RESEARCH REPORT

Václav Šmídl, Radek Hofman

Bayesian Methods for Optimization of Radiation Monitoring Networks

Particle Filtering Approach

2315	December 2011

MV ČR VG20102013018

1 Introduction

Release of radioactive material into the atmosphere is the last possible resort of any accident in a nuclear power plant. It is an extremely rare event, however with severe consequences for potentially many people living in proximity of the power plant. Awareness of radiation security has been increased after the Chernobyl accident, and almost every country is now equipped with monitoring network of on-line connected receptors continually measuring radiation levels. Initial configurations of the network were designed by experts using their experience. With increasing pressure on improvement on the network reliability and overall safety, many countries are considering expansion of the measurement network or it reconfiguration.

Hence, formal methodologies how to optimize the monitoring network (and mobile extensions) received recently significant attention in the literature, see e.g. [Baume et al., 2011, Abida et al., 2008, Melles et al., 2011] for the latest development.

In general, there are many aspects to consider when designing monitoring network. Some issues are now summarized:

Scale: large scale networks needs to be considered for severe accidents [Abida et al., 2008], while less severe accidents affect only direct neighbourhood of the power plant.

Network Type: different approaches has to be taken for optimization of stationary network which needs to consider long term statistics [Abida et al., 2008, Melles et al., 2011], and specific short-term scenario and mobile groups [Heuvelink et al., 2010, Abida and Bocquet, 2009]. The latter approach is to be applied when the radiation plume is already in the air and the task is to improve its monitoring as fast as possible.

Uncertainty model: Common tool for modeling spatial filed an it uncertainty is krieging, or regression krieging which is commonly used in optimization [Melles et al., 2011]. Another possibility is to represent the posterior by Gaussian density [Zidek et al., 2000]. An alternative model of uncertainty is Monte Carlo samples [Johannesson et al., 2004, Hiemstra et al., 2011] which has not been used in any optimization approach yet.

Loss function: defines criteria of optimization. One of the early criteria was information entropy [Caselton and Husain, 1980]. It has been extended to reflect cost of measurement installation in [Zidek et al., 2000]. However, its results are very sensitive to the introduced cost of the entropy which may be hard to quantify.

More intuitive option is to count success or failure of an assimilation procedure to correctly classify the affected area into the classes of radiation protection countermeasures, [Heuvelink et al., 2010].

Optimization restrictions: unrestricted forms of optimization via. simulated annealing were employed on the large scale, where accuracy of location is low hence there is still enough freedom to locate the station with respect to local conditions. Application of this approach to short term optimization, e.g. of mobile groups [Abida and Bocquet, 2009, Heuvelink et al., 2010], is less attractive since the freedom to choose location is limited.

In this paper, we are concerned with local scale modeling of less severe accident in the range of tens of kilometers from the nuclear power plant. Both the stationary and mobile groups will be discussed. The preferred model of uncertainty is the empirical density which will be assimilated with measurements using the sequential Monte Carlo methodology. We will discuss influence of various loss functions. In general, we will not consider free form optimization but only comparison

of preselected set of network configurations. The main reason for this are practical issues that needs to be taken into account such as: availability of power supply, maintenance, permission from owner of the land, etc. These concerns are too complex to optimize automatically. Therefore, these will be assessed by experts and the task is only to evaluate suitability of certain configuration over its competitors.

2 Methods

2.1 Decision Theory Framework

The key result of decision theory under uncertainty is formally simple [Berger, 1985]. If we are to choose which action, a^*, from a set $\mathcal{A} = \{a_1, \ldots, a_3\}$ is best, we are to choose such action that minimizes the expected value of the loss function

$$a^* = \arg \min_{a \in \mathcal{A}} E_{p(x|d)}(L(a, X)|D), \tag{1}$$

where
   X is the potential outcome of the action,
   $L(a, X)$ is a function mapping the space of all actions and outcomes to the real axis,
   D are the measured data,
   $E_{p(X|D,a)}()$ is the operator of expected value $E_{p(X|D)}(L(a, X)) = \int p(X|D, a)L(a, X) \, dX$, and
   $p(X|D)$ is the probability of realization of a specific value of X given realization of data D and action a.
   This formalizm is rather general and its results will differ with different choices of the loss function L and/or different reprezentations of uncertainty. Indeed, many existing solutions may be interpreted as various choices of these two factors. For example, entropy minimization techniques choose logarithm of the probability density function to be their loss function. In particular, Zidek et al. [2000] considered X to be spatial distribution of the pollutant with Gaussian distributed density. Different scenarios of optimization are defined by their parameterizations.

Stationary network the set of actions $\mathcal{A}$ contains all possible configurations of the measuring sites described by their spatial coordinates. The space may contain networks of various sizes. The uncertainty space X contains all possible realizations of weather conditions and release scenarios. The available data, D, are weather measurements recorded over a period of time. An example of such optimization is described in [Abida et al., 2008].

Mobile sensors the set of sensors is usually fixed and the action space $\mathcal{A}$ contains possible trajectories of the mobile sensors. The uncertainty space, X, contains parameters of the release and weather conditions. The measured data D contains measurements of the actual weather and readings from the stationary monitoring network.

In this paper, we will focus on the mobile sensor scenario, however, the methodology can be easily applied to optimization of the stationary network.

