# On-line optimization of radiation monitoring network using mobile units

Václav Šmídl, Radek Hofman

23.1.2012

# Introduction

In case of an accident in a nuclear powerplant we need measured radiological data in order to perform data assimilation and estimate the spatio-temporal radiation situation on terrain.

- Up to now we assumed:
  - radiation monitoring network (RMN) comprises only of stationary measuring devices and its configuration is invariable in time
  - in Czech conditions: RMN is rather sparse and placed close to nuclear powerplants $\Rightarrow$ the number of available observations is small
- New assumptions:
  - there are some mobile units (cars, drones, .. whatever) equiped with dose-rate detectors
  - put in operation after the release start
  - moving with a given speed
  - their movement is restricted to some trajectorie (e.g. roads in case of cars)

# Introduction

- Our goal is to:
  1. Develop a new data assimilation algorithm (based on particle filtering) capable to take into account mobile devices and optime their trajectories in order to obtain as much infromation as possible given physical constraints (speed of mobile devices, their trajectoris, etc.).
  2. As a second goal can be considered development of a suitable algorithm for designig of stationary monitoring network (the best way of further extending the current RMN) using long-term meteorlogical and release statistics in order to maximize their capability of notice unusual radiation levels.

# Available approaches and tools

- Tools for spatial modeling:
  - krieging or regression krieging (Melles et al., 2011)
  - Gaussian models (Zidek et. al, 2000)
  - Monte Carlo (Johannesson, 2005; Hiemstra, 2011)
- Loss functions defining the criteria of optimzation
  - simulated annealing (Huevelink et. al, 2010)
  - entropy optimization

# Selected methods

We treat the problem as a decision task, where we have to choose an action, $a^\star$, from a set $\mathcal{A} = \{a_1, \ldots, a_n\}$ in order to minimize the expected value of a 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 $a$.

# Selected methods

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.

# Selected methods

- **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.

# Selected methods - Entropy optimization

The purpose of the locating new mobile measuring station is to reduce uncertainty in the estimated parameters. The main idea follows from the following equality:

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

where $\lambda$ are "candidates" on observation locations and $Z$ are expected values of data measured in these locations. A common assumption (Zidek et. al, 2000) is that the total entropy $H(X, Z(\lambda)|D)$ is constant for all locations $\lambda$ since we are not adding any new information. This is true only under assumption that the entropy of measurements is independent of its location $\lambda$ and $X$. Rewriting (2) as

$$\begin{aligned} H(X, Z|D, \lambda) &= H(Z|X, \lambda, D) + H(X|D, \lambda), & (3) \\ &= H(Z|X, \lambda, D) + H(X|D), & (4) \end{aligned}$$

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

# Navigation of mobile sensors using etropy loss

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

$$H(X, Z|D, \lambda) = \int p(X, Z) \log p(X, Z)\, dX\, dZ, \tag{5}$$

$$H(Z|D, \lambda) = \int p(Z) \log p(Z)\, dZ. \tag{6}$$

We approximate estimated pdfs using empirical distributions given by a particle filter:

$$p(X, Z|D, \lambda) = \sum_i w^{(i)} p(Z|X, \lambda) \delta(X - X^{(i)}). \tag{7}$$

$$p(Z|D, \lambda) = \sum_i w^{(i)} p(Z|X^{(i)}, \lambda). \tag{8}$$

Substituting (7) into (5)–(6) we obtain:

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

Given that the measurements are normally distributed,

$$p(Z|X^{(i)}, \lambda) = \mathcal{N}(\mu_j^{(i)}, (\gamma\mu_j^{(i)})^2), \tag{12}$$

evaluation of (9) for (12) is relatively simple, since

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

However, evaluation of (10) is a complex integral which is hard to evaluate. We propose three different approximations.

**1. Gaussian approximation** of the mixture (7) obtained by moment matching:

$$\begin{aligned} p(Z|D,\lambda) &= \mathcal{N}(\mu_z, \Sigma_z), \\ \mu_z &= \sum_i w^{(i)} \mu^{(i)}, \\ \Sigma_z &= \left[\sum w^{(i)} \mu^{(i)} (\mu^{(i)})'\right] - \mu_z \mu_z'. \end{aligned} \tag{13}$$

Then, the entropy (10) is approximated by the entropy of the Gaussian distribution

$$H(Z|D,\lambda) \approx \tilde{H}(Z|D,\lambda) = \frac{1}{2} \log \left[(2\pi e)^k \, |\Sigma_z|\right], \tag{14}$$

where $k$ is dimensionality of the covariance $\Sigma_z$.

