RESEARCH REPORT

R. Hofman, P. Pecha, V. Šmídl

Data assimilation methods for early and late phase of radiation accident

A comprehensive review with examples of methods examined within solution of the grant projet VG20102013018

No. 2312   13. prosince 2011

Abstract

The task of the decision support in the case of a radiation accident is to provide up-to-date information on the radiation situation, prognosis of its future evolution and possible consequences. The reliability of predictions can be significantly improved using data assimilation, which refers to a group of mathematical methods allowing an efficient combination of observed data with a numerical model. The report concerns application of the advanced data assimilation methods in the field of radiation protection. We focus on assessment of off-site consequences in the case of a radiation accident when radionuclides are released into the environment.

In this report we present a comprehensive review of data assimilation evaluated for purposes of inclusion into the data assimilation and decision support system ASIM developed within the grant project VG20102013018 provided by the Ministry of the Interior of the Czech Republic. Besides the well established data assimilation methods like function fitting or constant statistical methods we introduce a new approach to data assimilation in both the early and the late phases of a radiation accident. Data assimilation is understood here as a particular case of recursive Bayesian estimation. Instead of using traditional estimation methods for state-space models, we focused on sequential Monte Carlo methods, specifically particle filtering and marginalized particle filtering.

Firstly, data assimilation methodology for the early phase of an accident is described. It employs particle filtering with adaptive selection of proposal density for estimation of the most important variables describing the aerial propagation of radionuclides. The general methodology is applicable to all parametrized atmospheric dispersion models. It is demonstrated on a simulated release, where a bias of the basic meteorological inputs and the source term is corrected using inference of gamma dose measurements.

Secondly, for the purpose of data assimilation in the late phase, we extended the idea of marginalized particle filtering to analytically intractable approximate filters, e.g. ensemble filters. The result is a hybrid data assimilation methodology, where multiple ensemble filters are run in parallel. The methodology was applied for joint estimation of the spatial distribution of deposition on terrain and estimation of the speed of radionuclides removal due to environmental processes in a simulated release scenario.

Financial support has been provided from the project MV ČR VG20102013018.

Contents

List of Figures

List of Tables

List of Acronyms

ADM	Atmospheric Dispersion Model
EKF	Extended Kalman Filter
EnKF	Ensemble Kalma Filter
HARP	HAzardous Radioactivity Propagation
HIRLAM	HIgh Resolution Limited Area Model
IAEA	International Atomic Energy Agency
KF	Kalman Filter
MPF	Marginalized Particle Filter
NPP	Nuclear Power Plant
OSCAAR	Off-Site Consequence Analysis code for Atmospheric Releases in reactor accidents
PF	Particle Filter
RIMPUFF	RIsø Mesoscale PUFF model
RODOS	Real-time Online DecisiOn Support system for nuclear emergency management
RMSE	Root Mean Square Error
SIR	Sampling Importance Re-sampling
SIS	Sampling Importance Sampling
TDS	Tele-Dosimetric System
UKF	Unscented Kalman Filter
UTM	Universal Transverse Mercator
i.i.d.	independent, identically distributed (random variable)
pdf	probability density function

Chapter 1

Introduction

In the case of a radiation accident, the risk evaluation and the decision-making process focused on protecting the public have the highest priority. The task of the decision support is to provide reliable and up-to-date information on the radiation situation, prognosis of its future evolution and possible consequences. Knowledge of spatial distribution of radionuclides and prediction of the future evolution are essential for planning of effective countermeasures. Historically, accidents in nuclear facilities have revealed unsatisfactory level of preparedness and lack of adequate modeling tools. Great attention has been paid to this topic since the Chernobyl disaster [OVZ07]. Nowadays, decision makers dispose of complex computer systems intended to provide assistance to them throughout various phases of the accident, e.g., [PS00, PHP07, TNDM99].

During the last decades, a great progress has been made in our understanding the atmospheric dispersion and related natural phenomena. Despite all the effort, the stochastic nature of involved physical processes, the deficiencies in their mathematical conceptualization and particularly ignorance of the initial conditions prevent obtaining of accurate results. The only way how to attain satisfactory accuracy of the model forecasts is exploitation of observational data, which represent the only connection with the physical reality. Observations are often sparse in both time and space and it is not possible to get a complete picture of radiological situation based on monitoring data alone, especially during the first hours after the accident.

Data assimilation provides a framework for optimal combination of numerical model predictions and the available observational data [Kal03]. It makes possible to consistently account for uncertainties in the model, its inputs and observations, and produces probabilistic answers which are more informative than those deterministic. Data assimilation is a compromise between the pure modeling approach on one hand and the data mining approach on the other hand. Nowadays, data assimilation arise in many scientific areas. The main fields of its application are meteorology, oceanography and hydrology [PX09, WZZ00]. This work addresses the problem of exploitation of advanced data assimilation methods in the field of radiation protection.

1.1 Introduction and terminology

Data assimilation results from the methods of objective analysis introduced in the middle of the 20th century in order to eliminate a subjective human factor in numerical weather prediction [Dal93]. It refers to a group of mathematical methods for estimation of a state of a dynamic system by the means of combining multiple sources of information, typically observational data with a numerical model of the system under investigation. We are concerned with 4-D data assimilation, where the assimilation is performed in time and space.

1.1.1 Classification of data assimilation methods

There are two basic approaches to data assimilation: (i) sequential assimilation, that only considers observation made in the past until the time of analysis, which is the case of real-time assimilation systems, and (ii) non-sequential, or retrospective assimilation, where observation from the future can be used, for instance in a reanalysis exercise. In this work, we focus on the first type of methods.

Another distinction can made between methods that are *intermittent* or *continuous* in time. In an intermittent method, observations can be processed in small batches, which is usually technically convenient. In a continuous method, observation batches over longer periods are considered, and the correction to the analyzed state is smooth in time, which is physically more realistic [BC02].

1.1.2 Data assimilation cycle

Data assimilation is performed in cycles, where each the assimilation cycle has two steps. Adopting the generally accepted data assimilation terminology unified in [ICGL97], the first step, the *data update*, can be described as follows: Given the model *forecast* (so called *background field*) and the observations, the data update produces their statistically optimal combination called *analysis*. It is an estimate of the current system state considered to be better both the standalone model forecast and the observations. Essentially, the analysis step tries to balance the uncertainty in the data and in the forecast. In the second step, the *time update*, the analysis is integrated forward in time using the model equations. This becomes the new forecast in the next assimilation cycle. Periodic updating of the model with observations should ensure that the model will not diverge from the physical truth.

Illustration of the sequential data assimilation process is in Figure 1.1: Let the system state be a one-dimensional continuous random variable estimated in discrete time steps. Observations available in discrete time steps represent a connection with the physical reality and can be understood as a noisy samples from the true state represented by the blue curve. Observations are denoted with squares and the green circles represent their uncertainty. In each time instance, the best state estimate—analysis denoted by asterisk—is produced on basis of current model forecast (plus sign)

and observations. The red and yellow circles represent the uncertainties of forecast and analysis, respectively. In the figure is schematically depicted that the forecast error is reduced in each time step after the data update (yellow dashed line). The red dashed line represents the time update step, when the analysis is advanced via the model forward in time.
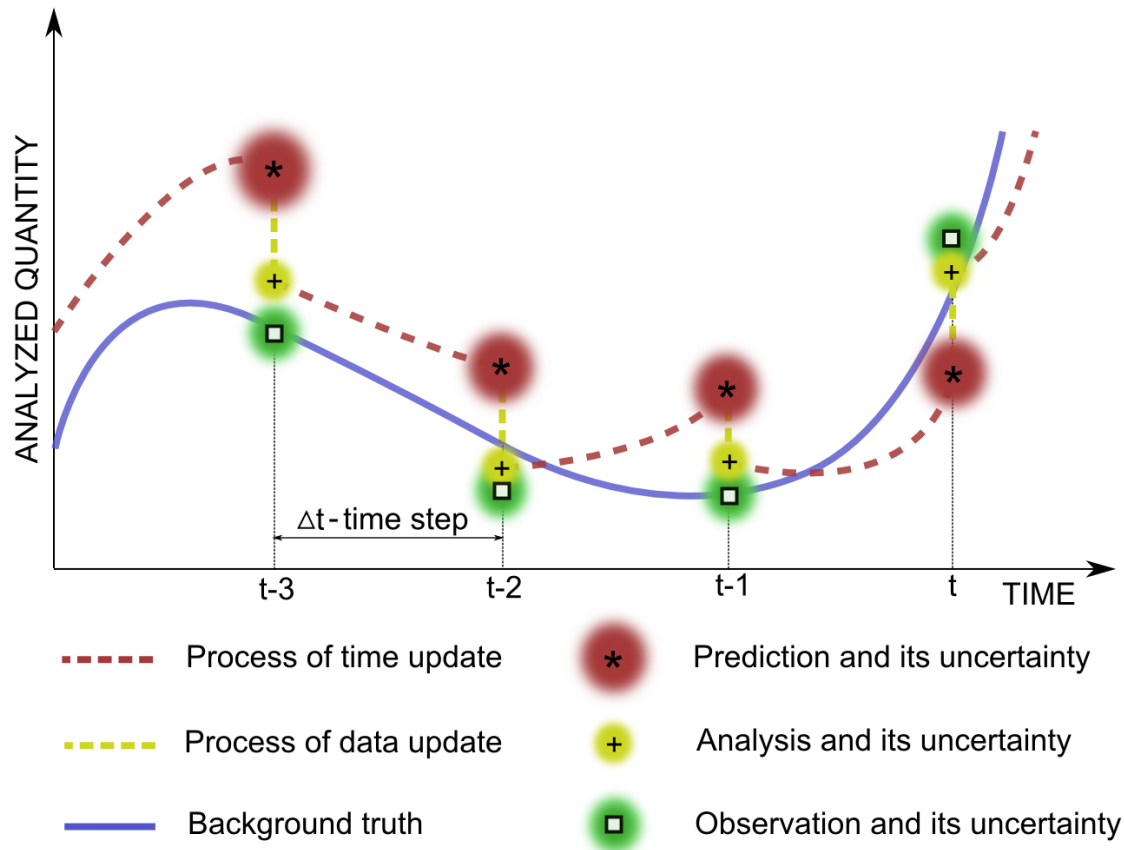
Figure 1.1: Illustration of basic principle of sequential data assimilation.

1.2   Data assimilation in radiation protection

We are concerned with application of data assimilation in the case of a severe radiation accident, when an accidental release of radionuclides into the environment occurred and it is likely to require at least partial implementation of countermeasures. The main objective of data assimilation is to estimate the true scale of the accident and predict its consequences in order to improve reliability of the decision support through different phases of the accident.

The time tract of an accidental release of radionuclides can be formally split into two consecutive phases:

Early phase begins when the radionuclides are released into the environment. We focus on atmospheric releases, when the effluent forms a radioactive plume advected by the wind field and dispersed by turbulent processes. The plume causes external irradiation from *cloudshine* and internal irradiation due to inhalation. Duration of this phase is from a few hours up to several days and let it formally ends when the plume leaves the area of interest. The main objectives of data assimilation in the early phase are (i) on-line estimation of radiation situation and its evolution and (ii) estimation of committed population doses.

Late phase covers latter stages of the accident and immediately follows after the early phase. After the plume passage, there is no more irradiation due to cloudshine, however, on the ground remains deposited radioactive material. It causes external irradiation from *groundshine* and internal irradiation from inhalation due to resuspension and ingestion. This phase ends when radiation levels resume to background values. The main objectives of data assimilation in the late phase are (i) identification of contaminated areas and (ii) estimation of radiation levels and the speed of the radionuclides removal for purposes of long-term forecasting. The estimates enter subsequent models of radionuclides propagation through the different compartments of the environment.

Data assimilation is potentially applicable in both phases, however, different physical processes, time scales etc., determine specific requirements on assimilation inputs and target fields of predictions. The key properties of the early and the late phase are summarized in Table 1.1. Summary of applicability of DA methods reviewed in this work is in Table. 1.2.

1.2.1 State of the art

1.2.1.1 Assimilation of Lagrangian particle models

Lagrangian particle model is a Monte Carlo dispersion model, where the spreading of pollutants is simulated using a large number of particles released from the source, each of them carrying some elemental activity. Trajectories of particles are given by a meteorological forecast entering the model. Random perturbations are added to the wind speed of the particles in order to simulate stochastic turbulent processes in the atmosphere. In this model, the three-dimensional space is divided into partial volumes. At each time step, movement of all the particles is traced and the activity concentration in each partial volume is obtained by summing up the activity assigned to particles within the volume. When a new set of observations is available, the assimilation procedure is performed as a modification of the number of particles in the partial volumes, e.g. [ZLLL07]. Between consecutive measurement updates, the redistributed particles are propagated forward in time by the meteorological forcing.

The advantage of Lagrangian models is their capability to account for many physical processes in a natural way. Their application in data assimilation allows for local

PHASES OF REACTOR ACCIDENT		
Property	**Early phase**	**Late phase**
Duration	This phase starts when the radionuclides are released into the environment and lasts from few hours to days until the plume leaves the area of interest	The phase begins immediately after the early phase and lasts until the radiation levels resume to background values (years)
Pathways of irradiation	(i) External irradiation from the cloud and the deposition (ii) Internal irradiation from inhalation	(i) External irradiation from deposition (ii) Internal irradiation from ingestion and inhalation of resuspended material
Typical countermeasures	Iodine prophylaxis, sheltering, evacuation	Food bans, relocation, amelioration (liming, fertilization) and decontamination of selected areas (ploughing, soil removal)
Possible sources of observations	(i) Radiation monitoring network (ii) Airborne and ground-based mobile groups	(i) Thermoluminiscent dosimetry network (ii) Airborne and ground-based mobile groups (iii) Laboratory analysis, monitoring
Main objectives of data assimilation in this phase	(i) On-line (step-wise) estimation of the affected area and committed doses. (ii) Determination of source term, correction of meteorological and other inputs	(i) Identification of contaminated area and radiation levels (ii) Estimation of the speed of the radionuclides removal for purposes of long-term forecasting

Table 1.1: Summary of the key properties of the early and the late phase of radiation accident.

Class of methods	Requirements:				Applicable to:		
	R1	R2	R3	R4	A1	A2	A3
Function fitting methods	x				x		
Empirical interpolation methods	x	x			x		
Constant statistical methods	x	x	x		x		
Minimization (optimizations) methods	x			x			x
Ensemble filters	x	x	x	x	x	x	x
Sequential MC methods	x	x	x	x	x	x	x

Table 1.2: Data assimilation methods. Legend: R1 - observations, R2 - background field, R3 - BF and/or observation error statistics, R4 - numerical model of state evolution. A1 - intermittent assimilation of spatially (2-D or 3-D) distributed quantity, A2 - continuous assimilation of spatially distributed quantity, A3 - estimation of dispersion model parameters (use with parameterized models - inverse modeling).

assimilation of the activity concentrations and thus the results better consider local variations in terrain, meteorology etc. The disadvantage is the fact, that a large number of particle trajectories must be computed to simulate a release using this type of model. Such an assimilation algorithm based on this approach model must be run on a supercomputer in order to meet the strict time constraints in the early phase.

1.2.1.2   Assimilation of parameterized models

A substantial reduction of the computational complexity can be reached by the use of a deterministic model parametrized by a set of control variables. The term *control variables* refers to a selected subset of inputs to the model and parameters influencing its result. The set is selected using the uncertainty and sensitivity studies performed with dispersion models, [EKT07, Rao05, TvTB07]. Given some particular values of control variables, concentration in air is calculated simply by evaluation of the model as a deterministic function of the variables. Contrary to Lagrangian particle models, direct assimilation of concentration values in the grid points is not possible with these models. Modification of the analytical shape of the plume would forbid its propagation in the next time step. Data assimilation is then formulated as an optimization of the control variables in order to reach the best correspondence of model forecast with available observations. These estimates may in turn re-enter atmospheric dispersion models, resulting in a greatly improved dose rate assessment. Parameters not included in the set of control variables are not treated as uncertain but they are initialized with a fixed value.

The most simple methods for optimization of the control variables are not probabilistic and minimize just a loss function measuring point-wise distance between model and observations. [EKT07] presented a simple assimilation scheme for tuning of the effective release height and the wind direction of the Gaussian plume model. This idea is more developed in [PH08], where a segmented version of the Gaussian plume model

[HPP08a] is used and the set of optimized control variables is extended to address their time variability. The advantage of this method is its simplicity and a potential for extension of the set of optimized control variables. The disadvantage is the fact that the method does not consider error statistics of the model and observations, contrary to variational methods, where the difference between the model forecast and the observations is weighted with appropriate error statistics. Assimilation schemes based on variational approach are described in [JKS$^+$05, KTAB09, QSI05]. Here, all the optimized control variables are treated as time invariant.

More advanced methods are based on sequential data assimilation. [DLM05] described extended Kalman filtering of the Gaussian plume. Here, the set of optimized control variables is restricted to the ratio of the release rate and the wind speed, the wind direction and the plume height. Similar assimilation scheme is proposed in [ATP$^+$04] describing assimilation of the RIMPUFF model [TNDM99]. A continuous release is with the RIMPUFF (RIsø Mesoscale PUFF) model approximated by a sequence of overlapping puffs. This allows inclusion of complex meteorological and other local characteristics. Control variables are radioactive inventories of partial puffs and the wind direction affecting spatial positions of the puffs within the computation domain. The number of control variables changes dynamically as new puffs are released and other puffs leave the domain. This assimilation methodology for the early phase is implemented in the RODOS (Real-time Online Decision Support System for nuclear emergency management), [PMG$^+$03].

Due to time pressure in the early phase, we will focus our attention to the sequential version of the Monte Carlo technique, which is also known as the particle filter [GSS93, DGA00]. It has been shown in [JHN04, HKvD11] that it provides results comparable to those obtained by the classical Monte Carlo Markov Chain algorithm.

1.2.1.3 Data assimilation in the late phase

The basic aspects of modeling and assimilation in the late phase are formulated in [GWW$^+$04]. Modeling in the late phase covers a broad range of disciplines focusing on different problems, e.g., contamination of arable soil and urban areas, contamination of water resources, propagation of radionuclides in the food chain, etc. In [YKM$^+$05], the method *iterations to optimal solution* is applied for assimilation of an aquatic model with observations of the Black Sea contamination after the Chernobyl accident. The details regarding this simple empirical interpolation method can be found in [Dal93].

In [Pal05], the ensemble Kalman filtering (EnKF) based data assimilation system for assimilation of the groundshine measurements with a radio-ecological model is described. The system is a part of the RODOS. EnKF introduced by [Eve94] is proposed here as the most promising approach for data assimilation in the late phase.

1.2.2   Evaluation of performance

The performance assessment of data assimilation methods is in the field of radiation protection problematic. The dispersion modeling of radioactive pollutants has its specific properties and the existing data sets from experiments with non-radioactive pollutants are not suitable. Since there is a lack of observational data sets from the real reactor accidents, the measurements used for validation of data assimilation methods are simulated using the *twin experiments* [EKT07]. It means, that the measurements are generated using a model of the system under investigation, initialized by some reference values. Observations are sampled from the model output fields in locations of the receptor points. From the theoretic point of view, the twin experiments are useful, because they make possible to evaluate assimilation performance against a known "background truth" and the convergence can be easily assessed. The method also provides a transparent tool for controlling of measurement error type and magnitude.

Chapter 2

Physical Models

Before we proceed to description of particular assimilation methods, we briefly introduce basic physical models needed for modeling of different physical phenomena in the early and the late phase of a radiation accident. The first part of this chapter is devoted to atmospheric dispersion modeling and the second part concerns basic radiological quantities. The focus is paid to development of an observations operator transforming spatially distributed activity in air into (time integrated) gamma dose rate.

2.1 Atmospheric dispersion modeling

Atmospheric dispersion modeling is the mathematical simulation of how air pollutants disperse in the ambient atmosphere. Dispersion models are computer codes solving equations describing the propagation of pollutants given the initial conditions, i.e., the meteorological conditions (wind speed and direction, precipitation) and the process conditions (heat capacity of the plume, terrain roughness, etc.) prevailing in the atmospheric boundary layer. Output from such a dispersion model is a 3-dimensional field of pollutant concentration in air. In the case of radioactive pollutants, the output is given in terms of activity concentration in air $[Bq\,m^{-3}]$.

Atmospheric dispersion models are basic tools for decision makers when assessing the atmospheric radionuclide releases. The models predict concentration of pollutants in the downwind directions from the source. Combined with the information on demography, the models can estimate expected exposure of population to ionizing radiation, and consequently, the health effects in terms of total committed doses. Nowadays, there exist various approaches to atmospheric dispersion modeling.

2.1.1 From advection to diffusion

In the following text in this chapter, the subscript i iterates over the set $\{1, 2, 3\}$ denoting the three spatial coordinates.

Following [Bar01], let us assume a release of a material into the atmosphere. If no chemical reactions and molecular diffusion are assumed, the concentration of material,

C, resulting from the release is given by the advection equation

$$\frac{\partial C}{\partial \tau} + \frac{\partial}{\partial s_i}(u_i C) = 0, \tag{2.1}$$

where $\mathbf{u} = (u_1, u_2, u_3)$ are the wind speed components in directions $\mathbf{s} = (s_1, s_2, s_3)$ and $C = C(\mathbf{s}, \tau)$ is a function describing concentration of the pollutant in space and time. As the actual wind speed is not known and it can not be incorporated into the equation, we assume that the wind speed at a certain time can be described according to the scheme

$$u_i = \overline{u_i} + u_i', \tag{2.2}$$

where $\overline{u}$ is the mean wind speed during a time period and the second term u' stands for a stochastic component, fluctuation of the wind due to the momentary turbulence. This scheme implies also fluctuation of the concentration

$$C = \overline{C} + C', \tag{2.3}$$

where $\overline{C}$ is the concentration taken over a time period and C' is a stochastic fluctuation. The stochastic fluctuation terms are assumed to have zero mean values

$$\overline{C'} = 0, \qquad \overline{u_i'} = 0. \tag{2.4}$$

Substituting (2.2, 2.3) into (2.1) yields

$$\frac{\partial C}{\partial \tau} + \frac{\partial}{\partial s_i}\left[(\overline{u_i} + u_i')(\overline{C} + C')\right] = 0. \tag{2.5}$$

Exploiting linearity of the differentiation operator $\partial/\partial s_i$ and averaging over time yields

$$\frac{\partial \overline{C}}{\partial \tau} + \frac{\partial}{\partial s_i}\left(\overline{u_i}\,\overline{C}\right) + \frac{\partial}{\partial s_i}\left(\overline{u_i' C'}\right) = 0, \tag{2.6}$$

where the terms $\overline{u_i'}C'$ and $\overline{C'}u_i'$ are zero due to (2.4).

To describe the mean stochastic turbulent flux term $\overline{u_i' C'}$, we introduce the eddy diffusivity coefficients K_i. Turbulent *diffusion* is a diffusion process by which substances are mixed in the atmosphere or in any fluid system due to eddy motion. As the turbulent fluxes $\overline{u_i' C'}$ can be measured only with fast-response instruments and it is difficult to treat them theoretically by analogy with the molecular case, the turbulent flux is commonly assumed to be directly proportional to the mean gradient

$$\overline{u_i' C'} = -K_i \frac{\partial \overline{C}}{\partial s_i}, \tag{2.7}$$

where K_i are *diffusivity coefficients* in units $m^2 s^{-1}$. The negative sign is included so that the flux is down the gradient, i.e., from the high values of C to the low values. The mean wind components and the mean concentration represent average values over

the time scale T_a and the corresponding spatial scale s_a. Typical values of T_a are a few minutes in magnitude. Fluctuation in smaller scales is assumed to be turbulent and is included in K_i . Substituting (2.7) into (2.6) gives

$$\frac{\partial \overline{C}}{\partial \tau} + \frac{\partial}{\partial s_i}\left(\overline{u_i}\,\overline{C}\right) = \frac{\partial}{\partial s_i}\left(K_i \frac{\partial \overline{C}}{\partial s_i}\right). \tag{2.8}$$

Expanding the second term on the left side of (2.8) and assuming that the atmosphere is incompressible,

$$\frac{\partial u_1}{\partial s_1} + \frac{\partial u_2}{\partial s_2} + \frac{\partial u_3}{\partial s_3} = 0,$$

the second term on the left side of (2.8) vanishes and we obtain the *advection-diffusion equation*

$$\frac{\partial \overline{C}}{\partial \tau} + \overline{u_i}\frac{\partial \overline{C}}{\partial s_i} = \frac{\partial}{\partial s_i}\left(K_i \frac{\partial \overline{C}}{\partial s_i}\right). \tag{2.9}$$

It describes the relationship between the spatial and the temporal behavior of concentration. When combined with appropriate initial conditions and boundary conditions, this equation forms the basis for the dispersion modeling and may be solved for various scenarios. In this equation, both the pollution transport by the advection and the diffusion due to the atmospheric turbulence are represented. However, (2.9) is not possible to solve analytically for completely general functional forms for the diffusivity coefficients K_i and the wind speed components u_i.

2.1.2   Classifications of air pollution models

Models vary considerably in their complexity, and may take account of different physical and chemical processes affecting the flow and transport. Different mathematical expressions can be derived to represent these atmospheric processes. Consequently, there is an enormous range of available atmospheric dispersion models. Comprehensive review of atmospheric dispersion methodology is given, e.g. by [HM06].

2.1.2.1   Box models

This is a simple model, largely based on the concepts of conservation of mass and conservation of energy. The treatment of transport is simplified, but the model is capable to include complex chemistry. The model evaluates mass balance of a given system using the conservation laws, where the particles of pollutant are transferred from one domain of the environment to another. Inside a domain, the air mass is assumed to be well mixed and concentration of the pollutant is assumed to be homogeneous. Boundaries of the domains are boxes. For every pollutant, we can write the mass balance equation:

$$\text{Input rate} = \text{Output rate} + \text{Transformation rate} + \text{Accumulation rate}$$

Depending on the physical and chemical interactions, some of the pollutants may pass through the system unchanged, some may accumulate within the system, while some may undergo chemical transformation or radioactive decay.

The simplicity of the model implies that it requires simple meteorological inputs and simple parametrization of the emission source. As it provides area-wide averages of concentration for a given region, the box model is a useful tool for screening purposes, where we need quick answers without any stress on accuracy. However, well-mixed and homogeneous conditions are sometimes unrealistic and the box models should not be used to calculate concentration in large areas, where the local changes must be reflected. For more detailed modeling we need more complex models continuously tracking the plume through the environment as it is advected by the wind, spread by diffusion, mixed by turbulence and reflected or channeled by surfaces such as the ground and the buildings [Bar01].

2.1.2.2 Lagrangian and Eulerian models

Both the *Lagrangian* and the *Eulerian* models solve the same advection-diffusion equation. The difference between Lagrangian and Eulerian approach to modeling consists in the different treatment of the frame of reference. The Lagrangian approach is based on studying the property of a particular fluid by following its trajectory. Lagrangian models are similar to the box models, where the region of air containing an initial concentration of pollutants is considered as a box [Gur08]. The box is considered to be advected with the flow and the model follows the trajectory of the box. It is said that an observer of a Lagrangian model follows along with the plume. The motion of air parcels is modeled as a superposition of the mean wind speed and a random perturbations simulating chaotic nature of the atmosphere. It is a random walk process indeed. Concentration is in the Lagrangian models evaluated in partial volumes (boxes) forming a 3-dimensional grid. Average concentration in a given grid cell is evaluated in a way that we sum up all the elemental concentrations associated with the particles in the cell. The main advantage of Lagrangian models is the capability to account for many physical processes in a natural way. They work well both for homogeneous and stationary conditions over the flat terrain and for inhomogeneous and unstable media conditions for the complex terrain. Particle dispersion model is an example of practical implementation of a Lagrangian model [ZLLL07].

In Eulerian modeling, we also track the movement of a hypothetical parcel of air, but we use a fixed frame of reference. The Eulerian approach is based on studying fluid property in a control volume at a fixed point in space, that is, the control volume is stationary and fluid moves through the control volume [Gur08]. It is said that an observer of an Eulerian model watches the plume go by. Eulerian models use 2-

dimensional and 3-dimensional grids for solving the differential equations, e.g. (2.9), where diffusion, transport, and removal of pollutant emission is simulated in each cell.

2.1.2.3   Gaussian models

Gaussian models are widely used in atmospheric dispersion modeling, and are often "nested" within Lagrangian and Eulerian models. They are based on a Gaussian distribution of concentration in the plume in vertical and horizontal directions under the steady state conditions [Zan90, HM06]. Gaussian models are popular, particularly for the following reasons:

- The Gaussian models represent a solution of (2.9) under some simplifying assumptions (e.g., constant wind and eddy diffusivity coefficients) and they are consistent with the random nature of the turbulence.

- Their simplicity allows for fast evaluation even with small computational resources. This is an essential property when we attempt to employ assimilation techniques based on Monte Carlo approach, when the model must be repeatedly run for many times.

- The analytical form of the Gaussian models allows for a good insight and a transparent evaluation of experimental results.

- The Gaussian models are easy to implement and they can be embedded into various forecasting and assimilation systems.

- Validity of the Gaussian models was satisfactorily verified for different meteorological conditions via comparison to the results of field tests with tracer releases, when the agreement of measured and modeled concentration was assessed, e.g. [CEE^{+}95].

Gaussian models are not designed to model dispersion under low wind conditions or at sites close to the source, i.e., at distances closer than 100m. It was found that these models over-predict concentrations in low wind conditions [HBHJ82].

Gaussian models—in their basic form—assume just the diffusion and advection of the pollutants. Modified versions of the Gaussian models are capable to include physical processes such as dry and wet deposition and radioactive decay [HPP08a]. We can distinguish two main variants of the Gaussian models. The *Gaussian plume model* assumes a continuous release when a plume in the downwind direction is formed under stationary conditions. The *Gaussian puff model* assumes a sudden instantaneous release when an expanding puff is formed.

2.1.2.4   Computational fluid dynamics models

Computational fluid dynamics models are able to deal with the fluid flux in a complex geometry by solving the Navier-Stokes equation and the continuity equation when the
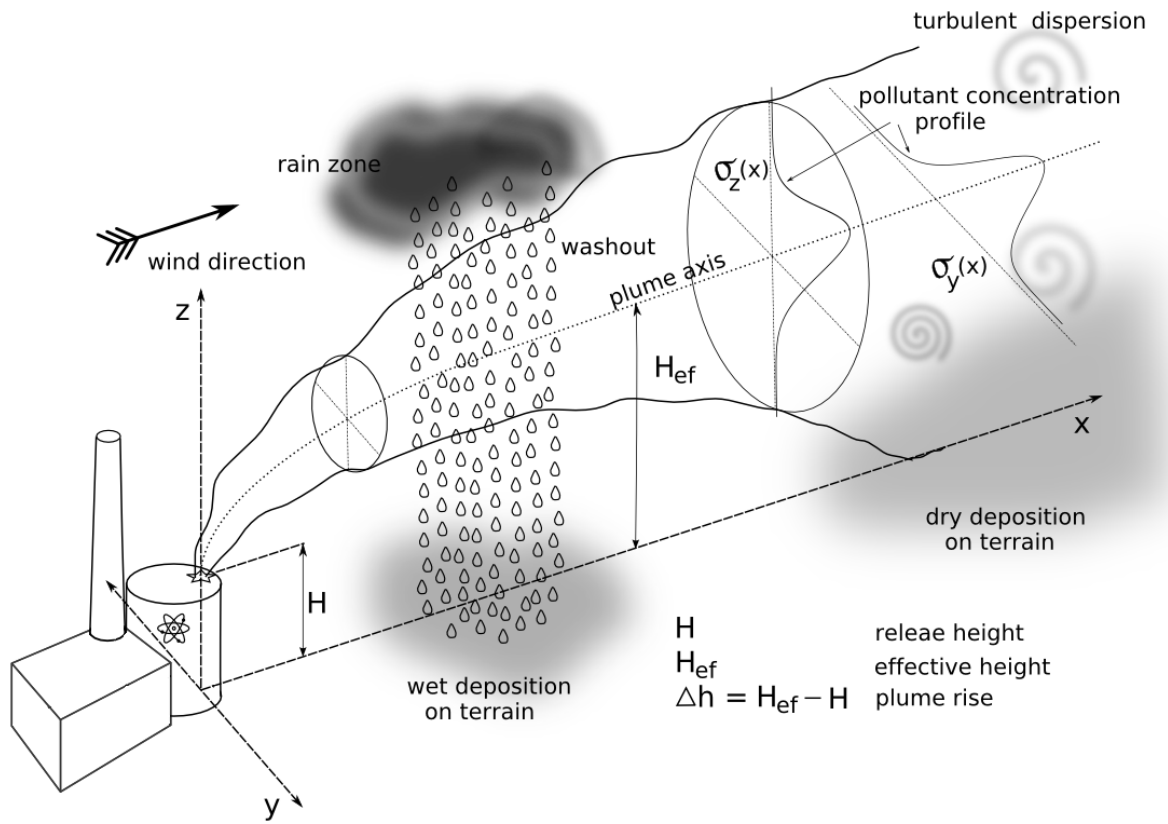
Figure 2.1: Illustration of the Gaussian plume model.

flow is idealized as a laminar flow [Gur08]. These two equations can be solved simultaneously using finite difference or finite volume methods. If the flow is turbulent, the Reynolds Navier-Stokes equation with the continuity and turbulence closure models is used for this case [TL72].

2.1.2.5 Summary

In this work, we focus on parametrized Gaussian models which are discussed in more detail in the following section.

2.1.3 Gaussian models

To simplify the notion, we omit the over-bars denoting the average values of C and u_i in the following text.

2.1.3.1 Gaussian plume model

Gaussian plume model is obtained as a 3-dimensional, time independent solution of (2.9) for a continuous source at the ground level, constant u_i and K_i, and a flat terrain.

Gaussian plume models gives us a steady-state solution, i.e., the model does not take into account the time required for the pollutant to travel to the receptor and describes concentration in a fully established plume under stationary meteorological conditions.

Let us assume a continuous point-source ground-based release of mass Q with time constant coefficients K_i and wind speed components u_i. The wind is assumed to blow along the direction of s_1-axis ($\mathbf{u} = (u_1, 0, 0)$) of an Eulerian coordinate system with the origin aligned with the source location. We can assume that the advection dominates diffusion in the downwind direction,

$$u_1 \frac{\partial C}{\partial s_1} >> \frac{\partial}{\partial s_1} \left(K_1 \frac{\partial C}{\partial s_1} \right).$$

The diffusion in this direction can be then neglected and the basic equation (2.9) modified for the steady-state conditions is

$$u_1 \frac{\partial C}{\partial s_1} = \frac{\partial}{\partial s_2} \left(K_2 \frac{\partial C}{\partial s_2} \right) + \frac{\partial}{\partial s_3} \left(K_3 \frac{\partial C}{\partial s_3} \right). \tag{2.10}$$

The boundary conditions for solution of (2.10) are:

1. $C \rightarrow 0$ for $r = \sqrt{s_1^2 + s_2^2 + s_3^2} \rightarrow +\infty$,

2. $C \rightarrow +\infty$ for $r \rightarrow 0$,

3. $K_3 \frac{\partial C}{\partial s_3} \rightarrow 0$ for $s_3 \rightarrow 0$,

4. $\int_0^{+\infty} \int_{-\infty}^{+\infty} u_1 C \, ds_2 ds_3 = Q$.

Here, Q defines a continuous point-source in terms of released mass per time (activity per time in the case of a radioactive release). The first two conditions represent our requirements on the concentration values at zero and infinite distances from the source. Condition 3 expresses the fact that we assume no sedimentation on the ground and condition 4 is the formalization of the law of conservation of the released mass (activity). The appropriate solution for constant wind and diffusivity coefficients is

$$C(\mathbf{s}) = \frac{Q}{4\pi s_1 \sqrt{K_2 K_3}} \exp \left[-\frac{u_1}{4 s_1} \left(\frac{s_2^2}{K_2} + \frac{s_3^2}{K_3} \right) \right]. \tag{2.11}$$

Since the values of diffusion coefficients is difficult to measure, we use *dispersion coefficients* defined as

$$\sigma_i = \sqrt{2 K_i \tau}. \tag{2.12}$$

The dispersion coefficients are usually functions of the atmospheric stability category and the downwind distance from the source. Substitution of (2.12) into (2.11) gives us a time independent formula for concentration in a developed Gaussian plume

$$C(\mathbf{s}) = \frac{Q}{2\pi u_1 \sigma_2 \sigma_3} \exp \left(-\frac{s_2^2}{2\sigma_2^2} \right) \exp \left(-\frac{s_3^2}{2\sigma_3^2} \right). \tag{2.13}$$

The coefficients σ_2 and σ_3 represent the standard deviations of the concentration in the cross-wind (s_2) and vertical (s_3) planes, respectively. Larger values of the coefficients mean broader distribution and consequently higher dilution of pollutant and smaller concentration.

2.1.3.2 Gaussian puff model

Contrary to the plume models, the puff models are applicable when mean wind vector transporting the puff is variable in space and time. Using the puff model, these variations can be directly incorporated in a numerical scheme whereby they are used to transport the center of the puff. Dispersion of the puff (Gaussian distribution) is evaluated using the concept of *virtual source* at each time step. The characteristic feature of these models is that the calculation of pollutant diffusion, transportation, and removal is performed in the Lagrangian frame of reference attached to a number of parcels as they are transported around the geographical region of interest. Approximation of a continuous release is reached by simultaneous propagation of multiple puffs. Under assumption of stationarity, spatially homogeneous flow of multiple puffs over the flat terrain represents the Gaussian plume model, which is demonstrated in [JPP03]. The overall concentration is evaluated as a superposition of the puffs [LGR77, Zan86].

The form of (2.9) describing advection and diffusion of a single puff is

$$\frac{\partial C}{\partial \tau} + u_1 \frac{\partial C}{\partial s_1} = \frac{\partial}{\partial s_1}\left(K_1 \frac{\partial C}{\partial s_1}\right) + \frac{\partial}{\partial s_2}\left(K_2 \frac{\partial C}{\partial s_2}\right) + \frac{\partial}{\partial s_3}\left(K_3 \frac{\partial C}{\partial s_3}\right). \qquad (2.14)$$

We solve (2.14) with the following boundary conditions:

1. $C \to 0$ for $r = \sqrt{s_1^2 + s_2^2 + s_3^2} \to +\infty$,

2. $C \to +\infty$ for $r \to 0$,

3. $K_3 \frac{\partial C}{\partial s_3} \to 0$ for $s_3 \to 0$,

4. $\int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C \, ds_1 ds_2 ds_3 = Q^i$.

Here, Q^i defines an instantaneous point-source in terms of released mass (activity in the case of a radioactive release). Analytical solution for the given case describing the concentration of pollutant everywhere in space is

$$C(\mathbf{s}, \tau) = \frac{Q^i}{8(\pi\tau)^{3/2}\sqrt{K_1 K_2 K_3}} \exp\left\{-\frac{1}{4\tau}\left[\frac{(s_1 - u_1\tau)^2}{K_1} + \frac{s_2^2}{K_2} + \frac{s_3^2}{K_3}\right]\right\}. \qquad (2.15)$$

After substitution of dispersion coefficients σ_i for the eddy diffusivity coefficients K_i we obtain equation of the Gaussian puff. Under general conditions, a puff located in $\mathbf{s}^P = (s_1^P, s_2^P, s_3^P)$ in time τ generates the concentration field

$$C(\mathbf{s}, \tau) = \frac{Q^i}{(2\pi)^{3/2}\sigma_1\sigma_2\sigma_3} \exp\left\{-\frac{1}{2}\left[\frac{\left(s_1^{\mathrm{R}}-s_1^{\mathrm{P}}\right)^2}{\sigma_1^2} + \frac{\left(s_2^{\mathrm{R}}-s_2^{\mathrm{P}}\right)^2}{\sigma_2^2} + \frac{\left(s_3^{\mathrm{R}}-s_3^{\mathrm{P}}\right)^2}{\sigma_3^2}\right]\right\},$$

$$(2.16)$$

at a receptor located at $\mathbf{s}^{\mathrm{R}} = (s_1^{\mathrm{R}}, s_2^{\mathrm{R}}, s_3^{\mathrm{R}})$. Similarly to the plume model, coefficients σ_1, σ_2 and σ_3 represent the standard deviations of the concentration in the downwind (s_1), cross-wind (s_2), and vertical (s_3) planes, respectively.

2.1.3.3 Parametrization of dispersion coefficients

Variability of temperature with altitude influences the turbulence characteristics and thus the dispersion of pollutants. The temperature in the atmosphere is governed by incident solar radiation, prevailing wind velocity, and percentage of cloud cover. Depending on the magnitude of these parameters, [Pas61] introduced the six stability classes named A, B, C, D, E, and F of the atmospheric turbulence. Class A denotes the most unstable or most turbulent conditions (the dispersion is higher), and class F the most stable or the least turbulent class (very low dispersion).

Besides the atmospheric stability category, dispersion coefficients are also dependent on travel time from the source and the type of terrain (urban, rural, etc.), e.g. [Gif76]. The comprehensive review can be found in [HBHJ82]. More advanced models apply Monin-Obukhov *similarity* theory and use the surface roughness length and the Monin-Obukhov length to determine the magnitude of dispersion, see [Che02, CPV$^+$04].

2.1.3.4 Elevated sources and reflections

The last exponential terms in (2.16) and (2.13) stand for exponential concentration profiles in the vertical direction. Let the terms be denoted as V,

$$V = \exp\left[-\frac{s_3^2}{2\sigma_3^2}\right]. \qquad (2.17)$$

In the most of real situations we assume, that the source is elevated over terrain in a height H. Moreover, if the effluent has a heat capacity or an initial vertical momentum, the height of the plume reaches so called *effective height* H_{ef}. It is a sum of the release height H and the height change due to the plume rise or subsidence ΔH,

$$H_{\mathrm{ef}} = H + \Delta H.$$

Vertical dispersion is usually assumed to be a growing function of the downwind distance, $\sigma_3 = \sigma_3(r)$. The form of (2.17) suggests that the vertical expansion of a puff or plume can be infinite. This is, of course, not possible in reality . The Gaussian distribution of the concentration is modified at greater distances from the source due to the effects of turbulent reflections from the surface. When the pollutant reaches the ground due to the vertical dispersion, the further spreading in vertical direction is not

possible and it is assumed that at the surface is the pollutant reflected without any loss. Reflection on the ground is modeled as a virtual source at the effective height H_{ef} below the ground. To account for the elevated source and the ground reflection we modify the vertical dispersion term as follows,

$$V_1 = \exp\left[-\frac{(s_3 - H_{\text{ef}})^2}{2\sigma_3^2}\right] + \exp\left[-\frac{(s_3 + H_{\text{ef}})^2}{2\sigma_3^2}\right].$$

Puff or plume can be also reflected from the top of *mixing layer.* In such a case, the vertical profile of the plume is bounded by the ground and the top of mixing layer at height H_{mix}. Theoretically, the number of reflections can be infinite. However, multiple reflections on the ground and at the top of mixing layer lead to vertical homogenization of concentration and numerical experiments proved that one virtual source below ground and one at the top of the boundary layer gives sufficient accuracy [Bar01],

$$
\begin{aligned}
V_2 &= \exp\left[-\frac{(s_3 - H_{\text{ef}})^2}{2\sigma_3^2}\right] + \exp\left[-\frac{(s_3 + H_{\text{ef}})^2}{2\sigma_3^2}\right] \\
&+ \exp\left[-\frac{(s_3 - 2H_{\text{mix}} - H_{\text{ef}})^2}{2\sigma_3^2}\right] + \exp\left[-\frac{(s_3 + 2H_{\text{mix}} - H_{\text{ef}})^2}{2\sigma_3^2}\right] \\
&+ \exp\left[-\frac{(s_3 - 2H_{\text{mix}} + H_{\text{ef}})^2}{2\sigma_3^2}\right] + \exp\left[-\frac{(s_3 + 2H_{\text{mix}} + H_{\text{ef}})^2}{2\sigma_3^2}\right]. \quad (2.18)
\end{aligned}
$$

The principle used for modeling of the reflections as a superposition of multiple plumes released from virtual sources is illustrated in Figure 2.2.
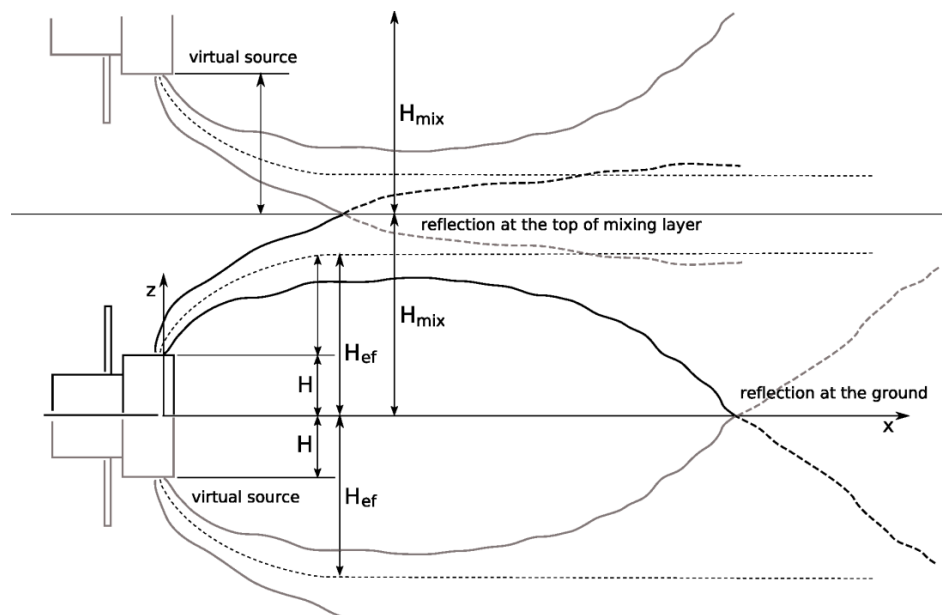
Figure 2.2: Illustration of the principle used for modeling of the reflections as a super-position of multiple plumes released from virtual sources below the ground and above the top of mixing layer .