2.2 Loss function

We will test various loss functions proposed in the literature: (i) entropy, (ii) misclassification of decisions.

2.2.1 Entropy optimization

The purpose of the locating new mobile measuring station is to reduce uncertainty in the estimated parameters. This vague idea can be formalized using the entropy principle [Caselton and Husain, 1980, Zidek et al., 2000]. The main idea follows from the following equality

$$H(X, Z|D, \lambda) = H(X|Z, \lambda, D) + H(Z|D, \lambda), \tag{2}$$

which is a well known identity on from entropy. A common assumption made e.g. in [Zidek et al., 2000] is that the total entropy $H(X, Z(\lambda)|D)$ is constant for all locations λ since we are not adding any new information. This is true only under assumption that the entropy of measurements is independent of its location λ and X. Rewriting (2) as

$$H(X, Z|D, \lambda) = H(Z|X, \lambda, D) + H(X|D, \lambda), \tag{3}$$
$$= H(Z|X, \lambda, D) + H(X|D), \tag{4}$$

we note that the entropy in X can not be changed by λ. However, the conditional entropy $H(Z|X, \lambda, D)$ can.

Example 1. Consider complete knowledge of the state variable $p(X) = \delta(X - \overline{X})$, and normal distributed measurement error with variance $\sigma(X, \lambda, D)$. Then the joint entropy is

$$H(X, Z|D, \lambda) = \int p(Z|\overline{X}, \lambda, D) \log p(Z|\overline{X}, \lambda, D) dZ$$
$$= \frac{1}{2} \log(2\pi|\sigma(\overline{X}, \lambda, D)|).$$

which is constant only for constant variance of observations σ.

In this work, we are concerned primarily with observations that have absolute value of their error relative to the observed value. Hence, we should not rely on this assumption and perform full optimization of $H(X|Z, \lambda, D) = H(X, Z|D, \lambda) - H(Z|D, \lambda)$.

2.2.2 Mutual information

In Hoffmann and Tomlin [2010], it is argued that optimal loss is the mutual information

$$I(Z; X) = H(Z) + H(X) - H(X, Z). \tag{5}$$

Since $H(X)$ is constant, (5) and (3) are equivalent.

2.2.3 Misclassification of decision

While various loss functions has been proposed in the literature, we follow [Heuvelink et al., 2010] and define loss function to be proportional to the number of incorrectly classified people in evacuation zones

$$L(\lambda, X) = \alpha I_{false\_positive} + \beta I_{false\_negative}, \tag{6}$$

where $I_{false\_positive}$ is the number of people incorrectly classified for evacuation, and $I_{false\_negative}$ is the number of people that are incorrectly classified to stay in the polluted area. It is computed as a sum over all inhabited places indexed by j:

$$I_{false\_positive} = \sum_j population_j \times (\hat{C}_j > \overline{C} \,\&\, C_j < \overline{C}),$$
$$I_{false\_negative} = \sum_j population_j \times (\hat{C}_j < \overline{C} \,\&\, C_j > \overline{C}).$$

4

Here, $\overline{D}$ denotes a threshold for the total accumulated dose of the pollutant.

For simplicity, we may split the area around the power plant into J districts, each representing a constant number of inhabitants, e.g. 100, which live approximately at given location, $\mathbf{i}_j$, $j = 1 \ldots J$, which denote the points of interest. The total absorbed dose in these localities will be represented by a vector $\mathbf{c} = [C(\mathbf{i}_1), C(\mathbf{i}_2), \ldots C(\mathbf{i}_J)]$.

$$I_{false\_positive} = \sum_{j=1}^{J} (\hat{\mathbf{c}}_j > \overline{C} \,\&\, \mathbf{c}_j < \overline{C}). \tag{7}$$

$$I_{false\_negative} = \sum_{j=1}^{J} (\hat{\mathbf{c}}_j < \overline{C} \,\&\, \mathbf{c}_j > \overline{C}). \tag{8}$$

and the loss function can be expressed in terms of expected value

$$
\begin{aligned}
L(a, X) &= L(\mathbf{c}(X, \lambda)). \\
\hat{\mathbf{c}}(\lambda) &= \mathsf{E}(\mathbf{c}(X, Z)|\lambda).
\end{aligned}
\tag{9}
$$

2.3 Atmospheric dispersion model

When the pollutant is released into the atmosphere, it forms a plume the shape of which is subject to wind and dispersion. For a perfectly known parameters of the release θ and a perfectly known weather conditions w (containing wind speed and direction, Pasquill's stability category, etc.), the shape of the plume can be very well approximated by the Gaussian plume model of a sequence of puff models [Thykier-Nielsen et al., 1999]. The proposed methodology will work with any type of dispersion model, however, we choose the classical puff model to perform our experiments. Specifically, the release is modeled by a sequence of puffs labeled $i = 1, \ldots, I$ where concentration of the pollutant in one puff at time τ is given by:

$$C_i(\mathbf{s}, \tau) = \frac{Q_i}{\sqrt{2}\pi^{3/2}\sigma_1\sigma_2\sigma_3} \exp\left[-\frac{(s_1 - l_{1,i,\tau})^2}{2\sigma_1^2} - \frac{(s_2 - l_{2,i,\tau})^2}{2\sigma_2^2} - \frac{(s_3 - l_{3,i,\tau})^2}{2\sigma_3^2}\right], \tag{10}$$

where $\mathbf{s}$ vector of spatial coefficients, $\mathbf{l}_i = [l_{1,i,\tau}, l_{2,i,\tau}, l_{3,i,\tau}]$ is vector of location of the puff center, $\sigma_1, \sigma_2, \sigma_3$ are dispersion coefficients, and Q_i is the released quantity in the ith puff.

