

# SCIENTIFIC REPORTS



OPEN

## Source localisation and its uncertainty quantification after the third DPRK nuclear test

Pieter De Meutter<sup>1,2,3</sup>, Johan Camps<sup>1</sup>, Andy Delcloo<sup>2,3</sup> & Piet Termonia<sup>2,3</sup>

The International Monitoring System is being set up aiming to detect violations of the Comprehensive Nuclear-Test-Ban Treaty. Suspicious radioxenon detections were made by the International Monitoring System after the third announced nuclear test conducted by the Democratic People's Republic of Korea (DPRK). In this paper, inverse atmospheric transport and dispersion modelling was applied to these detections, to determine the source location, the release term and its associated uncertainties. The DPRK nuclear test site was found to be a likely source location, though a second likely source region in East Asia was found by the inverse modelling, partly due to the radioxenon background from civilian sources. Therefore, techniques to indirectly assess the influence of the radioxenon background are suggested. In case of suspicious radioxenon detections after a man-made explosion, atmospheric transport and dispersion modelling is a powerful tool for assessing whether the explosion could have been nuclear or not.

The Comprehensive Nuclear-Test-Ban Treaty (CTBT) was opened for signature in 1996. Once the Treaty enters into force, it will ban atmospheric, underwater and underground nuclear explosions. The International Monitoring System is almost fully operational and already measures infrasound, hydroacoustic and seismic waves, together with certain airborne radionuclides. The radionuclide measurement system will consist of eighty ground stations measuring radioactive particulates. Forty of these stations will also be equipped with noble gas detectors, allowing the measurement of airborne concentrations of  $^{131m}\text{Xe}$ ,  $^{133m}\text{Xe}$ ,  $^{133}\text{Xe}$  and  $^{135}\text{Xe}$  (with half-lives of respectively 11.84 d, 2.20 d, 5.25 d and 9.14 h<sup>1</sup>). These radioactive xenon isotopes (hereafter radioxenon) are selected for monitoring due to their significant production after a nuclear explosion, their half-life and their likelihood to seep through the ground in case of an underground nuclear explosion. Observations from the International Monitoring System are made available to member states that have signed the CTBT in order to verify compliance with the Treaty.

In the past, the Democratic People's Republic of Korea (DPRK) announced several times to have conducted a nuclear test. Although the International Monitoring System has not yet been fully installed (90% to date<sup>2</sup>), these tests provide interesting cases for testing the capability to detect clandestine nuclear tests and to confirm announced nuclear tests. Seismic waves were detected after each DPRK nuclear test<sup>2</sup>. These observations allow the determination of the timing, location and magnitude of the explosion. In order to discriminate conventional explosions from nuclear explosions, a specific signature of radionuclides needs to be determined. If measured, atmospheric transport models can be used to verify whether these radionuclides originated from the nuclear test site. This task is complicated by two factors: (i) the DPRK nuclear tests were conducted underground, so that it is a priori not sure if and what fraction of radionuclides were released into the atmosphere; (ii) there is an ongoing background of radioxenon from civilian sources such as nuclear power plants and medical isotope production facilities, resulting in frequent detections by the International Monitoring System<sup>3</sup>.

The inverse problem in atmospheric transport and dispersion modelling (ATM) consists of finding source parameters such as release location, release height, release amount and release period, based on a set of observations. The (dis)agreement between these observations and their corresponding simulated values obtained from the selected source parameters, is quantified by a cost function or likelihood function. An overview of different inverse modelling methods has recently been published<sup>4</sup>. Examples of inverse atmospheric transport modelling of point sources include the calculation of the release amount of radionuclides from nuclear accidents, such as

<sup>1</sup>Belgian Nuclear Research Institute, Mol, 2400, Belgium. <sup>2</sup>Royal Meteorological Institute of Belgium, Brussels, 1180, Belgium. <sup>3</sup>Ghent University, Department of Physics and Astronomy, Ghent, 9000, Belgium. Correspondence and requests for materials should be addressed to P.D.M. (email: [pieter.de.meutter@sckcen.be](mailto:pieter.de.meutter@sckcen.be))

the Fukushima Dai-ichi nuclear power plant accident in 2011<sup>5–8</sup>, or modelling the release of ash from volcano eruptions like the Eyjafjallajökull in 2010<sup>9,10</sup> which severely disturbed air traffic above Europe. For these cases, inverse modelling is often followed by forward modelling using the best source parameters obtained from the inverse modelling, in order to know the concentration of a tracer at any location and any time. This complements the observations, which are typically sparse. In other cases, inverse modelling is used to locate possible sources (such as the source location of <sup>131</sup>I detections made in Europe in 2011<sup>11</sup>). Inverse atmospheric transport modelling is also used in other areas of research, such as the quantification of greenhouse gas fluxes in the context of climate science.

