Exploitation of particle filter and marginalized particle filter in radiation protection

Radek Hofman

hofman@utia.cas.cz

4 Mar 2010

Presentation outline

1. Data assimilation (DA) - how does it work?
2. Stages of radiation accident
3. Application of DA in early phase of radiation accident
   - Description of problem and objectives of DA
   - Description of modeling methodology
   - Description of assimilation methodology
   - Numerical example
   - Conclusion
4. Application of DA in late phase of radiation accident
   - Description of problem and objectives of DA
   - Description of modeling methodology
   - Description of assimilation methodology
   - Numerical example
   - Conclusion
5. Overall conclusion and future work

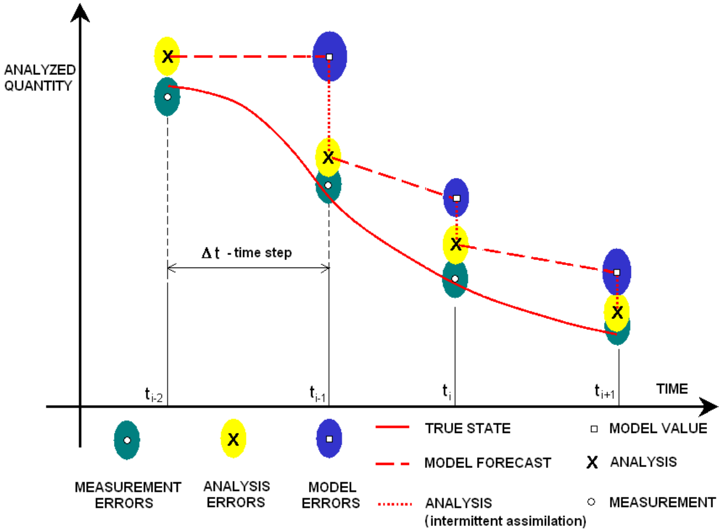
Why the data assimilation?

- The mathematical model could be wrong:
  - Wrong conceptualization of the physical reality
  - Erroneous inputs
  - Inherited errors (numerical errors, interpolation errors, ...)
  - Wrong meteorological forecast
  - ...
- The measurements could be also bad or insufficient:
  - Sparse measurement – we don't have enough information
  - Erroneous, but mostly assumed more precise than the model
  - Often undirect – need of transformation
  - ...

$\Rightarrow$ the combination of numerical model which can be evaluated in unlimited analysis points with few accurate measurements can improve our estimate (prediction)

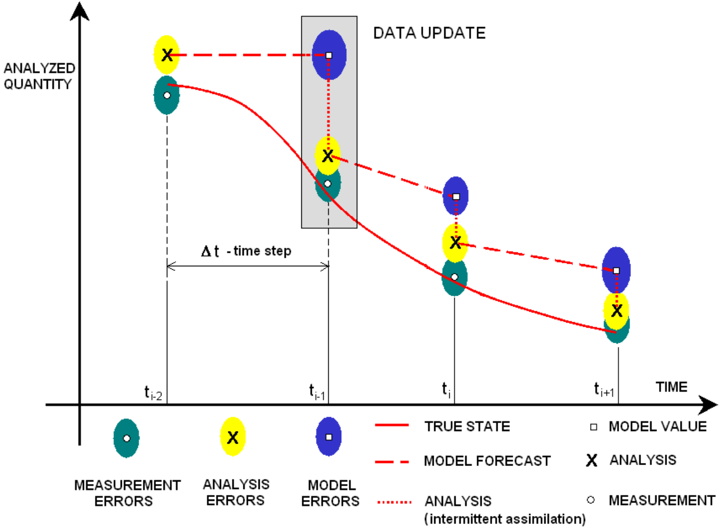
Data assimilation

Assimilation schema:



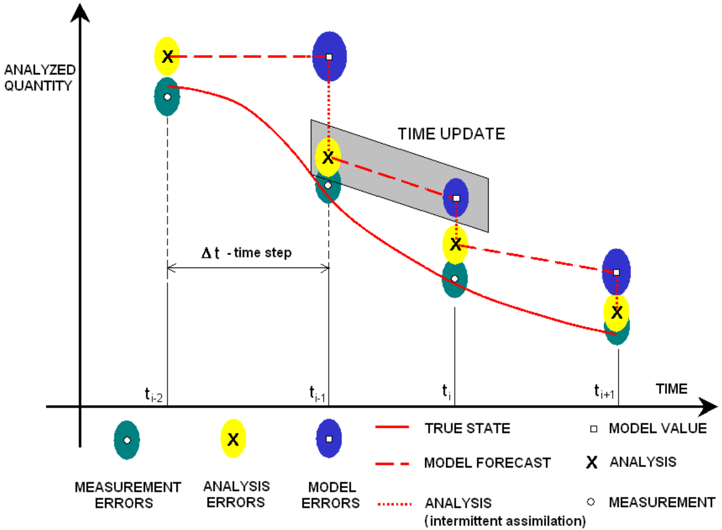
Data assimilation - *data update* step

Assimilation schema:



Data assimilation - *time update* step

Assimilation schema:



The methods of objective analysis

- Empirical methods
  - Interpolation methods
  - Cressman algorithm
  - SCM
- Bayesian methods
  - Optimal interpolation
  - Variational methods
    - 3D–Var
    - 4D–Var
  - Kalman filter
  - Sequential Monte Carlo methods – particle filters
  - Marginalized particle filter
  - Ensemble methods
  - hybrid filters
  - ...

General Bayesian approach to data assimilation

- Uncertainty is described in terms of probability density functions

The prior distribution $p(\mathbf{x}_0)$ is transformed into posterior pdf $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ using measurements $\mathbf{y}_{1:t} = \{\mathbf{y}_1, \ldots, \mathbf{y}_t\}$ by recursive repetition of the following steps:

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

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{\int p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})d\mathbf{x}_t}, \tag{2}$$

Evaluation of (1) and (2) involves integration over complex spaces and often it is computationally infeasible. Suboptimal solution can be found by the means of sequential Monte Carlo methods also known as particle filters.

Sequential Monte–Carlo methods (particle methods)

- Recursive estimation of the probability density function.
- Representation of probability density function as a set of particles and its associated weights.

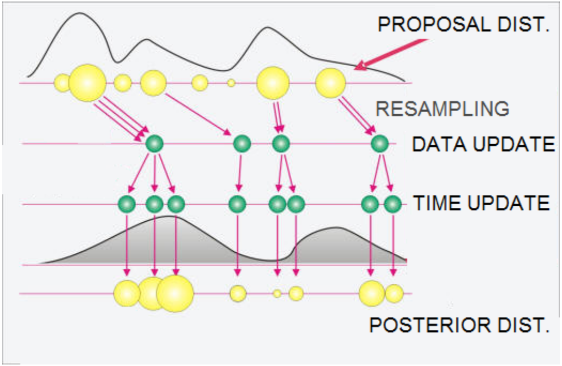
$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \sum_{i}^{M} \tilde{q}_t^{(i)} \delta(\mathbf{x}_t^n - \mathbf{x}_t^{n,(i)}), \quad \sum_{i}^{M} \tilde{q}_t^{(i)} = 1, \quad \tilde{q}_t^{(i)} \geq 0 \quad \forall i \tag{3}$$

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \{\tilde{q}_t^{(i)}, \mathbf{x}_t^{(i)}\}_{|i=1\ldots M} \tag{4}$$

We have to be able to generate random samples from complicated distributions.