**2. Semi-Gaussian** uses the approximation by the Gaussian approximation only inside the log function in (10)

$$H(Z|D,\lambda) \approx -\int \sum_i w^{(i)} p(Z|X^{(i)},\lambda) \log\left[\mathcal{N}(\mu_z,\Sigma_z)\right] dz.$$

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

$$= \frac{1}{2}\log|\Sigma_z| + \frac{1}{2}\left[\sum_i w^{(i)}\left((\mu^{(i)})'\Sigma_z^{-1}\mu^{(i)} + \mathrm{tr}(\Sigma^{(i)}\Sigma_z^{-1})\right) + \Sigma_z - \mu_z\Sigma_z^{-1}\mu_z\right],$$

where the first term is equal to (14) and the second term is its correction.

**3. Numeric integration:** where we first establish support of the integral using approximation (13) to be

$$z \in < \mu_z - 3\sqrt{\Sigma_z}, \mu_z + 3\sqrt{\Sigma_z} > . \tag{15}$$

This support is discretized into $M$ bins and integral (10) is approximated by

$$\tilde{H}(Z|D) = \sum_{z_m = z_{min}}^{z_{max}} (z_{m+1} - z_m) \left( \sum_i w^{(i)} p(z_m|X^{(i)}, \lambda) \log \left[ \sum_i w^{(i)} p(z_m|X^{(i)}, \lambda) \right] \right). \tag{16}$$

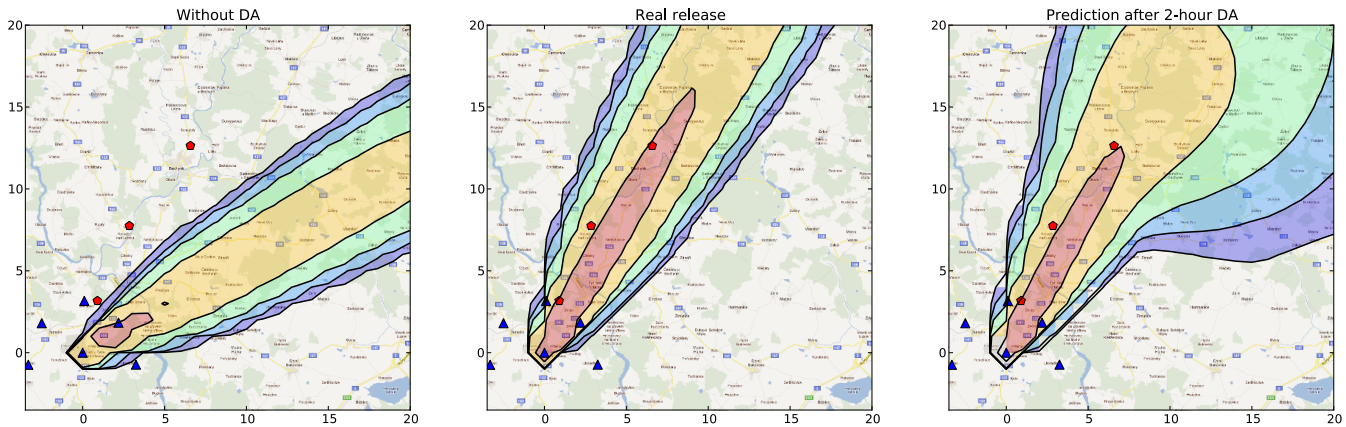
This technique does not scale well with increasing dimensionality of the measurements.

# Optimization - Simulation results

A hypothetical 1 hour long release of radionuclide $^{41}\mathrm{Ar}$ 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 using a twin experiment.

# Optimization - Simulation results

Particle filter with $N = 100$ was used to obtain estimates of the posterior. Estimates of the dose at time $t = 24$ given observations up to time $t = 12$ are displayed via their mean value:



# Optimization - Simulation results - single location

Evaluation of conditional entropy was performed on a rectangular grid of $\lambda = [\lambda_1, \lambda_2]$, $\lambda_1 =< 0, 20 >$km, $\lambda_2 =< 0, 20 >$km. The value of conditional entropy for each location of $\lambda$ 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.