In the past, radioxenon observations were linked to nuclear tests conducted by the DPRK<sup>12–14</sup>. Of particular interest are the specific combinations of <sup>131m</sup>Xe and <sup>133</sup>Xe that were detected seven to eight weeks after the third announced nuclear test in February 2013<sup>14</sup>. These detections were made in three samples taken at the station RN38 in Takasaki (Japan) between 7–9 April 2013 and in two samples taken at the station RN58 in Ussuriysk (Russia) between 12–13 April 2013. The radioxenon detections were attributed to a nuclear explosion, mainly based on their unusual occurrence compared to the station's record<sup>14</sup>. Atmospheric transport and dispersion modelling was used to calculate the source-receptor-sensitivities between the stations (or “receptors”) and possible sources. By calculating correlations between the source-receptor-sensitivities and the observed concentrations, so-called “Possible Sources Regions”<sup>15</sup> were constructed. The nuclear test site Punggye-ri was found to fall within these possible source regions<sup>14</sup>.

More complex ATM methods for assessing the third nuclear test conducted by the DPRK have been made using Bayesian methods to find source parameters<sup>16,17</sup>, and using an optimisation method to find the possible source regions<sup>18</sup>. All three studies identified the Punggye-ri nuclear test site as a likely source of the suspicious radioxenon detections. An automated methodology for principled computational inference for Bayesian source reconstruction was described and applied to the third nuclear test conducted by the DPRK<sup>17</sup>. In another study, a method for fusing information from the radionuclide verification regime with the seismic verification regime was proposed<sup>18</sup>, which method consists of assessing the cost function values at only those locations where seismic activity had been measured.

After the announced nuclear test in January 2016, elevated <sup>133</sup>Xe concentrations were measured by the International Monitoring System, compatible with a delayed release from the Punggye-ri nuclear test site<sup>19</sup>. However, due to the lack of detections of other radioxenon isotopes, it could not be excluded that the observed <sup>133</sup>Xe originated from civilian sources. In such cases, ATM is able to narrow down possible source regions and provide information on the release period and release amount. The resulting area, preferably as small as possible, depends on the meteorological conditions and the network configuration: more observations at different geotemporal locations will result in a smaller area<sup>20</sup>.

In this paper, the existing methodology<sup>19</sup> was extended to improve the estimation of the source location and to address the problem of the radioxenon background. Inverse long-range atmospheric transport and dispersion modelling was used to assess the source location of the radioxenon detections previously described<sup>14</sup>. The following two questions were addressed in particular: (i) how precise can we narrow down possible source regions and (ii) can these detections be linked to a delayed release from the Punggye-ri nuclear test site. An uncertainty quantification was provided by using unique meteorological ensemble data, which allowed to construct grid box probability maps for the source localisation.

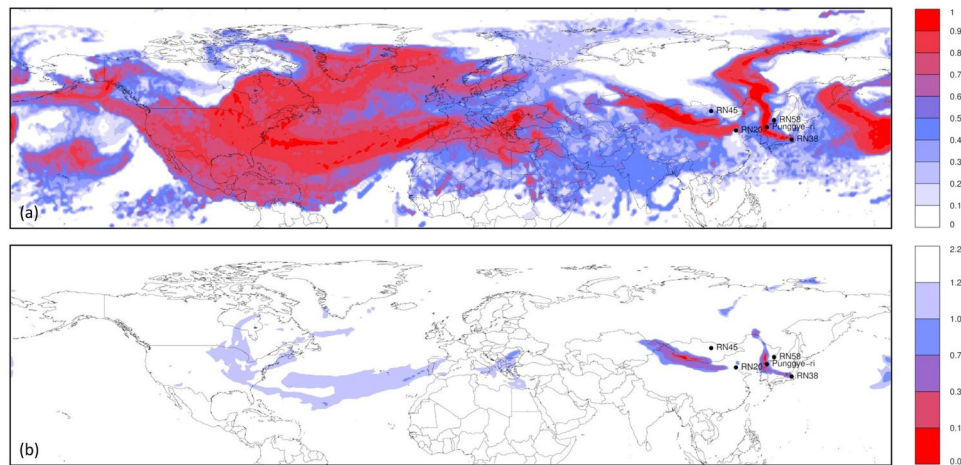
## Source localisation of the Xe-133 detections

In this section, the possible source locations of the selected <sup>133</sup>Xe samples will be determined. First, inverse modelling was applied over the northern hemisphere using deterministic weather data (Methods). From that assessment, a smaller region of interest was identified for which the inverse modelling was repeated using the ensemble weather data (Methods).