Sequential Monte–Carlo methods (particle methods)

Illustration of the particle filter idea



Resampling step avoids degradacy of particle ensemble

Marginalized particle filter

Marginalized particle filter estimates the probability density function by a combination of non–parametric and parametric density functions

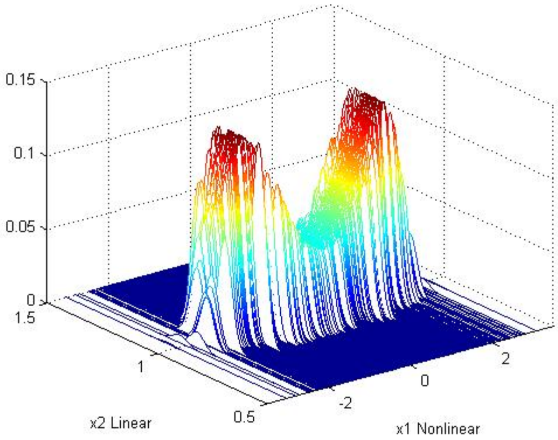
$$\mathbf{x}_t = \begin{bmatrix} \mathbf{x}_t^l \\ \mathbf{x}_t^n \end{bmatrix} \tag{5}$$

$$p(\mathbf{x}_t^l, \mathbf{x}_t^n | \mathbf{y}_{1:t}) = p(\mathbf{x}_t^l | \mathbf{x}_t^n, \mathbf{y}_{1:t}) p(\mathbf{x}_t^n | \mathbf{y}_{1:t}) \tag{6}$$

E.g. if the $\mathbf{x}_t^l$ can be treated as linear-Gaussian, the estimated density function is represented by a weighted sum of Gaussian distributions, where each particle has a Gaussian distribution attached to it

$$p(\mathbf{x}_t^l, \mathbf{x}_t^n | \mathbf{Y}_t) = \sum_{i=1}^{M} \tilde{q}_t^{(i)} \delta(\mathbf{x}_t^n - \mathbf{x}_t^{n,(i)}) N(\hat{\mathbf{x}}_t^{l,(i)}, \mathbf{P}_t^{(i)}) \tag{7}$$

Marginalized particle filter



Stages of a radiation accident

1. Pre-release phase
   - Something is going wrong with the nuclear reactor:)
2. Early phase (EP)
   - Failure of nuclear reactor followed by an aerial release of radionuclides (aerosols, noble gases)
   - After the release, there is a radioactive cloud passing over the terrain
   - EP lasts until the radioactive cloud leaves area of interest (e.g. 24hours)
   - Dominant pathways of irradiation: inhalation, cloudshine, groundshine
3. Late (post-emergency) phase (LP)
   - Radioactive material is deposited on the ground (the cloud left a "radioactive trace")
   - Dominant pathways of irradiation: groundshine, ingestion, inhalation from resuspension
   - LP lasts until the radiation levels resume to background values

Part I. - Application of DA in early phase

Early phase - problem formulation

- Assume an accident in a nuclear power plant followed by aerial release of radionuclides
- After the release, there is a radioactive cloud passing over the terrain
- The spatio-temporal distribution of radionuclides is modeled by the means of numerical dispersion models in order to determine appropriate countermeasures
- Output of such a model is a prediction of radiation situation given in terms of radiological quantities, eg. activity concentration in air $C(\mathbf{s}, t)$
- Modeling is supported by available measurements from radiation monitoring network designed to measure the γ-dose rate

Early phase - objectives of DA

We have got

- Atmospheric dispersion model C_{ADM} modeling activity concentration in air $C(\mathbf{s},t)$ in a set of grid points – vector $\mathbf{C}_t$
- Measurements of time integrated γ-dose rate at time t – vector $\mathbf{y}_t$
- ADM is a function of parameters and inputs $C_{\mathrm{ADM}} = C_{\mathrm{ADM}}(\boldsymbol{\Theta})$
- A group of most significant parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is modeled as random due to the stochastic nature of the background physics

We want to

- On-line estimate the state $\mathbf{x}_t = [\mathbf{C}_t, \boldsymbol{\theta}_t]^T$ as the cloud is passing over stationary measuring sites
- Use posterior distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ to predict future evolution of the radiation situation $p(\mathbf{x}_{t+k}|\mathbf{y}_{1:t})$

State evolution model

Evolution of the state is given by the transition pdf $p(\mathbf{x}_t|\mathbf{x}_{t-1})$:

$$\begin{aligned} p(\mathbf{x}_t|\mathbf{x}_{t-1}) &= p(\mathbf{C}_t,\, \boldsymbol{\theta}_t|\mathbf{C}_{t-1},\, \boldsymbol{\theta}_{t-1}) \\ &= p(\mathbf{C}_t|\mathbf{C}_{t-1},\, \boldsymbol{\theta}_t,\, \boldsymbol{\theta}_{t-1})p(\boldsymbol{\theta}_t|\mathbf{C}_{t-1},\, \boldsymbol{\theta}_{t-1}) \end{aligned} \quad (8)$$

Under the choice of atmospheric dispersion model $C_{\mathrm{ADM}}(\boldsymbol{\theta}_t)$ and its parameters $\boldsymbol{\theta}_t$, the evaluation of $\mathbf{C}_t$ is deterministic:

$$p(\mathbf{C}_t|\mathbf{C}_{t-1},\, \boldsymbol{\theta}_t,\, \boldsymbol{\theta}_{t-1}) = \delta(\mathbf{C}_t - C_{\mathrm{ADM}}(\boldsymbol{\theta}_t)) \quad (9)$$

Time evolution of $\boldsymbol{\theta}_t$ is given by the pdf $p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1})$. Invariant parameters $(\boldsymbol{\theta}_t = \boldsymbol{\theta}) \Rightarrow$ transition pdf is

$$p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) = \delta(\boldsymbol{\theta}_t - \boldsymbol{\theta}). \quad (10)$$

The process is initialized with prior pdf $p(\boldsymbol{\theta}_0)$.

State evolution model

We chose the Gaussian puff model (GPM) for the ADM:

$$C(\mathbf{s},\ t) = \frac{Q\, f_{\mathrm{D}}(t)\, R(t)}{(2\pi)^{\frac{3}{2}}\, \sigma_{s_1}\, \sigma_{s_2}\, \sigma_{s_3}} \exp\left\{ -\frac{1}{2} \left[\left(\frac{s_1 - ut}{\sigma_{s_1}} \right)^2 + \left(\frac{s_2}{\sigma_{s_2}} \right)^2 + \left(\frac{s_3}{\sigma_{s_3}} \right)^2 \right] \right\}, \tag{11}$$

where

- t is time index, Q is the total released activity in Bq and u is the wind speed
- Dispersion coefficients $\{\sigma_{s_i}\}|_{i=1,\,2,\,3}$ are functions of distance from the source
- Factor $f_{\mathrm{D}}(t)$ stands for radioactive decay, dry and wet deposition
- Term $R(t)$ accounts for homogenization of the vertical profile of concentration due to the reflections from the top of mixing layer and the ground

It is based on approximative solution of the three dimensional advection-diffusion equation

$$u\frac{\partial C}{\partial s_1} = \frac{\partial}{\partial s_1}\left(K_1 \frac{\partial C}{\partial s_1}\right) +, \frac{\partial}{\partial s_2}\left(K_2 \frac{\partial C}{\partial s_2}\right) + \frac{\partial}{\partial s_3}\left(K_3 \frac{\partial C}{\partial s_3}\right). \tag{12}$$

Measurement model

Measurements are assumed to be normally distributed and mutually independent given the state $\mathbf{x}_t$. Errors of measurements are set proportional to the their values with an offset term modeling the background radiation superposed to the actual dose measurements

$$\mathbf{y}_t \sim \mathcal{N}(\mathbf{D}_t, \boldsymbol{\Sigma}(\mathbf{D}_t)), \tag{13}$$

where $\mathbf{D}_t$ is a vector of measurements of *time integrated absorbed* γ*-dose in tissue* in all the measuring sites available in time t.

$$D_{i,t} = \int_{t-1}^{t} \sum_j \frac{K_j \, \mu_{a,j} \, E_{\gamma,j}}{\rho} \Phi_j(C(\mathbf{s}_{(i)}, \tau)) \, d\tau. \tag{14}$$

- $\{E_{\gamma,j}; j \in I_E\}$ is a set energy levels in a mixture of radionuclides
- Φ_j is effective dose of gamma rays
- $mu_{a,j}$ absorption coefficient
- K_j conversion coefficient
- ρ is the mass density of air
- summation is over assumed energy levels

Φ at a receptor located at $\tilde{\mathbf{s}} = (\tilde{s}_1,\, \tilde{s}_2,\, \tilde{s}_3)$ from a source of energy E_γ dispersed in air is

$$\Phi(\tilde{s}_1,\, \tilde{s}_2,\, \tilde{s}_3,\, E_\gamma) = \iiint \frac{f(E_\gamma) B(E_\gamma, \mu r) C(s_1,\, s_2,\, s_3)}{4\pi\, r^2}\, ds_1\, ds_2\, ds_3, \qquad (15)$$

