Atmospheric Environment 43 (2009) 6312–6327

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

# Targeting of observations for accidental atmospheric release monitoring

Rachid Abida [a,b], Marc Bocquet [a,b,*]

[a] Université Paris-Est, CEREA, Joint Laboratory École des Ponts ParisTech and EDF R&D, Champs-sur-Marne, France
[b] INRIA, Paris-Rocquencourt research centre, France

ARTICLE INFO

*Article history:*
Received 17 June 2009
Received in revised form
10 September 2009
Accepted 16 September 2009

*Keywords:*
Radionuclides dispersion
Accidental release modelling
Data assimilation
Inverse modelling
Network design

ABSTRACT

In the event of an accidental atmospheric release of radionuclides from a nuclear power plant, accurate real-time forecasting of the activity concentrations of radionuclides is acutely required by the decision makers for the preparation of adequate countermeasures. Yet, the accuracy of the forecasted plume is highly dependent on the source term estimation. Inverse modelling and data assimilation techniques should help in that respect. However the plume can locally be thin and could avoid a significant part of the radiological monitoring network surrounding the plant. Deploying mobile measuring stations following the accident could help to improve the source term estimation. In this paper, a method is proposed for the sequential reconstruction of the plume, by coupling a sequential data assimilation algorithm based on inverse modelling with an observation targeting strategy. The targeting design strategy consists in seeking the optimal locations of the mobile monitors at time $t+1$ based on all available observations up to time $t$.

The performance of the sequential assimilation with and without targeting of observations has been assessed in a realistic framework. It focuses on the Bugey nuclear power plant (France) and its surroundings within 50 km from the plant. The existing surveillance network is used and realistic observational errors are assumed. The targeting scheme leads to a better estimation of the source term as well as the activity concentrations in the domain. The mobile stations tend to be deployed along plume contours, where activity concentration gradients are important. It is shown that the information carried by the targeted observations is very significant, as compared to the information content of fixed observations. A simple test on the impact of model error from meteorology shows that the targeting strategy is still very useful in a more uncertain context.



## 1. Introduction

From the experience gained in European countries after the Chernobyl accident, and in accordance with the International Atomic Energy Agency (IAEA) recommendations (IAEA, 1980) as well as the European Union (EU) 87/600 Council Decision, many EU and/or IAEA member states have increased their efforts in operational systems for early warning and monitoring of radioactivity in the environment (see for instance Galmarini et al., 2008). In France, The Institute of Radiation Protection and Nuclear Safety (IRSN), whose primary mission is to alert the French Nuclear Safety Authority, operates an automatic gamma ambient dose rate monitoring network (Teleray) composed of 164 stations covering uniformly the country. This automatic gamma-dose monitoring system has been established to detect potential radioactive releases from foreign and domestic sources (in France, there are 20 civil nuclear facilities). In routine monitoring situation, measurements of ambient dose rate are transmitted together with data on local rainfall at 10-min intervals to the IRSN. Hourly and daily average measurements are calculated and stored in a database. If a critical threshold value of gamma ambient dose rate is exceeded, an alarm is raised immediately and the agency issues a warning to the national and local authorities. Recently, aiming to enhance its automatic monitoring system, the IRSN has decided to setup a network of new automatic aerosol stations (Descartes project). The instruments should be able to measure activity concentrations of various aerosol radionuclides. A study was carried out by Abida et al. (2008) and Saunier et al. (2009) to issue recommendations on the design of such monitoring network.

In case of a regional to large scale release of radioactive materials in the environment, Teleray and Descartes networks will play a central role in the monitoring strategy during the response phase of a nuclear accident. Yet, such monitoring networks are designed to cover the whole country. They will not be very informative in describing a radioactive release at near-field scale occurring at one of the country's nuclear power plants. So to cope with this issue,

* Corresponding author at: CEREA, Ecole des Ponts ParisTech, 6-8 avenue Blaise Pascal, Cité Descartes, Champs-sur-Marne, 77455 Marne la Vallée, France.
*E-mail address:* bocquet@cerea.enpc.fr (M. Bocquet).

1352-2310/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2009.09.029

many gamma-dose monitoring stations are installed in the vicinity of each nuclear facility, within a typical radius of 22 km.

Early decisions such as exhorting the population to seek shelter, taking iodine tablets, could be based directly on the reading of gamma-dose detectors or aerosol radionuclides monitors. The static or stationary networks such as Teleray and those installed in the vicinity of a nuclear power plant will be inappropriate for this purpose because the exposure pathways depend to a large extend on ground deposition. In the absence of rain and in case of a well circumscribed cloud, calculated values for the ground deposition may be available (Levi, 1991). However the Chernobyl accident has taught that in case of rainfall one must cope with a patchy pattern (Clark and Smith, 1988), which can be assessed only by portable and preferably mobile equipment. Furthermore, in the early phase of an accident, decision makers wish to rely on high resolution and accurate information about the spatial distribution of the radioactive materials in the local environment around the accident site in order to issue adequate protective countermeasures. Nevertheless, in general, the density of the monitoring stations in the vicinity of a nuclear power plant is not enough to cover homogeneously the contaminated area around the accident site. In such circumstances, an adaptive monitoring strategy will increase the measurements density during the accidental event by adding new in-situ measurements from mobile devices. These data can be then ingested into an assimilation scheme so as to improve the dispersion model predictions and to obtain a real-time estimation of the radioactive plume spread.

In 2008, the crisis centres of the IRSN and of Électricité De France (EDF) which operates most of France's nuclear power plants, have carried out together an exercise aiming to optimise the coordination between the two centres. For the first time they deployed a mobile monitoring network of 7 gamma-dose monitoring stations, plus 12 fictitious stations, around the site accident by helicopter, within a single day. In this exercise, the design of the mobile network was based only on the main wind information, population density, and Gaussian dispersion model forecast during the day of the experiment. Consequently, the design was not obtained from optimal spatial design considerations. Later on, a joint venture between EDF, AREVA and CEA, in charge of such a deployment, may own up to 20 mobile dose measurement instruments.

After the start of the release, the question of where the mobile measuring devices should be placed is naturally raised. The design problem is solved by deciding at time $t$ where mobile stations should be located at time $t+1$ using some design criterion. This criterion could make use of the combined information given by the previous observations and an analysis state of the plume derived from a dispersion model and the use of data assimilation techniques. A widely used design criterion is to consider the prediction error covariance at time $t+1$. In the spatial statistics community, the construction of time-varying designs is known as targeting design, dynamic design, adaptive design or sometimes as space-time dynamic design.

In the recent years, great efforts have been devoted to examining optimal spatial designs for environmental process (e.g. Cressie and Cassie, 1993; Fedorov and Hackl, 1997; Arbia and Lafratta, 1997; Nychka and Saltzman, 1998; Müller, 2007). However, given that most environmental processes are dynamic, static designs will not be as efficient as designs that are allowed to evolve over time. Wikle and Royle (1999) have used a space-time Kalman filter based on a measurement equation and a state equation in order to design a mobile network for the ozone monitoring in the Chicago area.

In the meteorological context, adaptive (or targeted) observations strategies in connection with a data assimilation system were used in order to reduce the uncertainty in the forecast of a numerical weather prediction system (e.g. Daescu and Navon, 2004; Carrassi et al., 2007). Berliner et al. (1999), Fourrié et al. (2006) (and references therein) have introduced several methodologies of adaptive observations that have been tested during the Fronts and Atlantic Storm-Track Experiment (FASTEX), or in the framework of The Observing System Research and Predictability Experiment (THORPEX). Although very successful, these techniques are of limited interest because of the number of adaptive observations compared to the constant flux of synoptical meteorological observations.

The present work is intended to investigate the possibility, in the event of a nuclear or radiological emergency, to improve our knowledge about radioactive materials spread around a nuclear power plant by using a targeting monitoring strategy feeding a sequential data assimilation scheme. This goes beyond the simpler objective consisting in trying to obtain the best mapping of the contaminant field. Quantitatively, the goal is to optimally reduce the uncertainty of the current analyses and predictions.

To do so, we will perform an observing system simulation experiment [OSSE], in a 100 km $\times$ 100 km domain around the Bugey nuclear power plant. In Section 2, the complete setup about the OSSE is described: the meteorological model, the chemistry and transport models, the existing monitoring network around the plant, the local meteorological conditions. A typical release rate sequence of an accident (the so-called *source term*) will be described. In Section 3, the data assimilation scheme, that relies on inverse modelling, is detailed. A discussion on error modelling is then needed. Examples of inverse modelling are given by twin experiments with realistic errors. In Section 4, our targeting scheme is presented. Details are given on the mathematical and numerical aspects. In Section 5, the targeting strategy is illustrated on several experiments built on the OSSE. The benefits of observation targeting are discussed. Conclusions are drawn in Section 6.

## 2. Setup

### 2.1. Meteorological input data

In this study, the mesoscale atmospheric model MM5 (Anthes and Warner, 1978) is used to drive a chemistry-transport model by an off-line coupling. For the present study, MM5 is implemented with three nested domains with horizontal resolution 18, 6 and 2 km. The outer first domain covers entirely France with other major parts of the surrounding countries. The second one is mainly fixed on the southeast region of France whereas the inner third domain covers an area of 40,000 km$^2$ around the Bugey nuclear site between 4°01′E, 44°55′N and 6°31′E, 46°41′N. This nesting design is depicted in Fig. 1. All domains have 100 $\times$ 100 grid points and contain 31 vertical $\sigma$ levels of which nine levels (pressure levels 990, 980, 970, 950, 925, 900 hPa) are chosen near to the surface. The initial meteorological conditions are taken from the 1° $\times$ 1° Final Analysis (FNL) data of the National Center for Environmental Prediction (NCEP). The model is initialised at 00:00 UTC on 01 Jun 2007 and 20 Dec 2007, then is integrated for one week for each of the two starting dates. At first we performed one week run with only two domains nesting for the outer domain and the second one (18 and 6 km horizontal resolution). Terrain elevation, and land use data are provided by the United States Geological Survey. For the second nest, these data are incorporated at a horizontal resolution of 0.9 km (for the second and third domains the topography data resolution is 0.9 km). The analysis nudging four-dimensional data assimilation (FDDA), which blends FNL analysis (the first guess) with observations coming from NCEP operational global surface and upper-air observations subsets, is performed on the first and seconds domains to produce additional mesoscale analyses.

### 2.2. The chemistry-transport model and parametrisations

The numerical chemical-transport model used in this study is Polair3D, part of the Polyphemus platform. It has been validated on

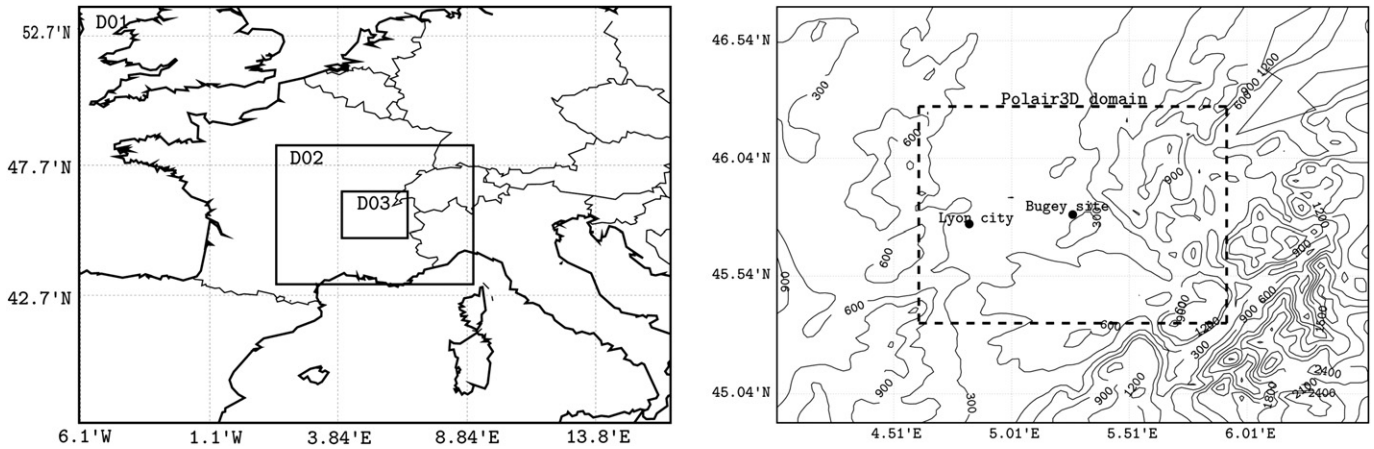


**Fig. 1.** MM5 nesting design: the three nested domains are represented on the left. The third domain D03 is enlarged on the right, where the domain chosen for the transport model is plotted, and where the local topography is shown.

the European Tracer Experiment, Algeciras and Chernobyl (Quélo et al., 2007). As discussed later, iodine-131 will be chosen as the species of interest representative of a typical dispersion in case of a nuclear accident. The model integrates the concentration field $c$ of $^{131}$I, following the transport equation

$$\frac{\partial c}{\partial t} + \mathrm{div}(\boldsymbol{u}c) = \mathrm{div}\left(\rho \boldsymbol{K} \nabla \left(\frac{c}{\rho}\right)\right) - \Lambda^s c - \Lambda^d c + \sigma \quad (1)$$

where $\Lambda^s$ is the wet scavenging coefficient, $\Lambda^d$ represents the radioactive decay and $\sigma$ is the point-wise source for $^{131}$I. $\boldsymbol{u}$ is the wind velocity, $\rho$ is the air density, and $\boldsymbol{K}$ is the matrix of turbulent diffusion, diagonal in practice. Later on, these fields will be obtained or diagnosed from MM5 simulations. The vertical component is given by $K_z$. Its parametrisation follows Louis (Louis, 1979). The horizontal component is of uniform and constant value $K_H$. All boundary conditions are taken null (in particular incoming mass fluxes are null). The boundary condition on the ground is

$$-K_z \nabla c \cdot \mathbf{n} = -v^{\mathrm{dep}} c \quad (2)$$

where $\mathbf{n}$ is the unitary vector upward oriented. $v^{\mathrm{dep}}$ is the dry deposition velocity of $^{131}$I.