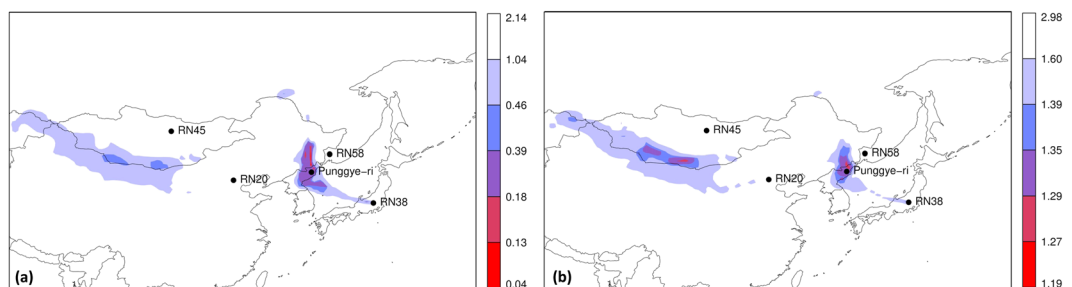
**Assessment over the northern hemisphere.** We started with assessing which areas in the northern hemisphere could be potential source locations, explaining the observations. The deterministic weather model was used with horizontal grid spacings of 1°. Two methods were used, the results of which are shown in Fig. 1. First, the correlation between the source-receptor-sensitivities and the observed <sup>133</sup>Xe activity concentrations were calculated for all simulation times and for each grid box. This corresponds to the correlation between observed and simulated activity concentrations assuming a single short release (the duration of the release corresponds to the output frequency of the source-receptor-sensitivity fields by Flexpart, which was 3 hours). The result is what is sometimes called the “Possible Source Region” or PSR<sup>15</sup>. Since we were interested in the possible source locations and not in the release timing, the grid box maximum correlation was taken over the full simulation period of 26 days. The result is shown in Fig. 1a.

Second, inverse modelling using the optimisation technique described in Methods was used to obtain a cost function value for each grid box. The inverse modelling was applied once for the full simulation period, so that a single cost function value was obtained for each grid box, in contrast to the method where correlations were calculated.

The correlation map (Fig. 1a) shows a few large distinct areas in East Asia, as well as a huge area spanning from the Middle East to a large part of North America. Many grid boxes had a high correlation at some point in the simulation period, so that the correlation did not seem to be a good way to distinguish likely source regions from other regions, unless when using a very high correlation coefficient as threshold. The cost function values resulting from the inverse modelling are shown in Fig. 1b. Two distinct areas can be seen: the first lies roughly over the DPRK and the second lies near the border between China and Mongolia. These two regions have also a very high correlation (Fig. 1a).



**Figure 1.** (a) Grid box maximum “Possible Source Region” (see text) for all simulation times for the northern hemisphere. (b) Minimum value of the cost function (Eq. 3 was used here) after inverse modelling was applied to each grid box separately. Maps were generated using the R statistical software<sup>44,45</sup>.

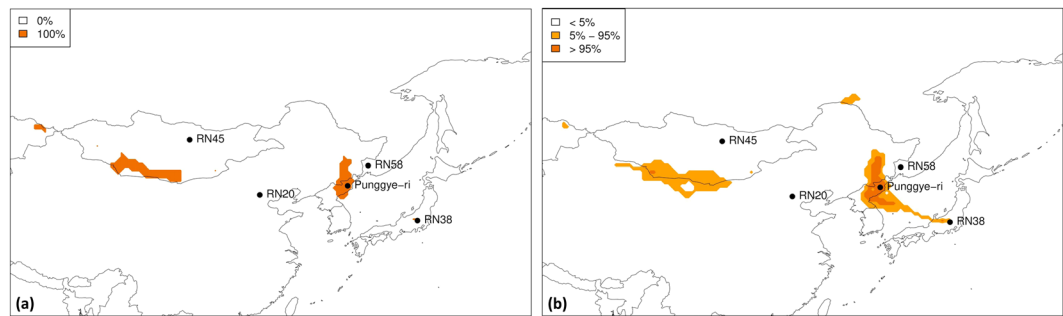


**Figure 2.** Grid box median cost function in the lowest model level using (a) Eq. 3 and (b) Eq. 4. The levels in the legend correspond to the 0, 0.0005, 0.001, 0.005, 0.01, 0.05 and 1 quantiles. Maps were generated using the R statistical software<sup>44,45</sup>.

When comparing both methods, it can be seen that the inverse modelling is better in discriminating likely source regions from other regions than the “Possible Source Regions” method: many grid boxes that had a high correlation, are unlikely according to the inverse modelling. One might wonder whether the difference could be attributed to the upper bound on the release term in the inverse modelling. However, the cost function maps obtained from applying inverse modelling with an upper bound increased by three orders of magnitude ( $Q_{max} = 10^{16} \text{ Bq/day}$ ) gave a roughly identical pattern (not shown).

**Focus on East Asia.** Based on the assessment over the northern hemisphere, we identified two regions of interest located in East Asia. To assess these regions further, we used the unique data from the Ensemble Data Assimilation system of ECMWF. This resulted in a set of 51 ensemble members or meteorological scenarios representing the uncertainty in the meteorological data (Methods). For each ensemble member, we performed a set of backward Flexpart runs to calculate the source-receptor-sensitivity fields with horizontal grid spacings of  $0.5^\circ$ . Next, we performed inverse modelling (Methods) to obtain a set of 51 cost function values and optimal release terms for each grid box, one for each ensemble member. The inverse modelling was repeated twice, each time using one of the two cost functions described in Methods (Eqs 3 and 4). Figure 2 shows the resulting two maps of cost function values. The grid box median value was taken over all 51 cost function values. Both patterns of the cost function values are similar, so that the possible source locations are robust for the choice of the cost function.

