions mentioned above are only partially relevant for the purpose of the continuous monitoring system described in Section 2.2. The location of the source is well known. Sensitivity studies of the parameters of the puff model (Twenhöfel, van Troost, and Bader 2007) identified the following parameters as the most significant: (1) magnitude of the release at each time step, (2) the wind speed and direction. Models for these two terms have been proposed in Johannesson, Hanley, and Nitao (2004) and Hiemstra, Karssenberg, and van Dijk (2011), respectively. While there are other unknown parameters, we consider only these three here to illustrate the complexity of the problem.

The choice of the probability density of the measurements is relatively easy since it can be deduced from 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

Probability densities of the state transition model (4) and the measurement model (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 manufacturers specify accuracy of the measuring devices only by their relative accuracy in percents of the measured value. This can be interpreted as two statistical moments of the distribution, for example,

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

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.

One way to turn these quantities into a density, $p(y_t)$, is by moment matching. The following parametric forms have the same moments as Equation (6):

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

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

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

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 of these 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 (Equation (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.

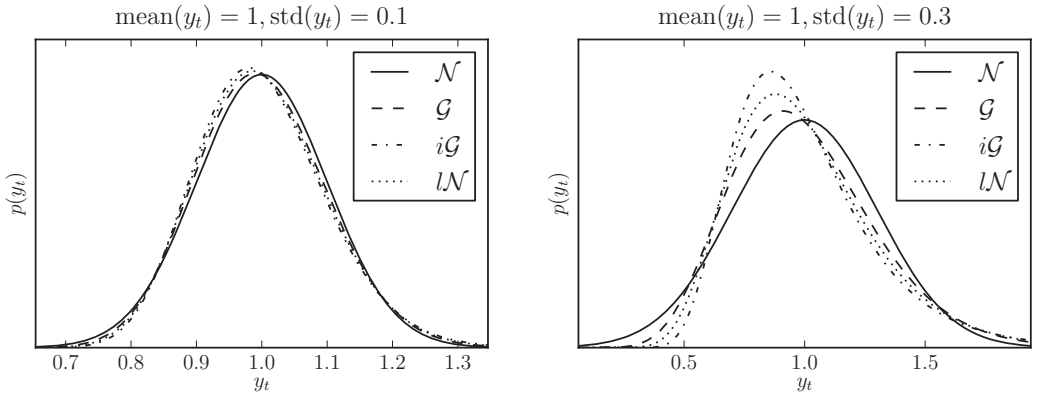

Figure 2. Comparison of different probability densities with equal first two moments. All probability densities have the same mean value $\mu_t = 1$, but different standard deviations $\gamma = 0.1$ (left) and $\gamma = 0.3$ (right).

3.2 Release Scenario

In previous work (Johannesson, Hanley, and Nitao 2004; Hiemstra, Karssenberg, and van Dijk 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 a Gamma density, the model of the first step is

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

followed by random walk

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

where γ_Q governs the spread of the random walk.

However, a 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 (Equations (10) and (11)), instead of their deterministic change after $t = 1$. 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, that is, the prior (Equation (10)) holds for all t. Moreover, we aim for completely uninformative prior, that is, for $\alpha_Q = 1$ and $\beta_Q \to 0$ for which (Equation (10)) approaches the improper Jeffreys' 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 and altitude, stored in three-dimensional vector. From this point, it is advected by the wind field (Equation (3)). While it is possible to obtain numerical weather forecasts from various sources, their accuracy at the power plant location is usually poor. An illustration of this fact is displayed in Figure 3 by comparing 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 parameterization of the calibration. We follow (Hiemstra, Karssenberg, and van Dijk 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 the 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 wind field are available (Monbet, Ailliot, and Prevosto 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

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

Matching of these moments to the densities (Equation (8)) and a truncated version of (Equation (7)) yields

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

The Gamma density was chosen for a_t since its variance is proportional to its mean value and allows for conjugate update. The standard deviation of the wind direction is constant and typically lower than 30°, the appropriate von Mises distribution can be very well approximated by the truncated Normal with support on the unit circle.

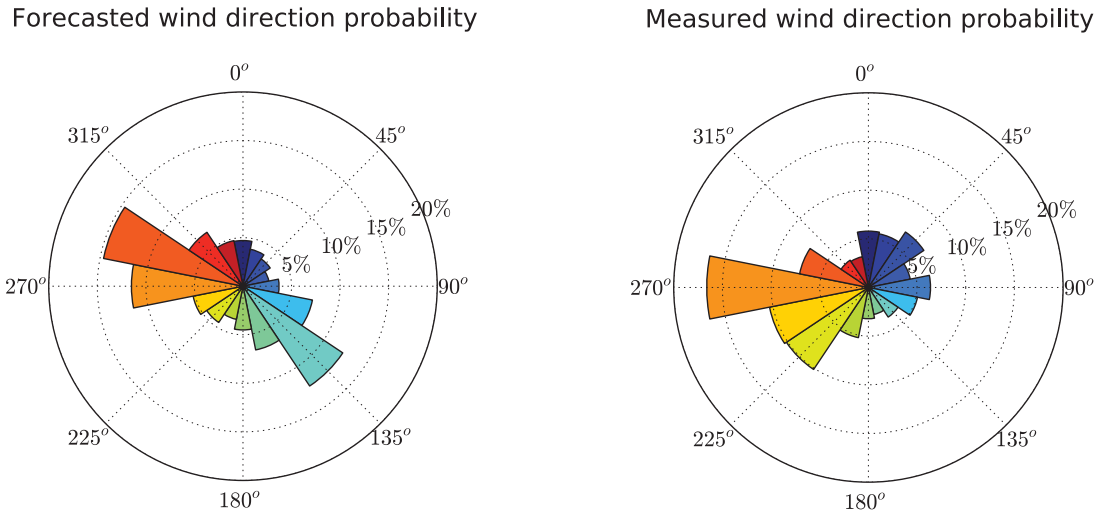

Figure 3. Comparison of histograms of the wind direction at the location of the power plant for year 2008 from the numerical model (left) and observed data (right).