The advection is implemented thanks to a third-order explicit scheme, with a Koren–Sweby flux limiter function. Because of the sharp gradients found, it is important that such a limiter be used. The diffusion scheme is integrated through an implicit second-order Rosenbröck scheme, with a three-step spatial scheme.

Iodine is deposited using a simple scheme with a constant deposition velocity: $v^{\mathrm{dep}} = 0.5$ cm s$^{-1}$. The parametrisation used in this study is Belot of the form $\Lambda^s = a p_0{}^b$, with $a = 8 \times 10^{-5}$ and $b = 0.8$. $p_0$ is the rain intensity, in mm h$^{-1}$ (Baklanov and Sorensen, 2001).

The transport model Polair3D is then integrated in the domain surrounding the Bugey site with a spatial resolution of 2 km by 2 km. The number of grid points in the domain simulation is $50 \times 50$ (100 km $\times$ 100 km) and the model has 11 vertical levels ranging from 0 to 3500 m.

### 2.3. Monitoring networks around the Bugey site

The nuclear power plant of Bugey belongs to EDF. EDF has the obligation to detect and measure any presence of radionuclides activity recorded in the immediate vicinity of the site. The gamma ambient dose rate is automatically measured at a total of 31 locations, all within a radius of 22 km from the nuclear power plant.

These locations are distributed in the following manner: 1) a network of 10 gamma-dose monitoring stations is setup in the closest border of the site over a range of less than 1 km from the site; 2) four additional gamma-dose monitoring stations are installed within a radius of 1 km around the power plant; 3) four gamma-dose stations are over a range of 6 km from the site; 4) a network of 13 monitoring gamma-dose stations are located within a range of 22 km from the site. Besides, three gamma-dose monitoring stations from the Teleray Network which is operated by the IRSN, are added to the 31 monitoring stations. These stations are located within a radius of 18 km from the site. Four other Teleray stations will be added to the set of monitoring stations within a radius 50 km from the site, as our dispersion domain encompasses them. Eventually, 38 monitoring stations are surrounding the site.

In routine monitoring situations, the measured gamma ambient dose rate are transmitted in real time at 10-min intervals to the IRSN. The frequency of data acquisition can be lowered down to a range of 1–5 min in an emergency situation. These monitoring stations are equipped with an alarm system. If a critical threshold value of gamma ambient dose rate is exceeded, the alarm is raised immediately and the IRSN issues a warning to the national and local authorities. Within a radius of 12 km around the Bugey power plant, 109 monitoring locations (PPI points) were identified by the IRSN as potential sites where the IRSN can rapidly deploy monitoring stations in order to operate supplementary measurements. These points are located in open terrains to be easily accessed. The PPI point and the monitoring stations are depicted on Fig. 2. Table 1 summarises additional information on the monitoring stations around the Bugey nuclear site.

During this study we have hypothesised that all the monitoring stations installed around the Bugey power plant are homogeneous, in the sense that measurements are performed with the same type of instrument, which gives volumetric activity concentrations over a range of $10^{-5}$–$10^{9}$ Bq m$^{-3}$ instead of gamma ambient dose. This assumption simplifies the problem considerably since the assimilation of ambient dose would additionally require a spatial model integration, and would require to properly account for contamination effects of dose instruments.

### 2.4. Local meteorological conditions at the Bugey nuclear power plant

The study area is focused around the Bugey nuclear power plant within a radius of 50 km (Fig. 1). The plant is located in the Ain plain at

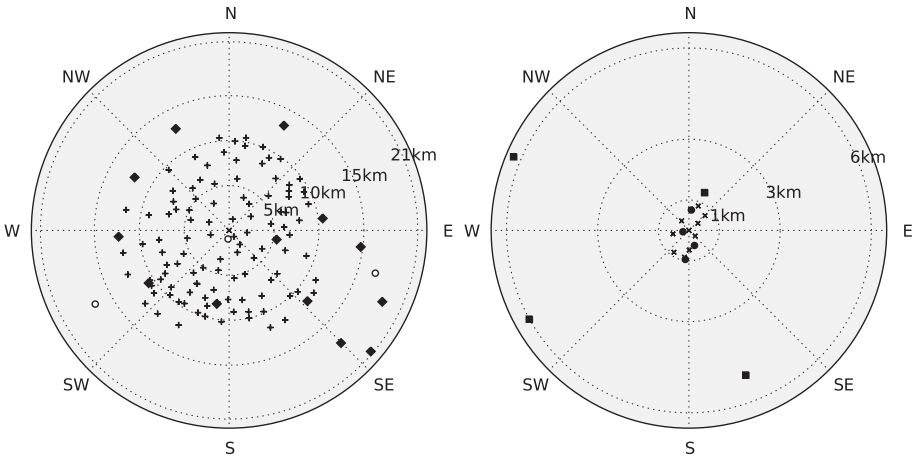


**Fig. 2.** Locations of gamma-dose monitoring stations and PPI points around the Bugey power plant. Diamonds (◆) are gamma-dose stations. Plus signs (+) correspond to PPI points (virtual stations). White disks (○) depict three gamma-dose stations from the Teleray network. On the right: enlargement emphasising the border sondes around the site. The squares (■) denote close gamma-dose stations. Crosses (×) mark the site border stations. Black bullets (●) correspond to balises within a radius of 1 km.

5°16′E, 45°47′N. It is 30 km from the Swiss border and about 35 km away from Lyon city in the western side. It is on the edge of Rhône river, from where the site gets its cooling water. Moreover, the site is inside a basin surrounded by mountains in the SW–NW (Massif central) and by the Alps and Jura mountains in the SE–E, NE respectively.

As a consequence of this complex topography around the site (within a radius of 100 km) the local meteorological conditions at Bugey site on mesoscale are non-stationary and non-homogeneous (Zhao et al., 1983). Therefore a wide range of dispersion conditions is possible on a single day scale. Also, at synoptic scale, many meteorological configurations could affect extensively the micro-meteorology around the site. A typical frequent situation is the Mistral wind, which blows from the northwest or north of France through the Rhône Valley river.

Fig. 3 depicts the windrose of Bugey nuclear power plant at 100 m height over the period 1995–2002. As we can notice, the wind is blowing mainly from two different directions: the first direction which is the most common wind direction is from the north–north-west (slightly more than 15% of time), the second one is from the south–southeast (more than 10% of time). The average wind speed values within these two directions are respectively of 5.1 m $s^{-1}$ and 5.3 m $s^{-1}$. The reason to have a wind direction confined essentially into two opposite directions is the dynamic effect of the mountains (Alpes and Massif Central) surrounding the site.

The efficiency of a targeting design following a radioactive release occurring at the Bugey site will be strongly constrained by local wind conditions. In case of strong wind speeds (more than 5 m $s^{-1}$) a radioactive leak from the site will move swiftly and reach the border of the targeted area within less than 3 h. Therefore, in such an extreme situation, the setup of a mobile network within this domain, may start losing its interest, because the plume may be out of reach before any effective deployment. Relying on the level of emergency preparedness of the IRSN, we estimated that the necessary intervention time needed for installing a mobile station (by sending vans or helicopters) within a radius of 50 km around the Bugey site is about 1 h. Conversely, at weak and normal wind speeds (0.5 m $s^{-1}$–2 m $s^{-1}$–4 m $s^{-1}$) the radioactive plume will stay longer inside the targeting area (more than 4 h) before reaching the border.

For the sake of demonstrating the potential of a targeting approach with respect to different local wind conditions, and by investigating meteorological data of a synoptical tower installed near the site, we have chosen two different meteorological situations around the Bugey nuclear site. The first lasts one week starting 1 June 2007. It is characterised by moderate to strong wind speeds. The peculiar feature of this situation is that the wind direction changes considerably, so that the radioactive plume rotates around the site. The second one lasts one week from 20 December 2007. The surface wind field speed is weak to moderate. Fig. 4 depicts temporal profile of the surface wind module (at 10 m height) calculated by MM5 model on the Bugey nuclear power plant.

### 2.5. Source term

The *source term* is a technical expression used to describe the accidental release of radioactive material from a nuclear facility to the environment. Not only are the levels of radioactivity released important, but also their distribution in time as well as their chemical and physical forms.

The estimation of a release scenario pattern over time is a complex task, depending on the nature of the accident occurring at a nuclear facility. It is characterised by a release rate as a function of time for each radionuclide and the parameters relevant for the dispersion in the environment (such as position, height, energy content and volumetric flow of the release). Aerosol size and chemical form have a direct impact on the deposition and plume

**Table 1**
Specifications of the monitoring stations around Bugey site.

| Name | Type | Unit | Range | Frequency | Number of stations |
|---|---|---|---|---|---|
| Balise Teleray | Genitron | Gamma-dose | $5.10^{-9}$–10 Gy $h^{-1}$ | 10 min | 7 |
| Sonde cloture | Genitron | Gamma-dose | $10^{-8}$–$10^{-2}$ Gy $h^{-1}$ | 10 min | 10 |
| Balise 1 km | Genitron | Gamma-dose | $10^{-8}$–$10^{-4}$ Gy $h^{-1}$ | 10 min | 4 |
| Balise 5 km | Genitron | Gamma-dose | $10^{-8}$–0.5 Gy $h^{-1}$ | 10 min | 4 |
| Sonde 10 km | Genitron | Gamma-dose | $10^{-8}$–$10^{-2}$ Gy $h^{-1}$ | 10 min | 13 |
| PPI point | – | – | – | – | 109 |

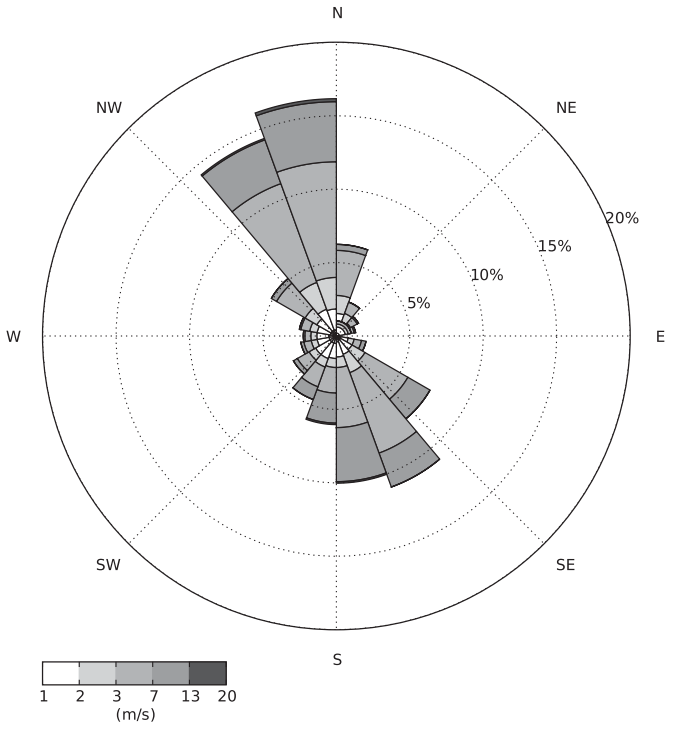


**Fig. 3.** Windrose at the Bugey nuclear power plant over the period 1995–2002.

depletion. Here in France, the emergency plans of IRSN contemplate two types of accidents. The first is a severe nuclear reactor accident consisting of a core meltdown, but without major hull breach. The second one, of lesser impact, is a breakdown of a vapour generating tube, leading to a preventive release into the atmosphere of overpressure through an exhaust pipe.

In this study we have considered a hypothetical scenario of a core meltdown occurring at the Bugey plant. From the radiological point of view, the radionuclides $^{131}$I, $^{137}$Cs are the most important radionuclides to consider, because they are responsible for most of the radiation exposure received by the general population. As mentioned earlier, species $^{131}$I was selected for this study.

Relying on data of emission rates (ground release) of different isotopes issued by the IRSN from internal estimates of the release of radioactivity to the environment under a hypothetical core meltdown accident at the Bugey site, a simplified temporal profile of $^{131}$I during one week of release is assumed to take the form depicted on Fig. 5.