2.2 Radiological quantities

In this section we briefly describe some essential radiological quantities used in this work and their relations.

2.2.1 Radioactive decay and radioactivity

Radioactive decay is a spontaneous nuclear transformation followed by an emission of ionizing particles. As the decay of an unstable nucleus is entirely random and it is not possible to predict when a particular atom will decay, it is described in terms of a continuous quantity $N(\tau)$, mean value of radioactive (undecayed) atoms in time τ.

Given a sample of a particular radioisotope, the number of decay events $(-dN(\tau))$ expected to occur in a small interval of time dt is proportional to the number of atoms present:

$$-\frac{dN(\tau)}{d\tau} = \lambda\,N(\tau).\qquad(2.19)$$

The negative sign indicates that the mean value of radioactive atoms $N(\tau)$ decreases with each decay event and λ is the proportionality constant known as the *decay constant*. Particular radionuclides decay at different rates, each having its own λ (there are more than 500 different nuclides). Solution of the first order differential equation (2.19) is a function

$$N(\tau) = N_0 \; \exp(-\lambda\,\tau), \tag{2.20}$$

where N_0 is the number of radioactive atoms in time $t = 0$.

The decay rate is denoted as *activity*, $A(\tau)$, which is a basic physical unit quantitatively describing physical phenomena *radioactivity*. Activity is a measure of the expected number of disintegrations per unit time [CLR02],

$$A(t) = -\frac{dN(\tau)}{d\tau}.$$

Since $A(\tau)$ is proportional to the mean number of atoms $N(\tau)$, we can write

$$A(\tau) = A_0 \; \exp(-\lambda\,\tau), \tag{2.21}$$

where A_0, analogously to N_0, is the activity in time $\tau = 0$. Although the radioactive decay is a discrete random process, the continuous exponential functions (2.20) and (2.21) are for large numbers of atoms (comparable to Avogadro's number in magnitude) a good approximation.

Mean lifetime τ^{m} of an atom before it undergoes the decay is inversely proportional to λ,

$$\tau^{\mathrm{m}} = \frac{1}{\lambda}.$$

Even more common is the use of physical half-life, $T_{1/2}$, which is the time needed to reduce the amount of radioactive material by a factor of 2. The formula

$$\frac{N(\tau)}{N_0} = \frac{1}{2} = \exp(-\lambda\,T_{1/2})$$

yields

$$T_{1/2} = \frac{\ln 2}{\lambda}. \tag{2.22}$$

Introducing the physical half-life (2.22) into (2.21), we get the *decay law*

$$A(\tau) = A_0 \; \exp\left(-\ln 2\,\frac{\tau}{T_{1/2}}\right). \tag{2.23}$$

The SI unit for radioactivity is *Becquerel* (Bq) and the activity is given in reciprocal seconds, s^{-1},

$$1 \text{ Becquerel (Bq)} = 1 \text{ (disintegration) s}^{-1}.$$

In the field of dispersion modeling of radionuclides, we evaluate *activity concentration in air*, which is a number of disintegration of a dispersed radionuclide in a unit volume per unit time, i.e., its unit is $Bq\,m^{-3}$. It is a quantity of particular importance because it can be used for evaluation other radiological quantities like deposition and doses.

2.2.2 Calculation of absorbed doses

The *absorbed dose* (also known as the total ionizing dose) is a measure of the energy deposited in a medium by ionizing radiation. It is equal to the energy deposited per unit mass of medium. Its unit $J\ kg^{-1}$ was given the special name Gray (Gy).

2.2.2.1 Absorbed dose from cloudshine

The *cloudshine* is external gamma radiation from a radioactive plume passing over the terrain. The simplest way of cloudshine dose rate calculation is based on approximation of the plume as a semi-infinite hemisphere with homogeneous concentration of radionuclides [RAC01, TNDL95]. Resulting formula for the gamma dose rate at a receptor R located at $\mathbf{s}^{\mathrm{R}} = (s_1^{\mathrm{R}}, s_2^{\mathrm{R}}, s_3^{\mathrm{R}})$ is

$$D_c(\mathbf{s}^{\mathrm{R}},\ \tau) = K\frac{E\,C(\mathbf{s},\ \tau)}{2\rho}, \tag{2.24}$$

where $\mathbf{s} = (s_1, s_2, s_3)$ is a spatial location; K is the dose rate conversion factor $[Gy\,kg\,eV]$; E is the gamma energy produced by decay of assumed radionuclide; $C(\mathbf{s},\ \tau)$ is radionuclide concentration $[Bq\,m^{-3}]$ in spatial location $\mathbf{s}$; and ρ is the air density. This formula assumes an equilibrium between the gamma energy released in the plume and that absorbed in the air. Approximation of a non-homogeneous plume (e.g. Gaussian) using the semi-infinite approach may lead to large errors. What is more, if the receptor point in not immersed in the radioactive cloud, the application of (2.24) is not well-founded at all.

The general expression for the effective flux of gamma rays at a receptor point $\mathbf{s}^{\mathrm{R}}$ from a source of ionizing radiation dispersed in air is according to [TNDL95], as follows,

$$\Phi(\mathbf{s}^{\mathrm{R}}, \tau, E) = \int_{\Omega} \frac{f(E)\,C(\mathbf{s},\ \tau)\,B(E, \mu r)\exp(-\mu r)}{4\pi r^2}d\mathbf{s}, \tag{2.25}$$

where $f(E)$ is the branching ratio to the specific energy E; B is the build up factor; μ is the linear attenuation coefficient; Ω is a spatial domain of integration; and $r = ||\mathbf{s}^{\mathrm{R}} - \mathbf{s}||$ is the distance of spatial locations $\mathbf{s}^R$ and $\mathbf{s}$. The build-up factor can be calculated from Berger's analytical formula

$$B(E, \mu\,r) = 1 + a\,\mu r\ \exp(b\,\mu r),$$

where coefficients μ, a and b depend on E. Energy dependent absorption coefficient μ_a is calculated using

$$\mu_a = \frac{\mu}{1 + \frac{a}{(1-b)^2}}.$$

The gamma dose rate from a mixture of nuclides emitting gamma radiation on different energy levels E^i, $i = 1, \ldots, N_E$, is

$$D_c(\mathbf{s}^R, \tau) = \sum_i^{N_E} \frac{K^i \, \mu \, E^i \, \Phi^i}{\rho}.$$

For a plume of a mono-energetic radionuclide dispersed in air emitting gamma radiation on a single energy level E, i.e. $f(E) = 1$, we obtain

$$D_c(\mathbf{s}^R, \tau) = \frac{K \, E \, \mu_a}{\rho} \int_\Omega \frac{C(\mathbf{s}, \tau) \, B(E_\gamma, \mu r) \exp(\mu r)}{4\pi r^2} d\mathbf{s}. \tag{2.26}$$

The time integrated gamma dose rate $\mathcal{D}_c(\mathbf{s}^R, \tau_1, \tau_2)$ integrated over the time interval (τ_1, τ_2) is defined as

$$\mathcal{D}_c(\mathbf{s}^R, \tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} D_c(\mathbf{s}^R, \tau) \, d\tau. \tag{2.27}$$

In the case of Gaussian models, concentration $C(\mathbf{s})$ is given by the analytical formulas (2.16) and (2.13). Specifically, the simplicity of the Gaussian puff model (2.16) allows for numerical evaluation of the integral in (2.26) on a compact support where the activity concentration is not negligible.

2.2.2.2 Absorbed dose from groundshine

The *groundshine* is the external gamma radiation from radioactive material deposited on the ground, trees, buildings etc. [GBEJ99]. Given deposition $SD(\mathbf{s}, \tau)$ in location $\mathbf{s}$ and time τ, the groundshine dose rate $D_g(\mathbf{s}, \tau)$ is calculated as follows,

$$D_g(\mathbf{s}, \tau) = DF_g \, SD(\mathbf{s}, \tau) \, SF, \tag{2.28}$$

where the coefficient of proportionality DF_g is the radionuclide-dependent integrated *dose rate conversion factor* for groundshine in units $Sv \, s^{-1}$ per $Bq \, m^{-2}$. Unit-less *shielding factor* SF is defined as

$$SF = \sum_i f_i \, SF_i,$$

where f_i is ith fraction of time spent in different places (indoor, outdoor, etc.) and $SF_i \in [0, 1]$ is the shielding factor at each place.

The time evolution of the deposition is modeled according to the standard formula

$$SD(\mathbf{s}, \tau) = SD(\mathbf{s}, 0) f_R(\tau) f_E(\tau), \tag{2.29}$$

where $SD(\mathbf{s}, 0)$ is initial deposition in time $\tau = 0$; $f_R(\tau)$ is a function taking into account *radioactive decay* (2.23) in terms on relative amount of undecayed material in time τ; and $f_E(\tau)$ a function taking into account decrease of radioactivity due to the environmental removal processes. Environmental removal is a general term referring to different processes causing radioactivity removal from terrain, e.g., radionuclide migration deeper into the soil, weathering, leaching.

An adequate description of the long term dynamics of radionuclides in soil is the most important factor in the correct estimation of the radioactive contamination of local agricultural and forest products, which has a major contribution to the exposure of the local population. There exist several computer codes implementing formulas (2.28) and (2.29). Their comprehensive review can be found in [TSA$^+$05]. In calculation of the groundshine dose, the main difference among the codes consists in different descriptions of $f_{\mathrm{E}}(\tau)$ and in application of different numerical values of SF and DF_g [IAE03].

We adopt groundshine dose model from Japanese code OSCAAR (Off-Site Consequence Analysis Code for Atmospheric Releases in reactor accidents) [Hom02]. Environmental removal is in (2.29) modeled using

$$f_{\mathrm{E}}(t) = d^{\mathrm{f}} \, \exp\left(-\ln 2 \, \frac{t}{T^{\mathrm{f}}}\right) + d^{\mathrm{s}} \, \exp\left(-\ln 2 \, \frac{t}{T^{\mathrm{s}}}\right), \tag{2.30}$$

where the rate of environmental decay is modeled as a superposition of two exponentials, fast and slow components with fractions $d^{\mathrm{f}}, d^{\mathrm{s}} > 0$; $d^{\mathrm{f}} + d^{\mathrm{s}} = 1$, and removal half-times $T^{\mathrm{f}}, T^{\mathrm{s}}$.

Illustration of the relative decrease of goudnshine dose from deposition of ^{134}Cs in time due to the radioactive decay and the environmental removal is in Figure 2.3. Radionuclide ^{134}Cs has the half-life of radioactive decay $T_{1/2} = 2.0648$ years. We see, that after 24 months, the exponential function describing the radioactive decay (green line) falls approximately to 0.5 of the initial value. The blue line represents the total ^{134}Cs removal due to the environmental removal (red line) and the radioactive decay. Values of coefficients in (2.30) were set to some reference values.
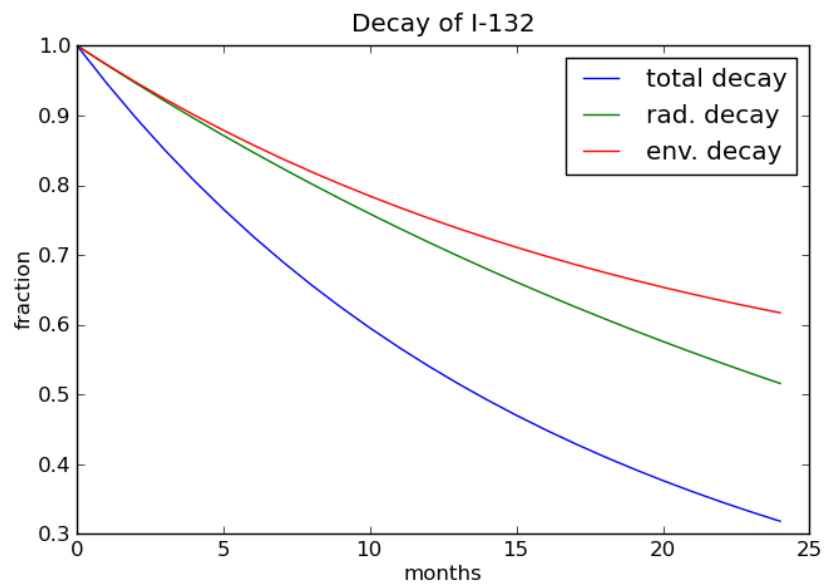
Figure 2.3: Illustration of the relative decrease of groundshine dose from deposition of ^{134}Cs in time due to the radioactive decay and the environmental removal.

Chapter 3

Function Fitting Methods

Principle of this class of methods is based on assumption, that the analysis can be expanded into a finite series of ordered mathematical basis functions with unknown expansion coefficients [Dal93]. This series is evaluated at the observation locations and a cost function measuring the distance between the observed values and the values given by the series is minimized. The common choice of the cost function is the squared distance between observations and the background field. Under this choice the data assimilation problem results in a linear relation for the unknown expansion coefficients, which can be determined upon solution of a systems of linear equations. The spatial analysis obtained is then obtained by evaluating the series at the set of analysis gridpoints forming an analysis grid.

Interpolation is the problem of approximating the value for a non-given point in some space when given some colors of points around (neighboring) that point.

3.1 Nearest-neighbor interpolation

Nearest-neighbor interpolation (also known as proximal interpolation, Zero-order interpolation or, in some contexts, point sampling) is a simple method of multivariate interpolation in one or more dimensions. The nearest neighbor algorithm selects the value of the nearest point and does not consider the values of neighboring points at all, yielding a piecewise-constant interpolant. The algorithm is very simple to implement and is commonly used (usually along with mipmapping) in real-time 3D rendering to select color values for a textured surface.

3.2 Bilinear interpolation

Bilinear interpolation is an extension of linear interpolation for interpolating functions of two variables (e.g., x and y) on a regular grid. Let $f = f(x, y)$ be a two-dimensional function known on a rectangular regular grid. Let's assume, that we want to find a value of $f(x, y)$ in an arbitrary point $P = (x_P, y_P)$. Let the grid points $Q_1 = (x_1, y_1)$, $Q_2 =$
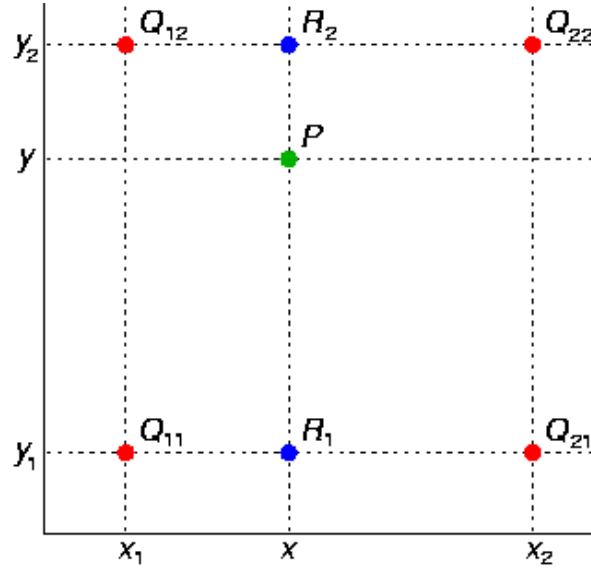
Figure 3.1: Illustration of principle of bilinear interpolation.

(x_2, y_2), $Q_3 = (x_3, y_3)$, $Q_4 = (x_4, y_4)$ be the points representing the grid-cell containing P, see Fig. 3.1.

The key idea is to perform linear interpolation first in one direction, and then again in the other direction. Although each step is linear in the sampled values and in the position, the interpolation as a whole is not linear but rather quadratic in the sample location:

$$f(R_1) \approx \frac{x - x_2}{x_1 - x_2} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}), \qquad \text{where} R_1 = (x, y_1),$$
$$f(R_2) \approx \frac{x - x_2}{x_1 - x_2} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}), \qquad \text{where} R_2 = (x, y_2).$$

Using $f(R_1)$ and $f(R_2)$, the approximation (estimate) of $f(P)$ is

$$f(P) \approx \frac{y_P - y_2}{y_1 - y_2} f(R_1) + \frac{y_P - y_1}{y_2 - y_1} f(R_2).$$

The general expression for bilinear interpolation formula according to scheme in Fig. 3.1 is as follows:

$$f(x, \ y) \approx \frac{f(Q_{11})}{(x_1 - x_2)(y_1 - y_2)}(x - x_2)(y - y_2) - \frac{f(Q_{21})}{(x_1 - x_2)(y_1 - y_2)}(x - x_1)(y - y_2) -$$

$$-\frac{f(Q_{12})}{(x_1 - x_2)(y_1 - y_2)}(x - x_2)(y - y_1) + \frac{f(Q_{22})}{(x_1 - x_2)(y_1 - y_2)}(x - x_1)(y - y_1).$$

If we choose a coordinate system, in which the four points where f is known, are $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, i.e., the function is known on a unit square, then the interpolation formula simplifies to
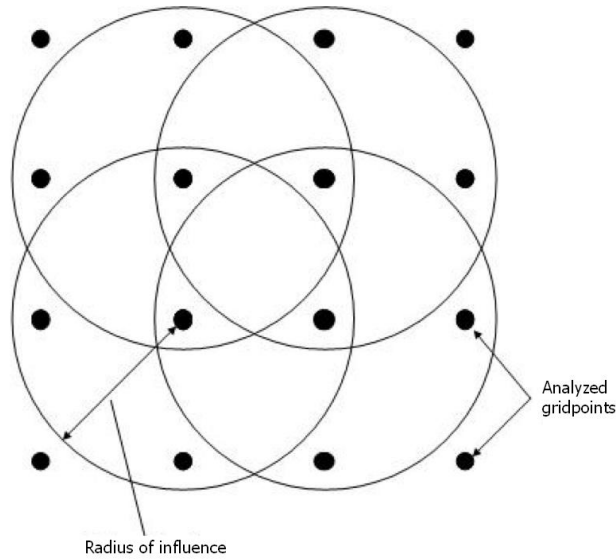
Figure 3.2: Illustration of principle of bilinear interpolation.

$$f(x,y) \approx f(0,0)(1-x)(1-y) + f(1,0)x(1-y) + f(0,1)(1-x)y + f(1,1)xy.$$

Or equivalently, in matrix form

$$f(x,y) \approx \begin{bmatrix} 1-x & x \end{bmatrix} \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1-y \\ y \end{bmatrix}.$$

3.3 Local polynomial fitting

Local polynomial fitting is another simple technique for spatial analysis. Contrary to the bilinear interpolation, polynomial fitting uses polynomial functions of a given order for approximation of an analyzed quantity given gridded values. Similarly to the bilinear interpolation, it is a local technique exploiting the data points close to the analyzed point. Since we are concerned with spatial analysis, the meaning of distance has the physical meaning of the real distance of two spatial locations in the Euclidean sense.

The principle is illustrated in Fig. 3.2. There are 16 analysis points in the figure - the points, where the value of an analyzed quantity is to be estimated. Coordinates of each of points are denoted with a vector $\mathbf{r}_i \in \mathbb{R}^2$, $i = 1,\ldots,16$. Around each point is drawn a circle called the radius of influence, R_i, and the area circumscribed is called the region of influence [Dal93]. The span of area of influence R_i determines the number of observations that influence the interpolated value of point $\mathbf{r}_i$. In other words, we assume that the analysis is at the point $\mathbf{r}_i$ is influenced by all the observations that lie within the region of influence of a particular gridpoint.

Following [Dal93], consider the spatial analysis of a function $f : \mathbb{R}^2 \to \mathbb{R}$. Define the local coordinate system (x, y) at the ith point such that $x = 0$, $y = 0$ at the gridpoint itself. Let the observation $f_o(x_k, y_k)$ fall into are of influence of the ith gridpoint, i.e., $x_k^2 + y_k^2 \leq R_i^2$, let the number of observations be K_i.

Define the analyzed quantity f within the ith region of influence as $f_a(x, y)$ and that it can be represented by a two-dimensional polynomial expansion of the form

$$f_a(x, y) = \sum_m \sum_n c_{mn} x^m y^n, \qquad (m + n \leq M), \qquad (m, n \geq 0), \qquad (3.1)$$

where the c_{mn} are the real expansion coefficients. The task is to determinate these coefficients using available observation $f_o(x_k, y_k)$, $k = 1, \ldots, K_i$. Given the coefficients, we can use the formula (3.1) to estimate value of f_a in an arbitrary location (x, y) within the region of influence, e.g. at the analysis gridpoint. Note that at the analysis gridpoint itself $x = y = 0$ and $f_a(0, 0) = c_{00}$. Form the following quadratic expression

$$I = \frac{1}{2} \sum_{k=1}^{K_i} \left[\sum_m \sum_n c_{mn} x^m y^n - f_o(x_k, y_k) \right]^2$$

by formal evaluation $f_a(x, y)$ at each of the observation stations in the ith region of influence. I in (3.2) is minimized by differentiating it with respect to each of the coefficients c_{mn} in turn and setting the results equal to zero:

$$\frac{\partial I}{\partial c_{mn}} = \sum_{k=1}^{K_i} x_k^m x_k^n \left[\sum_\mu \sum_\nu c_{\mu\nu} x_k^\mu y_k^\nu - f_o(x_k, y_k) \right] = 0, \qquad (3.2)$$

yielding

$$\sum_\mu \sum_\nu c_{\mu\nu} \sum_{k=1}^{K_i} x_k^{m+\mu} y_k^{n+\nu} = \sum_{k=1}^{K_i} x_k^m y_k^n f_0(x_k, y_k), \qquad (\mu + \nu \leq M)$$

for all $m + n \leq M$. μ and ν are iteration indices.

3.3.1 Cubic interpolation

Cubic interpolation is a special case of local polynomial fitting where we use all the polynomial terms up to order of 3. For sake o clarity, suppose that we want to interpolate on a unit square, where the function f_o and its derivatives $f_o^x = \frac{\delta f_o}{\delta x}$, $f_o^y = \frac{\delta f_o}{\delta x}$, $f_o^{x,y} = \frac{\delta^2 f_o}{\delta x \delta y}$ are known in corners $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. The interpolated surface can be then expressed using

$$f_a(x, y) = \sum_{m=0}^{3} \sum_{n=0}^{3} c_{mn} x^m y^n$$

for all $(x, y) \in \langle 0, 1 \rangle \times \langle 0, 1 \rangle$. The interpolation problem consists in determining 16 coefficients c_{mn}. Matching $f_a(x, y)$ with the function values yields four equations:

$$
\begin{aligned}
f_o(0, 0) &= f_a(0, 0) = c_{00}, \\
f_o(1, 0) &= f_a(1, 0) = c_{00} + c_{10} + c_{20} + c_{30}, \\
f_o(0, 1) &= f_a(0, 1) = c_{00} + c_{01} + c_{02} + c_{03}, \\
f_o(1, 1) &= f_a(1, 1) = \sum_{m=0}^{3} \sum_{n=0}^{3} c_{mn}.
\end{aligned}
$$

Likewise, eight equations for the derivatives in the x-direction and the y-direction:

$$
\begin{aligned}
f_o^x(0, 0) &= f_a^x(0, 0) = c_{10}, \\
f_o^x(1, 0) &= f_a^x(1, 0) = c_{10} + 2c_{20} + 3c_{30}, \\
f_o^x(0, 1) &= f_a^x(0, 1) = c_{10} + c_{11} + c_{12} + c_{13}, \\
f_o^x(1, 1) &= f_a^x(1, 1) = \sum_{m=1}^{3} \sum_{n=0}^{3} c_{mn} m.
\end{aligned}
$$

,

$$
\begin{aligned}
f_o^y(0, 0) &= f_a^y(0, 0) = c_{01}, \\
f_o^y(1, 0) &= f_a^y(1, 0) = c_{01} + c_{11} + c_{21} + c_{31}, \\
f_o^y(0, 1) &= f_a^y(0, 1) = c_{01} + 2c_{02} + c_{03}, \\
f_o^y(1, 1) &= f_a^y(1, 1) = \sum_{m=0}^{3} \sum_{n=1}^{3} c_{mn} n.
\end{aligned}
$$

And four equations for the cross derivatives xy:

$$
\begin{aligned}
f_o^{xy}(0, 0) &= f_a^{xy}(0, 0) = c_{11}, \\
f_o^{xy}(1, 0) &= f_a^{xy}(1, 0) = c_{11} + 2c_{21} + 3c_{31}, \\
f_o^{xy}(0, 1) &= f_a^{xy}(0, 1) = c_{11} + 2c_{12} + 3c_{13}, \\
f_o^{xy}(1, 1) &= f_a^{xy}(1, 1) = \sum_{m=1}^{3} \sum_{n=1}^{3} c_{mn} mn.
\end{aligned}
$$

The derivatives f_a^x, f_a^y, and f_a^{xy} in the expressions above are as follows:

$$f_a^x(x,y) \;=\; \sum_{m=1}^{3}\sum_{n=0}^{3} c_{mn} m x^{m-1} y^n,$$

$$f_a^y(x,y) \;=\; \sum_{m=0}^{3}\sum_{n=1}^{3} c_{mn} x^i n y^{n-1},$$

$$f_a^{xy}(x,y) \;=\; \sum_{m=1}^{3}\sum_{n=1}^{3} c_{mn} m x^{i-1} n y^{n-1}.$$

This procedure yields a surface $f_a(x,y)$ on the unit square which is continuous and with continuous derivatives. Bicubic interpolation on an arbitrarily sized regular grid can then be accomplished by patching together such bicubic surfaces, ensuring that the derivatives match on the boundaries. If the derivatives are unknown, they are typically approximated from the function values at points neighboring the corners of the unit square, e.g. using finite differences. Grouping the unknown parameters c_{mn} in a vector,

$$\mathbf{c} = \left[c_{00}, c_{10}, c_{20}, c_{30}, c_{01}, c_{11}, c_{21}, c_{31}, c_{02}, c_{12}, c_{22}, c_{32}, c_{03}, c_{13}, c_{23}, c_{33} \right]^{\mathrm{T}},$$

and

$$\mathbf{b} = \left[\, f_a(0,0)\ f_a(1,0)\ f_a(0,1)\ f_a(1,1)\ f_a^x(0,0)\ f_a^x(1,0)\ f_a^x(0,1)\ f_a^x(1,1)\ f_a^y(0,0)\ f_a^y(1,0)\ f_a^y(0,1)\ f_a^y(1,1)\ f_a^{xy}(0,0)\ f_a^{xy}(1,0)\ f_a^{xy}(0,1)\ f_a^{xy}(1,1) \,\right]$$

the problem can be reformulated into a system of linear equations $\mathbf{Ac} = \mathbf{b}$ where the inversion of A is:

$$\mathbf{A}^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 \\
-3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 \\
9 & -9 & -9 & 9 & 6 & 3 & -6 & -3 & 6 & -6 & 3 & -3 & 4 & 2 & 2 & 1 \\
-6 & 6 & 6 & -6 & -3 & -3 & 3 & 3 & -4 & 4 & -2 & 2 & -2 & -2 & -1 & -1 \\
2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
-6 & 6 & 6 & -6 & -4 & -2 & 4 & 2 & -3 & 3 & -3 & 3 & -2 & -1 & -2 & -1 \\
4 & -4 & -4 & 4 & 2 & 2 & -2 & -2 & 2 & -2 & 2 & -2 & 1 & 1 & 1 & 1
\end{bmatrix}.$$

3.4 Numerical examples - comparison of techniques for spatial interpolation
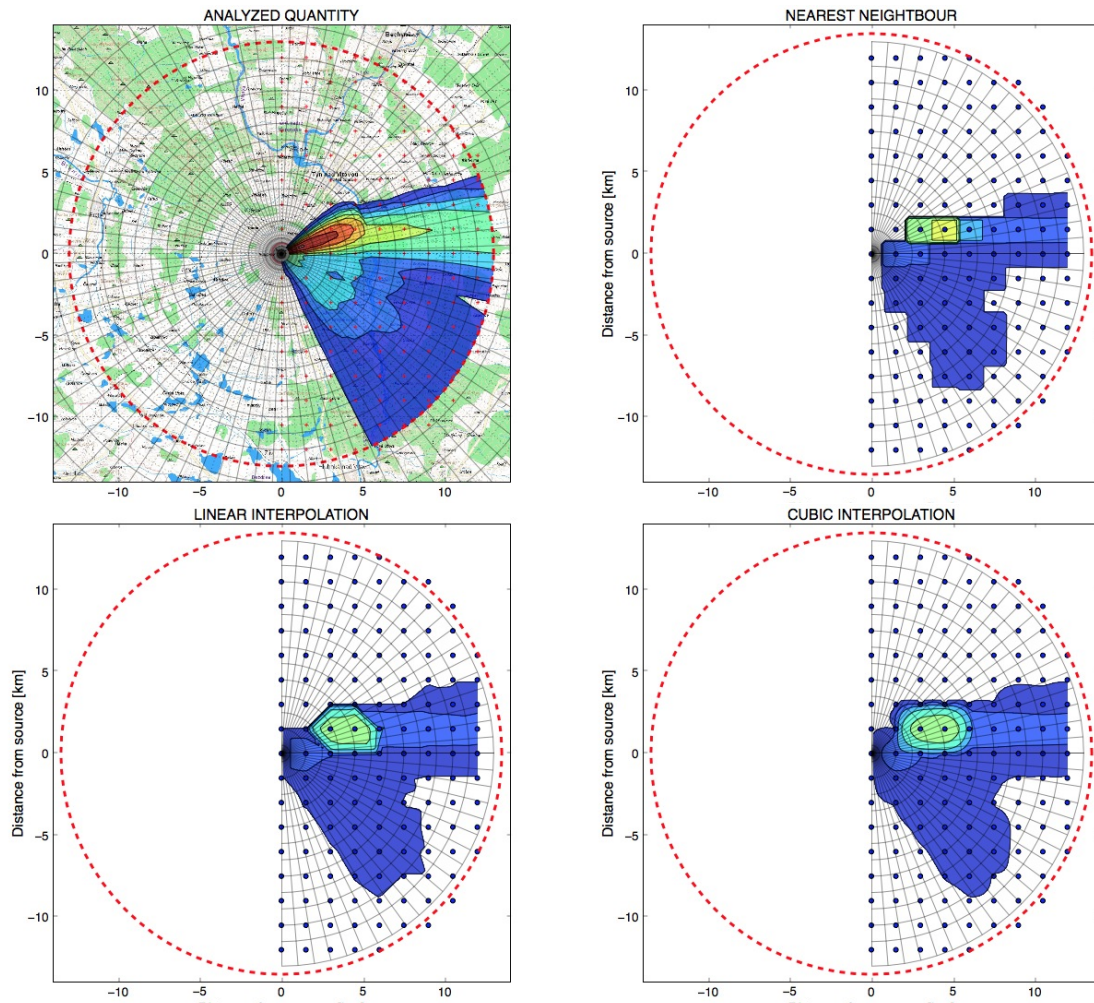
Figure 3.3: Interpolation results achieved with different interpolation methods. Top-left: the true deposition on terrain. Observations are sampled in locations delimited by red crosses. Interpolation results: Top-right: nearest neighbor; Bottom-left: bilinear interpolation; Bottom-right: bicubic interpolation.

Chapter 4

Empirical Interpolation Methods

Contrary to the function fitting methods relying only on observations, empirical interpolation methods use so called *background filed*, or *first guess* - a prior estimate of the analyzed quantity given by a mathematical model. In empirical interpolation, the background field is corrected using available observations. This class of methods thus represents the simplest inference methods merging multiple sources of information.

In case of a single observation, the basic principle of this class of methods can be expressed using formula

$$x_a = x_b + W(y - x_b), \tag{4.1}$$

where $\mathbf{x}_b$ is a forecast given by model - the *background field*, $\mathbf{x}_a$ the improved estimate - the *analysis*, $\mathbf{W}$ is an empirical *weight* and the term $\mathbf{x}_b - \mathbf{y}$ stands for the difference between model forecast and observed value - *residuum*. The difference between empirical interpolation methods is in the choice of $\mathbf{W}$ which determines how much the disagreement between the model and the observation influences the model.

In case of multiple observations and analysis on a grid, the principle is illustrated in Fig. 4.1. The analyzed point (a point from a polar network in this case) is denoted with the bold cross. The gray circle with radius R represent the area of influence. We assume that the analysis point is influenced with all the observations (denoted with plus-signs) in the area of influence. We process all the observation in a batch, the update equation (4.1) then becomes

$$x_a = x_b + \sum_{k=1}^{K} W_k(y_k - x_b),$$

where $\mathbf{y}_k$ is kth observation in area of influence corresponding to the analyzed point.

Empirical interpolation methods can be run in an assimilation cycle, where the analysis is gradually improved using a series of observations. Unser this iteration procedure the analysis $\mathbf{x}_a$ becomes the background for the next step:
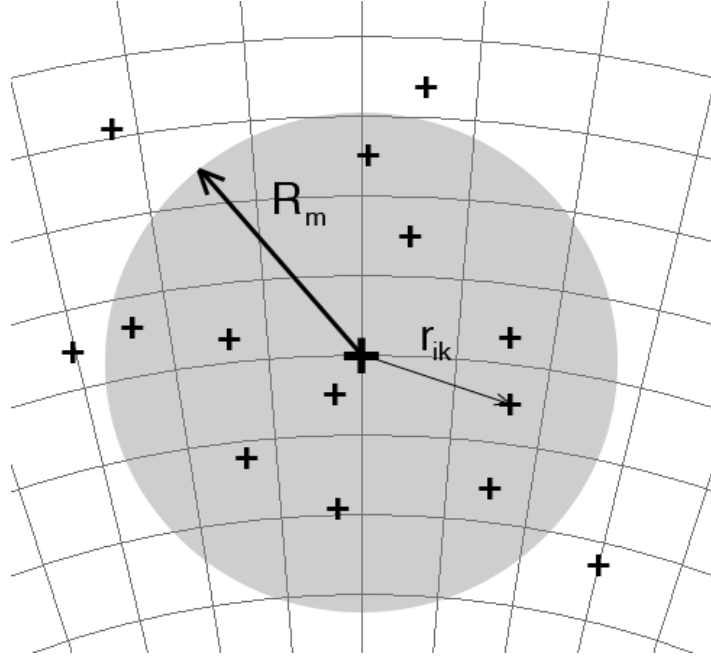
Figure 4.1: Illustration of principle of empirical interpolation methods.

$$x_{t+1}^a = x_t^b + \sum_{k=1}^{K} W_{k,t}(y_{k,t} - x_t^b),$$

where t is the iteration index (time index) and $\mathbf{y}_t$ is the observation corresponding to time t.

4.1 Cressman analysis

George Cressman developed the Cressman interpolation technique in 1959 [Coo92]. The technique interpolates station data to a user-defined latitude-longitude grid. Multiple passes are made through the grid at consecutively smaller radii of influence to increase precision. The radius of influence is defined as the maximum radius from a grid point to a station by which the observed station value may be weighted to estimate the value at the grid point. Stations beyond the radius of influence have no bearing on a grid point value. At each pass, a new value is calculated for each grid point based on its correction factor. This correction factor is determined by analyzing each station within the radius of influence. For each such station, an error is defined as the difference between the station value and a value arrived by interpolation from the grid to that station. A distance-weighted formula (shown below) is then applied to all such errors within the radius of influence of the grid point to arrive at a correction value for that grid point. The correction factors are applied to all grid points before the next pass
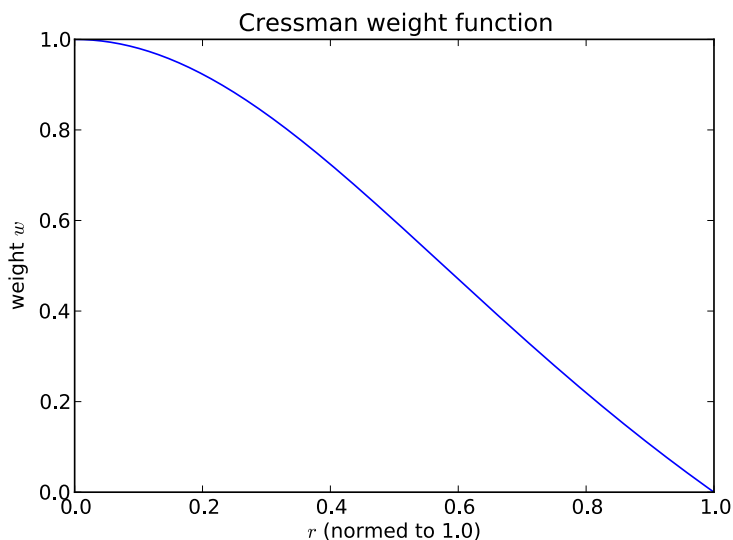
Figure 4.2: Weighting function of the Cressman analysis (R=1).

is made. Observations nearest the grid point carry the most weight. As the distance increases, the observations carry less weight. The Cressman function calculates the weights as follows:

$$W_k \quad = \quad \frac{R^2 - r_k^2}{R^2 + r^2}, \quad r^2 \leq R^2, \tag{4.2}$$

$$W_k \quad = \quad 0, \quad r_k^2 > R^2, \tag{4.3}$$

where R is the radius of influence and r_k is the distance between the observations location and the analyzed gridpoint.

As the radius of influence is tightened, results become more representative of the observed data. The analysis value at each gridpoint is calculated as the analysis value from the previous pass added to the sum of the products of the calculated weights and the difference between the actual station value and the interpolated background value at the station, divided by the sum of the weights.

Advantages:

- Simple and computationally fast (speed depends upon the number of scans).

- Generally more accurate than other simple methods such as linear interpolation.

Disadvantages:

- Can be unstable if grid density is higher than station density (i.e., more grid points than station data points).

- Sensitive to observational errors (random observation errors can generate unphysical features in analysis).

- Analysis may produce unrealistic extrema in the grid values, especially near the edges of the spatial domain.

- Does not account for the distribution of observations relative to each other.

- Consistency of the result with observations varies with observation (station) density.

- Optimum radii of influence have to be determined by trial and error.

4.2 Successive corrections methods (SCM)

SCM is a variant of Cressman analysis developed by Bergthorsson and Doos (1955) in Sweden. SCM is run in iterations, so it can be considered as a 4D data assimilation methods. Following [Kal03],iIn SCM assimilation cycle, the first estimate of the gridded field is given by the background field, $\mathbf{x}_0^a = \mathbf{x}_0^b$. After the first estimate, the following iterations are obtained by "successive corrections"

$$x_{t+1}^a = x_t^b + \frac{\sum_{k=1}^{K} w_{k,t}(y_{k,t} - x_t^b)}{\sum_{k=1}^{K} w_{k,t} + \epsilon^2},$$

where $y_{k,t}$is kth observation in the area of influence of the analyzed gridpoint in iteration t. This update is applied to all the analysis gridpoints in each iteration. The weights $w_{k,t}$ can be defined in different ways, e.g. using the Cressman's weights (4.2)(4.3). Different form of weight function proposed by Sasaki (1958) [Sas58] is

$$w_k = \exp\left(\frac{-r_k^2}{2R^2}\right).$$

In Fig. 4.3 we see that the weight falls to zero at some finite distance.

The radius of influence is allowed to vary with the iteration.

Although the SCM method is empirical, it is simple and economical, and it provides reasonable analyses. More details on SCM including derivation of its properties can be found in [Dal93].

4.2.1 Numerical example

We apply the SCM to on-step correction of a forecast of time integrated activity of I-131 with observations. TIC of I-131 is predicted on a polar network surrounding the source o pollution. Linear interpolation of TIC predicted in gridpoints of the polar network is in Fig. 4.4. Analysis is performed only on its subset denoted with the trapezoid.
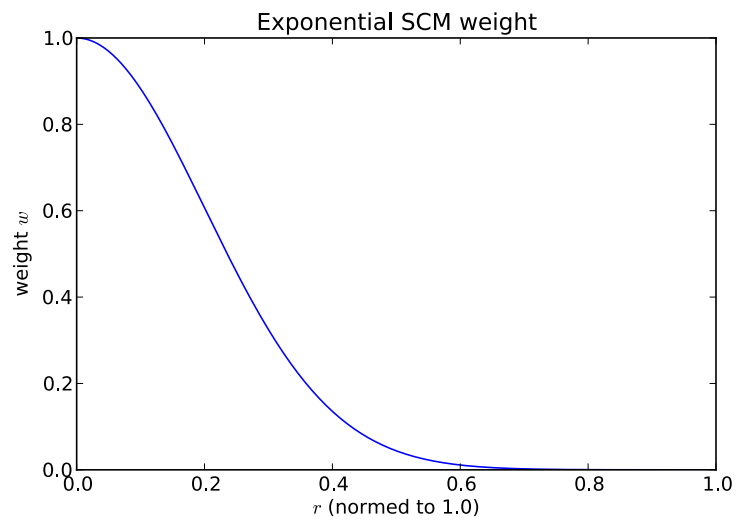
Figure 4.3: Weighting function proposed by Sasaki (1958) (R=0.2).
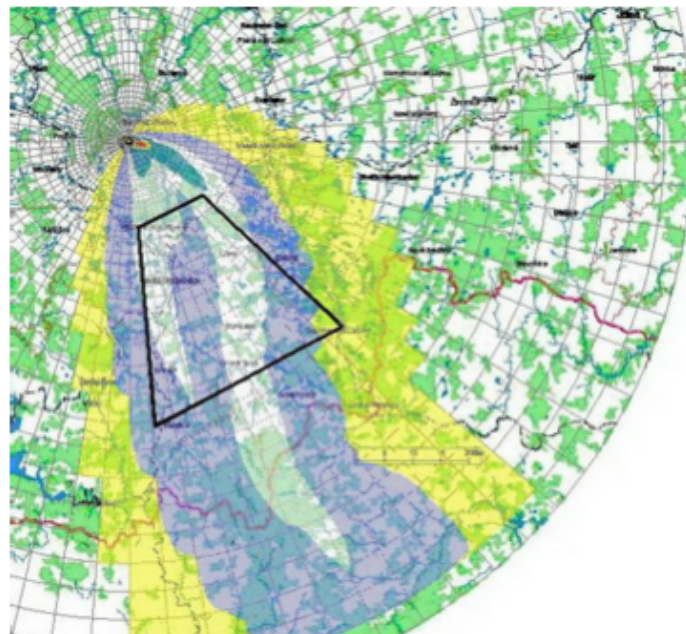
Figure 4.4: Linear interpolation of TIC predicted in gridpoints of the polar network.
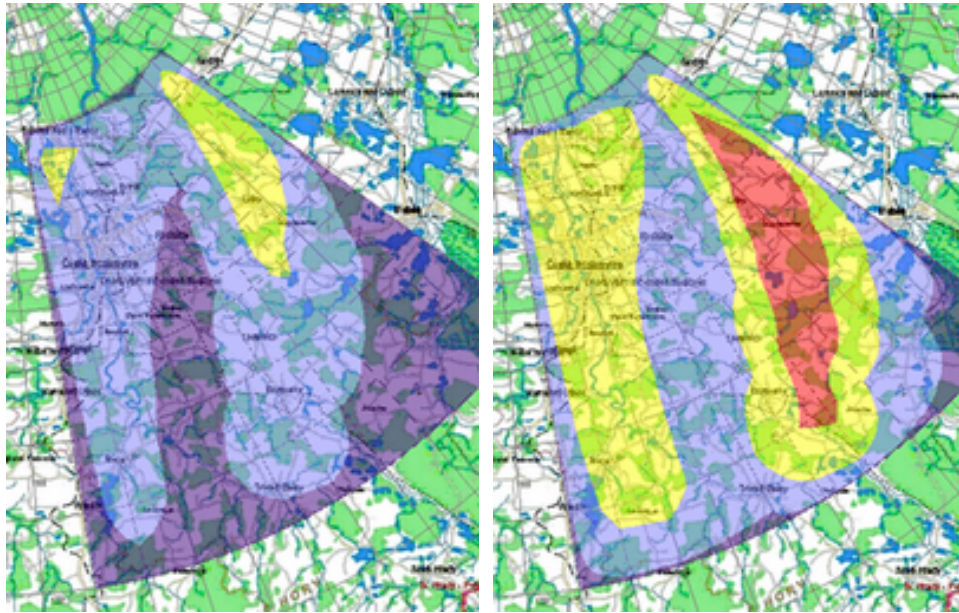
Figure 4.5: Background field (left) and the true spatial distribution of TIC used for simulation of observations (right).

In Fig. 4.5 we see the background field given on a polar network and the twin model representing the true spatial distribution of TIC.

Firstly, we illustrate the effect of choice of the radius of influence. In Fig. 4.6 we see assimilation result (analysis) using three observations after one step of SCM with different value of radius of influence: R=1, 3, 5, 10, 15 and 20 km. Observation locations are denoted with red dots. We see that the fact that all the observations are located in one place caused that the assimilation result differs from the twin model.

Secondly, we show that with more observations regularly covering the analyzed area we can achieve the good agreement of the analysis and the twin model. The result for radius of inference r=3km is in Fig. 4.7.