- $r^2 = (\tilde{s}_1 - s_1)^2 + (\tilde{s}_2 - s_2)^2 + (\tilde{s}_3 - s_3)^2$
- $f(E_\gamma)$ is the branching ratio to the specific energy
- μ is the attenuation coefficient of air
- $B(E_\gamma, \mu r)$ is the dose build-up factor given by Bergers analytical expression

$$B(E_\gamma, \mu\, r) = 1 + a\, \mu r\, \exp(b\, \mu r), \qquad (16)$$

- Coefficients μ, a and b depend on E_γ.
- Energy dependent absorption coefficient μ_a is calculated as

$$\mu_a = \mu \Big/ \left[1 + \frac{a}{(1-b)^2} \right]. \qquad (17)$$

- The form and simplicity of used Gaussian puff model (11) allows for numerical evaluation of integral (15) on a compact support where the concentration is not negligible
- If the radioactive plume is large compared to the mean free path of the γ-rays, then the semi-infinite cloud approximation of effective flux can be successfully used
- Integration is performed numerically by the means of Gaussian quadrature method

Parametrization of atmospheric dispersion model

A group of the most significant variables affecting the dispersion process (including meteorological inputs) was selected using available sensitivity and uncertainty studies performed on Gaussian dispersion models.

variable	physical effect	parametrization
Q	magnitude of release	$Q = \omega_t \, Q_0$
u	wind speed	$u = (1 + 0.1\,\xi_t)\,u_0 + 0.5\,\xi_t$
ϕ	wind direction	$\phi = \phi_0 + \Delta\phi,\ \Delta\phi = \psi_t\,(2\pi/80)$ rad.
$\sigma_{s_i}\big\|_{i=1,\,2}$	horizontal dispersion	$\sigma_{s_i} = \zeta_t\,\sigma_{s_{i0}}\big\|_{i=1,\,2}$

Table: Parametrization of selected variables and inputs to the ADM.

- All the random parameters are treated as time invariant: $\boldsymbol{\theta}_t(\omega_t, \xi_t, \psi_t, \zeta_t) = \boldsymbol{\theta}(\omega, \xi, \psi, \zeta)$.

In case of time horizon of several hours, the assumption of stationarity of the meteorological condition vanishes. Parametrization of the meteorological data has to be fragmented into shorter time intervals (usually hourly intervals) where the assumption of stationarity holds.

Numerical experiment

For purposes of numerical experiment was chosen assimilation scenario with an instantaneous release of ^{41}Ar:

- half life of decay of ^{41}Ar is 109.34 minutes $\Rightarrow$ the radioactive decay cannot be neglected
- γ radiation produced on a few energy levels
- we assume just the energy level $1293.57 keV$ with the branching ratio 99.1%.
- the rest being included in the 0.9% is neglected
- noble gas $\Rightarrow$ no deposition $\Rightarrow$ no ground shine $\Rightarrow$ measured dose is just from the cloud

Numerical experiment - measurements

- The measured quantity is the time integrated γ-dose from cloud shine of ^{41}Ar.
- The topology of measuring sites is similar to that of the Early Warning Network of the Czech Republic
- Nuclear power plant is surrounded by almost fifty stationary measuring sites capable to measure time integrated γ-dose

- Time horizon spans up to the 1 hour
- The data update step is performed every 10 minutes.

$\Rightarrow$ we performed 6 assimilation cycles consisting of time and data update steps

Numerical experiment - twin model

Numerical experiment is conducted as a twin experiment, where the measurements are simulated via a twin model and perturbed. The set of parameters $\boldsymbol{\theta}_{\mathrm{TWIN}}$ used for evaluation of the twin model simulating measurements is

$$\boldsymbol{\theta}_{\mathrm{TWIN}} = (0.72, -0.17, -8.3, 1.3). \tag{18}$$

variable	prior val.	param. value	true value
Q – released activity	1.0E+10Bq	0.72	7.2E+09Bq
u – wind speed	3.10m/s	-0.17	2.96m/s
ϕ – wind direction	310.0deg	-8.3	272.7deg
σ_{s_i} – horizontal disp.	$\sigma_{s_i} = \sigma_{s_i}(dist)\vert_{i=1,2}$	1.3	$\sigma_{s_i} = 1.3\,\sigma_{s_i}\vert_{i=1,2}$

Table: Values of variables of the initial model setting and the twin model.

Numerical experiment

Assimilation results are presented in the form of expected value of TIC with respect to the predictive densities at different time steps

$$TIC(\mathbf{s}) = \int_0^{t_{\mathrm{MAX}}} C(\mathbf{s}, \tau)\, d\tau \qquad (19)$$

on a rectangular grid of dimension 41×41 grid points with the grid step $1km$. The source of pollution is placed in the center of the grid.

parameter	**physical effect**	**pdf type**	**mean value**	**std. dev.**
ω_t	magnitude of release	log-normal	1.0	1.0 (3σ trunc.)
ξ_t	wind speed	uniform	0.0	1.0
ψ_t	wind direction	uniform	0.0	10.0
ζ_t	horizontal dispersion	log-normal	1.0	1.0 (3σ trunc.)

Table: Prior distributions of estimated parameters $\boldsymbol{\theta}_t = (\omega_t, \xi_t, \psi_t, \zeta_t)$.

Numerical experiment

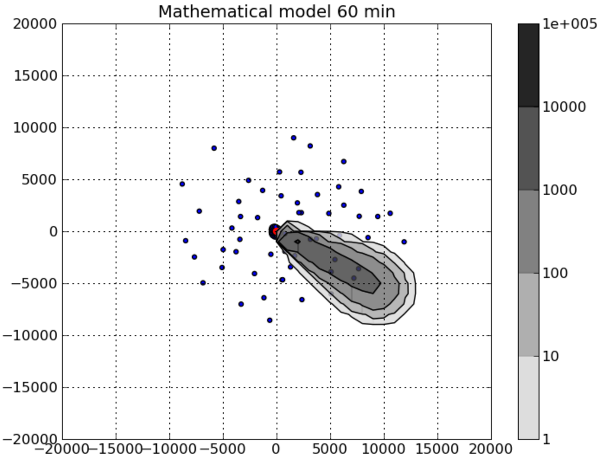


Figure: TIC evaluated by the atmospheric dispersion model without the data assimilation and with initial setting of variables $Q = Q_0$, $u = u_0$, $\phi = \phi_0$ and $\sigma_{s_i}|_{i=1,2} = \sigma_{s_{i0}}|_{i=1,2}$