In order to visualize the uncertainty captured by the ensemble of 51 cost function maps, a cost function threshold was first applied to discriminate possible source regions from other regions. The threshold was set at the 0.01 quantile of all grid box cost function values (the total number of grid boxes was 13,680). A possible source location probability map was then constructed by taking the percentage of members having a cost function below the threshold for each grid box, as was done in a previous study<sup>19</sup>. This is justified by the fact that in the Ensemble Data Assimilation system of ECMWF, each member can be considered equally likely, hence each member should have equal weighting. This is expected not to be the case for multimodel ensembles, where instead different weightings should be used depending on the member’s individual performance. The result for the inverse modelling using Eq. 3 is shown in Fig. 3. The left panel of Fig. 3 was obtained using the unperturbed meteorological data only (to better illustrate the added value of the ensemble versus the deterministic case), while the right panel was obtained using the full ensemble. The deterministic case shows two possible source regions (Fig. 3a). For the



**Figure 3.** Grid box probability map using the cost function in Eq. 3 for (a) the unperturbed member only and (b) the full ensemble. Maps were generated using the R statistical software<sup>44,45</sup>.

full ensemble, a discrimination was made between grid boxes for which at least 95% of the ensemble members agree that grid boxes are likely sources (dark orange area in Fig. 3b) or unlikely sources (white area in Fig. 3b). The likely sources regions were smaller when using the ensemble compared to the deterministic case, but a new region appeared where the ensemble members did not agree on whether a grid box was a likely source or not (yellow area in Fig. 3b). A notable difference between the deterministic and probabilistic case is the area near the Chinese-Mongolian border: the deterministic model showed that it is a likely source region, while the ensemble showed that it is associated with large uncertainties. The deterministic case can thus (i) be overly confident for certain grid boxes and (ii) perhaps worse, miss certain grid boxes that are likely source regions according to the ensemble.

### Influence of the $^{133}\text{Xe}$ background

The radioxenon background varies significantly with time and space, and modelling its daily value at a specific location is not trivial. Furthermore, although significant effort has been put into understanding the global radioxenon background<sup>21,22</sup>, it could be that not all local sources of radioxenon are known. In this paragraph, we describe the assumptions that we made on the background, but we do not attempt to model it explicitly due to the uncertainties described here.

The signal of interest for which we wanted to find the possible source regions are the five samples that were shown in a previous study to be exceptional with respect to the station's historic data due to the particular ratio of  $^{133}\text{Xe}$  and  $^{131\text{m}}\text{Xe}$ <sup>14</sup>. It was assumed that the signal of interest came from a single source (either a nuclear test or a civilian source). The background signal, defined as any other signal originating from one or multiple sources, was assumed to be significantly smaller than the signal of interest in the  $^{133}\text{Xe}$  samples containing more than 1  $\text{mBq}/\text{m}^3$ . In the other samples, having a  $^{133}\text{Xe}$  activity concentration ranging from below detection level up to a few tenths of  $\text{mBq}/\text{m}^3$ , the contribution from the background was assumed to be anything up to 100%.

In this section, we assess the influence of the assumed  $^{133}\text{Xe}$  background on the inverse modelling.

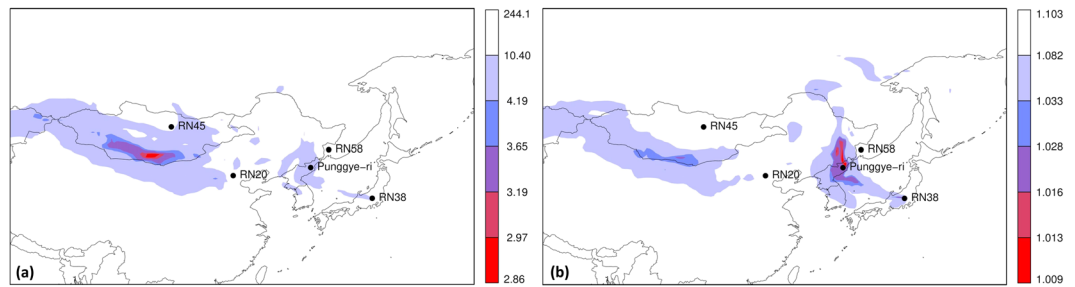
**Robustness of the cost function to the background.** We start by arguing that the cost functions used here (Eqs 3 and 4) have an inherent robustness against small background contamination of the samples. Indeed, both the mean square error and the correlation in Eq. 3 are resilient to small-scale perturbations on the radioxenon activity concentrations.

The geometric variance treats all observations equally, but can have similar resilience via the parameter  $\alpha$  in Eq. 4. This parameter  $\alpha$  was added in the formula to allow the inclusion of non-detections in the inverse modelling (having activity concentrations ranging from zero to just below detection level). At the same time, it allowed for the adjustment of the tightness of fit between the observed and simulated activity concentrations. If a low value for  $\alpha$  is chosen, observations with low activity concentrations should match equally well as observations with high activity concentrations. If a high value is attributed to  $\alpha$ , the inverse modelling will be dominated by the information from observations with high activity concentrations. The value of  $\alpha$  was set to 0.1  $\text{mBq}/\text{m}^3$ , slightly below the typical minimal detectable concentration. The effect of changing  $\alpha$  in Eq. 4 on the source localisation is shown in Fig. 4. For very small  $\alpha = 0.01 \text{ mBq}/\text{m}^3$ , and thus allowing a large influence of the background on the inverse modelling, the most likely source area is now found near the Chinese-Mongolian border. Although the area around the nuclear test site Punggye-ri is still distinguishable from other regions by a lower cost function, it is now a less likely source region (Fig. 4, left). When, on the other hand, a large  $\alpha = 1 \text{ mBq}/\text{m}^3$  was used, the inverse modelling now uses mostly information coming from samples with high activity concentration. The background will now have a small effect, with the risk of not using all relevant available information. The opposite effect can be seen on the cost function map: the area around the nuclear test site Punggye-ri is now the most likely source region (Fig. 4, right).