We observe, that for appropriate choice of radius of influence can SCM deliver good results, however the value of r is usually not known and must be experimentally tuned. This implies that SCM is not suitable for operational application regarding decision support in case of accident in a nuclear power plant.
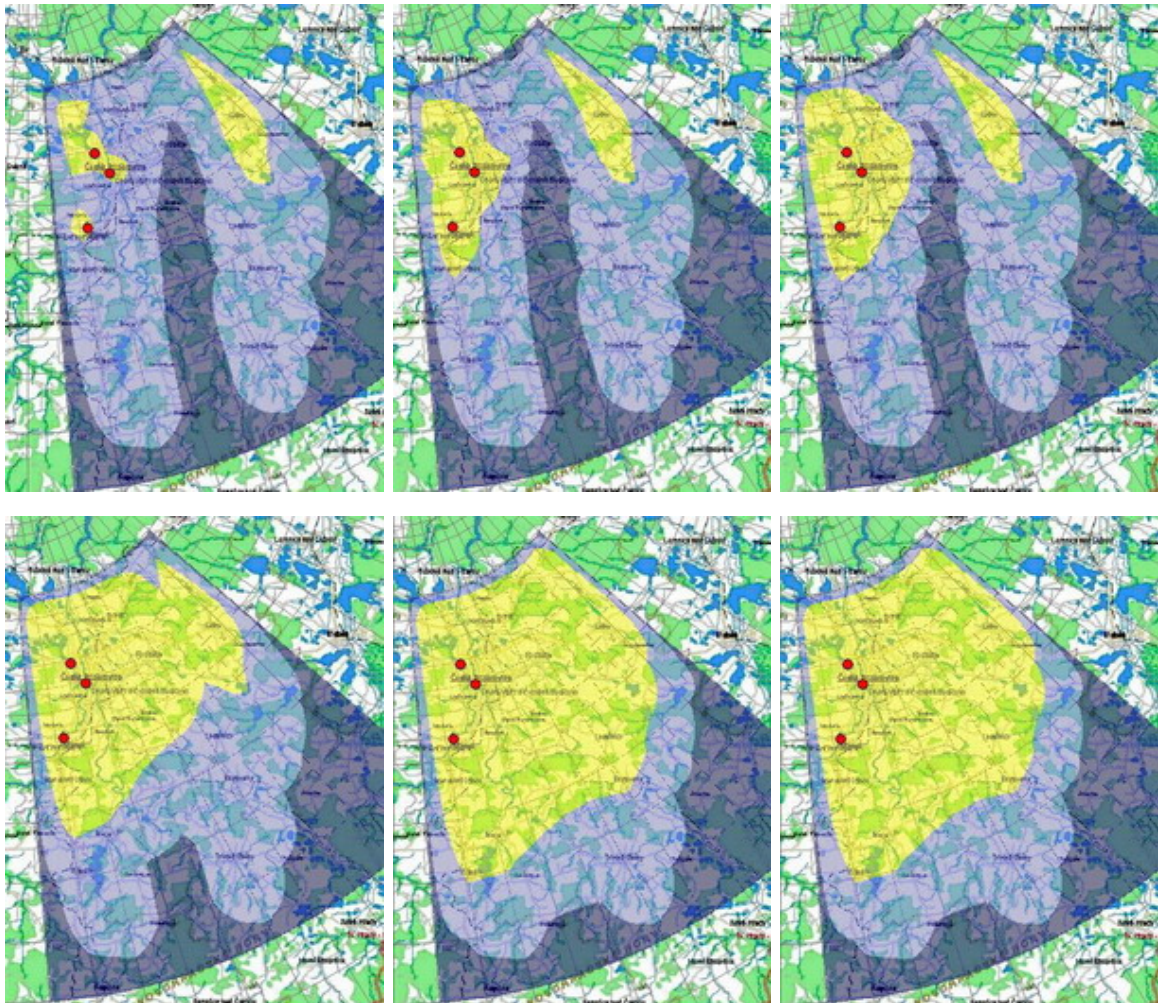
Figure 4.6: SCM: Illustration of choice of radius of influence. From left to right, top to down: R=1, 3, 5, 10, 15 and 20 km. Observation locations are denoted with red dots.
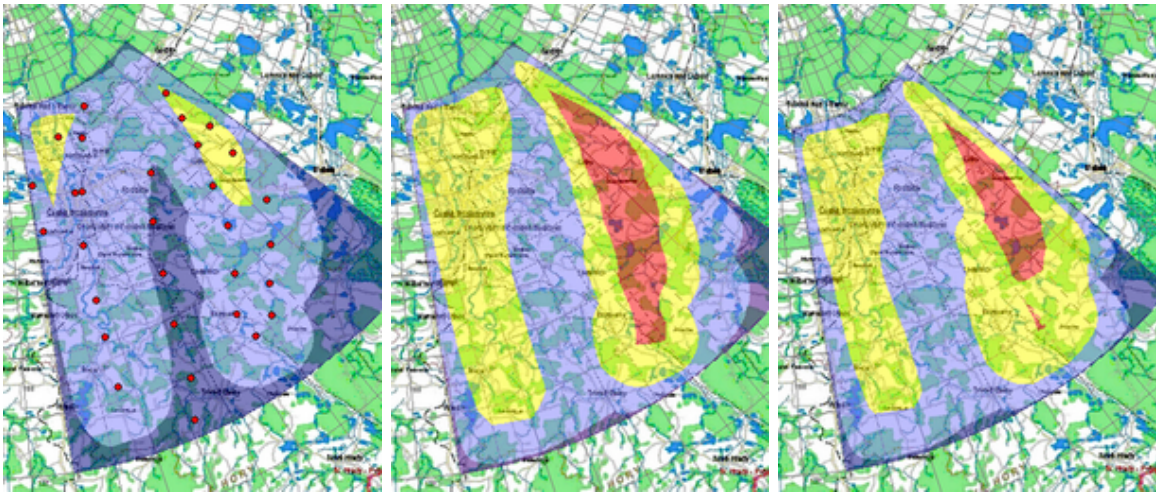
Figure 4.7: Illustration of SCM with more observations. From left to right: background field, twin model, assimilation result after one step of SCM with R=3km.

Chapter 5

Constant statistical methods

This class of methods comes out from the least squares, which was in environmental analysis (meteorology) popularized by Lev Gandin. Methods minimize the analysis variance—expected value of square of difference between analysis and the true value. Since this class of methods accounts for error statistics of the background field and the observations, the methods can be considered as statistical, however, we call them constant because the error statistics are not evolved in time using model equations.

5.1   Scalar least squares (LS) method

Following [Kal03], we illustrate the method on a one-dimensional example. This example will serve as an introduction to statistical estimation.

Let T_1 and T_2 be two independent observations of a scalar quantity T. Both the observations have errors ϵ_1 and ϵ_2, respectively. Let operator $\mathbb{E}[\cdot]$ denote the operator of expected value. We assume that our measuring instrument is unbiased, i.e.

$$\begin{aligned} \mathbb{E}[T - T_1] &= \mathbb{E}[\epsilon_1] = 0, \\ \mathbb{E}[T - T_2] &= \mathbb{E}[\epsilon_2] = 0. \end{aligned}$$

Let the variances of out observations be:

$$\mathbb{E}[\epsilon_1^2] = \sigma_1^2, \quad \mathbb{E}[\epsilon_2^2] = \sigma_2^2,$$

and let the errors by uncorellated, ii.e.,

$$\mathbb{E}[\epsilon_1 \epsilon_2] = 0.$$

We assume that the true value of T can be estimated as a linear combination of T_1 and T_2:

$$T_a = a_1 T_1 + a_2 T_2.$$

Since we want the analysis T_a to be unbiased, i.e. $\mathbb{E}[T_a] = \mathbb{E}[T]$, the coefficients a_1 and a_2 must fulfill the condition

$$a_1 + a_2 = 1. \tag{5.1}$$

The best estimate T_a of T minimizes the variance of T_a given by:

$$
\begin{aligned}
\sigma_a^2 &= \mathbb{E}[(T_a - T)^2] \\
&= \mathbb{E}[(a_1 T_1 + a_2 T_2 - (a_1 + a_2)T)^2] \\
&= \mathbb{E}[a_1^2(T_1 - T)^2 + 2a_1 a_2(T_1 - T)(T_2 - T) + a_2^2(T_2 - T)^2] \\
&= \mathbb{E}[a_1^2 \epsilon_1^2 + 2a_1 a_2 \epsilon_1 \epsilon_2 + a_2^2 \epsilon_2^2].
\end{aligned}
$$

Exploiting linearity of $\mathbb{E}[\cdot]$ we obtain

$$\sigma_a^2 = a_1^2 E(\epsilon_1^2) + 2a_1 a_2 E(\epsilon_1 \epsilon_2) + a_2^2 E(\epsilon_2^2) = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2. \tag{5.2}$$

Minimizing 5.2 with respect to constrain 5.1 using method of Langrange's multiplicators yields

$$
\begin{aligned}
\Phi &= \sigma_a^2 + \lambda(1 - a_1 - a_2) \\
&= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + + \lambda(1 - a_1 - a_2).
\end{aligned}
$$

Differentiation with respect to a_1 and a_2 give us:

$$\frac{\partial \Phi}{\partial a_1} = 2a_1 \sigma_1 - \lambda \equiv 0, \tag{5.3}$$

$$\frac{\partial \Phi}{\partial a_2} = 2a_2 \sigma_2 - \lambda \equiv 0. \tag{5.4}$$

Solving linear system (5.3)–5.4 for a_1, a_2 gives us

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \qquad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}, \tag{5.5}$$

or equivalently,

$$a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2}, \qquad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}.$$

These coefficients are called optimal weights and minimize expected variance of analysis σ_a. Substituting the coefficients into 5.2 we obtain

$$\sigma_a^2 = \frac{\sigma_2^4}{(\sigma_1^2 + \sigma_2^2)^2} + \frac{\sigma_1^4}{(\sigma_1^2 + \sigma_2^2)^2},$$

implying

$$\sigma_a^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2},$$

or equivalently,

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}.$$

5.1.1   Example of scalar LS application

In this sections we assume that T_1 is a forecast of the scalar quantity (given by a mathematical model) and T_2 is corresponding observation, i.e. $T_1 = T_b$, $T_2 = T_o$. It holds that

$$T_a = a_1 T_b + a_2 T_o.$$

Using (5.5) we can show that

$$T_a = (1 - \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}) T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} T_o,$$

yielding

$$T_a = T_b + W(T_o - T_b),$$

where the difference of background and observation $T_b - T_o$ is called *innovation* and W is the *optimal innovation weight*:

$$W = \sigma_b^2 (\sigma_b^2 + \sigma_o^2)^{-1}.$$

The variance of analysis is then

$$\begin{aligned}
\sigma_a^2 &= (\sigma_b^{-2} + \sigma_o^{-2})^{-1} \\
&= \sigma_b^2 \sigma_o^2 / (\sigma_b^2 + \sigma_o^2) \\
&= (1 - W)\sigma_b^2.
\end{aligned}$$

5.2 Optimal interpolation (OI)

OI is a multivariate generalization of the least squares for spatial analysis. Following [Kal03], let's assume a data assimilation problem of finding optimum analysis of a field of model variables aggregated in a vector $\mathbf{x}_a$, given a background field $\mathbf{x}_b$ available on a analysis grid in two- or three- dimensional physical space, and a set of p irregularly spaced observations $\mathbf{y}_o$.

The true value $\mathbf{x}_t$ is estimated using OI as

$$
\begin{aligned}
\mathbf{x}_t &\approx \mathbf{x}_a = \mathbf{x}_b + \mathbf{W}[\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b)], \\
\boldsymbol{\epsilon}_a &= \mathbf{x}_t - \mathbf{x}_a,
\end{aligned}
$$

where $\mathbf{x}_a$, $\mathbf{x}_b$ are vectors of analysis and background, respectively, $\mathbf{W}$ is the *gain matrix*, $\mathbf{y}_o$ is a vector of available observations. In previous sections we assumed that the analyzed quantity is the same as the observed quantity. This is not a general case, where the quantities can differ. Generally nonlinear *observation operator* $\mathcal{H}(\cdot)$ transforms the vectors from space of model into the space of observations and make modeled and measured data thus comparable.

5.2.1 Best linear unbiased estimate

First describe the construction of the best linear unbiased estimator (multidimensional linear regression). Assume two vectors

$$
\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.
$$

centered about their mean values: $E(\mathbf{x}) = 0$ a $E(\mathbf{y}) = 0$. We would like to find a matrix $\mathbf{W}$ such that

$$
\begin{aligned}
\mathbf{x}_t &\approx \mathbf{x}_a = \mathbf{W}\mathbf{y} \\
&= \mathbf{W}\mathbf{y} - \boldsymbol{\epsilon}_a.
\end{aligned}
$$

Here, $\boldsymbol{\epsilon}_a$ is the regression error and $\mathbf{W}$ is a matrix minimizing the mean square error $\mathbb{E}[\boldsymbol{\epsilon}_a^{\mathrm{T}}\boldsymbol{\epsilon}_a]$.

Explicit form of regression equation is

$$
x_i = \sum_{k=1}^{p} w_{ik} y_k - \epsilon_i.
$$

The mean square error over all the elements of $\mathbf{x}$ and $\mathbf{y}$ is given by

$$\sum_{i=1}^{n} \epsilon_i^2(t) = \sum_{i=1}^{n} [\sum_{k=1}^{p} w_{ik} y_k(t) - x_i(t)]^2,$$

and the derivatives with respect to the weight matrix components is

$$\frac{\partial \sum_{i=1}^{n}}{\partial w_{ij}} = 2[\sum_{k=1}^{p} w_{ik} y_k(t) - x_i(t)][y_j(t)],$$

$$= 2[\sum_{k=1}^{p} w_{ik} y_k(t) y_j(t) - x_i(t) y_j(t)].$$

In matrix form

$$\frac{\partial \boldsymbol{\epsilon}^T \boldsymbol{\epsilon}}{\partial w_{ij}} = -2\{[\mathbf{W}\mathbf{y}(t)\mathbf{y}^T(t)]_{ij} - [\mathbf{x}(t)\mathbf{y}^T(t)]_{ij}\}.$$

The optimal value of $\mathbf{W}$ minimizing expected value of $\boldsymbol{\epsilon}$ is obtained from

$$\mathbf{W}\mathbb{E}[\mathbf{y}\mathbf{y}^T] - \mathbb{E}[\mathbf{x}\mathbf{y}^T] = 0,$$

yielding *normal equations*

$$\mathbf{W} = \mathbb{E}(\mathbf{x}\mathbf{y}^T)[\mathbb{E}(\mathbf{y}\mathbf{y}^T)]^{-1}. \tag{5.6}$$

5.2.2   Optimal value of gain matrix W in OI

To derive the form of the optimal gain matrix, we define the background errors and the analysis errors as follows:

$$\begin{aligned} \boldsymbol{\epsilon}_b(x,\ y) &= \mathbf{x}_b(x,\ y) - \mathbf{x}_t(x,\ y), \\ \boldsymbol{\epsilon}_a(x,\ y) &= \mathbf{x}_a(x,\ y) - \mathbf{x}_t(x,\ y). \end{aligned} \tag{5.7}$$

Similarly, we define the observation errors

$$\boldsymbol{\epsilon}_o = \mathbf{y_o} - \mathcal{H}(\mathbf{x}_t).$$

The background and observation errors are assumed to be unbiased

$$\mathbb{E}[\boldsymbol{\epsilon}_b] = \mathbb{E}[\mathbf{x}_b] - \mathbb{E}[\mathbf{x}_t] = 0, \quad \mathbb{E}[\boldsymbol{\epsilon}_o] = \mathbb{E}[\mathbf{y}_o] - \mathbb{E}[\mathbf{y}_t] = 0,$$

and mutually uncorellated, i.e. $\mathbb{E}[\boldsymbol{\epsilon}_o \boldsymbol{\epsilon}_b^{\mathrm{T}}] = 0$.

Using the errors, we define the error covariance matrices of the analysis, background and observations respectively:

$$\mathbf{P}_a = \mathbf{A} = \mathbb{E}[\boldsymbol{\epsilon}_a \boldsymbol{\epsilon}_a^T],$$
$$\mathbf{P}_b = \mathbf{B} = \mathbb{E}[\boldsymbol{\epsilon}_b \boldsymbol{\epsilon}_b^T],$$
$$\mathbf{P}_o = \mathbf{R} = \mathbb{E}[\boldsymbol{\epsilon}_o \boldsymbol{\epsilon}_o^T].$$

Innovation vector is defined as the difference between the background and observation:

$$
\begin{aligned}
\mathbf{d} &= \mathbf{y}_o - H(\mathbf{x}_b) = \mathbf{y}_o - H(\mathbf{x}_t + (\mathbf{x}_b - \mathbf{x}_t)), \\
&= \mathbf{y}_o - H(\mathbf{x}_t) - \mathbf{H}(\mathbf{x}_b - \mathbf{x}_t) = \boldsymbol{\epsilon}_o - \mathbf{H}\boldsymbol{\epsilon}_b.
\end{aligned}
$$

If we write approximation

$$\mathbf{x}_t - \mathbf{x}_b = \mathbf{W}\mathbf{d},$$

we can use (5.6) to derive the optimal $\mathbf{W}$:

$$
\begin{aligned}
\mathbf{W} &= \mathbb{E}\{(\mathbf{x}_t - \mathbf{x}_b)[\mathbf{y}_o - H(\mathbf{x}_b)]^T\}(\mathbb{E}\{[\mathbf{y}_o - H(\mathbf{x}_b)][\mathbf{y}_o - H(\mathbf{x}_b)]^T\})^{-1} \\
&= \mathbb{E}[(-\boldsymbol{\epsilon}_b)(\boldsymbol{\epsilon}_o - \mathbf{H}\boldsymbol{\epsilon}_b)^T]\{\mathbb{E}[(\boldsymbol{\epsilon}_o - \mathbf{H}\boldsymbol{\epsilon}_b)(\boldsymbol{\epsilon}_o - \mathbf{H}\boldsymbol{\epsilon}_b)^T]\}^{-1},
\end{aligned}
$$

yielding

$$\mathbf{W} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1},$$

where $\mathbf{B}$ is error covariance matrix of background field, $\mathbf{R}$ is error covariance matrix of observations and $\mathbf{H}$ is matrix of a linear observation operator. If $\mathbf{H}$ is not linear, it must be linearized using its Jacobian, i.e. the elements of $\mathbf{H}$ are given by

$$h_{ij} = \frac{\partial H_i}{\partial x_j}.$$

The error covariance matrix of resulting analysis if given by

$$\mathbf{A} = (\mathbf{I} - \mathbf{W}\mathbf{H})\mathbf{B}.$$

5.2.3 Remarks

Usually, we assume that the error covariance matrix of observations has the form

$$
\mathbf{R} = \sigma_o^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},
$$

i.e., the observation error is the same for all observations.

A reasonable form of background error covariance matrix can be constructed as

$$\mathbf{B} = \sigma_b^2 \begin{pmatrix} 1 & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & 1 & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & 1 \end{pmatrix},$$

where function $\gamma_{ij} = \gamma_{ij}(\mathbf{r}_i - \mathbf{r}_j)$ is a homogeneous and isotropic correlation function. The covariances are artificially introduced into the matrix according to paradigm, that the covariance is proportional to the distance between the locations. Suitable functions modeling this relation can be for example the following:

$$\begin{aligned} \gamma_{ij} &= \exp\left(-\frac{r_{ij}^2}{L}\right), \\ \gamma_{ij} &= \left(1 + \frac{r_{ij}}{L}\right)\exp\left(-\frac{r_{ij}}{L}\right), \end{aligned} \tag{5.8}$$

where r_{ij} is the Euclidean distance between locations and L is a length scale parameter. More on construction of covariance matrices can be found in [GC99].

5.2.4 Uncertainty study - an alternative way for estimation of error covariance's using physical model

Uncertainty study performed on a physical model allows us to estimate its error covariance structure [IH05]. Assuming that the distribution of error is Gaussian, we have to determine the variance. It can be evaluated as a sample covariance of a set of realizations generated by multiple model runs with different initialization. Model inputs can be treated as random with given probability density function or they can be regarded as deterministic and set up to their "best expert estimate".

We illustrate this approach on an atmospheric dispersion model (ADM) having 14 parameters [HPP08b]. The significant difference between treatment of ADM input parameters (random vs. "best estimate") is demonstrated in the following example. Let's have two different release scenarios:

1. First scenario S-14: fourteen parameters of AMD is treated as random (intensity of release, horizontal dispersion, wind direction, dry and wet deposition, rain intensity etc.)

2. Second scenario S-1: the only parameter is random – the rain intensity between the fifth and sixth hour of release

In Fig. 5.1-left, prior mean value for S-14 is is visualized. In the same figure, right side, we see mean value for S-1. Each row or column of covariance matrix is covariance of value in certain spatial location with the rest. In Fig. 5.3 and 5.2 are shown

visualizations of prior covariance to S-14 and S-1, respectively. In the left half of both the figures is covariance of all the spatial locations with point [14; 5] (the point is close to the source under the plume axis, in front of the rain zone). On the right side is covariance of all the spatial locations with point [22; 25] (the point is inside the rain zone). Both the reference points are denoted by blue squares.

In Fig. 5.3 we see covariance of a location close to the source under the plume axis and another inside the rain zone (the circle–shaped area approximately in the middle of the trace). Rain makes depletion process more intensive and speeds up the deposition. It is noticeable, that the "shape" of covariance structure for both the locations corresponds with the "shape" of mean value. This is because of predominant influence of parameters having global effect (for example the intensity of release).

As in the second scenario S-1 is the only random parameter the rain intensity, for all the members of ensemble is the tract of release the same until the rain zone. It means, that the covariance of location before the rain zone with the rest of polar network is zero (as can be seen in the Fig. 5.2-left). On the other hand, there are strong covariances of the points inside and behind the zone. It represents analogy with the law of conservation of released activity (if we neglect radioactive decay). The more activity is depleted due to the rain, the less radioactivity can be depleted behind the rain zone (Fig. 5.3-right). Interesting is also the fact, that the area of locations with nonzero covariances is much smaller than in the first scenario. This is due to fact, that the wind direction was treated as a fixed value and not as a random parameter. It can be interpreted as the prior estimate based on S-1 says that the area affected by radioactive pollution is smaller than the area estimated upon S-14. The rain intensity is a typical example of a random parameters having local effect.

5.2.5 Numerical example

OI is demonstrated on assimilation of spatial distribution of deposition on terrain.

5.2.5.1 Influence of length-scale parameter **L**

In this example we compare assimilation results for two matrices **B** constructed using (5.8) with different values of length scale parameter L. In Fig. 5.4we see the background field and observations locations. Again, we assume that our computational domain is a subset of a polar network comprising of $12 \times 5 = 60$ computational nodes. In Fig 5.5 we see assimilation results for different values of L.

5.2.5.2 Influence of variance of background error

Here we examine the influence of model error σ_b^2. In Fig. 5.6 we see the background field and the only measurement assumed in this example. In Fig. 5.7 are assimilation results for different values of σ_b^2. Twin model is the same as in the previous example. The smaller the background field variance, the better correspondence of the assimilated result with the twin model.
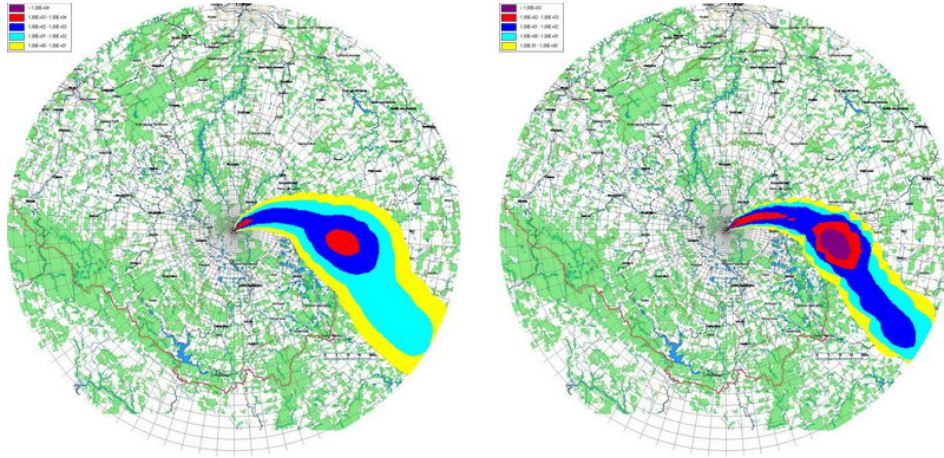
Figure 5.1: Mean values for S-14 (left) and S-1 (right).
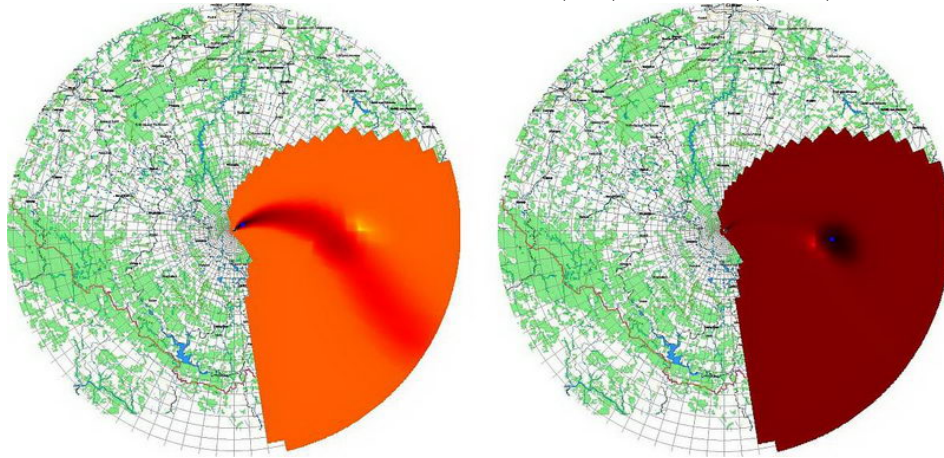
Figure 5.2: S-14: Covariance of selected points with the rest of polar network.
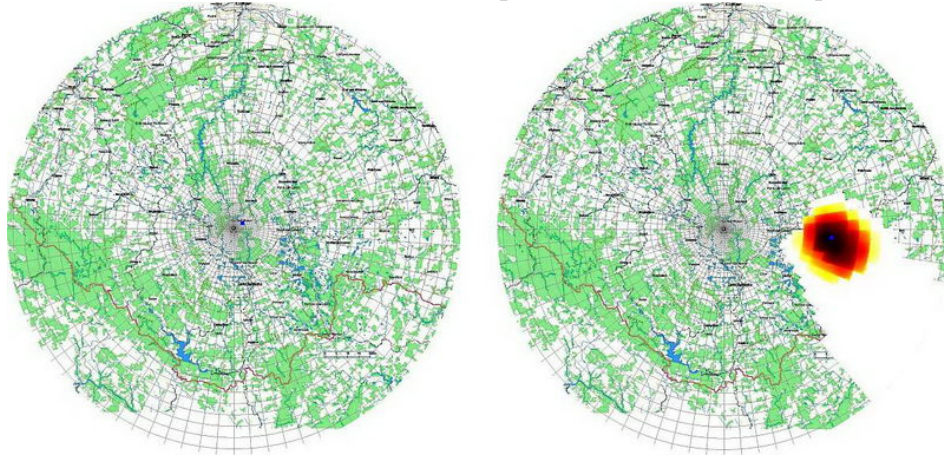
Figure 5.3: S-1: Covariance of selected points with the rest of polar network.
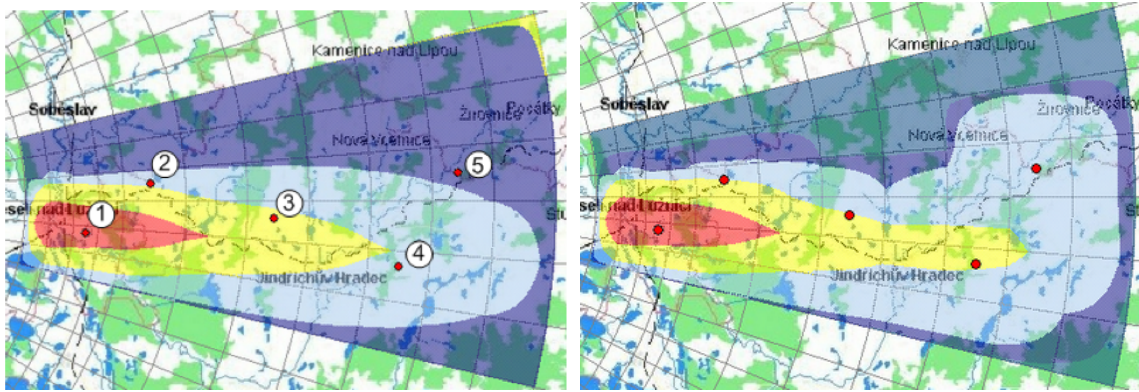
Figure 5.4: Background field and observations locations (left) and the twin model (right).
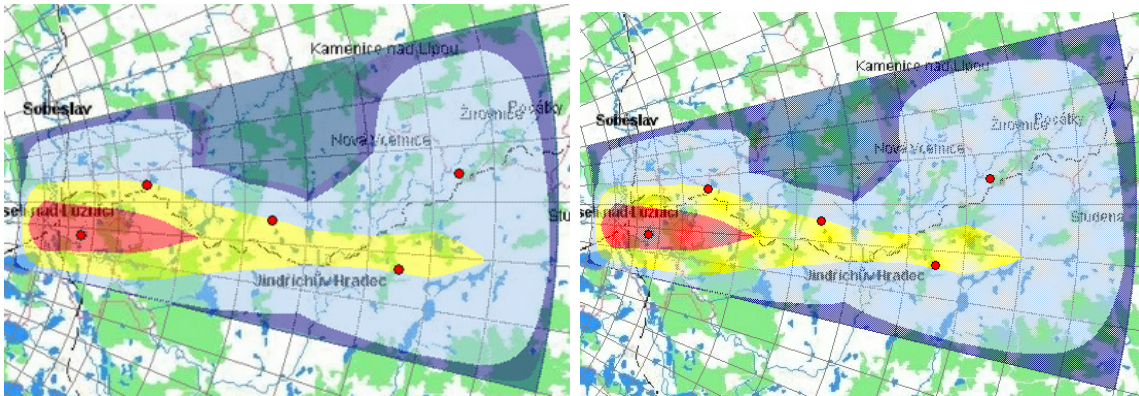
Figure 5.5: Assimilation results for different values of L. left: L=1.0E+05, right: L=1,0E+08.
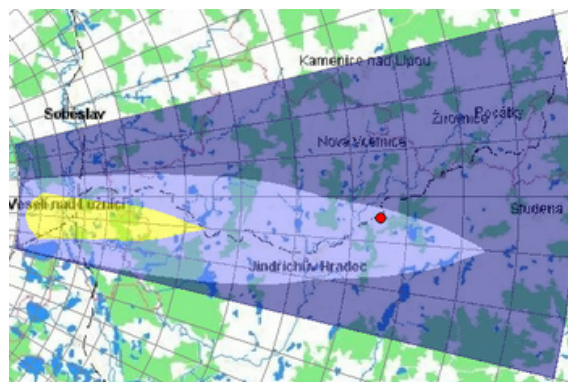
Figure 5.6: Background field and the only observation location.
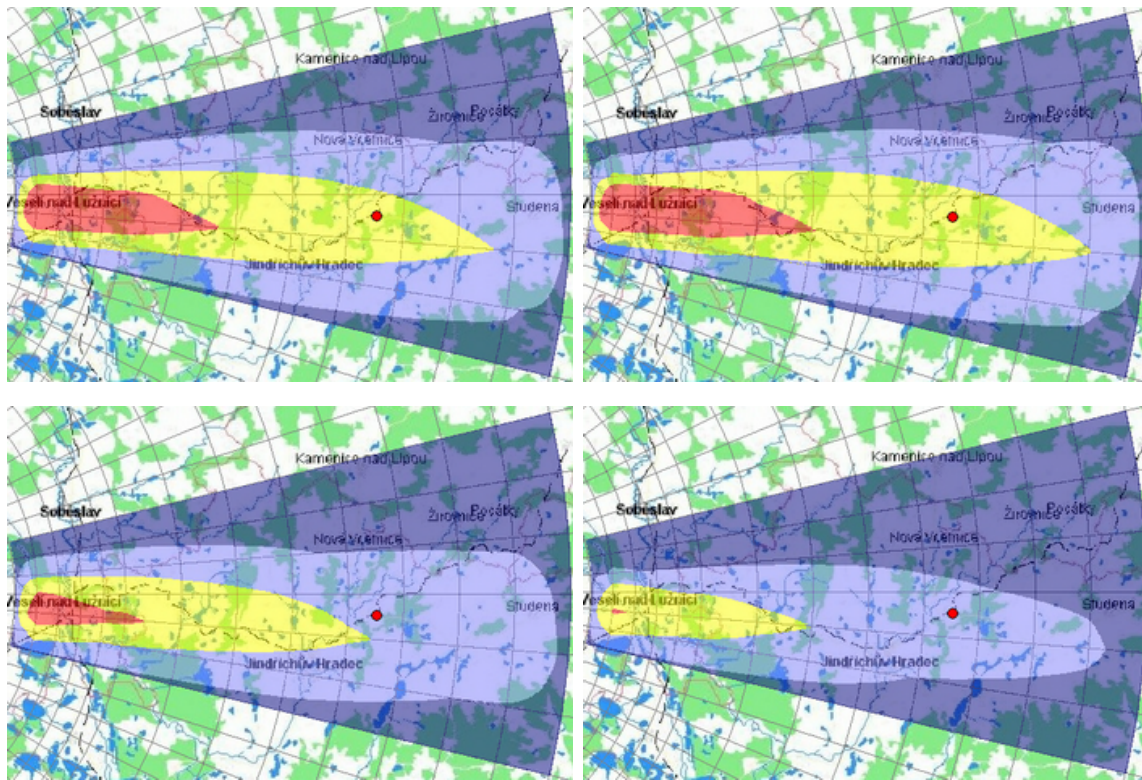
Figure 5.7: Background field and the only observation location. From top to bottom, from left to right: σ_b=1.0E+12, 1.0E+13, 1.0E+14, 1.0E+15. The smaller the background field variance, the better correspondence of the assimilated result with the twin model.
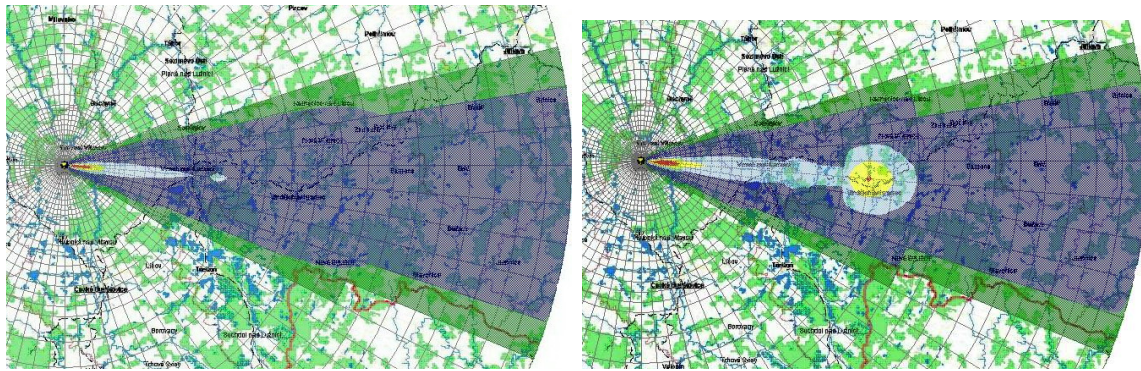
Figure 5.8: Background field and twin model.

5.2.5.3 Correction of misspecified deposition due to rain during the plume phase

In this experiment we assume an atmospheric release in the eastward direction, see Fig. 5.8. In the figure-top is the background field, the first estimate of the deposition given by an atmospheric dispersion model. In the figure-bottom we see the twin model used for simulation of observations. We see that the deposition much higher due to the rain zone between km 40 to 50 with rain intensity $2mm/hod$. We assume, that we sampled measurement just in the middle of the rain area. The measurement tells us, that the reposition was in reality much higher than tells us the dispersion model. We use OI to correct whole prediction using this measurement.

Assimilation results are in Fig. 5.9. We examine deposition values under the axis of the plume, i.e. just in the eastward direction. To demonstrate the advantages of the statistical methods we compare the results with results achieved using SCM - Fig. 5.9-top. We see that the assimilation procedure increases deposition in place of observation (according to selected radius of influence), however, after the rain zone the deposition remains the same as before assimilation procedure.

In Fig. 5.9-bottom we see results for OI. It performs much better, since it correctly accounts for decrease of deposited activity *behind* the rain zone. If the deposition is increased in the zone, it must best decreased behind the zone in order to fulfill the law of activity conservation. The relations between the background field gridpoints is encoded in the background field error covariance matrix (matrix S1 in Section 5.2.4), which is not used in SCM. SCM is "blind" to the local effect of wet deposition in the rain zone.
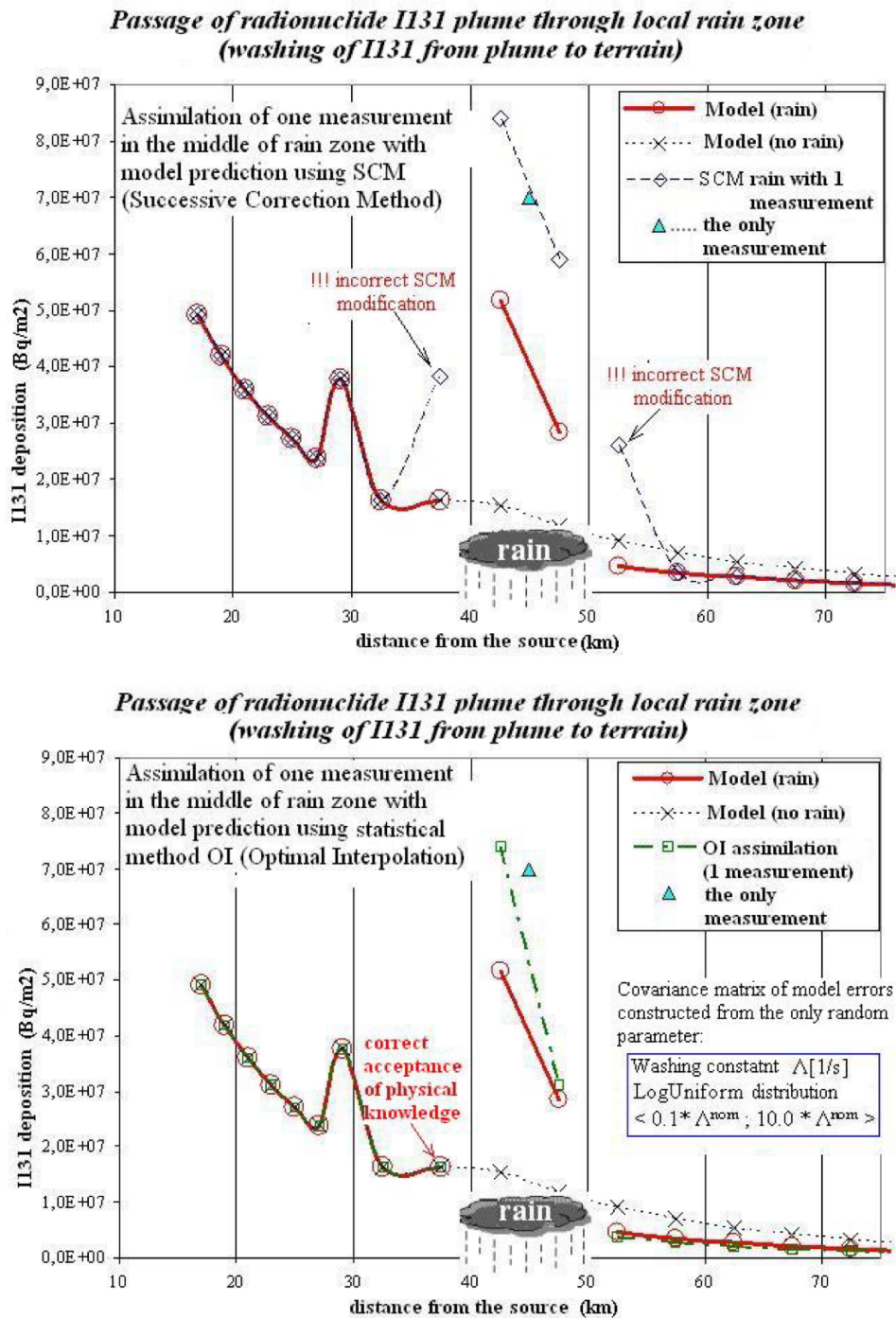
Figure 5.9: Assimilation results for the scenario with a rain zone. Top: results achieved using SCM. Bottom: results achieved using OI.

Chapter 6

Data assimilation as optimization task

6.1 Introduction

In this chapter we show that data assimilation can be also understood as an optimization task, where we optimize values of parameters of a parameterized model in order to obtain a good fit of observations. During assimilation we assume precise measurements and thus the procedure cannot be presented as pure statistical DA. On the other hand it requires proper environmental model which describes uncertainty propagation. Our model is based on segmented Gaussian plume model (SGPM) approach that can account approximately for dynamics of released discharges and short-term forecast of hourly changes of meteorological conditions. For near area from the source and constant meteorological conditions is used also simplified version of Gaussian straight-line propagation (GPM). Implemented numerical difference scheme enables to approximate simulations of important parent-daughter pair formation. The objective multi-dimensional function F of N variables (subjected to bounds) is minimized starting at initial estimate. Commonly used Nelder-Mead direct search or Powell minimization methods are tested here for elementary scenarios of accidental harmful discharges. Applicability bounds are examined for which satisfactory results at acceptable time of computation were achieved.

Even for the simplest formulation of atmospheric dispersion and deposition in terms of Gaussian straight-line propagation the model $\mathcal{M}$ is nonlinear. In the following paragraphs we shall concentrate on accidental radioactivity release into atmosphere and its further deposition on terrain. Approximation in terms of source depletion scheme accounts for removal mechanisms of admixtures from the plume due to radioactive decay and dry and wet deposition on terrain [PHP07]. Let us proceed directly to the examination of the resulting fields of radioactivity deposition of a certain nuclide on terrain. The output is assumed to be represented by vector Z having dimension equal to the number N of total calculating points in the polar grid (in our case N=2800, what means 80 radial sections and 35 concentric radial zones up to 100 km from the source of pollution). General expression for dependency of Z on model input parameters $\theta_1, \theta_2, \ldots, \theta_k$ can be formally written as

$$\mathbf{Z} = \mathcal{M}(\theta_1, \theta_2, \ldots, \theta_R). \tag{6.1}$$

Let there be R receptor points on terrain where the respective values are measured. Generally, the number of receptors is much lower then N and we meet the problem with rare measurements expressed by observation vector $\mathbf{Y} \equiv (y_1, y_2, \ldots, y_R)$. Positions of sensors generally differ from the points of calculation grid. W

Number K of input parameters is rather high (several tenth) and then for practical purposes only S of them are treated as random. Rest of them are assumed to be less important from viewpoint of uncertainty propagation through the model and we assign them their best estimated values. (6.1) has then the form

$$\mathbf{Z} = M(\theta_1, \theta_2, \ldots, \theta_S, \theta_{S+1}^b, \ldots, \theta_R^b)$$

In other words a certain number S of selected problem-dependent optimization parameters $\theta_1, \theta_2, \ldots, \theta_S^b$, are considered to be uncertain and subjected to fluctuations within some range. The function F is constructed as a sum of squares in the measurement points between the values of model predictions and values observed in terrain expressed as:

$$F(\theta_1, \theta_2, \ldots, \theta_S) = \sum_{r=1}^{R} \left[y_r - \mathcal{H}(\mathcal{M}(\theta_1, \theta_2, \ldots, \theta_S)) \right]^2. \tag{6.2}$$

Minimization algorithm searches a minimum of scalar function F of S parameters starting at an initial "best estimate". In brief glance, the test points $\theta_1, \theta_2, \ldots, \theta_S$ of the objective function F are arranged as a S-dimensional simplex and the algorithm tries to replace iteratively individual points with aim to shrink the simplex towards the best points. Further specific analysis concerns the resulting spatial fields of radioactivity deposition of a certain nuclide on terrain. Model predictions can be interpreted as Gaussian surface (or superposition of partial Gaussian extents) over the terrain. Our objective is to take into account both model predictions and available measurements incoming from the terrain and to improve spatial distribution of deposited radioactivity. We can imagine the iterative process of minimization of function F such consecutive adjustment of the resulting respond surface, always according to the new evaluation of the parameters $\theta_1, \theta_2, \ldots, \theta_S$. Thus, the predicted respond surface of results is gradually "deformed by permissible manipulations" directly driven by changes of problem-dependent optimization parameters θ . MT algorithm controls the procedure until the best fit of modified surface with observation values is reached. Important feature of the method insists in preservation of physical knowledge, because the new set of parameters evaluated by minimization algorithm always re-enters the entire nonlinear environmental model $\mathcal{M}$ according to (6.1).

6.2 Numerical examples

6.2.1 Practical implementation

Investigation of applicability of minimization assimilation technique was tested on so called "twin experiment". Lack of real observations is simulated by artificial generation of measurements. Moreover, if we use for this generation the same environmental model (e.g. for a fix one set of disturbed input parameters) we can examine the problem convergence issues. In application part of the paper the results of two simulation experiments TWIN1 and TWIN2 are illustrated. TWIN1 relates to release of nuclide I-131. Its further straight-line propagation and deposition on terrain is described according to simple scheme of straight-line Gaussian plume model. TWIN2 experiment deals with the problem of evolution of Cs-137 deposition on terrain during the plume phase. Minimization search is applied with more complicated but more realistic segmented model SGPM.

6.2.2 MT applied to simple Gaussian straight-line model

Accidental one-hour release of radionuclide I-131 with total radioactivity 1.28E+11Bq discharged into atmosphere from nuclear facility is analyzed. Release height is $100m$, propagation continues under constant meteorological conditions (straight-line propagation in direction North-East, mean plume velocity $1.6m/s$, Pasquill's category D of atmospheric stability, no rain). Atmospheric dispersion coefficients are calculated according to KFK-Jülich semi-empirical formulas.

