This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 662287.

EJP-CONCERT

European Joint Programme for the Integration of Radiation Protection Research

H2020 – 662287

D9.5 – Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modelling and ensemble calculations, and tools for ensemble calculation in emergency response

Lead Authors: Irène Korsakissok, Peter Bedwell, Erik Berge, Thomas Hamburger

With contributions from: WP1 members

Reviewer(s): CONCERT coordination team

Work package / Task	**WP9** **Task 9.1**	**ST 9.1.3**
Deliverable nature:	**Report**	
Dissemination level: (Confidentiality)	**Public**	
ontractual delivery date:	**30 October 2019**	
Actual delivery date:	**14 November 2019**	
Version:	**Final**	
Total number of pages:	**175**	
Keywords:	**CONFIDENCE, ensembles, dispersion, uncertainties, REM, Western Norway, probability maps, Fukushima, foodchain propagation, operational recommendations**	
Approved by the coordinator:	**M54**	
Submitted to EC by the	**M54**	

Introduction

Work Package 1 (WP1) of the CONFIDENCE project investigates the uncertainties linked to the pre and early release phase of an accidental release of radionuclides in the atmosphere. In such a situation, in the absence of environmental measurements of radioactivity, the assessment of the potential consequences and subsequent recommendations for the protection of the population (such as sheltering, evacuation or stable iodine intake) mostly rely on atmospheric dispersion calculations. In such an emergency context, uncertainties are very high and there is a dire need for speed so that actions to protect the population are undertaken as rapidly as possible. The aim of WP1 was to determine the main sources of uncertainties in these calculations, evaluate how they affect the endpoints used for decision making, and propose practical ways of taking them into account for emergency response in an operational context. This lead to three subtasks:

Task 1.1: analysing and ranking sources of uncertainties; this included the investigation of uncertainties related to meteorology, source term and model parameters. The outcome of this task carried out in the first year of the project was deliverable D1.1 (Mathieu et al. 2017);

Task 1.2: uncertainty propagation and analysis; this included the propagation of uncertainties through atmospheric dispersion models (up to atmospheric concentrations, deposition and doses, but without including food contamination) for hypothetical accident scenarios in Europe (the Radiological Ensemble Modelling-REM case study located in Borssele, and the Western Norway case), and for the Fukushima case.

Task 1.3: emergency response and dose assessment; the first subtask consisted in food chain uncertainty propagation, using one of the scenarios of task 1.2, and the second one consisted in issuing practical recommendations for the use of ensemble calculations in an emergency context.

The present deliverable D9.5 contains the results of tasks 1.2 and 1.3, concluding three years of work involving 13 participants from 8 countries. It is separated into five reports: the first three reports correspond to different case studies carried out in the framework of task 1.2, while the last two reports respectively contain the results of task 1.3.1 (food chain propagation) and task 1.3.2 (operational recommendations). These reports can be read independently, although those related to task 1.3 make use of ensemble calculations carried out in task 1.2. The deliverable D9.5 is therefore structured as follows:

- **D9.5.1** Ensemble calculation for hypothetical accident scenarios in Europe: the REM case studies (Korsakissok et al. 2019b);
- **D 9.5.2** Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the Western Norway case study (Berge et al. 2019);
- **D 9.5.3** Ensemble calculation for a past accident scenario: the Fukushima case study (Korsakissok et al. 2019a);
- **D9.5.4** Uncertainty propagation through a terrestrial food chain and dose model (Hamburger et al. 2019);
- **D 9.5.5** Guidelines for the use of ensembles in the description of uncertainty in atmospheric dispersion modelling: operational applications in the context of an emergency response (Bedwell et al. 2019).

References

Bedwell, P., I. Korsakissok, S. Leadbetter, R. Périllat, C. Rudas, J. Tomas and J. Wellings (2019). D 9.5.5 Guidelines for the use of ensembles in the description of uncertainty in atmospheric dispersion modelling: operational applications in the context of an emergency response. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Berge, E., H. Klein, M. Ulimoen, S. Andronopoulos, O. C. Lind, B. Salbu and N. Syed (2019). D 9.5.2 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the Western Norway case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Hamburger, T., F. Gering, I. Ievdin and S. Schantz (2019). Uncertainty propagation through a terrestrial food chain and dose model. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Korsakissok, I., R. Périllat, S. Andronopoulos, P. Astrup, P. Bedwell, E. Berge, A. Quérel, H. Klein, S. Leadbetter, O. Saunier, A. Sogachev, J. Tomas and M. Ulimoen (2019a). D 9.5.3 Ensemble calculation for a past accident scenario: the Fukushima case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Korsakissok, I., R. Périllat, G. Geertsema, H. De Vries, T. Hamburger, J. Tomas, S. Andronopoulos, P. Bedwell, T. Charnock, I. Ievdin, S. Leadbetter, T. Pazmandi, C. Rudas, R. Scheele, A. Sogachev, P. Szanto and J. Wellings (2019b). D 9.5.1 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the REM case studies. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Mathieu, A., I. Korsakissok, S. Leadbetter, S. Andronopoulos, P. Bedwell, J. Wellings, G. Geertsema, A. Jones, J. Tomas, H. De Vries, H. Klein, R. Périllat, K. Chevalier-Jabet, F. Stéphanie, S. Fougerolle, V. Créach, E. Cogez, T. Hamburger, T. Pazmandi, C. Rudas, A. Sogachev and P. Szanto (2017). D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project. Report D9.1 125. http://www.concert-h2020.eu/en/Publications

D9.5.1 – Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the REM case studies

Lead Authors: Hans De Vries, Gertie Geertsema, Irène Korsakissok, Susan Leadbetter, Raphaël Périllat, Rinus Scheele, Jasper Tomas

Based on calculations from: Spyros Andronopoulos, Poul Astrup, Peter Bedwell, Tom Charnock, Thomas Hamburger, Ievgen Ievdin, Susan Leadbetter, Tamás Pázmándi, Raphaël Périllat, Csilla Rudas, Andrey Sogachev, Péter Szántó, Joseph Wellings

Review(s): Peter Bedwell, Susan Leadbetter, Spyros Andronopoulos, Jasper Tomas, Csilla Rudas, Gertie Geertsema

Table of Contents

Introduction 7

Description of the REM case studies scenarios 7

Meteorology 7

Meteorological scenario 1: “easy case” 8

Meteorological scenario 2: “warm front, higher variability case” 10

Release scenarios and associated uncertainties 13

Atmospheric dispersion simulations setup and endpoints 18

Modelling set-up 18

Synthesis of the case studies 20

Endpoints 21

Ensemble results for REM1 – short release 22

Median of the ensembles 22

Maximum distance of threshold exceedance 23

Level of agreement for threshold exceedance 26

Ensemble results for REM2 – short release 30

Ensemble results for REM – long release 34

Conclusions 36

References 37

Introduction

One of the aims of the CONFIDENCE project is to understand, reduce and cope with the uncertainty of meteorological and radiological data and their further propagation in decision support systems, including atmospheric dispersion, dose estimation, food-chain modelling and countermeasure simulation models. Work package 1 (WP1) is focused on the modelling of uncertainties during the emergency phase, from meteorological and source term inputs, and applied to atmospheric dispersion and dose estimates. This report presents ensemble dispersion simulations performed by WP1 participants for hypothetical accident scenarios at the nuclear power plant (NPP) in Borssele (The Netherlands). These case studies are called the REM (Radiological Ensemble Modelling) scenarios.

The first part of the report summarizes briefly the two release scenarios (one short and one long release) and the two meteorological scenarios (one with small variability, one with larger variability) considered. Then, the following parts present the results for the different case studies.

Description of the REM case studies scenarios

In Figure 1 the location of the Borssele nuclear power plant (NPP) is shown. It is located at a latitude and longitude of 51.43 and 3.71 decimal degrees, respectively.

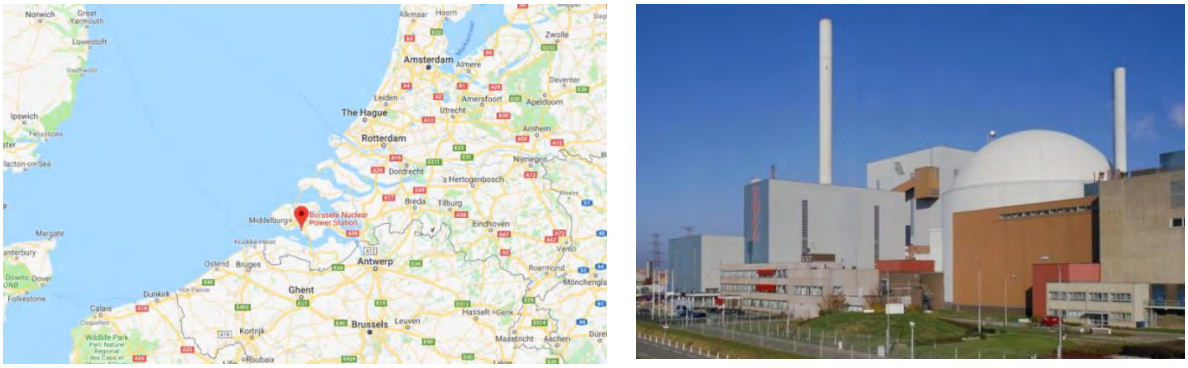


Figure 1: The Borssele nuclear power plant. Latitude, longitude: 51.43°, 3.71°

Meteorology

The meteorological data was provided by KNMI for the REM case study. The Harmonie-AROME model was used (Bengtsson et al. 2017). KNMI runs an operational and a semi-operational suite with different model versions using a Lambert Conformal coordinate system, with a horizontal resolution of 2.5 km. and a temporal resolution of one hour. The time-span of the data is 72 hours. The domain provided contained 300 x 300 points in the horizontal directions covering the area indicated by the black dots in Figure 2. The data contained 65 levels in the vertical direction up to a pressure level of 10 hPa corresponding to a height of approximately 31 km.

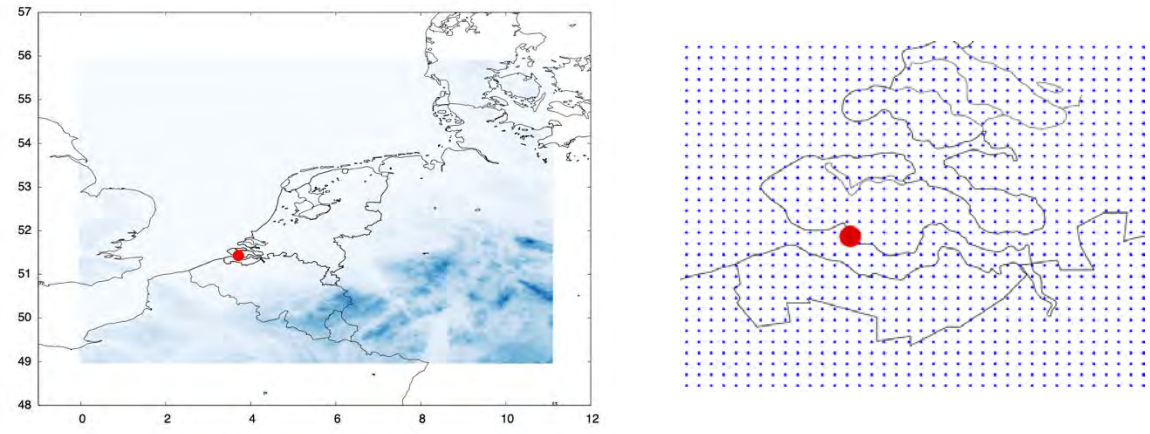


Figure 2: Meteorological domain for the REM case study. Blue dots show the horizontal resolution. Red dot indicates the location of the Borssele NPP.

KNMI constructed a Harmonie-AROME ensemble from 2 different versions of the meteorological model, with different turbulent schemes, and combined successive deterministic forecasts to create a hybrid lagged ensemble (Geertsema et al. 2019). The ensemble is hybrid in the sense that two different model versions are used; and lagged in the sense that successive forecasts are used. The reason to construct an ensemble in this way is that the KNMI Harmonie-AROME archive can be used to construct an ensemble where the spread of the resulting ensemble can represent a realistic ensemble spread and be used as a pilot for high resolution ensembles which start to become available. Each model version was used to construct 5 ensemble members with a forecast length of the required forecast length (Figure 3).

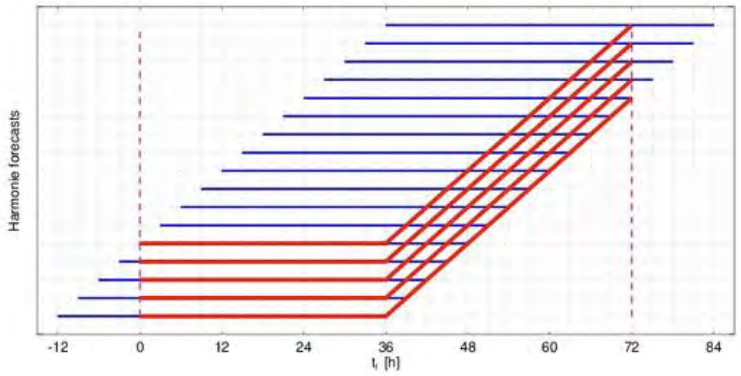


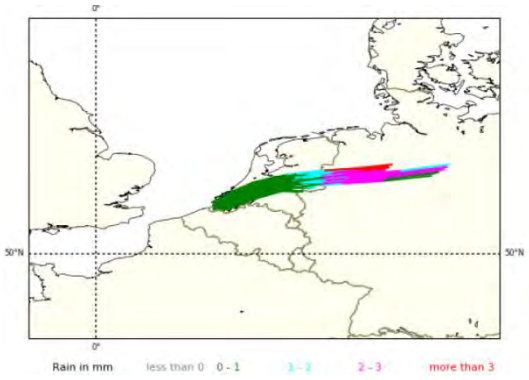
Figure 3 : Construction of a 5-member ensemble with a length of 72h. The blue lines indicate the original Harmonie-AROME 48h forecasts. The red lines indicate how each of the members is constructed.

From the start of the constructed ensemble (t_f=0), we use the forecast that starts there, and also the 4 forecasts that each started 3 hours earlier, for the first 36 hours for which they all overlap. From there, we use forecasts from successive runs to a maximum forecast of 72 hours for scenario 1 and 60 hours for scenario 2 (see below). Eventually, for all of the (hourly) steps, the date/time/forecast step values are changed to give 5 members that range from t_f=0–72 (resp. 60) for the same start date/time.

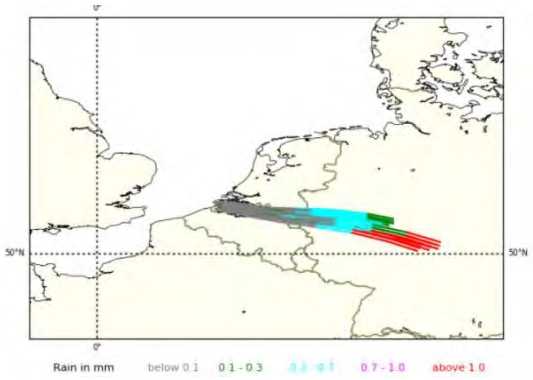
Meteorological scenario 1: "easy case"

The first scenario considered applies to a release on 11 Jan 2017. It was labelled "easy case" or "REM1", in the sense that the wind direction is well established (Figure 4). It is an interesting case in the sense that there is rain, which adds uncertainty to the scenario (depending on the release time, the plume may or may not be scavenged by rain). Trajectories starting at Borssele at different time steps provide an indication of the plume direction (Figure 4). The trajectory calculations are based on analysed weather only.

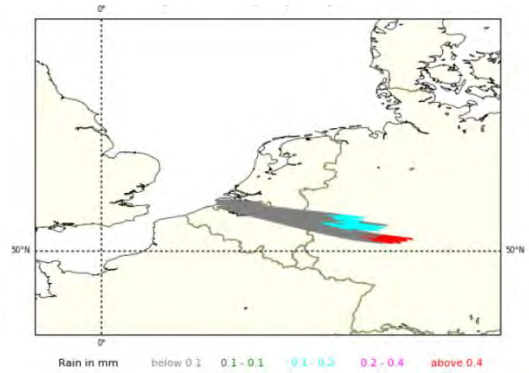
Therefore the trajectory results provided here are only meant as indicative information. The plots show the trajectories for locations at and near Borssele, starting at different heights ranging from 10m to 500m above surface (stack height is 60 meters). The calculation length for the trajectory is 6 hours.



(a) Release at 06:00 UTC



(b) Release at 12:00 UTC



(c) Release at 18:00 UTC

Figure 4 : Indicative 6-hours plume trajectories based on analysed weather as a function of height (between 10m and 500m), for a release on 11/01/17 at 06:00, 12:00 and 18:00 UTC, and associated rain cumulated on one hour (colours show the amount of rain. Trajectories plotted by KNMI.

Figure 4 indicates that the wind direction is and clearly shifting to more northerly through the course of the day, but with little uncertainty when considering a short release at 12:00 UTC. However, these are trajectories based on the analysed weather only, and it is interesting to have a look on the trajectories given by the ensemble, in order to infer the spread featured by the different members. For further analysis, trajectories were plotted for the 10 ensemble members, for a release of particles between ground level and 100 meters above ground level, and at 12:00UTC (Figure 5). This tends to confirm that for this case, there is little uncertainty in the wind direction. However, on full ensemble dispersion simulations, the effect of rain on the plume deposition may lead to a little more variability between the ensemble members.

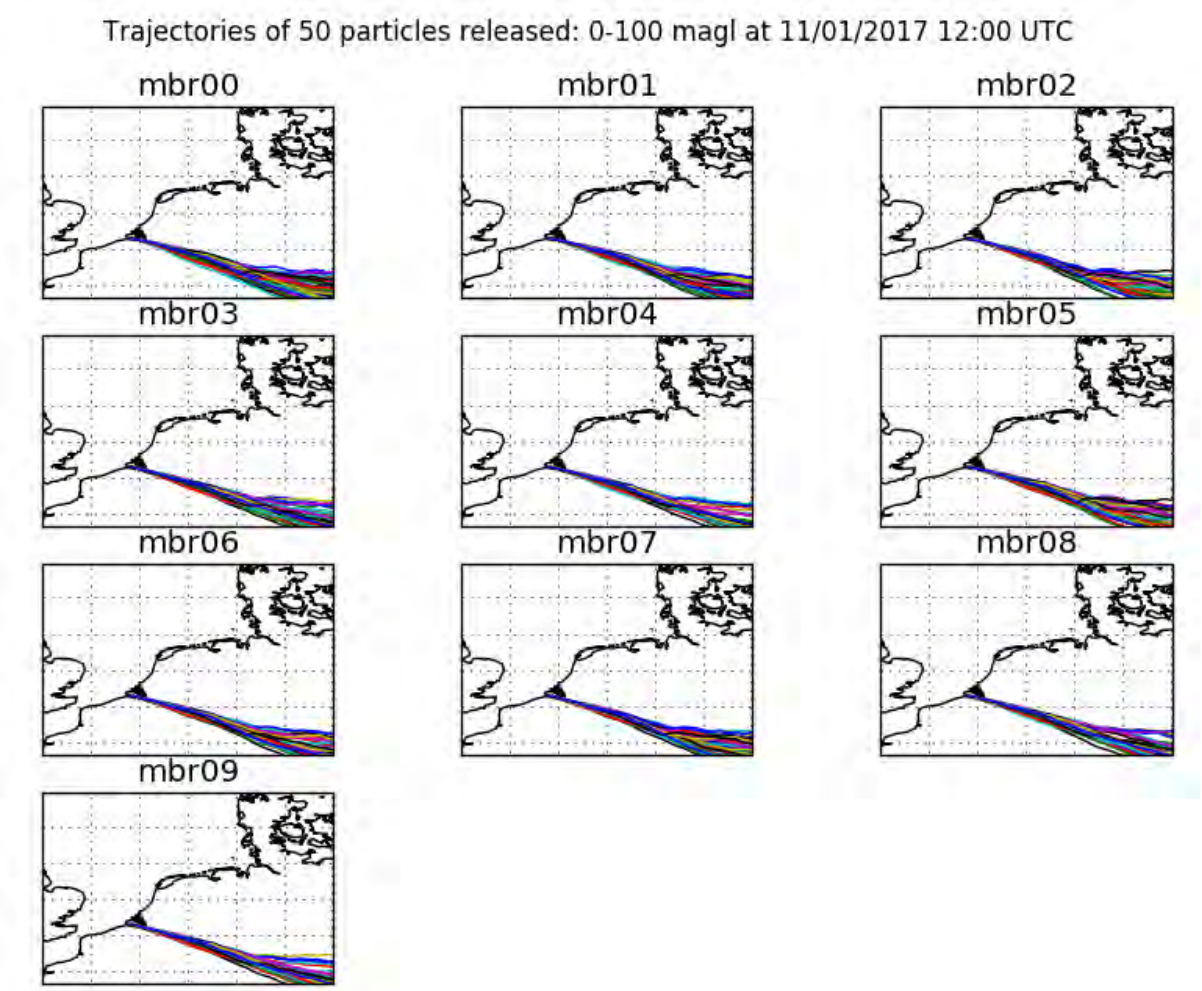
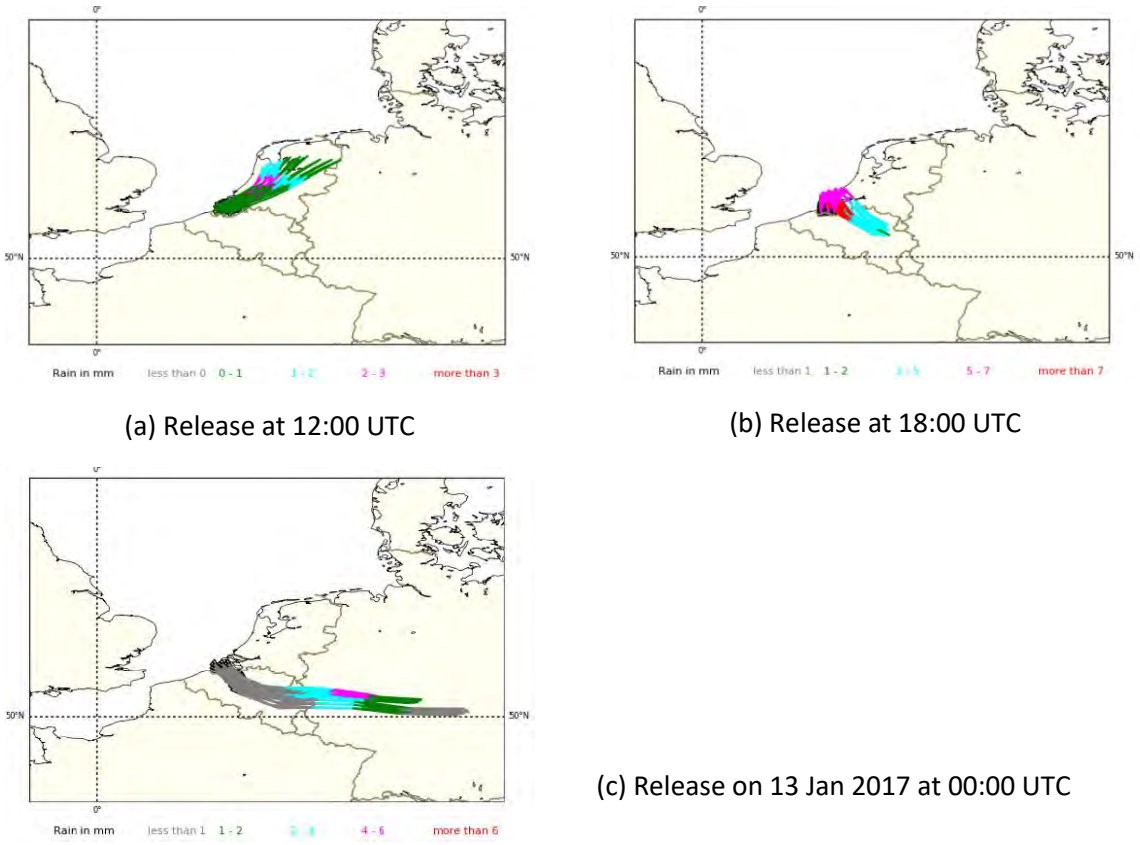


Figure 5 : trajectories for the 10 ensemble members for a release on 11 January 2017 at 12:00 UTC. Trajectories plotted by MetOffice.

Meteorological scenario 2: "warm front, higher variability case"

The second scenario considered is a weather situation with a warm front passage resulting in precipitation, wind direction and wind speed changes. It occurs on the day following scenario 1, that is, for a release on January 12, 2017.The trajectories are shown in Figure 6. The length of the trajectories is 6 hours. The colour codes denote the amount of precipitation. The trajectories start at different heights between 10 and 500 metres above surface. The trajectories start at the Borssele site and 8 locations close to the Borssele site, thus mimicking a (small) timing error in the model. The trajectories starting at 12 UTC move slowly to the North East, while the moving warm front catches up with the plume resulting in significant amount of precipitation (pink and red colours close to the release site). Trajectories starting at 18 UTC move faster to the East with precipitation close to the source. It was labelled "warm front – higher variability case" or "REM2".

The wind information from KNMI observations sites in the vicinity of the Borssele site is analysed. Three observation sites are within approximately 40 kilometres of the plant (stations 308, 310 and 340 in Figure 7). Meteorological information from three other sites located North-North East of the plant are also analysed. These 6 stations show the passage of the warm front over the area. The wind observations shown in Figure 7 clearly feature the warm front passage between 18:00 and 21:00 UTC on January 12, resulting in a change of wind speed and wind direction.



(a) Release at 12:00 UTC

(b) Release at 18:00 UTC

(c) Release on 13 Jan 2017 at 00:00 UTC

Figure 6 : Indicative 6-hours plume trajectories based on analysed weather as a function of height (between 10m and 500m), for a release on 12/01/17 at 12:00, 18:00 UTC, and 13/01/17 at 00:00 UTC and rain (cumulated in one hour). Trajectories plotted by KNMI.

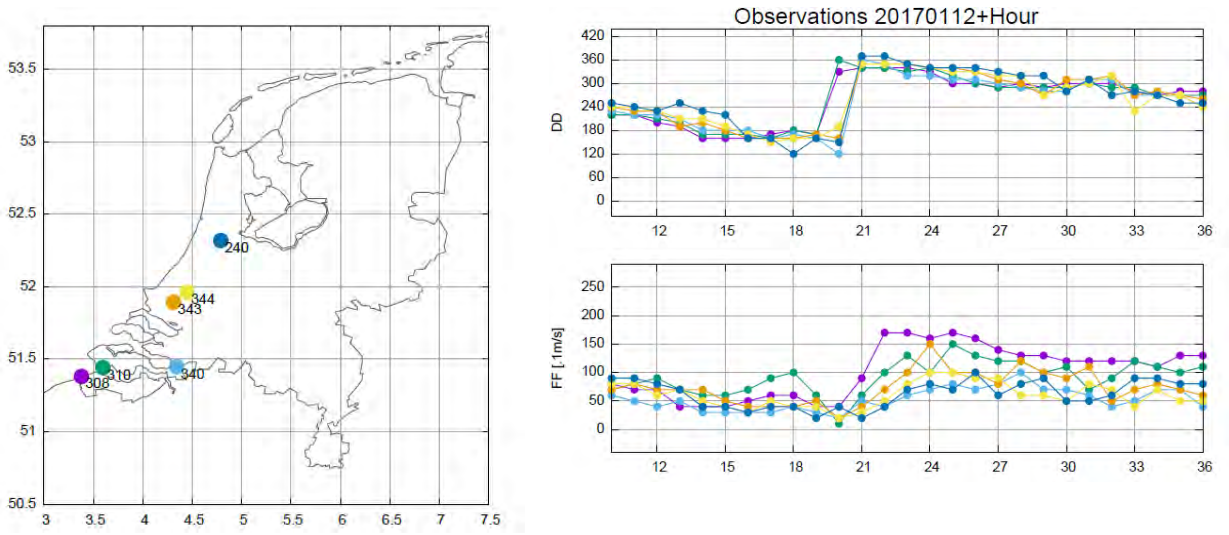


Figure 7 : Meteorological observations: position of the stations (left), wind direction (upper right) and wind speed at 10 m (lower right) averaged over previous 10 minutes, between T0=20170112 and T0+36 hours . The wind direction is given in meteorological convention (360=North, 90=East, 180=South 270=West). The colour code of the stations corresponds to the colours used on the right panel. The observations are from the KNMI archive.

As for the “easy case”, it is interesting to investigate whether the constructed ensemble features some variability in variables such as wind direction, wind speed, and rain, and how that variability translates into different plume trajectories. To answer the first question, meteograms were plotted (Figure 8). The ensemble spread is not very large during the first 36 hours (corresponding to the “easy case” described

earlier), but the variability increases after the passage of the warm front. This is particularly true for the 250-m wind speed and direction. There is also some variability in the rain.

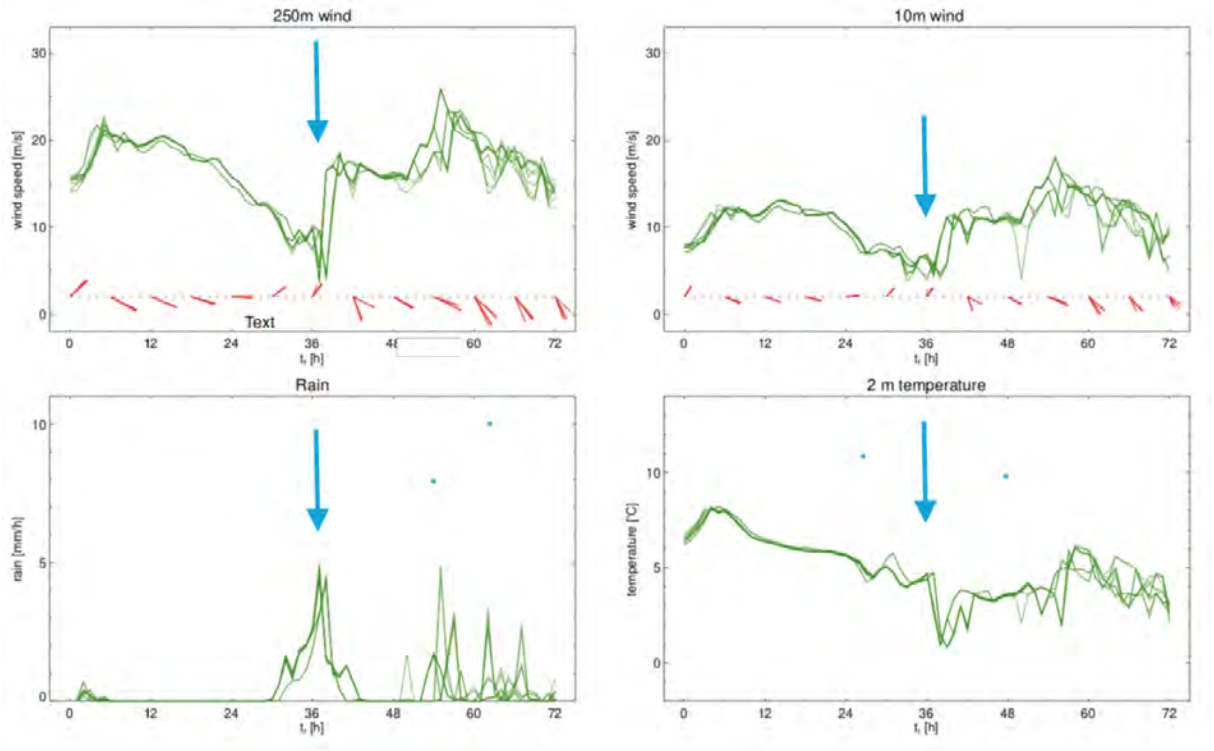


Figure 8 Harmonie ensemble meteogram. Analysis time is 2017011106. The blue arrows indicate the warm front passage, which is at 20170112 at 18:00 UTC (36 hours forecast). The green lines are the ensemble members, for 250-m and 10-m wind speed (upper left and right respectively), rain (lower left) and 2-m temperature (lower right). The red lines feature the wind direction (upper panel) at 250-m (left) and 10-m height (right). Meteograms from KNMI.

Finally, Figure 9 features the trajectories for the 10 ensemble members, for a release on 12 January 2017 at 18:00UTC. Here, there is more variability in plume trajectories than in Figure 5: some ensemble members feature trajectories toward North-North-East (e.g. members 02, 07) while others have completely shifted toward East-South-East (e.g. members 04 and 09). Some members show a turning wind during this period, with a few trajectories going in the Northern direction and most oriented toward the Southern direction (members 00, 01, 06). One may expect that the inclusion of rain in the full dispersion calculations may add further variability to the results.

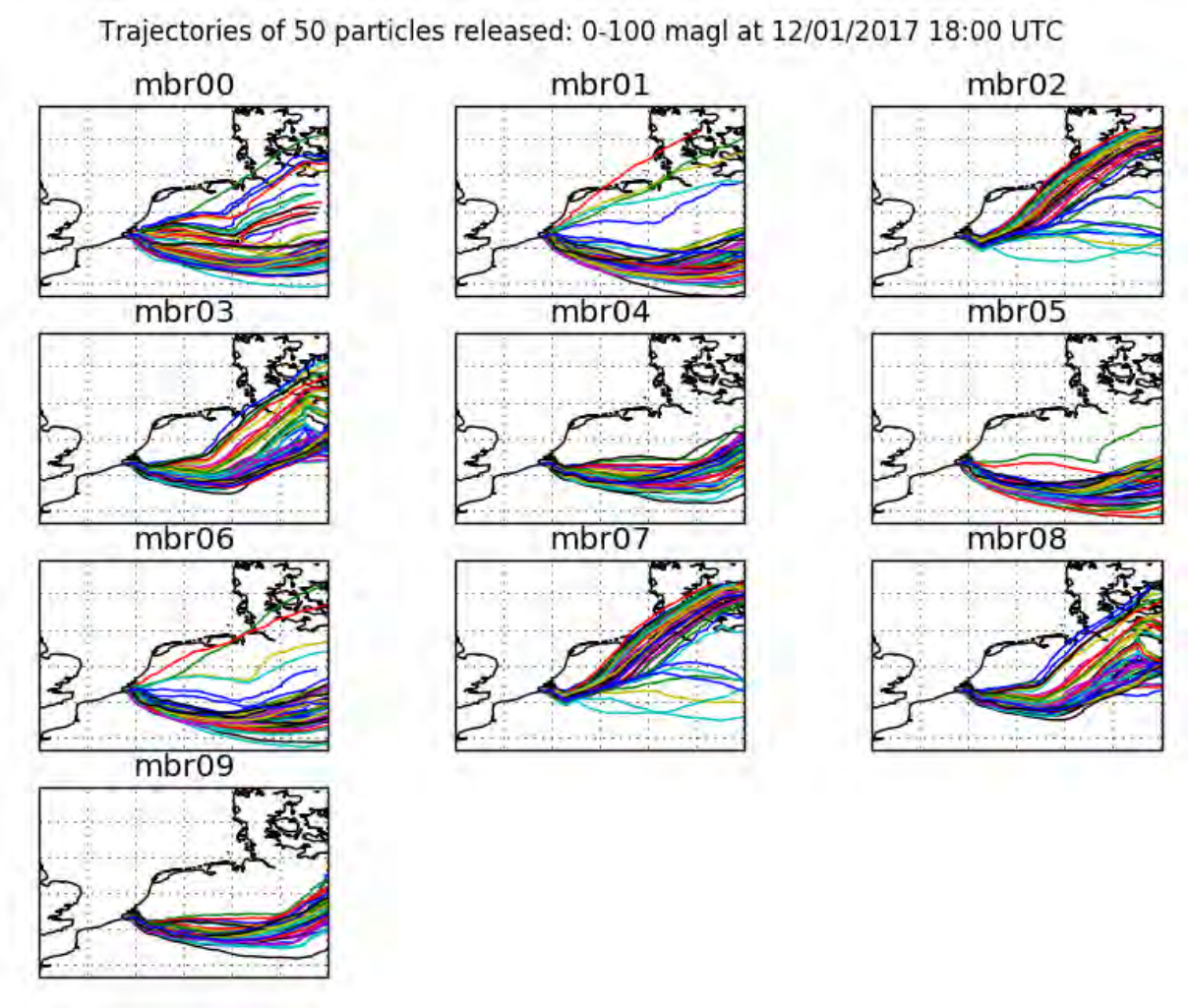


Figure 9 : trajectories for the 10 ensemble members for a release on 12 January 2017 at18:00 UTC. Trajectories plotted by MetOffice.

Release scenarios and associated uncertainties

Figure 10 schematically shows the timeline and the names of the events during an incident at a nuclear reactor involving a release of radioactivity. The horizontal arrows indicate the duration of the events and their dependencies. The variable positions of the events represent the uncertainties in their timing. Clearly, these uncertainties can complicate the assessment of the situation. As time passes, uncertainties may decrease due to additional information about the status of the reactor or through measurements.

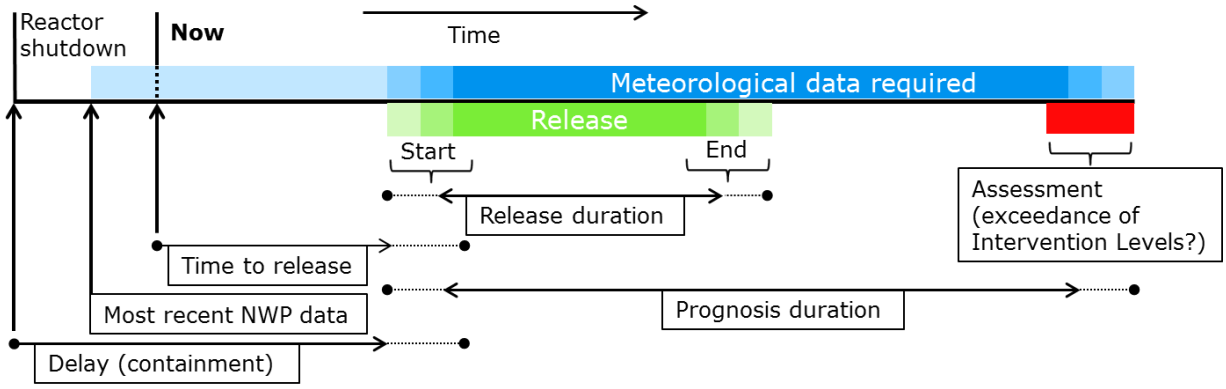


Figure 10: Illustrative timeline with the events during an incident at a nuclear reactor involving a release of radioactivity including uncertainties.

The pre-release and release phase are studied by considering two hypothetical release scenarios:

1. A release with duration of 4 hours; this is based on an accident which is anticipated to start in approximately 24 hours;
2. A release with duration of three days based on a 'Loss of Coolant Accident' (LOCA) scenario from the FASTNET project. There is no uncertainty on the timing of the start of the release. This scenario is representative of the uncertainties during the release phase (see Fig. 10). The construction of this ensemble of source terms was detailed in CONFIDENCE deliverable D9.1 (Mathieu et al. 2018).

Table 1: uncertainties considered in the two release scenarios: short release and long release

Short release	Long release
Time to release: 24 h +/- 6 h, equally distributed **Release duration**: 4 hours **Effective release height:** 50 m +/- 50 m, equally distributed **Released activity** : Factor 1/3 to 3 of Table 2	**Time to release**: 0 hours, no uncertainties **Release duration**: 3 days **Effective release height**: 50 m +/- 50 m, equally distributed **Released activity**: given by the spread of the ensemble of source terms

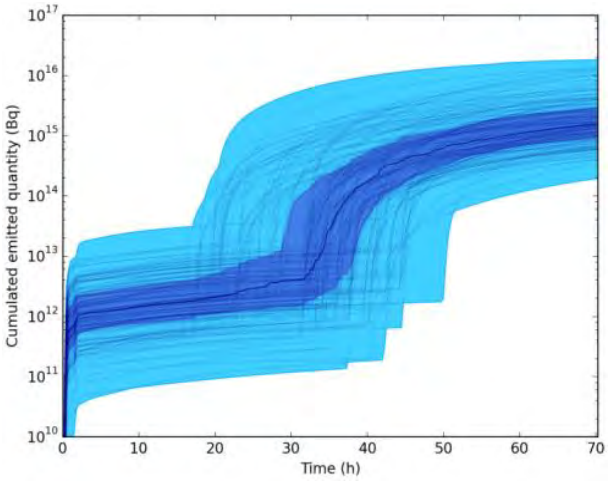
In Table 1 the source terms and associated uncertainties are summarized; the timing, release height and the released activity are shown for the two scenarios. The rationale behind these two source terms is that the "short release" is a very preliminary estimate, when little information is available in the pre-release phase, and the uncertainty on the release time is large. The long release is representative of uncertainties that stem from models used to represent reactors physics during severe accidents, when the accident scenario is known and there is no uncertainty on the release time (during the release phase for instance). These two source terms therefore represent different situations in emergency response. For example, the approximate time to release of 24 hours in the short release scenario might allow for direct countermeasures in the pre-release phase, such as evacuation and/or iodine prophylaxis, even though source term uncertainties are large. In the long release scenario, on the other hand, the release start now and depending on the onset of the scenario direct countermeasures might have been initiated already.

Details on the release scenarios

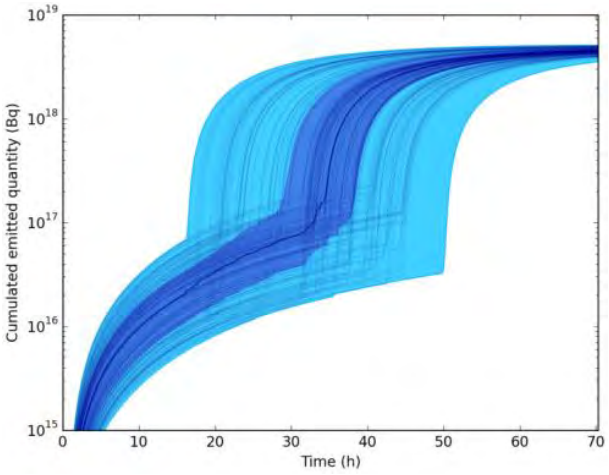
The nuclide composition for the short release is based on a constructed source term derived through interpolation of results from a safety study of the Borssele NPP. This source term is scaled to a hypothetical reactor with an electrical power output of 900 MW (to be consistent with the reactor considered in the long release scenario). For the short release the nuclide composition is based on a *Delay* time of 24 h. In addition, the uncertainty in the released activity is assumed to be within a factor of 1/3 to 3 of the reference values in Table 2. For the long release the uncertainty in the released activity was derived from the distribution of the source terms in the FASTNET project (Chevalier-Jabet 2019a). The median and maximum values are shown in the table for the same radionuclides as for the short release. However, the full source terms included over 200 radionuclides in total. The FASTNET ensemble of source terms was constructed using the severe accident code ASTEC, and several assumptions as inputs for the LOCA scenario. For instance, a 3-inches break was assumed, with an uncertainty of +/- 1 inch on the break size. In the same way, other uncertainties were taken into account, including the release rate (factor 10), leak rate (factor 2), iodine behaviour and concrete-corium interaction (10% error). These uncertainties were detailed in Mathieu et al. (2018) and Chevalier-Jabet (2019b). It resulted in an ensemble of 150 source terms for a three-day release. The particularity of the release kinetics is that, after the initial release start, there is a second, more important release, corresponding to the opening of the filtered venting containment system. For this second release, aerosols are supposed to be filtered but there is an important release of gaseous materials, in particular Iodine and Xenon, which explains why the released amounts for these radionuclides is much larger than, for instance, for Caesium (Table 2).

Table 2: The nuclide composition of the hypothetical source terms for the REM case study. The total released activity is shown for all nuclides in the source term. Only the eight nuclides in bold font were considered in the model runs.

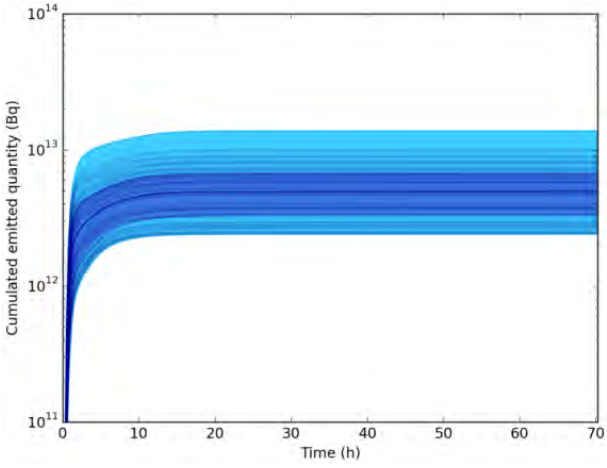
	Short release	**Long release**
	Borssele NPP scaled to 900 MWe: Total released activity [Bq]	*FASTNET 3 inches break:* Total released activity [Bq] median (max)
	Particle size: 1 µm. Iodine group: 1/3 particulate, 2/3 elemental	Iodine form: given by the ensemble of source terms
Nuclide		
Kr-85m	8.76E+15	6.65E+15 (2.89E+16)
Kr-85	1.53E+16	2.08E+16 (2.12E+16)
Kr-88	2.38E+15	2.36E+15 (1.62E+16)
Xe-133	**3.51E+18**	**4.97E+18 (5.25E+18)**
Xe-135	7.46E+17	8.99E+17 (1.64E+18)
I-131	**2.25E+16**	**1.08E+15 (1.972E+16)**
I-132	**2.84E+16**	**1.23E+15 (2.21E+16)**
I-133	2.15E+16	6.62E+14 (1.21E+16)
I-135	3.04E+15	6.47E+13 (1.19E+15)
Rb-88	1.97E+13	2.83E+13 (7.32E+13)
Sr-89	2.36E+15	3.46E+12 (2.89E+13)
Sr-90	2.19E+14	2.36E+11 (1.97E+12)
Y-90	1.78E+13	5.91E+09 (1.60E+11)
Zr-95	4.19E+14	1.33E+10 (4.40E+12)
Ru-103	3.80E+15	1.79E+13 (1.31E+14)
Ru-106	1.24E+15	6.71E+12 (4.93E+13)
Rh-106	1.24E+15	6.67E+14 (1.33E+18)
Te-131m	1.02E+15	8.51E+12 (3.05E+13)
Te-132	**1.37E+16**	**6.58E+13 (2.41E+14)**
Cs-134	**2.69E+15**	**1.30E+13 (6.02E+13)**
Cs-136	**6.37E+14**	**4.90E+12 (2.22E+13)**
Cs-137	**2.06E+15**	**8.80E+12 (4.06E+13)**
Ba-137m	**2.78E+14**	**1.37E+13 (8.45E+13)**
Ba-140	4.08E+15	8.46E+13 (8.30E+14)
La-140	4.47E+14	7.67E+12 (7.04E+13)
Pu-238	2.60E+11	1.65E+08 (4.97E+10)
Pu-241	3.19E+13	5.13E+09 (1.55E+12)
Cm-242	9.02E+12	1.82E+09 (5.98E+11)
Cm-244	1.02E+11	1.43E+08 (4.68E+10)



(a) ^{133}Xe



(b) ^{131}I



(c) ^{137}Cs

Figure 11 : release kinetics for the long release ensemble from the FASTNET project, for three radionuclides.

It should be noted that several simplifications have been made in this study for the sake of computational time. First, for the short release as well as the long release scenarios, only the eight radionuclides highlighted in bold in Table 2 have been used. This selection comes from previous Fukushima studies. Those were radionuclides considered having a significant contribution to the dose (Saunier et al. 2013). However, several significant contributors have been overlooked, especially Tellurium and, as far as agricultural contamination is concerned, ^{90}Sr for instance.

The second approximation was made for the long release scenario. The full ensemble of 150 source terms was not used. Two subsets of 5 and 10 source terms have been defined, in order to keep some properties of the ensemble (mean and standard deviation). The method used is detailed in (Bedwell et al. 2019). Most participants have used the 10 source terms for their long release calculations, except DTU and MTA EK who used the 5 source terms.

The impact of these two assumptions (reduction of the number of radionuclides and reduction of the number of ensemble members) was assessed by comparison to “full” calculations. This is presented in Bedwell et al. (2019).

Atmospheric dispersion simulations setup and endpoints

Table 3 summarizes the two different meteorological scenarios and two release scenarios that have been described. The short release scenario was used with both meteorological scenarios REM1 and REM2. Therefore, the corresponding case studies will be labelled hereafter as REM1 (or REM2)-short release or REM1-S (REM2-S) as summarized in Table 3. The long release scenario was used with the meteorological data provided for REM1 study; however, as the release lasts three days, it covers both REM1 and REM2 meteorological situations. The case study will therefore be named REM-long release or, to be concise, REM-L.

Table 3 : Synthesis of the meteorological and release scenarios, and corresponding short names of the case studies

Case studies	Short release	Long release (FASTNET release)
Meteorological scenario 1 (easy case) - REM1	REM1-S	REM-L
Meteorological scenario 2 (warm front) - REM2	REM2-S	

Modelling set-up

Table 4 summarizes the participants who ran simulations for the REM case study and the endpoints that were computed by the participants. It also presents the type of atmospheric dispersion model used: one Eulerian model, one Lagrangian particle model, one Lagrangian puff model and four Gaussian puff models.

Table 4 : Summary of participants, output variables computed, and type of atmospheric dispersion model. The green cells indicate the outputs provided by participants. The white cells correspond to outputs reconstructed by IRSN using activity, deposition, and IRSN's assumptions for dose calculation.

		Activity	Deposition	Dose rate	Dose	Type of model
France	**IRSN**	✓	✓	✓	✓	IdX – Eulerian
The Netherlands	**RIVM**	✓	✓	✓	✓	NPK-puff – Gaussian puff
Germany	**BfS**	✓	✓	✓	✓	RIMPUFF – Gaussian puff
UK	**Met Office/PHE**	✓	✓		✓	NAME – Lagrangian particle
Greece	**EEAE**	✓	✓	✓		DIPCOT – Lagrangian puff
Denmark	**DTU**	✓	✓			RIMPUFF – Gaussian puff
Hungary	**MTA EK**	✓	✓	✓	✓	SINAC – Gaussian puff

Apart from the type of model that differs from one participant to another, discrepancies in the outputs may stem from various physical parameterizations, notably those related to dry and wet deposition (Bedwell et al. 2018). Considering the importance of precipitation in the two meteorological scenarios considered, it is worth studying the difference in wet deposition schemes between the participants (Table 5). For most models, scavenging coefficients are of the form $\Lambda = aq^b$ where q is the rain intensity (mm.h^{-1}) and Λ is in s^{-1}. Factors a and b are given by the user and may depend on the height, with a distinction mixing layer / reservoir layer (RIVM), in-cloud or below cloud scavenging

(which is diagnosed through computing the cloud basis height) (IRSN and MetOffice), and/or on isotope families (DTU). For particles, the models used by EEAE and BfS compute the scavenging coefficient as function of the particle size and the rain intensity according to Baklanov and Sørensen (2001). The washout coefficient $\Lambda(s^{-1})$ is expressed as a polynomial function of particle radius *r* (µm) and rain-rate *q* (mm/h) as follows:

$$\begin{array}{cc} \Lambda(r,q) = a_0 q^{0.79} & r < 1.4\mu m \\ \Lambda(r,q) = (b_0 + b_1 r + b_2 r^2 + b_3 r^3) f(q) & 1.4\mu m < r < 10\mu m \\ \Lambda(r,q) = f(q) & r > 10\mu m \end{array}$$

with

$$f(q) = a_1 q + a_2 q^2$$

and

$$a_0 = 8.4 \cdot 10^{-5}, a_1 = 2.7 \cdot 10^{-4}, a_2 = -3.618 \cdot 10^{-6}$$

$$b_0 = -0.1483, b_1 = 0.3220133, b_2 = -3.0062 \cdot 10^{-2}, b_3 = 9.34458 \cdot 10^{-4}$$

For the calculations presented here, a particle size of 1µm was assumed. Therefore, the values presented in Table 5 for EEAE and BfS are given for this particle size. RIVM and IRSN have different values for in-cloud and below-cloud scavenging, but used a uniform value over the boundary layer height in this case. However, in practice, IRSN generated their ensembles using Monte Carlo perturbations for the REM1 and REM2 cases. Therefore, the reference value indicated in Table 5 was only used for the REM1-L case where simulations were compared for 10 meteorological members without additional perturbation (see REM1-S results below). For the "full" perturbations of REM1, see Table 7.

Table 5 : Scavenging coefficients used by the different participants

	Type of model	Scavenging scheme aq^b - Coefficients values	
		a ($h.mm^{-1}.s^{-1}$)	b
IRSN	ldX – Eulerian	Below cloud: $a=5\times10^{-5*}$ In cloud: $a=1\times10^{-4}$	b=1
RIVM	NPK-puff – Gaussian puff	Below cloud: $a=7.0\times10^{-5}$ In cloud: $a=7.0\times10^{-5*}$	For q < 1 mm/h, b=1, else b=0.8
BfS	RIMPUFF – Gaussian puff	Gases: $a=8\times10^{-5}$ Aerosols 1µm: $a=8.4\times10^{-5}$	Gases: b=0.6 Aerosols 1µm: b=0.79
MetOffice/PHE	NAME – Lagrangian particle	Below cloud: $a=8.4\times10^{-5}$ In cloud: $a=3.36\times10^{-4}$	b=0.79
EEAE	DIPCOT – Lagrangian puff	Gases: $a=8\times10^{-5}$ Aerosols 1µm: $a=8.4\times10^{-5}$	Gases: b=0.6 Aerosols 1µm: b=0.79
DTU	RIMPUFF – Gaussian puff	*Depends on the isotope family*	
MTA EK	SINAC – Gaussian puff	$a=1\times10^{-4}$	b=0.8

* Used in this study for all layers (no distinction between below-cloud and in-cloud)

The variability featured in Table 5 is consistent with the range of variation of an order of magnitude pointed out in Bedwell et al. (2018). However, in this study, when considering a rain rate of 1 mm/h and assuming the main process is below-cloud scavenging (which is true close to the source), the range of variation is

much lower: the scavenging coefficient Λ is within $[5\times10^{-5};8.4\times10^{-5}]s^{-1}$ with the exception of MTA-EK (1×10^{-4}). Farther from the source, when in-cloud scavenging may be included, Met Office may have higher deposition than the others with $\Lambda = 3.36\times10^{-4}s^{-1}$ for in-cloud scavenging (again, for a rain of 1 mm/h). Overall, it may be supposed that, all other things being equal, the highest deposition values should probably be obtained by Met Office and/or MTA EK.

Synthesis of the case studies

For the two short release scenarios (REM1-S and REM2-S) perturbations were applied to the "reference" release, to account for the uncertainties described in Table 1. The exact way these uncertainties should be taken into account was left to be decided by the participants, depending on their practical possibilities (especially in terms of computational resources). In practice, it resulted in between 10 and 650 simulations per release scenario, as detailed in Table 6.

Table 6 : source perturbations applied for the short release scenario, in the REM1 and REM2 case studies.

Participant	Number of simulations per release scenario	Source perturbations		
		Release height	Release time	Released quantity
IRSN	100 (Monte Carlo)	[0, 100m] uniform	[-6h, 6h] uniform	[1/3, 3] uniform
BfS	150	[0m, 50m, 100m]	T0 + [-6h, -3h, 0h, +3h, +6h]	[x1/3, x1, x3]*
Met Office/ PHE	90	[50m]	T0 + [-6h, 0h, +6h]	[x1/3, x1, x3]
EEAE	50	[50m]	T0 + [-6h, -3h, 0h, +3h, +6h]	[x1/3, x1, x3]*
MTA EK	150 (REM 1-S) 90 (REM 2-S)	[0m, 50m, 100m]	T0 + [-6h, -3h, 0h, +3h, +6h] T0 + [-6h, 0h, +6h]	[x1/3, x1, x3]*
RIVM	650	[0m, 25m, 50m, 75m, 100m]	[-6h, +6h] with a time step of 1 hour (13 steps)	[x1/3, x1, x3]*
DTU	10 (REM 1-S) 50 (REM 2-S)	[50m]	T0 + [-6h, -3h, 0h, +3h, +6h] (REM2-S)	[x1/3, x1, x3]*

*Perturbation applied a posteriori on the results

IRSN decided to apply perturbations to the model's physical parameters, namely vertical diffusion coefficients, dry deposition velocities and scavenging coefficients, to account for model uncertainties

described in (Bedwell et al. 2018). The ranges of variation for these parameters are given in Table 7. As it would not be possible to run "cross-simulations", that is, taking into account all combinations of parameters, IRSN used a Monte Carlo method, which consists of randomly drawing the uncertain parameters' values within the given range of variation; a set of N combinations is drawn, N being chosen by the user to ensure that there is sufficient convergence in the results (here, N=100). The Monte Carlo method, convergence and limitations are discussed in Bedwell et al. (2019). Besides, Périllat et al. (2017) already compared the use of "cross-simulations", using a meteorological ensemble and a set of source terms on the Fukushima case, and Monte Carlo method including the variation of physical parameters in the model. It concluded that meteorology and source terms were the main sources of uncertainties, and results were not significantly different using the "full" Monte Carlo ensembles in terms of ensemble's ability to encompass the observations. The main advantage of using Monte Carlo is reducing the number of simulations, while allowing to take into account more sources of uncertainties.

Table 7 : Perturbations applied by IRSN for the short release scenario, in the REM1 and REM2 case studies. The (x) denote a multiplicative coefficient applied to the reference value; the (+) is an additive perturbation applied to the reference value; no sign indicates that the parameter's range of variation is directly given by the table.

	Variable	Lower bound	Upper bound
Source term	Source term amplitude (×)	0.333	3
	Source term time shift (+)	- 6 H	+ 6 H
	Source term height	0 m	100 m
Model (physical parameters)	Dry deposition (iodine)	2.0×10^{-3} m/s	2.0×10^{-2} m/s
	Dry deposition (others)	5.0×10^{-4} m/s	5.0×10^{-3} m/s
	Scavenging coefficient a (in cloud)	1.0×10^{-5} h.mm^{-1}.s^{-1}	1.0×10^{-2} h.mm^{-1}.s^{-1}
	Scavenging coefficient a (below cloud)	1.0×10^{-5} h.mm^{-1}.s^{-1}	1.0×10^{-4} h.mm^{-1}.s^{-1}
	Scavenging coefficient b	0.5	1.0
	Vertical diffusion matrix Kz (×)	0.333	3

Endpoints

Some participants provided their own dose calculations (see Table 4). For those who provided only atmospheric dispersion results (air and deposition concentrations for all radionuclides), the dose calculations were made by IRSN, based on the air concentrations and deposition data provided. The dose calculations were derived for 1-year-old children, with no sheltering nor any other forms of protective action. The effective dose calculation includes the following pathways:

- External dose due to irradiation by radionuclides in the atmosphere (plume-shine),
- External dose due to the irradiation by radionuclides deposited on the ground (ground-shine),
- Internal dose due to plume inhalation.

It does not take into account dose resulting from food intake.

The outputs proposed here are maps of probability of threshold exceedance[1]. Instead of a single contour showing the impacted area (based on a single deterministic simulation), the probability maps are based on an ensemble of simulations and correspond to the probability that a given zone is contaminated above a given level. The reference levels chosen for this project are:

- 37 kBq/m^2 of Cs-137 deposition (Chernobyl reference level);
- 50 mSv inhalation thyroid dose for 1-year old child (IAEA reference level for iodine intake; IAEA (2011));
- 50 mSv effective dose for 1-year old (French reference level for evacuation).

The 555 kBq/m^2 reference level for Cs-137 deposition, initially chosen (Chernobyl reference level) is not shown since the threshold is too high and was not relevant for this case study. This is also the case for the 100 mSv effective dose (IAEA reference level for evacuation). Instead, for information, additional levels were considered:

- 10 kBq/m^2 deposition for Cs-137 and I-131;
- 10 mSv effective dose for 1-year old child;
- 10 mSv inhalation thyroid dose for 1-year old child.

Once the thresholds have been chosen, useful outputs for decision making, presented in this report, are the maximum distance and surface area affected by the threshold exceedance, and associated uncertainties. In addition, to compare several uncertainty assessments, a particular percentile may be chosen and drawn for each ensemble calculation. Once a given percentile has been chosen, it may be used for decision making in the same way as a deterministic simulation. Other outputs were already provided as examples in a previous deliverable (De Vries et al. 2019) and in conference presentations (Korsakissok et al. 2019).

Ensemble results for REM1 – short release

Results are shown for the short release, using the Harmonie-AROME ensemble of 10 members for REM1 scenario. The release time is on January 11th, 2017 at 12:00 UTC, the release height is 50 m, and the released quantities are those given in Table 2. Ensembles are presented for two configurations: (a) an unperturbed source term, therefore propagating only meteorological uncertainties, and (b) with source term perturbations as described in Table 6.

Thus, all participants have the same meteorological and, in the unperturbed case, release data. The dispersion model itself differs, as well as deposition-related data. This includes the type of model (see Table 4), the dry and wet deposition models, diffusion models (Bedwell et al. 2018), and other default hypotheses such as the initial source dilution. In addition, there may also be differences in dose calculation assumptions.

Median of the ensembles

Figure 12 shows the median of the seven ensembles for ^{137}Cs deposition at the end of the calculation. Here, there are no perturbations to the source term, and no additional perturbations to the models, in order to ensure that all participants use exactly the same inputs. Therefore, the observed differences are only due to different dispersion modelling assumptions, such as the description of the atmospheric dispersion process (Eulerian, Lagrangian particle, Lagrangian puff, Gaussian puff with different parametrizations), dry deposition schemes (dry deposition velocities fixed or calculated depending on particles properties and

[1] better referred to as “level of agreement maps”, since the uncertainties taken into account are partial and the inputs do not necessarily represent the same probability.

surface roughness), wet scavenging schemes (described in Table 5), but also modelling domain (DTU and EEAE used a smaller modelling domain than other participants) and eventual use of an interpolation scheme to map the model results to the defined grid locations. In particular, conversions between Cartesian and latitude/longitude grids may be a source of approximations. This figure illustrates that the modelling differences are far from negligible, even when using exactly the same input data and even when the same dispersion model is used (BfS and DTU). The scavenging schemes presented in Table 5 did not show significant differences between the participants and, indeed, it does not seem to be a primary source of variability. For instance, BfS and EEAE used exactly the same scheme but the deposition map of BfS shows higher values than that of EEAE. The high deposition shown by Met Office far from the source (large yellow area), however, may be explained by the in-cloud scavenging which induces larger deposition than for other participants.

The median is an example of ensemble output that can be used for decision making; however, other percentiles should be examined as well, in particular to determine the possible "worst cases". Here, for the sake of brevity, other percentiles are not shown. One can refer to De Vries et al. (2019) and Berge et al. (2019) for other example outputs.

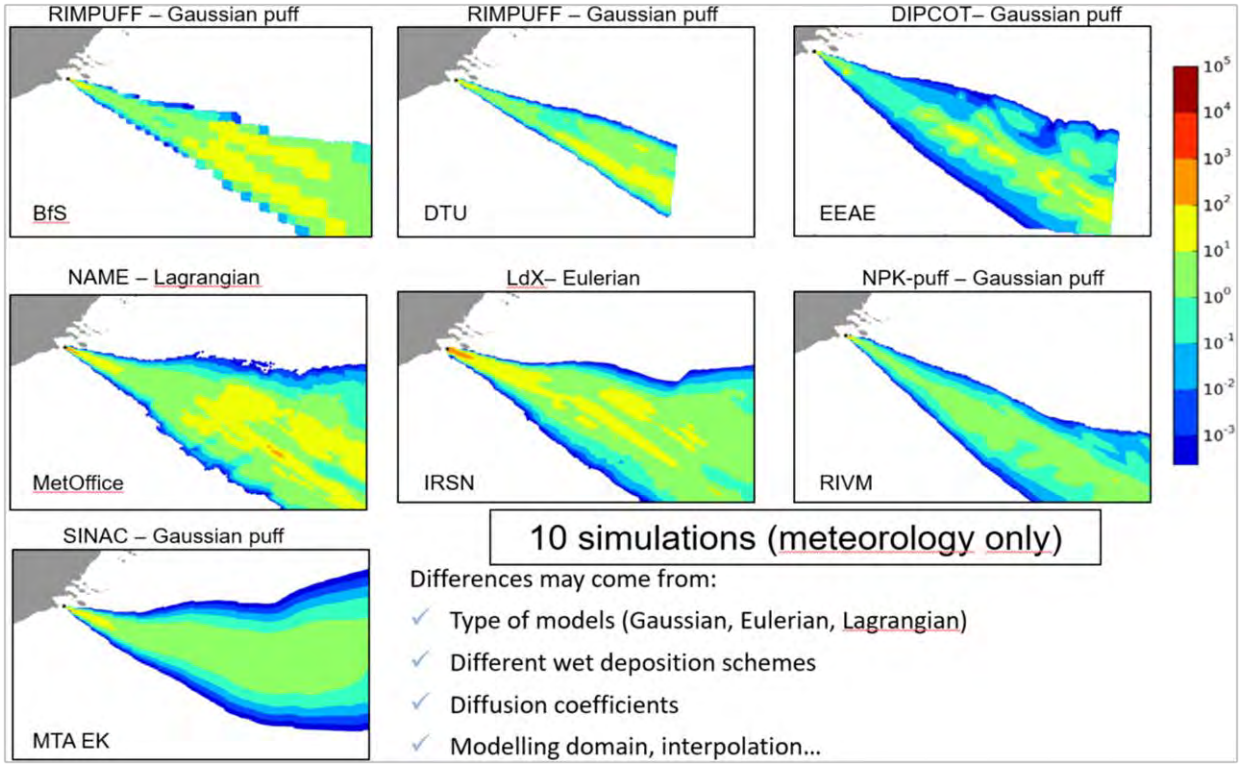


Figure 12 : median of the ^{137}Cs deposition in Bq/m^2 at the end of the release, for the seven ensembles. Only the meteorological uncertainties are taken into account (meteorological ensemble + reference source term without perturbation). Therefore, only the dispersion models differ between the participants.

Maximum distance of threshold exceedance

In Table 8, we compare the mean of the ensemble results for each participant, 24 hours after the beginning of the release. The variable of interest is the maximum distance from the source at which the threshold is exceeded, for several of the endpoints defined earlier. For each member of the ensemble, we determine the maximum distance above the given threshold (distance of the farthest grid point over the

threshold); the values given in Table 8 are the averaged values over the ensemble members (for each participant).

Table 8 : Distance (in km from the source) of the threshold exceedance averaged over the ensemble members, for the REM1 case study (short release), without source term perturbation and several variables of interest.

	Cs-137 Deposition		Inhalation thyroid dose		Effective dose	
	10 kBq/m²	37 kBq/m²	10 mSv	50 mSv	10mSv	50 mSv
BfS	543 km	447 km	142 km	40 km	16 km	4 km
DTU	417 km	403 km	97 km	37 km	11 km	0 km
EEAE	523 km	316 km	46 km	20 km	6 km	0 km
IRSN	558 km	389 km	209 km	79 km	26 km	1 km
Met Office/PHE	578 km	558 km	130 km	52 km	24 km	0 km
MTA EK	96 km	38 km	87 km	29 km	13 km	0 km
RIVM	543 km	15 km	40 km	19 km	7 km	0 km

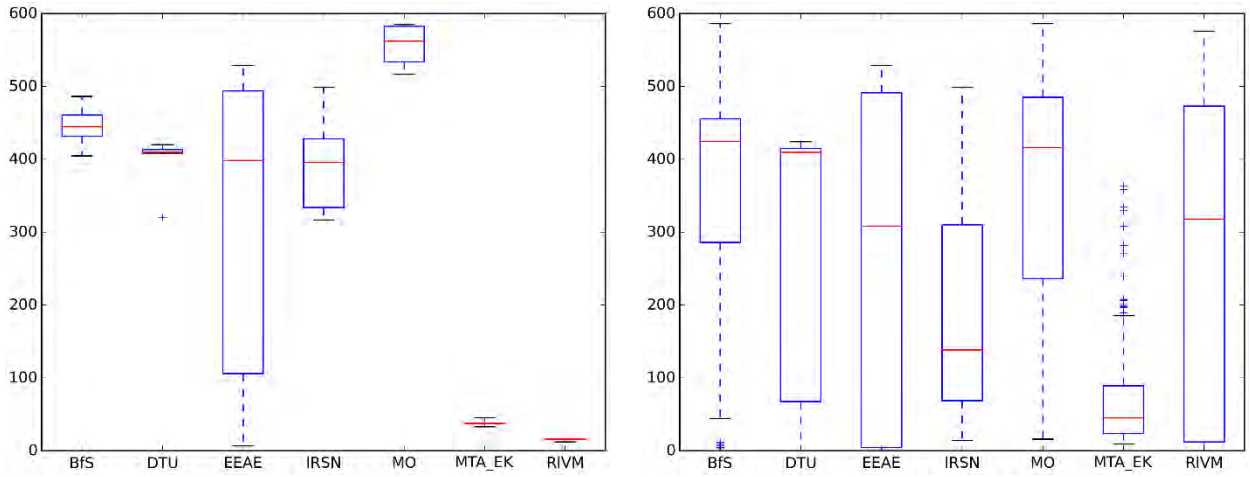
The values of Table 8 show a high variability between the participants. This variability comes from the type of model used (see Table 4). The first two lines correspond to the same model, the Gaussian puff model RIMPUFF. However, the modelling domain is not the same, which may explain the differences in the distance reached by ^{137}Cs deposition (the end of the modelling domain is reached by several members in these simulations). Differences may also come from the assumptions made in deposition and scavenging values; for instance, Met Office has the highest distance for deposition, which may be explained by the scavenging schemes discussed previously (Table 5). EEAE used a Lagrangian puff model with an additional (to the puff grow) 3-D random motion of the puffs which maybe overestimates turbulent diffusion, leading to lower distances; differences may come from assumptions concerning the diffusion model, scavenging coefficients and dry deposition velocities. Finally, IRSN's model is Eulerian, which is known to underestimate the values near the source but gives the highest distances in this case, except for ^{137}Cs deposition. This may come from the assumptions made in computing the vertical diffusion scheme during the preprocessing step, and/or in the choice of the vertical grid resolution, both of which may lead to overestimate concentrations in the first vertical layer. The inhalation thyroid dose should not be significantly impacted by a change in deposition scheme; however, there is an additional source of variability in the assumptions made for dose calculations. Here, the variability between participants is very large. For instance, the threshold of 50 mSv (that may be used to recommend stable iodine intake in some European countries) ranges between 19 km and 79 km. In the same way, the 10 mSv threshold on effective dose ranges between 7 and 26 km.

Table 9: Distance (in km from the source) of the threshold exceedance averaged over the ensemble members, for the REM case study (short release) with source term perturbation and several variables of interest.