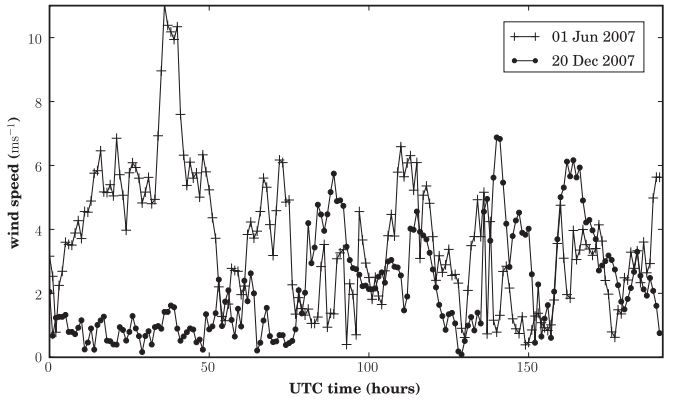


**Fig. 4.** Surface wind speed temporal profile for one week MM5 run at Bugey site, for the two meteorological configurations.

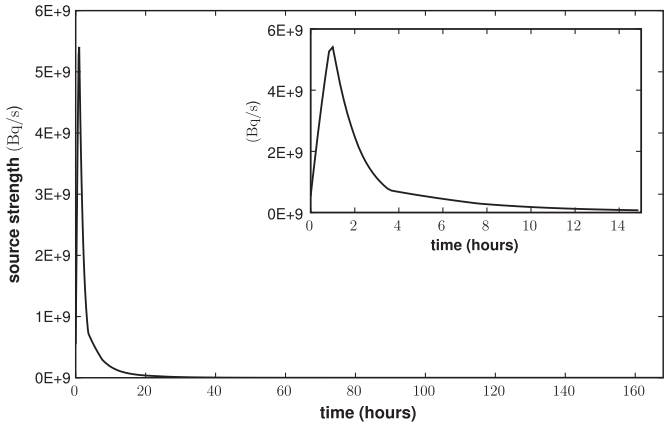


**Fig. 5.** A hypothetical source term temporal profile for $^{131}$I. The inset zooms on the first 14 h.

Since no hull breach is hypothesised, the release height is chosen to be 15 m. Heat plume rise is neglected here merely because it entails too many unknown parameters. Besides, the source height or local rise is not that significant for the demonstration of the paper. However this issue would matter if the plume rises swiftly above the boundary layer, since the information from ground detectors might not be enough for a reliable data assimilation.

## 3. Data assimilation

The targeting which will be developed should provide observations to be assimilated in the forecast of the plume. The main task of the assimilation is to identify the source and correctly estimate its release rate, as it is the main forcing of the plume (Politis and Robertson, 2004; Bocquet, 2007). Any other type of uncertainty (such as meteorological field errors) is secondary. That is why our inverse modelling procedure will be described first.

### 3.1. Inverse modelling

The source–receptor equation relates the activity concentration measurement $\mu_i$ (in Bq m$^{-3}$) to the source field $\sigma(x, t)$, thanks to the solution of Eq. (1). In a discretised form, it reads

$$\boldsymbol{\mu} = \boldsymbol{H}\boldsymbol{\sigma} + \boldsymbol{\varepsilon}. \tag{3}$$

$\boldsymbol{\mu} \in \mathbb{R}^d$ stands for the measurements vector. $\boldsymbol{\sigma} \in \mathbb{R}^N$ is the discretised source vector. We shall denote $\Delta t$ its discretisation time step. $\boldsymbol{H}$ is the Jacobian or source–receptor matrix. It embodies the model of dispersion as well as the observation operator. The vector error $\boldsymbol{\varepsilon} \in \mathbb{R}^d$ represents a combined model-representativity-instrumental error. Eventually, the inverse problem to be resolved is completely summarised by this equation. The computation of $\boldsymbol{H}$ will be detailed in the next section.

### 3.2. Control parameters and source–receptor matrix computing

The location of the source is in the centre of a grid cell which itself is located in the centre of the spatial grid of the experiment. Contrary to Krysta and Bocquet (2007) and Krysta et al. (2008), but similarly to Davoine and Bocquet (2007), the location of the release is supposed to be known since an alert is assumed to have been issued after the start of the release, or preventively. Therefore, the control parameter space is reduced to a one-dimensional grid along time.

The time resolution $\Delta t$ for source inversion must be defined according to the space and time scales of the phenomenon, and has

to be enough to reconstruct the source term with a reasonable precision. Yet, the precision inversion is constrained by both the availability and the quality of the observations.

Ideally, the time resolution could be chosen equal to the data acquisition time interval of the monitoring networks installed around the nuclear Bugey site. In an emergency situation, the frequency of data acquisition of the monitoring stations can be reduced to 1 min. However, to avoid that the inverse problem be ill-posed, the finest time resolution should still integrate a sufficient number of data.

The finest time resolution $\Delta t$ is defined here to be 10 min, which corresponds to the time interval data acquisition in a routine situation. Also, the time resolution must take into account the variability of the ambient meteorological conditions. It is therefore irrelevant to resolve the source term with a time interval much lower than 1 h, which is the frequency of MM5 input data. That is why a time resolution of 30 min is chosen here. Hence, the number of variables to be inverted over a period of one week will be $N = 336$. Yet, the number of measurements over the same period, gathered by a monitoring network of $n$ stations at 10 min interval, will be much greater than the number of unknowns, whatever their quality or their relevant information content to this inverse problem. Consequently, in our case the inverse problem turns out to be well-posed.

In general the matrix $\boldsymbol{H}$ can be computed by two methods. The first one is to calculate the source–receptor matrix row-by-row. In this case, one calculates the adjoint solution of each measurement $\mu_i$. That may be the case when the number of measurements is smaller compared to the number of variables to resolve. The second one is to proceed column-by-column. Each column is a solution of the transport equation for an elementary source represented in the canonical basis of $\mathbb{R}^N$. Since there is a greater number of measurements than unknown variables here, this latter method is retained. In our case, each elementary source has a temporal extent of $\Delta t = 30$ min. Thus, we have calculated the matrix $\boldsymbol{H}$ by integrating Polair3D model $N = 336$ times.

### 3.3. Solving the inverse problem

Using Gaussian assumptions on the errors and background information on the source, elementary Bayes inference leads to the following cost function (see for instance Krysta et al., 2006, in a very similar context):

$$\mathcal{L}(\boldsymbol{\sigma}) = \frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{H\sigma})^T \boldsymbol{R}^{-1}(\boldsymbol{\mu} - \boldsymbol{H\sigma}) + \frac{1}{2}\boldsymbol{\sigma}^T \boldsymbol{B}^{-1}\boldsymbol{\sigma}, \tag{4}$$

where $\boldsymbol{R}$ is the observation error covariance matrix. $\boldsymbol{B}$ is the background error covariance matrix which contains known prior statistical information on the source and is necessary to regularise ill-posed problems. Here, however, the information content of the observations is supposed to be sufficient to resolve the unknown variables. That is why the background term is dropped, and the cost function takes the simpler form

$$\mathcal{L}(\boldsymbol{\sigma}) = \frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{H\sigma})^T \boldsymbol{R}^{-1}(\boldsymbol{\mu} - \boldsymbol{H\sigma}), \tag{5}$$

which leads to the estimated source term $\overline{\boldsymbol{\sigma}}$ given by (generalised inverse)

$$\overline{\boldsymbol{\sigma}} = \left(\boldsymbol{H}^T \boldsymbol{R}^{-1}\boldsymbol{H}\right)^{-1} \boldsymbol{H}^T \boldsymbol{R}^{-1}\boldsymbol{\mu}, \tag{6}$$

with an error covariance matrix $\boldsymbol{R}$ which is assumed diagonal.

### 3.4. Error modelling

In the absence of specific measurement campaigns, assumptions must made on errors. In this study, two forms of errors are considered: Gaussian and non-Gaussian errors. In order to obtain *realistic* generated measurement sets which will be considered as the *truth*, we have firstly generated a set of synthetic measurements spanning over the experiment one week duration, and this, for all the monitoring stations around the Bugey site. This set is derived from the reference simulation, or via the noiseless source-measurement equation

$$\boldsymbol{\mu}^{\mathrm{s}} = \boldsymbol{H\sigma}^{\mathrm{t}}, \tag{7}$$

where $\boldsymbol{\sigma}^{\mathrm{t}}$ is the true source (see Fig. 5).

In order to take into account realistically important dispersion modelling errors, the concentrations measurements error are assumed to follow a log-normal law. More than instrumental errors, they are meant to account for representativeness errors. The perturbed observations are $\mu_i = r_i \mu_i^{\mathrm{s}}$, where $r_i$ follows a log-normal law:

$$r_i \sim \exp(\mathcal{N}(0, \lambda_l^2)). \tag{8}$$

where $\lambda_l$ is the standard deviation of the related Gaussian law $\mathcal{N}(0, \lambda_l^2)$. It is assumed here that the perturbed measurements have an unbiased median (as many overestimations as underestimations on average), whereas their mean is biased by a factor $\exp(\frac{1}{2}\lambda_l^2)$, which could easily be removed.

In the absence of information about the true errors, prior statistics are supposed to be additive and multiplicative Gaussian. They are respectively of the forms: $\boldsymbol{\mu}^{\mathrm{s}} + \mathcal{N}(\boldsymbol{0}, \lambda_g^2 \boldsymbol{I})$ and $\boldsymbol{\mu}^{\mathrm{s}} + \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$, where $\lambda_g$ is the standard deviation of the additive error, $\boldsymbol{I}$ is the identity matrix in observation space, and $\boldsymbol{R}$ is a diagonal error covariance matrix, whose entries $\chi_i$ are $(r\mu_i^{\mathrm{s}})^2$, with $r$ a factor to be chosen a priori. Note that in both cases, observation error is additive in the sense of Eq. (3). Multiplicative means here that the error standard deviation is a fraction of the measurement.

A Gaussian additive prior with a uniform variance ($\chi_i = \chi$) grants high confidence to greater measurements, and lesser confidence to smaller measurements. In that case the inverse problem estimate will not be sensitive to the variance $\chi$ (see Eq. (6)). A Gaussian multiplicative error prior was therefore chosen instead. The variances $\chi_i$ are chosen then to be a priori a fraction of the measurement $\mu_i^{\mathrm{s}}$.

Another possible prior, more consistent with our modelling of the perturbations, is to choose a log-normal one. However, this non-Gaussian approach will not be easily generalised to the targeting mathematical formalism. We have nevertheless performed inversion tests with the related prior. The corresponding multivariate probability density $p(\boldsymbol{\mu}|\boldsymbol{\sigma})$ associated with the observed measurements $\boldsymbol{\mu}$ given the source term $\boldsymbol{\sigma}$, follows:

$$p(\boldsymbol{\mu}|\boldsymbol{\sigma}) = \prod_{i=1}^{d} \frac{1}{\mu_i \lambda_l \sqrt{2\pi}} \exp\left(-\frac{(\ln\mu_i - \ln[\boldsymbol{H\sigma}]_i)^2}{2\lambda_l^2}\right). \tag{9}$$

Consequently, the cost function related to this likelihood is

$$\mathcal{L}(\boldsymbol{\sigma}) = \sum_{i=1}^{d} \frac{(\ln \mu_i - \ln[\boldsymbol{H\sigma}]_i)^2}{2\lambda_l^2}. \tag{10}$$

The retrieval of the source term $\boldsymbol{\sigma}$ is carried out through the minimisation of the cost function Eq. (10), using the L-BFGS-B limited-memory quasi-Newton minimiser (Liu and Nocedal, 1989).

To illustrate the choice of priors in the case of the log-normal generated errors (with a realistic $\lambda_l = 0.5$), a set of $10^5$ source inversions was performed. Fig. 6 shows the source profile retrieval density sampled over $10^5$ reconstructed sources, for several priors: Gaussian additive and uniform, Gaussian multiplicative and measurement-dependent, and log-normal. The relative error is set to 50% (i.e. $r = 0.5$ in rough accordance with $\lambda_l = 0.5$). The Gaussian additive inversions demonstrate a large overestimation of the true