Numerical experiment

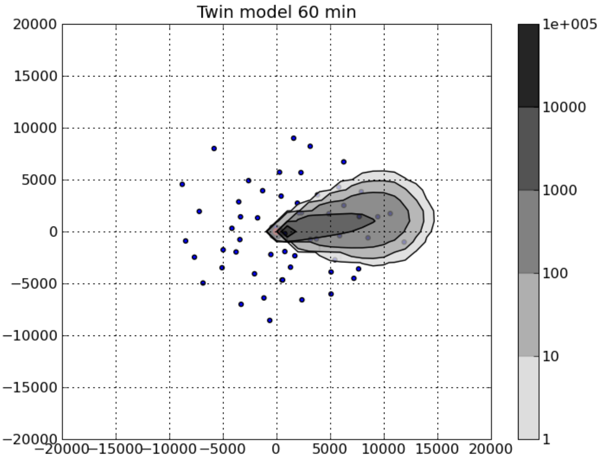


Figure: TIC evaluated by the twin model used for simulation of measurements

Numerical experiment

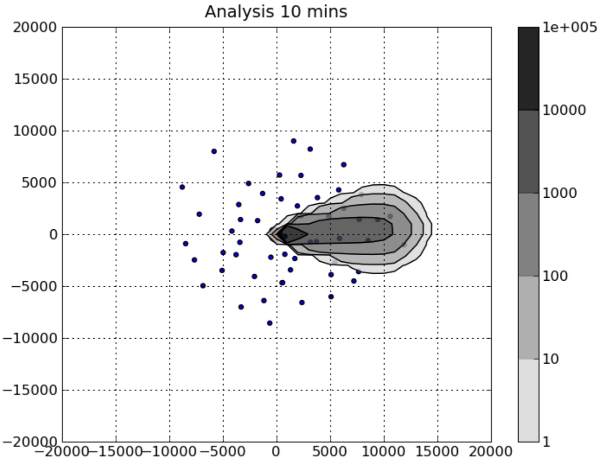


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_1$

Numerical experiment

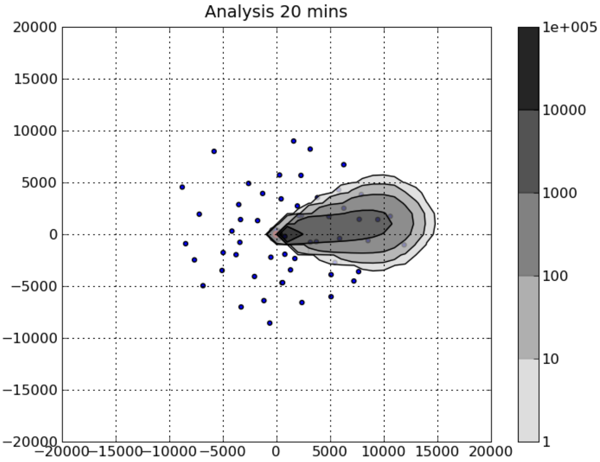


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_{1:2}$

Numerical experiment

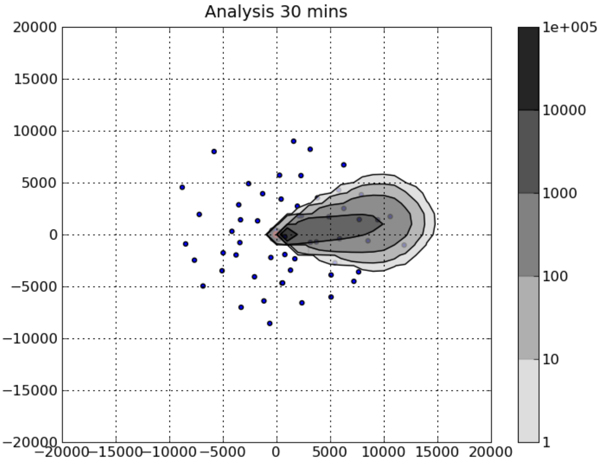


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_{1:3}$

Numerical experiment

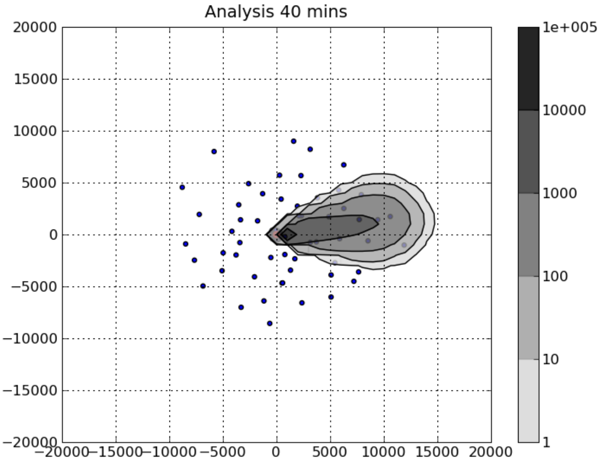


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_{1:4}$

Numerical experiment

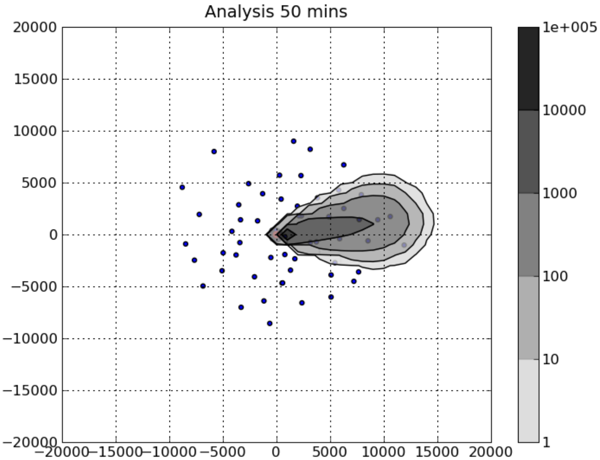


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_{1:5}$

Numerical experiment

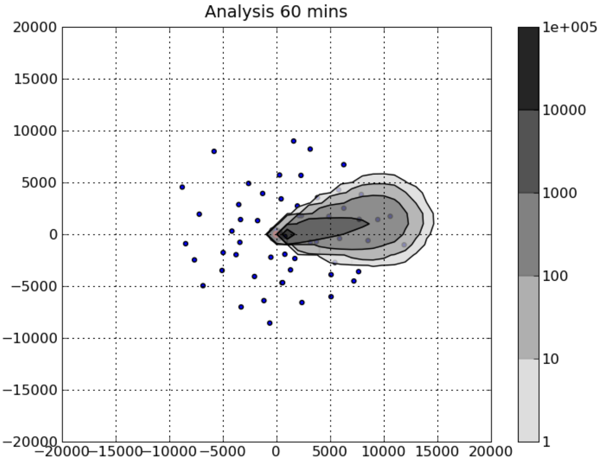


Figure: Expected value of prediction of TIC based on measurements $\mathbf{y}_{1:6}$

Numerical experiment

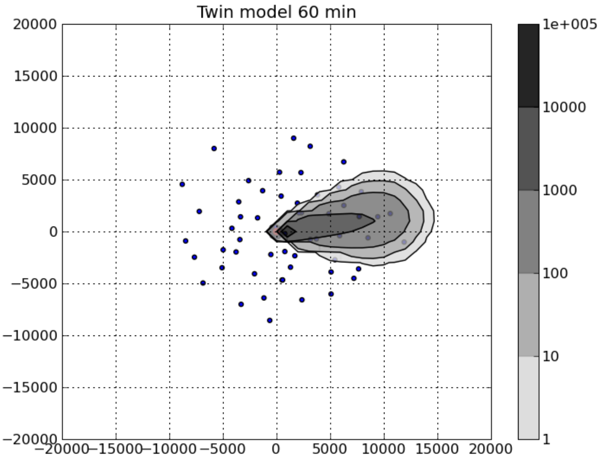


Figure: TIC evaluated by the twin model used for simulation of measurements

Early phase: Conclusion and future work

The presented scenario clearly illustrates the power of the method:

- Rapid assessment of the situation in case of an aerial release of radionuclides is crucial for planning of countermeasures
- Bayesian approach allows joint estimation of spatio-temporal distribution of activity and parameters of the dispersion model
- Assimilated estimate of the radiation situation on the terrain can be easily extended to predictions on an arbitrary horizon
- However, a lot of work is required to incorporate the method to the existing decision support systems

Future work is development of more realistic models of the state evolution and the measurements: mixture of radionuclides, extended set of uncertain variables, assumption ground-shine dose etc.

Part II. - Application of DA in late phase

Late phase - problem formulation

- The radioactive plume moving over terrain leaves a trace due to the deposition processes (dry and wet deposition)
- The deposited materials cause irradiation via groundshine and ingestion (radionuclides migrate via root system of plants into edible parts and causes internal irradiation to people and livestock)
- The knowledge of groundshine evolution is important in planning of long–term countermeasures
- We focus on ^{137}Cs (half–time of decay 30 years, detrimental long–term effect on health of population)
- Temporal distribution of spatially localized radionuclides is modeled by the means of numerical models
- Modeling is supported by measurements from mobile groups or from airborne γ spectrometry

Late phase - objectives of DA

We have got

- Parametrized model of time evolution of contamination on a computational grid - vector $\mathbf{d}_t$, some of its parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ are treated as random
- Measurements of γ-dose rate at time t – vector $\mathbf{y}_t$

We want to

- Spatial localization of the deposited material
- Estimate the state $\mathbf{x}_t = [\mathbf{d}_t, \boldsymbol{\theta}_t]^T$ from historical measurements $\mathbf{y}_{1:t}$
- Use posterior distribution $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ to predict future evolution of the radiation situation $p(\mathbf{x}_{t+k}|\mathbf{y}_{1:t})$