In the first step all input parameters are assumed to be represented by their best estimate values denoted by θ_i^b and then the corresponding output vector $\mathbf{Z}^b$ presents deterministic solution of deposited activity of selected nuclide on terrain. At the same time $\mathbf{Z}^b$ represents initial estimate for starting of minimization iterative search. In the second step we shall further reduce the number of parameters S from (6.2) to four parameters. Corresponding four uncertainties c_1, c_2, c_3, c_4 are introduced into the model according to scheme

$$\theta_i = c_i \theta_i^b, \quad \theta_i = \theta_i + c_i + f(\theta_i^b).$$

Specifically, their meaning, usage in the environmental code and real choice is given in Table 6.1.

The function $F(\theta_1, \theta_2, \ldots, \theta_S)$ now becomes $F(c_1, c_2, c_3, c_4)$ and minimization algorithm handles with 4-dimensional simplex. For purposes of construction of function F we have used slight modification of probabilistic version of existing environmental model HARP where original random inputs c_1, c_2, c_3, c_4 now play more general role of uncertainties characterized only by their range of possible fluctuations (see column 4 in Table 6.1). Minimization algorithm uses this constraints such lower and upper bounds for permissible manipulations with values of variables c_1, c_2, c_3, c_4 (see arrows in Figure

parameter	unit	uncertainty inside code	uncertainty bounds
θ_1: source release rate	$Bq.s^{-1}$	$Q = c_1 Q^b$	$c_1 \in \langle 0.1, 2.9 \rangle$
θ_2: horiz. dispersion	m	$\sigma_y(x) = c_2 \sigma_y^b(x)$	$c_2 \in \langle 0.1, 3.1 \rangle$
θ_3: wind direction	rad	$\varphi = \varphi^b + c_3 2\pi/80$	$c_3 \in \langle -5.0, 5.0 \rangle$
θ_4: dry depo. velocity	$m.s^{-1}$	$v_g = c_4 v_g^b$	$c_4 \in \langle 0.1, 4.0 \rangle$

Table 6.1: Introduction of uncertainties for four important input model parameters.
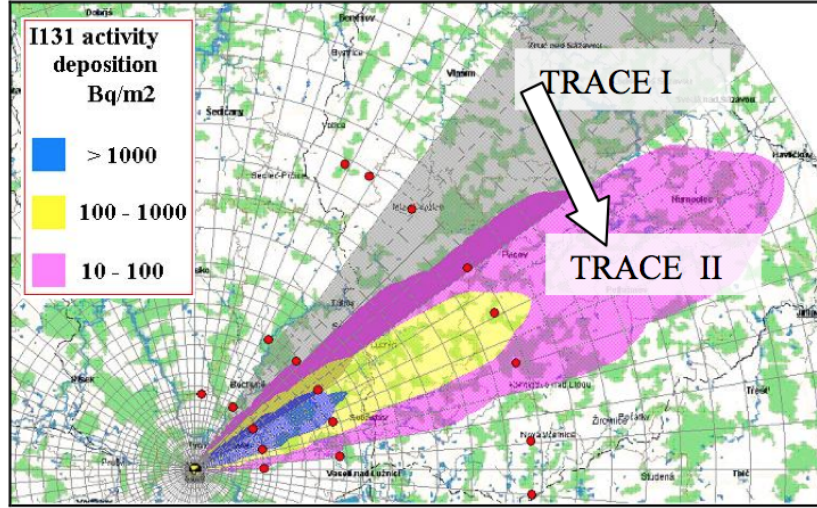
Figure 6.1: I-131 deposition levels $[Bq.m^{-2}]$ related to the end of plume progression. TWIN I experiment using Gaussian straight-line model. TRACE I and TRACE II are initial best estimate and resulting assimilation with simulated measurements (at red circles).

1). During TWIN experiments the observation vector $\mathbf{Y} \equiv (y_1, y_2, ..., y_R)$ is simulated artificially, the simplest way is utilization of the same environmental model $\mathcal{M}$.

Deterministic best estimate distribution $\mathbf{Z}^b$ generated on the polar calculation grid in original wind direction Sorig (North-East) is in Fig. 6.1 as TRACE I . It corresponds to the best estimate values $(c_1, c_2, c_3, c_4)^{\text{best}} = (1.0, 1.0, 1.0, 1.0)$. Selected positions of observations are labelled by red circles. For simulation of measurements in this red points we have selected a certain fixed quartet $(c_1, c_2, c_3, c_4)^{\text{obs}} = (1.73, 1.51, 4.00, 1.98)$. Artificially simulated measurements were generated using vector $\mathbf{Z}^{\text{obs}} = \mathcal{H} \left[\mathcal{M}((c_1, c_2, c_3, c_4)^{\text{obs}}) \right]$. Minimization algorithm in successive iterations j brings newly generated quartets $(c_1, c_2, c_3, c_4)^j$ closer and closer to the $(c_1, c_2, c_3, c_4)^{\text{obs}}$. Fast convergence of assimilated model predictions towards simulated observations has been found. 220 iterations are calculated during about 6 minutes and the following values has been found: $(c_1, c_2, c_3, c_4)^{j=220} = (1.731, 1.514, +4.003, 1.982)$. It demonstrates very good consent with "simulated" observations generated by $(c_1, c_2, c_3, c_4)^{\text{obs}}$. The results are illustrated in Fig. 6.1 as TRACE II isolines.

Original best estimate deposition on terrain (and at the same time initial guess

entering MT) is labelled as TRACE I. Deposition after 220 iterations is calculated as $\mathbf{Z}^{j=220} = \mathcal{H}\left[\mathcal{M}((c_1, c_2, c_3, c_4)^{\mathrm{obs}})\right]$ and its contours illustrates TRACE II. The assimilated respond surface TRACE II is at the same time practically identical with $\mathbf{Z}^{\mathrm{obs}}$ generated according to $\mathcal{M}((c_1, c_2, c_3, c_4)^{\mathrm{obs}})$ originally used for artificial simulations of measurements. The shapes of TRACE I and TRACE II reflect imposed changes in values of c_1^{best} to c_1^{obs} (higher nuclide discharge), c_2^{best} to c_2^{obs} (higher peripheral dispersion), c_3^{best} to c_3^{obs} (twist by 18°), c_4^{best} to c_4^{obs} (more intensive dry deposition causing steeper longitudinal gradient).

Direct search algorithm connected with Gaussian straight- line propagation model has proved fast convergence provided that the measurements are well positioned. Its applicability depends on validity and limitations of model itself. However, the TWIN 1 results support an idea of MT application for preliminary fleeting estimation in near distances and during constant meteorological conditions.

6.3   MT with more realistic SGPM environmental model

TWIN2 scenario is formulated in connection with segmented Gaussian plume scheme (model SGPM marked as MSGPM), which is much more complicated then straight- line spreading (our approach described in [PHP07]). The model synchronizes segmentation of release dynamics with hourly meteorological forecasts. The first two consecutive release segments of Cs-137 discharge (each with 1 hour duration) with released amount 2.0E+17Bq and 1.0E+17Bq has character of severe loss of coolant accident (LOCA) with partial fuel cladding rupture and fuel melting. Short-term meteorological forecast for the next 48 hours is provided by the Czech meteorological service. Then, for each hour since the release initiation there are available predictions of wind direction and speed, category of atmospheric stability according to Pasquill's classification and rain precipitation. Omitting other details, the TWIN II scenario covers period of the first 3 hours from the release start and we are declaring the following plan:

1. Each of the two segments is modeled up to third hour after the release start taking into account short-term hourly meteorological forecast. The situation just after 3 hours is given by superposition of both segments in their successive meteorological hourly phases (5 phases in total). Resulting best estimate fields are calculated in analogy with Equation (1) according to scheme

$$\mathbf{Z}_{3hours}^{\mathrm{best}} = \mathcal{H}\left[\mathcal{M}^{\mathrm{SGPM}}((c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53})^{\mathrm{best}})\right]$$

   and is illustrated in Fig. 6.2 as TRACE I.

2. Number of uncertainties is increased from four to five as c_1, c_2, c_3, c_4, c_5. c_5 stands for fluctuation of mean wind velocity. If we suppose wind direction and velocity fluctuations to be independent between hourly phases, then c_3 and c_5 split to 6 independent uncertainties c_{31}, c_{32}, c_{33} (for wind direction predicted for hours 1, 2 ,3) and c_{51}, c_{52}, c_{53} (for wind velocity predicted for hours 1, 2 ,3).
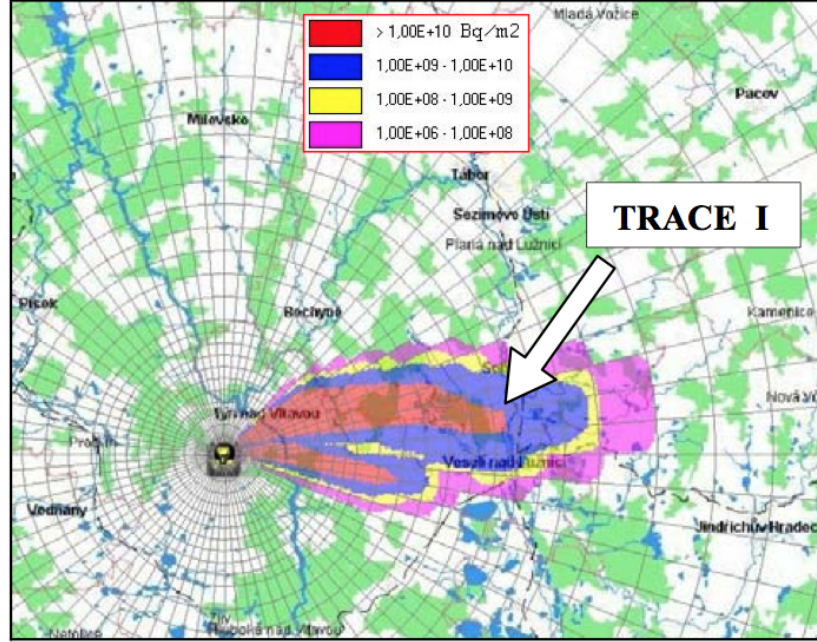
Figure 6.2: Nominal deposition of Cs-137 (just 3 hours after the release start).

3. We have simulated artificially fictive "observation surface" according to

$$\mathbf{Z}_{3hours}^{obs} = \mathcal{H}\left[\mathcal{M}^{\text{SGPM}}((c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53})^{\text{obs}})\right].$$

Let us suppose their incoming in one stroke just at hour 3 after the accident start. Let us state beforehand that assimilated TRACE II from Fig. 6.3 nearly corresponds with the "observation surface".

4. The main goal is to accomplish assimilation of the model predictions $\mathbf{Z}^{\text{best}}$ in compliance with 3hour measurements in analogy with (6.2) using BCPOL procedure of minimization.

Deposition of Cs-137 on terrain after 728 iterations is calculated as

$$\mathbf{Z}_{3hours}^{j=728} = \mathcal{H}\left[\mathcal{M}^{\text{SGPM}}((c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53})^{j=728})\right]$$

and its isolines illustrates in Fig. 6.3 a trail on terrain marked as TRACE II. The results represent a new distribution just at third hour after the release start, which is improved by observations. Minimization algorithm is initiated by the best estimate solution (TRACE I) and gradually approaches to the simulated observations. In short numerical summary, TWIN2 experiment required to prepare in advance sets of parameters $(c_1, c_2, c_{31}, c_{32}, c_{33}, c_4, c_{51}, c_{52}, c_{53})$ for:

	c_1	c_2	c_{13}	c_{23}	c_{33}	c_4	c_{15}	c_{25}	c_{35}
best	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
measurements	7.0	2.0	-4.0,	-5.0	-6.0	2.55	-0.5	-0.6	-0.7
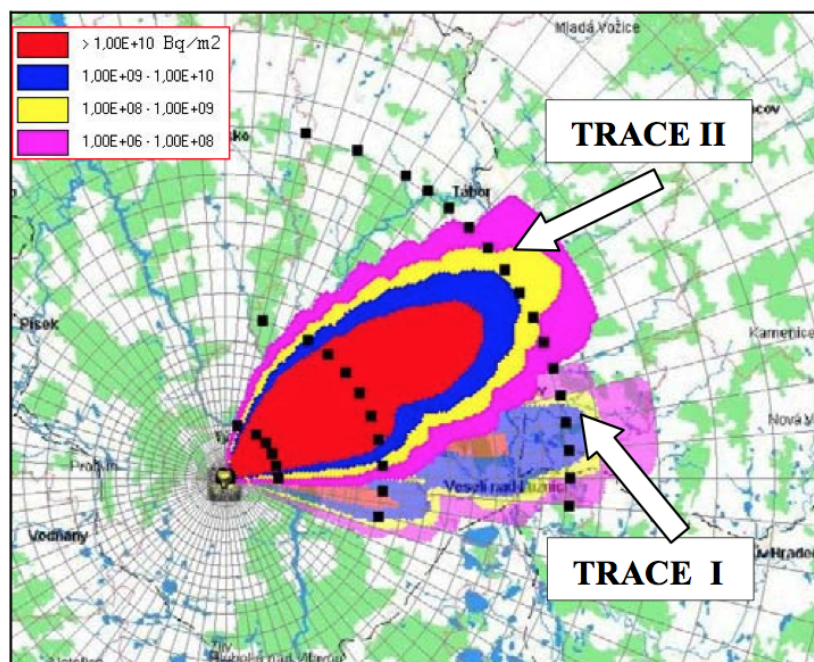
Figure 6.3: Assimilation of predicted deposition of 137Cs and simulated measurements just 3 hours after the release start; artificially simulated measurements in black squares.

Meaning of the parameters c_1 to c_4 is the same as described in Table 6.1. c_5 stands for uncertainty of the mean velocity of the plume. Further splitting to c_{i5}, $i = 1, 2, 3$, holds true for independent fluctuations of the mean velocity $\bar{u}_i$ forecasted for hours i. Uncertain $\bar{u}_i$ is then expressed according to $\bar{u}_i = \bar{u}_i^{\text{best}}(1 + 0.35 c_{i5})$, $c_{i5} \in \langle -1, 1 \rangle$.

TWIN II experiment took into consideration 9 optimization parameters with constructive idea to discriminate according to their global or local effect (introduced into the wind vector). The computation procedure is time consuming, but satisfactory convergence can be achieve. More detailed analysis related to the criterion of match between model and measurements (e.g. [ET04]) is so far pending.

6.4 Conclusion

Advantage of utilization of SGPM output fields as a fitting surface insists in preservation of physical knowledge of the model. Presented experience related to applicability of minimization techniques indicates that number of selected optimization parameters c_i should not be too high in order to avoid the poor convergence or even taking the wrong way (more sophisticated algorithms have to be searched). At this stage we recommend to consider five optimization parameters included in the TWIN II experiment (where wind velocity vector is global, it means no further splitting of c_3 to further c_{i3} and c_5 to c_{i5}) and link the 6th parameter c_6 representing uncertainty in precipitation intensity with local effect.

Presented minimization technique fits results on a certain specific situation. Any resulting effect (e.g. peripheral plume dispersion) usually depends on many other input random parameters. Thus, in no case the presented fitting technique should not be confused with parameter calibration. The problem of handling of real measurements still remains opened. Presented approach can play a specific role among empirical assimilation techniques, especially as fast and efficient software tool for analysis of possible discrepancies between the model predictions and observations incoming from terrain. The method is incorporated into the assimilation subsystem of the HARP code [HPP08a].

Realistic prediction of evolution of radiation situation during emergency gives decision makers necessary time on judgement and introduction of efficient urgent countermeasures on population protection. Reliable model predictions for the next hours in medium distances should account both for implementation of spatial meteorological forecast and development of new numerical techniques for time update of the trajectory models (e.g. how to propagate model for the next hours starting from assimilated results TRACE II in Fig. 6.3 Interventions introduced on the basis non-assimilated TRACE I could lead to fatal consequences on population health resulting from ill-anticipated impacted areas.

Chapter 7

Bayesian Methods

7.1 Identification of data assimilation with Bayesian estimation

Bayesian approach is based on quantifying uncertainty in statistical inference via probability density functions (pdfs). The importance of such approach is justified by the fact, that it facilitates a common-sense interpretation of statistical conclusions [Gel04].

If we think of the forecast and the analysis as of pdfs, the data assimilation can be understood as a particular case of recursive Bayesian estimation [Pet81]. In the Bayesian framework, the forecast and the analysis are represented by the *prior* pdf and *posterior* pdf, respectively. When no measurements are available, the pdf of the considered state must be rather wide to cover all possible realizations of the state. Each incoming measurement brings information about the "true" state, reducing the original uncertainty. In effect, with increasing measurements, the posterior pdf is narrowing down around the best possible estimate. From the Bayesian point of view, data assimilation is analogical to the problem of filtering, i.e., characterizing the distribution of the state of the hidden Markov model at the present time, given the information provided by all of the observations received up to the present time. Data update step of the assimilation cycle is implemented using *Bayes formula*.

7.2 Recursive Bayesian filtering

The task of data assimilation can be interpreted as a problem of *inference* of a discrete-time *stochastic process*:

$$\mathbf{x}_t \quad \sim \quad p(\mathbf{x}_t | \mathbf{x}_{t-1}), \tag{7.1}$$
$$\mathbf{y}_t \quad \sim \quad p(\mathbf{y}_t | \mathbf{x}_t). \tag{7.2}$$

Here, $\mathbf{x}_t \in \mathbb{R}^{N_x}$ is a vector known as the state variable, $\mathbf{y}_t \in \mathbb{R}^{N_y}$ is a vector of observations, t is the time index, and $p(\cdot|\cdot)$ denotes the conditional pdf of the variable.

State evolution model (7.1) describes the evolution of the state variables $\mathbf{x}_t$ over time, whereas the measurement model (7.2) explains how the measurements $\mathbf{y}_t$ relate to the state variables.

System given by (7.1)–(7.2) is rather general. It represents a *Markov process* of the first order, where realization of the process at time t contains all the information about the past, which is necessary to calculate its future behavior. In data assimilation we often restrict to its special case, where the explicit expressions for both the state model and the measurement model exist. This results in a discrete-time state-space models with additive noise represented by a set of difference equations [Jaz70]:

$$
\begin{aligned}
\mathbf{x}_t &= \mathcal{M}_t(\mathbf{x}_{t-1}) + \mathbf{w}_t, & (7.3) \\
\mathbf{y}_t &= \mathcal{H}_t(\mathbf{x}_t) + \mathbf{v}_t. & (7.4)
\end{aligned}
$$

The state transition operator $\mathcal{M}_t : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x}$ integrates the state forward to the next time step. The observation operator $\mathcal{H}_t : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y}$ transforms vectors from the state-space to the space of observations and makes them thus comparable with the observations. In environmental modeling, these operators represent our mathematical conceptualization of the physical reality under investigation. Vectors $\mathbf{w}_t$ and $\mathbf{v}_t$ with appropriate dimensions represent mutually independent noise processes of the model and the observations, respectively.

Formally, the prior distribution $p(\mathbf{x}_0)$ representing uncertainty of the forecast in time $t = 0$ is transformed into the 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 application of the data update and the time update:

1. Data update:

$$
p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p\left(\mathbf{y}_t|\mathbf{x}_t\right)p\left(\mathbf{x}_t|\mathbf{y}_{1:t-1}\right)}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{\int p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})d\mathbf{x}_t}, \tag{7.5}
$$

2. Time update:

$$
p(\mathbf{x}_{t+1}|\mathbf{y}_{1:t}) = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t})d\mathbf{x}_t. \tag{7.6}
$$

Given the prior pdf $p(\mathbf{x}_t|\mathbf{y}_{1:t-1})$ representing uncertainty in the forecast in time t, we use Bayes formula (7.5) and evaluate the posterior pdf $p\left(\mathbf{x}_t|\mathbf{y}_{1:t}\right)$ representing uncertainty in the analysis in time t. Likelihood function $p(\mathbf{y}_t|\mathbf{x}_t)$ is defined by the observation model (7.4). In recursive Bayesian filtering, we exploit the fact that if the prior pdf is properly chosen from a class *conjugate* to (7.2), the formula (7.5) yields a posterior pdf of the same type.

Chapman–Kolmogorov equation (7.6) [Jaz70] advances the the posterior $p\left(\mathbf{x}_t|\mathbf{y}_{1:t}\right)$ in time and produces the forecast in time $t + 1$ represented by the prior $p(\mathbf{x}_{t+1}|\mathbf{y}_{1:t})$. Pdf $p(\mathbf{x}_{t+1}|\mathbf{x}_t)$ is called the *state transition pdf* and represents model dynamics given by (7.3). Integration in (7.5)–(7.6) and everywhere else in this work is performed over the maximum support of the integrand, if not stated otherwise.

Using posterior $p\left(\mathbf{x}_t|\mathbf{y}_{1:t}\right)$, we can evaluate the expected value of a function $f(\cdot)$ of $\mathbf{x}_t$ integrable with respect to $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, [DDFG01]:

$$\mathbb{E}[f(\mathbf{x}_t)|\mathbf{y}_{1:t}] = \int f(\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t})d\mathbf{x}_t. \tag{7.7}$$

Evaluation of (7.5) and (7.6) may involve integration over complex spaces and in the most cases it is computationally infeasible. Thats the reason why were developed methods for solution of the problem under simplifying conditions or methods providing some sub-optimal, but still satisfactory, solution. In the following text we briefly review the basic approaches to solution of the sequential data assimilation problem.

7.3 Kalman filter

Kalman Filter (KF) [Kal60] gives us the optimal solution for the system (7.3)–(7.4) with linear dynamics (operators $\mathcal{M}_t$ and $\mathcal{H}_t$ are linear) and zero mean Gaussian white noise processes $\mathbf{w}_t$ and $\mathbf{v}_t$. The state transition pdf $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ and the likelihood function $p(\mathbf{y}_t|\mathbf{x}_t)$ then become of the Gaussian type:

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{M}_t\mathbf{x}_{t-1}, \mathbf{Q}_t), \quad p(\mathbf{y}_t|\mathbf{x}_t) = \mathcal{N}(\mathbf{H}_t\mathbf{x}_t, \mathbf{R}_t).$$

Here, $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a Gaussian pdf with mean value $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Matrices $\mathbf{M}_t \in \mathbb{R}^{N_x \times N_x}$ and $\mathbf{H}_t \in \mathbb{R}^{N_y \times N_x}$ are matrices of linear operators $\mathcal{M}_t$ and $\mathcal{H}_t$, respectively. Matrices $\mathbf{Q}_t$ and $\mathbf{R}_t$ are known covariance matrices of model error and measurement error, respectively, with appropriate dimensions:

$$\mathbf{Q}_t = \mathbb{E}\left[\mathbf{v}_t\mathbf{v}_t^{\mathrm{T}}\right], \quad \mathbf{R}_t = \mathbb{E}\left[\mathbf{w}_t\mathbf{w}_t^{\mathrm{T}}\right].$$

The analysis (posterior state estimate) is in the Kalman filter represented by mean value $\bar{\mathbf{x}}_{t|t}$ and covariance matrix $\mathbf{P}_{t|t}$ of the estimated filtering Gaussian distribution:

$$\bar{\mathbf{x}}_{t|t} = \mathbb{E}\left[\mathbf{x}_t|\mathbf{y}_{1:t}\right], \quad \mathbf{P}_{t|t} = \mathbb{E}\left[(\mathbf{x}_t - \bar{\mathbf{x}}_{t|t})(\mathbf{x}_t - \bar{\mathbf{x}}_{t|t})^{\mathrm{T}}|\mathbf{y}_{1:t}\right]. \tag{7.8}$$

Similarly, the forecast (prior state estimate) is represented with mean value $\bar{\mathbf{x}}_{t+1|t}$ and its covariance $\mathbf{P}_{t+1|t}$ of estimated predictive Gaussian distribution:

$$\bar{\mathbf{x}}_{t+1|t} = \mathbb{E}[\mathbf{x}_{t+1}|\mathbf{y}_{1:t}], \quad \mathbf{P}_{t+1|t} = \mathbb{E}\left[(\mathbf{x}_{t+1} - \bar{\mathbf{x}}_{t+1|t})(\mathbf{x}_{t+1} - \bar{\mathbf{x}}_{t+1|t})^{\mathrm{T}}|\mathbf{y}_{1:t}\right]. \tag{7.9}$$

The data update step of the KF assimilation cycled is given by the following equations:

$$
\begin{aligned}
\mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}}\left(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t\right)^{-1}, & (7.10) \\
\bar{\mathbf{x}}_{t|t} &= \bar{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_{1:t} - \mathbf{H}_t\bar{\mathbf{x}}_{t|t-1}\right), & (7.11) \\
\mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)^{\mathrm{T}} + \mathbf{K}_t\mathbf{R}_t\mathbf{K}_t^{\mathrm{T}} & (7.12) \\
&= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}, & (7.13)
\end{aligned}
$$

where $\mathbf{I} \in \mathbb{R}^{N_x \times N_x}$ is the identity matrix. We use the Kalman gain matrix $\mathbf{K}_t \in \mathbb{R}^{N_x \times N_y}$ for linear weighing of contributions given by the current observations $\mathbf{y}_t$ and the forecast to the resulting analysis. The analysis $\bar{\mathbf{x}}_{t|t}$ together with the posterior error covariance matrix $\mathbf{P}_{t|t}$ represent the sufficient statistics of the estimated posterior Gaussian pdf,

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \mathcal{N}(\bar{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t}).$$

The time update given by (7.14)–(7.15)

$$\begin{aligned}
\bar{\mathbf{x}}_{t+1|t} &= \mathbf{M}\bar{\mathbf{x}}_{t|t}, & (7.14) \\
\mathbf{P}_{t+1|t} &= \mathbf{M}_t\mathbf{P}_{t|t}\mathbf{M}_t^{\mathrm{T}} + \mathbf{Q}_{t+1}, & (7.15)
\end{aligned}$$

evaluates new prior pdf given by the forecast $\bar{\mathbf{x}}_{t+1|t}$ and its error covariance matrix $\mathbf{P}_{t+1|t}$,

$$p(\mathbf{x}_{t+1}|\mathbf{y}_{1:t}) = \mathcal{N}(\bar{\mathbf{x}}_{t+1|t}, \mathbf{P}_{t+1|t}).$$

The algorithm is initialized with prior estimates of the mean value $\bar{\mathbf{x}}_{0|-1}$ and covariance matrix $\mathbf{P}_{0|-1}$.

Generally, violation of assumptions on linearity of the model and normality of the noise terms leads to a suboptimal solution. The computationally cheaper form of the posterior error covariance matrix (7.13) should be used only for the optimal gain $\mathbf{K}_t$, otherwise it can cause a numerical instability.

7.3.1 Suboptimal solution for nonlinear model

Suboptimal modification of the KF algorithm for nonlinear $\mathcal{M}_t$ and $\mathcal{H}_t$ is called the *Extended Kalman Filter* (EKF) [WB95]. The EKF is based on assumption that local linearization of (7.3)–(7.4) may be sufficient description of nonlinearity. Given the $\mathcal{M}_t$ and $\mathcal{H}_t$ are differentiable functions, we can linearize them around the current estimates using the first terms in their Taylor series expansions:

$$\mathbf{M}_t \approx \left.\frac{\partial \mathcal{M}_t}{\partial \mathbf{x}}\right|_{\mathbf{x}=\bar{\mathbf{x}}_{t|t}}, \quad \mathbf{H}_t \approx \left.\frac{\partial \mathcal{H}_t}{\partial \mathbf{x}}\right|_{\mathbf{x}=\bar{\mathbf{x}}_{t+1|t}}. \tag{7.16}$$

Matrices $\mathbf{M}_t$ and $\mathbf{H}_t$ are used in the Kalman filter equations for advancing the posterior covariance matrix and during the data update step, respectively. Since the *Jacobians* (7.16) are dependent on the current state estimates, they must be recalculated at each time step.

If the functions $\mathcal{M}_t$ and $\mathcal{H}_t$ are highly nonlinear, the results of the EKF are rather poor. We can use expansions of higher orders or choose an alternative filtering methodology, e.g., the Unscented Kalman Filter [JU97] or an ensemble filter .

7.4 Ensemble filters

Since the propagation and storing of large covariance matrices is computationally demanding, formally correct KF and its variants are not suitable for high-dimensional problems commonly occurring in different geoscience applications, for instance, in meteorology [HMP$^+$05]. The idea of ensemble filtering was introduced by [Eve94]. Ensemble filters avoid explicit evolution of covariance by approximating the estimated pdf with an ensemble of states. It can be understood as a Monte Carlo approximation of the traditional KF.

7.4.1 Ensemble Kalman filter

In *Ensemble Kalman Filter* (EnKF), a small random ensemble of states is used to represent the estimated pdf. Similarly to the KF, the EnKF makes the assumption that all probability density functions involved are Gaussian.

Let $\mathbf{X}_{t|t-1}$ denote prior ensemble in time t,

$$\mathbf{X}_{t|t-1} = [\mathbf{x}_{t|t-1}^1, \mathbf{x}_{t|t-1}^2, \ldots, \mathbf{x}_{t|t-1}^M].$$

The prior estimate $\bar{\mathbf{x}}_{t|t-1}$ and prior covariance matrix $\mathbf{P}_{t|t-1}$ are approximated as sample mean and sample variance of $\mathbf{X}_{t|t-1}$, respectively:

$$
\begin{aligned}
\bar{\mathbf{x}}_{t|t-1} &\equiv \frac{1}{M} \sum_{i=1}^M \mathbf{x}_{t|t-1}^i, & (7.17)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{P}_{t|t-1} &\equiv \frac{1}{M-1} \sum_{i=1}^M \left(\mathbf{x}_{t|t-1}^i - \bar{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{x}_{t|t-1}^i - \bar{\mathbf{x}}_{t|t-1}\right)^{\mathrm{T}}. & (7.18)
\end{aligned}
$$

The posterior ensemble

$$\mathbf{X}_{t|t} = [\mathbf{x}_{t|t}^1, \mathbf{x}_{t|t}^2, \ldots, \mathbf{x}_{t|t}^M]$$

is given by the Bayesian data update, where each ensemble member is updated separately:

$$
\begin{aligned}
\mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}}\left(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t\right)^{-1}, & (7.19)\\
\mathbf{x}_{t|t}^i &= \mathbf{x}_{t|t-1}^i + \mathbf{K}_t\left(\mathbf{y}_{1:t}^i - \mathbf{H}_t\mathbf{x}_{t|t-1}^i\right), \quad i = 1, \ldots, M. & (7.20)
\end{aligned}
$$

A set of perturbed observation vectors $\mathbf{y}_t^i \sim \mathcal{N}(\mathbf{y}_t, \mathbf{R}_t)$, $i = 1, \ldots, M$, must be used to update the ensemble members in order to fulfill (7.12). It can be shown that if all the ensemble members were updated with the same observation vector $\mathbf{y}_t$ and the same gain $\mathbf{K}_t$, the posterior covariance will be

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)^{\mathrm{T}}. \tag{7.21}$$

Without the term $\mathbf{K}_t\mathbf{R}_t\mathbf{K}_t^{\mathrm{T}}$ is the posterior covariance systematically underestimated.

Using posterior ensemble $\mathbf{X}_{t|t}$, posterior estimate $\bar{\mathbf{x}}_{t|t}$ and covariance $\mathbf{P}_{t|t}$ are approximated with its sample mean and variance:

$$\bar{\mathbf{x}}_{t|t} \equiv \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_{t|t}^{i}, \tag{7.22}$$

$$\mathbf{P}_{t|t} \equiv \frac{1}{M-1}\sum_{i=1}^{M}\left(\mathbf{x}_{t|t}^{i}-\bar{\mathbf{x}}_{t|t}\right)\left(\mathbf{x}_{t|t}^{i}-\bar{\mathbf{x}}_{t|t}\right)^{\mathrm{T}}. \tag{7.23}$$

Advancing of the estimated Gaussian pdf approximated with the ensemble in time is achieved by simply advancing each ensemble member with the nonlinear forecast model $\mathcal{M}_t$:

$$\mathbf{x}_{t+1|t}^{i} = \mathcal{M}_t\left(\mathbf{x}_{t|t}^{i}\right), \quad i = 1,\ldots,M.$$

Since the time evolution of the posterior covariance is performed by evolution of an ensemble, the posterior covariance itself does not have to be stored.

What is more, since only $\mathbf{P}_{t|t-1}\mathbf{H}^{\mathrm{T}}$ and $\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}^{\mathrm{T}}$ are required during filter evaluation, the full prior covariance matrix $\mathbf{P}_{t|t-1}$ needs never to be calculated [Eve94]. We can directly calculate the terms occurring in the expression for the Kalman gain,

$$\mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}} = \frac{1}{M-1}\sum_{i=1}^{M}\left(\mathbf{x}_{t|t-1}^{i}-\bar{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{H}_t\mathbf{x}_{t|t-1}^{i}-\mathbf{H}_t\bar{\mathbf{x}}_{t|t-1}\right)^{\mathrm{T}},$$

$$\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}} = \frac{1}{M-1}\sum_{i=1}^{M}\left(\mathbf{H}_t\mathbf{x}_{t|t-1}^{i}-\mathbf{H}_t\bar{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{H}_t\mathbf{x}_{t|t-1}^{i}-\mathbf{H}_t\bar{\mathbf{x}}_{t|t-1}\right)^{\mathrm{T}}.$$

Covariance $\mathbf{P}_{t|t-1}$ is also used in the formula for predictive density of the observations,

$$p(\mathbf{y}_t|\mathbf{y}_{1:t-1}) = \mathcal{N}(\mathbf{H}_t\bar{\mathbf{x}}_{t|t-1}, \mathbf{Z}_t), \quad \mathbf{Z}_t = \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^{\mathrm{T}}+\mathbf{R}_t, \tag{7.24}$$

which corresponds to the standard predictive density of the Kalman filter [Pet81]. This quantity is often called *marginal likelihood* (marginalization is with respect to $\mathbf{x}_t$) and plays an important role in statistical model selection [Jef61].

7.4.1.1 Efficient implementation of EnKF

Following [Man06], let $\tilde{\mathbf{X}}_t = [\tilde{\mathbf{x}}_t^1,\ldots,\tilde{\mathbf{x}}_t^M] = [\mathbf{x}_t^1-\bar{\mathbf{x}}_t,\ldots\mathbf{x}_t^M-\bar{\mathbf{x}}_t]$ be an ensemble of deviations from the ensemble mean. $\tilde{\mathbf{X}}_t$ can be easily calculated using

$$\tilde{\mathbf{X}}_t = \mathbf{X}_t - \frac{1}{M}(\mathbf{X}_t\boldsymbol{\Psi}_{N_x\times1})\boldsymbol{\Psi}_{N_x\times1}^{\mathrm{T}},$$

where matrix $\boldsymbol{\Psi}_{N_x\times1}$ is a matrix of all ones of dimension $N_x \times 1$ and the expression $\frac{1}{M}(\mathbf{X}_t\boldsymbol{\Psi}_{N_x\times1})\boldsymbol{\Psi}_{N_x\times1}^{\mathrm{T}}$ stands for the ensemble, where all the members are equal to the

---

Algorithm 7.1 EnKF with perturbed observations.

1. Initialization. Generate a prior ensemble (background field):

$$\mathbf{X}_{0|-1} = \left[\mathbf{x}_{0|-1}^1, \ldots, \mathbf{x}_{0|-1}^M\right], \quad \mathbf{x}_{0|-1}^i \sim \mathcal{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0), \quad i = 1, \ldots, M.$$

2. EnKF data update:

   (a) Generate perturbed measurements:

   $$\mathbf{D}_t = \left[\mathbf{y}_t^1, \ldots, \mathbf{y}_t^M\right], \quad \mathbf{y}_t^i \sim \mathcal{N}(\mathbf{y}_t, \mathbf{R}_t), \quad i = 1, \ldots, N_y.$$

   (b) Calculate Kalman gain $\mathbf{K}_t$ for update of ensemble:

   $$\begin{aligned} \mathbf{S}_{t|t-1} &= \frac{1}{\sqrt{M-1}} \left(\mathbf{X}_{t|t-1} - \frac{1}{M}(\mathbf{X}_{t|t-1}\boldsymbol{\Psi}_{N_x \times 1})\boldsymbol{\Psi}_{N_x \times 1}^{\mathrm{T}}\right), \\ \mathbf{Z}_t^{-1} &= \mathbf{R}_t^{-1}\left[\mathbf{I} - \mathbf{H}\mathbf{S}_{t|t-1}\left(\mathbf{I} + (\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}}\mathbf{R}_t^{-1}\mathbf{H}_t\mathbf{S}_{t|t-1}\right)^{-1}(\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}}\mathbf{R}_t^{-1}\right], \\ \mathbf{K}_t &= \mathbf{S}_t(\mathbf{H}_t\mathbf{S}_t)^{\mathrm{T}}\mathbf{Z}_t^{-1}. \end{aligned}$$

   (c) Update the ensemble:

   $$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \mathbf{K}_t(\mathbf{D}_t - \mathbf{H}_t\mathbf{X}_{t|t-1}).$$

3. EnKF time update. Predict new ensemble:

$$\mathbf{x}_{t+1|t}^i = \mathcal{M}(\mathbf{x}_{t|t}^i), \quad i = 1, \ldots, M.$$

4. Set $t := t + 1$ and iterate from step 2.

---

mean values $\bar{\mathbf{x}}_t$ of the original ensemble $\mathbf{X}_t$. Covariance of $\mathbf{X}_t$ can be then evaluated using

$$\mathbf{P}_t = \frac{1}{M-1}\tilde{\mathbf{X}}_t\tilde{\mathbf{X}}_t^{\mathrm{T}}. \tag{7.25}$$

If we rewrite (7.25) as follows,

$$\mathbf{P}_t = \frac{1}{\sqrt{M-1}}\tilde{\mathbf{X}}_t\frac{1}{\sqrt{M-1}}\tilde{\mathbf{X}}_t^{\mathrm{T}} = \mathbf{S}_t\mathbf{S}_t^{\mathrm{T}},$$

the matrix $\mathbf{S}_t$ can be thought of as a square root of $\mathbf{P}_t$.

Bayesian update can be then formulated in a matrix form, where all the ensemble members are updated simultaneously and the square root $\mathbf{S}_{t|t-1}$ of $\mathbf{P}_{t|t-1}$ is used,

$$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \mathbf{S}_{t|t-1}(\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}}\left(\mathbf{H}_t\mathbf{S}_{t|t-1}(\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}} + \mathbf{R}_t\right)^{-1}(\mathbf{D}_t - \mathbf{H}_t\mathbf{X}_{t|t-1}).$$

Here, $\mathbf{D}_t = [\mathbf{y}_t^1, \ldots, \mathbf{y}_t^M]$ is the ensemble of perturbed observation vectors.

For a large number of data points, the inversion of the term $\mathbf{Z}_t = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t$ in (7.19) can be computationally demanding or even numerically unstable [Man06]. Given that the observations error covariance matrix $\mathbf{R}_t$ is diagonal, i.e, the observations are uncorrelated, we can use *Sherman–Morrison–Woodbury formula* [Hag89] for computation of $\mathbf{Z}_t^{-1}$:

$$(\mathbf{A}+\mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}. \tag{7.26}$$

Substituting $\mathbf{A} = \mathbf{R}_t$, $\mathbf{U} = \mathbf{H}_t\mathbf{S}_{t|t-1}$, $\mathbf{C} = \mathbf{I}$, $\mathbf{V} = \mathbf{H}_t\mathbf{S}_{t|t-1}$ into (7.26) yields formula for $\mathbf{Z}_t^{-1}$, where only the diagonal matrix $\mathbf{R}_t$ must be inverted:

$$\begin{aligned}
\mathbf{Z}_t^{-1} &= \left[\mathbf{R} + \mathbf{H}_t\mathbf{S}_{t|t-1}(\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}}\right]^{-1} \\
&= \mathbf{R}_t^{-1}\left[\mathbf{I} - \mathbf{H}_t\mathbf{S}_{t|t-1}\left(\mathbf{I} + (\mathbf{H}_t\mathbf{S}_{t|-1})^{\mathrm{T}}\mathbf{R}_t^{-1}\mathbf{H}_t\mathbf{S}_{t|t-1}\right)^{-1}(\mathbf{H}_t\mathbf{S}_{t|t-1})^{\mathrm{T}}\mathbf{R}_t^{-1}\right].
\end{aligned}$$

7.4.2 Ensemble square root filter

The *Ensemble square root filter* (EnSRF) is a revised version of the EnKF that eliminates the necessity to perturb the observations [WH02] and the posterior ensemble is formed deterministically. Deterministic methods were developed to address the problems related to sampling errors associated with the use of perturbed observations in stochastic analysis ensemble update methods [TAB+03]. [WH02] demonstrated that for an ensemble of a given size, the EnSRF is more accurate than the EnKF.

In the EnSRF, the data update step (7.20) is expressed as $N + 1$ equations for separate update of the ensemble mean $\bar{\mathbf{x}}_{t|t-1}$ and the deviations $\tilde{\mathbf{x}}_{t|t-1}^i$:

$$\begin{aligned}
\bar{\mathbf{x}}_{t|t} &= \bar{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{H}\bar{\mathbf{x}}_{t|t-1}\right), \\
\tilde{\mathbf{x}}_{t|t}^i &= \tilde{\mathbf{x}}_{t|t-1}^i - \tilde{\mathbf{K}}_t\mathbf{H}\tilde{\mathbf{x}}_{t|t-1}^i = \left(\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{H}\right)\tilde{\mathbf{x}}_{t|t-1}^i, \quad i = 1, \ldots, M.
\end{aligned}$$

Here, $\mathbf{K}_t$ is the traditional Kalman gain (7.19) and $\tilde{\mathbf{K}}_t$ is the gain used to update the deviations. In the EnKF, $\mathbf{K}_t = \tilde{\mathbf{K}}_t$ and deviations $\tilde{\mathbf{x}}_{t|t-1}$ are updated using

$$\tilde{\mathbf{x}}_{t|t}^i = \tilde{\mathbf{x}}_{t|t-1}^i - \mathbf{K}_t\left(\tilde{\mathbf{y}}_t^i - \mathbf{H}\tilde{\mathbf{x}}_{t|t-1}^i\right), \quad i = 1, \ldots, M, \tag{7.27}$$

where $\tilde{\mathbf{y}}_t^i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t)$ are the deviations of perturbed measurements from the mean $\mathbf{y}_t$[BvLE98].

[WH02] derived a formula for $\tilde{\mathbf{K}}_t$ that will results in an ensemble whose posterior error covariance satisfies (7.12). Substituting $\tilde{\mathbf{K}}_t$ into (7.21) and requiring the expression to be equal to the correct $\mathbf{P}_{t|t}$, we obtain the following equation for $\tilde{\mathbf{K}}_t$,

$$(\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{H})\mathbf{P}_{t|t-1}(\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{H})^{\mathrm{T}} = (\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{H})\mathbf{P}_{t|t-1}(\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{H})^{\mathrm{T}} + \tilde{\mathbf{K}}_t\mathbf{R}_t\tilde{\mathbf{K}}_t^{\mathrm{T}}, \tag{7.28}$$

which has a solution

$$\tilde{\mathbf{K}}_t = \mathbf{P}_{t|t-1}\mathbf{H}^{\mathrm{T}} \left[\left(\sqrt{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t} \right)^{-1} \right]^{\mathrm{T}} \left[\sqrt{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t} + \sqrt{\mathbf{R}_t} \right]^{-1}. \quad (7.29)$$

The fact that the evaluation of (7.29) involves square roots of error covariance matrices is the reason why the algorithm implementing the deterministic version of data update is called the ensemble square root filter. The matrix square roots in (7.29) are not unique and can be calculated by a commonly used factorization methods, e.g., Cholesky or singular value decomposition.

Given that the state evolution operator $\mathcal{M}_t$ is linear, the posterior ensemble mean $\bar{\mathbf{x}}_{t|t}$ and the posterior deviations from the mean $\tilde{\mathbf{x}}_{t|t}^i$ can be evolved separately

$$\begin{aligned} \bar{\mathbf{x}}_{t+1|t} &= \mathbf{M}_t\bar{\mathbf{x}}_{t|t}, \\ \tilde{\mathbf{X}}_{t+1|t} &= \mathbf{M}_t\tilde{\mathbf{X}}_{t|t}. \end{aligned}$$

Otherwise, the full ensemble $\mathbf{X}_t$ must be formed before the time update.

7.4.2.1 Sequential processing of observations

For an individual observation, i.e., when $\mathbf{H} \in \mathbb{R}^{1 \times N_x}$ and $\mathbf{R}_t \in \mathbb{R}$, the terms $\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}}$ and $\mathbf{R}_t$ reduce to scalars and (7.28) may be written as follows,

$$\frac{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}}}{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t}\tilde{\mathbf{K}}_t\tilde{\mathbf{K}}_t - \mathbf{K}_t\tilde{\mathbf{K}}_t - \tilde{\mathbf{K}}_t\mathbf{K}_t + \mathbf{K}_t\mathbf{K}_t = \mathbf{0}. \quad (7.30)$$

If the desired gain $\tilde{\mathbf{K}}_t$ is assumed to be linearly proportional to the original gain $\mathbf{K}$,

$$\tilde{\mathbf{K}}_t = \alpha_t\mathbf{K}_t, \quad (7.31)$$

where $\alpha \in \mathbb{R}$ is a constant, we obtain

$$\frac{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}}}{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t}\alpha^2\mathbf{K}_t\mathbf{K}_t - 2\alpha\mathbf{K}_t\mathbf{K}_t - \mathbf{K}_t\mathbf{K}_t = \mathbf{0}. \quad (7.32)$$

This yields a quadratic equation for α

$$\frac{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}}}{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t}\alpha^2 - 2\alpha\mathbf{K}_t\mathbf{K}_t + 1 = 0. \quad (7.33)$$

The equation has two roots. Since we want the deviations from the ensemble mean to be reduced in magnitude, i.e., to decrease posterior variance of the ensemble, and to maintain the same sign, the appropriate solution is

$$\alpha_t = \left(1 + \sqrt{\frac{\mathbf{R}_t}{\mathbf{HP}_{t|t-1}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t}} \right)^{-1}, \quad (7.34)$$

which is always between 0 and 1 [WH02].