	Cs-137 Deposition		Inhalation thyroid dose		Effective dose	
	10 kBq/m²	37 kBq/m²	10 mSv	50 mSv	10 mSv	50 mSv
BfS	505 km	415 km	160 km	36 km	15 km	4 km
DTU	417 km	403 km	97 km	37 km	11 km	0 km
EEAE	469 km	237 km	45 km	18 km	6 km	0 km
IRSN	409 km	180 km	228 km	90km	35 km	7 km
Met Office/PHE	483 km	361 km	130 km	55 km	28 km	5 km
MTA_EK	184 km	46 km	85 km	28 km	12 km	0 km
RIVM	482 km	222 km	37 km	18 km	7 km	1 km

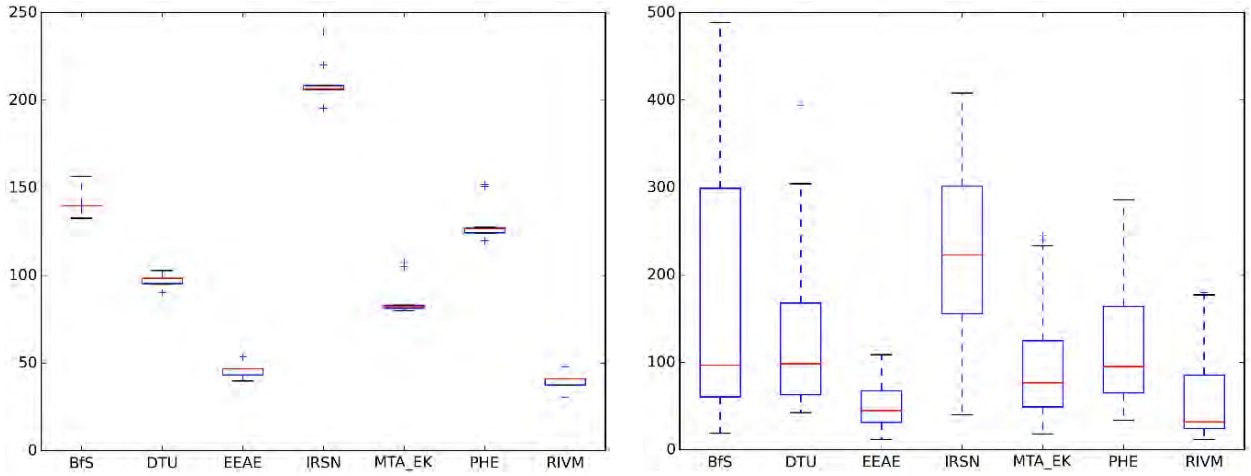
Maximum distances of threshold exceedance are then given for the case where source term perturbations are taken into account (Table 9). It can be noticed that the deposition distances are generally lower than those in Table 8, meaning that the source perturbations tend to decrease *on average* the deposition distance. It may be that the meteorological conditions for the perturbed release times lead to lower ground deposition (lower or no rain) than the initial release time. On the other hand, the distance for inhalation thyroid dose and effective dose thresholds tends to be similar or higher than the results without perturbation. If wet deposition is lower, then the plume is less depleted and it is consistent with a higher inhalation dose. For effective dose greater than 50 mSv, the distances were previously zero for IRSN and Met Office/PHE with the non-perturbed source term (in Table 4), while some perturbed simulations (for instance when multiplying the source term by 3) lead to non-zero distances, which explains the results in Table 9.

The previous results present an indicator useful for decision making, the “maximum distance of threshold exceedance”. However, this indicator was averaged over the ensembles and did not take into account the ensembles’ spreads. One of the questions is then whether the inter-model variability highlighted by these results is larger than each individual ensemble’s spread. In other words, how large is the inter-model (or inter-ensemble) variability compared to the uncertainties taken into account in the ensemble simulations? Figure 13 (for the deposition) and Figure 14 (for the inhalation thyroid dose) show that the variability across the seven ensembles is typically much greater than the individual ensemble model spread given by the standard deviation, when considering the meteorological uncertainties only. This is easily explained by the fact that the meteorological uncertainty is small in this case. On the contrary, when taking into account both meteorological and source uncertainties, the ensembles’ spreads are much larger, although there is still a very high variability between the medians of the seven ensembles (Figure 13(b) and Figure 14(b)). This illustrates the importance of taking into account other uncertainties besides meteorological data, including model-related uncertainties.



(a) Meteorological uncertainty only (b) Meteorology + ST perturbations

Figure 13: Ensemble mean of the maximum distance (in km) for the threshold exceedance of 37 kBq/m² of ^{137}Cs deposition for the seven participants, 24 hours after the beginning of the release, and associated standard deviations.



(a) Meteorological uncertainty only (b) Meteorology + ST perturbations

Figure 14: Ensemble mean of the maximum distance (in km) for the threshold exceedance of 10mSv for the inhalation thyroid dose for the seven participants, 24 hours after the beginning of the release, and associated standard deviations.

Level of agreement for threshold exceedance

The level of agreement maps for two participants are shown for illustration in Figure 15, with the 10-members simulation (left) and with the additional source perturbations (right). For these two participants, the maximum distance of threshold exceedance is lower with the perturbed simulations (503 km for BfS and 482 km for Met Office) than without (542 km and 578 km respectively). Figure 15 clearly shows that, although the area of highest probabilities is much smaller, the contaminated surface above the 5th percentile is much larger with the perturbed simulations. When taking into account different release times, different plume directions are taken into account (as shown by the trajectories Figure 4), which leads to a larger spread. Therefore, it is important to highlight that the maximum distance of threshold exceedance in itself is not a good indicator of the uncertainty: if the uncertainties are very large, then few members will agree with each other, leading to small probabilities and therefore, smaller distances. It should be associated with a map, in order to infer the possible contaminated areas, including those concerned by a low percentile.

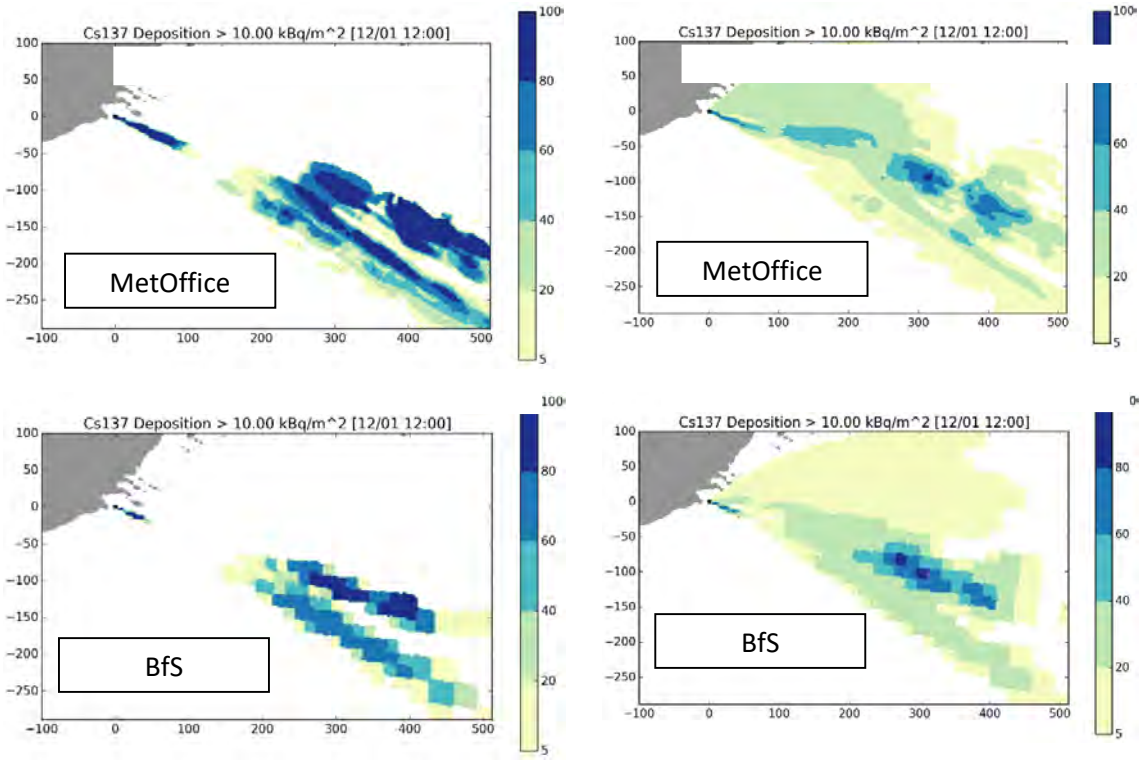


Figure 15 : Probability maps of a threshold exceedance of 10 kBq/m² for Cs-137 deposition, for a number of discrete bands of percentiles, without (left) and with (right) additional perturbations on the source term combined with the application of a meteorological ensemble.

In the remaining part of this chapter, only the results for the "full" ensembles, including source term uncertainties, are shown.

The levels of agreement of threshold of exceedance can be drawn for each ensemble, given an output variable and a threshold (Figure 16 and Figure 17). In Figure 16, the variability between the deposits of ^{137}Cs given by the participants is clear. It should be noted that, in this case, DTU did not perturb the release time; only the quantities were perturbed afterwards (Table 6), which explains the small area covered by the non-zero probabilities. All other participants show some probability of contamination in the north-eastern direction, due to a change in the release time. The pattern induced by scavenging, showing "hot spots" of deposition, is clear for most participants. Figure 17, representing the threshold exceedance of 10 mSv thyroid dose, shows a more continuous plume but still a large variability between the participants.

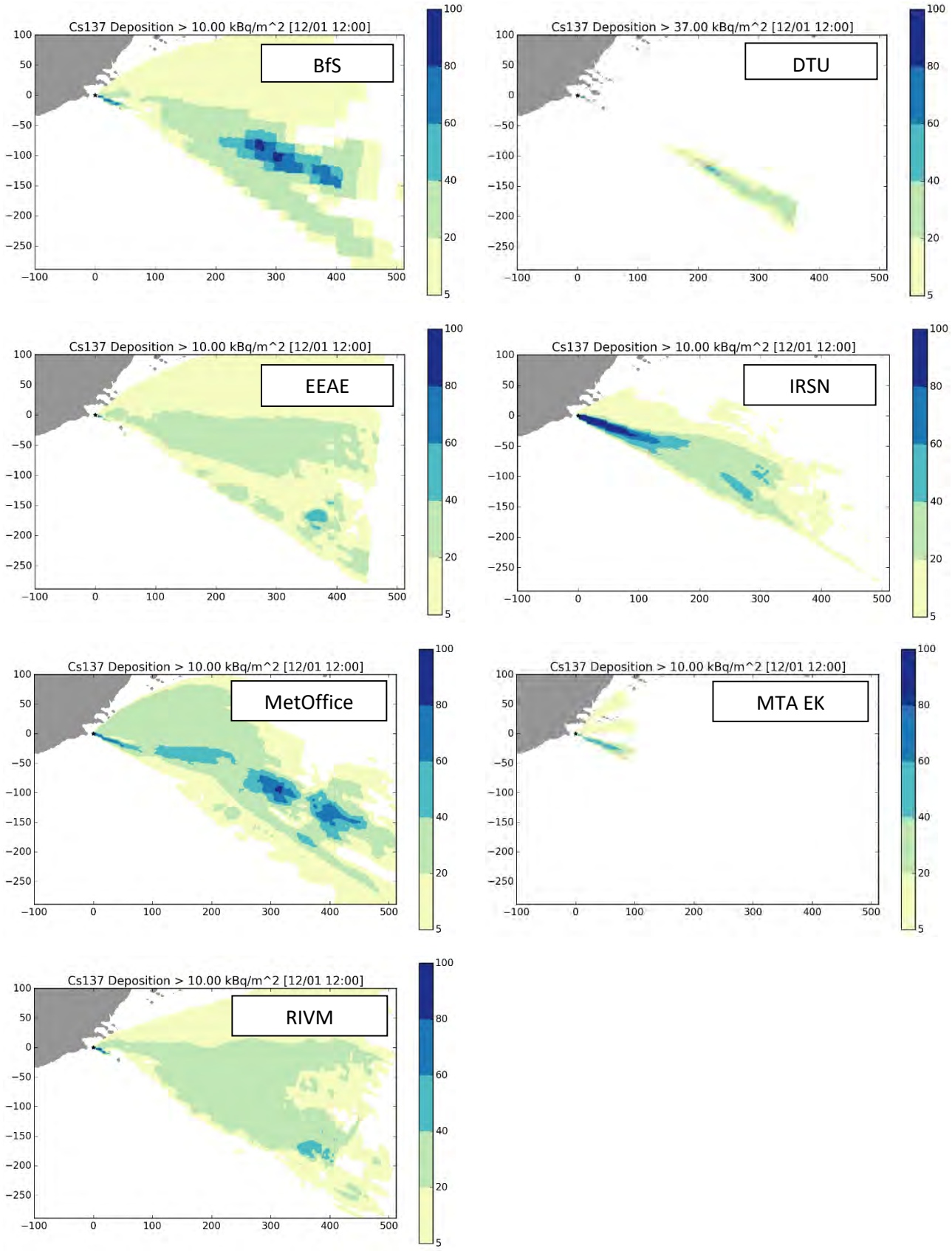


Figure 16 : Probability maps of a threshold exceedance of 10 kBq/m² for Cs-137 deposition, for a number of discrete bands of percentiles, for several participants.

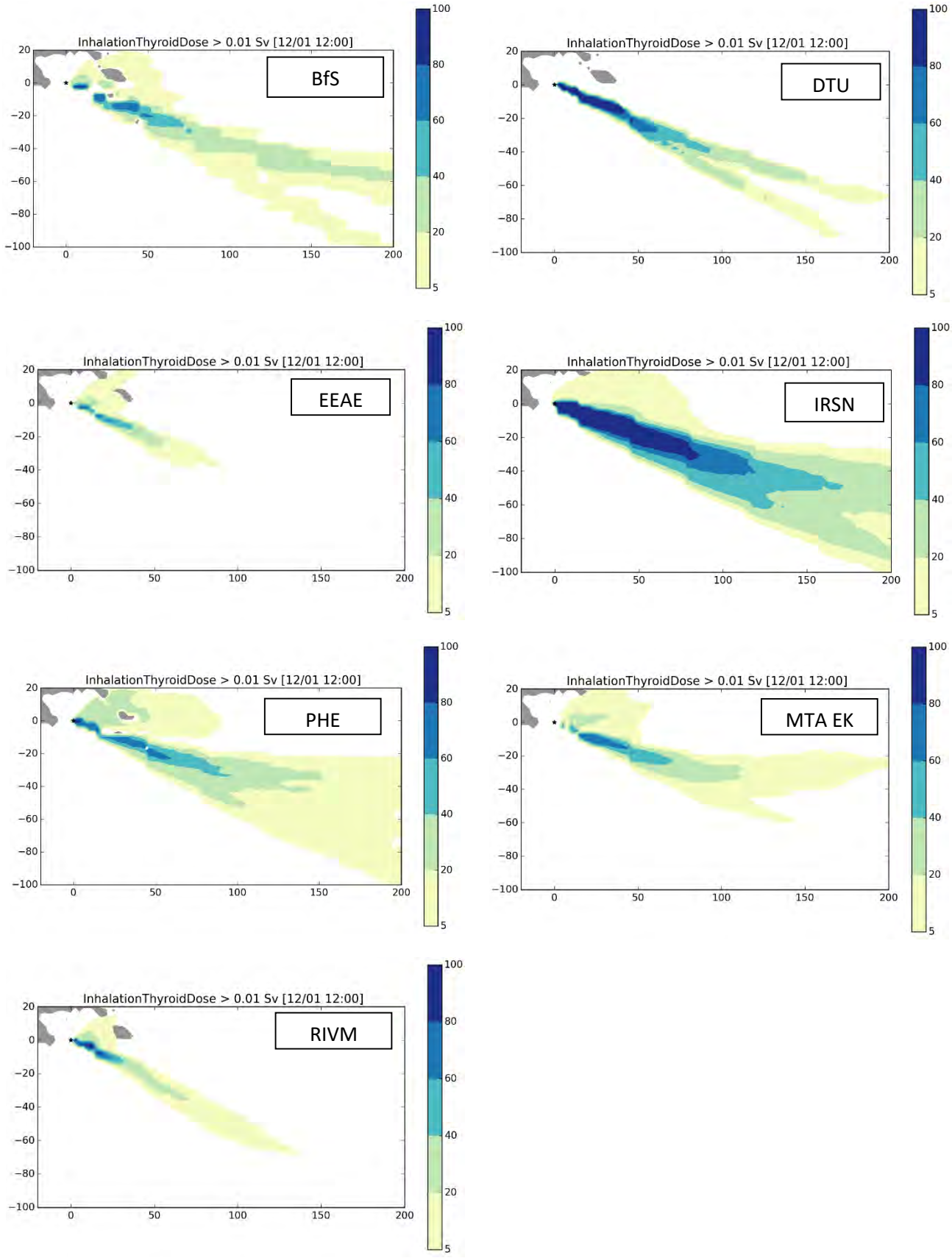


Figure 17 : Probability maps of a threshold exceedance of 10mSv for the inhalation thyroid dose, for a number of discrete bands of percentiles, for several participants, for the REM case study (short release) with source term perturbation and several variables of interest.

Ensemble results for REM2 – short release

This section briefly presents the results for the REM2-S case, corresponding to a release on January 12, at 18:00 UTC, with the same uncertainties on the release time, height and quantities as for the REM1-S case (provided in Table 6). Except Figure 18, both meteorological and source term uncertainties are taken into account. Figure 18 illustrates the inter-model variability by eliminating the differences in the source term perturbations between the participants: only the 10 meteorological members are used with the unperturbed source term. Therefore, all ensembles on this figure use strictly the same input data.

The median of the ensembles show that the deposition is globally higher close to the source than for REM1-S (Figure 18), which is explained by the presence of precipitation at the source location during the release time (Figure 6).The median plume trajectory is heading South-East and turning more East as the plume travels; the fact that some meteorological members may include trajectories heading North-East (Figure 9) is not reflected by the median maps. As for the REM1-S case, the median shows significant differences in the deposition between the participants. Beside interpolation and domain size artefacts, the plume spread is quite different, probably illustrating the variability between diffusion schemes. Again, wet diffusion schemes may explain some of the variability in the deposited amount of ^{137}Cs.

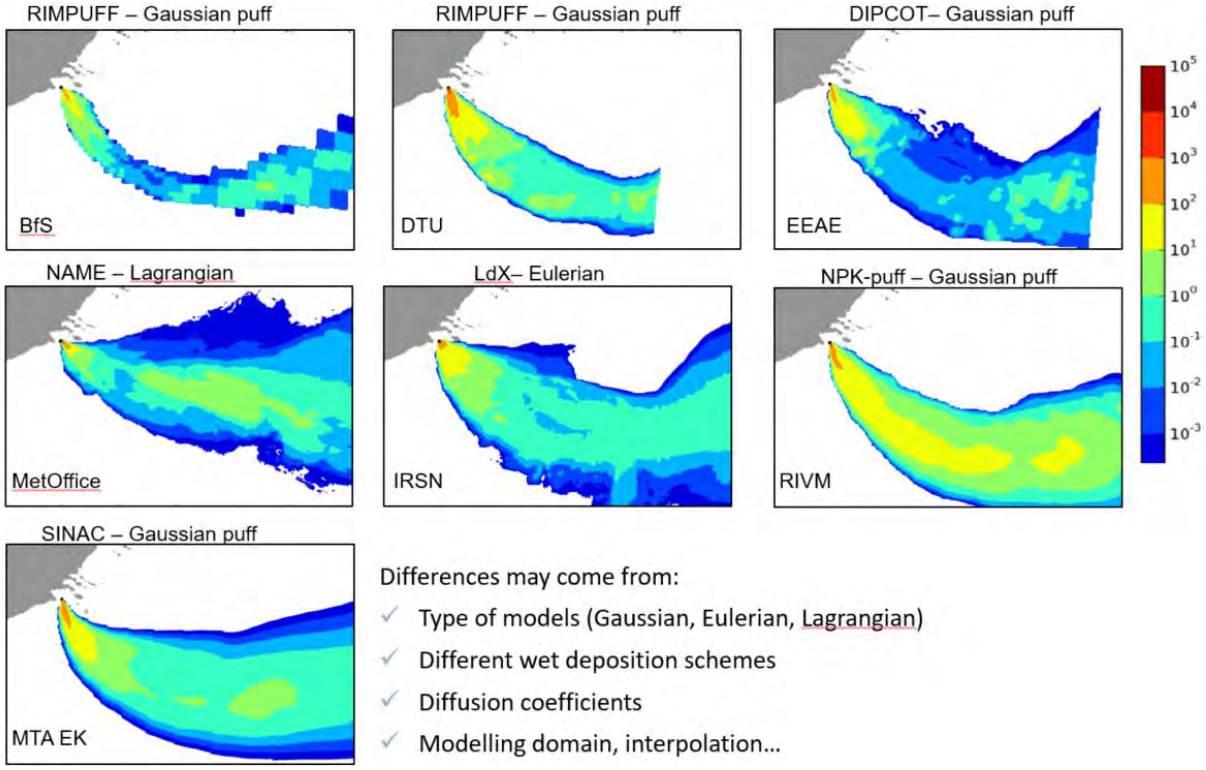


Figure 18 : median of the ^{137}Cs deposition in Bq/m^2 at the end of the release, for the seven ensembles. Only the meteorological uncertainties are taken into account (meteorological ensemble + reference source term without perturbation). Therefore, only the dispersion models differ between the participants. REM2-S case study.

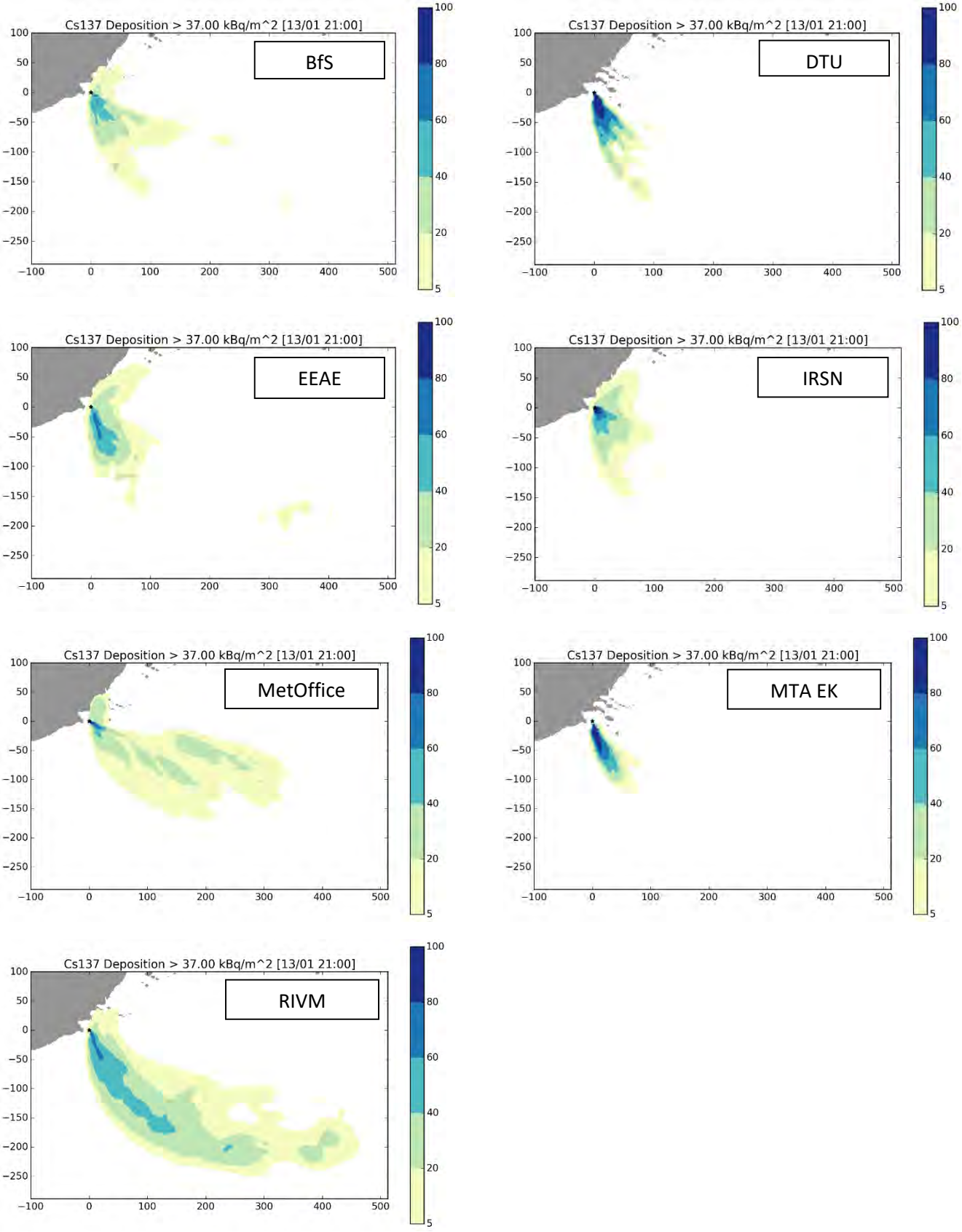


Figure 19 : Probability maps of a threshold exceedance of 37 kBq/m² for Cs-137 deposition, for a number of discrete bands of percentiles, for several participants. REM2 case study, short release.

The probability maps shown Figure 19 take into account both meteorological and source term variability. Here again, there are significant differences between the participants. In this case, in addition to the inter-model variability shown by Figure 18, they may also be caused by the various ways the source term uncertainties are taken into account. From the previous analysis of trajectories the plume may travel North according to several members; this is visible for several participants with a small lobe (low percentile) in this direction. In fact, precipitation exceeds 2mm/hr for many ensemble members and peaks as the wind direction changes between 18:00 UTC and 20:00 UTC (Figure 20).Thus, when the plume travels in that direction, the precipitation is so high that the plume is completely washed out before it has travelled very far, which explains the small distance of threshold exceedance.

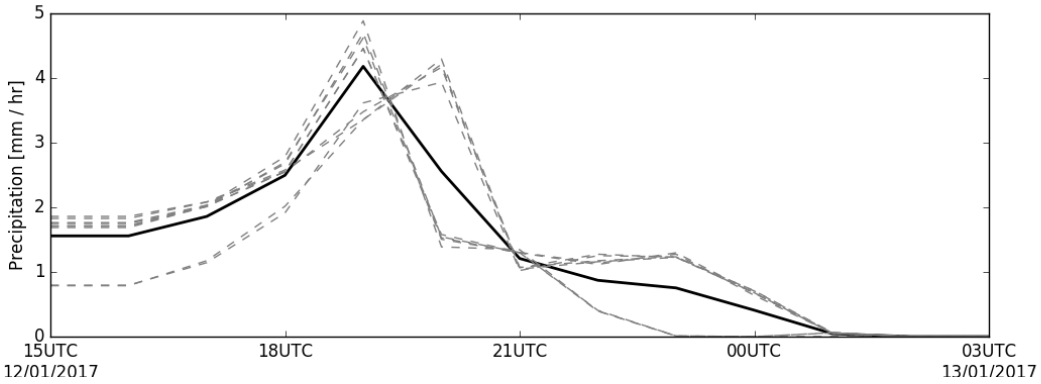
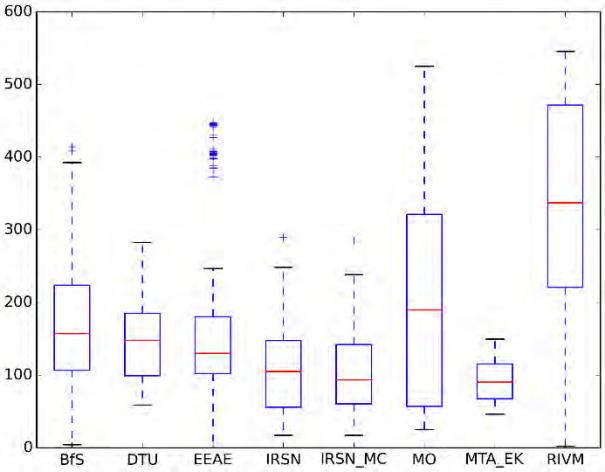
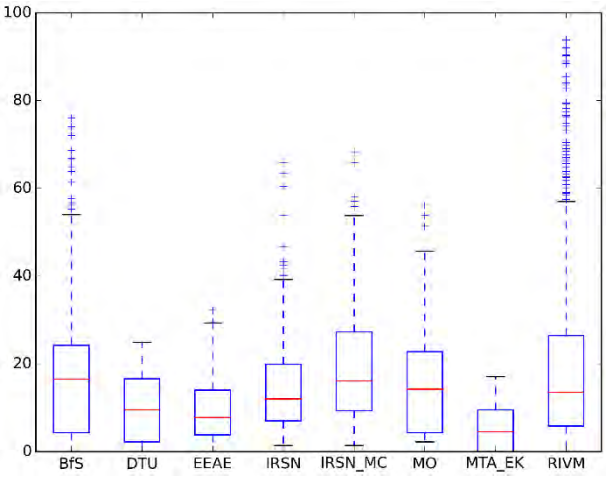


Figure 20 : precipitation (in mm/hr) as a function of time at the source location for the different ensemble members.

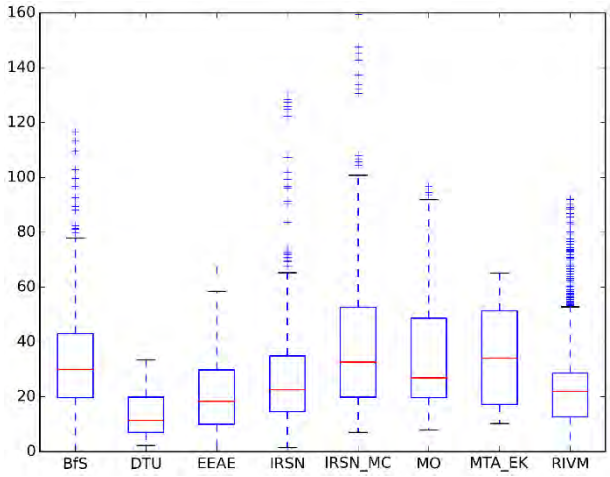
This variability in the areas affected by the contamination is not totally reflected when considering the maximum distance of threshold exceedance only (Figure 21(a)), which shows globally good consistency between the participants, both in the median and spread, except Met Office and RIVM which overestimate the deposition compared to other participants. The distances are globally lower than in the REM1-S case, which may be due to large precipitation close to the source, as mentioned earlier, as well as a convective situation favouring rapid mixing of the pollutants. As far as effective and inhalation thyroid dose are concerned, (Figure 21(b) and (c)), there is also globally a good consistency between the participants, when looking at the 25-75th percentile range; however, there are outliers (represented by blue crosses on the figures) for some ensembles, notably RIVM and BfS, which correspond to much larger distances of threshold exceedance: while the average distances for 10 mSv effective dose is below 20 kilometres, some particular members of these ensembles predict a threshold exceedance as far as 80 kilometres or more. This raises the issue of taking into account “extreme” values in case of emergency. In practice, such uncertainty assessment might be used for decision making as a way to prioritize the countermeasures, based on the probabilities associated with a given area. Countermeasures might be ordered in high-probability areas, and prepared in low-probability ones while waiting for more precise evaluations. Indeed, the uncertainties provided here are representative of the state of knowledge at a given time (in particular, uncertainty in the release time is still high) and may be reduced as the accidental sequence progresses (Figure 10). Finally, in Figure 21, IRSN presents two sets of results: those named simply “IRSN” represent cross-simulations, using only source term and meteorological perturbations, while “IRSN_MC” stands for “IRSN Monte Carlo” and represent the Monte Carlo ensemble designed as described in Table 7. The difference between the two results is therefore representative of the uncertainty related to physical parameters (diffusion and deposition). It does enhance a little the ensemble’s spread, particularly in the case of effective and inhalation dose; however, this is clearly of second order compared to meteorological and source term uncertainties.



(a) Maximum distance of threshold exceedance of ^{137}Cs deposition for a level of 37 kBq/m^2



(b) Maximum distance of threshold exceedance of effective dose for a level of 10 mSv



(c) Maximum distance of threshold exceedance of inhalation thyroid dose for a level of 50 mSv

Figure 21: Ensemble mean of the maximum distance for the threshold exceedance of (a) 37 kBq/m² of ^{137}Cs deposition, (b) 10 mSv for effective dose and (c) 50 mSv for inhalation thyroid dose, for the seven participants, 24 hours after the beginning of the release, and associated standard deviations. REM2 case study, short release.

Ensemble results for REM – long release

The long release begins on January 11th at 06:00 UTC and lasts three days. Therefore, it includes both weather situations described in the REM1 and REM2 scenarios. In particular, the time evolution of the release shown by Figure 11 indicates that there is a second, higher release for noble gas and gaseous Iodine (but not Caesium, as particles are filtered) corresponding to the opening of the containment venting system. There is a large uncertainty on the timing of this second release, but it occurs more or less 30 hours after the beginning of the release, which corresponds to January 12th at 12:00 UTC. Therefore, for a large part of the simulations, this second release corresponds to the REM2 situation described earlier in this report. Concerning the source terms, four participants (DTU, EEAE, IRSN and Met Office) used a reduced ensemble of ten source terms, designed to be representative of the original 150 ST ensemble's spread (as explained in (Bedwell et al. 2019), while the other two (BfS and MTA EK) used an ensemble of 5 source terms (designed in a similar way). Otherwise, inputs are the same (no additional perturbation for IRSN).Figure 22 features the probability maps associated with the REM-L simulations. There are, indeed, several areas potentially impacted by the plume, and in some cases, different "lobes" can be distinguished, indicating different plume directions. Here, only Iodine deposition is shown, as the release for Caesium was negligible due to the filtering, and no significant threshold exceedance could be displayed. The choice of using 5 source terms instead of 10 does not seem to be of primary importance for this kind of output: the map obtained by BfS (5 ST) is very similar, for instance, to that provided by DTU (10 ST), which is consistent with the fact that they use the same dispersion model (RIMPUFF). The results of those two participants were not so similar in the REM1-S and REM2-S cases, which was due to the difference in the way the two participants had perturbed the source term (Table 6). MTA EK have the highest deposition, which may partly due to the high scavenging coefficient, but also to other features related to dispersion, since inhalation thyroid dose is also very high (as can be seen on Figure 23(b)).Figure 23 shows the maximum distance of threshold exceedance for the different participants, for (a) a threshold of 10 kBq/m^2 for ^{131}I and (b) a threshold of 10 mSv for inhalation thyroid dose. While there is a significant variability in the ensembles' median, Figure 23(a) shows that there is a good consistency in the estimation of the overall uncertainty, represented by the blue boxes, between the participants, at least four of them (EEAE, IRSN, DTU and MetOffice).

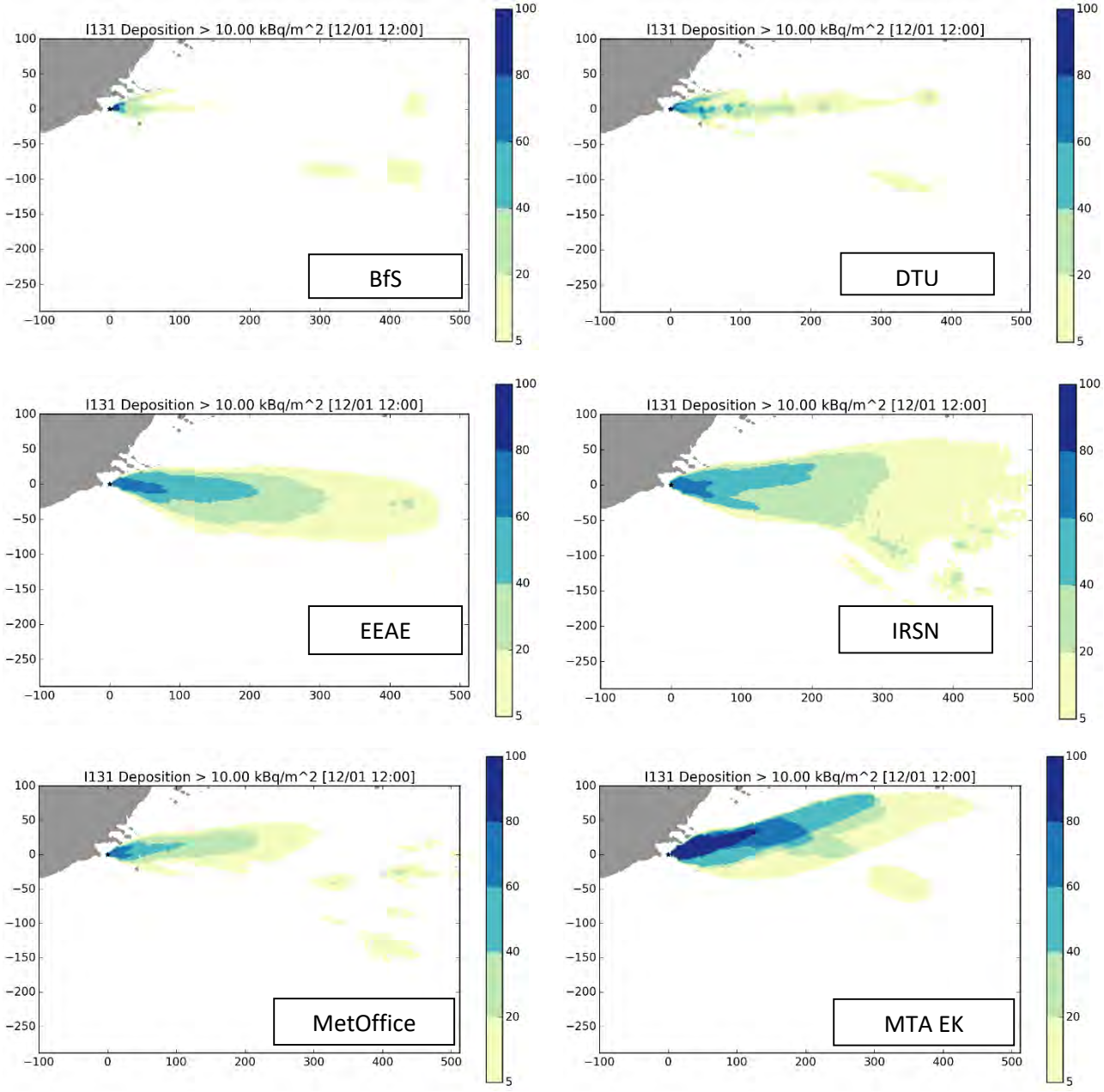
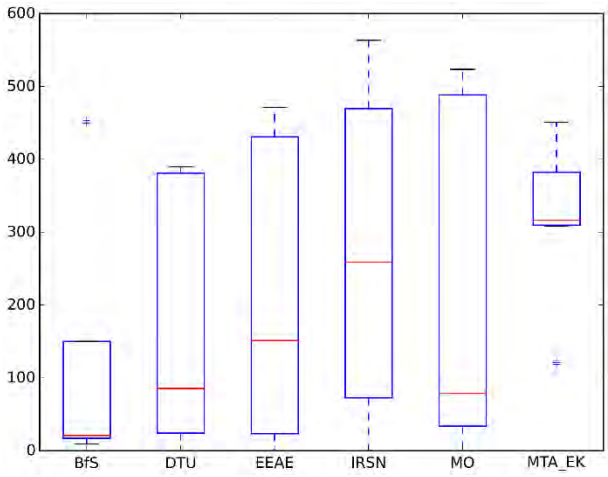
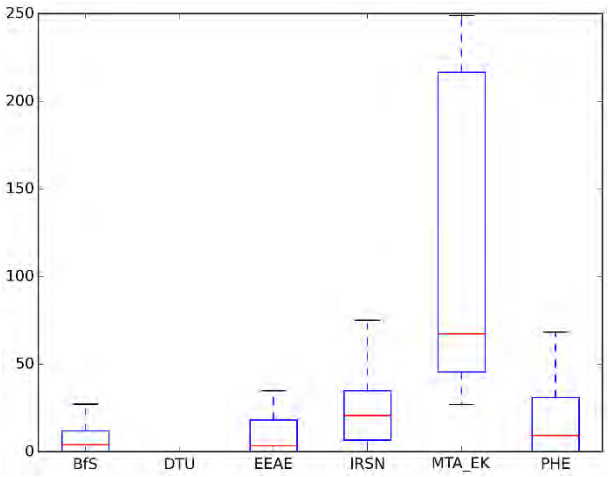


Figure 22: Probability maps of a threshold exceedance of 10 kBq/m² for I-131 deposition, for a number of discrete bands of percentiles, for several participants.



(a) Maximum distance of threshold exceedance of ^{131}I deposition for a level of 10 kBq/m^2



(d) Maximum distance of threshold exceedance of inhalation thyroid dose for a level of 10 mSv

Figure 23: Ensemble mean of the maximum distance for the threshold exceedance of (a) 10 kBq/m² of ^{131}I deposition, (b) 10 mSv for inhalation thyroid dose, for six participants, 24 hours after the beginning of the release, and associated standard deviations.

Conclusions

Three hypothetical accident scenarios that took place at Borssele (The Netherlands) have been investigated. They consisted of two different meteorological scenarios, REM1 with a small meteorological variability and REM2 with a larger variability, and two source term scenarios. The short release was used with both REM1 and REM2 scenarios, while the long release (3 days) covered both meteorological situations. This resulted in very large datasets (several hundreds of simulations carried out by six to seven participants for three cases). To our knowledge, this provides a unique database; usually, inter-comparisons such as the ones carried out here are focused on deterministic simulations.

For the two meteorological scenarios with the short release (REM1-S and REM2-S), seven participants carried out ensemble dispersion simulations, propagating both meteorological and source term uncertainties. Six participants provided results for the long release. Several endpoints were derived. A particular attention was devoted in this report to the maximum distance of threshold exceedance, for deposition, effective dose and inhalation thyroid dose. Several reference levels of interest for decision making were used. In addition to this variable, agreement maps were shown, that is, maps showing the proportion of ensemble members above a given threshold on each grid point.

The results showed that, when considering only meteorological uncertainty, there was a large variability in the outputs provided by the participants, especially when considering the maximum distance of threshold exceedance. In this case, the inter-model variability was clearly larger than the uncertainty linked to meteorological data (especially for the REM1 case). However, when considering source term uncertainties along with meteorological ensemble, the variability between the participants was of less importance, although not negligible. In particular, when considering exactly the same source term perturbations (for the long release), there was a good consistency between the participants' uncertainty estimates; the ensembles' spread mostly encompass the inter-model variability. These findings highlight the importance of properly taking into account all primary sources of uncertainties, especially the source term uncertainties and not only those stemming from the meteorological forecasts.

Another finding of this study was the sensitivity of the maximum distance of threshold exceedance to the input uncertainties. When more uncertainties are taken into account, the maximum distance (averaged

over the ensemble) tends to be reduced. In other words, if the uncertainties are large, ensembles' members tend to be in less good agreement with each other. This results in a larger surface covered by "low agreement" (e.g. 5th percentile) but numerical indicators averaged over the ensemble, such as the maximum distance of threshold exceedance, tend to be reduced. Therefore, care should be taken to not only use such a numerical output, but to include graphical representation of the uncertainties as well, such as the agreement maps (or maps provided for different percentiles), and/or maps given by one or several well-chosen members. The choice of the appropriate threshold, although usually driven by decision making, should also be made with care, as the outputs may highly differ depending on the threshold.

References

Baklanov, A. and J. H. Sørensen (2001). "Parameterization of radionuclide deposition in atmospheric long-range transport modelling." Phys. Chem. Earth (B) **26**(10): 787-799.

Bedwell, P., I. Korsakissok, S. Leadbetter, R. Périllat, C. Rudas, J. Tomas and J. Wellings (2019). D 9.5.5 Guidelines for the use of ensembles in the description of uncertainty in atmospheric dispersion modelling: operational applications in the context of an emergency response. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Bedwell, P., S. Leadbetter, J. Tomas, S. Andronopoulos, I. Korsakissok, R. Périllat, A. Mathieu, G. Geertsema, H. Klein, H. De Vries, T. Hamburger, T. Pazmandi, C. Rudas and A. Sogachev (2018). D 1.1.4 Guidelines detailing the range and distribution of atmospheric dispersion model input parameter uncertainties. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Bengtsson, L., U. Andrae, T. Aspelien, Y. Batrak, J. Calvo, W. d. Rooy, E. Gleeson, B. Hansen-Sass, M. Homleid, M. Hortal, K.-I. Ivarsson, G. Lenderink, S. Niemelä, K. P. Nielsen, J. Onvlee, L. Rontu, P. Samuelsson, D. S. Muñoz, A. Subias, S. Tijm, V. Toll, X. Yang and M. Ø. Køltzow (2017). "The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System." Monthly Weather Review **145**(5): 1919-1935.

Berge, E., H. Klein, M. Ulimoen, S. Andronopoulos, O. C. Lind, B. Salbu and N. Syed (2019). D 9.5.2 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the Western Norway case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Chevalier-Jabet, K. (2019a). Design and use of Bayesian networks for the diagnosis/prognosis of severe nuclear accidents, EU program for research and innovation H2020 - FAST Nuclear Emergency Tools (FASTNET) project. Report FASTNET-DATA-D2.3.

Chevalier-Jabet, K. (2019b). Source term prediction in case of a severe nuclear accident. 5th NERIS workshop. Roskilde, Denmark.

De Vries, H., G. Geertsema, I. Korsakissok, R. Périllat, J. Tomas, R. Scheele, S. Andronopoulos, P. Astrup, P. Bedwell, T. Charnock, T. Hamburger, I. Ievdin, S. Leadbetter, T. Pazmandi, C. Rudas, A. Sogachev, P. Szanto and J. Wellings (2019). D9.4 Published sets of probability maps of threshold exceedance for scenarios, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project. Report D9.4 20. http://www.concert-h2020.eu/en/Publications

Geertsema, G., H. De Vries and R. Sheele (2019). High resolution meteorological ensemble data for CONFIDENCE research on uncertainties in atmospheric dispersion in the (pre-)release phase of a nuclear accident. 19th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

IAEA (2011). Criteria for use in preparedness and response for a nuclear or radiological emergency: general safety guide. IAEA safety standards series No. GSG-2. Vienna, Austria, Jointly sponsored by the FAO, IAEA, ILO, PAHO, WHO. **STI/PUB/1467**.

Korsakissok, I., S. Andronopoulos, P. Astrup, P. Bedwell, K. Chevalier-Jabet, H. De Vries, G. Geertsema, F. Gering, T. Hamburger, H. Klein, S. Leadbetter, A. Mathieu, T. Pazmandi, R. Périllat, C. Rudas, A. Sogachev, P. Szanto, J. Tomas, C. Twenhöfel and J. Wellings (2019). Comparison of ensembles of atmospheric dispersion simulations: lessons learnt from the confidence project about uncertainty quantification. 19th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Mathieu, A., I. Korsakissok, R. Périllat, K. Chevalier-Jabet, F. Stephani, S. Fougerolle, V. Créach, E. Cogez and P. Bedwell (2018). D 1.1.3 Guidelines describing source term uncertainties. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Périllat, R., I. Korsakissok, V. Mallet, A. Mathieu, T. Sekiyama, M. Kajino, K. Adachi, Y. Igarashi and D. Didier (2017). Using meteorological ensembles for atmospheric dispersion modelling of the Fukushima nuclear accident. 18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Bologna, Italy.

Saunier, O., A. Mathieu, D. Didier, M. Tombette, D. Quélo, V. Winiarek and M. Bocquet (2013). "An inverse modeling method to assess the source term of the Fukushima nuclear power plant accident using gamma dose rate observations." Atmos. Chem. Phys. Discuss. **13**(6): 15567-15614.

D 9.5.2 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the Western Norway case study

Lead Authors: Erik Berge, Heiko Klein,
With contribution from: Magnus Ulimoen, Spyros Andronopoulos, Ole Christian Lind, Brit Salbu, Naeem Syed

Review(s): Irène Korsakissok

Table of Contents

Introduction 41

Meteorological data of the WN-case. 42

Source term of the WN-case. 43

The dispersion models DIPCOT and SNAP 44

Results 46

Uncertainties in the meteorological data 46

Results from the dispersion model runs 50

Base run of the dispersion of the plume from the explosion 50

Maps of uncertainty in accumulated concentrations 51

Case 1: Emission scenario (3) and 10 EPS members 51

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios 52

Case 3: 10 EPS members and 5 emission scenarios. 52

Maps of uncertainty in total accumulated depositions 55

Case 1: Emission scenario (3) and 10 EPS members 55

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios 55

Case 3: 10 EPS members and 5 emission scenarios. 55

Time-series of total activity at the location of the source and Vikedal 59

Accumulated concentrations 59

Case 1: Emission scenario (3) and 10 EPS members 59

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios 60

Total accumulated deposition 61

Case 1: Emission scenario (3) and 10 EPS members 61

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios 62

Summary of the uncertainty estimates due to meteorology, emission and dispersion model. 63

Conclusions 65

References 66

Introduction

The number of nuclear vessels in Norwegian territorial waters, including visits of allied vessels to Norwegian ports, has increased the last years. This poses an enhanced risk of a nuclear accident along the Norwegian coastline with a following release and dispersion of radioactive material into air and water.

In the Work Package 1 of the CONFIDENCE project, several hypothetical accident scenarios in Europe were investigated, and uncertainty propagation through atmospheric dispersion models was carried out. The case study presented in this report deals with a hypothetical accident in a nuclear vessel at 58.5°N and 4°E approximately 100 km off the Norwegian coast (see Figure 24). This case study has been investigated by two dispersion models. The case is named “the Western Norway case” (WN/case) and the aim of this study is to quantifying the uncertainties due to meteorology, the source term and the choice of dispersion model. The scenario of the hypothetical accident has been selected in close collaboration with the Centre for Environmental Radioactivity (CERAD CoE), and the Norwegian Radiation and Nuclear Safety Authorities (DSA). The two dispersion models applied to this case are: The SNAP (SNAP-Severe Nuclear Accident Program) model developed and operated by the Norwegian Meteorological Institute (see Bartnicki et al., 2011) and the DIPCOT model (DIPCOT – Dispersion over Complex Terrain) developed and operated by the Greek Institute of Nuclear Technology and Radiation (Andronopoulos et al., 2009).

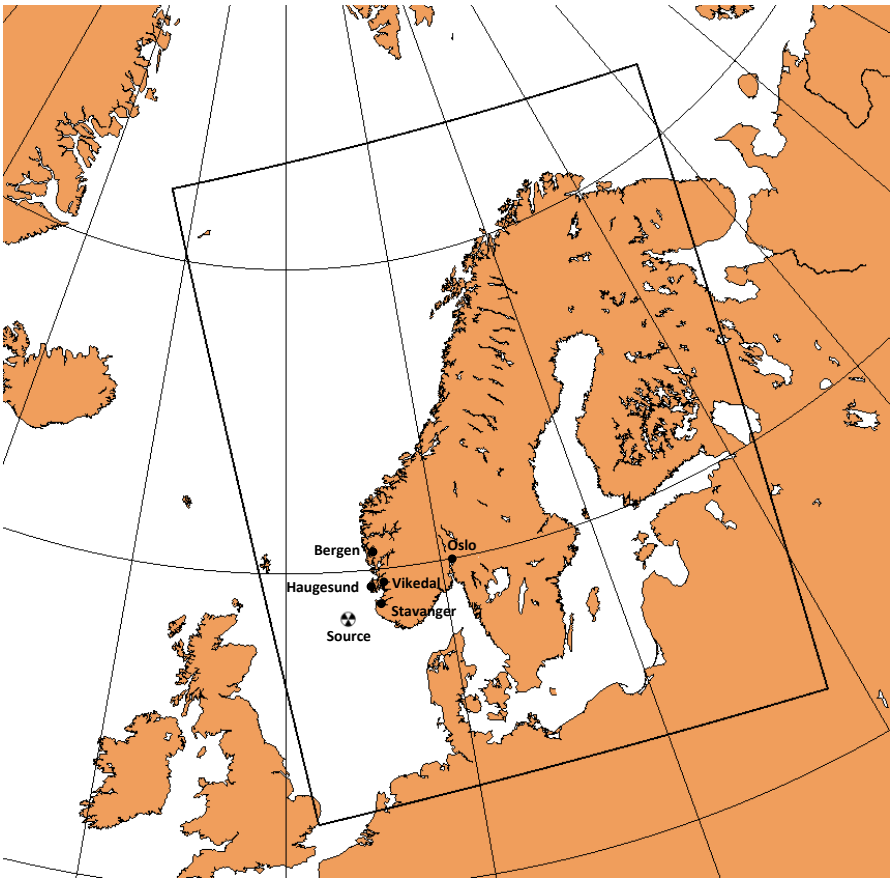


Figure 24. The location of the hypothetical source, five cities/towns and the domain of the meteorological calculation area (inside the frame).

Meteorological data of the WN-case.

The meteorological data is obtained from the operational weather prediction model of the Norwegian Meteorological Institute (see Müller et al., 2017). The model includes an Ensemble Prediction System (EPS) with 9 ensemble members in addition to the main deterministic run (member 0). A description of the ensemble version of the model can be found in Frogner et al. (2019). The model is run with 2.5 km horizontal grid spacing and 60 levels in the vertical, and the modelling domain spans an area of 739*949 grid-points. The operational domain applied to the WN/case is shown in Figure 25.

The meteorological data available for the WN-case are from a meteorological forecast starting at 06 UTC 16.03.2017. The potential accident takes place at 09 UTC and the focus of the dispersion calculations are for the 24 h period from 09 UTC 16.03.2017 to 09 UTC 17.03.2017. The available simulation period of the meteorological model is from 06 UTC 16.03.2019 to 00 UTC 19.03.2017 (+66 hours).

An overview of the weather pattern during the 24 hours of focus on the dispersion, is given in

Figure 25. At the start of the accident (09 UTC 16.03.2017), we encounter a surface low pressure near the western boundary of the domain east of Iceland. In the area of the source southwesterly winds up to about 10 m/s at 10 m height above the sea level, are seen. After 24 hours the low pressure is located outside the coast of central Norway, and the wind in the source region and over Western Norway has veered from southwest to west together with a strengthening of the wind speed. A key issue of this work is to analyze the uncertainty in the dispersion of radioactivity due to uncertainties in the meteorological parameters. The uncertainty associated with the meteorology as characterized by the 10 ensemble members, is discussed in section 4.

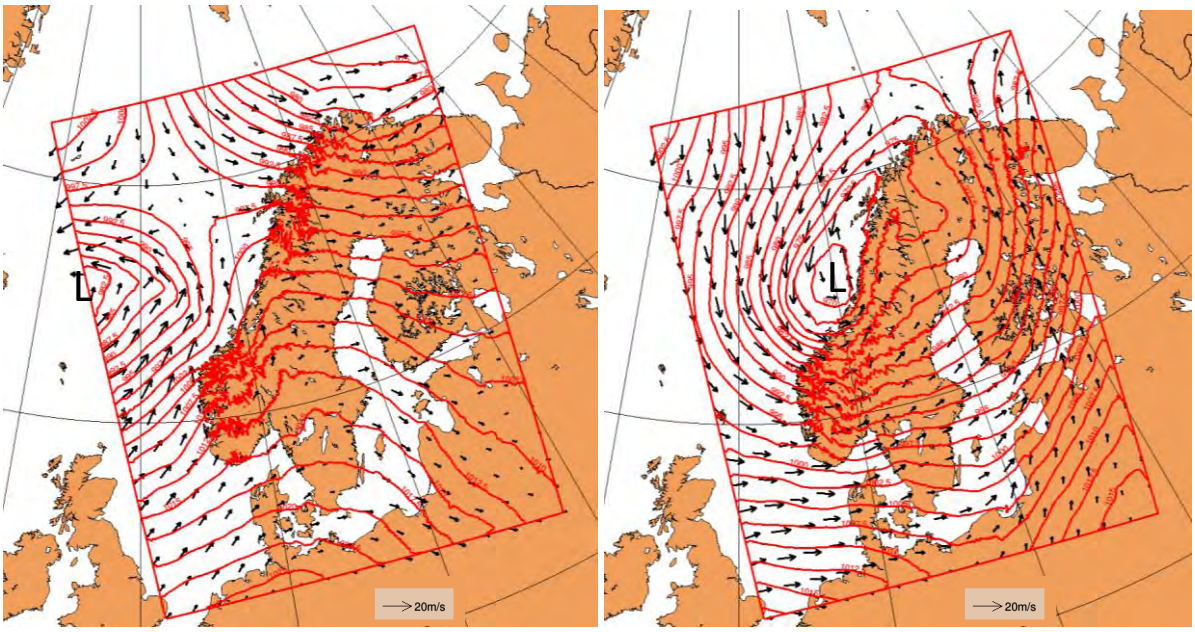


Figure 25. Mean sea level pressure (red lines every 2.5 hPa) and near surface wind speed ca. 10 m above the surface (black arrows) at 09 UTC 16.03.2017 (left panel) and 24 hours later at 09 UTC 17.03.2017 (right panel) for member 0 (deterministic) run of the meteorological ensemble.

Source term of the WN-case.

The WN-case simulates releases (total release of $5.41 \cdot 10^{17}$ Bq) from a hypothetical loss of coolant accident (LOCA) on a marine object (e.g., the KLT40S floating power plant) with an initial explosion at 09 UTC 16.03.2017, and then a subsequent fire. The number of different radionuclides, the particle size distribution, particle density and the total released activity for each particle are given in Table 10. A particle size distribution, corresponding to a log-normal distribution, is assumed for UO_2, U_3O_8 and RuO_2 (RuOxid), while a single size is assumed for ^{137}Cs, ^{134}Cs, ^{144}Ce, fly ash, and ^{131}I. It is further assumed that the emission from the accident lasts for seven hours. During the first hour the emissions are evenly distributed between 20 m and 500 m above sea level. This assumption accounts for the plume rise during the first most intensive part of the explosion and fire. After this, it is assumed that the emissions are evenly distributed between 20 m and 100 m. We apply five different temporal scenarios of the emissions with the constraint that the total integrated release of activity is the same for all five scenarios. The scenarios are:

(1) Constant release per hour during 7 hours

(2) 25 % of the emissions released during the first hour 75 % during the last 6 hours

(3) 50 % of the emissions released during the first hour 50 % during the last 6 hours

(4) 75 % of the emissions released during the first hour 25 % during the last 6 hours

(5) 90 % of the emissions released during the first hour 10 % during the last 6 hours.

The fractional releases for each hour and for the 5 scenarios are shown in Figure 26.

Table 10. Particle density, size distribution and total release during the 7 hours for each type of radioactive particle. The UO_2 and U_3O_8 particles contain the isotopes ^{144}Ce, ^{137}Cs, ^{134}Cs, ^{90}Sr, ^{89}Sr, ^{91}Y, ^{95}Zr, ^{95}NB and ^{238}Pu.

Nuclide	Density	0.5 µm	1 µm	5 µm	10 µm	20 µm	50 µm	Total released activity (Bq)
UO_2	10.96 g/cm^3	6 %	21 %	23 %	18 %	17 %	15 %	$1.9523 \cdot 10^{17}$
U_3O_8	8.3 g/cm^3	6 %	21 %	23 %	18 %	17 %	15 %	$1.9523 \cdot 10^{17}$
^{137}Cs and ^{134}Cs	2.3 g/cm^3	100 %						$9.5 \cdot 10^{15}$
^{144}Ce	7.2 g/cm^3	100 %						$1 \cdot 10^{16}$
RuOxid	7.0 g/cm^3	6 %	21 %	23 %	18 %	17 %	15 %	$6 \cdot 10^{16}$
Fly ash	2.3 g/cm^3			100 %				$6.95 \cdot 10^{16}$
^{131}I (Particle and gas phase)	2.3 g/cm^3 (Particles), 0.0113 g/cm^3 (Gas)	100 % (Particles)						$1.64 \cdot 10^{15}$
Sum								$5.41 \cdot 10^{17}$

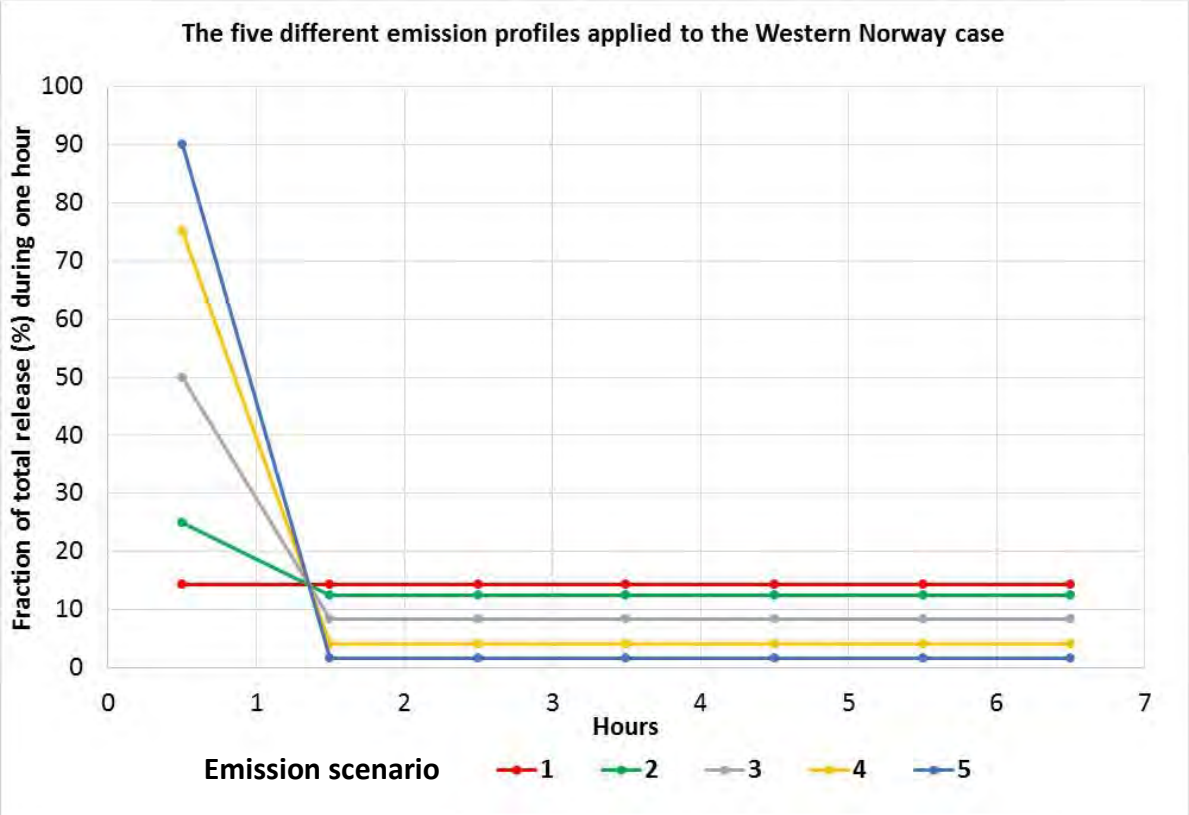


Figure 26. Fraction (in percent) of the total released activity for the 5 different emission scenarios.

The dispersion models DIPCOT and SNAP

The DIPCOT model is a Lagrangian particle/puff model. In the present runs the puff version is applied in which the concentrations in the x, y and z directions are distributed according to Gaussian distribution functions from the center of the puff (see Andronopoulos et al., 2009, for further explanations). The model calculates the displacement of the puffs by the mean wind during a time step from the 3-D velocity field. Horizontal and vertical diffusion are calculated by random velocity increments and the Gaussian distribution functions. The local concentration in a point is found by summing the contribution from all puffs within a selected radius of influence where most of the puff-load is included (99.98 %). The model outputs concentrations at 1 m height.

The SNAP model is a Lagrangian particle model. The model carries particles with the 3-D velocity field in the same way as the puffs in the DIPCOT model. However, the horizontal and vertical diffusion are calculated from a random walk method in which different turbulent length scales are applied for horizontal and vertical diffusion (see Bartnicki et al., 2011). Local concentrations are found by averaging over all particles within a grid-volume. The model outputs concentrations at ca. 30 m height, which corresponds to the second lowest model level.

In Table 11 and Table 12 some key characteristics and parameterizations of the two models are presented. We see that dose calculations are only available from the DIPCOT model. In addition to the differences in handling the vertical and horizontal diffusion (see Table 12), we also note that in-cloud scavenging is not included in DIPCOT, while SNAP includes in-cloud scavenging below approximately 2.5 km. The particle size distributions are also handled differently in the two models.

Table 11. Model type, input and output data.

Model name	Model type	MET data	Source	Model output	Dose calculations
SNAP	Lagrangian particle	Harmonie 2.5 km, 10 EPS members	5 different release scenarios	1. Activity concentration in air (30 m above surface) 2. Time-integrated activity concentration in air (30 m above surface) 3. Dry deposition (rate and accumulated) 4. Wet deposition (rate and accumulated)	No
DIPCOT	Lagrangian puff	Harmonie 2.5 km, 10 EPS members	5 different release scenarios	1. Activity concentration in air (1 m above surface) 2. Time-integrated activity concentration in air (1 m above surface) 3. Dry deposition (rate and accumulated) 4. Wet deposition (rate and accumulated)	1. Gamma radiation dose-rate in air from anisotropic cloud and ground 2. Gamma radiation dose- in air from anisotropic cloud and ground

Table 12. Model parameterizations.

Model name	Horizontal diffusion	Vertical diffusion	Dry deposition	In-cloud scavenging	Sub-cloud scavenging	Particle size treatment
SNAP	Random walk, horizontal diffusion is characterized by a horizontal length scale. Different values above and below Z_{PBL}.	Random walk, vertical diffusion characterized by a vertical length scale. Different values above and below Z_{PBL}.	Particles: Constant dry deposition velocity for r< 10 um. For r > 10 um gravitational settling is included. Constant dry deposition velocity for gases and particles smaller than r=0.05 um.	For stratiform clouds: Scavenging coefficients as for sub-cloud scavenging. For convective clouds: Scavenging independent of particle size. In-clouds scavenging only consider below model level = 0.67 (approx.. 2.5 km). The probability of being hit by precipitation in a grid-square is included.	Scavenging coefficients dependent on particle size and precipitation intensity. The probability of being hit by precipitation in a grid-square is included.	Separate runs for the different sizes of Table 10 (optional)
DIPCOT	Random velocity increments of puff movement (Langevin equation with parameters depending on atmospheric stability) and increasing $\sigma_{x,y}$ of puffs	Random velocity increments of puff movement (Langevin equation with parameters depending on atmospheric stability) and increasing σ_z of puffs	Particles: dry deposition velocity calculated as function of particle diameter, mass density and ground roughness Gas: as particles but with very small density and diameter	No	Wet deposition velocity calculated as function of particle size and rain intensity (empirical relationship)	All sizes in the same run; Each nuclide has been distributed in groups according to Table 10 and calculated separately; In post-processing groups of the same nuclide have been aggregated.

Results

In the analyses of the results of the WN-case, special attention is given to a small Village, Vikedal, in the County Rogaland at the western side of on the Norwegian mountains (see Figure 24). At Vikedal there is large agricultural and marine (fish farm) production, and the location has been selected for studying risks and uncertainties of radionuclides in both terrestrial and aquatic systems. In addition to Vikedal, we also pay special attention to the source area.

Uncertainties in the meteorological data

A literature review on the importance of meteorological parameters on the uncertainty in dispersion calculations was carried out in WP1-D9.1 of Confidence (see Wellings and Bedwell, 2018). The three highest ranked meteorological parameters contributing to uncertainty were wind direction, wind speed and precipitation. Parameters such as atmospheric stability and mixing height had a lower ranking, but it was noted that these parameters could be important especially in the early phase of the dispersion. In their study it was underlined that the importance of the meteorological parameters for the uncertainty could vary largely depending on the weather situation.

We have used the 10 ensemble members to calculate the probability distribution of the meteorological parameters. In Figure 27, we present an example of the spatial probability distributions of exceeding 1 mm and 20 mm precipitation during the 24 h period from 09 UTC 16.03.2017 to 09 UTC 17.03.2017 since the wet scavenging of particles is important for the deposition of radioactive nuclides in western Norway during this event. From the left panel we note that a probability larger than 90 % of at least 1 mm (orange color) covers a large area. Also, areas with less than 10 % probability for 1 mm or more covers a large area, while the area with a probability between 10 and 90 % is rather small. From the right panel we see a very distinct area in western Norway with high probability (above 90 %) of receiving more than 20 mm in 24 hours. Outside this region, the probability is small (less than 10 %) of receiving more than 20 mm. For this particular case, the meteorological ensembles agree rather well on the spatial extent of accumulated precipitation during the 24 hours period.

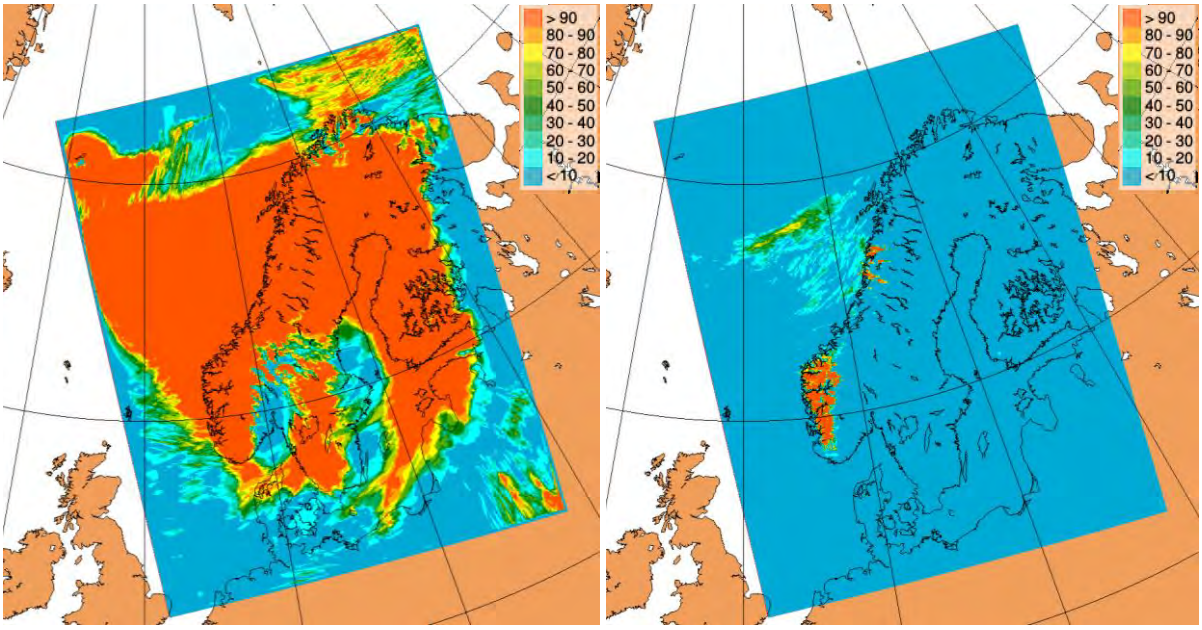


Figure 27. Probability of more than 1 mm precipitation in 24 hours (left panel) and 20 mm (right panel). Calculations for the period 09 UTC 16.03.2017 to 09 UTC 17.03.2017.

In the following, time-series of wind, precipitation, surface heat fluxes and boundary layer heights are given for all 10 EPS members at the location of the source and at Vikedal. The values are selected from the nearest grid-square in the meteorological model representing averages over an area of 2.5 km • 2.5 km. The figures also show the 10, 50 and 90 percentiles of the 10 EPS members. Note that the time-series start at 06 UTC, three hours before the emission incident.

The wind direction (Figure 28) at the lowest model level (ca. 10 m height) vary with approximately 5 degrees or less between the 10th and 90th percentile during most of the simulation at the source. One exception is the first hours for which the spread is larger and up to 20-30 degrees. At Vikedal the spread (difference between 10th and 90th percentile) in the wind directions is also small up to about 18 UTC. After this time the spread is increasing to ca. 30 degrees at the end of the run. As shown in Figure 25, the low pressure system is moving eastward quite far north of the emission area, implying a relatively uniform weather situation in the vicinity of the source which is commensurate with a rather small variability in the horizontal wind directions. At Vikedal, the larger variability after about 18 UTC may be due to the complex local terrain features giving rise to a larger sensitivity compared to the offshore source area.

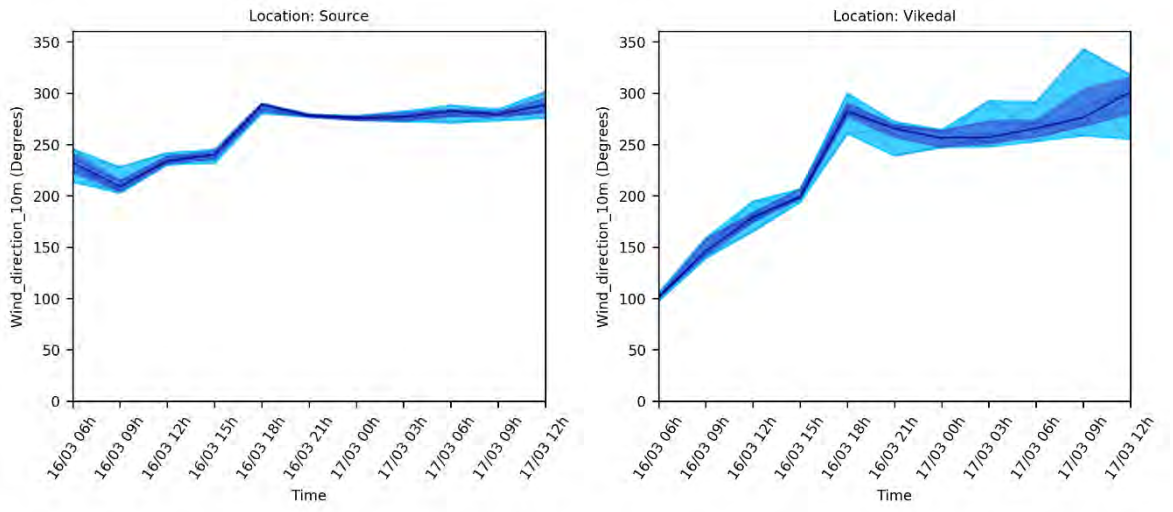


Figure 28. Time-series of wind direction (ca. 10 m) for the 10 EPS runs. Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile. Results for the location of the source (left panel) and Vikedal (right panel).