Ground exposure model

$$SD(k,\ t) = SD(k,\ t=0) \cdot R(t) \cdot E(t) \tag{20}$$

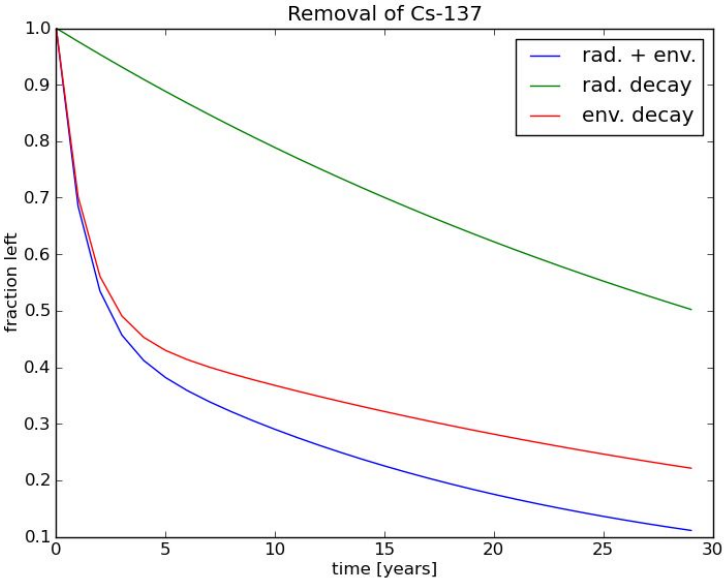
$$R(t) = exp(-\ln 2 \cdot \frac{t}{T_y}) \tag{21}$$

$$E(t) = d_f \cdot \exp(-\ln 2 \cdot \frac{t}{T_f}) + d_s \cdot \exp(-\ln 2 \cdot \frac{t}{T_s}) \tag{22}$$

$$D(k,\ t) = SD(k,\ t)DF \tag{23}$$

$D(k,\ t)$	dose rate on day t after deposition of a radionuclide [$Sv\ s^{-1}$]
$SD(k,t)$	deposition of the radionuclide in location k at time t [$Bq\ m^{-2}$]
$R(t)$	factor to account for radioactive dacay occuring between the deposition and t
$E(t)$	factor to account for the environmental decay of groundshine
DF	dose-rate conversion factor for groundshine [$Sv\ s^{-1}\ perBq\ m^{-2}$]
T_y	half-life of radioactive decay [s]
d_f, d_s	fractions of fast and slow decay terms, respectively ($d_f + d_s = 1$)
T_f, T_s	half-life for fast and slow components, respectively [s]

Modeling in the late phase



State space formulation

$\mathbf{d}_t$ - column vector of $SD(k,\ t)\ \forall k \in K$, where K is set of all the spatial locations on our computational grid

$$\mathbf{d}_t = \frac{E(t)R(t)}{E(t-1)R(t-1)}\mathbf{d}_{t-1} \tag{24}$$

$$\mathbf{d}_t = \mathbf{M}_t\mathbf{d}_{t-1}, \text{ where } \mathbf{M} = \tfrac{E(t)R(t)}{E(t-1)R(t-1)}\mathbf{I}$$

$\Rightarrow$ vector $\mathbf{d}_t$ can be evolved by a linear model as a Markovian process

Let $\mathbf{d}_t^{\mathrm{TRUE}}$ is the true unobserved state:

$$\mathbf{d}_t^{\mathrm{TRUE}} = \mathbf{M}_t\mathbf{d}_t^{\mathrm{TRUE}} + \boldsymbol{\xi}_t, \text{ where } \boldsymbol{\xi}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_t)$$
$$\mathbf{y}_t = \mathbf{H}_t\mathbf{d}_t^{\mathrm{TRUE}} + \boldsymbol{\epsilon}_t, \text{ where } \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$$

Kalman filter

If the initial state $\mathbf{d}_0^{\mathrm{TRUE}}$, and the noise vectors ξ_t and ϵ_t are all assumed to be mutually independent, we can use the Kalman filter to do the assimilation.

$$\mathbf{d}_t^f \equiv \mathbf{d}_{t|t-1}, \quad \mathbf{P}_t^f \equiv \mathbf{P}_{t|t-1} \text{ - } \textit{forecast}$$
$$\mathbf{d}_t^a \equiv \mathbf{d}_{t|t}, \quad \mathbf{P}_t^a \equiv \mathbf{P}_{t|t} \text{ - } \textit{analysis}$$

$$\mathbf{d}_t^f = \mathbf{M}_t \mathbf{d}_{t-1}^a \tag{25}$$

$$\mathbf{P}_t^f = \mathbf{M}_t \mathbf{P}_{t-1}^a \mathbf{M}_t^T + \mathbf{Q}_t \tag{26}$$

$$\mathbf{K}_t = \mathbf{P}_t^f \mathbf{H}^T [\mathbf{H}\mathbf{P}_t^f \mathbf{H}^T + \mathbf{R}_t]^{-1} \tag{27}$$

$$\mathbf{d}_t^a = \mathbf{d}_t^f + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}\mathbf{d}_t^f) \tag{28}$$

$$\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}_t^f(\mathbf{I} - \mathbf{K}_t\mathbf{H})^T + \mathbf{K}_t\mathbf{R}_t\mathbf{K}_t^T (\mathbf{P}_t^a = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}_t^f) \tag{29}$$

$\mathbf{H}$ - observation operator, cov. mats. $\mathbf{Q}_t$ and $\mathbf{R}_t$ must be known!!!

How to estimate model error?

If the model error covariance matrix in not known (our case), we can:

1. Neglect the model error and treat the model as perfect - this leads to *UNDERESTIMATION* of the model error and thus discrimination of the information provided by measurements

$$\mathbf{P}_t^f = \mathbf{M}_t \mathbf{P}_{t-1}^a \mathbf{M}_t^T + \mathbf{Q}_t \rightarrow \mathbf{P}_t^f = \mathbf{M}_t \mathbf{P}_{t-1}^a \mathbf{M}_t^T$$

2. In case of multiple models being available we can estimate the error somehow (e.g. use Bayesian model averaging)
3. Use KF and a parametrized form of model error covariance matrix and estimate its parameters from measurements
4. Use Ensemble Kalman filter (EnKF)
5. ...

Ensemble KF

ensemble: $\mathbf{d}^f_{t(1)}, \mathbf{d}^f_{t(2)}, \dots, \mathbf{d}^f_{t(M)}$

$$\mathbf{d}^f_{t(k)} = \mathbf{M}\mathbf{d}^a_{t-1(k)} \tag{30}$$

$$\mathbf{P}^f \approx \frac{1}{M-1}\sum_{k=1}^{M}(\mathbf{d}^f_{t(k)} - \bar{\mathbf{d}}^f_t)(\mathbf{d}^f_{t(k)} - \bar{\mathbf{d}}^f_t)^T, \quad \bar{\mathbf{d}}^f_t = \frac{1}{M}\sum_{k=1}^{M}\mathbf{d}^f_{t(k)} \tag{31}$$

$$\mathbf{K}_t = \mathbf{P}^f_t\mathbf{H}^T[\mathbf{H}\mathbf{P}^f_t\mathbf{H}^T + \mathbf{R}_t]^{-1} \tag{32}$$

Random perturbations are added to the observations to obtain observations $\mathbf{y}_{t(1)}, \mathbf{y}_{t(2)}, \dots, \mathbf{y}_{t(M)}$ for each independent cycle

$$\mathbf{d}^a_{t(k)} = \mathbf{d}^f_{t(k)} + \mathbf{K}_t(\mathbf{y}_{t(k)} - \mathbf{H}\mathbf{d}^f_{t(k)}) \tag{33}$$

It has been proven that an observational ensemble is required for update otherwise $\mathbf{P}^a_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H})\mathbf{P}^f_t$ is not satisfied

Alternatives: Ensemble Square Root KF (avoids measurements perturbation), Local Ensemble Transform KF, ...

Inflation factor

Ensemble variance can be still underestimated because of the sampling errors and insufficient number of ensemble members. The inflation factor is used to replace ensemble members according to:

$$\forall k \in M: \; \mathbf{d}_{t(k)} = \rho(\mathbf{d}_{t(k)} - \bar{\mathbf{d}}_t^f) + \bar{\mathbf{d}}_t^f,$$

where ρ tunes the magnitude of inflation

Formulation of assimilation scenario

- We want to estimate joint pdf of the state $\mathbf{x}_t = [\mathbf{d}_t, d_s, \rho]$ comprised of the deposition vector $\mathbf{d}_t$ (groundshine dose), environmental decay parameter d_s and the inflation factor ρ
- We exploit the fact that $\mathbf{d}_t$ can be treated as linear-Gaussian given the model parameter d_{s_t} and inflation factor ρ_t and thus can be marginalized out of the state $\mathbf{x}_t$

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = p(\mathbf{d}_t, d_s, \rho|\mathbf{y}_{1:t}) = p(\mathbf{d}_t|d_s, \rho, \mathbf{y}_{1:t})p(d_s, \rho|\mathbf{y}_{1:t}) \quad (34)$$

$$\underbrace{p(\mathbf{d}_t, d_s, \rho|\mathbf{y}_{1:t})}_{MPF} = \underbrace{p(\mathbf{d}_t|d_s, \rho, \mathbf{y}_{1:t})}_{EnKF}\underbrace{p(d_s, \rho|\mathbf{y}_{1:t})}_{PF} \quad (35)$$

How to estimate parameters $\boldsymbol{\theta} = (d_s, \rho)$?