In the case of non-correlated observations, when the observation error covariance matrix $\mathbf{R}_t$ is diagonal, the observations $\mathbf{y}_t$ can be processed sequentially, one at a time. Updating the deviations from ensemble mean with $\tilde{\mathbf{K}}_t$ given by (7.31) and (7.34) ensures the posterior error covariance to be equal to (7.12). Algorithm of EnSRF with sequential processing of observations is summarized in Algorithm 7.2.

Algorithm 7.2 EnSRF with sequential processing of observations.

1. Initialization. Generate a prior ensemble (background field):

$$\mathbf{X}_{0|-1} = \left[\mathbf{x}_{0|-1}^1, \ldots, \mathbf{x}_{0|-1}^M\right], \quad \mathbf{x}_t^i \sim \mathcal{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0), \quad i = 1, \ldots, M,$$

   and set $t := 0$.

2. EnSRF data update. For $j = 1, \ldots, N_y$:

   (a) Calculate Kalman gain $\mathbf{K}_{j;t}$ for update of ensemble mean with jth observation

   $$\mathbf{S}_{t|t-1} = \frac{1}{\sqrt{M-1}}\left(\mathbf{X}_t - \frac{1}{M}(\mathbf{X}_t\boldsymbol{\Psi}_{N_x \times 1})\boldsymbol{\Psi}_{N_x \times 1}^{\mathrm{T}}\right),$$

   $$\mathbf{K}_{j;t} = \frac{\mathbf{S}_{t|t-1}(\mathbf{H}_i\mathbf{S}_{t|t-1})^{\mathrm{T}}}{\mathbf{H}_i\mathbf{S}_{t|t-1}(\mathbf{H}_i\mathbf{S}_{t|t-1})^{\mathrm{T}} + \mathbf{R}_{j;t}},$$

   where $\mathbf{R}_{j;t}$ is variance of j-th observation and $\mathbf{H}_j \in \mathbb{R}^{1 \times n}$ is the corresponding observation operator.

   (b) Calculate Kalman gain $\tilde{\mathbf{K}}_{j;t}$ for update of deviations from the mean with j-th observation:

   $$\tilde{\mathbf{K}}_{j;t} = \alpha_{j;t}\mathbf{K}_{j;t},$$

   $$\alpha_{j;t} = \left[1 + \sqrt{\frac{\mathbf{R}_{j;t}}{\left(\mathbf{H}_j\mathbf{S}_{t|t-1}\right)\left(\mathbf{H}_j\mathbf{S}_{t|t-1}\right)^{\mathrm{T}} + \mathbf{R}_{j;t}}}\right]^{-1}.$$

   (c) Update of ensemble mean $\bar{\mathbf{x}}_{t|t-1}$ and departures from the mean $\tilde{\mathbf{X}}_{t|t-1}$

   $$\bar{\mathbf{x}}_{t|t} = \bar{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{H}\bar{\mathbf{x}}_{t|t-1}\right)$$

   $$\tilde{\mathbf{x}}_{t|t}^i = \tilde{\mathbf{x}}_{t|t-1}^i - \tilde{\mathbf{K}}_t\mathbf{H}\tilde{\mathbf{x}}_{t|t-1}^i, \quad i = 1, \ldots, M.$$

3. EnSRF time update. Predict new ensemble according to:

$$\mathbf{x}_{t+1|t}^i = \mathcal{M}(\mathbf{x}_{t|t}^i), \quad i = 1, \ldots, M.$$

4. Set $t := t + 1$ and iterate from step 2.

7.4.3   Sampling error issues in ensemble filtering

7.4.3.1   Ensemble Inflation

For a finite-sized ensemble, there is a sampling error in the estimation of forecast error covariance matrix (7.18). The theoretical exact forecast error covariance obtained from an infinite-sized ensemble differs from any obtained from a finite-sized ensemble of $M \in \mathbb{N}$ members [WH02]. Implication of this fact is, that in ensemble-based assimilation systems, the forecast error is systematically underestimated. Information brought by new measurements is then penalized because the measurement error seems to be relatively higher to the underestimated forecast error. Filter becomes too confident in the forecast and the divergence may occur. This effect can be observed particularly in small ensembles. Multiplicative *ensemble inflation* is a method for artificial increase of the model forecast error variance [AA99]. The inflation is used to replace the forecast ensemble according to:

$$\mathbf{x}^i \rightarrow \Delta \left(\mathbf{x}^i - \bar{\mathbf{x}}^i \right) + \bar{\mathbf{x}}^i, \quad i = 1, \ldots, M \tag{7.35}$$

with *inflation factor* Δ slightly greater than 1. From (7.35) is obvious that the mean value of the ensemble remains unchanged but its variance is increased.

7.4.3.2   Localization of covariance

The sampling error introduced by the finite ensemble size also causes *spurious covariances* in the estimated forecast error covariance matrix. Techniques of *covariance localization* filter out the small and noisy covariances and reduce the impact of the observations on remote state variables. In spatial data assimilation, where the state vector usually represent values of a quantity on a computational grid, the distance between states and observation simply denotes the real geographical distance between the grid points and the place of observation. Localization of a covariance matrix can be performed by using the *Schur product* of a localization matrix with the covariance matrix [GC99]. Schur product is an element-by-element matrix multiplication: the Schur product $\mathbf{A} \circ \mathbf{B}$ of matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$ is matrix $\mathbf{C} \in \mathbb{R}^{m \times n}$, where $C_{ij} = A_{ij} B_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.
   More specifically, we modify the formula for the Kalman gain (7.19) to be

$$\mathbf{K}_t = (\boldsymbol{\rho} \circ \mathbf{P}_{t|t-1})\mathbf{H}_t^{\mathrm{T}} \left(\mathbf{H}_t(\boldsymbol{\rho} \circ \mathbf{P}_{t|t-1})\mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t \right)^{-1}, \tag{7.36}$$

where $\boldsymbol{\rho}$ is a localization matrix [HM01]. Localization matrices are constructed by the means of correlation functions. Maximum of such a function reached at the observation location is 1 and the function typically decreases monotonically to zero at some finite distance from the observation location. The rate of correlation decrease with distance is given by the *length-scale parameter l*. Let $||D_{ij}||$ be the Euclidean distance between the observation location i and the grid points j. Then the example of a localization function
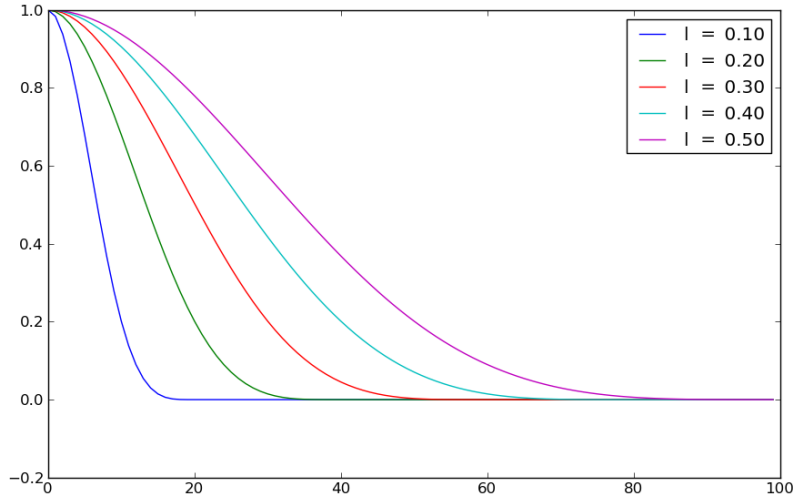
Figure 7.1: Correlation function given by (7.37) with different values of the length-scale parameter l.

is the compactly supported, 5th order piecewise rational function $\Omega(\sqrt{10/3}l, ||D_{ij}||)$ suggested by [GC99]:

$$\Omega(a, b) = \begin{cases} -\frac{1}{4}\left(\frac{b}{a}\right)^5 + \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 - \frac{5}{3}\left(\frac{b}{a}\right)^2 + 1, & \text{if } 0 \le b \le a; \\ \frac{1}{12}\left(\frac{b}{a}\right)^5 - \frac{1}{2}\left(\frac{b}{a}\right)^4 + \frac{5}{8}\left(\frac{b}{a}\right)^3 + \frac{5}{3}\left(\frac{b}{a}\right)^2 - 5\left(\frac{b}{a}\right) + 4 - \frac{2}{3}\left(\frac{a}{b}\right), & \text{if } a < b \le 2a; \\ 0, & \text{if } b > 2a, \end{cases}$$
(7.37)

where a and b correspond to $\sqrt{10/3}l$ and $||D_{ij}||$, respectively. Function given by (7.37) is similar to the Gaussian distribution in shape but is has a compact support. It is a homogeneous and isotropic correlation function, it means that it has the same behavior in all direction and the rate of correlation decrease is also invariant to translation of observation location in space. Correlation function given by (7.37) with different values of l is visualized in Figure 7.1.

Illustration of localization effect on spatial data is illustrated in Figure 7.2. In Figure 7.2 (a), a contour plot of spatial covariance of a point denoted with the red circle with the rest of the polar computational grid is visualized. In Figures 7.2 (b)–(d), we see the resulting covariance after application of the localization using a localization matrices given by (7.37) with different values of the length-scale parameter l.

7.5 Particle filter

Particle filtering (PF) refers to a group of methods further generalizing the Bayesian update problem for non-Gaussian pdfs. It includes a range of Monte Carlo techniques
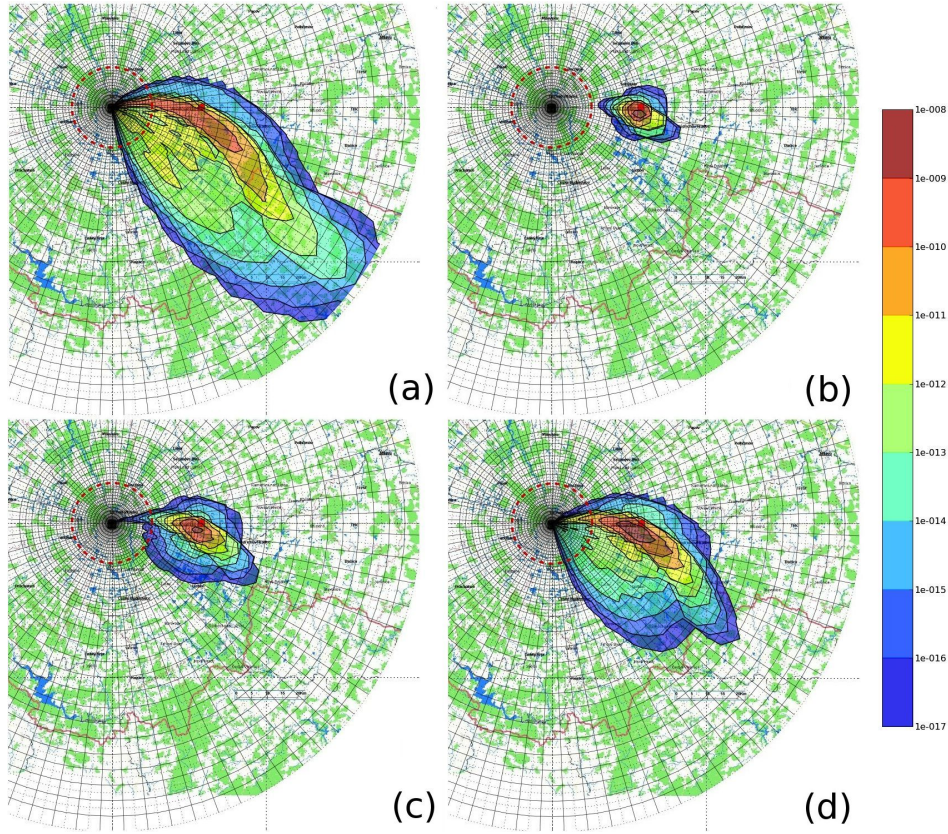
Figure 7.2: Illustration of the effect of covariance localization; (a) contour plot of spatial covariance (without localization) of a point denoted with the red circle with the rest of the polar computational grid; (b)–(d) resulting covariance after application of the localization using localization matrices given by (7.37) with increasing value of the length-scale parameter l.

for generating an empirical approximation of posterior $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ of a state trajectory $\mathbf{x}_{1:t} = (\mathbf{x}_1, \ldots, \mathbf{x}_t)$,

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \approx \frac{1}{N} \sum_{i=1}^{N} \delta\left(\mathbf{x}_{1:t} - \mathbf{x}_{1:t}^{(i)}\right). \tag{7.38}$$

Here, $\mathbf{x}_{1:t}^{(i)}, i = 1, \ldots, N$, are i.i.d..samples from the posterior $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ and $\delta(\cdot)$ is the Dirac δ-function. It comes out from the method of Monte Carlo integration. Expected value of an arbitrary function $f(\cdot)$ of $\mathbf{x}_{1:t}$ integrable with respect to $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ can be then approximated as

$$\mathbb{E}[f(\mathbf{x}_{1:t})|\mathbf{y}_{1:t}] = \int f(\mathbf{x}_{1:t})p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})d\mathbf{x}_{1:t} \approx \frac{1}{N} \sum_{i=1}^{N} f\left(\mathbf{x}_{1:t}^{(i)}\right), \tag{7.39}$$

and the rate of convergence of this approximation is independent of the dimension of the integrand [DDFG01].

In most of real applications we are not able to sample directly from the exact posterior, however, we can draw samples from a chosen *proposal density* (importance function) $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$:

$$
\begin{aligned}
p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) &= \frac{p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \\
&\approx \frac{p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})}\frac{1}{N}\sum_{i=1}^{N}\delta\left(\mathbf{x}_{1:t}-\mathbf{x}_{1:t}^{(i)}\right).
\end{aligned}
\tag{7.40}
$$

Approximation (7.40) can be written in a form of the *weighted* empirical distribution,

$$
p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N}w_t^{(i)}\delta\left(\mathbf{x}_{1:t}-\mathbf{x}_{1:t}^{(i)}\right),
\tag{7.41}
$$

$$
w_t^{(i)} \propto \frac{p(\mathbf{x}_{1:t}^{(i)}|\mathbf{y}_{1:t})}{q(\mathbf{x}_{1:t}^{(i)}|\mathbf{y}_{1:t})}.
\tag{7.42}
$$

Under this *importance sampling* procedure [RK08], the true posterior distribution needs to be evaluated point-wise only, since (7.41) can be normalized trivially via a constant $c = \sum_{i=1}^{n}w_t^{(i)}$.

In the following text, we will show how to recursively update a pdf given as a weighted empirical distribution. Following [RAG04], suppose we have a set of samples approximating posterior $p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})$ at time t − 1 and a new vector of measurements $\mathbf{y}_t$. We wish to approximate $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ with a new set of samples. If the proposal density is chosen to factorize such that

$$
q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) = q(\mathbf{x}_t|\mathbf{x}_{1:t-1},\mathbf{y}_{1:t})q(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1}),
\tag{7.43}
$$

then the new samples $\mathbf{x}_{1:t}^{(i)} \sim q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ can be obtained by augmenting each of the existing samples $\mathbf{x}_{1:t-1}^{(i)} \sim q(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})$ with the new state $\mathbf{x}_t^{(i)} \sim q(\mathbf{x}_t|\mathbf{x}_{1:t-1},\mathbf{y}_{1:t})$. Using the chain rule and the Bayes formula, $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ can be written in terms of $p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})$, $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ and $p(\mathbf{y}_t|\mathbf{x}_t)$, as follows:

$$
\begin{aligned}
p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) &= \frac{p(\mathbf{y}_t|\mathbf{x}_{1:t},\mathbf{y}_{1:t-1})p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \\
&= \frac{p(\mathbf{y}_t|\mathbf{x}_{1:t},\mathbf{y}_{1:t-1})p(\mathbf{x}_t|\mathbf{x}_{1:t-1},\mathbf{y}_{1:t-1})p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \\
&= \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})} \tag{7.44} \\
&\propto p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1}) \tag{7.45}
\end{aligned}
$$

By substituting (7.43) and (7.44) into (7.42), (7.42) may be written in the following

recursive form:

$$
\begin{aligned}
w_t^{(i)} &\propto \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})p(\mathbf{x}_{1:t-1}^{(i)}|\mathbf{y}_{1:t-1})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{1:t-1}^{(i)},\mathbf{y}_{1:t})q(\mathbf{x}_{1:t-1}^{(i)}|\mathbf{y}_{1:t-1})} \\
&\propto w_{t-1}^{(i)}\frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{1:t-1}^{(i)},\mathbf{y}_{1:t})}.
\end{aligned}
\tag{7.46}
$$

Furthermore, if the proposal density is chosen as follows,

$$
q\left(\mathbf{x}_t|\mathbf{x}_{1:t-1},\mathbf{y}_{1:t}\right) = q\left(\mathbf{x}_t|\mathbf{x}_{t-1},\mathbf{y}_t\right)
\tag{7.47}
$$

then the proposal density becomes only dependent on the $\mathbf{x}_{t-1}$ and $\mathbf{y}_t$. This is particularly useful in the common case when only an estimate of the marginal posterior $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ is required at each time step. It means, that only samples $\mathbf{x}_t^{(i)}$ need to be stored [RAG04] and the marginal posterior density $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ can be approximated as

$$
p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^{(i)}\delta\left(\mathbf{x}_t - \mathbf{x}_t^{(i)}\right),
\tag{7.48}
$$

$$
w_t^{(i)} \propto w_{t-1}^{(i)}\frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)},\mathbf{y}_t)}.
\tag{7.49}
$$

Using the particles, the mean value $\bar{\mathbf{x}}_{1:t}$ and the covariance $\mathbf{\Sigma}_{1:t}$ of the posterior approximation (7.41) can be calculated as follows,

$$
\bar{\mathbf{x}}_{1:t} = \sum_{i=1}^{N} w_t^{(i)}\mathbf{x}_{1:t}^{(i)},
\tag{7.50}
$$

$$
\mathbf{\Sigma}_{1:t} = \sum_{i=1}^{N} w_t^{(i)}\left[\left(\mathbf{x}_{1:t}^{(i)} - \bar{\mathbf{x}}_{1:t}\right)\left(\mathbf{x}_{1:t}^{(i)} - \bar{\mathbf{x}}_{1:t}\right)^{\mathrm{T}}\right].
\tag{7.51}
$$

The scheme for sequential evaluation of the weight with incoming observations is referred to as *sampling-importance-sampling* (SIS) [ADFDJ03]. Besides the appropriate choice of the proposal density, successful application of the PF requires more steps, namely implementation of a re-sampling algorithm, which avoids degeneracy of the weights.

7.5.1 Degeneracy problem and re-sampling

The variance of weights (7.46) increases during their recursive evaluation. The increase has a harmful effect on the accuracy and leads to the weights degeneracy, which is a common problems with the SIS particle filter [RAG04]. The weights degeneracy means, that after certain number of recursive steps, all but one particle have negligible

normalized weight which implies sample impoverishment and loss of diversity of the particles. In the SIS framework, weight degeneracy is unavoidable and has negative effects. Computational time must be spent on propagation of particles with negligible weights whose contribution to the approximation of $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ is small.

A suitable measure of degeneracy of an algorithm is the effective sample size N_{eff} [RAG04], which can be estimated using normalized weights $w_t^{(i)}$ as follows:

$$\overline{N_{\text{eff}}} = \frac{1}{\sum_{i=1}^{N} (w_t^{(i)})^2}, \tag{7.52}$$

When all the weight are approximately of the same value—ideally $w_t^{(i)} = 1/N$, $i = 1, \ldots, N$—then $\overline{N_{\text{eff}}} = N$. If there is a particle j such that $w_t^{(j)} = 1$, and $w_t^{(i)} = 0$ for all $i \neq j$, then $\overline{N_{\text{eff}}} = 1$. Small values of $\overline{N_{\text{eff}}}$ indicate a severe degeneracy of particle weights and the particles should be re-sampled.

Re-sampling is a method for elimination of the particles with low importance weights and copying of those samples with high importance weights. Reproduction of the best particles brings more focus on the promising parts of the state-space. During re-sampling, a random measure $\{\mathbf{x}_t^{(i)}, w_t^{(i)}\}_{i=1}^N$ is replaced with $\{\mathbf{x}_t^{(i)*}, 1/N\}_{i=1}^N$ with uniform weights [RAG04]. Re-sampling is not deterministic. The new set of particles and weights is generated in a way that the probability of sampling a particle $\mathbf{x}_t^{(j)}$ from discrete approximation of $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ is given by its normalized importance weight $w_t^{(j)}$:

$$\Pr\left(\mathbf{x}_t^{(i)*} = \mathbf{x}_t^{(j)}\right) = w_t^{(j)}, \quad i = 1, \ldots, N. \tag{7.53}$$

The resulting sample is an i.i.d. sample from the discrete approximation of density $p(\mathbf{x}_t|\mathbf{y}_{1:t})$, where the weights of all the particles are uniform.

Illustration of the basic idea behind the re-sampling is in Figure 7.3. The piecewise-constant blue line denotes the cumulative weight $\sum_i w^{(i)}$ of $N = 10$ particles. Particles with high weights have a high probability being re-sampled. The higher the weight $w_t^{(i)}$, the longer the interval

$$I = \left[\sum_{s=1}^{i-1} w_t^{(s)}, \sum_{s=1}^{i} w_t^{(s)}\right), \quad i = 1, \ldots, N, \tag{7.54}$$

and the higher the probability that random samples $u_i \sim \mathcal{U}[0,1)$, denoted with dashed lines, will be from I. In the figure, particle 1 was copied twice, particle 2 once, particle 5 for three times, particle 6 once, particle 8 twice and particle 10 once. These 10 samples with uniform weights represent the re-sampled empirical density.

Example of a re-sampling algorithm is the *systematic re-sampling* given in Algorithm 7.3, where we have to sample only one number from $\mathcal{U}[0,1)$. Modification of the SIS algorithm with re-sampling is called *sampling-importance-resampling* (SIR), see Algorithm 7.4. More on re-sampling algorithms can be found, e.g., in [DC05].
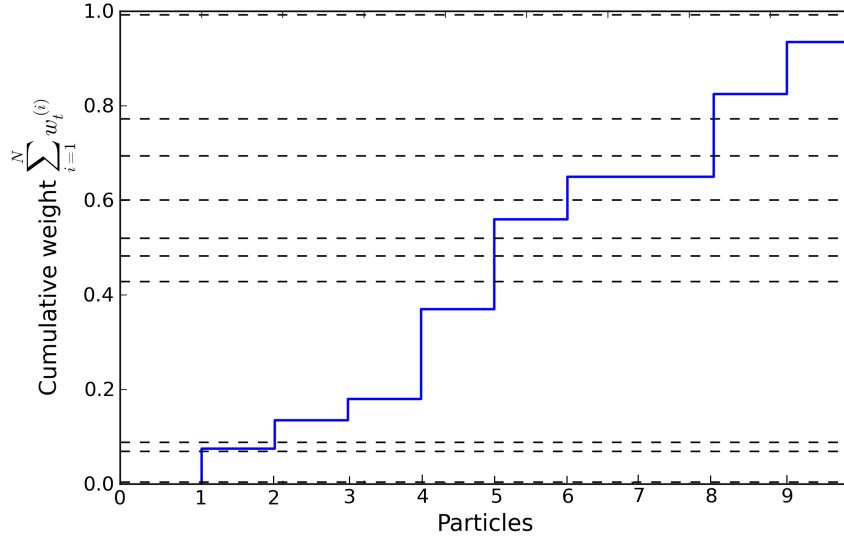
Figure 7.3: Illustration of basic principle of re-sampling in PF. The piecewise-constant blue line denotes the cumulative weight $\sum_i w^{(i)}$ of $N = 10$ particles. The higher the weight $w^{(i)}$, the longer the interval I (7.54) and the higher the probability that random samples $u_i \sim \mathcal{U}[0,1)$, denoted with dashed lines, are from I. Particle 1 was copied twice, particle 2 once, particle 5 for three times, particle 6 once, particle 8 twice and particle 10 once.

7.5.2 Choice of proposal density

The choice of proposal density $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ plays a crucial role in particle filtering. There is no easy prescription for choosing a good proposal density, nevertheless, we can summarize its typically desirable properties [OB92]:

- It should have convenient Monte Carlo properties, i.e., it should be easy to draw random samples from it.

- The tails of $q(\cdot)$ should not be sharper than the tails of $p(\cdot)$. Otherwise, approximation of $p(\cdot)$ may have a large variance or even fail to converge.

- Proposal density $q(\cdot)$ should mimic $p(\cdot)$ well.

The most straightforward choice of the proposal density in the recursive scheme (7.46) is the state transition pdf

$$q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)}, \mathbf{y}_t) = p(\mathbf{x}_t|\mathbf{x}_{t-1}). \tag{7.55}$$

Under the choice (7.55), the recursion of weights (7.48) is given by

$$w_t^{(i)} \propto w_{t-1}^{(i)} p(\mathbf{y}_t|\mathbf{x}_t^{(i)}). \tag{7.56}$$

However, this popular choice is rather sub-optimal.

Algorithm 7.3 Systematic re-sampling.

1. Generate N increasingly ordered numbers

$$u_j = \frac{(j-1) + \tilde{u}}{N}, \quad j = 1, \ldots, N,$$

where $\tilde{u}$ is sampled from uniform distribution $\mathcal{U}(0,1)$.

2. Produce new set of particles. Particle $\mathbf{x}_t^{(i)}$ is copied n_i-times, where

$$n_i \text{ is the number of } u_k \in \left[\sum_{s=1}^{i-i} w_t^{(s)}, \sum_{s=1}^{i} w_t^{(s)} \right).$$

More advanced approaches are based on adaptive selection of the proposal, where we assume a parameterized form of the proposal in time t and estimate its parameters using weights $w_t^{(i)}$ [ADFDJ03, ADH10]. The approach can be extended for estimation of the proposal of the whole state trajectory up to time t. From the re-estimated proposal $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ we can generate new population of trajectories $\mathbf{x}_{1:t}^{(i)}$ and recompute the weights from the beginning up to time t. For instance, given that the proposal density is assumed to be a multidimensional Gaussian pdf, we use the weights for estimation of its mean and covariance matrix. This example is a simple choice and more elaborated proposal densities can be constructed, e.g., a parametrized proposal in the form of a Gaussian mixture estimated using the EM algorithm [DLR+77].

The PF algorithm with an adaptive selection of the proposal density is summarized in Algorithm 7.5. In Step 2, weights $w_t^{(i)}$ are computed using up-to-now observation $\mathbf{y}_{1:t}$. The weights are used for re-estimation of the proposal density $q(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ in Step 3. Up-to-now trajectories of particles are discarded and new trajectories are sampled from the corrected proposal in Step 4. The new trajectories are augmented with the forecasts of the state values for the next time step evaluated using the state transition density. In Step 5, we increment the time index and proceed to Step 2, where the weights are recomputed using the new state trajectories.

7.5.3 Practical evaluation of weights

7.5.3.1 Evaluation of weighs in logarithmic scale

In practice, it is beneficial to evaluate non-normalized weights $\tilde{w}_t^{(i)}$ in a logarithmic scale. In the following text, let

$$\ln \tilde{w}_t^{(i)} = \ln(\tilde{w}_t^{(i)}), \quad i = 1, \ldots, N.$$

Before normalization, we subtract the value of the maximum weight $\tilde{w}_t^{\max}$ from all the weights, as follows,

---

Algorithm 7.4 Sampling–importance–re-sampling algorithm (particle filter).

1. Initialization. For $i = 1, \ldots, N$ initialize particles $\mathbf{x}_{0|-1}^{(i)} \sim p(\mathbf{x}_0)$ and set $t := 0$.

2. PF data update: Evaluate the importance weights

$$\tilde{w}_t^{(i)} = w_{t-1}^{(i)} p\left(\mathbf{y}_t | \mathbf{x}_{t|t-1}^{(i)}\right), \quad i = 1, \ldots, N,$$

   and normalize $w_t^{(i)} = \tilde{w}_t^{(i)} / \sum_{j=1}^{N} \tilde{w}_t^{(j)}$.

3. Re-sampling: Evaluate estimate of effective sample size $\overline{N_{\text{eff}}}$. If $\overline{N_{\text{eff}}} < N_{\text{Thr}}$, where N_{Thr} is a given threshold, sample N particles, with replacement, according to

$$\Pr\left(\mathbf{x}_{t|t}^{(i)} = \mathbf{x}_{t|t-1}^{(j)}\right) = w_t^{(j)}, \quad i = 1, \ldots, N,$$

   and set uniform weights $w_t^{(i)} = \frac{1}{N}, \quad i = 1, \ldots, N$.

4. PF time update: Predict new particles according to

$$\mathbf{x}_{t+1|t}^{(i)} \sim p\left(\mathbf{x}_{t+1|t} | \mathbf{x}_{t|t}^{(i)}\right), \quad i = 1, \ldots, N.$$

5. Set $t := t + 1$ and iterate from step 2.

Algorithm 7.5 Particle filter with adaptive selection of proposal density.

1. Initialization. For $i = 1, \ldots, N$ initialize particles $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0)$ and set $t := 0$.

2. Evaluation of normalized weights $w_t^{(i)}$

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

3. Adaptive selection of proposal density $q(\mathbf{x}_{1:t} | \mathbf{y}_{1:t})$ for the next time step using weights $w_t^{(i)}$ and particle trajectories $\mathbf{x}_{1:t}^{(i)}$.

4. Sample new trajectories from for the next time step:

   (a) $\mathbf{x}_{1:t+1}^{(i)} \sim q(\mathbf{x}_{1:t+1} | \mathbf{y}_{1:t})$, where $q(\mathbf{x}_{1:t+1} | \mathbf{y}_{1:t}) = q(\mathbf{x}_{1:t} | \mathbf{y}_{1:t}) p(\mathbf{x}_{t+1} | \mathbf{x}_t)$.

   (b) Reset particle weights $w_t^{(i)} = 1/N$, $i = 1, \ldots, N$.

5. Set $t := t + 1$ and iterate from step 2.

$$\ln \tilde{w}_t^{(i)*} = \ln \tilde{w}_t^{(i)} - \ln \tilde{w}_t^{\max} = \ln \left(\frac{\tilde{w}_t^{(i)}}{\tilde{w}_t^{\max}}\right),$$

where $\ln \tilde{w}_t^{\max} = \max_{i \in N} \ln \tilde{w}_t^{(i)}$.

This procedure ensures, that the weights are better scaled and the overall algorithm is numerically more stable. After exponentiation, the maximum weight is equal to 1 and there is always at least one particle which has a reasonable weight after normalization. The procedure does not affect the resulting normalized weights:

$$w_t^{(i)*} = \frac{\tilde{w}_t^{(i)*}}{\sum_{j=1}^N \tilde{w}_t^{(j)*}} = \frac{\frac{\tilde{w}_t^{(i)}}{\tilde{w}_t^{\max}}}{\sum_{j=1}^N \frac{\tilde{w}_t^{(i)}}{\tilde{w}_t^{\max}}} = \frac{\frac{\tilde{w}_t^{(i)}}{\tilde{w}_t^{\max}}}{\frac{1}{w_t^{\max}} \sum_{j=1}^N \tilde{w}_t^{(j)}} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}} = w_t^{(i)}.$$

7.5.3.2 Effective evaluation of Gaussian likelihood

Let the likelihood function given by the observation model (7.4) be of a multidimensional Gaussian type, $p(\mathbf{y}_t|\mathbf{x}_t) = \mathcal{N}(\mathcal{H}_t(\mathbf{x}_t), \mathbf{Z}_t)$, more specifically

$$p(\mathbf{y}_t|\mathbf{x}_t) \quad = \quad (2\pi)^{-\frac{N_y}{2}} (\det \mathbf{Z}_t)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left((\mathbf{y_t} - \mathcal{H}_\mathbf{t}(\mathbf{x_t}))^\mathrm{T} \mathbf{Z}_t^{-1}(\mathbf{y_t} - \mathcal{H}_\mathbf{t}(\mathbf{x_t}))\right)\right].$$

A logarithmic weight is then

$$\ln \tilde{w}_t \quad = \quad -0.5\left(\ln \det \mathbf{Z}_t + \mathbf{v}_t^\mathrm{T} \mathbf{Z}^{-1} \mathbf{v}_t\right).$$

Let $\mathbf{F}_t$ be the lower triangular Cholesky factor of $\mathbf{Z}_t$. It can be shown [GVL96], that it holds

$$\ln \tilde{w}_t = \ln \prod_{i=1}^{N_y} (\mathbf{F}_t[i,i])^2 + ||\mathbf{F}_t^{-1} \mathbf{v}_t||^2, \tag{7.57}$$

where $\mathbf{F}_t[i,i]$, $i = 1, \ldots, N_y$, are diagonal elements of $\mathbf{F}_t$ and $||\cdot||$ denotes the Euclidean 2-norm.

7.6 Marginalized particle filter

The main advantage of importance sampling is its generality. Particle filters are capable of approximating an arbitrary density via empirical density at the price of high computational cost, which is prohibitive in high-dimensional problems. This obstacle can be overcome in the cases, where the structure of the model (7.3)–(7.4) allows analytical marginalization over a subset, $\mathbf{x}_t^c$, of the full state vector

$$\mathbf{x}_t = \left[\begin{array}{c} \mathbf{x}_t^c \\ \mathbf{x}_t^p \end{array}\right]. \tag{7.58}$$

Using the chain rule and the factorization (7.58), the posterior $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ has the form

$$p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}) = \underbrace{p(\mathbf{x}_{1:t}^c|\mathbf{x}_{1:t}^p, \mathbf{y}_{1:t})}_{\text{analytical filter}} \underbrace{p(\mathbf{x}_{1:t}^p|\mathbf{y}_{1:t})}_{\text{PF}}, \tag{7.59}$$

where $p\left(\mathbf{x}_t^c|\mathbf{x}_{1:t}^p, \mathbf{y}_{1:t}\right)$ is analytically tractable, while $p\left(\mathbf{x}_{1:t}^p|\mathbf{y}_{1:t}\right)$ is not [DDFG01, SGN05], and we use particle filter for its approximation. This technique is referred as Rao-Blackwellization [DDFG01]. We replace the term $p\left(\mathbf{x}_{1:t}^p|\mathbf{y}_{1:t}\right)$ in (7.59) by a weighted empirical distribution, in analogy to (7.40), yielding

$$p\left(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}\right) \approx \sum_{i=1}^n w_t^{(i)} p\left(\mathbf{x}_{1:t}^c|\mathbf{x}_{1:t}^{p,(i)}, \mathbf{y}_{1:t}\right) \delta\left(\mathbf{x}_{1:t}^p - \mathbf{x}_{1:t}^{p,(i)}\right), \tag{7.60}$$

$$w_t^{(i)} \propto \frac{p\left(\mathbf{x}_{1:t}^{p,(i)}|\mathbf{y}_{1:t}\right)}{q\left(\mathbf{x}_{1:t}^{p,(i)}|\mathbf{y}_{1:t}\right)}. \tag{7.61}$$

Note that now we only have to sample from the space of $\mathbf{x}_t^p$. Recursive evaluation is achieved by application of the Bayes rule

$$p\left(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}\right) \propto p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p\left(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1}\right), \tag{7.62}$$

and substitution of (7.60) in place of $p(\mathbf{x}_{1:t}|\mathbf{y}_{1:t})$ and $p(\mathbf{x}_{1:t-1}|\mathbf{y}_{1:t-1})$. Comparing elements in the summations on both sides of equation (7.62), we obtain:

$$w_t^{(i)} \propto \frac{p\left(\mathbf{y}_t, \mathbf{x}_t^{p,(i)}|\mathbf{x}_{1:t-1}^{p,(i)}, \mathbf{y}_{1:t-1}\right)}{q\left(\mathbf{x}_t^{p,(i)}|\mathbf{x}_{1:t-1}^{p,(i)}, \mathbf{y}_{1:t}\right)}w_{t-1}^{(i)}, \tag{7.63}$$

$$p(\mathbf{y}_t, \mathbf{x}_t^{p,(i)}|\mathbf{x}_{1:t-1}^{p,(i)}, \mathbf{y}_{1:t-1}) = \int p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}^c|\mathbf{x}_{1:t-1}^{p,(i)}, \mathbf{y}_{1:t-1})d\mathbf{x}_t^c d\mathbf{x}_{t-1}^c. \tag{7.64}$$

The requirement of analytical tractability of integrations in (7.64) is always fulfilled when (7.1) contains a linear-Gaussian part, [SGN05], giving rise to the marginalized particle filter (MPF) with the Kalman filter. Resulting approximation of the posterior pdf (7.60) then becomes a weighted sum of Gaussian pdfs

$$p\left(\mathbf{x}_{1:t}|\mathbf{y}_{1:t}\right) \approx \sum_{i=1}^n w_t^{(i)} \mathcal{N}\left(\mathbf{x}_t^c; \bar{\mathbf{x}}_t^{c,(i)}, \mathbf{P}_t^{(i)}\right) \delta\left(\mathbf{x}_{1:t}^p - \mathbf{x}_{1:t}^{p,(i)}\right),$$

where $\bar{\mathbf{x}}_t^{c,(i)}$ and $\mathbf{P}_t^{(i)}$ are mean values and covariance matrices of Gaussian distributions $\mathcal{N}(\mathbf{x}_t^c; \bar{\mathbf{x}}_t^{c,(i)}, \mathbf{P}_t^{(i)})$ attached to particles $\mathbf{x}_t^{p,(i)}$.

Using the results from Appendix A, the minimum mean square error estimates of the expected value $\hat{\mathbf{x}}_t^c$ and covariance $\mathbf{P}_t$ of the resulting posterior mixture $p(\mathbf{x}_t^c|\mathbf{x}_{1:t}^p, \mathbf{y}_{1:t})$ are given, as follows:

$$\hat{\mathbf{x}}_t^c = \sum_{i=1}^N w_t^{(i)} \bar{\mathbf{x}}_t^{c,(i)}, \tag{7.65}$$

$$\mathbf{P}_t = \sum_{i=1}^N w_t^{(i)} \left[\mathbf{P}_t^{(i)} + \left(\bar{\mathbf{x}}_t^{c,(i)} - \hat{\mathbf{x}}_t^c\right)\left(\bar{\mathbf{x}}_t^{c,(i)} - \hat{\mathbf{x}}_t^c\right)^{\mathrm{T}}\right]. \tag{7.66}$$

Illustration of a 2-dimensional state $\mathbf{x} = [x_1, x_2]^\mathrm{T}$ estimated using MPF is in Figure (7.4). Resulting joint estimate of the posterior pdf is in the linear-Gaussian part of the state, x_1, estimated using the optimal Kalman filter and approximated using a particle filter in the non-linear and/or non-Gaussian part of the full state vector, x_2.

Tractable solution also exists for discrete-variable models [TBF05] and models based on conjugate statistics [SOvG10]. However, the range of models amenable to this approach is still rather small and does not contain any models suitable for large-scale and non-linear problems.
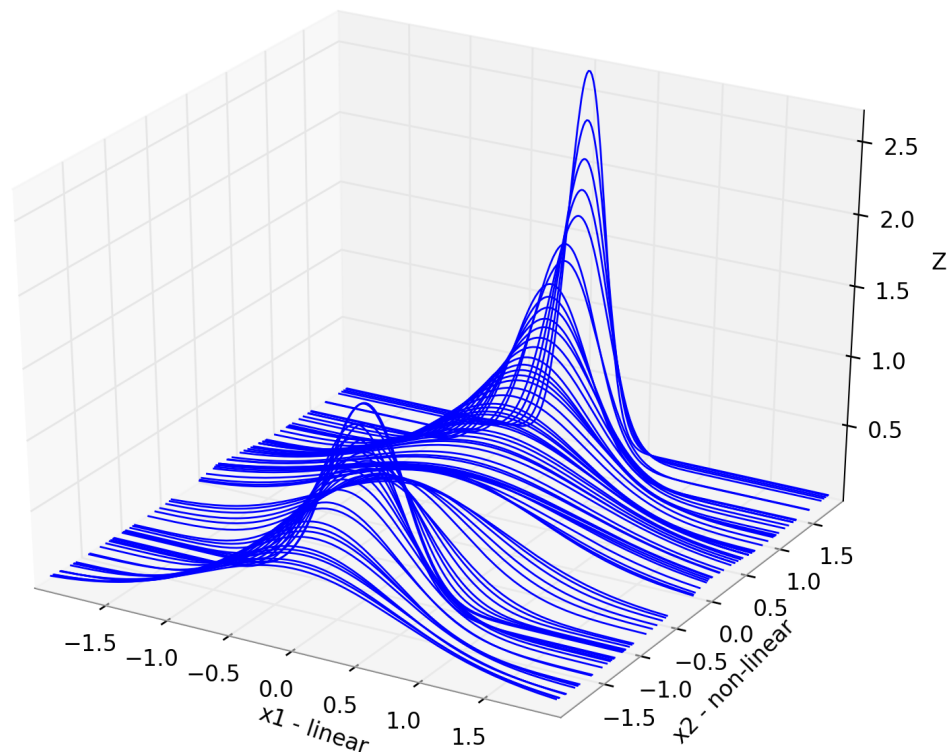
Figure 7.4: Illustration of a 2-dimensional state $\mathbf{x} = [x_1, x_2]^\mathrm{T}$ estimated using MPF. Resulting joint estimate of the posterior pdf is in the linear-Gaussian part of the state, x_1, estimated using the optimal Kalman filter and approximated using a particle filter in the non-linear and/or non-Gaussian part of the full state vector, x_2.

Chapter 8

Framework for Tuning of Ensemble Filters

8.1 Modification of MPF algorithm for suboptimal conditional filters

Since the optimal Kalman filter is not suitable for large scale and non-linear problems, it would advantageous to substitute it with an approximate filter within MPF. We propose to relax the requirement of exact marginalization in (7.63) and replace it by an approximation. We note that given numerical values of $\mathbf{x}_t^{p,(i)}, \mathbf{x}_{t-1}^{p,(i)}$, equation (7.64) is equivalent to the normalizing constant of a Bayesian filter (7.5). Hence, any Bayesian filter that is capable of evaluating its normalizing constant can be used to approximate (7.64). What results is an algorithm equivalent to marginalized particle filtering where the analytical Kalman filters are replaced by approximate conditional filters. Specifically, the following filters interact via Algorithm 8.1:

1. Conditional filter on variable $\mathbf{x}_t^c$, treating $\mathbf{x}_t^p$ as an observation, i.e.,

$$p(\mathbf{x}_t^c|\mathbf{x}_{1:t}^p, \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t, \mathbf{x}_t^p|\mathbf{x}_t^c, \mathbf{x}_{1:t}^p)p(\mathbf{x}_t^c|\mathbf{x}_{1:t-1}^p, \mathbf{y}_{1,t-1})}{p(\mathbf{y}_t, \mathbf{x}_t^p|\mathbf{x}_{1:t-1}^p, \mathbf{y}_{1,t-1})}. \tag{8.1}$$

2. Particle filter on variable $\mathbf{x}^p$, that handles sampling from the proposal density $q(\mathbf{x}_t^p|\mathbf{x}_{t-1}^p, \mathbf{y}_{1:t})$ and re-sampling. Each particle is attached to one conditional filter.

In this general form, the algorithm is rather unspecific. This is due to the fact that arbitrary conditional filters can be combined with arbitrary particle filtering approaches. Therefore, we consider Algorithm 8.1 to represent a framework for designing specific filtering variants. The word framework is used to distinguish this approach from the analytical MPF.

The key property of the MPF is partitioning of the state variable into two parts. In the original exact formulation, the choice of partitioning is fully determined by

---

Algorithm 8.1 General MPF framework.

1. Generate initial particles $\mathbf{x}_0^{p,(i)}$, $i = 1, \ldots, N$, and set initial statistics of all associated conditional filters.

2. For each new data record $\mathbf{y}_t$ do:

   (a) Sample new value of particles $\mathbf{x}_t^{p,(i)}$, $i = 1, \ldots, N$, and update statistics of all associated conditional filters via (8.1).

   (b) Compute weights (7.63) of all particles and their associated conditional filters.

   (c) If the number of efficient particles, $\overline{N_{\text{eff}}}$, is lower than the chosen threshold, re-sample the particles.

---

tractability of the Bayes rule (7.5). Finding a partitioning in the context of environmental modeling where the state variables typically obey the same equations is harder. However, the relaxed formalization of Section 7.6 allows to interpret $\mathbf{x}_t^p$ not as a partition of the full state but rather as an augmentation of the original state $\mathbf{x}_t^c$ by nuisance parameters. What results is a framework for on-line tuning of existing filters.

The general algorithm of tuning is described in Algorithm 8.1, specific variants arise for the following choices:

1. Choose a preferred variant of the conditional filter (e.g. a variant of ensemble filter) estimating $\mathbf{x}_t^c$,

2. Choose tuning parameters of interest, $\mathbf{x}_t^p$, use them to augment the original state $\mathbf{x}_t^c$ via the chosen model of their evolution, $p(\mathbf{x}_t^p|\cdot)$,

3. Choose a proposal density $q(\mathbf{x}_t^p|\cdot)$, e.g., the evolution model, $q(\mathbf{x}_t^p|\cdot) \equiv p(\mathbf{x}_t^p|\cdot)$.