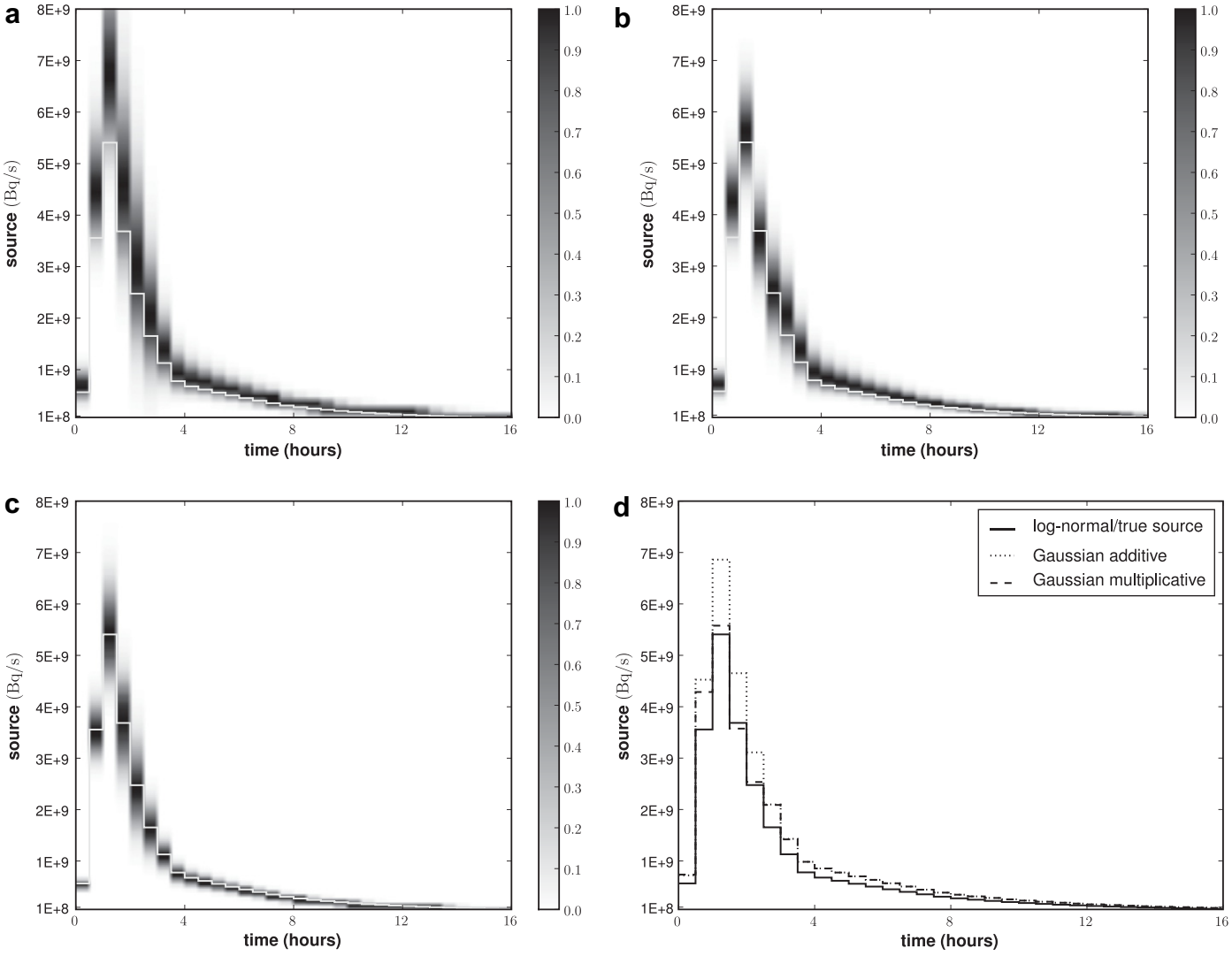


**Fig. 6.** The top left panel (a) displays the result of $10^5$ inversions with log-normal perturbations ($\lambda_l = 0.5$) and a Gaussian additive and uniform prior. The top right panel (b) displays the result of $10^5$ inversions with log-normal perturbations and a Gaussian multiplicative measurements-dependent prior. The bottom left panel (c) displays the result of $10^5$ inversions with log-normal perturbations and a log-normal prior. In each case, the density of the inversion results is displayed (shades of grey), the true source is represented by a light grey line. On the bottom-right panel (d), are plotted the median inversion estimates for these three configurations and the true source.

source, consistent with the existence of the mean bias. The Gaussian multiplicative inversions demonstrates a moderate overestimation of the true source. Ultimately we will choose the latter for our prior since the overestimation is less pronounced. Because it is Gaussian, it remains easily analytically tractable, contrary to the log-normal prior case. Note that for emergency centres, a moderate overestimation in the data assimilation analysis is not a flaw because the main danger is to underestimate the risk.

### 3.5. Plume data assimilation

By using a sequential data assimilation approach, data from the Bugey surrounding network can be assimilated in order to improve, not only the source term, but also the dispersion model predictions and to obtain a real-time estimation of the dispersion plume.

In this section the data assimilation scheme used to reconstruct the plume is described. It has been tested using partial or all observations on the fixed Bugey network. The assimilation interval is noted $\Delta t_{\mathrm{f}}$, and we took it to be the same as the forecast interval. It has also been tested for several values $\Delta t_{\mathrm{f}} = 1, 2, 3$ h.

In an emergency situation, and at a current time $t_n$ the sequential approach algorithm reads:

1. *Preprocessing step*: All available data in the interval $[t_n - \Delta t_{\mathrm{f}}, t_n]$ are collected. Screening and quality control should be applied to this set of measurement in order to eliminate for example the artificial gamma-dose rates. This does not arise in our case because we supposed that all observations data are *true*. Subsequently, the elementary solutions corresponding to this time interval are computed using the chemistry and transport model. Since the source time resolution is half an hour, then we have to compute $2\Delta_{\mathrm{f}}$ elementary solutions and store them.
2. *Analysis step*: We construct the source–receptor matrix $\boldsymbol{H}_{t_n}$ associated to the stationary monitoring network. This matrix is computed column-by-column using all the elementary solutions calculated from $t_0$ till $t_n$. The measurements vector $\boldsymbol{\mu}_n$ containing all the perturbed observations between $t_0$ and $t_n$ are built up. Then an estimation of the source term $\boldsymbol{\sigma}_n$ is obtained by using Eq. (6).
3. *Forecast step*: A forecast of $\Delta t_{\mathrm{f}}$ hours from $t_n$ to $t_{n+1} = t_n + \Delta t_{\mathrm{f}}$ is performed using the chemistry and transport model. For this,

we carry out a forward simulation from $t_0$ to $t_{n+1}$ forced by the source $\boldsymbol{\sigma}_n$ between $t_0$ and $t_n$, and by the last retrieved component of $\boldsymbol{\sigma}_n$, between $t_n$ and $t_{n+1}$. This means that we supposed that the source term is persistent. This hypothesis sounds more adequate in case of an accidental radioactive releases. Indeed, given the shape of typical radiological source terms, persistence mostly leads to an overestimation of the radionuclides released. This is considered a safer assumption by risks agencies, such as the IRSN.

The performance of data assimilation will be assessed later altogether with the targeting scheme.

## 4. Targeting

This section develops a formalism for the targeting of observations suited to the emergency context presented in the introduction.

### 4.1. Targeting theory

We have seen that all monitoring stations lie within a radius of 22 km from the nuclear facility site. Thus, in case of emergency, this fixed network can become less informative on the evolution of the plume in regions away from the source site. This situation could occur if the wind fields are strong enough and advect swiftly radioactive materials far from the source.

In such situations, deploying mobile measuring devices constitutes an adequate monitoring strategy that allows to follow the spatiotemporal evolution of the plume. Here, the collected measurements from the mobile network and the static monitoring network will be both assimilated in order to enhance our knowledge about the radioactive plume. This raises the question as to where the mobile monitors should be located at time $t+1$ given the information based on measurements collected up to time $t$. The targeting design consists in seeking the spatial locations at time $t+1$ that optimise some design criterion.

In the present context, a targeting design schedule needs a decision time $t_1$. Given all available information from $t_0 = 0$ the starting time of the accident up to time $t_1$, we decide how to collect additional measurements at time $t_2 > t_1$ so as to reduce the uncertainty on the future state of the radioactive plume. This decision should be taken according to the design criterion. At time $t_2$, the design will be carried out. Mobile measuring devices will have been deployed, and the new observations will be assimilated to get a better forecast, following Section 3. From $t_2$ on, a forecast step is performed till $t_3 > t_2$. But $t_2$ is also a decision time for a deployment at $t_3$. The schedule goes on following these steps till the end of the experiment.

### 4.2. Nature of the design criterion

In this study our targeting design problem is defined in order to have both an efficient estimation of the source term at time $t_2$ and to get the most precise plume forecast. As suggested earlier, a necessary condition to get a reasonable plume forecast is to get a good estimate of the source term. Thus the design criterion may consist in reducing the uncertainty on the source term analysis, that is minimising a measure of the source error covariance matrix (rather than the forecast error covariance matrix at time $t_3$).

The alternative is to perform the design in order to reduce the uncertainty on the plume activity concentrations at time $t_2$ (or possibly reduce a forecast uncertainty at time $t_3$). Because the source term and the concentration field are linearly related, the source error covariance matrix can easily be related to the concentration error covariance matrix.

Let $\mathcal{H}$ be the total Jacobian matrix linking the source to the concentration field on the domain computed over the whole POLAIR3D domain grid. The source–receptor matrix associated to both static and mobile networks is denoted by $\boldsymbol{H}$. Let $\overline{\boldsymbol{\sigma}}$ denotes the reconstructed source term from the set of measurements, and $\boldsymbol{\sigma}^t$ be the *true* source term. Denote respectively by $\overline{\boldsymbol{c}}$ and $\boldsymbol{c}^t$ the analysed field activity concentrations and its true state. The covariance matrix error of the source term and of the plume concentrations are given respectively by:

$$\boldsymbol{P}^{\sigma} = \mathrm{E}\left[(\overline{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^t)(\overline{\boldsymbol{\sigma}} - \boldsymbol{\sigma}^t)^T\right],\ \boldsymbol{P}^{c} = \mathrm{E}\left[(\overline{\boldsymbol{c}} - \boldsymbol{c}^t)(\overline{\boldsymbol{c}} - \boldsymbol{c}^t)^T\right] \tag{11}$$

where E indicates the expectation, so that

$$\boldsymbol{P}^{c} = \mathcal{H}\boldsymbol{P}^{\sigma}\mathcal{H}^T + \boldsymbol{Q} \tag{12}$$

where the $\boldsymbol{Q}$ term denotes the model error covariance matrix, which is neglected in this study. This equation tells how the uncertainty on the source term propagates to the concentrations space via the full Jacobian $\mathcal{H}$ operator. Note that $\boldsymbol{P}^{c}$ depends on the station locations through $\boldsymbol{P}^{\sigma}$ (which depends on $\boldsymbol{H}$). The total Jacobian matrix $\mathcal{H}$, that maps $\boldsymbol{P}^{\sigma}$ to $\boldsymbol{P}^{c}$ does not add extra dependence on the network. Therefore, apart from the impact of model error that is possibly design dependent, a good criterion based on the source analysis would qualitatively yield a good design on the concentration field analysis. Obviously, the precise correspondence between the two criteria depends on how dispersion is amplifying the uncertainty.

### 4.3. Decision time analysis

At the design decision time (say $t_1$), the future plume status at $t_2$ and even more the future source term on $[t_1, t_2]$ are unknown, as well as the future observations. We can only forecast them, or make an assumption for the source (persistence here).

Using all available information from $t_0$ up to $t_1$ through the data assimilation process, our best knowledge about the source term is given by a Gaussian probability distribution,

$$\boldsymbol{\sigma}_{t_1^+} \sim \mathcal{N}\left(\overline{\boldsymbol{\sigma}}_{t_1^+}, \boldsymbol{P}^{\sigma}_{t_1^+}\right), \tag{13}$$

where the $+$ sign means at time $t_1$ after the analysis. We now define a forecasted source $\boldsymbol{\sigma}^{\star}_{t_2}$ which is equal to the analysed source term $\overline{\boldsymbol{\sigma}}_{t_1^+}$ on $[t_0, t_1]$, and is defined on $[t_1, t_2]$ using the persistence assumption. A plume forecast is then carried out from $t_1$ to $t_2$ by integrating the transport model from the time $t_0$ until $t_2$ and using this source. A complete set of observations is then obtained: all observations up to $t_1$ and the forecasted fixed observations in $[t_1, t_2]$. Hence, the forecasted source covariance matrix error at $t_2$ is given by

$$\boldsymbol{P}^{\star\sigma}_{t_2} = (\boldsymbol{H}^T_{t_2}(\boldsymbol{R}^{\star}_{t_2})^{-1}\boldsymbol{H}_{t_2})^{-1}, \tag{14}$$

where $\boldsymbol{H}_{t_2}$ is the source–receptor matrix for these observations and the source term up to time $t_2$. The future targeted observations at $t_2$ are not included yet. $\boldsymbol{R}^{\star}_{t_2}$ is a measurement covariance matrix error. This matrix combines the errors obtained from the actual measurements available from $t_0$ to $t_1$ and from the forecasted concentrations from $t_1$ to $t_2$, and mapped to the locations of the monitoring stations. Recall that the measurement errors standard deviations are taken as a prior fraction of the measurements amplitude, and that $\boldsymbol{R}^{\star}_{t_2}$ is assumed diagonal. The $\star$ symbol qualifies a parameter related to a *forecasted* analysis. Note that with an inversion resolution of $\Delta t$, the matrix $\boldsymbol{H}_{t_2}$ is made of $(t_2 - t_0)/\Delta t$ elementary solutions (columns). In principle this source–receptor matrix is also forecasted for the $[t_1, t_2]$ part. But, in the absence of model error, the forecasted $\boldsymbol{H}^{\star}_{t_2}$ coincide with $\boldsymbol{H}_{t_2}$, that is effectively obtained at $t_2$ before the inclusion of the targeting observations.

The (forecasted) targeted observations $\boldsymbol{\mu}^{\star}_{t_2}$ taken from mobile monitors that will be installed at time $t_2$, abide to

$$\boldsymbol{\mu}^{\star}_{t_2} = \boldsymbol{h}\boldsymbol{\sigma}^{\star}_{t_2} + \boldsymbol{\varepsilon} \tag{15}$$

where $\boldsymbol{h}$ is a sub-matrix of the Jacobian matrix $\mathcal{H}$ corresponding to the additional observations, and $\boldsymbol{\varepsilon}$ is the measurement error vector, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{r}^{\star}_{t_2})$. The error covariance matrix of the targeted observation $\boldsymbol{r}^{\star}_{t_2}$ is assumed diagonal (uncorrelated observations). Note that each line of matrix $\boldsymbol{h}$ corresponds to a targeted observation.