Figure 29 shows the wind speed at ca. 10 m height. In the period from the start of the release at 09 UTC until 12 UTC the differences in wind speed between the 10th and 90th percentiles are less than 10 % at the location of the source. After 12 UTC this difference is increasing to ca. 30 % toward the end of the release. At Vikedal we also note small spread in the wind speed until 12 UTC. After this the difference between the 10th and 90th percentiles increases and vary between a factor of about 1.5 to 2.0 for the rest of the forecasting period.

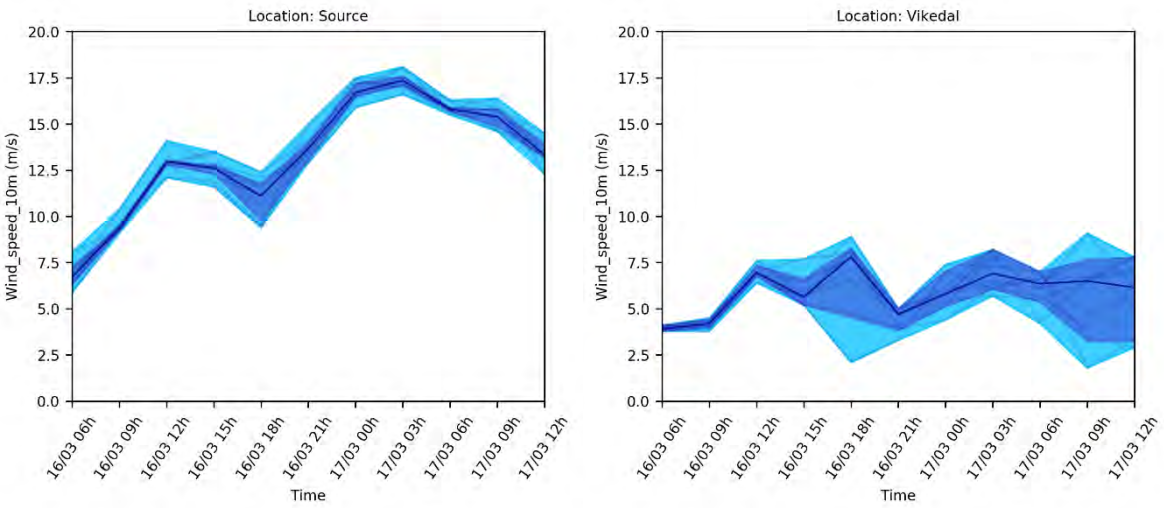


Figure 29. Time-series of wind speed (ca. 10 m) for the 10 EPS runs. Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile. Results for the location of the source (left panel) and Vikedal (right panel).

Only small precipitation amounts are encountered at the source during the fire (see Figure 30). The 3 hours accumulated precipitation has a maximum at 18 UTC 16.03.2019. At this time the 10th percentile is close to zero while the 90th percentile is about 1 mm. The reference member 0 is close to the 90 precentile level (not shown). Vikedal receives more precipitation and the 3 hours accumulated precipitation peaks at 18 UTC. The ensemble members are consistent in both the onset and duration of the precipitation at Vikedal. However, a factor of about two difference is seen between the 10th and 90th percentiles at 18 UTC.

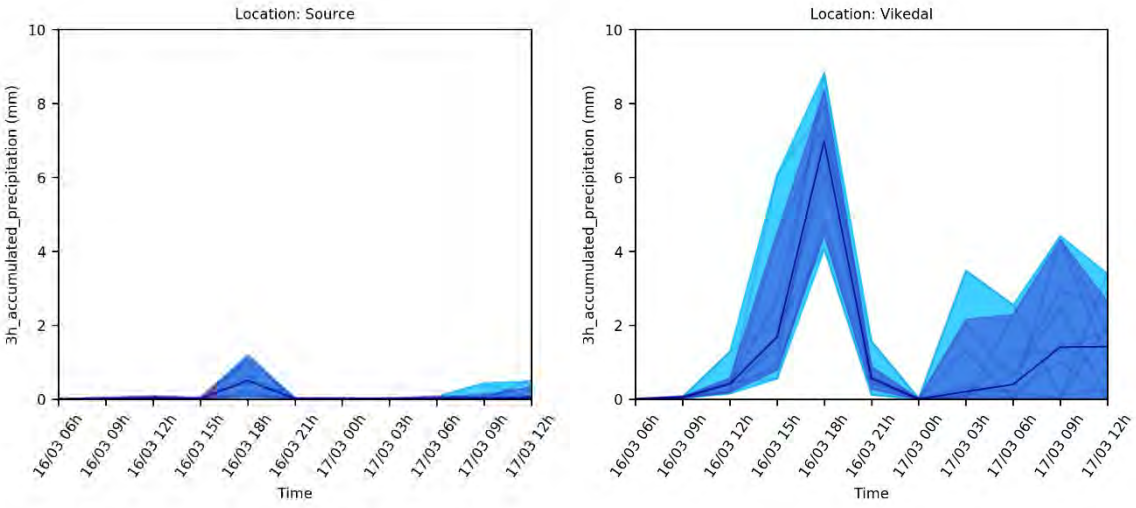


Figure 30. Time-series of 3 h accumulated precipitation for the 10 EPS runs. Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile. Results for the location of the source (left panel) and Vikedal (right panel).

Figure 31 presents the vertical eddy transport of sensible heat. The ensembles are very consistent in both magnitude and direction of the heat transport (negative values represent upward heat transport while positive values represent downward transport). The difference between the 10th and 90th percentiles is about 5.0 W/m^2 during the first 12 hours at both locations. Towards the end of the

forecast the difference increases to about 30 W/m^2 at Vikedal. The increase is smaller at the offshore location of the source.

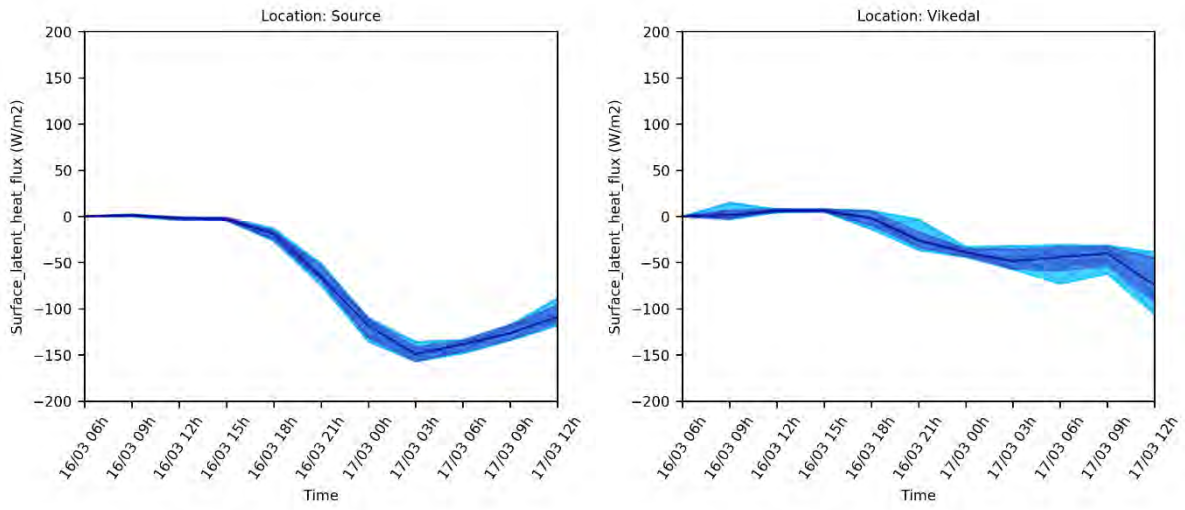


Figure 31. Time-series of vertical surface heat flux for the 10 EPS runs. Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile. A negative value means upward directed heat transport, while a positive value means downward transport. Results for the location of the source (left panel) and Vikedal (right panel).

The boundary layer height (BLH) is calculated according to Bartnicki et al. (2011). The method searches in the vertical for the level of a critical value (=1.8) of the gradient Richardson number. The height between the surface and this level becomes the BLH. Note that the BLH is not allowed to be smaller than 200 m. BLH (Figure 32) shows a larger uncertainty than the other parameters presented above. The difference between the 10th and 90th percentiles is varying from ca. 100 m to ca. 400 m. The uncertainty shows little dependence on forecast length and we also note that the uncertainty is, during several periods of the 24 hour forecast, larger at the offshore location of the source compared to Vikedal. This is in contrast to the other parameters, which become more uncertain with increasing forecast time, and more uncertain at Vikedal compared to source.

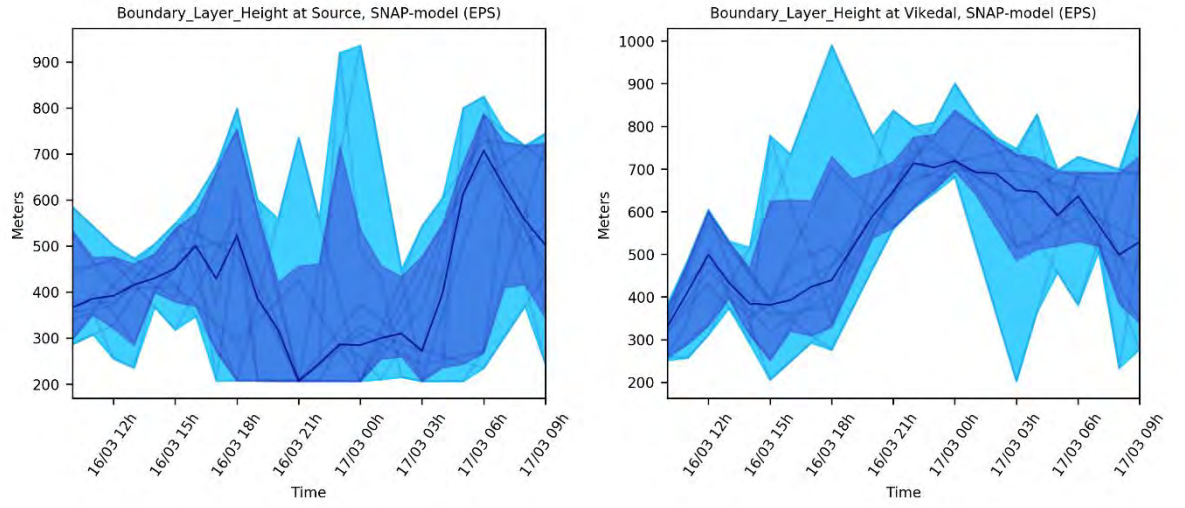


Figure 32. Time-series of boundary layer height (as estimated according to Bartnicki et al, 2011) for the 10 EPS runs. Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile. Results for the location of the source (left panel) and Vikedal (right panel).

Results from the dispersion model runs

This section presents the uncertainties in the dispersion of the radionuclides estimated from the ensemble meteorology, variations in the source term and applications of the two different dispersion models DIPCOT and SNAP.

Base run of the dispersion of the plume from the explosion

First we view the dispersion of the plume from the base run of the fire (Figure 33). The results are based on the deterministic meteorological ensemble (member 0) and emission scenario (3) in which 50 % of the emissions are released during the first hour and 50 % during the last 6 hours. The figure shows the concentrations of ^{137}Cs after 6 hours and after 12 hours, given as Bq/m^3 at 30 m height height above ground for SNAP and at 1 m height for DIPCOT. From the figures we see that the plume is first heading northeast then more northerly along western Norway during the first 6 hours. Between 6 and 12 hours the plume moves more in the easterly direction.

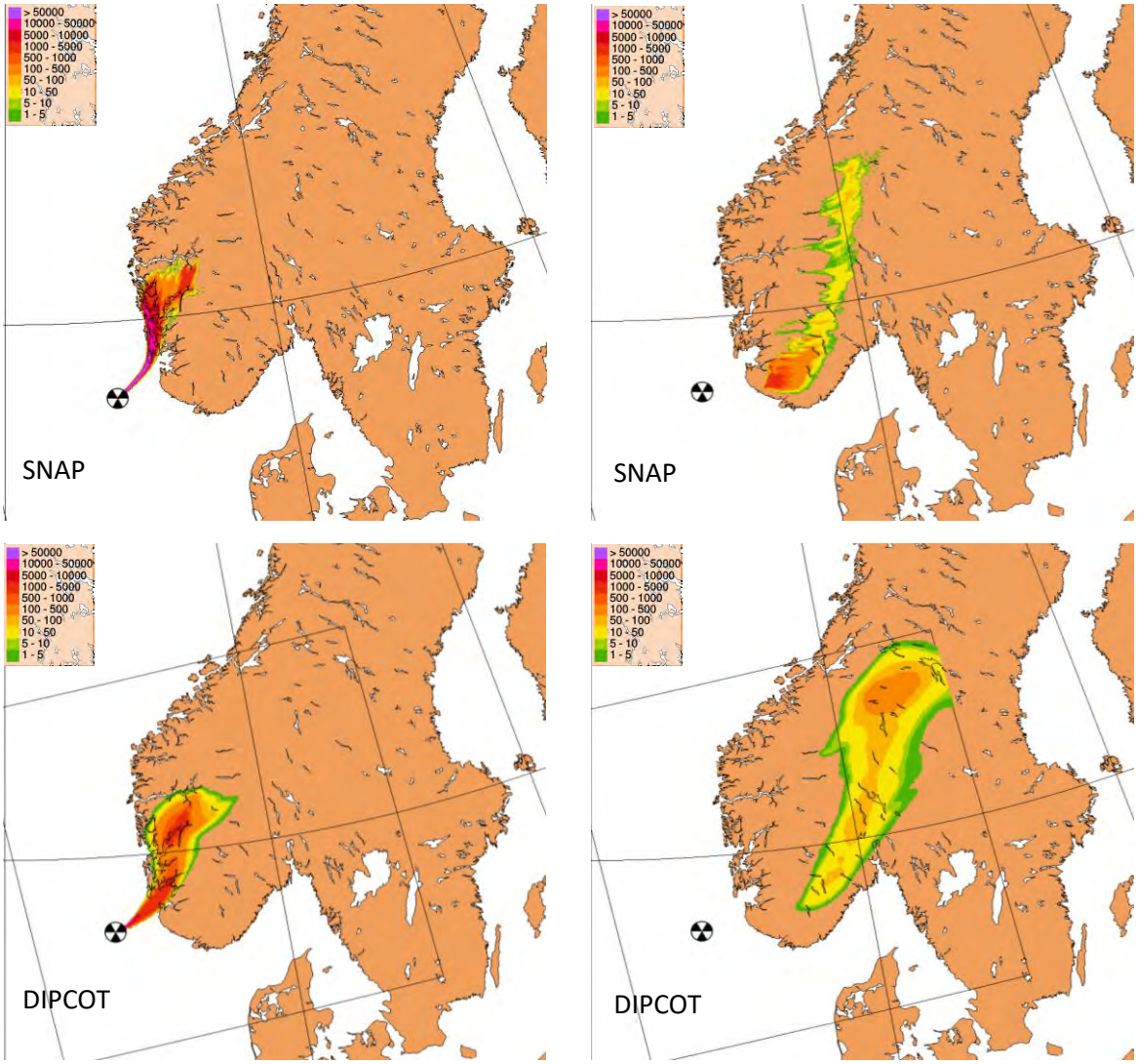


Figure 33. Location of the plume after 6 hours and 12 hours. Concentrations given as activity of Cs137 (Bq/m^3. Upper panel SNAP model and lower panel DIPCOT model. Dispersion run based on EPS member 0, and emission scenario 3.

The SNAP model gives a narrower plume with higher maximum concentrations near the source after 6 hours compared to DIPCOT. The width of the SNAP surface plume is approximately 10 km at the border between the sea and the land area after 6 hours, while the width of the DIPCOT plume is

approximately twice (20 km). The SNAP surface plume also has a slightly more northerly path compared to the DIPCOT plume in the area between the source and the coastline.

After 12 hours the concentrations are much lower in both model. The shape of the plumes are qualitatively similar, however, the DIPCOT surface plume advances faster toward east and north than the SNAP plume. We see that after 12 hours SNAP has the maximum surface concentrations in southwest Norway (east of Stavanger), while DIPCOT has the maximum in central Norway and Sweden approximately 600 km northeast of the SNAP maximum. Some of the differences in the concentrations may be due to the difference in height level of the concentrations from the two models (SNAP 30 m and DIPCOT 1 m). Close to the source this may give rise to substantial differences. A detailed analysis of both vertical and horizontal transport and concentration patterns together with the deposition fields, is needed to fully understand the differences between the two models.

The differences in the plumes calculated with the two models are also seen when the plume passes Vikedal (Figure 34). The plume arrives about three hours earlier in DIPCOT and lasts for a longer period compared to SNAP. However, the peak concentration of the 50th percentile is about a factor 4 lower in the DIPCOT model.

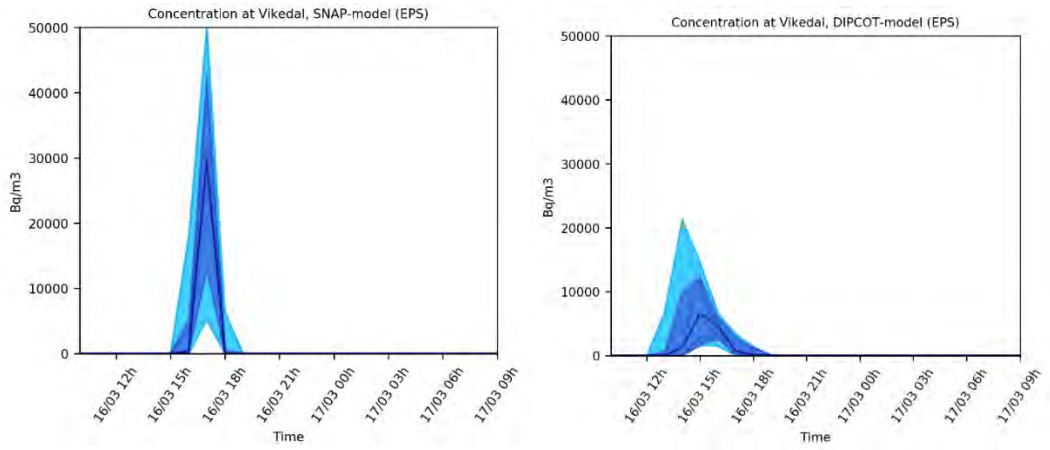


Figure 34. Temporal evolution of the concentration levels at Vikedal calculated by the SNAP model (left) and the DIPCOT model (right) for the 10 EPS members and emission scenario (3). Light blue area encompasses the max to min values, dark blue areas the 10th percentile to 90th percentile, while the thick solid line represents the 50th percentile.

Maps of uncertainty in accumulated concentrations

In this section we present results of total activity from all radionuclides in accumulated surface concentration and the probability of exceeding threshold levels in accumulated surface concentration. The maps are presented for the three different cases: Case 1: Reference emission scenario (3) and 10 EPS members, Case 2: Reference EPS member 0 and the 5 emission scenarios and Case 3: Combining all EPS members and emission scenarios.

Case 1: Emission scenario (3) and 10 EPS members

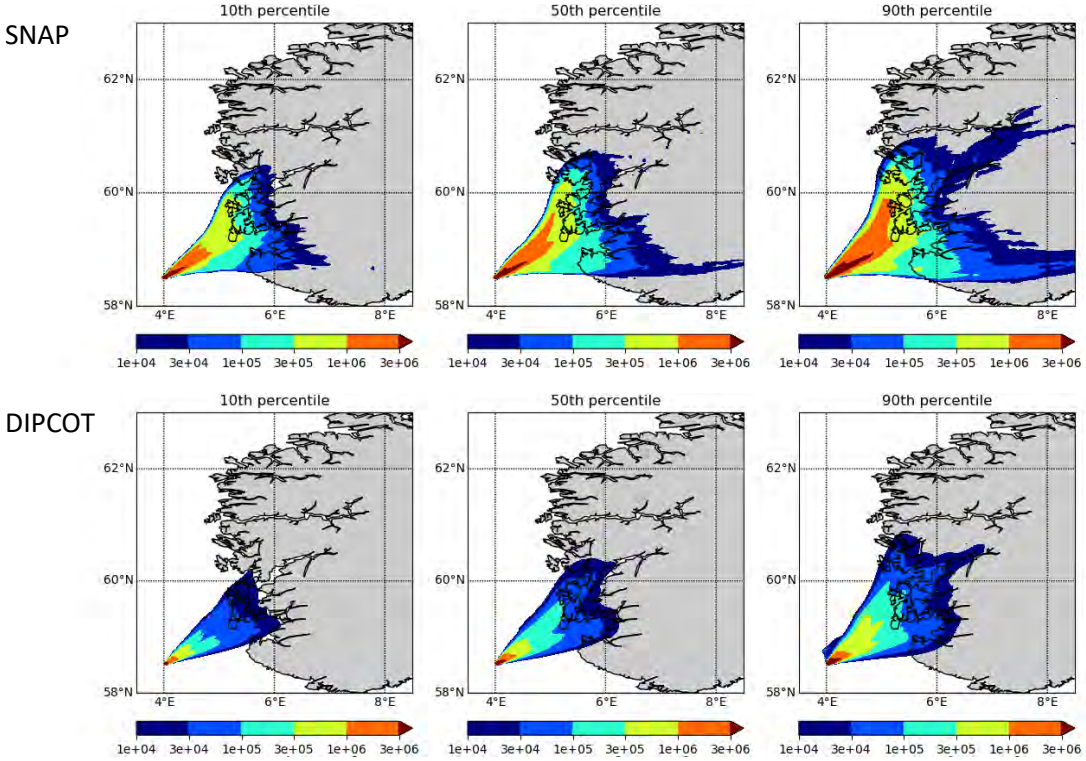
In Figure 35 are shown the 10th, 50th and 90th percentile of accumulated concentration (two upper panels) and the probability of exceeding 10^4 Bq·hr/m^3, 10^5 Bq·hr/m^3 and 10^6 Bq·hr/m^3 (two lower panels) for the 24 hours period, based on the 10 EPS members and the reference emission scenario (3). Qualitatively the affected areas are similar in the two models, but the SNAP model yields higher accumulated concentrations than the DIPOT model. The areas with probability of exceeding 10^4, 10^5 and 10^6 Bq.hr/m^3 are therefore also larger in SNAP compared to DIPCOT.

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios

The results are presented in Figure 36. Both for SNAP and DIPCOT the areas encompassed by the percentiles and the probabilities like case 1 (Figure 35). Some smaller differences are seen, especially for the low concentrations, which are extended further to the east in DIPCOT for this case compared to case 1. We see that the spread between 10th and 90th percentiles also are quite similar in case 1 and 2 for both models.

Case 3: 10 EPS members and 5 emission scenarios.

Figure 37 shows results from a combination of all EPS members with the 5 emission scenarios which yields 50 different model runs. The results differ little from case 1 and 2, but we note that the 10th percentile concentrations are a little lower while the 90th percentile concentrations are a little higher in SNAP compared to case 1 and 2. The 10th percentile concentrations are a little lower also for DIPCOT, while the 90 percentile concentrations are lower than in case 2. DIPCOT case 1 and 3 yield very similar results.



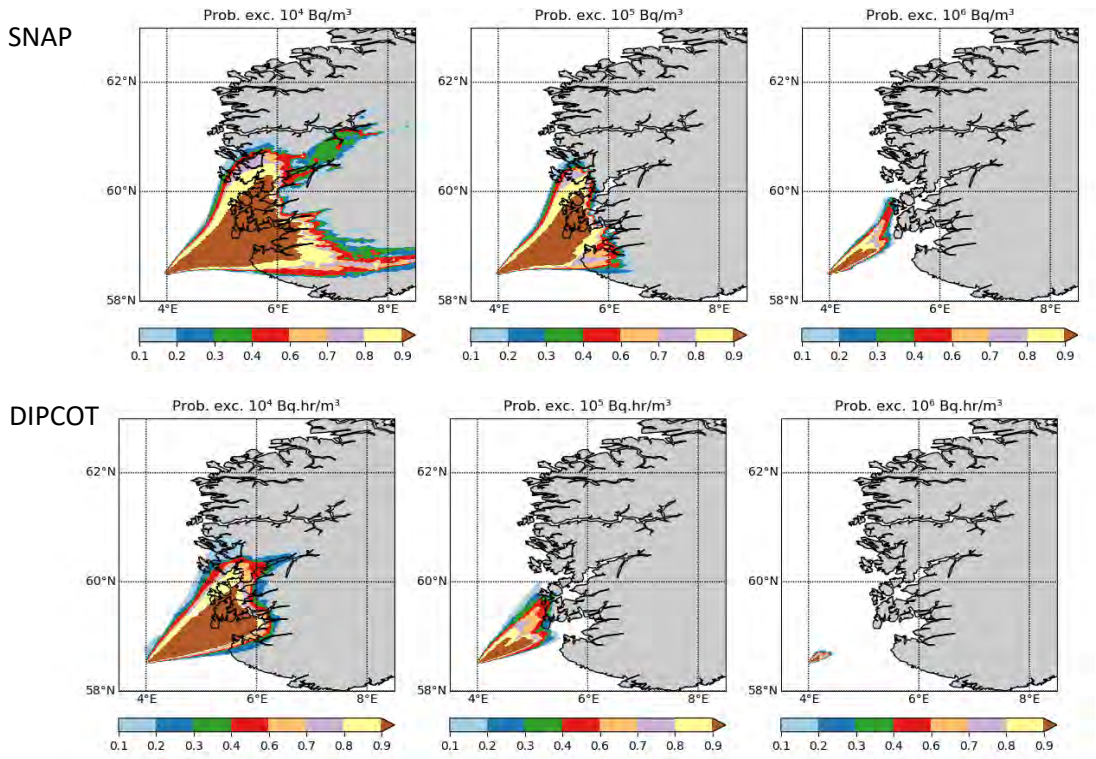
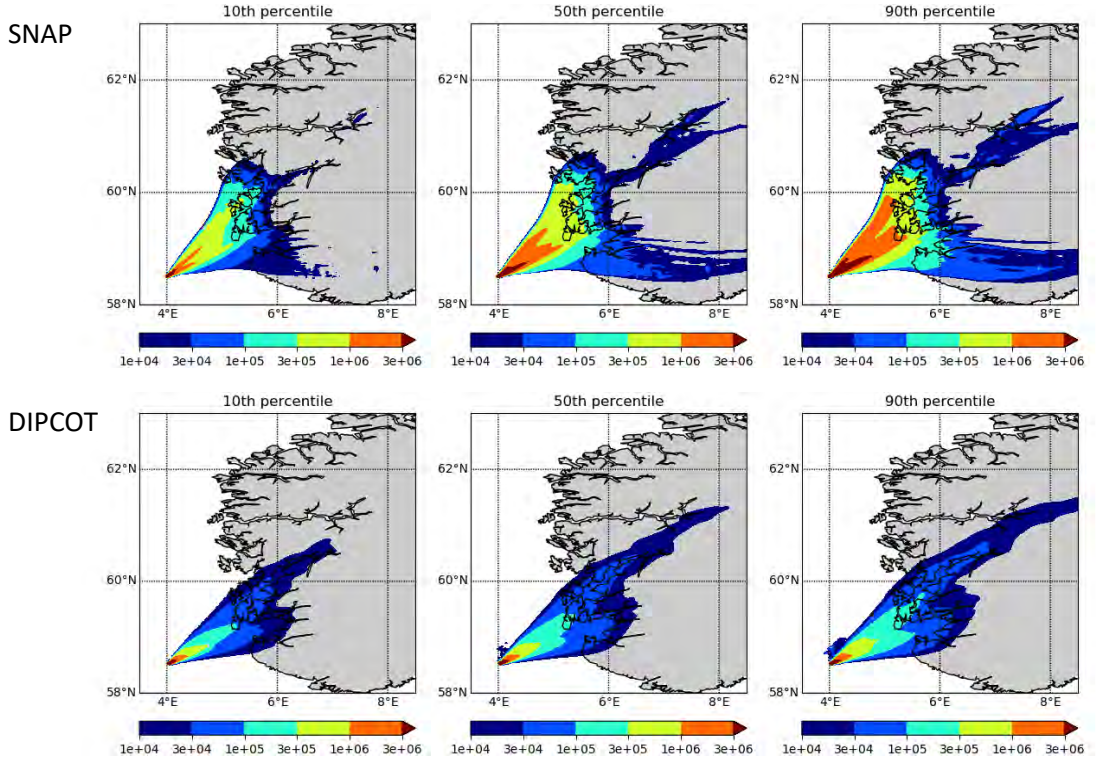


Figure 35. 10 EPS members and emission scenario (3). 10th, 50th and 90th percentile of accumulated concentrations in air (Bq.hr/m³) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq.hr/m³, 10^5 Bq·hr/m³ and 10^6 Bq·hr/m³ of accumulated concentration (two lower panels). Data for the 24 hours forecast period.



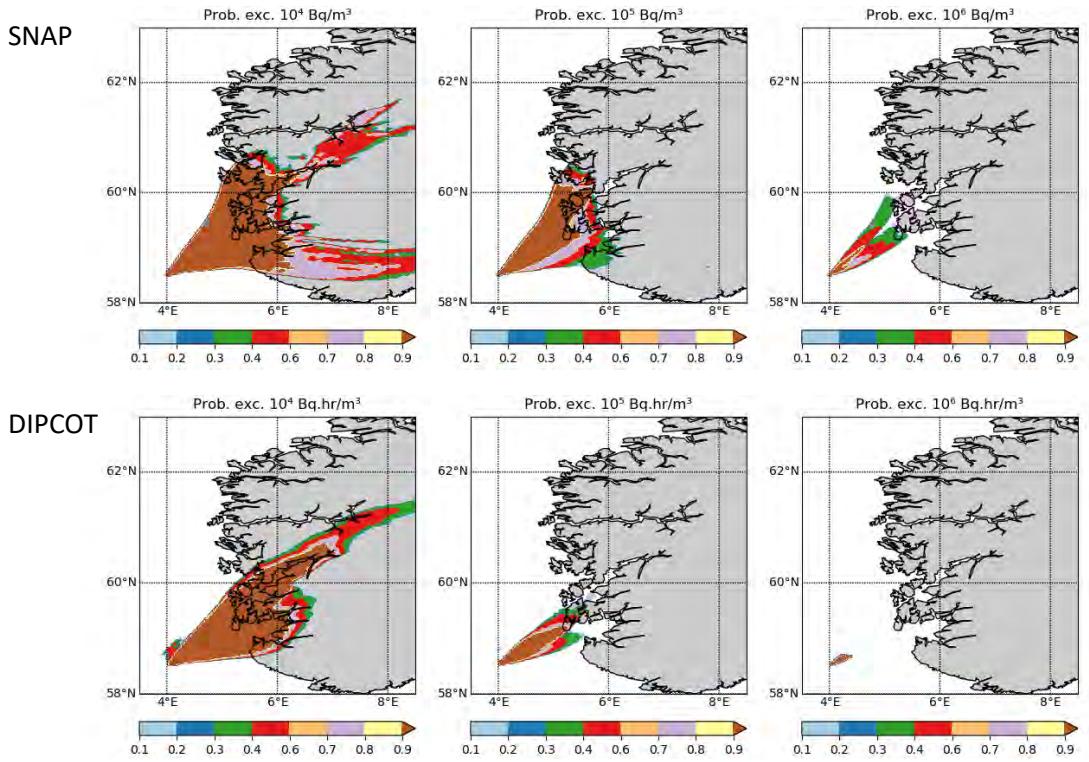
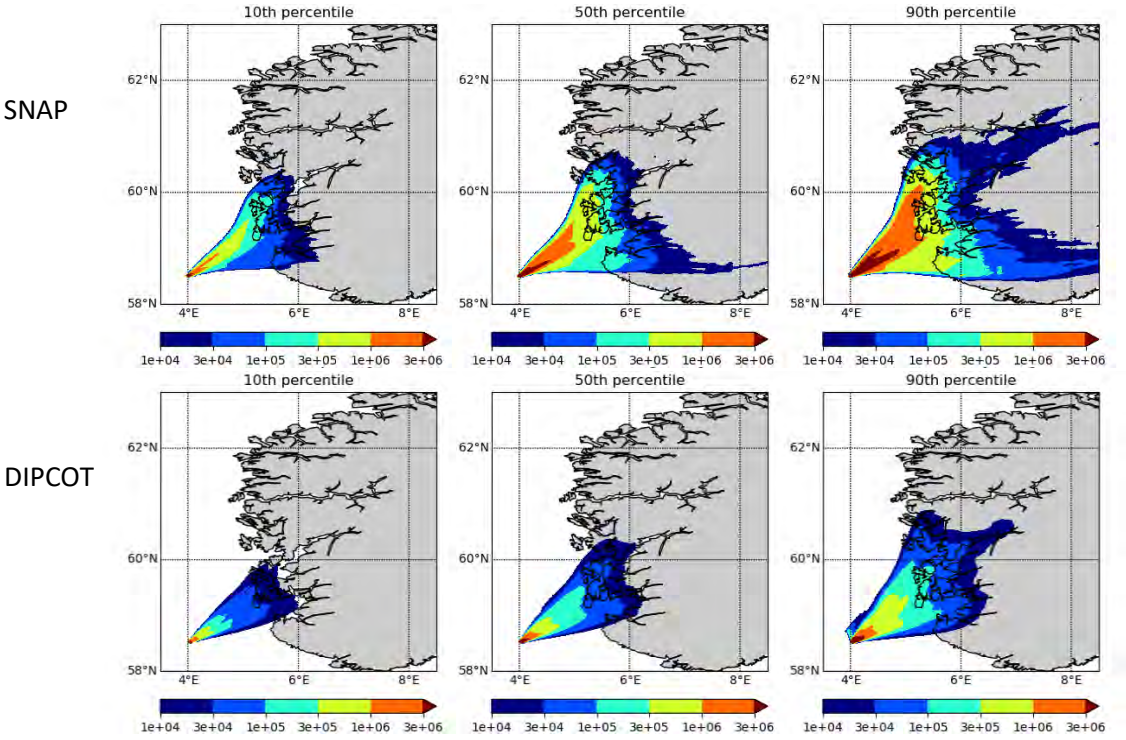


Figure 36. EPS member 0 (deterministic) and 5 emission scenarios. 10th, 50th and 90th percentile of accumulated concentrations in air (Bq.hr/m³) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq·hr/m³, 10^5 Bq·hr/m³ and 10^6 Bq.hr/m³ of accumulated concentration (two lower panels). Data for the 24 hours forecast period.



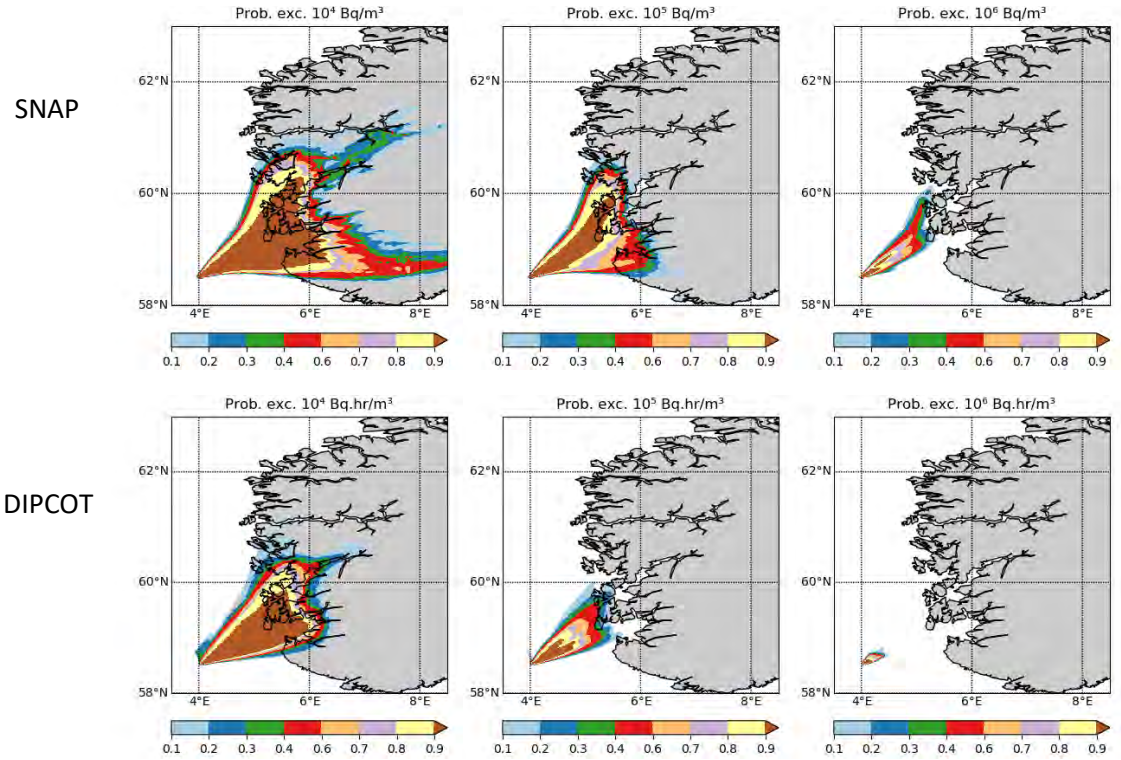


Figure 37. 10 EPS member and 5 emission scenarios. 10th, 50th and 90th percentile of accumulated concentrations in air (Bq.hr/m³) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq·hr/m³, 10^5 Bq·hr/m³ and 10^6 Bq·hr/m³ of accumulated concentration (two lower panels). Data for the 24 hours forecast period.

Maps of uncertainty in total accumulated depositions

In this section we present results of total activity from all radionuclides in total accumulated surface deposition together with the probability of exceeding threshold levels in the total accumulated deposition.

Case 1: Emission scenario (3) and 10 EPS members

Also for accumulated depositions we focus on maps of the 10th, 50th and 90th percentiles and on probabilities of exceeding 10^4 Bq/m², 10^5 Bq/m² and 10^6 Bq/m². From the 10 EPS members combined with emission scenario (3) (Figure 38). The deposition has a maximum near the source and on the western side of the Norwegian mountains. This is more pronounced in DIPCOT than in SNAP. We also note that the deposition from DIPCOT is larger, particularly in the area of large rain amounts which coincides with the plume from the fire (see Figure 27). The area with high probabilities of exceeding both 10^5 Bq/m² and 10^6 Bq/m² is somewhat larger especially toward east in the DIPCOT model.

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios

Qualitatively, the maps showing percentiles and probabilities (Figure 39) are similar to case 1. Also in this case the DIPCOT model results in more depositions compared to the SNAP model. The differences in the 10th and 90th percentiles are also qualitatively similar for case 1 and 2 although locally differences can be seen. We also note that the spatial probability distribution is more “sharp edged” compared to case 1.

Case 3: 10 EPS members and 5 emission scenarios.

The 10th, 50th and 90th percentile maps and the probability maps (Figure 40) are very similar to case 1 for both models. The differences between the 10th and 90th percentiles of total accumulated deposition are also very similar to the two other cases.

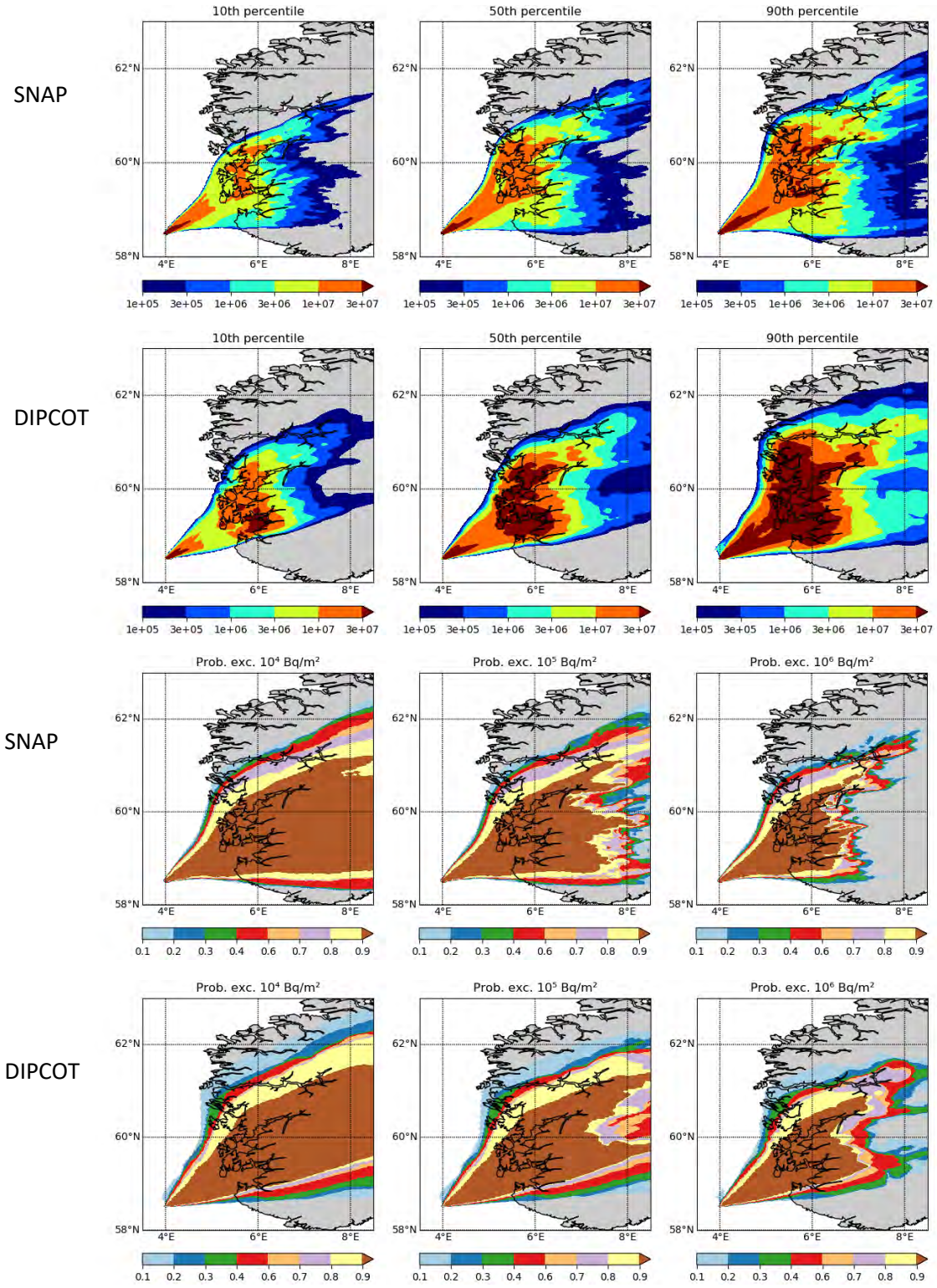


Figure 38. 10 EPS members and emission scenario (3). 10th, 50th and 90th percentile of accumulated deposition (Bq/m²) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq/m², 10^5 Bq/m² and 10^6 Bq/m² of accumulated deposition (two lower panels). Data for the 24 hours forecast period.

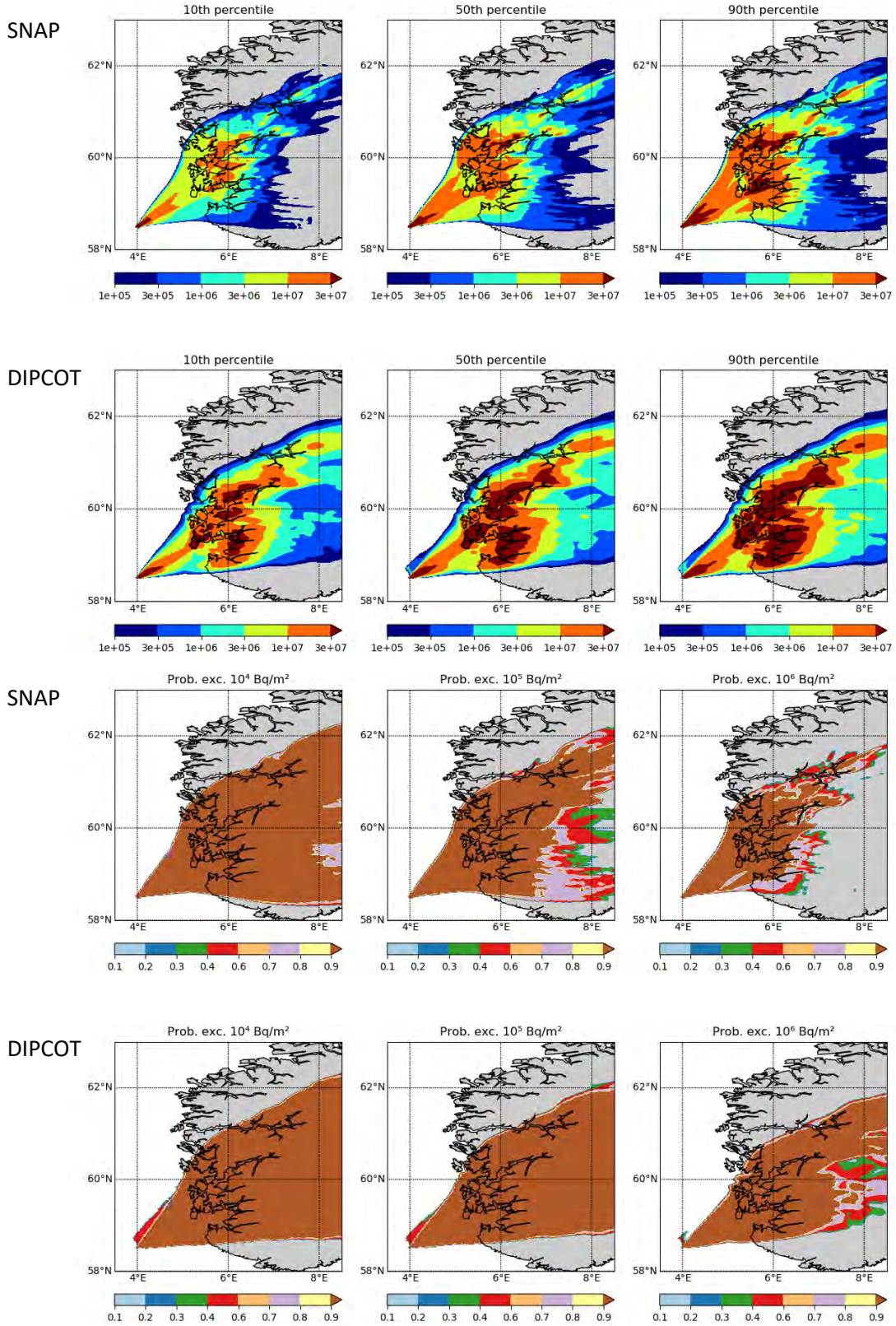


Figure 39. EPS member 0 and 5 emission scenarios. 10th, 50th and 90th percentile of accumulated deposition (Bq/m²) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq/m², 10^5 Bq/m² and 10^6 Bq/m² of accumulated deposition (two lower panels). Data for the 24 hours forecast period.

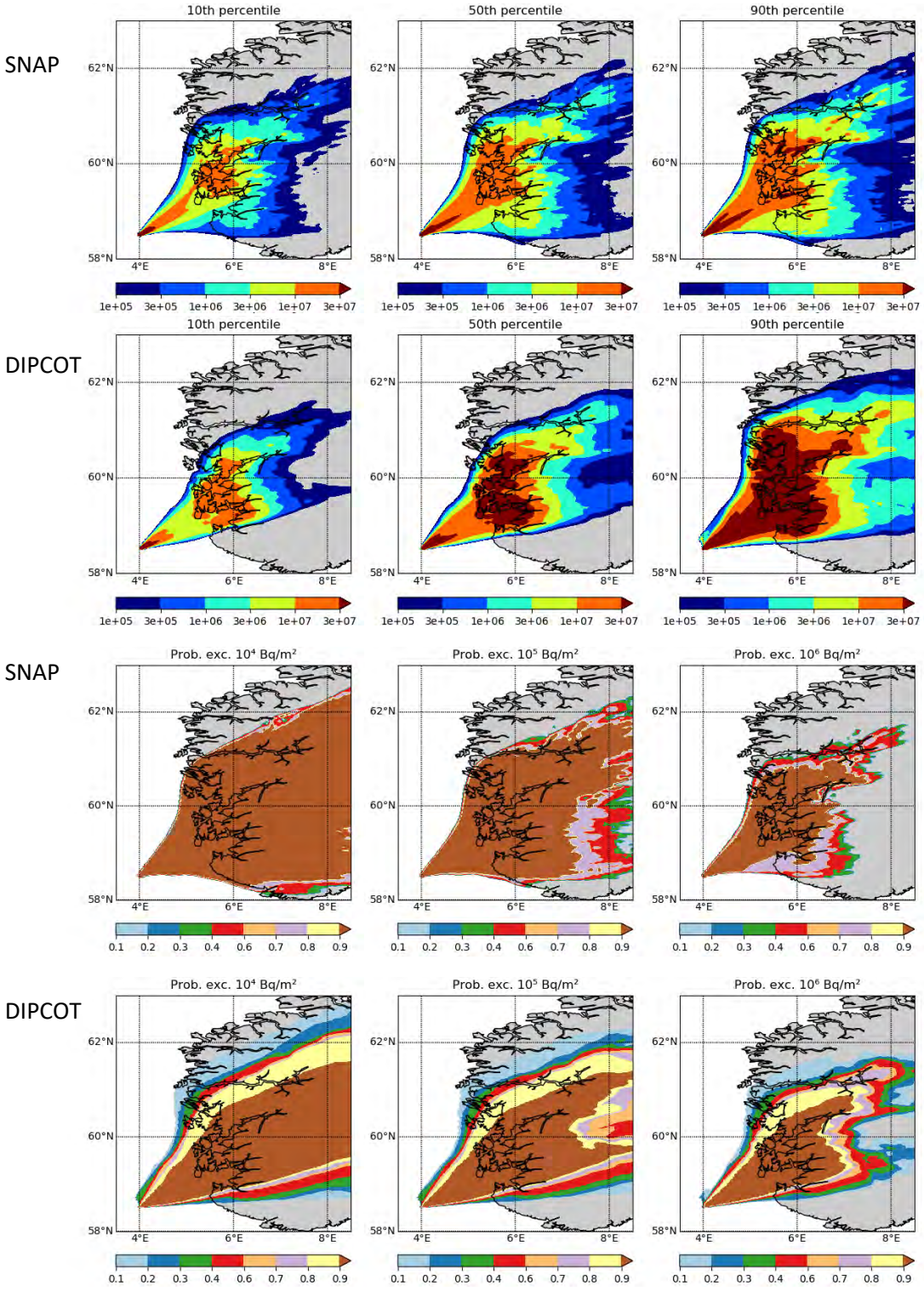


Figure 40. 10 EPS members and 5 emission scenarios. 10th, 50th and 90th percentile of accumulated deposition (Bq/m²) for SNAP and DIPCOT (two upper panels) and, probability of exceeding 10^4 Bq/m², 10^5 Bq/m² and 10^6 Bq/m² of accumulated deposition (two lower panels). Data for the 24 hours forecast period.

Time-series of total activity at the location of the source and Vikedal

In addition to the maps discussed in the previous sections we also discuss times-series of the total activity in accumulated concentrations and deposition at the location of the source and Vikedal.

Accumulated concentrations

Case 1: Emission scenario (3) and 10 EPS members

In Figure 41 and Figure 42 the temporal evolution of the accumulated concentrations of the grid-square (2.5 km·2.5 km) covering the source region is given. In SNAP (Figure 41) the accumulated concentrations increase monotonically for all ensemble members. The difference between the 10th and 90th percentiles is rather constant and approximately 0.2 Bq·hr/m^3 during the release. In percent of the 50th percentile the spread is dropping from nearly 100 % after the first hour to about 10 % after 7 hours. One of the ensembles (EPS member 6) give however quite much larger values than the other members. This may be due to the difference in wind direction for this member during the accident (see Figure 28 upper panel). A measure of the uncertainty in the accumulated concentration is the ratio between the 90th and 10th percentiles. For SNAP this number is 1.15 while for DIPCOT it is 3.5. Much larger spread in the ensemble runs is encountered in the DIPCOT simulations at the location of the source. Also, the 50th percentile accumulated concentration level is higher in DIPCOT compared to SNAP.

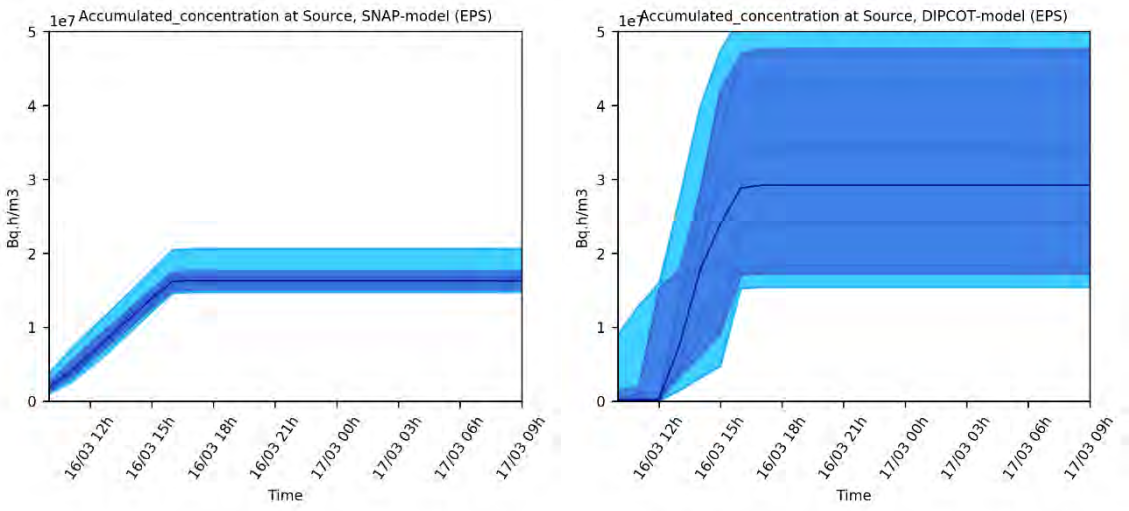


Figure 41. Time-series of accumulated concentration for the 10 EPS members at the location of the source for the SNAP model (upper panel) and the DIPCOT model (lower panel).

At Vikedal the two models behave more similarly (see Figure 42), still the differences are quite large. The plume arrives approximately 3 hours earlier in DIPCOT compared to SNAP. The ratio between the 90th and 10th percentiles is 2.3 for SNAP and 3.7 for DIPCOT, while the ratio between (SNAP 50th percentile)/(DIPCOT 50th percentile) is 2.3. The accumulated concentration level calculated from ensemble member 0 (not shown) is close to 25000 Bq·hr/m^3 in both models. I.e. if the reference meteorology and emission scenario is used the two models will give rather similar accumulated concentrations, but this should be regarded as rather coincidental in this case.

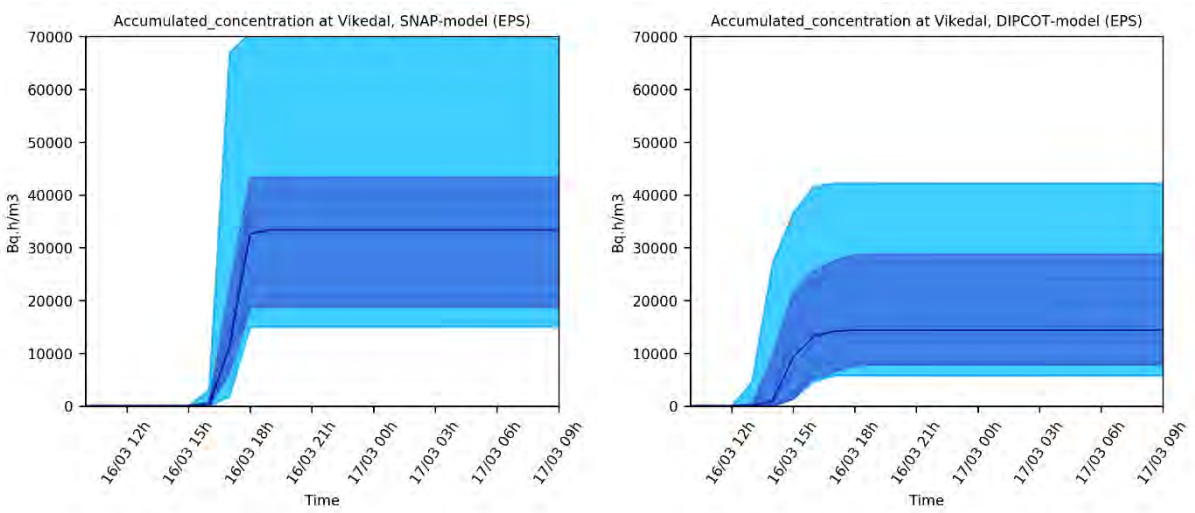


Figure 42. Time-series of accumulated concentration for the 10 EPS members at Vikedal for the SNAP model (upper panel) and the DIPCOT model (lower panel).

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios

In Figure 43 and Figure 44 results from the 5 emission scenarios are shown applying the member 0 meteorology. Although the temporal developments are quite different between the two models at the source, the final accumulated concentration levels are similar. The difference between the 10th and 90th percentiles is 5.4 in SNAP and 2.7 in DIPCOT at the end of the release. At Vikedal the difference between the 10th and 90th percentiles is 5.1 in SNAP and 2.6 in DIPCOT after the plume has passed. The largest spread due to the emission scenarios is encountered in the SNAP simulations. As noted for case 1, the reference member 0 meteorology gives very similar accumulated concentrations from the two models at Vikedal, corresponding well with the case 2 50th percentile being similar at Vikedal.

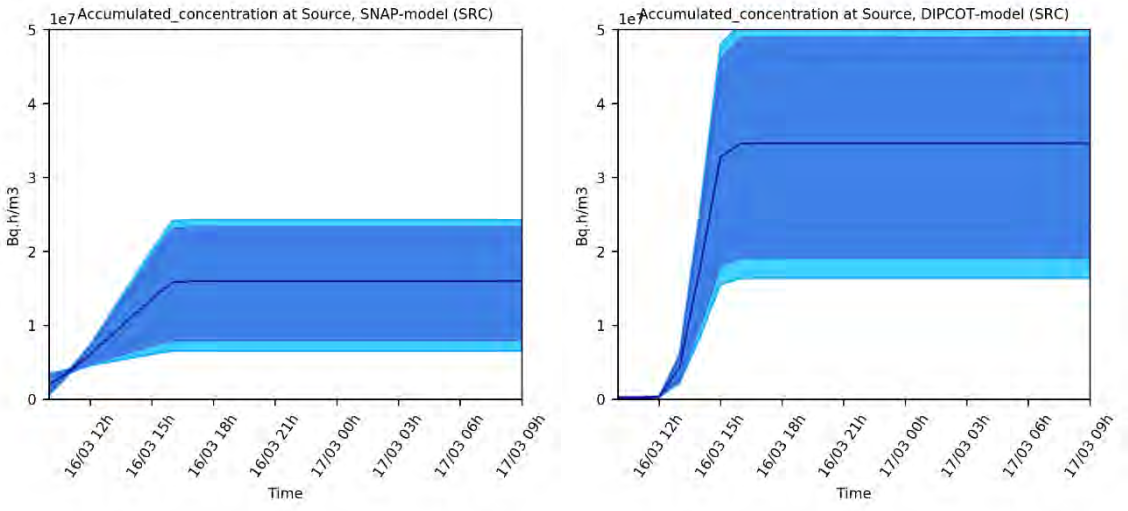


Figure 43. Time-series of accumulated concentration for the 5 emission scenarios EPS member 0 at the Source. SNAP model (left) and DIPCOT model (right).

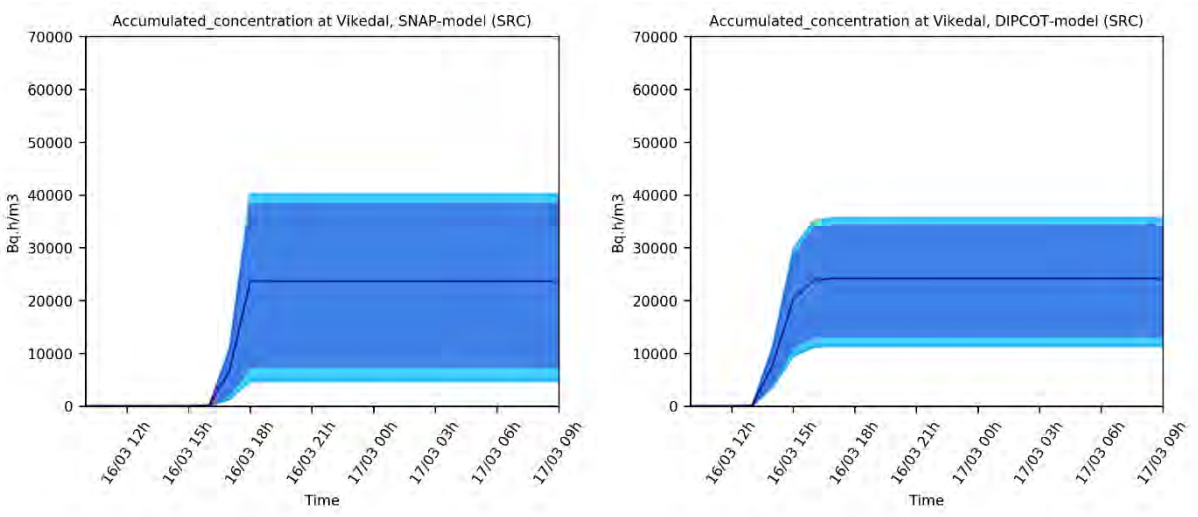


Figure 44. Time-series of accumulated concentration for the 5 emission scenarios EPS member 0 at Vikedal. SNAP model (left) and DIPCOT model (right).

Total accumulated deposition

Case 1: Emission scenario (3) and 10 EPS members

Figure 45 and Figure 46 show the time-series of this case. As for accumulated concentrations, the accumulated deposition increases monotonically at the source for all ensemble members in SNAP. The ratio of the 90th and 10th percentiles is about 1.2 at the end of the release. The corresponding number for DIPCOT is 3.5. We also note that at the source the deposition in DIPCOT is much smaller than in SNAP. The ratio (SNAP 50th percentile total deposition)/(DIPCOT 50th percentile total deposition) is 5.3 at the end of the release.

The deposition at Vikedal starts 1 hour earlier in SNAP for all except two EPS member compared to DIPCOT, even though the plume at the surface arrives later in SNAP for most members. This could be due to an earlier arrival of the plume at higher altitudes in combination with wet deposition. From

Figure 30 we see that this coincides with the precipitation taking place at Vikedal at 12 UTC. The ratio of the 90th and 10th percentiles of the total deposition is 2.9 in SNAP and 2.1 in DIPCOT. At Vikedal the accumulated deposition after the passage of the plume, is larger in DIPCOT than SNAP, the ration (SNAP 50th percentile total deposition)/(DIPCOT 50th percentile total deposition) being 0.33.

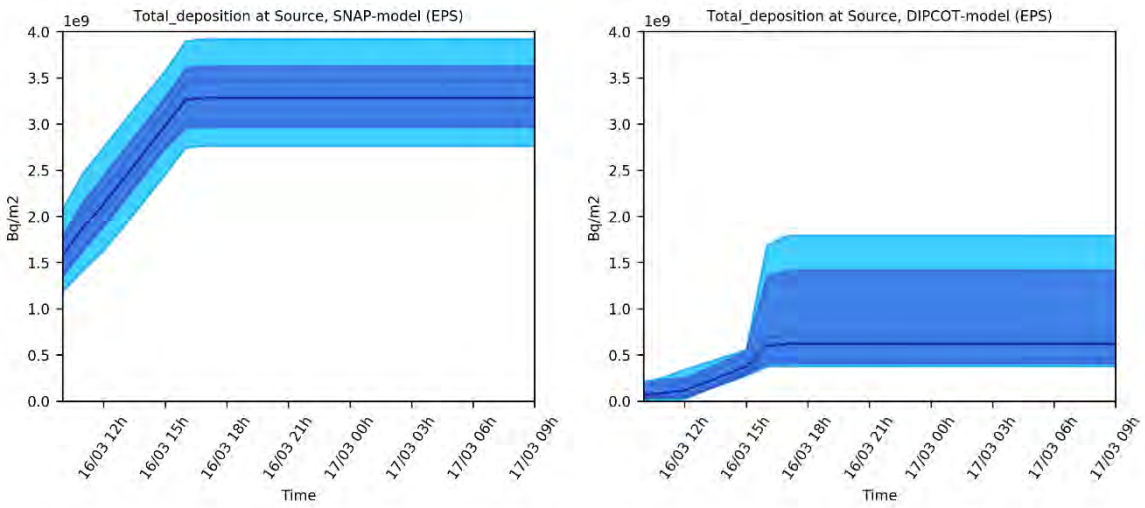


Figure 45. Time-series of accumulated deposition for the 10 EPS members at the location of the source. SNAP model (upper panel) and DIPCOT model (lower panel).

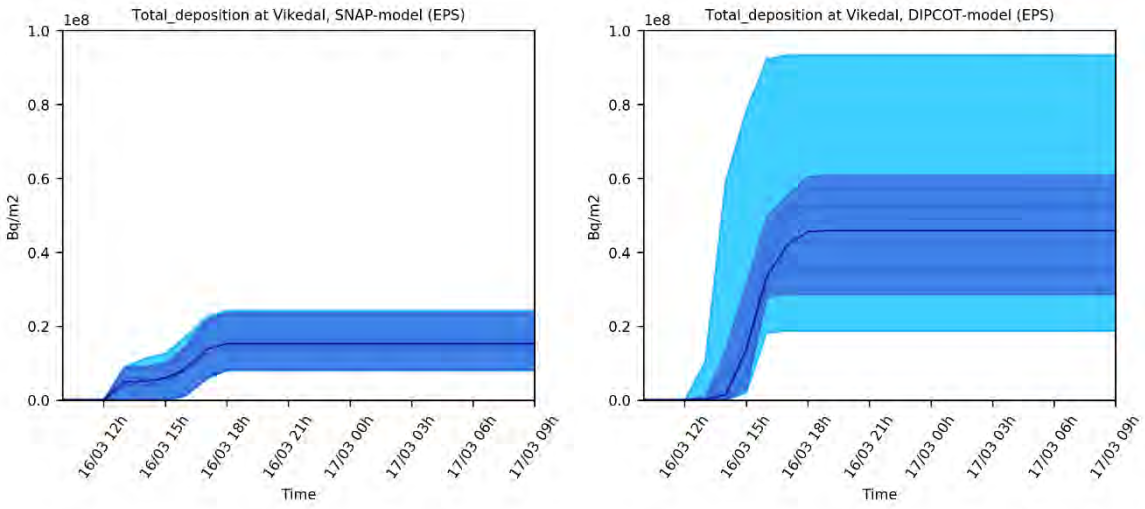


Figure 46. Time-series of accumulated deposition for the 10 EPS members at Vikedal. SNAP model (left) and DIPCOT model (right).

Case 2: EPS member 0 (deterministic) and 5 emissions scenarios

The time-series of this case are shown in Figure 47 and Figure 48. At the source the ratio of the 90th and 10th percentiles are about 1.1 for SNAP and 1.8 for DIPCOT at the end of the release. We note much higher deposition in SNAP compared to DIPCOT, the ratio (SNAP 50th percentile total deposition)/(DIPCOT 50th percentile total deposition) being 6.9. At Vikedal the ratio of the 90th and 10th percentiles are 2.1 for SNAP and 2.6 for DIPCOT. The ratio (SNAP 50th percentile total deposition)/(DIPCOT 50th percentile total deposition) is 1.3.

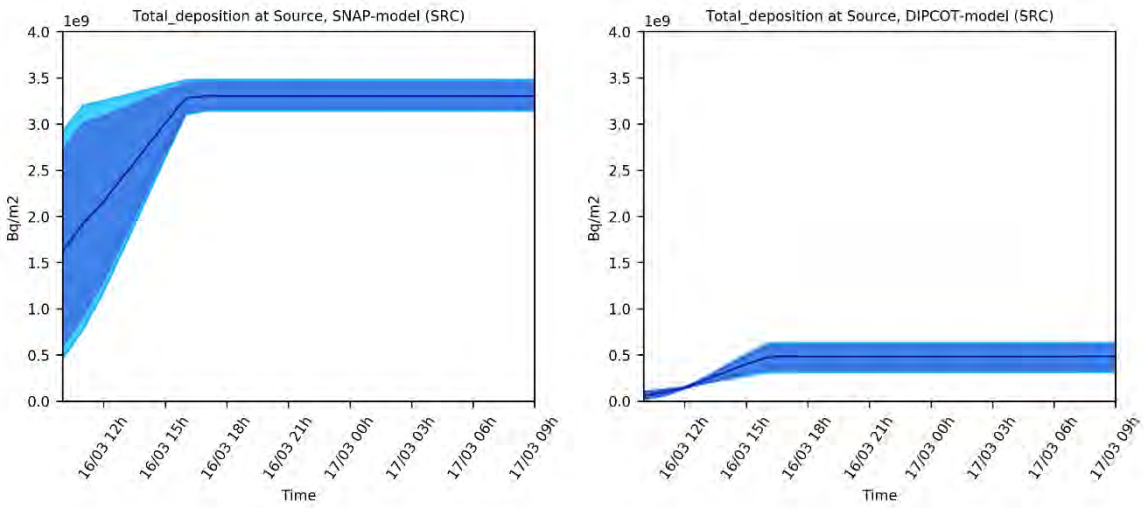


Figure 47. Time-series of accumulated deposition for the EPS member 0 and the 5 emission scenarios at the location of the source. SNAP model (upper panel) and DIPCOT model (lower panel).

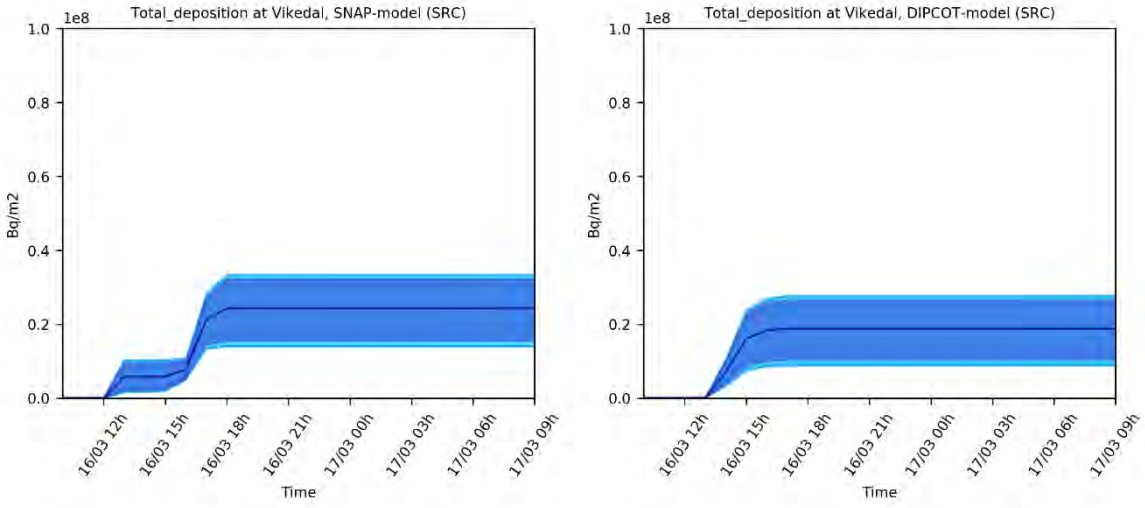


Figure 48. Time-series of accumulated deposition for the EPS member 0 and the 5 emission scenarios at Vikedal. SNAP model (upper panel) and DIPCOT model (lower panel).