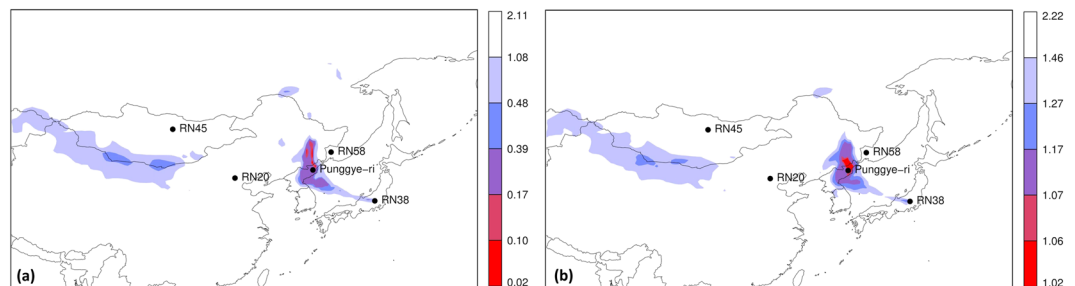
The value of  $\alpha$  should be chosen based on the minimum detectable concentration and the typical background concentrations and is thus dependent on the tracer of interest. Its best value for  $^{133}\text{Xe}$  likely ranged from 0.1  $\text{mBq}/\text{m}^3$  to 1  $\text{mBq}/\text{m}^3$ .

**Setting low  $^{133}\text{Xe}$  activity concentrations to zero.** The sensitivity to the background was further assessed by setting all  $^{133}\text{Xe}$  observations to zero, except for the five suspicious samples described in a previous study<sup>14</sup> and for the elevated  $^{133}\text{Xe}$  measurement of 0.7  $\text{mBq}/\text{m}^3$  taken at RN38 just after the three consecutive





**Figure 4.** Grid box cost function in the lowest model level using Eq. 4 for different values for  $\alpha$ : (a)  $\alpha = 0.01 \text{ mBq/m}^3$  and (b)  $\alpha = 1 \text{ mBq/m}^3$ . The levels in the legend correspond to the 0, 0.0005, 0.001, 0.005, 0.01, 0.1 and 1 quantiles. Only the unperturbed member was used. Maps were generated using the R statistical software<sup>44,45</sup>.



**Figure 5.** As Fig. 2, but with certain observations set to zero (see text) and for the unperturbed member only. As before, (a) was obtained using Eq. 3 and (b) using Eq. 4. Maps were generated using the R statistical software<sup>44,45</sup>.

samples containing particular combinations of  $^{131m}\text{Xe}$  and  $^{133}\text{Xe}$ . The resulting cost function maps are shown in Fig. 5. It shows that the main features remain, which is not surprising since the cost functions are not very sensitive to low activity concentrations as discussed in the previous subsection. However, (i) the area near the Chinese-Mongolian border seems now a less likely source and (ii) the area of possible sources around DPRK is now more concentrated around the Punggye-ri nuclear test site.

**Inverse modelling using a subset of the observations.** We now assess the degree of consistency among subsets of observations, meaning that all subsets of observations can be explained by one and the same single source. Due to the existence of the civilian radionuclide background, one might expect that observations will generally not be consistent. However, in the previous paragraphs we concluded that small inconsistencies do not significantly influence the inverse modelling source localisation. Here we selected 500 random subsets containing 75% of the observations, and applied inverse modelling on each of these subsets, using the unperturbed member only.

Such analysis could also be used to check for robustness against the case that a few samples contain wrong information, in the sense that the observed radionuclide is the signature of other sources than the source of interest. Indeed, using combinatorics, we can calculate the probability of not drawing  $n$  observations when a fraction  $frac = 75\%$  of observations are drawn out of a total of  $N = 57$  observations, and repeating this procedure  $ntries = 500$  times:

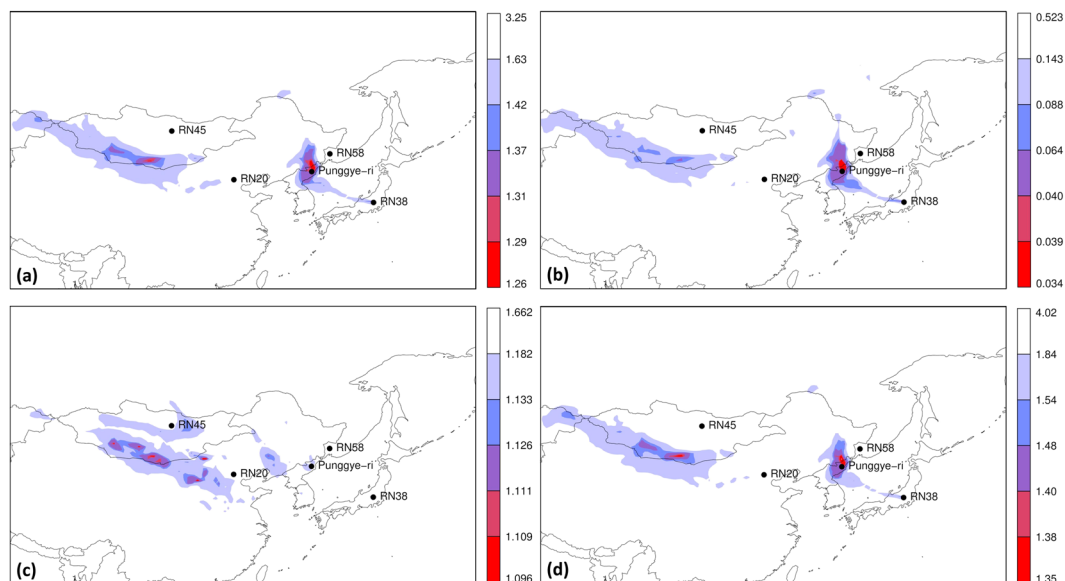
$$p(n; N, frac, ntries) = 1 - \left( 1 - \frac{(N-n)!(N-N\,frac)!}{N!(N-n-N\,frac)!} \right)^{ntries} \quad (1)$$

The probabilities for different values of  $n$  are given in Table 1. Note that with this setup, it is almost certain that any combination of three observations will not be selected at least once.

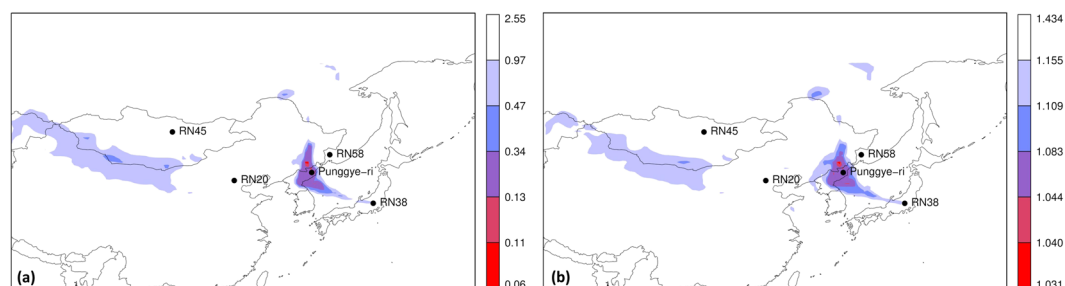
For each of the 500 subsets of observations, inverse modelling was applied. Equation 4 was used, since both cost functions gave similar results and the inverse modelling was slightly faster compared to Eq. 3. In order to interpret the resulting 500 cost function maps, we plotted the grid point minimum, maximum, median and standard deviation of the cost function value in Fig. 6 (as such, it could well be that in Fig. 6, different grid boxes make use of different subsets of observations). The grid box median of the cost function (Fig. 6a) is very similar to the calculation using all observations (Fig. 2b). This confirms indeed that the inverse modelling was consistent for subsets of observations, and that small perturbations from the background did not strongly influence the results. When plotting the grid box standard deviation of the cost function (Fig. 6b), the two possible source regions are shown to be robust for selecting different subsets of observations, especially the region around the Punggye-ri test site. The grid box minimum and maximum cost function values were also plotted (Fig. 6c and d). Figure 6c shows that it is possible to find a combination of observations for which a pattern was found significantly differing from

$n$	$p$
1	$\approx 1$
2	$\approx 1$
3	0.998
4	0.72
5	0.21
6	0.04

**Table 1.** Probability of not drawing  $n$  observations ( $p$ -values were calculated using Eq. 1).



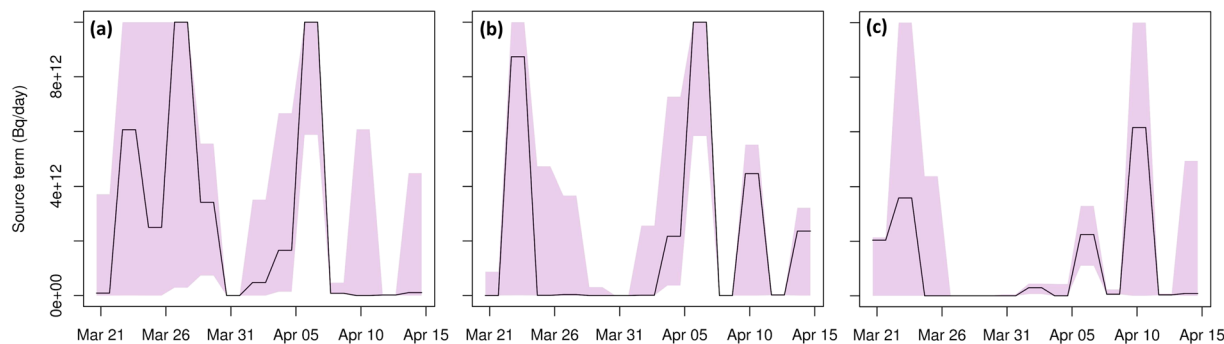
**Figure 6.** From 500 cost function maps obtained from different subsets of observation, the following statistics are plotted: (a) grid box median of the cost function; (b) grid box standard deviation of the cost function; (c) grid box minimum cost function; (d) grid box maximum cost function. The levels in the legend correspond to the 0, 0.0005, 0.001, 0.005, 0.01, 0.05 and 1 quantiles. Maps were generated using the R statistical software<sup>44,45</sup>.