3.4 Measurement Model

Measurement of the local concentration of the pollutant $C(\mathbf{s}, \tau)$ considered in Johannesson, Hanley, and Nitao (2004); Haupt et al. (2009) requires a large and expensive device. A more affordable Geiger counter typically measures a time integral of radiation gamma dose rates which is measured in sieverts per hour [Sv/h]. The RMN 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 the 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 the computational cost of such an operation. Evaluation of the dose rate for a more complex dispersion model would be even more costly.

If a single puff (Equation (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, Avila, and Cervantes 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. The 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 Equation (2); $k = (\mu - \mu_a)\mu^{-1}$ where μ is a linear attenuation coefficient (Raza, Avila, and Cervantes 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, that is, a sequence of all puffs, contributes to the jth sensor by dose

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

Transformation of the activity concentration $C(\mathbf{s},\ \tau)$ into the time integrated dose rate $y_{Q,j,t}$ via Equations (15)–(17) is in the simplest possible form and it is still highly nonlinear. An analytical solution is not available and the dose (Equation (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 (Equation (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

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

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

$$p(y_{j,t}|\mathbf{x}_{1:t}) = i\mathcal{G}\big(\gamma_y^{-2} + 2,\ \big(\gamma_y^{-2} + 1\big)(y_{nb,j} + y_{Q,j,t})\big), \tag{19}$$

since it is conjugate with the Gamma density of the released activity Q_t, Equation (10). 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 month of May 2010 in Figure 4 (right). Since the quantiles of the measurements displayed in Figure 4 (left), correspond well to the model (Equation (19)), with $y_{Q,j,t} = 0$, and $\gamma_y = 0.2$, we consider $y_{nb,j}$ to be known

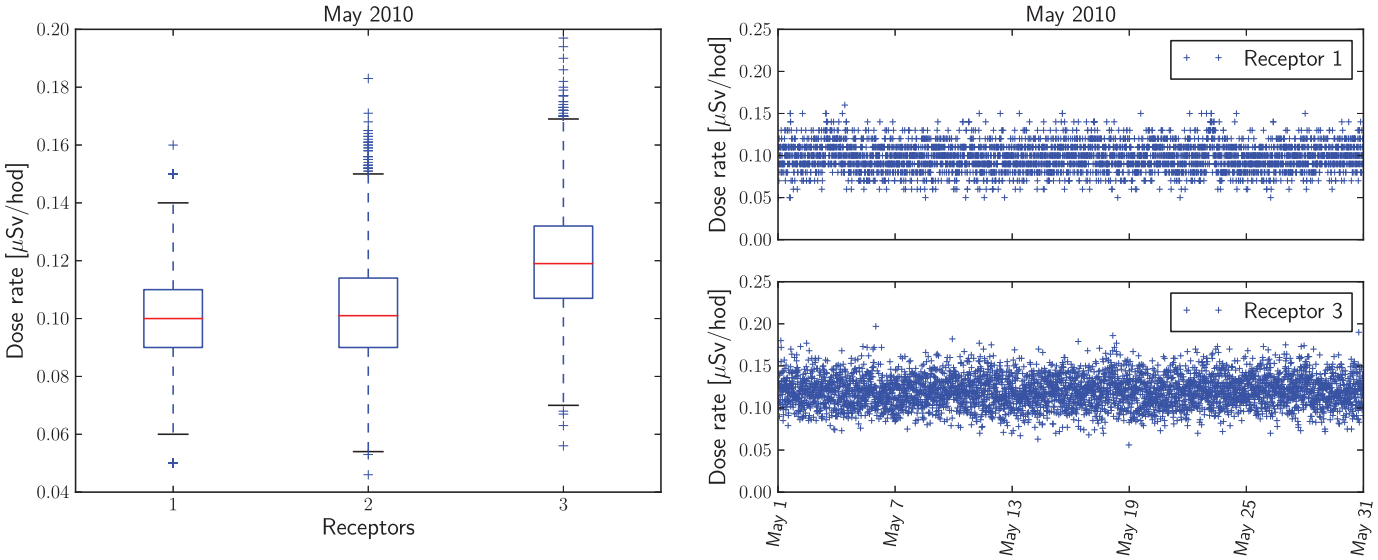

Figure 4. Natural background measured by selected sensors of the radiation network in May 2010 displayed via their quantiles in the form of a boxplot (left) and raw data (right).

and fixed at the average value on the sensor over the previous year, that is, fluctuation of the natural background is considered to be negligible.