Summary of the uncertainty estimates due to meteorology, emission and dispersion model.

A summary of the uncertainties in accumulated concentrations and total deposition is given in Figure 49 and Figure 50. Note that in the summarizing figures we have included the locations Stavanger and Haugesund (see Figure 24) in addition to the source area and Vikedal discussed in the previous section.

The ratio (90^{th} percentile/10^{th} percentile) of the 10 EPS runs with scenario (3) emissions (left panel in Figure 49), is ranging from about 1 to 3 in the source area. At the other locations the ratio ranges from about 2 to 5. When we consider the same ratio for EPS member 0 combined with the 5 emission scenarios (see right panel of Figure 49), we also find a range from about 1 to 3 in the source area, and from 2 to 5 for the three other locations. For the 10 EPS runs the highest ratios are found for the DIPCOT model, while SNAP has the highest ratios for the 5 emissions scenario runs. At least for the WN-case, DIPCOT appears more sensitive to the spread in meteorological data than SNAP, while SNAP

is more sensitive to the spread in the emission data than DIPCOT. We should however not generalize this finding, since it may be dependent on both the weather pattern of the WN-case and the specification of the source term.

In the left panel of Figure 50, we show the ratio of the 50th percentiles of the two models for accumulated concentrations and total deposition for the 10 EPS and 5 emission scenarios while letting the numerator being the maximum value of SNAP and DIPCOT and the denominator the minimum value. This gives a ratio that is always equal to or larger than 1. In the right panel of Figure 50 we calculate the fractional difference (SNAP 50th percentile – DIPCOT 50th percentile)/(0.5·(SNAP 50th percentile + DIPCOT 50th percentile). When this ratio is positive SNAP yields the highest values, while it is negative when DIPCOT yields the highest value.

The ratios of the 50th percentiles vary from about 1 (both models give nearly the same result) up to 9. Note that at the border of the plume this ratio may be large since one model may have very small concentrations of depositions. The largest ratios are seen at Haugesund and Stavanger, while smaller values are encountered at Vikedal (ranging from 1 to 3). From the right panel of Figure 50 it is clearly seen that SNAP has a tendency of higher accumulated concentrations while DIPCOT has larger total deposition as we already have seen from the maps presented in the previous section.

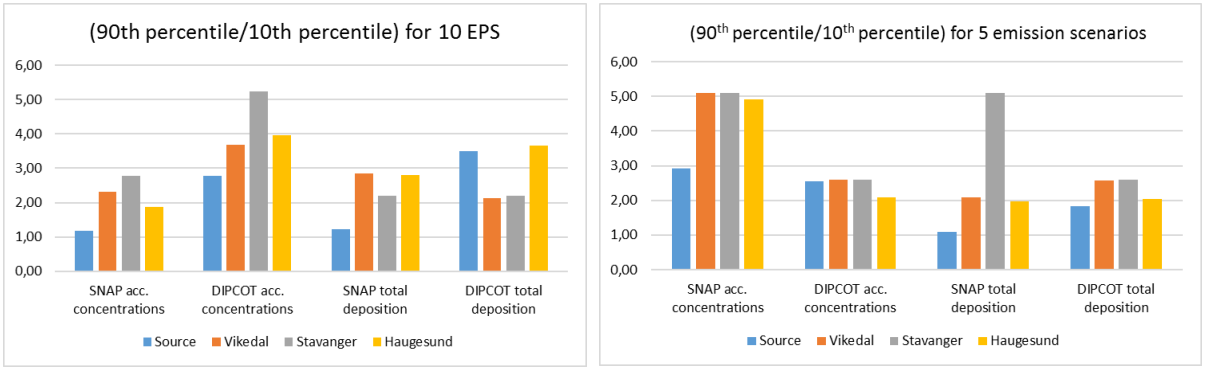


Figure 49. Ratio of (90th percentile/10th percentile) for the 10 EPS members and emission scenario (3) (left) and EPS member 0 and 5 emission scenarios (right)

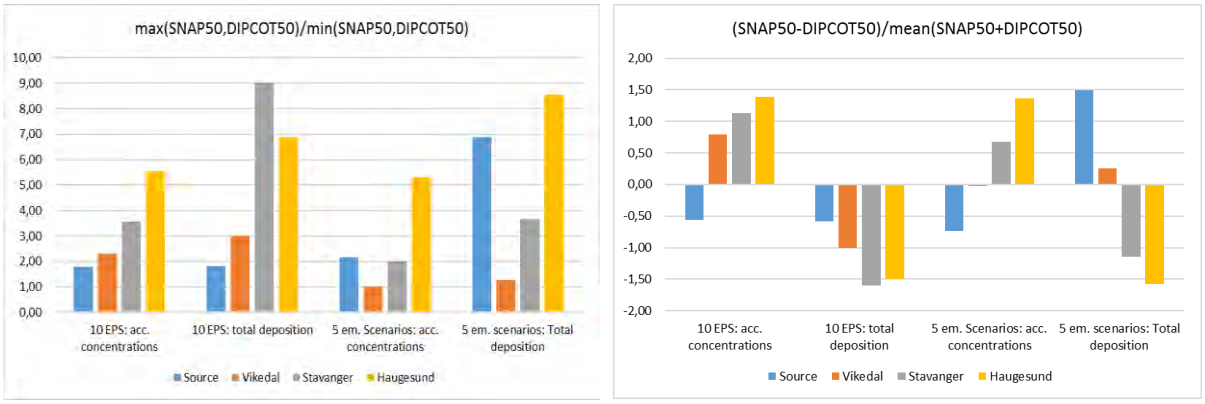


Figure 50. Ratio of the maximum and minimum 50th percentile (left) and fractional difference (right).

Conclusions

Uncertainties are an inherent part of the dispersion of radionuclides from an accidental release. Three important sources of uncertainties are the meteorology, the emission and the dispersion model. In the present study, we have investigated these sources of uncertainties for a hypothetical accident in a nuclear vessel approximately 100 km off the Norwegian west coast.

We investigate the uncertainties by (1) using of the meteorological ensemble prediction system (EPS) of the Norwegian Meteorological Institute with 10 EPS members, (2) 5 emission scenarios, and (3) the two dispersion models SNAP (Severe Nuclear Accident Program) developed and operated by the Norwegian Meteorological Institute and DIPCOT (Dispersion over Complex Terrain) developed and operated by the Greek Institute of Nuclear Technology and Radiation. The two models are run with the same meteorological and emission input data.

The spread in the dispersion of the radionuclides is expressed by the ratio between the 90^{th} percentile and 10^{th} percentile for accumulated concentrations and depositions. For four locations (including the source area) we encounter a spread ranging from close to 1 to about 3 in the source area and from about 2 to 5 at three other locations. The spread due to meteorology and the emission is similar. Considering the ratio of the 50^{th} percentiles of the two models, we find uncertainties of the same order of magnitude as the uncertainties arising from meteorology and emissions. The results therefore indicate that the uncertainties arising from the formulation of the dispersion model may be as important as uncertainties due to meteorology and emissions.

The maps of accumulated concentrations and depositions are qualitatively similar for the two models. However, we see that the dispersion calculations of SNAP yield higher accumulated concentration levels compared to DIPCOT for quite large areas, while DIPCOT yields larger total depositions. The probability of exceeding certain threshold levels of accumulated concentrations therefore become higher in SNAP compared to DIPCOT, while the opposite is the case for the deposition. We also find that the direction, speed of movement and spatial extension of the radioactive plume from the accident vary considerably from one model to the other.

We recommend a more in depth analysis of the differences in the concentrations and depositions of radionuclides of the two models. 3-D analyses of the dispersion, in particular close to the source region, are needed to better understand the differences in the transport. In addition, studies of the wet and dry deposition processes and the treatment of the particle size distributions are also recommended.

References

Andronopoulos, S., Davakis, E., and Bartzis, J. G. 2009. RODOS-DIPCOT Model Description and Evaluation RODOS(RA2)-TN(09)-01. Athens, Greece.

Jerzy Bartnicki, Hilde Haakenstad and Øystein Hov. 2011. Operational SNAP Model for Remote Applications From NRPA. Report no. 12-2011, ISSN: 0332-9879, Norwegian Meteorological Institute, Oslo, Norway.

Frogner, I-L., Singleton, A. T., Køltzow, M.A.Ø., Andrae, U. 2019. Convection-permitting ensembles: Challenges related to their design and use. Q J R Meteorol Soc. 2019; 145 (Suppl. 1): 90– 106. https://doi.org/10.1002/qj.3525

Müller, M., Homleid, M., Ivarsson, K-I., Køltzow, M. A. Ø., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M. and Vignes, O. 2017. AROME-MetCoOp. A Nordic Convective+Scale Operational Weather Prediction Model. Weather and Forecasting. DOI: 10.1175/WAF-D-16-0099.1.

Wellings and Bedwell, P. With contributions from: Leadbetter, S., Tomas, J., Andronopoulos, S,. Korsakissok, I., Périllat, R., Mathieu, A., Geertsema, G., Klein, H., de Vries, H., Hamburger, T., Pázmándi, T., Rudas, C., Sogachev, A. and Szántó, P. 2018. Guidelines detailing the range and distribution of atmospheric dispersion model input parameter uncertainties. Deliverable D 9.1.4. EU Confidence Project.

D 9.5.3 – Ensemble calculations for the atmospheric dispersion of radionuclides. The Fukushima case study: ensemble results and indicators to assess the quality of ensembles

Lead Authors: Irène Korsakissok, Raphaël Périllat

With contributions from: Spyros Andronopoulos, Poul Astrup, Peter Bedwell, Erik Berge, Arnaud Quérel, Heiko Klein, Susan Leadbetter, Olivier Saunier, Andrey Sogachev, Jasper Tomas, Magnus Ulimoen

Table of Contents

68

Introduction 69

Simulations setup 69

- Meteorological data 69
- Source terms 70
- Outputs and observation data 73

Ensemble results 73

- Results for ^{137}Cs air concentration 74
- Results for gamma dose rates 86
- Statistical indicators 91

Conclusions 100

Acknowledgement 101

References 101

Introduction

The hypothetical case studies have provided a large database to compare uncertainty assessment provided by different participants (Berge et al. 2019; Korsakissok et al. 2019). The objectives of this comparison were to provide a benchmark in the way uncertainties are taken into account, and to compare the variability between the ensembles (i.e. due to different models and different ways of taking into account the input uncertainties) with the input uncertainties, represented by the ensembles' spreads. In other words, one of the aims of the projects was to determine whether the variability between participants' uncertainty assessments was negligible or not, compared with the global uncertainty. The rationale behind that is that, while inter-model variability might be significant when looking at deterministic outputs, it may be of less importance as far as uncertainties estimations are concerned. Results have shown an important variability between the participants, even when the input uncertainties were the same, although this variability is not of the same order of magnitude as, for instance, source term uncertainties. The questions of how to present the outputs, how to choose the thresholds, percentiles, were also widely investigated. However, to properly evaluate the different ensemble results and possibly the models / approaches used to take into account the input uncertainties, there is the need for comparison with environmental observations. In this regard, the Fukushima case study provides a unique opportunity to complete and shed a new light on the case studies conducted within task 1.2 of WP1. More precisely, it helps infer whether the so-called "inter-model variability" (which, more precisely, should be named "inter-ensemble" to account for the differences in the ensembles' construction as well as between the models) is significant with regard to model-to-measurements comparisons, as well as measurement uncertainty. For instance, if all ensembles are very bad at predicting the observations variability with time at all stations, then the most important issue is to improve the knowledge on input uncertainties. If the performance of all ensembles is similar, then the inter-model variability is probably negligible in a real case. A related question is whether taking into account the combined meteorological and source term uncertainties is sufficient to account for uncertainties in atmospheric dispersion simulations, when comparing them with observations. If the ensembles are globally capable of encompassing the observations, then no significant source of uncertainties was left out. Finally, one of the aims of this report is to present several statistical and graphical indicators of an ensemble's performance by comparison to radiological observations. Probabilistic indicators represent the ability of an ensemble to properly encompass the observations. As a complement, the last part of this report presents deterministic indicators applied to the ensembles' median, which is a possible way of using ensemble results: in complement to the uncertainty estimation, the ensembles' median may perform better than deterministic "best estimates".

Simulations setup

Meteorological data

The ECMWF Integrated Forecast System (cycle 45r1) (Molteni et al. 1996; Leutbecher and Palmer 2008) was used to create an ensemble forecast for the period 00:00UTC 14 to 00:00UTC 17 March 2011. Initial perturbations were constructed from an ensemble data-assimilation with singular vector perturbations and model uncertainty is taken into account using a stochastic physics scheme. The ensemble consists of one run starting from unperturbed initial conditions (the control forecast) and 50 from perturbed initial conditions. Data were extracted over a horizontal domain of 32°N to 43°N and 134°E to 148°E, at a resolution of 0.2° (red region in Figure 51). 36 vertical levels were selected from the

surface to 148hPa with more levels close to the surface and the temporal resolution of the data was hourly.

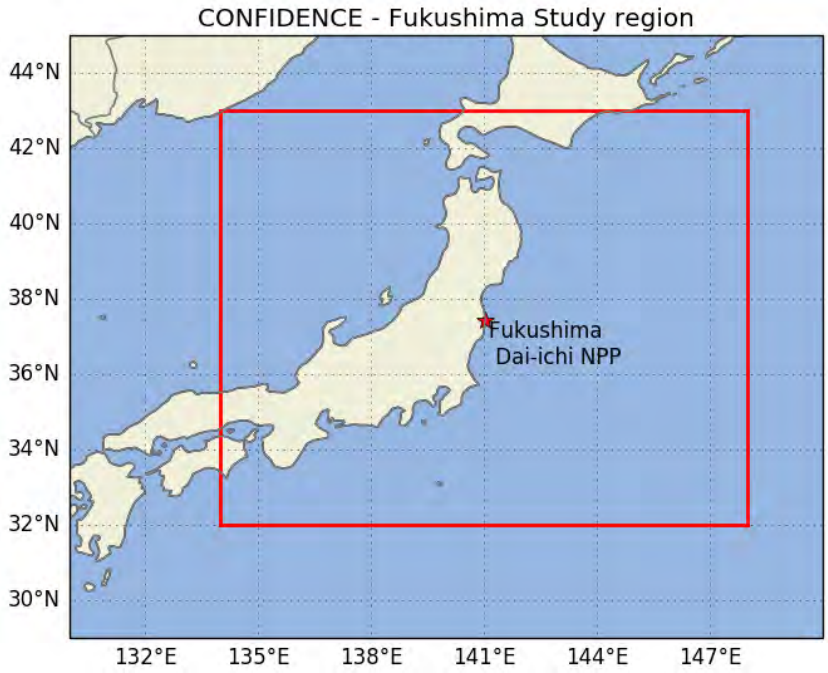


Figure 51 : Area covered by ECMWF data.

Source terms

There are two main families of methods for estimating releases (Mathieu et al. 2018).

1. An approach based on reactor physics and the knowledge of the state of the power plant. To date, there is no complete source term in the literature obtained from this approach which includes a kinetic release model for different radionuclides.

2. Methods coupling the radiological measurements in the environment and the atmospheric dispersion simulations to infer release rates. Inevitably, the quality of the source term correlates to the accuracy of the meteorological fields used, to the kind of atmospheric transport model used and to the relevance of the measurements, since an event can only be reconstructed if it is observed. (Mathieu et al. 2018) distinguish between methods called "simplified" and "inverse" methods. The former are manual or semi-automatic and are based on a limited set of measurements. The inverse modelling techniques are more operational automatic methods based on mathematically rigorous approaches.

Saunier et al. (2016) assessed the realism of several previously published source terms by comparing dispersion model output using these source terms to a dataset of air concentration measurements (the suspended particulate matter – SPM – dataset described below and used in this study) which were not used in these release estimations, having been made public only in 2015. Overall the agreement between the models and measurement is poor and no source term appears better than the others. Saunier et al. (2016) also presented initial attempts to use SPM data to estimate releases and show that these data better reproduce the temporal variability of the air concentration.

Table 13 synthesises the source terms reconstructed using either “simplified” or “inverse” methods. The nine source terms retained for this study are highlighted in green. Some source terms were not selected, either because they did not significantly differ from another source term (to avoid introducing bias in our ensemble, it is better to have source term that are significantly different), or

because they were not deemed relevant due to the spatial and temporal scale of interest of our study, or, because insufficient information was available. Figure 52 shows the release rate of ^{137}Cs for the nine source terms during the 3-day period of interest. Although they are all a posteriori source terms, based on observations, there is still a high variability within the source terms, especially when looking at the kinetics.

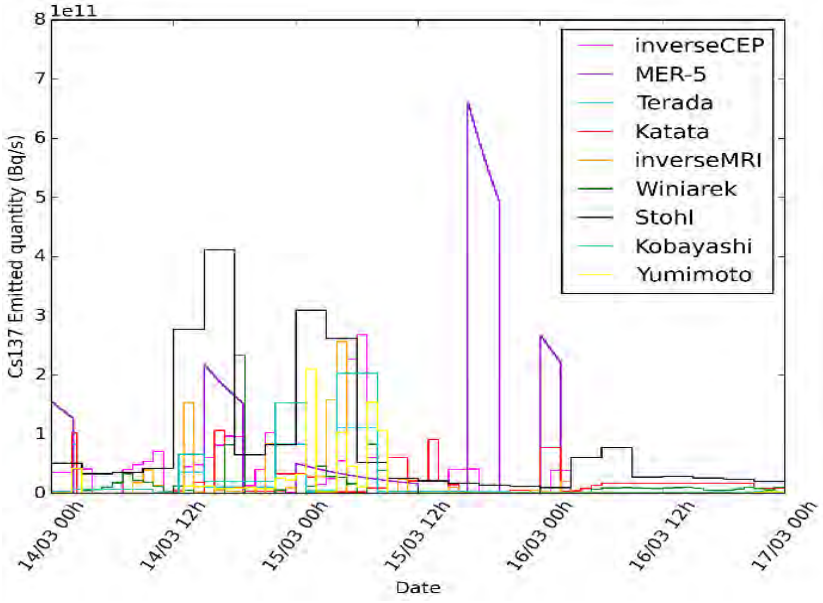


Figure 52: Release rates during the simulated period for the 9 source terms

Table 13 Source terms estimated from measurements in the environment, from (Mathieu et al. 2018). The source terms used in this study are highlighted in green. FDNPP stands for Fukushima Daiichi Nuclear Power Plant.

No.	Source term	Total ^{137}Cs (PBq)	Method	Observations used	Model / meteorology used
1	Chino et al. (2011) JAEA	13	Simplified	Concentrations in the air over Japan. Dose rates close to FDNPP. Facility events.	GEARN (Lagrangian) / GSM (Global Spectral Model)
2	Stohl et al. (2011)	35.7	Inverse	Concentrations in the air over the entire Northern Hemisphere. Facility events.	Flexpart (Lagrangian) / ECMWF 0.18°, 1° and GFS 0.5°
3	Winiarek et al. (2012)	10-19	Inverse	Concentrations in the air over Japan and North America.	Polair3D (Eulerian) / ECMWF 0.25°
4	Terada et al. (2012) JAEA	8.7	Simplified	Concentrations in the air over Japan. Dose rates close to FDNPP. Facility events.	GEARN (Lagrangian) / GSM + MM5
5	Mathieu et al. (2012)	20.6	Simplified	Concentrations in the air over Japan. Dose rates close to FDNPP. Facility events.	Gaussian puff (pX) and Eulerian (ldX) / ECMWF 0.125°
6	Saunier et al. (2013a)	15.5	Inverse	Dose rates in Japan.	ldX (Eulerian) / ECMWF 0.125°
7	Kobayashi et al. (2013) JAEA	13	Simplified	Depositions over the sea. Concentrations in air over Japan. Dose rates close to the facility. Facility events.	SEA-GEARN (ocean) + GEARN (Lagrangian) / GSM + MM5
8	Hirao et al. (2013)	9.6	Simplified	Concentrations in the air. Daily depositions by Prefecture.	LPRM (Lagrangian) / MM5
9	Winiarek et al. (2014)	11.6-19.3	Inverse	Total deposition over Japan. Concentrations in the air over Japan.	Polair3D (Eulerian) / WRF
10	Achim et al. (2014)	10.8	Simplified	Concentrations in the air of CTBTO stations. Dose rate at FDNPP. Facility event.	Flexpart (Lagrangian) / WRF + NCEP (GFS)
11	Katata et al. (2015) JAEA	14.1	Simplified	Concentrations in the air over Japan. Dose rates close to FDNPP. Depositions over the sea.	GEARN Lagrangian / GSM + MM5
12	Yumimoto et al. (2016)	8.12	Inverse	Total deposition over Japan.	CMAQ (Eulerian) / WRF
13	Saunier et al. (2016)	8.1	Inverse	Concentrations in the air over Japan including SPM data.	ldX (Eulerian) / MRI 0.03°
14	Liu et al. (2017)	7-35	Inverse	Concentrations in the air over Japan including SPM data.	Polair3D / WRF

Outputs and observation data

In this work several endpoints are considered. We make use of the hourly concentrations of ^{137}Cs retrieved by Tsuruta et al. (2014) and described in Oura et al. (2015). The air activity concentration is given in Bq/m^3. The data was obtained from the air quality automated monitoring network measuring SPM on filter tapes. They were retrieved too late to detect short-lived radionuclides, but they give crucial information on the temporal variation of ^{137}Cs concentration close to ground level, therefore indicating the passage of different plumes. Figure 53 shows the spatial distribution of the 108 stations; unfortunately, the distribution is not uniform, especially over the mostly contaminated area (in grey). The dataset came with an estimation of the associated measurement uncertainty, mostly of the order of 0.1-0.5 Bq/m^3. In this study, it will be neglected compared to the uncertainty associated with atmospheric dispersion simulations.

The gamma dose rate measurements were provided by automated stations as early as March 11th, with a 10-minutes time step. The dose rate measurements are described in (Saunier et al. 2013b) and most of them are available in the IAEA database (IAEA 2012). There are 88 stations spread over Japan, although the spatial coverage is heterogeneous (Figure 53). The dose rate readings have the advantage of measuring the contribution from all radionuclides, including the short-lived species that could not be detected by the other monitoring systems. They are composed of two parts: the direct plume contribution ("cloud-shine") and the gamma-ray emitted by radionuclides deposited on the ground ("ground-shine"). Hence, the cloud-shine is usually responsible for peak values observed during the plume passage, whereas ground-shine corresponds to a lasting contribution that lingers after the plume has left the area and decreases because of the radioactive decay.

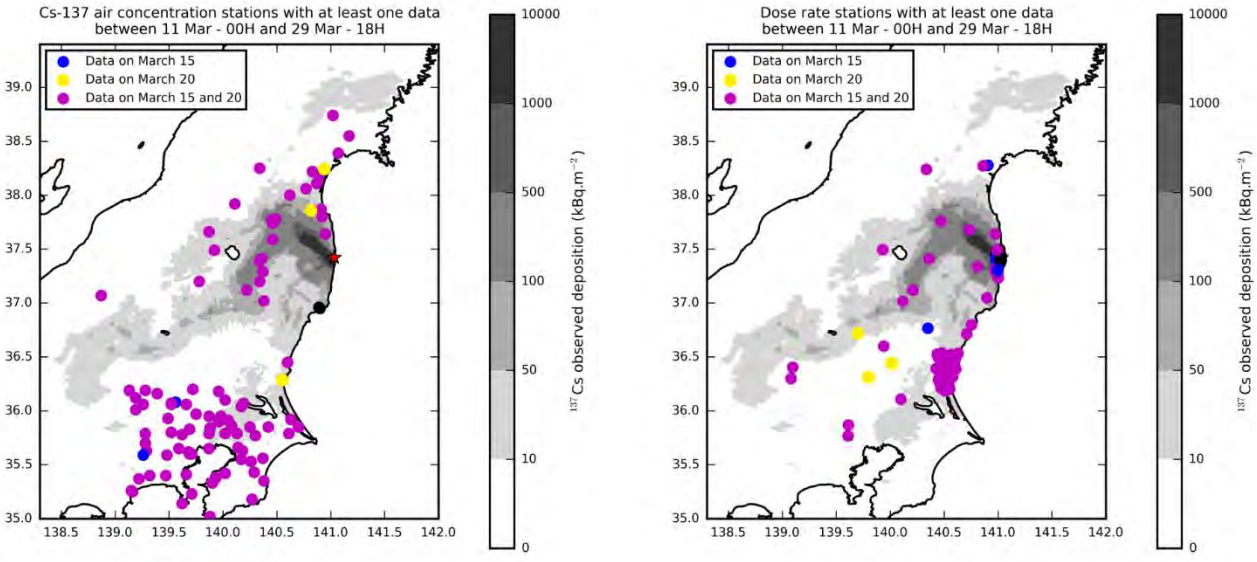


Figure 53 : Localization of ^{137}Cs activity concentration stations (left) and dose rate stations (right) over Japan. The grey shades show the deposition observations at the end of the Fukushima accident.

Ensemble results

The objective of the study is to use the meteorological ensemble (51 members) and the nine source terms and propagate them through the atmospheric dispersion models, in order to compare the output results with the observations. This amounted to many simulations, ranging from 102 to 459, depending on the way the input uncertainties were taken into account (Table 14). Among the participants, 3 of them used all combinations of meteorological members and source terms (459

simulations). The Met Office used randomly 2 different source terms for each meteorological member, which resulted in 100 simulations. IRSN used Monte Carlo sampling and randomly selected meteorological members, source terms, with additional physical perturbations on scavenging coefficients, deposition velocities and vertical diffusion. 200 samples were used. The ranges of variations are given in (Korsakissok et al. 2019).

Participant		Type of model	Number of simulations (perturbations)
France	**IRSN**	ldX – Eulerian	200 simulations (MC)
The Netherlands	**RIVM**	NPK-puff – Gaussian puff	459 simulations (9 ST x 51 MET)
UK	**Met Office**	NAME – Lagrangian particle	102 simulations (2 ST drawn randomly among 9 x 51 MET)
Greece	**EEAE**	DIPCOT – Lagrangian puff	459 simulations (9 ST x 51 MET)
Denmark	**DTU**	RIMPUFF – Gaussian puff	459 simulations (9 ST x 51 MET)
Norway	**NMI MET**	SNAP – Lagrangian particle	459 simulations (9 ST x 51 MET)

Table 14 : Summary of participants, name, type of atmospheric dispersion model used and number of simulations for the Fukushima case.

The observations used are time series of ^{137}Cs air concentrations and gamma dose rates on stations. Time series of the variables of interest at stations where observations are available provide a useful view of the ensembles' performance. While it is difficult to give an overview of the performance on more than a hundred stations, we selected stations located in different areas that were deemed representative of the results. Results are shown first for ^{137}Cs air concentration, then for gamma dose rate stations. In practice, some stations may be selected or be given a larger weight, depending on the contamination episode / area that is deemed more important for the case. For instance, one may decide that representing uncertainties close to the source (i.e. for highest activity or dose rate values) is more important than having a good representation of all uncertainties, including stations with very small values. This will be further discussed in the section devoted to statistical indicators.

Results for ^{137}Cs air concentration

Figures show the observations of ^{137}Cs concentration in Bq/m^3 (in red) and the ensemble results (in blue). The dark blue line is the median of the ensemble, and the dark shade of blue represents the 25-75th percentiles. The lighter shade of blue is the outer range of the ensemble (percentiles 0-25 and 75-100). Simulations were made from March 14th at 00:00 UTC to March 17 at 00:00 UTC, but results are only compared from March 14th at 12:00 UTC, so that any plume released earlier than March 14th at 00:00 UTC would likely to be gone from the stations locations.

North-western area and Abukuma valley

The north-western area is the zone featuring the highest contamination, as shown by the airborne deposition measurements (shades of grey in Figure 54(a)). This contamination was due to wet deposition that occurred on March 14-15, when a significant release, combined with turning winds and a rain episode, induced a significant wash-out. This contamination occurred only within a few hours.

Fukushima city

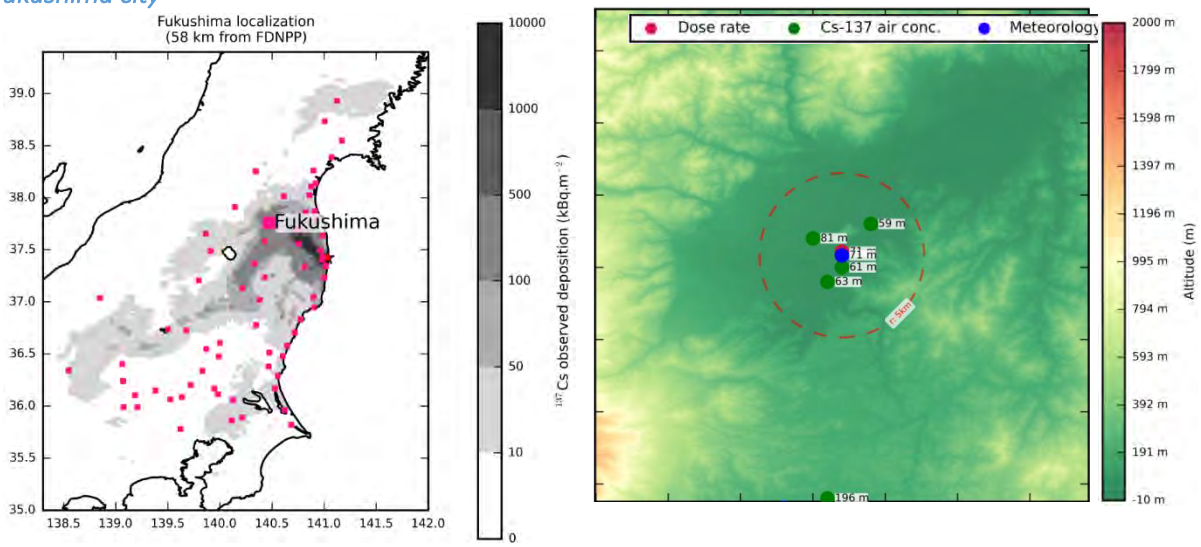
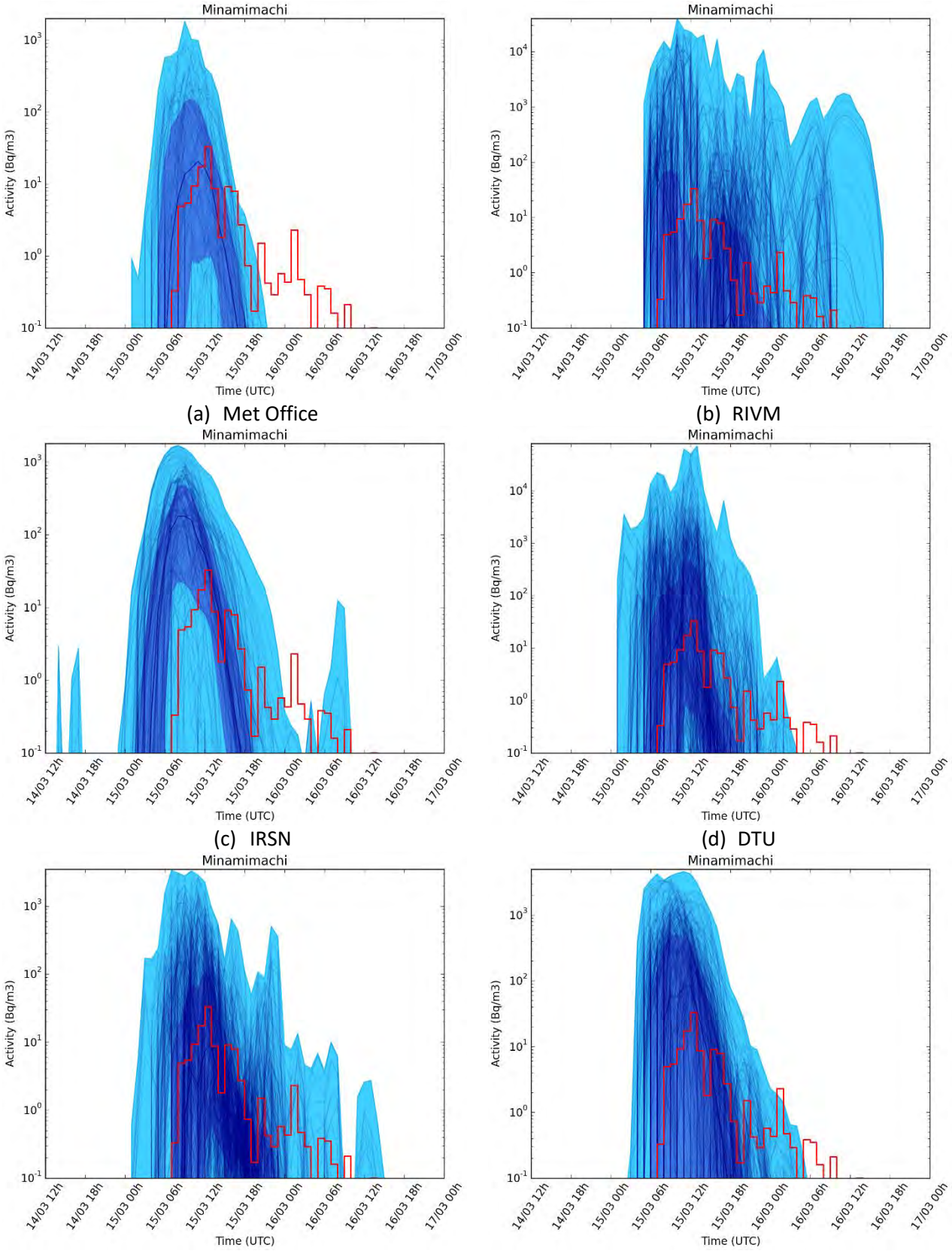


Figure 54 : Location of Fukushima city (a); topography of the area around Fukushima and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)

Fukushima city is located 58 kilometres from the FDNPP (Fukushima Daiichi Nuclear Power Plant), at the furthest point of the high deposition north-western area (Figure 54(a)). Along with gamma dose rate and meteorological measurements, there are four stations of ^{137}Cs measurements within 5 kilometres from each other (Figure 54(b)): Furukawa, Minamimachi, Moriai and Sugitsuma-cho. The four stations show very similar observations, with a plume passage at the same time and a maximum observed concentration of 30 Bq/m^3. However, there are some differences between them, with the plume passage lasting longer in some cases. This is illustrated below with two stations: Minamimachi and Sugitsuma-cho: at Minamimachi, there are activity measurements from March 15th at 06:00 UTC to March 15th at 06:00 UTC, whereas at Sugitsuma-cho, there is no detection after March 15th at 18:00 UTC (except a small peak later, on March 16th at 12:00). Some participants' ensemble predicted a plume duration too short to encompass all observations at Minamimachi (Met Office and, to a lesser extent, DTU and NMI MET) but have a good agreement with observations for Sugitsuma-cho; on the other hand, RIVM and IRSN correctly encompass the observations at Minamimachi but predict activity concentrations where there are no observations in the case of Sugitsuma-cho. For all participants, the modelled concentrations are almost or exactly the same at the two stations. The observed discrepancies between the two stations are probably due to local topography effects that are not represented by the dispersion models, especially with the meteorological resolution used.

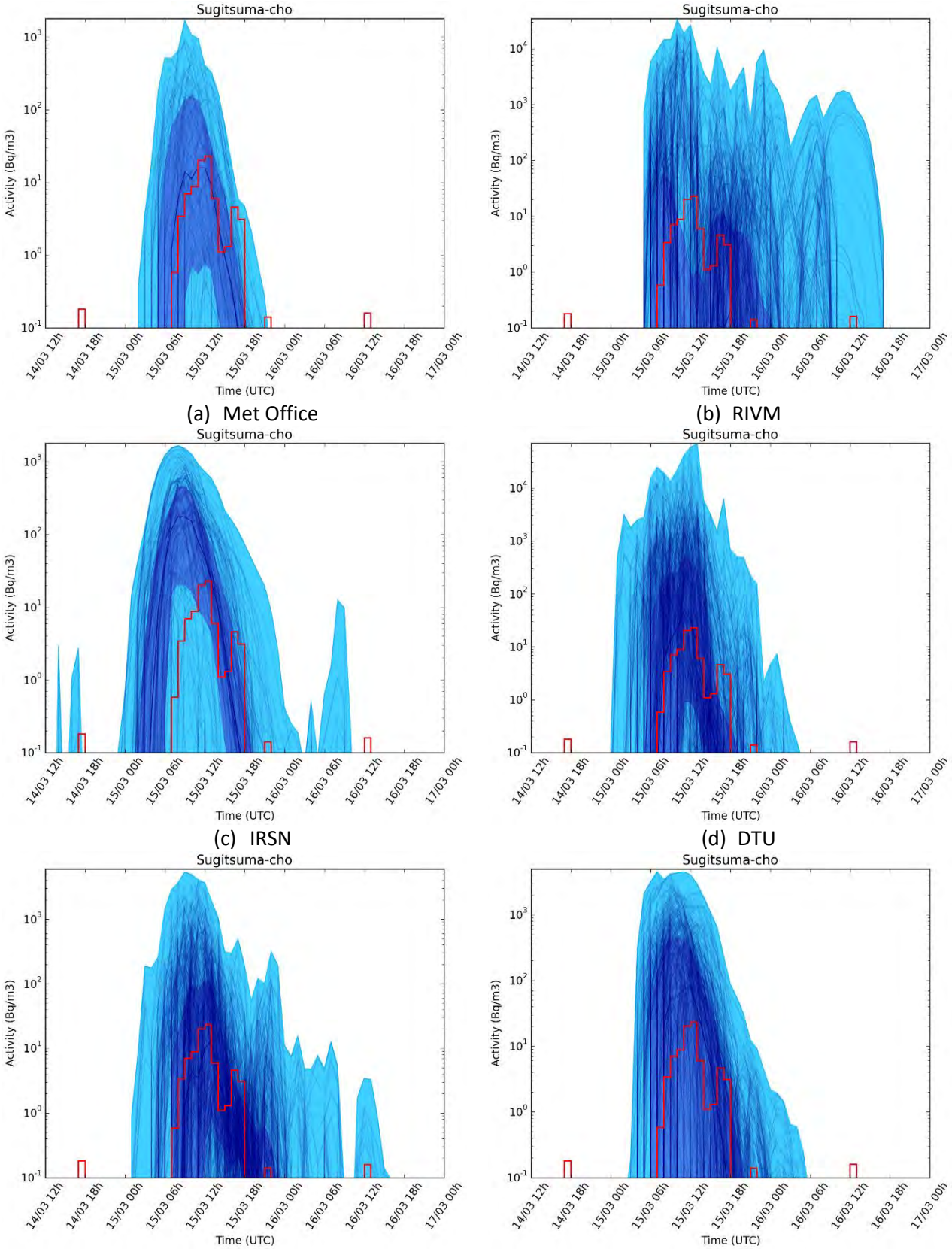


(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 55 : Ensemble results of ^{137}Cs activity concentration at station Minamimachi, in Fukushima city (58 km NW from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 56 : Ensemble results of ^{137}Cs activity concentration at station Sugitsuma-cho, in Fukushima city (58 km NW from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.

Shirakawa

Shirakawa city is located 77 kilometres south-west of the FDNPP, in a region where ^{137}Cs deposit ranges between 50 and 150 Bq/m2 (Figure 57(a)), in the Abukuma valley, surrounded by mountains north-west and south-west from the city (Figure 57(b)).

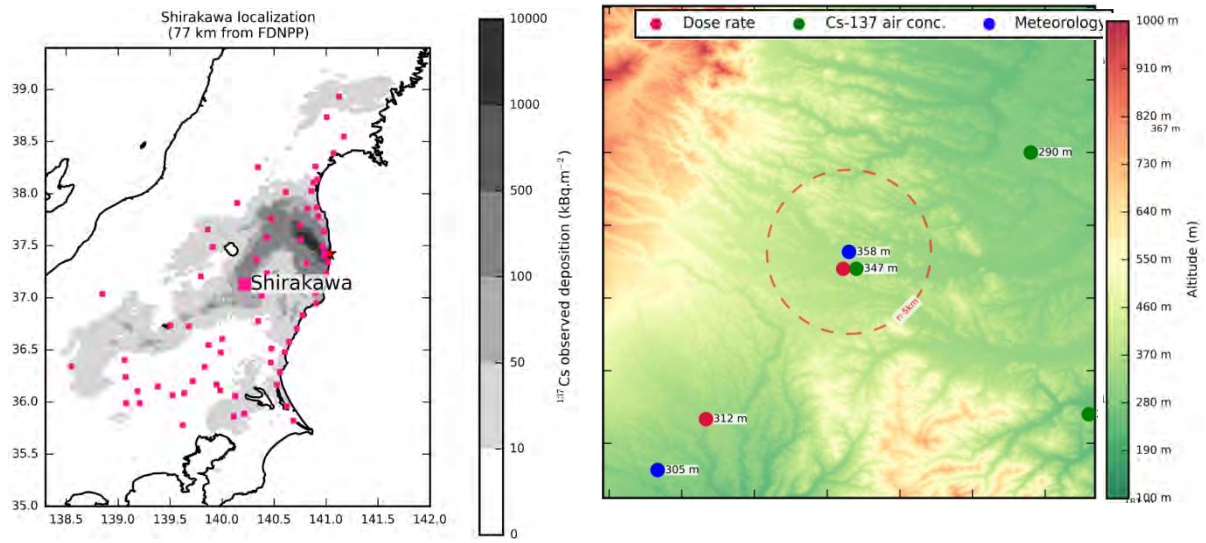


Figure 57 : Location of Shirakawa (a); topography of the area around Shirakawa and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)

On March 15th, the maximum observed activity concentration of ^{137}Cs is over 100 Bq/m^3. The plume is observed from March 15th at 03:00 to March 15th at 16:00, with light rain between 06:00 and 12:00 and more rain between 12:00 and 16:00. Therefore, the deposition corresponds to the scavenging of the plume, mostly due to light rain (Quérel et al, 2016). The ability of a model (or an ensemble of simulations) to correctly predict the deposition in this area is therefore correlated to (a) correctly predicting the plume passage, not only with the maximum value but also with the "tail" of the plume, and (b) accurately predicting the rain, including light rain which is usually difficult to forecast with meteorological models only. Figure 58 shows that the ECMWF ensemble is able to encompass most of the rain observations on the 15th March, including the possibility to have rain earlier than the observation time. Indeed, as Japanese rain gauges have a step of 0.5 mm/h, it is probable that rain with intensity lower than 0.5 mm/h participated in the plume wash-out without being detected by the gauge.

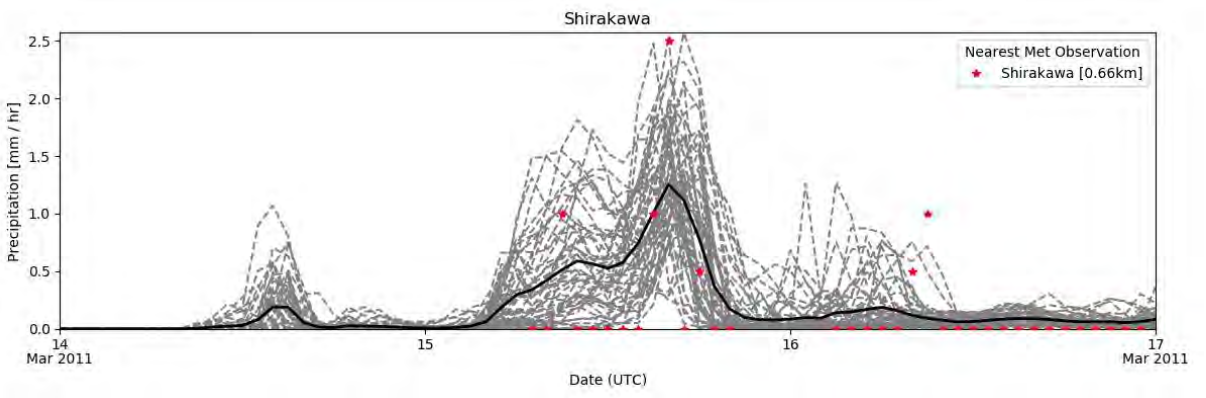
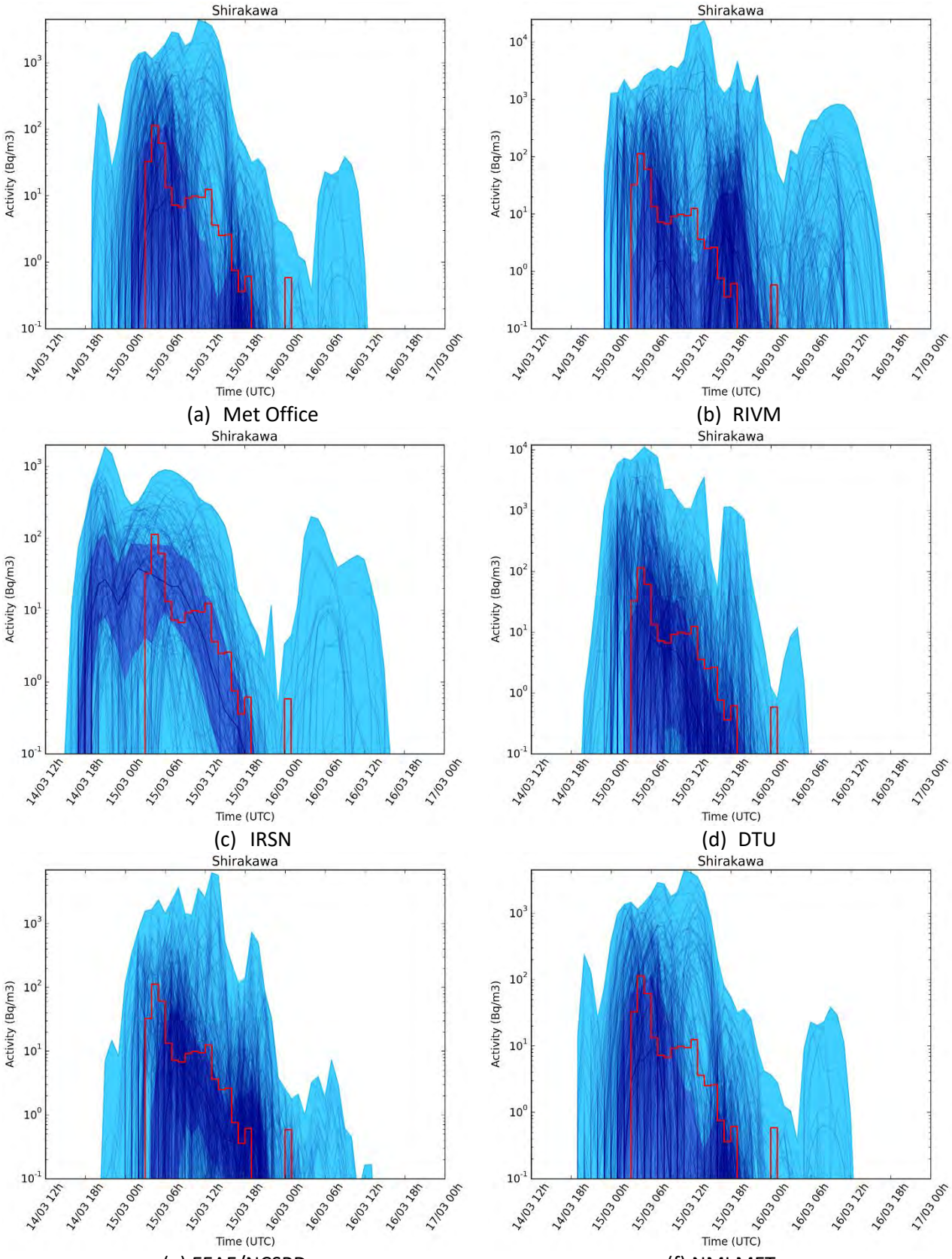


Figure 58 : Ensemble precipitation forecast by the meteorological ensemble at Shirakawa, and comparison with nearest meteorological observations available (in red). The distance to the meteorological observations is given in brackets.



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 59 : Ensemble results of ^{137}Cs activity concentration at station Shirakawa (77 km SW from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.

Aizuwakamatsu

This station is located within 100 kilometres of the source, close to the highly contaminated north-western area (Figure 60). The contamination in this area occurred mostly within a few hours during March 15th, due to scavenging by rain. The first peak corresponding to the plume passage on March 15th is correctly reproduced by all participants, with a very large spread corresponding to a high overestimation of some members. However, the second peak on March 16th at 05:00, corresponding to a latter release, is only reproduced by one participant, RIVM, and by very few IRSN members. Therefore, the presence or not of this peak is not due to meteorological or source term data. Rather, the station may be located at the edge of the plume and, depending on diffusion schemes, some simulated plumes may "touch" the station while most do not.

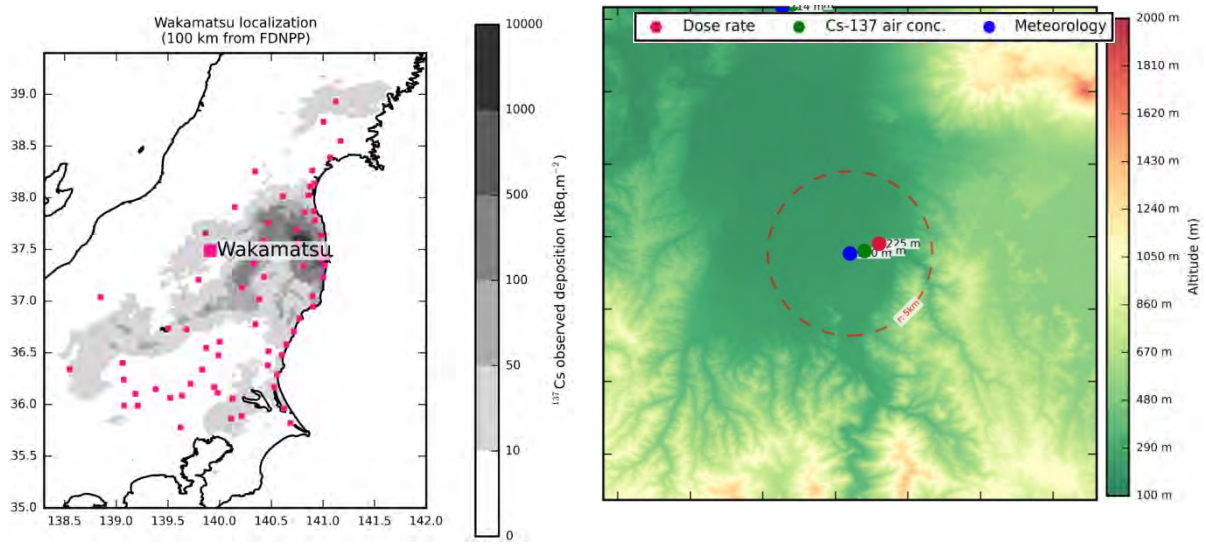
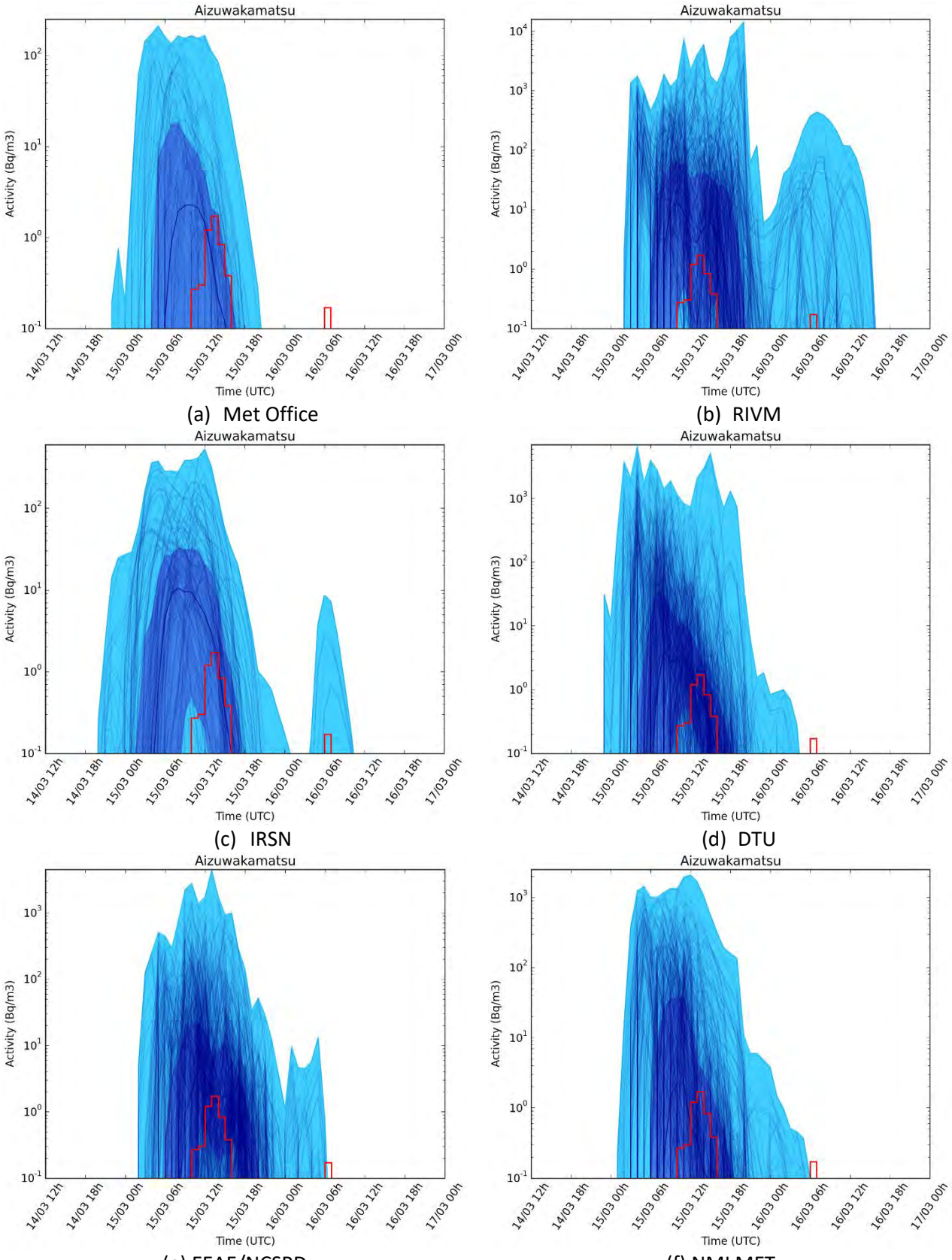


Figure 60 : Location of Aizuwakamatsu (a); topography of the area around Aizuwakamatsu and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 61 : Ensemble results of ^{137}Cs activity concentration at station Aizuwakamatsu (98 km west from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.

South coast and Kanto area

Kamisu-yokose (Tonosho)

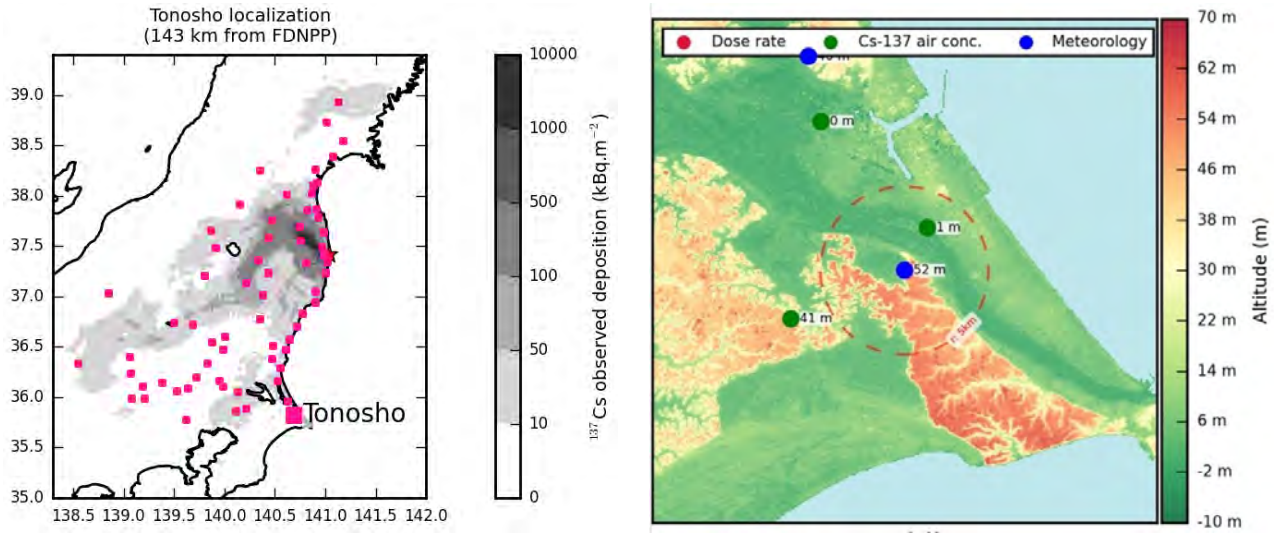


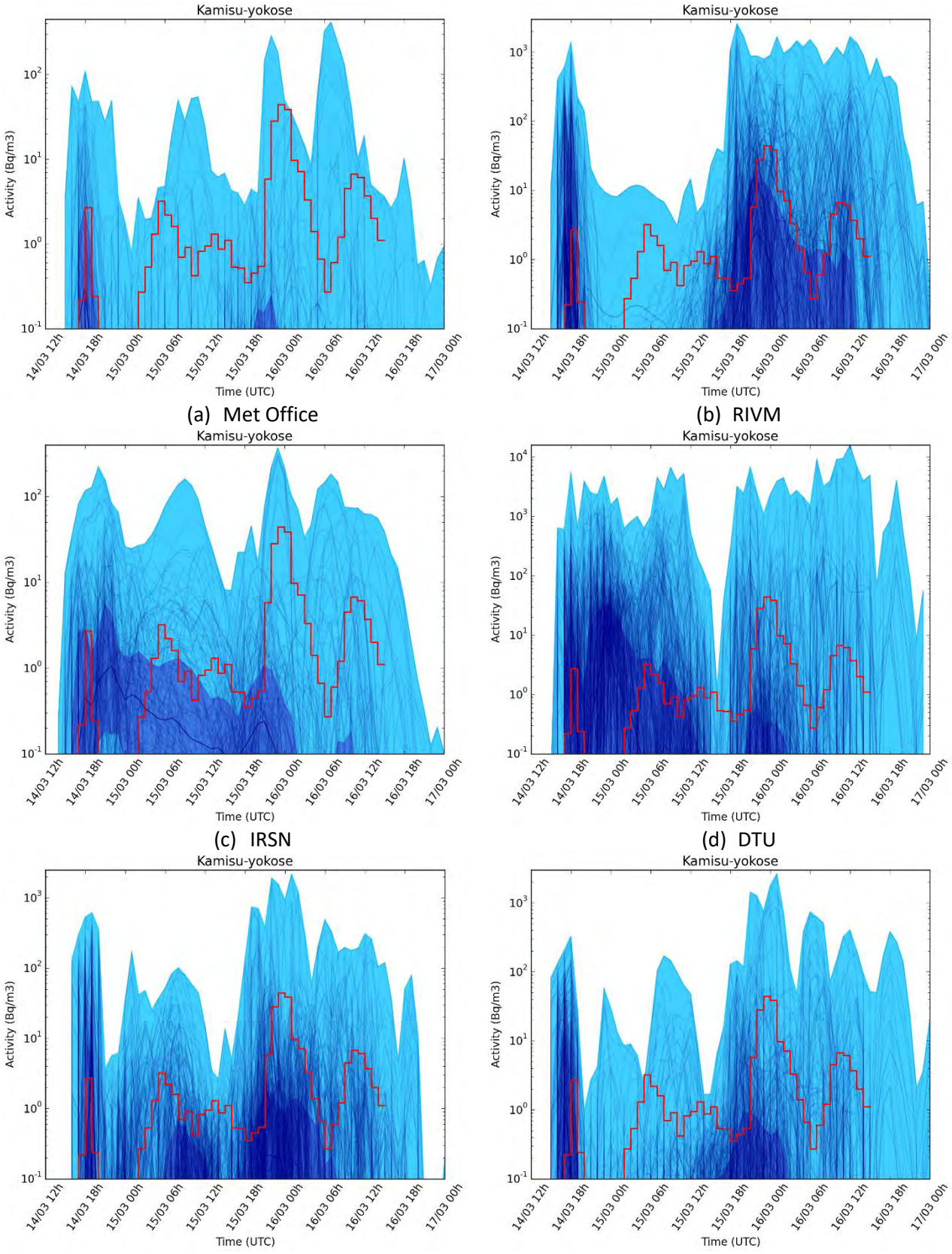
Figure 62 : Location of Tonosho and Kamisu-yokose (a); topography of the area around Tonosho and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)

The meteorological station of Tonosho and the ^{137}Cs concentration station Kamisu-yokose are located about 140 kilometres south from FDNPP, along the coast (Figure 62). The main plume is detected by this station. There are four different plumes: (1) on March 14th at 18:00 UTC there is a small peak, (2) on March 15th between 03:00 and 21:00 UTC (both with values well below 10 Bq/m^3), (3) on March 15th between 21:00 UTC and March 16th at 06:00 UTC, with concentrations over 40 Bq/m^3, and (4) on March 16th between 09:00 and 15:00 UTC, with concentrations lower than 10 Bq/m^3. Of all four plumes, only the 3rd concurs with rain observations and is therefore eligible for undergoing wet deposition.

Here, all ensembles globally encompass the observations, with the exception of IRSN's ensemble which is a bit late for the 3rd episode. However, when looking at the darker blue, which represents the 25-75th percentile envelope, tendencies are different:

- RIVM and NMI MET represent reasonably well the 1st and 3rd plumes, with some underestimation of plume 2;
- DTU and IRSN represent plumes 1 and 2 (between March 14th at 12:00 to March 15th at 12:00) well except that plume 2 is observed longer than what the ensembles feature;
- Met Office only represents plume 1, and show a very small peak for plume 3;
- EEAE represents well plumes 1 and 3, and underestimates plume 2;
- Only RIVM models a peak for plume 4.

This highlights the importance of taking into account all members of the ensemble, and not focusing only on the 25-75th percentile. It also illustrates the variability between the different ensembles, due to the ensembles constructions and the different models and physical parameterizations.



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 63 : Ensemble results of ^{137}Cs activity concentration at station Kamisu-yokose (140 km South from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.

Abiko

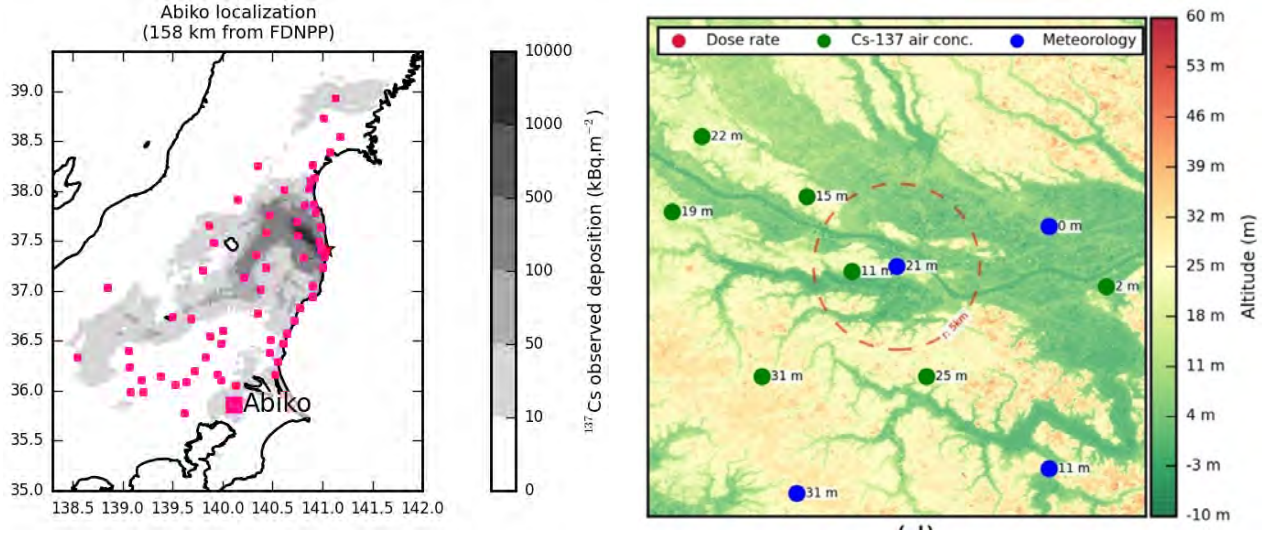
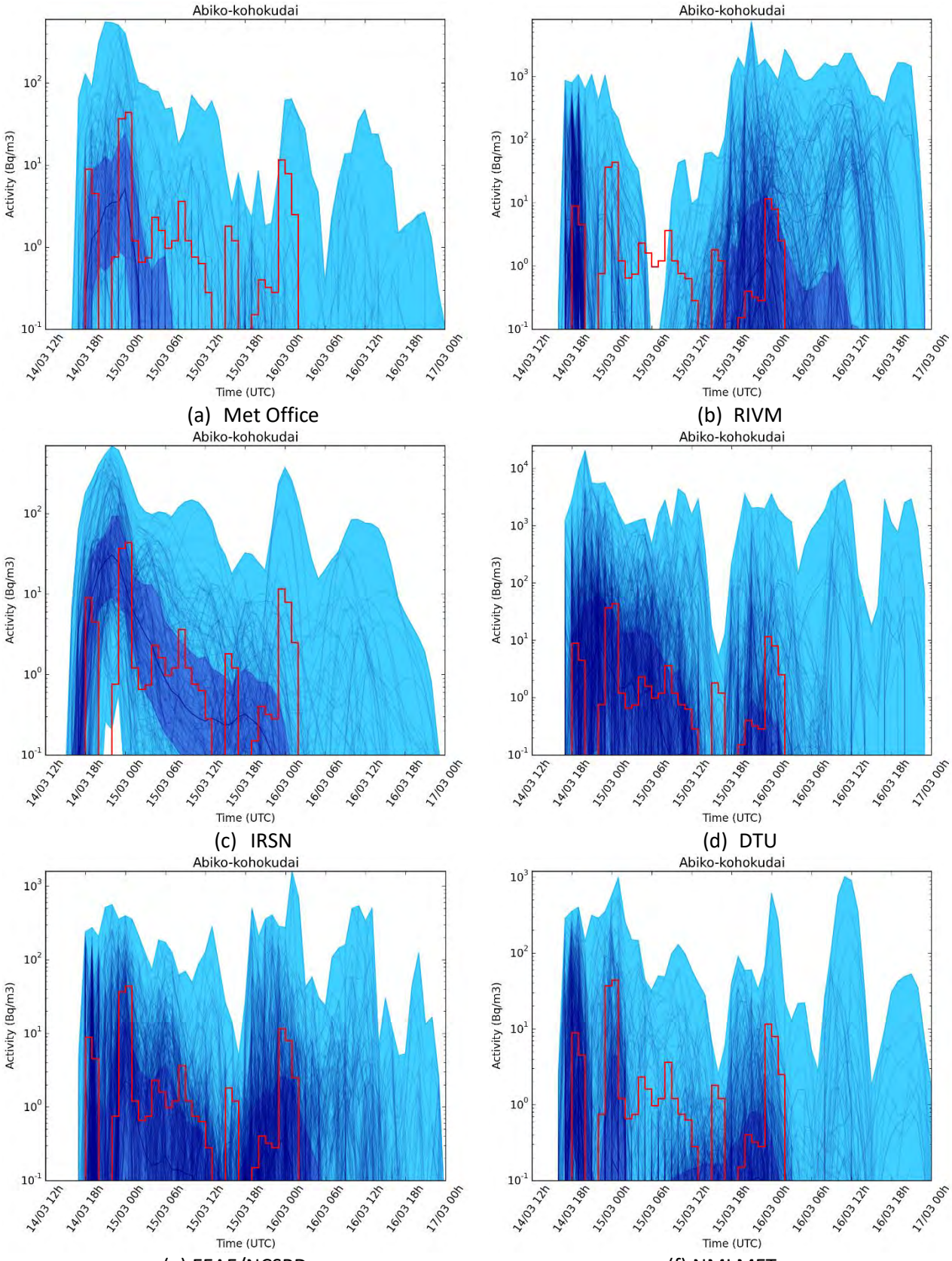


Figure 64 : Location of Abiko (a); topography of the area around Abiko and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)

Abiko is located in the Kanto plain, 158 kilometres south-west of FDNPP, in an urban area surrounded by croplands (Figure 64). It is located in a zone with ^{137}Cs deposition between 10 and 50 Bq/m^2. On March 14-16th, there are four different plumes, the main being between March 14th at 22:00 UTC and March 15th at 03:00 UTC (over 40 Bq/m^3). No rain observations coincide with the passage of these plumes, therefore only dry deposition occurs on these dates; the deposition featured on the map is likely to come from wet deposition on March 20th (Quérel et al, 2016).



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

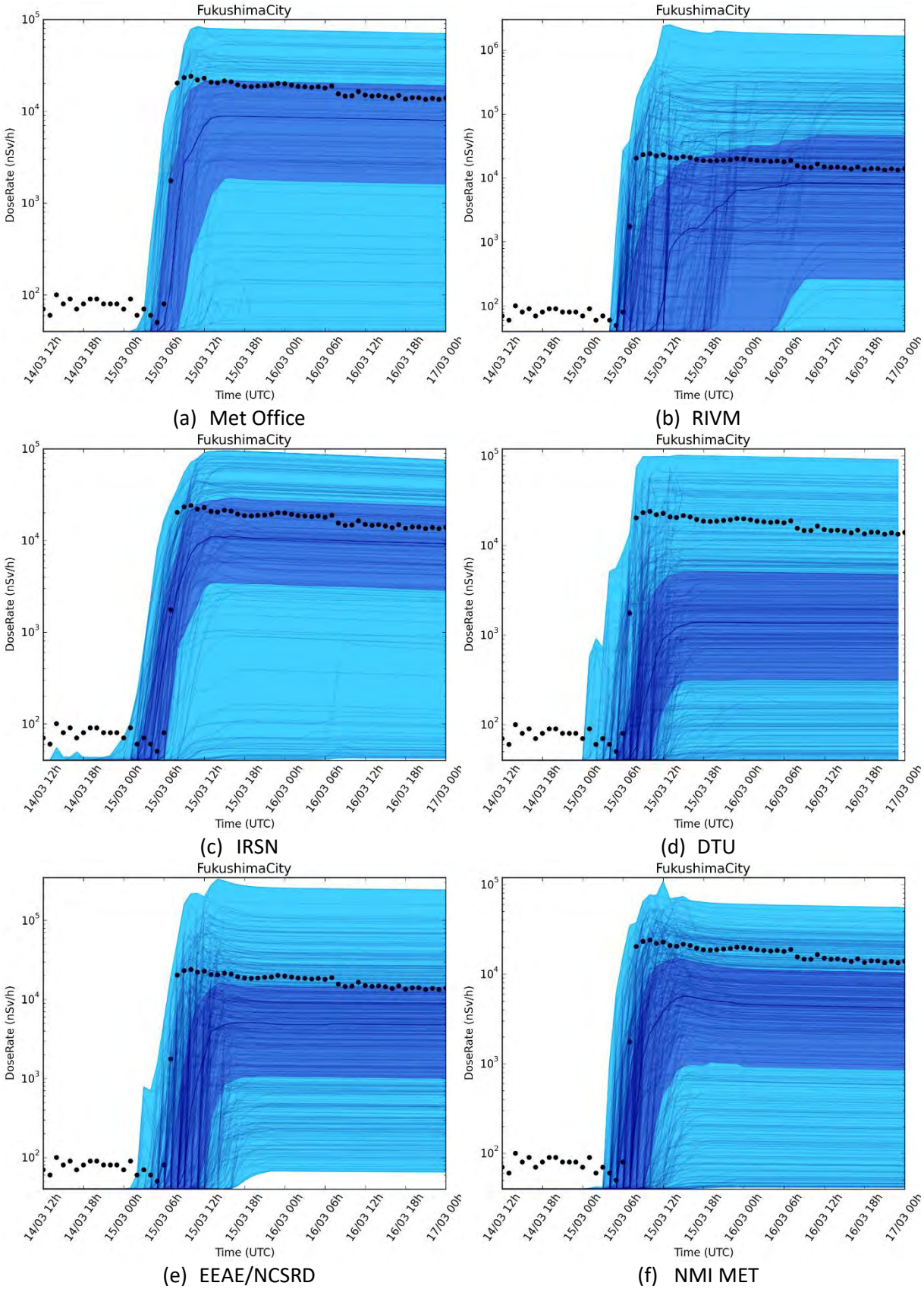
Figure 65 : Ensemble results of ^{137}Cs activity concentration at station Abiko (158 km South from FDNPP) for the Fukushima case study for several project participants, compared with observations (red line). The y-axis are not the same on all figures.

