



Nordic nuclear safety research

NKS-470

ISBN 978-87-7893-565-6

Source CHAracterizatiOn accounting for
meTeorological unCertainties
(SOCHAOTIC) –
final report

Jens Havskov Sørensen¹(co-ordinator), Henrik Feddersen¹
Kasper Skjold Tølløse¹
Andreas Uppstu²
Heiko Klein³, Magnus Ulimoen³
Lennart Robertson⁴
Jan Pehrsson⁵
Bent Lauritzen⁶
Agnieszka Hac-Heimburg⁷, Henrik Roed⁷
Einar Améen⁸, Naeem Ul Syed⁸
Anna Maria Blixt Buhr⁹, Jonas Lindgren⁹
Tuomas Peltonen¹⁰

¹Danish Meteorological Institute (DMI)

²Finnish Meteorological Institute (FMI)

³Norwegian Meteorological Institute (MET Norway)

⁴Swedish Meteorological and Hydrological Institute (SMHI)

⁵PDC-ARGOS

⁶Technical University of Denmark (DTU)

⁷Danish Emergency Management Agency (DEMA)

⁸Norwegian Radiation and Nuclear Safety Authority (DSA)

⁹Swedish Radiation Safety Authority (SSM)

¹⁰Radiation and Nuclear Safety Authority (STUK)

March 2023

Abstract

In recent years, events have occurred in which radionuclides were detected by filter stations in Europe without knowledge on the origin of those radionuclides. In such cases, there is a need to locate potential release sites. However, if the release site is actually known, or if a potential release site has been localized by inverse methods, then there is an additional need to estimate the release rates from this location as a function of time for the various radionuclides detected.

While in the SLIM NKS project, methodologies were developed to localize an unknown source of radionuclides dispersed in the atmosphere, the SOCHAOTIC project has developed methodologies, suited for operational use, by which characterization of the source, whose location is known, can be derived, i.e. to estimate the temporal release profiles of the radionuclides detected.

For operational use, nuclear decision-support systems should be extended with modules handling and analysing such monitoring data automatically, and conveying the data together with the geographical coordinates of the release point to the national meteorological centre accompanied by a request to estimate the temporal evolution of the release rates.

A number of case studies have been selected, viz. the ETEX-1 and the October 2017 case of Ru-106 in Europe as well as an artificial case. Methods for estimation of the temporal release profiles have been developed, and they have been applied to the selected cases by using the DERMA, MATCH, SILAM and SNAP atmospheric dispersion models. The methods have been applied both by using deterministic numerical weather prediction (NWP) model data and ensemble-statistical NWP model data derived by setting up and running the non-hydrostatic high-resolution Harmonie model for the selected cases. Finally, an interface to a nuclear decision-support system, ARGOS, is described.

Key words

nuclear emergency preparedness, atmospheric dispersion modelling, source characterization, inverse modelling, concentration measurements, gamma dose measurements, uncertainty

Source CHAracterizatiOn accounting for meTeorological unCertainties (SOCHAOTIC) – final report

**Final report of the NKS-B SOCHAOTIC activity
(Contract: AFT/B(22)1)**

Jens Havskov Sørensen¹ (co-ordinator)

Henrik Feddersen¹

Kasper Skjold Tølløse¹

Andreas Uppstu²

Heiko Klein³

Magnus Ulimoen³

Lennart Robertson⁴

Jan Pehrsson⁵

Bent Lauritzen⁶

Agnieszka Hac-Heimburg⁷

Henrik Roed⁷

Einar Améen⁸

Naeem Ul Syed⁸

Anna Maria Blixt Buhr⁹

Jonas Lindgren⁹

Tuomas Peltonen¹⁰

¹Danish Meteorological Institute (DMI)

²Finnish Meteorological Institute (FMI)

³Norwegian Meteorological Institute (MET Norway)

⁴Swedish Meteorological and Hydrological Institute (SMHI)

⁵PDC-ARGOS

⁶Technical University of Denmark (DTU)

⁷Danish Emergency Management Agency (DEMA)

⁸Norwegian Radiation and Nuclear Safety Authority (DSA)

⁹Swedish Radiation Safety Authority (SSM)

¹⁰Radiation and Nuclear Safety Authority (STUK)

Table of contents

Introduction	5
Atmospheric Dispersion Models	6
Danish Emergency Response Model of the Atmosphere (DERMA)	6
Method Employed for Source Characterization	6
Bayesian approach.....	8
Prior probability distribution	9
Likelihood	10
Usage of meteorological ensemble forecasting.....	11
Additional considerations.....	11
Multi-scale Atmospheric Transport and Chemistry model (MATCH)	12
Method Employed for Source Characterization	12
System for Integrated modeLLing of Atmospheric coMposition (SILAM).....	14
Method Employed for Source Characterization	14
Severe Nuclear Accident Program (SNAP)	15
Method Employed for Source Characterization	15
Semi-infinite Gamma Dose Model.....	17
ETEX-1	20
Sampling network	20
Synoptic situation	21
23 October, 1994	21
24 October, 1994	22
25 October, 1994	22
26 October, 1994	23
Tracer release	23
Results of DERMA	23
Results of MATCH	26
Results of SILAM	29
Results of SNAP.....	30
The Autumn 2017 Case of Ru-106.....	33
Results of DERMA	34
Results of MATCH	37
Results of SILAM	39
Results of SNAP.....	40
Mayak.....	40
	3

Niiar.....	41
Artificial Case of a Release from the Loviisa NPP	42
Results of DERMA	43
Discussion of use of gamma dose rates.....	49
Results of MATCH	50
Results of SILAM	53
Results of SNAP.....	54
Nord Stream leakage of methane in September 2022	57
ARGOS and Source Characterization	62
Concentration Measurements in ARGOS	62
Request for Source Characterization Calculation from ARGOS	63
Request interface for SL and SC modelling	64
Result interfaces for SL and SC	64
Types of measurements	65
Permanent gamma monitors and filter stations	65
Summary, Conclusions and Outlook.....	67
References	68
Appendix A Description of the source inversion approaches for the MATCH modelling system	73

Introduction

In recent years, several events have occurred in which radionuclides were detected by radiological filter stations in a number of European countries without any knowledge of the origin of those radionuclides. In such cases, there is a need to locate potential release sites by inverse dispersion modelling techniques. However, if the release site is actually known, or if a potential release site has been localized by inverse methods, then there is an additional need to estimate the release rates as a function of time from this location for the various radionuclides detected.

Two kinds of radiological monitoring networks exist, viz. filter and gamma stations. Filter stations detect radionuclide-specific time-average concentrations with high accuracy and low thresholds; however, with long averaging periods (typically one and seven days). Gamma stations produce time-average radiation doses in which the radionuclides are lumped together, and with lower accuracy and higher thresholds; however, with short averaging periods (down to 10 minutes).

While in the SLIM NKS project (Sørensen *et al.*, 2022), methodologies have been developed to localize an unknown source of radionuclides dispersed in the atmosphere and detected by a radiological monitoring network, the SOCHAOTIC project has developed methodologies, by which a characterization of the source, whose location is known, can be derived, i.e. to estimate the temporal release profiles of the radionuclides detected.

For operational use, nuclear decision-support systems should be extended with modules handling and analysing such monitoring data automatically, e.g. by employing the European Radiological Data Exchange Platform (EURDEP), and conveying the data together with the geographical coordinates of the release point to the national meteorological centre accompanied by a request to estimate the temporal evolution of the release rates.

In the course of SOCHAOTIC, a number of case studies have been selected, namely the ETEX-1 and the October 2017 case of Ru-106 in Europe as well as an artificial case produced by running a dispersion model forward and calculating average concentrations at filter stations and gamma dose rates at nearby gamma stations. Methods for estimation of the temporal release profiles have been developed and described in the present report, and they have been applied to the selected cases by using the DERMA, MATCH, SILAM and SNAP atmospheric dispersion models. The methods have been applied both by using deterministic numerical weather prediction (NWP) model data and ensemble-statistical NWP model data. The ensemble data have been derived by setting up and running the non-hydrostatic high-resolution Harmonie model for the selected cases. An additional section is prepared by SMHI for the Nord Stream leakage of methane in September 2022. In Appendix A, a full description of the SMHI contribution is given in addition to what is presented in the main report. Finally, an interface to the nuclear decision-support system ARGOS is described.

Atmospheric Dispersion Models

Danish Emergency Response Model of the Atmosphere (DERMA)

The Danish Emergency Response Model of the Atmosphere (DERMA) (Sørensen *et al.*, 2007; Sørensen, 1998) is a comprehensive numerical regional and meso-scale atmospheric dispersion model developed at the Danish Meteorological Institute (DMI). The model is used operationally for the Danish nuclear emergency preparedness, for which the Danish Emergency Management Agency (DEMA) is responsible (Hoe *et al.*, 2002). Besides, the model is employed for veterinary emergency preparedness (Sørensen *et al.*, 2000; 2001; Mikkelsen *et al.*, 2003; Gloster *et al.*, 2010a; 2010b), where it is used for assessment of airborne spread of animal diseases, e.g. foot-and-mouth disease. DERMA may also be used to simulate atmospheric dispersion of chemical substances, biological warfare agents and ashes from volcanic eruptions, and it has been employed for probabilistic nuclear risk assessment (Lauritzen *et al.*, 2006; 2007; Baklanov *et al.*, 2003; Mahura *et al.*, 2003; 2005).

The main objective of DERMA is to predict the dispersion of a radioactive plume and the accompanied deposition. However, the model may also be used in situations where increased levels of radioactivity have been measured but no information is revealed on a radioactive release. In such cases, inverse (adjoint) modelling may be applied whereby potential sources of radioactivity may be localised and release rates estimated.

The three-dimensional model is of Lagrangian type making use of a hybrid stochastic particle-puff diffusion description, and it is currently capable of describing plumes at downwind distances up to the global scale (Sørensen *et al.*, 1998). The model utilizes aerosol size dependent dry and wet deposition parameterisations as described by Baklanov and Sørensen (2001).

Currently, DERMA makes use of analysed and forecasted meteorological data of various deterministic versions at DMI of the NWP model Harmonie (Bengtsson *et al.*, 2017) covering North-western Europe, Greenland and the Faeroes, and from the global model developed and operated by the European Centre for Medium-range Weather Forecasts (ECMWF). Further, DERMA utilizes the COMEPS ensemble prediction system, which is based on the Harmonie model.

DERMA is interfaced with the Accident Reporting and Guidance Operational System (ARGOS) (Hoe *et al.*, 1999; 2002), a PC based nuclear decision-support system developed by the Prolog Development Center (PDC). The integration of DERMA with the ARGOS system is effectuated through automated online digital communication and exchange of data between the ARGOS system and the DMI High Performance Computing (HPC) facility.

Method Employed for Source Characterization

The method description is adapted from a recent study (Tølløse and Sørensen, 2022). In that study, we developed the method and applied it to the same Loviisa release case that we consider in this project. However, we considered an idealized case, where the same meteorological data, dispersion model and gamma dose rate model were used for both creating the data set and for conducting the source term estimation, i.e. model errors were negligible.

The temporal release profile of the various different radionuclides involved is estimated by using filter and gamma station measurement data and employing an atmospheric dispersion model.

Assume an overall start time t_0 and end time t_N of the release, and separate the release period $[t_0, t_N]$ in time bins (j), e.g. of one- or three-hour duration, and for each of these assume a unit release of each of the considered radionuclides (i), see Figure 1. For all considered test cases, we assume a point release near the surface of the Earth. Thus, some generalization of the method is needed if this assumption does not hold.

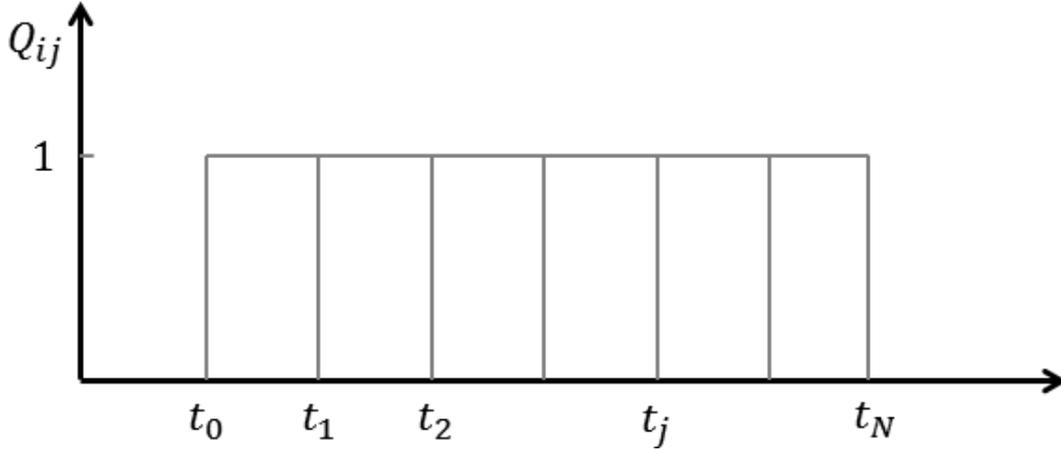


Figure 1 Time-binned unit release of radionuclide i .

For each release time bin j , and for each radionuclide i , the atmospheric dispersion model is run forward in time calculating average activity concentrations \bar{c}_{ijk} , where the k -index corresponds to the location and time of a specific measurement at a filter station. Also, the instantaneous activity concentrations c_{ijk} and deposition values d_{ijk} are calculated, where the κ -index corresponds to the location and time of a specific measurement at a gamma station. By using a semi-infinite gamma dose model, the contributions $\gamma_{ijk} = \gamma_{ijk}(c_{ijk}, d_{ijk})$ to the gamma dose rates are calculated.

For each radionuclide i , a linear combination of the time-binned releases, with non-negative coefficients λ_{ij} , is assumed. For a given set of coefficients λ_{ij} , the predicted average concentration is calculated at filter stations,

$$C_{ik}^m = \sum_j \lambda_{ij} \bar{c}_{ijk},$$

and predicted gamma dose rates at gamma stations,

$$\Gamma_{\kappa}^m = \sum_{ij} \lambda_{ij} \gamma_{ijk},$$

where the superscript m indicate that these are model predictions.

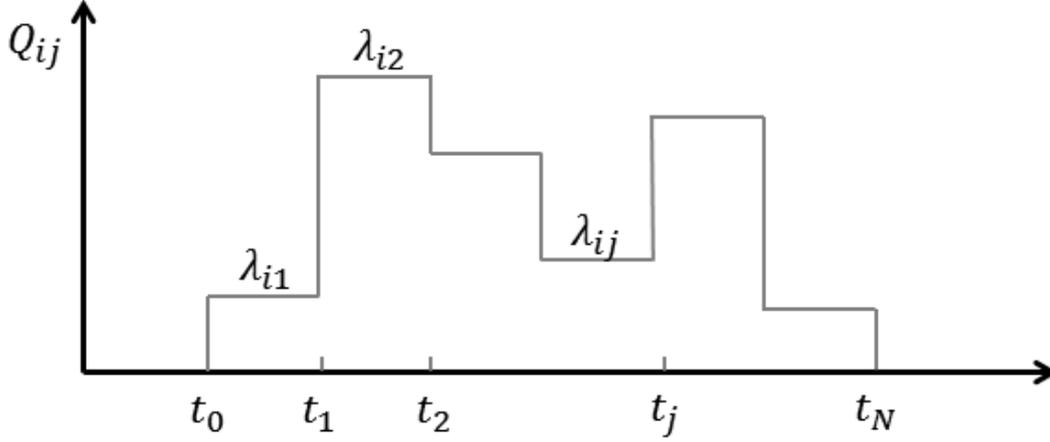


Figure 2 Estimated release rate of radionuclide i as function of time.

The corresponding piece-wise constant temporal profile of the source term is calculated for each radionuclide i , see Figure 2.

Bayesian approach

The probability distribution for λ is obtained by Bayes' theorem

$$P(\lambda|\mathbf{C}^o, \mathbf{\Gamma}^o, I) = \frac{P(\lambda|I) P(\mathbf{C}^o, \mathbf{\Gamma}^o|\lambda, I)}{P(\mathbf{C}^o, \mathbf{\Gamma}^o|I)} \propto P(\lambda|I)P(\mathbf{C}^o, \mathbf{\Gamma}^o|\lambda, I),$$

where I is any available background information about the release, and \mathbf{C}^o and $\mathbf{\Gamma}^o$ are the observed concentrations and gamma dose rates, respectively, indicated by the superscript o . $P(\lambda|\mathbf{C}^o, \mathbf{\Gamma}^o, I)$ is the *posterior* probability distribution for the coefficients λ , $P(\lambda|I)$ is the *prior* probability distribution for the coefficients, and $P(\mathbf{C}^o, \mathbf{\Gamma}^o|\lambda, I)$ is the likelihood (the probability of observing $(\mathbf{C}^o, \mathbf{\Gamma}^o)$ given a proposed set of coefficients, λ). Finally, $P(\mathbf{C}^o, \mathbf{\Gamma}^o|I)$ is called the evidence and is a λ -independent normalization constant.

For a high-dimensional λ , it is not computationally feasible to systematically explore all parts of the parameter space. Instead, the posterior probability distribution can be sampled using Markov Chain Monte Carlo methods (MCMC) such as Metropolis-Hasting or Gibbs (Hastings, 1970; Casella and George, 1992). The basic idea behind these methods is to iteratively propose source term models and *accept* them with a probability proportional to $P(\lambda|\mathbf{C}^o, \mathbf{\Gamma}^o, I)$. This way, areas with high probability are more likely to be explored, and therefore $P(\lambda|\mathbf{C}^o, \mathbf{\Gamma}^o, I)$ can be estimated even for a high-dimensional λ . When a sufficient number of models are accepted, the posterior statistics can be extracted from the selection of accepted models. In this study, we use the Hamiltonian Monte Carlo (HMC) method No U-Turn Sampling (NUTS), cf. Hoffman and Gelman (2014), which is implemented in the python library PyMC3 (Salvatier *et al.*, 2016). HMC methods generally have an advantage over random-walk based MCMC methods, because the model proposals are not generated by a random walk but instead based on an estimate of a functional form of $P(\lambda|\mathbf{C}^o, \mathbf{\Gamma}^o, I)$. An additional advantage of the NUTS algorithm is that it has built-in adaptive step sizing, which in practice means that almost no parameter tuning is necessary.

In this study, we started by using the default parameters, but in order to ensure convergence, the target acceptance rate was increased from the default 0.8 to 0.99. Aside from this,

everything was kept at the default values; two simultaneously running chains, each with 1000 tuning steps and 1000 draws from the target distribution. Thus, this provides a total of 2000 realizations of the posterior probability distribution. For further details on the NUTS parameters, see (Hoffman and Gelman; Salvatier *et al.*, 2016).

Prior probability distribution

Defining useful prior probability distributions for the release rates is challenging, since the magnitude of the release is unknown. To allow for variation over several orders of magnitude while ensuring non-negative values, we use log-normal prior distributions. Assuming a normal distributed variable $x \sim \mathcal{N}(\mu, \sigma)$, then the variable $z = e^x \sim \text{Lognormal}(\mu, \sigma)$ is log-normal distributed with parameters μ and σ . Thus, these denote the mean and standard deviation of x and not the log-normal distributed variable z . The prior probability distribution for the coefficients λ_{ij} can be written as:

$$P(\lambda_{ij}|I) = \text{Lognormal}(\mu_i, \sigma_i),$$

where μ_i and σ_i are parameters to be determined for the specific radionuclide. For all three cases, we use $\sigma_i = 1/2 \log(10^3)$. Thus, $\mu_i \pm 2\sigma_i$ includes six orders of magnitude for each release rate and thereby acts as a broad uninformative prior. Further, μ_i should ideally be chosen such that it is a few orders of magnitude smaller than the actual release rate. The reason for this is that for time bins where data do not constrain the release rate sufficiently, the algorithm will tend to sample the prior distributions. Thus, by ensuring that μ_i is small compared to the actual release, the posterior distribution will favor low values, if data do not “suggest” otherwise.

Since the magnitude of the release is unknown a priori, it is not trivial to choose μ_i . However, a test run can be performed to find the approximate magnitude of the release, and then μ_i can be determined. It is important to note that $P(\lambda_{ij}|I)$ will still be a broad uninformative prior distribution, but it will slightly prefer lower values if there is no information in data.

The numeric value of μ_i depends on the units assumed for each unit release: we use [ng/s] for ETEX and [Bq/s] for the Ru-106 case and the Loviisa case. For ETEX and the Ru-106 case, we use $\mu_i = \log(10^7)$, and for the Loviisa case, we use $\mu_i = \log(10^8)$.

Further, for the Loviisa case, we can use information about the core inventory to reduce the parameter space by imposing correlations between release rates of certain radionuclides, inspired by the method by Saunier *et al.* (2013; 2020). For example, two different isotopes of the same element will largely behave similarly during a release. Thus, if the half-lives of two such isotopes are long compared to the duration of the release and if there is no significant ingrowth from other processes, the ratio of the release rates between two isotopes can be assumed constant and equal to the ratio of the amounts in the core inventory. For example, Cs-134 and Cs-137 have half-lives of approximately 2 and 30 years, respectively. Thus, the ratio of their activity concentrations in the core inventory can be considered constant during the release. Accordingly, based on the amounts of the two isotopes in the core, we can assume the ratio of their release rates to be constant.

For other isotope pairs, it is necessary to take into account the difference in half-lives, in order to set realistic constraints on the release rates. In this case, knowing the amount of the two

isotopes at the time of SCRAM gives one limit for the isotopic ratios, while estimating the activity concentration n hours later will provide an estimate of the other limit, assuming no significant ingrowth. An example is the isotope pair I-131 and I-133, which has half-lives of approximately 8 days and 20.8 hours, respectively. Based on their activity concentrations in the core at the time of the accident, we have $\frac{q_{I-133}}{q_{I-131}} < 2.1$. Assuming that the duration of the main release is less than 24 hours, we can determine the other limit. Due to radioactive decay during these 24 hours, the amount of I-133 is decreased by a factor of 0.45, while we assume that the amount of I-131 is unchanged due to its relatively long half-life. Thus, a lower limit can be determined $\frac{q_{I-133}}{q_{I-131}} > 0.9$. Following this approach, we determine the following constraints:

$$\frac{q_{Cs-134}}{q_{Cs-137}} = 1.4, \quad 0.001 < \frac{q_{I-132}}{q_{I-131}} < 1.5, \quad 0.9 < \frac{q_{I-133}}{q_{I-131}} < 2.1 \quad \text{and} \quad 0.15 < \frac{q_{I-135}}{q_{I-131}} < 2.0.$$

where q denotes the release rates. For Cs-134 and Cs-137, this effectively means that only one release rate needs to be determined instead of two, and that the combined set of measurements of the two isotopes can be used. For the other isotope pairs, the constraints allow us to define log-normal distributions with upper and lower bounds, which depend on the release rate of one of the other nuclides. Let λ_{mj} and λ_{nj} be the coefficients for two release rates, which are related by the flexible constraints $r_{\text{lower}} < \lambda_{nj} / \lambda_{mj} < r_{\text{upper}}$. Then, the prior probability distribution for λ_{mj} will be defined as the log-normal distribution described above, while the prior probability distribution for λ_{nj} can be written as:

$$P(\lambda_{nj} | I, \text{constraints}) \propto \begin{cases} P(\lambda_{nj} | I) & \text{if } r_{\text{lower}} < \frac{\lambda_{nj}}{\lambda_{mj}} < r_{\text{upper}} \\ 0 & \text{otherwise} \end{cases}.$$

Likelihood

The likelihood is the probability of observing the set of measurements $(\mathbf{C}^o, \mathbf{\Gamma}^o)$ given a proposed source term, λ . The likelihood is evaluated by assuming a probability distribution for the residuals $C_{ik}^o - C_{ik}^m(\lambda_{ij})$ and $\Gamma_{\kappa}^o - \Gamma_{\kappa}^m(\lambda_{ij})$.

In this study, we use a log-normal likelihood, which is less sensitive to outliers than the Gaussian distribution and automatically gives higher weight to measurements/predictions of low values. This makes it useful when dealing with measurement values over several orders of magnitude, cf. e.g. Liu *et al.* (2017).

One practical challenge when dealing with log-normal distributions is that only positive values are allowed, while the physical quantity may in principle be zero. For the gamma dose rates, this is not an issue, since we add background radiation to the modelled measurements, thereby ensuring that values are always positive. For the air concentration measurements, on the other hand, modelled predictions may be zero, while the measured predictions may be below the detection limit.

Assume that for a given measurement, C_{ik}^o , the detection limit is ϵ_{ik} . To avoid zero-values, we use these altered observations and model predictions $\widetilde{C}_{ik}^o = \max(C_{ik}^o, \epsilon_{ik})$ and $\widetilde{C}_{ik}^m = \max(C_{ik}^m, \epsilon_{ik})$. For the ETEX case, we use the detection limit $\epsilon_{ik} = 0.01 \text{ ng/m}^3$, and for both the Ru-106 and Loviisa cases, we use $\epsilon_{ik} = 0.1 \text{ mBq/m}^3$, which is the order of magnitude of

the average detection limit in the Ru-106 data set. These altered forms have the additional benefit that they provide a theoretically sound way of using non-detections, since these will only contribute to the likelihood, when the modelled concentration is above the detection limit. Thus, there is no risk of falsely interpreting a low value as a zero. The likelihood is given as:

$$P(\widetilde{\mathbf{C}}^o, \mathbf{\Gamma}^o | \boldsymbol{\lambda}, I) = \prod_{ik} \text{Lognormal}(\widetilde{C}_{ik}^m, \sigma_f) \prod_{\kappa} \text{Lognormal}(\Gamma_{\kappa}^m, \sigma_g),$$

where σ_f and σ_g are related to the uncertainty of the measurements as well as the unknown model errors. Since model errors are unknown, these are treated as nuisance parameters, i.e. they are kept as free parameters and sampled by the Monte Carlo algorithm. In practice, a wide uniform distribution has been used as prior distribution for the nuisance parameters $\sigma_f, \sigma_g \sim U(0,100)$, which allows for a broad range of shapes of log-normal distributions.

As discussed in the Loviisa result section, we found that in case very limited filter measurement data are available, a Gaussian likelihood may be better suited.

Usage of meteorological ensemble forecasting

Assuming that the ensemble members are independent, the p 'th ensemble member will also result in an independent set of modelled measurements $(C_{ikp}^m, \Gamma_{\kappa p}^m)$. One can combine two independent probability densities simply by multiplication, $P(C|A, B) = P(C|A) P(C|B)$. Thus, by assuming independence, the resulting likelihood becomes:

$$P(\widetilde{\mathbf{C}}^o, \mathbf{\Gamma}^o | \boldsymbol{\lambda}, I) = \prod_p \left[\prod_{ik} \text{Lognormal}(\widetilde{C}_{ikp}^m, \sigma_f) \prod_{\kappa} \text{Lognormal}(\Gamma_{\kappa p}^m, \sigma_g) \right].$$

In addition, we run the algorithm individually with each meteorological ensemble. This will give an indication of how much the different source term estimates differ from each other.

Additional considerations

The result is sensitive to the choice of t_0, t_N as well as the width of the time bins. In case of an accident at a nuclear power plant, the time of the onset of the accident (scram) may be known, which provides a natural choice for t_0 . During the early phase of an accident, the release may not be over, when a source term estimation is first attempted. Thus, t_N may be set to the present time. However, it takes some time for the material to reach the measurement station, which means that one should probably exclude the latest time bins, thereby allowing the signal from the latest included time bins to reach the nearest measurement locations.

Further, it is challenging to determine the ideal bin width Δt_{bin} a priori, since this will most likely depend on several different parameters: the coarseness of the measurement network, the duration of the measurements observations, the distance travelled from release location to measurement stations, as well as the weather situation. Ideally, we would always choose a high temporal resolution, but due to the limited measurement capabilities, this leads to a larger risk of having an underdetermined problem. In practice, we found that the amount of data available limits the choice of Δt_{bin} , and that this will therefore be case dependent. See the selected values of t_0, t_N and Δt_{bin} in the result section for each case.

Multi-scale Atmospheric Transport and Chemistry model (MATCH)

The Multi-scale Atmospheric Transport and Chemistry model (MATCH) (Robertson *et al.*, 1999) is multi-purpose Eulerian chemical transport model (CTM) developed by the SMHI. The model is used for emergency application such as nuclear and natural events (volcanos), aerosol dynamics and optics (Andersson *et al.*, 2015), complex chemistry, and data assimilation (Robertson and Langner, 1998; Kahnert, 2008; Kahnert, 2018). The MATCH model is used operationally for chemical forecasts in CAMS (Copernicus Atmospheric Monitoring Service) and for SSM (Swedish Radiation Safety Authority) serving the ARGOS system needs (Hoe *et al.*, 1999; 2002). Other applications are studies for air quality and health issues in climate projections. In most applications MATCH is used as a limited-area model on various possible scales, but also for global applications.

The MATCH model is basically an Eulerian model but for emergency applications a Lagrangian particle model is used in the near field of the emission location.

A wide range of possible driving meteorological data is applicable like analyses and forecasts from HARMONIE, IFS (ECMWF) and WRF.

Method Employed for Source Characterization

In these applications we have applied so called Poor-man source inversion. This is a simplification of the complete 4Dvar where only one iteration is made followed of bias correction. In principle we run the adjoint model forced by the measurements where also adjoint gamma-dose may come in. We then pick up the adjoint solution for the column associated with the defined source location. A time series of such columns are then retrieved. The concept applies well for a Eulerian model. We now assume that these columns represent the source profile with an unknown scale factor to be determined. This is derived by running the forward model given the source suggestion from the adjoint run. When comparing the model and measurements we assume that there is only a scale factor other than implicitly one that is missing in a quadratic error sense.

Bringing this into a mathematical language we may first consider the transport model in a general form,

$$x_{t+1} = M_t x_t + P q_t$$

where x_t is the model state, M_t is the transport equations and P an injector matrix to place the source column q_t at the right location in the model grid. We now assume a cost function of the following form,

$$J(x_t) = \sum_t^T \frac{1}{2} (q_t - q_{bt})^T B^{-1} (q_t - q_{bt}) + \frac{1}{2} (H x_t - y_t)^T O^{-1} (H x_t - y_t)$$

that represents quadratic penalties for deviation from an initial source column guess and a penalty for deviation from the observations. Here the observation operator H is important, that maps the model grid values to the observation locations and in addition, when valid, perform gamma dose transform of the model nuclide specific model state. If we then bring the model as above a strong constraint we mat write,

$$J(x_t) = \sum_t^T \frac{1}{2} (q_t - q_{bt})^T B^{-1} (q_t - q_{bt}) + \frac{1}{2} (Hx_t - y_t)^T O^{-1} (Hx_t - y_t) + \lambda_{t+1}^T (Mx_t + Pq_t - x_{t+1})$$

Here λ represents the co-state or the adjoint state. The cost function will have minimum where the gradient with respect to the model state, x_t , is zero that lead us to the adjoint transport model,

$$\lambda_t = M_t^T \lambda_{t+1} + H^T O^{-1} (Hx_t - y_t)$$

It is worth to make short stop here and conclude that the adjoint state is integrated backward in time driven by the adjoint model equations (M_t^T) and forced by the deviation between model and measurements, transferred through the transpose (adjoint) of the observation operator, that in turn may include adjoint gamma dose rates.

We could from above also derive the gradient with respect to a specific source column q_t ,

$$\nabla J_{q_t} = P^T \lambda_t + B^{-1} (q_t - q_{bt})$$

that is derived from the adjoint state, where P^T turns in to an extractor of a column that belongs to the source location. For a full variational assimilation the second term on the right side will be important, while in the Poor-man case this will disappear. We may stepwise update the source columns,

$$q_t^{n+1} = q_t^n - \alpha \nabla J_q$$

where α is a scaling factor. The source columns may then be incrementally updated by forward and adjoint runs. The Poor-man approach does then simplify by just making one iteration,

$$\begin{aligned} q_t^0 &\equiv q_{bt} = 0 \\ q_t^{n+1} &= -\alpha \nabla J_q = -\alpha \lambda_t \\ \alpha &= ? \end{aligned}$$

where α initially assumed to be unit. One forward run then will determine how good this assumption appeared to be. Then setting up quadratic cost function with the unknown factor α ,

$$J(\alpha) = \sum_t^T (\alpha Hx_t - y_t)^T (\alpha Hx_t - y_t)$$

and just using Newton-Raphson algorithm we could iteratively determine best α ,

$$\alpha^{n+1} = \alpha^n - \nabla J(\alpha) / J''(\alpha)$$

$$\alpha^{n+1} = \alpha^n - \sum_t H x_t^T (\alpha H x_t - y_t) / \sum_t (H x_t)^T H x_t$$

This procedure may be iterated a couple of turns with a forward model run in between.

One element to consider is that the source term derived above is not directly the source intensity but the concentration that the source intensity would lead to. Converting from concentrations to source intensities could then in the final output be a bit arbitrary.

The gamma dose operator is derived from Heinonen (2017) and references therein with tabulated gamma energies and range parameters. A generic gamma dose stencil is for each nuclide and model layer derived integrating in a surrounding of the cell coordinate then assumed applicable to any grid cell coordinate. The adjoint is the transpose of this generic gamma dose operator.

System for Integrated modeLling of Atmospheric coMposition (SILAM)

SILAM (System for Integrated modeLling of Atmospheric coMposition, , last access: 5 Jan 2021) is an offline 3D chemical transport model (Sofiev *et al.*, 2015). SILAM features a mass-conservative positive-definite advection scheme that makes the model suitable for long-term runs. The model can be run at a range of resolutions starting from a kilometre scale in a limited-area up to a global coverage. The vertical structure of the modelling domain consists of stacked layers starting from the surface. The layers can be defined either in z- or hybrid sigma-pressure coordinates. The model can be driven with a variety of NWP (numerical weather prediction) or climate models.

The model is used for emergency-response applications and includes radioactive transformation mechanism and various passive tracers.

Method Employed for Source Characterization

The source characterization is performed through optimization of the time-slot specific model emission. The cost function of the optimization is the sum of the squared model errors, to which an additional regularization term can be added. Here, the regularization term is set to be proportional the sum of the squares of the differences between the emission coefficients of consecutive time slots. A suitable trade-off between the smoothness of the optimized temporal evolution of the emission and the model to measurement error is found through an L-curve method, which utilizes a plot of the model error versus the residual error. The method can be used to effectively remove the length of the emission time slot from the degrees of freedom of source characterization. In other words, a short time slot of one hour can be applied, with the regularization helping to avoid overfitting in terms of temporal complexity.

The optimization procedure itself is based on a standard application of the L-BFGS-B algorithm, as applied within the Python package *scipy*. As the method is computationally light, at least for the cases studied here, a brute force method of iterative application of random initial guesses spanning up to ten orders of magnitude is applied to ensure that good optimization results are achieved. Although the vertical emission profile is fixed, the impacts profile alterations can be studied through separate simulation runs.

Severe Nuclear Accident Program (SNAP)

The Norwegian Meteorological Institute (MET-Norway) is responsible for modelling atmospheric dispersion of radioactive debris in the event of a nuclear emergency related to a nuclear accident or detonation. An additional task of the MET-Norway in a nuclear emergency is to identify unknown sources of radiation indicated by elevated levels of measurement. The basic tool used by the MET-Norway for such events is the Severe Nuclear Accident Program (SNAP) (Bartnicki *et al.*, 2011; Klein and Bartnicki, 2018).

The SNAP model was developed at the MET-Norway in 1994 as a Lagrangian particle model. The present version is fully operational at the MET-Norway and takes into account atmospheric transport and deposition of gases, noble gases and particles of different size and density emitted during nuclear accidents or explosions. SNAP can also be run remotely by experts from the Norwegian Radiation and Nuclear Safety Authority (DSA) where the Norwegian Crisis Committee is located.

Once released into the air, radioactive gases and particles are subject to advection, turbulent diffusion and deposition (dry and wet). In the SNAP calculations, the advection process is immediately followed by the diffusion process. A random walk approach is used to parameterise horizontal and vertical diffusion. When large and dense particles are released, gravitational settling is more effective than vertical diffusion, and this process is taken into account. The SNAP model has been used both for simulations of historical events, e.g. nuclear detonations in Novaya Zemlya, Chernobyl Accident (Bartnicki *et al.*, 2016), and real-time simulations, e.g. the Fukushima accident. It was tested in the ETEX-1 experiment and showed good agreement with observations (Saltbones *et al.*, 1998). SNAP is the dispersion model currently used by the MET-Norway in the Center of Excellence: CERAD CoE.

Method Employed for Source Characterization

The method for the source characterization follows the approach chosen by (Eckhard *et al.*, 2008), (Tichy *et al.*, 2020) and (Brodtkorb *et al.*, 2023). The temporal release characteristics of the release is estimated by using the observations from air filter stations and gamma stations. The release is assumed consisting of releases at disjoint time bins t_j in $[t_0, t_n]$ where each bin has a duration of one hour. These time bins are illustrated by Figure 1. The SNAP model is run forwards assuming a constant unit release of the radionuclides for each time bin t_j from a predetermined location, such as a suggested facility following the methods developed during the SLIM project. Each model run is then co-located with observations to form the source-receptor matrix M_{ij} .

$$M_{ij} = q_j(o_i)$$

where q_j is the result of the unit release at each location and o_i is the location of the observation. The release characteristics is described by the linear system

$$y = Mx$$

Where y describes the real observations at the positions o_i , and x are the different releases. The matrix M does not take into account uncertainties from the dispersion model or the atmospheric conditions and which is included by an additional term ΔM giving:

$$y = (M_0 + \Delta M)x$$

Finding a suitable $M_0 + \Delta M$ can be challenging and we solve it by the approach provided by (Eckhardt *et al.*, 2008) and (Brodtkorb *et al.*, 2023). First, we start from suitable a-priori emissions x_a , which usually are given from source-term expert or visual observations (e.g. fire, inventory fraction) at the source. The model-results of this a priori is y_a . By defining $\tilde{x} = x - x_a$ and $M\tilde{x} = \tilde{y}$ we obtain a least-squares term:

$$J_1 = ||M\tilde{x} - \tilde{y}||$$

which when solved should be close to our a priori solution. To allow for errors in observations (e.g. representativeness error), we add a matrix σ_o with observation errors on the diagonals, providing an additional least squares correction term:

$$J_1 = ||\sigma_o^{-1}(M\tilde{x} - \tilde{y})||$$

Similarly, we can add an uncertainty to our a priori (σ_x) resulting in another term:

$$J_2 = ||\sigma_x^{-1}(x - x_a)|| = ||\sigma_x^{-1}\tilde{x}||$$

This σ_x is a matrix with the errors of the a-priori on the diagonal. We started by setting the a-priori errors to $\frac{1}{2}$ of the a-priori values, or a minimal value in case of 0, dividing it by a scalar tuning factor ϵ_a between 0.1 and 1 with:

$$\sigma_x = \frac{1}{\epsilon_a} \cdot \frac{1}{2} x_a$$

To avoid sharp gradients, we introduce an additional smoothing term to our matrix:

$$J_3 = \epsilon ||D\tilde{x}||$$

The matrix D is a second derivate matrix with -2 on the diagonal and 1 on the first off-diagonals. ϵ is a scalar value for the smoothness, which we set to 1e-4.