**Figure 7.** As Fig. 2, but for the  $^{131m}\text{Xe}$  observations. Maps were generated using the R statistical software<sup>44,45</sup>.

previous patterns. However, this pattern corresponded to the case where the largest three  $^{133}\text{Xe}$  observations were omitted from the inverse modelling, while it was assumed that these contain most information on the event of interest. As such, it should not surprise that a different pattern was found. When requiring that a grid box should comply with all subsets of observations, which can be obtained by taking the grid box maximum of the cost functions (Fig. 6d), we found a pattern roughly identical to the median (Fig. 6a).

**Source localisation of the Xe-131m detections.** The inverse modelling was repeated using 57  $^{131m}\text{Xe}$  observations (the same period and stations as for the  $^{133}\text{Xe}$  observations as described in Methods were used). The cost function map is shown in Fig. 7. Not surprisingly, a pattern similar to Fig. 2 can be observed. Indeed, the same meteorological data were used, consequently the source-receptor-sensitivities are similar (the only difference comes from the radioactive decay since  $^{131m}\text{Xe}$  and  $^{133}\text{Xe}$  have different half-lives). Furthermore, the observations of  $^{131m}\text{Xe}$  and  $^{133}\text{Xe}$  are strongly correlated: the Pearson correlation coefficient is 0.86 for these 57



**Figure 8.** Release profile (black solid line) obtained from the unperturbed member for the Punggye-ri nuclear test site: (a)  $^{133}\text{Xe}$  release profile using unadapted  $^{133}\text{Xe}$  observations, (b)  $^{133}\text{Xe}$  release profile using modified  $^{133}\text{Xe}$  observations and (c)  $^{131m}\text{Xe}$  release profile using unchanged  $^{131m}\text{Xe}$  observations. The shadings represent the 0.025 and 0.975 quantiles of the full ensemble. The figure was made using the R statistical software<sup>44</sup>.

observations. Figures 2 and 7 differ notably in the area around the Punggye-ri nuclear test site, which is the single most likely source region in Fig. 7, for both cost functions Eqs 3 and 4.

**Hypothetical radioxenon release profile at the Punggye-ri nuclear test site.** In this section, it is assumed that the radioxenon signal of interest originated from a delayed release at the Punggye-ri nuclear test site. The corresponding radioxenon release profile was calculated for three sets of observations: (i) unmodified  $^{133}\text{Xe}$  observations, (ii) the  $^{133}\text{Xe}$  observations changed as discussed previously (where most  $^{133}\text{Xe}$  observations were set to zero) and (iii) unmodified  $^{131m}\text{Xe}$  observations. The resulting three release profiles are shown in Fig. 8. The full ensemble was used, from which uncertainty bounds were created. Eq. 3 was used as cost function, though similar features were found when using Eq. 4 (not shown).

The release profile obtained using unmodified observations (Fig. 8a) shows a possible release between 21 and 30 March 2013, but the associated uncertainties are very large. Around 27 and 29 March, most ensemble members predicted a significant release, although the exact amount released showed quite some spread among the ensemble members. Between 6 and 7 April, a sharp release was found with less uncertainty. The ensemble revealed that the release could have started a few days earlier. Finally, the ensemble also showed the possibility of releases around 10 April and 14 April but with a large degree of uncertainty.

When using modified observations as discussed previously with the aim to minimise the effect of the background, we see that the releases between 26 and 30 March 2013 had decreased significantly. The uncertainty bounds obtained from the ensemble suggests that the only significant release occurred between 4 and 7 April. The additional release in Fig. 8(a) compared to Fig. 8(b) is the result of the inverse modelling trying to fit the simulated activity concentrations with the observed low activity concentrations. However, due to the construction of the cost function, these have a small impact on the cost function. As a result, large differences in the release term can lead to only small differences in the cost function. The additional release should therefore not be interpreted too strictly, and the release profile obtained from the modified observations, showing a single robust release, is by far the most interesting feature. Note that the uncertainty bounds from the ensemble are helpful for carefully interpreting the additional releases.

The  $^{131m}\text{Xe}$  release profile is shown in Fig. 8(c). As for the  $^{133}\text{Xe}$  release profile, a significant release is predicted around 6–7 April. Note that this release is lower than for the  $^{133}\text{Xe}$  release profile. This is expected since at the time of the potential release, the  $^{131m}\text{Xe}$  inventory is still smaller than the  $^{133}\text{Xe}$  inventory due to the different cumulative fission yields<sup>23</sup>.