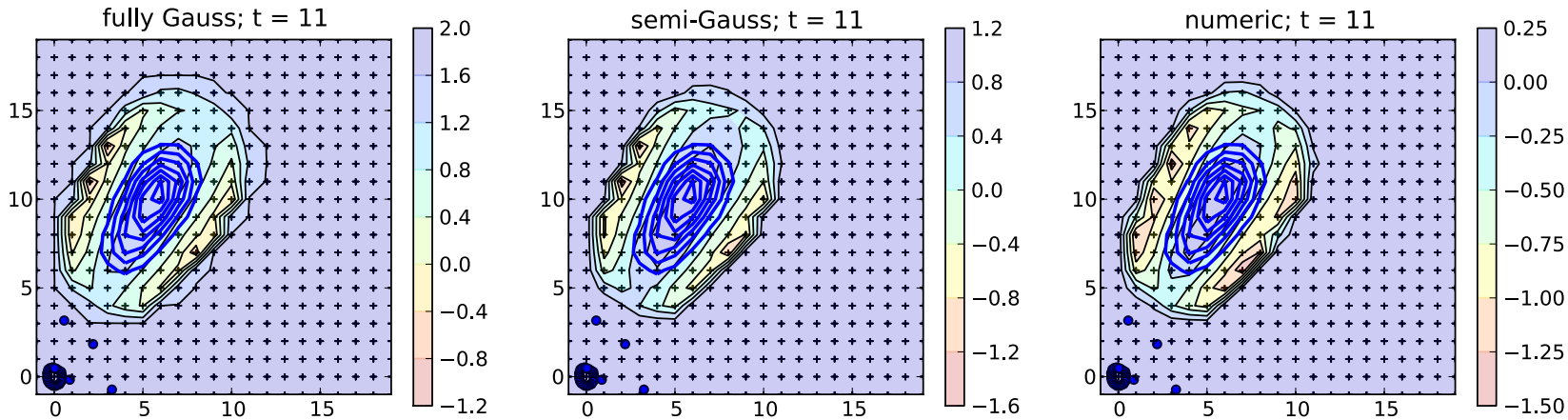



Figure: Contour plot of estimated conditional entropy on the grid of $\lambda$ for three approximations: Gaussian, semi-Gaussian and numeric, repectively.

# Optimization - Simulation results - two locations, one fixed

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

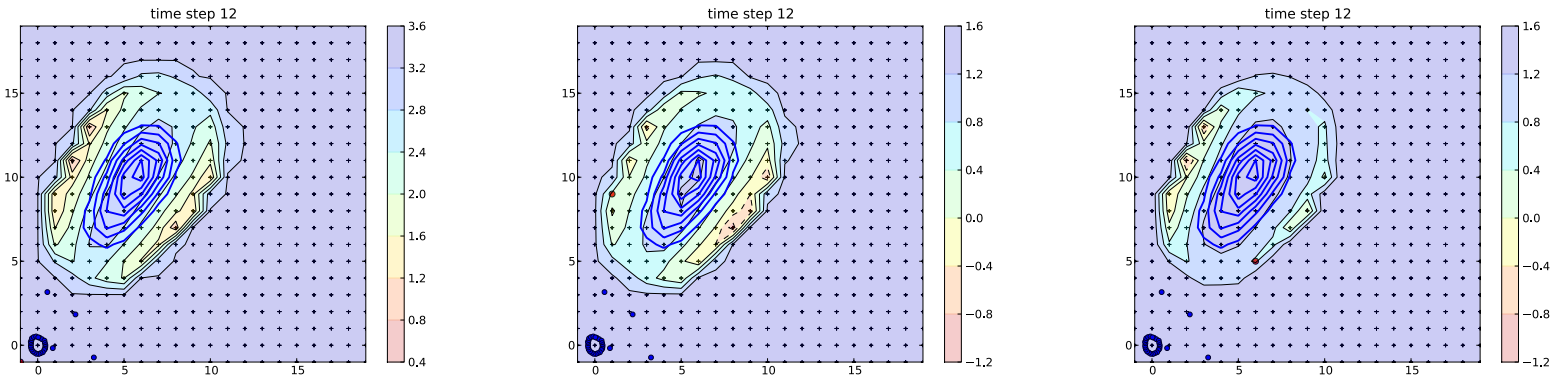



Figure: Entropy of the second measurement location for three selected locations of the first measurement (denoted by red dot)..

# Optimization - Simulation results - two mobile groups

The release is governed by meandering wind field:

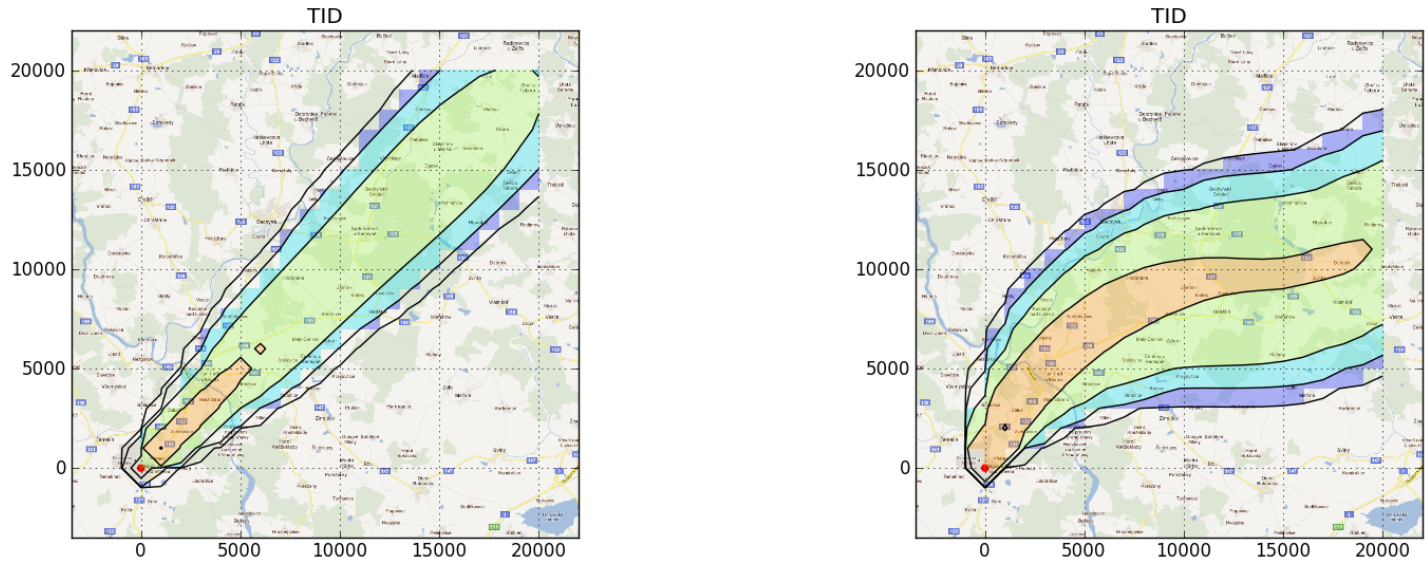
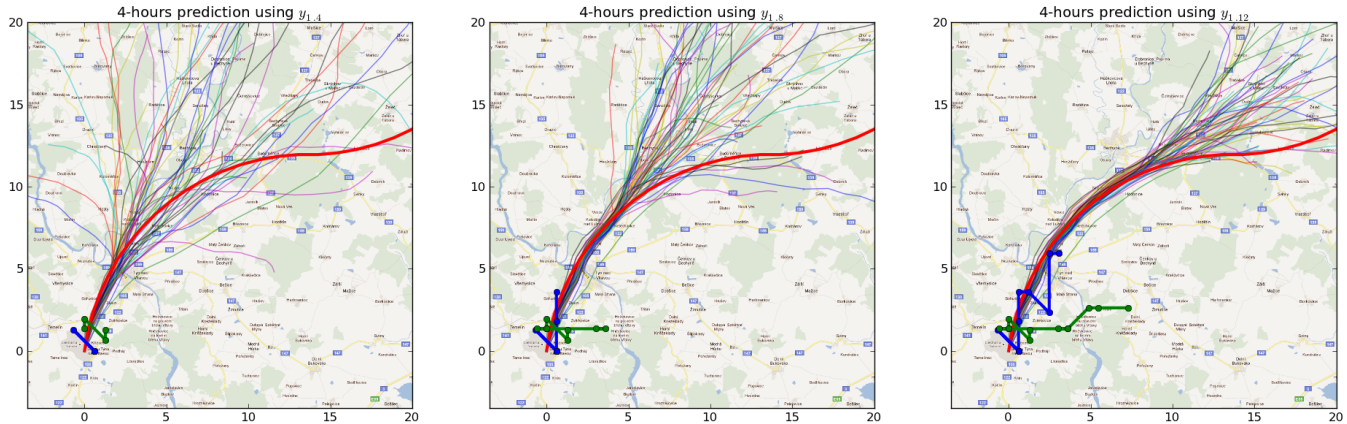