The complete problem is formulated as a Tikhonov problem with the equation to solve for given by:

$$x = x_a + [M^T \sigma_o^{-2} M + \sigma_x^{-2} + \epsilon D^T D]^{-1} M^T \sigma_o^{-2} \tilde{y}$$

There are still a number of free parameters, i.e. the observation errors σ_o , the smoothing factor ϵ , the choice of the a-priori and the errors of the apriori. In this work we only tuned the a priori and the a priori error ϵ_a .

The solution can arrive at negative emissions which is unrealistic. In these cases, the corresponding σ_x a-priori uncertainties are reduced and thus pushing the results to be closer to the a-priori. This procedure is re-run iteratively until a solution with only positive terms has been arrived at.

Semi-infinite Gamma Dose Model

The gamma ray flux at point \mathbf{r} originating from the in-air or deposited concentration of a nuclide emitting radiation at energy E is calculated from the expression (CERC, 2012)

$$\Phi(\mathbf{r}, E) = \int \frac{f(E) c(\mathbf{r}') B(E, \mu(E)|\mathbf{r} - \mathbf{r}'|) \exp(-\mu(E)|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|^2} d\mathbf{r}', \quad (1)$$

where $c(\mathbf{r}_0)$ is the concentration of the nuclide, $f(E)$ is the branching ratio, $B(E, \mu(E)|\mathbf{r} - \mathbf{r}_0|)$ is the build-up factor, and $\mu(E)$ is the linear attenuation coefficient. In the dose rate calculator, the build-up factor is based on Berger's expression, i.e.

$$B(E, \mu(E)|\mathbf{r} - \mathbf{r}'|) = 1 + a(E)\mu|\mathbf{r} - \mathbf{r}'| \exp(b(E)\mu(E)|\mathbf{r} - \mathbf{r}'|), \quad (2)$$

where $a(E)$ and $b(E)$ are fitting constants that are readily available as tabulated data (CERC, 2012). In the atmosphere, $\mu(E)$ is not a constant, but varies as a function of air density, which in turn mostly depends on the height above sea level. $\mu(E)$ is calculated from the mass attenuation coefficient $\mu_0(E)$ through $\mu(E) = \mu_0(E)\rho_{\text{air}}$, where ρ_{air} is the average air density between the source and point \mathbf{r} . In the dose rate model, the air density is simply taken from a standard atmosphere.

Berger's expression and the corresponding fitting constants are strictly speaking not valid when the attenuation coefficient depends on \mathbf{r} , but as the air density is varying only slowly as function of altitude compared to the scale of attenuation of gamma radiation in the atmosphere, the resulting error is expected to be small. The benefit of the method is that the dose rate calculator can be applied for any height above surface, even for typical flight altitudes, where the attenuation coefficient is only a fraction of the value at the surface.

The effective dose rate D at energy E is calculated from $\Phi(\mathbf{r}, E)$ based on

$$D(\mathbf{r}, E) = C(E) \mu_{\text{abs}}(E) E \Phi(\mathbf{r}, E), \quad (3)$$

where μ_{abs} is an energy-dependent absorption coefficient and $C(E)$ a conversion factor from dose rate to effective dose rate.

The dose rate calculator approximates the radioactive cloud to be semi-infinite in the horizontal plane, which is a reasonable approximation for standard dispersion model output, as the horizontal size of the computational cell of a dispersion model is typically much larger than the relevant length scale of the attenuation. In the dose rate calculator, to simplify the calculations and to increase the computational speed, the height above the surface is approximated to equal the height above the sea level, which is a reasonable approximation for most parts of the world. If the dispersion model output is given in terms of hybrid levels, the levels are converted to height using the air density of a standard atmosphere. These approximations allow for performing the integration of Eq. 1 through pre-computed integration weights $w_{\text{dr},i}(E)$, i.e.

$$\Phi(h_{\text{dr}}, E) = f(E) \sum_i w_{\text{dr},i}(E) c_i, \quad (4)$$

where h_{dr} is the requested height for the dose rate calculation, i is the layer index of the dispersion model output, and c_i is the concentration of the nuclide inside layer i . The weights w_i are calculated using cylindrical coordinates (h, ρ, φ) from

$$w_{\text{dr},i}(E) = \int_{h_{i,1}}^{h_{i,2}} \int_0^{\rho_{\text{max}}} \frac{\rho B(E, \mu(E)r(h, \rho)) \exp(-\mu(E)r(h, \rho))}{2r(h, \rho)^2} d\rho dh, \quad (5)$$

where

$\rho_{\text{max}} \rightarrow \infty$, and we have defined $r \equiv ((h - h_{\text{dr}})^2 + \rho^2)^{\frac{1}{2}}$. $h_{i,1}$ and $h_{i,2}$ are the lower and upper bounds of layer i , respectively. In the expression, the dependency on the polar angle φ has already been integrated out. However, the lack of an analytical solution prevents from doing that for the radial distance ρ .

Currently, instead, the radial part of the integral is computed numerically, with ρ_{max} set to 1500 m. While performing the numerical integration, the lengths of the integration steps in the h and ρ directions are optimized based on the magnitude of μ . The integration weights are calculated separately for in-air and deposited concentrations of nuclides.

Because of the approximation of taking the air density from a standard atmosphere, Eq. 5 needs to be solved only at the start of the dose rate calculation, but independently for all output heights h_{dr} and gamma emission energies E . However, tabulated values for $\mu_0(E)$, $a(E)$ and $b(E)$ exist only for specific energies, and in its two-dimensional form, the integral is also somewhat heavy computationally. Thus it is beneficial to perform the calculation of the vertical integration weights for a restricted number of different energies, with interpolation of the weights applied for all intermediate emission energies. The weights $w_{\text{dr},i}$ are computed for all the energies listed in Table 1, rather than for all of the emission energies of all of nuclides in the output of the dispersion model. Each emission at energy E_{ems} is set to contribute to the total emission through a linear combination of the closest pre-computed values, i.e. weights for different pre-computed energies can be 3 defined as

$$w_{\text{ems-}} = \frac{\exp(E_{\text{ems}}) - \exp(E_-)}{\exp(E_+) - \exp(E_-)} f(E_{\text{ems}}) \quad (6)$$

$$w_{\text{ems+}} = (1 - w_{\text{ems-}}) f(E_{\text{ems}}), \quad (7)$$

if the growth of the gamma flux rate as function of energy is approximated to be exponential. E_- is the closest pre-computed energy value below the emission energy E_{ems} and E_+ is the closest pre-computed energy value above it. $w_{\text{ems-}}$ and $w_{\text{ems+}}$ are the corresponding weights assigned for the precomputed energies. Final energy-dependent weights w_{ems} are acquired by summing the individual weights obtained for all the gamma emission energies of a specific nuclide. By combining Eqs. 4, 6, and 7 we end up with the expression for the total gamma ray dose rate $D_n(h_{\text{dr}})$ originating from the concentration of nuclide n , i.e.

$$D_n(h_{\text{dr}}) = \sum_{i,j} C(E_j) \mu_{\text{abs}}(E_j) E_j w_{\text{ems},j} w_{\text{dr},i}(E_j) c_i, \quad (8)$$

where the index j runs over all the energies in Table 1 and i over all the vertical layers of the dispersion model output. As discussed, the weights $w_{\text{ems},j}$ and $w_{\text{dr},i}(E_j)$ need to be calculated

only at the start of the run. The emission energies E_{ems} and the corresponding branching ratios $f(E_{\text{ems}})$ are obtained from a database maintained by STUK.

Table 1. Tabulated values of the parameters of the dose rate calculation (CERC, 2012). $\mu_{0\text{m}}$ is the linear attenuation coefficient at sea level.

E (MeV)	$\mu_{0\text{m}}$ (1/m)	a	b	μ_{abs} (Gy m ²)	C (Sv/Gy)
0.01	0.623	0.025	-0.0464	7.43×10^{-16}	0.00296
0.015	0.187	0.0947	-0.0484	3.12×10^{-16}	0.0183
0.02	0.0893	0.2652	-0.0463	1.68×10^{-16}	0.0543
0.03	0.0411	1.055	-0.0192	0.721×10^{-16}	0.191
0.05	0.0253	3.498	0.0729	0.323×10^{-16}	0.557
0.065	0.0226	4.209	0.1169	0.278×10^{-16}	0.63
0.1	0.0195	4.033	0.1653	0.371×10^{-16}	0.765
0.2	0.0159	2.678	0.1678	0.856×10^{-16}	0.703
0.5	0.0112	1.748	0.1014	2.38×10^{-16}	0.689
1.0	0.00821	1.269	0.0559	4.47×10^{-16}	0.732
1.5	0.00668	1.040	0.0338	6.12×10^{-16}	0.765
2.0	0.00574	0.891	0.0215	7.50×10^{-16}	0.791
4.0	0.00398	0.5879	0.0022	12.0×10^{-16}	0.850
10	2.65e-3	0.3113	-0.0194	23.1×10^{-16}	0.935

ETEX-1

After the Chernobyl accident in April 1986 and the adoption of the Convention on Early Notification of a Nuclear Accident (IAEA, 1986), the International Nuclear Safety Advisory Group (INSAG) of the International Atomic Energy Agency (IAEA) recommended inter alia that the IAEA should, in collaboration with the World Meteorological Organisation (WMO), review and intercalibrate the models of atmospheric transport of radionuclides over short and long distances and of radionuclide deposition on terrestrial surfaces, and establish a database for validation studies of these models.

Following this recommendation, the joint IAEA/WMO Atmospheric Transport Model Evaluation Study (ATMES) was initiated in November 1986. The objective of ATMES was to compare the evolution of the radioactive cloud (I-131 and Cs-137) with the evolution predicted by mathematical models for atmospheric dispersion, using as input only the estimated source term of the Chernobyl accident.

The ATMES suffered, however, from a number of weaknesses regarding lack of monitoring data and large uncertainties regarding the source term. Therefore, it was decided to carry out a tracer experiment in Europe. The sponsoring organisations were the European Commission (EC), the World Meteorological Organization (WMO) and the International Atomic Energy Agency (IAEA), and later joined by the US Department of Energy (USDOE).

The experiment was named ETEX, European Tracer Experiment (Graziani *et al.*, 1998; ETEX web-site, 2019). It was designed to test the readiness of interested services to respond in the case of an emergency, to organise the tracer release and compile a data set of measured air concentrations and to investigate the performance of long-range atmospheric transport and dispersion models using that data set. In total, thirty-six organisations around the world were involved in the project.

Sampling network

The sampling network consisted of 168 ground-level sampling stations in western and eastern Europe. National meteorological services hosted the samplers at a number of WMO synoptic stations over their territory. Thus, ETEX could take advantage of this existing network, which is homogeneously distributed throughout Europe and linked to the WMO.

A final number of 168 sampling stations were selected, almost all located at existing WMO stations. Three samplers were located in the North Sea: one on a Dutch oil platform, the other two on gas platforms. The average spacing between two sampling stations in the resulting configuration was about 80 km.

Each station was labelled with one or two letters identifying the Country where it was located, and numbered sequentially.

It was planned to start the sampling operations at each station about 6 hours before the expected time of tracer arrival to obtain contemporaneous measurements of the tracer background levels and to ensure that the plume arrival was not missed. Each station was designed to sample over a period of 72 consecutive hours (24 three-hour samples), with sampling starting time progressively delayed from West to East. The stations closest to the source started sampling 3 hours before the release start; the most distant stations ended sampling 90 hours after the release start. The sampling network is depicted in Figure 3.

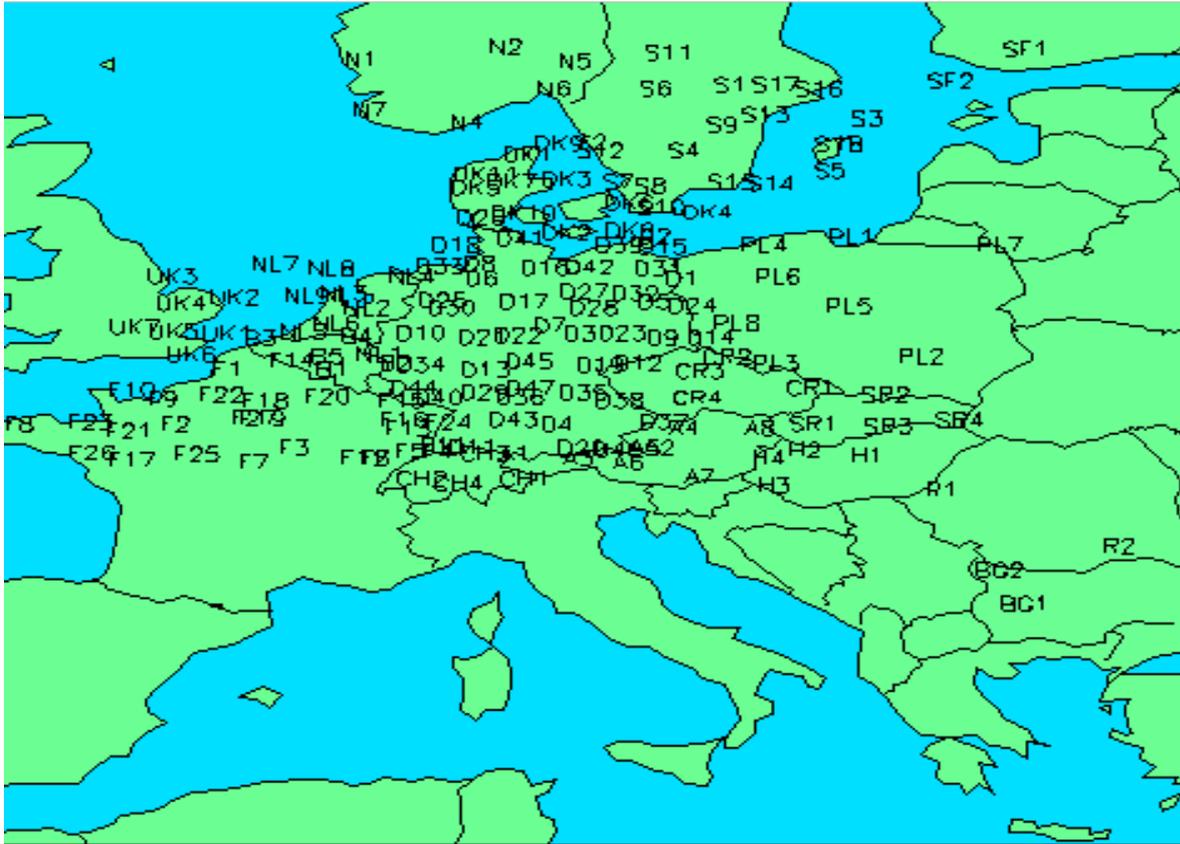


Figure 3 ETEX Sampling network (ETEX web-site, 2019).

Synoptic situation

Weather predictions suggested the following conditions on Sunday 23 October, 1994:

- the presence of a rather strong West to South-westerly flow, advecting the tracer during the experiment over several tracer stations
- no centre of high- or low-pressure, and no extending ridges or troughs, would have passed close to the release site
- no frontal systems would have passed the release site shortly before, during or after the release

Therefore, on Friday 21 October, 1994, the alert procedure was started.

23 October, 1994

A deep low, 975 hPa, to the East of Scotland was slowly moving north, maintaining a strong south-westerly flow over the release-site (Rennes). The advected air was unstable, with showers, some accompanied by thunder and squall-lines. Similar observations could be made from satellite pictures. The 12:00 UTC radio sounding of Brest showed a temperature profile which was unstable with respect to moist air, allowing the development of shower clouds up to about 28000 ft. Also the radio sounding of Paris showed an unstable atmosphere but with lower water vapour content. At both locations, the upper winds were Southwest and rather strong. The release started at 16:00 UTC.

24 October, 1994

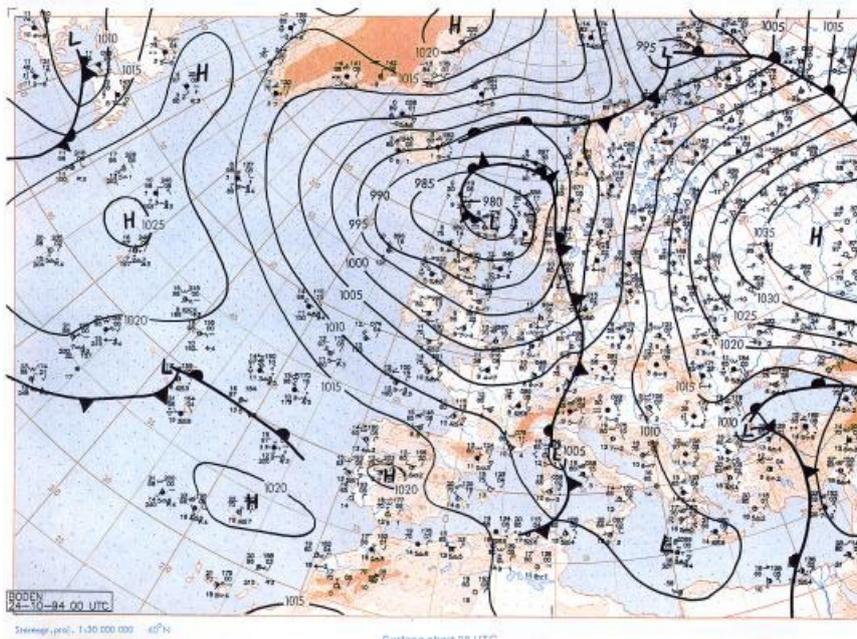


Figure 4 24 October, 1994, 00 UTC (ETEX web-site, 2019).

There was still an unstable flow over the release site and the advection area. However, because of the northerly movement of the controlling low over the North Sea, the wind in the advection area decreased. The expected cold front was to be seen south of Ireland.

25 October, 1994

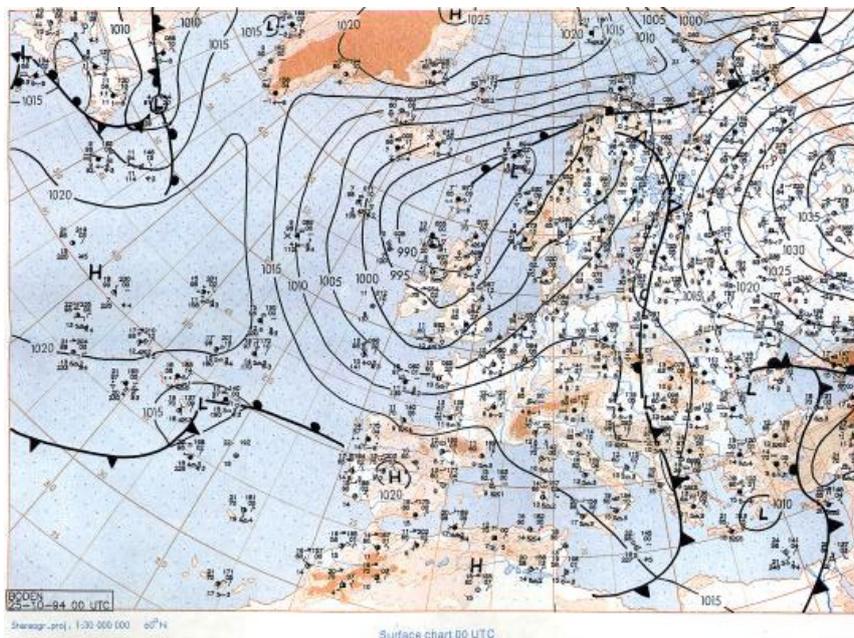


Figure 5 25 October, 1994, 00 UTC (ETEX web-site, 2019).

The further deepening of the mentioned cold front had not developed. The system was to be seen as a minor secondary low, at 52°N, 5°E. The cold front over The Netherlands, an instability front with showers, had the pressure pattern of a trough. The wind was backing more to the south with the approach of the front during the day, and after the passage of this front the wind was veering to the Southwest.

26 October, 1994

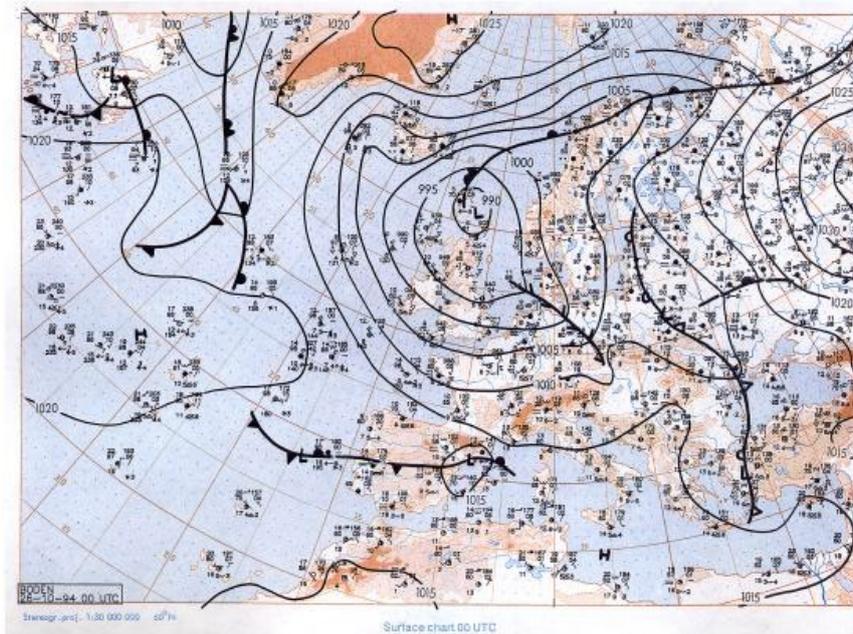


Figure 6 26 October, 1994, 00 UTC (ETEX web-site, 2019).

There was still a complex low-pressure system over the North Sea and Scotland. Shower weather with a wind tending to veer a bit over Western Europe towards west-southwest. The high-pressure cell over the Black Sea indicated that southerly winds could block any further movements of the tracer cloud towards the East.

Tracer release

Atmospheric tracers were released in the form of a homogeneous air stream containing a few percent of perfluoromethylcyclohexane (PMCH) tracer. The gas stream passed through a small chimney where the gas was released at the top.

The release started at 16:00 UTC on October 23, 1994, and lasted 11 hours and 50 minutes. 340 kg of the non-depositing inert gas PMCH (perfluoromethylcyclohexane) were released from Monterfil (48°03'30"N, 2°00'30"W) at an average flow rate of 8.0 g/s.

Results of DERMA

The results are based on all available measurements, both non-zero measurements and non-detections, which are used as described in the method section. The selected start time, t_0 , is September 23, 1994, 00:00 UTC, i.e. 16 hours before the actual release began, while end time, t_N , is September 25, 1994, 00:00 UTC, i.e. 20 hours after the release ended. Further, we use time bins of 3-hour duration, i.e. $\Delta t_{\text{bin}} = 10800$ s.

First, Figure 7 shows the results based on the deterministic meteorological forecast. From this figure, we see that the timing of the release is reasonably well estimated and the estimated released amount is roughly 160-225 kg, while the true release was 340 kg. We see that there is an additional peak at around 35 hours, 5-10 hours after the actual release ended.

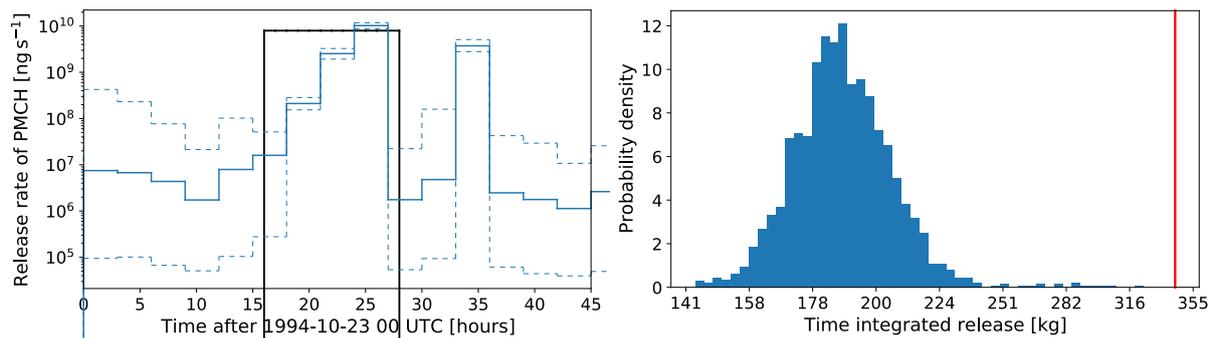


Figure 7 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability density for time integrated release. The red vertical line indicates the true release.

Next, Figure 8 shows the modelled air concentrations as function of observations.

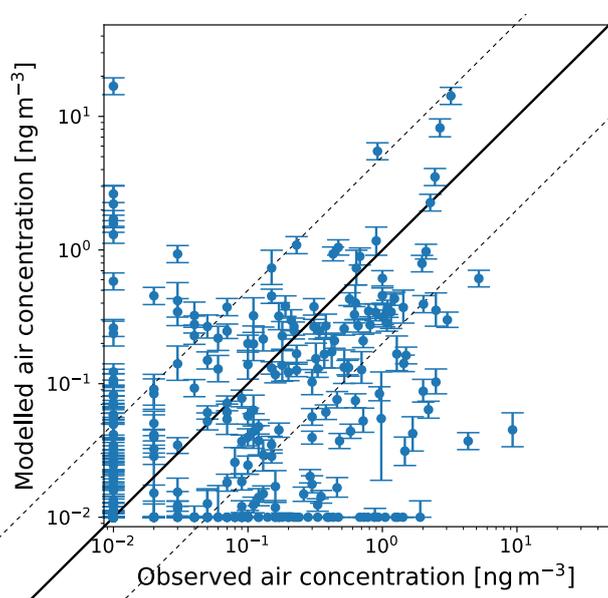


Figure 8 Modelled concentrations as function of observations. The solid black line shows perfect correlation, while the dashed black lines indicate a factor of 5 between model and observation.

Next, Figure 9 shows the results based on the meteorological ensemble forecast. Figure 9 shows all 21 ensemble members in the same plot. This shows that all estimated source terms agree that the release peaks at around 21-27 hours, while most source terms also have a second peak at within the period 30-39 hours. The estimated time integrated release varies from roughly 100 kg to 400 kg.

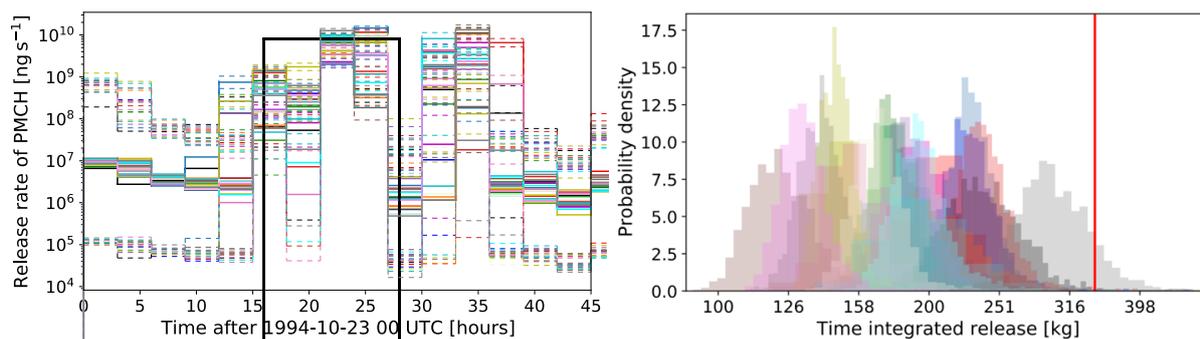


Figure 9 Left: Estimated release rates. The solid colored lines show the medians for the different ensemble members, and the dashed colored lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability density for time integrated releases for all ensemble members. The colors correspond to the colors of the release rates. The red vertical line indicates the true release.

Finally, Figure 10 shows the results based on using all the meteorological ensembles in the same source term inversion as described in the method section. Figure 10 thus shows a single source term based on all 21 meteorological ensemble members. Qualitatively, this does not differ significantly from the other results. On the other hand, it does not seem to be the best estimate either. This makes sense, since this source term somehow must combine the information from all the ensemble members. The best source term estimate, on the other hand, should be the one that uses the “best” ensemble member.

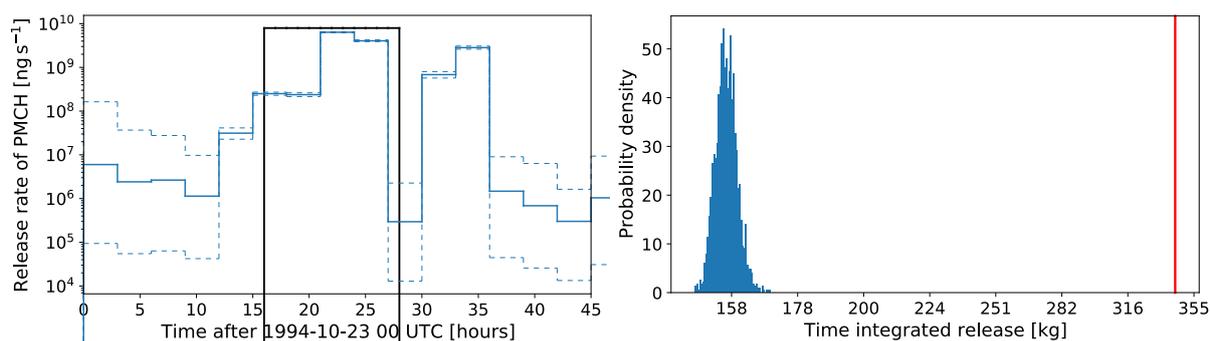


Figure 10 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability density for time integrated release. The red vertical line indicates the true release.

Results of MATCH

This experiment was run with ERA5 weather data on 0.2 degree resolution covering the period 12UTC 23 October to 00UTC 26 October, 1994.

A subset of 50 out of 930 non-zero observations were extracted by first restrict to observations valid after the release ended (9 UTC 24 to 00 UTC 26, October). The subset was then extracted by so called weighted bootstrapping with withdrawal, that implies that a random selection will favour larger measured values, and that an observation only could be selected once. The very ten largest observation in the set were though explicitly selected to avoid being randomly excluded. Locations for the selected stations are shown in Figure 11.

The retrieved vertical source distribution is shown in Figure 12 (top) and total release rate (bottom). The timing of the retrieved release very much catches the true release period, although 86 kg of the release is retrieved less than the 340 kg in reported emitted. Anyhow the transport pattern is well caught as shown in Figure 13 and Figure 14. In Figure 15 finally we present scatter diagrams for dependent and independent observations.

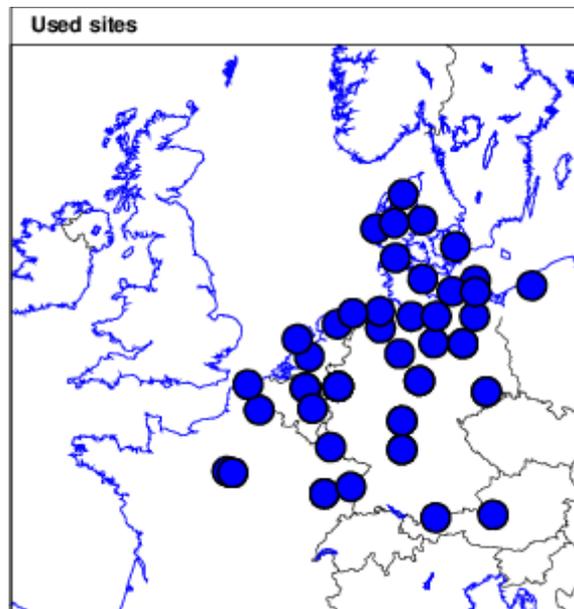


Figure 11 The sites for the 50 observations used for the source inversion.

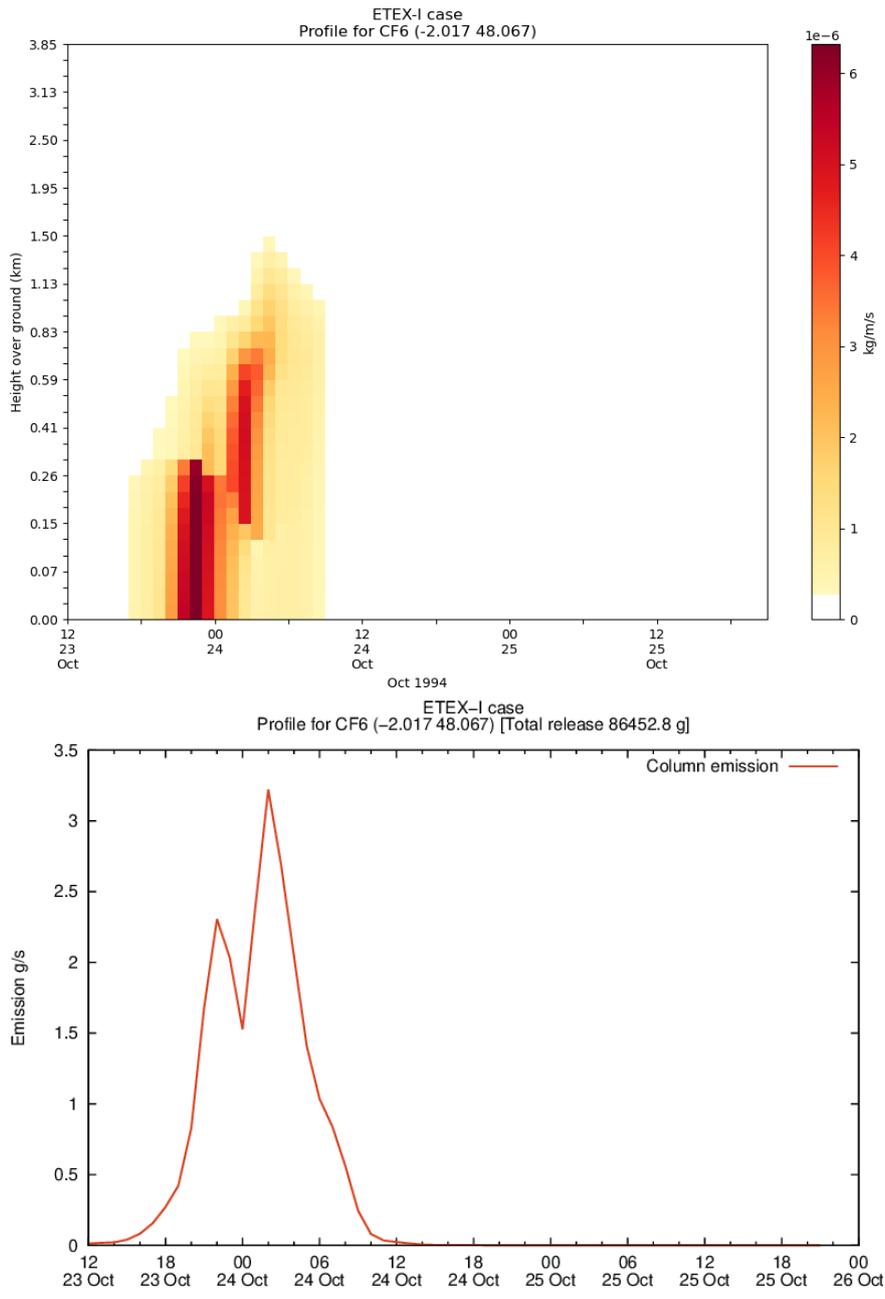


Figure 12 Retrieved source profile given as vertical extension (top) and emission rates (bottom).

ETEX case Mon 24 Oct 1994 09:00 UTC ETEX case Mon 24 Oct 1994 09:00 UTC

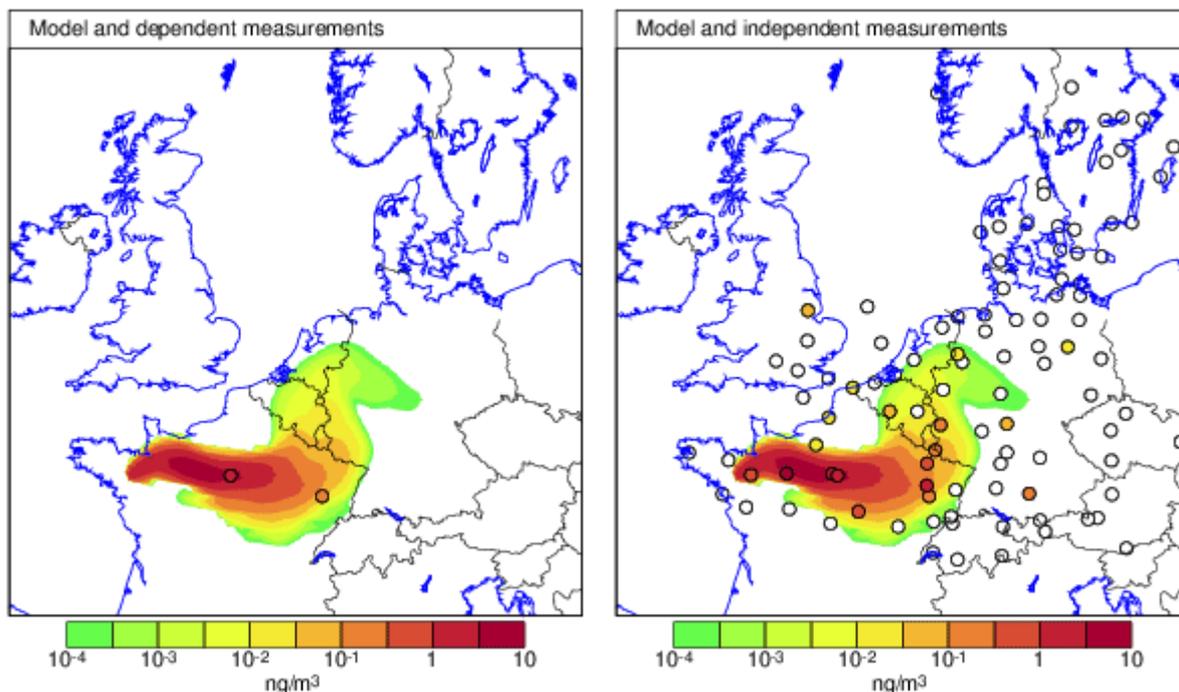


Figure 13 Assimilated ETEX-I plume at 09 UTC 24 October, 1994, together with the observations used (left) and all observations (right).

ETEX case Tue 25 Oct 1994 09:00 UTC ETEX case Tue 25 Oct 1994 09:00 UTC

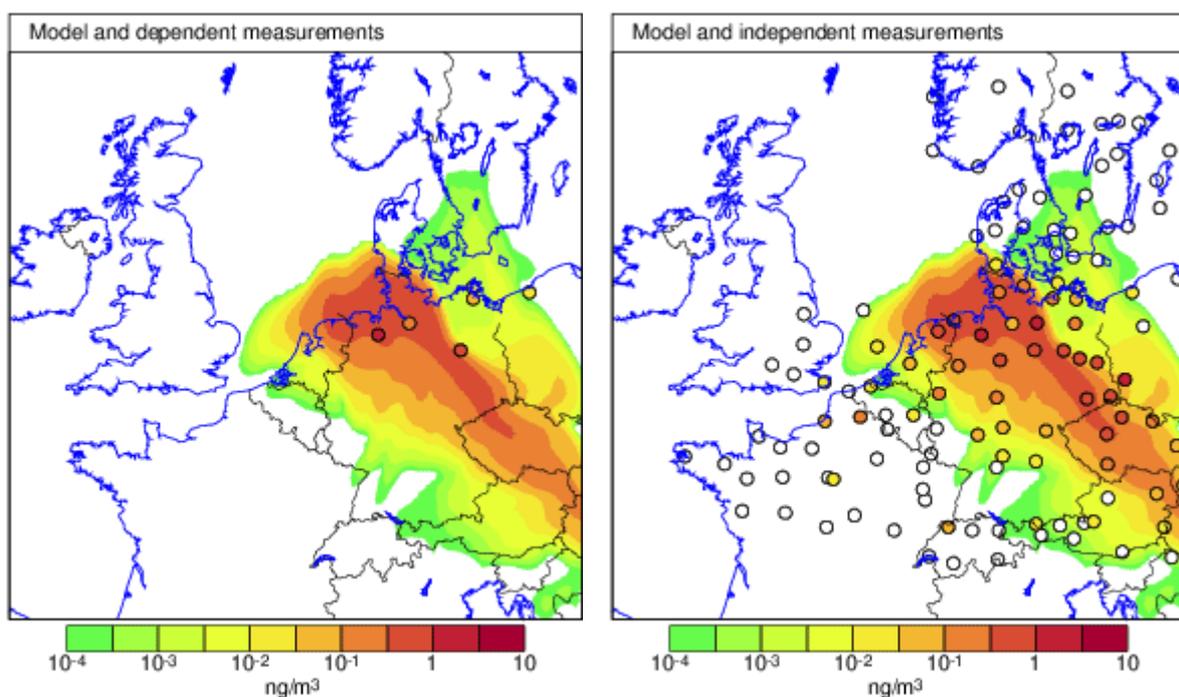


Figure 14 Assimilated ETEX-I plume at 09 UTC 25 October, 1994, together with the observations used (left) and all observations (right).