2.4 State space model

The key uncertainty is with the released dose $Q_i, i = 1, \ldots, I$ and the weather conditions ϕ and v. We have chosen to model the uncertainty in the wind field as corrections of the numerical weather prediction:

$$
\begin{aligned}
v_t(\mathbf{s}) &= \tilde{v}_t(\mathbf{s})a_t, \tag{11} \\
\phi_t(\mathbf{s}) &= \tilde{\phi}_t(\mathbf{s}) + b_t, \tag{12}
\end{aligned}
$$

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 prediction model at time t. Correction of the wind field forecast is then achieved by estimation of a_t and b_t using random walk model on their time evolution.

We assume that we measure the radiation dose, wind speed and wind direction at meteostation near the power plant.

The full state of the system then

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

and observation vector at time t is

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

The measurements $y_{2,t}$ and also $z_{j,t}$ are assumed to have normal distribution around value $\mu_j^{(i)}$ predicted by the ith dispersion model $\mu_j^{(i)} = dispersion\_model(X^{(i)}, \lambda_j)$. Standard deviation of the measurements is proportional to the mean

$$p(Z|X^{(i)}, \lambda) = \mathcal{N}(\mu_j^{(i)}, (\gamma \mu_j^{(i)})^2), \tag{13}$$

while the measurements are assumed to be uncorrelated. I.e. the vector of measurements

$$
\begin{aligned}
p(\mathbf{z}|X^{(i)}, \lambda) &= \mathcal{N}(\mu^{(i)}, \Sigma^{(i)}), \\
\mu^{(i)} &= [\mu_1^{(i)}...], \\
\Sigma^{(i)} &= \mathrm{diag}(\gamma \mu^{(i)})^2.
\end{aligned}
$$

2.5 Data Assimilation via Sequential Monte Carlo

We assume that all uncertainty is modeled by empirical probability density function

$$p(X_t|\mathbf{D}_t) \approx \sum_{i=1}^{n} w_t^{(i)} \delta(X_t - X_t^{(i)}), \tag{14}$$

where $X_t^{(i)}$ is a sample of the state space trajectory $X_t = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t]$. Assimilation of the measured data is then achieved via sampling-importance-resampling procedure, where the weights can be computed recursively,

$$w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(y_t|x_t) p(x_t|x_{t-1})}{q(x_t)}.$$

3 Navigation of mobile sensors

3.1 Entropy loss

For the entropy loss function, we are to evaluate relation (3), where

$$H(X, Z|D, \lambda) = \int p(X, Z) \log p(X, Z) \, dX \, dZ, \tag{15}$$

$$H(Z|D, \lambda) = \int p(Z) \log p(Z) \, dZ, \tag{16}$$

Under the chosen approximation (14), the joint and the marginal densities are

$$p(X, Z|D, \lambda) = \sum_i w^{(i)} p(Z|X, \lambda) \delta(X - X^{(i)}). \tag{17}$$

$$p(Z|D, \lambda) = \sum_i w^{(i)} p(Z|X^{(i)}, \lambda). \tag{18}$$

Substituting (17) into (15)–(16) we obtain:

$$
\begin{aligned}
H(X,Z|D,\lambda) &= \int \sum_i w^{(i)} p(Z|X^{(i)},\lambda)\delta(X-X^{(i)}) \log\left[\sum_i w^{(i)} p(Z|X^{(i)},\lambda)\delta(X-X^{(i)})\right] dZ\, dX. \\
&= \int \sum_i w^{(i)} p(Z|X^{(i)},\lambda) \log\left[p(Z|X^{(i)},\lambda)\right] dZ \\
&= \sum_i w^{(i)} H(Z|X^{(i)}) \quad (19)
\end{aligned}
$$

$$
\begin{aligned}
H(Z|D,\lambda) &= \int \sum_i w^{(i)} p(Z|X^{(i)},\lambda) \log\left[\sum_i w^{(i)} p(Z|X^{(i)},\lambda)\right] dZ. \quad (20) \\
&= \sum_i w^{(i)} \int p(Z|X^{(i)},\lambda) \log\left[\sum_i w^{(i)} p(Z|X^{(i)},\lambda)\right] dZ. \quad (21)
\end{aligned}
$$

Note that evaluation of (19) for (13) is relatively simple, since

$$
\begin{aligned}
H(z|X^{(i)},\lambda) &= \frac{1}{2}\log(2\pi) + \log\left(\gamma\prod_k \mu^{(i)}_{j,k}\right) \\
H(X,Z|D,\lambda) &= \frac{1}{2}k\log(2\pi e) + k\log\gamma + \sum_i w^{(i)} \sum_k \log(\mu^{(i)}_{j,k}).
\end{aligned}
$$

However, evaluation of (20) is a complex integral which is hard to evaluate.