Results for gamma dose rates

Gamma dose rate monitoring stations measure the gamma radiation coming from the air and the ground nearby the station. Therefore, they include the deposition from plumes released before March 14th, which was not modelled by our studies. Although, deposition that occurred earlier than this date is negligible compared to the contamination that occurred during March 14-16th, it appears that, in many cases, an offset corresponding to the background radiation should be added to the simulations in order to compare the ensembles with the observations. In the following results, an offset equal to the minimum value for each station on the period March 14th at 12:00 UTC and March 15th at 15:00 UTC was added to the modelled values. For some stations (e.g. Iwaki, Figure 70) there was no contamination due to the accident before March 14th, and this value represents the background noise coming from cosmic and terrestrial radiations. For others (e.g. Fukushima, Figure 66), this value is higher due to previous contamination of this area.

Fukushima city

Fukushima city is located in the highly contaminated north-western area (Figure 54). The deposition in this area occurred within a few hours on March 15th, due to the scavenging by rain of the plume that reached the station between 06:00 and 15:00 UTC, as demonstrated by air concentration measurements (Figure 55 and Figure 56). This episode was particularly difficult to simulate by models, since it is a conjunction between turning winds, rainfall and plume passage within a few hours ((Korsakissok et al. 2013); (Mathieu et al. 2018)). The gamma dose rate observations shown in Figure 66 are typical of a wet deposition pattern: there is a first increase that corresponds to the beginning of the plume passage and/or rain, but no significant decrease after the plume has left the area, due to the important contribution of radionuclides deposed on the ground to the total measured gamma dose rate. It should also be noted that, in the case of Fukushima, the initial increase of the gamma dose rate corresponds to the beginning of the rain (between 03:00 and 07:00 UTC on March 15th) and is likely due to the scavenging of a plume in altitude, since concentrations at ground level are not high enough to explain the deposition values and subsequent gamma dose rate measured (Quérel et al. 2016). This is consistent with the fact that Fukushima city is surrounded with high mountains, and the plume trajectory is likely to have crossed some of them. This is still another pitfall for modelling this episode, especially with a low resolution meteorological dataset. In this respect, results presented in Figure 66 are correct, since all ensembles encompass the observations. This means that most ensemble members manage to more or less reproduce this wet deposition episode, although there is an expected high variability between them. When looking at the dark blue lines, representing the 25-75th percentile, globally the observations are underestimated. The timing of the gamma dose rate increase is well encompassed by the models, which is usually difficult for deterministic simulations, as light rains are not well forecasted (Mathieu et al. 2018). This is allowed by the meteorological ensemble which has a sufficient variability in the rain timing (Figure 67).



(a) Met Office (b) RIVM

(c) IRSN (d) DTU

(e) EEAE/NCSRD (f) NMI MET

Figure 66 : Ensemble results of dose rate (nSv/h) at Fukushima city (60 kilometres north-westof FDNPP) for the Fukushima case study for several project participants, compared with observations (black dots). The y-axis are not the same on all figures.

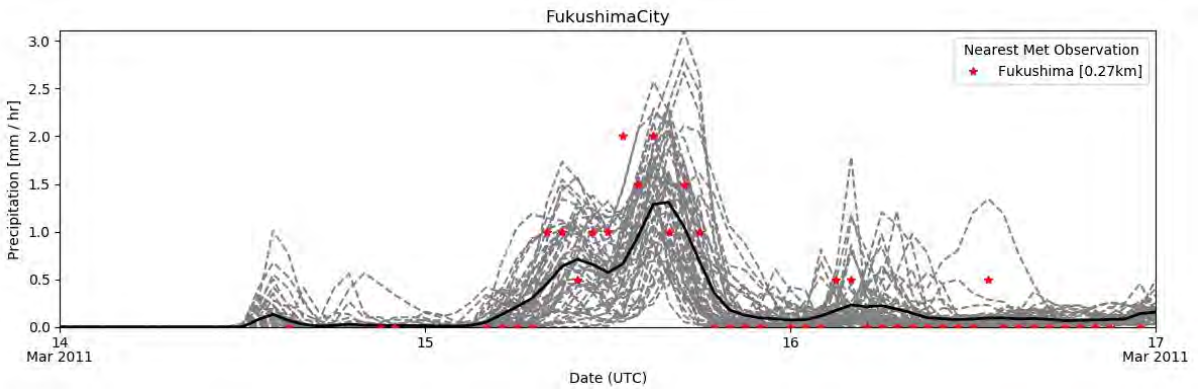


Figure 67: Ensemble precipitation forecast by the meteorological ensemble at Fukushima city, including meteorological observations at the nearest station (in red).

Iwaki

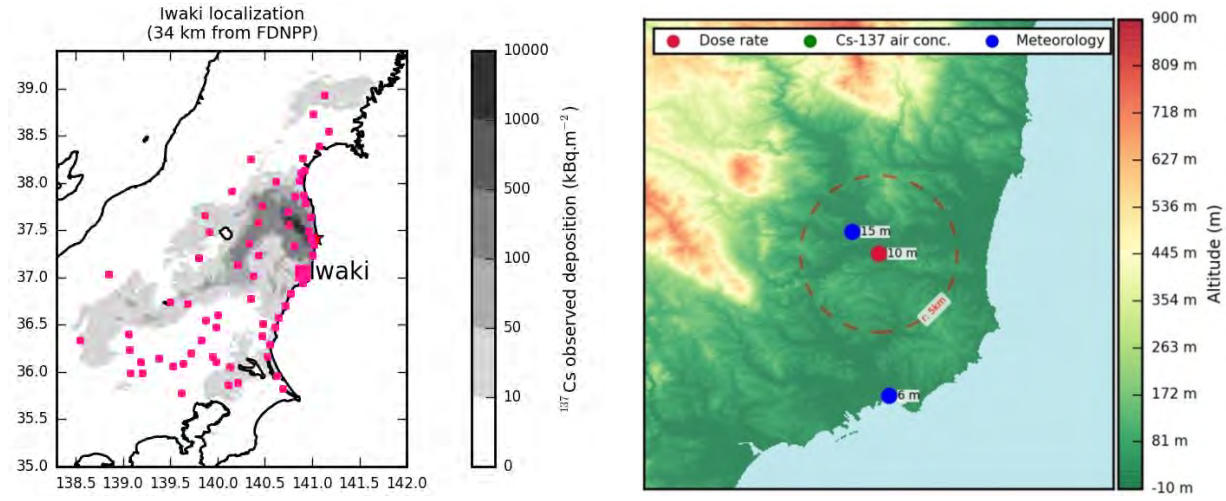


Figure 68 : Location of Iwaki (a); topography of the area around Iwaki and measuring stations (b). Courtesy of A. Quérel (Quérel et al. 2016)

Iwaki station is located 34 kilometres south of FDNPP, along the coast (Figure 68). Here, gamma dose rate observations on March 14-16th are typical of a plume passage without wet deposition, with a large peak during the plume passage and then a much lower gamma dose rate, due to dry deposition. There is observed precipitation from 16:00 to 22:00 UTC on March 15th, but at this moment, the plume that was scavenged contained a significantly lower concentration of radioactivity than during the initial peak that occurred between 16:00 and 23:00 UTC on March 14th. For this initial peak, the gamma dose rate is as high as 22 mSv/h, whereas it is about 5 mSv/h during the rain episode on March 15th. However, it appears that the meteorological ensemble forecasts rain in the night between March 14th and March 15th (Figure 69).

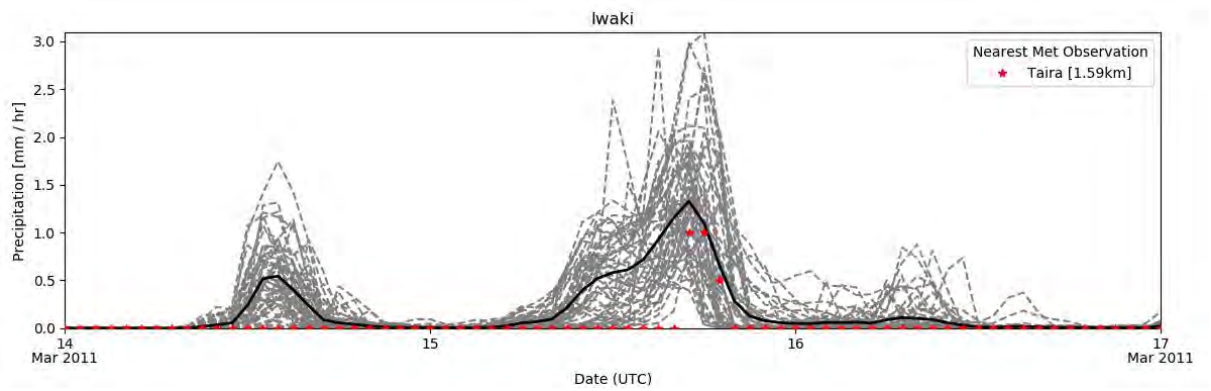


Figure 69: Ensemble precipitation forecast by the meteorological ensemble at Iwaki, and comparison with nearest meteorological observations available (in red).

This explains why the ensemble results for gamma dose rates feature a behaviour similar to the one showed at Fukushima city, with a large ground shine due to wet deposition (Figure 70). This also explains why the 25-75th percentiles have a tendency to overestimate the gamma dose rates after the first peak. Globally the ensembles encompass the observations, but it is difficult to see whether the peaks occurring on March 15th are only forecasted by some ensemble members. In summary, most members may be correct, not because they are accurately simulating the peaks, but because there is already a large variability in the gamma dose rate due to deposition after the initial increase of gamma dose rate.

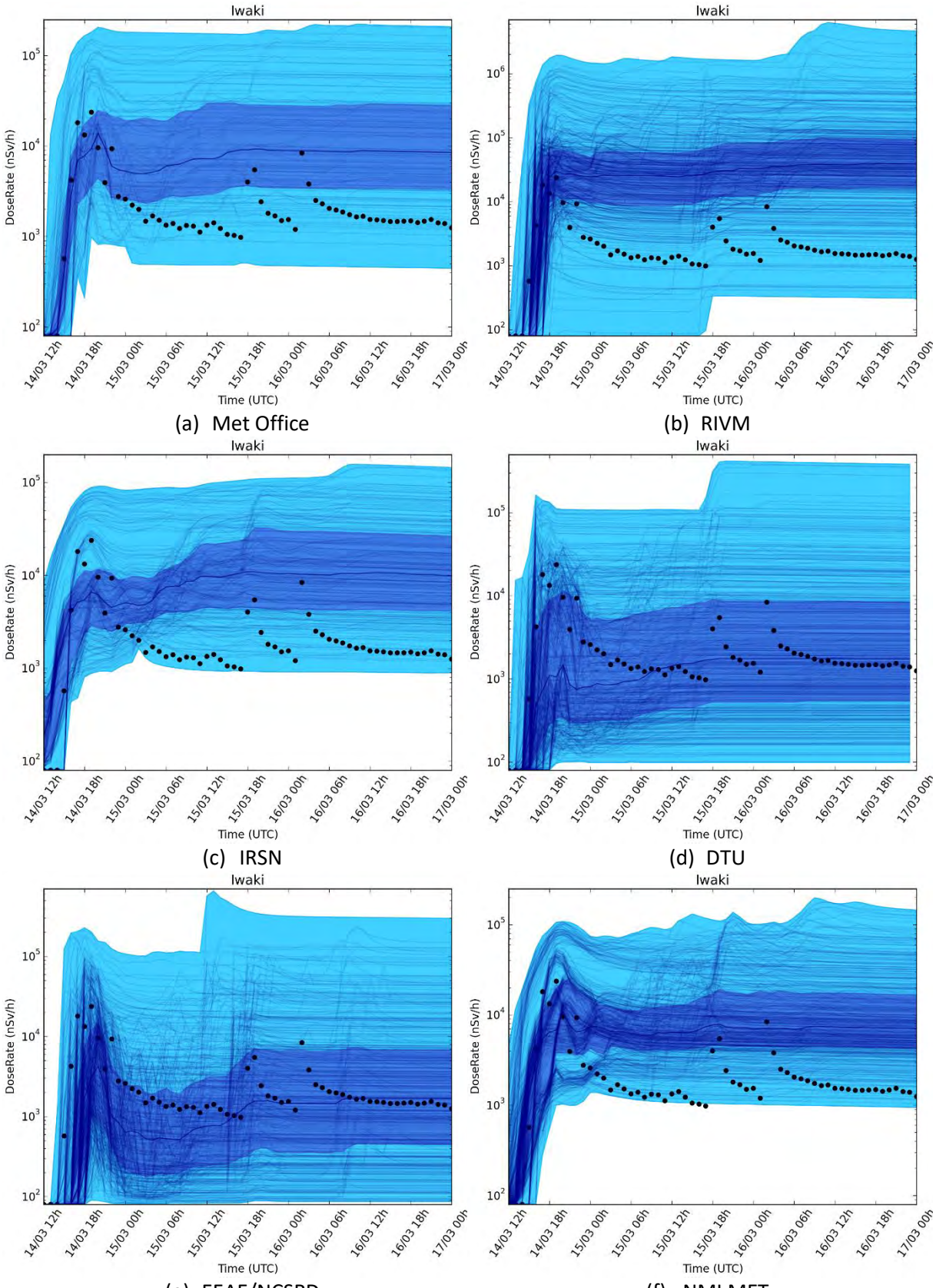


Figure 70 : Ensemble results of dose rate (nsv/h) at Iwaki (42 kilometres south-west of FDNPP) for the Fukushima case study for several project participants, compared with observations (black dots). The y-axis are not the same on all figures.

Statistical indicators

The Fukushima case allows for the presentation and discussion of additional outputs and indicators, not designed for decision making, but oriented toward ensemble verification. These indicators are briefly presented in this section.

Rank histograms

In a synthetic way, the rank histogram is a good method to estimate if an ensemble is representative of the observed uncertainty. The rank of an observation is determined by counting how many members of the ensemble are below this observation (Figure 71). Therefore, observations of rank zero are below all members, etc. The rank histogram shows the number of observations of each rank (ranks range between zero and the number of members in the ensemble), counted on all stations and time steps. Theoretically, a "perfect" rank histogram would be flat, indicating no bias in the model results.

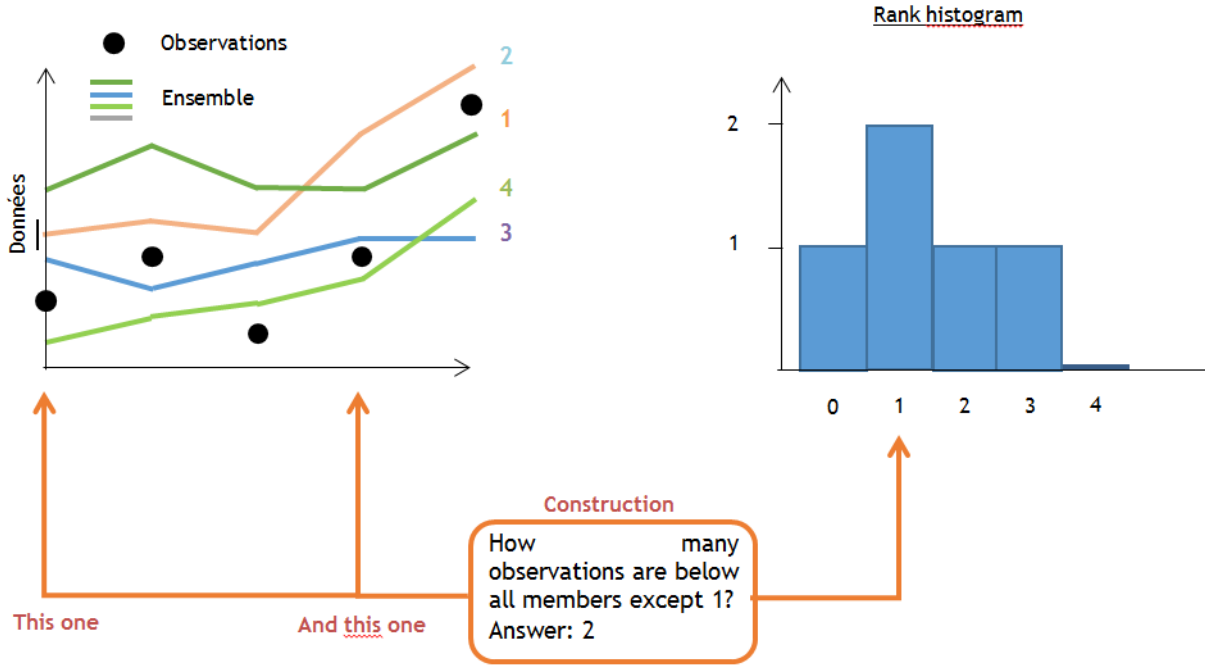


Figure 71 : Construction of a rank histogram.

The rank histograms of ^{137}Cs air concentrations (Figure 72) show a bias towards underestimation (the right-side bar is higher) for all participants. This was not obvious when looking at time series, although at some stations, there are peaks not encompassed by the ensemble (mostly on March 16th). It should also be noted that the look of the diagram is slightly different when there are more members in the ensemble; for instance, for DTU and RIVM the x-axis ranges from 0 to 460 while for Met Office there are only 102 simulations; therefore, the "bins" of the histogram are wider for the latter, since the number of observations is the same for all participants. It should also be noted that these rank diagrams are much better than those presented in a previous study with IRSN's model ldX, on the same dataset, with an ECMWF at a cruder spatial resolution (Le et al. 2017). However, those rank diagrams were made on the full duration of the accident (3 weeks).

An important point related to the rank diagram is whether the observation dataset is truly representative of what we are trying to simulate with the ensemble. For instance, Figure 53 shows that there is a lot of air concentration stations located in the southern area, while few of them are in the highly contaminated north-western area. The rank diagram gives the same weight to all observations, regardless of their actual value. Therefore, the rank diagram is not necessarily representative of the ensembles' performance, for instance, for the high contamination area. The use of a threshold on the

observations and/or the selection of a suitable subset of stations might be a useful complement for further analysis, with the drawback that enough observations have to be left to provide a useful rank diagram.

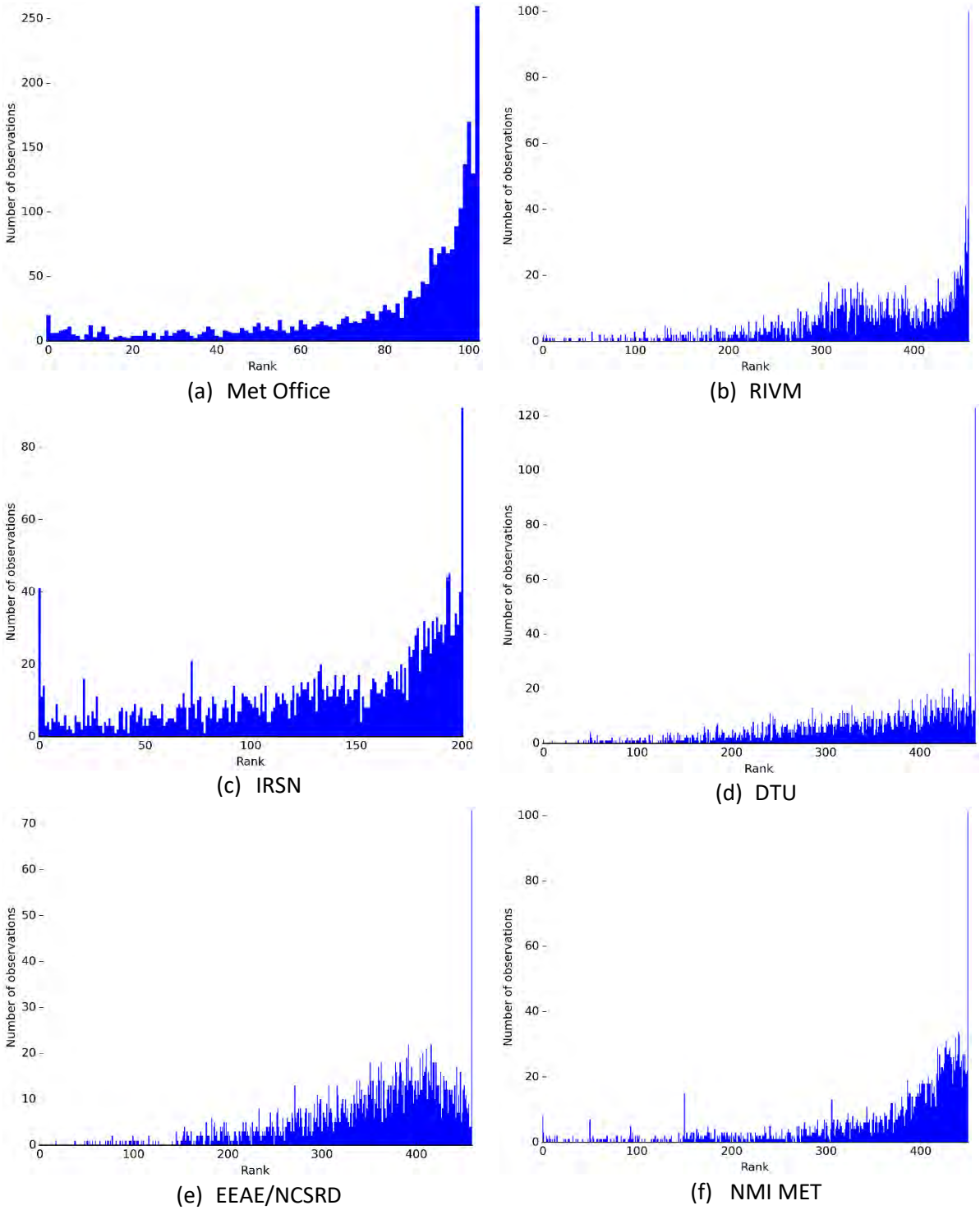


Figure 72 : Rank histogram for observations of ^{137}Cs air concentrations on all stations, for 6 participants.

The rank diagrams for gamma dose rates are shown Figure 73. Here, the rank diagrams are good although not totally flat. However, there is a problem of redundancy in the observations than for air concentrations. Indeed, for each station, the time series of gamma dose rate due to deposition gives several times the same information, the only difference between one time step and the next being the

radioactive decay. This explains the presence of some peaks corresponding to a large number of observations that have the same rank.

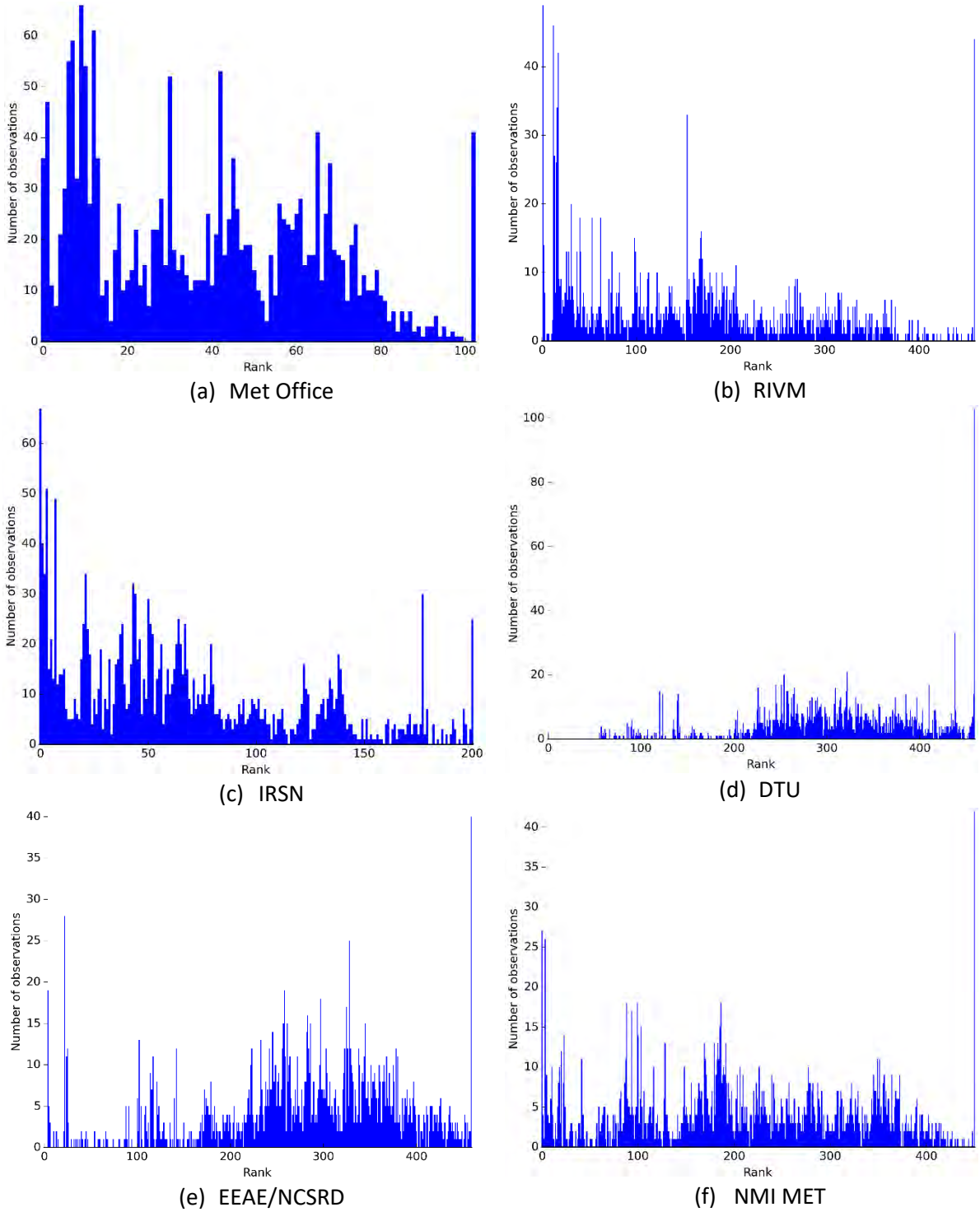


Figure 73 : Rank histogram for observations of gamma dose rates on all stations, for 6 participants.

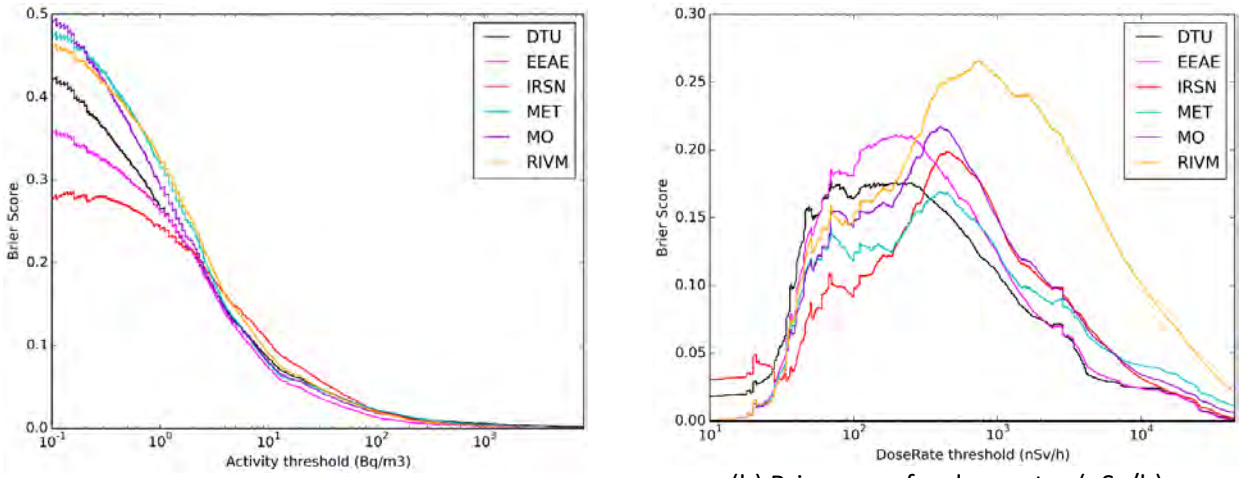
Brier score and CRPS / DRPS

The Brier score is defined as follows:

$$\mathrm{BS} = \frac{1}{N}\sum_{t=1}^{N}(p_i - o_i)^2$$

Where p_i is the probability of exceeding a given threshold according to the ensemble (between 0 and 1) and o_i is the outcome, given by the observation (0 if the threshold is not exceeded, 1 otherwise). Therefore, the closer to zero the score is, the better the ensemble is at accurately forecasting the probability of threshold exceedance.

The Brier score can be plotted for different thresholds. The 'mean' of all Brier scores for a threshold's range of variation is called the "continuous rank probability score" (CRPS). It can be obtained by calculating the area below the brier curve, divided by the length of the threshold variation range. In our case, since the variation over the threshold is discrete, the area below the curve is not calculated by an integral, but as a sum over several "bins" that discretize the x-abscissa. Then, the associated score is called the "discrete rank probability score" (DRPS). The Brier score tends toward zero when the threshold increases: for a high enough threshold, all observations as well all simulated values will be below the chosen value, reflecting the lack of events above the threshold rather an actual skill of the ensemble. In the case of a variable such as the ambient gamma dose rate, the Brier score for a very low threshold also tends toward zero, since the background radiation ensures that all observed and simulated values will be above this threshold. This is not necessarily the case for air activity concentrations: the values cannot go below zero, but there can be many occurrences where observations are equal to zero but some members give a positive value, which may result in a high Brier score.



(a) Brier score for air concentration of ^{137}Cs (Bq/m3) (b) Brier score for dose rates (nSv/h)

Figure 74 : Brier score for 6 participants, for ^{137}Cs concentrations (left) and dose rates (right).

The Brier scores for air concentrations (Figure 74(a)) are consistent between the 6 participants. For air concentrations, there is some variability between the participants for lower thresholds (between 0.1 and 1 Bqm^3), where the score ranges from 0.28 to 0.5. Above 1 Bq/m^3, which corresponds to a significant observation value, all results are very close to each other. As far as dose rates are concerned, there is a peak in Brier score, which is reached around 200 nSv/h (EEAE) to 1000 nSv/h (RIVM). Some participants are better when low dose rates are considered (mostly due to dry deposition, or small wet

deposition), others are better for high dose rates, i.e., forecasting peaks of dose rates. For instance, EEAE and DTU have the highest Brier scores (meaning that they have more difficulty in predicting threshold exceedance) for low thresholds, but are among the best scores for thresholds above 1000 nSv/h. Conversely, DTU and, to a lesser extent, IRSN and Met Office, have very good scores for low thresholds and perform less well for high thresholds. Looking at the DRPS, which represents the value averaged over all thresholds, gives an overview of the overall performance of the ensembles. Table 15 shows that the best DRPS (i.e. closest to zero) are obtained by EEAE for air concentration and dose rates, although that was not necessarily visible on the figures, which are drawn in logarithmic scale.

Table 15: DRPS for air concentration and dose rates, for all participants. The best scores are highlighted in green.

DRPS	DTU	EEAE/NCSRD	IRSN	NMI MET	Met Office	RIVM
Air Concentration (× 10^{-3})	4.01	1.54	2.36	2.56	2.22	2.01
Dose Rate (× 10^{-2})	2.98	2.80	3.66	3.04	3.71	11.46

Skill-spread diagram

The skill-spread diagram consists in comparing the "skill" of a reference simulation by comparison to observations, to the "spread" of the ensemble. The rationale behind this is that the e-del-to-observations discrepancy is linked to the model uncertainties, which is precisely what is supposed to be represented by the spread of the ensemble (Jones et al. 2019; Kaufmann and Rüdisühli 2019). Therefore, comparing the Root Mean Square Error (RMSE) of the "deterministic" simulation (the "skill") with the ensemble's spread given by its standard deviation is a way to evaluate whether the uncertainties are properly taken into account in the ensemble construction. The question arises how to properly determine the reference simulation to which the ensemble's spread will be compared. One possibility is to use the ensemble median (or mean) as a reference, which was done here. Indeed, it may be possible to use the control member of the meteorological ensemble (if available) for a reference dispersion simulation, but there is no "reference" source term, and source terms are responsible for a huge part of the overall uncertainty. A good skill-spread diagram would be close to the y = x axis.

Figure 75 shows the skill-spread diagrams of all participants for ^{137}Cs air activity concentrations on all stations. It globally shows that the ensembles' spreads have a tendency to underestimate the error of the ensemble median, especially for high concentrations. This may be related to the fact that the rank diagrams for this variable show a bias (Figure 72). The skill-spread diagrams for dose rates (not shown) feature the same tendency to underestimate the error of the median. However, this indicator is to be taken with care, as the use of a "reference" simulation (as done in meteorology with the control member) would have much more sense. One may, for instance, prefer to draw skill-spread diagrams for each individual source term separately (not shown here).

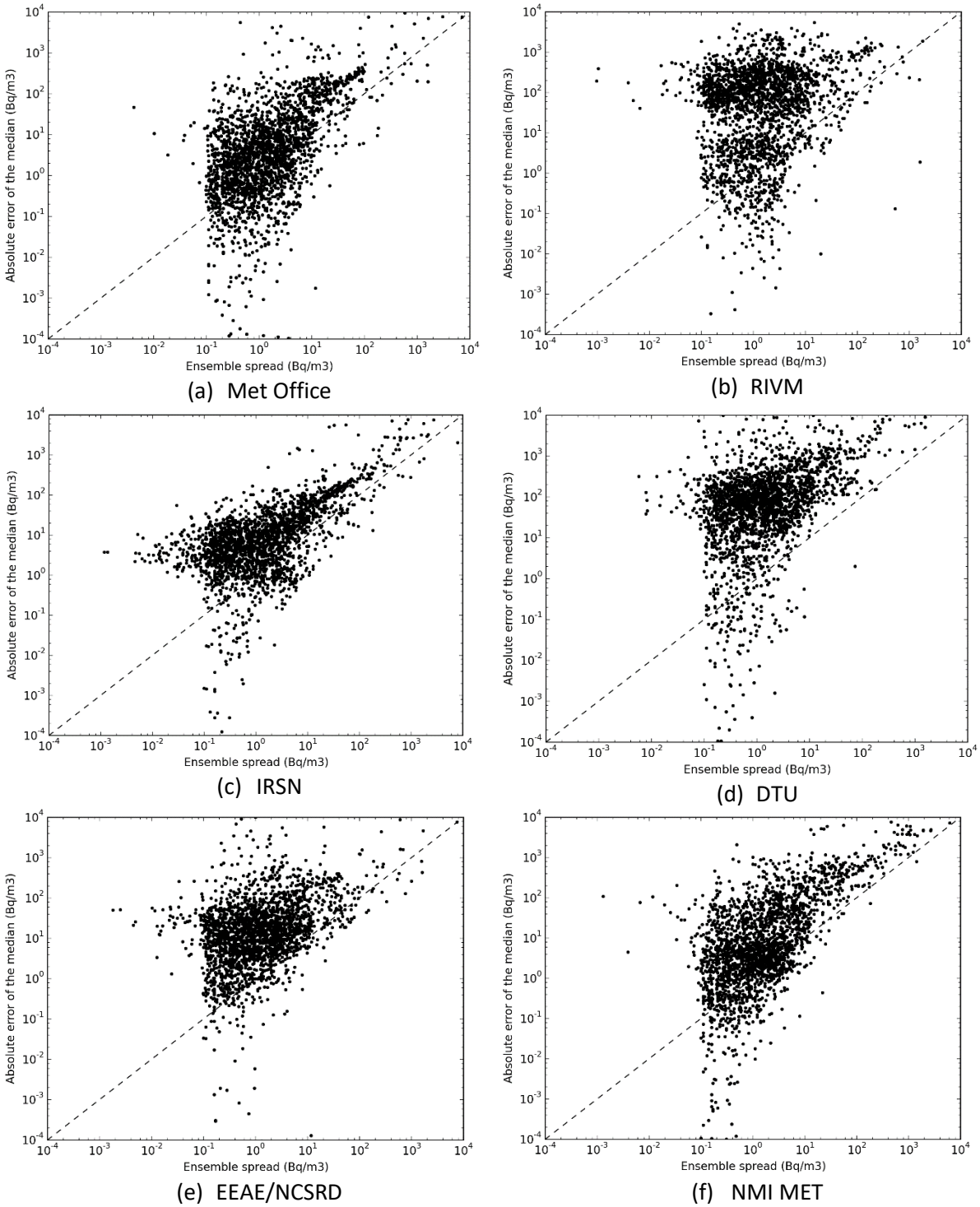
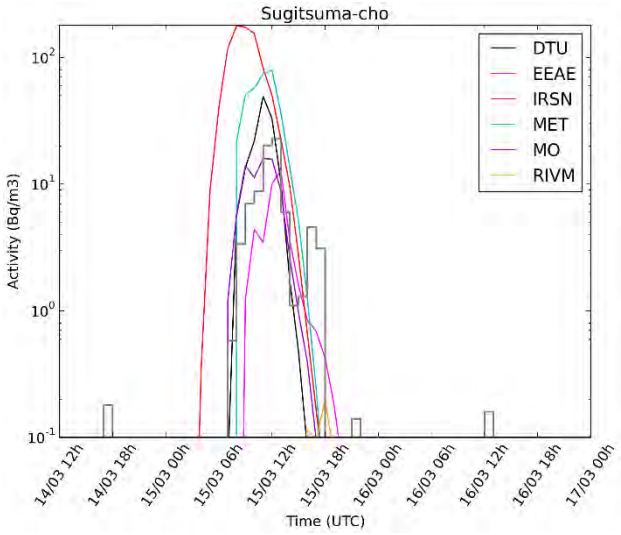


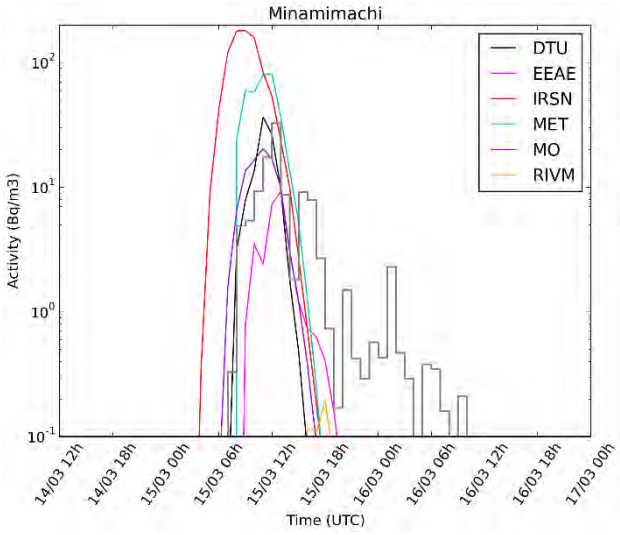
Figure 75 : Skill-spread diagram for observations of ^{137}Cs air concentrations on all stations, for 6 participants.

Performance evaluation of the ensembles median

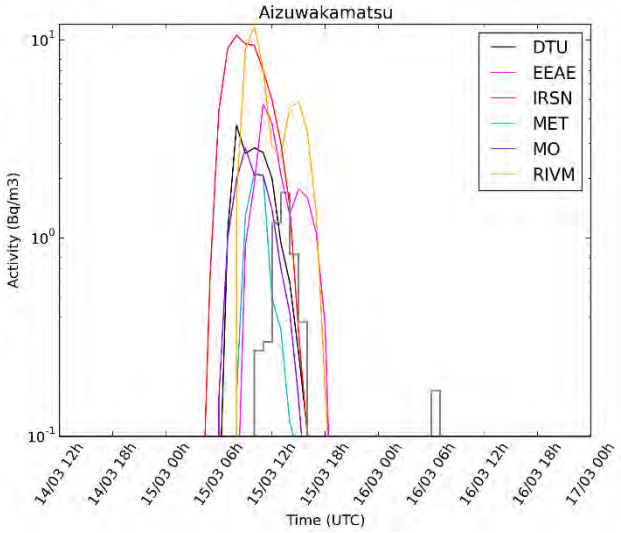
In addition to the ensemble scores described in the previous sections, it is interesting to compute some deterministic model-to-observations statistical scores. For this, the ensemble median may be used, as a "deterministic" simulation resulting from taking into account all uncertainties. It is interesting to determine whether these scores may be higher than those of a "best estimate".



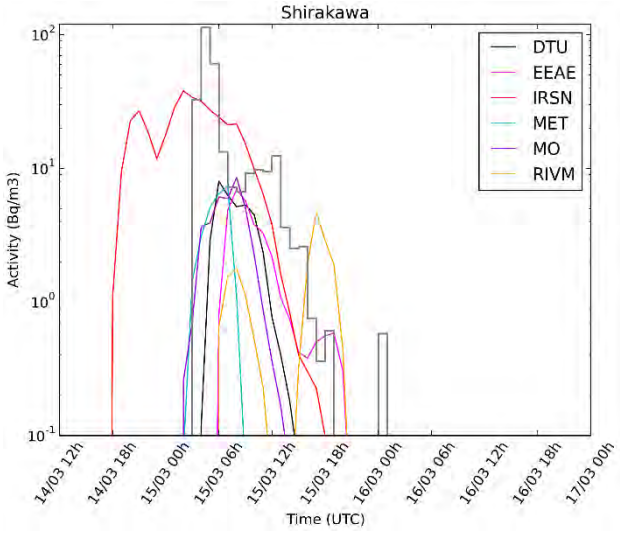
(a) Sugitsuma-cho



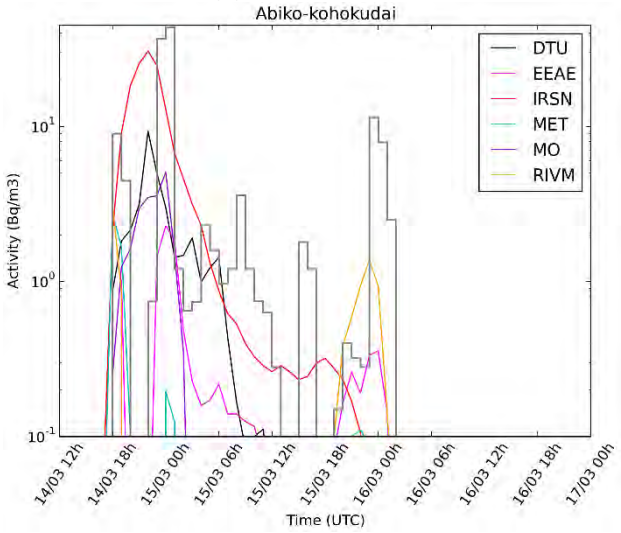
(b) Minamimashi



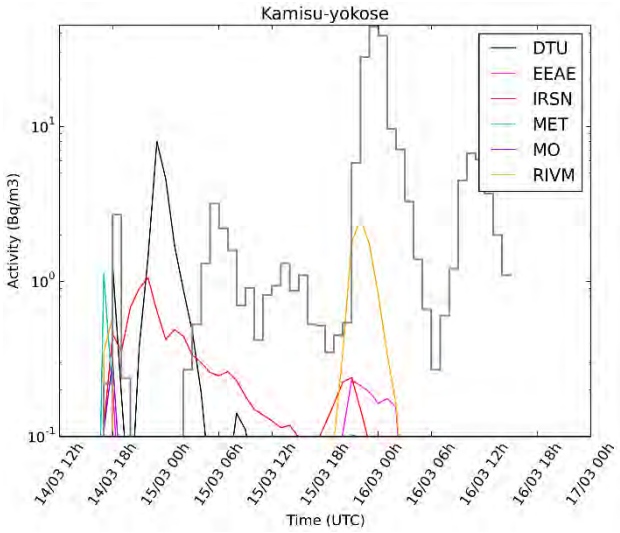
(c) Aizuwakamastsu



(d) Shirakawa



(e) Abiko-kohokudai



(f) Kamisu-yokose

Figure 76 : Median of ^{137}Cs activity at 6 stations for the Fukushima case study for several project participant, compared with observations (in grey).

Figure 76 shows the different stations for air activity concentrations that have already been presented in the previous sections. Here, the observations are in grey, and the median of the six participants' ensembles are given in various colours. As noted before, the observations are well represented at Sugitsuma-cho by most participants, but not at Minamimashi, while the two stations are very close to each other (in the same meteorological cell), denoting the influence of subgrid-scale processes not taken into account by the simulations. The ensembles' medians are better at representing the high values of stations located in the north-western area (Figures (a) to (d)) than low values farther south (Figures (e) and (f)). Globally the RIVM ensemble tends to have a very low median compared to the other ensembles on these stations, although it is one of the most widespread ensembles with a good ensemble score on air concentrations.

The statistical scores of the ensemble medians obtained by the six participants are computed in this section. The chosen scores are the Root Mean Square Error (RMSE), Pearson Correlation Coefficient (Corr), Figure of Merit in Time (FMT), and FAC2 (respectively FAC5) which represent the proportion of simulated values that are within a factor 2 (respectively 5) of the observations. These scores are defined below.

If $(o_i)_{i\in[1,N]}$ is the series of all scalar observations and $(s_i)_{i\in[1,N]}$ the values of the simulation corresponding to those observations, then we can define the RMSE and the Pearson correlation by:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(s_i - o_i)^2}$$

$$\mathrm{Corr} = \frac{\sum_{i=1}^{N}(s_i - \overline{s_i})(o_i - \overline{o_i})}{\sqrt{\sum_{i=1}^{N}(s_i - \overline{s_i})^2}\sqrt{\sum_{i=1}^{N}(o_i - \overline{o_i})^2}}$$

Also, FAC *N* is defined as the percentage of s_i which are located between $N \times o_i$ and o_i/N.

The Figure of Merit in Time (FMT) is a score based on a threshold. If $\boldsymbol{A}_m$ and $\boldsymbol{A}_p$ are the ensembles of time steps where respectively the measurement and the prediction are above the fixed threshold, and if $|\,.\,|$ is the cardinality of an ensemble (i.e. the number of elements in the ensemble), then:

$$\mathrm{FMT} = 100 \times \frac{|\boldsymbol{A}_m \cap \boldsymbol{A}_p|}{|\boldsymbol{A}_m \cup \boldsymbol{A}_p|}$$

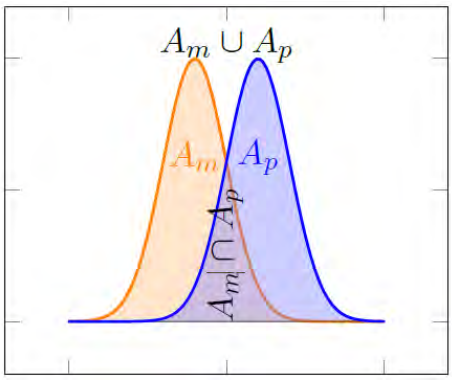


Figure 77 : Illustration of the Figure of Merit in Time between two time series. The abscissa usually represents the time, while the ordinate represents the variable of interest (here, concentration or dose rate). From Le et al. (2019).

In practice, the FMT represents the ability of a simulated time series to overlap the time series of the observations (Figure 77). If, for instance, a peak in the observations is reproduced by the model with an incorrect timing, and there is no overlap of the peaks, then FMT will be zero. In general, a "perfect" model-to-observations comparison (assuming there are no errors) would be given by RMSE=0, Corr=1, FMT=1, FAC2=FAC5=1. The RMSE represents the ability of the model to be unbiased, but gives a large weight to high values, as it is not normalized. The correlation and FMT both represent the ability of the model to correctly represent the temporal variation of the data, regardless of the magnitude of the observations. FAC2 and FAC5 represent the ability of the model to give close values to the observations, but grant the same weight to all observations. The selection of a particular subset of stations could be discussed with regard to whether it is more important to accurately represent the observations close to the source, farther from the source, or in particular regions.

It should be noted that, here, the observation errors are not taken into account, since they are considered negligible compared to the simulations' uncertainties. Table 16 shows the statistical indicators of all participants' median for ^{137}Cs air concentrations, and Table 17 for ambient gamma dose rates. Here, the threshold for FMT is chosen so that all values are above the threshold (i.e. zero). It is interesting to note that the "best score" (highlighted in green) is not obtained by the same participant, depending on the statistical score considered. Overall, four out of six participants obtain the "best score" at least for one score and one variable, denoting that no particular model and configuration is significantly better than the others. Some may be better at forecasting the plume passage, others at simulating deposition; some may be better close to the source, others at farther locations.

Table 16: Ensemble median scores for ^{137}Cs air concentration results. The best score is highlighted in green.

Air concentration scores	DTU	EEAE/NCSRD	IRSN	NMI MET	Met Office	RIVM
RMSE (Bq/m^3) (× 10^2)	5.57	1.83	2.07	1.90	2.54	1.93
Correlation	0.23	0.36	0.22	0.45	0.23	0.26
FMT (%)	41.7	46.5	63.8	26.3	32.6	28.9
FAC 2 (%)	10.3	11.8	18.7	4.9	8.5	6.6
FAC 5 (%)	24.5	26.4	41.7	13.2	18.8	16.2

Table 17: Ensemble median scores for ambient gamma dose rate results. The best score is highlighted in green.

Dose rate scores	DTU	EEAE/NCSRD	IRSN	NMI MET	Met Office	RIVM
RMSE (nSv/h) (× 10^3)	4.95	4.85	4.92	6.70	4.81	9.23
Correlation	0.33	0.34	0.37	0.15	0.38	0.04
FMT (%)	81.3	79.3	97.8	93.5	97.2	89.7
FAC 2 (%)	37.7	38.6	41.4	41.4	43.4	20.1
FAC 5 (%)	74.1	73.3	77.5	83.8	77.1	47.7

It is interesting also to compare these scores with those obtained by "best estimate" configurations. For instance, in Saunier et al. (2016), statistical scores are computed using IRSN's model ldX, various source terms and meteorological data. These scores are computed on the three week period of the Fukushima accident; nonetheless, it may be noted that for ECMWF deterministic data at 0.125° resolution, the FAC5 obtained on air concentration ranges from 25 to 32 depending on the source term used (all of those being included in the simulations made in the present study). Therefore, it is probable that all source terms bring some information and contribute to improving the ensemble median, leading to a better score (for the same dispersion model *ldX*) for the ensemble than for the "best estimate" with any separate source term. It is also interesting to note that the "best" ensemble, in

terms of encompassing all observations, which corresponds to EEAE if CRPS/DRPS is to be chosen as an indicator, is not necessarily the "best" median in terms of scores.

Also, it is worth mentioning that among the source terms presented in Table 13, two of them were constructed using IRSN's model *IdX*. This may bias the results in favour of this model, although the meteorological data used were not the same: ECMWF deterministic meteorology was used to inverse gamma dose rates in Saunier et al. (2013a), while a high-resolution meteorological data provided by the Japan Meteorological Research Institute was used to inverse air concentrations in (Saunier et al. 2016). Therefore, there is no clear evidence that IRSN's ensemble should perform better than others with these source terms. Besides, another study using a similar ensemble of source term and a meteorological ensemble from ECMWF with a more crude resolution, and carried out with the IdX model on the 3 weeks of the accident, showed far lower performance on the air concentration stations (Le et al. 2017; Le et al. 2019).

Conclusions

The Fukushima case study shows a good consistency between the participants' ensembles performance. When using the ECMWF meteorological ensemble and the nine source terms, without additional perturbations, the ensembles show a very large spread and encompass the observations, although there is a bias toward underestimation. Four out of six participants carried out cross-simulations with all possible combinations of source terms and meteorological members (DTU, NMI, RIVM and EEAE), and therefore, the differences between their results only stem from the use of different dispersion models and configurations. Additional perturbations on physical parameters introduced by IRSN do not seem to significantly change the ensembles' spread.

This report presented several indicators to evaluate the performance of an ensemble, that is, its ability to represent the overall uncertainty. The rank diagram is a synthetic way to evaluate how well an ensemble can encompass the observations variability. A good rank diagram is a necessary, but not sufficient condition for an ensemble to be deemed correct. In practice, in meteorology for instance, rank diagrams are never flat, since the sources of uncertainties are never totally and properly taken into account. In that respect, the rank diagrams for air concentrations and gamma dose rates presented here seem to present a good performance, despite a bias on air concentrations. Other indicators are the Brier score and the DRPS; the Brier score represents the ability of the ensemble to forecast the probability of exceeding a given threshold, and the DRPS is its average over all thresholds. The drawback of this score is that it is difficult to interpret in itself, but it is useful to compare several ensembles. Here, the ensembles performance seem good but show some significant variability, denoting that the inter-model variability described in Korsakissok et al. (2019) still has its importance here. Finally, the skill-score diagrams represent the ability of the ensemble to represent the model-to-observation error, which itself is a measure of the uncertainties associated to the simulation. The drawback, here, is that a "reference" simulation has to be defined. Although the meteorological ensemble has a control member, there is no "reference" source term. The skill-score diagrams shown in this report were drawn using the ensembles' medians as reference. They tend to show an underestimation of the modelling errors, meaning that not all sources of uncertainties are taken into account. However, further investigation should be made by taking into account the different source terms separately.

Finally, the performance of the ensembles' medians was assessed by comparison to the observations, using traditional deterministic statistical scores. It is interesting to note that, depending on the score and the variable of interest (gamma dose rate or air concentrations), the "best" ensemble was not the

same. Four out of six ensembles obtained the best score at least for one indicator and one variable. It is also worth mentioning that these scores seem to be better than those obtained with the same model and deterministic configurations with different meteorological data and source terms. Although this finding has to be further analysed, it seems to infer that there is an additional value provided by the ensemble median by comparison to the "best estimates" given by deterministic simulations. In other words, even configurations that are not optimum in themselves provide useful information and can improve the overall results.

Future work includes refining the meteorological ensembles. In particular, here, a single 3-day forecast was used, but, to simulate the three-week period of the accident, it would be necessary to combine several ensemble forecasts together, which raises the issue of temporal continuity between the ensemble members. Other aspects of future research would include working on reducing the number of ensemble members (a topic addressed in Bedwell et al. (2019)) and better taking into account all sources of uncertainties, including those highlighted by the variability between the participants.

Acknowledgement

We acknowledge use of ECMWF's computing and archive facilities for the re-running of the ensemble prediction system.

References

Achim, P., M. Monfort, G. Le Petit, P. Gross, G. Douysset, T. Taffary, X. Blanchard and C. Moulin (2014). "Analysis of Radionuclide Releases from the Fukushima Dai-ichi Nuclear Power Plant Accident Part II." Pure and Applied Geophysics **171**(3): 645-667.

Bedwell, P., I. Korsakissok, S. Leadbetter, R. Périllat, C. Rudas, J. Tomas and J. Wellings (2019). D 9.5.5 Guidelines for the use of ensembles in the description of uncertainty in atmospheric dispersion modelling: operational applications in the context of an emergency response. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Berge, E., H. Klein, M. Ulimoen, S. Andronopoulos, O. C. Lind, B. Salbu and N. Syed (2019). D 9.5.2 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the Western Norway case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Chino, M., H. Nakayama, H. Nagai, H. Terada, G. Katata and H. Yamazawa (2011). "Preliminary estimation of release amounts of ^{131}I and ^{137}Cs accidentally discharged from the Fukushima Daiichi Nuclear Power Plant into the atmosphere." Journal of nuclear science and technology **48**(7): 1129-1134.

Hirao, S., H. Yamazawa and T. Nagae (2013). "Estimation of release rate of iodine-131 and cesium-137 from the Fukushima Daiichi nuclear power plant." Journal of nuclear science and technology **50**(2): 139-147.

IAEA. (2012). "Fukushima monitoring database." from https://iec.iaea.org/fmd/search_by_dataset.aspx.

Jones, A., S. Leadbetter and B. Evans (2019). Assessing the performance of a global met ensemble for radiological dispersion. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Katata, G., M. Chino, T. Kobayashi, H. Terada, M. Ota, H. Nagai, M. Kajino, R. Draxler, M. C. Hort, A. Malo, T. Torii and Y. Sanada (2015). "Detailed source term estimation of the atmospheric release for the Fukushima Daiichi Nuclear Power Station accident by coupling simulations of an atmospheric dispersion model with an improved deposition scheme and oceanic dispersion model." Atmos. Chem. Phys. **15**(2): 1029-1070.

Kaufmann, P. and S. Rüdisühli (2019). Dispersion model uncertainty simulation using a limited area ensemble model. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Kobayashi, T., H. Nagai, M. Chino and H. Kawamura (2013). "Source term estimation of atmospheric release due to the Fukushima Dai-ichi Nuclear Power Plant accident by atmospheric and oceanic dispersion simulations." Journal of nuclear science and technology **50**(3): 255-264.

Korsakissok, I., A. Mathieu and D. Didier (2013). "Atmospheric dispersion and ground deposition induced by the Fukushima Nuclear power plant accident : a local-scale simulation and sensitivity study." Atmospheric Environment **70**: 267-279.

Korsakissok, I., R. Périllat, G. Geertsema, H. De Vries, T. Hamburger, J. Tomas, S. Andronopoulos, P. Bedwell, T. Charnock, I. Ievdin, S. Leadbetter, T. Pazmandi, C. Rudas, R. Scheele, A. Sogachev, P. Szanto and J. Wellings (2019). D 9.5.1 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the REM case studies. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Le, N. B. T., I. Korsakissok, V. Mallet, A. Mathieu, R. Périllat and D. Didier (2019). "Uncertainty study on atmospheric dispersion simulations using meteorological ensembles with a Monte Carlo approach, applied to the Fukushima nuclear accident." Atmospheric Environment(To be submitted).

Le, N. B. T., R. Périllat, I. Korsakissok, V. Mallet, A. Mathieu and D. Didier (2017). Uncertainty studies on an atmospheric dispersion model with Monte Carlo method. 18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bologna, Italy.

Leutbecher, M. and T. N. Palmer (2008). "Ensemble forecasting." J. Comp. Phys **227**: 3515-3539.

Liu, Y., J.-M. Haussaire, M. Bocquet, Y. Roustan, O. Saunier and A. Mathieu (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.

Mathieu, A., M. Kajino, I. Korsakissok, R. Périllat, D. Quélo, A. Quérel, O. Saunier, T. T. Sekiyama, Y. Igarashi and D. Didier (2018). "Fukushima Daiichi–derived radionuclides in the atmosphere, transport and deposition in Japan: A review." Applied Geochemistry **91**: 122-139.

Mathieu, A., I. Korsakissok, D. Quélo, J. Groëll, M. Tombette, D. Didier, E. Quentric, O. Saunier, J.-P. Benoit and O. Isnard (2012). "Atmospheric dispersion and deposition of radionuclides from the Fukushima Daiichi nuclear power plant accident." Elements **8**: 195-200.

Molteni, F., R. Buizza, T. N. Palmer and T. Petroliagis (1996). "The ECMWF Ensemble Prediction System: methodology and validation." Q. J. R. Meteorol. Soc. **122**: 73-119.

Oura, Y., M. Ebihara, H. Tsuruta, T. Nakajima, T. Ohara, M. Ishimoto, H. Sawahata and W. Nitta (2015). "A Database of Hourly Atmospheric Concentrations of Radiocesium (134Cs and 137Cs) in Suspended Particulate Matter Collected in March 2011 at 99 Air Pollution Monitoring Stations in Eastern Japan." Journal of Nuclear and radiochemical sciences **15**(2): 1-12.

Quérel, A., D. Quélo, Y. Roustan, A. Mathieu, M. Kajino, T. Sekiyama, K. Adachi, D. Didier, Y. Igarashi and T. Maki (2016). Fukushima: Lessons Learned on Wet Deposition from a Combined Analysis of Radiation Dose Rate and Volume Activity Measurements of 137Cesium. Goldschmidt, Yokohama.

Saunier, O., A. Mathieu, D. Didier, M. Tombette, D. Quélo, V. Winiarek and M. Bocquet (2013a). "An inverse modeling method to assess the source term of the Fukushima nuclear power plant accident using gamma dose rate observations." Atmos. Chem. Phys. Discuss. **13**(6): 15567-15614.

Saunier, O., A. Mathieu, D. Didier, M. Tombette, D. Quélo, V. Winiarek and M. Bocquet (2013b). "An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations." Atmos. Chem. Phys. **13**(22): 11403-11421.

Saunier, O., A. Mathieu, T. Sekiyama, M. Kajino, K. Adachi, M. Bocquet, T. Maki, T. Higarashi and D. Didier (2016). A new perspective on the Fukushima releases brought by newly available 137Cs air concentration observations and reliable meteorological fields. 17th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Budapest (Hungary).

Stohl, A., P. Seibert, G. Wotawa, D. Arnold, J. F. Burkhart, S. Eckhardt, C. Tapia, A. Vargas and T. J. Yasunari (2011). "Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition." Atmospheric Chemistry and Physics Discussions **11**(10): 28319-28394.

Terada, H., G. Katata, M. Chino and H. Nagai (2012). "Atmospheric discharge and dispersion of radionuclides during the Fukushima Dai-ichi Nuclear Power Plant accident. Part II: verification of the source term and analysis of regional-scale atmospheric dispersion." Journal of Environmental Radioactivity **112**: 141-154.

Tsuruta, H., Y. Oura, M. Ebihara, T. Ohara and T. Nakajima (2014). "First retrieval of hourly atmospheric radionuclides just after the Fukushima accident by analyzing filter-tapes of operational air pollution monitoring stations." Sci. Rep. **4**.

Winiarek, V., M. Bocquet, N. Duhanyan, Y. Roustan, O. Saunier and A. Mathieu (2014). "Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations." Atmospheric Environment **82**(0): 268-279.

Winiarek, V., M. Bocquet, O. Saunier and A. Mathieu (2012). "Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant." J. Geophys. Res. **117**(D5): D05122.

Yumimoto, K., Y. Morino, T. Ohara, Y. Oura, M. Ebihara, H. Tsuruta and T. Nakajima (2016). "Inverse modeling of the 137Cs source term of the Fukushima Dai-ichi Nuclear Power Plant accident constrained by a deposition map monitored by aircraft." Journal of Environmental Radioactivity **164**: 1-12.

D 9.5.4 Uncertainty propagation through a terrestrial food chain and dose model

Lead Authors: Thomas Hamburger, Florian Gering, Ievgen Ievdin, Stefan Schantz

Review(s): Gertie Geertsema, Irène Korsakissok

Table of Contents

Introduction 106

Model setup 106

Decision support system JRODOS 107

Atmospheric dispersion model (RIMPUFF) 107

Terrestrial Food Chain and Dose Module (FDMT) 107

Endpoints 108

Ensemble setup 108

Meteorological and source term uncertainties (MET+ST) 109

Uncertainties introduced to the food chain (FDMT) 110

Uncertainties in meteorology, source term and food chain (MET+ST+FDMT) 111

Results 112

Conclusions 114

References 114

Appendix A 116

Introduction

The task of Subtask 1 within Task 1.3 of Work package 1 (WP1) within the CONFIDENCE project is to analyse the propagation of uncertainties from ensemble dispersion simulations through a terrestrial food chain and dose model. Ensemble atmospheric dispersion calculations were performed in Task 1.2 of WP1. The ensemble calculations were based on hypothetical accident scenarios at the Borssele nuclear power plant (Netherlands). The scenarios as well as methods and results of the ensemble calculations are described in detail in CONFIDENCE deliverables 9.4 (Geertsema et al. 2019) and 9.5.1 (Korsakissok et al. 2019). The endpoints of Task 1.2 covered activity concentrations in air as well as deposited activity concentrations of the released radionuclides, and dose calculations through the pathways of cloud and ground shine (external dose), and inhalation (internal dose). Subtask 1 (Task 1.3) will extend the available endpoints of the ensemble calculations to activity concentrations in the food chain, i.e. feedstuffs and foodstuffs, as well as the internal dose through ingestion. Uncertainties that may occur in the food chain are added to the existing uncertainties that are introduced by meteorological uncertainties and uncertainties in the source term. Ensemble dispersion calculation input to Subtask 1 (Task 1.3) is taken from the first REM scenario using a set of 10 meteorological ensemble members and a subset of five source term ensemble member from the long release scenario (FASTNET). For detailed information on the input for the ensemble members see CONFIDENCE deliverable D9.1. Within this deliverable, in particular, uncertainties related to meteorology are described in Leadbetter et al. (2018) and those related to source terms in (Mathieu et al. 2018). Further description of the meteorological scenarios used for the REM case can be found in Geertsema et al. (2019) and Korsakissok et al. (2019). Further description of the design of the source term ensemble may be found in Chevalier-Jabet (2019).

Model setup

The basic model setup allows the calculation of activity concentrations in the air as well as on several surfaces and in consumption products. Following the calculation of activity concentrations the organ doses and effective dose through different pathways (cloud shine, ground shine, inhalation, and ingestion) is estimated.

The principle of the model chain is illustrated in Figure 78. Atmospheric dispersion calculations based on meteorological data and source term information on the release are performed to calculate the atmospheric transport of the released radionuclides. The deposition of the airborne radionuclides on different surfaces is modelled in the next step. The transport within the human food chain is modelled according to the model parameter settings. Subsequently, organ doses and effective dose to the population are modelled using the available pathways of cloud shine, ground shine, inhalation, and ingestion.

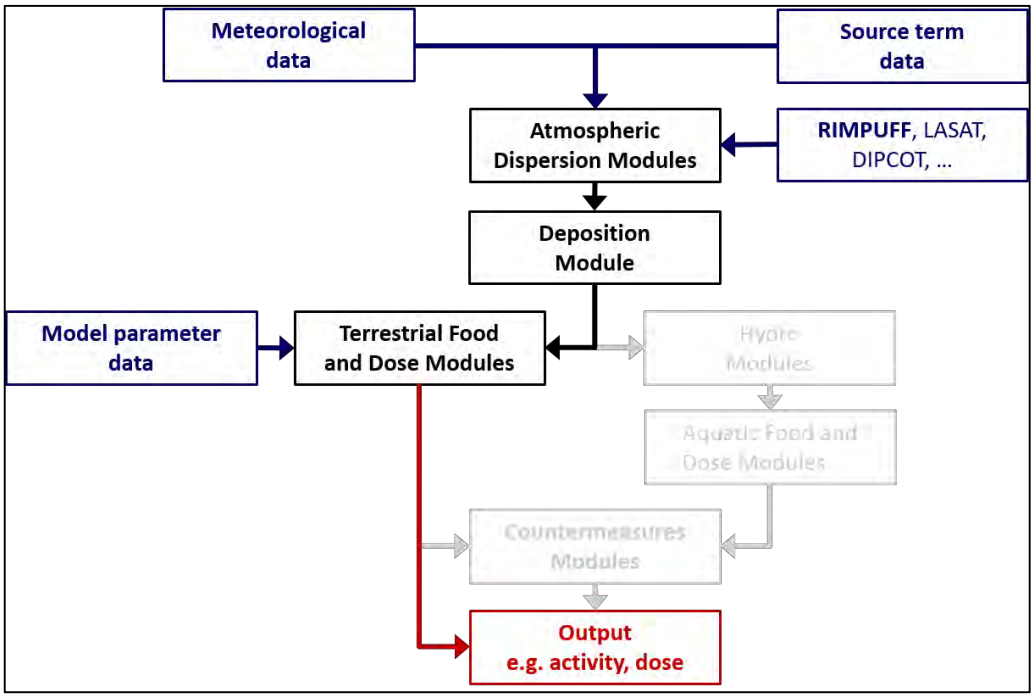


Figure 78 Setup of model chain and input parameter for the calculation of activity concentrations and dose resulting from a release of radionuclides. Blue: input data, black: used model chain, red: results, grey: available but not used model chain.

The different tools and modules used in the study to propagate the uncertainties from dispersion calculation through the food chain and to the dose are briefly described below.

Decision support system JRODOS

The Real-time Online Decision Support System (RODOS) is operated at BfS - as well as in numerous other European countries - in order to perform dispersion and dose calculations in the event of a nuclear accident or other radionuclide releases into the environment and to assess the potential consequences. RODOS was developed after the Chernobyl accident and supported by the European Commission's Research and Technological Development Framework Programmes (Ehrhardt 1997; Raskob et al. 2012). The resent development of the Java-based RODOS version is named JRODOS.

The JRODOS program serves as a framework for several models for data processing, physical modelling and user interactions. The model tasks include the processing of meteorological input data, the calculation of the dispersion in the atmosphere, the simulation of the transfer of radionuclides within the human food chain, and the estimation of radiation exposure experienced by the population.

Atmospheric dispersion model (RIMPUFF)

The RIMPUFF model (Risø Mesoscale PUFF Model) is a Lagrangian, mesoscale, atmospheric puff dispersion model that calculates the activity and doses of airborne radionuclides considering parameters of puff diffusion, wet and dry deposition (Thykier-Nielsen et al. 1999). Time-variable releases are represented by a series of Gaussian puffs where each puff represents the amount of release within a certain time interval. The model can be used in terrains with moderate orography for a range of up to several hundred kilometres from the release site.

Terrestrial Food Chain and Dose Module (FDMT)

The FDMT (Food Chain and Dose Module for Terrestrial Pathways) is the module in the JRODOS system to simulate the transfer of radioactive material in food chains, and to assess the doses to the

population via all relevant exposure pathways - internal exposure via inhalation and ingestion, external exposure from cloud and ground shine (Gering and Müller 2004; Müller et al. 2004). The main input parameters for FDMT are derived from the atmospheric dispersion calculations. They comprise the near ground activity concentration in air, the deposited activity concentration on the ground, the amount of precipitation and the date, i.e. season, of deposition. The different transfer steps of radionuclides through the food chain are shown in Figure 79.

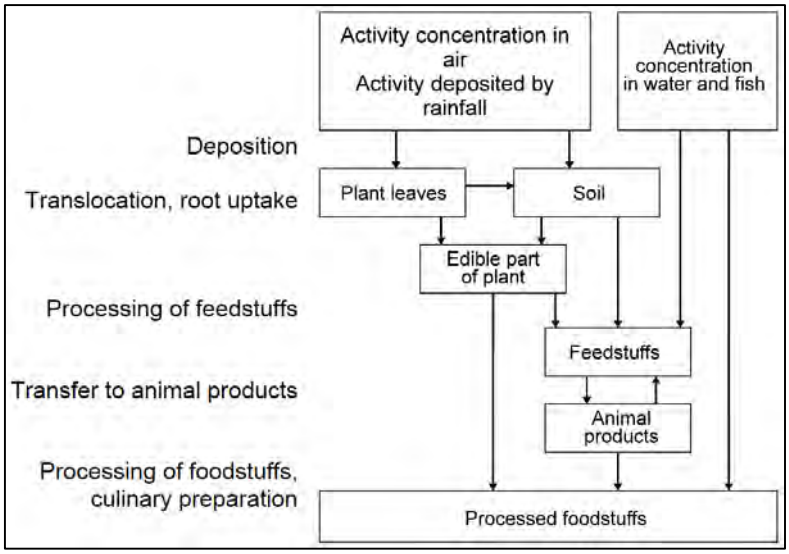


Figure 79 Steps of food chain transfer calculations (from Müller et al. (2004)).