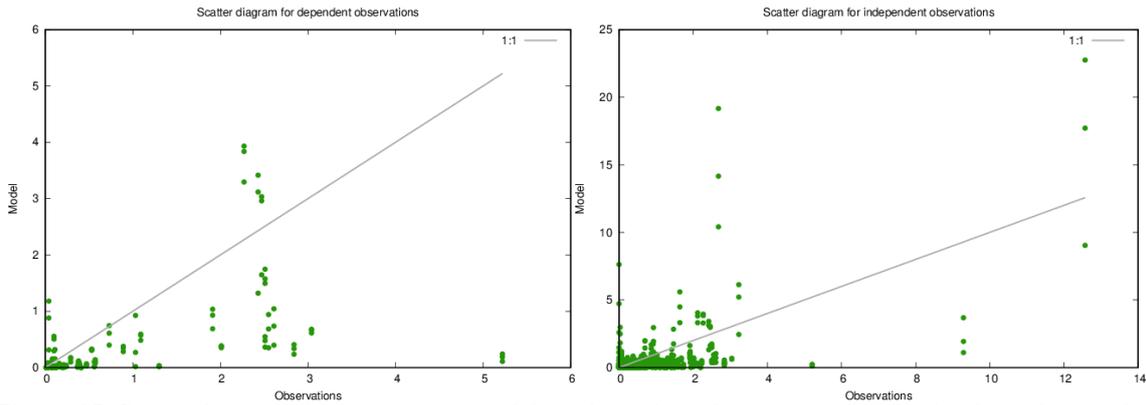


Figure 15 Scatter diagram observations vs. model for dependent observations (left) and independent (right).

Results of SILAM

The source term retrieval was performed through optimizing scaling coefficients for a linear combination of one-hour long emission slots by minimizing the root-mean-square error (RMSE) between modelled values and observations, with an added cost term penalizing non-smooth solutions. The optimization was performed for every ensemble member separately. The model ensemble was run with a vertical emission profile of 8 to 50 m, as it represents a realistic estimation of the actual source profile, although already based on ERA5 results, this may lead to an underestimation of the source.

The results of the source term retrieval are displayed in Figure 16, which shows the 10th, 50th and 90th percentiles for the full ensemble, as well as the same for the ten ensemble members with the smallest RMSEs with regard to the observations. By calculating the gradient of the RMSE with respect to changes in the retrieved emission profile, it was ensured that the entire temporal profile displayed in the figure is sensitive to the observed values.

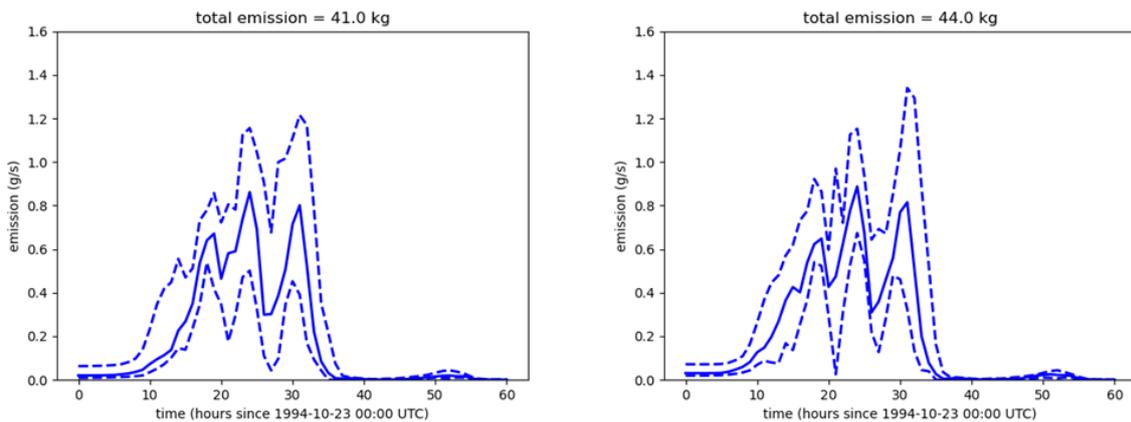


Figure 16 Retrieved emission from the SILAM model based on the full ensemble (left) and the best ten ensemble members in terms of the RMSE (right). The solid lines indicate the median emission and the dashed lines the 10th and 90th percentiles. The indicated total emissions are based on the median emission curve.

While the temporal window of the release is obtained with a similar accuracy as with the deterministic ERA5 data, the magnitude of the release is underestimated even more strongly than with the ERA5 data. Despite penalizing less smooth solutions in a similar way as with the ERA5 data, there are also more fluctuations in the ensemble median result. The negative bias of the retrieved emission is mostly due to the emission plume not properly reaching the correct measurement stations in any member of the ensemble. As shown with the deterministic ERA5 results, this is partly corrected if a vertical emission profile extending to a

higher elevation is selected (or even one that assumes full initial mixing in the atmospheric boundary layer), but the effect may also be accidental, as optimally the SILAM model itself should decide on the amount of vertical mixing.

Based on the temporal median curve, the estimated total emission is 41 kg for the full ensemble and 44 kg for the ten best performing members. However, this method may underestimate the emissions due to the fact that the emissions may peak at different times for different members. Computed over the totals for all individual ensemble members, the 10th, 50th and 90th percentiles for the total emission are 35 kg, 45 kg and 53 kg, respectively, and 40 kg, 46 kg and 57 kg for emission of the ten best performing members. The ensemble member with the smallest RMSE corresponds to a total emission of 63 kg.

Results of SNAP

The SNAP model was first run in forward mode to examine if the model is able to match the observations. A setup with wide initial spread in the entire grid-cell and with full mixing within the atmospheric boundary layer gave realistic results as can be seen in Figure 17 with only a slight overestimation of the model results with an NMB of 0.23.

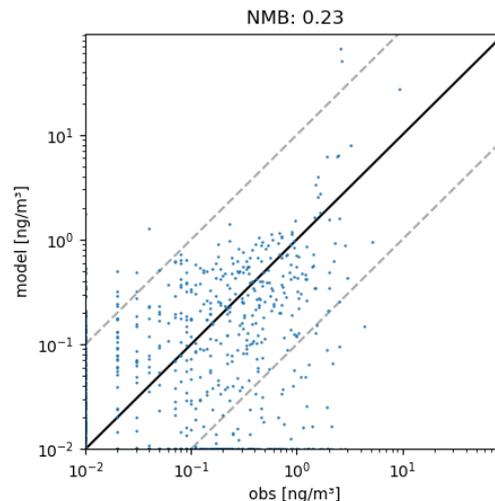


Figure 17 Scatterplot SNAP model results vs observations with known emissions giving a slight overestimation of the model results. The gray lines indicate a factor 10 over- and underestimation.

As a second test, we ran the model with the known a priori emissions. Depending on the factor ϵ_x , i.e. the confidence factor for the a-priori, we retrieved total emissions of 206 g/s for high confidence (1) to 106 g/s for medium confidence (0.4). For lower confidence, the results showed lower releases rates, but indicated many of the iterations steps lead to negative emissions and were thus have disregarded. A reduction of total emissions is to be expected as the model showed a positive bias compared to real emissions.

To test sensitivity of the model to the a priori term, we additionally ran tests with initial guesses which did not correspond to the real a priori. From the NKS-SLIM project, the results showed highest likelihood of emissions between 1994-10-23T11:00 and 1994-10-24T11:00, which we used as start and end time for our modified a priori. We scaled our emission using three different constant a priori emission rates: 1 g/s, 10 g/s and 100 g/s.

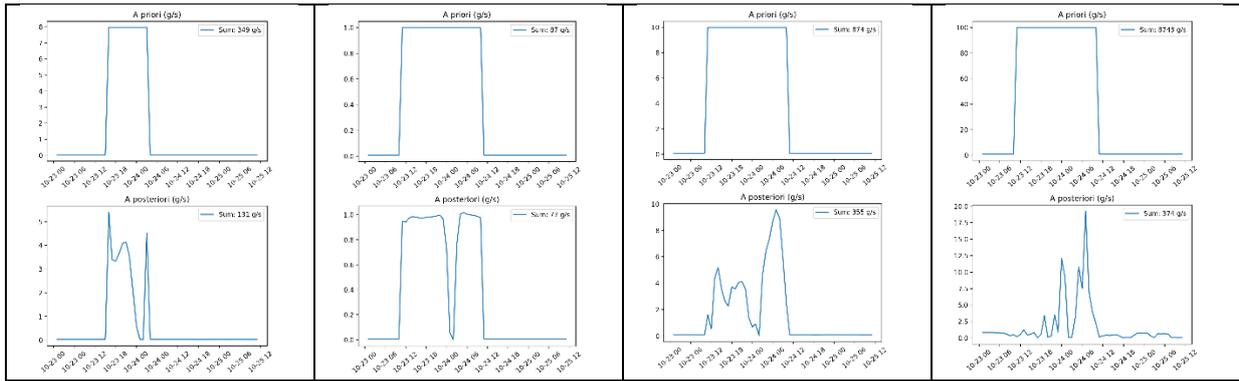


Figure 18 Results of SNAP inversion with the real emission as a priori and with 3 a priori with longer emissions in three orders of magnitudes. All a posterioris have been selected at ϵ_x of 0.5.

The results of the inversion can be seen in Figure 18 and gives in the three case 1, 10, 100 g/s total emissions of 77 kg, 355 kg and 374 kg, compared to the real emissions of 340 kg. The method is not able to increase a priori emissions that are too low compared to real emissions, but manages to reduce emissions when a high initial emission rate is used. The temporal profile corresponds most to the real release best for the 10 g/s emissions a priori, although the real constant time-profile was not estimated. The a priori with the largest overestimation tends to be less smooth with short bursts of emissions reaching up to 20 g/s.

The method was run with all ensemble members of the meteorological ensemble and the inversion results have been computed for each ensemble-member separately and results for mean, 70th and 90th percentiles of the ensembles provided in Figure 19. The results with a priori of 10 g/s emissions over a longer time-span provided best results for the total emissions, although the emissions are offset by 6 hours to a later time and the total amount is still somewhat lower than the true total release.

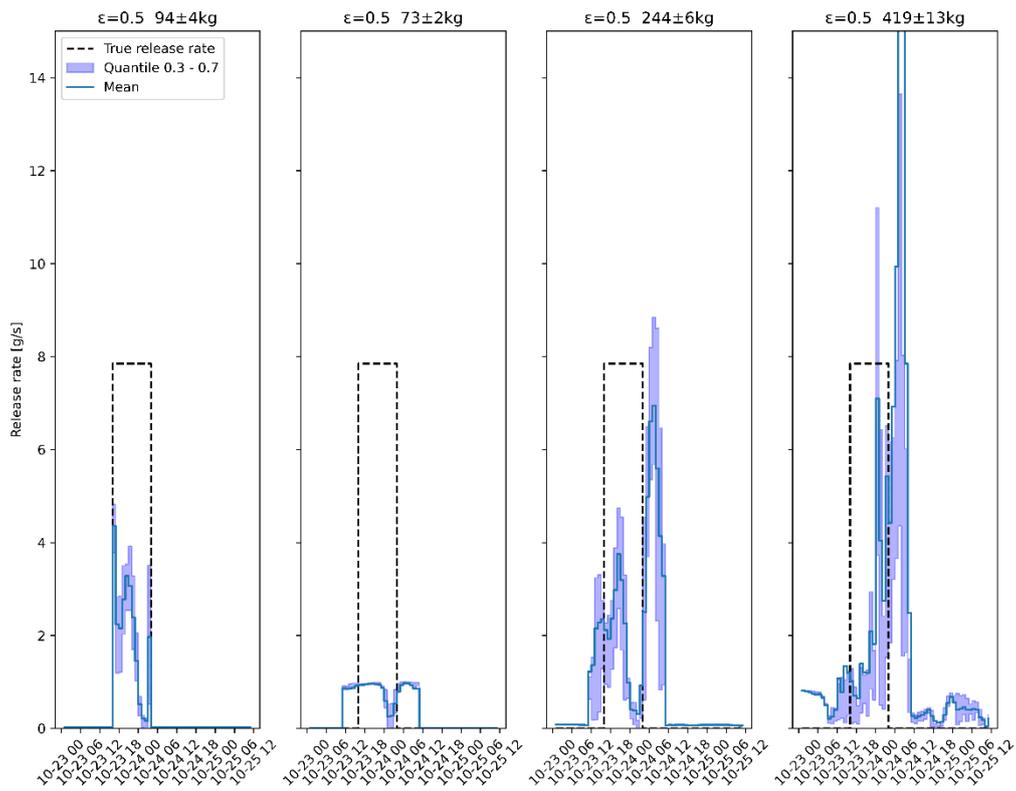


Figure 19 Ensemble results with true a priori, and a priori with 1 g/s, 10 g/s and 100 g/s.

The Autumn 2017 Case of Ru-106

During the period 3–6 October 2017, the Incident and Emergency Centre of the International Atomic Energy Agency (IAEA) was informed by Member States that low concentrations of Ru-106 were measured in high-volume air samples in Europe. The detected isotopes did not contain any other radionuclides (e.g. other fission products such as Cs-137) and were at levels far below those requiring public protective actions, however not accounting for the anticipated higher concentrations at the release site. Corresponding data and information were obtained from the IAEA (2017a, 2017b). The data comprise 387 measurements of Ru-106, some of which correspond to levels below minimum detectable activities. The data are time-average concentrations corresponding to varying time periods of up to seven days, cf. Figure 20 below.

From a meteorological point of view, seven days can be a long time with potentially a number of meteorological phenomena such as front passages taking place at the release site within the period. Possible sampling scenarios include evenly distributed low concentrations at the station site throughout the sampling period, or brief high concentrations corresponding to a narrow plume passing over the site in a short while. Therefore, such measurement data should possibly be discarded in a localization study. The discarded data can, however, be used for verification purposes.

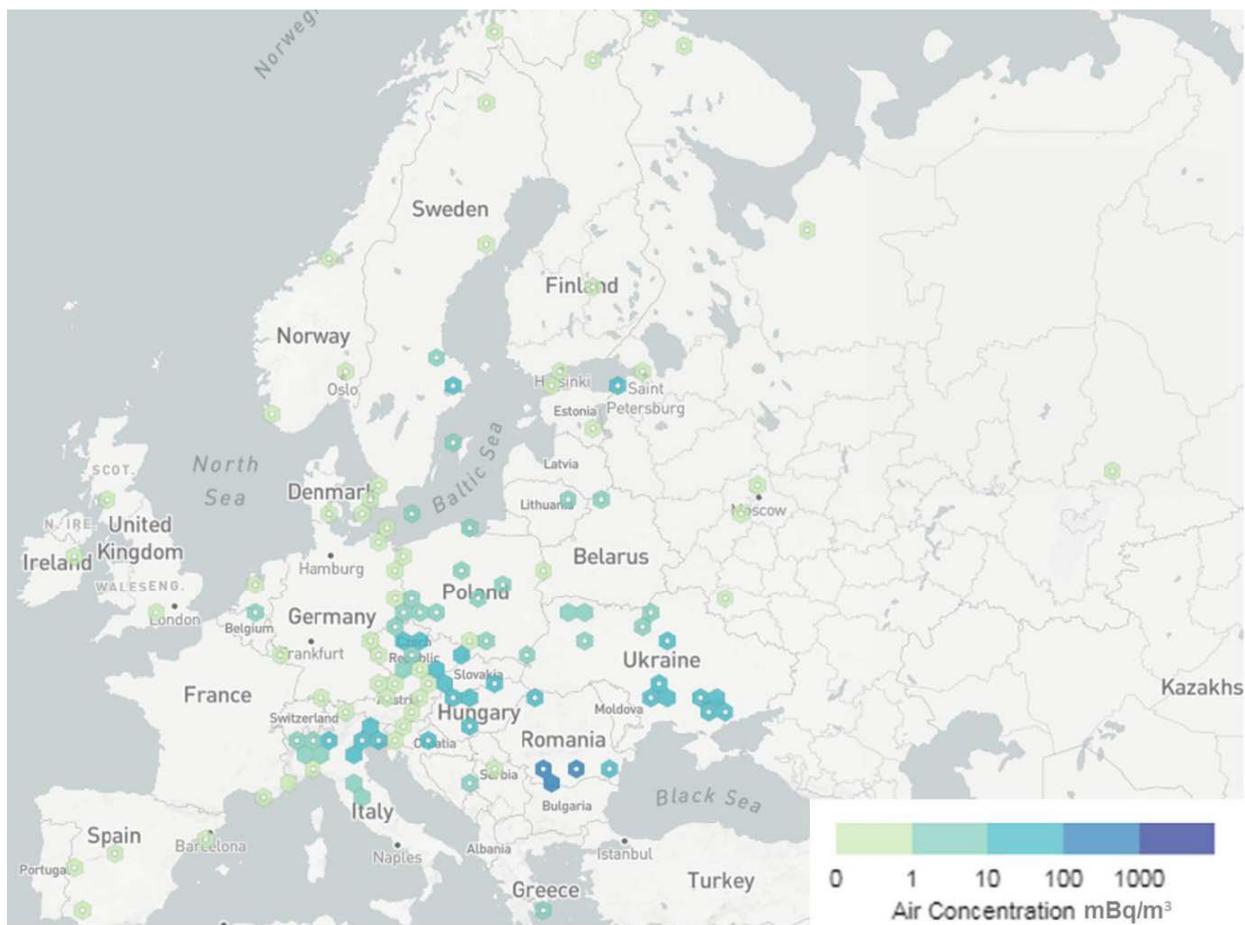


Figure 20 Locations where concentrations of Ru-106 in the air have been reported to the IAEA (IAEA, 2017a; 2017b). The measurements were taken during different sampling periods ranging from daily to weekly.

Results of DERMA

The results are based on all available measurements, both non-zero measurements and non-detections. Measurements were obtained from Masson *et al.* (2019). The source location is assumed to be the Mayak nuclear facility, and the start time and end time were selected based on estimates from previous studies, cf. e.g. Tølløse *et al.* (2021). The selected start time, t_0 , is September 23, 2017, 00:00 UTC, while end time, t_N , is September 27, 2017, 00:00 UTC. Further, we use time bins of 6-hour duration, i.e. $\Delta t_{\text{bin}} = 21600$ s.

First, we show the result based on the deterministic meteorological forecast, see Figure 21.

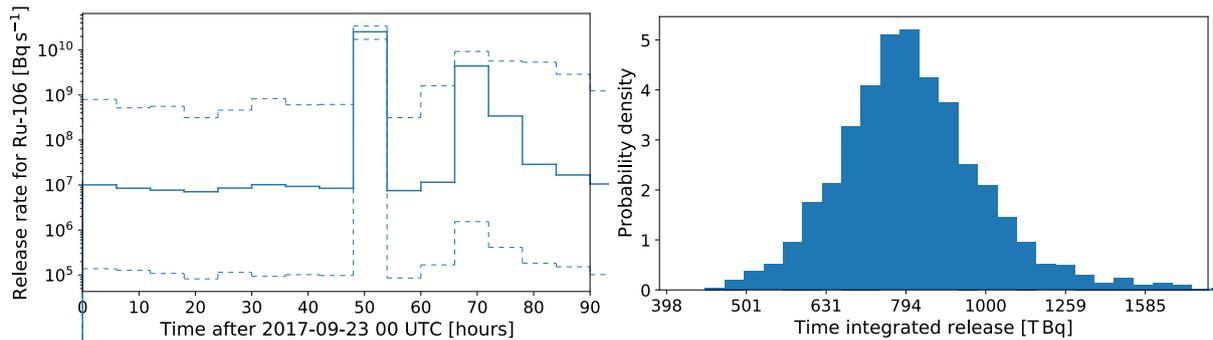


Figure 21 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. Right: Probability density for time integrated release.

The resulting source term has a large peak after about 48 hours, which is at 0 UTC on September 25, 2017. Further, there is a release, although much more uncertain, from 18 UTC, September 25 to 6 UTC, September 26. Both the timing and magnitude of the release is similar to estimates by previous studies, cf. e.g. Masson *et al.* (2019), Le Brazidec *et al.* (2020) and Tølløse *et al.* (2021).

Next, Figure 22 shows the modelled air concentrations as function of observations.

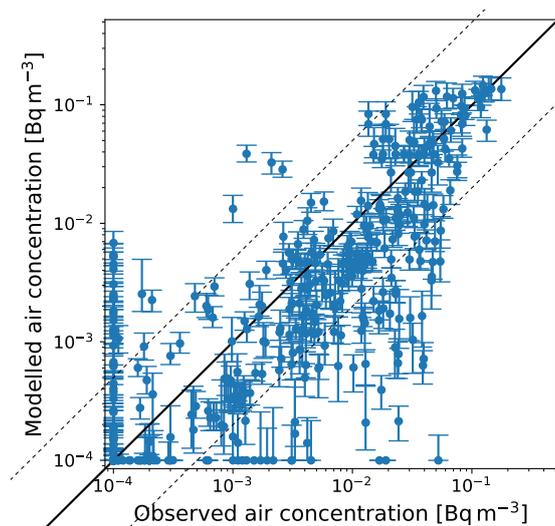


Figure 22 Modelled concentrations as function of observations. The solid black line shows perfect correlation, while the dashed black lines indicate a factor of 5 between model and observation.

Next, we show the results based on the meteorological ensemble forecast. Figure 23 shows all 21 ensemble members in the same plot. From this figure, we see that the estimated source terms depend strongly on the meteorological ensemble used. This is true for both the time of the peak of the release and the magnitude of the release. As opposed to the ETEX case where all source terms looked qualitatively similar, the source terms for this case vary significantly. This makes sense considering the long travel distance from Mayak to most filter stations as well as the low temporal resolution of measurement data compared to ETEX.

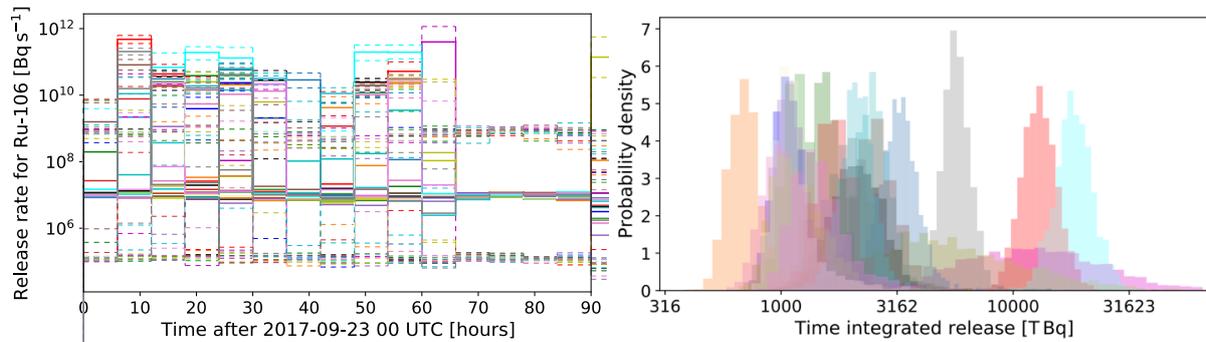


Figure 23 Left: Estimated release rates. The solid colored lines show the medians for the different ensemble members, and the dashed colored lines show the 10th and 90th percentiles. Right: Probability density for time integrated releases for all ensemble members. The colors correspond to the colors of the release rates.

Next, we show the results based on using all the meteorological ensembles in the same source term inversion. Figure 24 thus shows a single source term based on all 21 meteorological ensemble members.

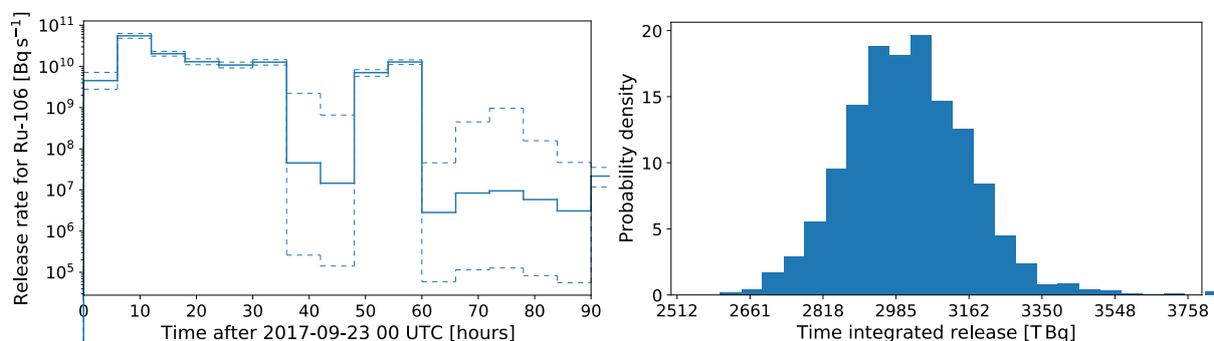


Figure 24 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. Right: Probability density for time integrated release.

The estimated source term in Figure 24 is quite different from any of the source term estimates based on single ensemble members. We suspect that the different ensemble members constrain release rates in different time bins, and therefore the resulting release profile is constrained for most time bins. However, if the different ensemble members predict the same release at different times, this may lead to an overestimation of the source. Thus, when the meteorological uncertainties are large, this may not be the ideal way of making use of the meteorological ensemble prediction. Instead, one should perhaps look at the individual predicted source terms to get an idea about the possible source terms and the uncertainty of these estimates.

In addition, Figure 25 shows the result when excluding the four ensemble members that give the worst fit with data, determined by looking at the maximum likelihood, which is a measure of how well model predictions match the data.

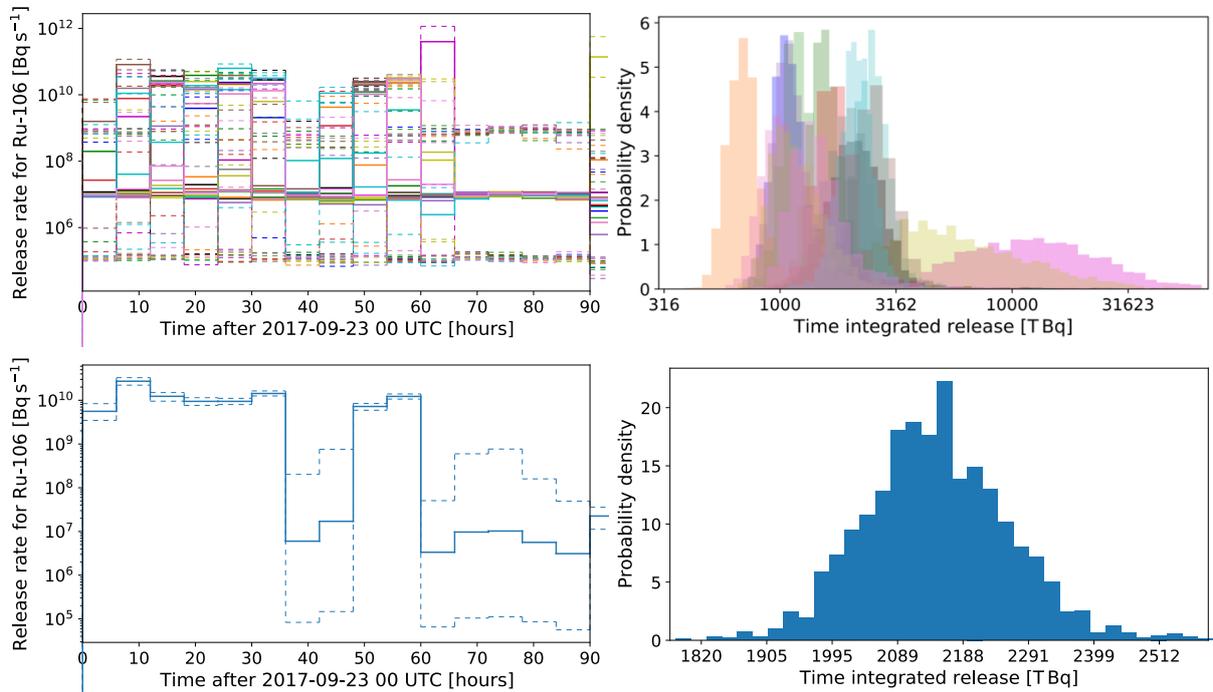


Figure 25 Upper figures are like Figure 23 but excluding the four members with lowest maximum likelihood. Lower figures are like Figure 24 but excluding the four members with lowest maximum likelihood.

The overall shape of the release profile in Figure 25 is very similar to that in Figure 24. However, we see that the magnitude of the release is predicted to be somewhat smaller. Again, the release profile is quite different from the release rates based on the individual ensemble members.

Finally, Figure 26 shows the ensemble member with the highest likelihood.

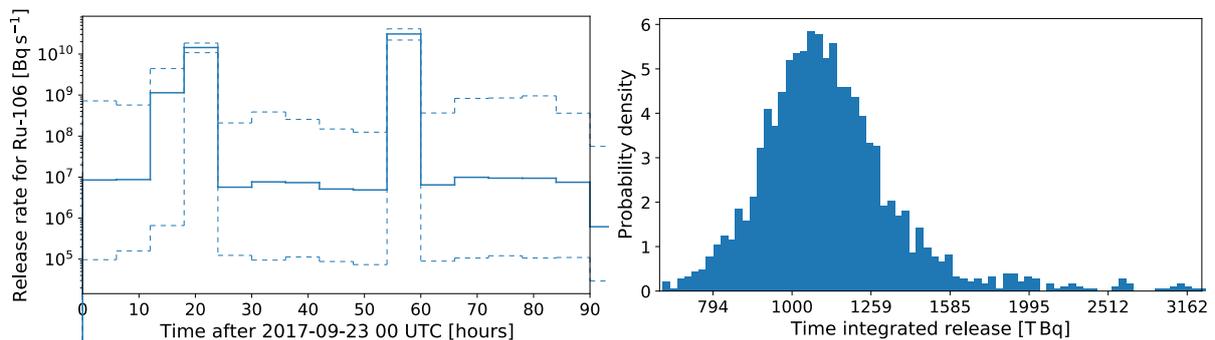


Figure 26 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. Right: Probability density for time integrated release.

Although the different source terms are not exactly identical, they all suggest that there was a release during the first half of September 25, while most also predict that part of the release occurred on September 23.

Results of MATCH

The met data used were IFS data on 0.2 degrees for the period 25 September to 5 October, 2017.

From out of 131 measurements with up to 36 hours sampling time a reduced set of 30 measurements were randomly selected by weighted bootstrapping (with withdrawal). The 10 largest measured values were taken before the random selection. Remaining measurements were used for validation. Figure 27 shows the sites for the selected measurements, and sites for validation.

The retrieved source profile for the Mayak location is shown in Figure 28 and Figure 29. The rather deep source may to some extent originate from long-range backward adjoint transport where the adjoint plume will be vertically distributed during the course of the transport. A total amount of 2×10^{12} Bq is retrieved as total activity released.

Figure 30 shows scatter plots for dependent and independent observations. A rather good fit to independent observations as well is seen.

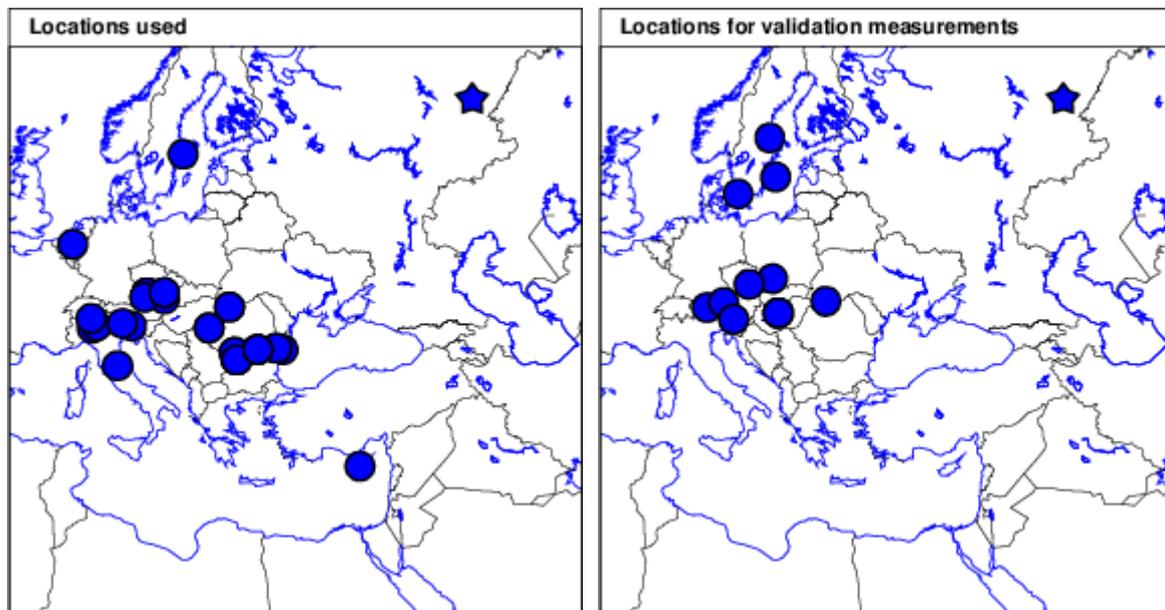


Figure 27 Sites for the reduced set of 30 measurements (left) and sites for validation measurements.

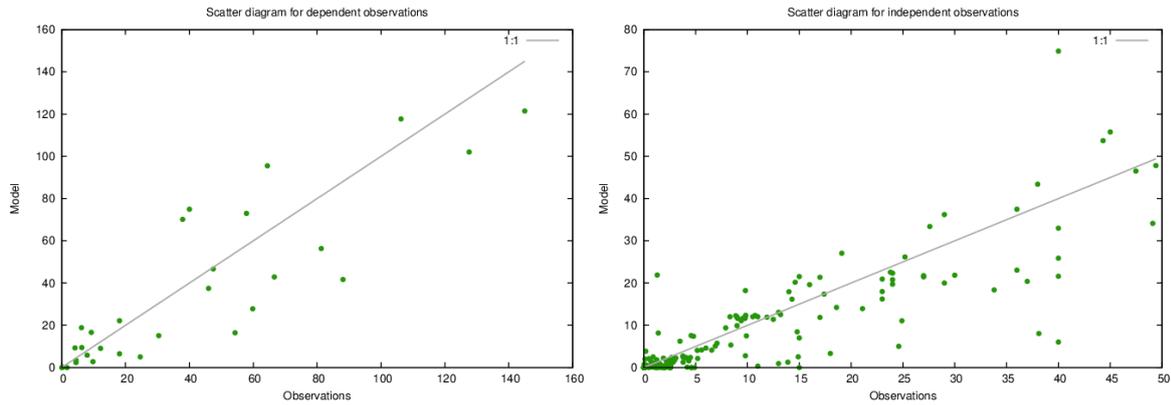


Figure 30 Scatter plots for observations vs model with dependent measurements (left) and independent measurements (right).

Results of SILAM

The ensemble results for Ru-106 emission at Mayak were calculated in a similar way as for the ETEX case. When calculating the gradient of the RMSE with respect to the retrieved emission profile, it was noted that for every ensemble member, the measurements show zero sensitivity to emissions that occur before 2017-09-24 09:00 UTC. Thus the results presented here do not include estimations for emissions occurring before that point in time.

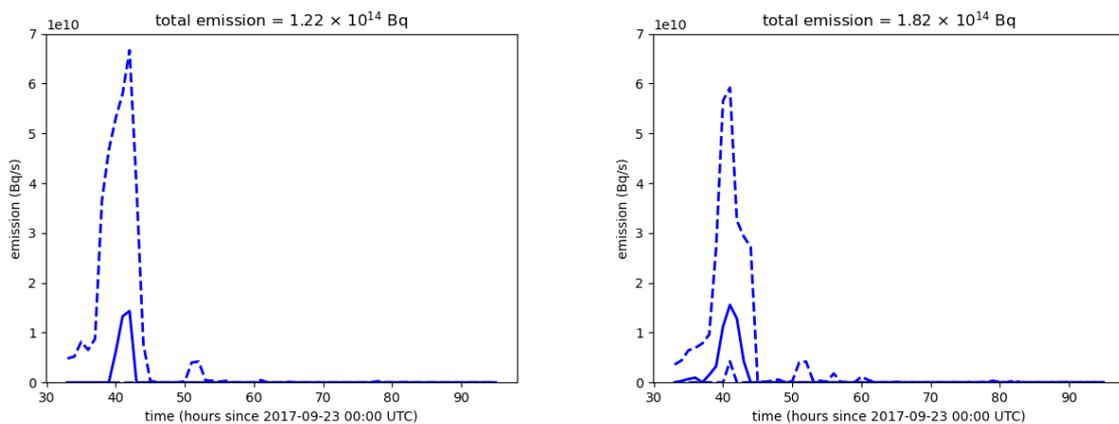


Figure 31 Retrieved temporal emission profile at Mayak from the SILAM model for the full ensemble (left) and for the ten best ensemble members in terms of RMSE (right). The solid lines indicate the median emission and the dashed lines the 10th and 90th percentiles. The indicated total emissions are based on the median emission curves.

Figure 31 shows the retrieved 10th, 50th and 90th emission percentiles based on the full ensemble and the ten best performing ensemble members in terms of the RMSE. Based on the temporal median curve, the estimated total emission is 122 TBq for the full ensemble and 182 TBq for the ten best performing members. However, computed over the totals for all individual ensemble members, the 10th, 50th and 90th percentiles for the total emission are 250 TBq, 380 TBq and 800 TBq, respectively, and 280 TBq, 330 TBq and 520 TBq for the total emission of the ten best performing members.

Results of SNAP

To investigate the source characteristics following the Ruthenium episode, the two likely release locations of Mayak and NIIAR were chosen. For both release scenarios the expected time of release was set between 2017-09-25 T12:00 and 2017-09-27 T00:00 as determined by the NKS-SLIM project.

The inversion-method used with the SNAP model has input files for observations requiring constant measurement times and same end-times of measurements. This is the case for the ETEX case and for the Loviisa case, but is not true for the highly irregular measurements of ^{106}Ru . Since it was considered complicated and error-prone to adapt a new input-file format to the inversion-method, we rather put the measurements in 12 h bins as a best effort approach.