Therefore, we propose the following simplifications

Gaussian approximation of the mixture (17) obtained by moment matching:

$$
\begin{aligned}
p(Z|D,\lambda) &= \mathcal{N}(\mu_z,\Sigma_z), \quad (22) \\
\mu_z &= \sum_i w^{(i)}\mu^{(i)}, \\
\Sigma_z &= \left[\sum w^{(i)}\mu^{(i)}(\mu^{(i)})'\right] - \mu_z\mu_z'.
\end{aligned}
$$

Then, the entropy (20) is approximated by the entropy of the Gaussian distribution

$$
H(Z|D,\lambda) \approx \tilde{H}(Z|D,\lambda) = \frac{1}{2}\log\left[(2\pi e)^k \,|\Sigma_z|\right], \quad (23)
$$

where k is dimensionality of the covariance Σ_z.

The final entropy is then

$$
\begin{aligned}
H(X|Z) &\approx \frac{1}{2}k\log(2\pi e) + k\log(\gamma) + \sum_i w^{(i)} \sum_k \log(\mu^{(i)}_{j,k}) - \tilde{H}(Z|D,\lambda), \\
&= k\log(\gamma) + \sum w^{(i)} \sum_k \log(\mu^{(i)}_j) - \frac{1}{2}\log|\Sigma_z|, \quad (24) \\
&= k\log\gamma + \sum_i w^{(i)} \sum_k \log(\mu^{(i)}_{j,k}) - \frac{1}{2}\log|\Sigma_z|, \quad (25)
\end{aligned}
$$

Semi-Gaussian uses the approximation by the Gaussian approximation only inside the log function in (20)

$$
\begin{aligned}
H(Z|D,\lambda) &\approx -\int \sum_i w^{(i)} p(Z|X^{(i)},\lambda) \log\left[\mathcal{N}(\mu_z, \Sigma_z)\right] dz. \\
&= -\sum_i w^{(i)} \int p(Z|X^{(i)},\lambda)\left[-\frac{1}{2}\log|\Sigma_z| - \frac{1}{2}(z-\mu_z)\Sigma_z^{-1}(z-\mu_z)\right] dz. \\
&= \frac{1}{2}\log|\Sigma_z| + \frac{1}{2}\sum_i w^{(i)}\left[\int p(Z|X^{(i)},\lambda)(z-\mu_z)\Sigma_z^{-1}(z-\mu_z)dz.\right], \\
&= \frac{1}{2}\log|\Sigma_z| + \frac{1}{2}\sum_i w^{(i)} \int p(Z|X^{(i)},\lambda)(z'\Sigma_z^{-1}z - \mu_z\Sigma_z^{-1}z - z'\Sigma_z^{-1}\mu_z + \mu_z'\Sigma_z^{-1}\mu_z)dz. \\
&= \frac{1}{2}\log|\Sigma_z| + \frac{1}{2}\sum_i w^{(i)}\left[(\mu^{(i)})'\Sigma_z^{-1}\mu^{(i)} + \mathrm{tr}(\Sigma^{(i)}\Sigma_z^{-1}) - \mu_z\Sigma_z^{-1}\mu^{(i)} - (\mu^{(i)})'\Sigma_z^{-1}\mu_z + \mu_z'\Sigma_z^{-1}\mu\right. \\
&= \frac{1}{2}\log|\Sigma_z| + \frac{1}{2}\left[\sum_i w^{(i)}\left((\mu^{(i)})'\Sigma_z^{-1}\mu^{(i)} + \mathrm{tr}(\Sigma^{(i)}\Sigma_z^{-1})\right) + \Sigma_z - \mu_z\Sigma_z^{-1}\mu_z\right],
\end{aligned}
$$

where the first term is equal to (23) and the second term is its correction.

Numeric integration: where we first establish support of the integral using approximation (22) to be

$$
z \in\ <\mu_z - 3\sqrt{\Sigma_z}, \mu_z + 3\sqrt{\Sigma_z}>. \tag{26}
$$

This support is discretized into M bins and integral (20) is approximated by

$$
\tilde{H}(Z|D) = \sum_{z_m=z_{min}}^{z_{max}} (z_{m+1} - z_m)\left(\sum_i w^{(i)} p(z_m|X^{(i)},\lambda) \log\left[\sum_i w^{(i)} p(z_m|X^{(i)},\lambda)\right]\right). \tag{27}
$$

This technique does not scale well with increasing dimensionality of the measurements.

Also, techniques like importance sampling and Gauss Hermite quadrature can be used.

3.2 Misclassification loss

Expected value of the misclassification loss (6) may be computed as

$$
\begin{aligned}
\mathsf{E}(L(X,Z,\lambda)|\lambda) &= \int p(X,Z|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \int p(Z|X,\lambda)p(X|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \int p(Z|X,\lambda)p(X|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \sum w^{(i)} \int p(Z|X^{(i)},\lambda)L(\hat{C}(\lambda),X^{(i)})dZ, \\
&= \sum w^{(i)} L^{(i)}(\lambda), \tag{28} \\
L^{(i)}(\lambda) &= \mathsf{E}(L(\hat{C}(\lambda),X^{(i)})) \tag{29}
\end{aligned}
$$

Note that (28) is a sum of contributions from each particle (29), where each contribution is an integral over Z.

Algorithm 1 Evaluation of misclassification loss
- For each particle i do

  1. generate K samples of fictitious measurements $Z^{(k)}$,

  2. Compute weights (33), expected value $\hat{\mathbf{c}}$, and $L(\hat{C}(Z^{(k)}), X^{(i)})$,

  3. Sum all loss functions to obtain (30).

- Compute the loss function via (28).