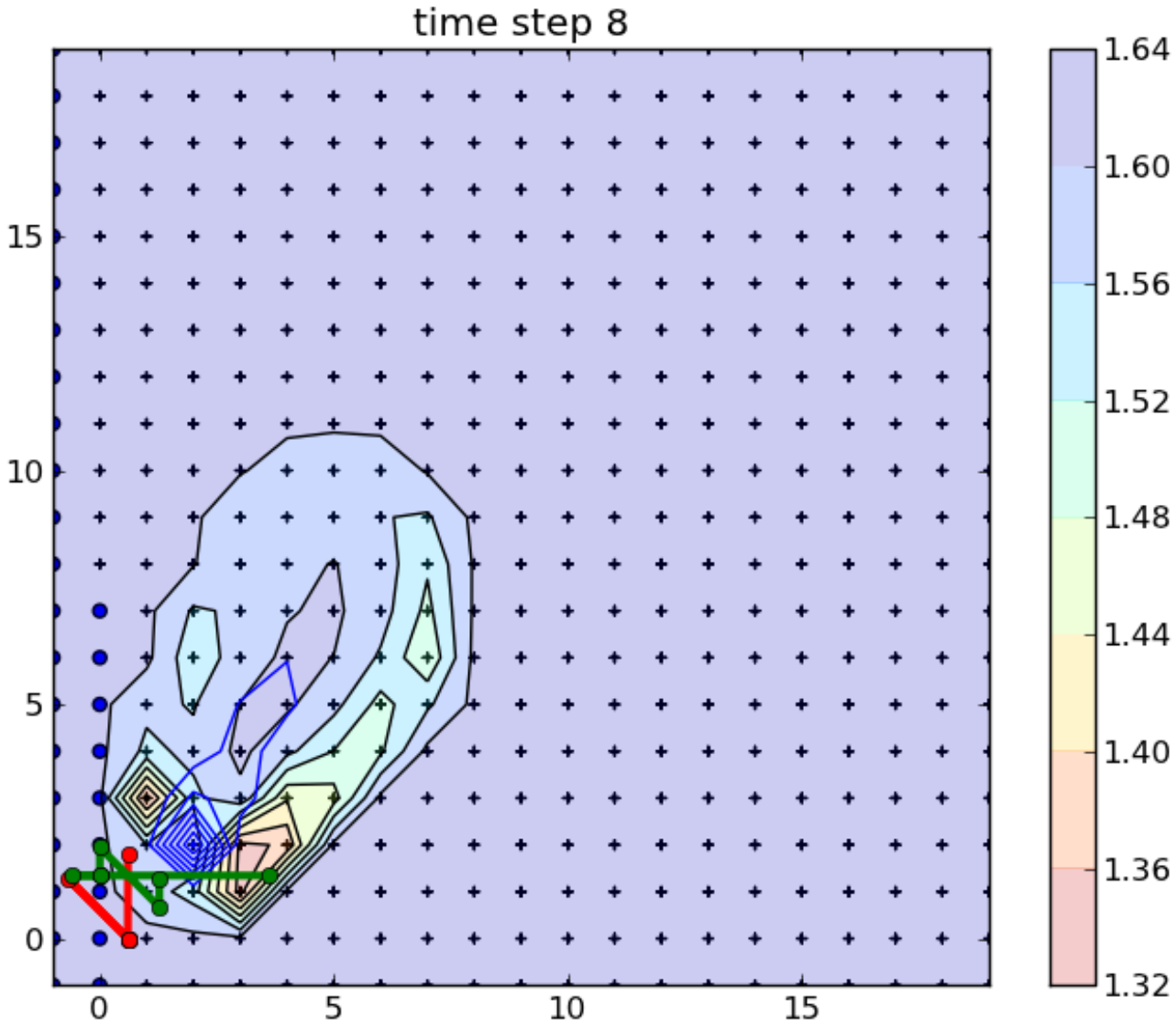

Figure: Release given by meteorological forecast (left) and the twin model representing real release (right).

# Optimization - Simulation results - two mobile groups

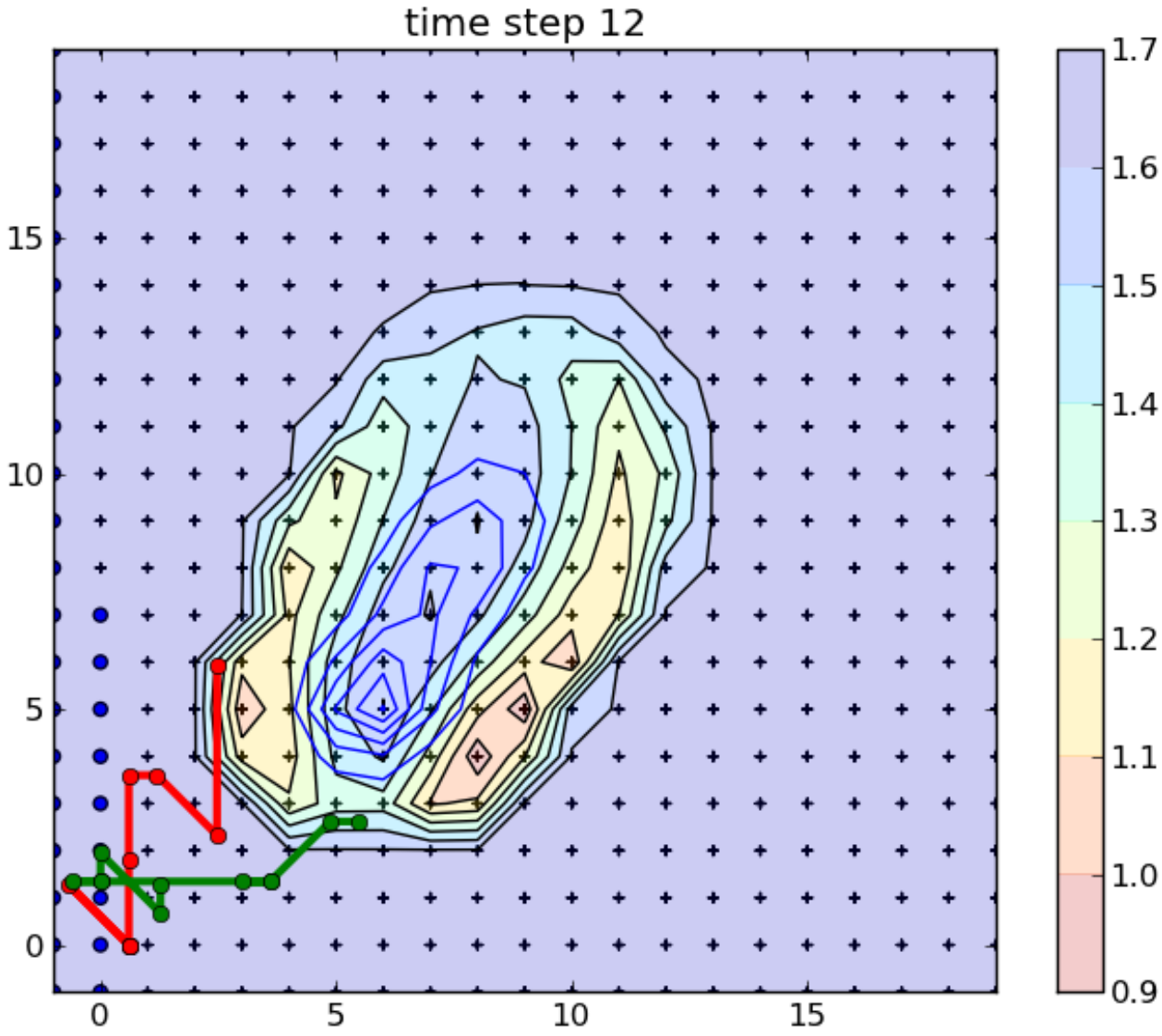
We assume two mobile groups released after the release start. Tha sampling period of mobile groups is 10 minutes. Their action radius is given by their speed.