The observations represent the only source of information about forecast error. When the measurements are available, we can evaluate *observed–minus–forecasted residuals*:

$$\mathbf{v}_t = \mathbf{y}_t - \mathbf{H}_t \bar{\mathbf{d}}_t^f \tag{36}$$

$$E[\mathbf{v}_t \mathbf{v}_t^T] = \mathbf{H}_t \mathbf{P}_t^f \mathbf{H}_t^T + \mathbf{R}_t = \mathbf{S}_t \tag{37}$$

- Parametrization of model error by $\boldsymbol{\theta}_t$ implies parametrization of the residual error $\mathbf{S}_t = \mathbf{S}_t(\boldsymbol{\theta}_\mathbf{t})$

Estimation of model error covariance

- We assume $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{S}_t(\boldsymbol{\theta}_t))$
- We want to assign weights to all $\boldsymbol{\theta}_t$ according to available observations in each time step

The weights are given by the likelihood function:

$$p(\mathbf{v}_t|\boldsymbol{\theta}_t) = (2\pi)^{-\frac{n}{2}}(\det \mathbf{S}_t)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{v}_t^T \mathbf{S}_t^{-1} \mathbf{v}_t)\right] \qquad (38)$$

Maximizing (38) is equivalent to maximizing $f(\boldsymbol{\theta}_t)$ - the logarithm of (38)

$$f(\boldsymbol{\theta}_t) = -\log\det \mathbf{S}_t - \frac{1}{2}(\mathbf{v}_t^T \mathbf{S}_t^{-1} \mathbf{v}_t) \qquad (39)$$

For given $\boldsymbol{\theta}_t$, let $\mathbf{S}(\boldsymbol{\theta}_t) = \mathbf{G}\mathbf{G}^T$ and $\mathbf{s} = \mathbf{G}^{-1}\mathbf{v}$, where $\mathbf{G}$ is lower triangular Cholesky factor of $\mathbf{S}$:

$$f(\boldsymbol{\theta}_t) = -\log \prod_{i=1}^{n} g_{ii}^2 - ||\mathbf{s}||_2^2$$

Details of assimilation scenario

- Computational grid is a subset of polar grid, total number of analyzed points is 520
- Measurements comprise a rectangular grid with the grid step 1.5km, total number of measurement locations is 136
- Time step is 30 days
- Observation operator is a nearest neighbor interpolation operator
- The value of d_s used for simulation of measurements was set to 0.56
- Each EnKF attached to a particle has 30 members and the total number of particles is 50 $\Rightarrow$ we run 50 EnKFs for different combinations of (d_s, ρ) and evaluate its weights upon agreement with measurements

Simulation of measurements

- Simulated measurements are sampled from a twin model

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \boldsymbol{\epsilon}_t; \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \mathbf{R}_t)$$

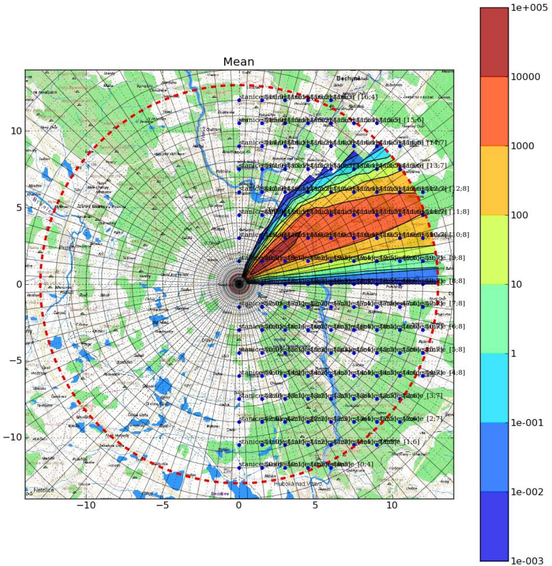
$$\mathbf{R}_t = diag(\epsilon_{t;1}, \epsilon_{t;2}, \ldots, \epsilon_{t;p})$$

$$\epsilon_{t;i} = C\mathbf{H}_i\mathbf{x}_t + K$$

$$\mathbf{x}_t = \mathbf{M}_t\mathbf{x}_{t-1}$$

- C is the coefficient of proportionality and K is the offset term simulating background radiation
- Measurements are assumed mutually independent

Twin model used for simulation of measurements



Selection of the prior ensemble

Initial uncertainty must be embodied into the prior distribution given by the ensemble - *background field*

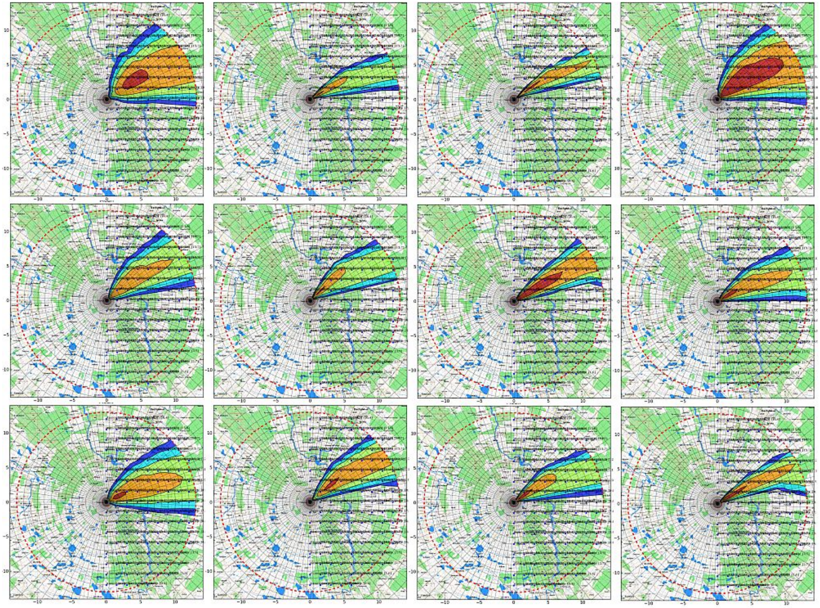
- Spatial localization of the ensemble members should be in accordance with the physical reality
- In the variability of the prior ensemble should be embodied all the uncertainty regarding the release scenario (wind speed and direction, magnitude of release etc.)

Selection of the prior ensemble

- The initial condition in the late phase are given by the radiation situation on terrain at the end of the early phase $\Rightarrow$ we can employ atmospheric dispersion model to determine initial conditions for the late phase
- Parameters of the atmospheric dispersion model

ADM_1 – intenzity of release
ADM_2 – horizontal dispersion
ADM_3 – horizontal fluctuation of wind dir.
ADM_4 – dry deposition of elem.
ADM_5 – dry depozition of aero.
ADM_6 – elution of elem. iodine
ADM_7 – elution of aero.
ADM_8 – advection speed of plume
ADM_9 – wind profile
ADM_{10} – vertical dispersion
ADM_{11} – mixing layer height
ADM_{12} – heat flux
ADM_{13} – precipitation intensity