Endpoints

The full model chain of dispersion calculations and FDMT results in the following available endpoints:

- Air concentration and dry and wet deposition
  - I-131, I-132, T-132, Xe-133, Cs-134, Cs-136, Cs-137, Ba-137m
- Gamma dose rate
- Projected dose (1y and lifetime)
  - 5y old, 10y old, adults
  - Effective, RBM, thyroid
  - Ingestion, inhalation, all pathways, all pathways w/o ingestion
- Concentration in feedstuff (Cs, I)
  - Grass, hay corn, beet
- Concentration in foodstuff (Cs, I)
  - Leafy veg., potatoes, winter wheat, milk, beef

Ensemble setup

The propagation of uncertainties through the model chain in form of ensemble members was handled in form of three different ensemble model setups to evaluate the differences in the uncertainties introduced by different parameters. The model setups are:

- Meteorological and source term uncertainties (“MET+ST”)
- Uncertainties introduced to the food chain (“FDMT”)
- Uncertainties in meteorology, source term and food chain (“MET+ST+FDMT”)

Each of the model setups results in an ensemble with 50 ensemble members. The 5 source term ensemble member are taken from a subset of the long release scenario from FASTNET, described in (Mathieu et al. 2018). An example of Cs-134 and I-131 release rates is given in Figure 80.

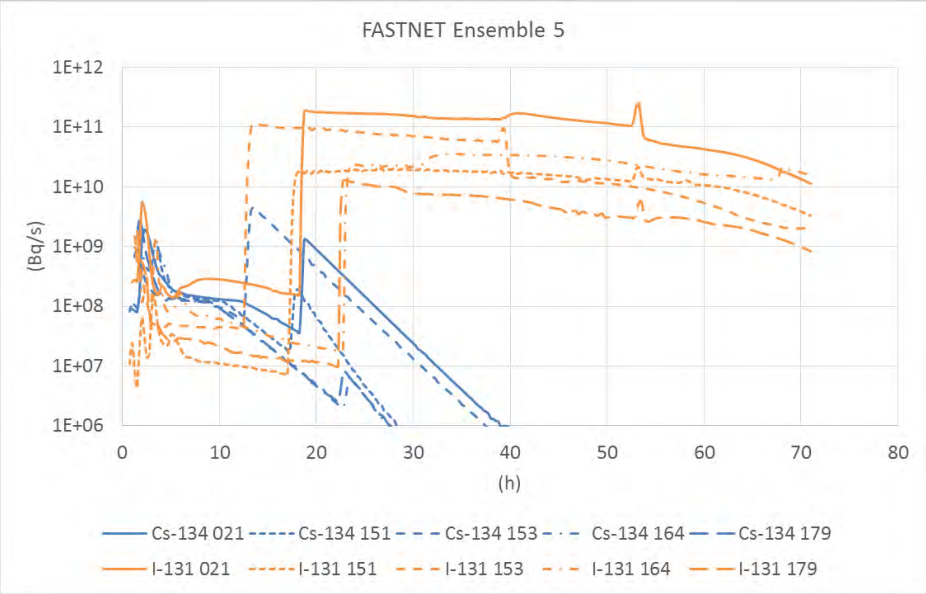


Figure 80 Release rates of Cs-134 and I-131 for the 5 member source term ensemble. The source term ensemble members are labeled as 021, 151, 153, 164, and 179.

Meteorological and source term uncertainties (MET+ST)

The first model setup uses the uncertainties introduced in Task 1.2 of WP1. 10 meteorological ensemble member and five source term ensemble member result in a total amount of 50 ensemble members. Each one of the ensemble members is passed through the model chain, including the atmospheric dispersion calculation and the calculation of the endpoints of FDMT. All model parameters in FDMT are set to their default values and do not change within this model setup. A schematic representation of the model chain is given in Figure 81.

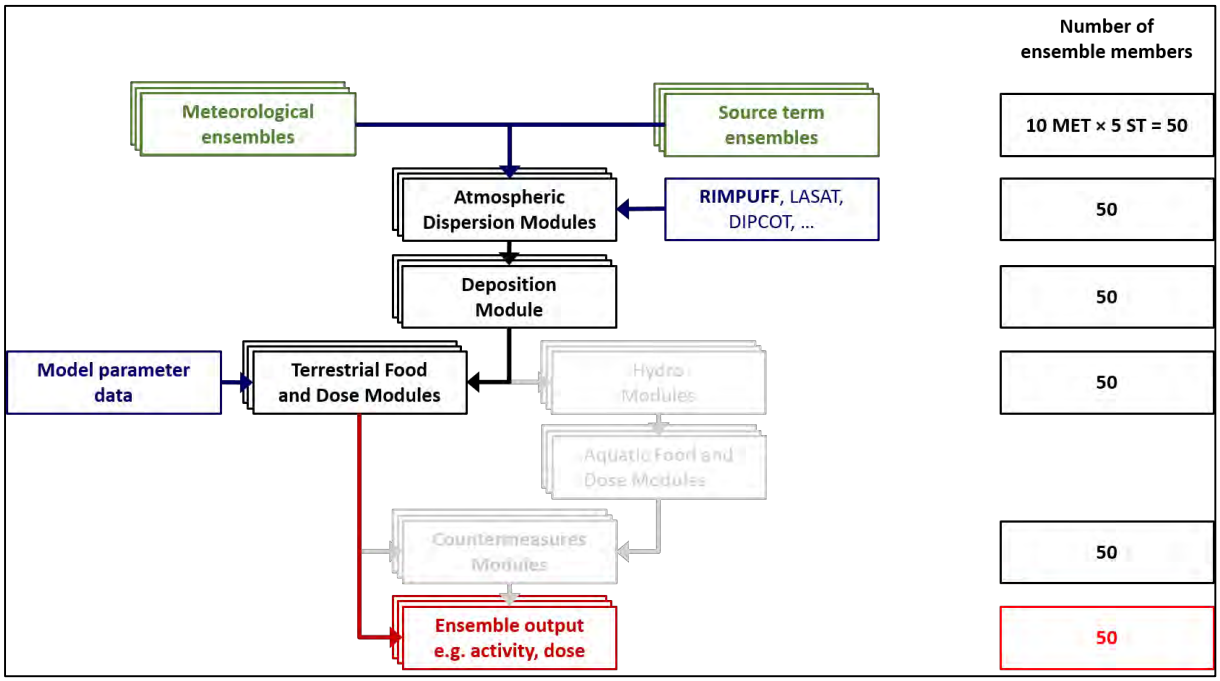


Figure 81 Model setup for propagation of meteorological and source term uncertainties.

Uncertainties introduced to the food chain (FDMT)

The model setup "FDMT" uses only one member of the meteorological ("mbr00") and source term ("021") ensemble as a basis for the atmospheric dispersion calculation. Therefore, only model parameter for the FDMT model are randomly applied to the model chain. 50 representations of the varied model parameters are used, giving 50 ensemble member as output (Figure 82). The variable parameter as well as their ranges and distribution functions are given in the tables in Appendix A. Several parameter values and ranges were updated.

The variability of the following parameter were taken from CONFIDENCE deliverable 9.13 (Brown et al. 2019):

- Retention coefficient of radionuclides on different plant types (Table 19)
- Weathering rate from plants (Table 20)
- Soil-plant transfer factor for radionuclides (Table 21)
- Transfer coefficients to different animal products (Table 26)

Data on harvest times (Table 22, Table 23) were derived from phenological data collected by Germany's National Meteorological Service (DWD) (DWD 2019). Historical data of the yield of crops (Table 24) was derived from the online database of the German Federal Statistical Office (Destatis 2019). Data for food consumption in Germany were used to model the intake of foodstuff (MRI (2008), Table 32).

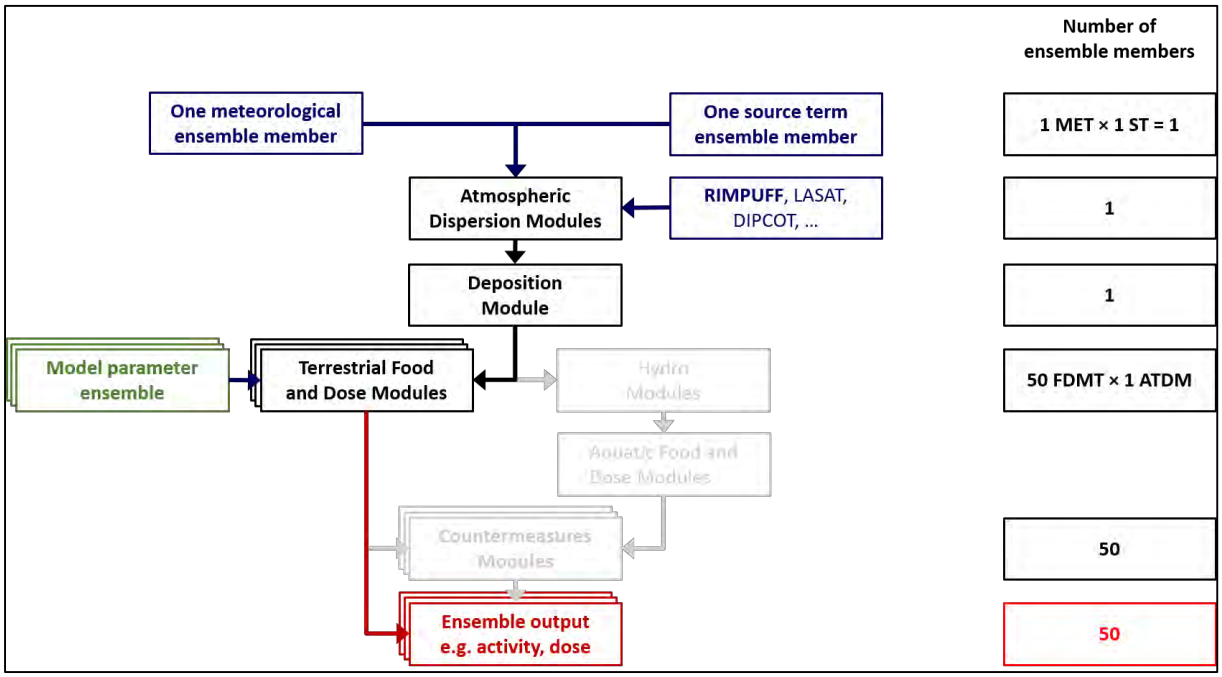


Figure 82 Model setup for propagation of FDMT uncertainties.

Uncertainties in meteorology, source term and food chain (MET+ST+FDMT)

This model setup combines the uncertainties of the meteorological and source term ensemble and the FDMT ensemble. To avoid a too large number of resulting ensemble members and to optimize model run-time and post-processing each one of the 50 representations of the FDMT model parameter from the "FDMT" model setup was randomly assigned to one of the 50 representations of the dispersion calculations of the "MET+ST" model setup. Hence the total number of ensemble results remained at 50 ensemble members (Figure 83).

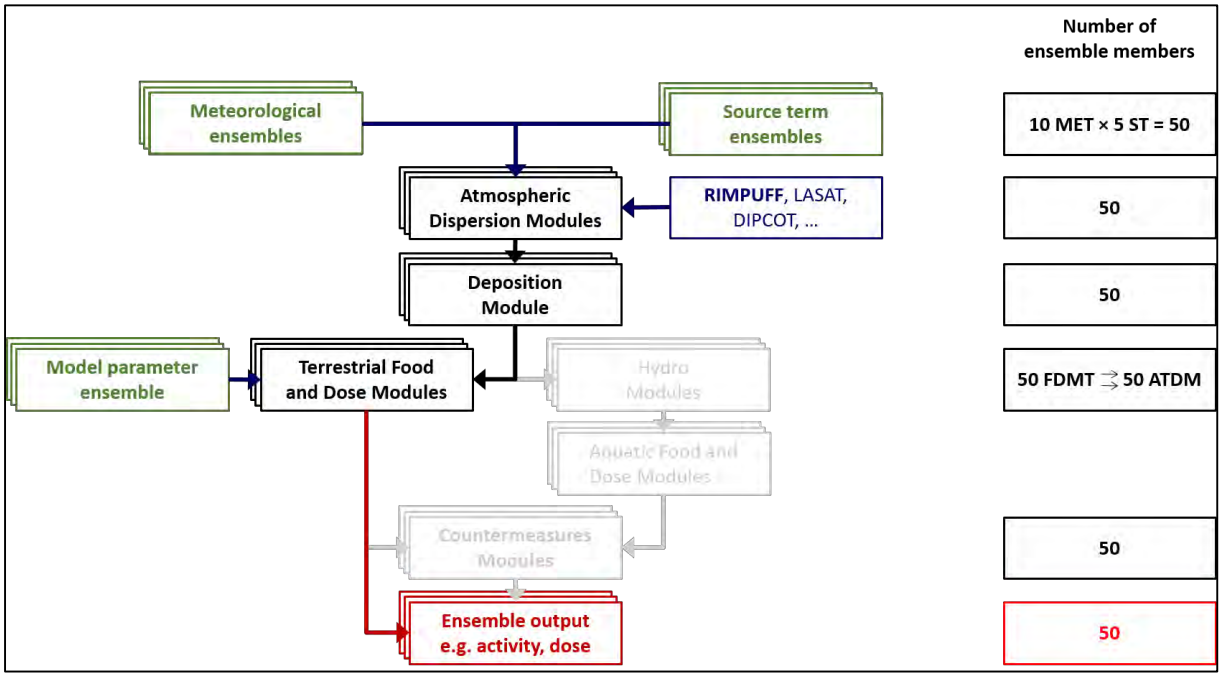


Figure 83 Model setup for propagation of meteorological, source term and FDMT uncertainties.

Results

The results of the uncertainty propagation of atmospheric dispersion calculation ensembles through an ensemble of food chain calculations is presented below. Maps for areas where a threshold value for Cs in leafy vegetables of 1250 Bq/kg is exceeded are shown in Figure 84 and Figure 85 as one example for activity concentrations in foodstuff. All ensemble model setups described above ("MET+ST", "FDMT", "MET+ST+FDMT") are considered. Figure 84 shows the area where all ensemble member exceed the threshold value (red) and the area where at least one ensemble member exceeds the threshold value (blue). The ratio of the areas can be taken as a measure of the variability within the ensemble of the respective model setup. The ratios in this example are as follows:

- "MET+ST": 16 km²/596 km² ≈ 0.03
- "FDMT": 40 km²/208 km² ≈ 0.19
- "MET+ST+FDMT": 8 km²/692 km² ≈ 0.01

Although the variability within the "FDMT" ensemble is significant by itself and adds variability to the overall "MET+ST+FDMT" ensemble, the variability of the meteorological and source term ensemble dominates the overall variability.

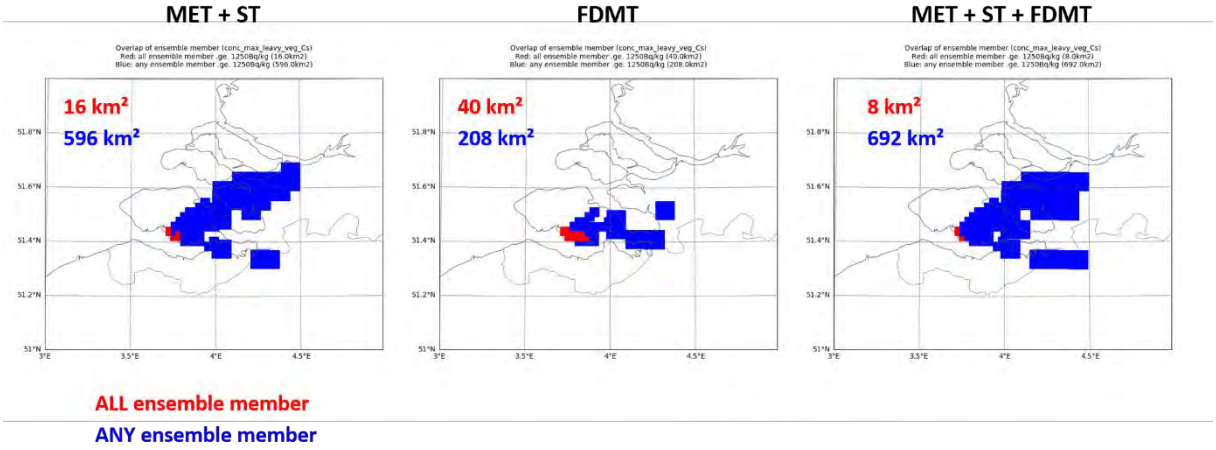


Figure 84 Area covered by ensemble members for a threshold value for leafy vegetables Cs > 1250 Bq/kg. The red area shows areas where the threshold value was exceeded for all ensemble members, the blue area shows areas where the threshold value was exceeded for at least one ensemble member.

Figure 85 shows the results of the same parameter and ensemble runs as Figure 84. The colour scale represents the frequency, i.e. fraction, in each grid cell of how many ensemble member show an exceedance of the threshold value for Cs in leafy vegetables. Again, lowest variability can be observed for the ensemble model setup "FDMT". The dominating variability comes from the meteorological and source term ensemble.

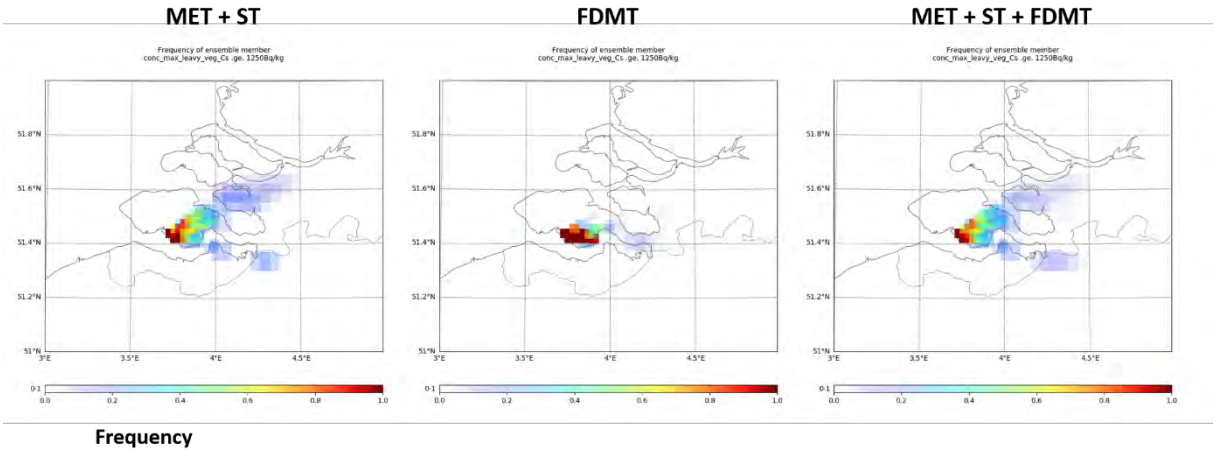


Figure 85 Frequency maps for a threshold value for leafy vegetables Cs > 1250 Bq/kg.

The hypothetical accident scenario takes place in January. The effect of the released radionuclides on the activity concentration in foodstuff in winter is less compared to a release in spring or summer. In winter the deposition on plants happens before the growth phase while in spring or summer the radionuclides are deposited on the plants during the growth season or directly on the foliage. Hence, ensemble calculations were shifted by 6 months to July to consider deposition on existing crops. The input of the dispersion calculations remained the same while the season was shifted for FDMT only.

A comparison of the results between January and July for threshold values for Iodine and Caesium in leafy vegetables and in milk is listed in Table 18. The value to indicate the variability within one ensemble is the interquartile range (IQR). It is normalized to the 75th percentile to be able to compare the results amongst the different parameters.

Table 18 Relative interquartile range for areas where threshold values in foodstuff are exceeded. The interquartile range is normalized to the 75th percentile.

IQR (%) (P75-P25)/P75 * 100	January			July		
	MET + ST	FDMT	MET + ST + FDMT	MET + ST	FDMT	MET + ST + FDMT
Leafy veg. Cs > 1250 Bq/kg	52	31	59	53	32	56
Leafy veg. I > 2000 Bq/kg	80	42	92	81	42	86
Milk Cs > 1000 Bq/kg	-	-	-	75	100	85
Milk I > 500 Bq/kg	77	64	82	83	38	94

For all parameter, except Caesium in milk, the results of the previous comparisons are supported; lowest variability for the ensemble setup “FDMT” and dominating variability for ensembles with meteorological and source term uncertainties. The normalized IQR values for Caesium in milk are an exception as the threshold value of 1000 Bq/kg is not exceeded in any grid cell (January) or only few grid cells (July).

Conclusions

The analysis of the propagation of uncertainties through a food chain model showed that the dominating variability amongst an ensemble, and therefore the uncertainty, is introduced by the meteorological input data and the source term data. The uncertainties of food chain parameter introduce an additional, though not dominating variability.

References

Brown, J. E., R. Avila, C. L. Barnett, N. A. Beresford, A. Hosseini, O.-C. Lind, D. H. Oughton, D. Perez, B. Salbu, H. C. Teien and H. Thørring (2019). Improving models and learning from post-Fukushima studies. EJP-CONCERT European Joint Programme for the Integration of Radiation Protection Research - CONFIDENCE Project. Report.http://www.concert-h2020.eu/en/Publications

Chevalier-Jabet, K. (2019). Source term prediction in case of a severe nuclear accident. 5th NERIS workshop. Roskilde, Denmark.

Destatis. (2019). "GENESIS-Online Database." 2019, from https://www-genesis.destatis.de/genesis/online.

DWD. (2019). "CDC-Portal (Climate Data Center)." 2019, from https://www.dwd.de/DE/leistungen/cdc_portal/cdc_portal.html.

Ehrhardt, J. (1997). "The RODOS System: Decision Support for Off-Site Emergency Management in Europe." Radiation Protection Dosimetry **73**(1-4): 35-40.

Geertsema, G., H. De Vries and R. Sheele (2019). High resolution meteorological ensemble data for CONFIDENCE research on uncertainties in atmospheric dispersion in the (pre-)release phase of a nuclear accident. 19th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Gering, F. and H. Müller (2004). User Guide for the Terrestrial Food Chain and Dose Module FDMT of RODOS PV6.0, GSF - Institut für Strahlenschutz. Report

Korsakissok, I., R. Périllat, G. Geertsema, H. De Vries, T. Hamburger, J. Tomas, S. Andronopoulos, P. Bedwell, T. Charnock, I. Ievdin, S. Leadbetter, T. Pazmandi, C. Rudas, R. Scheele, A. Sogachev, P. Szanto and J. Wellings (2019). D 9.5.1 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the REM case studies. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Leadbetter, S., S. Andronopoulos, P. Bedwell, G. Geertsema, A. Jones, I. Korsakissok, J. Tomas and H. De Vries (2018). D1.1.1 Using ensemble meteorological forecasts to represent meteorological uncertainty in dispersion models. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Mathieu, A., I. Korsakissok, R. Périllat, K. Chevalier-Jabet, F. Stephani, S. Fougerolle, V. Créach, E. Cogez and P. Bedwell (2018). D 1.1.3 Guidelines describing source term uncertainties. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

MRI (2008). Nationale Verzehrs Studie II Ergebnisbericht, Teil 2, Max Rubner-Institut, Bundesforschungsinstitut für Ernährung und Lebensmittel. Report.https://www.bmel.de/DE/Ernaehrung/GesundeErnaehrung/_Texte/NationaleVerzehrsstudie_Zusammenfassung.html

Müller, H., F. Gering and G. Pröhl (2004). Model Description of the Terrestrial Food Chain and Dose Module FDMT in RODOS PV6.0. RODOS Report - Decision Support for Nuclear Emergencies, GSF - Institut für Strahlenschutz. Report RODOS(RA3)-TN(03)06

Raskob, W., D. Trybushnyi, I. Ievdin and M. Zheleznyak (2012). "JRODOS: Platform for improved long term countermeasures modelling and management." Radioprotection **46**(6): S731-S736.

Thykier-Nielsen, S., S. Deme and T. Mikkelsen (1999). Description of the Atmospheric Dispersion Module RIMPUFF. RODOS Report - Decision Support for Nuclear Emergencies, Risø National Laboratory. Report

Appendix A

Table 19 Retention coefficient of plants (mm)
Normal distribution

PLANT TYPE	CS		I		SR	
	Median	SD	Median	SD	Median	SD
GRASS (INT.)	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
GRASS (EXT.)	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
MAIZE	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
MAIZE BULBS	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
POTATOES	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
BEETS	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
BEET LEAVES	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
WINTER BARLEY	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
SPRING BARLEY	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
WINTER WHEAT	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
SPRING WHEAT	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
RYE	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
OATS	2.00E-01	2.00E-01	1.00E-01	1.00E-01	4.00E-01	4.00E-01
LEAFY VEG.	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
ROOT VEG.	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
FRUIT VEG.	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
FRUITS	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01
BERRIES	3.00E-01	3.00E-01	1.50E-01	1.50E-01	6.00E-01	6.00E-01

Table 20 Weathering rate describing exponential decrease of activity on plants due to wind, rain etc. (1/d)
Log-normal distribution

ELEMENT	MEDIAN	SD
CS	5.00E-02	1.90E-02
I	7.50E-02	2.90E-02
SR	5.00E-02	1.90E-02

Table 21 Transfer factor soil/plant (kg soil/kg fresh plant), all soil types
Log-normal distribution

PLANT TYPE	CS		I		SR	
	Median	SD	Median	SD	Median	SD
GRASS (INT.)	1.21E-01	1.80E-01	9.90E-02	3.10E-02	3.74E-01	2.64E-01
GRASS (EXT.)	2.42E-02	2.64E-02	9.90E-02	3.10E-02	3.74E-01	2.64E-01
MAIZE	3.00E-02	2.75E-02	2.75E-02	4.50E-02	2.48E-01	1.90E-01
MAIZE BULBS	1.05E-02	1.08E-02	2.66E-04	5.32E-04	1.12E-01	1.16E-02
POTATOES	2.10E-02	2.52E-02	2.10E-02	2.52E-02	5.04E-02	4.62E-02
BEETS	1.20E-02	1.76E-02	2.08E-03	1.92E-01	2.40E-01	2.24E-01
BEET LEAVES	1.09E-02	1.92E-02	2.08E-03	1.92E-03	2.40E-01	2.24E-01
WINTER BARLEY	6.61E-02	1.31E-01	1.22E-04	2.44E-03	1.57E-01	1.65E-01
SPRING BARLEY	6.61E-02	1.31E-01	1.22E-04	2.44E-03	1.57E-01	1.65E-01
WINTER WHEAT	6.69E-02	1.32E-01	1.23E-04	2.46E-03	1.58E-01	1.67E-01
SPRING WHEAT	6.69E-02	1.32E-01	1.23E-04	2.46E-03	1.58E-01	1.67E-01
RYE	6.61E-02	1.31E-01	1.22E-04	2.44E-03	1.57E-01	1.65E-01
OATS	6.61E-02	1.31E-01	1.33E-04	2.44E-03	1.57E-01	1.65E-01
LEAFY VEG.	1.70E-02	2.10E-02	1.60E-03	2.90E-03	1.90E-01	1.80E-01
ROOT VEG.	1.20E-02	1.76E-02	2.08E-03	1.92E-03	2.40E-01	2.24E-01
FRUIT VEG.	3.50E-03	7.50E-03	3.50E-03	7.50E-03	4.90E-02	9.00E-02
FRUITS	2.25E-03	3.30E-03	1.80E-03	1.80E-03	3.75E-03	2.85E-03
BERRIES	2.90E-03	3.30E-03	2.90E-03	3.30E-03	5.50E-02	6.90E-02

Table 22 Begin and end of harvest of plants (Julian days)
Triangular distribution

PLANT TYPE	BEGIN			END		
	Min	MPV	Max	Min	MPV	Max
GRASS (INT.)	83	115	126	153	168	283
HAY (INT.)	98	122	136	167	197	306
GRASS (EXT.)	83	115	126	153	168	283
HAY (EXT.)	98	122	136	167	197	306
MAIZE	218	234	245	266	297	322
MAIZE BULBS	217	251	273	296	322	346
POTATOES	146	154	166	271	297	303
BEETS	190	211	230	331	337	355
BEET LEAVES	190	211	230	331	337	355
WINTER BARLEY	146	170	185	208	233	276
SPRING BARLEY	121	185	194	232	266	280
WINTER WHEAT	156	191	203	228	251	301
SPRING WHEAT	177	193	200	258	270	288
RYE	160	185	181	195	264	292
OATS	167	192	206	232	265	293
LEAFY VEG.	106	121	136	289	304	319
ROOT VEG.	167	182	197	304	319	334
FRUIT VEG.	177	185	189	264	276	289
FRUITS	124	144	158	241	293	333
BERRIES	127	149	161	195	214	295

Table 23 End of first harvesting period (Julian days)
Triangular distribution

PLANT TYPE	MIN	MPV	MAX
GRASS (INT.)	127	140	152
HAY (INT.)	137	152	166
GRASS (EXT.)	127	140	152
HAY (EXT.)	137	152	166
MAIZE	246	256	265
MAIZE BULBS	274	285	295
POTATOES	167	219	270
BEETS	231	281	330
BEET LEAVES	231	281	330
WINTER BARLEY	186	197	207
SPRING BARLEY	195	213	231
WINTER WHEAT	204	216	227
SPRING WHEAT	201	229	257
RYE	182	188	194
OATS	207	219	231
LEAFY VEG.	137	213	288
ROOT VEG.	198	251	303
FRUIT VEG.	190	227	263
FRUITS	159	200	240
BERRIES	162	178	194

Table 24 Yield of plants (kg/m^2)
Triangular distribution

PLANT TYPE	MIN	MPV	MAX
GRASS (INT.)	7.30E-01	8.60E-01	9.30E-01
HAY (INT.)	7.30E-01	8.60E-01	9.30E-01
GRASS (EXT.)	7.30E-01	8.60E-01	9.30E-01
HAY (EXT.)	7.30E-01	8.60E-01	9.30E-01
MAIZE	3.94E+00	4.38E+00	4.56E+00
MAIZE BULBS	7.10E-01	8.50E-01	9.50E-01
POTATOES	3.02E+00	3.87E+00	4.38E+00
BEETS	5.01E+00	7.56E+00	1.02E+01
BEET LEAVES	2.00E+00	3.00E+00	4.00E+00
WINTER BARLEY	5.60E-01	6.20E-01	7.00E-01
SPRING BARLEY	4.30E-01	4.80E-01	5.20E-01
WINTER WHEAT	6.50E-01	7.20E-01	7.90E-01
SPRING WHEAT	5.10E-01	5.40E-01	5.90E-01
RYE	4.00E-01	5.10E-01	5.80E-01
OATS	4.20E-01	4.60E-01	5.10E-01
LEAFY VEG.	2.80E-01	2.51E+00	4.37E+00
ROOT VEG.	8.90E-01	2.63E+00	3.00E+00
FRUIT VEG.	4.30E-01	1.70E+00	3.29E+00
FRUITS	5.50E-01	1.03E+00	2.86E+00
BERRIES	4.10E-01	4.50E-01	4.60E-01

Table 25 Time grid for feeding rates (Julian days)
Triangular distribution

ANIMAL PRODUCT	FEED. TIME #1			FEED. TIME #2			FEED. TIME #3			FEED. TIME #4		
	Min	MPV	Max	Min	MPV	Max	Min	MPV	Max	Min	MPV	Max
COW MILK	101	111	121	121	131	141	284	294	304	304	314	324
SHEEP MILK	101	111	121	121	131	141	284	294	304	304	314	324
GOATS MILK	101	111	121	121	131	141	284	294	304	304	314	324
COW BEEF	101	111	121	121	131	141	284	294	304	304	314	324
LAMB	101	111	121	121	131	141	284	294	304	304	314	324

Table 26 Transfer factor feed / animal product (d/kg)
Log-normal distribution

ANIMAL PRODUCT	CS		I		SR	
	Median	SD	Median	SD	Median	SD
COW MILK	6.10E-03	6.30E-03	9.10E-03	7.00E-03	1.50E-03	8.10E-04
COW BEEF	3.00E-02	2.30E-02	1.20E-02	1.50E-02	2.10E-03	2.20E-03
LAMB	8.67E-01	3.85E-01	7.78E-02	6.49E-02	2.58E-03	1.08E-03

Table 27 Feeding rates for animals (kg/d), grass (int.)
Uniform distribution

ANIMAL PRODUCT	FEED. TIME #1			FEED. TIME #2			FEED. TIME #3			FEED. TIME #4			FEED. TIME #5		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
COW MILK				4.90E+01	7.00E+01	9.10E+01	4.90E+01	7.00E+01	9.10E+01						
SHEEP MILK				6.00E+00	9.00E+00	1.20E+01	6.00E+00	9.00E+00	1.20E+01						
GOATS MILK				9.00E+00	1.30E+01	1.70E+01	9.00E+00	1.30E+01	1.70E+01						
COW BEEF				6.00E+01	7.00E+01	8.00E+01	6.00E+01	7.00E+01	8.00E+01						
BULL BEEF	2.50E+01	2.80E+01	3.10E+01												
VEAL	2.30E+00	2.90E+00	3.50E+00												
PORK	2.40E+00	3.00E+00	3.60E+00												
LAMB				4.00E+00	5.00E+00	6.00E+00	4.00E+00	5.00E+00	6.00E+00						
CHICKEN	8.00E-02	9.00E-02	1.00E-01												
EGGS	8.00E-02	9.00E-02	1.00E-01												

Table 28 Feeding rates for animals (kg/d), hay (int.)
Uniform distribution

ANIMAL PRODUCT	FEED. TIME #1			FEED. TIME #2			FEED. TIME #3			FEED. TIME #4			FEED. TIME #5		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
COW MILK	1.00E+01	1.40E+01	1.80E+01							1.00E+01	1.40E+01	1.80E+01	1.00E+01	1.40E+01	1.80E+01
SHEEP MILK	1.20E+00	1.80E+00	2.40E+00							1.20E+00	1.80E+00	2.40E+00	1.20E+00	1.80E+00	2.40E+00
GOATS MILK	1.80E+00	2.60E+00	3.40E+00							1.80E+00	2.60E+00	3.40E+00	1.80E+00	2.60E+00	3.40E+00
COW BEEF	1.20E+01	1.40E+01	1.60E+01							1.20E+01	1.40E+01	1.60E+01	1.20E+01	1.40E+01	1.60E+01
BULL BEEF	3.60E+01	4.00E+01	4.40E+01												
VEAL	8.00E+00	1.00E+01	1.20E+01												
PORK	6.40E+00	8.00E+00	9.60E+00												
LAMB	8.00E-01	1.00E+00	1.20E+00							8.00E-01	1.00E+00	1.20E+00	8.00E-01	1.00E+00	1.20E+00
CHICKEN	1.80E-01	2.00E-01	2.20E-01												
EGGS	1.80E-01	2.00E-01	2.20E-01												

Table 29 Feeding rates for animals (kg/d), grass (ext.)
Uniform distribution

ANIMAL PRODUCT	FEED. TIME #1			FEED. TIME #2			FEED. TIME #3			FEED. TIME #4			FEED. TIME #5		
	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max	Min	Median	Max
COW MILK	5.00E+01	7.50E+01	1.00E+02	5.00E+01	7.50E+01	1.00E+02	5.00E+01	7.50E+01	1.00E+02	5.00E+01	7.50E+01	1.00E+02	5.00E+01	7.50E+01	1.00E+02
SHEEP MILK	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00
GOATS MILK	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00	4.00E+00	6.00E+00	8.00E+00
COW BEEF	6.40E+01	7.50E+01	8.60E+01	6.40E+01	7.50E+01	8.60E+01	6.40E+01	7.50E+01	8.60E+01	6.40E+01	7.50E+01	8.60E+01	6.40E+01	7.50E+01	8.60E+01
BULL BEEF															
VEAL															
PORK															
LAMB	3.20E+00	4.00E+00	4.80E+00	3.20E+00	4.00E+00	4.80E+00	3.20E+00	4.00E+00	4.80E+00	3.20E+00	4.00E+00	4.80E+00	3.20E+00	4.00E+00	4.80E+00
CHICKEN															
EGGS															

Table 30 Biological halflifes for animal products (d), compartment #1
Uniform distribution

ANIMAL PRODUCT	CS			I			SR		
	Min	Median	Max	Min	Median	Max	Min	Median	Max
COW MILK	1.00E+00	1.50E+00	2.00E+00	4.00E-01	7.00E-01	1.00E+00	2.00E+00	3.00E+00	4.00E+00
SHEEP MILK	1.00E+00	1.50E+00	2.00E+00	4.00E-01	7.00E-01	1.00E+00	2.00E+00	3.00E+00	4.00E+00
GOATS MILK	1.00E+00	1.50E+00	2.00E+00	4.00E-01	7.00E-01	1.00E+00	2.00E+00	3.00E+00	4.00E+00
COW BEEF	2.50E+01	3.00E+01	3.50E+01	5.00E+01	1.00E+02	1.50E+02	5.00E+00	1.00E+01	1.50E+01
BULL BEEF	4.00E+01	5.00E+01	6.00E+01	5.00E+01	1.00E+02	1.50E+02	5.00E+00	1.00E+01	1.50E+01
VEAL	2.50E+01	3.00E+01	3.50E+01	5.00E+01	1.00E+02	1.50E+02	5.00E+00	1.00E+01	1.50E+01
PORK	3.00E+01	3.50E+01	4.00E+01	5.00E+01	1.00E+02	1.50E+02	5.00E+00	1.00E+01	1.50E+01
LAMB	1.50E+01	2.00E+01	2.50E+01	5.00E+01	1.00E+02	1.50E+02	5.00E+00	1.00E+01	1.50E+01
CHICKEN	1.50E+01	2.00E+01	2.50E+01	5.00E+01	1.00E+02	1.50E+02	2.00E+00	3.00E+00	4.00E+00
EGGS	2.00E+00	3.00E+00	4.00E+00	4.00E-01	7.00E-01	1.00E+00	1.50E+00	2.00E+00	2.50E+00

Table 31 Biological halflifes for animal products (d), compartment #2
Uniform distribution

ANIMAL PRODUCT	CS			I			SR		
	Min	Median	Max	Min	Median	Max	Min	Median	Max
COW MILK	1.00E+01	1.50E+01	2.00E+01				5.00E+01	1.00E+02	1.50E+02
SHEEP MILK	1.00E+01	1.50E+01	2.00E+01				5.00E+01	1.00E+02	1.50E+02
GOATS MILK	1.00E+01	1.50E+01	2.00E+01				5.00E+01	1.00E+02	1.50E+02
COW BEEF							5.00E+01	1.00E+02	1.50E+02
BULL BEEF							5.00E+01	1.00E+02	1.50E+02
VEAL							5.00E+01	1.00E+02	1.50E+02
PORK							5.00E+01	1.00E+02	1.50E+02
LAMB							5.00E+01	1.00E+02	1.50E+02
CHICKEN							5.00E+01	1.00E+02	1.50E+02
EGGS							1.50E+01	2.00E+01	2.50E+01

Table 32 Food consumption rates of average humans (g/d)
Triangular distribution

FOODSTUFF	AGE 1			AGE 5			AGE 10			AGE 15			AGE ADULTS		
	Min	MPV	Max	Min	MPV	Max	Min	MPV	Max	Min	MPV	Max	Min	MPV	Max
WINTER WHEAT FLOUR	34	45	56	128	170	213	131	174	217	124	164	205	124	165	206
RYE FLOUR	11	15	18	42	56	70	43	57	72	41	54	68	41	55	68
OATS	3	4	6	13	17	22	13	18	22	13	17	21	13	17	21
POTATOES	17	23	30	51	67	87	60	78	101	72	95	123	62	81	105
LEAFY VEGS.	56	73	88	59	76	91	59	77	93	62	81	98	91	119	143
ROOT VEGS.	20	27	32	21	27	32	21	27	33	22	29	34	32	42	50
FRUIT VEGS.	12	15	18	29	38	46	30	39	46	31	41	49	46	59	71
FRUITS	80	200	313	60	150	235	61	154	240	60	151	235	64	160	250
BERRIES	0	0	0	20	28	36	17	24	30	9	13	16	10	14	18
FRESH MILK	187	314	562	122	204	365	124	208	372	99	165	296	77	129	231
BEEF COW	1	1	2	7	13	21	8	15	25	9	17	29	10	18	30
BEEF BULL	1	2	3	14	26	44	17	31	52	19	35	58	20	36	61
VEAL	0	0	0	1	1	2	1	1	2	1	1	2	1	1	2
PORK	1	3	4	28	51	85	33	60	101	38	68	115	39	71	120
EGGS	0	0	0	9	12	16	10	14	19	9	12	17	14	19	26
BEER	0	0	0	0	0	0	0	1	2	6	25	55	34	142	313

D9.5.5 – Guidelines for the use of ensembles in the description of uncertainty in atmospheric dispersion modelling: operational applications in the context of an emergency response

Lead Author: Peter Bedwell

With contributions from: Irène Korsakissok, Susan Leadbetter, Raphaël Périllat, Csilla Rudas, Jasper Tomas, Joseph Wellings

Reviewer(s): Damien Didier, Gertie Geertsema

Table of Contents

Table of Contents 127
Introduction 128
Objective 128
Scope 128
Operational methods used currently to describe uncertainty 131
Clustering: reducing the number of members in an ensemble approach 133
Automated clustering of ensembles using computational algorithms 134
Manual clustering of ensemble members a priori – on the basis of wind direction and the timing of the release 139
Sampling: reducing the number of members in an ensemble approach 152
Reducing model run time for a single ensemble member 153
Alternative temporally and spatially resolved meteorological ensembles 154
Cheaper dispersion modelling schemes & modified model domains 154
Reducing the number of radionuclides modelled 164
Computational resource 165
Operational ensemble results & output 166
Uncertainty indicators 166
Presentation of ensemble output 167
Recommendations 167
Reducing the run time for a single ensemble member 167
Reducing the number of ensemble members used in the simulations 168
Additional technical recommendations to save computational time 169
Uncertainty representation for operational purposes 169
Future work 169
Reducing model run time for a single ensemble member 169
Reducing the number of members in an ensemble approach by an appropriate selection method 170
Use of dispersion field studies to assess the quality of ensemble simulations 170
Improvement of input uncertainties, combining a priori knowledge and observational data 171
References 172

Introduction

Mathieu et al. (2018a) presented guidelines ranking uncertainties for atmospheric dispersion modelling. These guidelines introduced the concept of the application of ensembles in the description of atmospheric dispersion model uncertainty and provided estimates of ranges and distributions of uncertainties on atmospheric dispersion model input parameters. By way of hypothetical (Berge et al. 2019; Korsakissok et al. 2019b) and historical (Korsakissok et al. 2019a) scenarios, these uncertainties were propagated through the chain of atmospheric dispersion and radiological assessment models based on an ensemble approach. The work performed thus far in the study has provided insight into the prioritisation of input parameter uncertainties, the relative magnitudes of model endpoint uncertainties, methods for analysing uncertainty and techniques for presenting uncertainty. However, it was apparent from the exercise undertaken to propagate uncertainties through the chain of atmospheric dispersion and radiological assessment models that the time taken (ranging from several hours to a few tens of hours) to complete model runs was not appropriate for an emergency response.

As a result, to ensure an operational method, there is a requirement to reduce the number of ensemble members and/or reduce model run time for a single ensemble member, without significant detriment to the model endpoints derived, the uncertainty estimated and the radiation protection advice inferred.

Objective

The primary objective of this work is to recommend approaches that enable a measure of uncertainty to be obtained within the timeframe of one hour. In this case, uncertainty is related to the process of atmospheric dispersion modelling. The new approaches could lead to improvements in the provision of radiological protection advice in an emergency response.

Scope

This report details a range of (tiered) recommendations from more simplistic to more complex approaches, commensurate with the range of capabilities, computational resources and staffing resources of different organisations. The operational goals of an organisation may be inhibited by their capacity to implement different levels of complexity of recommendations. Recommendations range from what can be achieved relatively easily to what is merely aspirational. Operational goals may range from having the capability to demonstrate that uncertainty exists and that there is the potential for alternative scenarios to occur, to quantifying uncertainty and estimating the likelihoods of different scenarios occurring.

It is recognised here that there exist numerous different approaches to estimating and describing uncertainty. This study is focused on ensemble approaches, and methods which can be used to simplify and/or improve the application of an ensemble approach.

Furthermore, this report considers only uncertainties that originate within the atmospheric dispersion modelling process (including source term uncertainties). It does not consider uncertainties that originate within other parts of the radiological assessment chain of models, for example foodchain models.

The use of ensemble meteorological forecasts to represent meteorological uncertainty in atmospheric dispersion models was assessed in Leadbetter et al. (2018). Whilst it was found that the uncertainty in some variables is well captured, e.g. wind and precipitation, the uncertainty in other variables is less

well captured, e.g. boundary layer height. Furthermore, meteorological ensembles are designed to model uncertainty at scales (spatial and temporal) which may not be optimum for atmospheric dispersion modelling. These limitations are recognised, and work to develop meteorological ensembles for a broader range of applications (such as dispersion modelling) is strongly encouraged, but the development of recommendations describing preferential improvements to the meteorological ensembles themselves is not the purpose of this report.

It might be expected that determining uncertainty would be simpler over relatively small spatial scales because of a greater propensity for relatively uniform meteorological conditions in the horizontal plane, relatively more dependency on near-ground meteorological variables and uncertainties not having had an opportunity to accumulate along a trajectory. Conversely, it might be expected that determining uncertainty would be more complex at a local scale because of relatively strong environmental concentration gradients and a relatively high influence of potentially unknown variables such as release height, building effects, or particle size distribution of the release. Therefore, the ease of determination of uncertainty associated with relatively local scale model endpoints (such as evacuation, sheltering and iodine prophylaxis countermeasures) and relatively large scale model endpoints (notably food restriction protective actions) is likely to be very much scenario dependent.

This study will focus on uncertainty which exists in the emergency phase (pre-and early release phase) of an accident, before measurements become available, and where it is expected that uncertainties will be highest, and decision making most critical. From a decision maker's perspective, the protective actions evacuation, sheltering and iodine prophylaxis are likely to take priority and consideration of food restrictions is likely to be of secondary importance (however it is recognised that this may vary from one country to another; for example, in France there is growing pressure to provide radiological protection advice in respect of food restrictions in the emergency phase). Thus, more emphasis will be placed on recommendations for an ensemble approach that aids the development of radiation protection advice in respect of evacuation, sheltering and iodine prophylaxis.

Partial uncertainty estimation is deemed acceptable as long as major sources of uncertainty (source term and meteorology) are included, and omitted uncertainties (notably deep uncertainties i.e. within the time and data available to support the emergency management process, there is little chance of getting agreement on their evaluation or quantification) are documented. Implicit in the discrete nature of an ensemble approach is that the entire spread of uncertainty will not be described.

It seems likely that inter-model variability (including internal dispersion model uncertainty) will be of second order in the pre-release phase when uncertainties on the source term (and potentially the meteorology) are large, therefore source term and meteorological uncertainties will be the primary considerations in this report. It is recognised that inter-model variability might become significant at a later stage (during the release and more so post-release) when source term and meteorological uncertainties will tend to be significantly reduced.

Leadbetter et al. (2018) and Galmarini et al. (2004b) highlight that an ensemble approach could consist of a number of different dispersion models with a single (or even variable) meteorological model(s). It is thought that such an approach would be challenging, notably from a logistical perspective, requiring significant inter-agency collaboration, which has yet to be demonstrated for deterministic model runs in an emergency, let alone an ensemble approach. Therefore, operationalising a multi dispersion model ensemble approach is not explored in the remainder of this report. However, were such an approach to be considered, reflection of the inclusion and omission of different models (by type, and extent of validation and verification, for example) and potential weighting of the model results (based on how representative and revered the model is deemed to be, and not biased as a result of multiple users applying the same model) would be beneficial.

Recommendations of methods for the analysis (by the scientific community) and presentation (to the decision-making community) of uncertainty resulting from atmospheric dispersion modelling will not be a primary consideration in this report. However, it is recognised that recommending one or more ensemble modelling approaches, and the subsequent analysis and presentation of uncertainty, are interlinked. For example, there is a desire to minimise the number of ensemble members, but this in turn reduces the model output generated and therefore endpoints in the form of percentiles (and likelihoods) may become of less value. Where necessary, such aspects will be discussed.

This report does not seek to recommend one type of meteorological ensemble over another or to advise where to retrieve data to populate the meteorological and source term ensembles. Leadbetter et al. (2018) summarise the range of meteorological ensembles, including initial condition perturbation ensembles (multi model, model physics, multi parameter and stochastic physics approaches) and time lagged ensembles. However, the decision regarding which meteorological ensemble to employ will, for the majority of organisations (with a remit of providing advice in the event of a radiological accident) be driven by availability, at least in the short to medium term. Likewise, the decision regarding which source term ensemble members to consider will be driven by accessibility, most notably to expert knowledge, especially that held by staff at the pertinent (nuclear power plant) sites, and to key data sources, such as that derived in the FASTNET project (Mathieu et al. 2018) and (Chevalier-Jabet 2019b).

Approaches recommended in these guidelines should ideally replicate the “full” uncertainty. Here “full” uncertainty is considered to be that described by an ensemble approach including all meteorological and source term ensemble members (as considered in investigative work previously undertaken in this study), and termed here, “full configuration”. A “reduced configuration” is an ensemble approach including a subset of the meteorological and/or source term ensemble members. It is important that for a “reduced” ensemble (member) approach the respective radiation protection advice does not differ from that based on the “full” ensemble approach. In this study the reduced configuration is compared with the previously performed full configuration, evaluating whether there remains a suitable representation (or significant departure) of the uncertainty and whether model run time has decreased (to approximately one hour or less). Approaches are described in these guidelines that do not reflect the “full” uncertainty. Application of such approaches may be of value if limitations are transparent and the implications are qualified.

This section of the report utilises investigative model runs performed previously in this study, primarily on the basis of hypothetical scenarios termed “REM1” and “REM2” where “REM” is an acronym of Radiological Ensemble Modelling. For a detailed description of REM1 and REM2, see Korsakissok et al. (2019b). However, in summary, the principal difference between “REM1” and “REM2” is the release time and date. Furthermore, the meteorological conditions associated with the two scenarios differ significantly. For both “REM1” and “REM2” scenarios the full configuration meteorological ensemble comprises of 10 members. However, the full configuration source term ensemble comprises of differing numbers of members depending on which participant in Work Package 1 (WP1) undertook the respective model runs; therefore descriptions of individual applications of source term ensembles are summarised in the main body of this section of the report. Unless specified otherwise, the ensemble simulations used in this report were made with the short release scenario.

The remainder of this section of the report considers operational methods used currently to describe uncertainty, possible clustering and sampling methods for reducing the number of members in an ensemble approach, investigative work for reducing atmospheric dispersion model run time for a single ensemble member, additional considerations including computational resource and ensemble output, and concluding with recommendations for operationalising an ensemble approach and potential future work.

Operational methods used currently to describe uncertainty

Many organisations with a remit of providing radiological protection advice currently account for uncertainty within their assessments of a radiological accident by way of a manual and piecemeal sensitivity analysis of input variables within the dispersion model, based on expert judgement. Whilst this is of some value, this approach is limited (by the range of input variables which can easily be considered, by neglecting the interdependencies between input variables, and potentially by the labour-intensive nature of the approach, for example).

Some organisations therefore consider more than a single possible outcome (somewhat akin to that described by French et al. (2016) to describe the uncertainty. Thus, as well as a deterministic (best guess) simulation, a worst case based on "reasonably conservative" assumptions may also be considered. The intention is that this provides the decision maker with a better understanding of the potential for alternative outcomes, and helps ensure that suitable protective actions are advised. The worst case may comprise of a source term corresponding to no mitigating actions being taken, forecast meteorological conditions resulting in the most acute consequences, and conservative assumptions in respect of atmospheric dispersion model input parameters (such as those describing the parameterisation of wet and dry deposition). This approach is relatively simple and quick to implement (specifically, model estimates can be made available for the purposes of decision making within one hour) and provides an estimated upper bound of the potential impacts, including risks to public health. Where little or no consideration of uncertainty is made, this method is likely to be an improvement. However, the method has several limitations. Firstly, it is not possible to determine a single set of conditions that will result in the worst case in all situations; much depends on the model endpoint and uncertainty descriptor (as demonstrated in Table 33 and Table 34). Secondly, the provision of very conservative estimates (with no expression of graduated risk) may hinder decision making (as observed by French et al., 2016) and result in protective actions which do more harm than good. Considering an "optimistic" scenario, and one or two relatively conservative scenarios, alongside the aforementioned deterministic and worst case scenarios, may help develop a more balanced view of risk and provide a better basis for decision making.

The determination of a single set of conditions that will result in the worst case in all, or at least a subset of, situations has been investigated. The REM2 scenario was considered alongside a 90-member ensemble. The 90-member ensemble was comprised of 9 different source terms (3 release times and 3 release magnitudes) in conjunction with 10 different descriptions of the meteorology. A range of model endpoints were estimated, including the number of people affected by, the area affected by, and the maximum distance of, three different protective actions (evacuation, sheltering and administration of stable iodine) implemented in accordance with the UK's Emergency Reference Levels (ERLs) (Nisbet 2019). Further model endpoints were estimated, including the area affected and the mass or volume affected by banned green vegetables and milk, implemented in accordance with the EC's Maximum Permitted Levels (MPLs) (CEC 2016). For each model endpoint considered, the three worst case ensemble members were determined, where the first numeric value in Table 33 and Table 34 represents a source term (1-9) and the second numeric value represents a description of the meteorology (1-10). Whilst ensemble member 4,4 was frequently the worst case in respect of the lower ERL for sheltering and administration of stable iodine, in total 12 different worst case ensemble members were observed depending on the model endpoint and uncertainty descriptor concerned. Of the ten meteorological scenarios considered, eight featured in a worst case scenario for at least one of the endpoints. Furthermore, no single ensemble member consistently appeared in the top three worst cases for the majority of model endpoints. Similar results were observed for the REM1 scenario, but for purposes of brevity have not been included here. The results detailed in Table 33 and Table 34 highlight the difficulty in identifying a worst case scenario a priori.

Table 33. Worst case source term ensemble member and meteorological ensemble member (represented by numeric values) for a range of protective action model endpoints and uncertainty descriptors

	Primary		Secondary		Tertiary	
	Upper ERL	Lower ERL	Upper ERL	Lower ERL	Upper ERL	Lower ERL
Number evacuated	5,1	4,3	5,6	4,8	5,2	4,1
Area evacuated	5,1	5,3	5,6	5,8	5,2	5,4
Max distance evacuated	5,2	6,6	5,7	6,10	5,1	5,6
Number sheltered	4,3	4,4	4,8	4,9	4,1	6,6
Area sheltered	5,3	4,4	5,8	4,9	5,4	6,6
Max distance sheltered	6,3	4,4	6,9	4,9	6,8	6,6
Number stable iodine	4,8	4,4	4,3	4,9	5,1	4,5
Area stable iodine	5,1	4,4	5,6	4,5	6,9	4,10
Max distance stable iodine	6,6	6,9	6,10	6,10	5,6	6,3

Table 34. Worst case source term ensemble member and meteorological ensemble member (represented by numeric values) for a range of food restriction model endpoints and uncertainty descriptors

	Primary	Secondary	Tertiary
	EC MPL	EC MPL	EC MPL
Area green veg restricted	6,7	6,5	6,10
Mass green veg restricted	5,7	5,2	5,4
Area milk restricted	4,9	4,4	6,6
Volume milk restricted	4,9	4,4	6,6

For the REM1 case the meteorological conditions differed little between ensemble members, however for the REM2 case the meteorological conditions were much more variable. Therefore the variability in the meteorological conditions between ensemble members was not thought to be significant when explaining why the worst case ensemble member was seen to be so variable across model endpoints. The assessment applied a Lagrangian particle atmospheric dispersion model, NAME (Jones et al. 2007), alongside temporally and spatially varying meteorological data, resulting in non-uniform plume footprints. Thus, a dose contour associated with the upper ERL may be a different shape from the dose contour associated with the respective lower ERL (even for a single ensemble member). Furthermore, two different ensemble members may result in very similar dose contours associated with the lower ERL, but very different dose contours associated with the respective upper ERL. In particular, in the case of “numbers of people” (and also food production), both of which are spatially inhomogeneous, the movement of a contour, even by a small amount, can make a large difference, e.g. if near a town (are an agriculturally productive region). This is less applicable to “area affected” and “maximum distance”, but these can still be associated with different worst case ensemble members, especially if plumes are irregular in shape.

There have been only a relatively small number of examples of operational applications of ensemble approaches for describing uncertainty in the provision of advice in the event of a radiological accident. The Nordic Nuclear Safety Research (NKS) projects, termed MUD (Sørensen et al. 2014), FAUNA (Sørensen et al. 2016), MESO (Sørensen et al. 2017) and most recently AVESOME (Sørensen et al. 2019) have made significant advances in understanding the derivation and presentation of uncertainty, and this understanding has been employed in a Decision Support System called ARGOS. ARGOS applies a source term and meteorological ensemble approach, where substantial computational resource is used to overcome the significant computational expense resulting from large numbers of model simulations across large numbers of ensemble members. The NKS projects have also investigated how

efficiency savings in an ensemble approach can be made. Such efficiency savings include the application of a Latin Hypercube Sampling (LHS) approach to reduce the number of atmospheric dispersion simulations in respect of all dispersion dependent input parameters (including a meteorological ensemble). This is described by Sigg and Grahn (2019). Also included is a post-processing method for non-dispersion related properties (such as the magnitude of release and radioactive decay), akin to that described by Draxler et al. (2015). This is achieved by assuming a unit release of all emitted radionuclides.

A LHS approach requires knowledge of the parameter uncertainties and their probability distributions, which can be ascertained from Bedwell et al. (2018a). One challenge is to sample a suitable number of parameter combinations to cover the most important uncertainty features; Sigg and Grahn (2019) quote that it is typically necessary to consider 50-100 simulations (relatively few compared to brute force Monte Carlo Methods). LHS is a stratified sampling approach; Sigg and Grahn (2019) acknowledge that such an approach risks missing important parameter combinations.

Sigg and Grahn (2019) recommend that, where possible, source term ensemble members should be accompanied by a priori probabilities, and respective weighting should be employed when analysing and presenting the associated uncertainty. Care is required, as the probabilities are relative and not absolute. The same can be said for the meteorological ensemble members which are deemed to have an equal likelihood of occurrence, but the actual probability of occurrence is unknown. Furthermore, the possibility of receiving a priori source term ensemble member probabilities should be explored.

Clustering: reducing the number of members in an ensemble approach

A challenge when applying an ensemble of atmospheric dispersion simulations in the context of emergency response is to consider an appropriate number of ensemble members, ensuring that the ensemble contains enough members to accurately represent the underlying uncertainty, whilst minimising the number of members to ensure a computationally feasible approach. Additionally, an understanding is required a priori which subset of ensemble members should be retained in order to maintain the description of the uncertainty expressed by the "full" ensemble, whilst prior knowledge of the uncertainty propagation is limited. Clustering similar ensemble members is a method to achieve this.

The European Centre for Medium-Range Weather Forecasts have been producing Ensemble Prediction System (EPS) cluster products operationally since 1992 (Ferranti and Corti 2011), and there are a range of different clustering approaches (Ferranti 2017). Such clustering tends to be based on aspects relating to the synoptic flow (for example grouping by similarity between 500 hPa geopotential fields) because the application is weather forecasting. As a result, the broad method of such approaches may be applicable, but the specific details are not applicable here (for example an ensemble of atmospheric dispersion simulations may be clustered by 10 m wind fields, but is unlikely to be clustered by 500 hPa geopotential fields).

Klonner (2013) undertook a Masters thesis at the University of Vienna titled, "Clustering ECMWF ENS ensemble predictions to optimise FLEXPART plume dispersion ensembles". The aim of this study was to determine one or more suitable single-model multiple input ensemble methods for application in the estimation of forecast uncertainty within an operational volcanic ash monitoring and forecasting service. Klonner (2013) considered three different clustering analysis methods (Ward's method, K-means method and Partitioning Around Medoids method). The work concluded that reduced ensembles can be used for practical dispersion applications and K-means and Ward's methods were

recommended.(Galmarini et al. 2004a; Galmarini et al. 2004b; Galmarini et al. 2004c) describes and evaluates a multi-model ensemble dispersion method, based on the simultaneous analysis of several model simulations by means of ad-hoc statistical treatments and parameters. These references are of significant value for better understanding clustering and its application in dispersion modelling. The methods outlined could usefully be considered for the estimation of uncertainty within the provision of radiological protection advice following an accidental release to atmosphere. A caveat is that the applications demonstrated are specific to long range atmospheric dispersion modelling, however Galmarini et al. (2004a) recognises that the method can also be applied to short-range dispersion and weather fields.

A number of methods of clustering have been considered and are outlined below. Where practical and where time allowed hypothetical scenarios detailed in Korsakissok et al. (2019b) were re-run with proposed revisions to the ensemble approach (by way of clustering) and then the "reduced configuration" was compared with the previously performed "full configuration", assessing whether there remains a suitable representation (or significant departure) of the uncertainty and whether model run time has been decreased (to approximately one hour or less).

Pictorial representations of the combined ensemble results for the "full" and "reduced" configurations have been compared qualitatively based on expert judgement (and little or no statistical analysis has been applied). Whether such analysis and subjective acceptance criteria is adequate going forward should be reviewed, and consideration should be made of the timeliness and robustness of such a process. In some cases human interpretation can be beneficial, for example weather forecasters still use expert judgement in the interpretation of meteorological model output. But it is likely that expert judgement will be better informed if some kind of quantitative analysis is undertaken and fed into the process (see the section titled, "Uncertainty indicators", below for details of possible methods).

Automated clustering of ensembles using computational algorithms

A source term ensemble

The FASTNET ensemble of source terms (as described by Mathieu et al. (2018) and Chevalier-Jabet (2019b)) comprised of 197 members, each corresponding to a different hypothetical release scenario. In the process of uncertainty estimation, ideally all potential source terms would be considered (alongside all meteorological members of an ensemble), but a model simulation time of one hour or less could not be met. To reduce computation time the source term ensemble was reduced to ten members.

The goal was to determine ten source terms which represent the spread characterised by the original 197 members. This was trialled using principal component analysis (PCA). PCA was applied to data generated by the concatenation of Cs-137 and I-131 source term time series (and normalised to ensure that their influence within the PCA process was equivalent). Each time series represented 600 values. One concatenated source term and 600 temporal values resulted in 600 variables. Principal component analysis (PCA) is a statistical procedure that converts a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (such that the first principal component has the largest possible variance i.e. accounts for as much of the variability in the data as possible, and each succeeding component in turn has the highest variance possible given the preceding components). Five components were selected from the PCA results (Figure 86). Thus, each of the 197 source terms of the ensemble was defined as a point in a five-dimensional space (corresponding to the first five components). Note that the y axis in Figure 86 represents "new" uncorrelated variables called principle components (they are not variables as we would recognise) and are unitless. The x axis in Figure 86 and Figure 87 represents the number of timesteps and are unitless.

To aid in the optimisation of the solution and identify ten points which suitably represent the spread of the 197 points in this five-dimensional space, the Kennard-Stone algorithm was applied. These ten

points correspond to ten source terms as depicted in Figure 87 and Figure 88. The full method, applying PCA and the Kennard-Stone algorithm, can be used to identify a suitable subset of source term ensemble members in advance of performing the atmospheric dispersion modelling.

Note that in Figure 89, Figure 90, Figure 91 and Figure 92, the light blue region depicts the maximum to minimum range, the darker blue region depicts the 20th to 80th percentile range and the darkest blue region depicts the 40th to 60th percentile range. The dark line is the 50th percentile, while the dashed line is the mean.

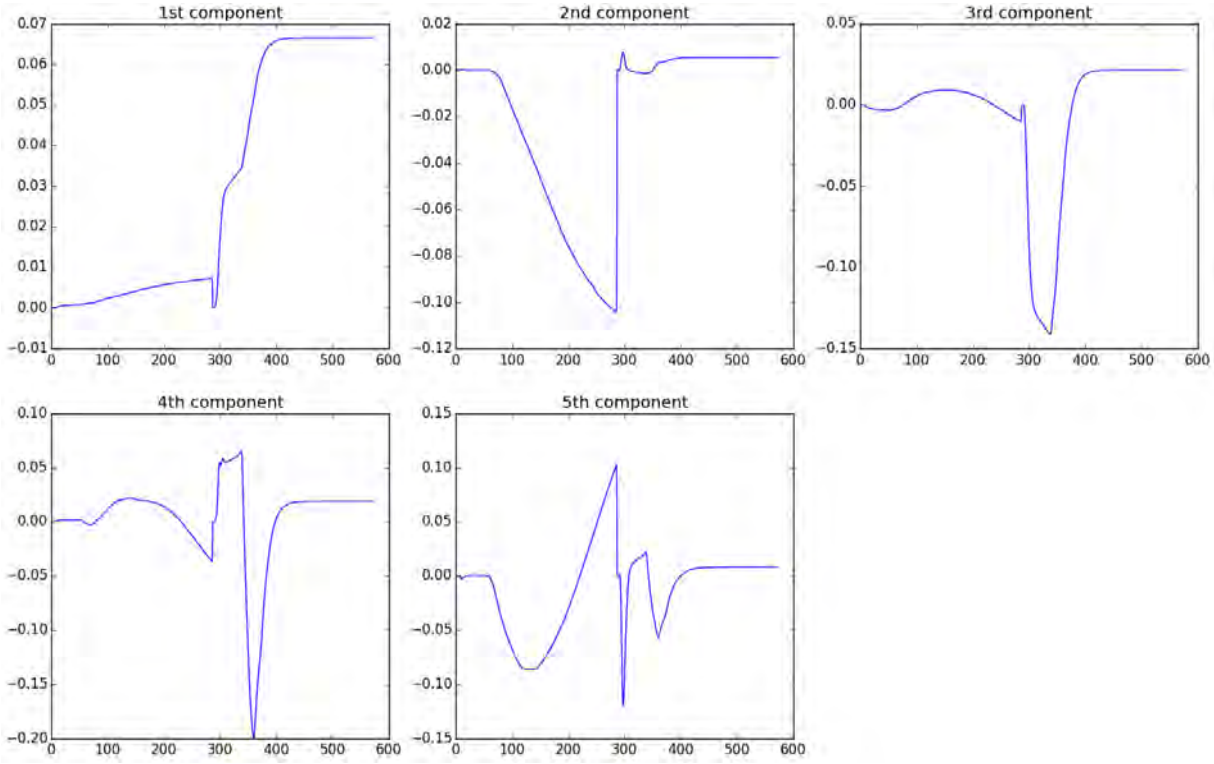
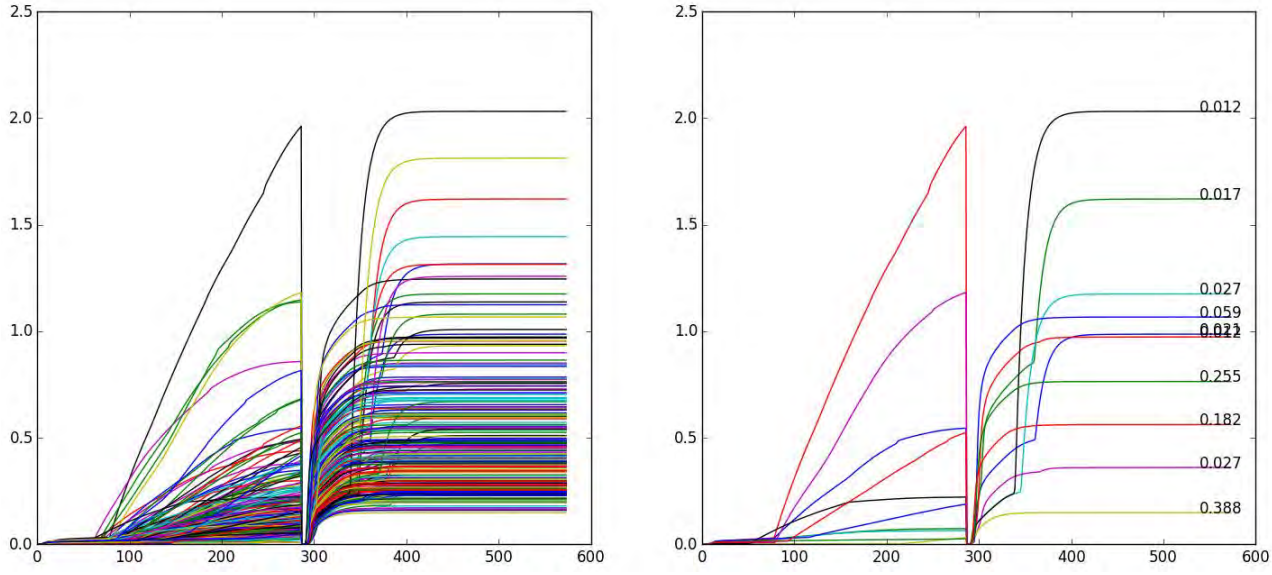
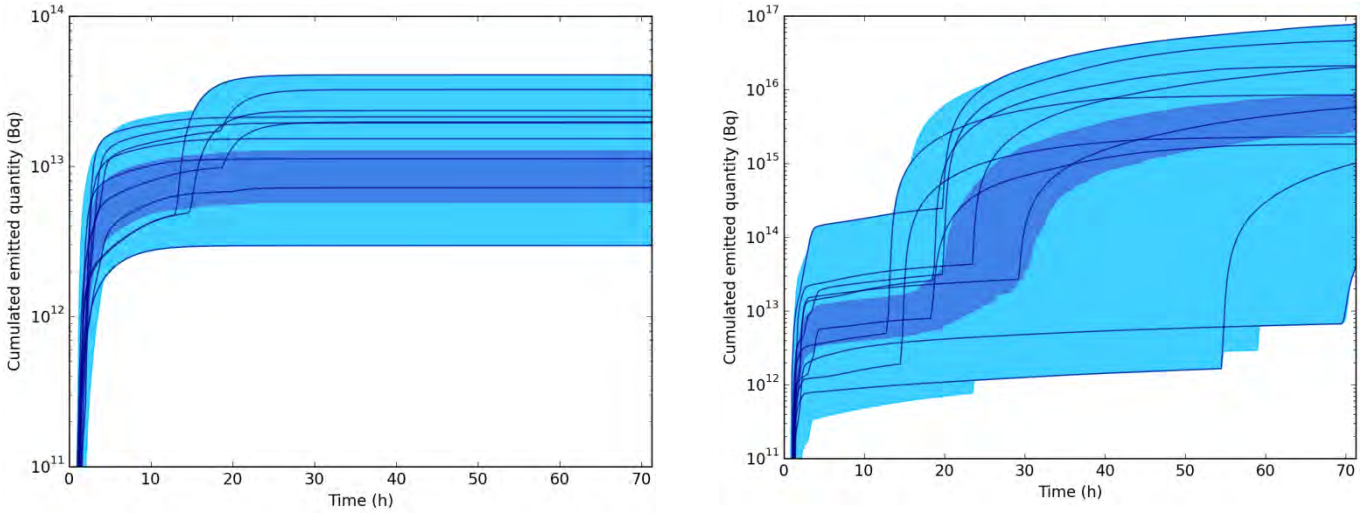


Figure 86. Five selected principal components as a result of PCA on the (normalised) concatenated Cs-137 and I-131 source term time series.