Anemometers accuracy is typically also available in terms of relative error of the wind speed, v_t, and constant error of the wind direction, ϕ_t, that is,

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

where $\mathbf{s}_{\text{meteo}}$ is the location of the meteostation. From all choices of possible densities (Equations (7)–(9)) we select the inverse Gamma form (9) for v_t as a conjugate model to (Equation (14)). Probability density of the wind direction ϕ_t is considered as Normal with fixed variance (Equation (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, that is,

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

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

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

The parameter evolution model is then composed of all considered models

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = p(a_t|a_{t-1})p(b_t|b_{t-1})p(Q_t), \tag{20}$$

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

$$p(\mathbf{y}_t|\mathbf{x}_t) = p(v_t, \phi_t|a_t, b_t)\prod_{j=1}^{J} p(y_{Q,j,t}|Q_{t-K:t}, \mathbf{l}_{t-I:t}). \tag{21}$$

This state-space model (Equations (20) and (21)) 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 (Equation (18)), since the majority of the parameters in (Equation (15)) is known for specific radionuclides. A 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 prerelease 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 where the pollutant is captured in the containment for a 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 article, we set these parameters to their typical values for an aggregated parameter called Pasquill's category of stability (Hanna, Briggs, and Hosker 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 (Equation (19)). 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, for example, by unmanned aerial vehicles (Šmídl and Hofman 2013).

4. SEQUENTIAL MONTE CARLO ESTIMATION

The model (Equations (4) and (5)) is fully specified by probability density functions (Equations (20) and (21)). 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):

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}, \tag{22}$$

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

where $p(\mathbf{x}_1|\mathbf{y}_0)$ is the prior density, and $\mathbf{y}_{1:t} = [\mathbf{y}_1, \dots, \mathbf{y}_t]$ denotes the set of all observations. The integration in Equation (23), and elsewhere in this article, is over the whole support of all involved probability density functions.

Equations (22) and (23) are analytically tractable only for a limited set of models. The most notable example is the linear Gaussian model for which Equations (22) and (23) are equivalent to the Kalman filter. For other models, (22)–(23) need to be evaluated approximately. One option is sequential Monte Carlo (e.g., Gordon, Salmond, and Smith 1993; Doucet, de Freitas, and Gordon 2001) which provides an approximation of the posterior density by a weighted empirical density

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^{(i)}\delta\left(\mathbf{x}_{1:t} - \mathbf{x}_{1:t}^{(i)}\right), \tag{24}$$

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 (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 that this approximation can be evaluated for an arbitrary model

(Equations (4) and (5)) given a suitable proposal density, $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$, yielding

$$w_t^{(i)} \propto \frac{p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}. \tag{25}$$

An important property of Equation (25) 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. This is prevented by 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, de Freitas, and Gordon 2001).

4.1 Choice of the Proposal Density

The proposal density is often the determining factor in the computational efficiency of the particle filter and was heavily studied for this purpose. The optimal proposal density is (Doucet, de Freitas, and Gordon 2001)

$$\begin{aligned} q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) &= q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) q(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1}), \\ q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) &= \frac{p(x_t|x_{t-1}) p(y_t|x_t)}{\int p(x_t|x_{t-1}) p(y_t|x_t) dx_t}. \end{aligned} \tag{26}$$

However, evaluation of the integral in Equation (26) is computationally intractable and Equation (26) is helpful only as a theoretical concept. The goal is to approximate Equation (26) as closely as possible, with many approaches for 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, Salmond, and Smith (1993), which is often called the *bootstrap* approximation. The main advantage of this choice is simplicity of the resulting algorithm.
- local linearization of Equation (26) via a Taylor expansion (Pitt and Shephard 1999; Doucet, Godsill, and Andrieu 2000), which is also known as the Laplace 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 classical Monte Carlo methods (Oh and Berger 1992; Rubinstein and Kroese 2004) and has been used in sequential Monte Carlo in Cornebise, Moulines, and Olsson (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, Hanley, and Nitao (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 our 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 proposed model and common conditions of its use. Specifically, the RMNs were designed to measure quantities that are essential in prediction of the consequences of the release. 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. Considering the measured wind direction as the actual one would cause deviations between the predicted and the measured radiation dose on sensors. These deviations would increase with increasing distance of the sensor from the location of the meteostation.

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 (Equation (26)) is then conditionally independent as follows:

$$\begin{aligned} q^{[1]}(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t) = {} & q(a_t|a_{t-1}, v_t) q(b_t|b_{t-1}, \phi_t) \\ & \times q(Q_t|\mathbf{y}_{1:t}, \mathbf{l}_{1:t}) q(\mathbf{l}_t|\mathbf{l}_{t-1}, a_t, b_t). \end{aligned} \tag{27}$$

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 makes use of the latest measurements which is necessary to follow rapid changes in any of the state variable. The second stage relaxes the assumption of conditional independence and provides a proposal based on all measured quantities. The second stage can be repeated several times in the sense of the method of population Monte Carlo (PMC; Cappé et al. 2004). These methods are based on repetitive runs of the importance sampling, each time with a different proposal. Each run of the importance sampling produces a *population* of the particles. The key improvement is in using the statistics of the previous population to adjust parameters of the proposal function for the next run. Formally, the full set of N particles is composed of particles from M populations, each of $n^{[m]}$ particles, $m = 1, \ldots, M$, $\sum_{m=1}^{M} n^{[m]} = N$. The proposal function is in the form $q(\mathbf{x}_t|\theta)$, with parameter θ, and is adapted in each population by replacing parameter θ with its actual estimate, giving $q(\mathbf{x}_t|\hat{\theta}^{[m]})$. Various modifications of the basic method have been proposed; we will study three of them.

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 afterward. 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: the idea of deterministic mixture sampling is used to increase computational efficiency of the scheme (Cornuet et al. 2012). Each parametric proposal generated by the previous

populations is interpreted as a component of a mixture density

$$q_{\mathrm{AMIS}}^{[m]}(\mathbf{x}_t|\theta) = \sum_{k=1}^{m} \frac{n^{[k]}}{\sum_{k=1}^{m} n^{[k]}} q\left(\mathbf{x}_t|\hat{\theta}^{[k]}\right), \tag{28}$$

and the weights (Equation (25)) are reevaluated after each population. The estimates of the parameters $\hat{\theta}^{[m+1]}$ are evaluated using all reweighted samples. However, proof of convergence of this approach 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, Pudlo, and Sedki 2012). It still uses the mixture density (Equation (28)), but the parameter estimates are based only on the samples from the previous population. As a consequence, it is sufficient to recompute the weights using Equation (28) only after the last population.