As a priori, we again choose 3 different orders of magnitude of emissions ($1\text{e}9\text{ Bq/s}$, $1\text{e}10\text{ Bq/s}$, $1\text{e}11\text{ Bq/s}$) over the 36 h period mentioned above. The other setup of the inversion is identical to the ETEX-1 case.

Mayak

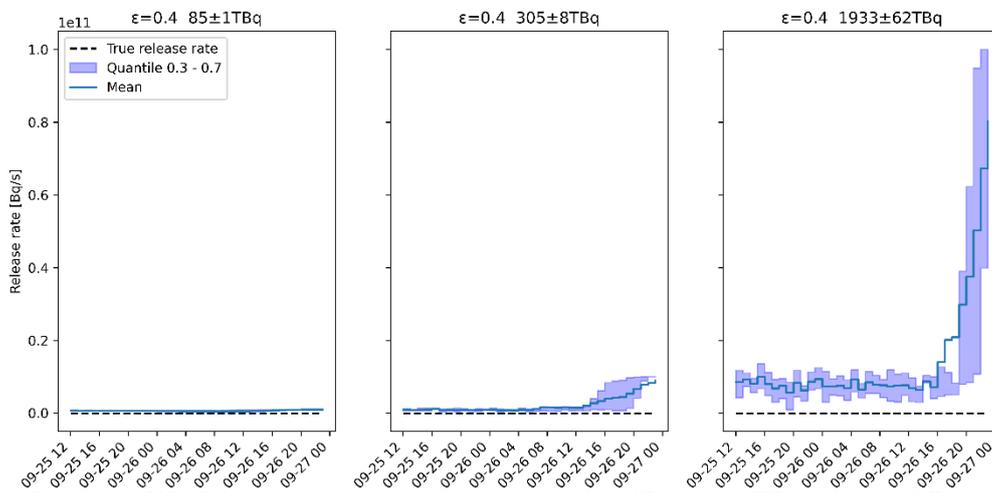


Figure 32 Results of SNAP-Mayak inversion with a a priori of $1\text{e}9\text{Bq/s}$ (130 TB) (left), $1\text{e}10\text{ Bq/s}$ (1300 TB) (middle) and $1\text{e}11\text{ Bq/s}$ (13000 TB) (right)

The results of the Mayak ensemble inversion is included in Figure 32 for three a priori with total releases of 130 TB, 1300 TB and 13000 TB resulting in total a posteriori emissions of 85 TB, 305 TB and 1933 TB. Since the inversion-method rather removes than increases emissions, the first a priori may be considered unrealistically low. Since the last a priori has been heavily reduced, while the second a priori is only slightly reduced (per time step), the results of the middle scenario were considered the most reasonable. The release has been moved rather to the end of the pre-selected time, i.e. between 2017-09-26 16:00 and 2017-09-17 00:00, which is later than expected by other participants of this work, but the total release of 305 TBq is comparable with the DERMA and literature results.

It should be noted that not all member of the ensemble gave reasonable results. In two cases, member 7 and 13, the algorithm could not provide a solution with non-negative terms within 50 iterations and those members have not been used in the results.

Niiar

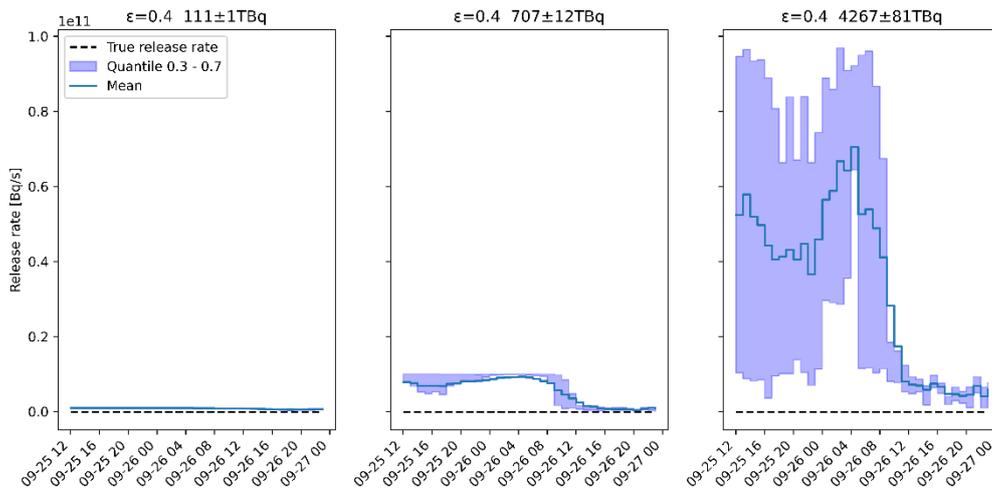


Figure 33 Results of SNAP-Niiar inversion with a priori of $1e9$ Bq/s (130 TB) (left), $1e10$ Bq/s (1300 TB) (middle) and $1e11$ Bq/s (13000 TB) (right)

The results for the Niiar inversion can be seen in Figure 33 with total releases of 130 TB, 1300 TB and 13000 TB resulting in total a posteriori emissions of 111 TB, 707 TB and 4267 TB. With the same argumentation as for the ETEX-1 case and the Mayak case, the first a priori needed a emission increase, which does not work well with this inversion method and the last a priori needed too strong reductions, leaving the middle results, i.e. 707 TBq released mainly between 2017-09-25 16:00 and 2017-09-16 08:00 as most likely release.

The inversion-algorithm failed also in the Niiar case to calculate reasonable emissions terms for some members. Members 0, 1, 7, 12, 13 and 19 had to be removed, leaving the ensemble with only 14 members for quantile, mean and standard-deviation calculations.

Artificial Case of a Release from the Loviisa NPP

A crude method was applied to select the most important nuclides to capture the pre-set of a large scale event by means of gamma dose rates at nearby gamma monitoring stations and national air filter stations.

The source term from which the selection of nuclides is derived is based on the outcome of calculated release sequence using the MELCOR computer code with an input deck for unit 4 at Ringhals NPP. The selected source term is a representative source term for a core melt event without functioning mitigation systems. The initial event is a total loss of all power systems without battery back-up, in addition to all steam-driven systems. Source terms from this type of event are connected with large uncertainties. Therefore, the selected source term is used to represent a large release for all Swedish nuclear power units in operation (see Johansson *et al.*, 2017). The filtered containment venting system is assumed disconnected and instead comprise an exhaust pathway from the reactor containment. It was postulated that the exhaust pathway was open at the time of melt-through of the reactor vessel.

The MELCOR outcome (proportions of mass of the core inventory) has been translated into the accumulated released activity per nuclide including all 285 nuclides in the declared core inventory for unit 4 at Ringhals NPP. Almost 200 nuclides were selected and the released activity per unit of time was corrected for decay and ingrowth for the time period between the scram and the time of the release.

The time evolution is given in 1 hour time steps starting at the onset of the accident (time of the scram) and the following 12 hours, intended to represent the first part of release to undergo subsequent detection by the gamma monitoring stations and, if possible, capture by the air filter stations (see Figure 34 below).

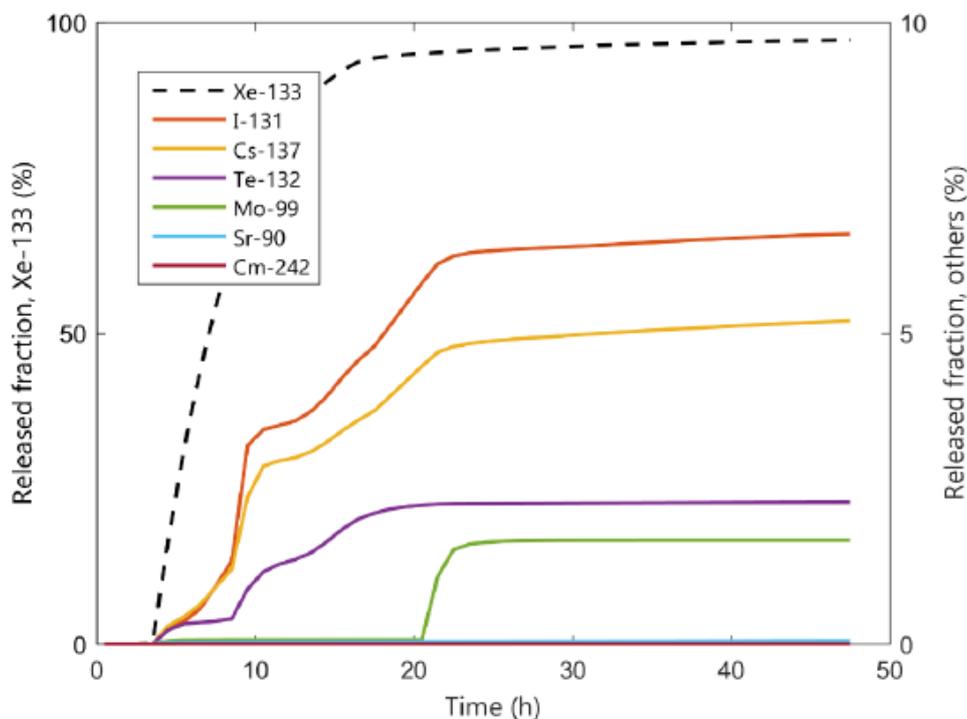


Figure 34 Released fraction of activity for selected nuclides in the case of a postulated event without functioning mitigation systems. Figure taken from Appendix 3 of Johansson *et al.* (2017).

The total dose rate at the gamma monitoring stations would reflect both the contribution from cloud and ground since the stations are not shielded from activity deposited on the ground. Over time the contamination of the station itself would also contribute to the measurement. The accumulated activity per nuclide for the 12 hour time period was multiplied with the nuclide specific cloud dose gamma factor, values taken from Eckerman and Leggett (2013). Here, no considerations were made regarding the differences among the amount of released nuclides subject to dispersion and plume depletion during the transport from release to the station. Furthermore, no corrections were made for decay and ingrowth during the plume transport as implied by the crude methodology employed. The accumulated released activity was multiplied with the nuclide specific ground deposition gamma factor (Eckerman and Leggett, 2013) for all nuclides except noble gases as a rough approximation for the ground contribution); the ground deposition gamma factors selected are those for 1 cm soil entrainment to simulate surface roughness.

The accumulated released activity for the 12 hour time period, excluding the noble gases, was combined with the information on the energy and yield (i.e. photon energy per nuclide transformation for each nuclide (Eckerman and Ryman, 1993) as a first approximation) to mimic the air filter station signals. No account was taken for decay and ingrowth for the period of time between the capture of the release in the filter and the subsequent gamma detector measurement.

Based on the investigation, a set of 11 nuclides was chosen to represent important nuclides: Kr-88*, Xe-133*, Xe-135*, Xe-135m*, Cs-134#, Cs-137, I-131#, I-132*#, I-133#, I-135# and Te-132. The list consist of top 5 for the gamma monitoring station (denoted with *), top 5 for the air filter stations (denoted with #) representing more than 90% of the dose rate contribution in the first 12 hours of the postulated event with the crude methodology employed. Moreover, two nuclides from the top 10 list, Cs-137 and Te-132, were included since they represent key nuclides as seen from historical releases.

The artificial scenario consisting of simulated filter station and gamma station measurements was derived by predicting the atmospheric dispersion of radionuclides from a 9-hour release at the Loviisa nuclear power plant starting at 8 UTC on 2021-09-22. The FMI atmospheric dispersion model SILAM was applied to the release scenario described above using Harmonie NWP model forecast data of 5 km horizontal resolution and hourly time resolution thereby providing 48 hours of hourly average concentration values at filter stations and gamma doses at gamma stations by using the gamma dose model described above in section Semi-infinite Gamma Dose Model. Note that in a real case, the filter stations are likely to measure average concentrations at longer time periods, e.g. 12 or 24 hours, or even a week.

By using the artificially generated filter station and gamma station data, the task for the atmospheric dispersion models now is to estimate the time profiles of the release of the various radionuclides detected by the filter stations involved.

Results of DERMA

For this case, we used the instantaneous concentration values at the locations of the five available filter stations to compute averages over 24 hours, from 8 UTC every day to 8 UTC the next day. If the average concentration is below the threshold value 0.1 mBq/m^3 , the measurement is interpreted as a non-detection. The reasoning for this is to obtain a realistic dataset. However, the result is that each of the five measurement stations provide only two

measurements in the period. Thus, there are only ten 24-hour measurements of each particle, and most of these are non-detections. Thus, only very limited filter measurement data are available for this case. The selected start time, t_0 , is September 22, 2021, 05:00 UTC, i.e. the time of SCRAM, while end time, t_N , is September 24, 2021, 00:00 UTC. Further, we use time bins of 6-hour duration, i.e. $\Delta t_{\text{bin}} = 21600$ s.

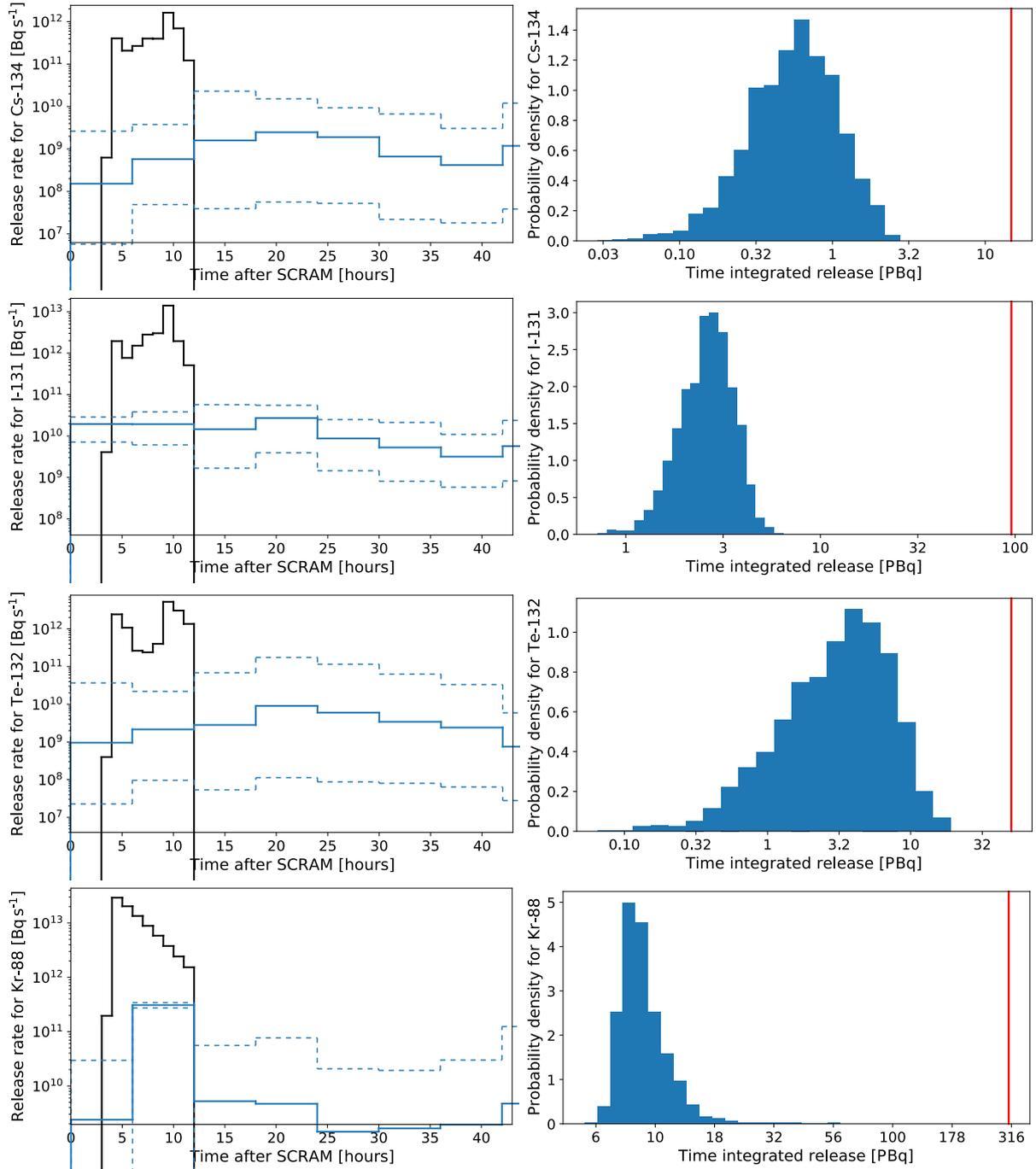


Figure 35 Left: Estimated release rates. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability densities for time integrated release. The red vertical lines indicate the true release.

First, we show the results based on the deterministic meteorological forecast obtained by including both filter and gamma measurements. We include 10 of the 11 described

radionuclides, excluding Xe-135m due to its short half-life. As examples, we show the result for the radionuclide Cs-134, I-131, Te-132 and Kr-88.

From Figure 35, we see that the estimated release rates do not really resemble the true release rates. Only the release rate of Kr-88 seems to reasonably capture the timing of the release. Further, the magnitudes of all releases are underestimated. When examining this, we found that the gamma dose rates predicted by DERMA+ARGOS are larger than the simulated observations by about 2 orders of magnitude on average. We have included a discussion on the matter in the discussion section below.

Next, Figure 36 shows the modelled air concentrations as function of observations. Here, we see that the modelled air concentration measurements are underestimated (for high values), while the magnitude of modelled gamma dose rates reasonably correspond to observations. Again, we discuss this matter in the section below.

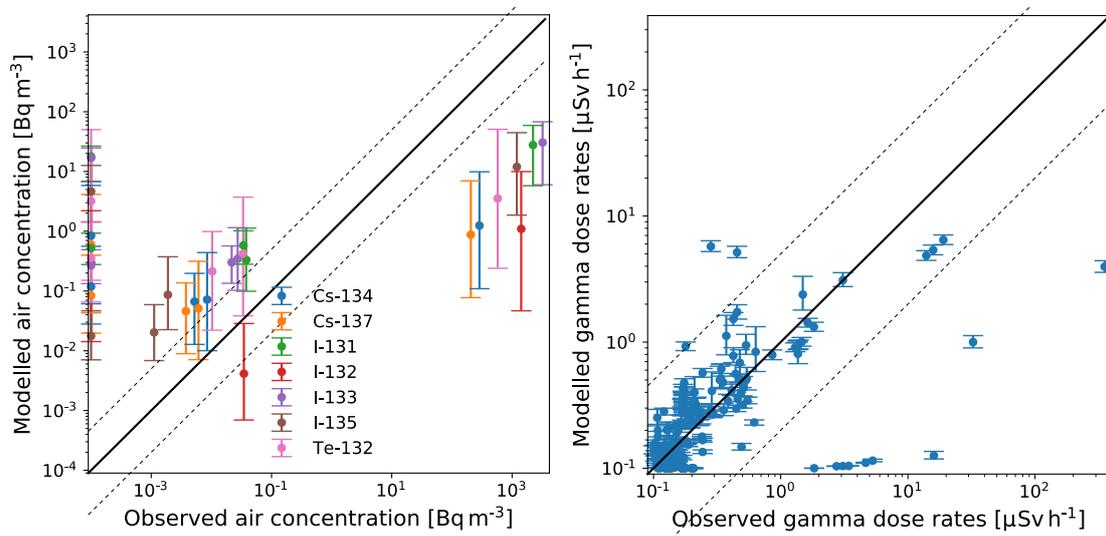


Figure 36 Left: Modelled concentrations as function of observations. Right: Modelled gamma dose rates as function of observations. The solid black line shows perfect correlation, while the dashed black lines indicate a factor of 5 between model and observation.

Based on the attempt above, we have decided to exclude the gamma dose rates. Further, we found that when only limited data are available, a Gaussian likelihood is better suited, since it gives high weight to the few measurements of high values. Thus, we use the likelihood:

$$P(\widetilde{\mathbf{C}}^o | \boldsymbol{\lambda}, I) = \prod_{ik} \text{Gaussian}(\widetilde{C}_{ik}^m, \sigma_f).$$

In Figure 37, we show the results obtained by this approach. As examples, we show the result for the radionuclide Cs-134 and I-131 and Te-132. The release profiles for the noble gasses, on the other hand, cannot be determined when excluding the gamma measurements.

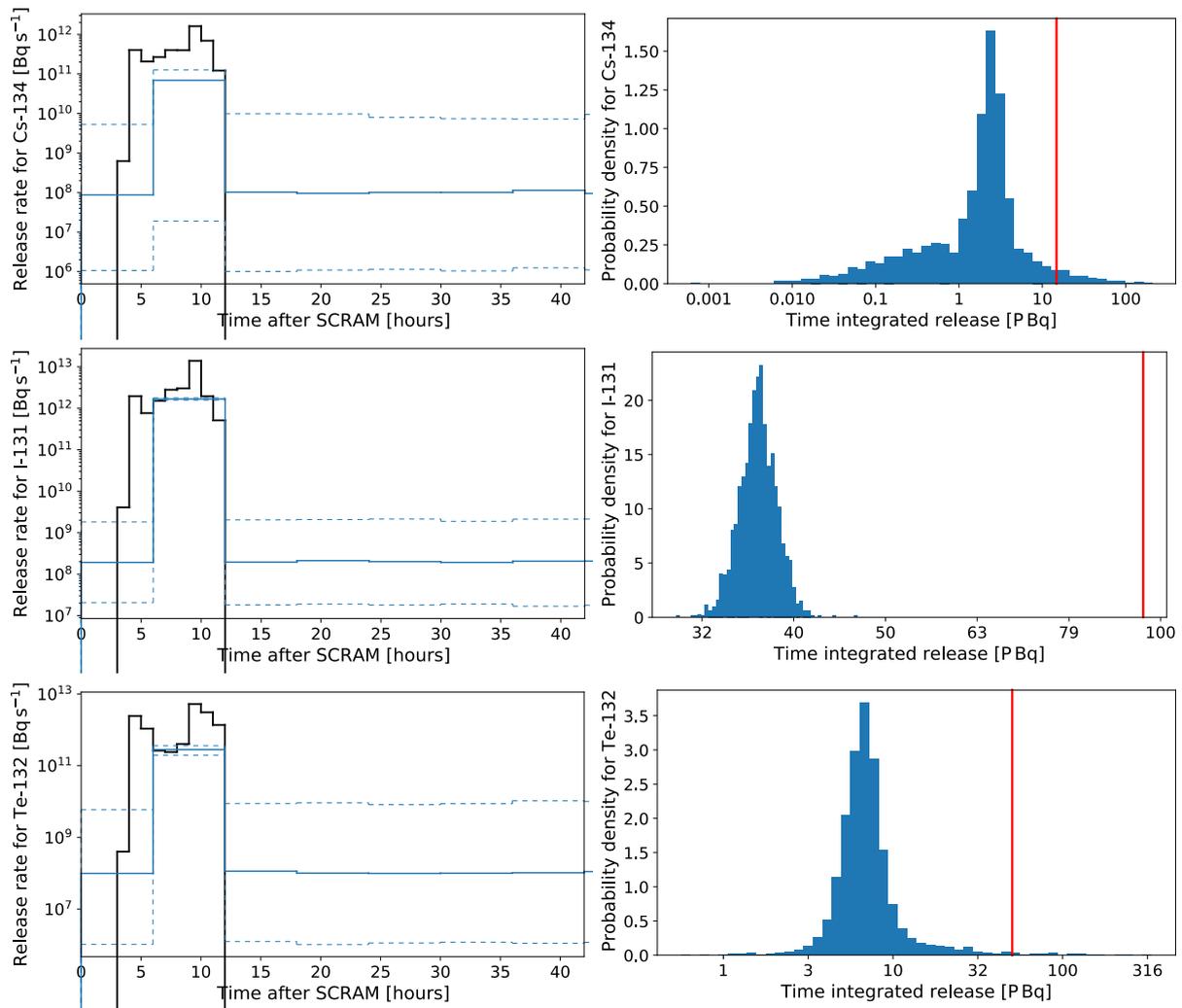


Figure 37 Left: Estimated release rates. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability densities for time integrated release. The red vertical lines indicate the true release.

For all three isotopes, the timing of the release is reasonably captured. However, the magnitude of the release is systematically underestimated.

Next, Figure 38 shows the modelled air concentrations as function of observations.

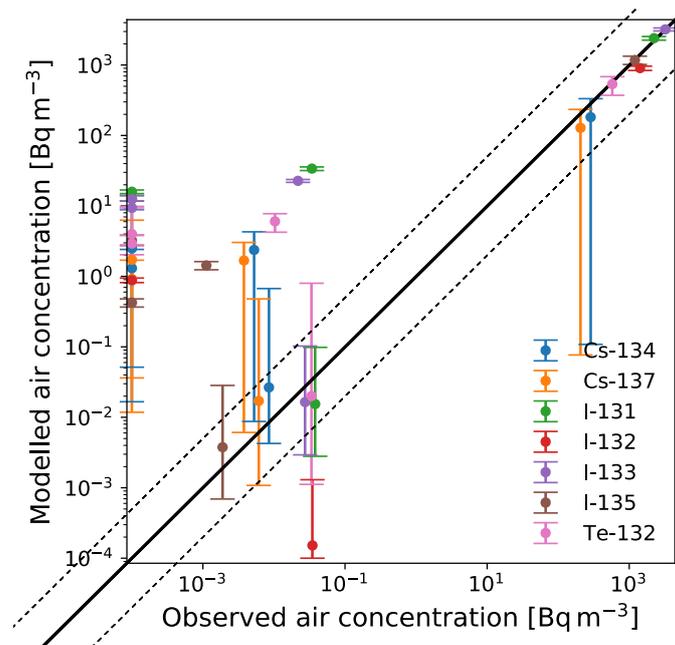


Figure 38 Modelled concentrations as function of observations. The solid black line shows perfect correlation, while the dashed black lines indicate a factor of 5 between model and observation.

We see that the high measurements are all very well estimated. This makes sense, since the Gaussian likelihood gives high weight to these, and because there is one measurement per particle, which is much higher than the remaining; this single measurement completely determines the release rates.

Finally, we show the results based on the meteorological ensemble forecast (again only including filter measurements). Further, some of the estimated source terms essentially just look like the prior distribution and these have been excluded:

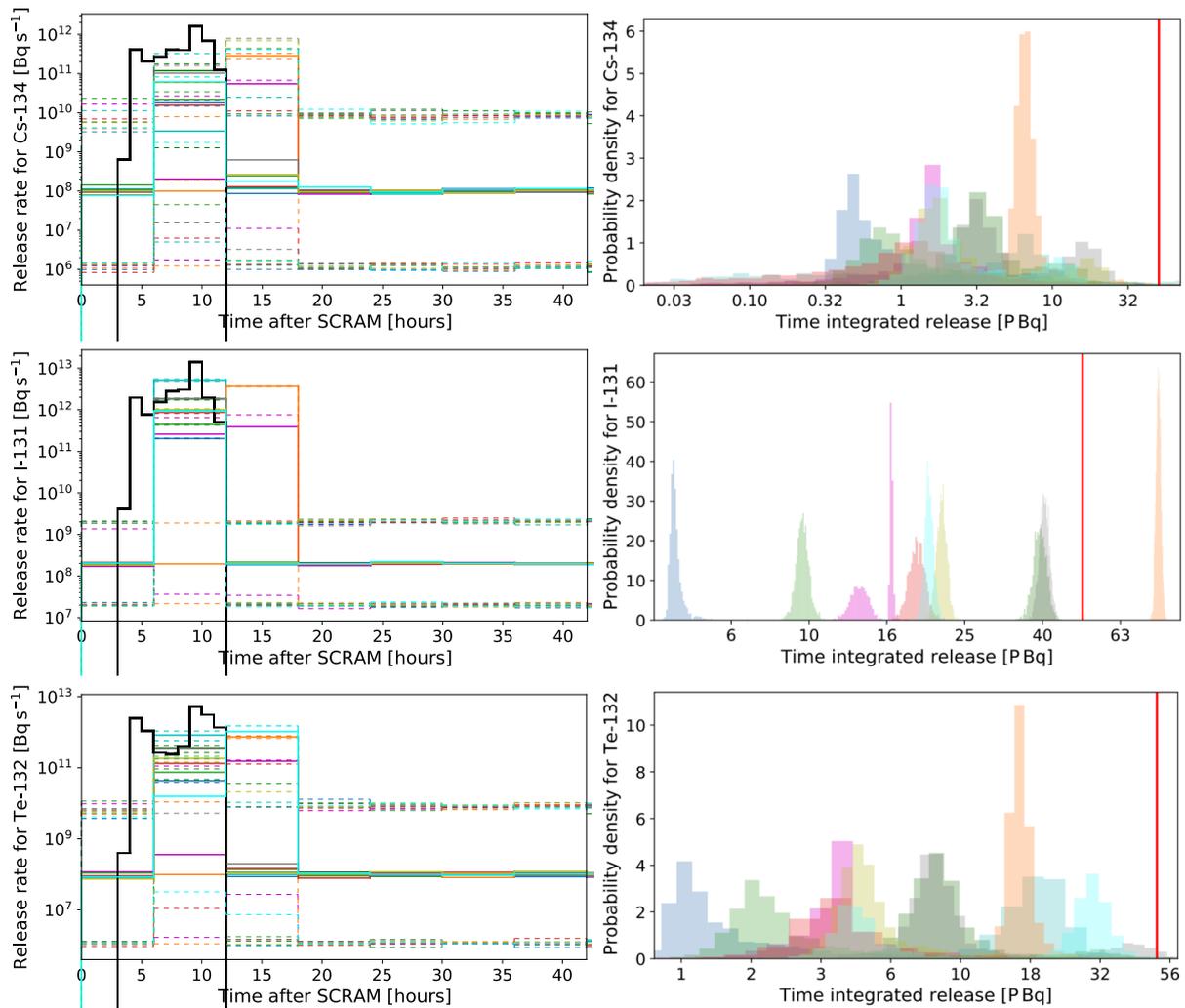


Figure 39 Left: Estimated release rates. The solid colored lines show the medians for the different ensemble members, and the dashed colored lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability density for time integrated releases for all ensemble members. The colors correspond to the colors of the release rates. The red vertical line indicates the true release.

From Figure 39, we see that the estimated released amounts vary significantly. The timing of the release is determined to either 6-12hours or 12-18hours for all members, which is reasonable. However, measurements from nearby gamma dose rate stations should give a better estimate of the release period in practice (even if they are not very useful for source term estimation).

Finally, Figure 40 shows the results based on using all the meteorological ensemble members in the same source term inversion. From this figure, we see that the resulting source term when using all ensemble members is actually worse than the results from individual ensemble members. Thus, this indicates that when the source terms from the different ensemble members are too different, this is probably not the best way to use the ensemble predictions; or it is an indication that the meteorological uncertainty plays a significant role in this case.

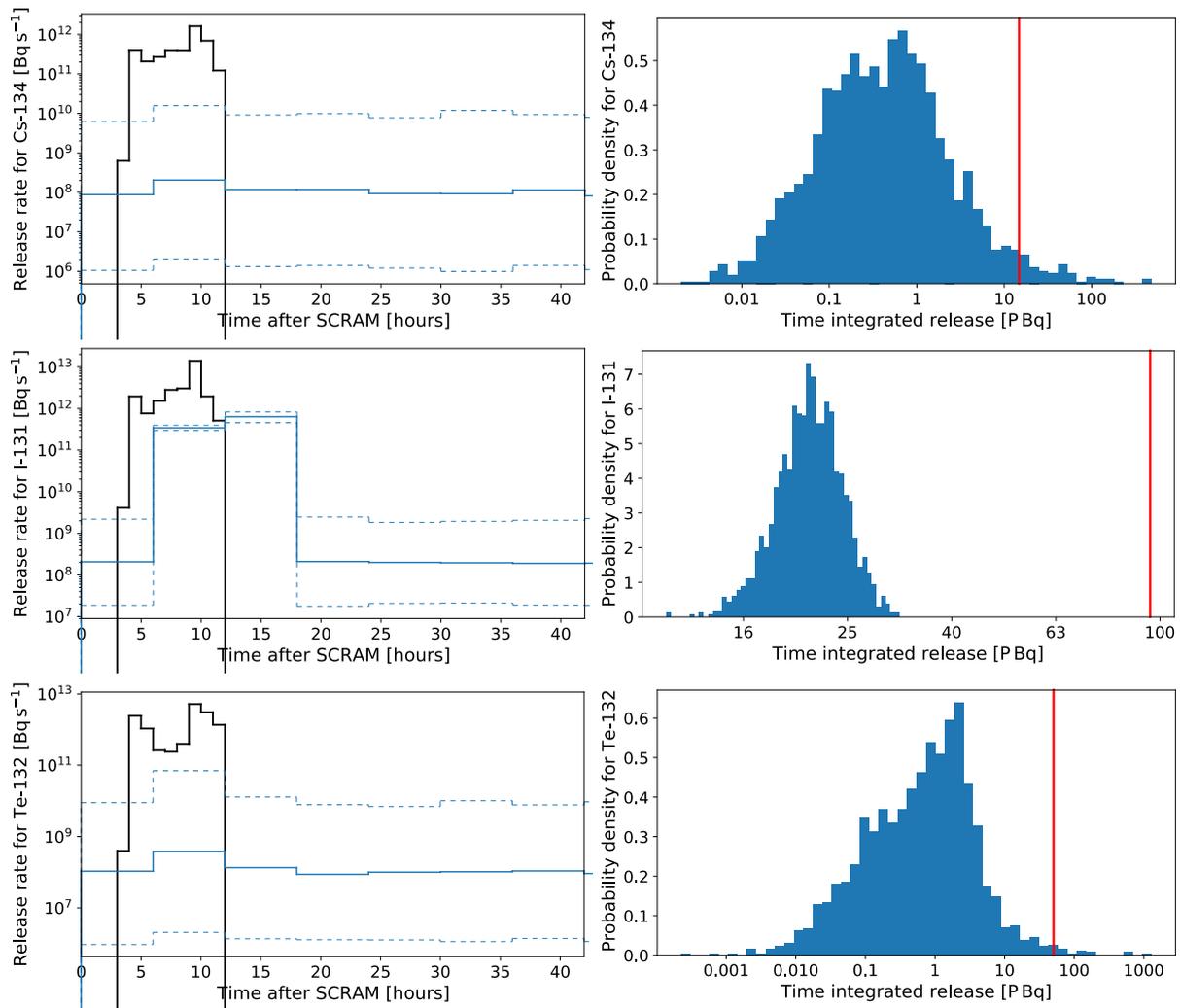


Figure 40 Left: Estimated release rate. The solid blue line shows the median of the posterior distribution, and the dashed blue lines show the 10th and 90th percentiles. The black line shows the true release rate. Right: Probability density for time integrated release. The red vertical line indicates the true release.

Discussion of use of gamma dose rates

In a recent study, we examined the use of gamma dose rates based on an idealized case, where the same meteorological prediction, dispersion model and gamma dose rate model were used for both creating the data set and for the source term estimation (Tølløse and Sørensen, 2022). In this study, we showed that the gamma dose rates are useful for source term estimation in the case, where model errors are negligible.

However, in the study presented in this report the “measured” gamma dose rates are produced by the SILAM gamma dose rate model, while for the source term inversion, we use a combination of the Danish version of the Harmonie model, the DERMA model, and the gamma dose rate model implemented in ARGOS. Figure 41 shows both the simulated observations from the SILAM model, and our modelled gamma dose rates when imposing the true source term. We have subtracted the background observations and removed non-detections for this comparison.

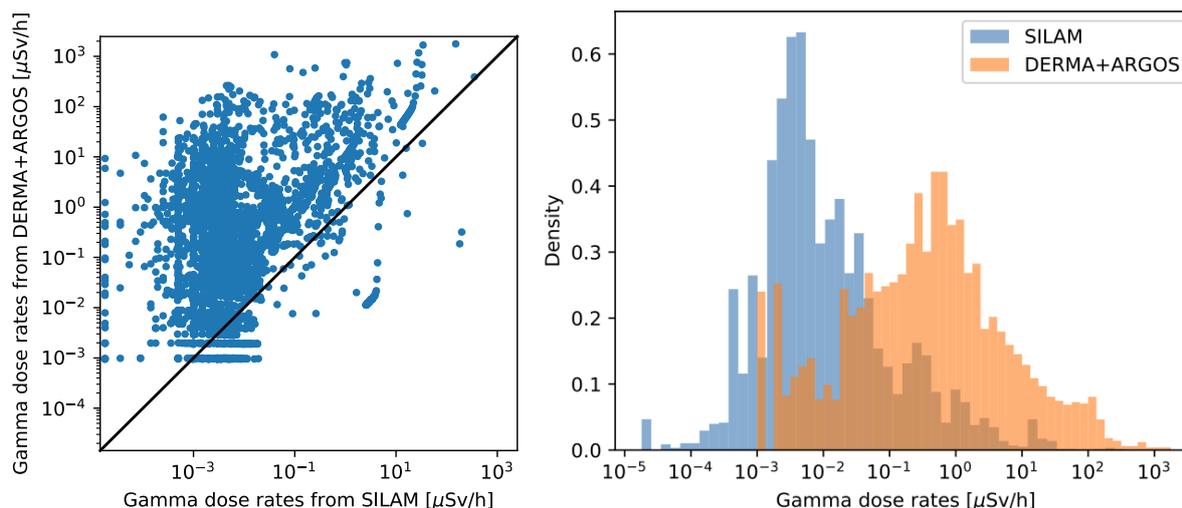


Figure 41 Left: gamma dose rates from DERMA+ARGOS vs gamma dose rates from SILAM. Right: histograms showing gamma dose rates from both models.

From Figure 41, we clearly see why we would not expect the gamma dose rates to be useful for source term estimation for this specific problem. From the histograms in the right plot, we see that the gamma dose rates predicted by the DERMA+ARGOS models are on average roughly two orders of magnitude larger than those from SILAM. However, it does not seem to be systematic, since the left plot indicates that there is no clear correlation between the two sets of observations.

Since the two data sets are produced using different meteorological predictions, dispersion models and gamma dose rate models, it is not possible to say with certainty, why the two data sets differ so much. However, a likely explanation is the meteorological forecasts, since the two data sets are produced using models with different resolution (2.5 km for the Danish and 5 km for the Finnish). The higher resolution could allow local-scale weather phenomena such as land breezes/sea breezes to be resolved, which may significantly impact the concentration pattern (and thus the radiation pattern).

Results of MATCH

The met-data used for this case is operational IFS data on rotated lat-long grid with 0.1 degree resolution.

We retrieved the source profiles from both synthetic filter data and modified gamma doses at EURDEP locations. For the filter station we used all hourly provided data, while we for the gamma doses we first only considered measurements above 5 micro Sv/h and then made an extraction to 50 measurements by weighted bootstrapping. Figure 42 shows the sites used in the both cases. A bit though odd to find gamma dose sites over the Finish bay.

For retrieval of nuclides from gamma doses we assumed three nuclides, I-132, Kr-88 and Xe-135, to be the dose carrier. In the following we present retrievals of Kr-88 as one nuclide appearing in both data sets. A complicating factor is that the modelled gamma doses only consider air cloud contribution while the observed gamma doses may also include ground shine.

We found some issues regarding filter data where the Kotka site dominates the scene and very much defines the solution. Moreover, we had to slightly move the Loviisa power plant somewhat to the east to get hit by the adjoint plume from Kotka. The Figure 43 shows the retrieval of Kr-88 profile using filter data. The retrieved release considering Kr-88 is here in between 12 and 18 UTC 22 September, 2021, with the total activity released for this nuclide of about 1.8×10^{16} Bq.