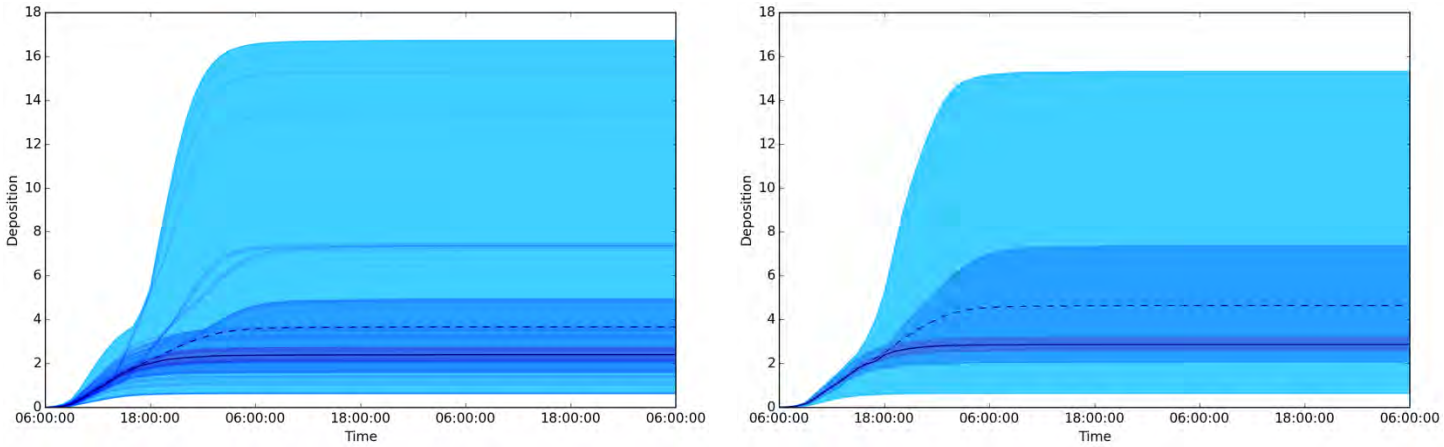
(a) (b)

Figure 87. Source term time series corresponding to the concatenation of Cs-137 and I-131 source terms and normalised for (a) all 197 source term ensemble members and (b) 10 selected source term ensemble members alongside their associated weighting.



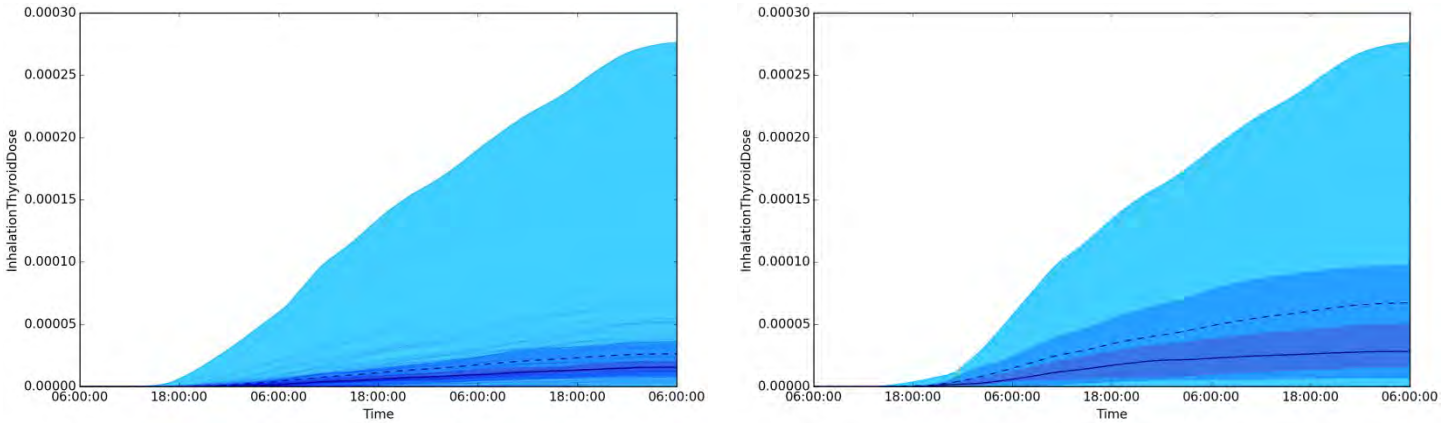
(a) Cs-137 (b) I-131

Figure 88. Selection of 10 source term ensemble members (illustrated by the thin blue lines), representing the spread characterised by all 197 members of the full FASTNET ensemble of source terms (illustrated by the light blue shading).



(a) (b)

Figure 89. Cs-137 deposition concentration ensemble results averaged over the model domain for (a) a full ensemble including 197 source terms and (b) a reduced ensemble including 10 source terms.



(a) (b)

Figure 90. Inhalation thyroid dose ensemble results averaged over the model domain for (a) a full ensemble including 197 source terms and (b) a reduced ensemble including 10 source terms.

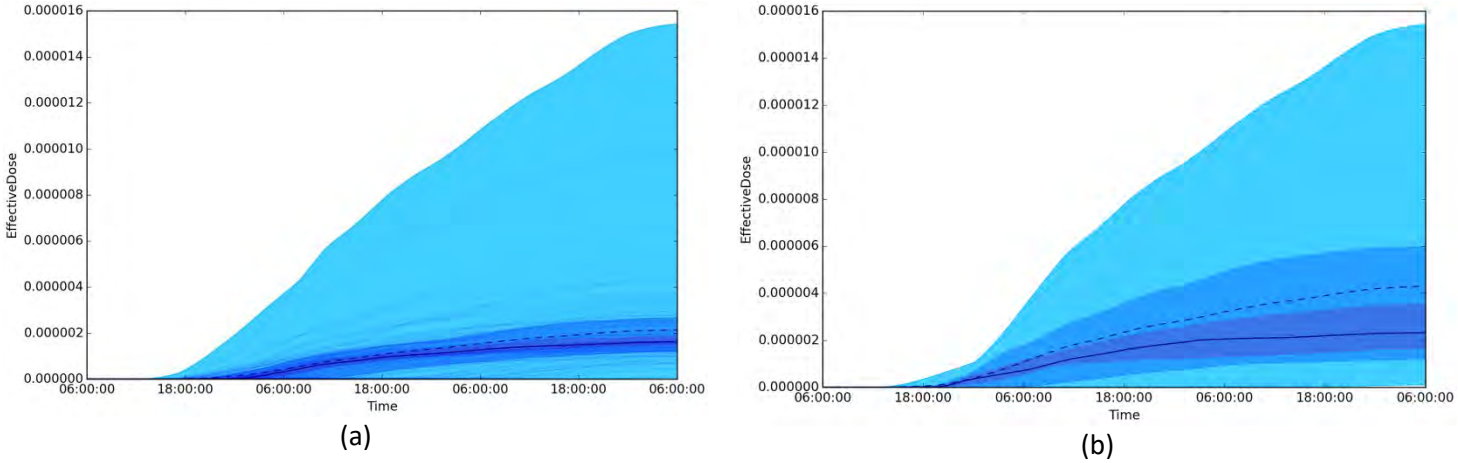


Figure 91. Effective dose ensemble results averaged over the model domain for (a) a full ensemble including 197 source terms and (b) a reduced ensemble including 10 source terms.

Table 35. Ensemble mean maximum distance (km) of threshold exceedance for both the full ensemble of 197 source terms and for the reduced ensemble of 10 source terms.

	Cs-137 Deposition		I-131 Deposition	Inhalation thyroid dose		Effective dose
	10 kBq/m²	37 kBq/m²	10 kBq/m²	10 mSv	50 mSv	10 mSv
10 ST	0.0	0.0	10.6	30.0	18.8	11.3
197 ST	0.8	0.0	13.6	27.2	9.7	1.9

Note however that a reduced selection of source terms does not necessarily replicate the results derived on the basis of the full ensemble, even if the selection has been intelligently made. The fact that source terms have been selected to provide a uniform coverage over the dataset does not automatically imply that the smaller selection respect the probabilistic repartition of the ensemble. In this instance the mean and the median of the ensemble members are not preserved in the reduced configuration. The general behaviour of the ensemble results was similar for both the full and reduced configurations (see Figure 89, Figure 90 and Figure 91), but discrepancies were evident upon closer inspection. For example, there was an overestimation of the mean Cs-137 and I-131 deposition concentration (averaged over the domain and at the end of the simulation) by approximately 30% in the reduced configurations. For dose estimates this effect is more pronounced, with overestimates of approximately 60% and 50% for estimates of inhalation thyroid dose and effective dose (averaged over the domain and at the end of the simulation), respectively. For estimates of the ensemble mean distances of threshold exceedance doses in Table 35 the same effect was observed.

To avoid biases in results and conserve features of the uncertainty described by the full ensemble (such as the ensemble mean), a weighting can be applied to each of the 10 source terms. The weightings should be derived by way of an optimisation algorithm. Subsequently the weightings could be applied as part of a Monte Carlo approach.

A limitation of PCA is that it is not well suited to temporally variable data. Such datasets may result in a relatively large number of (important) components. An alternative approach would be to consider self-associative methods which determine non-linear uncorrelated variables of the source terms instead (of linear uncorrelated variables of the source terms, as produced using a PCA approach). Self-

associative methods would therefore result in an improved description of the source terms and a more informed selection of the subset of source terms in a reduced ensemble.

A meteorological ensemble

For a meteorological ensemble, clustering by way of PCA is possible but not practical, because meteorological datasets contain a large number of variables, with each variable having multiple dimensions, and therefore the number of components which would need to be considered would be too great.

To simplify the problem, only those meteorological variables deemed to be most important were selected: wind speed and direction (by way of zonal and meridional wind) and the boundary layer height. Precipitation was not included for reasons detailed below.

The full meteorological ensemble used in the REM1 case study comprised of 10 members and the conditions across the different meteorological ensemble members in the REM1 case (considered) varied relatively little (Geertsema et al. 2019a; Korsakissok et al. 2019b). The clustering approach applied here was based on an algorithm which compares the variance of the original ensemble to the variance of all different subsets (for N members from one to nine) of the full ensemble and identifies which combination of 1 to N members results in the most comparable variance. In this case, five members were found to suitably represent the full meteorological ensemble (Figure 92). Only inhalation thyroid dose ensemble results (averaged over the model domain) are presented here for purposes of brevity. However, a similar level of agreement was observed between the full and reduced ensembles for both deposition concentration and effective dose.

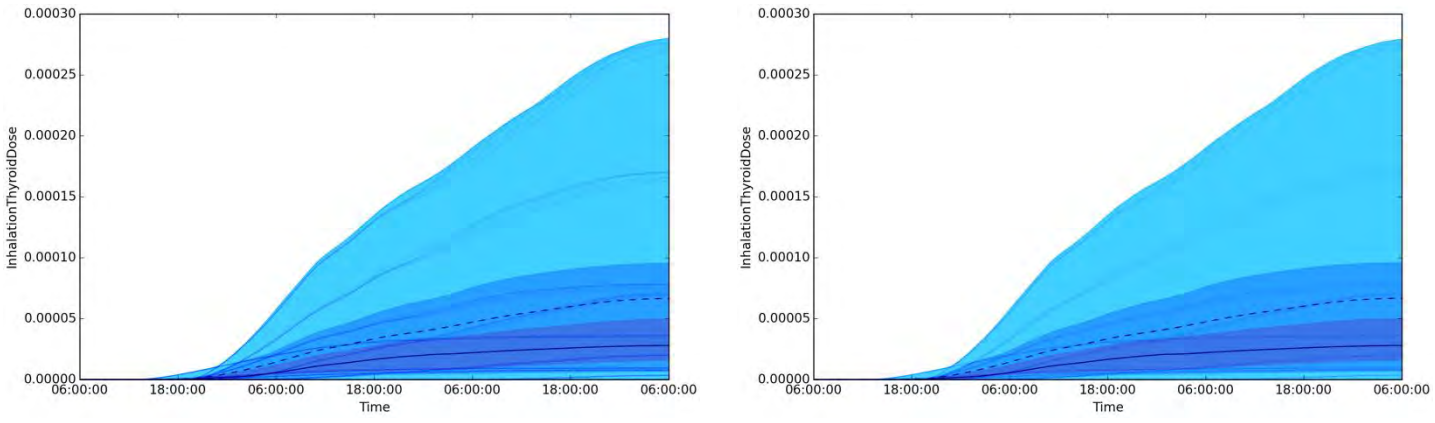


Figure 92. Average inhalation thyroid dose ensemble results for the REM1 case for (a) 10 meteorological members and (b) only 5 selected meteorological members.

This clustering approach was also tested using the REM2 case (Geertsema et al. 2019b; Korsakissok et al. 2019b), which features more variability in meteorological conditions as a function of time for a single ensemble member and more spread in the meteorological conditions across the range of ensemble members for single points in time (most significantly for wind direction). Again, five meteorological ensemble members were found to suitably represent the full ensemble.

The determination of five members to suitably represent the original ensemble for both REM1 and REM2 cases (where the meteorological conditions and the respective uncertainty differed markedly), were in part due to the manner by which the HARMONIE-AROME meteorological ensembles were

constructed. These ensembles were “time-lagged” and constructed using five different forecast lead times and two different versions of the physical model. The two model versions resulted in similar model output (at least, relative to other uncertainties) and, therefore, ensemble members were effectively paired (in this case, five pairs) and the variance was conserved by retaining one member in each pair. It would be of value to test this clustering approach with a meteorological ensemble constructed using an initial condition perturbation method. Preliminary comparison of the 10 ENS members with HARMONIE-AROME time series for the wind speed at 10 meter height show that the spread is comparable during the first 30 hours.

It should be highlighted that whilst this approach seeks to preserve the variance of the meteorological ensemble, it does not necessarily remove the ensemble members that are comparable to one another. A further limitation of this method is that the optimum number of ensemble members cannot be identified easily before performing the atmospheric dispersion modelling, so this method is not well suited to operational use.

Note that the inclusion of precipitation (and the large numbers of zero values associated with this variable) in the aforementioned approach would have adversely skewed the results. However, precipitation can have a significant impact on atmospheric dispersion (most notably deposition, but also potentially plume depletion), and therefore it is likely that it should be accounted for. To include precipitation in a clustering method, an alternative must be considered.

Manual clustering of ensemble members a priori – on the basis of wind direction and the timing of the release

A fast and simple approach is to consider a meteorological variable that is (in general) most influential in determining the key endpoints; the wind direction at the release location was considered here. This proposed method incorporates the assessment of wind direction meteograms and subsequent selection of representative atmospheric dispersion simulation ensemble members before performing the atmospheric dispersion modelling. The approach was presented in Tomas and Twenhöfel (2019).

This methodology was applied to scenarios REM1 and REM2 to illustrate its potential. Two endpoints were considered:

- The accumulated deposition concentration of I-131
- The time-integrated activity concentration in air (TIAC) of I-131

For accumulated deposition concentration, a threshold level of 5000 Bq m^{-2} was considered, representative of the intervention level for a grazing ban in the Netherlands. TIACs of I-131 is a suitable proxy for thyroid dose resulting from inhalation and can be used to assess whether iodine tablets should be provided to the public. Here, a threshold level of 5000 Bq h m^{-3} was chosen, such that considerable distances are achieved.

REM1 case

For the REM1 case, most of the overall uncertainty was associated with the uncertainty in the timing of the release. Figure 93a shows the wind direction at the release height for the REM1 case. The meteorological forecasts agree well, suggesting low uncertainty. However, because the wind direction changes over time during the course of the expected release (specifically from -6 to +6 hours), the uncertainty in the release duration has a significant effect on estimated atmospheric dispersion model endpoints. Because all meteorological forecasts contain this change in wind direction, a single meteorological forecast (Figure 93b) is used in the dispersion simulations.

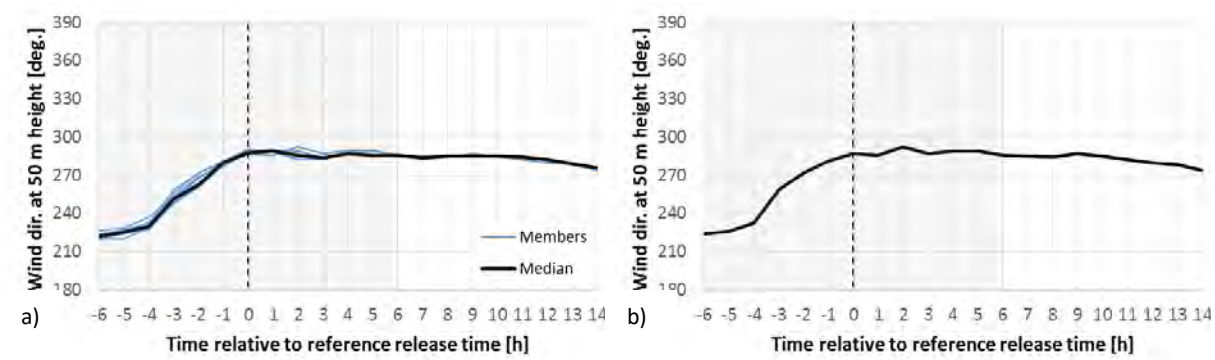


Figure 93. Meteograms of the temporal evolution of the wind direction at a reference release height (50 m) and source location for the REM1 case for (a) all meteorological members, and (b) (a priori) selection of 1 member.

Figure 94 shows contours of the "level of agreement" between the dispersion simulation members exceeding 5000 Bq m^{-2} deposition concentration of I-131. A (maximum) level of agreement of 1 indicates that all members of the ensemble predict an exceedance of the given threshold. Figure 94a illustrates the results based on all dispersion simulation members (a "full" ensemble of 650 members comprising of five release heights ranging from 0 to 100 m, thirteen release times spanning a period of twelve hours, and ten meteorological ensemble members), whilst Figure 94b illustrates the results based on a single meteorological forecast (ensemble of 65 members). The similarity between the figures is clear, indicating that the meteorological uncertainty is negligible compared to the uncertainty in the release moment. Figure 95 is analogous to Figure 94 but presents the level of agreement between results exceeding 5000 Bq h m^{-3} TIAC of I-131. The same conclusion can be drawn as for Figure 94.

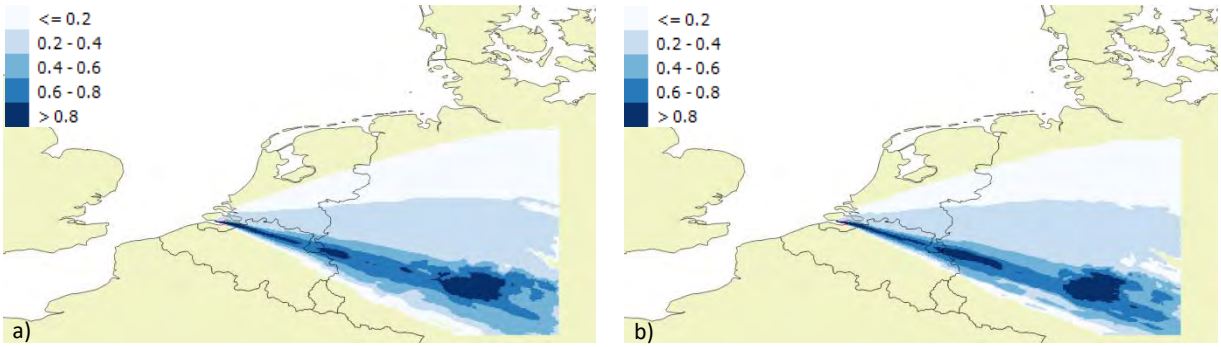


Figure 94. Level of agreement between results exceeding 5000 Bq m^{-2} deposition concentration of I-131 for (a) a full ensemble (650 members & 100%); (b) using a single meteorological member (65 members & 10%). Results of the REM1 case were derived by RIVM for 2017-11-12 12:00 UTC (24 hours after the reference release time).

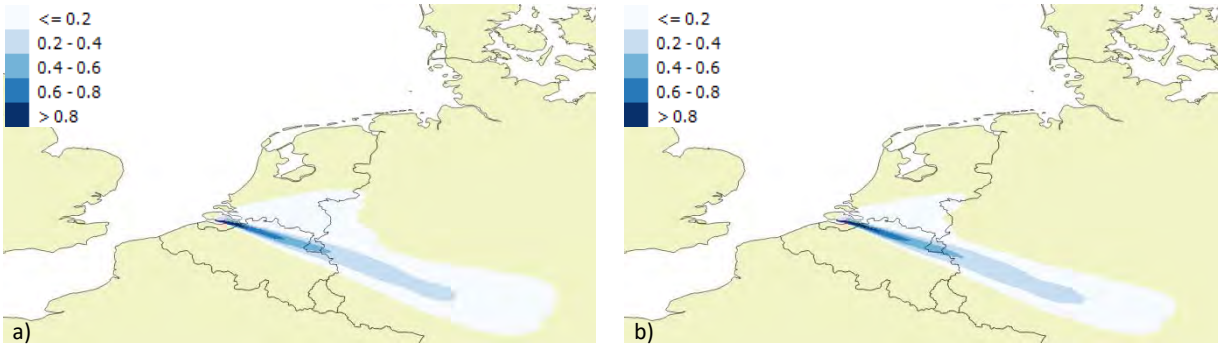


Figure 95. Level of agreement between results exceeding 5000 Bq h m^{-3} time integrated activity concentration in air of I-131 for (a) a full ensemble (650 members & 100%); (b) using a single meteorological member (65 members & 10%). Results of the REM1 case were derived by RIVM for 2017-11-12 12:00 UTC (24 hours after the reference release time).

REM2 case

For the REM2 case, the uncertainty in the meteorology plays a much larger role than for the REM1 case, as there is much less agreement between the meteorological members (Figure 96a). In the first seven hours of the period (from -6 to +1 hours) when the release may occur, the wind direction differs by up to 135 degrees between members. However, after a front has passed, agreement between ensemble members significantly improves.

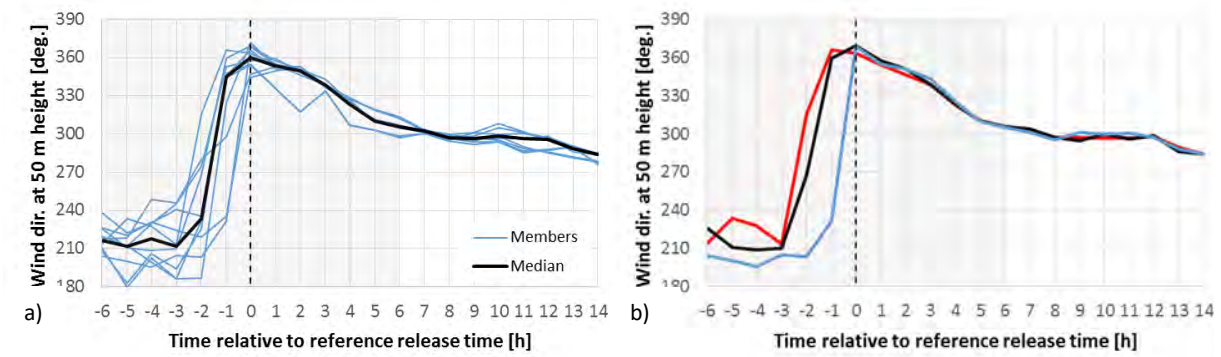


Figure 96. Meteograms of the temporal evolution of the wind direction at a reference release height (50 m) and source location for the REM2 case for (a) all meteorological members, and (b) chosen members that are representative of the minimum (blue), maximum (red), and median (black) values.

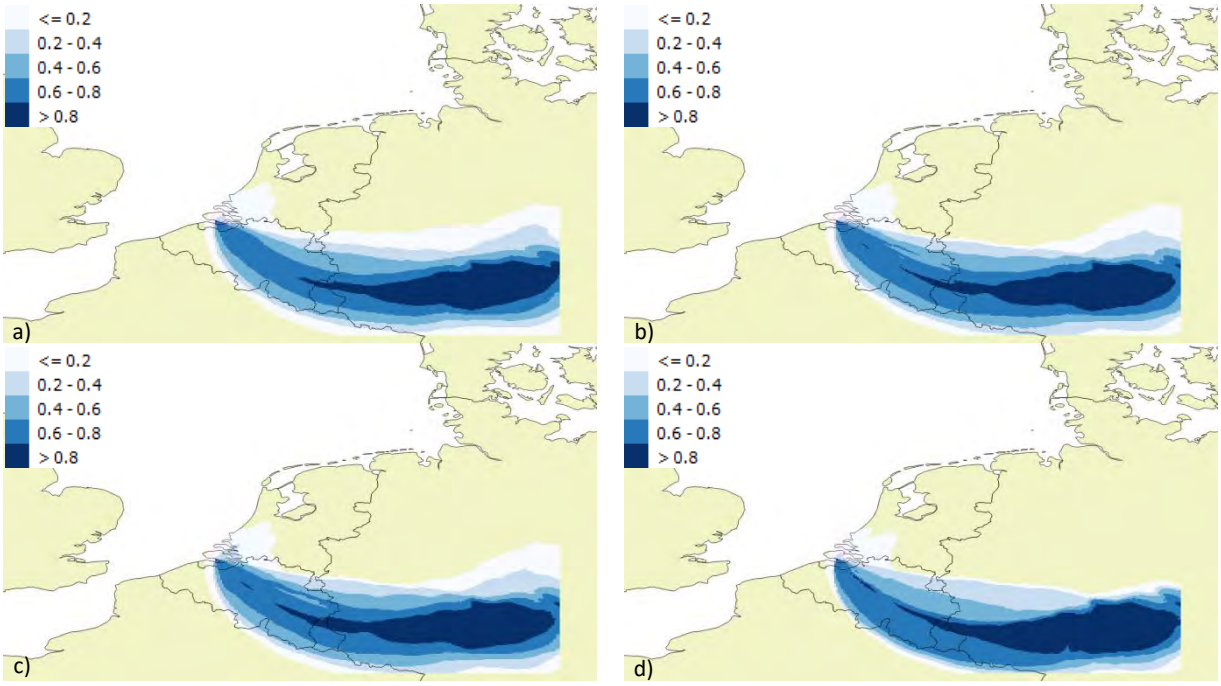


Figure 97. Level of agreement between results exceeding 5000 Bq m^{-2} deposition concentration of I-131 for (a) a full ensemble (650 members & 100%), (b) clustering 1 (380 members & 58%), (c) clustering 2 (280 members & 43%), and (d) clustering 3 (135 members & 21%). Results of the REM2 case were derived by RIVM for 2017-11-13 21:00 UTC (24 hours after the reference release time).

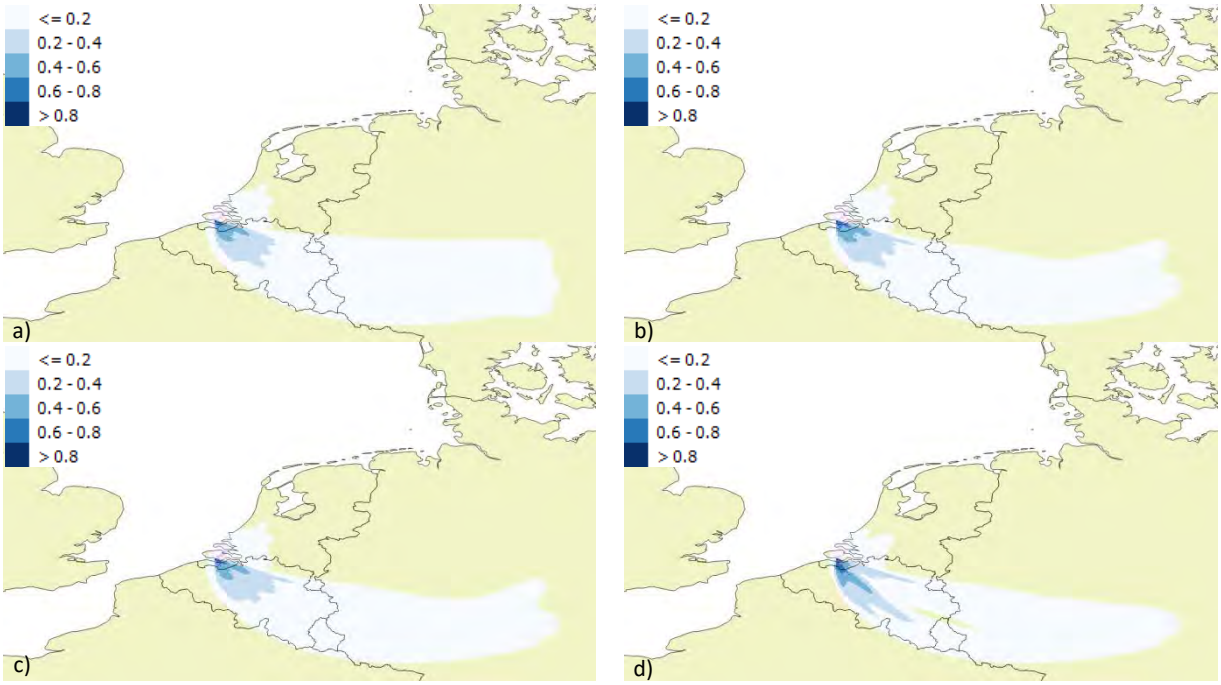


Figure 98. Level of agreement between results exceeding 5000 Bq h m^{-3} time integrated activity concentration in air of I-131 for (a) a full ensemble (650 members & 100%), (b) clustering 1 (380 members & 58%), (c) clustering 2 (280 members & 43%), and (d) clustering 3 (135 members & 21%). Results of the REM2 case were derived by RIVM for 2017-11-13 21:00 UTC (24 hours after the reference release time).

Reduction of the number of ensemble members: clustering 1

The dispersion simulation members with a release time of -6h to +0h (inclusive) relative to the reference release time continue to apply all (ten) meteorological members due to the large uncertainty

in the wind direction. However, the uncertainty in the wind direction in the period +1h to +6h (inclusive) relative to the reference release time is low. Therefore, a reduction in dispersion ensemble members can be achieved by considering only one meteorological member in that period. A meteorological member similar to the median value would be the most representative, depicted in Figure 96b by the black line. The "level of agreement" between the dispersion simulation members used to derive deposition concentration and TIAC are shown in Figure 97 and Figure 98, respectively. Figure 97a and Figure 98a show the results for all dispersion calculation members ('full' ensemble of 650 members), whilst Figure 97b and Figure 98b show the results based on the proposed reduction (an ensemble of 380 members). Clearly the reduction in dispersion calculation members hardly affected the outcome, whilst the size of the ensemble was reduced to 58% of the original.

Reduction of the number of ensemble members: clustering 2

This approach includes the reduction of the number of ensemble members described previously, but also applies a further reduction. Because the wind direction (on average) changes little between -6h and -4h relative to the reference release time, a further reduction can be achieved by applying the -6h runs to the -5h and -4h time stamps. This reduces the number of dispersion calculation members to 280, i.e. 43% of the original ensemble size. For this reduced ensemble, Figure 97c and Figure 98c show the "level of agreement" between the dispersion simulation members used to derive deposition concentration and TIAC, respectively. The reduced ensemble does show minor local differences compared to the original. However, in terms of distances and affected area, the outcome does not change significantly.

Reduction of the number of ensemble members: clustering 3

This approach includes the reduction of the number of ensemble members described in the initial revision ("clustering 1"), but also applies a further reduction. This approach does not consider the full spread in meteorological members (i.e. does not consider all ten meteorological forecasts) in the period -6h to 0h inclusive, but approximates the variation in the members by considering members that are representative of the minimum, maximum and median values of the wind direction (see Figure 96b). This reduces the number of dispersion simulation members to 135, i.e. 21% of the original ensemble size. For this ensemble, Figure 97d and Figure 98d show the "level of agreement" between the dispersion calculation members used to derive deposition concentration and TIAC, respectively. This method of reducing the size of the ensemble does affect the outcome, as can be seen in both the maps of deposition concentration and TIAC. With only three meteorological forecasts in the period -6h to +0h, the spatial spread is relatively well captured on a larger scale when considering the level of agreement between results exceeding a threshold deposition concentration for use when providing advice in relation to a grazing ban (and is likely to apply more broadly to food ban). However, the spatial spread is less well captured on a smaller scale when considering the level of agreement between results exceeding a threshold time-integrated activity concentration in air for use when providing advice in relation to the administration of stable iodine. As a result, such a reduction in ensemble members seems unsuitable for local protective action advice, such as sheltering, evacuation and administration of stable iodine, but further work is required to determine if such a reduction could be applied in the provision of food restriction advice. Thus, for food restriction advice it seems that ten meteorological ensemble members is sufficient (and perhaps fewer than ten may suffice?). But ten meteorological ensemble members may not be sufficient to capture a suitable representation of the uncertainty in the wind direction for local protective action advice, and further work considering a larger number of meteorological ensemble members is required before any firm conclusions can be drawn.

Note that this ensemble approach considered five different release heights in the range 0 – 100 m. It is likely that the number of ensemble members could be reduced further in this case by reducing the number of release heights considered; however, this was not the focus of this investigation.

Computation time

Model simulations for all 650 ensemble members took approximately 13 hours (on eight threads of a 4-core Intel® Core™ i7-6700T Processor 8M Cache). The reduction in simulation time was approximately proportional to the reduction in the number of ensemble members. Specifically, 280 members and 135 members were computed in 5-6 hours and 2-3 hours, respectively. Further time would also be required for the analysis of model results, including the generation of plots of level of agreement.

Visualisation of Meteorological Data

When making decisions about how to cluster a meteorological ensemble it is useful to visualise the ensemble meteorological data, which can be plotted as a gridded field or as a time series (sometimes known as a meteogram). Where the region of implementation of protective actions is small (a few km to a few tens of km) it is reasonable to assume that the meteorology at the release location is representative of the whole area of interest and thus to visualise the meteorology simply as a time series. However, where the region of interest is larger, or the meteorological dataset is spatially variable (for example, in complex terrain), gridded meteorological data should also be visualised.

For the purposes of describing uncertainty arising from the process of atmospheric dispersion modelling, time series of wind direction (at a default height of 10 m, and as a function of height) is likely to be of significant value. Time series of wind speed, precipitation (and precipitation rate) and boundary layer height may also be of notable value. Less commonly considered, but still of potential value, are cloud cover (as a percentage), Pasquill Stability (as a number which varies continuously, but which is not an explicit output of all meteorological models and is dependent on the method applied) and 1/(Obukhov Length). In most cases the meteorological variable is represented by a line; however, plots can be presented as a boxplot, wind direction can be represented by arrows, lines or wind barbs within a series of circular plots, and precipitation by a column (or histogram style) format. In addition, plots can include the ensemble mean (as a line), ensemble median (as a line), percentile ranges (as shading) and a deterministic simulation (as a line).

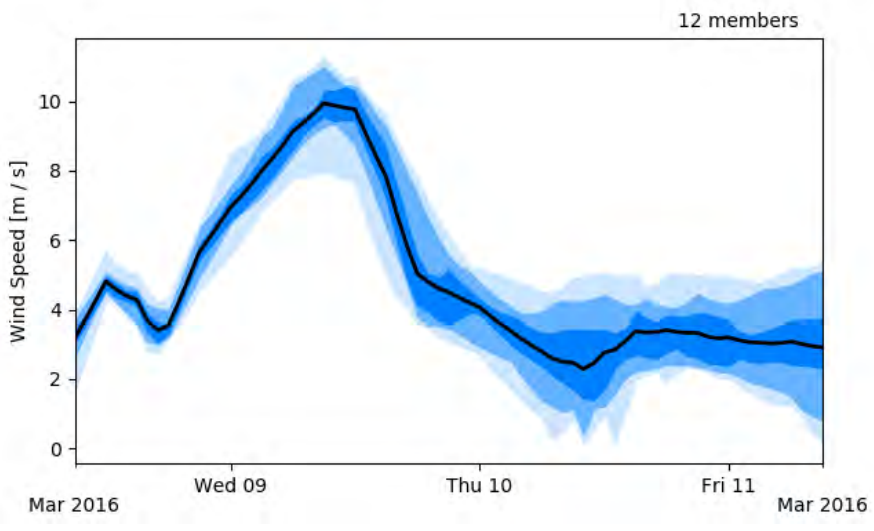


Figure 99. Meteogram of wind speed (m s^{-1}), including the mean wind speed depicted by a solid black line, the 0th, 10th, 25th, 75th, 90th and 100th percentiles depicted by different degrees of blue shading.

Meteorological variables are often correlated so it may be useful to plot several meteorological variables together (and in such instances consistency of format is likely to ease interpretation). Methods of presentation should be considered, including the ability to differentiate between individual ensemble members, and if a time series line plot is used to describe wind direction, the y axis should be based on the midpoint reflecting the dominant wind direction rather than simply running from 0 to 360 degrees. Context may be beneficial, for example indicating the degree of spread relative to historical climatology and thereby assessing whether the "current" level of uncertainty is large or small (perhaps by way of simple colour coding). Also, the number of ensemble members considered should be detailed on a meteogram.

Furthermore, meteorological observations could be considered in an effort to assess the uncertainty in a meteorological ensemble, and could even be used to reduce the number of considered ensemble members, but such an approach would be somewhat limited due to time constraints.

Visualisation of ensemble members indicates where significant confidence and significant uncertainty exist in the (forecast) meteorology as a function of time. Plots of meteorological data are likely to be of significant value in developing an early understanding of the uncertainty associated with the fundamental meteorological parameters and may be of significant value in developing a method for reducing the number of meteorological ensemble members.

Criteria for clustering ensemble members

The example of clustering based on wind direction alone highlights a potential approach, but this approach has been demonstrated for only two scenarios. There are an infinite number of possible scenarios, and exploration of a large but discrete number of scenarios would be required before more general conclusions can be drawn on the suitability of the approach and criteria for clustering. It is clear that identifying criteria for clustering ensemble members is complex. This section of the report suggests some possible methods and considers some key factors for determining clustering criteria.

It helps to frame the problem. In the first instance urgent protective actions (evacuation, sheltering and administration of stable iodine tablets) are considered. The implementation or planned implementation of these actions tend to be of the order of a few to a few tens of kilometres from a release. Over such spatial scales it can be assumed that the meteorological conditions will vary relatively little as a function of space and therefore the meteorological conditions at the release location, and the respective meteograms, only need be considered.

Bedwell et al. (2018b) investigate the extent to which the uncertainty on atmospheric dispersion models' various input variables has an effect on the models' outputs, and in particular to determine which input variables are most influential in the sense of giving rise to the largest uncertainties on those outputs. The primary input variables identified were source term (i.e. the magnitude of the release) and wind direction, where the latter was considered in the example of clustering based on wind direction (and the timing of the release). However, Bedwell et al. (2018b) also recognise that this is very scenario dependent. The second category of input variables included parameters that will often (but not always) be influential and that are capable of having a large influence in some circumstances, including plume rise, release height, wind speed and release timing (or time shift). The third category of input variables included parameters that are sometimes not relevant at all, but that can be very influential in some circumstances, notably precipitation. The fourth category of input variables included parameters that will often have some influence, but usually only a moderate one, including release duration, atmospheric stability class and mixing height. Further categories of progressively

diminishing influence were detailed but the associated input variables were deemed unlikely to be significant when determining criteria for clustering.

The example of clustering based on wind direction, adds to the evidence that wind direction must be included as part of any criteria. This is corroborated by (Haywood et al. 2010) who identified that wind direction is a key factor in determining imprecision in radiological emergency response assessments, and is substantiated by Figure 100 which illustrates the significant uncertainty (and variability) in wind direction (in the first one and two hours after the start of a release) across all considered wind speeds, but especially at low wind speeds (Lablans and Rijkoort 1974).

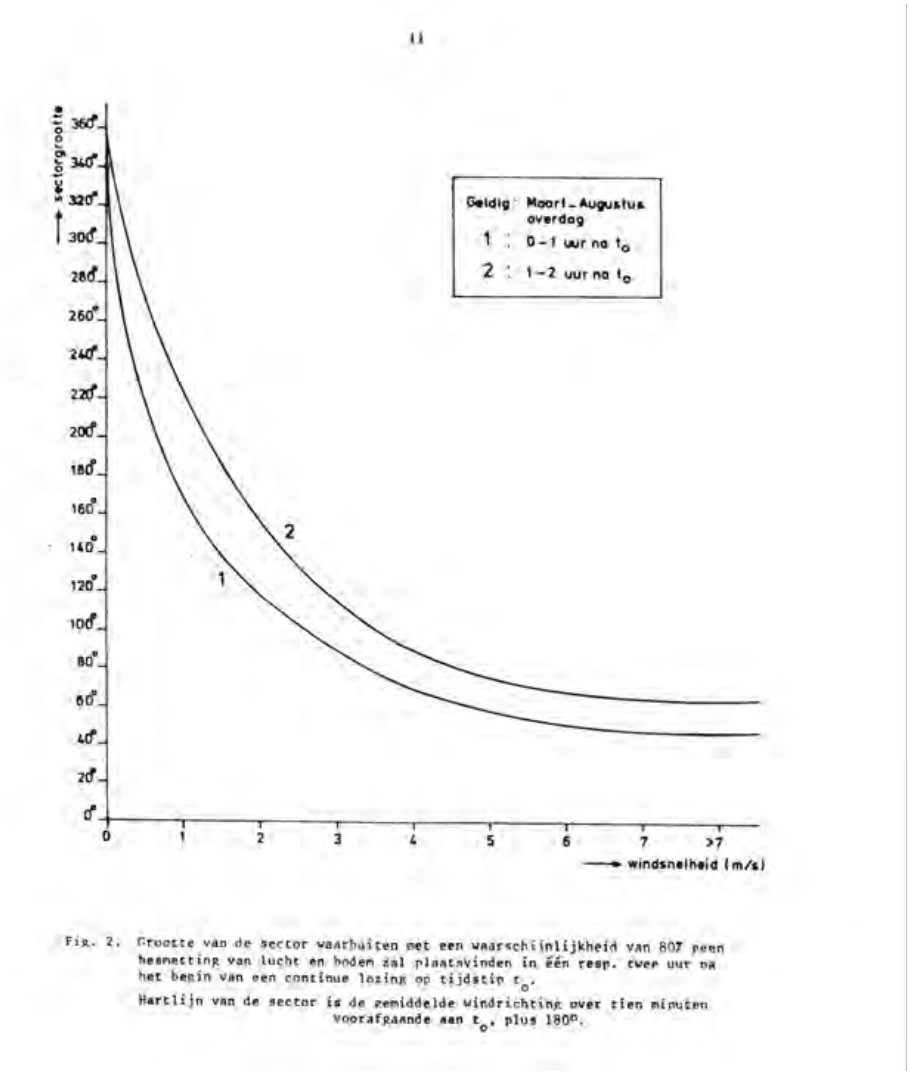
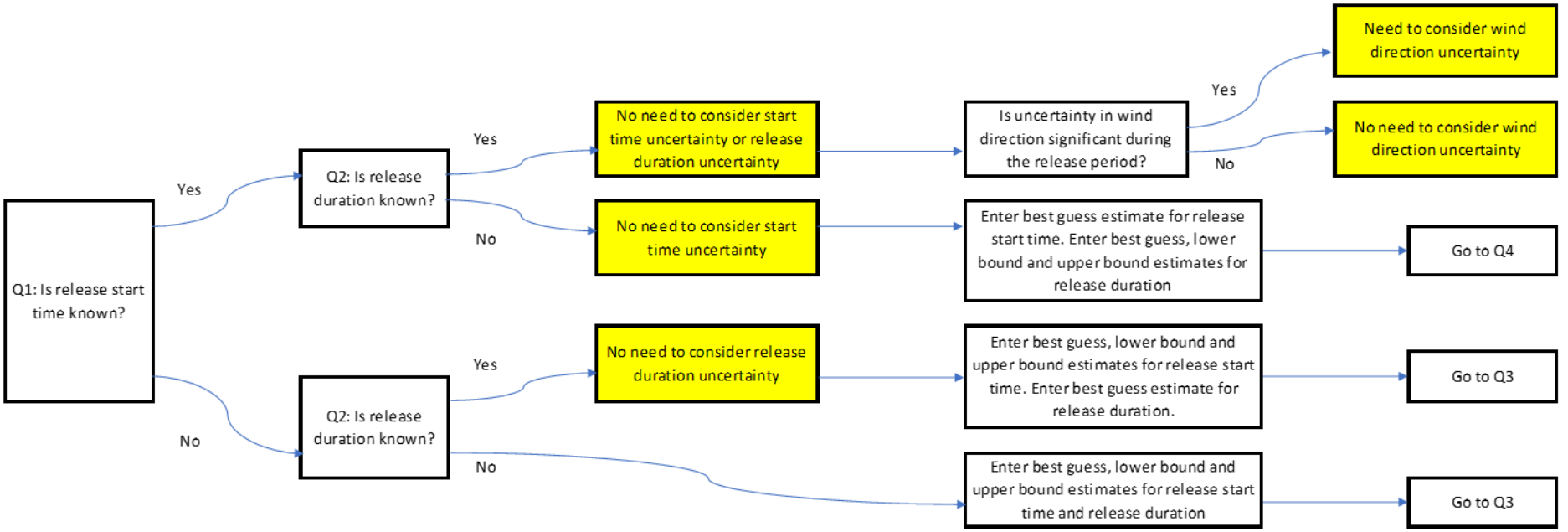


Fig. 2. Grootte van de sector waarbuiten met een waarschijnlijkheid van 80% geen besmetting van lucht en boden zal plaatsvinden in één resp. twee uur na het begin van een continue lozing op tijdstip t_o.
Hartlijn van de sector is de gemiddelde windrichting over tien minuten voorafgaande aan t_o, plus 180°.

Figure 100. Size of the sector (representing the uncertainty and variability in wind direction) beyond which no contamination of the air (or ground) will occur (with an 80% probability) in (1) 1 hour and (2) 2 hours after the start of a release (at time = t_0). The centerline of the sector is the wind direction averaged over 10 minutes before t_0. Data derived on the basis of measurements of noxious gases (from the Netherlands). Applicable to daytime for the period March-August (a second figure, not included for purposes of brevity, was applicable to the period September-February, and at nighttime from March-August). X-axis: wind speed in m/s. Y-axis: sector size in degrees.



Continued on next page…

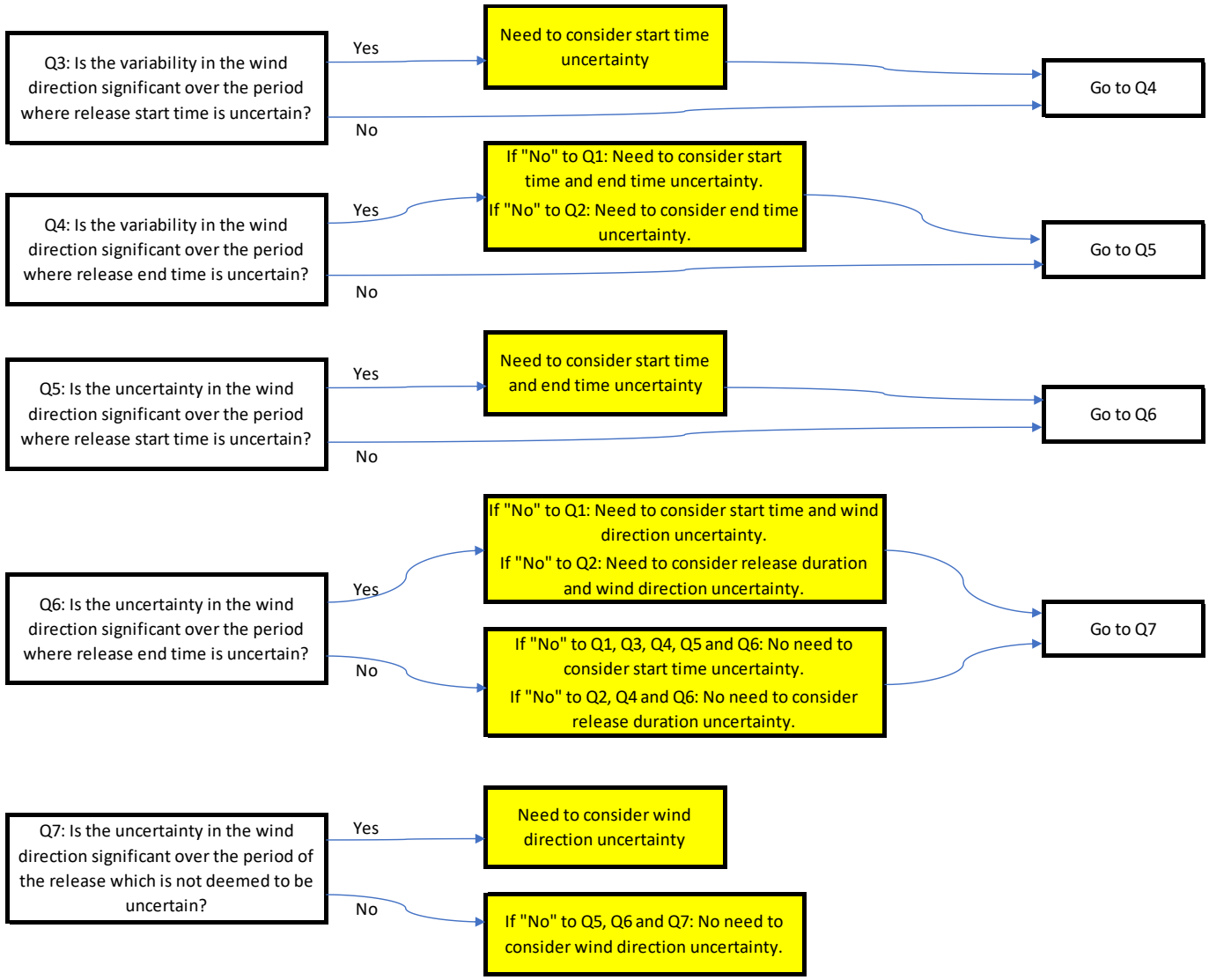


Figure 101. The steps required in a decision-making process for whether or not wind direction (and release duration, start and end time) uncertainty need to be taken into account (ignoring the fact that there may be other uncertainties that need to be taken into account)

As in the manual clustering approach described previously, it is assumed here that clustering is performed on the basis of only one meteorological variable: wind direction. The relationship between meteorological variables and time is such that source term uncertainties (notably release start time and release end time) must be considered in parallel. It is likely that in any one scenario where uncertainty is assessed, that other meteorological variables will also be of significance; however, to simplify the problem, only wind direction has been considered. The flow diagram in Figure 101 represents decision-making steps for determining whether or not the uncertainty relating to a meteorological variable need be considered by way of an ensemble. Figure 101 is a simplified example and assumes that where there exists significant uncertainty and/or variability in the wind direction, this remains significant for the entire period (but as demonstrated in Figure 96 this may not be the case). Source term data and meteorological data would need to be input into the process. Input of source term data could be a manual step or could be automated by reading source term files. Meteorological data could be input using meteograms. An algorithm would then interrogate the data to determine which uncertainties need to be considered and which can be omitted.

However, the challenge of identifying suitable criteria for clustering remains (specifically, what is "significant" in terms of wind direction variability and uncertainty has yet to be defined). Haywood et al. (2010) considers the extent to which protective actions are estimated to vary as a result of a +/- 10 degree and +/- 20 degree shift in wind direction. It is evident in the study that for uncertainty in wind direction greater than 10 degrees the impact on the respective model endpoints is significant. Clustering criteria could be based on all of the ensemble members being within a range of 10 degrees. However, in order to determine robust criteria for clustering (on the basis of wind direction) significantly more results (in published literature) would be required.

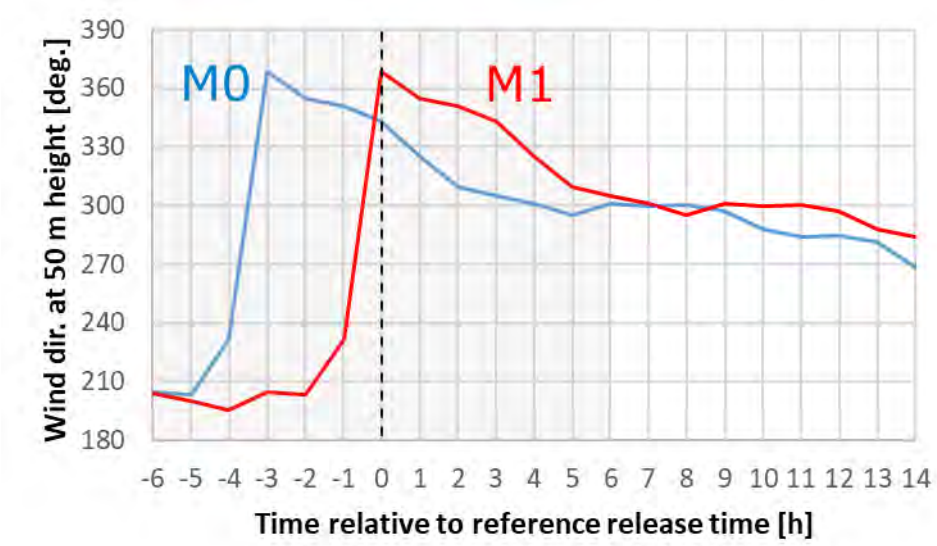


Figure 102. Example of a pair of ensemble members with similar wind direction profiles, which may occur as a result of uncertainty in the timing the passage of a (cold or warm) front (for example)

Clustering could be based on a correlation criteria, whereby the profiles (e.g. of the wind direction) are similar but the timings may differ, as illustrated in Figure 102. It could be imagined that such clustering could be performed when there exists uncertainty in the timing of the passage of a (cold or warm) front. In Figure 102 the two wind direction profiles are identical for illustrative purposes (i.e. $M0_{t-3h}$ = $M1_{t+0h}$) but in reality, they would never be identical (i.e. $M0_{t-3h} \approx M1_{t+0h}$). A potential method would assess the correlation of each pair of ensemble members, and when two members are suitably correlated (according to a criterion), they are clustered. There are however a number of caveats. Firstly, do such consistencies in the profiles of two or more ensemble members (frequently) occur across the range of key meteorological variables; a large number of model runs would have to be

performed to investigate this. Secondly, do circumstances exist whereby ensemble members x and y meet the criterion for the correlation, ensemble members y and z meet the criterion for the correlation, but ensemble members x and z do not meet the criterion for the correlation, and if so, is there a solution to this problem.

The criteria for a priori clustering could be determined by some form of statistical analysis. It is suggested that an assessment of the sensitivity of a model endpoint of choice to a change in one or more input variables could be undertaken. For example, for wind direction at the source location a large number of simulations could be undertaken which suitably scope the range of possible meteorological conditions and source terms. For source terms, perhaps the model runs could be simplified to focus on an I-131, Cs-137, actinide and noble gas release. Model runs could be performed within PSA software (such as PACE (Charnock et al. 2013)), which are designed to perform many simulations (cyclically) sampling over meteorological data, to ensure that weather conditions (preferably over a dataset of multiple years) representative of both annual and diurnal variability are considered. For each simulation, an assessment of the change in a model endpoint (e.g. sheltering area) as a result of a change in wind direction could be performed. For example, for all the model results where the change in wind direction is less than or equal to +/-10 degrees, what is the respective mean or median and 95th percentile or 100th percentile sheltering area. The sensitivity of a model endpoint to a change in one meteorological parameter may be intrinsically linked to a change in another meteorological parameter. Therefore, for each simulation (described previously), an assessment of the change in a model endpoint should not only consider the preceding change in wind direction, but the preceding changes in all relevant meteorological parameters (highlighted below). To evaluate a suitable range of meteorological scenarios and the interdependencies of key meteorological parameters, many model runs would need to be performed, perhaps beyond the number of model runs that is practical. Were such an approach feasible and to be employed, a possible alternative could be to “map” any one scenario to a comparable scenario from a database of many scenarios.

There is a requirement to determine the level of uncertainty in the model endpoint that is deemed acceptable, i.e. a degree of uncertainty in the model endpoint below which the uncertainty is deemed negligible (possibly on the basis of the variance or a percentile value). If uncertainty is deemed negligible, all resultant radiation protection advice can be based on a deterministic model run. The respective input variable uncertainty can then be used as a threshold value, below which the input variable concerned is omitted from the clustering process. It is worth highlighting that the determination of (for example) “uncertainty in sheltering area by more than +/-x%” may differ from country to country, where different countries are happy to accept (or live with) different levels of uncertainty; so country A may accept +/-5% and country B may accept +/-10% uncertainty in sheltering area (for example).

If an assessment were to be performed for all key input parameters, (general) statements (on the basis of the mean or median values) can be made such as: +/- half an hour in release duration introduces an uncertainty in sheltering area of +/-5%, which is similar to the effect of +/-10 degrees in wind direction at the source location. Then, when clustering, the criterion can be defined as (for example) each (a priori) key input variable may affect the sheltering area by no more than +/-5%.

To this point meteorological uncertainty in respect of only one meteorological variable (wind direction) has been considered. However, it may be necessary to consider the uncertainty in respect of other meteorological variables, notably wind speed, boundary layer depth, atmospheric stability and precipitation.

It is evident in Figure 100 that relatively low wind speeds tend to be associated with significant (wind direction) variability and uncertainty, and under such conditions wind speed is likely to be an important

variable within a meteorological ensemble. Furthermore, where wind speed was variable as a function of time (over the period of the potential release), including a spell of relatively high wind speeds (for the REM2 case), it was observed that the uncertainty indicator "maximum distance above a given threshold" was markedly variable. Therefore, it is evident that relatively large wind speed conditions can also be an important variable within a meteorological ensemble, however this may be dependent on the uncertainty indicator being applied.

Boundary layer (BL) height is likely to be an important variable within a meteorological ensemble when there exists a union between the respective uncertainty and release height uncertainty, and when significant BL uncertainty coincides with a release inside a shallow (e.g. nocturnal) BL.

Atmospheric stability may be an uncertain variable within a meteorological ensemble for a release which occurs around sunrise or sunset (when there may be a significant shift in stability over a relatively short period of time) and such uncertainty may be amplified by further uncertainty in the release start time and/or release duration (Kendell et al. 2019). There may also exist large variability across meteorological ensemble members for a coastal site.

To investigate the importance of precipitation within a meteorological ensemble (for model endpoints relating to protective actions) model runs for the REM2 case (using the Met Office's NAME model) were repeated but assuming no precipitation occurred (i.e. modelled by "switching" off wet deposition). For this scenario it was found that precipitation uncertainty was not a significant factor when considering the implementation of evacuation, sheltering and administration of stable iodine tablets (on the basis of the UK ERLs). For each of the three protective actions, the total number of people affected, the total area affected and the maximum distance affected were all estimated. For these model endpoints the maximum was primarily considered but minimum, mean and a range of percentile values were also derived (over the different ensemble members). Across the nine different model endpoints (and two different ERLs), the maximum values derived for model runs with and without precipitation were mostly in agreement within a factor of two (only for total number of people evacuated on the basis of the lower ERL and total number of people sheltered on the basis of the upper ERL was the difference a factor of 2-3). There was however greater disparity for some of the other statistical endpoints and significant disparity for food restriction endpoints. Further analysis, investigating the spread of the ensemble results (for example the variance) would be of value. Furthermore, only a single scenario has been considered here; many more scenarios would need to be considered before firm conclusions could be drawn regarding whether precipitation uncertainty need be considered within a meteorological ensemble. In contrast, for a deterministic model run precipitation uncertainty *is* likely to be a significant factor.

Consideration of a manual clustering approach of ensemble members a priori has been based solely on uncertainty attributed to wind direction. Whilst in some scenarios this may suffice, it is clear that in other scenarios there would be a need to consider additional meteorological variables (such as wind speed, boundary layer height, atmospheric stability and possibly precipitation). This is clearly an area for further work.

Clustering criteria may also need to include a minimum threshold (as well as a maximum threshold) number of ensemble members, depending on the analysis and presentation method of the ensemble results. If percentiles or fraction of agreement between the different ensemble members (e.g. for an exceedance of an environmental concentration threshold value) are used, a minimum number of ensemble members will be required to ensure that the results are not statistically noisy. This minimum number is also likely to depend on the level of presentation detail. Verification of proposed methods for simplifying an ensemble will be critical to ensure that the distribution remains suitably representative.

Sampling: reducing the number of members in an ensemble approach

The full configuration (or baseline) number of ensemble members considered in the REM case study varied by WP1 participant, ranging from 50 to 650. A uniform sampling method (across all considered ensemble members) was typically employed, however in one instance a Monte Carlo approach was employed.

Sampling can be used to reduce the number of ensemble members as demonstrated by Sigg and Grahn (2019) in an application of a Latin Hypercube Sampling (LHS) approach.

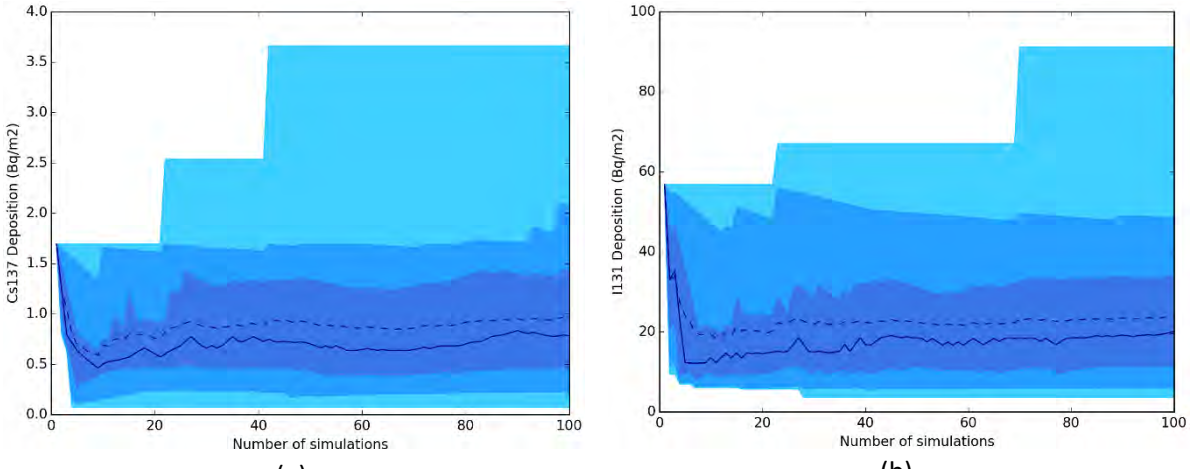
There are a large range of sampling approaches detailed in literature. Such sampling approaches could be used to sample a full ensemble or could be used to sample a clustered ensemble. A commonly recognised approach is the Monte Carlo approach, and a possible application is discussed below.

Using a Monte Carlo approach

A Monte Carlo ensemble simulation could be performed by randomly sampling meteorology and source terms, and therefore avoid considering all possible combinations.

The number of simulations required for Monte Carlo analysis depends on the probability distribution (or profile) of the output and in operational use it mainly depends on the computation resources available (because a greater degree of sampling is required where the distribution is more complex). When considering a meteorological ensemble alongside a source term ensemble it is possible to significantly reduce the number of model runs without any detriment to the results obtained, as long as a suitable number of Monte Carlo's simulations are employed.

A method to verify if a suitable number of simulations has been chosen is to analyse the ensemble results following each addition of a further simulation. At IRSN, the Monte Carlo method was used in the REM cases. One hundred simulations were carried out, scoping a combination of meteorological ensembles, source term ensembles and uncertainty attributed to atmospheric dispersion model input parameters. Several model endpoints (including Cs-137 and I-131 deposition concentration, inhalation thyroid dose and effective dose) were considered (Figure 103) in an effort to identify the optimum number of simulations whereby the model results sufficiently converge. It is clear from Figure 103 that (for this scenario, in combination with the chosen range of uncertainties) fewer simulations are required to derive a stable ensemble mean and median compared to the ensemble maximum and that the description of the uncertainty does not change significantly beyond approximately 70 simulations.



(a) (b)

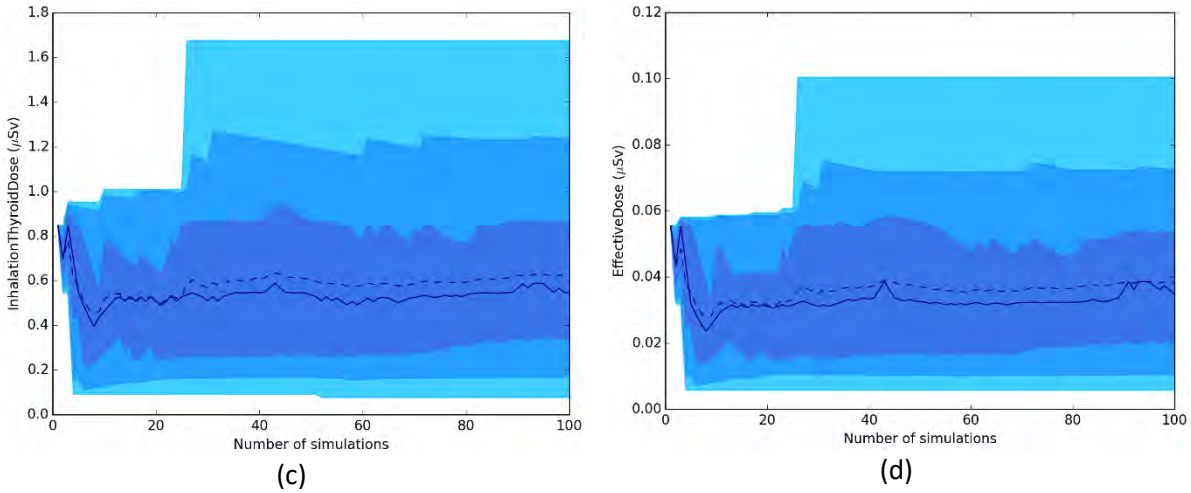


Figure 103. (a) Cs-137 deposition concentration, (b) I-131 deposition concentration, (c) inhalation thyroid dose and (d) effective dose (spatially) averaged over the model domain and integrated to 24 hours after the start of the release, for varying numbers of simulations using a Monte Carlo sampling method.

Attribution of probabilities to ensemble members

Probabilities can be considered as part of a sampling approach, but not in the attribution of probabilities to derived model endpoints. For example, meteorological ensembles are deemed to have an equal likelihood of occurrence. If such ensembles are clustered, the representative ensemble member of the cluster can be weighted accordingly (e.g. if a total of 10 ensemble members are clustered into Group A with 4 members and Group B with 6 members, the representative ensemble member of Group A can be attributed a weighting of 0.4 and the representative ensemble member of Group B can be attributed a weighting of 0.6). These weightings can then be used to appropriately "bias" the sampling method. However, these are relative rather than absolute probabilities i.e. the actual probability of occurrence (of the meteorology of an ensemble member, for example) is unknown and therefore the probability associated with the derived model endpoints cannot be inferred. Furthermore, the probability associated with the derived model endpoints cannot truly be inferred because all of the uncertainties will never be known.

Reducing model run time for a single ensemble member

As previously described, the computation time required to perform an array of atmospheric dispersion model runs based on a full (meteorological and/or source term) ensemble can be significant. In the investigative part of this project, model run time ranged from a few hours to a few days, exceeding the timescale that would be acceptable for provision of radiological protection advice in an emergency response. Model users are constantly searching for a balance between the quality of results and model run time. This search becomes even more important when considering an ensemble approach, where multiple model runs are performed. In the case of an ensemble of model runs, the balance between quality of results and model run time may not be the same as it is in the case of a deterministic run. Further investigation of this is now described.

Where practical, and where time allowed, hypothetical scenarios from (Korsakissok et al. 2019b) were re-run with proposed revisions to the model ensemble. The revisions were made by simplifying aspects of the model setup. The results of the reduced configuration were then compared with the previously performed full configuration in order to assess whether the reduced configuration still allowed a suitable representation of the uncertainty. Model run time for the reduced configuration was also noted, most notably to see whether it had decreased to approximately one hour or less.

Pictorial representations of the combined ensemble results for the full and reduced configurations have been compared qualitatively based on expert judgement (rather than any detailed statistical analysis). Consideration would need to be given as to whether such subjective acceptance criteria are feasible and appropriate in practice.

Alternative temporally and spatially resolved meteorological ensembles

Meteorological data files can be very large and therefore the handling and reading of meteorological data can be computationally expensive and time consuming. Furthermore, some atmospheric dispersion models require meteorological data to be pre-processed before it can be used, adding further time to the dispersion modelling process. Consideration of different resolutions of meteorological data should be explored, to determine if there are significant impacts in terms of the "accuracy" of the results and computational time.

It is thought that the effect of meteorological resolution on atmospheric dispersion simulations is likely to vary significantly depending on model type. Eulerian models (such as IRSN's IdX model) will be affected to a greater degree than Gaussian or Lagrangian models. Therefore, a comprehensive investigation should consider a range of model types.

In an assessment of the temporal and/or spatial resolution of an ensemble, it is suggested that a relatively fine meteorological ensemble is used as the baseline case and the resolution is reduced by averaging over grid cells. Were an Eulerian model to be used in such an assessment, mass consistency would need to be ensured.

Alternatively, a comparison considering two different meteorological datasets with differing resolutions could be performed. Périllat et al. (2017) considered two different meteorological ensembles for atmospheric disperison modelling of the Fukushima Daiichi Nuclear Power Plant accident. The first ensemble was derived by the Meteorological Research Institute (MRI) of the Japan Meteorological Agency, with a relatively fine horizontal resolution (3 km), vertical resolution (8 levels between 0-1000m) and temporal resolution (1 hour). The second ensemble used was developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), with a coarser horizontal resolution (0.25 degrees), vertical resolution (5 levels between 0-1000m) and temporal resolution (3 hours). Despite the difference in resolution of the meteorological datasets employed, the results (gamma dose rates and Cs-137 deposition concentrations) were similar (with both ensembles giving rise to over-dispersed consequences). However, it was recognised that the cruder resolution of the ECMWF meteorological ensemble was likely to result in poorer agreement with observations. The limitation of this comparative approach is that there are likely to be other variables at play, making it difficult to identify the contribution of resolution to any observed differences in results.

Cheaper dispersion modelling schemes & modified model domains

In this section of the report, simplifications to the dispersion modelling schemes (notably modifying the number of puffs or particles and the frequency of the puff/particle calculations) and adjustments to the model domain and output resolution are considered. The impact on the representation of the uncertainty and model run time is assessed. Two models have been considered. Firstly, a Gaussian Puff model implemented in the SINAC programme system developed by the Centre for Energy Research (MTA EK); and secondly, the UK Met Office's Lagrangian Particle model, NAME. The consideration of two different types of model provides an opportunity to investigate whether modifications that produce the greatest benefits vary.

A Gaussian Puff approach

The calculation process in SINAC is performed in three steps which follow each other consecutively. The first step (denoted "Initialisation") of the calculation process is to read the configuration file detailing the simulation parameters, read the list of the output points and characteristics, read the required meteorological data, and reserve memory space for the result variables. The second step

(denoted “Dispersion”) includes the dispersion calculation and storing of the temporal and spatial states of the puffs. The third step (denoted “Results”) includes the calculation of model endpoints and writing the results to output files. In an ensemble approach the calculation process is performed consecutively for each ensemble member (rather than in parallel).

The REM2 case was considered, assuming ten meteorological ensemble members and a single source term.

The primary reference case assumptions include:

- number of puffs released: 16 (with no subsequent splitting of the puffs)
- puff propagation time resolution: 4 minutes
- model domain: 767 x 765 km

The following assumptions were also made: 90,000 receptor points; a maximum simulation distance of 600 km; a maximum simulation time of 36 hours; an hourly output time resolution for estimates of environmental concentrations and a 24-hourly output time resolution for estimates of dose and dose rate; and a time limit for model output produced of 24 hours.

Computation times are presented for each of the three primary steps (initialisation, dispersion and results) and the total computation time, which includes the three aforementioned primary steps (x10). Model results are accompanied by an estimate of the standard deviation.

Number of model puffs

In the reference simulation, 16 puffs were released equally over a 4 hour release duration. Modifying only the number of released puffs to 4 and 64 affected the run time (detailed in Table 36). As expected, the total run time of the simulations with fewer puffs was lower, and with more puffs was higher than for the reference case. This was due to the time taken to write results to output files. Overall, reducing the number of released puffs from 16 to 4 decreased the run time from 76 to 47 minutes.

The mean maximum distance and the mean surface area (Table 37 and Table 38, respectively) of different threshold exceedances across three different model endpoints for three releases described by 4, 16 and 64 puffs, are approximately within one standard deviation of each other (except for mean surface area for exceeding an inhalation thyroid dose of 10 mSv and mean surface area for exceeding an effective dose of 1 mSv). The results are remarkably similar as a function of numbers of puffs released (except for Cs-137 deposition with a threshold exceedance of 10 kBq/m² assuming 4 puffs). The results for the different numbers of puffs are relatively similar because of the relatively small puff propagation time assumed (four minutes). It is noteworthy that for the majority of assessed quantities, the results computed assuming 4 puffs were higher than those calculated with 16 and 64 puffs. Furthermore, considering 64 puffs instead of 16 is not justified (because the results are almost identical). Further work could consider the impact of assuming numbers of puffs released between 4 and 16.