A problematic part of this approach is the choice of the initial value of the parameters, $\hat{\theta}^{[0]}$. From this point of view, the proposal density (Equation (27)) from the first 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 (Equation (19)) for the likelihood of the measured wind speed, the Gamma transition models (Equation (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 (Equation (14)). The first two factorized densities in Equation (27) are then

$$q(a_t|a_{t-1}, v_t) \propto p(v_t|a_t)p(a_t|a_{t-1}) = \mathcal{G}(k_a, \theta_a), \tag{29}$$
$$q(b_t|b_{t-1}, \phi_t) \propto p(\phi_t|b_t)p(b_t|b_{t-1}) = \mathcal{N}(\mu_b, r_b), \tag{30}$$

with parameters

$$\begin{aligned} k_{a,t} &= \gamma_a^{-2} + \gamma_v^{-2} + 2, \\ \theta_{a,t} &= \left[\gamma_a^{-2} a_{t-1}^{-1} + \tilde{v}_t v_t^{-1}\left(\gamma_v^{-2} + 1\right)\right]^{-1}, \\ \mu_{b,t} &= r_{b,t}\left[\sigma_b^{-2} b_{t-1} + \sigma_\phi^{-2}(\phi_t - \tilde{\phi}_t)\right], \\ r_{b,t} &= \left(\sigma_b^{-2} + \sigma_\phi^{-2}\right)^{-1}. \end{aligned} \tag{31}$$

Derivation of $q(Q_t|y_{1:m}, \mathbf{s}_{1:t})$ is more demanding since the likelihood (Equation (19)) is not conjugate. We propose to follow the Laplace method (Kass and Raftery 1995), based on local linearization via the Taylor expansion. The conditional distribution on the dose Q_t is then

$$q(Q_t|y_{1:m,t}, \mathbf{s}_{1:t}) = t\mathcal{N}\left(\mu_Q, \sigma_Q^2, \langle 0, \infty\rangle\right), \tag{32}$$

where algorithm for evaluation of moments μ_Q, σ_Q is described in the supplementary material available online.

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

$$q(\log a_t, b_t, \log Q_t|\mu_\theta, \Sigma_\theta) = \mathcal{N}(\mu_\theta, \Sigma_\theta), \tag{33}$$

which requires 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 (24) and $q(\mathbf{x}_t|\theta)$. For parametric forms from the exponential family, the minimum can be obtained analytically. Specifically, for the Normal distribution, $q(\mathbf{x}_t|\theta) = \mathcal{N}(\mu_\theta, \Sigma_\theta)$:

$$\hat{\mu}_\theta = \sum_i w_i \mathbf{x}_t^{(i)}, \quad \hat{\Sigma}_\theta = \sum_i w_i \mathbf{x}_t^{(i)}\left(\mathbf{x}_t^{(i)}\right)' - \hat{\mu}_\theta \hat{\mu}_\theta'. \tag{34}$$

Note however, that for an extremely low number of effective samples, $n_{\mathrm{eff}} = (\sum_i (w_t^{(i)})^2)^{-1}$, 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 so-called geometric parameter estimation (Kulhavý 1996). Specifically, Equation (34) is the maximum likelihood estimate which is sensitive to the lack of data. Therefore, we propose to replace it by Bayesian version of geometric estimation (Kulhavý 1996, chap. 2):

$$p(\theta|\mathbf{x}_t, w_t) \propto p(\theta) \exp\left(-n_{\mathrm{eff}}\, \mathrm{KL}(p_{\mathrm{emp}}(\mathbf{x}_t|\mathbf{y}_{1:t}),\ p(\mathbf{x}_t|\theta))\right). \tag{35}$$

Here, $p_{\mathrm{emp}}(\cdot)$ denotes the approximate posterior (24), and $\mathrm{KL}(\cdot,\cdot)$ is the Kullback–Leibler divergence. Substituting a Normal distribution (Equation (34)) into Equation (35), we obtain a nonstandard form of a conjugate Bayesian update of its parameters. The conjugate prior for the Normal likelihood is in the form of Gaussian–inverse-Wishart

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

with statistics $\hat{\mathbf{x}}^{[0]}$, $\gamma^{[0]}$, $\nu^{[0]}$, $\Lambda^{[0]}$. Posterior statistics in the sense of Equation (35) are

$$\begin{aligned} \gamma_t^{[m]} &= \frac{\gamma^{[0]}}{1 + \gamma^{[0]} n_{\mathrm{eff}}}, \\ \hat{\mathbf{x}}_t^{[m]} &= \hat{\mathbf{x}}^{[0]} + \gamma_t n_{\mathrm{eff}}\left(\hat{\mu}_\theta - \hat{\mathbf{x}}^{[0]}\right), \\ \nu_t^{[m]} &= \nu^{[0]} + n_{\mathrm{eff}}, \\ \Lambda_t^{[m]} &= \Lambda^{[0]} + \Big[n_{\mathrm{eff}}(\hat{\Sigma}_\theta + \hat{\mu}_\theta \hat{\mu}_\theta' - \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t') \\ &\quad + \frac{1}{\gamma^{[0]}}\left(\hat{\mathbf{x}}^{[0]}\left(\hat{\mathbf{x}}^{[0]}\right)' - \hat{\mathbf{x}}_t \hat{\mathbf{x}}_t'\right)\Big]. \end{aligned} \tag{36}$$