At decision time $t_1$, the forecasted update of $\boldsymbol{\sigma}$ at time $t_2$, including the assimilation of the forecasted targeted observations, follows

$$\boldsymbol{\sigma}^{\star}_{t_2^+} \sim \mathcal{N}\left(\overline{\boldsymbol{\sigma}}^{\star}_{t_2^+}, \boldsymbol{P}^{\star\sigma}_{t_2^+}\right). \tag{16}$$

Standard Bayesian analysis leads to

$$\begin{aligned}\boldsymbol{P}^{\star\sigma}_{t_2^+} &= \left((\boldsymbol{P}^{\star\sigma}_{t_2})^{-1} + \boldsymbol{h}^T(\boldsymbol{r}^{\star}_{t_2})^{-1}\boldsymbol{h}\right)^{-1} \\ &= \boldsymbol{P}^{\star\sigma}_{t_2} - \boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\left(\boldsymbol{r}^{\star}_{t_2} + \boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\right)^{-1}\boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2},\end{aligned} \tag{17}$$

$$\overline{\boldsymbol{\sigma}}^{\star} = \overline{\boldsymbol{\sigma}}^{\star}_{t_2} + \boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\left(\boldsymbol{r}^{\star}_{t_2} + \boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\right)^{-1}\left(\boldsymbol{\mu}^{\star}_{t_2} - \boldsymbol{h}\overline{\boldsymbol{\sigma}}^{\star}_{t_2}\right). \tag{18}$$

It may seem that the stationary observations, then the adaptive ones were sequentially assimilated. But it is actually equivalent to assimilating them altogether, thanks to the assumed independence of the observation errors. It is judicious to carry out the splitting because, in the optimisation scheme (Section 4.6), Eq. (20) is often evaluated with different mobile stations configurations. Thanks to this splitting, the inverse operator is defined in the smaller space of the adaptive observations rather than in the full observation space. This is the key to the numerical efficiency of the scheme.

### 4.4. Quantitative design criterion

The targeting design problem consists in finding optima of functions of $\boldsymbol{P}^{\star\sigma}_{t_2^+}$. Let $\Psi$ be a scalar convex decreasing function of any non-singular non-negative symmetric matrix. Many choices are possible in the context of optimal design (Fedorov and Hackl, 1997) such as

- The determinant function, which is called D-optimality. From Eq. (17) we obtain

$$\begin{aligned}\ln\det\left(\boldsymbol{P}^{\star\sigma}_{t_2^+}\right) &= \ln\det(\boldsymbol{P}^{\star\sigma}_{t_2}) \\ &\quad - \ln\det\left(\boldsymbol{I}_p + (\boldsymbol{r}^{\star}_{t_2})^{-1}\boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\right).\end{aligned} \tag{19}$$

  Thus the D-optimality criterion maximises the quantity $\ln\det(\boldsymbol{I}_p + (\boldsymbol{r}^{\star}_{t_2})^{-1}\boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T)$. By using Eq. (12), and in the absence of model error, one checks that $\det(\boldsymbol{P}^{c}) = \det(\boldsymbol{P}^{\sigma})\det(\mathcal{H}^T\mathcal{H})$. Since the matrix $\mathcal{H}^T\mathcal{H}$ does not depend on the design, we can then conclude that one can reason equivalently on the error covariance matrix of either the source or the activity concentrations.

- The trace function, or A-optimality. Taking the trace of Eq. (17) one obtains

$$\begin{aligned}\mathrm{Tr}\left(\boldsymbol{P}^{\star\sigma}_{t_2^+}\right) &= \mathrm{Tr}(\boldsymbol{P}^{\star\sigma}_{t_2}) \\ &\quad - \mathrm{Tr}\left(\left(\boldsymbol{r}^{\star}_{t_2} + \boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_2}\boldsymbol{h}^T\right)^{-1}\boldsymbol{h}(\boldsymbol{P}^{\star\sigma}_{t_2})^2\boldsymbol{h}^T\right).\end{aligned} \tag{20}$$

  Hence minimising $\mathrm{Tr}(\boldsymbol{P}^{\star\sigma}_{t_2^+})$ is equivalent to maximising the second term in this equation. Note that $\mathrm{Tr}(\boldsymbol{P}^{c}) = \mathrm{Tr}(\boldsymbol{P}^{\sigma}\mathcal{H}^T\mathcal{H})$. In this case design with respect to the source term and to the field plume concentrations are not equivalent.

In our study we will focus on designs obtained from A-optimality (and to a lesser extent D-optimality).

### 4.5. Plume data assimilation with targeting

In Section 3 we have introduced a data assimilation scheme without adaptive observation strategy. The new one incorporates this additional strategy. The algorithm is an augmented version of the original data assimilation one. The major difference is that, at decision time $t_n$, a *forecasted analysis* is performed according to the previous section. It leads to the maximisation of $\mathrm{Tr}(\boldsymbol{P}^{\star\sigma}_{t_2}) - \mathrm{Tr}(\boldsymbol{P}^{\star\sigma}_{t_2^+})$, using Eq. (20). As a result, the optimal locations of the targeted observations are obtained. Then the data assimilation scheme is performed between $t_1$ and $t_2$, using the fixed and targeted observations on $[t_1, t_2]$, leading to a posterior uncertainty of $\boldsymbol{P}^{c}_{t_2^+}$.

### 4.6. Simulated annealing algorithm

With a number of mobile network configurations increasing exponentially with both the number of mobile devices and the number of potential sites to welcome these devices, we have chosen an optimisation based on a stochastic algorithm. Following Abida et al. (2008) where details can be found, a swap simulated annealing algorithm is adopted. At each iteration, an ungauged station is selected while a previously selected one is removed. This network configuration is then evaluated through the cost function $J_{\mathrm{proposal}}$. If it is lower than the current network $J_{\mathrm{current}}$, it is accepted. If its cost is higher, it is still accepted with a probability that depends on a Boltzmann weight of the form $e^{-\Delta J/T}$, with $\Delta J = J_{\mathrm{proposal}} - J_{\mathrm{current}}$ the cost function difference between the proposal and the current configurations. $T$ is a fictitious temperature parameter. In order to avoid being trapped in local minima, the temperature $T$ is decreased, and the system is progressively cooled down.

The optimisation time depends on one hand on the algorithmic complexity, and on the other hand linearly on the time taken to evaluate the cost function. The complexity depends on the number of combinations which is of the order of the number of mobile devices (10 in the following) that can be dispatched on $50 \times 50$ potential sites. This leads to roughly $10^{27}$ configurations which is still tractable and leads to a typical optimisation time for one targeting decision of 5 to 10 min on a single processor machine.

## 5. Results and discussion

This section is aiming to show how an adaptive observation strategy connected to a sequential data assimilation scheme could reduce the uncertainty in the estimation of the present and future state of a radioactive plume.

### 5.1. Targeting experiments

We have considered that around the Bugey nuclear site, the static monitoring network is composed of 21 off-site gamma-dose stations (13 stations denoted by diamonds, 8 stations denoted by black circles and squares, referring to Fig. 2). It is also assumed that observations coming from this monitoring network are available every 30 min rather than every 10 min. This represents only a fraction of the theoretical capabilities of the current network. We have chosen to do so in order to strengthen the relative impact of targeted observations but, most of all, because the fixed observations would unlikely be all available and quality-controlled in case of crisis. This should also compensate for the fact that many stations are gamma dose instruments. They do not measure directly activity concentrations and are prone to contamination effects, so that the

information content of those observations is weaker for the data assimilation system.

Following Section 3.4, an observation is generated from the perfect model, and is multiplicatively perturbed by a log-normal law with $\lambda_l = 0.5$. The data assimilation system assumes these observations have Gaussian errors with a relative standard deviation of 50%, according to the error modelling of Section 3.4.

The forecast horizon for the sequential data assimilation scheme is chosen to be $\Delta t_f = 2$ h. According to the targeting decision time analysis described in Section 4, the first design time $t_1$ (or decision time) is taken to be 1 h after the start of the accident. So the first analysis window for the sequential updating algorithm ranges from $t_0 = 0$ to $t_1 = 1$ h. The number of mobile stations is set to 10 stations, delivering one targeted observation each. First targeted observations will be carried out at time $t_2 = t_1 + \Delta t_f$. When the analysis window reaches time $t_2$, the observations from both the static and adaptive networks will be assimilated. From this will be carried out a new design for time $t_2 + \Delta t_f$. During the complete trajectory of the sequential updating process, the time-lag between the targeting decision time and the targeted observations time (when mobile observations are carried out) is taken to be equal to the forecast window.

Herein, we have explored the two meteorological conditions of 20th December 2007 and 1st June 2007 (weak to moderate wind, and moderate to strong wind). Both A-design and D-design criteria have been investigated. Here only the results obtained for the A-design will be described.

For each studied case (December and June 2007) we have performed a sequential reconstruction of the plume field by (i) assimilating only the observations from the static monitoring network without any targeting strategy and (ii) assimilating both the stationary observations and the optimally targeted observations during the complete window of the assimilation process. In order to highlight the performance provided by assimilating optimal targeted observations, and so as to make (i) and (ii) comparable, we have computed on top of case (i) an ensemble of data assimilation trajectories by adding 10 observations selected randomly at each decision time.

### 5.2. Statistical indicators

In order to compare between the reference plume and the reconstructed plume obtained through the sequential data assimilation scheme (with or without a targeting strategy) we have merely used two statistical indicators following (Bocquet, 2007): the correlation coefficient $\rho$ and a specific figure of merit indicator fm. They are defined respectively at some time $\tau$ as follows:

$$\rho = \sum_{h\in S} \frac{[\overline{\boldsymbol{c}}]_h [\boldsymbol{c}^t]_h}{\sqrt{\left(\sum_{h\in S} [\overline{\boldsymbol{c}}]_h^2\right)\left(\sum_{h\in S} [\boldsymbol{c}^t]_h^2\right)}}, \tag{21}$$

$$\mathrm{fm} = \frac{\sum_{h\in S} \min\left([\overline{\boldsymbol{c}}]_h, [\boldsymbol{c}^t]_h\right)}{\sum_{h\in S} \max\left([\overline{\boldsymbol{c}}]_h, [\boldsymbol{c}^t]_h\right)}, \tag{22}$$

where $S$ denotes the set of the spatial grid cell indices, $\boldsymbol{c}^t$ is the true concentration field at time $\tau$, and $\overline{\boldsymbol{c}}$ represents the reconstructed plume field at the same time.

Note that in the definition of the correlation coefficient, the reference and the estimated fields are not centred about their means, in view of the fact that the mean value in an accidental release context is not a relevant indicator.

In addition to these two indicators we have considered also the root mean square error (rmse) essentially to compare between the source term reference and the estimated source at some time $\tau$. The rmse indicator is given by:

$$\mathrm{rmse} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left([\boldsymbol{\sigma}^t]_i - [\overline{\boldsymbol{\sigma}}]_i\right)^2}, \tag{23}$$

where $\boldsymbol{\sigma}^t$, $\overline{\boldsymbol{\sigma}}$ are respectively the source reference and the estimated source at time $\tau$. $N$ denotes the size of the vector source.

### 5.3. Locations of the mobile stations

As a first result, we have noticed that, for the two meteorological cases, A-designs with respect to the source term or with respect to the plume concentration field lead qualitatively to similar designs. In particular, both designs dispatch stations near the plume contour so as to control the spread of the radioactive cloud. The plume boundaries also correspond to high gradients of activity concentrations, where the activity concentration field is uncertain. This can be explained by the fact that since we are estimating a temporally decreasing source term, observations coming from the static network from $t_0$ to $t_2 = 3$ h and targeted measurements taken at $t_2 = 3$ h are enough to reasonably estimate the magnitude of the source term.

Some discrepancies are noticeable between the two types of A-design especially near the release site. More precisely, the design obtained with respect to the source term always allocates a few stations to the source neighbourhood. Obviously, this can be explained by the fact that an accurate estimation of the source term requires more monitoring stations close to the release site, in addition to the stationary stations close-by. Conversely, the concentration uncertainty based design puts mobile stations on the plume boundary only.

It is noticeable that the approach used by Jiang et al. (2007) to optimise the locations of 25 mobile devices by minimising a cost function counting the expected weighted sum of false positives and false negatives areas (relatively to a fixed threshold), leads to the same result: the 25 mobile devices are found located close to the plume boundaries. A sequential updating algorithm using a static Kalman Filter that relies only on a measurement equation was used by Herrera and Pinder (2005) to define a targeting strategy for the monitoring of contaminated ground-water aquifers. Each time, the algorithm chooses a single observation that reduces the most the total error variance of the plume estimate. Since stationary stations near the release site are not considered, the algorithm starts putting stations near the source, before choosing many others close to the plume boundary.

On the practical and applicative side, the fact that the targeted observations are located on the boundary of the radioactive plume has advantages. Measurement points will be reachable and more easily performed with terrestrial equipments outside the higher activity concentration zone, although past deposition should also be taken into account. Furthermore, it's worth noting that in case of a substantial accident, the stationary stations surrounding the site can become contaminated, leading to incorrect measurements. Yet, the observations derived from mobile devices placed inside the plume boundaries will be of good quality.