To illustrate the dominance of Kotka for the filter data, Figure 44 shows the retrieval where Kotka is left out. The signal does then mainly arise from Kiruna and Umeå filter stations. The retrieved emission profile is in time about 12 hours earlier and with significant lower total activity released.

The retrieval of the release profile of Kr-88 derived from gamma dose measurements is shown in Figure 45 where the release is proposed to be lasting from 12 UTC 22 September to 06 UTC 23 September. Here the three nuclides I-132, Kr-88 and Xe-135 are proxy for the nuclides not included. The retrieved total activity of Kr-88 is then amplified in relation to the filter data to 1.4×10^{17} Bq. The retrieved activities released for I-132 and Xe-135 are 2.2×10^{17} and 3.0×10^{16} Bq, respectively, that in total lead to 3.8×10^{17} Bq.

Figure 46 show scatter diagram of gamma doses for observations vs model with dependent and independent observations. The dependent the model does fairly well for lower magnitudes of the data while some high observations are not included in the source inversion. For independent observations the scatter is not all to convincing. The modelled gamma doses only consider air cloud contribution while the observed gamma doses also included ground shine.

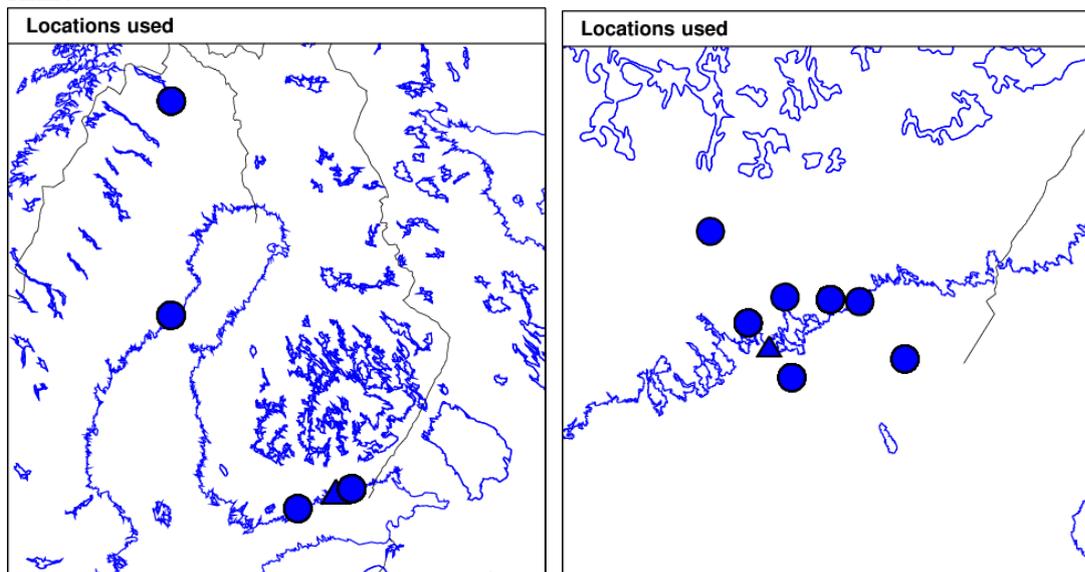


Figure 42 Locations for filter station data (left) and for gamma dose data (right). The Loviisa NPP location is plotted with a filled triangle.

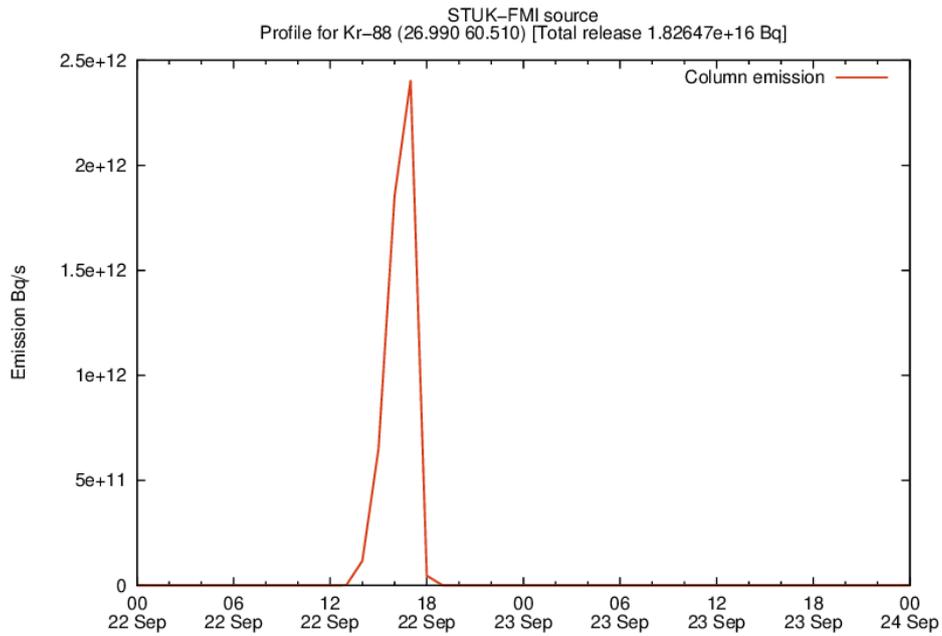


Figure 43 Retrieved source profile for Kr-88 with filter data for Loviisa NPP showing a rather short emission pulse.

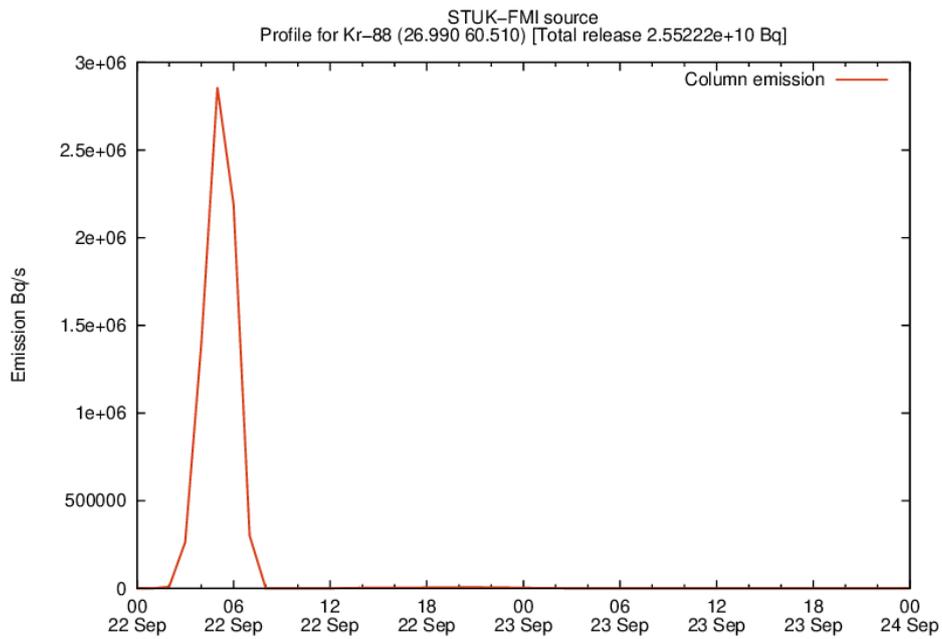


Figure 44 Retrieved source profile with **filter data** for Loviisa where the site Kotka is excluded. The retrieved profile is shifted somewhat earlier than in the retrievals above, and with significant lower total activity.

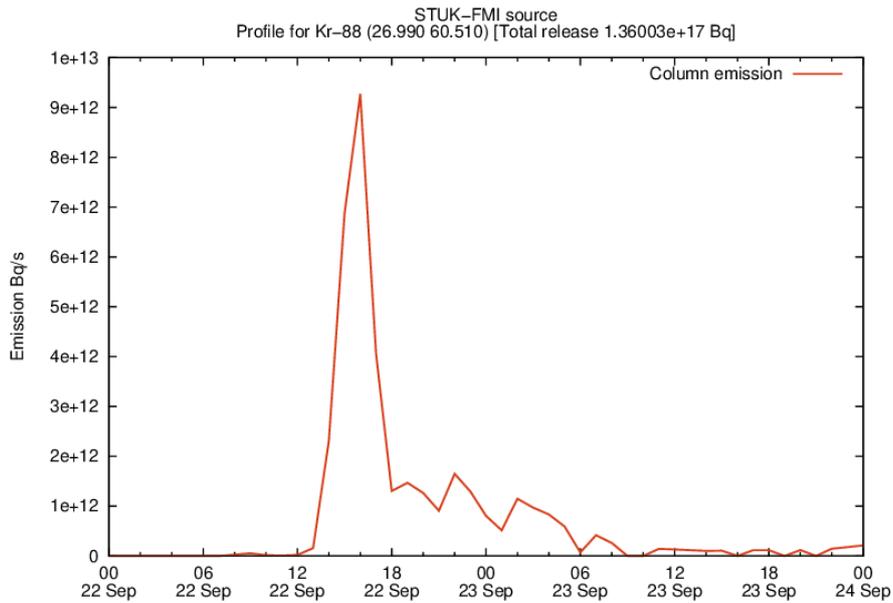


Figure 45 Retrieved source profile for Kr-88 using gamma dose measurement and I-132, Kr-88 and Xe-135 as dose carriers.

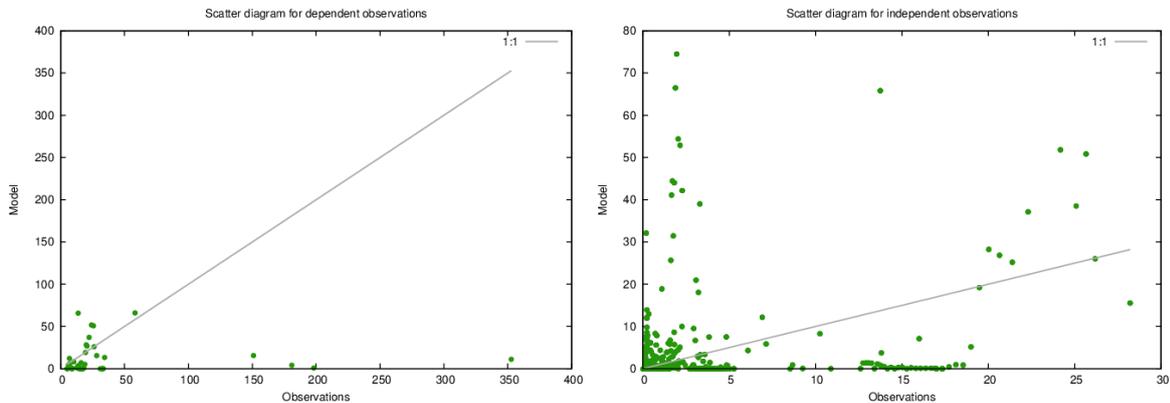


Figure 46 Scatter diagram for gamma dose observations vs model gamma doses for dependent observations (left) and independent observations (right).

Results of SILAM

The source retrieval was based on using the simulated dose rate observations, as due to the nature of the observation method, hourly dose rate observations are normally readily available, unlike hourly in-air concentrations. The emission was set to occur between 25 and 30 m in elevation. A background dose rate of 0.1 $\mu\text{Sv/h}$ is added to the model results. Half of the emitted nuclides of the original simulation of the accident, i.e. I-131, I-132, Kr-88, Cs-137, and Xe-135 were selected as dose rate carriers, emitted evenly in terms of activity. Thus the retrieved results refers to emission of this specific species cocktail, as dose rate measurements do not provide sufficient data for treating the nuclides separately.

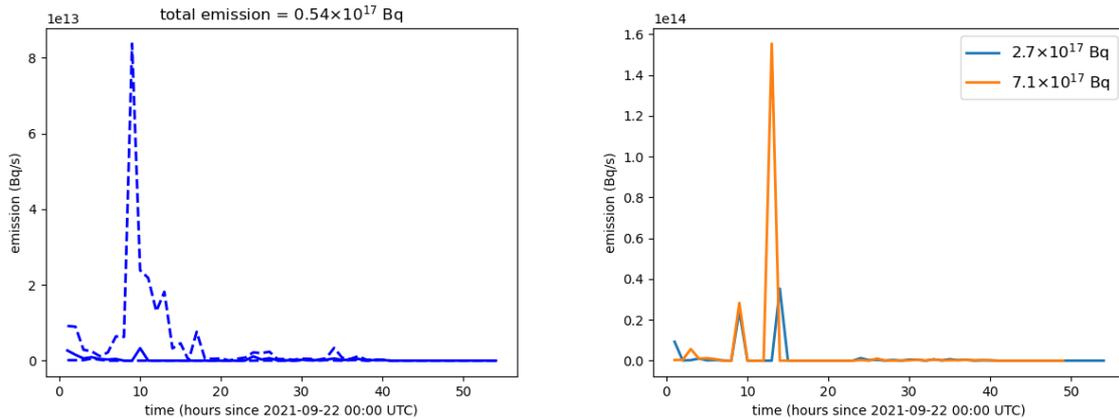


Figure 47 Retrieved profile of the emission at Loviisa NPP from the SILAM model. Left: emission profile for the full ensemble. The solid lines indicate the median emission and the dashed lines the 10th and 90th percentiles. The indicated total emission is based on the median emission curve. Right: emission profiles based on the two best performing members of the ensemble.

In contrary to the ETEX and Ru-106 cases, the performance of the individual ensemble members with respect to the simulated observations is much more uneven for the Loviisa case. While for the other ensemble members the optimized emission yields an average RMSE with respect to the observations of about 4 $\mu\text{Sv/h}$ or more and a correlation below 0.15, similar to the simulation with the ERA5 data, two members correspond to an RMSE of about 2.6 $\mu\text{Sv/h}$ and a correlation of about 0.75. Thus in Figure 47 results for the full ensemble as well as the two best performing members are presented separately. For the full ensemble, the median temporal emission curve yields a total emission of about 54 Pbq, calculated over the optimized emission totals of the ensemble members, the 10th, 50th and 90th percentiles are 160 PBq, 300 PBq and 830 PBq, respectively.

Results of SNAP

For the Loviisa case, SNAP was first run in forwards mode with the real release and the SNAP concentration and deposition fields were adapted to gamma-doserates using the FMI doserate converter. The results can be seen in Figure 48 showing a negative bias for the air-concentrations. In addition, none of the air-concentrations corresponded within a factor 10 between the model and the observations. The observations show a few values larger than 100 Bq/m³ at a station close by the source (Kotka) where the SNAP model fails to resolve any concentration values. In Figure 49 the SNAP concentrations are plotted at the time when SILAM predicts the maximum concentrations (> 5000 Bq/m³) at Kotka. Both stations at Helsinki and Kotka are not hit by the SNAP concentration fields, likely due to parameterization within the dispersion model which provides less diffusion than is necessary. In addition, the chosen meteorological condition at the time of the release where difficult to model with winds changing direction almost 180 degree within the first few hours. With models differing so largely at the few selected stations, an inversion was not feasible and provided results which were not realistic.

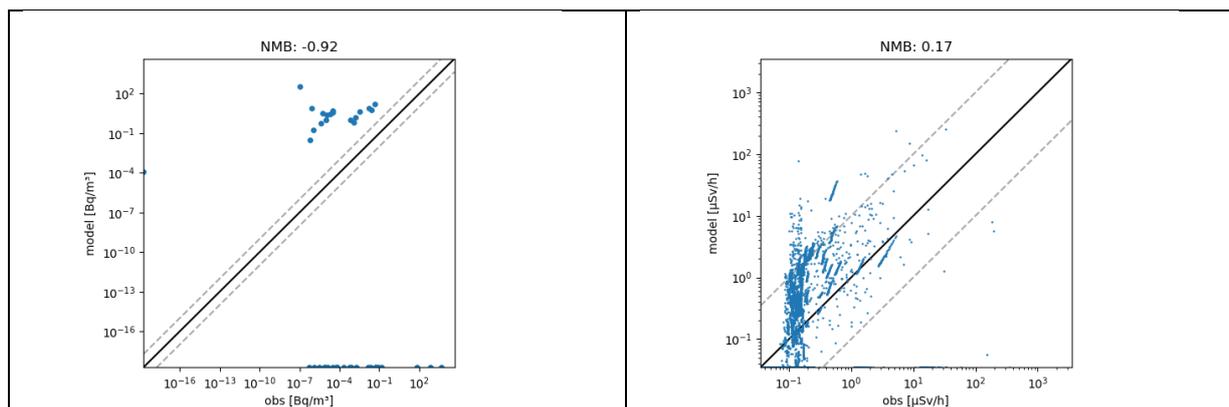


Figure 48 Tempo-spatial scatterplot for the hourly observations in the Loviisa case. Cs-137 concentrations (left), gamma-dose rates (right).

As the observations included many gamma-stations compared to air filter stations, results of dose rates provided an improved bias of 0.17 and a reasonable spread of results mostly within a factor 10 for the gamma-dose rates (see Figure 48 (right)) and inversion using gamma-dose rates were attempted.

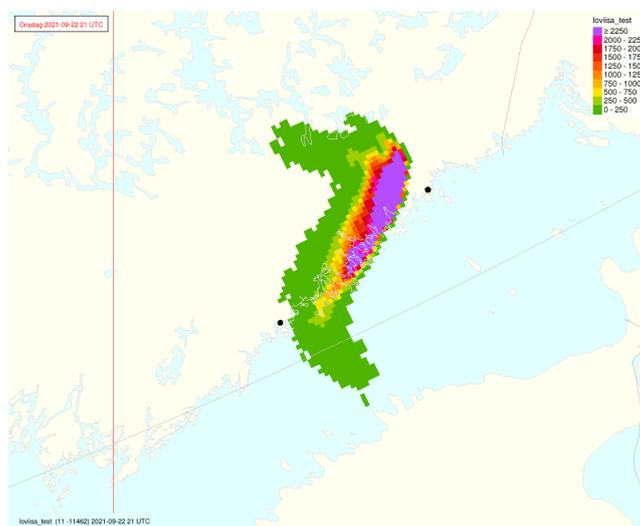


Figure 49 SNAP concentration fields with closest approach to Kotka and Helsinki (black dots).

For the gamma dose rate inversion SNAP was run forward in the same way as for the ETEX and Ruthenium case with unit-emissions in all radionuclides for all time steps between 2021-09-22 00:00 and 2021-09-23 00:00. The inversion method has been tried with the following a priori: true release; all nuclides with $1e+16$, $1e+17$ and $1e+18$ Bq/h. Bias ranges from 1.65 for the true release, 3.59 for $1e+16$, 58.4 for $1e+17$ up to 1875 for the release of $1e+18$. Figure 50 shows the method is not able to decrease the a priori to a sufficient degree. For the Cs-134 isotope we see some reductions based on the meteorological ensemble member, although the a posteriori is not decreased to a sufficient level as to approach the true release rate. For most other components, such as Xe-133 as in Figure 51 we do not see any changes from the a priori. The scatter plot of Figure 52 show some correlation between observations and the model, although the uncertainty is too low.

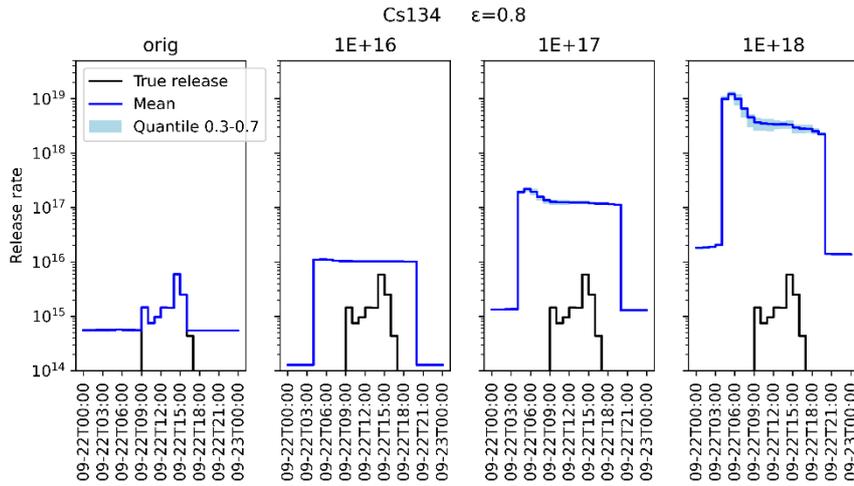


Figure 50 Temporal release characteristics in the Loviisa case for Cs-134 using the SNAP inversion method.

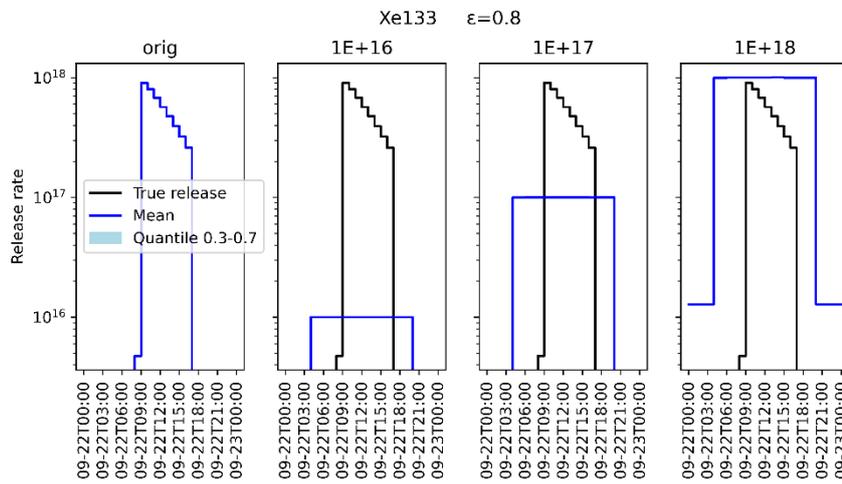


Figure 51 Temporal release characteristics in the Loviisa case for Xe-133 using the SNAP inversion method.

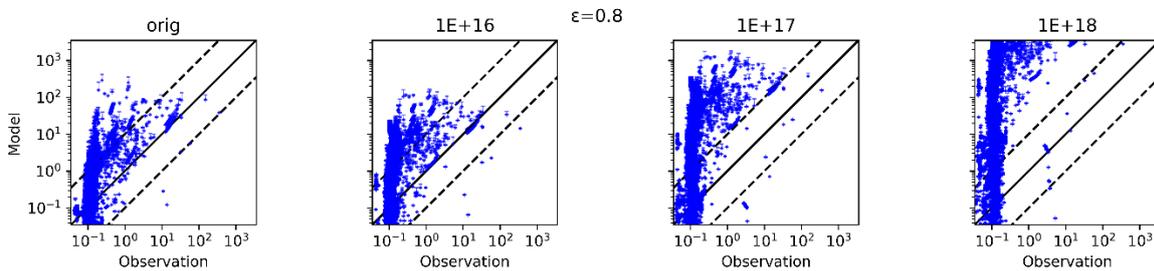


Figure 52 Scatter plot of gamma observations for the SNAP inversion method. Blue dots show the mean of all ensemble members with error bars providing the 0.3-0.7 quantiles. Full black line show perfect correspondence, dashed lines show a factor 10.

Nord Stream leakage of methane in September 2022

In Appendix A, a full description is made of the source localisation and inversion described for the MATCH system at SMHI. This is based on Bayesian inference where the prior distribution is derived from the adjoint footprint of the observations used and a full source profile derived by 4Dvar data assimilation. Here, we present the case of methane (CH_4) release from broken Nord Stream pipe lines.

In late September 2022, the pipe lines Nord Stream I and II were broken close to the Danish island Bornholm. The locations are known and the source rate of methane (CH_4) was estimated to 2.3×10^7 g/h (Reuters, 2022). The network for measuring greenhouse gases (ICOS, 2022) have two Swedish sites recording a pulse from the broken pip-lines, Norunda and Hyltemossa, as shown in Figure 53 and Figure 54, where the locations of the stations are shown in Figure 55.

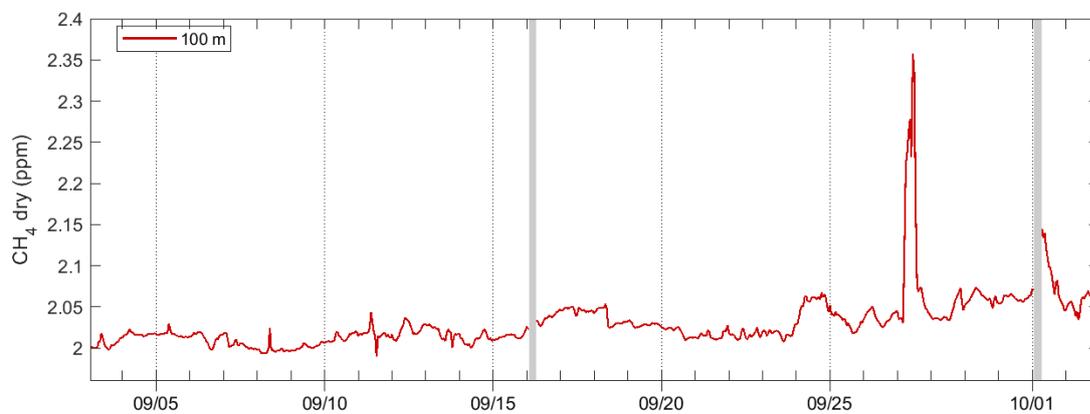


Figure 53 Time-series of measured methane (CH_4) at Norunda station north of Stockholm. The peak at 27-28 September were used for source inversion.

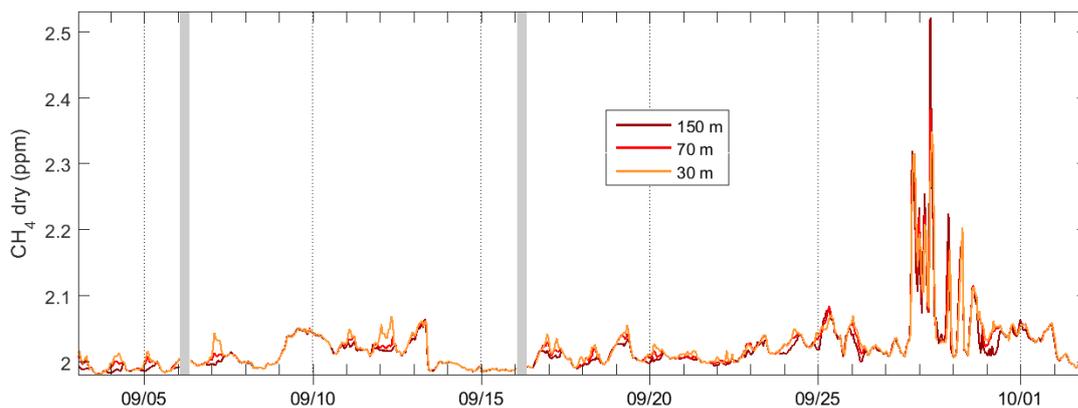


Figure 54 Time-series of measured methane (CH_4) at Hyltemossa station in southern Sweden. The peak around 28-29 September were used for source inversion.

ICOS sites Norunda and Hyltemossa

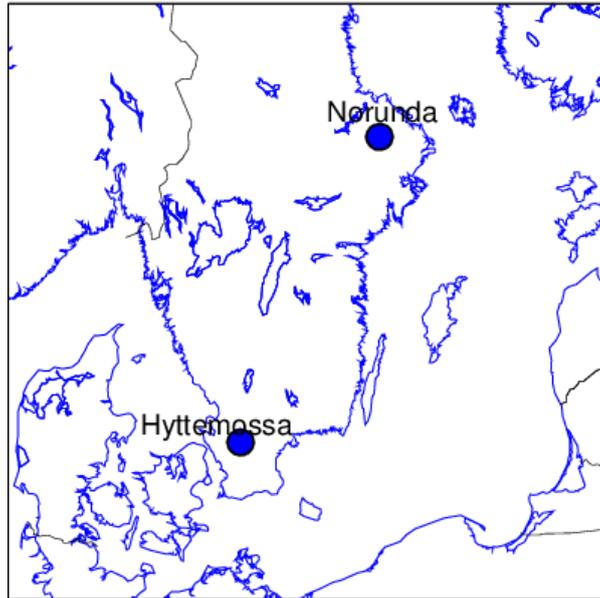


Figure 55 Locations of the two stations used for CH₄ source inversion at Norunda and Hyltemossa.

Figure 56 shows the footprint from these measurements (time-integrated adjoint run) from which uniform and Gaussian priors are derived. We have here restricted the priors to only be valid for sea areas. This is partly related to the lack of observational information, where then more knowledge is put into the prior. Figure 57 shows the sampling of uniform and Gaussian distributions in a Bayesian context, and Figure 58 shows the resulting posteriors for the highest ranked locations. The locations of the sources are in this analysis slightly to the west of Bornholm for the both priors while the known broken pipe lines are slightly to the east of Bornholm.

CH₄ case of Nord Stream broken pipes, September 2022 Footprint and priors for uniform and Gaussian distributions

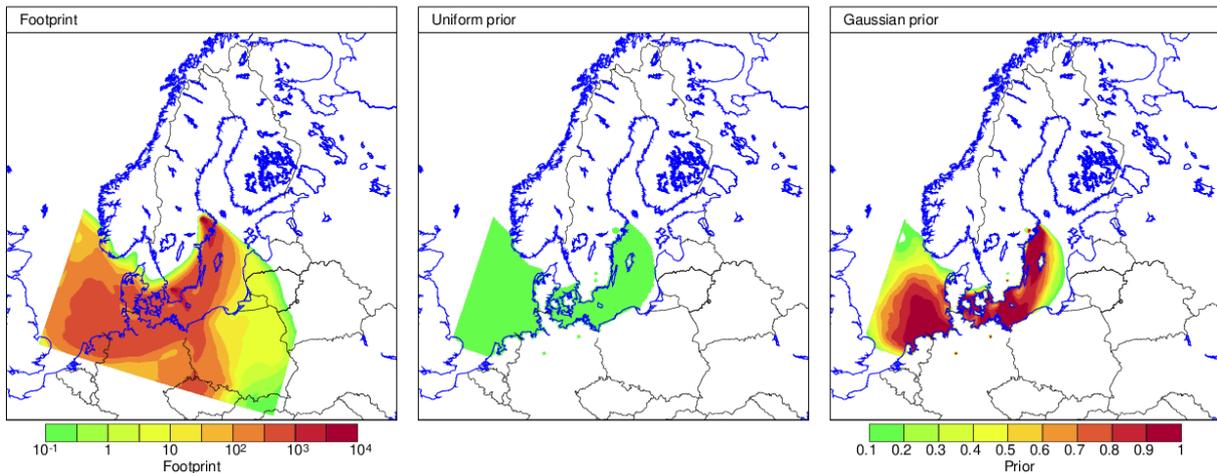


Figure 56 Footprint (left) of the methane (CH₄) peaks in Figure 53 and Figure 54 from which two priors are derived, uniform prior (middle) and Gaussian (right).

CH₄ case of Nord Stream broken pipes, September 2022
Priors for uniform and Gaussian distributions

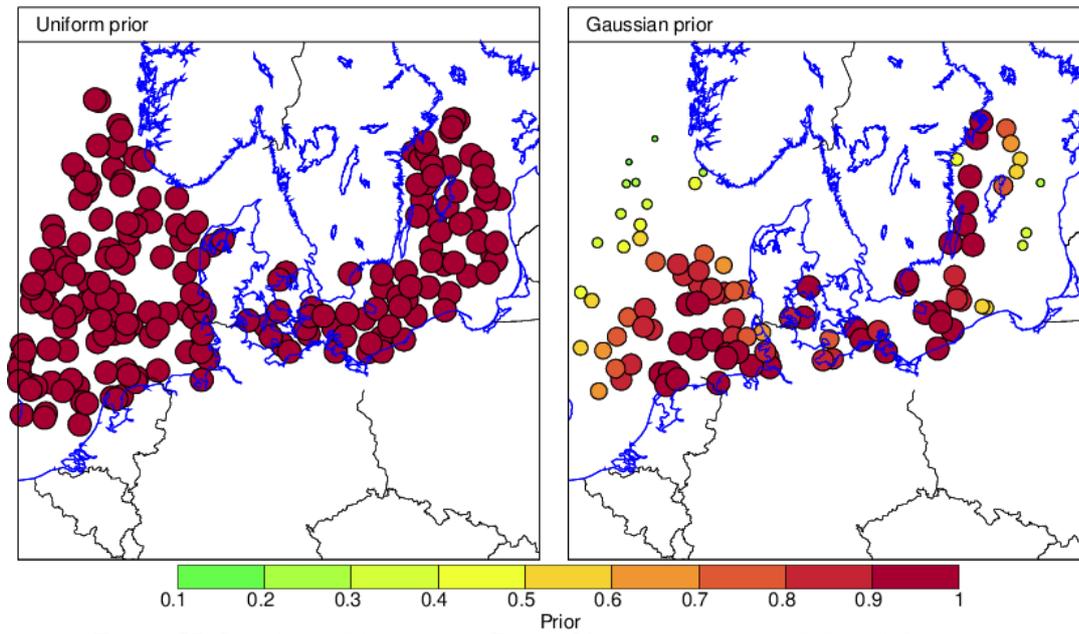


Figure 57 Sampling of the priors in Figure 56 for uniform prior (left) and Gaussian prior (right).

CH₄ case of Nord Stream broken pipes, September 2022
Posteriors for uniform and Gaussian priors

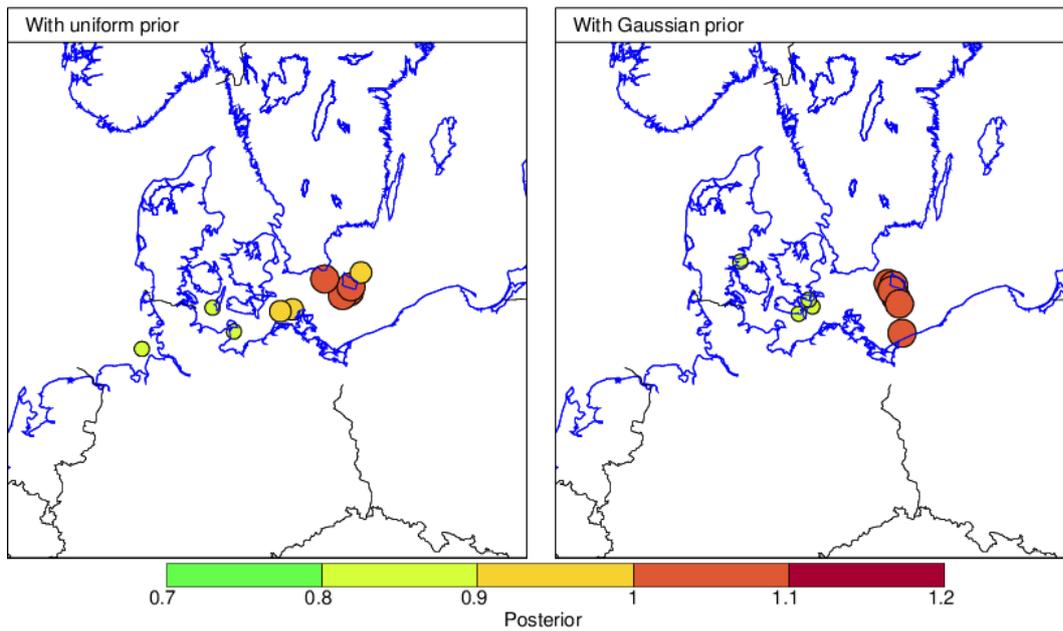


Figure 58 The most promising source locations given by the posterior for uniform prior (left) and Gaussian prior (right).

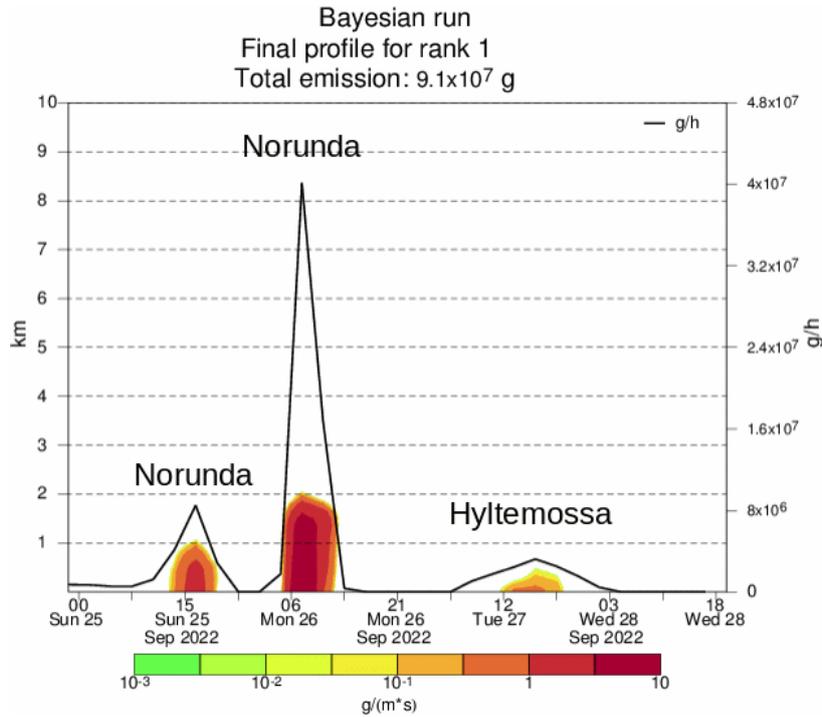


Figure 59 Retrieved source profile at the most likely position in Figure 58. The origin of the pulses are shown. Note that one the part of the release seen by Norunda and Hyltemossa are visible in the time profile. See also Figure 60.

The source profile that arise by the used measurements, for the most likely location, is shown in Figure 59 that also arrive at source rates in the same order as the reported ($\approx 2 \times 10^7$ g/h). In this source term, only those parts of the source that could be verified by the measurements are captured. By assuming the background error matrix in 4Dvar to have correlation in time (see Appendix A, Section 6.4) the retrieved source will be continuous as shown in Figure 60. Note that the magnitude of the source rate does not change by adding time correlation.

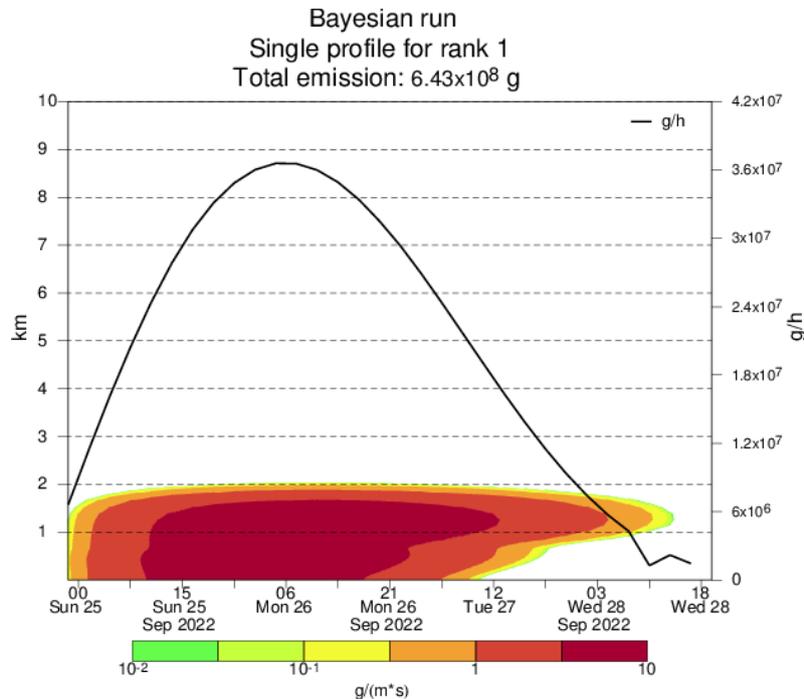


Figure 60 As for Figure 59 by where the background source error is assumed to have temporal correlation. The temporal correlation acts as a smoothing function filling the gap between the pulses shown in for Figure 59.