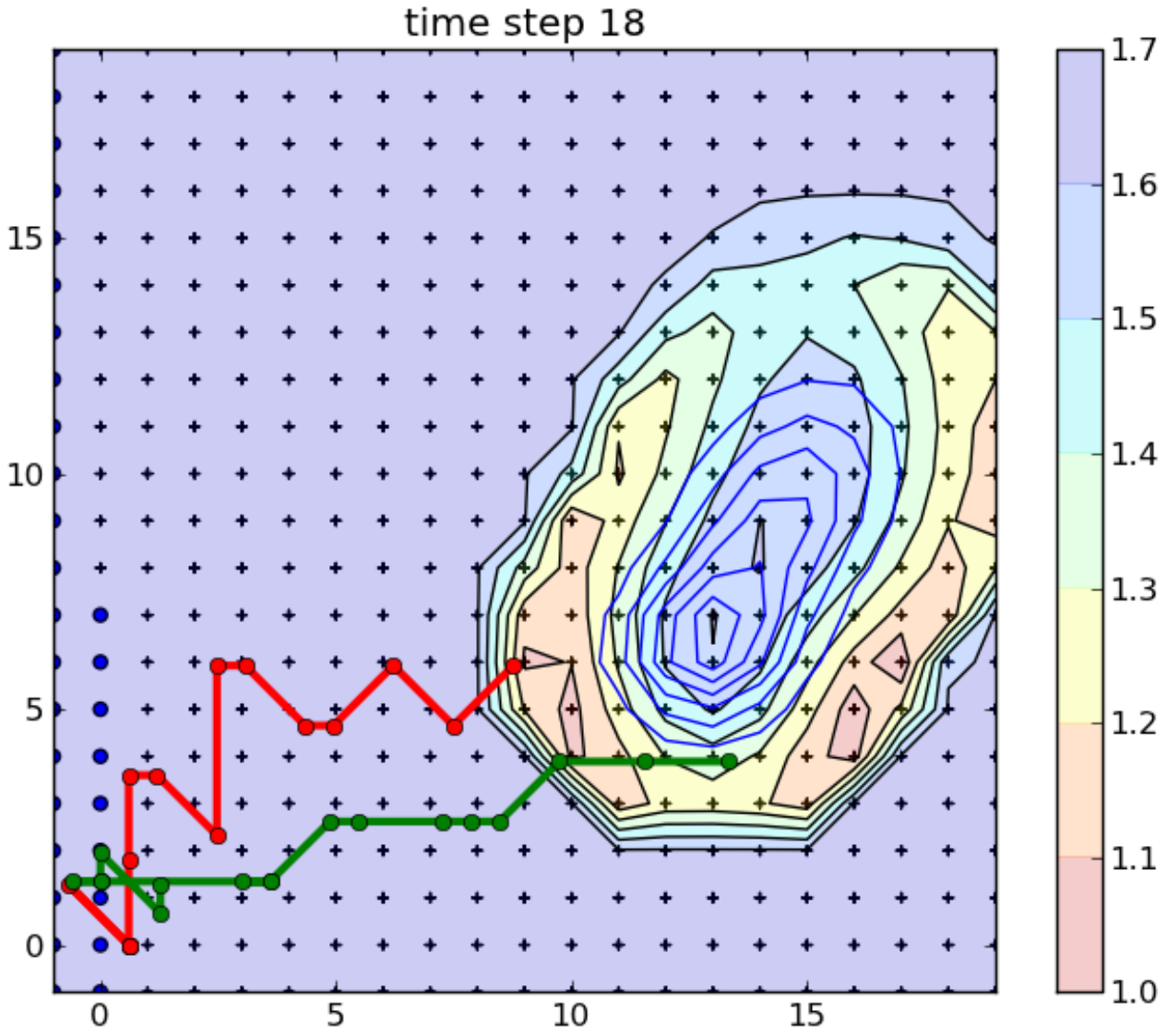
# Optimization - Simulation results - two mobile groups