Different choices in each of the points above will lead to different properties of the resulting filter. The number of possible combinations of these choices is enormous, and finding guidelines for the best option in a given application context is a task for further research. In some applications, a physically motivated evolution model of $\mathbf{x}_t^p$ may be found, while heuristic or expert-chosen models may be more appropriate in others.

8.2   Estimation of inflation factor, observation error variance, and length-scale parameter

We focus on ensemble filters, which are suitable for large-scale problems arising in spatial data analysis. Generally, ensemble methods tend to underestimate model error, which can significantly decrease filtering performance or even result in divergence. The techniques for compensating small ensemble issues described in Section 7.4.3 have tuning parameters that typically need to be set-up experimentally. Substantial effort has been put into on-line estimation of the inflation factor alone, [And07a], or in tandem

with the observation error [LKM09]. In this Section, we approach the same problem using the MPF framework with the following specific choices.

1. We have chosen the EnSRF with multiplicative inflation (7.35) as our conditional filter.

2. The unknown tuning parameters regarding model error are: the time-variant inflation factor Δ_t and the time-varying length-scale parameter l_t of the covariance localization function. We can include also different types of parameters like magnitude of observation error, r_t, for all observations, i.e., $\mathbf{R}_t = r_t\mathbf{I}$. The augmentation of the state vector is then $\mathbf{x}_t^p = [\Delta_t, r_t, l_t]^{\mathrm{T}}$.

3. The proposal density is chosen as $p(\mathbf{x}_t^p|\mathbf{x}_{t-1}^p)$.

Augmentation $\mathbf{x}_t^p$ is evolved using

$$p(\mathbf{x}_t^p|\mathbf{x}_{t-1}^p) = p(\Delta_t|\Delta_{t-1})p(r_t|r_{t-1})p(l_t|l_{t-1}),$$

where, evolution of the parameters is modeled by truncated Gaussian random walks,

$$\begin{aligned}
p(\Delta_t|\Delta_{t-1}) &= t\mathcal{N}(\Delta_{t-1}, \sigma_\Delta^2, [1,\infty]), \\
p(r_t|r_{t-1}) &= t\mathcal{N}(r_{t-1}, \sigma_r^2, [0,\infty]), \\
p(l_t|l_{t-1}) &= t\mathcal{N}(l_{t-1}, \sigma_l^2, [0,\infty]).
\end{aligned} \tag{8.2}$$

Scalar parameters σ_Δ, σ_r, and σ_l denote the spread of the random walks, respectively. Non-negativity of all considered parameters motivates truncation of support of the random walks.

Under the choice of proposal density (7.63) reduces to

$$w_t^{(i)} = w_{t-1}^{(i)}p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathbf{x}_t^{p,(i)}), \tag{8.3}$$

where $p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathbf{x}_t^{p,(i)})$ defined by (7.24) is now explicitly conditioned on the unknown parameters:

$$\begin{aligned}
p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathbf{x}_t^{p,(i)}) \quad &\propto \quad \det\left(\mathbf{Z}_t(\mathbf{x}_t^{p,(i)})\right)^{-\frac{1}{2}} \times \\
&\times \exp\left[-\frac{1}{2}\left(\mathbf{y}_t - \mathbf{H}\bar{\mathbf{x}}_{t|t-1}^{c,(i)}\right)^{\mathrm{T}}\mathbf{Z}_t^{-1}(\mathbf{x}_t^{p,(i)})\left(\mathbf{y}_t - \mathbf{H}\bar{\mathbf{x}}_{t|t-1}^{c,(i)}\right)\right]. \quad (8.4)
\end{aligned}$$

Resulting algorithm defines an adaptation scheme related to other approaches used in the literature. Specifically, (8.4) is the same equation that was used for maximum likelihood estimation of covariance parameters [Dee95]. Maximization of this function is achieved, e.g., via simplex methods [MH00]. In our approach, (8.4) serves as a likelihood function for Bayesian estimation of the tuning parameters, $\mathbf{x}_{p,t}$. The variance of the random walk then models our belief in time-variability of the tuning parameters. In

the special case of stationary parameters, (i.e., $\sigma_\Delta = \sigma_r = \sigma_l = 0$), Algorithm 8.1 is reduced to a parallel run of N ensemble filters, each of which is accumulating the product of $p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathbf{x}_t^{p,(i)})$ in each step. After several hundreds of steps, majority of the weights will converge to zeros and one of them will converge to one. Such behavior is known as sample impoverishment in the particle filtering literature. The convergence of probability mass to a single point may be useful for finding the best tuned values in off-line phase. However, this degeneracy is undesirable for on-line application, and non-zero variances of random walks (8.2) have to be used.

For non-stationary parameters, each of the N filters follows a random walk of the tuning parameters. The re-sampling operation removes filters that diverged into unlikely regions, and replaces them by copies of the filters with parameters that are more likely. The area of higher likelihood is then explored by more filters in detail. This of course requires to run N ensemble filters in parallel which is computationally expensive. However, the key advantage of this approach is that it is able to optimize non-convex and multi-modal likelihood functions.

8.3   Simulation studies

8.3.1   Lorenz-96 model

To demonstrate versatility of the method, we test the MPF approach in the Lorenz-96 model by [LE98] which has been widely used in simulation studies. The model is given by

$$\frac{dx_j}{dt} = x_{j-1}(x_{j+1} - x_{j-2}) - x_j + F, \tag{8.5}$$

where F is the model forcing and x_j are variables forming a cyclic chain. We define $x_{-1} = x_{J-1}$, $x_0 = x_J$ and $x_{J+1} = x_1$ to make (8.5) meaningful for all values of $j = 1 \ldots J$. We use 40 variables, and $F = 8$ for the strength of forcing. The model (8.5) can be integrated forward with the fourth-order Runge–Kutta scheme. The system is computationally stable for step of 0.05 non-dimensional units, which is also the step of the analysis. All the experiments are performed as twin experiments.

8.3.2   Stationary parameters

To create a baseline for comparison of adaptive tuning strategies, we performed parallel run of EnSRFs for fixed values of Δ, l selected on a rectangular grid, as in [WH02]. From Bayesian point of view, this setup corresponds to estimation of stationary parameters:

$$p(\Delta, l|\mathbf{y}_{1:t}) \propto p(\Delta, l) \prod_{j=1}^{t} p(\mathbf{y}_j|\mathbf{y}_{1:j-1}, \Delta, l) \tag{8.6}$$

where $p(\Delta, l)$ is a prior probability density on discrete values of Δ, l at the grid points, which is uniform, and $p(\mathbf{y}_j|\mathbf{y}_{1:j-1}, \Delta, l)$ is given by (8.4). For numerical stability, (8.6)

| | RMSE | | | marginalized log-likelihood | | |
rank	inflation	localization	mean value	inflation	localization	sum
#1	1.04	7.0	0.2074	1.05	7.0	-2079401
#2	1.04	6.0	0.2075	1.05	8.0	-2079483
#3	1.05	8.0	0.2076	1.04	6.0	-2079513

Table 8.1: Best stationary choices of nuisance parameters for EnSRF according to two criteria.

is often computed in logarithmic scale where the product is replaced by the sum of marginal log-likelihoods.

The observation data are generated from the perfect model scenario where the "true" state was generated by integrating the Lorenz-96 model (8.5) for 100000 steps and observations are generated from the true state by addition of zero-mean Gaussian noise with variance $r = 1$. The analysis was performed by the EnSRF with covariance localization constructed using a compactly supported fifth-order piecewise rational function given by (7.37) with length-scale parameter l. The results of a simulation experiment with 132 EnSRFs with 15 ensemble members, $r = 1$, $\Delta = [1.0, 1.1, \ldots, 1.10]$, and $l = [0, 1, \ldots, 11]$ are displayed in Figure 8.1 in two modalities. First, the traditional RMSE is computed for each couple of parameters,

$$\text{RMSE} = \frac{1}{99000} \sum_{t=1000}^{100000} \sqrt{\frac{1}{40}(\mathbf{x}_t - \hat{\mathbf{x}}_t^c)^{\text{T}}(\mathbf{x}_t - \hat{\mathbf{x}}_t^c)}, \tag{8.7}$$

for the MPF algorithm. Second, the sum of marginal log-likelihoods (8.6) within the same time intervals is displayed for illustration.

Note that the contours of the marginal log-likelihood (8.6) correspond closely to the contours of the RMSE. This suggests that the marginal likelihood $p(y_\tau | y_{1:\tau-1}, \Delta, l)$ is a good measure to optimize for the best RMSE in the cases where the true state values are not known. The three best choices within each modality are given in Table 8.1. Note that two choices—$\Delta = 1.05, l = 8$ and $\Delta = 1.04, l = 6$—are in the top three for both criteria. The relative differences are rather small, however, in terms of normalized posterior probability (8.6) the best parameters in Table 8.1 are e^{82} times more likely that the second best.
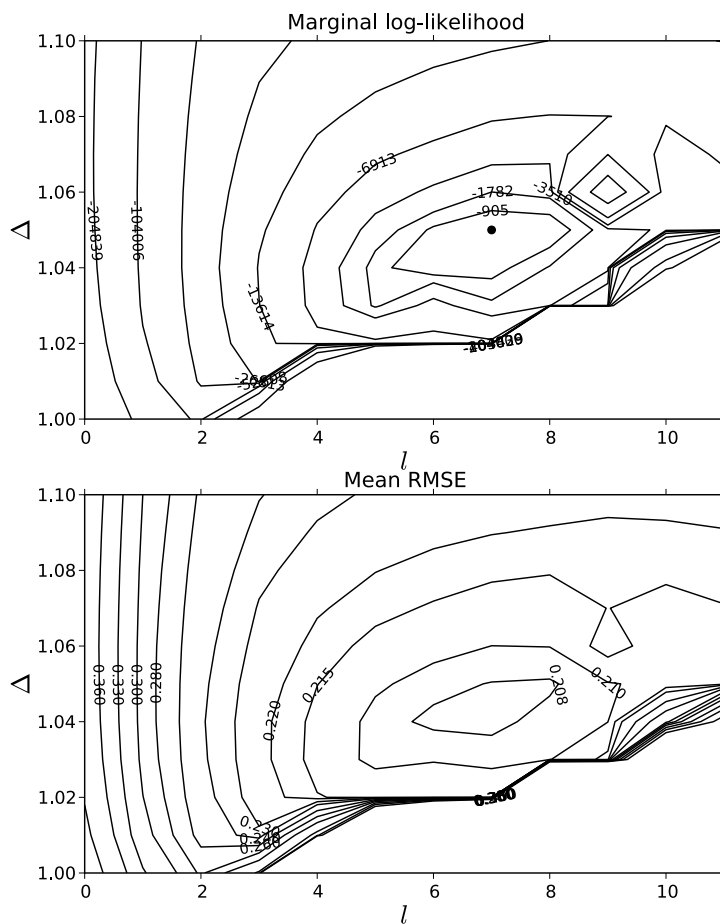
Figure 8.1: Estimation of stationary parameters on rectangular grid. Top: marginal log-likelihood (8.6) of the tuned parameters Δ, l. Labels of the contour lines denote difference from the maximum which is marked by a circle. Bottom: Time average of RMSE (8.7).

8.3.3 Adaptive estimation in perfect model scenario

The same observation data used for estimation of stationary parameters were used to estimate the time-varying parameters in two different scenarios:

Scenario (i): fixed $r_t = 1.0$, estimated Δ_t, l_t, i.e., $\mathbf{x}_t^p = [\Delta_t, l_t]^\mathrm{T}$,

Scenario (ii): estimated r_t, l_t, Δ_t, i.e., $\mathbf{x}_t^p = [\Delta_t, r_t, l_t]^\mathrm{T}$.
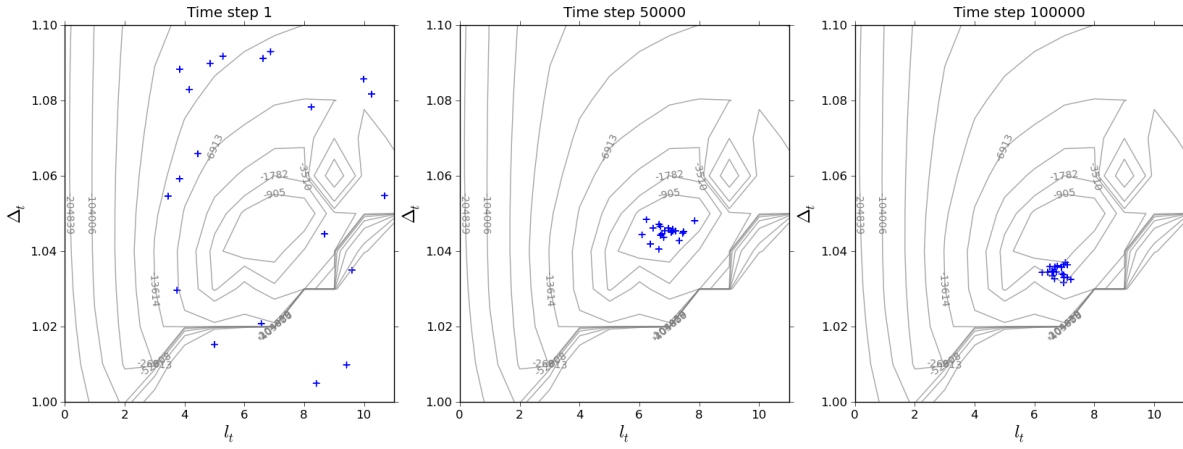
Figure 8.2: Position of the particles $\Delta_t^{(i)}, l_t^{(i)}$ denoted by '+' on the background of stationary contours of marginal log-likelihood.

The particle filter is using multinomial re-sampling [GSS02] and $\overline{N_{\text{eff}}} = 0.8N$. The variances of random walks (8.2) for the MPF algorithm were chosen as

$$
\begin{aligned}
\sigma_\Delta &= 0.01\Delta_{t-1}^{(i)} + 0.0001, \\
\sigma_l &= 0.01 l_{t-1}^{(i)} + 0.0001, \\
\sigma_r &= 0.005 r_{t-1}^{(i)} + 0.0001,
\end{aligned}
\tag{8.8}
$$

The prior density of the tuning parameters is chosen as uniform on support $p(\sigma_{y,0}) = U(0.1, 4)$, $p(\Delta_0) = U(1.0, 1.10)$, $p(l_0) = U(0.11, 11.11)$.

Results of simulations for different number of particles and both scenarios are displayed in Tables 8.2 and 8.3, respectively. Using particles and their weights, estimates of the parameters are evaluated using

$$
\bar{\mathbf{x}}_t^p = \sum_{i=1}^{n} w_t^{(i)} \mathbf{x}_t^{p,(i)}.
$$

In accordance with [WH02, LKM09], we ignore the first 1000 steps and report the results only for the subsequent steps. Spatial distribution of the particles for Δ_t, l_t in scenario (i) at time steps $t = 1, 50000, 100000$ is displayed in Figure 8.2 on the background of contours for the stationary marginal log-likelihood from Fig 8.1. We note that alignment of the particles in the middle of the stationary contour at $t = 50000$ is a coincidence, in majority of all other time steps the cloud is a bit off the stationary optimum.

As expected, the RMSE is steadily decreasing with increasing number of particles for both considered scenarios. Note that for $N = 10$ and higher, the MPF filter achieves better performance than the best-tuned filter. In the more complex scenario of tuning all three parameters, Table 8.3 the MPF algorithm achieves only negligible increase of the RMSE over the scenario with known observation variance r.

Results for scenario (i)							
N	mean RMSE	std RMSE	mean $\bar{\Delta}_t$	std $\bar{\Delta}_t$	mean $\bar{l}_t$	std $\bar{l}_t$	mean $\bar{r}_t$	std $\bar{r}_t$
5	0.2089	0.0521	1.0334	0.0080	5.6942	0.9072	1.0	0.0
10	0.2071	0.0510	1.0337	0.0064	6.3431	0.9229	1.0	0.0
20	0.2065	0.0523	1.0331	0.0061	6.2933	0.5005	1.0	0.0

Table 8.2: Adaptive tuning of $\mathbf{x}_t^p = [\Delta_t, l_t]^{\mathrm{T}}$ and the resulting analysis RMSE error, averaged over assimilation steps between $t = 1000$ and $t = 100000$, std denotes standard deviation of the estimates from the mean over time.

Results for scenario (ii)							
N	mean RMSE	std RMSE	mean $\bar{\Delta}_t$	std $\bar{\Delta}_t$	mean $\bar{l}_t$	std $\bar{l}_t$	mean $\bar{r}_t$	std $\bar{r}_t$
5	0.2094	0.0521	1.0317	0.0075	5.6154	0.7289	1.0059	0.0215
10	0.2072	0.0509	1.0354	0.0058	6.7455	0.9541	1.0031	0.0230
20	0.2064	0.0511	1.0355	0.0055	6.7182	0.9202	1.0018	0.0193

Table 8.3: Adaptive tuning of $\mathbf{x}_t^p = [\Delta_t, r_t, l_t]^{\mathrm{T}}$ and the resulting analysis RMSE error, averaged over assimilation steps between $t = 1000$ and $t = 100000$, std denotes standard deviation of the estimates from the mean over time.

We note that good performance of the adaptive tuning was achieved for as low as 10 particles. This result is especially promising since it suggests that even more challenging assimilation scenarios can be handled at comparable computational complexity. Addition of one extra tuning parameter in second scenario had negligible impact on the performance.

8.3.4 Model with random perturbations

For comparison with [LKM09], we tested the MPF algorithms on data simulated with model (8.5) with additive random perturbations

$$\frac{dx_j}{dt} = x_{j-1}(x_{j+1} - x_{j-2}) - x_j + F + \alpha e_t, \tag{8.9}$$

where e_t is Gaussian distributed noise with zero-mean and unit variance. The observed data were generated using model (8.9) with $\alpha = 4$ for 100000 steps. The same setup of the EnSRF as in the previous experiments was used, including the same initial conditions. Results for estimation for the 100000 steps are displayed in Table 8.4. Since parameter α is stationary, the time evolution of the parameter estimates $\bar{\Delta}_t, \bar{l}_t$ using MPF is reaching the stationary values in Table 8.4 after the initial convergence period.

Note that the additive noise was compensated by higher values of $\bar{\Delta}_t$ and lower values of $\bar{l}_t$ than that of the perfect model, Table 8.3. This is in agreement with findings of

Results for imperfect model scenario, $\alpha = 4$							
mean RMSE	std RMSE	mean Δ_t	std Δ_t	mean l_t	std l_t	mean $\bar{r}_t$	std $\bar{r}_t$
0.3563	0.0542	1.1944	0.0230	3.0810	0.8198	1.0055	0.0618

Table 8.4: Adaptive tuning of $\mathbf{x}_t^p = [\Delta_t, l_t, r_t]^\mathrm{T}$ for system with random model errors (8.9) with variance $\alpha = 4$. The resulting analysis RMSE error is averaged over assimilation steps between $t = 1000$ and $t = 100000$, time averages of parameter estimates are displayed in tandem with standard deviation of the estimates from the mean over time. Both algorithms were run with $N = 10$ and $\overline{N_{\mathrm{eff}}} = 0.8N$.

[LKM09] and also expected because α is increasing the background covariance (reflected by higher inflation) and decreasing correlation between elements of the state vector (reflected by lower length-scale).

For testing the tracking properties of the MPF algorithms, we have designed a scenario with time varying α_t according to a triangular profile. Posterior densities of the parameters obtained using the MPF algorithm are displayed in Fig. 8.3. For this experiment, we increased the variances of random walks (8.2) to

$$
\begin{aligned}
\sigma_\Delta &= 0.01\Delta_{t-1}^{(i)} + 0.001, \\
\sigma_l &= 0.01l_{t-1}^{(i)} + 0.001, \\
\sigma_r &= 0.01r_{t-1}^{(i)} + 0.01.
\end{aligned}
$$

This experiment confirms the trend of increasing Δ_t and decreasing l_t with increasing α_t. Note that when α returns to the stationary values, so do the estimates of the tuning parameters.

8.4 Summary

The purpose of this chapter is to present marginalized particle filtering (also known as Rao-Blackwellized filtering) as an attractive tool for research of data assimilation methods in environmental modeling and especially for tuning of ensemble filters.

The method is based on partitioning of the state (or unknown parameters) into two parts: (i) unknowns estimated by a conditional filter, and (ii) unknowns estimated by a particle filter. The original MPF assumes that the conditional filter is analytically tractable, which allows to prove advantages over a pure particle filter. In this chapter, we propose to replace analytical filters by ensemble filters. The resulting algorithm is loosing its theoretical advantages, however it allows to address the problem of tuning of ensemble filters. We have shown that the number of particles needed to achieve acceptable performance is rather low, for example 10 particles are sufficient to achieve on-line tuning of the inflation factor and the length-scale parameter in the EnSRF for the 40-dimensional Lorenz-96 model. Furthermore, we have shown that the approach
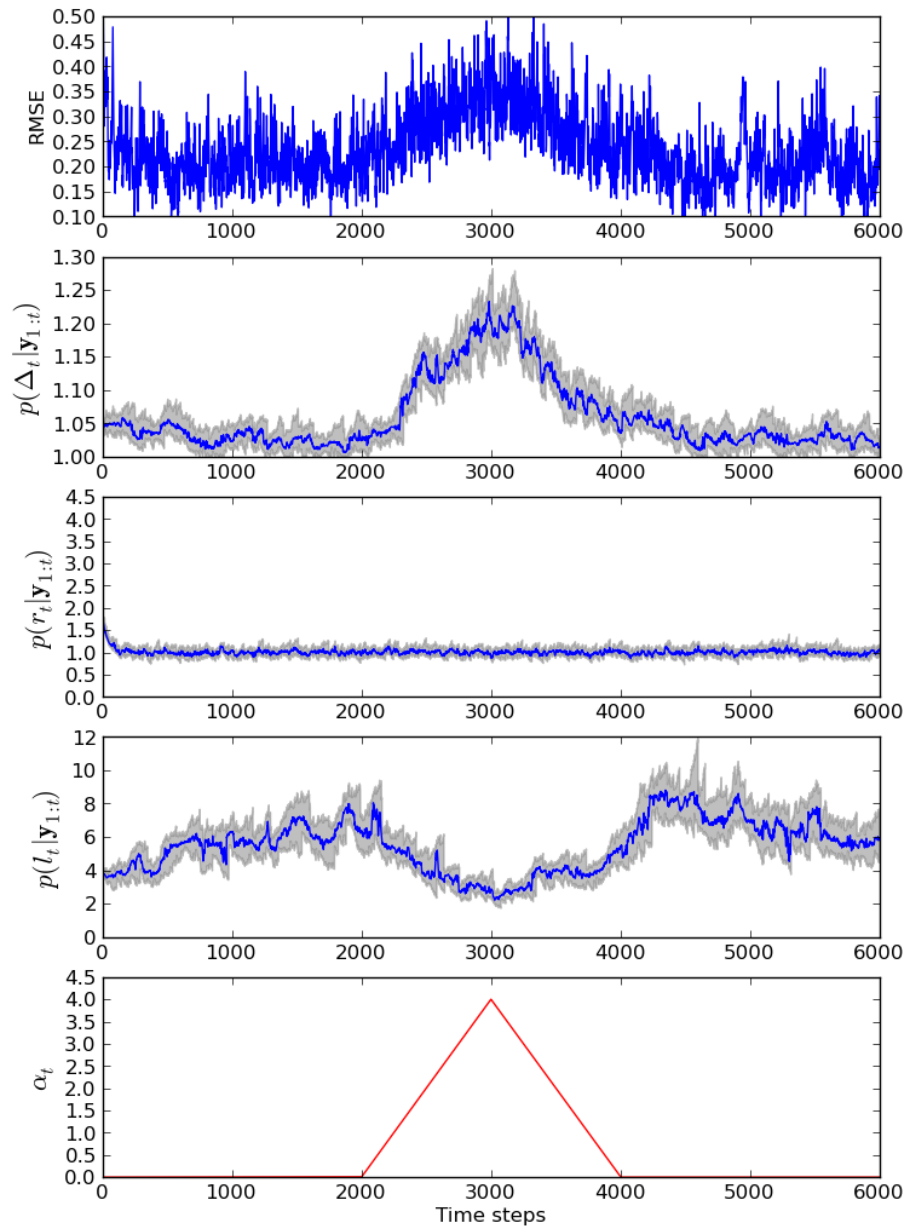
Figure 8.3: Estimation of system (8.9) with time-varying α_t of triangular profile displayed at the bottom. Posterior densities of the parameters are displayed via their mean value (blue line) and region between minimum and maximum value of the particles (gray area).

easily extends to estimation of unknown variance of the observation error and potentially any other tuning parameters. Once again, 10 particles were sufficient to achieve performance comparable to that of the best-tuned filter.

Computational cost of the MPF framework is high since it requires running N filters in parallel. We expect that advantages of parallel evaluation of ensemble filters over adaptation of a single ensemble filter will become apparent in even more demanding scenarios. Computational complexity may prevent its operational use, however, it may be an important tool for gaining insight into the ensemble filters in the same spirit as in [And07b].

The potential of the framework has been demonstrated on on-line tuning of the ensemble filters. However, it is not the only scenario where it can be used. Since posterior density of the MPF is a mixture of Gaussians, the approach may be adapted for estimation of Gaussian mixture filters that have been studied, e.g. by [BSN03]. More work is required to discover full potential of the method. The open problems include justified design of suitable models of parameter evolution and approximations reducing the computational cost of the MPF. However, the existence of the exact solution allows to design the necessary computational simplifications to resemble its behavior.

More experiments and a heuristic simplification suitable for certain assimilation scenarios can be found in [ŠH11].

Chapter 9

Application of Bayesian DA in the Early Phase

9.1 Problem statement

Assume an accident in a nuclear power plant followed by an atmospheric release of radionuclides. After the release, there is a radioactive plume moving over the terrain. Urgent protective measures must be introduced as soon as possible to protect the public from the harmful effects of ionizing radiation. These are planned with regards to expected exceeding of regulatory radiation limits given by the law. Decision making regarding countermeasures is supported with radiological measurements from terrain. However, particularly in the first hours of the accident, the measurements are sparse and it is not possible to base prognoses of radiation situation just upon them solely. For determination of affected areas and estimation of radiation levels in a wider scale, atmospheric dispersion models (ADMs) are used. Given a meteorological forecast and values of other important control variables of an ADM, the model evaluates a *prediction* of spatial and temporal distribution of radionuclides on a computational grid in terms of activity concentration in air. It is an important radiological quantity which can be used for calculation of other quantities like deposition and doses. Under the term *control variables* we understand a set of inputs to the model, which parameterize initial conditions and important physical processes influencing the spreading of the pollutants, e.g., information on the source term (composition and magnitude of the release and its dynamics) and meteorological inputs.

In reality, our knowledge of the release conditions is limited. Typically, meteorological inputs are set using a numerical prediction model and other control variables are set with expert-provided values. This subjective choice can introduce significant errors into the resulting predictions. What is more, the chaotic nature of the atmosphere makes impossible to obtain accurate results using a model of a finite complexity and evaluated with finite computational resources. In other words, inaccurate model initialized with erroneous inputs can not provide reliable predictions. Relying on them can lead in fatal errors in the countermeasures planning. Inherent uncertainty of the problem is not

the only factor making the forecasting in the early phase challenging. There are also strict time constraints caused by the problem dynamics and the urgent need of reliable information. There must be made a compromise between the time spent on evaluation of the dispersion model and its accuracy.

Using data assimilation we can tune uncertain model inputs in a way that the model output fits the available measurements. The number of variables is potentially large but the most important subset can be identified for a specific scenario and a given ADM [ET04, PHP07]. Recent advances in this field have shown great promise in improving model performance through optimal calculation of emission and meteorological inputs by a systematic comparison of observations and modeled concentrations. These improved estimates may in turn be used as inputs to long- and short-range atmospheric dispersion models, resulting in greatly improved efficiency of the countermeasures.

In this chapter, we describe a new data assimilation method based on particle filtering for estimation of important control variables of a parameterized ADM.

9.2 Proposed data assimilation methodology

We propose a new data assimilation methodology based on particle filtering for reduction of uncertainty in atmospheric dispersion modeling during the early phase of a radiation accident. We focus on the parametrized ADMs, where selected control variables are treated as random and we attempt to select their most plausible values in consecutive time steps using available measurements.

9.2.1 State evolution model

Parametrized ADMs can be understood as deterministic functions of the control variables $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$. It means, that all the uncertainty is assumed to be in values of the variables, not in the parametrization itself. Trajectory $\boldsymbol{\theta}_{1:t}$ represents values of control variables of the model up to time t and fully determines its propagation. Vector $\boldsymbol{\theta}_t$ aggregates values of control variables used for model propagation between time instances t and $t+1$. Physics behind the dispersion modeling motivates us to distinguish between two types of control variables, where each type must be treated differently:

Mutable control variables: Values of mutable control variables can—and are expected to—change in respective time steps. Typically, control variables describing meteorological conditions must be treated as mutable in order to correctly simulate stochastic fluctuations of the wind field and other atmospheric phenomena.

Immutable control variables: Values of immutable control variables must the same along the whole state trajectory. Typical representative is the magnitude of release in the case of an instantaneous releases. As the initial magnitude of release affects the deposition, doses and other radiological quantities during the whole

propagation of the plume, its variation would violate the law of activity conservation. Neglecting the radioactive decay, the integral of activity over time and space must be equal to the initial value in all time steps. In context of the classical estimation theory the immutable control variables denote the *stationary parameters*.

We want to estimate the state trajectory $\boldsymbol{\theta}_{1:t}$—from the Bayesian point of view, evaluate the posterior $p(\boldsymbol{\theta}_{1:t}|\mathbf{y}_{1:t})$—in successive time steps $t = 1, 2, \ldots$. Let the state $\boldsymbol{\theta}_t$ be comprised of two parts, the immutable variables $\boldsymbol{\eta}_t$ and the mutable variables $\boldsymbol{\nu}_t$:

$$\boldsymbol{\theta}_t = \begin{bmatrix} \boldsymbol{\eta}_t \\ \boldsymbol{\nu}_t \end{bmatrix}. \tag{9.1}$$

We assume that $\boldsymbol{\eta}_t$ and $\boldsymbol{\nu}_t$ are mutually independent. Since the immutable variables are not allowed to change during the model propagation, we can evolve only the mutable part of $\boldsymbol{\theta}_t$ and the state transition pdf has then the form:

$$\begin{aligned} p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) &= p(\boldsymbol{\eta}_t, \boldsymbol{\nu}_t|\boldsymbol{\eta}_{t-1}, \boldsymbol{\nu}_{t-1}) \\ &= p(\boldsymbol{\eta}_t|\boldsymbol{\eta}_{t-1})p(\boldsymbol{\nu}_t|\boldsymbol{\nu}_{t-1}) \tag{9.2} \\ &= \delta(\boldsymbol{\eta}_t - \boldsymbol{\eta}_1)p(\boldsymbol{\nu}_t|\boldsymbol{\nu}_{t-1}). \tag{9.3} \end{aligned}$$

9.2.2 Observation operator

Measurements are assumed to be normally distributed with covariance matrix $\mathbf{R}_t$ and mutually independent given the state trajectory $\boldsymbol{\theta}_{1:t}$,

$$\mathbf{y}_t \sim \mathcal{N}\left(\mathcal{H}(\boldsymbol{\theta}_{1:t}), \mathbf{R}_t\right), \tag{9.4}$$

where $\mathcal{H}(\cdot)$ is an observation operator. It performs two tasks. Firstly, the observation operator relates measured radiological quantity and an output quantity given by the dispersion model. Secondly, it performs spatial interpolation in the case that the computational and receptor grids are not aligned. Given a radiological quantity evaluated on a computational grid $\mathcal{S}_t$, observation operator yields a vector of measurements $\mathbf{y}_t \in \mathbb{R}^{N_y}$ evaluated in a set of receptor points $\mathcal{S}_t^{\mathrm{R}} = \{\mathbf{s}_{1,t}^{\mathrm{R}}, \ldots, \mathbf{s}_{N_y,t}^{\mathrm{R}}\}$. Generally, the set $\mathcal{S}_t^{\mathrm{R}}$ can vary between time steps. This would be of a particular importance in the case of measurements provided by the moving mobile groups. In the case of a stationary radiation monitoring network we can treat the observation operator as time invariant, i.e., $\mathcal{H}_t = \mathcal{H}$.

9.2.3 Data assimilation algorithm

The posterior pdf is approximated using particle filter,

$$p(\boldsymbol{\theta}_{1:t}|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} w_t^{(i)} \delta\left(\boldsymbol{\theta}_{1:t} - \boldsymbol{\theta}_{1:t}^{(i)}\right), \quad w_t^{(i)} \propto \frac{p\left(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t}\right)}{q\left(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t}\right)}, \tag{9.5}$$

Particles are represented with trajectories $\boldsymbol{\theta}_{1:t}^{(i)}$ parameterizing N simultaneously propagated dispersion models.

If the proposal density

$$q\left(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t}\right) = q\left(\boldsymbol{\eta}_1^{(i)}|\mathbf{y}_{1:t}\right) q\left(\boldsymbol{\nu}_{1:t}^{(i)}|\mathbf{y}_{1:t}\right)$$

is badly chosen, the performance of the filter would be rather poor. Application of sequential evaluation of weights would results in a computationally ineffective scheme, where the computational resources would be wasted on propagation of particles with small weights. Enormous number of particles would be needed to achieve a good performance.

Significant improvements can be achieved by application of the adaptive proposal selection methodology described in Section 7.5.2, where the proposal density $q(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t})$ is re-estimated in respective time steps using the weights $w_t^{(i)}$. Drawbacks of this approach is the fact that the dispersion models and the weights $w_t^{(i)}$ must be always recomputed because of the immutable variables. However, this adaptive procedure guarantees that the trajectories with low weights are discarded and a new population of trajectories is sampled from the regions of the state-space determined by particles with high weights. In other words, the sequential update of the proposal pdf suppresses the effect of sample impoverishment. The resulting algorithm is summarized in Algorithm 7.5.

9.2.4 Evaluation of radiological quantities of interest

Usually, we are not interested just in the estimates of control variables but also in radiological quantities evaluated by the dispersion model. Let the radiological quantity of interest be a continuous function $C(\mathbf{s}, \tau, \boldsymbol{\theta})$ of spatial coordinates $\mathbf{s} = (s_1, s_2, s_3)$; time since the release star τ; and the control variables $\boldsymbol{\theta}$. For computational reasons, $C(\mathbf{s}, \tau, \boldsymbol{\theta})$ is discretized in both spatial and temporal domains. Let $\mathbf{c}_t \in \mathbb{R}^{N_c}$ be a vector aggregating values of $C(\mathbf{s}, \tau, \boldsymbol{\theta})$ evaluated in an ordered set of spatial location $\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_{N_c}\}$ forming a computational grid in time $\tau = \Delta_\tau t$:

$$\mathbf{c}_t = \begin{bmatrix} C(\mathbf{s}_1, \Delta_\tau t, \boldsymbol{\theta}_{1:t}) \\ \vdots \\ C(\mathbf{s}_{N_x}, \Delta_\tau t, \boldsymbol{\theta}_{1:t}) \end{bmatrix}.$$

Here, Δ_τ is the time step length and t is time step index. In the following text, $\mathbf{c}_t \equiv C(\boldsymbol{\theta}_{1:t})$.

Using the particles $\mathbf{c}_t^{(i)} = C(\boldsymbol{\theta}_{1:t}^{(i)})$, $i = 1, \ldots N$, the mean value $\bar{\mathbf{c}}_t$ and the covariance $\boldsymbol{\Sigma}_t^{\mathbf{c}}$ of the radiological quantity evaluated by the dispersion models can be at each time step computed using the posterior (9.5), as follows,

$$\bar{\mathbf{c}}_t = \sum_{i=1}^{N} w_t^{(i)} \mathbf{c}_t^{(i)}, \quad \boldsymbol{\Sigma}_t^{\mathbf{c}} = \sum_{i=1}^{N} w_t^{(i)} \left[\left(\mathbf{c}_t^{(i)} - \bar{\mathbf{c}}_t \right) \left(\mathbf{c}_t^{(i)} - \bar{\mathbf{c}}_t \right)^{\mathrm{T}} \right]. \tag{9.6}$$

9.3 Application to Gaussian puff model

In this section we describe application of the proposed methodology to the assimilation of the Gaussian puff model with the time integrated gamma dose rate measurements.

9.3.1 Parametrization of Gaussian guff model

The idea behind the methodology allows for estimation of an arbitrary set of control variables. However, we restrict to the parametrization of the following physical quantities identified as the most influencing the resulting dose rates: magnitude of instantaneous release Q^i, wind speed $u = |\mathbf{u}|$ and wind direction ϕ. Using location parameters $Q^{i,\star}, u_t^{\star}, \phi_t^{\star}$ and control variables $\boldsymbol{\theta}_t = (\omega_t, \xi_t, \psi_t)^{\mathrm{T}}$, the three physical quantities are parametrized:

1. **Parametrization of magnitude of release Q^i:**

   The overall magnitude of release must be treated as time invariant. It is parameterized using multiplicative immutable control variable $\omega_t \in \mathbb{R}^+$ as follows:

   $$Q = \omega_t Q^{i,\star}. \tag{9.7}$$

2. **Parametrization of wind speed u:**

   In contrast to the overall magnitude of release, the wind direction is assumed to be variable in time. This assumption is in agreement with the stochastic nature of the atmospheric flow. It is parametrized using control variable $\xi_t \in \mathbb{R}$ as follows:

   $$u_t = (1 + 0.1\xi_t)u_t^{\star} + 0.5\xi_t. \tag{9.8}$$

   We can see, that $u_t = u_0$ given $\xi_t = 0$.

3. **Parametrization of wind direction ϕ:**

   Wind direction is also assumed to be variable over time and homogeneous over the whole calculation domain at a time. It is parametrized using control variable $\psi_t \in \mathbb{R}$ as follows:

   $$\phi_t = \phi_t^{\star} + \psi_t. \tag{9.9}$$

From the parameterization is evident that the wind field in assumed to be homogeneous in the whole computational domain at a time. This simplifying assumption is reasonable in the case of a short-range dispersion modeling, where the wind field in not likely to change dramatically in space. Location parameters $u_t^\star$ and $\phi_t^\star$ represent the wind speed and the wind direction given by a meteorological forecast and $Q^{i,\star}$ is the initial estimate of source term based on the safety parameters of a NPP. Similar parameterizations can be constructed for some other physical processes involved, e.g., magnitude of vertical and horizontal dispersion, dry and wet deposition.

The state vector $\boldsymbol{\theta}_t$ is evolved using a transitional pdf

$$p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) = \delta(\omega_t - \omega_1)p(\xi_t|\xi_{t-1})p(\psi_t|\psi_{t-1}), \tag{9.10}$$

and the process is initialized with a prior pdf

$$p(\boldsymbol{\theta}_1) = p(\omega_1)p(\xi_1)p(\psi_1). \tag{9.11}$$

9.3.2 Observation operator

Given the control variables and other inputs, the model (2.16) evaluates activity concentration in air in $Bq\,m^{-3}$. The observation operator converting the concentration to the time integrated gamma dose in Gy is defined by (2.26)–(2.27).

9.3.3 Evaluation of weights

The weights are evaluated using

$$
\begin{aligned}
w_t^{(i)} &\propto \frac{p(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t})}{q(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t})} = \frac{p(\mathbf{y}_t|\boldsymbol{\theta}_t^{(i)})p(\boldsymbol{\theta}_t^{(i)}|\boldsymbol{\theta}_{t-1}^{(i)})p(\boldsymbol{\theta}_{1:t-1}^{(i)}|\mathbf{y}_{1:t-1})}{q(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t})} \\
&\propto \frac{\prod_{j=1}^t p(\mathbf{y}_j|\boldsymbol{\theta}_j^{(i)})p(\boldsymbol{\theta}_j^{(i)}|\boldsymbol{\theta}_{j-1}^{(i)})}{q(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t})}.
\end{aligned} \tag{9.12}
$$

For computational reasons, the weights are evaluated in the logarithmic scale according to Section 7.5.3 and the product in (9.12) becomes a sum of logarithms

$$\ln w_t^{(i)} = \sum_{j=1}^t \left[\ln p(\mathbf{y}_j|\boldsymbol{\theta}_j^{(i)}) + \ln p(\boldsymbol{\theta}_j^{(i)}|\boldsymbol{\theta}_{j-1}^{(i)}) \right] - \ln q(\boldsymbol{\theta}_{1:t}^{(i)}|\mathbf{y}_{1:t}).$$

The normality of the observation model (9.4) determines the likelihood functions $p(\mathbf{y}_j|\boldsymbol{\theta}_j^{(i)})$ to be

$$p(\mathbf{y}_j|\boldsymbol{\theta}_j^{(i)}) = (2\pi)^{-\frac{N_y}{2}}(\det \mathbf{R}_j)^{-\frac{1}{2}} \exp\left[-0.5 \left(\mathbf{y}_j - \mathcal{H}(\boldsymbol{\theta}_j^{(i)}) \right)^{\mathrm{T}} \mathbf{R}_j^{-1} \left(\mathbf{y}_j - \mathcal{H}(\boldsymbol{\theta}_j^{(i)}) \right) \right].$$

Matrix $\mathbf{R}_j$ is the covariance matrix of observations $\mathbf{y}_j$. Since the observations are assumed to be conditionally independent given $\boldsymbol{\theta}_j$, covariance matrix $\mathbf{R}_j$ is diagonal and the observations can be processed sequentially using

$$p(\mathbf{y}_j|\boldsymbol{\theta}_j^{(i)}) \;=\; \prod_{k=1}^{N_y} p(y_{k,j}|\boldsymbol{\theta}_j^{(i)}) = (2\pi)^{-\frac{N_y}{2}} \prod_{k=1}^{N_y} \sigma_{k,j}^{-1} \exp\left[-0.5\left(y_{k,j} - \mathcal{H}_k(\boldsymbol{\theta}_j^{(i)})\right)^2 / \sigma_{k,j}^2\right].$$

Here, $\mathcal{H}_k$ is a reduced observation operator evaluating just time integrated gamma dose rate in location $\mathbf{s}_k^{\mathrm{R}} \in \mathcal{S}^{\mathrm{R}}$, and $\sigma_{k,j}^2 = \mathbf{R}_j[k,k]$, $k = 1, \ldots, N_y$, is the kth diagonal element of $\mathbf{R}_j$.

9.3.4 Adaptive selection of proposal density

We apply the adaptive proposal selection procedure described in Section 7.5.2. Let the proposal density be normally distributed. Using conditional independence of control variables we can write:

$$\begin{aligned}
q(\boldsymbol{\theta}_{1:t}|\mathbf{y}_{1:t}) &= q(\omega_1, \xi_1, \ldots, \xi_t, \psi_1, \ldots, \psi_t|\mathbf{y}_{1:t}) \\
&= \mathcal{N}(\omega_1; \bar{\omega}_1, \Sigma_t^\omega)\mathcal{N}(\boldsymbol{\xi}_t; \bar{\boldsymbol{\xi}}_t, \boldsymbol{\Sigma}_t^{\boldsymbol{\xi}})\mathcal{N}(\boldsymbol{\psi}_t; \bar{\boldsymbol{\psi}}_t, \boldsymbol{\Sigma}_t^{\boldsymbol{\psi}}).
\end{aligned} \tag{9.13}$$

Vectors $\boldsymbol{\xi}_t = (\xi_1, \ldots, \xi_t)^{\mathrm{T}}$ and $\boldsymbol{\psi}_t = (\psi_1, \ldots, \psi_t)^{\mathrm{T}}$ aggregate wind speed and wind direction in time steps $1, \ldots, t$. Their dimensions thus increase in time.

At each time step, moments of Gaussian pdfs in (9.13), mean values $\bar{\omega}_t \in \mathbb{R}$, $\bar{\boldsymbol{\xi}}_t \in \mathbb{R}^t$, $\bar{\boldsymbol{\psi}}_t \in \mathbb{R}^t$ and corresponding variance Σ_t^ω and diagonal covariance matrices $\boldsymbol{\Sigma}_t^{\boldsymbol{\xi}}$, $\boldsymbol{\Sigma}_t^{\boldsymbol{\psi}}$, can be estimated independently from the weights and the particles $\boldsymbol{\theta}_{1:t}^{(i)} = (\omega_1^{(i)}, \boldsymbol{\xi}_t^{(i)}, \boldsymbol{\psi}_t^{(i)})^{\mathrm{T}}$:

$$\begin{aligned}
\bar{\omega}_1 &= \sum_{i=1}^{N} w_t^{(i)} \omega_1^{(i)}, \quad \Sigma_t^\omega = \sum_{i=1}^{N} w_t^{(i)}\left(\omega_1^{(i)} - \bar{\omega}_1\right)^2, \\
\bar{\boldsymbol{\xi}}_t &= \sum_{i=1}^{N} w_t^{(i)} \boldsymbol{\xi}_t^{(i)}, \quad \boldsymbol{\Sigma}_t^{\boldsymbol{\xi}}[j,j] = \sum_{i=1}^{N} w_t^{(i)}\left(\boldsymbol{\xi}_t^{(i)}[j] - \bar{\boldsymbol{\xi}}_t[j]\right)^2, \quad j = 1, \ldots, t, \\
\bar{\boldsymbol{\psi}}_t &= \sum_{i=1}^{N} w_t^{(i)} \boldsymbol{\psi}_t^{(i)}, \quad \boldsymbol{\Sigma}_t^{\boldsymbol{\psi}}[j,j] = \sum_{i=1}^{N} w_t^{(i)}\left(\boldsymbol{\psi}_t^{(i)}[j] - \bar{\boldsymbol{\psi}}_t[j]\right)^2, \quad j = 1, \ldots, t.
\end{aligned}$$

9.4 Numerical experiment

In numerical experiment we assume an instantaneous release of radionuclide ^{41}Ar with half-life of decay 109.34 minutes. Radionuclide ^{41}Ar was chosen for two reasons: Firstly, since ^{41}Ar is a noble gas, there is no deposition and consequently no groundshine. We

need to calculate only the gamma dose rate from cloudshine. Secondly, according to Tables of Radioactive Isotopes [BFS86], the radionuclide ^{41}Ar emits gamma radiation on energy level $1293.57 keV$ with branching ration 99.1%. Generally, for calculation of the gamma dose rate we need to assume all the energy levels and their branching ratios specific to the given radionuclide. In the case of ^{41}Ar we can neglect the other energy levels within the remaining branching ratio 0.9% without any significant loss of accuracy. Both these facts substantially simplifies gamma dose rate calculations and makes the experiment more transparent. Since we simulate a release of a noble gas, the deposition is not calculated here.