The Nord Stream case illustrates both the importance of the prior in Bayesian inference and the background error formulation of the source errors in 4Dvar. Assuming potential source locations for both land and sea will not put the source at the vicinity of Bornholm (for these few observations). Then limiting of the prior just to sea areas is meant to illustrate that expert judgement may be needed to confine the solution, especially when only a few observations are available. The observations do by nature only catch a fraction of the source that did hit the measurement sites. For continuous sources the background error in the 4Dvar enables filling the gap between the pulses (seen by the observations) when the background error is assumed correlated in time.

ARGOS and Source Characterization

The Long-Range dispersion model interface in ARGOS has been developed in close cooperation with the different model providers through a number of years. The default interface is capable of handling forward deterministic Atmospheric Dispersion Modelling. As well, interfaces have been developed for specific modelling needs such as handling ensemble calculations (developed in cooperation with DMI) and Adjoint modelling results (developed in cooperation with SMHI and SSM). Likewise, new interfaces will have to be developed in order to handle Source Localization (SL) as well as Source Characterization (SC) based on various measurements of radionuclide air concentration and gamma doses. The implications of such interfaces will be discussed in this section.

Concentration Measurements in ARGOS

The ARGOS-DSS features several different options for visualising different kinds of radiological measurements. An example is shown below where a plot of European Monitoring Stations is presented in ARGOS. Station data are imported using the EURDEP-protocol, see Figure 61.

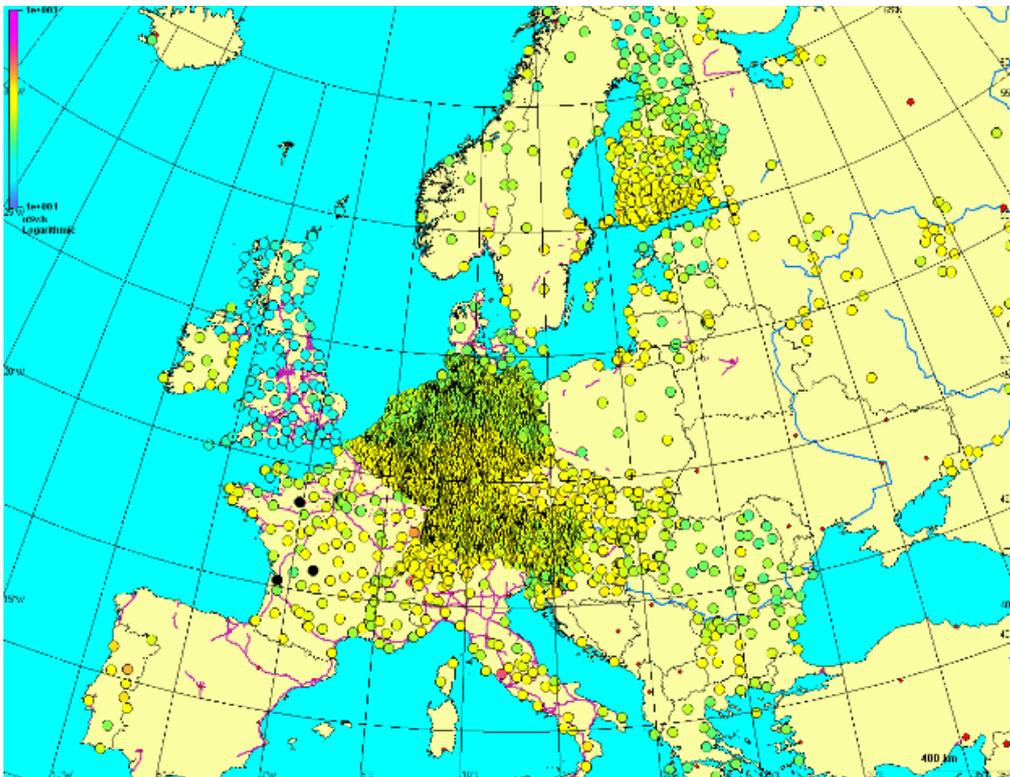


Figure 61 European Monitoring Stations presented in ARGOS.

The typical output from these types of Permanent Monitoring Stations is a dose rate; normal unit $\mu\text{Sv/h}$.

Likewise, ARGOS is capable of importing and presenting data from Air Sample Stations – again importing data using the EURDEP-protocol. Typical output from these Air Sample Stations is an air concentration; normal unit Bq/m^3 (per nuclide).

Request for Source Characterization Calculation from ARGOS

The existing Request dialog in ARGOS today is focused on doing forward Atmospheric Dispersion modelling; giving the user options for selecting a release point (a reactor) and a release description (source term – or ensemble of source terms) and a release time, see Figure 62.

Atmospheric Dispersion: Request Run

Service: MLDP1

Run ID: Test

Reactor Name: RINGHALS-1

Source term type: Model

Model source term: Nordic

Output Timestep [h]: 3

Start Time [UTC]: 17-dec-2019 10:21

Grid Size [km]: Native

Coordinates

Lon: 12*6*30 Lat: 57*15*23

Coordinate System: WGS84

Weather Data

NWP Manual Setup

Mode

Test

Save Cancel Send Request

Figure 62 Forward atmospheric dispersion modelling request dialog in ARGOS.

For Source Localization and Characterization, the needs are quite different from (normal) forward ADM. The user needs to provide a (number of) detection(s) of time-average air concentrations to be part of the request sent to the model, and for source characterization potentially also gamma measurements. As ARGOS already today has a module for presenting different measurements in the system (see section above) it would be natural to base the GUI for requesting SL or SC on the existing GUI for selecting measurement data to be visualized in the system. Due to the number of different types of measurement data, the GUI for selection is rather complicated, see Figure 62.

The needs for SL and SC can be narrowed down to these three parameters:

- The area of interest – the area from which the system should select measurement data
- The time frame of interest – the time period from which the system should select measurement data
- The type of measurements

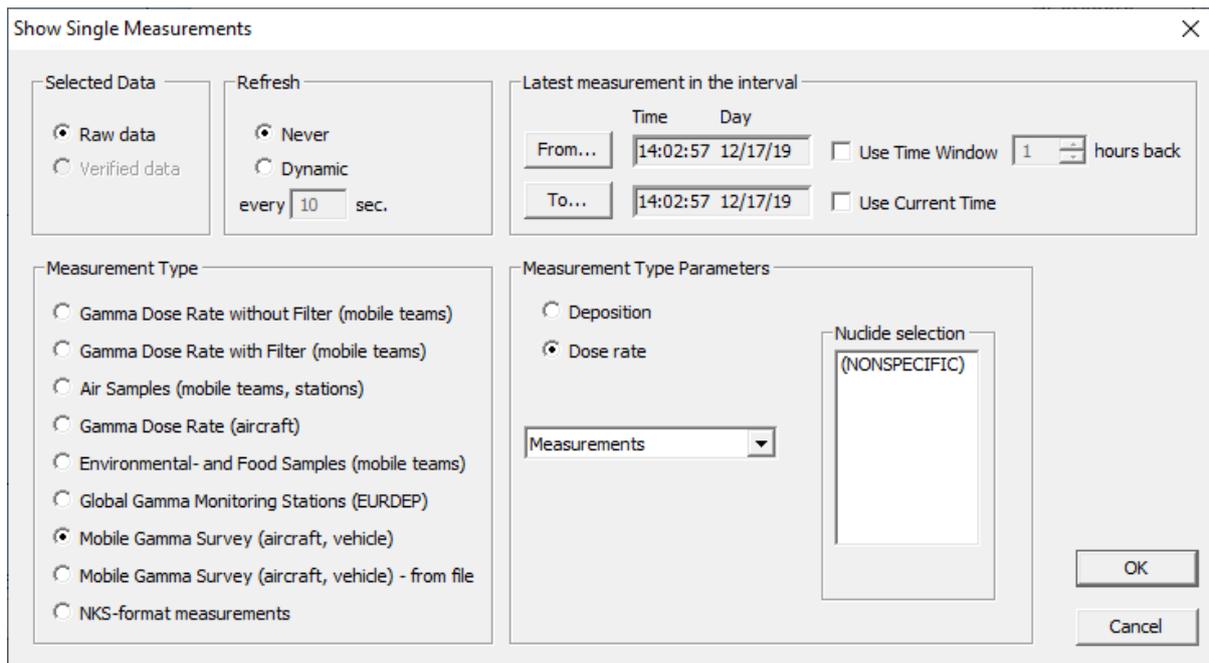


Figure 63 GUI for selecting measurement data to be visualized in the system.

Request interface for SL and SC modelling

The request interface should be able to at least provide information from ARGOS to the SL and SC models about

- General
 - List of nuclides
 - List of measurements
 - Position of source if known
- Per measurement
 - Position of measurement
 - Time (time frame) of measurement
 - Air concentration (Bq/m³) per nuclide
 - Detection limits (per nuclide)
 - Gamma dose rates

In case of too few measurements, ARGOS should be able to receive the reply back from the national meteorological service that the problem is undetermined, and to communicate this to the ARGOS user.

Result interfaces for SL and SC

The result of SL will resemble ordinary dispersion modelling; however, here the results shown on a map will be probability densities for the location of the source. For SC, the model result will to a high degree resemble a source term description, thus it would be quite efficient to reuse the interface for source term description from ARGOS to the ADM models as the interface from SC to ARGOS.

Types of measurements

As can be seen from the GUI for visualising measurements in ARGOS above, ARGOS can handle quite many types of measurements. In order to simplify the GUI, we will restrict the selection for SLC to be limited to:

- Permanent Monitoring Stations – dose rates,
- Air Sampler Stations – air concentrations,

or a combination of the two.

It has to be considered how to distinguish between “lack of data” from a station and “below detection limit measurements”. Of course, in the radiological domain there will always be some background radiation.

Permanent gamma monitors and filter stations

In Europe, many radiological filter station measurements are taken once a week. However, in special cases it is possible to change to daily measurements. From a meteorological perspective, a week can be a long time covering a number of different meteorological phenomena taking place over the station site within the period. Additionally, a week could be longer than the accidental release. Thus, if the intended use of the measurements among other things is to assist in locating the release point or to characterize the source, daily measurements are of much greater value. The filters are changed manually, and in most cases sent by regular mail for analysis. This implies a delay in retrieving the measurement data, and it means that short-lived radionuclides have decayed when the filter is measured.

The detection limit depends first of all on the amount of air drawn through the filter. Thus, the pump efficiency and the measurement period are key parameters for the detection limits, and therefore detection limits vary across Europe. In addition, the presence of many radionuclides on a filter makes it difficult to measure concentrations accurately.

There are currently no international agreements on routine distribution of filter station concentration measurements. For the October 2017 case of Ru-106 measurements in Europe, the IAEA collected the available filter station data (IAEA, 2017a,b).

The European Radiological Data Exchange Platform (EURDEP), see <https://eurdep.jrc.ec.europa.eu/Entry/>, is a network for the exchange of radiological monitoring data between most European countries. Currently, EURDEP is used for the European automatic gamma monitoring network which does not provide activity concentrations, only gamma dose rates. However, EURDEP might be used also for filter station measurements which could be very helpful in future events.

In comparison with filter-stations, the gamma monitoring network in Europe is much denser and reports automatically at high frequency, e.g. hourly, all of which make such data attractive for an operational nuclear DSS. However, in order to use the gamma-monitoring data for inverse modelling it is beneficial that measurements of nuclide-specific average activity concentrations are provided. Further, the measurement sensitivity is several orders of magnitude worse than for filter stations. Thus, gamma monitoring results are most likely only useful for source localization at the early phase of a nuclear accident. They might, however, be of value for the characterization of the release.

The Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) includes a monitoring network of 80 radiological stations measuring radioactive particles, around half of them also noble gasses. Near real time access to these data will be very helpful for locating or characterizing an unknown release of radionuclei.

All in all, there is a need for automated and uniform handling of filter station measurements and for a network or an organization to take care of the collection of data and providing them in an organized way for Emergency Management Organizations and DSS systems.

Summary, Conclusions and Outlook

For a source, with a known geographic location, of radionuclides which have been accidentally released to the atmosphere, it is important to estimate the release rates for the dose-contributing nuclides. In the course of the SOCHAOTIC NKS project, various methods have been developed for estimation of the release employing measurements of activity concentrations at filter stations as well as gamma dose measurements at gamma monitors. The methods have been applied to three cases, viz. the European Tracer Experiment (ETEX), the 2017 case of Ru-106 in Europe, and an artificial case of a release from the Loviisa nuclear power plant in Finland.

For the selected cases, deterministic numerical weather prediction model data have been obtained from the global model of the European Centre for Medium-Range Weather Forecasts (ECMWF) and both deterministic and ensemble-statistical data from the Harmonie high-resolution non-hydrostatic model. By employing the atmospheric dispersion models DERMA, MATCH, SILAM, and SNAP, four different approaches have been applied in order to estimate the release rates for the selected cases. However, qualitatively these methods have provided quite similar results, and it has therefore not been possible to prioritize these methods.

The previous NKS projects MUD (Sørensen *et al.*, 2014), FAUNA (Sørensen *et al.*, 2016) and AVESOME (Sørensen *et al.*, 2019) have demonstrated that inherent meteorological uncertainties play an important role for the atmospheric dispersion model results; see also (Sørensen *et al.*, 2020). In the SOCHAOTIC project, we have examined the uncertainties of the source estimation arising from the inherent meteorological uncertainties. This was carried out by applying the methodologies developed in the first year of SOCHAOTIC to an ensemble of numerical weather prediction model results. Subsequently, ensemble-statistical methods were applied to quantify the uncertainties of the estimated release profiles.

References

- Andersson, C., Bergström, R., Bennet, C., Robertson, L., Thomas, M., Korhonen, H., Lehtinen, K.E.J., and Kokkola, H. 2015. MATCH-SALSA – Multi-scale Atmospheric Transport and CHEMistry model coupled to the SALSA aerosol microphysics model – Part 1: Model description and evaluation. *Geosci. Model Dev.* **8**, 171–189. doi:10.5194/gmd-8-171-2015
- Baklanov, A. and J. H. Sørensen. Parameterisation of radionuclide deposition in atmospheric dispersion models. *Phys. Chem. Earth* **26** (2001) 787–799
- Baklanov, A., A. Mahura and J. H. Sørensen. Methodology for Prediction and Estimation of Consequences of Possible Atmospheric Releases of Hazardous Matter: ‘Kursk’ Submarine Study. *Atmos. Phys. Chem.* Vol. **3** (2003) 747–762
- Bartnicki, J., H. Haakenstad, and Ø. Hov. Operational SNAP model for remote applications from NRPA. *Norwegian Meteorological Institute Report* (2011) 12–2011
- Bartnicki, J., I. Amundsen, J. Brown, A. Hosseini, Ø. Hov, H. Haakenstad, H. Klein. Atmospheric transport of radioactive debris to Norway in case of a hypothetical accident related to the recovery of the Russian submarine K-27. *Journal of environmental radioactivity* **151** (2016) 404–416
- Bengtsson, L., U. Andrae, T. Aspelien, Y. Batrak, J. Calvo, W. de Rooy, E. Gleeson, B. H. Sass, M. Homleid, M. Hortal, K.-I. Ivarsson, G. Lenderink, S. Niemelä, K. P. Nielsen, J. Onville, L. Rontu, P. Samuelsson, D. S. Muñoz, A. Subias, S. Tijn, V. Toll, X. Yang, and M. Ø. Kjøltzow. The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System. *Monthly Weather Review* (2017) **145** No. 5 <https://doi.org/10.1175/MWR-D-16-0417.1>
- Brodtkorb, A., Benedictow, A., Klein, H., Kylling, A., Nyiri, A., Valdebenito, A., Sollum, E., and Kristiansen, N. Estimating volcanic ash emissions using retrieved satellite ash columns and inverse ash transport modelling using VolcanicAshInversion v1.2.1. Submitted to *Geosci. Model Dev.*
- Casella, G., and George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, **46**(3), 167-174
- Eckerman, K.F. and J.C. Ryman. External exposure to radionuclides in air, water, and soil, Federal Guidance Report 12 (1993)
- Eckerman, K.F. and R.W. Leggett. DCFPAK 3.02 (Dose Coefficient Data File Package), Oak Ridge National Laboratory (2013)

Eckhardt, S., Prata, A. J., Seibert, P., Stebel, K., and Stohl, A.: Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modelling. *Atmos. Chem. Phys.*, 8, 3881–3897 (2008) <https://doi.org/10.5194/acp-8-3881-2008>

ETEX web-site (2019), <https://rem.jrc.ec.europa.eu/etex/>

Gloster, J., A. Jones, A. Redington, L. Burgin, J. H. Sørensen, R. Turner. International approach to atmospheric disease dispersion modelling. *Veterinary Record* 03 (2010a) **166** (12):369. DOI:10.1136/vr.166.12.369a

Gloster, J., A. Jones, A. Redington, L. Burgin, J. H. Sørensen, R. Turner, P. Hullinger, M. Dillon, P. Astrup, G. Garner, R. D'Amours, R. Sellers and D. Paton. Airborne spread of foot-and-mouth disease – model intercomparison. *Veterinary Journal* **183** (2010b) 278–286

Graziani, G., Klug, W., and Nodop, K. Real-time Long-range Dispersion Model Evaluation of the ETEX First Release. **EUR 17754 EN** (1998) 92-828-3657-6 European Commission

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57 (1), 97–109

Heinonen, S. (2017) CALCULATION OF GAMMA-RAY DOSE RATE FROM AIRBORNE AND DEPOSITED ACTIVITY. X-ray and gamma-ray decay data standards for detector calibration and other applications. https://www-nds.iaea.org/xgamma_standards/

Hoe, S., J. H. Sørensen and S. Thykier-Nielsen. The Nuclear Decision Support System ARGOS NT and Early Warning Systems in Some Countries around the Baltic Sea. In: Proceedings of the 7th Topical Meeting on Emergency Preparedness and Response, September 14–17, 1999, Santa Fe, New Mexico, USA

Hoe, S., H. Müller, F. Gering, S. Thykier-Nielsen and J. H. Sørensen. ARGOS 2001 a Decision Support System for Nuclear Emergencies. In: Proceedings of the Radiation Protection and Shielding Division Topical Meeting, April 14–17, 2002, Santa Fe, New Mexico, USA

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1), 1593-1623

IAEA. Convention on Early Notification of a Nuclear Accident. (1986) <https://www.iaea.org/publications/documents/treaties/convention-early-notification-nuclear-accident>

IAEA. Updated Technical Attachment Status of Measurements of Ru-106 in Europe, 2017-10-20, 12:00 UTC, prepared by the Incident and Emergency Centre of the IAEA. Available through national radiation protection authorities. (2017a)

IAEA, Status of Measurements of Ru-106 and Ru-103 in Europe at 2017-11-24, 15:00 UTC, prepared by the Incident and Emergency Centre of the IAEA. Available through national radiation protection authorities. (2017b)

ICOS. Icos Sweden is the national contribution a joint European effort for European coverage aiming to provide accessible, high-quality data to improve our understanding of greenhouse gas sinks and emissions. 2022. <https://www.icos-sweden.se>

Johansson, Jan, Peder Kock, Jonas Boson, Simon Karlsson, Patrick Isaksson, Jonas Lindgren, Elisabeth Tengborn, Anna Maria Blixt Buhr, Ulf Bäverstam. Review of Swedish emergency planning zones and distances. SSM report (2017)
<https://www.stralsakerhetsmyndigheten.se/en/publications/reports/radiation-protection/2017/201727e/>

Kahnert, M., Variational data analysis of aerosol species in a regional CTM: Background error covariance constraint and aerosol optical observations operators, *Tellus*, 60B, 753-770, (2008)

Kahnert, M. (2018). Information constraints in variational data assimilation. *Q. J. R. Meteorol. Soc.* 144, 2230-2244

Klein, H., and Bartnicki, J. (2018). The Severe Nuclear Accident Programme, Concert Air2 Bulletin, Issue 24, 2018, page 4

Lauritzen, B., A. Baklanov, A. Mahura, T. Mikkelsen and J. H. Sørensen. K-model description of probabilistic long-range atmospheric transport in the Northern Hemisphere. *Atmos. Environ.* **40** (2006) 4352–4369

Lauritzen, B., A. Baklanov, A. Mahura, T. Mikkelsen and J. H. Sørensen. Probabilistic risk assessment for long-range atmospheric transport of radionuclides. *J. Envir. Radioactivity* **96** (2007) 110–115

Le Brazidec, J. D., Bocquet, M., Saunier, O., & Roustan, Y. (2020). MCMC methods applied to the reconstruction of the autumn 2017 Ruthenium-106 atmospheric contamination source. *Atmospheric Environment: X*, 6, 100071

Liu, Y., Haussaire, J. M., Bocquet, M., Roustan, Y., Saunier, O., & Mathieu, A. (2017). Uncertainty quantification of pollutant source retrieval: comparison of Bayesian methods with application to the Chernobyl and Fukushima Daiichi accidental releases of radionuclides. *Quarterly Journal of the Royal Meteorological Society*, 143(708), 2886-2901

Mahura, A., A. Baklanov and J. H. Sørensen. Methodology for evaluation of possible consequences of accidental atmospheric releases of hazardous matter. *Radiat. Prot. Dos.* **103** (2003) 131–139

- Mahura, A. G., A. A. Baklanov, J. H. Sørensen, F. L. Parker, V. Novikov, K. Brown, K. L. Compton 2004: Assessment of Atmospheric Transport and Deposition Patterns Related to Russian Pacific Fleet Operations. *Environmental Monitoring and Assessment* **101** (2005) 261–287
- Masson, O., Steinhauser, G., Zok, D., Saunier, O., Angelov, H., Babić, D., & Zorko, B. (2019). Airborne concentrations and chemical considerations of radioactive ruthenium from an undeclared major nuclear release in 2017. *Proceedings of the National Academy of Sciences*, *116*(34), 16750-16759
- Mikkelsen, T., S. Alexandersen, H. Champion, P. Astrup, A. I. Donaldson, F. N. Dunkerley, J. Gloster, J. H. Sørensen and S. Thykier-Nielsen. Investigation of Airborne Foot-and-Mouth Disease Virus Transmission during Low-Wind Conditions in the Early Phase of the UK 2001 Epidemic. *Atmos. Chem. Phys. Disc.* **3** (2003) 677–703
- PDC-ARGOS. <http://www.pdc-argos.com>
- Robertson, L. and Langner, J. 1998. Source function estimate by means of variational data assimilation applied to the ETEX-1 tracer experiment. *Atmos. Environ.* **32**, 4219-4225
- Robertson, L., Langner, J. and Engardt, M., An Eulerian limited-area atmospheric transport model, *J. Appl. Met.* **38**, 190-210, 1999
- Saltbones, J., A. Foss, and J. Bartnicki. Norwegian Meteorological Institute's real-time dispersion model SNAP (Severe Nuclear Accident Program): Runs for ETEX and ATMES II experiments with different meteorological input. *Atmospheric Environment* **32** no. 24 (1998) 4277–4283
- Salvatier J., Wiecki T.V., Fonnesbeck C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2**:e55 DOI: 10.7717/peerj-cs.55
- Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., and Genikhovich, E. Construction of the SILAM Eulerian atmospheric dispersion model based on the advection algorithm of Michael Galperin. *Geosci. Model Devel.* (2015) **8** 3497–3522
- Sørensen, J.H., Bartnicki, J., Blixt Buhr, A.M., Feddersen, H., Hoe, S.C., Israelson, C., Klein, H., Lauritzen, B., Lindgren, J., Schönfeldt, F., Sigg, R. Uncertainties in atmospheric dispersion modelling during nuclear accidents. *J. Environ. Radioact.* **222** (2020) 1-10
<https://doi.org/10.1016/j.jenvrad.2020.106356>
- Sørensen, J. H. Sensitivity of the DERMA Long-Range Dispersion Model to Meteorological Input and Diffusion Parameters. *Atmos. Environ.* **32** (1998) 4195–4206
- Sørensen, J. H., D. K. J. Mackay, C. Ø. Jensen and A. I. Donaldson. An integrated model to predict the atmospheric spread of foot-and-mouth disease virus. *Epidemiol. Infect.* (2000) **124**, 577–590

Sørensen, J. H., C. Ø. Jensen, T. Mikkelsen, D. Mackay and A. I. Donaldson. Modelling the atmospheric spread of foot-and-mouth disease virus for emergency preparedness. *Phys. Chem. Earth* **26** (2001) 93–97

Sørensen, J. H., A. Baklanov and S. Hoe. The Danish Emergency Response Model of the Atmosphere. *J. Envir. Radioactivity* **96** (2007) 122–129

Sørensen, J. H., B. Amstrup, H. Feddersen, U. S. Korsholm, J. Bartnicki, H. Klein, P. Wind, B. Lauritzen, S. C. Hoe, C. Israelson, and J. Lindgren. Meteorological Uncertainty of atmospheric Dispersion model results (MUD). **NKS-307** ISBN 978-87-7893-385-0, http://www.nks.org/en/nks_reports/view_document.htm?id=111010212220490 (2014)

Sørensen, J.H., B. Amstrup, H. Feddersen, J. Bartnicki, H. Klein, M. Simonsen, B. Lauritzen, S. C. Hoe, C. Israelson and J. Lindgren. Fukushima Accident: UNcertainty of Atmospheric dispersion modelling (FAUNA). **NKS-360** ISBN 978-87-7893-444-4 (2016), <http://www.nks.org/scripts/getdocument.php?file=111010213440189>

Sørensen, J. H., Schönfeldt, F., Sigg, R., Pehrsson, J., Lauritzen, B., Bartnicki, J., Klein, H., Hoe, S. C., and Lindgren, J. Added Value of uncertainty Estimates of SOURCE term and Meteorology (AVESOME) – final report. NKS-420, ISBN 978-87-7893-509-0 (2019) http://www.nks.org/en/nks_reports/view_document.htm?id=111010214696230

Sørensen, J.H., H. Feddersen, K.S. Tølløse, R. Kouznetsov, M. Sofiev, H. Klein, M. Ulimoen, L. Robertson, J. Pehrsson, B. Lauritzen, D. Bohr, A. Hac-Heimburg, C. Israelson, A.M. Blixt Buhr, J. Lindgren, T. Karhunen, and T. Peltonen, Source Localization by Inverse Methods (SLIM). NKS-456, ISBN 978-87-7893-549-6 (2022) https://www.nks.org/en/nks_reports/view_document.htm?id=111010214697785

Tichý, O., Ulrych, L., Smidl, V., Evangelidou, N. and Stohl, A. On the tuning of atmospheric inverse methods: comparisons with the European Tracer Experiment (ETEX) and Chernobyl datasets using the atmospheric transport model FLEXPART. *Geosci. Model Dev.*, **13**, 5917–5934 (2020) <https://gmd.copernicus.org/articles/13/5917/2020/>

Tølløse, K. S., Kaas, E., and Sørensen, J. H. (2021). Probabilistic Inverse Method for Source Localization Applied to ETEx and the 2017 Case of Ru-106 including Analyses of Sensitivity to Measurement Data. *Atmosphere*, **12**(12), 1567

Tølløse, K. S., and Sørensen, J. H. (2022). Bayesian Inverse Modelling for Probabilistic Multi-Nuclide Source Term Estimation Using Observations of Air Concentration and Gamma Dose Rate. *Atmosphere*, **13**(11), 1877.

Appendix A Description of the source inversion approaches for the MATCH modelling system

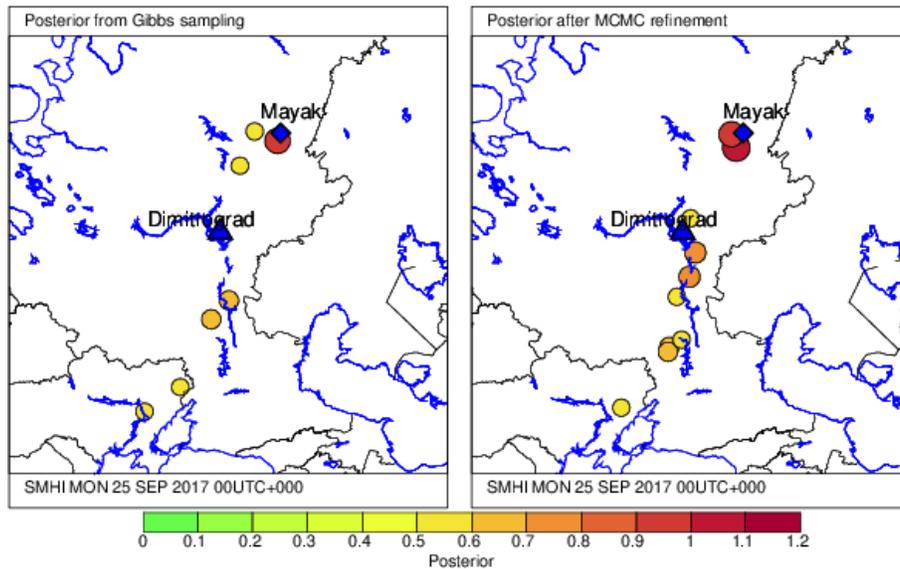
Here follows a thorough description of the source inversion approaches for the MATCH modelling system.

SMHI contribution to SOCHAOTIC NKS contract AFT/B(19)6

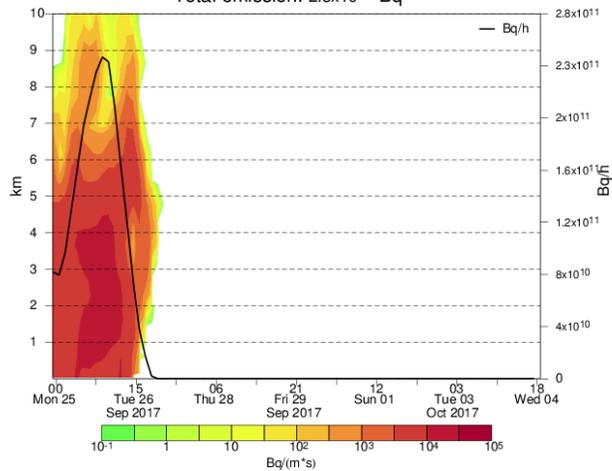
Lennart Robertson, SMHI

January 16, 2023

Posterior before and after MCMC refinement



Ru-106 case 2017
Ensemble profile for rank 1
Total emission: 2.3×10^{12} Bq



Contents

1 Introduction	77
2 Cost function, the Poor-man approach and hit-rate	78
2.1 Cost function	78
2.2 The Poor-man assimilation approach	79
2.3 Hit-rate	80
3 Bayesian inference	81
4 Components of the Bayesian inference	82
4.1 Prior distribution	82
4.2 Likelihood	83
4.3 α probability	84
4.4 Detection probability	84
5 MCMC refinement	85
6 4D variation source inversion	85
6.1 Change of emission variable	85
6.2 The background error	86
6.3 Positive definite solution	86
6.4 Background error correlated in time	87
7 Results for the ¹⁰⁶Ru case in 2017	88
7.1 The most promising source locations	88
7.2 Source profile and information gain	89
7.3 The source locations suggested by different met ensemble members	90
8 Ensemble variational assimilation of source profiles	91
9 What if?	93
10 Outlook A: ¹³³I case in 2016 - handling of zero measurements	95

11 Outlook B: CH₄ case from broken Nord Stream pipe-lines in the Baltic	96
12 Conclusions	101
13 Acknowledgement	102
A The ¹⁰⁶Ru case in 2017	102
B Measurements used from the ¹³³I case in 2016	105

Abstract

Through the NKS contracts AFT/B(19)6 and AFT/B(20)3 we have developed means for source localisation and source profile inversion. The contracts have enabled refinement of already existing methodology and new developments.

The outcome from these contracts are from our perspective Bayesian inference for source localisation and improved variational source profile inversion. The latter have implied building a generic library for variational data assimilation. The resulting tools are contained in a pre-operational environment.

1 Introduction

The NKS SOCHAOTIC project follows upon the SLIM project that both are devoted to various aspects of source inversion from given sets of observation material (contracts AFT/B(19)6 and AFT/B(20)3). While SLIM were more oriented towards source localisation SOCHAOTIC is devoted to source inversion where the location is assumed known. The projects had also a component of using ensemble weather predictions.

We have in this presentation chosen to span over both of these project then covering source localisation as well as source profile inversion, that of course are connected.

In Section 2 we describe the cost function that leads to the adjoint transport model and adjoint gradient for the source term update. We present the so called “Poor-man” variational approach that is an approximation and a short-cut to the full variational evaluation but proper enough Bayesian inference.

This is followed by a description of our implementation of Bayesian inference in Section 3, where the meaning of each element is illustrated in Section 4 using the ^{106}Ru case in 2017. We then refine the Bayesian inference by Markov Chain Monte Carlo (MCMC) described in Section 5, and the full variational source profile inversion is described in Section 6.

Results for the ^{106}Ru case from Bayesian inference, MCMC and variational source inversion are shown in Section 7. The most prominent candidate appears from this analysis to be Mayak, where also a nuclear waste treatment facility is located, with a total release of 2×10^{12} Bq.

Exploring the use of ensemble weather predictions is made by an extension to ensemble variational source profile inversion is shown in Section 8

In Section 9 we ask to what level we could have identified the source location during the course of the event.

Finally we make two extensions: first to a ^{133}I case in 2016 in Section 10 where we discuss the impact of zero measurements, and secondly the Nord Stream CH_4 leakage in September 2022 in Section 11.

The described methodology is developed within the framework of the MATCH modelling system at SMHI [Robertson et al., 1999].

2 Cost function, the Poor-man approach and hit-rate

2.1 Cost function

We shall describe variational source inversion by means of a cost function. We first define our transport model operating from time-level 0 to T forced by a set of source columns $Q = Q_1, Q_2, \dots, Q_T$ at the position r ,

$$x_{t+1} = L_t x_t - \Lambda_t x_t + F_r Q_{t+1} \quad (1)$$

$$x_0 = 0 \quad (2)$$

where the different elements are:

- x_t : model state (mixing ratio of activity, Bq/kg air),
- L_t : forward transport model (that includes the meteorology),
- Λ_t : depletion processes (deposition, decay etc.),
- Q_{t+1} : time dependent source column,
- F_r : injector of the source column into the grid at location r .

We now introduce a cost function that is a measure of the distance between the source term and some background and the distance between the model and the measurements,

$$J = \sum_{t=1}^T \frac{1}{2} (Q_t - Q_{t,b})^T B^{-1} (Q_t - Q_{t,b}) + \frac{1}{2} (H_t x_t - d_t)^T O_t^{-1} (H_t x_t - d_t) \quad (3)$$

where $Q_{t,b}$ is some background emissions, d_t are the measurements, H_t the observation operator bringing the model to the observations, and B and O are the background and observation covariance matrices, respectively, both assumed diagonal. The observation operator may just be an interpolation from the grid points to the observation sites, but may contain transformations such as gamma-dose conversion, and imply averaging in time when observations are sampled over longer time periods than the model time resolution (one hour). However the latter is our normal case we have simplified the presentation by assuming model and observations to have the same time resolution. The first term in Eqn 3 with the square difference between the source and background acts as a balancing term not to over-fit the observations.

The cost function Eqn 3 should be solved having the model (Eqn 1) as a strong constraint,

$$J = \sum_{t=1}^T \frac{1}{2} (Q_t - Q_{t,b})^T B^{-1} (Q_t - Q_{t,b}) + \frac{1}{2} (H_t x_t - d_t)^T O_t^{-1} (H_t x_t - d_t) + \lambda_{t+1}^T (L_t x_t - \Lambda_t x_t + F_r Q_{t+1} - x_{t+1}) \quad (4)$$

where λ_{t+1} is the Lagrangian or adjoint state. The minimum of the cost function in Eqn 4 is found where the gradient with respect to the model state, x_t , is zero and we arrive at,

$$\lambda_t = (L_t^T - \Lambda_t)\lambda_{t+1} + H_t^T O^{-1}(H_t x_t - d_t) \quad (5)$$

$$\lambda_T = 0 \quad (6)$$

which is the adjoint model (L_t^T) integrated backwards in time and forced by the innovations (model minus observations). H_t^T is the adjoint of the observation operator that will in the case of gamma dose measurements project the gamma dose innovation back to the contributing nuclides. The initial condition 6 does also come with zero boundaries [Le Dimet and Talagrand, 1986]. The depletion operator Λ_t is assumed diagonal and it is worth noting that this process acts symmetrical in forward and adjoint mode. In forward mode the depletion reduce the impact of the emissions, while in adjoint mode the depletion reduce the impact of the observations.

The gradient of the cost function Eqn 4 with respect to the source term Q_t is given by,

$$\nabla J(Q_t) = F_r^T \lambda_t + B_{t,t}^{-1}(Q_t - Q_{t,b}) \quad (7)$$

where F_r^T is extracting a column at location r of the 3D adjoint state λ_t , plus a friction term with the departure from a background source. The observations contribute to this gradient through the adjoint state, λ_t . The source column could iteratively be updated by an approximation of the Newton-Raphson iteration method,

$$Q_t^{n+1} = Q_t^n - \alpha \nabla J(Q_t) \quad (8)$$

where each new source term will feed back to the innovation and thus an iterated adjoint solution. Formally α is the inverse Hessian of the cost function. In full variational assimilation this is approximated by conjugate gradient methods [Gilbert and Lemaréchal, 1989] but as shown in next section this could be a scalar as we apply in the Poor-man assimilation approach.

2.2 The Poor-man assimilation approach

The Poor-man approach were proposed by Robertson and Langner [1998] and restrict the variational solution to one iteration of the emission update Eqn 8. The assumption is that nothing is known about the source, except for the location, and emissions are thus set to zero, $Q_t = Q_b = 0$, and the model state will consequently be zero at all time levels. By taking one step in Eqn 8 we follow the steepest descent towards emissions from the assumption of no sources. Starting with $\alpha = 1$ our first source term estimate is,

$$Q_t^{fg} = F_r^T \lambda_t^{fg} \quad (9)$$

where the first guess adjoint solution is given by,

$$\lambda_t^{fg} = (L_t^T - \Lambda_t)\lambda_{t+1}^{fg} + H^T O^{-1} d_t \quad (10)$$

From making a forward run using this first-guess emissions we may derive α that fit the best to the observations in quadratic sense,

$$J(\alpha) = \frac{1}{2} \sum_{t=1}^T (\alpha H_t x_t^{fg} - d_t)^T O_t^{-1} (\alpha H_t x_t^{fg} - d_t) \quad (11)$$

here x_t^{fg} is the solution given the source set Q^{fg} . Some care is needed to determine the initial value of α . Consequently we arrive for the Poor-man source term at,

$$Q_t = \alpha Q_t^{fg} \quad (12)$$

In summary the Poor-man approach is using the time-variation given by the steepest descent gradient scaled to fit to the observations.