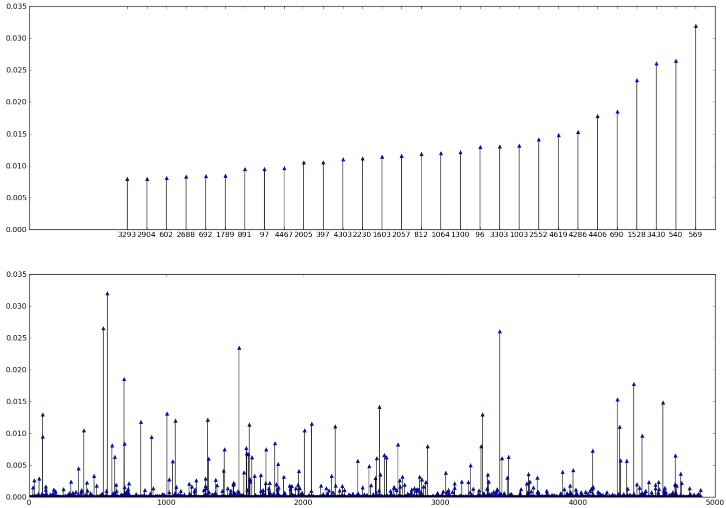
Selection of the prior ensemble



Selection of the prior ensemble

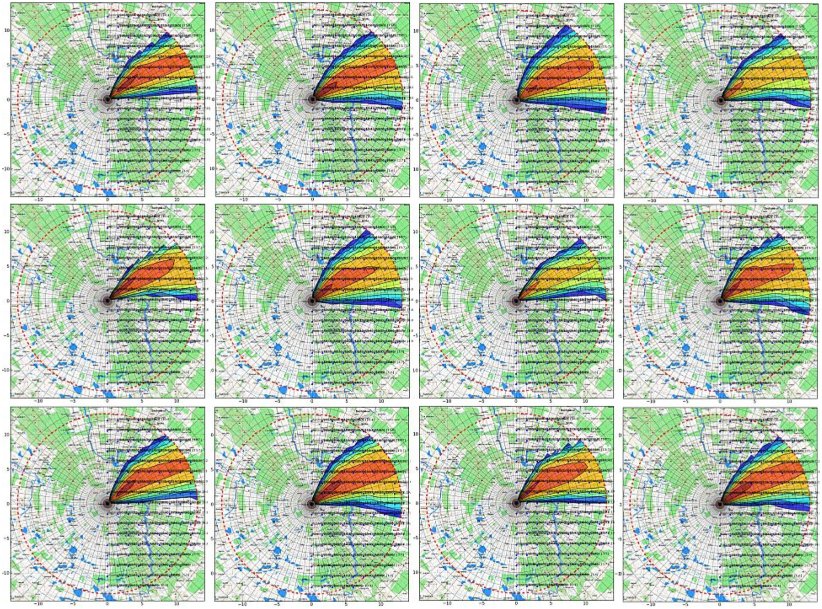
- In the variability of the prior pre-ensemble is embodied all the uncertainty regarding the release scenario (wind speed and direction, magnitude of release etc.)
- Twin model is not included in the prior pre-ensemble
- From the initial pre-ensemble of size 5000 we select members which are in best accordance with the measurements in time $t = 0$
- This performs spatial localization of the release given and accounts for the uncertainty in the parameters of the release scenario

Selection of the prior ensemble



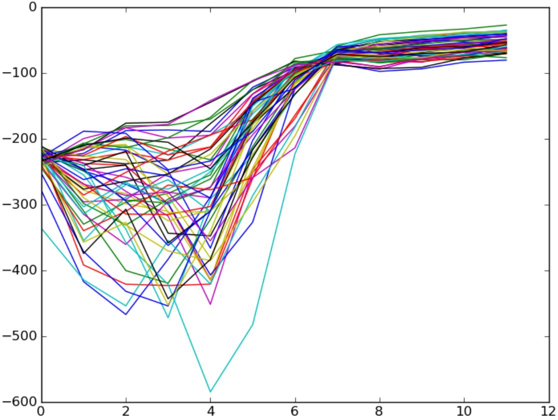
Weights of pre-ensemble members

Selection of the prior ensemble

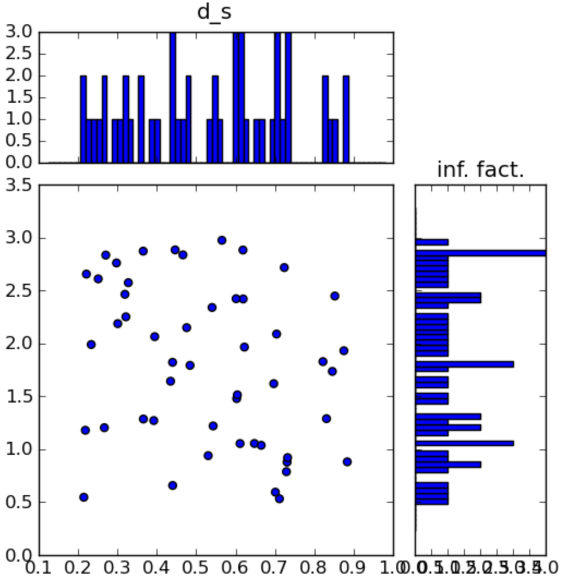


Late phase - numerical experiment

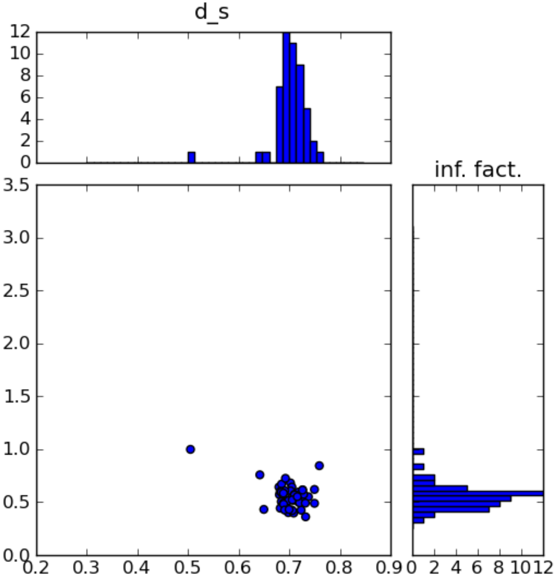
Weights $f(\boldsymbol{\theta}_t) = -\log\det \mathbf{S}_t - \frac{1}{2}(\mathbf{v}_t^T \mathbf{S}_t^{-1} \mathbf{v}_t)$ of all particle filter members in all time steps. Each single line denotes one Ensemble KF.