3.2.1 Importance sampling

Since (29) is an expected value, we may use the importance sampling procedure with $p(Z|X^{(i)}, \lambda)$ as its importance function and drawing K random trials $Z^{(k)}, k = 1 \ldots K$. The final approximation of (29) is then:

$$\mathsf{E}(L(\hat{C}(\lambda), X^{(i)})) \quad \approx \quad \sum \frac{p(Z|X^{(i)}, \lambda)}{p(Z|X^{(i)}, \lambda)} L(\hat{C}(Z^{(k)}), X^{(i)}) = \sum_{k=1}^{K} L(\hat{C}(Z^{(k)}), X^{(i)}), \qquad (30)$$

$$L(\hat{C}(Z^{(k)}), X^{(i)}) \quad = \quad \alpha \sum_{j=1}^{J} (\hat{\mathbf{c}}_j > \overline{C} \,\&\, \hat{\mathbf{c}}_j < \overline{C}) + \beta \sum_{j=1}^{J} (\hat{\mathbf{c}}_j < \overline{C} \,\&\, \hat{\mathbf{c}}_j > \overline{C}) \qquad (31)$$

$$\hat{\mathbf{c}} \quad = \quad \sum_{l} C(X^{(l)}) \tilde{w}^{(l)}, \qquad (32)$$

$$\tilde{w}^{(l)} \quad = \quad w^{(k)} \frac{p(Z^{(k)}|X^{(l)}\lambda)}{\sum_{m} w^{(m)} p(Z^{(k)}|X^{(m)}, \lambda)}. \qquad (33)$$

Note that (33) is the same formula as in the update of the particle filter (). In this case however, the measurement $Z^{(k)}$ is fictitious.

3.2.2 Fast importance sampling

While evaluation of loss functions for each particle may be most accurate, it is also computationally demanding. An alternative is to use

$$\mathsf{E}(L(X, Z, \lambda)|\lambda) \quad = \quad \int \sum w^{(i)} p(Z|X^{(i)}, \lambda) L(\hat{C}(\lambda), X^{(i)}) dZ,$$

$$q(Z, i) \quad = \quad \sum w^{(i)} p(Z|X^{(i)}, \lambda). \qquad (34)$$

By drawing K random couples $\{i^{(k)}, Z^{(k)}\}$ we may approximate the whole loss function by the same functions as in (30)–(33), with i replaced by $i^{(k)}$.

3.2.3 Gauss Hermite quadrature

Integration of expected value of a Gaussian distribution can be easily converted into the conditions of Gauss Hermite quadrature [Abramowitz and Stegun, 1972],

$$
\begin{aligned}
\int \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2}\left(\frac{z-\mu_z}{\sigma}\right)^2)L(z)dz &= \left| x = \frac{z-\mu}{\sqrt{2}\sigma}, dz = 2\sigma \right|, \\
&= \sqrt{2}\sigma \int \frac{1}{\sigma\sqrt{2\pi}} \exp(-x^2)L(\sqrt{2}\sigma x + \mu), \\
&= \pi^{-\frac{1}{2}} \int \exp(-x^2)L(\sqrt{2}\sigma x + \mu), \\
&\approx \pi^{-\frac{1}{2}} \sum_{k=1}^{K} w_k L(\sqrt{2}\sigma x_k + \mu),
\end{aligned}
$$

where w_k and x_k are Gauss Hermite coefficients. For example for $K = 6$, the values were obtained using http://www.efunda.com/math/num_integration/findgausshermite.cfm:

	$k = 1$					$k = 6$
x_k	-2.350604973	-1.335849074	-0.4360774119	0.4360774119	1.335849074	2.350604973
w_k	0.0045300099	0.1570673203	0.7246295952	0.7246295952	0.1570673203	0.0045300099

Multivariate integration with respect to z_1 and z_2 can be achieved by transformation of the coordinates to independent variables and application of the quadrature rules for each axis. However, the computation cost grows exponentially with each dimension. Therefore, this approach is unsuitable for higher dimensions.

3.3 Evaluation of the loss on horizon

In scenario with constricted trajectories of the mobile measurements—such as mobile cars on a route—we may need to examine loss function on a horizon of multiple time step ahead. Formally, we still seek optimization of (1), however with notable extensions:

1. Then, the future measurements of the stationary network are also unknown and the stationary network sites become part of the Z vector. This considerably increases its dimension, and complicates evaluation of the integrals.

2. The state X now contains also future values of the particles. This can be easily evaluated and does not substantially increase the computational complexity. However, the accuracy of the prediction is limited.

4 Results

A hypothetical 1 hour long release of radionuclide ^{41}Ar with half-life of decay 109.34 minutes was simulated. Bayesian filtering is performed in time steps $t = 1, \ldots, 18$, with sampling period of 10 minutes. This sampling period was chosen to match the sampling period of the radiation monitoring network which provides measurements of time integrated dose rate in 10-minute intervals. The same period was assumed for the anemometer. The simulated release started at time $t = 1$ with release activity $Q_{1:6} = [1, 5, 4, 3, 2, 1] \times 1e16\,Bq$.