In the following, we present only the A-design criterion results obtained with respect to the concentration field. According to Subsection 4.4, and Eq. (12) together with Eq. (17), the cost function of this design criterion is formulated as

$$J(\boldsymbol{h}) = \mathrm{Tr}\left(\left(\boldsymbol{r}^{\star}_{t_n} + \boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_n}\boldsymbol{h}^T\right)^{-1}\boldsymbol{h}\boldsymbol{P}^{\star\sigma}_{t_n}\mathcal{H}^T\mathcal{H}\boldsymbol{P}^{\star\sigma}_{t_n}\boldsymbol{h}^T\right). \tag{24}$$

This cost function measures the uncertainty reduction in the concentration field after adding new observations mapped by the sub-matrix $\boldsymbol{h}$ chosen from the global Jacobian matrix $\mathcal{H}$. Therefore it should be maximised.

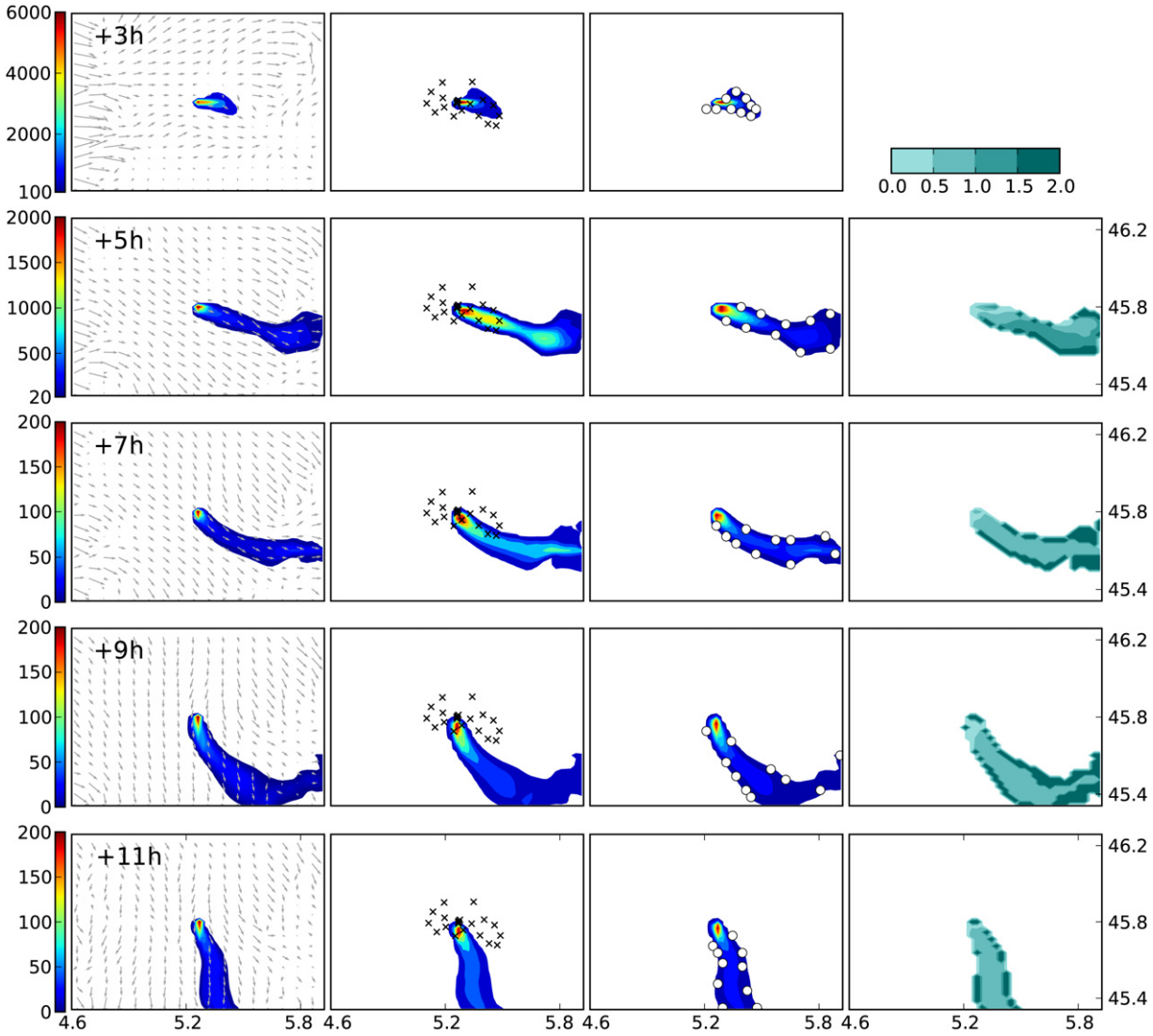


**Fig. 7.** Sequential plume reconstruction for strong wind conditions. The first column gives the reference plume knowing the true source term. The second column gives the forecasted plume from an analysis assimilating only the stationary observations. In the third is the forecasted plume from an analysis assimilating simultaneously the stationary and the targeted observations. The unit is Bq $m^{-3}$. The evolution of the plume is given at times +3 h, +5 h, +7 h, +9 h and +11 h. The fixed stations are depicted by cross marks. The mobile stations are marked with white disks. The fourth column indicates the contrast factor between the forecasts with and without targeting.

Figs. 7 and 8 show the sequential plume reconstruction of radionuclide $^{131}$I under strong wind and weak to moderate wind conditions, respectively. The dynamical evolution of the cloud is given at times +3, +5, +7, +9 and +11 h after the start of the release occurring at time $t_0 = 00$ UTC. For both figures, the first column presents the reference plume knowing the true source term. The wind field on the ground at 10 m high is superimposed. The second column gives the forecasted plume obtained by the forecast of the analysis at time $t - \Delta t_f$, which was obtained by assimilating only the observations gathered by the fixed monitoring network around the site. The fixed network is displayed on these panels. The third column shows the targeted plume resulting from the forecast of the analysis from an assimilation at time $t - \Delta t_f$ of both stationary and targeted observations. The locations of these targeted observations are drawn on the plume field. Each deployment at one of these dates $t$, was computed based on all available observations from $t_0$ to $t - \Delta t_f$, in order to carry out a forecast from $t$ to $t + \Delta t_f$ by assimilating the new targeted observations obtained at time $t$. Also, the first targeted observations are performed at time $t = +3$ h. So the top panels of the second and third columns represent the same forecasted plume from the same analysis with the observations obtained from the static network at $t = +1$ h. A fourth column displays the contrast $2|[\overline{\boldsymbol{c}}_{\text{with}}]_h - [\overline{\boldsymbol{c}}_{\text{without}}]_h|/([\overline{\boldsymbol{c}}_{\text{with}}]_h + [\overline{\boldsymbol{c}}_{\text{without}}]_h)$ between the field forecasts, with and without targeting, for each grid cell $h$. The relative improvement due to targeting is stronger on the plume contour, where adaptive observations are located.

### 5.4. Calm and strong wind conditions

For the strong wind conditions case, and within the 10-h window after the beginning of the release, the radioactive plume is moving from the west to the south following the clockwise rotation of the local wind around the Bugey power plant. The plume is mainly advected downwind. The crosswind and vertical dispersion is much smaller compared to the downwind transport. Consequently, the contaminated area is strongly confined along downwind direction and reaches swiftly the domain borders around the site, within a radius of 50 km. Notice that as the plume is moving around the site, many of the static monitoring stations notably those that are outside a radius of 4 km, do not catch the plume.

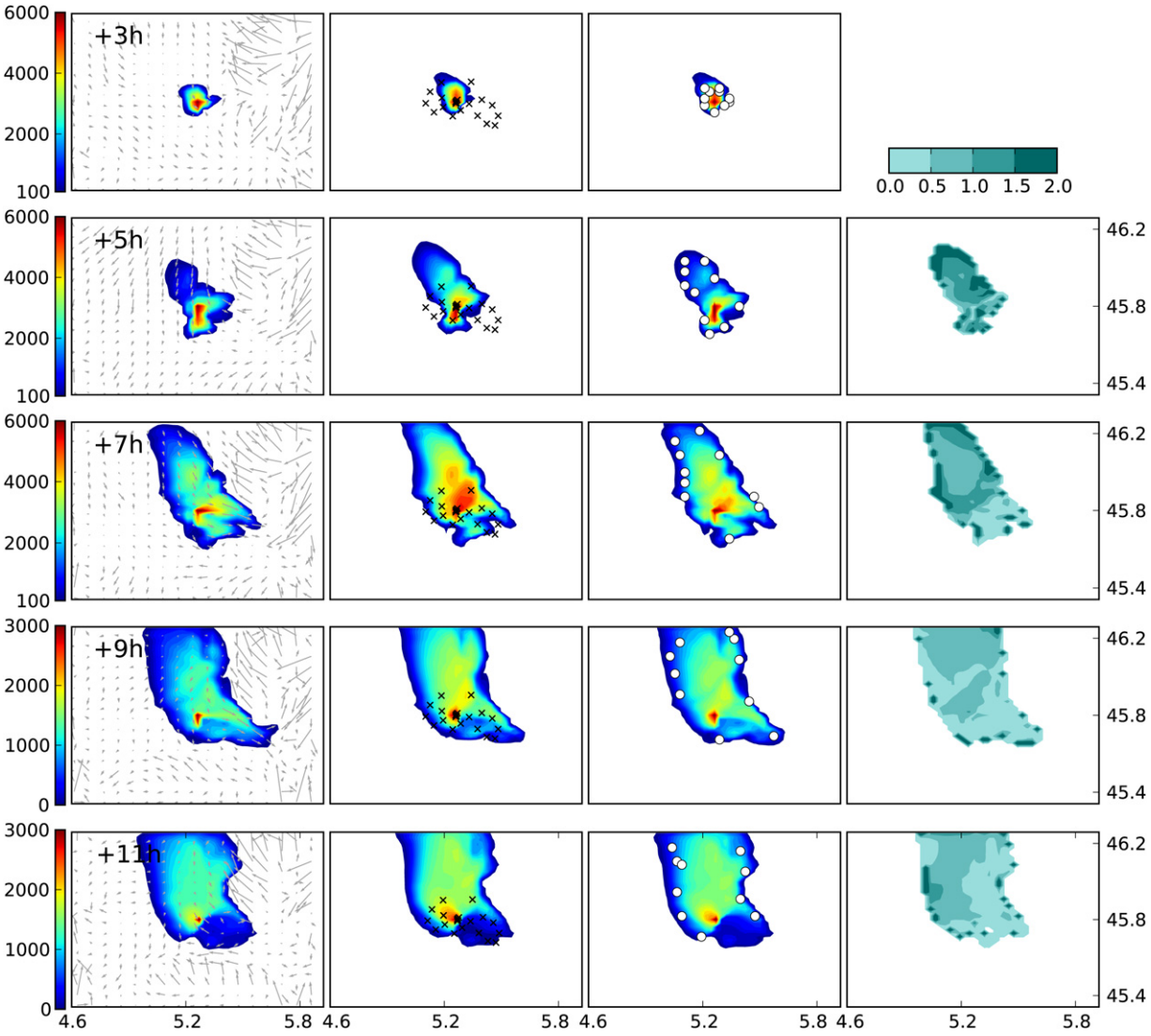


**Fig. 8.** Sequential plume reconstruction for weak wind conditions. The conventions and units are the same as in Fig. 7.

Hence when the wind intensity is strong, deploying mobile measurement could provide significant information as compared to the static network.

As for the weak to moderate wind conditions, the circulation of the wind field around the accident site is much more complex than in the first case. Within 10 km from the plant, the wind remains weak to moderate. As a result, the radioactive cloud diffuses in the vicinity of the plant, leading to large concentration values in the plant surroundings. Far away from the site, essentially in the East and South directions, the evolution of the plume is dynamically blocked by mountains (The Alps). The spread of the radioactive plume is much larger than in the first case, and the contaminated area is getting bigger and bigger with time. Contrary to the first case, the plume is observed by almost all the fixed stations. In addition, because the winds are weak in the vicinity of the plant, the contaminated air parcels close to the release source will be observed several times before leaving the site. As a result, the source term will be much more characterised than in the first case, in which the contaminated parcels near the source are transported away quickly by the wind. The statistical indicators will later confirm that the reconstruction performance is indeed better than in the strong wind conditions case. A comparable inverse modelling experiment though at continental scale led to similar results to Krysta and Bocquet (2007) on the strong and calm meteorological conditions.

### 5.5. Quantitative performance of the targeting scheme

Let us now compare between the reference plume and the plume analyses for the two wind conditions. Globally the boundary of the plume, which is defined up to some reference activity, is well forecast. However the uncertainty coming from the wind field is neglected in the sequential updating approach. Its presence, as in real conditions, could degrade this performance. This model error issue will be discussed later. The main discrepancies with the truth are observed nearby the source and along the central axis of the plume. In fact, the reconstruction of the plumes from the static observations, or by assimilating static and mobile observations simultaneously, leads to an overestimation of the ground activity concentrations during the whole experiment duration (7 days). This excess of activity is due mainly to the source persistence hypothesis, and also to the log-normal perturbation method, which is asymmetric. However note that, near the plume boundaries where most of the adaptive observations are located, this overestimation is strongly reduced.

As expected, the performance of the sequential reconstruction is improved by assimilating the targeted observations. The retrieval of

the plume including the targeted observations is in better agreement with the reference plume.

Fig. 9 gives the total uncertainty estimated by $\mathrm{Tr}(\boldsymbol{P}^{c}_{t_n^+})$ during the whole experiment. The reduction of uncertainty

$$J(\boldsymbol{h}) = \mathrm{Tr}(\boldsymbol{P}^{\star c}_{t_n}) - \mathrm{Tr}\left(\boldsymbol{P}^{\star c}_{t_n^+}\right) \tag{25}$$

is also plotted at targeting times $t_n = 3h + n\cdot\Delta t_f$, for $n = 0, 1, 2, \ldots, 82$.