2.3 Hit-rate

We will here discuss hit-rate or source-receptor relation. This is by nature derived individually for each observation considered. The hit-rate for observation k is derived by the adjoint formulation forced by a unit response at this observation site over the time the observation is valid,

$$\lambda_{t,k}^* = (L_t^T - \Lambda_t)\lambda_{t+1,k}^* + H_{t,k}^T \quad (13)$$

We will then end up with a catalogue of hit-rates for each observation from which we could get the full adjoint formulation,

$$\lambda_t = \sum_{k=1}^N \lambda_{t,k}^* \sigma_k^{-2} (H_{t,k} x_t - d_{t,k}) \quad (14)$$

that is a summation over all observations at some time-level t , where σ_k is the observation error. In turn the forward model integration at the observation point k is derived by a summation over time-steps,

$$H_{t,k} x_t = \sum_{t=t}^T Q_t^T (F_r^T \lambda_t^*) \quad (15)$$

The relations in Eqns 14 and 15 were first suggested by Marchuk [1995] and represents the Lagrangian duality of adjoint operators. The notation hit-rate may be understood by viewing the depletion process, Λ_t . If the depletion is switched off the hit-rate will increase and the emissions needed to hit the observations could be reduced in comparison to when the depletion is on. In Eqns 14 and 15 we have expressed the model at the observation point as an issue of one single model time-step. Normally, however, are the observations covering several model time-steps. The formalism will still hold as the source-receptor relation will be inherent in the hit-rate.

3 Bayesian inference

The Bayesian inference imply spampling a prior probability distribution, adding observations and arriving at a prosterior probability distribution from which we could deduce the source most likely [Monache et al., 2008, Keats et al., 2008, Smith and Gelfand, 1992]. We will consider the following elements for our Baysian inference,

- r , the source location,
- α , the Poor-man α parameter,
- D , detectability,
- d_1, \dots, d_T , the observations

The motivation is as follows: the location r and observations d are obvious elements in the inference, with the Poor-man α we may favour locations with larger hit-rate (lower α) that otherwise score similar toward the observations. The detectability may be understood by the prior distribution that may include locations downstream of one or several observations. Detectability is meant to penalise such locations having in mind that the observations may be wrong.

This lead us to the following Bayesian formula,

$$p(r, \alpha, D | d_1, \dots, d_T) \propto p(d_1, \dots, d_T | r, \alpha) p(\alpha | r) p(D | r) p(r) \quad (16)$$

On the left had side we find the posterior, the ability for the observations to see our parameters, and on the right hand side is the likelihood, the ability of our parameters to see the observations (that require forward model evaluation), followed by the probability of the α , D , and finally the prior of the location r it self. We have dropped an unknown normalisation constant as we only interested in the distribution of the posterior not its absolute value. It could also be noted for the right-hand side that α and D are assumed uncorrelated.

We solve the Bayesian formula in 16 from the right to the left by sampling a prior distribution and the implications of the different elements are shown in Section 4 having the ^{106}Ru in 2017 as example (Appendix A). In Bayesian inferences the prior is what we may know before the observations and the posterior when the observations are added.

4 Components of the Bayesian inference

We are here continuing the description in Section 3 sorting out the various elements in Eqn 16 having the ^{106}Ru case (Appendix A) as example.

4.1 Prior distribution

We propose to derive the prior from the adjoint gradient fed by the observations integrated over time (footprint). Figure 54 shows the time integrated adjoint columns (left) normalised to yield unit maximum. This gives a signal of where or where not we may expect a potential source to be found. The gradient is not a probability suggestion as such but we may still process this to provide a useful prior distribution. In the middle panel we have deduced the prior by setting equal probabilities in an area that enclose values of the integrated gradient down to 1% its maximum value. We have then abandoned any possible variation within the prior domain. A way to gain some more information from the integrated gradient is shown in the rightmost panel in Figure 54 where we have applied a normal quantile transform of the gradient into a Gaussian distribution with moderate spread [Bogner et al., 2012]. This prior will in a soft way favour solutions in the main stream of the integrated gradient.

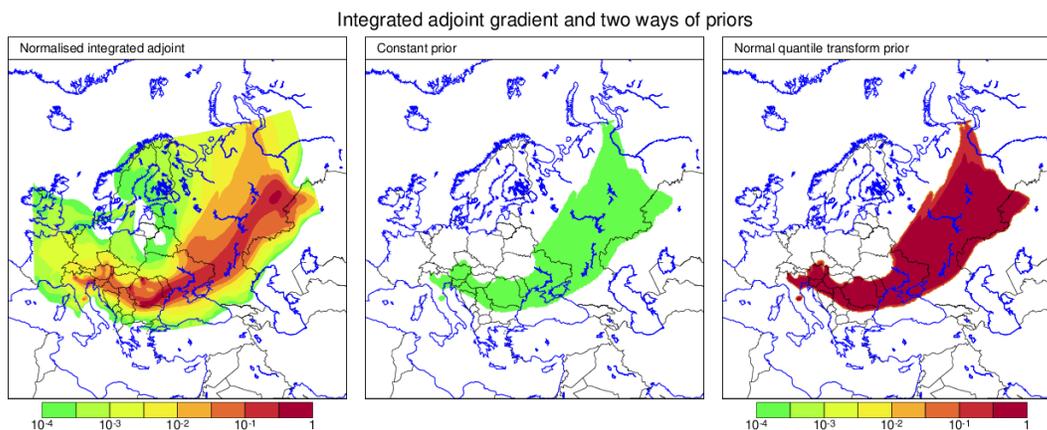


Figure 54: Normalised time integrated adjoint gradient columns (left) from where we have deduced two priors, uniform (middle) and Gaussian by normal quantile transform (right).

Figure 55 shows Gibbs sampling for 100 points from the two priors in Figure 54, uniform distributed prior (left) and normal quantile transform prior to the (right). For each of the sampled points we derive a Poor-man source term for which we will get the α (see Section 2), the likelihood, and the detectability D . Figure 56 shows the different components in our Bayesian formula that are described in the following.

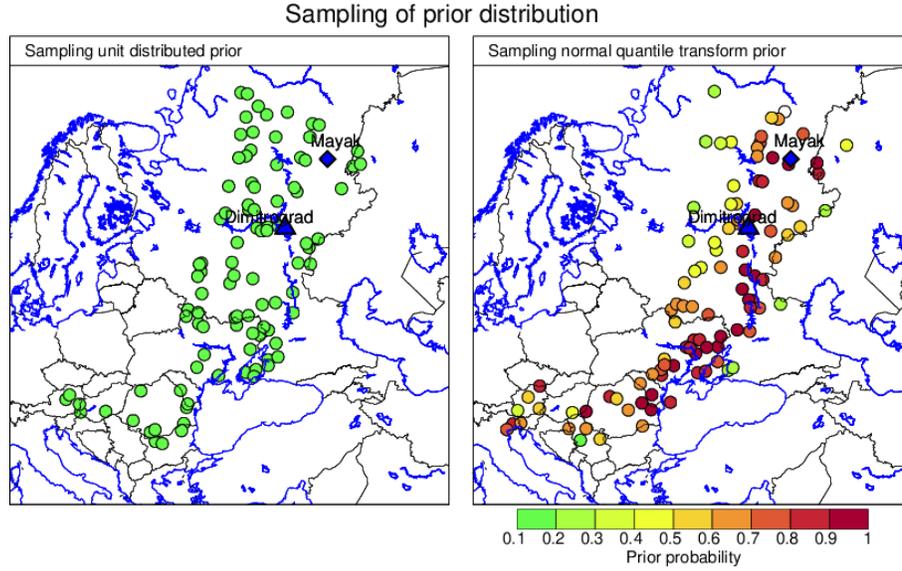


Figure 55: Sampling of unit distributed prior (left), and normal quantile transform prior (right).

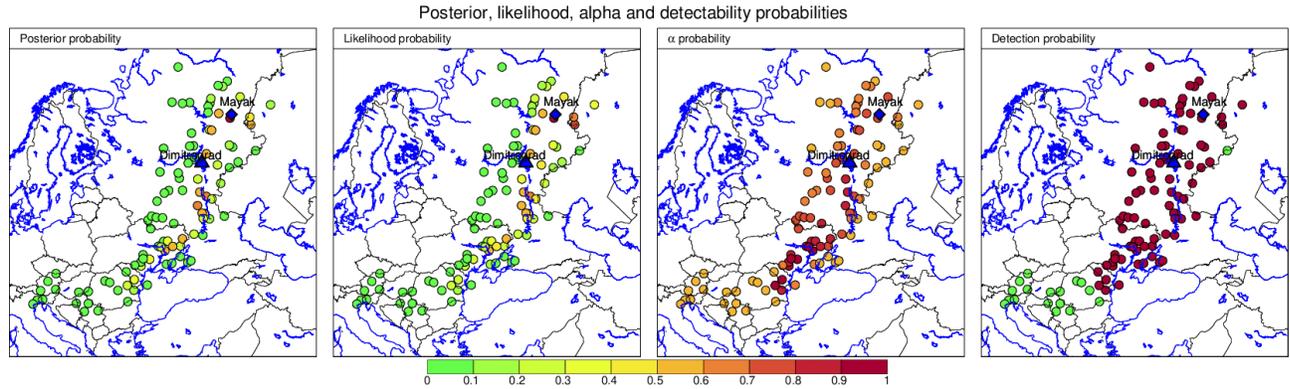


Figure 56: The various components in the Bayes formula in Eqn 16, posterior, likelihood, α and detection probabilities for each selected source location. The likelihood do as expected dominate the posterior output. The contributing distributions are individually normalised to 1.

4.2 Likelihood

The likelihood represents the ability of the model to see the observations within the range of the uncertainty of the observations, and we have defined this as Gaussian distributed,

$$p(d_1, \dots, d_T | r, \alpha) \propto \prod_{t=1}^T e^{-\frac{1}{2}(H_t x_t - d_t)^T O^{-1} (H_t x_t - d_t)} \tag{17}$$

The likelihood involves a forward model evaluation that connect our model parameters with the observations. The forward model at the observation sites ($H_t x_t$) are solved by using the hit-rate and the current source estimate (see Eqn 15).

4.3 α probability

The Poor-man α is a measure of the integrated hit-rate. Larger α is a compensating scaling for lower hit-rate in order to reach up to the measurements. We aim at favour source locations with lower α (thus higher inverse) but as seen from Figure 57 the inverse α ranges in more than 6 magnitudes. This will be a too strong measure and we apply normal quantile transform that take us from left to right in Figure 57 in order to limit the penalty of low hit-rate.

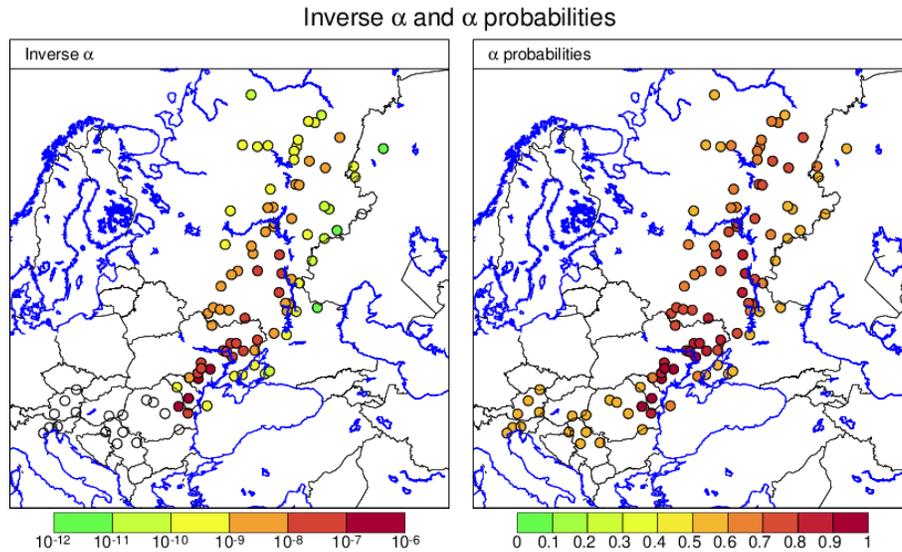


Figure 57: Invers Poor-man α (left) and the α probabilities (right). The right panel is a normal quantile transform of the left one in order to yield a less dramatic measure.

4.4 Detection probability

The detectability probability is simply the number of observations seen by the source location (through the hit-rate) over the total number of observations. As seen from Figure 56 (rightmost panel) there are proposed source locations to the south-west of the domain with lower detectability. If there are some dominating measurements a source location in the vicinity may score well despite being down stream of several observations. The detectability is aimed to penalise such source location. This may however not be the case in the ^{106}Ru case 2017 example where the likelihood is low for source locations with low detectability (Figure 56, second panel).

5 MCMC refinement

It would be just by chance that random selection from the prior would have hit the location that maximize the posterior distribution, and we therefore need to refine the sampling. This is done for the most promising candidates where the surroundings is search by means of Markov Chain Monte Carlo [Grasserman, 2003, Walsh, 2004, Vrugt, 2016]. Figure 58 shows such random walk and deduced particle filter standard deviation around the mean. The latter is there to illustrate additional use of the random walk.

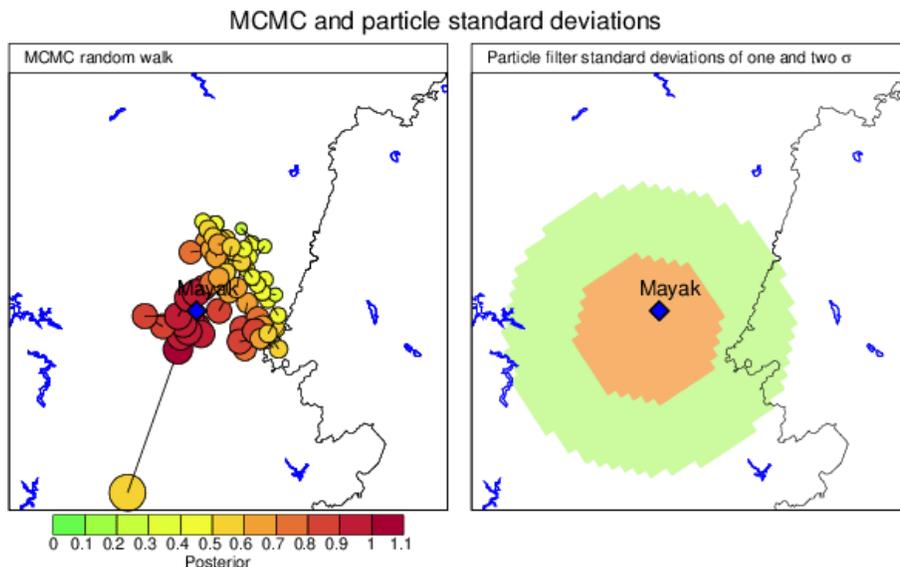


Figure 58: Example of Markov chain random walk (left), and from which we could deduce particle filter standard deviation around the mean for one and two standard deviations (σ) (right).

6 4D variation source inversion

In Section 2 we made a short-cut to the cost function (Eqn 3) arriving at the Poor-man approach and end up omitting the background source term. In a full variational assimilation we have to account for this element. In the following we describe some steps we have to take for the implementation of the variational scheme: change to virtual emissions, background error estimate, and measures to ensure positive emission solutions.

6.1 Change of emission variable

We first restrict us to view the background covariance matrix to be diagonal and with constant error variance, $B = \sigma_b^2 I$ (see an extension in Section 6.4). The standard way of solving the costs function such as Eqn 3 is to change variable to a virtual source term, v_t [Courtier et al., 1993], in our case defined as,

$$\sigma_b v_t = Q_t - Q_{t,b} \quad (18)$$

that converts the cost function to,

$$J = \sum_{t=1}^T \frac{1}{2} v_t^T v_t + \frac{1}{2} (H_t x_t - d_t)^T O_t^{-1} (H_t x_t - d_t) \quad (19)$$

We may identify that the virtual emission, v_t , has zero mean and unit error variances, and is thus an emission space that simplifies minimisation. The gradient with respect to the virtual emission, v_t , is then,

$$\nabla J(v_t) = v_t + \sigma_b F_r^T \lambda_t \quad (20)$$

6.2 The background error

The next part to define is the background error, σ_b , where we will get some help from the Hessian, the second derivative of J . This represents the inverse relative analysis covariance matrix (relative to the background errors) or the information matrix, that is given by,

$$J''(v_t) = 1 + \sigma_b (A_1 A_2 \dots A_T) (A_1 A_2 \dots A_T)^T \sigma_b \quad (21)$$

$$A_t = F_r^T \sum_{k=1}^N \lambda_{t,k}^* \sigma_o^{-1} \quad (22)$$

where A_t summarise over all observation hit-rates at time-level t . Note that the information gain is dependent on the hit-rate but not on the observation values as such. This may look a bit more complicated than it will be in practice. We will only look into the trace of J'' (the diagonal) and moreover sorting out the element that appears at maximum. If we then define a maximum allowed information gain (left hand of Eqn 21) the background error, σ_b , remains to be the only unknown. We have set the maximum allowed information gain (σ_b^2/σ_o^2) to 2.

6.3 Positive definite solution

Variational data assimilation is solved by assuming linear processes where the minimisation is free to suggest negative emissions. Negative emission may also be as good as positive ones to explain the observations in quadratic sense. We therefore need to explicitly tell the minimisation that we dislike negative emissions. One way is to redefine the source term in the model equation 1,

$$x_{t+1} = L_t x_t - \Lambda_t x_t + F_r S_+(Q_{t+1}) \quad (23)$$

where S_+ is an element-wise softplus function, $\ln(1 + e^x)$, which derivative s_+ is the sigmoid function, $1/(1 + e^{-x})$, appearing in the gradient of the cost function with respect to the virtual emissions,

$$\nabla J(v_t) = v_t + s_+(Q_t)\sigma_b F_r^T \lambda_t \quad (24)$$

where v_t relates to Q_t from Eqn 18. The nature of the modifications of the model equation and gradient are shown in Figure 59. The implication for the forward model is that for suggested positive emissions these will be returned untouched, while negative suggested emissions are returned as zero. For the gradient negative suggested emissions will switch off the gradient. This implies that the forward model will always be fed with positive or zero emissions while we do not allow the gradient to direct towards negative emissions. The non-linearity introduced comes with more iterations during the minimization.

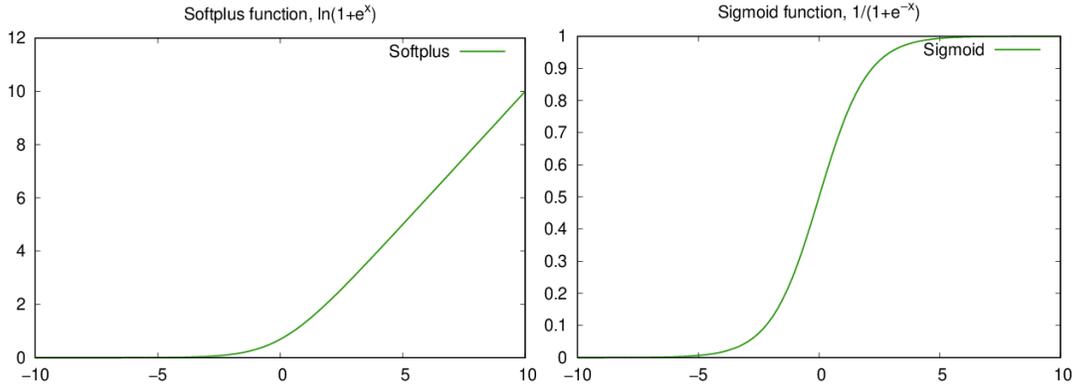


Figure 59: The softplus function (left) and the sigmoid (right), used to ensure positive emissions in the variational data assimilation.

6.4 Background error correlated in time

In the above we have assumed the background error matrix to be diagonal. We may though make an extension where the errors are correlated in time. This is used in Section 11 where we discuss the Nord Stream case. We then describe the background error matrix as following,

$$B = \sigma_b D D^T \sigma_b \quad (25)$$

where D is a square-root diffusion operator in time. The virtual emission (Eqn 18) will then appear somewhat differently,

$$\sigma_b D v = Q - Q_b \quad (26)$$

where the suffix t is removed as all time-steps are involved, and consequently the gradient is filtered by the diffusion,

$$\nabla J(v) = v + D^T \sigma_b F_r^T \lambda \quad (27)$$

The idea behind is that we assume the time correlations to be Gaussian and that the Gaussian equation is a solution to the diffusion equation (given some constraints) [Elbern et al., 2007]. To make the diffusion square-rooted is straight forward. For clarity we have included the transpose of the diffusion operator but the diffusion operator is symmetric.

7 Results for the ^{106}Ru case in 2017

The results presented are mostly based on the unperturbed member of the HARMONIE forecasts provided by DMI on a resolution of 0.2° . The normal quantile transform prior is used giving somewhat more priority to the mayor flow direction. An outlook is made to some other of the met ensemble member in Section 7.3 in terms of posterior of most promising source locations.

7.1 The most promising source locations

Figure 60 shows the posterior for the most promising source locations from the first Gibbs sampling and after MCMC refinement, respectively using the unperturbed HARMONIE forecast. The MCMC refinement works as expected were specifically a couple of locations to the west of Mayak have moved in close to Mayak. The MCMC also moved sampled locations south of Dimitrograd closer rising the probability for this location.

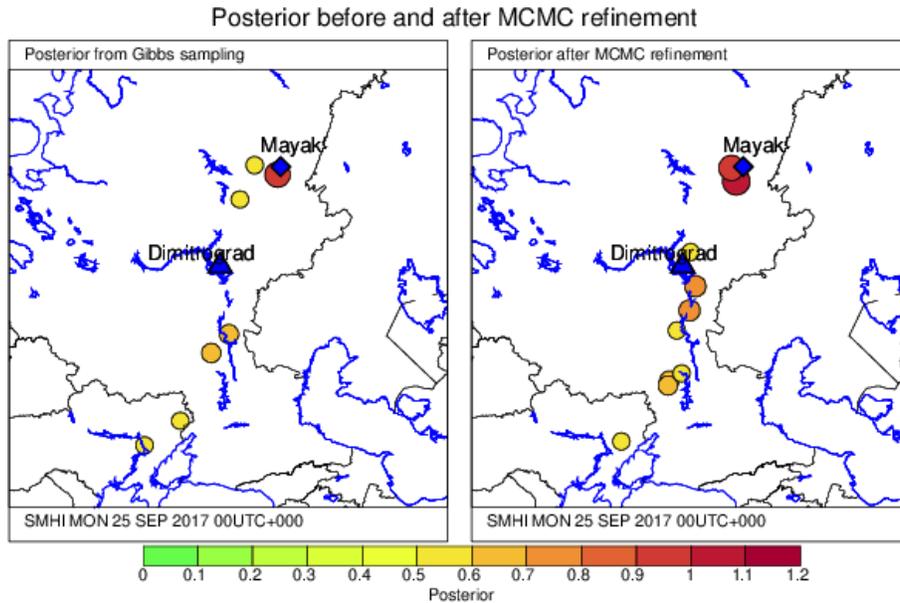


Figure 60: The most promising source locations from the Gibbs sampling (left) and the MCMC refinement (right).

7.2 Source profile and information gain

Figure 61 shows results from variational source inversion for the first guess, the final variational source, and the information gain, respectively. The first guess is the Poor-man source term and the information gain is the trace of the Hessian of J (J'') from which we have derived the source background error, $\sigma_b = 12 \times 10^3$ mBq/m³, given by an observation error $\sigma_o = 1$ mBq/m³. The assimilated source has a peak near surface around 15 UTC Tuesday 26 September. The model source is however spread over three days from 25 to 27 September and propagates up to 10 km during the course of long range inverse transport. We have thus not put any constraints on the expected vertical extent, just given the model free hands to suggest what best fit the observations. The total emission needed to explain the observations is 2×10^{12} Bq. We should realise that the measurements used are up to 1 day of sampling that are not able to return a distinct pulse if that could have been the case.

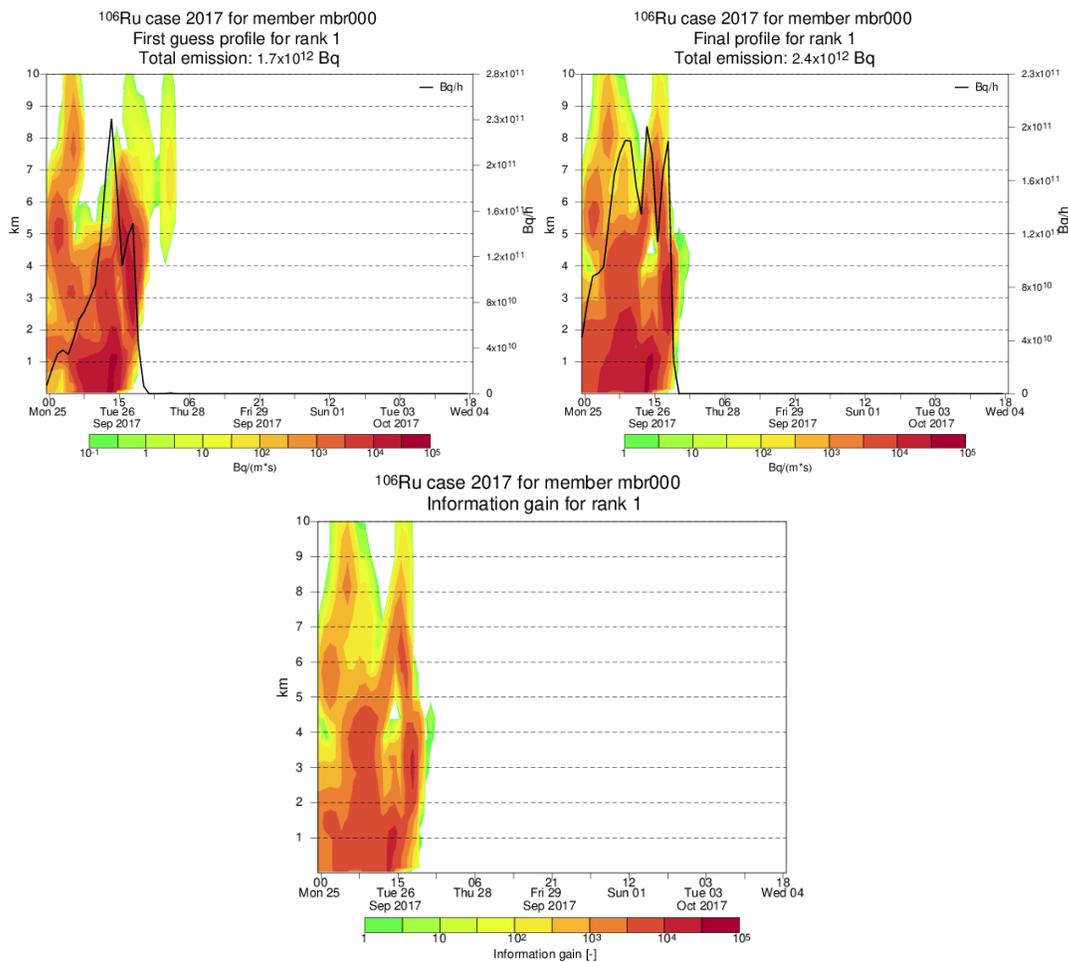


Figure 61: First guess emission profile for the ^{106}Ru case in 2017 (upper left), assimilated emission profile (upper right, and information gain (lower middle).

7.3 The source locations suggested by different met ensemble members

The results above are from using the unperturbed met forecasts. Figure 62 shows posterior for the most promising locations using the met ensemble members 5, 10, 15, 20, together with the unperturbed. The ensemble members are more or less giving the same picture as the unperturbed. The ensemble members 10 and 20 do also pinpoint Dimitrograd as the second most likely. This also true for the unperturbed met forecast even-though not as strong. The ensemble member numbering is a bit unambiguous as consecutive forecasts charing the same member number are not connected except for the unperturbed member.

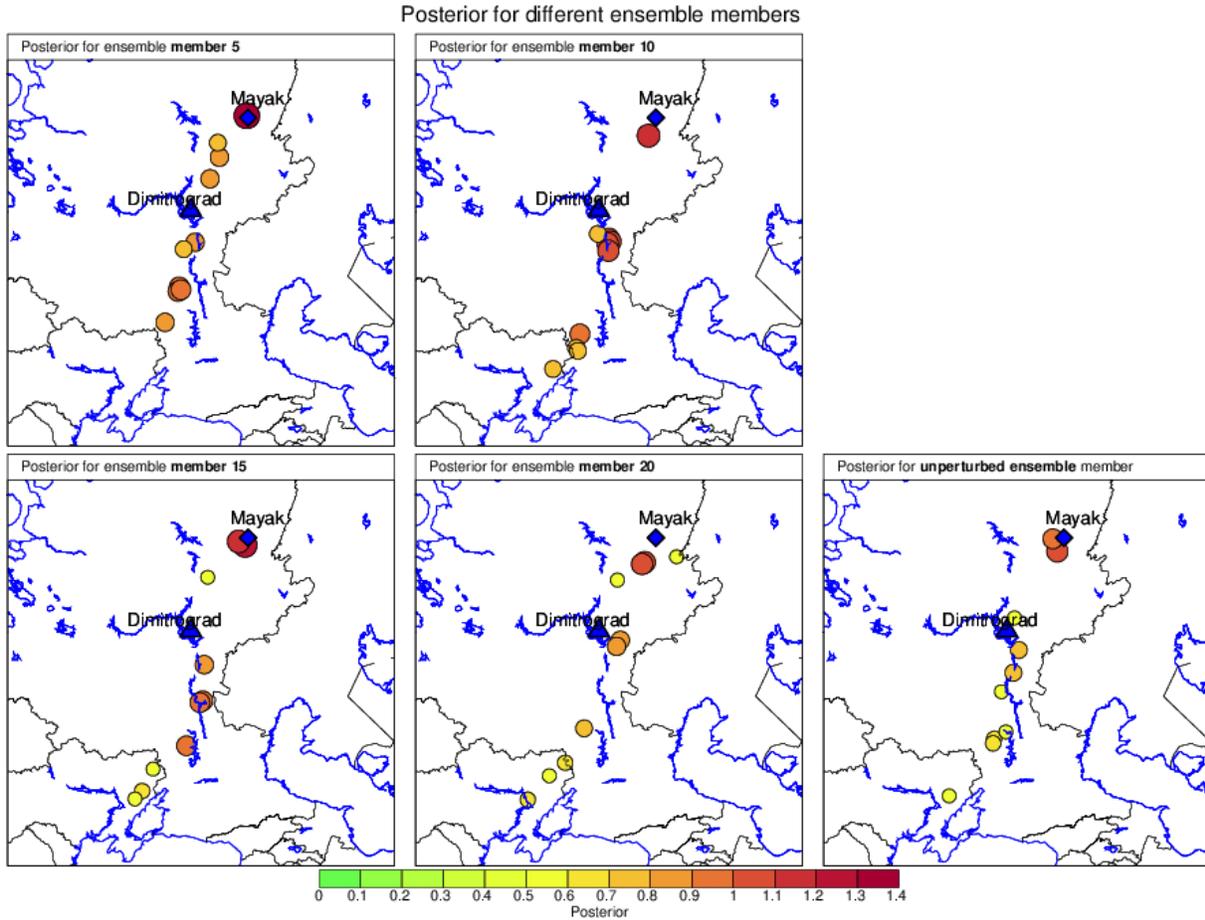


Figure 62: Posterior of the most promising locations for the ensemble members 5, 10, 15, 20 and the unperturbed member.

8 Ensemble variational assimilation of source profiles

In the above we have shown how to perform source inversion for a single met forecast or single ensemble member. We may though by a slight tweak of the variational scheme solve for a number of ensemble members simultaneously. We first rewrite the model in Eqn 1 to denote different ensemble members, $i = 1 \dots M$,

$$x_{t+1,i} = L_{t,i}x_{t,i} - \Lambda_{t,i}x_{t,i} + F_r Q_{t+1,i} \text{ for } i = 1, \dots, M \quad (28)$$

where the model operator $L_{t,i}$ is ensemble specific as we use different weather-data, as well as for the depletion, $\Lambda_{t,i}$, where deposition processes are dependent on the meteorology. The trick we do is to define a cost function where the observation part uses the ensemble mean, and we add a constraint of fulfilling all the used ensemble models,

$$J = \sum_{t=1}^T \left(\frac{1}{2} (Q_t - Q_{t,b})^T B^{-1} (Q_t - Q_{t,b}) + \frac{1}{2} (H_t \bar{x}_t - d_t)^T O_t^{-1} (H_t \bar{x}_t - d_t) \right. \\ \left. \sum_{i=1}^M \lambda_{t+1,i}^T (L_{t,i} x_{t,i} - \Lambda_{t,i} x_{t,i} + F_r Q_{t+1,i} - x_{t+1,i}) \right) \quad (29)$$

where \bar{x}_t is a weighed ensemble mean,

$$\bar{x}_t = \sum_{i=1}^M w_i x_{t,i} \quad (30)$$

We have chosen to have the weights, w_i , equally set to $1/M$, but other weighting could be possible, like having stronger weight on the unperturbed member. The adjoint solution for the member i is then,

$$\lambda_{t,i} = (L_{t,i}^T - \Lambda_{t,i}) \lambda_{t+1,i} + w_i H^T O^{-1} (H_t \bar{x}_t - d_t) \quad (31)$$

which means that ensemble source update is driven by its share of the ensemble mean deviation from the observations. Figure 63 shows an example of what could be expected together with the ensemble standard deviation that illustrates the disagreements among the ensembles.

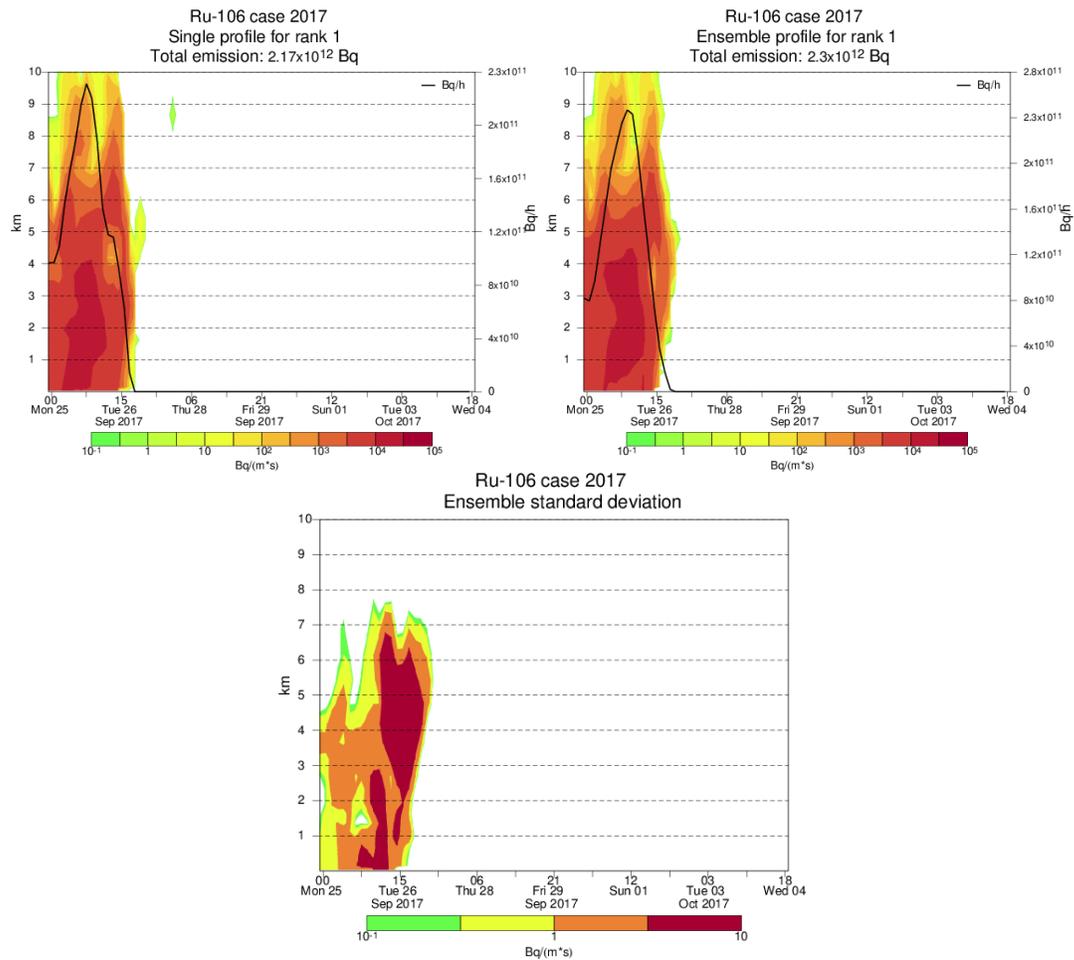


Figure 63: Demonstration of ensemble variational source inversion for a location close to Mayak. Shown are the single source inversion using unperturbed met data (upper left), assimilation where also perturbed ensembles contribute through the ensemble mean (upper right), and the ensemble standard deviation (bottom). The perturbed members used are 5, 10, 15 and 20.

9 What if?

During an event we are first served with some early measurements while more follow during the course of the event. International cooperation may imply additional measurement as time goes by. We may then ask how well this proposed approach would have answered the source inversion issue at different dates. Figures 64 and 65 show measurements used up to 1 and 2 September, respectively, for values and number of measurements, respectively. Some of the sites have measurements both for 1 and 2 September. Figure 66 shows the impact on localisation for these two datasets. Having just observations up to 1 September is not sufficient to point out neither Dimitrograd nor Mayak, while when adding observations from 2 September Mayak become clearly pointed out. The conclusion is that we are dependent on the observations available and that the pointed out source locations may change as

the event proceeds.

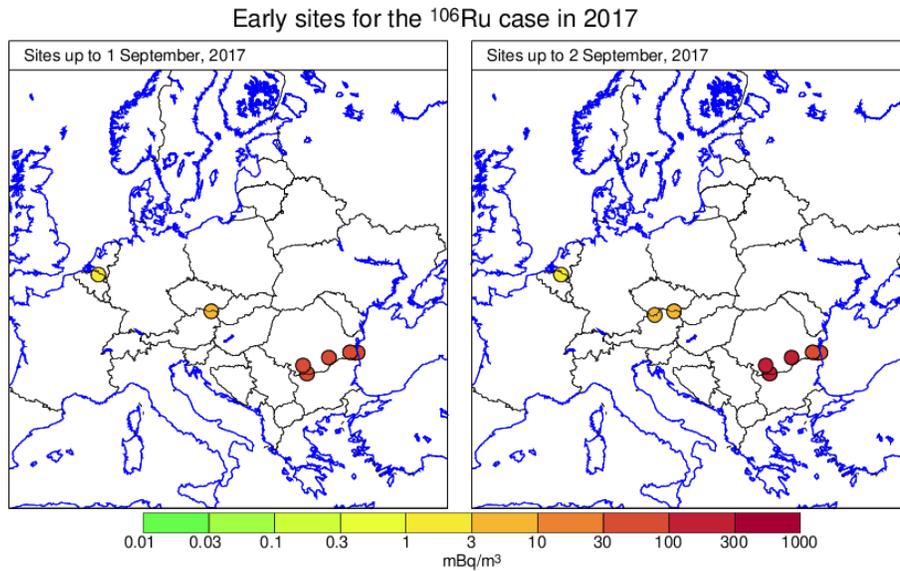


Figure 64: Values for selected observations up to 1 September (left) and up to 2 September, 2017 (right).