# Optimization - Simulation results - two mobile groups



# Optimization - Simulation results - two mobile groups



# Misclassification loss

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

$$L(\lambda, X) = \alpha I_{false\_positive} + \beta I_{false\_negative}, \tag{17}$$

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 $j$:

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

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

# Misclassification loss

For simplicity, we may assume to split the area around the power plant into $J$ districts, each representing a constant number of inhabitants, e.g. 1000, which live approximately at given location, $\mathbf{i}_j$, $j = 1 \dots J$, which denote the points of interest. The total absorbed dose in these localities will be represented by a vector $\mathbf{c} = [C(\mathbf{i}_1), C(\mathbf{i}_2), \dots C(\mathbf{i}_J)]$.

$$I_{false\_positive} = \sum_{j=1}^{J} (\hat{\mathbf{c}}_j > \overline{C} \,\&\, \hat{\mathbf{c}}_j < \overline{C}). \tag{18}$$

$$I_{false\_negative} = \sum_{j=1}^{J} (\hat{\mathbf{c}}_j < \overline{C} \,\&\, \hat{\mathbf{c}}_j > \overline{C}). \tag{19}$$

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

$$\begin{aligned} L(a, X) &= L(\hat{C}(\lambda),) \\ \hat{\mathbf{c}}(\lambda) &= \mathrm{E}(\mathbf{c}(X, Z)|\lambda). \end{aligned} \tag{20}$$

# Misclassification loss

Expected value of the misclassification loss (17) may be computed as

$$
\begin{aligned}
\mathrm{E}(L(X,Z,\lambda)|\lambda) &= \int p(X,Z|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \int p(Z|X,\lambda)p(X|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \int p(Z|X,\lambda)p(X|\lambda)L(\hat{C}(\lambda),X)dXdZ, \\
&= \sum w^{(i)} \int p(Z|X^{(i)},\lambda)L(\hat{C}(\lambda),X^{(i)})dZ, \\
&= \sum w^{(i)} L^{(i)}(\lambda), \qquad (21) \\
L^{(i)}(\lambda) &= \mathrm{E}(L(\hat{C}(\lambda),X^{(i)})) \qquad (22)
\end{aligned}
$$

Note that (21) is a sum of contributions from each particle (22), where each contribution is an integral over $Z$.

# Misclassification loss - importance sampling

Since (22) 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 (22) is then:

$$\mathrm{E}(L(\hat{C}(\lambda), X^{(i)})) \approx \sum \frac{p(Z|X^{(i)}, \lambda)}{p(Z|X^{(i)}, \lambda)} L(\hat{C}(Z^{(k)}), X^{(i)}) = \sum_{k=1}^{K} L(\hat{C}(Z^{(k)}), X^{(i)}), \quad (23)$$

$$L(\hat{C}(Z^{(k)}), X^{(i)}) = \alpha \sum_{j=1}^{J} (\hat{\mathbf{c}}_j > \overline{C} \,\&\, \hat{\mathbf{c}}_j < \overline{C}) + \beta \sum_{j=1}^{J} (\hat{\mathbf{c}}_j < \overline{C} \,\&\, \hat{\mathbf{c}}_j > \overline{C}) \quad (24)$$

$$\hat{\mathbf{c}} = \sum_{l} C(X^{(l)}) \tilde{w}^{(l)}, \quad (25)$$

$$\tilde{w}^{(l)} = w^{(k)} \frac{p(Z^{(k)}|X^{(l)}\lambda)}{\sum_m w^{(m)} p(Z^{(k)}|X^{(m)}, \lambda)}. \quad (26)$$

Note that (26) is the same formula as in the update of the particle filter (). In this case however, the measurement $Z^{(k)}$ is fictitious.

# Misclassification loss - 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

$$\begin{aligned} \mathrm{E}(L(X,Z,\lambda)|\lambda) &= \int \sum w^{(i)} p(Z|X^{(i)},\lambda) L(\hat{C}(\lambda), X^{(i)}) dZ, \\ q(Z,i) &= \sum w^{(i)} p(Z|X^{(i)},\lambda). \end{aligned} \tag{27}$$

By drawing $K$ random couples $\{i^{(k)}, Z^{(k)}\}$ we may approximate the whole loss function by the same functions as in (23)–(26), with $i$ replaced by $i^{(k)}$.