Note that Equation (36) can approach Equation (34) arbitrarily close by for a very flat prior. However, an 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 min. Radionuclide ^{41}Ar was

Algorithm 1 Population Monte Carlo estimation for the continuousmonitoring 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, for example, $n^{[m]} = \frac{N}{M}$.
At each time t do:

1. Collect measurements $\mathbf{y}_t$.
2. For each particle, $i = 1, \ldots n^{[1]}$ do:
   (a) Update parameters $k_{a,t}, \theta_{a,t}, \mu_{b,t}, r_{b,t}$ using Equation (31) and sample new values of $a_t^{(i)}, b_t^{(i)}$ from Equations (29) and (30).
   (b) Compute new locations of all puff centers $\mathbf{l}_t^{(i)}$ using Equations (12), (13), and (3).
   (c) Evaluate radiation dose coefficients $c_{i,t}$ for each sensor using Equation (15).
   (d) Update parameters $\mu_{Q,t}, \sigma_{Q,t}$ of proposal density (Equation (32)) and sample new values $Q_t^{(i)}$.
3. Evaluate statistics $\hat{\mu}_t, \hat{\Sigma}_t$ (34) for $n^{[1]}$ particles and set $\hat{\mathbf{x}}^{[0]} = \hat{\mu}_t$, $\Lambda^{[0]} = \nu^{[0]}\hat{\Sigma}_t$.
4. For each population $m = 1, \ldots, M-1$ do:
   (a) Evaluate weights $w_t^{(i)}$ (Equation (25)). PMC & mAMIS: only for the last population using proposal (Equation (33)). AMIS: for all previous samples using Equation (28). 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 Equations (36) and (34), 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 Equation (33)with $\hat{\mu}_\theta^{[m]}, \hat{\Sigma}_\theta^{[m]}$.
5. Evaluate weights $w_t^{(i)}$ (Equation (25)). PMC: for i from the last population. AMIS & mAMIS: for all samples using Equation (28).

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 min, matching the sampling period of the RMN which provides measurements of a time integrated dose rate in 10-min 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 substeps $\Delta t/5$. Spatial integration in Equation (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, $\gamma_v = 0.1$ 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 their forecasts 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 $\gamma_a = 0.2$ 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 9 km. 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 (Equation (11)) for the released activity is questionable. To verify if the data are sufficiently informative to justify the use of the temporally uncorrelated prior (Equation (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, Karssenberg, and van Dijk 2011).

Results for 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. Note that the number of efficient particles has significant impact on the spread of the posterior density (Figure 5, left) and consistency of the estimate for repeated runs of the same procedure in a Monte Carlo study (right). The following proposal densities are compared: *bootstrap*, particles are drawn from the transitional density (Equation (20)); *Laplace*, where all N particles are drawn from the first stage, that is, $N = n^{[1]}$; 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 then 9 adaptive populations of 100 particles; and mAMIS, with the same setup as AMIS. The bootstrap filter was run with the transitional density (Equation (11)). All other filters used the temporally uncorrelated prior (Equation (10)) with uninformative parameters, $\alpha_Q = 1$, $\beta_Q = 0$. The use of the transition proposal in the population-based methods yields results nearly identical to those with the temporally uncorrelated prior.

All methods have nearly identical execution time, since the most expensive operation is evaluation of the moments of the likelihood (Equation (19)) via Equations (15)–(17) which is done N times for all methods. All methods were implemented in the C language and their execution time was measured. The number of effective samples per second of execution time is displayed in Figure 6 (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 $n_{\text{eff}} = 1$, 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 5 (right); hence, the accuracy of the point estimate may be sufficient for some tasks. However, from $t = 3$, all particles have the same value and all quantiles of the posterior lie on the same line in Figure 5 (bottom left). This lack of reliable uncertainty bounds makes the method similar to the optimization methods mentioned in the introduction. The