Table 36. Mean run time (per ensemble member) for each step of the calculation process and the total run time over all ensemble members and all steps of the calculation process, calculated assuming variable numbers of released puffs (reference simulation indicated in bold font)

Mean time of each step	Number of puffs		
	4	**16**	64
Initialisation [s]	44	**45**	41
Dispersion [s]	27	**26**	27
Results [s]	213	**386**	1004
Total run time for 10 ensemble members [s]	2840	**4570**	10720

Table 37. Maximum distance (in km from the source) of threshold exceedance for the mean of the ensemble results, for several model endpoints and for three simulation cases.

Mean Max Distance [km]	Cs-137 Deposition		Effective dose		Inhalation thyroid dose	
	10 kBq/m²	37 kBq/m²	1 mSv	5 mSv	10 mSv	50 mSv
4 puffs	164 ± 82	102 ± 9	37 ± 7	12 ± 2	82 ± 6	35 ± 6
16 puffs	**132** ± 12	**94** ± 10	**34** ± 6	**9** ± 4	**79** ± 5	**33** ± 6
64 puffs	131 ± 12	94 ± 10	33 ± 6	11 ± 2	78 ± 5	32 ± 6

Table 38. Surface area (in km²) of threshold exceedance for the mean of the ensemble results, for several model endpoints and for three simulation cases.

Mean Surface Area [km²]	Cs-137 Deposition		Effective dose		Inhalation thyroid dose	
	10 kBq/m²	37 kBq/m²	1 mSv	5 mSv	10 mSv	50 mSv
4 puffs	3634 ± 513	1679 ± 213	221 ± 30	24 ± 9	1069 ± 48	208 ± 28
16 puffs	**3630** ± 270	**1590** ± 238	**188** ± 27	**20** ± 11	**1009** ± 26	**180** ± 24
64 puffs	3642 ± 229	1574 ± 262	190 ± 20	23 ± 9	993 ± 39	182 ± 22

Cs-137 time integrated activity concentration in air and deposition concentration time series averaged over the whole spatial model domain are presented in Figure 104 and Figure 105, respectively. It is evident that the time integrated activity concentration in air time series were similar for the three simulations, and (for this scenario and uncertainty indicator at least) that assuming 4 puffs represents the spread and median sufficiently.

In the case of modelled deposition concentration, the differences are more visible. The deposition concentrations for cases with 16 and 64 puffs are less (including the median, 25th and 75th percentiles) than for the case with 4 puffs. However, the results for 16 and 64 puffs look very similar.

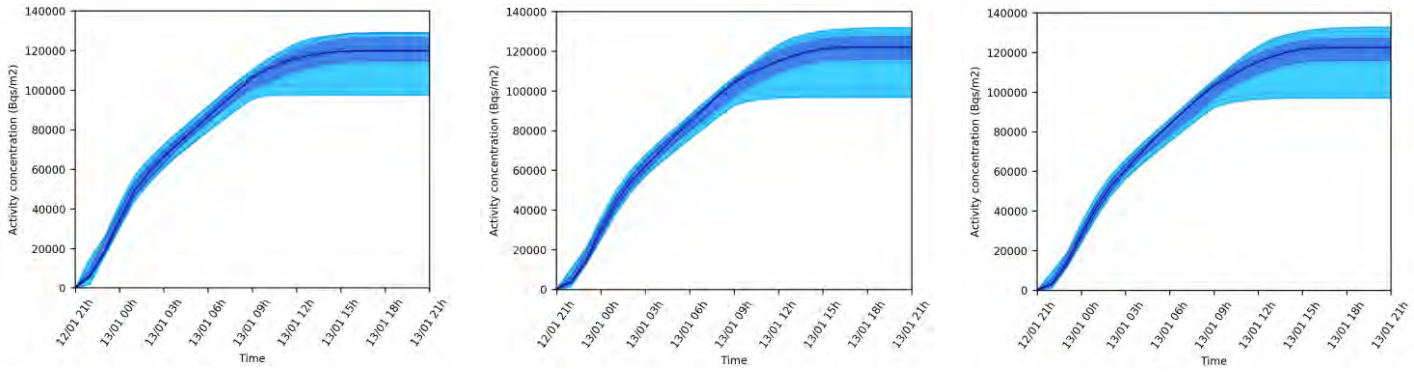


Figure 104. Time series of Cs-137 time integrated activity concentration in air averaged over the model domain for 10 ensemble members calculated with a release of (a) 4 puffs, (b) 16 puffs and (c) 64 puffs. The median is illustrated by a dark blue line, and the region between the 25th and 75th percentiles is illustrated by dark blue shading.

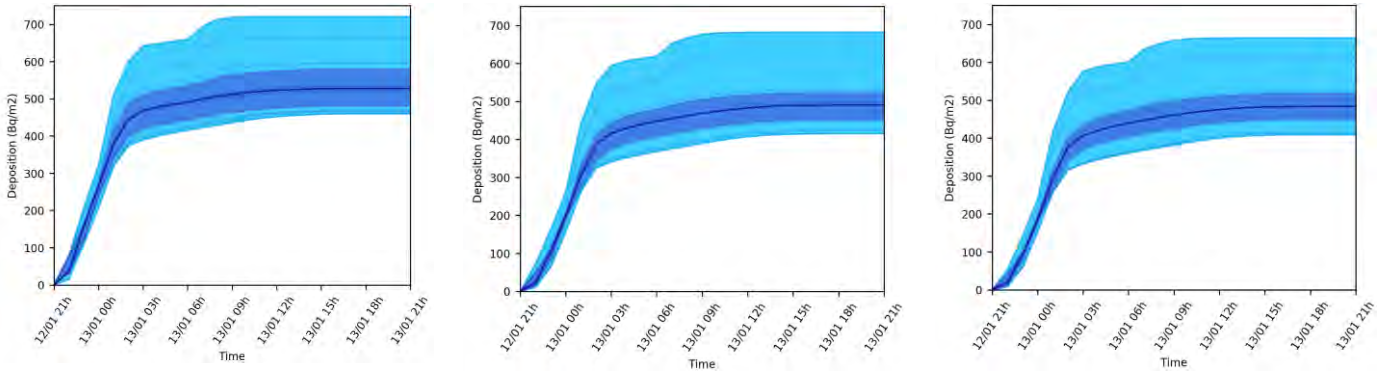


Figure 105. Time series of Cs-137 deposition concentration averaged over the model domain for 10 ensemble members calculated with a release of (a) 4 puffs, (b) 16 puffs and (c) 64 puffs. The median is illustrated by a dark blue line, and the region between the 25th and 75th percentiles is illustrated by dark blue shading.

Puff propagation time resolution

In the reference simulation, the time resolution of puff propagation was 4 minutes. Thus, the position, the properties and the deposition from each puff was calculated and stored with a temporal resolution of 4 minutes. Performing the dispersion with an 8 and 16 minute temporal resolution resulted in calculations of fewer puff states and thus reduced simulation times (Table 39); specifically, the time for calculating and writing the results was reduced. The effect of a lower puff propagation time resolution on the time taken to undertake initialisation and dispersion calculations was negligible. Reducing the puff propagation time resolution to one half and one quarter resulted in a 20-25 minute reduction in run time. If further investigative model runs were deemed necessary the "Richardson extrapolation" could be applied (using the tabulated results) to predict model endpoint values further into the tails of the distribution (e.g. for puff propagation time resolutions of less than 4 minutes and greater than 16 minutes).

The mean maximum distance of exceedance of 10 kBq/m² and 37 kBq/m² Cs-137 deposition concentration decreases with a coarser time resolution, whereas the distances for dose exceedance increase slightly with a coarser time resolution (Table 40). Almost all the dose exceedance distances for the different cases fall within the standard deviation of each other.

Table 39. Mean run time (per ensemble member) for each step of the calculation process and the total run time over all ensemble members and all steps of the calculation process, calculated assuming different puff propagation time resolutions (reference simulation indicated with bold font).

Mean time of each task	**Time resolution of puff propagation**		
	4 min	8 min	16 min
Initialisation [s]	**45**	46	44
Dispersion [s]	**26**	26	26
Results [s]	**386**	257	235
Total run time for 10 ensemble members [s]	**4570**	3290	3050

Table 40. Maximum distance (in km from the source) of threshold exceedance for the mean of the ensemble results, for several model endpoints and for three simulation cases.

Mean Max	**Cs-137 Deposition**		**Effective dose**		**Inhalation thyroid dose**	
Distance [km]	10 kBq/m²	37 kBq/m²	1 mSv	5 mSv	10 mSv	50 mSv
4 min	**132 ± 12**	**94 ± 10**	**34 ± 6**	**9 ± 4**	**79 ± 5**	**33 ± 6**
8 min	124 ± 11	80 ± 11	36 ± 6	12 ± 2	82 ± 6	35 ± 6
16 min	110 ± 8	62 ± 10	38 ± 6	15 ± 3	86 ± 8	39 ± 3

Table 41. Surface area (in km²) of threshold exceedance for the mean of the ensemble results, for several model endpoints and three simulation cases.

Mean Surface	**Cs-137 Deposition**		**Effective dose**		**Inhalation thyroid dose**	
Area [km²]	10 kBq/m²	37 kBq/m²	1 mSv	5 mSv	10 mSv	50 mSv
4 min	**3630 ± 270**	**1590 ± 238**	**188 ± 27**	**20 ± 11**	**1009 ± 26**	**180 ± 24**
8 min	2910 ± 282	1006 ± 238	194 ± 31	22 ± 6	1069 ± 31	188 ± 29
16 min	2110 ± 127	527 ± 141	181 ± 30	25 ± 9	1099 ± 67	181 ± 24

The mean surface area of exceedance for Cs-137 deposition concentration also decreases with a coarser time resolution (Table 41). The exceedance surface areas for dose were generally more uniform as a function of time resolution (Table 41); however, the results for 10 mSv inhalation thyroid dose show

a tendency to increase for a coarser time resolution. The 1 and 5 mSv effective dose and 50 mSv inhalation thyroid dose exceedance surface area results were within each other's standard deviation for the three time resolutions considered. Further work could consider the impact of assuming a puff propagation time resolution of 2 minutes (for example).

The Cs-137 time integrated activity concentration in air time series for a 4, 8 and 16 minute time resolution are detailed in Figure 106. Results for a 4 and 8 minute time resolution look similar, but they both differ with the results for a 16 minute time resolution.

The Cs-137 deposition time series (Figure 107) look very different for all considered temporal resolutions. The spread of ensemble results and the median both decrease with decreasing time resolution. The 25th and 75th percentile curves also vary significantly as a function of temporal resolution.

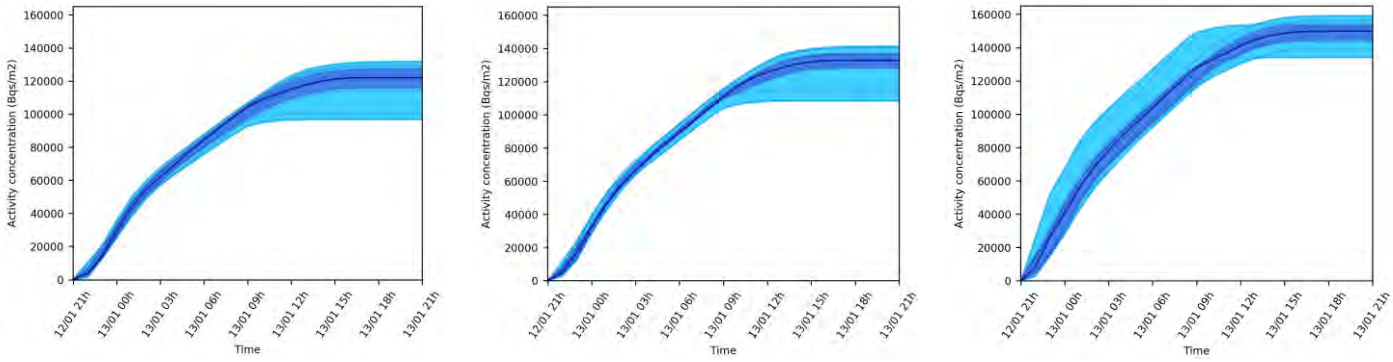


Figure 106. Time series of Cs-137 time integrated activity concentration in air averaged over the model domain for 10 ensemble members calculated with a puff propagation time resolution of (a) 4 minutes, (b) 8 minutes and (c) 16 minutes. The median is illustrated by a dark blue line, and the region between the 25th and 75th percentiles is illustrated by dark blue shading.

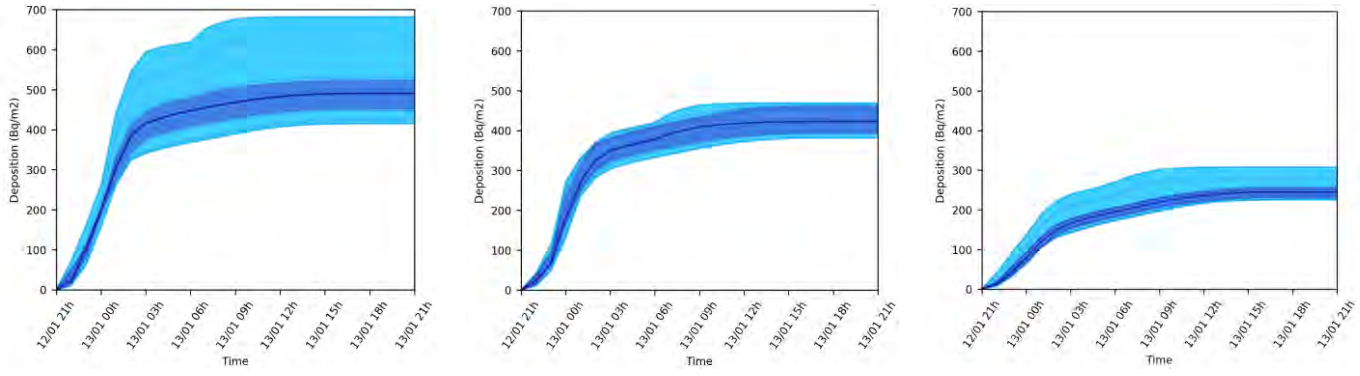


Figure 107. Time series of Cs-137 deposition concentration averaged over the model domain for 10 ensemble members calculated with a puff propagation time resolution of (a) 4 minutes, (b) 8 minutes and (c) 16 minutes. The median is illustrated by a dark blue line, and the region between the 25th and 75th percentiles is illustrated by dark blue shading.

Model domain

In the reference simulation, a spatial domain of 767 km by 765 km (alongside a 300 x 300 grid of receptor points) was assumed. Two further simulations were performed, modifying the domain to 510 km by 510 km (with 200 x 200 receptor points), and to 254 km by 254 km (with 100 x 100 receptor points) as depicted in Figure 108.

The simulation time was reduced significantly when considering smaller model domains (Table 42). By reducing the domain to 2/3 and 1/3 of its original size, the run time was reduced to approximately 37% and 12% of the original value, respectively. The major reduction was achieved in the calculation of results, because there were significantly fewer receptor points in the smaller domains (and perhaps also less of an I/O bottleneck). The time of initialisation was also reduced, because fewer receptor points had to be read from the configuration file.

It is important that the model domain is not reduced in size to such an extent that the model endpoints of interest are no longer fully described (which may be difficult to determine a priori). In this investigation, the mean maximum distances and surface areas of threshold exceedance were within the bounds of the smaller domains for all the ensembles and therefore the estimated values were identical irrespective of the considered modifications to the spatial area. For the purposes of brevity, these results are not detailed here.

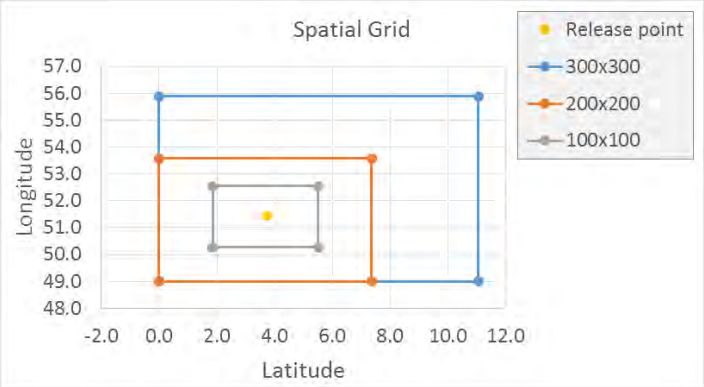


Figure 108. The model spatial domain and release location.

Table 42. Mean run time and total run time for 10 ensemble members calculated with different spatial domains and numbers of output receptor points (reference simulation indicated with bold font).

Mean time of each task	Spatial domain [km^2]		
	767 x 765	510 x 510	254 x 254
Initialisation [s]	**45**	9	1
Dispersion [s]	**26**	27	27
Results [s]	**386**	133	28
Total run time for 10 ensemble members [s]	**4570**	1690	560

Model output resolution

Modifying the temporal resolution of the time integrated activity concentration in air and deposition concentration output from 1 hour to 2, 4 and 12 hours achieved a reduction in total run time of 22%, 26% and 36%, respectively (Table 43). The largest contribution to simulation time and the largest reduction in simulation time was for the calculation of results and writing output files.

Varying the temporal resolution did not affect the estimated mean maximum distance and surface area values of threshold exceedance (of Cs-137 deposition concentration, effective dose and inhalation thyroid dose) and therefore for the purposes of brevity, these results are not detailed here. Time series of Cs-137 deposition concentration (averaged over the model domain) were plotted and appear in Figure 109. The different temporal resolutions of the output are reflected in the different degrees of smoothness of the curves, but otherwise the three plots are very similar.

Table 43. Mean run time and total run time for 10 ensemble members calculated whilst varying the temporal output resolution (reference simulation indicated with bold font)

Mean time of each task	Time resolution of output			
	1 hour	2 hours	4 hours	12 hours
Initialisation [s]	**45**	42	45	41
Dispersion [s]	**26**	26	26	27
Results [s]	**386**	289	269	224
Total run time for 10 ensemble members [s]	**4570**	3570	3400	2920

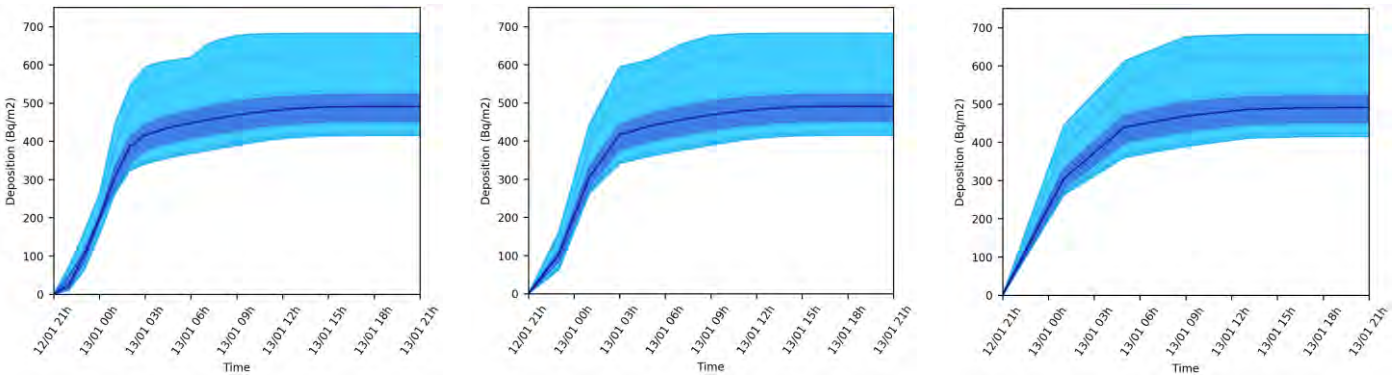


Figure 109. Time series of the Cs-137 deposition concentration averaged over the model domain for 10 ensemble members calculated with output time resolution of (a) 1 hour, (b) 2 hours and (c) 4 hours. The median is illustrated by a dark blue line, and the region between the 25th and 75th percentiles is illustrated by dark blue shading.

A Lagrangian Particle approach

NAME is the UK Met Office's atmospheric dispersion model and was applied here for the REM2 case, initially assuming ninety ensemble members in total (ten meteorological ensemble members and nine source term ensemble members, comprising of three release start times and three release magnitudes).

The primary reference case assumptions include:

- number of particles released: 1,000,000 (50% to model iodine; 50% to model all other radionuclides)
- particle propagation time resolution (sync time): 1 minute
- output temporal frequency: 1 hour (for estimates of environmental concentration)
- model domain: 0 to 11.063 degrees longitude and 49 to 55.877 degrees latitude (~760 km x 760 km)

Note also that 90,000 output locations, a four hour release duration, a maximum simulation time of 48 hours, and a time limit for model output produced of 24 hours, were all assumed. All NAME model runs were performed on 8 cores and running 10 members in parallel.

On the basis of the assumptions described above (the reference case), the NAME model took approximately 11 hours to complete the 90 model runs (where each model run represented a single ensemble member). The time spent in individual sections of the model code are detailed in Figure 110. It is evident that the majority of the computational time was spent processing the model particle loops (denoted in light blue in Figure 110).

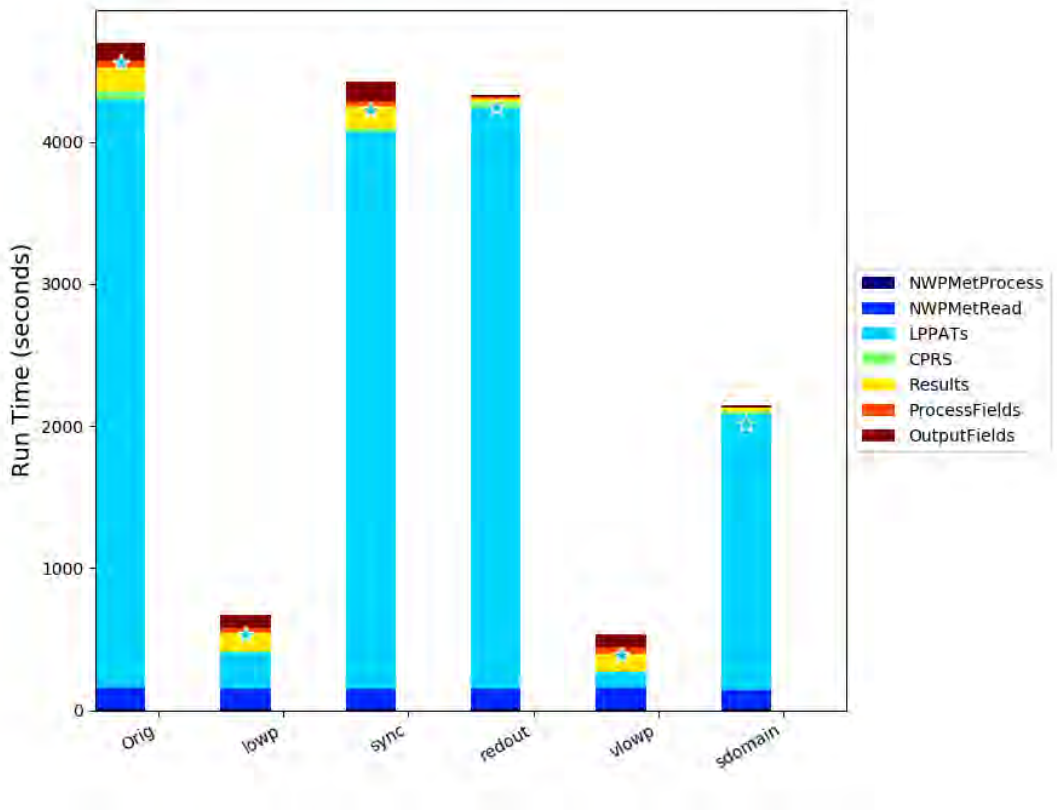


Figure 110. Reference case NAME model run times per ensemble member. On the x axis Orig = the reference case, lowp = a reduced number of model particles, sync = a reduced particle propagation time resolution, redout = reduction in output, vlowp = very few particles, and sdomain = smaller model domain. In the key NWPMetProcess = processing NWP met data, NWPMetRead = reading NWP met data, LPPATs = the main loop evolving particles over time steps, CPRS = the particle update loop that calculates Lagrangian fields, Results = the processing and output of results, ProcessFields = the processing of fields, and OutputFields = the outputting of fields.

Model run setup simplifications and results analysis

In an attempt to reduce the model run time of the reference case (denoted as "Orig" in Figure 110), three separate modifications were made to the model setup. Firstly, the number of model particles were decreased by a factor of 25 to 40,000 ("lowp" in Figure 110). This reduced the run time per ensemble member to approximately 8 minutes. Extrapolating to the full ensemble of 90 members (noting that 10 ensemble members were being run in parallel) implies a full computation time of just over one hour. A further reduction in the number of model particles released was considered (only 1000 in total), termed "vlowp" in Figure 110.

The maximum distance at which 10 kBq m^{-2} of Cs-137 deposition concentration was exceeded and the total area over which 10 kBq m^{-2} of Cs-137 deposition concentration was exceeded were estimated and plotted in Figure 111 and Figure 112, respectively. It is evident from both figures that the reduction of the number of particles from 1,000,000 to 40,000 does not affect the observed spread nor the median surface area value, but it does affect the median maximum distance value. The box plots in Figure 111 and Figure 112 adhere to a standard format (https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51).

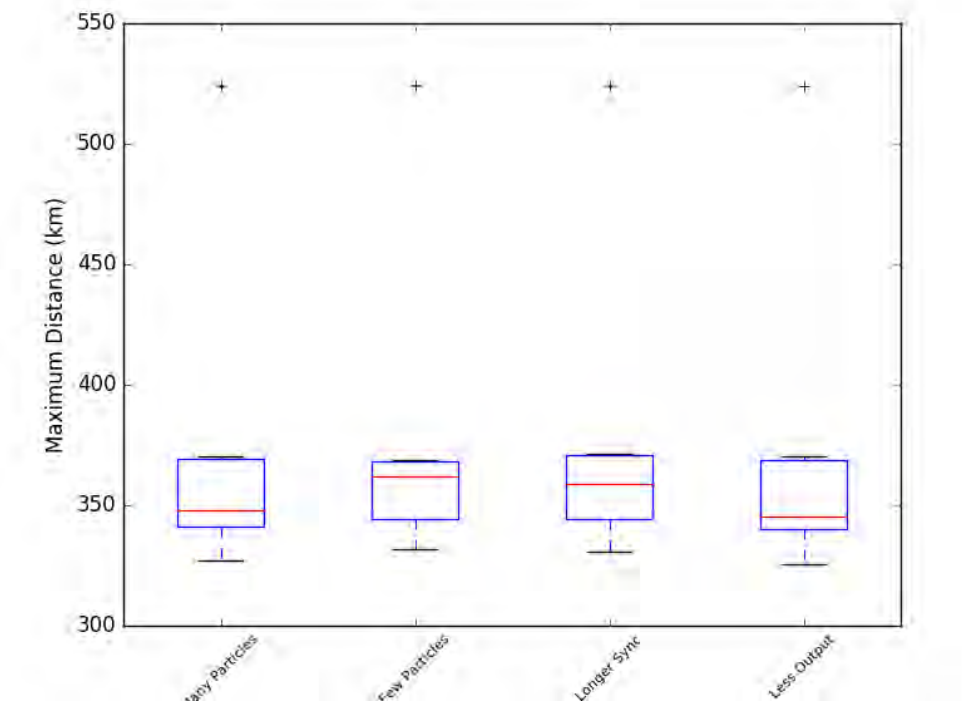


Figure 111. The maximum distance at which 10 kBq m^{-2} of Cs-137 deposition concentration was exceeded (where the red line represents the median value, the blue box represents the 25th to 75th percentiles, the T-bars represents a value of quartile 1 minus 1.5 x interquartile range and quartile 3 plus 1.5 x interquartile range and the '+' represents outliers beyond the range described by the T-bars). Many particles = 1,000,000, few particles = 40,000, longer sync time = 2 minutes, less output = 2 hour frequency.

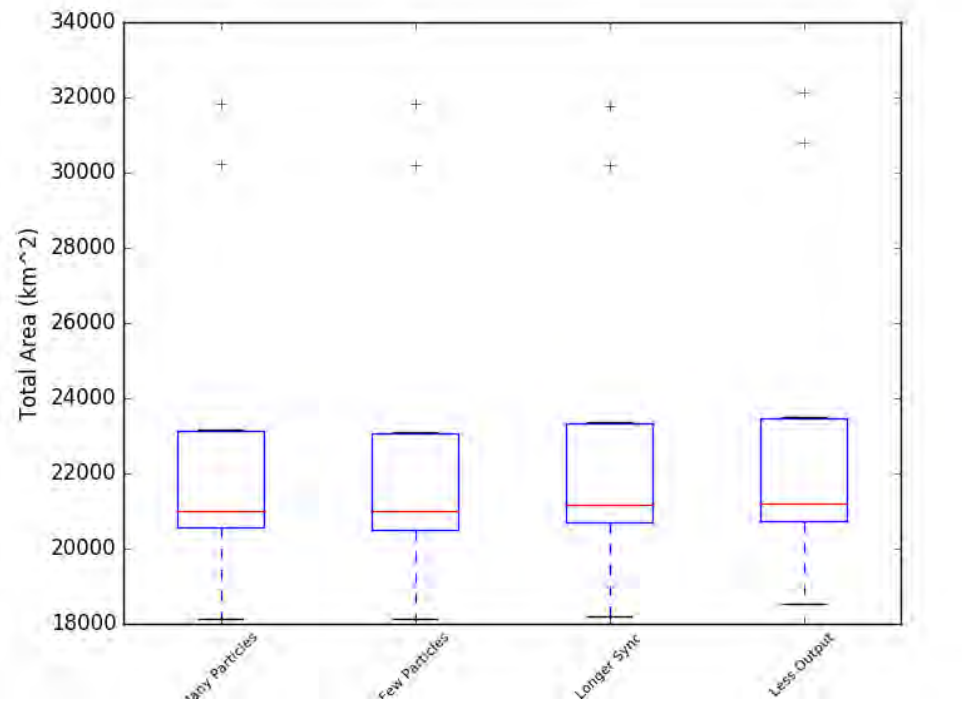


Figure 112. The surface area at which 10 kBq m^{-2} of Cs-137 deposition concentration was exceeded (where the red line represents the median value, the blue box represents the 25th to 75th percentiles, the T-bars represents a value of quartile 1 minus 1.5 x interquartile range and quartile 3 plus 1.5 x interquartile range and the '+' represents outliers beyond the range described by the T-bars). Many particles = 1,000,000, few particles = 40,000, longer sync time = 2 minutes, less output = 2 hour frequency.

Three other methods were also employed in an effort to reduce run time. Firstly, particle propagation time resolution (or sync time - the internal Lagrangian timestep) was increased from 1 minute to 2 minutes (termed "sync" in Figure 110). Secondly, the amount of output was reduced by outputting fields every 2 hours instead of every hour (termed "redout" in Figure 110). And thirdly, the model domain was reduced from 760 km x 760 km to 200 km x 200 km (termed "sdomain" in Figure 110). These modifications resulted in speed increases of 6%, 4% and 55% respectively, again with no significant impact on estimated Cs-137 deposition concentration exceedances. The effect of all three simplifications to the model run setup have not been considered collectively, but it is thought that if all three simplifications were applied together, the model run time would be reduced to less than one hour.

Nested grids (where the spatial grid resolution decreases with increasing distance from the release) were not considered here but such an approach has the potential to significantly reduce run time by capping the required number of model particles, especially when modelling two (or more) model endpoints over two (or more) different spatial scales (e.g. evacuation and food restriction areas).

Thus far, only results reflecting integrated model endpoints have been presented. However, the results for hourly environmental concentrations were similar, with no significant impact on estimated ensemble median values and ensemble spread as a result of the described simplifications in model run setup, most notably when applying significantly fewer model particles (although, as with any Lagrangian Particle approach, continual reduction in the number of model particles eventually results in statistically noisy model results).

Cheaper dispersion modelling schemes and modified model domains: findings

It was evident that for a Gaussian Puff model, the computationally expensive element of the calculation process was the calculation of model endpoints and writing of results, whereas for a Lagrangian Particle model, the computationally expensive element of the calculation was processing the model particle loops within the atmospheric dispersion modelling process. Despite this, the major efficiency savings were the same irrespective of the model being applied, notably modifying the number of model puffs or particles, and reducing the size of the model domain. The challenge with respect to both of these forms of efficiency savings, is identifying the optimum number of model puffs or particles and the optimum size of the domain a priori (before the tipping point is reached, where any further reductions will have a detrimental impact on ensemble results). A single upper bound value determined (derived on the basis of many model runs across many different model scenarios) could be applied but will be limited because of the scenario dependent nature of these variables. A possible solution is to base the model domain on a scoping run using a simple Gaussian plume model, estimating the likely distance of interest, based on conservative meteorological conditions (notably wind speed and direction) and the most conservative magnitude of the release, inferred from the meteorological and source term ensemble, respectively. An alternative approach would be to begin by focusing on the key model endpoints, which are likely to be those used in the provision of protective action advice (rather than food restriction advice). In the UK at least, it is thought that a model domain up to 50 km from the release location will almost always be sufficient in such cases. It is worth noting that whilst there will be a significant reduction in the computation cost as a result of reducing the model domain, this will be tempered somewhat if the meteorology being applied is still representative of the full domain.

Whilst reducing the temporal resolution of model output was observed to result in only moderate computational savings, such improvements may still be of value given that there is likely to be significant urgency in the provision of radiation protective advice in an emergency response. It is unclear whether the optimum approach would be to perform an initial run with a low temporal resolution (e.g. 24 or 48 hours) for estimating primary endpoints, and then a subsequent run with a higher temporal resolution (e.g. hourly) for supplementary information, or to produce all such model output first time by way of a single set of model simulations.

There are also efficiencies which can be made in pre- and post-processing, for example in respect of meteorological data and model output.

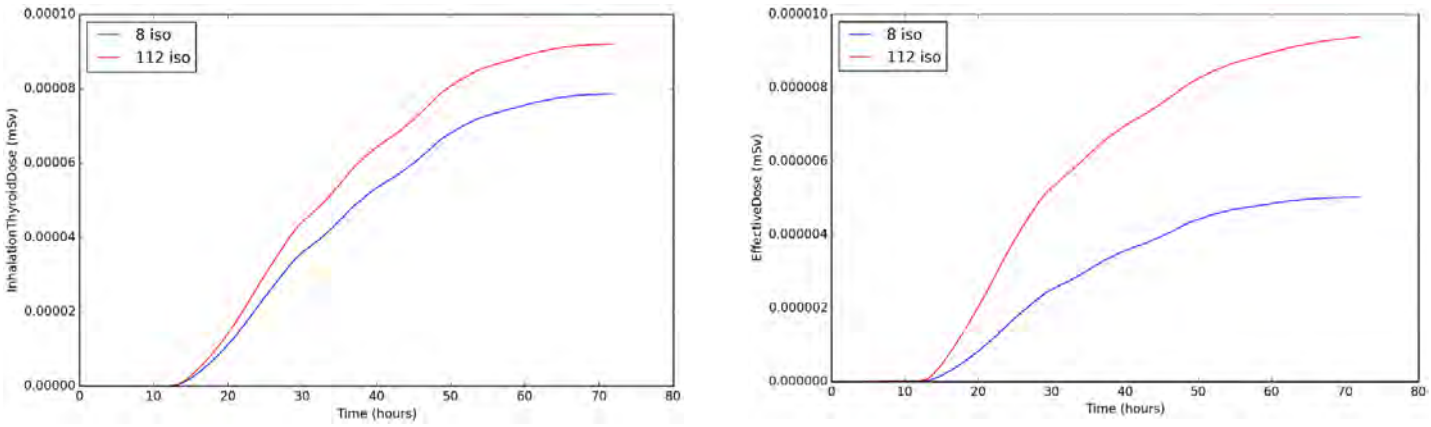
Reducing the number of radionuclides modelled

Two possible approaches are highlighted here, the first based on a unit release of all emitted radionuclides (Sigg and Grahn 2019) and a subsequent post-processing step for non-dispersion related properties (such as the magnitude of release and radioactive decay). A possible variation on this first approach would be to use a unit release of a subset of pseudo radionuclides, where each is representative in terms of chemical and physical characteristics of groups of radionuclides. The second approach is based on modelling radionuclides explicitly but reducing the number considered to only those expected to make a significant contribution to the model endpoints derived.

Reducing the number of modelled radionuclides to only those which are significant

Depending on the atmospheric dispersion model, reducing the number of radionuclides considered can drastically reduce computation time. An ensemble of source terms was derived from the European Project FASTNET (Chevalier-Jabet 2019a; Chevalier-Jabet 2019b), and considered in one of the scenarios identified in this project (Mathieu et al. 2018), the REM case with “long release” (Korsakissok et al. 2019b). The whole source terms contain 112 radionuclides. An ensemble simulation on the basis of the full FASTNET source term can take between two and three days to complete.

There are a number of different possible approaches for reducing the number of radionuclides. One such approach is based on the descriptions of source terms of previous nuclear accidents in the literature. In the case of the Fukushima Daiichi NPP accident, source term reconstructions were performed to determine the source term properties of a few radionuclides, mostly ^{137}Cs, and ^{133}Xe in a few cases, and thereafter isotopic ratios were assumed to determine other radionuclides such as ^{131}I and ^{132}I/^{132}Te. In total, a list of eight radionuclides (Cs-137, Cs-136, Cs-134, I-131, I-132, Xe-133, Ba-137m and Te-132) were identified (Saunier et al. 2013). This list of radionuclides and the full FASTNET source term were subsequently considered in the REM case study (Figure 113).



(a) (b)

Figure 113. Time series of (a) inhalation thyroid dose and (b) effective dose spatially averaged over the model domain for simulations with 112 and 8 radionuclides

It is apparent (from Figure 113) that there is a significant difference in the average dose estimated for the full set and for a subset of radionuclides, especially in respect of effective dose (where there was a difference of nearly 50%). Note that this comparison was performed for only one ensemble member (and not the entire ensemble of source terms) because of the very large computation time to perform a full simulation. The differences observed in Figure 113 are a result of the different source term compositions (i.e. the different radionuclide ratios) between the Fukushima and FASTNET source terms.

An alternative method to select radionuclides for consideration is to perform a simple dose assessment using a Gaussian plume approach to describe the atmospheric dispersion. Such an approach could be based on a single receptor point on the plume centreline (assuming constant meteorological conditions). A single dose integration period could be assumed and the necessary exposure pathways would need to be considered. Dose as a function of radionuclide would be derived and ranked by contribution. Any radionuclides contributing less than a specified percentage could be removed. Alternatively, only those which contribute (in order of ranking) to a pre-defined percentage (e.g. 90%) of the dose remain for further assessments. The benefit of the latter approach is the ability to scale the dose after the full dispersion calculation (e.g. by 1/0.9). Such scaling is substantiated by Figure 113 where the difference between the dose based on the full and reduced source terms can be approximated by a constant factor.

Computational resource

Additional computational resource, as employed in The Nordic Nuclear Safety Research (NKS) projects (by way of a super computer) will speed up the processing of a large number of ensemble model runs and reduce computational time, but at a financial cost. However, consideration of the utilisation of current computational resource should also be made. For instance, computational time can be reduced by launching several simulations in parallel on different cores of a computer or on different computers (which may not be made available for the purposes of research, but may be made available operationally in the event of an emergency response). However, reductions in computation time can sometimes be somewhat offset by an increase in the individual time of a simulation as a result of memory RAM issues. For example at IRSN, atmospheric dispersion simulations were performed using a long range model called *IdX*, and the subsequent dose assessments were performed using a model called *consx.* A single simulation in *IdX* and *consx* took approximately 1 hour 30 minutes and 30 minutes, respectively. When 15 simulations were launched in parallel on several cores of a 24-core-

computer, the cumulative time taken for *ldX* and *consx* simulations was 3 hours and 4 hours, respectively.

Operational ensemble results & output

Uncertainty indicators

The uncertainty indicator (or uncertainty descriptor) is the method used to analyse and present the derived uncertainty. For numerical and tabulated values, this may be in the form of the area and/or number of people affected and the maximum distance above a given threshold dose (or environmental concentration). However, it was found for the REM2 case that considering such values on their own can be limited and worse still, misleading. For example, the magnitudes of the areas, numbers of people and maximum distances can be similar but the spatial coverage may be very different. Furthermore, it was observed in this study that the application of the mean or a percentile (other than the 100th) of the maximum distance affected above a given threshold dose was prone to outliers and misleading results (notably the maximum distance can decrease when more uncertainty is considered). However, it is thought that the maximum (100th percentile) of the maximum distances is not likely to suffer from such idiosyncrasies. For percentiles to have a statistical meaning, there must be a sufficient number of simulations. Preferably, more than a hundred simulations should be included, such that one percent is represented by at least one ensemble member. If fewer members are considered, then it may be more desirable to use minimum and maximum values as an indicative range of variation. However, care should be taken in the use of the terms “minimum” and “maximum”, which may be misleading, implying that all uncertainties have been taken into account and that there is no possibility for observations to be outside those bounds.

Alternatively, evaluations of uncertainty could consider using the figure of merit in space, FMS, (i.e., area-based threat score) evaluated against the ensemble control. The FMS is defined as the ratio of the intersection of the areas, and the union of the areas determined as a result of the consideration of a range of ensemble members. This concept could be extended to the figure of merit of population which may be of value for radiation protection advice.

$$FMS = 100\frac{A_1 \cap A_2}{A_1 \cup A_2}$$

In general, a high value of the FMS corresponds to confidence in model results and relatively little uncertainty. Whilst a low FMS implies relatively poor confidence in model results and relatively large uncertainty, the extent may vary for the same FMS value. For example, two sets of ensemble results of threshold exceedance may have the same FMS, but the former may be very similar in shape and shifted in space, and the latter may differ significantly in shape but shifted relatively little in space. Therefore, where only numerical / tabulated values are derived the production of FMS and one or more of the area and/or number of people affected and the maximum distance (above a given threshold dose) would be beneficial.

However, the preferred option is to evaluate numerical / tabulated values together with a graphical representation of uncertainty. Examples include level of agreement plots (e.g. Figure 94 which depicts exceedance of a threshold environmental concentration for assessing spatial variability) and percentile plots (e.g. Figure 89 which depicts environmental concentration averaged over the model domain for assessing temporal variability).

Presentation of ensemble output

As discussed previously, model ensemble output should be presented in terms of likelihood (or agreement) maps and not probability maps. It was also considered that contours on maps are more readily misinterpreted and that the use of “heat maps” is preferred.

Deterministic model runs are likely to continue to be applied beyond the time that ensembles are introduced into an operational framework. For the ECMWF meteorological ensemble, the first ensemble member has the same initial conditions as the deterministic run and differs only as a result of a coarser resolution. However, the variable resolution may result in challenges in terms of integrating, interpreting and presenting the two sets of model output (for example higher resolution forecasts may include additional physical phenomena e.g. showers (whether or not accurately represented), that lower resolution forecasts cannot represent). Such challenges are likely to amplify if the deterministic model is not one of the members of the ensemble (i.e. no ensemble member has the same initial conditions as the deterministic model run).

Recommendations

Note that recommendations made here are in accordance with the scope of the work described previously. Further details of any recommendations can be found in the relevant sections of the report above.

Recommendation 1: First and foremost, it is recommended to take into account uncertainties in atmospheric dispersion simulations for emergency response. To achieve this goal, it is recommended that an ensemble approach be used operationally to describe uncertainty in the provision of advice in the event of a radiological accident. This ensemble should take into account the main uncertainties, at least those linked to the meteorological and source term data.

It was apparent from the exercise undertaken to propagate uncertainties through the chain of atmospheric dispersion and radiological assessment models that the time taken to complete ensemble runs was not appropriate for an emergency response. As a result, to ensure an operational method, two approaches were investigated in this report: (1) reducing the model run time for a single ensemble member and (2) reducing the number of ensemble members.

Of these two possibilities, the first one is to be privileged. It is thought, as was illustrated in this report, that a crude model configuration may be sufficient to obtain a good description of the main uncertainties, since meteorological data and source term uncertainties are of first order compared to those stemming from the models. This method may suffice, especially if relatively few ensemble members are considered in the full ensemble configuration. However, more members may be needed to correctly represent the main uncertainties. Therefore, methods to optimize the number of members used in ensemble calculations are also proposed below.

Reducing the run time for a single ensemble member

Recommendation 2: It is recommended that efficiency savings in model setup are identified and implemented in an effort to reduce model run time for a single ensemble member (without significant detriment to the model endpoints derived, the uncertainty estimated and the radiation protection advice inferred). Some approaches for that purpose are suggested below.

Efficiency savings highly depend on the model used, but could focus on the following issues: (a) reducing the number of model puffs or particles, or, in the case of an Eulerian model, using a coarse resolution, and (b) reducing the size of the model domain. It is recommended that the size of the model

domain is based on a scoping run using a simple Gaussian plume model or alternatively begin by focusing on the key model endpoints (notably sheltering, evacuation and the administration of stable iodine protective actions) and assuming a bounding value (e.g. 50 km from the release location).

Reducing the radiological specific computation is also of value. Two approaches are recommended. Firstly, applying a unit release of a subset of pseudo radionuclides, where each is representative in terms of chemical and physical characteristics of groups of radionuclides. The second approach is based on modelling radionuclides explicitly on the basis of a computationally inexpensive approach (for example a simple Gaussian plume model) and then reducing the number considered to only those expected to make a significant contribution to the model endpoints derived.

Reducing the number of ensemble members used in the simulations

If a large number of ensemble members are considered there may be benefits of applying a method to reduce the number of ensemble members (without detriment to the uncertainty estimated).

Recommendation 3: It is recommended to optimize the number of ensemble members so that the ensemble is still representative of the full uncertainty. Two approaches may be explored to achieve this goal (and may be used together):

1. clustering the (meteorological and/or source term) ensemble members, to select members that may be representative of a subset of the ensemble;
2. using sampling methods, to avoid running dispersion calculations for all different possibilities (especially when several sources of uncertainties are taken into account).

A manual clustering approach of meteorological ensemble members combined with the uncertainty in the timing of the release could be applied. This approach requires the production of meteograms, which can be used to identify which meteorological parameters (including wind direction) are uncertain and/or variable over the (potentially uncertain) timing of the release, and whether a single ensemble member can represent a subset or the full set of meteorological ensemble members over the period of the release. In addition to the use of meteograms, it is recommended to focus on efforts to cluster on the basis of model output(s) from sets of atmospheric dispersion model runs performed in a very simplified mode, for example using particle trajectories, only using single site meteorological conditions, using a different more simplified model altogether and/or significant simplifications such as removing the assessment of deposition for scenarios where the exists large confidence in no precipitation (where the latter will only be suitable for advice in respect of certain protective actions).

Recommendation 4: It is recommended to work on advanced clustering methods with meteorological offices responsible for providing the data in case of emergency, so that the particular needs for this application are well understood by data providers; if possible, tailor-made ensembles should be specifically designed. If possible, clustering methods should consider the dispersion model output variables and not only meteorological variables.

Recommendation 5: In an effort to reduce the number of simulations performed accounting for all atmospheric dispersion model uncertainties (notably a source term ensemble, a meteorological ensemble and possibly atmospheric dispersion model input parameter uncertainties) a sampling approach is recommended to suitably reduce computational time. Latin Hypercube Sampling (LHS) and Monte Carlo approaches are potential methods. An approach to verify if a suitable number of simulations has been chosen is to analyse the ensemble results following each addition of a further simulation until model results sufficiently converge. Probabilities can be considered as part of a sampling approach (but not in the attribution of probabilities to derived model endpoints).

Additional technical recommendations to save computational time

Recommendation 6: Whilst making efficiencies savings in pre- and post-processing, in respect of meteorological data and model output, is likely to result in only moderate computational savings, such improvements may still be of value given that there is likely to be significant urgency in the provision of radiation protective advice in an emergency response. For instance, meteorological pre-processing may be run automatically as soon as meteorological data are available, to be ready should the need for dispersion simulations arise.

Recommendation 7: Additional computational resource will speed up the processing of a large number of ensemble model runs and reduce computational time, but at a financial cost. However, consideration of the utilisation of current computational resource should also be made, by way of launching several simulations in parallel on different cores of a computer or on different computers.

Uncertainty representation for operational purposes

Recommendation 8: It is recommended that a range of uncertainty indicators (the method used to analyse and present the derived uncertainty) are considered. Such uncertainty indicators should comprise of a combination of numerical (tabulated) values and graphical representations. Uncertainty indicators in the form of numerical and tabulated values may include surface area and number of people affected and maximum distance above a given threshold dose (or environmental concentration) and figure of merit in space (FMS) and figure of merit of population (whilst analysis on the basis of FMS was not referenced in this report, this method was used widely within the project by some WP1 participants). Uncertainty indicators in graphical form may include level of agreement plots and/or percentile plots. It was found in this study that a spread of uncertainty indicators and presentation methods are vital to fully understand model endpoint uncertainties which are spatially and temporally complex.

Future work

During the course of this study, a number of areas for potential further investigation were identified. These have been categorised below. In most cases, more detail of the suggested work can be found in the relevant sections of the report above.

Reducing model run time for a single ensemble member

From this study it was found that modifications to the number of model puffs or particles, and reducing the size of the model domain resulted in major efficiency savings in model run time. A future step is to determine a methodology for simplifying ensemble model runs. Such a methodology should also focus on the reduction in the time taken to process and output model results, which can be minimized by focusing on the primary model outputs required, and should focus on opportunities to reduce the time taken to read meteorological data. Different spatial and temporal resolutions of meteorological data could be considered in an investigation of the effect on the accuracy of the results and computational time. Such an investigation may wish to consider a range of model types (Gaussian, Lagrangian, Eulerian), as the effect of meteorological resolution on atmospheric dispersion simulations may depend on model type.

The use of an emulator may be considered to save considerable computational time; an emulator is a mathematical approximation of the physical model that and runs in a matter of seconds, or less. Emulators use interpolation techniques such as kriging, to reproduce scalar results of a more advanced physical model. Emulators have already been used in the field of atmospheric dispersion modelling for volcanic ash (Harvey et al. 2018) and for radiological applications, in the study of the Fukushima accident (Girard et al. 2015). For such a tool to be used operationally, it would be necessary to

construct emulators for a wide range of pre-calculated source terms for several accident scenarios and meteorological situations. In the case of short-range consequences when meteorological conditions may be considered homogeneous, it would be feasible to design an emulator that would instantly give an operational endpoint (for instance, maximum distance of threshold exceedance) for a wide range of values for wind direction, wind speed, rain intensity and boundary layer height, among other parameters; one emulator per scalar endpoint and per accident scenario would be needed. It has yet to be investigated whether emulation using a database of 3-D meteorological fields would be feasible and interesting for operational purposes.

Reducing the number of members in an ensemble approach by an appropriate selection method

If the number of members have to be reduced, any future work should perhaps focus on efforts to cluster on the basis of model output(s) from sets of atmospheric dispersion model runs performed in a very simplified mode. This report deals mainly with clustering of ensemble members based on uncertainty associated with wind direction, which is clearly a requirement of any clustering criteria, but it would also be beneficial to investigate the inclusion of other meteorological variables (such as wind speed, boundary layer height, atmospheric stability and precipitation) as part of any clustering criteria.

Further work assessing the ability of different methods to suitably sample ensemble members from a full ensemble configuration would be of value. For example, methods relating to adaptive sampling and Latin Hypercube Methods may be explored. Methods proposed in the AVESOME project, using a unit source term in the dispersion calculations and combining the results during the post-processing phase to represent the real release, may also be investigated (Sigg and Grahn 2019; Sørensen et al. 2019). And investigation whether the work of Klonner (2013) and Galmarini et al (2004a, 2004b and 2004c) could be applied to the estimation of uncertainty within the provision of radiological protection advice following an accidental release to atmosphere would be worthy.

Whilst a preliminary evaluation of a priori clustering on the basis of statistical analysis (notably an assessment of the sensitivity of a model endpoint of choice to a change in one or more input variables) has been undertaken, further work would be required to determine if such an approach is practical and of merit. If statistical methods are to be used, it is recommended that self-associative methods (which determine non-linear uncorrelated variables of the source terms) are investigated to reduce the number of source term ensemble members for emergency response. However, for a meteorological ensemble, this is not practical, because of the large number of variables involved.

Use of dispersion field studies to assess the quality of ensemble simulations

To validate the estimation of uncertainty for operational purposes, it is necessary to evaluate uncertainty estimation by comparison to environmental observations. This was partly done in the CONFIDENCE project by using the Fukushima environmental data. However, there is a need to go further in this area, in particular by using measurements at a local scale, which is of prime interest for emergency response.

For model-to-data comparison, several existing case studies may be used, depending on the objective and scale of interest. Meteorological ensembles of sufficient resolution should be designed for the case studies to be relevant:

- the Cabauw experiment (Piters et al. 2012), for a short-distance range, flat terrain, for which a HARMONIE-AROME ensemble may be constructed;

- the ^{85}Kr measurements around La Hague reprocessing plant, for a range between 5 and 30 kilometres, with complex terrain (Connan et al. 2011; Connan et al. 2013; Connan et al. 2014; Korsakissok et al. 2016); to be fully relevant for the purposes of ensemble modelling, the setting up of a larger number of sensors and a new measurement campaign may be required;
- at longer distances, the ETEX experiment (Van dop et al. 1998) might be considered.

Other relevant case studies could be considered, such as the NERIS ADM experiment (Camps et al. 2018) at short range and the CTBTO inter-comparison at long distances (Maurer et al. 2018). The use of meteorological ensembles for the latter case has already been explored by (De Meutter et al. 2016).

In the above case studies, the source terms were perfectly known, which allows focusing on meteorological and model uncertainties. If source term uncertainties were to be included, cases such as the ^{106}Ru release (Masson et al. 2019; Saunier et al. 2019) and the accident at the Fukushima Daiichi NPP could be considered. The work carried out in this project relating to the Fukushima accident scenario only considered a 3 day period, but future work may include representing the uncertainties of the whole accidental sequence, which would allow the use of deposition measurements that were made at the end of the releases.

Improvement of input uncertainties, combining a priori knowledge and observational data

To improve the estimation of uncertainty, it is necessary to develop efficient methods combining prior knowledge of the uncertainties with observational data.

This could include work on meteorology, source term, use of an ensemble of models, and more advanced inverse methods to infer the input uncertainties from the use of measurements. Investigations into improving the a priori knowledge of uncertainties could include:

- working on calibration methods to enhance the spread of meteorological ensembles; a way to compensate for ill-represented uncertainties in weather forecasts which is already used by meteorological offices (Christensen et al. 2017; Kaufmann and Rüdisühli 2019); as discussed in Leadbetter et al. (2018)), whilst uncertainty in variables such as wind and precipitation are well captured in a meteorological ensemble, the uncertainty in other variables, such as boundary layer height are not;
- setting up a framework of discussion with experts of source term estimation to improve knowledge of source term uncertainties; this could take place through a link with other European projects such as the MUSA project that deals with the uncertainties of severe accidents (https://cordis.europa.eu/project/rcn/223700/factsheet/en);
- better taking into account model-related uncertainties, perhaps connecting with the ENSEMBLE project (http://ensemble.jrc.ec.europa.eu/), for instance.

In addition, more advanced inverse methods may be used to infer input uncertainties from the use of measurements. In particular, Bayesian methods could be used to reconstruct the probability density function of input parameters (Lucas et al. 2016; Lucas et al. 2017). These methods could be fully integrated in inverse modelling systems, as they allow reconstruction of the source term (pdf) as well as other inputs.

References

Bedwell, P., S. Leadbetter, J. Tomas, S. Andronopoulos, I. Korsakissok, R. Périllat, A. Mathieu, G. Geertsema, H. Klein, H. De Vries, T. Hamburger, T. Pazmandi, C. Rudas and A. Sogachev (2018a). D 1.1.4 Guidelines detailing the range and distribution of atmospheric dispersion model input parameter uncertainties. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Bedwell, P., J. Wellings, S. Leadbetter, J. Tomas, S. Andronopoulos, I. Korsakissok, R. Périllat, A. Mathieu, G. Geertsema, H. Klein, H. De Vries, T. Hamburger, F. Gering, T. Pazmandi, P. Szanto, C. Rudas, A. Sogachev and C. Twenhöfel (2018b). D 1.1.5 Guidelines for ranking uncertainties in atmospheric dispersion. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Berge, E., H. Klein, M. Ulimoen and S. Andronopoulos (2019). D 9.5.2 Ensemble calculation for hypothetical accident scenarios in Europe: the Norway case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Camps, J., F. Menneson, W. Raskob, T. Peltonen, H. Walter, M. Pattantyùs-Ábrahám, S. Andronopoulos, M. Montero Prieto and C. Trueba Alonso (2018). The NERIS near-range atmospheric dispersion modelling experiment. 4th NERIS workshop. Dublin, Ireland.

CEC (2016). "Council Regulation (Euratom) 2016/52 of 15 January 2016 laying down maximum permitted levels of radioactive contamination of food and feed following a nuclear accident or any other case of radiological emergency, and repealing Regulation (Euratom) No 3954/87 and Commission Regulations (Euratom) No 944/89 and (Euratom) No 770/90." Official Journal of the European Union.

Charnock, T. W., A. Bexon, J. C. Sherwood, N. A. Higgins and S. J. Field (2013). PACE: A geographic information system based level 3 probabilistic accident consequence evaluation program. ANS PSA 2013 International Topic Meeting on Probabilisitic Safety Assessment and Analysis. Columbia, SC, American Nuclear Society.

Chevalier-Jabet, K. (2019a). Design and use of Bayesian networks for the diagnosis/prognosis of severe nuclear accidents, EU program for research and innovation H2020 - FAST Nuclear Emergency Tools (FASTNET) project. Report FASTNET-DATA-D2.3.

Chevalier-Jabet, K. (2019b). Source term prediction in case of a severe nuclear accident. 5th NERIS workshop. Roskilde, Denmark.

Christensen, H. M., S.-J. Lock, I. M. Moroz and T. N. Palmer (2017). "Introducing independent patterns into the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme." Quarterly Journal of the Royal Meteorological Society **143**(706): 2168-2181.

Connan, O., C. Leroy, F. Derkx, D. Maro, D. Hébert, P. Roupsard and M. Rozet (2011). "Atmospheric dispersion of an elevated release in a rural environment: Comparison between field SF6 tracer measurements and computations of Briggs and ADMS models." Atmospheric Environment **45**(39): 7174-7183.

Connan, O., K. Smith, C. Organo, L. Solier, D. Maro and D. Hébert (2013). "Comparison of RIMPUFF, HYSPLIT, ADMS atmospheric dispersion model outputs, using emergency response procedures, with 85Kr measurements made in the vicinity of nuclear reprocessing plant." Journal of Environmental Radioactivity **124**: 266-277.

Connan, O., L. Solier, D. Hébert, D. Maro, M. Lamotte, C. Voiseux, P. Laguionie, O. Cazimajou, S. Le Cavelier, C. Godinot, M. Morillon, L. Thomas and S. Percot (2014). "Near-field krypton-85 measurements in stable meteorological conditions around the AREVA NC La Hague reprocessing plant: estimation of atmospheric transfer coefficients." Journal of Environmental Radioactivity **137**: 142-149.

De Meutter, P., J. Camps, A. Delcloo, B. Deconninck and P. Termonia (2016). "On the capability to model the background and its uncertainty of CTBT-relevant radioxenon isotopes in Europe by using ensemble dispersion modeling." Journal of Environmental Radioactivity **164**: 280-290.

Draxler, R., D. Arnold, M. Chino, S. Galmarini, M. Hort, A. Jones, S. Leadbetter, A. Malo, C. Maurer, G. Rolph, K. Saito, R. Servranckx, T. Shimbori, E. Solazzo and G. Wotawa (2015). "World Meteorological

Organization's model simulations of the radionuclide dispersion and deposition from the Fukushima Daiichi nuclear power plant accident." Journal of Environmental Radioactivity **139**: 172-184.

Ferranti, L. (2017). Clustering Techniques and their applications at ECMWF.

Ferranti, L. and S. Corti. (2011). "New clustering products." ECMWF Newsletter 127, from https://www.ecmwf.int/sites/default/files/elibrary/2011/17442-new-clustering-products.pdf.

French, S., N. Argyris, H. Layton, J. Smith, S. M. Haywood and M. Hort (2016). Presenting Uncertain Information in Radiological Emergencies. ADMLC/2014/01. Report. https://admlc.files.wordpress.com/2014/05/admlc_final_report_v1-0.pdf

Galmarini, S., R. Bianconi, R. Addis, S. Andronopoulos, P. Astrup, J. C. Bartzis, R. Bellasio, R. Buckley, H. Champion, M. Chino, R. D'Amours, E. Davakis, H. Eleveld, H. Glaab, A. Manning, T. Mikkelsen, U. Pechinger, E. Polreich, M. Prodanova, H. Slaper, D. Syrakov, H. Terada and L. Van der Auwera (2004a). "Ensemble dispersion forecasting—Part II: application and evaluation." Atmospheric Environment **38**(28): 4619-4632.

Galmarini, S., R. Bianconi, W. Klug, T. Mikkelsen, R. Addis, S. Andronopoulos, P. Astrup, A. Baklanov, J. Bartniki, J. C. Bartzis, R. Bellasio, F. Bompay, R. Buckley, M. Bouzom, H. Champion, R. D'Amours, E. Davakis, H. Eleveld, G. T. Geertsema, H. Glaab, M. Kollax, M. Ilvonen, A. Manning, U. Pechinger, C. Persson, E. Polreich, S. Potemski, M. Prodanova, J. Saltbones, H. Slaper, M. A. Sofiev, D. Syrakov, J. H. Sorensen, L. Van der Auwera, I. Valkama and R. Zelazny (2004b). "Can the confidence in long range atmospheric transport models be increased? The pan-european experience of ensemble." Radiation Protection Dosimetry **109**(1-2): 19-24.

Galmarini, S., R. Bianconi, W. Klug, T. Mikkelsen, R. Addis, S. Andronopoulos, P. Astrup, A. Baklanov, J. Bartniki, J. C. Bartzis, R. Bellasio, F. Bompay, R. Buckley, M. Bouzom, H. Champion, R. D'Amours, E. Davakis, H. Eleveld, G. T. Geertsema, H. Glaab, M. Kollax, M. Ilvonen, A. Manning, U. Pechinger, C. Persson, E. Polreich, S. Potemski, M. Prodanova, J. Saltbones, H. Slaper, M. A. Sofiev, D. Syrakov, J. H. Sørensen, L. V. d. Auwera, I. Valkama and R. Zelazny (2004c). "Ensemble dispersion forecasting—Part I: concept, approach and indicators." Atmospheric Environment **38**(28): 4607-4617.

Geertsema, G., H. De Vries and R. Scheele (2019a). High resolution meteorological ensemble data for CONFIDENCE research on uncertainties in atmospheric dispersion in the (pre-) release phase of a nuclear accident. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Geertsema, G., H. De Vries and R. Sheele (2019b). High resolution meteorological ensemble data for CONFIDENCE research on uncertainties in atmospheric dispersion in the (pre-)release phase of a nuclear accident. 19th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Girard, S., I. Korsakissok, V. Mallet and A. Mathieu (2015). "Emulation and Sobol' sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident." Journal of Geophysical Research **accepté pour publication**.

Harvey, N. J., N. Huntley, H. F. Dacre, M. Goldstein, D. Thomson and H. Webster (2018). "Multi-level emulation of a volcanic ash transport and dispersion model to quantify sensitivity to uncertain parameters." Nat. Hazards Earth Syst. Sci. **18**(1): 41-63.

Haywood, S. M., P. Bedwell and M. C. Hort (2010). "Key factors in imprecision in radiological emergency response assessments using the NAME model." Journal of Radiological Protection **30**(1): 23-36.

Jones, A., D. Thomson, M. Hort and B. Devenish (2007). The UK Met Office's next generation atmospheric dispersion model, NAME III. 27th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application. Berlin, Springer.

Kaufmann, P. and S. Rüdisühli (2019). Dispersion model uncertainty simulation using a limited area ensemble model. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Klonner, R. (2013). Clustering ECMWF ENS ensemble predictions to optimise FLEXPART plume dispersion ensembles (Master's Thesis), University of Vienna

Korsakissok, I., M. Contu, O. Connan, A. Mathieu and D. Didier (2016). Validation of the Gaussian puff model pX using near-field krypton-85 measurements around the AREVA NC La Hague reprocessing plant: comparison of dispersion schemes. HARMO 17. Budapest.

Korsakissok, I., R. Périllat, S. Andronopoulos, P. Bedwell, E. Berge, A. Quérel, H. Klein, S. Leadbetter, O. Saunier, A. Sogachev, J. Tomas and M. Ulimoen (2019a). D 9.5.3 Ensemble calculation for a past accident scenario: the Fukushima case study. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Korsakissok, I., R. Périllat, G. Geertsema, H. De Vries, T. Hamburger, J. Tomas, S. Andronopoulos, P. Bedwell, T. Charnock, I. Ievdin, S. Leadbetter, T. Pazmandi, C. Rudas, R. Scheele, A. Sogachev, P. Szanto and J. Wellings (2019b). D 9.5.1 Ensemble calculations for the atmospheric dispersion of radionuclides. Hypothetical accident scenarios in Europe: the REM case studies. D9.5 Guidelines for the use of ensemble calculations in an operational context, indicators to assess the quality of uncertainty modeling and ensemble calculations, and tools for ensemble calculation for use in emergency response, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Lablans, W. N. and P. J. Rijkoort (1974). "De betekenis van de veranderlijkheid van de windrechting voor schattingen van de ligging van het besmette gebied na ongevallen met schadelijke stoffen (The significance of the variability of the wind alignment for estimates of the location of the contaminated area after accidents with harmful substances). Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meteorological Institute). Wetenschappelijk Rapport (WR) 74-12 (Scientific Report 74 12).".

Leadbetter, S., S. Andronopoulos, P. Bedwell, G. Geertsema, A. Jones, I. Korsakissok, J. Tomas and H. De Vries (2018). D1.1.1 Using ensemble meteorological forecasts to represent meteorological uncertainty in dispersion models. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Lucas, D. D., A. Gowardhan, P. Cameron-Smith and R. L. Baskett (2016). "Impact of meteorological inflow uncertainty on tracer transport and source estimation in urban atmospheres." Atmospheric Environment **143**: 120-132.

Lucas, D. D., M. Simpson, P. Cameron-Smith and R. L. Baskett (2017). "Bayesian inverse modeling of the atmospheric transport and emissions of a controlled tracer release from a nuclear power plant." Atmospheric Chemistry and Physics (Online): Medium: ED; Size: p. 13521-13543.

Masson, O., G. Steinhauser, D. Zok, O. Saunier, H. Angelov, D. Babić, V. Bečková, J. Bieringer, M. Bruggeman, C. I. Burbidge, S. Conil, A. Dalheimer, L.-E. De Geer, A. de Vismes Ott, K. Eleftheriadis, S. Estier, H. Fischer, M. G. Garavaglia, C. Gasco Leonarte, K. Gorzkiewicz, D. Hainz, I. Hoffman, M. Hýža, K. Isajenko, T. Karhunen, J. Kastlander, C. Katzlberger, R. Kierepko, G.-J. Knetsch, J. Kövendiné Kónyi, M. Lecomte, J. W. Mietelski, P. Min, B. Møller, S. P. Nielsen, J. Nikolic, L. Nikolovska, I. Penev, B. Petrinec, P. P. Povinec, R. Querfeld, O. Raimondi, D. Ransby, W. Ringer, O. Romanenko, R. Rusconi, P. R. J. Saey, V. Samsonov, B. Šilobritienė, E. Simion, C. Söderström, M. Šoštarić, T. Steinkopff, P. Steinmann, I. Sýkora, L. Tabachnyi, D. Todorovic, E. Tomankiewicz, J. Tschiersch, R. Tsibranski, M. Tzortzis, K. Ungar, A. Vidic, A. Weller, H. Wershofen, P. Zagyvai, T. Zalewska, D. Zapata García and B. Zorko (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.

Mathieu, A., I. Korsakissok, R. Périllat, K. Chevalier-Jabet, F. Stephani, S. Fougerolle, V. Créach, E. Cogez and P. Bedwell (2018). D 1.1.3 Guidelines describing source term uncertainties. D9.1 Guidelines ranking uncertainties for atmospheric dispersion, European Joint Program for the Integration of Radiation Protection Research CONCERT - CONFIDENCE project.

Maurer, C., J. Baré, J. Kusmierczyk-Michulec, A. Crawford, P. W. Eslinger, P. Seibert, B. Orr, A. Philipp, O. Ross, S. Generoso, P. Achim, M. Schoeppner, A. Malo, A. Ringbom, O. Saunier, D. Quèlo, A. Mathieu, Y. Kijima, A. Stein, T. Chai, F. Ngan, S. J. Leadbetter, P. De Meutter, A. Delcloo, R. Britton, A. Davies, L. G. Glascoe, D. D. Lucas, M. D. Simpson, P. Vogt, M. Kalinowski and T. W. Bowyer (2018). "International challenge to model the long-range transport of radioxenon released from medical isotope production to six Comprehensive Nuclear-Test-Ban Treaty monitoring stations." Journal of Environmental Radioactivity **192**: 667-686.

Nisbet, A. F. (2019). Public Health Protection in Radiation Emergencies. PHE-CRCE Series. Chilton, UK, Public Health England. Report PHE-CRCE-049.

Périllat, R., I. Korsakissok, V. Mallet, A. Mathieu, D. Didier, T. Sekiyama, M. Kajino, K. Adachi and Y. Igarashi (2017). Using meteorological ensembles for atmospheric dispersion modelling of the Fukushima nuclear accident. 18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Bologna, Italy, 9-12 October 2017.

Piters, A. J. M., K. F. Boersma, M. Kroon, J. C. Hains and M. Van Roozendael (2012). "The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results." Atmospheric Measurement Techniques **5**(20): 457-485.

Saunier, O., A. Mathieu, D. Didier, O. Masson and D. l. B. J. (2019). "Radioactive Ruthenium from an Undeclared Major Nuclear Release in 2017, Part 2: Atmospheric Modelling and Source Reconstruction." Proceedings of the National Academy of Sciences **Accepted for publication**.

Sigg, R. and H. Grahn (2019). Effective and efficient combination of weather and radioactive source ensembles. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Sørensen, J. H., B. Amstrup, T. Bøvith, H. Feddersen, R. Gill, M. Sørensen, F. Vejen, P. Astrup, N. Davis, B. Lauritzen, S. Cordt Hoe, J. E. Dyve and P. Lindahl (2017). MEteorological uncertainty of ShOrt-range dispersion (MESO), NKS. Report NKS-380. http://www.nks.org/en/nks_reports/view_document.htm?id=111010214043891

Sørensen, J. H., B. Amstrup, H. Feddersen, J. Bartnicki, H. Klein, M. Simonsen, B. Lauritzen, S. C. Hoe, C. Israelson and J. Lindgren (2016). Fukushima Accident: UNcertainty of Atmospheric dispersion modelling (FAUNA), NKS. Rapport NKS-360. Report.

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 (2014). Meteorological Uncertainty of atmospheric Dispersion model results (MUD), Nordic Nuclear Safety Research (NKS). Report NKS-307. http://www.nks.org/en/nks_reports/view_document.htm?id=111010212220490

Sørensen, J. H., F. Schönfeldt, R. Sigg, J. Pehrsson, B. Lauritzen, J. Bartnicki, H. Klein, S. Cordt Hoe and J. Lindgren (2019). Added Value of uncertainty Estimates of SOurce term and Meteorology (AVESOME), NKS. Report NKS-420. http://www.nks.org/en/nks_reports/view_document.htm?id=111010214696230

Tomas, J. and C. Twenhöfel (2019). The effects of uncertainties on the radiological assessment in the (pre-) release phase of a nuclear accident. 19th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium.

Van dop, H., R. Addis, G. Fraser, F. Girardi, G. Graziani, Y. Inoue, N. Kelly, W. Klug, A. Kulmala, K. Nodop and J. Pretel (1998). "ETEX: A European tracer experiment; observations, dispersion modelling and emergency response." Atmospheric Environment **32**(24): 4089-4094.