# Misclassification loss - Gauss Hermite quadrature

Integration of expected value of a Gaussian distribution can be easily converted into the conditions of Gauss Hermite quadrature.

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

where $w_k$ and $x_k$ are Gauss Hermite coefficients.

# Misclassification loss - simulation results

We use the same scenario and the same grid as in case of entropy for single points. The values of the misclassification loss obtained by the Gauss-Hermite quadrature rule are displayed in figure 4. The points of interests $\mathbf{i}_j$ are all points on the $\lambda$ grid. To study sensitivity 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 $k_c = \{20, 1, 1/100\}$.

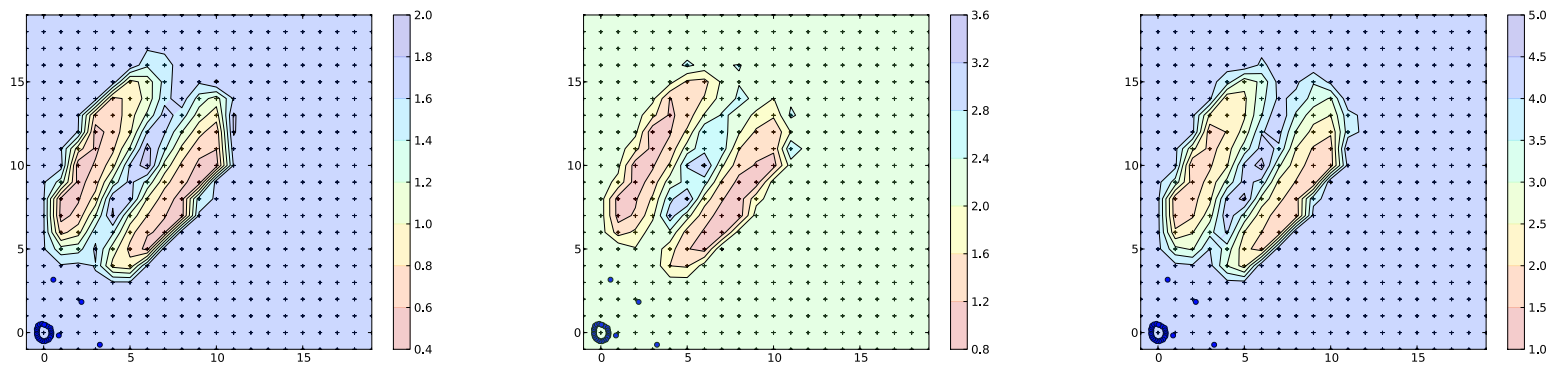


Figure: Evaluation of the misclassification loss function for various locations $\lambda$.

# Misclassification loss - simulation results

Sensitivity with respect to the selected point of interest was tested by selecting only one point in $\mathbf{i}_1$. Results are displayed in Figure 5.

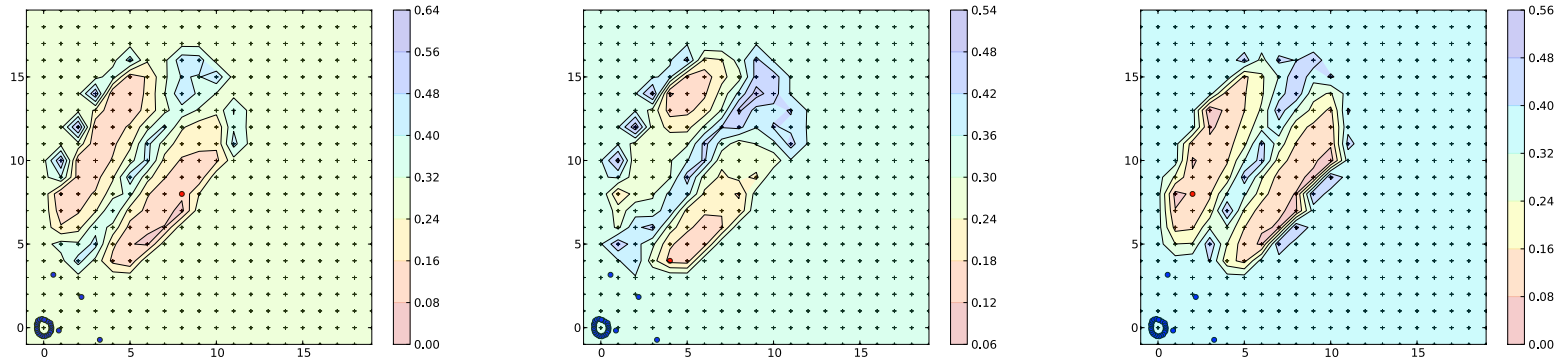



Figure: Comparison of contour plot of the misclassification loss for different points of interest. Point of interest $\mathbf{i}_1$ is denoted by red dot.

The results are very similar to those obtained for the entropy loss. However, evaluation of the misclassification loss is much more computationally demanding $\Rightarrow$ focus of efficient and parallel implementation of the code.