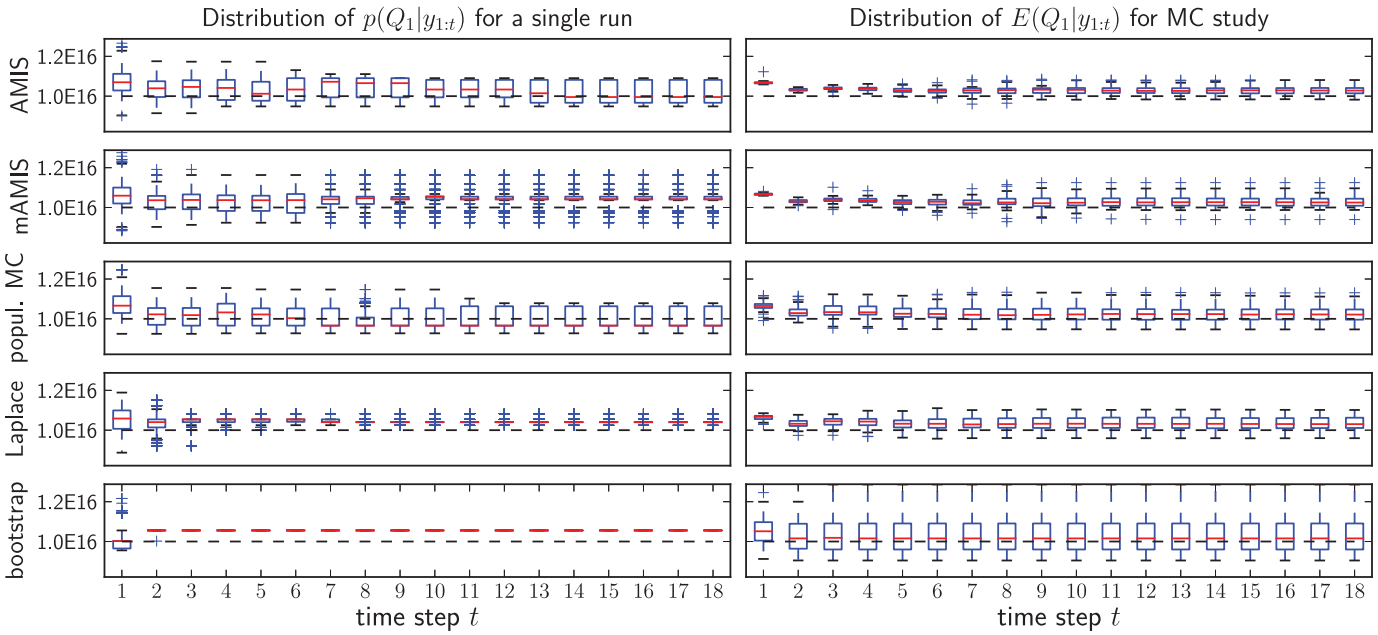

Figure 5. Release with constant release rate, known start and duration. Left: estimates of the released activity of the first puff, Q_1 using data available at times $t = 1, \ldots, 18$, from a typical run of the SMC estimation. Dashed horizontal line denotes simulated value of Q_1. Right: distribution of the expected value of Q_1 at times $t = 1, \ldots, 18$ for 100 Monte Carlo trial runs of the SMC estimation.

Laplace method yields a more efficient proposal, however the estimated n_{eff} can be also very low, as in time $t = 6$ (Figure 6, left). This is not surprising, since the Laplace approximation is not recommended 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 RMN registers only measurements of the natural radiation background and the proposals (Equations (29) and (30)) for the wind speed and the direction becomes optimal (Equation (26)).

Accuracy of the estimates provided by all methods can be visualized in Figure 5 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 in all 18 steps (Figure 5, top left), and the method yields the most consistent results as demonstrated by lowest variance of the mean of Q_1 in the Monte Carlo study (Figure 5, top right). This suggests that the potential theoretical drawback of the AMIS method—that is, the possibility of a biased estimate—does not occur, or that it is negligible in this application. The best performance of the AMIS method can be contributed to its ability to achieve the highest number of effective particles. Note that it never falls below 100 (Figure 6, left), which sharply contrasts with that of the bootstrap proposal which falls as low as one effective particle.

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 Equation (10). This model

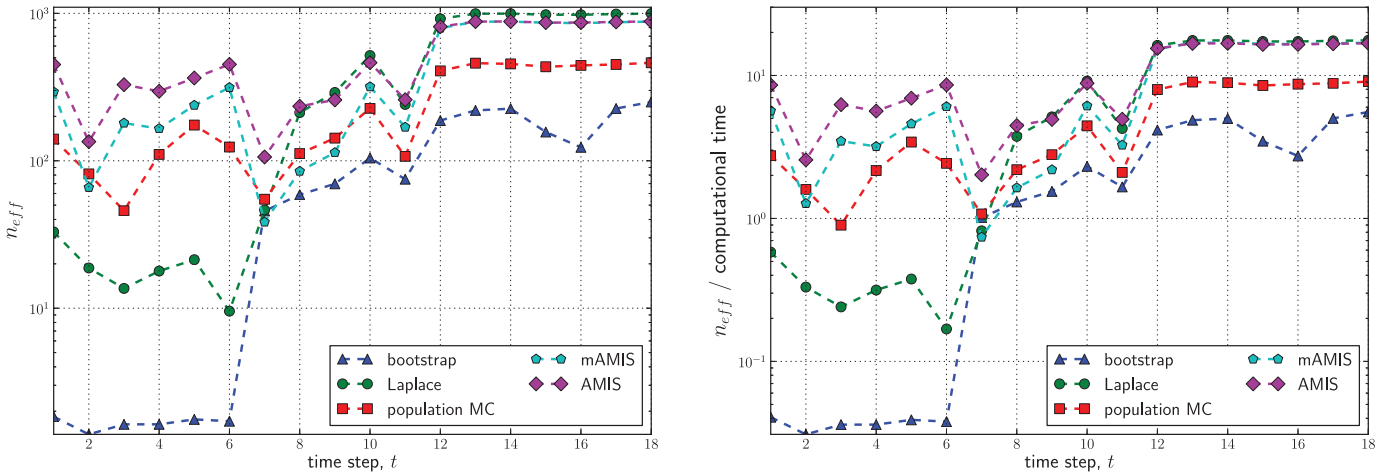

Figure 6. Release with constant release rate, known start, and duration. Left: average number of effective particles over 100 Monte Carlo trials. Right: average number of effective particles per one second of execution time over 100 Monte Carlo trials.