The total uncertainty is decreasing due to the gain of information from fixed and targeted observations that overwhelm the increase in the source unknowns. Beyond a few days, it reaches a plateau which corresponds to a drop in the information gain. This is understood by the fact that the contribution of targeted observations is more important in the beginning of the experiment as there are relatively less available observations. With time, the number of observations derived both from static and mobile stations increases. So the sequential reconstruction is getting gradually less and less sensitive to the assimilation of new targeted observations, until reaching a saturation regime in which the reduction of uncertainty becomes practically unchanged. Then the assimilation strategy becomes unnecessary. Notice that for the calm wind conditions, the saturation regime is reached after 3 days. However, in strong wind conditions the regime is reached nearly after 6 days.

In insert of Fig. 9 is plotted the following ratio:

$$\mathrm{QF} = \frac{\mathrm{Tr}(\boldsymbol{P}^{\star c}_{t_{n+1}}) - \mathrm{Tr}(\boldsymbol{P}^{\star c}_{t^+_{n+1}})}{\mathrm{Tr}(\boldsymbol{P}^{c}_{t_n^+}) - \mathrm{Tr}(\boldsymbol{P}^{c}_{t^+_{n+1}})}. \tag{26}$$

As before, $\boldsymbol{P}^{c}_{t_n^+}$ and $\boldsymbol{P}^{c}_{t^+_{n+1}}$ are the analysis error covariance matrix error at time $t_n$ and $t_{n+1}$ after both fixed and targeted observations have been assimilated. $\boldsymbol{P}^{\star c}_{t_{n+1}}$ is the forecasted analysis error covariance matrix error at time $t_{n+1}$ after fixed forecasted observations have been assimilated and before the assimilation of targeted observations. The difference $\mathrm{Tr}(\boldsymbol{P}^{\star c}_{t_{n+1}}) - \mathrm{Tr}(\boldsymbol{P}^{\star c}_{t^+_{n+1}})$ is a measure of the total reduction of uncertainty (gain of information) obtained from the targeted observations only. (Actually it is only a forecasted reduction of uncertainty, which is easily accessible in the implementation, but should be close to the actual one.) The difference $\mathrm{Tr}(\boldsymbol{P}^{c}_{t_n^+}) - \mathrm{Tr}(\boldsymbol{P}^{c}_{t^+_{n+1}})$ measures the total reduction of uncertainty between two analysis times. It is the composition of three terms: an increase of uncertainty due the new unknown variables of the source between $t_n$ and $t_{n+1}$ (four scalars here), a reduction of uncertainty due to the assimilation of fixed observations, and the reduction of uncertainty due to the assimilation of targeted observations. This ratio is therefore a measure of the quality of the targeted observations. In this experiment, recall that the ratio of targeted observation to the total number of observations is 10–94. Yet the QF factor is very significantly above 40%, indicating that the targeting observations have a much higher informational content. It is especially high in the first 40 h after the beginning of the accident. As expected earlier, QF is higher on average for the strong wind configuration.

In Fig. 10 the total reconstructed activity within a radius of 50 km from the accident site is shown at each date $t_n$. The total activity is obtained by summing the activity in all the 11 vertical layers of the model domain. As expected, the reconstructed activity obtained by the targeting strategy lies between the activity of the reference and the reconstructed activity by assimilating only the stationary observations. Besides the total activity for strong wind conditions is much more important than in the case of calm wind conditions.

Fig. 11 gives the rmse and $\rho$ indicators computed with respect to the source term, for strong and calm wind conditions. The grey shaded area correspond to an ensemble of 100 trajectories obtained by assimilating at each time $t_n$ 10 targeted observations chosen randomly, rather than optimally. For the two cases we notice that the rmse is decreasing gradually with time as expected, and the rmse trajectory obtained by using targeting strategy performs systematically better than the trajectory reconstructed only by assimilating the static observations. Moreover, these two trajectories form an envelope for the trajectories based on random targeted stations. Indeed, adding new observations, even randomly distributed, always increases the available information content, from a very poor content to an optimal one (optimal targeting). The same logic applies to the correlation curves. Here also, and for both cases, the effect of the saturation regime discussed previously is obvious.

At the first targeting time $t = +3$ h the rmse value is about $10^4$ Bq$^{-1}$, and in the end of the sequential assimilation process, it gets the value of 100 Bq s$^{-1}$ for the strong wind conditions and about 20 Bq s$^{-1}$ for the case of weak wind conditions. Whatever the case, the correlation coefficient between the true source and the

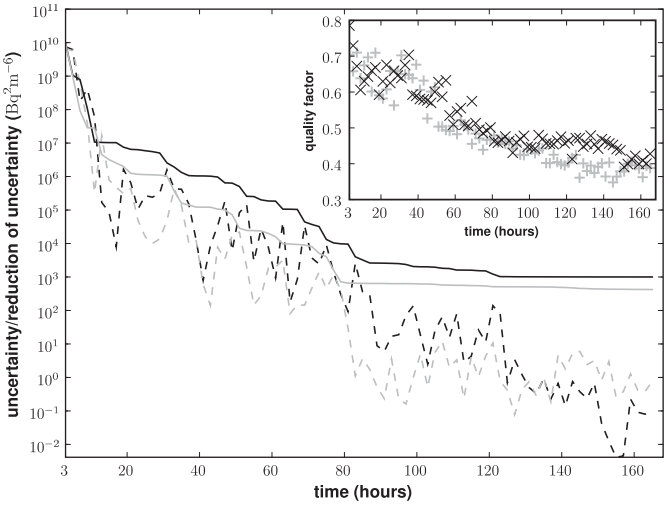


**Fig. 9.** The main figure shows the absolute uncertainty and the reduction of uncertainty as a function of time. The full lines show the absolute uncertainty, given by the trace of the total analysis error covariance matrix $\mathrm{Tr}(\boldsymbol{P}^{c}_{t_n^+})$. The dashed lines correspond to the reduction of uncertainty $J(\boldsymbol{h})$. The light grey lines are related to the calm meteorological conditions, while the dark ones are related to the strong wind conditions. In the inset is plotted the relative reduction of uncertainty due to targeting over time interval $[t_n, t_{n+1}]$, QF, in percentage. The + are related to the calm winds, while the × are related to the strong winds.

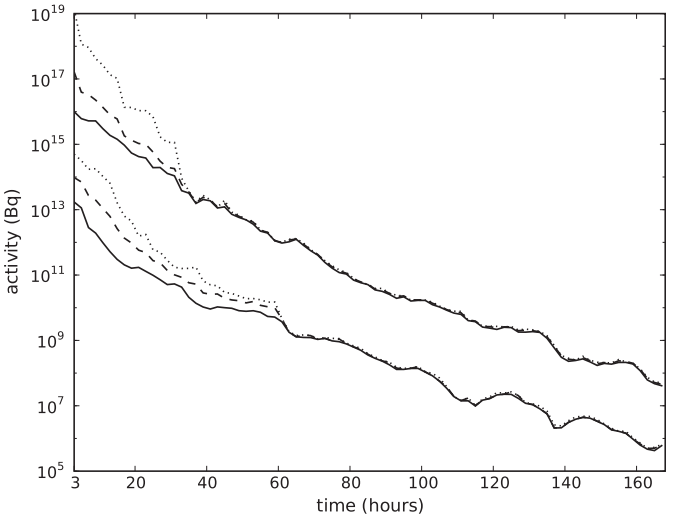


**Fig. 10.** Reconstructed activity around the site. The full line gives the total activity computed from the reference plume. The dotted line gives the activity from the reconstructed plume by using only the stationary observations. The dashed line gives the activity from the reconstructed plume using stationary and fixed observations together. The three bottom lines give the total retrieved activity profile for strong wind conditions. The three top lines correspond to the calm wind conditions.

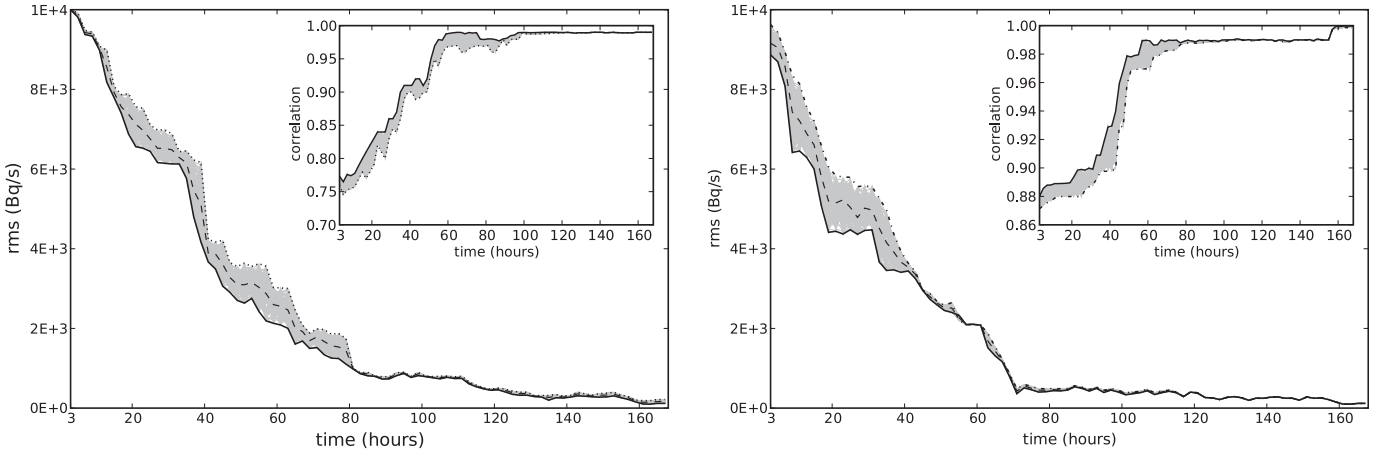


**Fig. 11.** Source root mean square error for strong (left) and calm (right) wind conditions. The solid line gives the rmse between the true source and the reconstructed source using the stationary and targeted observations. The dotted line gives the rmse between the true source and the reconstructed source using only the static observations. In the shaded grey area are plotted 100 trajectories with random targeted observations. The dashed line gives the median of these trajectories. The nested figure displays the correlation coefficient.

reconstructed source takes quickly values close to 1. This means that the decreasing shape of the source term is correctly resolved. Therefore the correlation coefficient may not be a very stringent statistical indicator to quantify the performance of analysed fields. The figure of merit indicator fm is more thorough, because it allows a comparison between the magnitude of the reference plume and the reconstructed plume, at each point of the domain.

To shed some light on the optimality of the targeting algorithm, we have computed $10^4$ assimilation trajectories (against 100 previously). For a given trajectory, 10 targeted observations are selected randomly and are assimilated at each time $t_n = 3\ \mathrm{h} + n \cdot \Delta t_f$, and so on until reaching the end of the assimilation window. So there are $10 \times 83 = 830$ targeted observations in total in a complete trajectory, where 83 is the number of targeting times. This random ensemble was binned into 100 bins, for each date. This allows to compute a (marginal) probability density map. Fig. 12 represents the maps of probability density obtained respectively for strong wind conditions and calm wind conditions.

For each density map, three trajectories are plotted. The dashed line is the optimal trajectory obtained by assimilating the mobile observations, selected by minimising the total error variance with respect to the concentration field. The full line is optimal with respect to the total error variance of the source term. The dotted line is the optimal trajectory by assimilating only the stationary observations.

First of all, the targeted observations chosen with respect to the concentration field covariance error matrix perform better than those selected with respect to the source term, and this for the whole experiment window. Although expected, it was not guaranteed since the figure of merit and the minimum variance criteria are not identical.

Note that the figure of merit indicator is gradually increasing with time as the number of assimilated observations is growing, until reaching a saturation regime in which the values of fm become close to 1, similarly to what was obtained by Bocquet (2007). In any case, the optimised trajectories are above the trajectory obtained by assimilating only the static observations.

The assimilation trajectories, with or without optimal targeting, shape an envelope for the ensemble. Note that the mode trajectory of the ensemble is located well in the centre of this envelope. The optimal deployment appears as a rare event. This stresses the importance of the optimisation. Again, the correlation coefficient between the reconstructed plume and the reference takes values near to 1 during the whole experiment window. In contrast, the figure of merit indicator is increasing slightly from values about 0.1 to the values near to 1 at the end of the experiment.

Finally, the density plot of Fig. 12 demonstrates that the optimal targeting strategy is not trivial since it is in the wake of the distribution.