Data assimilation is performed in time steps $t = 1, \ldots, 18$. Duration of time step Δ_τ is set to 10 minutes. This step length was chosen because we assume that the radiation monitoring network provides measurements of the time integrated gamma dose rate in 10-minute intervals [DNTW09].

9.4.1   Computational and observational grids
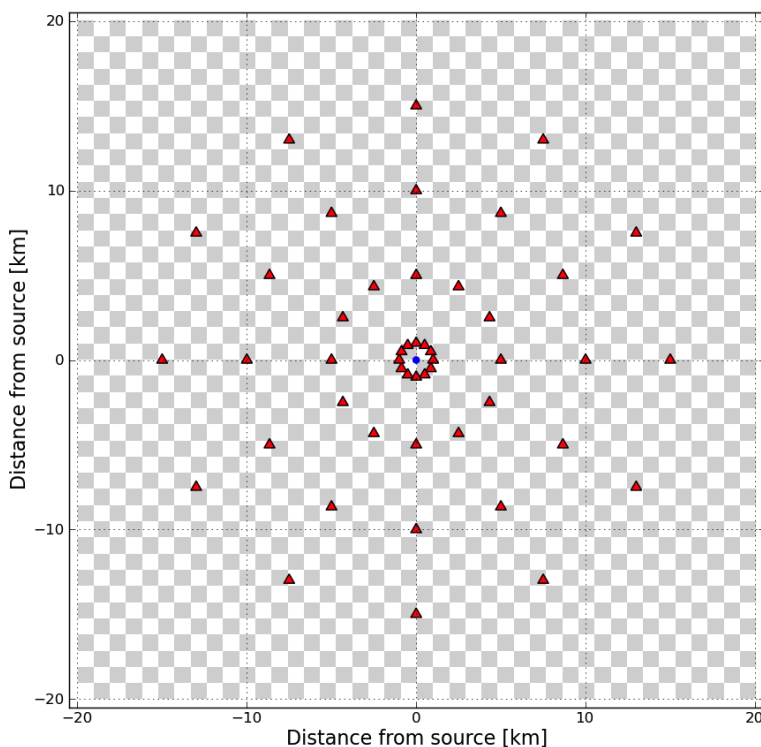
Figure 9.1: Illustration of computational grid and monitoring network.

The computational domain is delimited with a square centered at the location of the Czech nuclear power plant Temelin and with the side length 20km. The area is regularly covered with a rectangular grid with the grid step 1km. The total number of grids points is $41 \times 41 = 1681$. This area includes the emergency planning zone delimited with a

COORDINATES OF RECEPTORS								
Rec. no.	x $[m]$	y $[m]$	Rec. no.	x $[m]$	y $[m]$	Rec. no	x $[m]$	y $[m]$
1	0	1000	17	4330	-2500	33	-8660	-5000
2	500	866	18	2500	-4330	34	-10000	0
3	866	500	19	0	-5000	35	-8660	5000
4	1000	0	20	-2500	-4330	36	-5000	8660
5	866	-500	21	-4330	-2500	37	0	15000
6	500	-866	22	-5000	0	38	7500	12990
7	0	-1000	23	-4330	2500	39	12990	7500
8	-500	-866	24	-2500	4330	40	15000	0
9	-866	-500	25	0	10000	41	12990	-7500
10	-1000	0	26	5000	8660	42	7500	-12990
11	-866	500	27	8660	5000	43	0	-15000
12	-500	866	28	10000	0	44	-7500	-12990
13	0	5000	29	8660	-5000	45	-12990	-7500
14	2500	4330	30	5000	-8660	46	-15000	0
15	4330	2500	31	0	-10000	47	-12990	7500
16	5000	0	32	-5000	-8660	48	-7500	12990

Table 9.1: Coordinates of the receptors comprising radiation monitoring network in the numerical example.

circle of radius 13km centered at the power plant. The zone delimits potential accident site, where public is expected to be put in risk if exposed to the radioactive plume.

The measurements are assumed to come from a stationary radiation monitoring network. Let the network comprises of 48 receptors placed in the four circular bands of radii 1km, 5km, 10km, and 15km. In reality, the gamma dose rate receptors comprising the monitoring network would be placed in settled areas, e.g., in towns and villages within the zone. However, since we investigate the properties of the algorithm, the configuration of receptor points is fully justifiable for our purposes. The receptors closest to the center denote the receptors placed in the area of the power plant, the tele-dosimetric system (TDS) on fence of the power plant. Schematic illustration of the monitoring network and the computational grid is in Figure 9.1. The computational points are represented with edges of the "chess board" and the receptors are denoted with the red triangles. Coordinates of the receptors are in Table 9.1.

9.4.2 Simulation of observations

Numerical experiment is performed as a twin experiment, where the measurements are generated using a twin model and perturbed with a random noise. The twin model is a dispersion model initialized with a set of inputs defining unknown conditions of the "real" release. Convergence of the dispersion model initialized with a set of nominal

PARAMETERS OF NOMINAL MODEL AND TWIN MODEL									
		Immutable variables			Mutable variables				
		Magnitude of release			Wind speed			Wind direction	
Hour	t	$Q^{i,\star}$	ω_1^{r}	$Q^{i,\mathrm{r}}$	$u_t^{\star}$	ξ_t^{r}	u_t^{r}	$\phi_t^{\star}$	ψ_t^{r}	ϕ_t^{r}
1	1	5.0E+14	5.0	2.5E+15	2.0	0.71	2.5	270.0	0.0	270.0
1	2				2.0	0.71	2.5	270.0	10.0	280.0
1	3				2.0	0.71	2.5	270.0	20.0	290.0
1	4				2.0	0.71	2.5	270.0	30.0	300.0
1	5				2.0	0.71	2.5	270.0	40.0	310.0
1	6				2.0	0.71	2.5	270.0	50.0	320.0
2	7				2.0	0.71	2.5	280.0	50.0	330.0
2	8				2.0	0.71	2.5	280.0	60.0	340.0
2	9				2.0	0.71	2.5	280.0	70.0	350.0
2	10				2.0	0.71	2.5	280.0	60.0	340.0
2	11				2.0	0.71	2.5	280.0	50.0	330.0
2	12				2.0	0.71	2.5	280.0	40.0	320.0
3	13				2.0	0.71	2.5	290.0	20.0	310.0
3	14				2.0	0.71	2.5	290.0	10.0	300.0
3	15				2.0	0.71	2.5	290.0	0.0	290.0
3	16				2.0	0.71	2.5	290.0	-10.0	280.0
3	17				2.0	0.71	2.5	290.0	-20.0	270.0
3	18				2.0	0.71	2.5	290.0	-30.0	260.0

Table 9.2: Parameters of nominal and twin model. Nominal values $Q^{i,\star}, u_t^{\star}, \phi_t^{\star}$ of physical quantities treated as uncertain. "Real" values $Q^{i,\mathrm{r}}, u_t^{\mathrm{r}}, \phi_t^{\mathrm{r}}$ of the quantities used for simulation of measurements. Sought values of variables $\boldsymbol{\theta}_t^{\mathrm{r}} = (\omega_1^{\mathrm{r}}, \xi_t^{\mathrm{r}}, \psi_t^{\mathrm{r}})$ transforming the nominal values into the real values using parameterizations (9.7)–(9.9).

inputs to the twin model is then assessed.

The nominal values of the wind speed and the wind direction are hourly meteorological forecast from a numerical weather prediction model. The true values of the wind direction are assumed to change every 10 minutes. The true wind speed is constant and systematically higher than the nominal values. Summary of setting of the nominal and the twin model is in Table 9.2:

1. Nominal values $Q^{i,\star}, u_t^{\star}, \phi_t^{\star}$ given by an expert and the meteorological forecast (locations parameters).

2. "Real" values of physical quantities $Q^{\mathrm{i,r}}, u_t^{\mathrm{r}}, \phi_t^{\mathrm{r}}$ used for simulation of measurements.

3. Sought values of control variables $\boldsymbol{\theta}_t^{\mathrm{r}} = (\omega_1^{\mathrm{r}}, \xi_t^{\mathrm{r}}, \psi_t^{\mathrm{r}})$ transforming the nominal values into the real values using parameterizations (9.7)–(9.9).
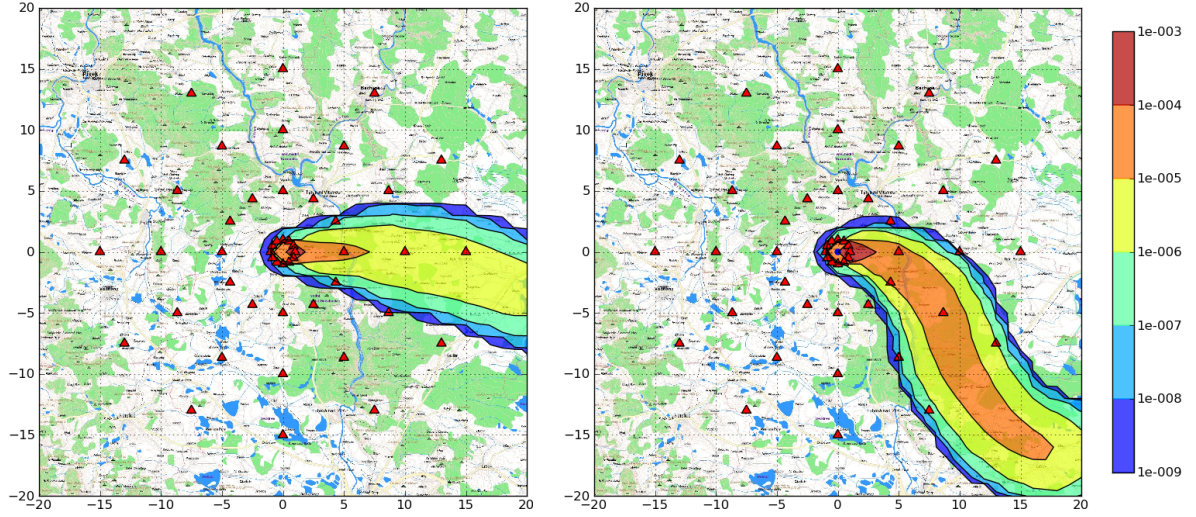
Figure 9.2: Time integrated gamma dose evaluated for the first three hours of the release with the nominal model (left) and the twin model (right).

In the experiment we expect convergence of the estimated control variables $\bar{\boldsymbol{\theta}}_t$ (9.6) to $\boldsymbol{\theta}_t^{\mathrm{r}}$. In Figure 9.2 we see the time integrated dose evaluated for the first three hours of the release with the nominal model (left) and the twin model (right). Radiation monitoring network $\mathcal{S}^{\mathrm{R}}$ is denoted with the red triangles. Measurements $\mathbf{y}_t$ are sampled during the twin model propagation in 10-minute intervals according to $\mathbf{y}_t \sim \mathcal{N}(\mathcal{H}(\boldsymbol{\theta}_{1:t}), \mathbf{R}_t)$, where the observation operation $\mathcal{H}$ is given by (2.26)–(2.27). Covariance matrix $\mathbf{R}_t$ is a diagonal matrix, where the standard deviations of elements of $\mathbf{y}_t$ are linearly proportional to measurements:

$$\mathbf{R}_t[j, j] = (0.1\mathbf{y}_t[j] + 1.0\mathrm{E} - 20)^2 . \tag{9.14}$$

Here, $\mathbf{R}_t[j, j]$ and $\mathbf{y}_t[j]$ denote the jth diagonal element of $\mathbf{R}_t$ and the jth element of $\mathbf{y}_t$, respectively. Particles are initialized with values of control variables sampled from prior pdf (9.11), where

$$\begin{aligned}
p(\omega_1) &= log\mathcal{N}(0.5, 0.25), \\
p(\xi_1) &= \mathcal{U}(-2, 2), \\
p(\psi_1) &= \mathcal{U}(-22.5, 22.5).
\end{aligned}$$

Control variables are evolved using transitional pdf (9.10), where

$$\begin{aligned}
p(\xi_t|\xi_{t-1}) &= \mathcal{N}(\xi_{t-1}, \sigma_\xi^2), \quad \sigma_\xi = 0.4, \\
p(\psi_t|\psi_{t-1}) &= \mathcal{N}(\psi_{t-1}, \sigma_\psi^2), \quad \sigma_\psi = 2.5.
\end{aligned}$$

9.4.3 Results

We run the assimilation algorithm with $N = 3000$ particles for 18 steps covering the first three hours of the release.

In Figure 9.3, the nominal model, the twin model, and the assimilated model are compared. The results are visualized in terms of $\mathcal{D}^c$ integrated from time step 0 up to time steps 6, 12, and 18, respectively. We see that the nominal model without the data assimilation would predict doses smaller in magnitude and also the affected areas would be misspecified. We can observe that the expected values of $\mathcal{D}^c$ up to time step $t = 12$ well approximate $\mathcal{D}^c$ evaluated by the twin model. During the last six assimilation steps we observe a misfit of the wind direction. This is due to the lack of measurements in the area where the puffs (particles) were during the third hour of their propagation, see Figure 9.3 (bottom-right).
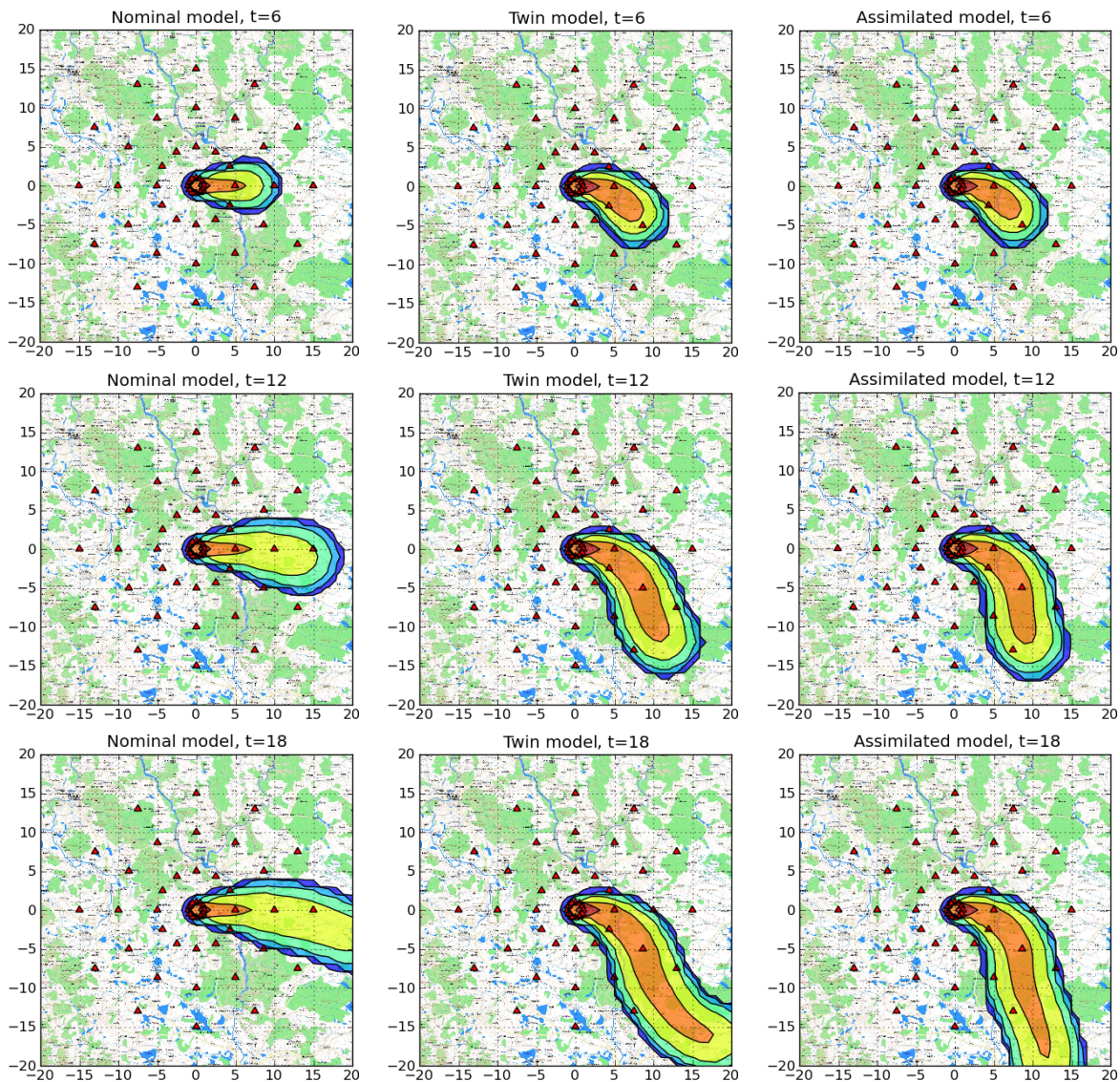
Figure 9.3: Comparison of the nominal model, the twin model and the assimilated model in time steps $t = 6, 12, 18$. The misfit of the wind direction in the last hour is due to the lack of measurements in the area.

In Figure 9.4, estimates $\bar{Q}^i$, $\bar{u}_t$, $\bar{\phi}_t$ (9.6) evaluated using nominal values $Q^{i,\star}$, $u_t^\star$, $\phi_t^\star$, weights $w_t^{(i)}$ and the states $\boldsymbol{\theta}_t^{(i)}$ of the re-computed state trajectories $\boldsymbol{\theta}_{1:t}^{(i)}$ in each time step are visualized. We see how the estimated values of magnitude of release, wind speed and wind direction (red lines) approach the values used for simulation of measurements (green lines). Values of the physical quantities used for propagation of the puffs are denoted with the blue dots. Only those particles with nonzero weights are visualized. During the first five steps, the magnitude of release is correctly recognized and stays tuned for the remaining time steps. Convergence to the correct wind speed and wind directions is more rapid. The lack of measured information during the last six time steps cause, that the weights are approximately equal and the variances of the estimates increase. With the non-informative weights, the algorithm does not have enough information to correctly estimate the wind direction.

In Figure 9.5 we see how the expected values $\bar{\boldsymbol{\theta}}_t$ of control variables (red lines) approach the true values of control variables $\boldsymbol{\theta}_t^r$ (green lines) used for simulation of measurements. The gray bands denote the maximal and minimal values of particles in each time step.

Time evolution of time integrated doses $\mathcal{D}^c$ at selected receptor locations is visualized in Figure 9.6. There is a good agreement between the doses generated by the twin model and the assimilated model. In the case of receptors 40 and 41 we observe a disagreement due to misspecification of the wind direction caused by the lack of monitoring data.

9.5 Summary

This chapter has developed and demonstrated a new methodology for data assimilation of gamma dose rate measurements with modeled activity concentration air. We propose to use sequential Monte Carlo methods for estimation the most important parameters of a dispersion model and thus improve the correspondence of the model output with the measurements. The methodology is based on simultaneous propagation of multiple dispersion models initialized with different inputs. The resulting algorithm seeks for the most plausible values of these parameters (here referred as control variables) using particle filtering with adaptive selection of the proposal density. Adaptive proposal selection makes the algorithm more efficient, because the trajectories of particles are sampled from a promising subspace of the full state-space. The presented form of the proposal density is a rather simple choice and more elaborated approaches can be constructed, e.g., a parametrized proposal in the form of a Gaussian mixture.

Introduced Bayesian methodology has very interesting properties suitable for the solved scenario. The probabilistic aspect of the solution optimally combines a likely answer with uncertainties of the available data. Since the uncertainty is accounted for, the physical parameters of the model are the best parameters possible, not in the sense of exact match, but because they lead to the best representation of the true system, given the assumptions that were used to build the model. The corrected parameters may in turn be used as input to long- and short-range atmospheric dispersion models,
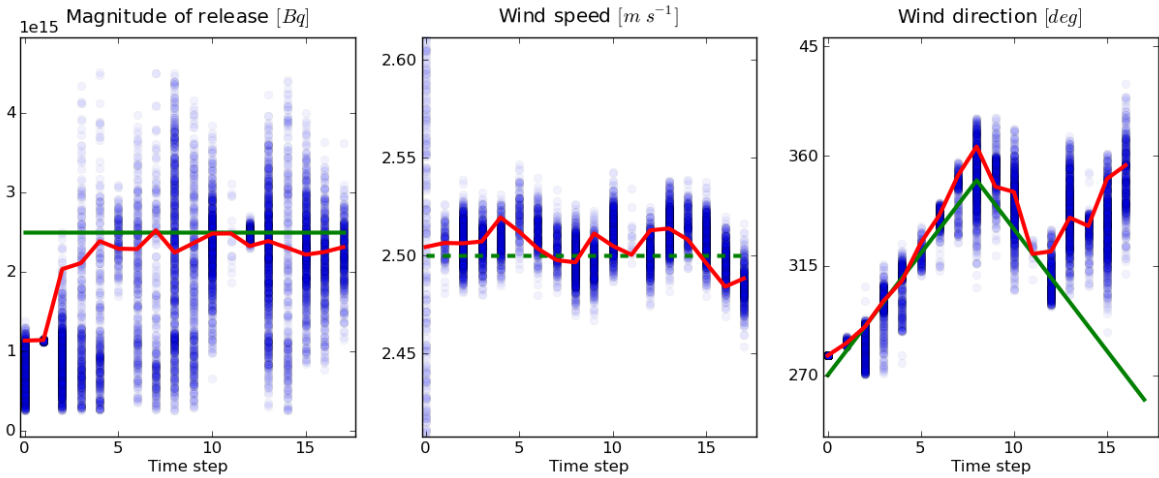
Figure 9.4: Estimated values of magnitude of release, wind speed and wind direction. Green lines: values used for simulation of measurements; red lines: estimated values of physical quantities; blue dots: values of particles with nonzero weights.
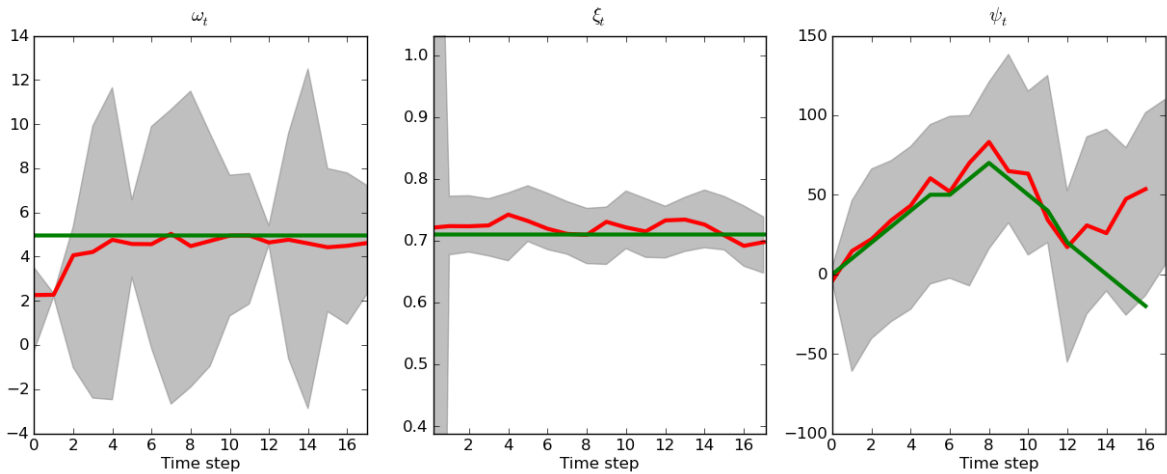
Figure 9.5: Estimated values of magnitude of release, wind speed and wind direction. Green lines: values used for simulation of measurements; red lines: estimated values; gray areas: regions between minimum and maximum values of the particles.
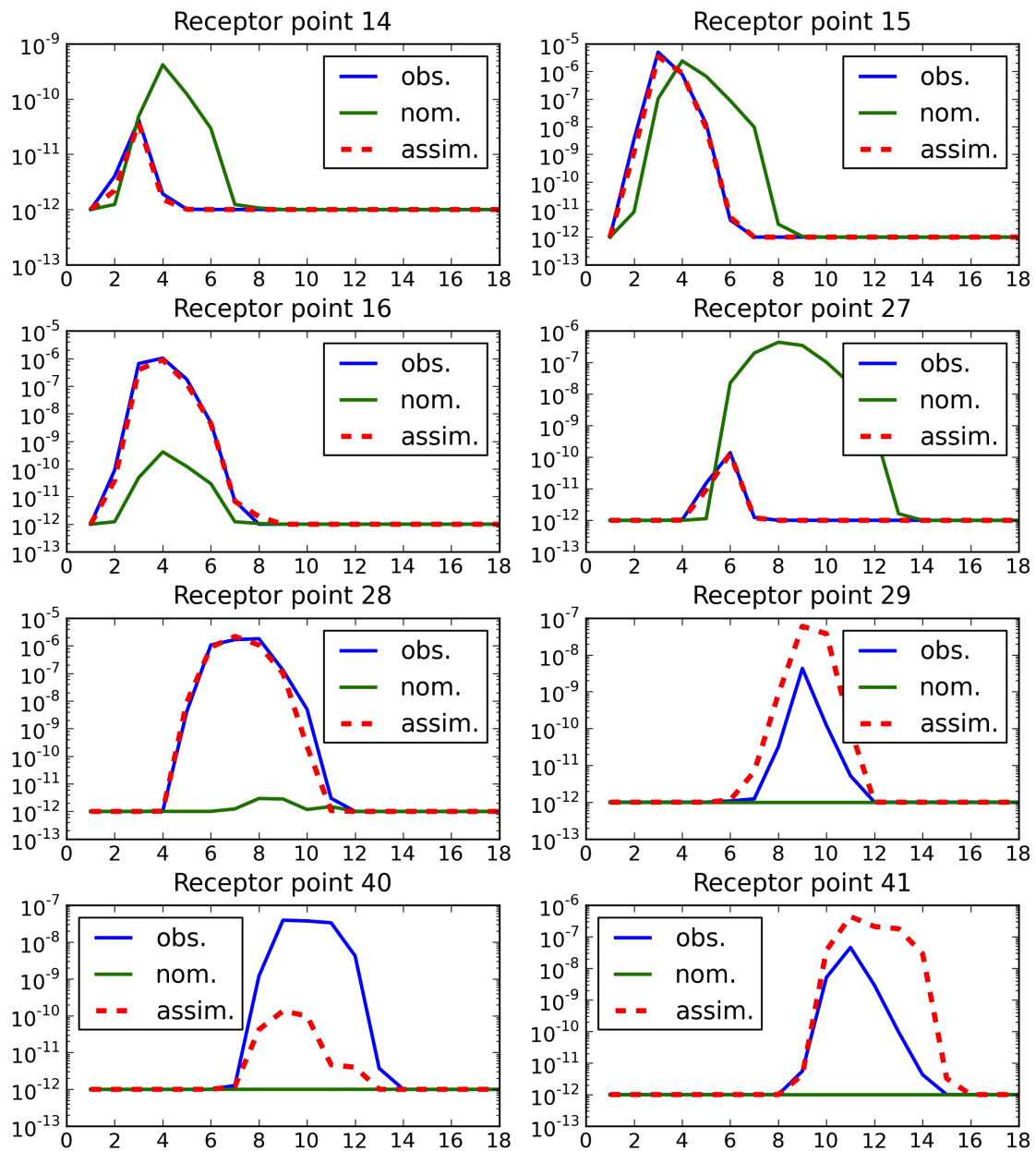
Figure 9.6: Time evolution of time integrated doses $\mathcal{D}^c$ given by the twin model (blue solid lines), the nominal model (green solid lines), and the assimilated model (red dashed lines) at receptor locations 14, 15, 16, 27, 28, 29, 40 and 41 (see Table 9.1).

resulting in greatly improved dose rate assessment.

The algorithm was demonstrated on estimation of magnitude of release, wind speed, and wind direction of a Gaussian puff model. Since the measuring of concentration itself is not technically feasible, nonlinear observation operator for transformation of the activity concentration in air into the time integrated gamma dose rate was implemented. Selected control variables were successfully estimated and the assimilated dose rates were close to the dose rate from the twin model using a sparse observational grid. The algorithm performed well in a meandering wind field, which is particularly important under low-wind conditions. The extension of the algorithm to account for different physical effects is straightforward, however, we have to consider computational demands regarding intensive sampling during the particle filtering assimilation procedure.

Chapter 10

Application of Bayesian DA in the Late Phase

10.1 Problem statement

Under the term *late phase* we understand the time period after the release of radioactive material when the atmospheric transport (and subsequently the deposition) of radionuclides has finished. During the late phase, there is no more irradiation from the cloud but the deposited radioactive material causes external irradiation from groundshine and internal irradiation from inhalation of the re-suspended material. What is more, radionuclides migrate through the root system and foliage of plants into their edible parts and thus can cause internal irradiation of people and livestock when eaten. This phase extends over a period of several weeks or many years, depending on the magnitude and type of initially deposited radionuclides.

From the point of view of radiation protection, the attention is focused on the long-term monitoring of the radiation levels and modeling of its further transport towards human body through the food chain. We are concerned with the deposition modeling. Determination of the spatial and temporal distribution of radionuclides on terrain and the rate of radionuclides removal is crucial for planning of the late phase countermeasures. These regard agriculture, foodstuffs production and water-resources management [PFM93].

In this chapter we propose a new data assimilation method based on the MPF framework developed in Chapter 8 for joint estimation of a spatially distributed radiological quantity and a set of parameters concerning the process of radioactivity removal.

10.2 Data assimilation scenario

During the late phase, the main aim of monitoring in to obtain a comprehensive picture of contamination of the environment [GWW+04]. The most significant monitored radiological quantities are:

- The external dose rate from deposition (*groundshine*),

- The spectrum of deposited radionuclides and the relative deposition on different surfaces,

- The contamination of reference foodstuffs.

We focus on the groundshine dose modeling, specifically, on determination of the spatial distribution of radionuclides and the groundshine dose time evolution. We approach the problem using the data assimilation, where the groundshine dose measurements are related to the predictions of deposition via groundshine dose evolution model. Here, the highest uncertainty consists (i) in the initial displacement of the deposition and (ii) in the rate of groundshine dose mitigation due to radionuclides removal and migration processes. The initial deposition displacement is fully determined by the plume depletion during the early phase. This means that the uncertainty regarding aerial pollution propagation must be considered. What regards the issue of groundshine mitigation, two dominant processes—radioactive decay and environmental removal—must be modeled. The latter is given by parametrized formula (2.30). Generally, the parameterization of environmental removal depends on many factors, including the place of model application. A reasonable approach is to treat the parameters as random variables and attempt for their estimation using available radiological measurements.

In this work, the following objectives of the groundshine dose modeling in the late phase are addressed:

1. The estimation of initial deposition displacement using available measurements referring to the beginning of the late phase.

2. The estimation of radiation levels in the contaminated areas and the prediction of its time evolution.

3. The estimation of the speed of radionuclides removal.

The refined estimates of the spatio-temporal distribution of radioactivity and its time evolution can in turn be used for long-term predictions.

Data assimilation in the late phase has its own specifics. The key differences compared to the data assimilation in the early phase are as follows:

High state dimension: The state vector contains the values of deposition in an ordered set of spatial points. To achieve a good spatial resolution, we want to estimate the deposition values on a dense computational grid. This means that we have to employ an estimation methodology suitable for large-scale problems, e.g., ensemble filtering.

Time constraints: With respect to the dynamics of the radionuclides transport in the late phase and its duration, the time constrains are not so strict as in the early phase. Particularly in the case of a retrospective analysis, the time constraints are of a minor interest.

Higher amount of available data: We can assume that the number of available measurements is higher than in the early phase. Monitoring strategies in the late phase combine various monitoring systems in an operational way to meet the requirements of off-site emergency management. We can assume that the observations come from the airborne gamma spectrometry. It is a powerful tool capable of rapid mapping of contamination levels in a broader scale, which was demonstrated during the Chernobyl accident [GWW$^+$04].

With respect to the above specific properties, the data assimilation strategy must be appropriately chosen. We propose to use the marginalized particle filtering framework developed in Chapter 8. Here, the uncertain parameters of the environmental removal are estimated using the particle filter and the spatio-temporal distribution of the deposition given the parameters is estimated using the ensemble square root filter. What results is a hybrid filter, where N weighted ensemble filters are simultaneously run. Overall rate of environmental removal in the considered area is estimated using the particle filter and the ensemble filter account for local characteristics.

For numerical reasons, the calculations are performed in terms of deposition. Transformation of the deposition into the groundshine dose (2.28) can be simply done using the dose rate conversion factor.

10.3 Proposed data assimilation methodology

Let $\mathbf{d}_t \in \mathbb{R}^{N_\mathbf{d}}$ be a vector of deposition values in a set of computational points $\mathcal{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_{N_\mathbf{d}}\}$. Let $\boldsymbol{\theta}_t \in \mathbb{R}^{N_\theta}$ aggregate radionuclides removal rate parameters and parameters influencing magnitude and structure of model error in the ensemble filter. We use marginalized filtering framework introduced in Chapter 8, where we substitute $\mathbf{x}_t^c \equiv \mathbf{d}_t$, $\mathbf{x}_t^p \equiv \boldsymbol{\theta}_t$ in factorization (7.58), yielding the state vector

$$\mathbf{x}_t = \left[\begin{array}{c} \mathbf{d}_t \\ \boldsymbol{\theta}_t \end{array} \right]. \tag{10.1}$$

Parameters $\boldsymbol{\theta}_t$ are estimated using the particle filter and the deposition field $\mathbf{d}_t$ is estimated using the conditional ensemble square root filter. The resulting posterior is of the form given by (7.60).

10.3.1 State evolution model

Model $\mathcal{M}(\mathbf{d}, \boldsymbol{\theta}, \boldsymbol{\epsilon})$ describing the evolution of deposition $\mathbf{d}$ is a non-linear function of parameters $\boldsymbol{\theta}$ and zero-mean mutually independent random noise $\boldsymbol{\epsilon}$,

$$\mathbf{d}_t = \mathcal{M}(\mathbf{d}_{t-1}, \boldsymbol{\theta}_t, \boldsymbol{\epsilon}_t).$$

From (2.29) follows that given particular values of $\boldsymbol{\theta}_t$, the model becomes linear in $\mathbf{d}_t$ and we can construct a linear state evolution model for fixed $\boldsymbol{\theta}_{t-1}$ and $\boldsymbol{\theta}_t$ represented with a matrix $\mathbf{M}$,

$$\mathbf{d}_t = \mathbf{M}(\boldsymbol{\theta}_t)\mathbf{d}_{t-1} + \boldsymbol{\epsilon}_t,$$
$$\mathbf{M}_t = \frac{f_{\mathrm{R}}(t)f_{\mathrm{E}}(t,\boldsymbol{\theta}_t)}{f_{\mathrm{R}}(t-1)f_{\mathrm{E}}(t-1,\boldsymbol{\theta}_{t-1})}\mathbf{I}. \qquad (10.2)$$

Determination of spatio-temporal distribution of deposition and the speed of its removal can be interpreted as estimation of the augmented state $\mathbf{x}_t$ comprised of the vector $\mathbf{d}_t$ of the deposition values on a grid and the vector $\boldsymbol{\theta}_t$ of variables parameterizing the removal speed. Since we attempt for estimation of the model error, its parameterization is also included into $\boldsymbol{\theta}$. The state is evolved using transitional pdf

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

where the time evolution of $\boldsymbol{\theta}_t$ is assumed to be dependent just on its previous value,
$$p(\boldsymbol{\theta}_t|\mathbf{d}_{t-1},\boldsymbol{\theta}_{t-1}) = p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}).$$
The time evolution of $\mathbf{d}_t$ is given by a Gaussian pdf

$$\begin{aligned} p(\mathbf{d}_t|\mathbf{d}_{t-1},\boldsymbol{\theta}_t,\boldsymbol{\theta}_{t-1}) &= \mathcal{N}\left(\bar{\mathbf{d}}_{t|t-1},\mathbf{P}_{t|t-1}\right), \qquad (10.3) \\ \bar{\mathbf{d}}_{t|t-1} &= \mathbf{M}_t\bar{\mathbf{d}}_{t-1|t-1}, \qquad (10.4) \end{aligned}$$

where $\bar{\mathbf{d}}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$ are predictive statistics evaluated by the time update step of the ensemble filter. Since the parameterization of model error is included in $\boldsymbol{\theta}_t$, the predictive error covariance matrix $\mathbf{P}_{t|t-1}$ is a function of $\boldsymbol{\theta}_t$.

10.3.2 Observation operator

Groundshine dose measurements given on an observational grid $\mathcal{S}_t^{\mathrm{R}} = \{\mathbf{s}_{1,t}^{\mathrm{R}},\ldots,\mathbf{s}_{N_y,t}^{\mathrm{R}}\}$ are aggregated in a vector $\mathbf{y}_t \in \mathbb{R}^{N_{\mathbf{y}}}$. The measurements assumed to be normally distributed with covariance matrix $\mathbf{R}_t$ and mutually independent given the state $\mathbf{x}_t$,

$$p(\mathbf{y}_t|\mathbf{x}_t) = \mathcal{N}\left(\mathbf{H}\mathbf{d}_{t|t-1},\mathbf{R}_t\right). \qquad (10.5)$$

Covariance $\mathbf{R}_t$ describes the instrumental error of the measuring device and the linear observation operator $\mathbf{H}$ relates the deposition with the groundshine dose. In the case that the observational and the computational grids are not aligned, the operator also performs a spatial interpolation using the bilinear interpolation.

10.3.3 Evaluation of weights

The normality of the observation model (10.5) implies the likelihood function used for evaluation of weights in the particle filter has the form

$$p(\mathbf{y}_t|\mathbf{y}_{1:t-1}, \mathbf{x}_t^{(i)}) \propto \det\left(\mathbf{Z}_t(\boldsymbol{\theta}_t^{(i)})\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\mathbf{v}_t^{(i)}\right)^{\mathrm{T}}\mathbf{Z}_t^{-1}(\boldsymbol{\theta}_t^{(i)})\mathbf{v}_t^{(i)}\right], \qquad (10.6)$$

where $\mathbf{v}_t^{(i)} = \mathbf{y}_t - \mathbf{H}\bar{\mathbf{d}}_{t|t-1}^{(i)}$, $\mathbf{Z}_t^{(i)} = \mathbf{H}\mathbf{P}_{t|t-1}^{(i)}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_t$. Note, that the covariance matrices $\mathbf{Z}_t^{(i)}$ are explicitly conditioned on vectors $\boldsymbol{\theta}_t^{(i)}$ corresponding to particles.

Since the covariance square roots are stored in EnSRFs, it would be beneficial to avoid evaluation of the full covariance matrices. This can be achieved using the results of Appendix B. Let $\mathbf{F}_t^{(i)}$ and $\mathbf{S}_{t|t-1}^{(i)}$ be square roots of covariance matrices $\mathbf{Z}_t^{(i)}$ and $\mathbf{P}_{t|t-1}^{(i)}$, respectively:

$$\mathbf{Z}_t^{(i)} = \mathbf{F}_t^{(i)}\left(\mathbf{F}_t^{(i)}\right)^{\mathrm{T}}, \quad \mathbf{P}_{t|t-1}^{(i)} = \mathbf{S}_{t|t-1}^{(i)}\left(\mathbf{S}_{t|t-1}^{(i)}\right)^{\mathrm{T}}.$$

Using the result of Appendix B and substituting $\mathbf{B} = \mathbf{S}_{t|t-1}^{(i)}$, $\mathbf{C} = \mathbf{H}_t$, and $\mathbf{D} = \mathbf{R}_t$, it follows that

$$\left[\begin{array}{c}\left(\mathbf{F}_t^{(i)}\right)^{\mathrm{T}} \\ \mathbf{0}\end{array}\right] = \mathbf{T}\left[\begin{array}{c}\left(\mathbf{S}_{t|t-1}^{(i)}\right)^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}} \\ \mathbf{R}^{\frac{\mathrm{T}}{2}}\end{array}\right],$$

where $\mathbf{T} \in \mathbb{R}^{N_\mathbf{y} \times N_\mathbf{y}}$. Having upper triangular Cholesky factor $\mathbf{F}_t^{(i)}$ of $\mathbf{Z}_t^{(i)}$, we can use (7.57) for computing the particle weights in logarithmic scale.

10.4 Numerical experiment

We focus on modeling of the groundshine in the zone of emergency planning during the first two years after a severe reactor accident. In our numerical experiment we assume a deposition of radionuclide ^{134}Cs with half-life of decay $T_{1/2} = 2.0648$ years.

We want to perform a retrospective analysis using historical measurements sampled in one-month intervals for the time period of two years. The augmented state vector is of the from (10.1), where $\mathbf{d}_t$ accounts for spatio-temporal distribution of the deposition and $\boldsymbol{\theta}_t$ describes the rate of radionuclides removal and the magnitude of model error. For given $\boldsymbol{\theta}_t$, deposition $\mathbf{d}_t$ is estimated with the ensemble square root Kalman filter with multiplicative inflation of covariance (7.35) and evolved using the model (10.3).

Since the assumed time period is much shorter than the expected half-life of slow component of the environmental removal, we do not treat T^{s} in (2.30) as random and set it with a fixed value. We estimate just the half-life of the fast component T^{f} and its fraction d^{f}. Fraction of the slow component d^{s} is fully determined by the binding condition $d^{\mathrm{f}} + d^{\mathrm{s}} = 1$. Parameters $\boldsymbol{\theta}_t = (d_t^{\mathrm{f}}, T_t^{\mathrm{f}}, \Delta_t)^{\mathrm{T}}$ are evolved using transitional pdf

$$p(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) = p(d_t^{\mathrm{f}}|d_{t-1}^{\mathrm{f}})p(T^{\mathrm{f}}_t|T^{\mathrm{f}}_{t-1})p(\Delta_t|\Delta_{t-1}), \qquad (10.7)$$

where the evolution of its elements is modeled by random walk processes

$$p(d_t^f|d_{t-1}^f) = \mathcal{N}(d_{t-1}^f, \sigma_d^2), \quad \sigma_d = 0.05 d_{t-1}^f + 0.01, \tag{10.8}$$
$$p(T_t^f|T_{t-1}^f) = \mathcal{N}(T_{t-1}^f, \sigma_T^2), \quad \sigma_T = 0.05 T_{t-1}^f + 0.01,$$
$$p(\Delta_t|\Delta_{t-1}) = t\mathcal{N}(\Delta_{t-1}, \sigma_\Delta^2, [1, \infty]), \quad \sigma_\Delta = 0.05\Delta_{t-1} + 0.01. \tag{10.9}$$

Proposal density of the particle filter is chosen as (10.7) yielding the formula for recursive evaluation of weights

$$w_t^{(i)} \propto p(\mathbf{y}_t|\boldsymbol{\theta}_t^{(i)})w_{t-1}^{(i)}, \quad i = 1, \ldots, N.$$

10.4.1 Computational and observational grids

The deposition is estimated on a polar grid covering the eastern half of the zone of emergency planning using total number $N_\mathbf{y} = 136$ observations. Rectangular observational grid with $N_\mathbf{d} = 520$ grid points has the grid step 1.5km. Illustration of the observational and computational grid is in Figure (10.1). The fact that the polar computational grid and the rectangular observational grid are not aligned means that the observation operator must interpolate modeled values into the locations of receptor points. We implemented a linear observation operator $\mathbf{H}$ using the bilinear interpolation.

10.4.2 Estimation of prior distribution of deposition

Let us recall, that the initial conditions in the late phase are fully determined by the plume trajectory and plume depletion during the early phase. For simulation of the release we use atmospheric dispersion model from the environmental code HARP (HAzardous Radioactivity Propagation) [PHP07]. It is a segmented Gaussian plume model capable for simulation of many physical processes.

We performed an extensive Monte Carlo simulation to consistently account for uncertainty connected with the plume propagation. The simulation covered a broad range of possible release scenarios, where 14 parameters $\boldsymbol{v} = (v_1, \ldots, v_{14})^\mathrm{T}$ of the dispersion model were treated as random. The list of parameters selected upon uncertainty study performed with the model is in Table (10.1). The total number of 5000 model realization were computed using inputs $\{\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(5000)}\}$ sampled from a joint prior pdf $p(\boldsymbol{v})$.