Values of the measurements were simulated as random draws from measurement model (13) with parameters given by the dispersion model with the "true" parameters.
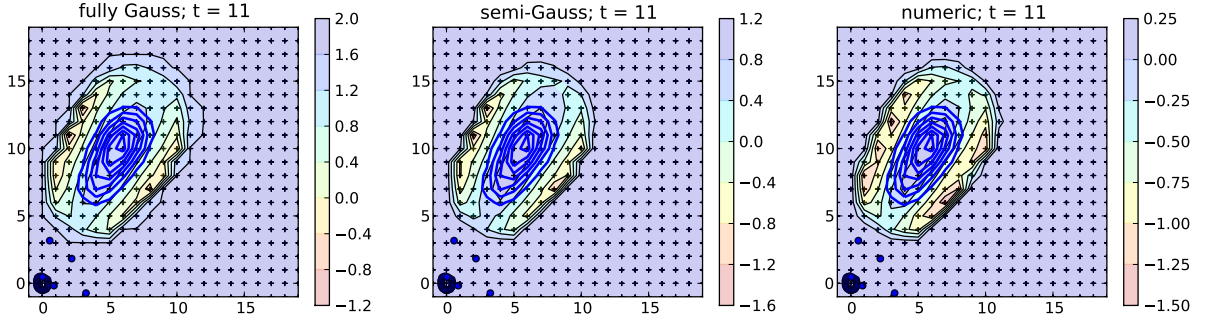
Figure 1: Contour plot of estimated conditional entropy on the grid of λ for three approximations: Gaussian, semi-Gaussian and numeric, repectively.
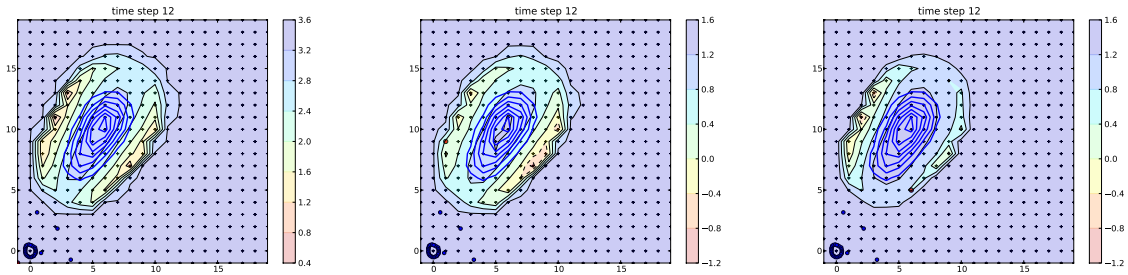

Figure 2: Entropy of the second measurement location for three selected locations of the first measurement (denoted by red dot)..

4.1 Assimilation method

Particle filter with $N = 100$ was used to obtain estimates of the posterior. Estimates of the dose at time $t = 12$ given observations up to time $t = 12$ are displayed in figure x, via their mean value.

4.2 Entropy

Evaluation of conditional entropy was performed on a rectangular grid of $\lambda = [\lambda_1, \lambda_2]$, $\lambda_1 = \langle 0, 10 \rangle$km, $\lambda_2 = \langle 0, 10 \rangle$km.

The value of conditional entropy for each location of λ is displayed in Fig. 1. Note that their differences are negligible. However, computational requirements of their evaluation differ significantly. The Gaussian approximation is the easiest to compute, hence it will be used in further experiments.

An extension of the experiment was to consider a fixed position of one mobile sensor at location λ_{fix} and optimize position the second one. The results are in figure 2.

4.3 Misclassification loss

Evaluation of the misclassification loss was performed on the same grid of λ as for the entropy.

The values of the misclassification loss obtained by the Gauss-Hermite quadrature rule are displayed in figure 3. The points of interests $\mathbf{i}_j$ are all points on the λ grid. To study sensitivity
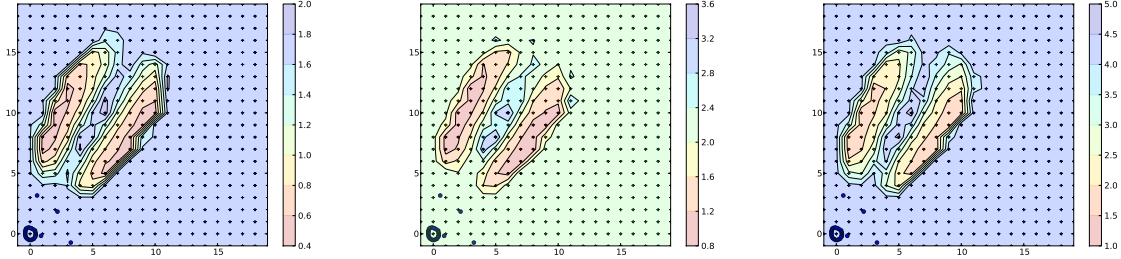
11