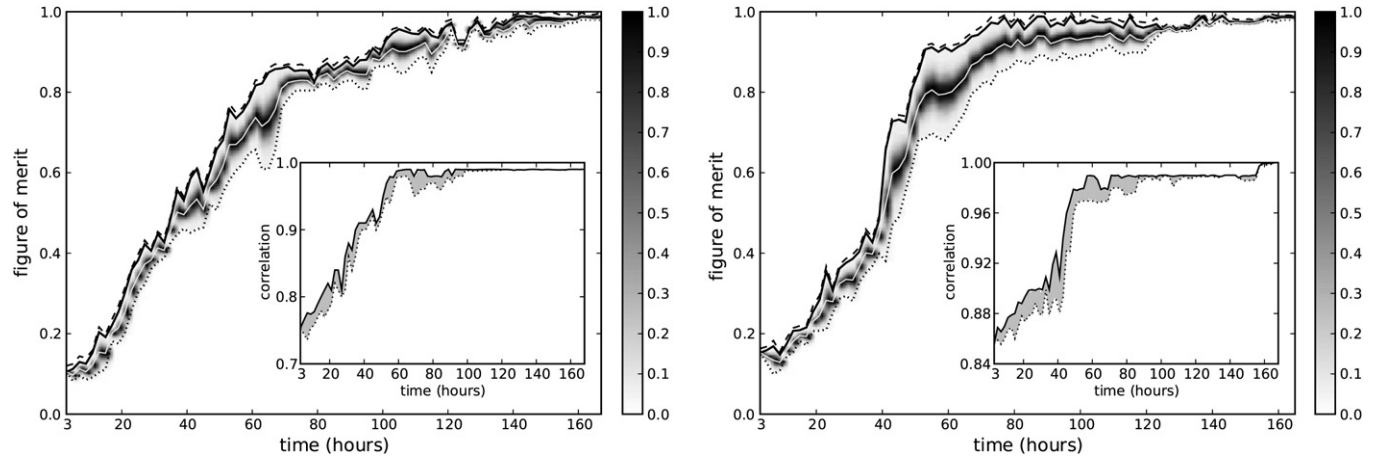


**Fig. 12.** Figure of merit: strong (left) and calm (right) wind conditions. The light grey line gives the median of the $10^4$ random assimilation trajectories. In insert is plotted the correlation coefficient. Details are given in the text.

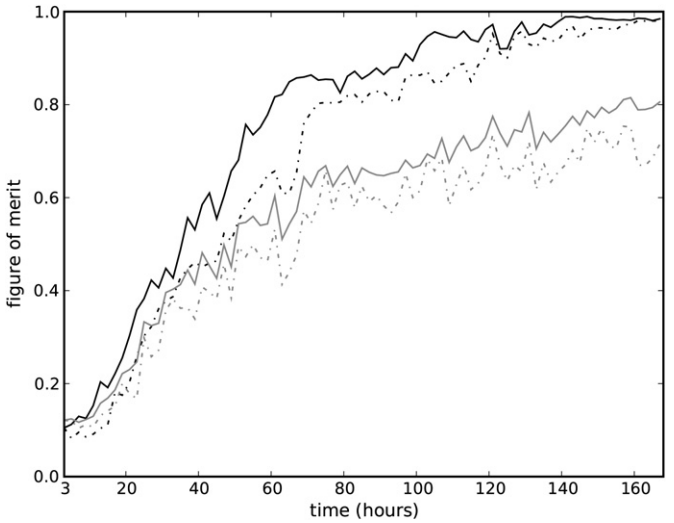


**Fig. 13.** Figure of merit of the data assimilation schemes with and without model error ($\mathbf{Q} = \mathbf{0}$) in the strong wind conditions. The dark curves are related to data assimilation with a perfect model, while the grey curves are related to data assimilation with a time shift in the meteorological fields, as an Ersatz for a possible source of model error. The full lines correspond to the case of data assimilation with targeting, while dashed curves correspond to data assimilation without targeting.

### 5.6. Impact of meteorological conditions errors

In such accidental release data assimilation scheme, the first source of errors is the uncertainty on the source term. This source of errors should be partially removed by inverse modelling. The second source of errors in magnitude is the uncertainty in the input wind field data, and on the other meteorological fields used in the diagnostic of the vertical turbulent diffusion. For pollutant such as iodine-131, this is the main contribution to model error (Quélo et al., 2007) after the uncertainty on the source.

Taking into account these errors is still a difficult topic that goes beyond the aim of this paper. However, in order to assess how this whole study would be affected by these errors, we have carried out the following test.

The Jacobian matrix $\mathcal{H}$ has been perturbed by shifting back the meteorological fields by 15 min. The shifted meteorological fields are what we assumed as the best available estimate and are provided to the chemistry and transport model. Then, we have performed a sequential plume reconstruction by assimilating only the stationary observations. Fig. 13 gives the assimilation trajectory of this reconstruction. Model error seriously decreases the performance of sequential data assimilation in the long term (beyond 40 h in our example). What is important here, is that the targeting strategy still improves significantly data assimilation in a proportion which is comparable to the perfect model study, as can be observed in Fig. 13. Therefore, in presence of model error, the targeting strategy should still be relevant, even though the targeting scheme becomes suboptimal.

## 6. Conclusions

In the event of an accident on a nuclear power plant resulting in the dispersion of radioactive materials into the environment, decision makers need to rely on high resolution and accurate information on the spatial distribution of the radioactive contamination around the accident site, in order to issue adequate protective countermeasures. Fixed monitoring networks in the vicinity of nuclear power plants are designed to provide an early warning and/or to provide a first, rough assessment of the release magnitude. The primary goal of this paper was to explore the possibility of an optimal deployment of mobile stations at mesoscale, in the event of a release. The targeting of observations is optimal when maximising the information content of observations forwarded to the plume data assimilation system operated after the release.

For that, the chemical-transport model Polair3D forced by MM5 meteorological input data was used to build an OSSE in a radius of 50 km around the Bugey nuclear power plant. An hypothetical dispersion scenario of a core meltdown nuclear accident occurring at the Bugey site, was performed under strong and calm local wind conditions.

The modelling of the uncertainties in the observations, possibly non-Gaussian, was discussed, and its impact on inverse modelling of the source term was studied. A realistic set of measurements, with significant log-normal errors, was generated to drive a sequential data assimilation based on inverse modelling. The proposed targeting strategy of observations, to be used to enhance the sequential data assimilation, was based on minimising the total error variance with respect to either the plume field, or to the source term.

As hoped for, the targeted observations that are assimilated improve clearly the performance of the plume estimation. The targeted observations are found being located near the plume boundaries, where strong gradients of the activity concentration field hold. The lateral spread of the plume is much more characterised by the assimilated targeted observations, while the source and the plume magnitude estimates are controlled by the observations from the fixed network surrounding the plant. On a practical side, that should make easier the monitoring of the plume boundary by mobile terrestrial equipments (such as vans), rather than performing in-situ measurement near the source release, with high levels of contamination.

It was shown that the difference between optimal targeting strategies based on the plume concentrations or on the source is very faint, though qualitative differences in the positioning of the mobile devices could be noticed.

The optimal reduction of uncertainty was also shown to be very significant as compared to a random deployment of the same set of devices. The information content of targeted observations was proven to be far better than the information content of fixed observations, during the whole release and especially at the beginning: the targeted observations represent 30–70% of the information gain, though they stand for only 10% of all observations.

As for model error stemming from uncertainties in the meteorological input data, a simple test was carried out. The wind field was shifted by 15 min so as to perturb the source–receptor matrix. As expected, the performance of the plume data assimilation is degraded, especially in the long term. However, it was shown on this example that the targeting scheme is still very beneficial to the reconstruction of the source and plume.

A natural extension of this study is to properly take into account model error of meteorological origin, which is identified as the second source of uncertainty after the source term uncertainty. This work can also be extended by assimilating targeted observations driven from an aerial mapping at higher altitudes, namely above the mixing boundary layer (e.g. 1500 m), or at low altitude in order to characterise well the three-dimensional aspect of the plume. This would especially be relevant if the local plume rise is significant, lowering the importance of ground monitoring.

## Acknowledgements

R. Abida acknowledges partial financial support from the IRSN. The authors thank A. Mathieu and O. Isnard for useful discussions and the monitoring networks data, W.G. Müller for his support and advice, L. Wu and O. Saunier for interesting discussions. This paper is also

a contribution to the INSU/LEFE-ASSIM project 2007-21: "Application of advanced data assimilation techniques to the dispersion of accidental release of pollutants in case of emergency".

The authors thank A. Carrassi for his in-depth review as well as another reviewer.

## References

Abida, R., Bocquet, M., Vercauteren, N., Isnard, O., 2008. Design of a monitoring network over France in case of a radiological accidental release. Atmos. Environ. 42, 5205–5219.

Anthes, R.A., Warner, T.T., 1978. Development of hydrodynamic models suitable for air pollution and mesometeorological studies. Mon. Weather Rev. 106, 1045–1078.

Arbia, G., Lafratta, G., 1997. Evaluating and updating the sample design in repeated environmental surveys: monitoring air quality in padua. J. Agric. Biol. Environ. Stat. 2 (4), 451–466.

Baklanov, B., Sorensen, J.H., 2001. Parameterisation of radionuclides deposition in atmospheric long-range transport modelling. Phys. Chem. Earth Pt B. 26, 787–799.

Berliner, L.M., Lu, Z.Q., Snyder, C., 1999. Statistical design for adaptive weather observations. J. Atmos. Sci. 56, 2536–2552.

Bocquet, M., 2007. High resolution reconstruction of a tracer dispersion event: application to ETEX. Q. J. R. Meteorol. Soc. 133, 1013–1026.

Carrassi, A., Trevisan, A., Uboldi, F., 2007. Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system. Tellus A 59, 101–113.

Clark, M.J., Smith, F.B., 1988. Wet and dry deposition of Chernobyl releases. Nature 332, 245–249.

Cressie, N.A.C., Cassie, N.A., 1993. Statistics for Spatial Data. Wiley, John & Sons, New-York.

Daescu, D.N., Navon, I.M., 2004. Adaptive observations in the context of 4D-Var data assimilation. Meteorol. Atmos. Phys. 85, 205–226.

Davoine, X., Bocquet, M., 2007. Inverse modelling-based reconstruction of the Chernobyl source term available for long-range transport. Atmos. Chem. Phys. 7, 1549–1564.

Fedorov, V.V., Hackl, P., 1997. Model-oriented Design of Experiments, Number 125 in Lecture Notes in Statistics. Springer Verlag.

Fourrié, N., Marchal, D., Rabier, F., Chapnik, B., Desroziers, G., 2006. Impact study of the 2003 North Atlantic THORPEX regional campaign. Q. J. R. Meteorol. Soc. 132, 275–295.

Galmarini, S., Bianconi, R., de Vries, G., Bellasio, R., 2008. Real-time monitoring data for real-time multi-model validation: coupling ENSEMBLE and EURDEP. J. Environ. Radioact. 99, 1233–1241.

Herrera, G.S., Pinder, G.F., 2005. Space-time optimization of groundwater quality sampling networks. Water Resour. Res. 41, W12407. doi:10.1029/2004WR003626.

IAEA: Atmospheric Dispersion in Nuclear Power Plant: a Safety Guide, Safety Series, No. 50-SG-S3, 1980.

Jiang, Z., de Bruin, S., Heuvelink, G.B.M., Twenhöfel, C.J.W. Optimization of mobile radioactivity monitoring networks. In: Proceedings of the Fifth International Symposium on Spatial Data Quality, Enschede, The Netherlands, 13–15 June, 2007.

Krysta, M., Bocquet, M., 2007. Source reconstruction of an accidental radionuclide release at European scale. Q. J. R. Meteorol. Soc. 133, 529–544.

Krysta, M., Bocquet, M., Sportisse, B., Isnard, O., 2006. Data assimilation for short-range dispersion of radionuclides: an application to wind tunnel data. Atmos. Environ. 40, 7267–7279.

Krysta, M., Bocquet, M., Brandt, J., 2008. Probing ETEX-II data set with inverse modelling. Atmos. Chem. Phys. 8, 3963–3971.

Levi, H.W., 1991. Radioactive deposition in Europe after the Chernobyl accident and its long-term consequences. Ecol. Res. 6 (2), 201–216.

Liu, D.C., Nocedal, J., 1989. On the limited memory method for large scale optimization. Math. Program. B. 45 (3), 503–528.

Louis, J.F., 1979. A parametric model of vertical eddy fluxes in the atmosphere. Boundary-Layer Meteorol. 17, 197–202.

Müller, W.G., 2007. Collecting Spatial Data: Optimum Design of Experiments for Spatial Data, 3rd Rev. Physica-Verlag Heidelberg, Germany.

Nychka, D., Saltzman, N., 1998. Design of air quality monitoring networks, chapter 4 of case studies in environmental statistics. In: Nychka, D., Piegorsch, W., Cox, L.H. (Eds.), Springer Lecture Notes in Statistics, Number 132. Springer Verlag, New-York, pp. 51–76.

Politis, K., Robertson, L., 2004. Bayesian updating of atmospheric dispersion after a nuclear accident. J. R. Stat. Soc. C-APP 53, 583–600.

Quélo, D., Krysta, M., Bocquet, M., Isnard, O., Minier, Y., Sportisse, B., 2007. Validation of the Polyphemus platform on the ETEX, Chernobyl and Algeciras cases. Atmos. Environ. 41, 5300–5315.

Saunier, O., Bocquet, M., Matthieu, A., Isnard, O., 2009. Model reduction via principal component truncation for the optimal design of atmospheric monitoring networks. Atmos. Environ. 43, 4940–4950.

Wikle, C.K., Royle, J.A., 1999. Space-Time dynamic design of environmental monitoring networks. J. Agric. Biol. Environ. Stat. 4, 489–507.

Zhao, M., Ming, Z., Panofsky, H.A., Ball, R., 1983. Wind profiles over complex terrain. Boundary-Layer Meteorol. 25 (3), 221–228.