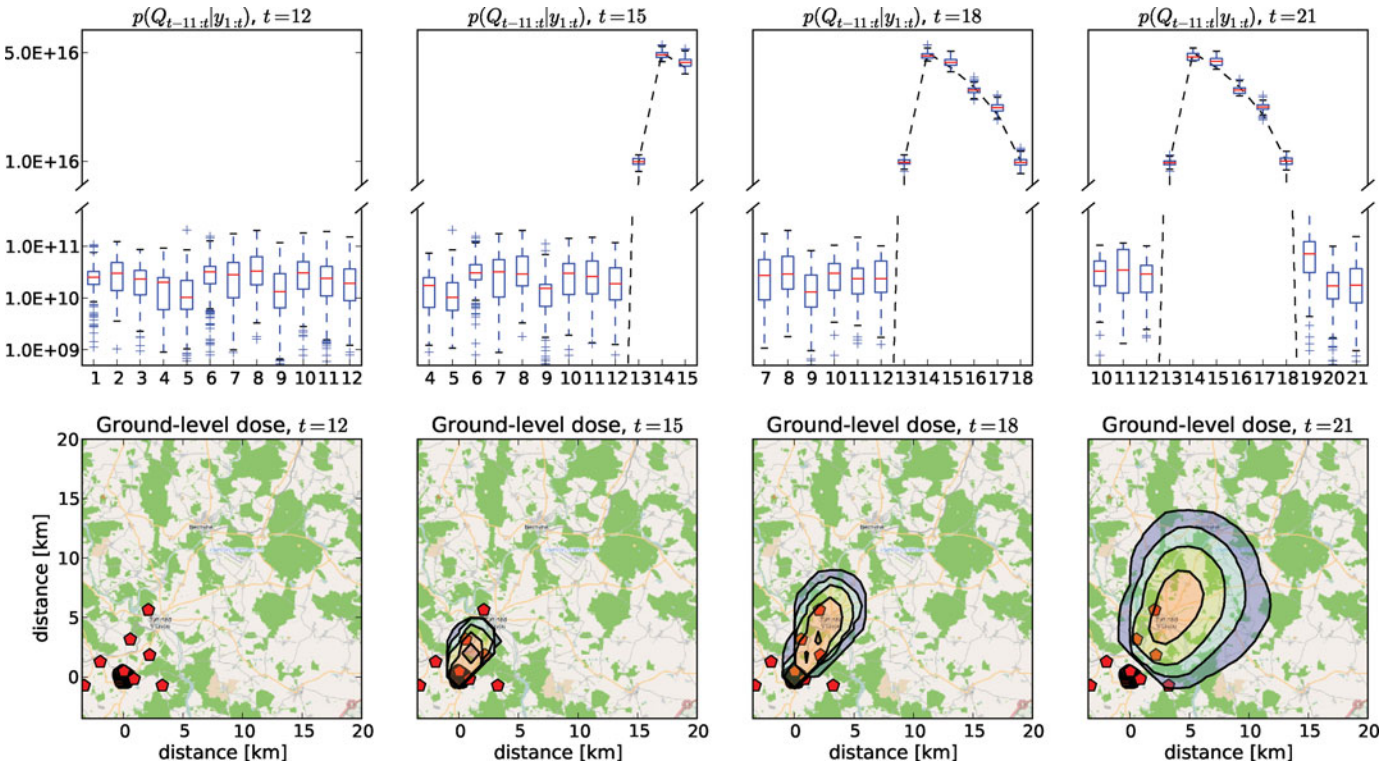

Figure 7. Continuous monitoring of a sudden release of activity. Top row: Boxplot of estimated activity of the last 12 (hypothetically) released puffs, and its comparison to the simulated values (dashed line). Bottom row: contour plots of the mean value of the ground level gamma dose at each point on 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.

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 allow 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 python code and data for this experiment is provided online as a supplementary material. 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 a 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 min 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 the 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

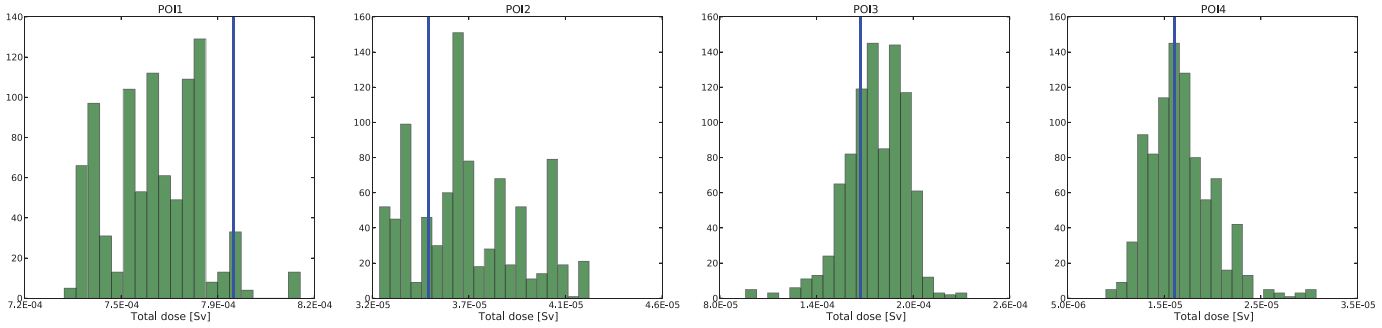

Figure 8. Histograms of the total committed dose from a radiation accident in selected POI (locations of the POI are displayed in Figure 1). The vertical line denotes the simulated value.

were estimated before the release. This confirms informativeness of data from the first ring of the RMN.

The posterior densities (Equation (24)) encode useful information which needs to be presented to the decision makers. One of the most valuable supporting materials from the human point of view is 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, leading to contours of its own, or alternatively a histogram of the quantity of interest at the POI, 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 RMNs. The model is highly nonlinear 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 a combination of ideas of local linearization, population Monte Carlo, and information geometric parameter estimation. The new proposal density significantly improved efficiency of the filter that allows us to evaluate much more informative estimates in the same computational time. The new proposal also allowed us to use a noninformative 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 with inexpensive hardware. Moreover, efficient proposals may allow us to estimate more parameters of interest, for example, the dispersion coefficients or Pasquill's category of atmospheric stability, in the future.