Figure 3: Evaluation of the misclassification loss function for various locations λ.
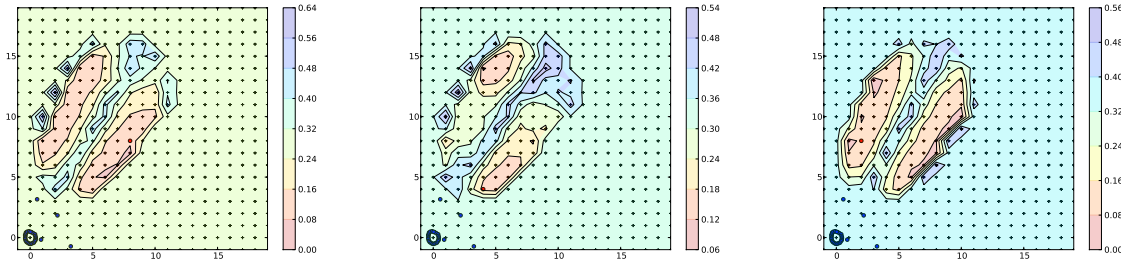

Figure 4: Comparison of contour plot of the misclassification loss (Guss Hermite method) for different points of interest. Point of interest $\mathbf{i}_1$ is denoted by red dot.

of the result to the chosen threshold, we set

$$\overline{C} = k_c \sum_j \hat{\mathbf{c}}(\mathbf{i}_j)$$

for various choices of $k_c = \{20, 1, 1/100\}$. Note that difference between these values is in the scale, however, the shape of the loss is extremely similar.

Sensitivity with respect to the selected point of interest was tested by selecting only one point in $\mathbf{i}_1$. Results computed by the Gauss Hermite method are displayed in Figure 4. Even for a single point of interest the loss function exhibits symmetry around the center axis of pollution spreading. The results are very similar to those obtained for the entropy loss. However, evaluation of the misclassification loss is much more computationally demanding. Therefore, in the simple case presented here, optimization of the entropy loss would yield the expected results for the least computational effort.

The main disadvantage of the Monte Carlo method is its sensitivity to the realization of the random integration points. However, results obtained by the Monte Carlo simulation are rather robust to the realization, see Fig 5. Moreover, they are meaningful even for a single realization $K = 1$. This is caused by the fact that the particles are resampled, i.e. there are copies of the same particle in the ensemble. Hence, different realizations of one point for p copies of the same particle has the same effect as p realizations for the single copied particle. This may be very advantageous property for multivariate integrations.

Advantages of the misclassification loss may become apparent on more demanding assimilation scenarios. Insensitivity of misclassification loss to its parameters (the selected threshold and points of interest) can be explained by small variability of the particle population such that sup-
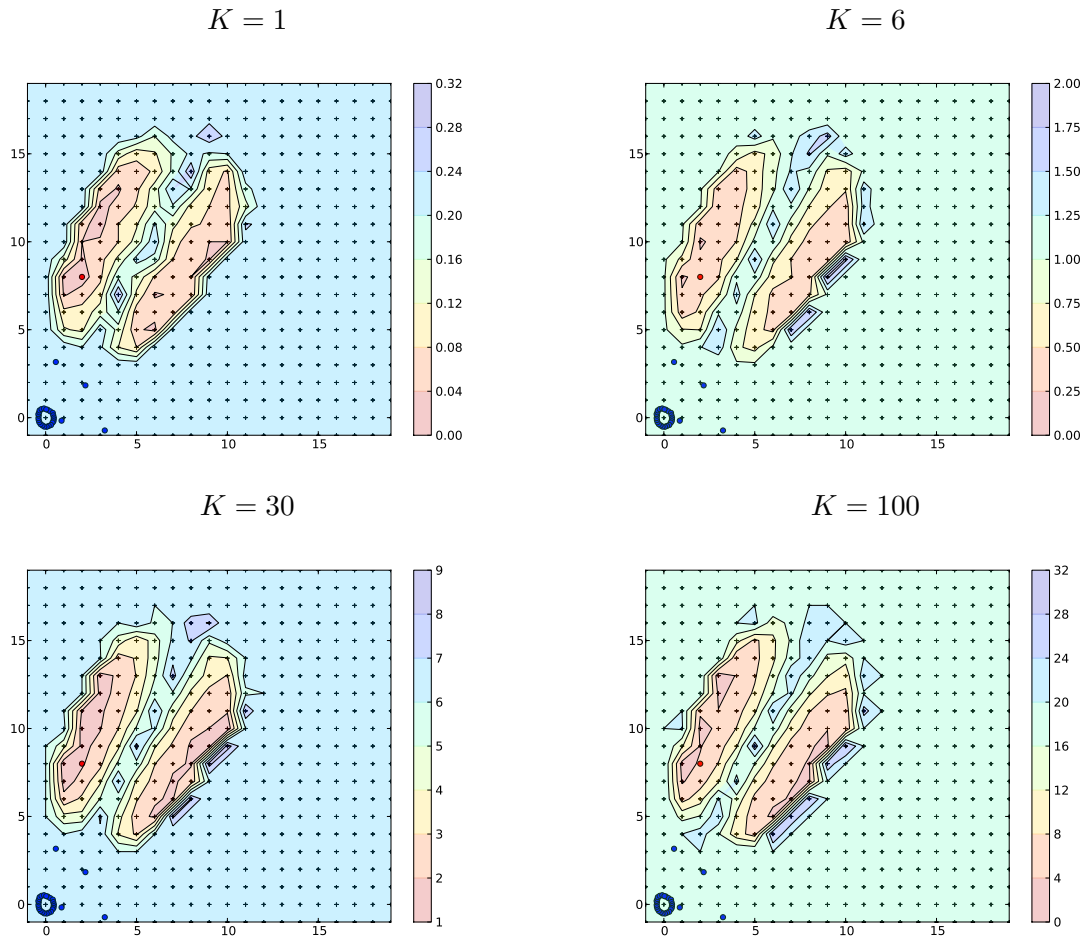
Figure 5: Repetition of the first experiment in Fig. 4 using the Monte Carlo method for different number of Monte Carlo trials.