Let $\{\mathbf{d}_0^{(1)}, \ldots, \mathbf{d}_0^{(5000)}\}$ be deposition vectors generated using the dispersion model initialized with different sets of input parameters $\{\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(5000)}\}$. Prior pdf $p(\mathbf{d}_0) = \mathcal{N}(\bar{\mathbf{d}}_0, \boldsymbol{\Sigma}_0^\mathbf{d})$ represents an empirical distribution of the initial deposition:

$$\bar{\mathbf{d}}_0 = \frac{1}{5000}\sum_{i=1}^{5000}\mathbf{d}_0^{(i)}, \quad \boldsymbol{\Sigma}_0 = \frac{1}{4999}\sum_{i=1}^{5000}\left(\mathbf{d}_0^{(i)} - \bar{\mathbf{d}}_0\right)\left(\mathbf{d}_0^{(i)} - \bar{\mathbf{d}}_0\right)^\mathrm{T}.$$
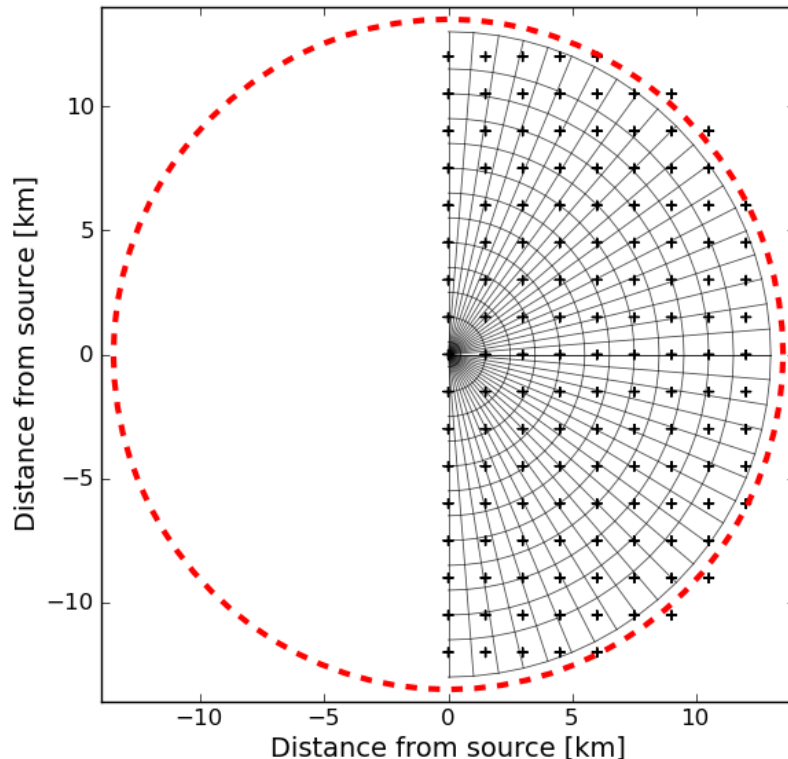
Figure 10.1: Illustration of computational and observational grids. We focus on the eastern half of the zone of emergency planning.

Param.	Physical effect	Param.	Physical effect
v_1	intensity of release	v_8	advection speed of plume
v_2	horizontal dispersion	v_9	wind profile
v_3	horizontal fluctuation of wind dir.	v_{10}	vertical dispersion
v_4	dry deposition of elements	v_{11}	mixing layer height
v_5	dry deposition of aerosols	v_{12}	heat capacity of the effluent
v_6	elution of elemental. iodine	v_{13}	precipitation intensity
v_7	elution of aerosols	v_{14}	time shift of precipitation

Table 10.1: Parameters treated as uncertain during Monte Carlo sampling of candidates on ensemble members.

10.4.3   Simulation of observations

Measurements are simulated from the deposition vectors $\mathbf{d}_t^{\text{twin}}$, $t = 1, \ldots, 24$, evaluated using (2.29) and an initial vector $\mathbf{d}_0^{\text{twin}}$ computed with the HARP model with initialized with a set of inputs $\boldsymbol{v}^{\text{twin}}$. The speed of environmental removal in the twin model is determined by the fast component with fraction $d^{\text{f}} = 0.54$ and half-life $T^{\text{f}} = 1.20$ years.

Measurement vectors $\mathbf{y}_t$ are sampled from $\mathbf{d}_t^{\text{twin}}$ in receptor locations and perturbed with zero-mean Gaussian noise according to

$$\mathbf{y}_t \sim \mathcal{N}(\mathbf{H}\mathbf{d}_t^{\text{twin}}, \mathbf{R}_t). \tag{10.10}$$

The standard deviation of the observation error is a linear function of the measured values $\mathbf{y}_t^{\text{twin}} = \mathbf{H}\mathbf{d}_t^{\text{twin}}$,

$$\mathbf{R}_t[j,j] \quad = \quad (0.1\mathbf{y}_t^{\text{twin}}[j] + 500)^2, \quad j = 1, \ldots, N_{\mathbf{y}}.$$

Deposition generated using the twin model is in Figure 10.2 (top).

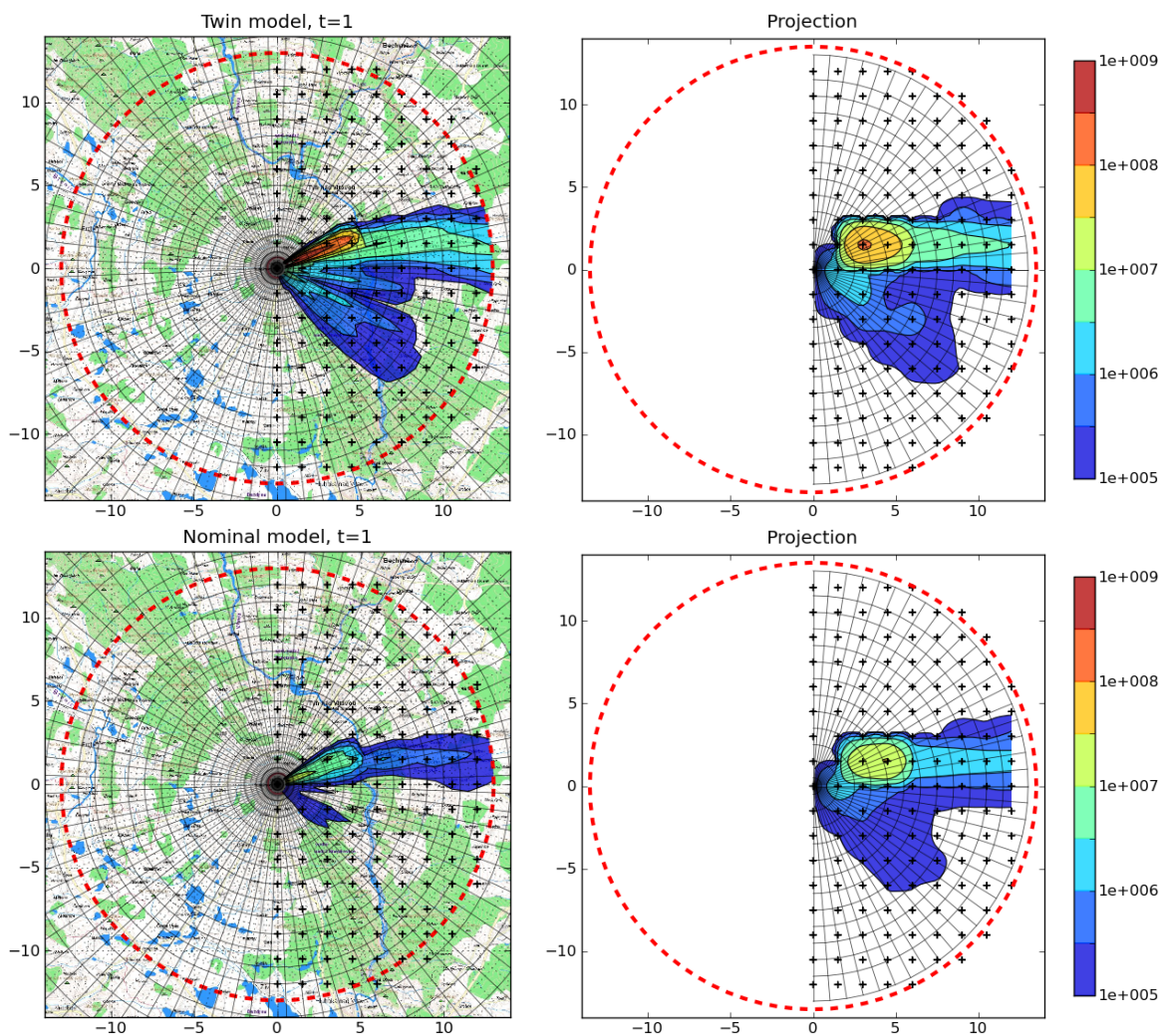Figure 10.2: Left column: Visualization of initial deposition of the twin model $\mathbf{d}_0^{\text{twin}}$ (top) and of the initial ensemble mean (bottom). Right column: Interpolation of $\mathbf{d}_0^{\text{twin}}$ projected into the space of observations using (10.10) (top) and interpolation of the initial ensemble mean projected into the space of observations using (10.10) (bottom).

10.4.4 Selection of prior ensemble

EnSRFs within the MPF framework are initialized with a prior ensemble. This ensemble represents our initial belief on distribution of the deposition at the beginning of the late phase. In Section 10.4.2 we estimated the prior distribution of the deposition using 5000 realizations generated with the HARP system. However, to achieve a better agreement of the initial ensemble with the spatial localization of the deposition, we can use measurements $\mathbf{y}_0$ referring to the beginning of the late phase and select those realizations from the set $\{\mathbf{d}_0^{(1)}, \dots, \mathbf{d}_0^{(5000)}\}$, which are in the best agreement with the measurements. This procedure increases the representativeness of the prior ensemble. Realizations are weighted with the likelihood function

$$p(\mathbf{y}_0|, \mathbf{d}_0^{(i)}) \quad \propto \quad \det\left(\mathbf{Z}_0\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\mathbf{y}_0 - \mathbf{H}\bar{\mathbf{d}}_0^{(i)}\right)^{\mathrm{T}} \mathbf{Z}_0^{-1} \left(\mathbf{y}_0 - \mathbf{H}\bar{\mathbf{d}}_0^{(i)}\right)\right],$$

$$\mathbf{Z}_0 \quad = \quad \mathbf{H}\mathbf{\Sigma}_0^{\mathbf{d}}\mathbf{H}^{\mathrm{T}} + \mathbf{R}_0.$$

Resulting weights are in Figure 10.3. We used the multinomial re-sampling and selected 20 realizations which were included into the prior ensemble. 2-dimensional visualization of the prior ensemble mean is in Figure 10.2 (bottom).

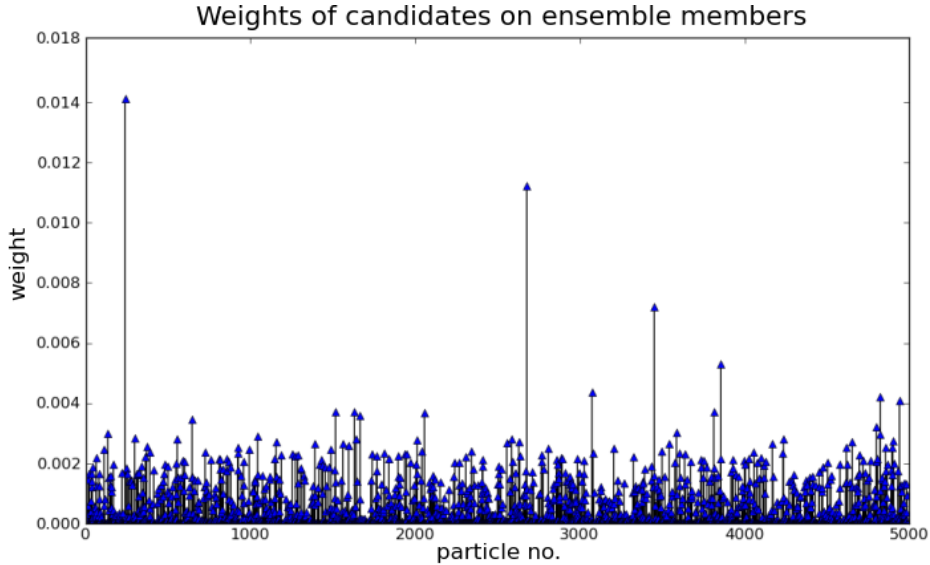Figure 10.3: Weights of candidates $\mathbf{d}_0^{(i)}$ on members of initial ensemble.

10.4.5 Results

We run the data assimilation algorithm with 100 particles for 24 steps, i.e., 100 ensemble square root filters were run using different parameters $\boldsymbol{\theta}_t^{(i)}$ in respective time steps. In Figure 10.4 we see the deposition values for a randomly selected receptor location. We
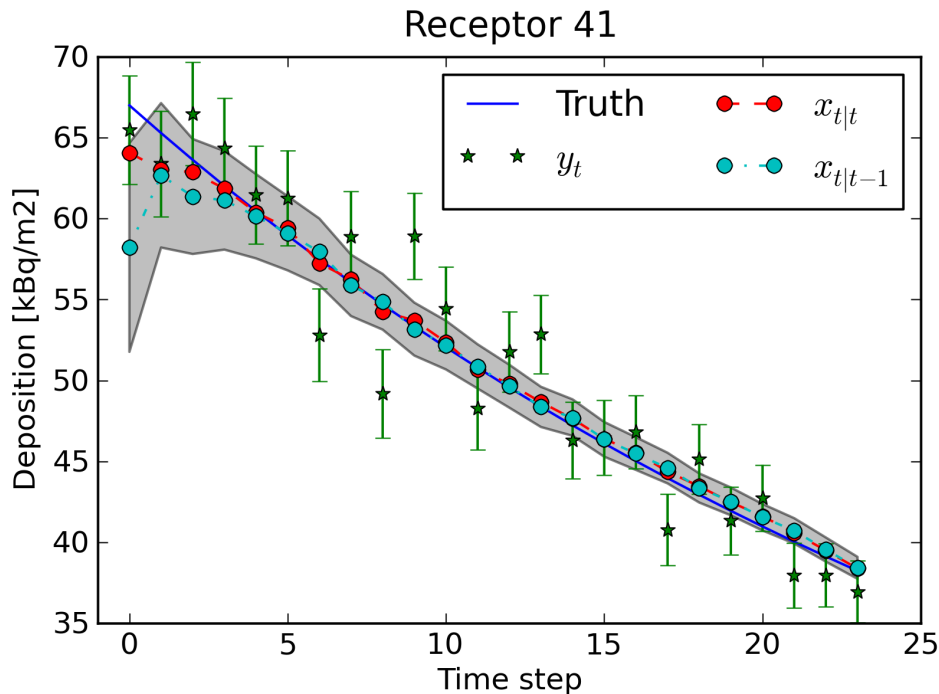
Figure 10.4: Results for a selected observation location. Blue line: true deposition (twin model); green dots: measurements and their standard deviations; cyan dots: prior estimates (forecasts); gray area: standard deviation of the forecast error; red dots: posterior estimates.

see that the posterior values (red dots) becomes identical with the true deposition (blue line) given by the twin model. Also that the variance of estimates (gray area) decreases with time.

In Figure 10.5 we see spatial visualization of the data assimilation results for time step 0, 8, 16, and 24. The nominal model, the twin model and the assimilated model are compared in respective time steps. We see that the ensemble localization procedure perform well and the initial estimate of affected area is similar to that given by the twin model. Already after the first assimilation cycle we obtain a good agreement between the assimilated model and the twin model. In the remaining time steps we observe, how the assimilated model approaches the twin model. To achieve a good agreement in all the computational points, more data assimilation cycles would be needed.

Besides the time evolution of the deposition we also estimate the speed of radionuclides removal given by the fast component of the environmental removal. Estimated parameters d^f and T^f are assumed to be time invariant. Comparison of the "true" parameters used for the simulation of measurement and the average values of the estimates is in Table 10.2. We observe a good correspondence of the estimates and the true values. The estimates can be in turn used as an input into the subsequent predictive models regarding further transport of radionuclides through the environment.

Figure 10.5: Visualization of assimilation results in time steps 0, 8, 16, and 24. Nominal model (prior mean), twin model (measurements) and assimilated model (posterior mean) are in first, second and third column, respectively. Color scale is the same as in Figure 10.2.

Parameter	"True" values	Mean estimated values
d^{f}	0.54	0.57
T^{f}	1.20	1.26

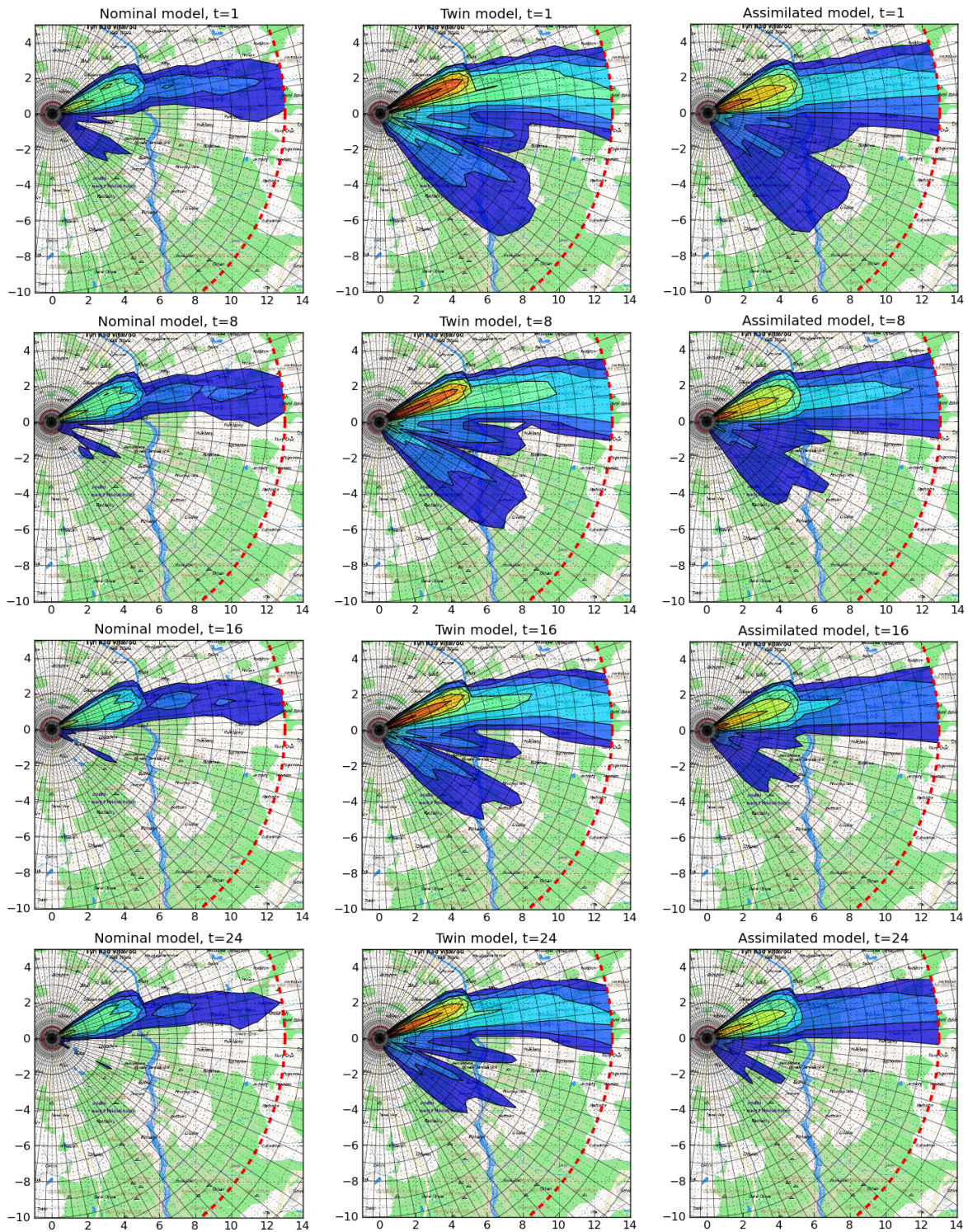Table 10.2: Comparison of "true" parameters used for simulation of measurements with those estimated using the data assimilation procedure.

The mean value of the estimated multiplicative inflation factor used for correction of model error was 2.27 which indicate the fact that the spread of the initial ensemble was heavily underestimated.

10.5   Summary

This chapter has addressed the problem of data assimilation in the late phase of a radiation accident. We were concerned with the task of data assimilation of the forecasted spatio-temporal distribution of deposition with the groundshine dose measurements.

In [Pal05], ensemble Kalman filtering has been identified as the most promising approach for this task. We developed this idea further and applied the MPF framework introduced in Chapter 8 for estimation of spatio-temporal distribution of the groundshine in tandem with the speed of environmental removal. In the proposed data assimilation method, particle filter approximates the posterior pdf of the global speed of radionuclides removal, whereas the conditional ensemble square root filter accounts for local variations in the deposition field reflected in groundshine dose measurements.

The performance of the method was demonstrated on a twin experiment, where the groundshine dose evolution model was assimilated with groundshine dose measurements. A release of ^{134}Cs was simulated with the atmospheric dispersion model embedded in the decision support system HARP. We performed an extensive Monte Carlo simulation of the possible release scenarios in order to account for uncertainty regarding the plume propagation during the early phase. Deposition fields resulting from 20 most plausible release scenarios were used for initialization of the ensemble filters attached to particles. The estimated deposition field approached that evaluated using the twin model. The rate of environmental removal also was correctly identified. The results are promising, however, the operational applicability remains to be demonstrated on real deposition data.

Chapter 11

Conclusion

This research has addressed the application of data assimilation methods in the field of radiation protection. We focused on exploitation of advanced data assimilation methods, namely particle filtering and marginalized particle filtering for assessment of radiation situation in the early and the late phase of a radiation accident. Respective data assimilation methodologies were formalized using consistent Bayesian framework and notation.

The research has demonstrated that the in case of continuous data assimilation, e.g. during the early phase, particle filtering approach, when applied in this area, provides useful insights into the problem and results in improved versatility over more traditional approaches, e.g. Kalman filtering. Data assimilation systems based on particle filtering have a potential to be used for real-world emergency response in the near future. We have also demonstrated, that simple data assimilation techniques, like optimal interpolation of funtion fitting methods, can be successfully applied to special scenarios of intermittent assimilation.

Since no data from real reactor accident were available, all experiments were performed as twin experiments. In the twin experiments, measurements are simulated using a model of the system under investigation and perturbed with a random noise. The convergence of the estimated values to the known "background truth" can be then easily assessed.

Appendix A

Mean value and covariance of a mixture of Gaussians

Assume a weighted mixture of n Gaussian distributions $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^{(i)}, \mathbf{P}^{(i)})$ of a random variable $\mathbf{x}$ with mean values $\boldsymbol{\mu}^{(i)}$ and covariance $\mathbf{P}^{(i)}$, $i = 1, \ldots, n$:

$$p(\mathbf{x}) = \sum_{i=1}^{n} w^{(i)} \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^{(i)}, \mathbf{P}^{(i)})$$

The mean $\mathbb{E}[\mathbf{x}]$ and covariance $\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])^2]$ of $\mathbf{x}$ can be calculated as follows:

$$
\begin{aligned}
\mathbb{E}[\mathbf{x}] &= \int \mathbf{x} p(\mathbf{x}) d\mathbf{x} \\
&= \sum_{i=1}^{n} w^{(i)} \int \mathbf{x} \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^{(i)}, \mathbf{P}^{(i)}) d\mathbf{x} \\
&= \sum_{i=1}^{n} w^{(i)} \boldsymbol{\mu}^{(i)}, \\
\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])^2] &= \int (\mathbf{x} - \mathbb{E}[\mathbf{x}])^2 p(\mathbf{x}) d\mathbf{x} \\
&= \sum_{i=1}^{n} w^{(i)} \int (\mathbf{x} - \mathbb{E}[\mathbf{x}])^2 \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^{(i)}, \mathbf{P}^{(i)}) d\mathbf{x} \\
&= \sum_{i=1}^{n} w^{(i)} \int \left[(\mathbf{x} - \boldsymbol{\mu}^{(i)}) + (\boldsymbol{\mu}^{(i)} - \mathbb{E}[\mathbf{x}]) \right]^2 \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}^{(i)}, \mathbf{P}^{(i)}) d\mathbf{x} \\
&= \sum_{i=1}^{n} w^{(i)} \left[\mathbf{P}^{(i)} + (\boldsymbol{\mu}^{(i)} - \mathbb{E}[\mathbf{x}])^2 \right] \\
&= \sum_{i=1}^{n} w^{(i)} \mathbf{P}^{(i)} + \sum_{i=1}^{n} w^{(i)} (\boldsymbol{\mu}^{(i)} - \mathbb{E}[\mathbf{x}])^2.
\end{aligned}
$$

Appendix B

Computation of time update using square roots

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{C} \in \mathbb{R}^{m \times n}$, $\mathbf{D} \in \mathbb{R}^{n \times n}$ be real matrices. Following [Sim06], let us suppose that we can find an orthogonal matrix $\mathbf{T} \in \mathbb{R}^{2n \times 2n}$ such that

$$\begin{bmatrix} \mathbf{A}^{\mathrm{T}} \\ \mathbf{0} \end{bmatrix} = \mathbf{T} \begin{bmatrix} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} \\ \mathbf{D}^{\frac{\mathrm{T}}{2}} \end{bmatrix} \tag{11.1}$$

$$= \begin{bmatrix} \mathbf{T_1} & \mathbf{T_2} \end{bmatrix} \begin{bmatrix} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} \\ \mathbf{D}^{\frac{\mathrm{T}}{2}} \end{bmatrix} = \begin{bmatrix} \mathbf{T_1} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} + \mathbf{T_2} \mathbf{D}^{\frac{\mathrm{T}}{2}} \end{bmatrix}, \tag{11.2}$$

where $\mathbf{T_1}, \mathbf{T_2} \in \mathbb{R}^{n \times n}$. From orthogonality of $\mathbf{T}$ follows that

$$\mathbf{T}^{\mathrm{T}} \mathbf{T} = \begin{bmatrix} \mathbf{T_1}^{\mathrm{T}} \\ \mathbf{T_2}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{T_1} & \mathbf{T_2} \end{bmatrix} = \begin{bmatrix} \mathbf{T_1}^{\mathrm{T}} \mathbf{T_1} & \mathbf{T_1}^{\mathrm{T}} \mathbf{T_2} \\ \mathbf{T_2}^{\mathrm{T}} \mathbf{T_1} & \mathbf{T_2}^{\mathrm{T}} \mathbf{T_2} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

and consequently

$$\mathbf{T_1}^{\mathrm{T}} \mathbf{T_2} = \mathbf{T_2}^{\mathrm{T}} \mathbf{T_1} = \mathbf{0}, \tag{11.3}$$

$$\mathbf{T_1}^{\mathrm{T}} \mathbf{T_1} = \mathbf{T_2}^{\mathrm{T}} \mathbf{T_2} = \mathbf{I}, \tag{11.4}$$

where $\mathbf{0}$ and $\mathbf{I}$ are zero and identity matrices, respectively. Using (11.2) we can write

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{\mathrm{T}} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_t \mathbf{B} & \mathbf{D}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{T_1}^{\mathrm{T}} \\ \mathbf{T_2}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{T_1} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} + \mathbf{T_2} \mathbf{D}^{\frac{\mathrm{T}}{2}} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{H}_t \mathbf{B} & \mathbf{D}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{T_1}^{\mathrm{T}} \mathbf{T_1} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} + \mathbf{T_1}^{\mathrm{T}} \mathbf{T_2} \mathbf{D}^{\frac{\mathrm{T}}{2}} \\ \mathbf{T_2}^{\mathrm{T}} \mathbf{T_1} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} + \mathbf{T_2}^{\mathrm{T}} \mathbf{T_2} \mathbf{D}^{\frac{\mathrm{T}}{2}} \end{bmatrix}.$$

From identities (11.3)–(11.4) follows

$$\mathbf{A} \mathbf{A}^{\mathrm{T}} = \mathbf{C} \mathbf{B} \mathbf{T_1}^{\mathrm{T}} \mathbf{T_1} \mathbf{B}^{\mathrm{T}} \mathbf{C}_t^{\mathrm{T}} + \mathbf{D}^{\frac{1}{2}} \mathbf{T_2}^{\mathrm{T}} \mathbf{T_2} \mathbf{D}^{\frac{\mathrm{T}}{2}}$$

$$= \mathbf{C} \mathbf{B} \mathbf{B}^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} + \mathbf{D}^{\frac{1}{2}} \mathbf{D}^{\frac{\mathrm{T}}{2}}.$$

So if we can find an orthogonal matrix $\mathbf{T} \in \mathbb{R}^{2n \times 2n}$ such that (11.1) is fulfilled, then matrix $\mathbf{A}$ is equal to transpose of square root of $\mathbf{Z} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$. We can use various methods to find the orthogonal matrix and resulting square root of $\mathbf{Z}$, e.g., Householder transform, Gram-Schmidt orthogonalization, or Givens rotations [GVL96]. Different methods give us different square roots of $\mathbf{Z}$. From QR decomposition we obtain a upper triangular factor of the Cholesky decomposition of $\mathbf{Z}$.

Bibliography

[AA99]   J.L. Anderson and S.L. Anderson, *A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts*, Monthly Weather Review **127** (1999), 2741–2758.

[ADFDJ03]   C. Andrieu, N. De Freitas, A. Doucet, and M.I. Jordan, *An introduction to MCMC for machine learning*, Machine learning **50** (2003), no. 1, 5–43.

[ADH10]   C. Andrieu, A. Doucet, and R. Holenstein, *Particle Markov chain Monte Carlo methods*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **72** (2010), no. 3, 269–342.

[And07a]   J.L. Anderson, *An adaptive covariance inflation error correction algorithm for ensemble filters*, Tellus A **59** (2007), no. 2, 210–224.

[And07b]   _____, *Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter*, Physica D: Nonlinear Phenomena **230** (2007), no. 1-2, 99–111.

[ATP+04]   P. Astrup, C. Turcanu, RO Puch, C.R. Palma, and T. Mikkelsen, *Data assimilation in the early phase: Kalman filtering RIMPUFF*, Risø NL (2004).

[Bar01]   R. Barratt, *Atmospheric dispersion modelling: an introduction to practical applications*, Earthscan, 2001.

[BC02]   F. Bouttier and P. Courtier, *Data assimilation concepts and methods march 1999*, Meteorological training course lecture series. ECMWF (2002).

[BFS86]   E. Browne, R.B. Firestone, and V.S. Shirley, *Table of radioactive isotopes.*

[BSN03]   T. Bengtsson, C. Snyder, and D. Nychka, *Toward a nonlinear ensemble filter for high-dimensional systems*, Journal of Geophysical Research **108** (2003), no. D24, 8775.

[BvLE98]   G. Burgers, P.J. van Leeuwen, and G. Evensen, *Analysis scheme in the ensemble Kalman filter*, Monthly Weather Review **126** (1998), 1719–1724.

[CEE+95]   DJ Carruthers, HA Edmunds, KL Ellis, CA McHugh, BM Davies, and DJ Thomson, *Atmospheric Dispersion Modelling System (ADMS): Comparisons with data from the Kincaid experiment*, International Journal of Environment and Pollution **5** (1995), no. 4, 382–400.

[Che02]   N.P. Cheremisinoff, *Handbook of air pollution prevention and control*, Butterworth-Heinemann, 2002.

[CLR02]   G.R. Choppin, J.O. Liljenzin, and J. Rydberg, *Radiochemistry and nuclear chemistry*, Butterworth-Heinemann, 2002.

[Coo92]   W. A. Cooper, *The analysis of observations with applications in atmospheric science*, National Center for Atmospheric Research, 1992.

[CPV+04]   AJ Cimorelli, SG Perry, A. Venkatram, JC Weil, RJ Paine, RB Wilson, RF Lee, WD Petersand, RW Brode, and JO Paumier, *AERMOD: description of model formulation (EPA–454/R–03–004)*, US Environmental Protection Agency, 2004.

[Dal93]   R. Daley, *Atmospheric data analysis*, Cambridge Univ Pr, 1993.

[DC05]   R. Douc and O. Cappé, *Comparison of resampling schemes for particle filtering*, Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005. ISPA 2005., IEEE, 2005, pp. 64–69.

[DDFG01]   A. Doucet, N. De Freitas, and N. Gordon, *Sequential Monte Carlo methods in practice*, Springer Verlag, 2001.

[Dee95]   D.P. Dee, *On-line estimation of error covariance parameters for atmospheric data assimilation*, Monthly weather review **123** (1995), no. 4, 1128–1145.

[DGA00]   A. Doucet, S. Godsill, and C. Andrieu, *On sequential Monte Carlo sampling methods for Bayesian filtering*, Statistics and computing **10** (2000), no. 3, 197–208.

[DLM05]   M. Drews, B. Lauritzen, and H. Madsen, *Analysis of a Kalman filter based method for on-line estimation of atmospheric dispersion parameters using radiation monitoring data*, Radiation protection dosimetry **113** (2005), no. 1, 75.

[DLR+77]   A.P. Dempster, N.M. Laird, D.B. Rubin, et al., *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society. Series B (Methodological) **39** (1977), no. 1, 1–38.

[DNTW09]   H. Dombrowski, S. Neumaier, IMG Thompson, and F. Wissmann, *EURADOS intercomparison 2006 to harmonise European early warning dosimetry systems*, Radiation Protection Dosimetry **135** (2009), no. 1, 1.

[EKT07]   H. Eleveld, Y.S. Kok, and C.J.W. Twenhöfel, *Data assimilation, sensitivity and uncertainty analyses in the Dutch nuclear emergency management system: a pilot study*, International Journal of Emergency Management **4** (2007), no. 3, 551–563.

[ET04]   H. Eleveld and CJW Twenhöfel, *Improving the reliability of the prognosis of atmospheric dispersion using data assimilation*, Ninth International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 2004, pp. 1–4.

[Eve94]   G. Evensen, *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, Journal of Geophysical Research **99** (1994), 10–10.

[GBEJ99] V. Golikov, M. Balonov, V. Erkin, and P. Jacob, *Model validation for external doses due to environmental contaminations by the Chernobyl accident*, Health physics **77** (1999), no. 6, 654.

[GC99]   G. Gaspari and S.E. Cohn, *Construction of correlation functions in two and three dimensions*, Quarterly Journal of the Royal Meteorological Society **125** (1999), no. 554, 723–758.

[Gel04]   A. Gelman, *Bayesian data analysis*, CRC press, 2004.

[Gif76]   F.A. Gifford, *Turbulent diffusion-typing schemes: A review*, Nuclear Safety **17** (1976), no. 1, 68–86.

[GSS93]   N.J. Gordon, D.J. Salmond, and A.F.M. Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*, vol. 140, IEE, 1993, pp. 107–113.

[GSS02]   NJ Gordon, DJ Salmond, and AFM Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*, Radar and Signal Processing, IEE Proceedings F, vol. 140, IET, 2002, pp. 107–113.

[Gur08]   B.R. Gurjar, *Air Pollution: Health and Environmental Impacts*, CRC, 2008.

[GVL96]   G.H. Golub and C.F. Van Loan, *Matrix computations*, Johns Hopkins Univ Pr, 1996.

[GWW+04] F. Gering, W. Weiss, E. Wirth, R. Stapel, P. Jacob, H. Müller, and G. Pröhl, *Assessment and evaluation of the radiological situation in the late phase of a nuclear accident*, Radiation protection dosimetry **109** (2004), no. 1-2, 25.

[Hag89]   W.W. Hager, *Updating the inverse of a matrix*, SIAM review **31** (1989), no. 2, 221–239.

[HBHJ82]   SR Hanna, GA Briggs, and RP Hosker Jr, *Handbook on atmospheric diffusion*, Tech. report, National Oceanic and Atmospheric Administration, Oak Ridge, TN (USA). Atmospheric Turbulence and Diffusion Lab., 1982.

[HKvD11]   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 (2011), 6149–6157.

[HM01]   P.L. Houtekamer and H.L. Mitchell, *A sequential ensemble Kalman filter for atmospheric data assimilation*, Monthly Weather Review **129** (2001), no. 1, 123–137.

[HM06]   N.S. Holmes and L. Morawska, *A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available*, Atmospheric Environment **40** (2006), no. 30, 5902–5928.

[HMP⁺05]   P.L. Houtekamer, H.L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, *Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations*, Monthly Weather Review **133** (2005), no. 3, 604–620.

[Hom02]   T. Homma, *OSCAAR Development and Applications*, International MACCS user group I (2002), 57.

[HPP08a]   R. Hofman, P. Pecha, and E. Pechová, *A simplified approach for solution of time update problem during toxic waste plume spreading in atmosphere*, Proc. of 12th Inter. Conf. on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, HARMO-12, Hrvatski Meterološki Časopis, 2008, pp. 510–515.

[HPP08b]   R. Hofman, P. Pecha, and E. Pechová, *A simplified approach for solution of time update problem during toxic waste plume spreading in atmosphere*, 10-th In. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, (Cavtat, HR, 06.10.2008-10.10.2008), 2008.

[IAE03]   IAEA, *Testing of environmental transfer models using Chernobyl fallout data from the Iput River catchment area, Bryansk Region, Russian Federation*, 2003.

[ICGL97]   K. Ide, P. Courtier, M. Ghil, and A.C. Lorenc, *Unified notation for data assimilation: Operational, sequential and variational*, J. Meteor. Soc. Japan **75** (1997), no. 1B, 181–189.

[IH05]   J. S. Irwin and S. R. Hanna, *Characterizing uncertainty in plume dispersion models*, Int. Journal of ennvironment and Pollution, No. 1-4, 2005, pages 16 - 25 (2005).

[Jaz70]   A.H. Jazwinski, *Stochastic processes and filtering theory*, Academic Pr, 1970.

[Jef61]   H Jeffreys, *Theory of probability*, Oxford University Press, 1961.

[JHN04]   G. Johannesson, B. Hanley, and J. Nitao, *Dynamic bayesian models via monte carlo–an introduction with examples*, Tech. report, Lawrence Livermore National Laboratory, 2004.

[JKS$^+$05] H.J. Jeong, E.H. Kim, K.S. Suh, W.T. Hwang, M.H. Han, and H.K. Lee, *Determination of the source rate released into the environment from a nuclear power plant*, Radiation protection dosimetry **113** (2005), no. 3, 308.

[JPP03]   Y.R. Jung, W.G. Park, and O.H. Park, *Pollution dispersion analysis using the puff model with numerical flow field data*, Mechanics research communications **30** (2003), no. 4, 277–286.

[JU97]   S.J. Julier and J.K. Uhlmann, *A new extension of the Kalman filter to nonlinear systems*, Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, vol. 3, 1997, p. 26.

[Kal60]   R.E. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering **82** (1960), no. 1, 35–45.

[Kal03]   E. Kalnay, *Atmospheric modeling, data assimilation, and predictability*, Cambridge Univ Pr, 2003.

[KTAB09]   I.V. Kovalets, V. Tsiouri, S. Andronopoulos, and J.G. Bartzis, *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** (2009), no. 8, 3511–3521.

[LE98]   E.N. Lorenz and K.A. Emanuel, *Optimal sites for supplementary weather observations: Simulation with a small model*, Journal of the Atmospheric Sciences **55** (1998), 399–414.

[LGR77]   FL Ludwig, LS Gasiorek, and RE Ruff, *Simplification of a Gaussian puff model for real-time minicomputer use*, Atmospheric Environment (1967) **11** (1977), no. 5, 431–436.

[LKM09]   H. Li, E. Kalnay, and T. Miyoshi, *Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter*, Quarterly Journal of the Royal Meteorological Society **135** (2009), no. 639, 523–533.

[Man06]   J. Mandel, *Efficient implementation of the ensemble Kalman filter*, University of Colorado at Denver and Health Sciences Center, 2006.

[MH00]   H.L. Mitchell and PL Houtekamer, *An adaptive ensemble Kalman filter*, Monthly Weather Review **128** (2000), no. 2, 416–433.

[OB92]   M.S. Oh and J.O. Berger, *Adaptive importance sampling in Monte Carlo integration*, Journal of Statistical Computation and Simulation **41** (1992), no. 3, 143–168.

[OVZ07]   Y. Onishi, O.V. Voitsekhovich, and M.J. Zheleznyak, *Chernobyl–what have we learned?: the successes and failures to mitigate water contamination over 20 years*, Springer, 2007.

[Pal05]   C.R. Palma, *Data assimilation for off-site nuclear emergency management, report RODOS(RA5)-RE(04)-01*, 2005.

[Pas61]   F. Pasquill, *The estimation of the dispersion of windborne material*, Meteorol. Mag **90** (1961), no. 1063, 33–49.

[Pet81]   V. Peterka, *Bayesian system identification*, Automatica **17** (1981), no. 1, 41–53.

[PFM93]   G. Pröhl, W. Friedland, and H. Müller, *Potential reduction of the ingestion dose after nuclear accidents due to the application of selected countermeasures*, Radiation Protection Dosimetry **50** (1993), no. 2, 359.

[PH08]   P. Pecha and R. Hofman, *Fitting of segmented gaussian plume model predictions on measured data*, Proc. of 22th European Simulation and Modelling Conference ESM'2008, Le Havre, France, 2008.

[PHP07]   P. Pecha, R. Hofman, and E. Pechová, *Training simulator for analysis of environmental consequences of accidental radioactivity releases*, Proc. of 6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia, 2007.

[PMG+03]   C.R. Palma, H. Madsen, F. Gering, R. Puch, C. Turcanu, P. Astrup, H. Müller, K. Richter, M. Zheleznyak, D. Treebushny, et al., *Data assimilation in the decision support system RODOS*, Radiation protection dosimetry **104** (2003), no. 1, 31.

[PS00]   J. Päsler-Sauer, *Description of the atmospheric dispersion model ATSTEP*, Report of project RODOS (WG2)-TN (99)-11 (2000).

[PX09]   S.K. Park and L. Xu, *Data assimilation for atmospheric, oceanic and hydrologic applications*, Springer Verlag, 2009.

[QSI05]   D. Quelo, B. Sportisse, and O. Isnard, *Data assimilation for short range atmospheric dispersion of radionuclides: a case study of second-order sensitivity*, Journal of environmental radioactivity **84** (2005), no. 3, 393–408.

[RAC01]   S.S. Raza, R. Avila, and J. Cervantes, *A 3-D Lagrangian (Monte Carlo) method for direct plume gamma dose rate calculations*, Journal of Nuclear Science and Technology **38** (2001), no. 4, 254–260.

[RAG04]   B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*, Artech House Publishers, 2004.

[Rao05]   K.S. Rao, *Uncertainty analysis in atmospheric dispersion modeling*, Pure and Applied Geophysics **162** (2005), no. 10, 1893–1917.

[RK08]   R.Y. Rubinstein and D.P. Kroese, *Simulation and the Monte Carlo method*, Wiley-Interscience, 2008.

[Sas58]   Y. Sasaki, *An objective analysis based on the variational method*, J. Meteor. Soc. Japan **36** (1958), no. 3, 77–88.

[SGN05]   T. Schön, F. Gustafsson, and P.J. Nordlund, *Marginalized particle filters for mixed linear/nonlinear state-space models*, IEEE Transactions on Signal Processing **53** (2005), no. 7, 2279–2289.

[ŠH11]   V. Šmídl and R. Hofman, *Marginalized particle filtering framework for tuning of ensemble filters*, Monthly Weather Review **139** (2011), 3589–3599.

[Sim06]   D. Simon, *Optimal state estimation: Kalman, H_∞ and nonlinear approaches*, John Wiley and Sons, 2006.

[SOvG10]   S. Saha, E. Özkan, V. Šmídl, and F. Gustafsson, *Marginalized particle filters for Bayesian estimation of noise parameters*, Proc. of the 13th International Conference on Information Fusion (Edinburgh, UK), 2010.

[TAB+03]   MK Tippett, JL Anderson, CH Bishop, TM Hamill, and JS Whitaker, *Ensemble square-root filters*, Mon. Weather Rev **131** (2003), 1485–1490.

[TBF05]   S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*, MIT press, 2005.

[TL72]   H. Tennekes and J.L. Lumley, *A first course in turbulence*, MIT press, 1972.

[TNDL95]   S. Thykier-Nielsen, S. Deme, and E. Láng, *Calculation Method for Gamma Dose Rates From Gaussian Puffs*, Risø National Laboratory, 1995.

[TNDM99]   S. Thykier-Nielsen, S. Deme, and T. Mikkelsen, *Description of the atmospheric dispersion module RIMPUFF*, Risø National Laboratory, 1999.

[TSA⁺05]   KM Thiessen, TG Sazykina, AI Apostoaei, MI Balonov, J. Crawford, R. Domel, SV Fesenko, V. Filistovic, D. Galeriu, T. Homma, et al., *Model testing using data on Cs-137 from Chernobyl fallout in the Iput River catchment area of Russia*, Journal of environmental radioactivity **84** (2005), no. 2, 225–244.

[TvTB07]   CJW Twenhöfel, MM van Troost, and S. Bader, *Uncertainty analysis and parameter optimization in early phase nuclear emergency management*, 2007.

[WB95]   G. Welch and G. Bishop, *An introduction to the Kalman filter*, University of North Carolina at Chapel Hill (1995).

[WH02]   J.S. Whitaker and T.M. Hamill, *Ensemble data assimilation without perturbed observations*, Monthly Weather Review **130** (2002), 1913–1924.

[WZZ00]   B. Wang, X. Zou, and J. Zhu, *Data assimilation and its applications*, Proceedings of the National Academy of Sciences of the United States of America **97** (2000), no. 21, 11143.

[YKM⁺05]   S. Yuschenko, I. Kovalets, V. Maderich, D. Treebushny, and M. Zheleznyak, *Modelling the radionuclide contamination of the black sea in the result of chernobyl accident using circulation model and data assimilation*, Radioprotection **40** (2005), 685–691.

[Zan86]   P. Zannetti, *A new mixed segment-puff approach for dispersion modeling*, Atmospheric Environment (1967) **20** (1986), no. 6, 1121–1130.

[Zan90]   _____, *Air pollution modeling*, Van Nostrand Reinhold, 1990.

[ZLLL07]   D.Q. Zheng, J.K.C. Leung, B.Y. Lee, and H.Y. Lam, *Data assimilation in the atmospheric dispersion model for nuclear accident assessments*, Atmospheric environment **41** (2007), no. 11, 2438–2446.