SUPPLEMENTARY MATERIAL

Data and code: Data and code used in the experiment in Section 5.4. The code is implemented in the Python language using numpy libraries. Further speedup of the code execution can be achieved by translation of the code to the C language via the cython project. However, this is optional. (zip file)

Derivations: Laplace approximation of the shifted inverse gamma likelihood and gamma prior. (pdf file)

ACKNOWLEDGMENT

Support of grant MV ČR VG20102013018 is gratefully acknowledged.

[Received December 2011. Revised October 2013.]

REFERENCES

Anderson, J., and Anderson, S. (1999), "A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts," *Monthly Weather Review*, 127, 2741–2758. [514]

Bartzis, J., Ehrhardt, J., French, S., Lochard, J., Morrey, M., Papamichail, K., Sinkko, K., and Sohier, A. (2000), "RODOS: Decision Support for Nuclear Emergencies," in *Recent Developments and Applications in Decision Making*, eds., S. Zanakis, G. Doukidis, and C. Zopounidis, Dordrecht, The Netherlands: Kluwer, pp. 379–394. [526]

Cappé, O., Guillin, A., Marin, J., and Robert, C. (2004), "Population Monte Carlo," *Journal of Computational and Graphical Statistics*, 13, 907–929. [521]

Cervone, G., Franzese, P., and Grajdeanu, A. (2010), "Characterization of Atmospheric Contaminant Sources Using Adaptive Evolutionary Algorithms," *Atmospheric Environment*, 44, 3787–3796. [514]

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. [515,521]

Cornuet, J. M., Marin, J. M., Mira, A., and Robert, C. P. (2012), "Adaptive Multiple Importance Sampling," *Scandinavian Journal of Statistics*, 39, 798–812. [515,521]

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. [514,517]

Doucet, A., de Freitas, N., and Gordon, N. (eds.) (2001), *Sequential Monte Carlo Methods in Practice*, New York: Springer. [520,521]

Doucet, A., Godsill, S., and Andrieu, C. (2000), "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, 10, 197–208. [521]

Dowson, D., and Wragg, A. (1973), "Maximum-Entropy Distributions Having Prescribed First and Second Moments," *Information Theory, IEEE Transactions*, 19, 689–693. [517]

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. [523]

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. [514]

Gifford, F. (1976), "Turbulent Diffusion-Typing Schemes: A Review," *Nuclear Safety*, 17, 68–86. [515]

Golub, G., and Welsch, J. (1969), "Calculation of Gauss Quadrature Rules," *Mathematics and Computation*, 23, 221–230. [523]

Gordon, N., Salmond, D., and Smith, A. (1993), "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE*, 140, 107–113. [515,520,521]

Hanna, S., Briggs, G., and Hosker, Jr, R. (1982), "Handbook on Atmospheric Diffusion," Technical Report, National Oceanic and Atmospheric Administration, Oak Ridge, TN: Atmospheric Turbulence and Diffusion Lab. [515,520]

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. [514,519]

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. [514,517,518,523]

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. [515]

Jeffreys, H. (1961), *Theory of Probability* (3rd ed.), Oxford, UK: Oxford University Press. [518]

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, 308–313. [514]

Johannesson, G., Hanley, B., and Nitao, J. (2004), "Dynamic Bayesian Models via Monte Carlo–An Introduction With Examples," Technical Report, Lawrence Livermore National Laboratory. [514,516,517,518,521]

Kass, R. E., and Raftery, A. E. (1995), "Bayes Factors," *Journal of American Statistical Association*, 90, 773–795. [521,522,524]

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. [518]

Kennedy, M., and O'Hagan, A. (2001), "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society*, Series B, 63, 425–464. [520]

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. [514]

Kulhavý, R. (1996), *Recursive Nonlinear Estimation: A Geometric Approach*, Berlin: Springer. [515,521,522]

Marin, J., Pudlo, P., and Sedki, M. (2012), "Consistency of the Adaptive Multiple Importance Sampling," *arXiv:1211.2548*. [522]

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. [518]

Oh, M., and Berger, J. (1992), "Adaptive Importance Sampling in Monte Carlo Integration," *Journal of Statistical Computation and Simulation*, 41, 143–168. [521]

Pecha, P., and Hofman, R. (2011), "Construction of Observational Operator for Cloudshine Dose From Radioactive Cloud Drifting Over the Terrain," in *Proceedings of the 14th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*. [519,523]

Peterka, V. (1981), "Bayesian Approach to System Identification," in *Trends and Progress in System Identification*, ed. P. Eykhoff, Oxford: Pergamon Press, pp. 239–304. [520]

Pitt, M., and Shephard, N. (1999), "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 590–599. [515,521]

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. [515]

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. [519]

Rubinstein, R., and Kroese, D. (2004), *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*, Berlin: Springer. [521,522]

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. [514,517]

Šmídl, V., and Hofman, R. (2013), "Tracking of Atmospheric Release of Pollution Using Unmanned Aerial Vehicles," *Atmospheric Environment*, 67, 425–436. [520]

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, 89–100. [519]

Twenhöfel, C., van Troost, M., and Bader, S. (2007), "Uncertainty Analysis and Parameter Optimization in Early Phase Nuclear Emergency Management," Technical Report, RIVM. [514,517]

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*, 45, 2944–2955. [514]

Zannetti, P. (1990), *Air Pollution Modeling*, New York: Van Nostrand Reinhold. [515]