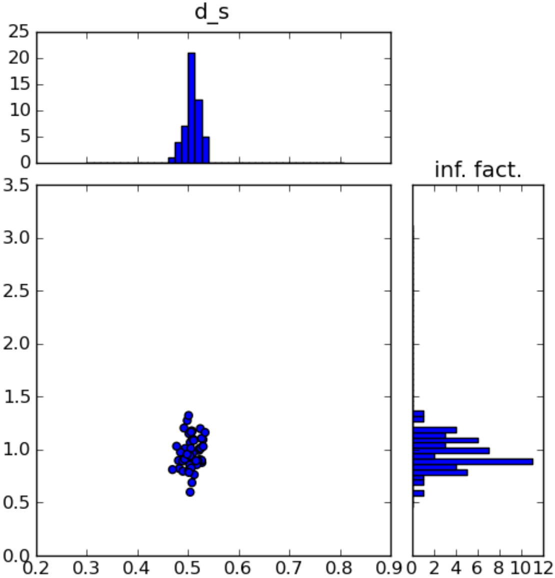
Late phase - numerical experiment



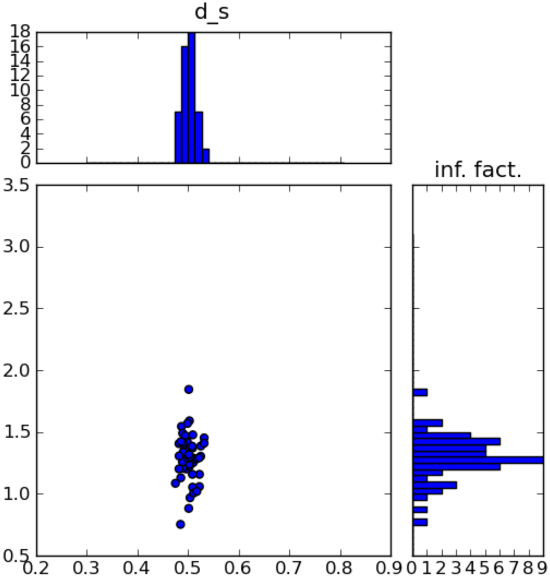
Late phase - numerical experiment



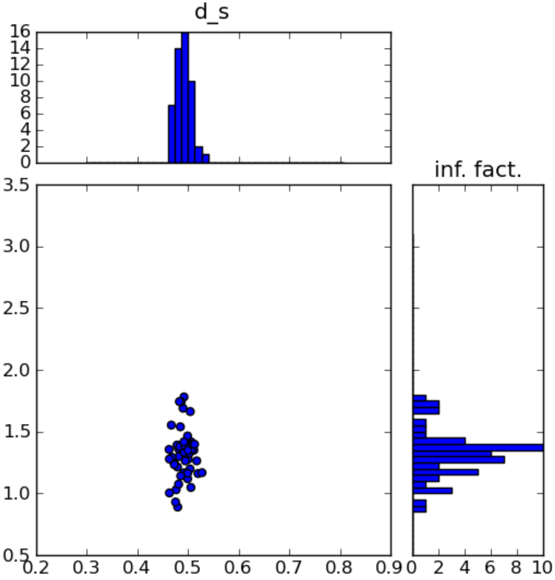
Late phase - numerical experiment



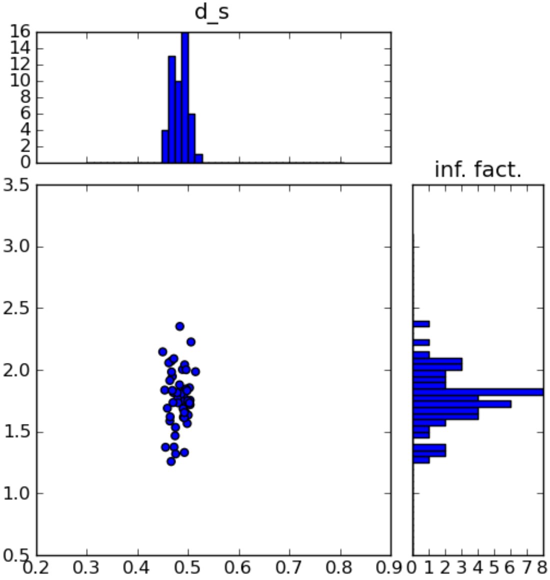
Late phase - numerical experiment



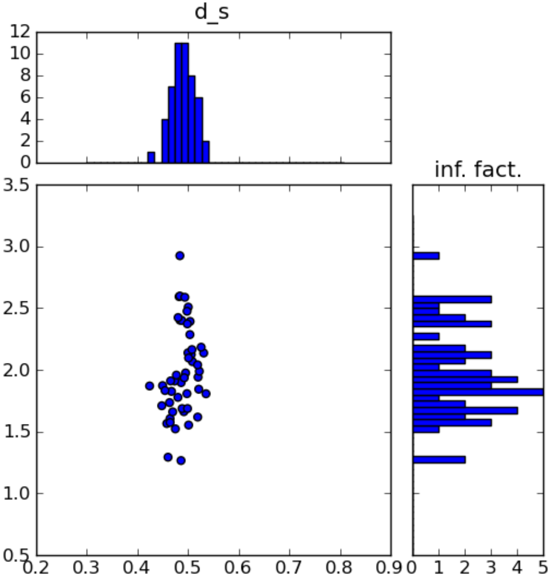
Late phase - numerical experiment



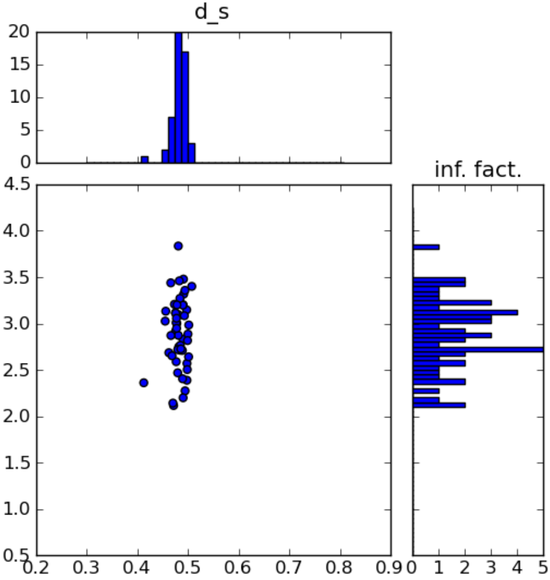
Late phase - numerical experiment



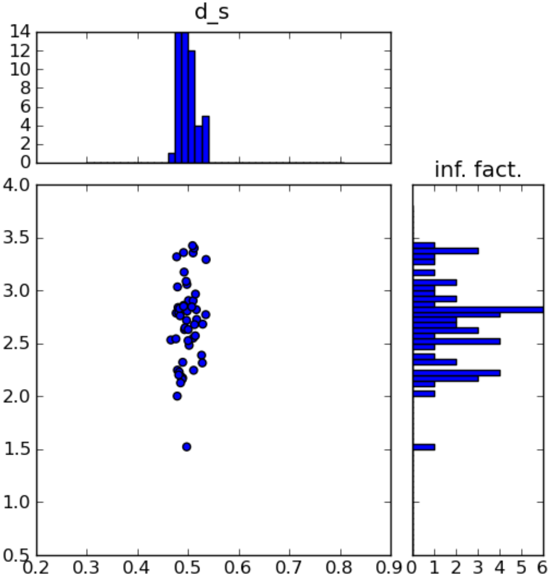
Late phase - numerical experiment



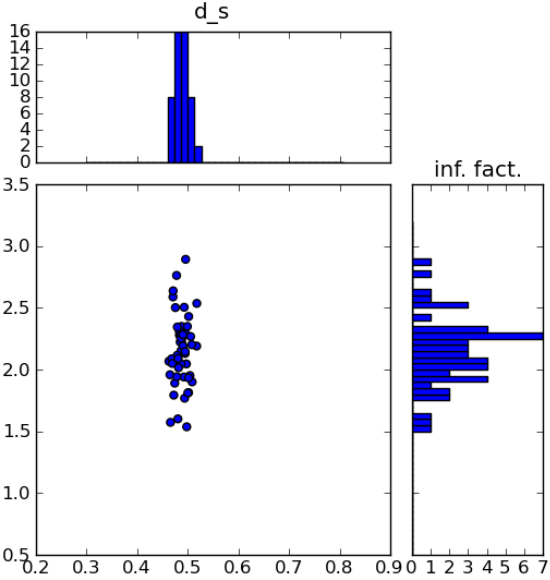
Late phase - numerical experiment



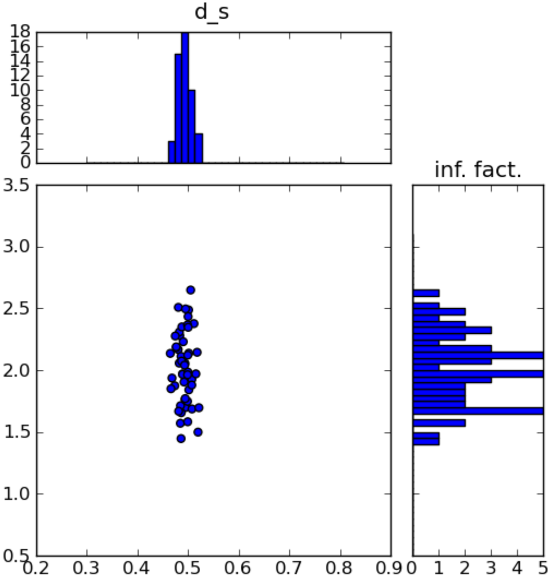
Late phase - numerical experiment



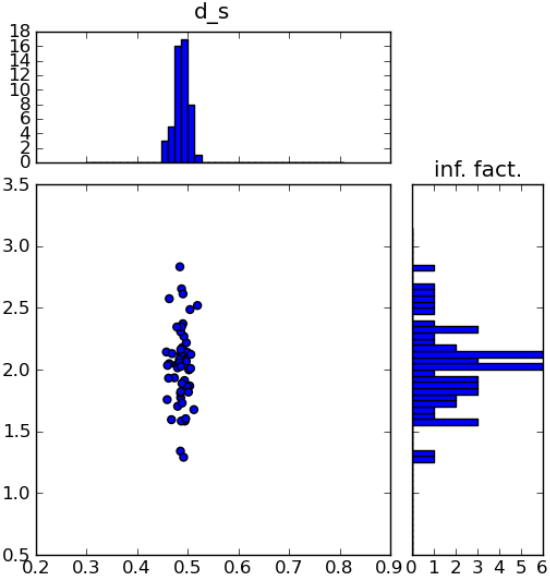
Late phase - numerical experiment



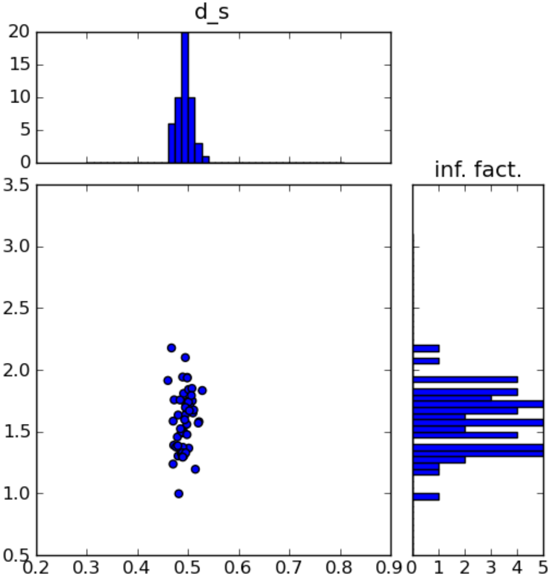
Late phase - numerical experiment



Late phase - numerical experiment



Late phase - numerical experiment



Thank you