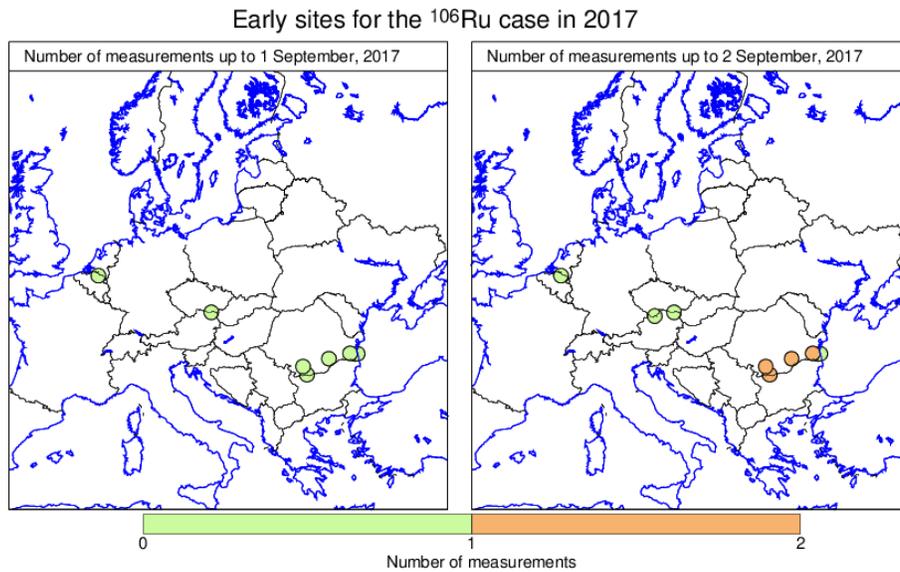


Figure 65: Number of measurements for selected observations up to 1 September (left) and up to 2 September, 2017 (right).

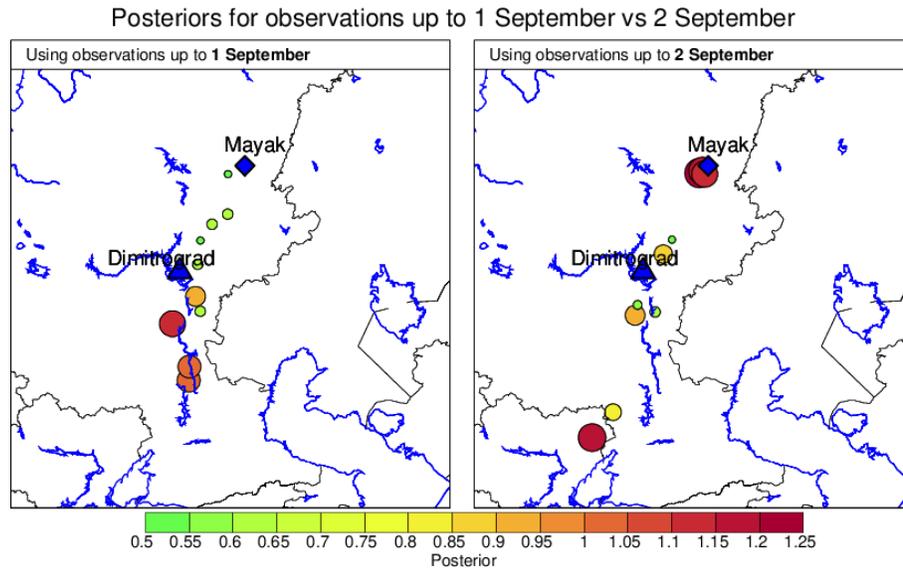


Figure 66: Impact on source localisation using observations for one day up to 1 September (left) and for two days up to 2 September (right).

10 Outlook A: ^{133}I case in 2016 - handling of zero measurements

A case of zero measurements pose a specific demand on the source inversion. The zero measurements may indicate by their adjoint influence area (footprint) where the source should not be expected. We suggest to bring this into the prior then excluding these areas already when sampling the prior. Figure 67 shows maximum measurements for a ^{133}I case in 2016, except for sites Helsinki, Ivalo and Sodankylä that have zero measurement marked. The site Helsinki did however not have zero measured for all times (see Appendix B).

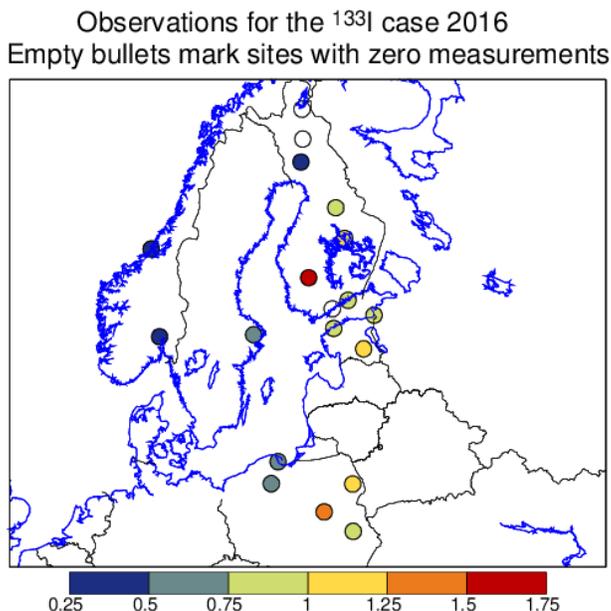


Figure 67: ¹³³I measurements where maximum values are shown except for Helsinki, Ivalo and Sodankylä that show zero measurements.

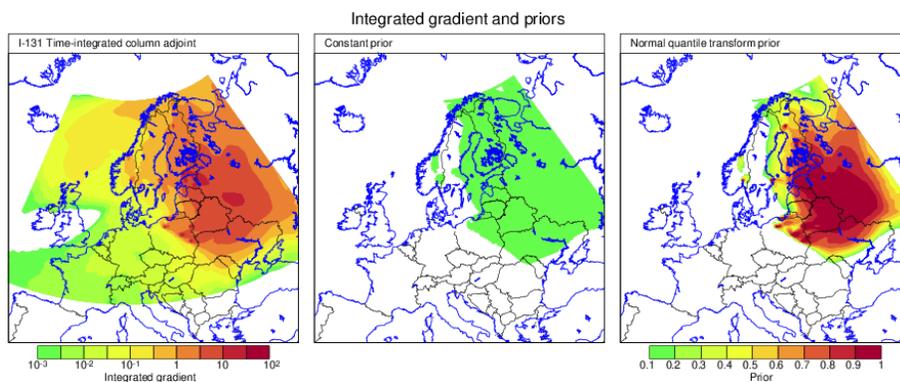


Figure 68: Time integrated adjoint gradient columns (left) for the ¹³³I case from where we have deduces two priors, unified distributed (middle) and smooth Gaussian by normal quantile transform (right). Note that the footprint of zero measurements are excluding western part of the domain from the priors.

11 Outlook B: CH₄ case from broken Nord Stream pipe-lines in the Baltic

In late September 2022 the pip-lines Nord Stream I and II were broken close to the Danish island Bornholm. The locations are known and the source rate of CH₄ was estimated to 2.3×10^7 g/h [Reuters, 2022]. The network for measuring greenhouse gases [ICOS, 2022] have two Swedish sites recording a pulse from the broken pip-lines, Norunda and Hyltemossa, as

shown in the Figures 69 and 70, where the locations of the stations are shown in Figure 71.

Figure 72 shows the footprint from these measurement (time-integrated adjoint run) from which uniform and Gaussian priors are derived. We have here restricted the priors to only be valid for sea areas. This is partly related to the lack of observational information, where then more knowledge is put into the prior. Figure 73 shows the sampling of uniform and Gaussian distributions and Figure 74 shows the resulting posteriors for the highest ranked locations. The locations of the sources are in this analysis slightly to the west of Bornholm for the both priors while the know broken pip-lines are slightly to the east of Bornholm.

The source profile that arise by the used measurements, for the most likely location, is shown in Figure 75, that also arrive at source rates in the same order as the reported ($\approx 2 \times 10^7$ g/h). In this source term only those parts of the source that could be verified by the measurements are captured. By assuming the background error matrix to have correlation in time (see Section 6.4) the retrieved source will be continuous as shown in Figure 76. Note that the magnitude of the source rate does not change by adding time correlation.

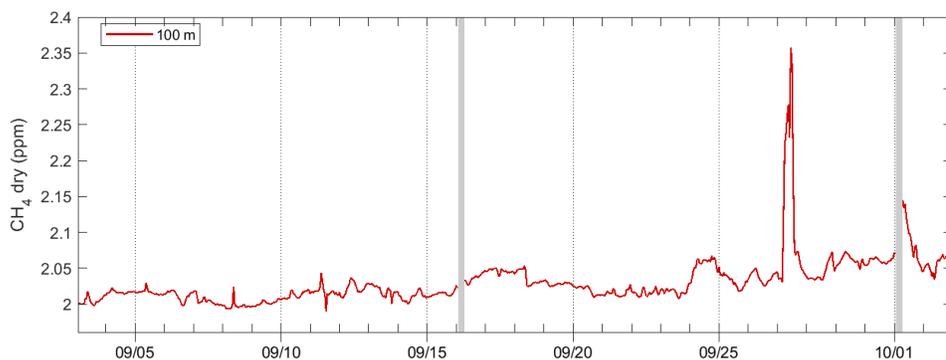


Figure 69: Time-series of CH_4 in ppm at Norunda (Sweden), 17.66E, 60.03N, September 2022.

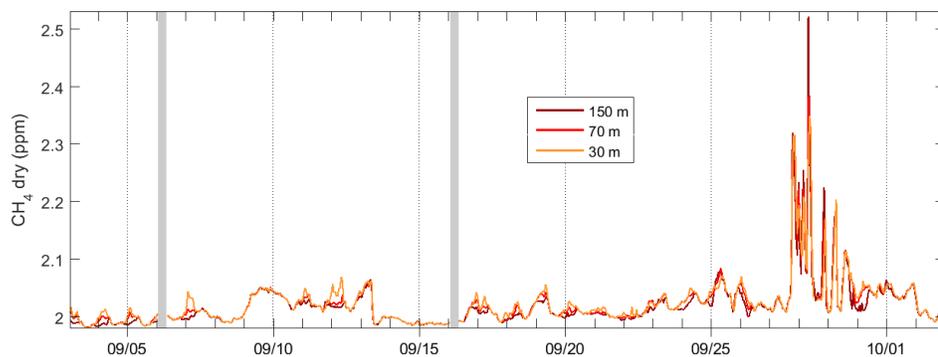


Figure 70: Time-series of CH_4 in ppm at Hyltemossa (Sweden), 13.42E, 56.8N, September 2022.

ICOS sites Norunda and Hyltemossa

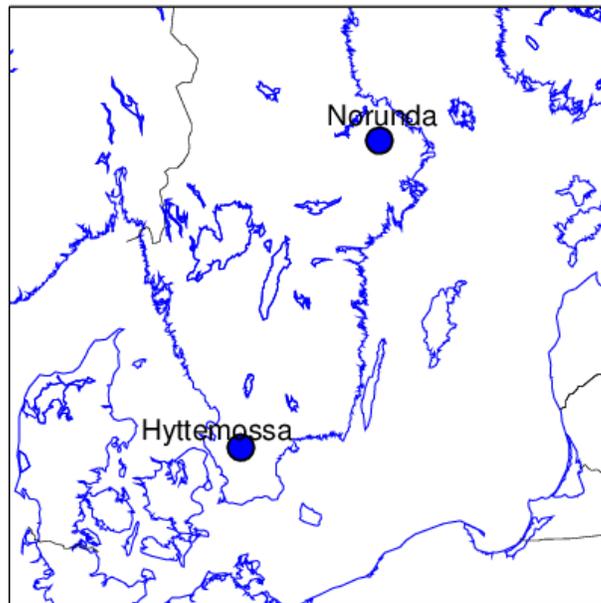


Figure 71: Locations of the stations Norunda and Hyltemossa.

CH₄ case of Nord Stream broken pipes, September 2022 Footprint and priors for uniform and Gaussian distributions

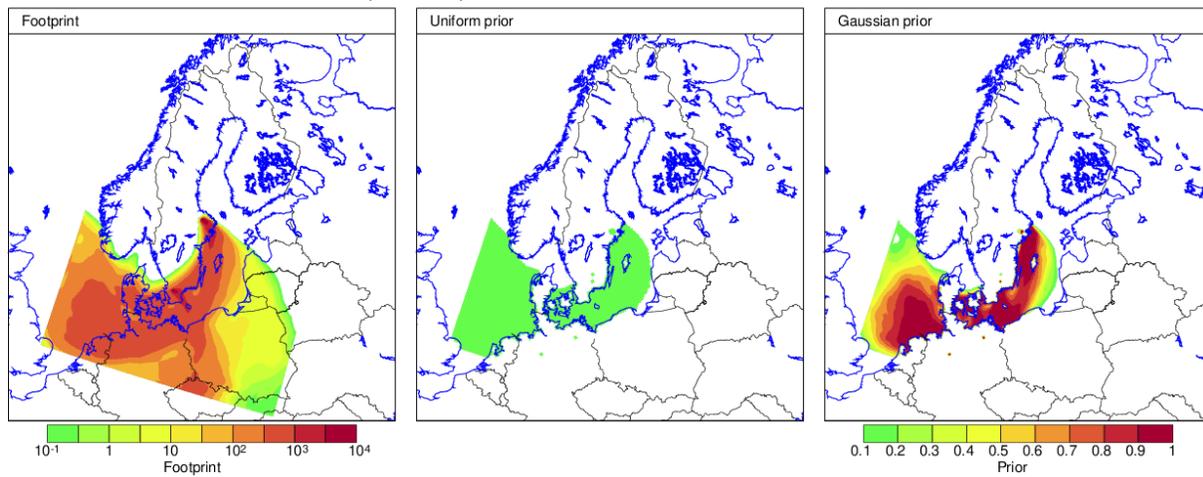


Figure 72: Footprint (left), uniform prior (middle) and Gaussian prior (right). The priors are restricted to sea areas.

CH₄ case of Nord Stream broken pipes, September 2022
Priors for uniform and Gaussian distributions

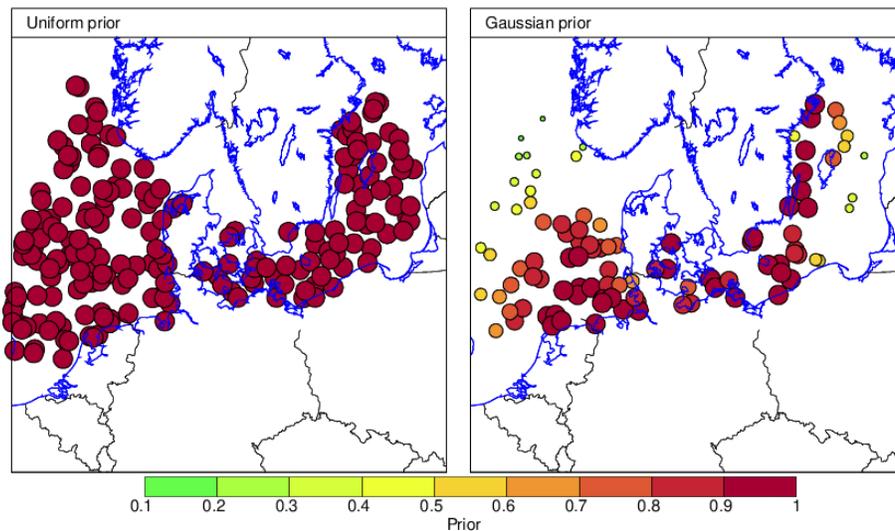


Figure 73: Sampled uniform prior (left) and Gaussian prior (right).

CH₄ case of Nord Stream broken pipes, September 2022
Posteriors for uniform and Gaussian priors

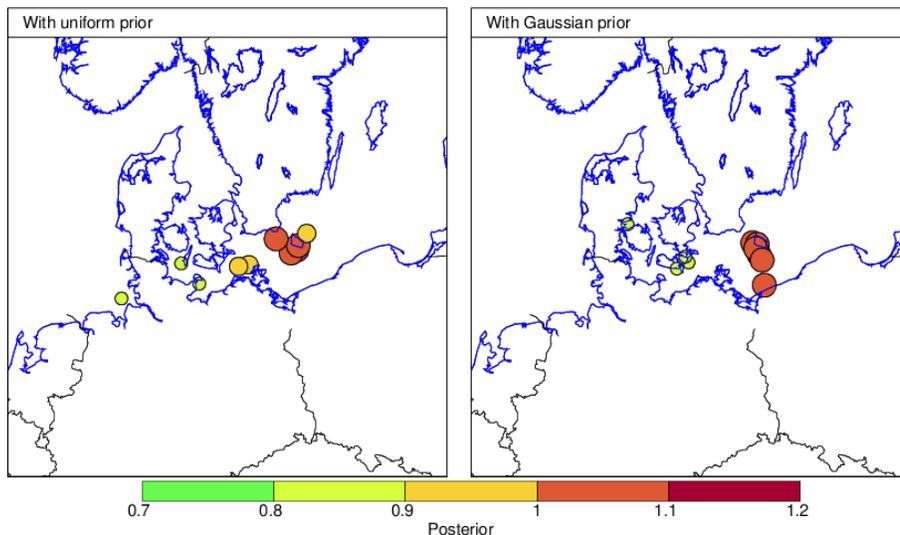


Figure 74: Potential source locations from uniform prior (left) and Gaussian prior (right).

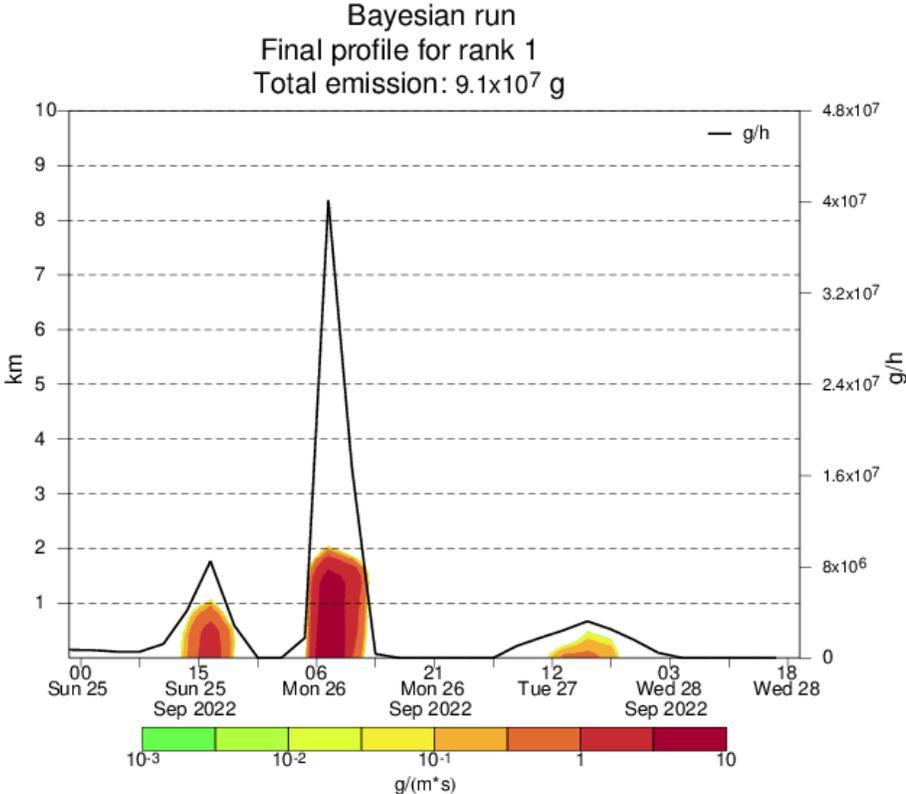


Figure 75: Source profile for the most likely source location close to Bornholm. The first two peaks arise from the Norunda site and the third from Hyltemossa.

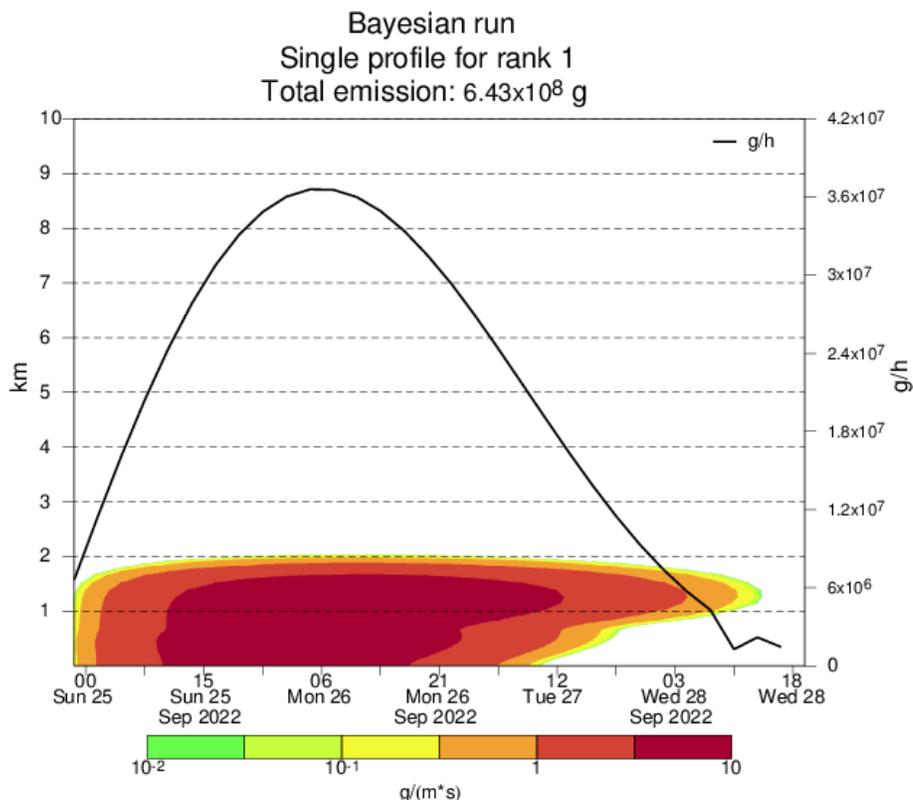


Figure 76: Source profile for the most likely source location close to Bornholm where the background error matrix is assigned correlation in time.

12 Conclusions

Bayesian inference is a strong tool for source localisation in combination with MCMC refinements and variational source profile inversion. The Poor-man approach open up for simplification making massive Bayesian sampling durable. The methodology did, by the selected observations used, quite clearly point out Mayak as the possible source for the ^{106}Ru event in 2017 with a release amount of 2×10^{12} Bq. The reason and location for the event is however so far not disclosed.

We have also shown that the message may be less clear during the course of the event. Having only early measurements at hand does not point out Mayak as the most likely location, while adding another day of measurements does.

Measured zero values are important as “invalid” source locations could be excluded from the solution. We propose to put this into the prior distribution and thus already when sampling the prior exclude locations that will have strong impact on measurement that should be explained as zero.

The Nord Stream case with leakage of CH_4 from broken pipe-lines in late September 2022 is well captured both in location and in magnitude of the source rate. This case with continuous source opened up to show that background error matrix with time correlation enable

the source inversion to catch a contentious source.

13 Acknowledgement

We wish to acknowledge NKS for their support to this work through the contracts AFT/B(19)6 and AFT/B(20)3.

A The ^{106}Ru case in 2017

In 2017 a case with elevated ^{106}Ru was recorded at a rising number of filter stations over Europe [IAEA, 2017]. The source of the event is still not known. ^{106}Ru arise specifically during nuclear waste treatment that limits potential site origins and potential candidates were the waste treatment plants in Dimitrograd or Mayak (Russia).

We have restricted the case to a period from 25 September to the 5 October, 2017. From the 320 number of measurement we have selected 30 with sampling periods up to 24 hours (Figures 77 and 78). The selection was made by first taken the 10 largest measured values (with the limit on sampling time), and then random selection using weighted boot-strapping for the remaining 20 measurements, where the weights were the measured values. The limit to 30 measurements were taken for computational load reasons.

The forcing meteorological data were provided by DMI for HARMONIE ensemble prediction forecasts (REF?). The provided weather data resolution was 5 km on Lambert projection but we have extrapolated to 0.2 degree resolution in rotated latitude-longitude projection.

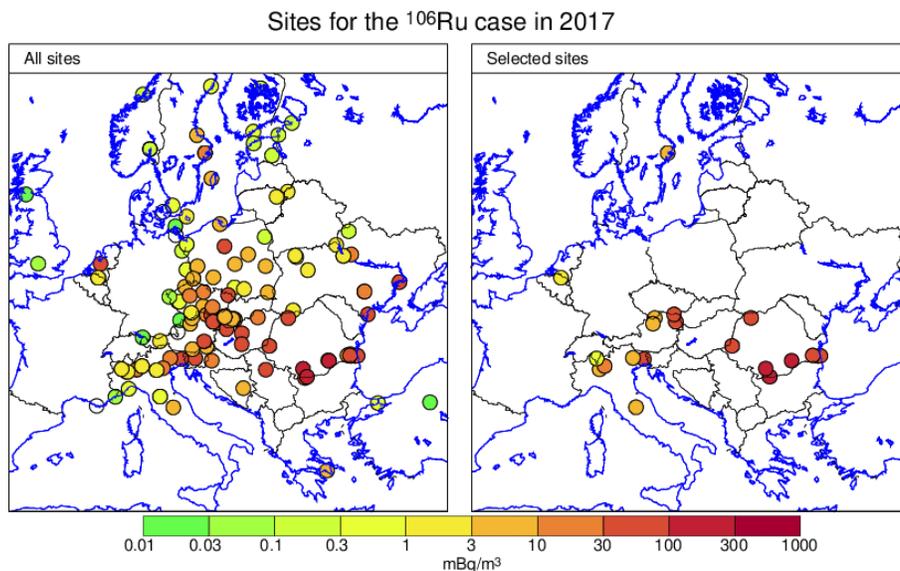


Figure 77: All measurements compiled by IAEA for the period 25 September to 13 October 2017 (left) and the ones used in this study for the period 25 September for 5 October 2017 (right).

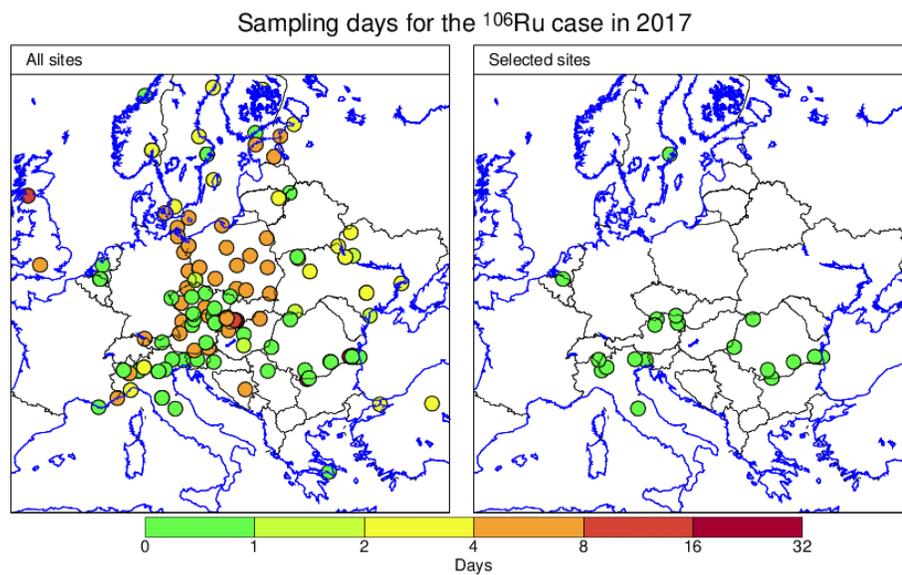


Figure 78: Sampling time in days for the various measurements. The selected measurements are all with sampling times within one day.

Table 1: Measurements used for the ^{106}Ru case, mBq/m^3 .

Site	Longitude	Latitude	Value	Start	End	Duration (h)
Constanta	28.64	44.16	88.1	2017 09 29 00	2017 09 30 00	24
Cernavoda	28.04	44.33	81.28	2017 09 29 00	2017 09 30 00	24
Bechet	23.96	43.79	66.6	2017 09 29 00	2017 09 30 00	24
Craiova	23.8	44.34	59.82	2017 09 29 00	2017 09 30 00	24
Bucuresti	26.15	44.4	37.91	2017 09 29 00	2017 09 30 00	24
Mol	5.1	51.22	1.71	2017 09 29 07	2017 09 30 07	24
Laa_a/d_Thaya	16.4	48.74	4.51	2017 09 29 19	2017 09 30 20	25
Cernavoda	28.04	44.33	57.89	2017 09 30 00	2017 10 01 00	24
Bechet	23.96	43.79	127.59	2017 09 30 00	2017 10 01 00	24
Craiova	23.8	44.34	106.27	2017 09 30 00	2017 10 01 00	24
Bucuresti	26.15	44.4	145.01	2017 09 30 00	2017 10 01 00	24
Leopoldschlag	14.51	48.62	4.51	2017 09 30 13	2017 10 01 13	24
Bechet	23.96	43.79	18.11	2017 10 01 00	2017 10 02 00	24
Baia_Mare	23.59	47.66	47.48	2017 10 01 00	2017 10 02 00	24
Bucuresti	26.15	44.4	18.11	2017 10 01 00	2017 10 02 00	24
Arad	21.32	46.19	64.42	2017 10 01 00	2017 10 02 00	24
Leopoldschlag	14.51	48.62	4.2	2017 10 01 13	2017 10 02 16	27
Laa_a/d_Thaya	16.4	48.74	40.01	2017 10 01 21	2017 10 02 10	13
BERGAMO	9.68	45.69	12.2	2017 10 02 00	2017 10 03 00	24
MILANO	9.23	45.48	6.31	2017 10 02 00	2017 10 03 00	24
Arad	21.32	46.19	30.41	2017 10 02 00	2017 10 03 00	24
PERUGIA	12.34	43.09	6.2	2017 10 02 00	2017 10 03 00	24
Nicosia	33.37	35.17	0.05	2017 10 02 00	2017 10 03 00	24
Stockholm_(SEP63)	17.96	59.39	9.81	2017 10 02 09	2017 10 03 09	24
Wien	16.49	48.26	46.01	2017 10 02 09	2017 10 03 09	24
Linz	14.29	48.27	8.01	2017 10 02 10	2017 10 03 10	24
UDINE	13.24	46.07	54.3	2017 10 03 00	2017 10 04 00	24
Baia_Mare	23.59	47.66	24.61	2017 10 03 00	2017 10 04 00	24
BELLUNO	12.25	46.17	9.31	2017 10 03 00	2017 10 04 00	24
Unknown	8.95	46.16	0.32	2017 10 04 00	2017 10 05 00	24

B Measurements used from the ^{133}I case in 2016

Measurements provided by SSM for a ^{133}I case in 2016. The hour of the measurements is set to 00 UTC that may not be fully true.

Table 2: Measurements used for the ^{133}I case in 2016, $\mu\text{Bq}/\text{m}^3$. The zero measurement are marked in **bold**.

Site	Longitude	Latitude	Value	Start	End	Duration (h)
Warszawa	21.0	52.22	1.3	2016 10 10 00	2016 10 17 00	168
Bialystok	23.22	53.07	1.04	2016 10 10 00	2016 10 17 00	168
Gdynia	18.53	54.52	0.56	2016 10 10 00	2016 10 17 00	168
Lublin	22.57	51.23	0.9	2016 10 10 00	2016 10 17 00	168
Torun	17.86	53.68	0.61	2016 10 10 00	2016 10 17 00	168
Narva-Joseuu	28.05	59.46	0.88	2016 10 15 00	2016 10 23 00	192
Harku	24.6	59.4	0.797	2016 10 15 00	2016 10 23 00	192
Toravere	26.46	58.26	1.09	2016 10 16 00	2016 10 23 00	168
Osteras (Norway)	10.35	60.0	0.4	2016 10 16 00	2016 10 24 00	192
Orland (Norway)	9.6	63.67	0.4	2016 10 16 00	2016 10 24 00	192
Helsingfors	24.9	60.2	1.1	2016 10 16 00	2016 10 18 00	48
Ivalo	27.5	68.65	0.0	2016 10 17 00	2016 10 24 00	168
Sodankyla	26.58	67.42	0.0	2016 10 17 00	2016 10 25 00	192
Loviisa	26.3667	60.3667	1.0	2016 10 17 00	2016 10 25 00	192
Stockholm	18.1	59.8	0.7	2016 10 17 00	2016 10 22 00	120
Kajaani	27.75	64.22	1.0	2016 10 17 00	2016 10 24 00	168
Kuopio	27.66	62.88	1.1	2016 10 17 00	2016 10 24 00	168
Lakiala	23.6	61.7	1.6	2016 10 17 00	2016 10 25 00	192
Rovaniemi	25.7	66.5	0.4	2016 10 17 00	2016 10 24 00	168
Helsingfors	24.9	60.2	0.0	2016 10 18 00	2016 10 20 00	48
Helsingfors	24.9	60.2	1.2	2016 10 20 00	2016 10 22 00	48
Helsingfors	24.9	60.2	4.0	2016 10 22 00	2016 10 24 00	48

References

- K. Bogner, F. Pappenberger, and H. L. Cloke. Technical note: The normal quantile transformation and its application in a flood forecasting system. *Hydrol. Earth Syst. Sci.*, 16: 1085–1094, 2012. doi: 10.5194/hess-16-1085-2012.
- P. Courtier, J. Derber, R. Errico J.-F. Louis, and T. Vukisevic. Important literature on the use of the adjoint, variational methods and the kalman filter in meteorology. *Tellu*, 45A: 342–357, 1993.
- H. Elbern, A. Strunk, H. Schmidt, and O. Talagrand. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, 7: 3749–3767, 2007.
- J.C. Gilbert and C. Lemaréchal. Some numerical experiments with variable storage quasi-newton algorithms. *Mathematical Programming*, 45:407–435, 1989.
- Paul Grasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.
- IAEA. *Updated Technological Attachment: Status of measurements of Ru-106 in Europe*. IAEA, 2017.
- ICOS. Icos sweden is the national contribution a joint european effort for european coverage aiming to provide accessible, high-quality data to improve our understanding of greenhouse gas sinks and emissions., 2022. URL <https://www.icos-sweden.se>.
- Andrew Keats, Eugene Yee, Fue-Sang Lien, and Real D’Amours. Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. *Journal of Wind Engineering*, 96:1895–1816, 2008.
- F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, (38A):97–110, 1986.
- Guri I. Marchuk. *Adjoint equations and analysis of complex systems*. Kluwer Academic Publisher, Boston, 1995.
- Luca Delle Monache, Julie K. Lundquist, Branko Kosovic, Gardar Johannesson, Kathleen M. Dyer, Roger D. Aines, Fotini K. Chow, Rich D. Belles, William G. Hanley, Shawn C. Larsen, Gwen A. Loosmore, John J. Nitao, Gayle A. Sugiyama, and Philip J. Vogt. Bayesian inference and markov chain monte carlo sampling to reconstruct a contaminant source on a continental scale. *J. Appl. Meteorol.*, 47:2600–2613, 2008.
- Reuters. Nord Stream rupture may mark biggest single methane release ever recorded, U.N. says, 2022. URL <https://www.reuters.com/world/europe/nord-stream-rupture-may-mark-biggest-single-methane-release-ever-recorded-un-2022-09-30>
- Lennart Robertson and Joakim Langner. Source function estimate by means of variational data assimilation applied to the ETEX-I experiment. *Atmospheric Environment*, 32(24): 4219–4225, 1998.
- Lennart Robertson, Joakim Langner, and Magnuz Engardt. An eulerian limited-area atmospheric transport model. *Applied Meteorology*, 38(2):190–210, 1999. doi: 10.1175/1520-0450(1999)038<0190:AELAAT>2.0.CO;2.
- A.F.M. Smith and A.E. Gelfand. Bayesian statistics without tears: A sampling-resampling perspective. *The American Statistician*, 42 (2):84–88, 1992.

- Jasper A. Vrugt. Markov chain monte carlo simulation using the dream software package: Theory, concepts, and matlab implementation. *Environmental Modelling & Software*, 75: 273–314, 2016. doi: 10.1016/j.envsoft.2015.08.013.
- B. Walsh. Markov chain monte carlo and gibbs sampling, 2004. URL <http://alumni.soe.ucsc.edu/~mds/Gibbs.pdf>.

Acknowledgements

NKS conveys its gratitude to all organizations and persons who by means of financial support or contributions in kind have made the work presented in this report possible.

Disclaimer

The views expressed in this document remain the responsibility of the author(s) and do not necessarily reflect those of NKS. In particular, neither NKS nor any other organisation or body supporting NKS activities can be held responsible for the material presented in this report.

Title	Source CHAracterizatiOn accounting for meTeorological unCertainties (SOCHAOTIC) – final report
Author(s)	Jens Havskov Sørensen ¹ (co-ordinator) Henrik Feddersen ¹ Kasper Skjold Tølløse ¹ Andreas Uppstu ² Heiko Klein ³ Magnus Ulimoen ³ Lennart Robertson ⁴ Jan Pehrsson ⁵ Bent Lauritzen ⁶ Agnieszka Hac-Heimburg ⁷ Henrik Roed ⁷ Einar Améen ⁸ Naeem Ul Syed ⁸ Anna Maria Blixt Buhr ⁹ Jonas Lindgren ⁹ Tuomas Peltonen ¹⁰
Affiliation(s)	¹ Danish Meteorological Institute (DMI) ² Finnish Meteorological Institute (FMI) ³ Norwegian Meteorological Institute (MET Norway) ⁴ Swedish Meteorological and Hydrological Institute (SMHI) ⁵ PDC-ARGOS ⁶ Technical University of Denmark (DTU) ⁷ Danish Emergency Management Agency (DEMA) ⁸ Norwegian Radiation and Nuclear Safety Authority (DSA) ⁹ Swedish Radiation Safety Authority (SSM) ¹⁰ Radiation and Nuclear Safety Authority (STUK)
ISBN	978-87-7893-565-6
Date	March 2023
Project	NKS-B / SOCHAOTIC
No. of pages	110
No. of tables	3
No. of illustrations	90
No. of references	54

Abstract
max. 2000 characters

In recent years, events have occurred in which radionuclides were detected by filter stations in Europe without knowledge on the origin of those radionuclides. In such cases, there is a need to locate potential release sites. However, if the release site is actually known, or if a potential release site has been localized by inverse methods, then there is an additional need to estimate the release rates from this location as a function of time for the various radionuclides detected.

While in the SLIM NKS project, methodologies were developed to localize an unknown source of radionuclides dispersed in the atmosphere, the SOCHAOTIC project has developed methodologies, suited for operational use, by which characterization of the source, whose location is known, can be derived, i.e. to estimate the temporal release profiles of the radionuclides detected.

For operational use, nuclear decision-support systems should be extended with modules handling and analysing such monitoring data automatically, and conveying the data together with the geographical coordinates of the release point to the national meteorological centre accompanied by a request to estimate the temporal evolution of the release rates.

A number of case studies have been selected, viz. the ETEX-1 and the October 2017 case of Ru-106 in Europe as well as an artificial case. Methods for estimation of the temporal release profiles have been developed, and they have been applied to the selected cases by using the DERMA, MATCH, SILAM and SNAP atmospheric dispersion models. The methods have been applied both by using deterministic numerical weather prediction (NWP) model data and ensemble-statistical NWP model data derived by setting up and running the non-hydrostatic high-resolution Harmonie model for the selected cases. Finally, an interface to a nuclear decision-support system, ARGOS, is described.

Key words

nuclear emergency preparedness, atmospheric dispersion modelling, source characterization, inverse modelling, concentration measurements, gamma dose measurements, uncertainty