13

pression of some weights leads to selection of a very small number of them. Further experiments are necessary to confirm this hypotheses.

5 Scenarios of application

The presented methods for evaluation of expected loss function can be applied in various scenarios of radiation protection. In the early phase, they can be used to navigate unmanned aerial vehicle (UAV) to track the radioactive plume. In the late phase, they can be used to evaluate suitability of pre-selected routes for monitoring cars.

5.1 UAVs

In the presence of UAVs, we may expect that they are capable of flying already in the early phase, when the cloud is still over the terrain. We assume that the UAVs are already in the air, and their measurements were used to assimilate the results using the particle filter (14). The task is to select a new direction where the UAV should fly.

- The space of actions $\mathcal{A}$ is then a set of angles $< 0°, 360° >$ and velocities $\langle v_{min}, v_{max} \rangle$ for each UAV.

- In this context, we do not consider planning on long horizon. Only immediate actions are considered.

- The UAVs are not autonomous and are supposed to communicate with the coordination center. Hence, the center solves the optimization problem for them.

- We suppose that the speed of the UAV is small compared to the sampling period. Hence, we may neglect the consequences of the action of the neigbouring UAV. The action space is then naturally decoupled and each UAV is optimized independently. Note however, that since each UAV communicate with the center, actions of all UAV will shape the loss function to optimize.

- The advantage of UAV is that it is well capable of flying even close to the radiation cloud. Hence, we will use the parameter evolution model in this case.

5.2 Mobile groups

The current state-of-the art technology for mobile measurement is the mobile car units. These units are supposed to follow a one of the pre-selected routes that is given by the current emergency regulations. While the number of available units is limited, the task is to select which car takes what route.

- The space of actions $\mathcal{A}$ is then a Cartesian product of set of routes $\mathcal{R}^V, \mathcal{R} \in \{r_1, r_2, \dots r_R\}$ and V is the number of available mobile units.

- The planing horizon is rather long, since we must consider the whole length of the route which takes approximately hours.

- The units also communicate with the center, however, we do not assume frequent re-routing of the mobile groups.

- Since the mobile groups are human operated, the risk of irradiation must be considered as a part of the loss function.

6 Discussion and Conclusion

The task of best acquisition of informative data is addressed. It has been formulated as a problem of decision making under uncertainty. The key factor in this task is the chosen loss function. We have tested two principal loss functions: entropy and misclassification loss. These were computed for the chosen approximation by the empirical distribution. Exact evaluation of these functions is computationally expensive, therefore, we have tested various computational simplifications. W have found that both loss functions and their approximations yield comparable results. Therefore, the follow-up research of their application to navigation of UAVs and mobile car units will be based on the most simple methods, such as the Gaussian approximation of the entropy loss function.

References

R. Abida and M. Bocquet. Targeting of observations for accidental atmospheric release monitoring. *Atmospheric Environment*, 43(40):6312–6327, 2009.

R. Abida, M. Bocquet, N. Vercauteren, and O. Isnard. Design of a monitoring network over france in case of a radiological accidental release. *Atmospheric Environment*, 42(21):5205–5219, 2008.

M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 1972.

OP Baume, A. Gebhardt, C. Gebhardt, GBM Heuvelink, and J. Pilz. Network optimization algorithms and scenarios in the context of automatic mapping. *Computers & Geosciences*, 37 (3):289–294, 2011.

J.O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, 1985.

W.F. Caselton and T. Husain. Hydrologic networks: Information transmission. *Journal of the Water Resources Planning and Management Division*, 106(2):503–520, 1980.

GBM Heuvelink, Z. Jiang, S. De Bruin, and CJW Twenhöfel. Optimization of mobile radioactivity monitoring networks. *International Journal of Geographical Information Science*, 24(3):365–382, 2010.

P.H. Hiemstra, D. Karssenberg, and A. van Dijk. Assimilation of observations of radiation level into an atmospheric transport model: A case study with the particle filter and the etex tracer dataset. *Atmospheric Environment*, pages 6149–6157, 2011.

G.M. Hoffmann and C.J. Tomlin. Mobile sensor network control using mutual information methods and particle filters. *Automatic Control, IEEE Transactions on*, 55(1):32–47, 2010.

G. Johannesson, B. Hanley, and J. Nitao. Dynamic bayesian models via monte carlo–an introduction with examples. Technical report, Lawrence Livermore National Laboratory, 2004.

S.J. Melles, G.B.M. Heuvelink, C.J.W. Twenhöfel, A. van Dijk, P.H. Hiemstra, O. Baume, and U. Stöhlker. Optimizing the spatial pattern of networks for monitoring radioactive releases. *Computers & Geosciences*, 37(3):280 – 288, 2011.

S. Thykier-Nielsen, S. Deme, and T. Mikkelsen. Description of the atmospheric dispersion module RIMPUFF. Technical report, Riso National Laboratory, 1999.

J.V. Zidek, W. Sun, and N.D. Le. Designing and integrating composite networks for monitoring multivariate gaussian pollution fields. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(1):63–79, 2000.