**Summary and conclusions.** Several weeks after the third announced nuclear test conducted by the Democratic People’s Republic of Korea, particular combinations of  $^{131m}\text{Xe}$  and  $^{133}\text{Xe}$  had been measured at two noble gas stations that are part of the International Monitoring System to verify compliance with the Comprehensive Nuclear-Test-Ban Treaty<sup>14</sup>. In this paper, the origin of this radioactive xenon has been assessed using long-range atmospheric transport and dispersion modelling. Unique meteorological ensemble data were obtained by running the latest version of the Ensemble Data Assimilation system of ECMWF for the period March–April 2013. Atmospheric transport and dispersion calculations were performed with Flexpart in backward mode to calculate the source-receptor sensitivity fields for 57 noble gas observations from the International Monitoring System. Inverse modelling was used to search for an optimal source term in each grid box of the domain that matched best with the observations (“best” being quantified by two alternative cost functions Eqs 3 and 4). Using the adjoint approach, there was no need to rerun the atmospheric transport model during the optimisation. The ensemble allowed to quantify the source location uncertainty and release uncertainty.

The source localisation of the  $^{133}\text{Xe}$  detections showed two distinct possible source regions: one at the Chinese-Mongolian border, and one around the Punggye-ri nuclear test site (Fig. 2). The results were robust for the two alternative cost functions that were used in this study. The correlation between the source-receptor-sensitivity fields and the observed activity concentrations (sometimes called the “Possible Source Regions” product) turned out to be poor in discriminating possible source regions from other regions when the timing of the release is not

specified (Fig. 1). The ensemble allowed to construct grid-pointwise probability maps such as Fig. 3. This provided additional valuable information compared to the deterministic result.

A particular challenge for the radioxenon verification part of the CTBT is the existence of a radioxenon background from civilian sources. The influence of the  $^{133}\text{Xe}$  background on the source location was assessed indirectly. The assessment suggests that the possible source region around the Chinese-Mongolian border should at least partly be attributed to the  $^{133}\text{Xe}$  background (partly, since the specific meteorology for that period and the noble gas station network configuration also influence the size of possible source regions). This was observed when (i) more weight was given to the observations having a high  $^{133}\text{Xe}$  activity concentration (Fig. 4b) and (ii) certain detections were changed to non-detections (Fig. 5). Furthermore, the results were consistent among different subsets of observations, such that the findings are not negatively influenced by a few samples containing  $^{133}\text{Xe}$  from sources other than the source of interest. This discussion highlights the importance of having a thorough understanding of the background, so that sound assumptions can be made when assessing the influence of the background. The parameter  $\alpha$  in Eq. 4 was kept constant in this study, but it can readily be extended to be station-specific or even sample-specific (representing the fact that different noble gas stations have different radioxenon backgrounds or samples have different minimum detectable concentrations).

The source localisation was repeated for the  $^{131m}\text{Xe}$  detections. The resulting possible source locations were similar to those for  $^{133}\text{Xe}$ . However, the area around the Punggye-ri nuclear test site was now the single most likely source region (Fig. 7).

When a delayed  $^{133}\text{Xe}$  release from the Punggye-ri nuclear test site is assumed, a sharp release around 6–7 April 2013 is observed (Fig. 8). The additional releases with large uncertainties according to the ensemble were partly removed by using the modified observations to minimise the effect of the background.

Currently, all four radioxenon isotopes are rarely (if at all) measured in a single sample by the International Monitoring System, making a routine discrimination of samples based on the discrimination line<sup>24</sup> not feasible. However, this technique might become useful in the future since the noble gas detection systems are further being improved, making it more likely that all four radioxenon isotopes will be measured simultaneously<sup>25,26</sup>. To date, the detection of radioxenon combined with atmospheric transport modelling provides most feasibly evidence to confirm the nuclear nature of a man-made explosion. The source localisation of both  $^{131m}\text{Xe}$  and  $^{133}\text{Xe}$  indicate that a delayed release from the Punggye-ri nuclear test can explain the observations, and the inverse modelling greatly confines possible source regions.

## Methods

**Radioxenon observations.** Radioxenon detections and non-detections made by the International Monitoring System were used for the inverse modelling. In order to confine the possible source regions of the radioxenon detections as much as possible, many observations should be used, since each detection or non-detection carries some information. However, detections can be contaminated by the radioxenon background from civilian emitters. It is challenging to estimate the correct radioxenon contribution from civilian emitters for every individual sample<sup>27,28</sup> (especially if not all civilian sources of radioxenon in the region are known). This highlights the importance of understanding the radioxenon background and reducing radioxenon emissions from civilian sources<sup>29</sup>. With the above considerations in mind, we selected all  $^{133}\text{Xe}$  observations taken between 5 April 2013 and 15 April 2013 for the IMS noble gas stations RN20, RN38, RN45 and RN58 (see Fig. 2 for their location). In total, 57 observations were used for the inverse modelling (Supplementary Information).

The IMS noble gas stations RN20 and RN45 are equipped with the Spalax system<sup>30,31</sup> and measure radioxenon activity concentrations every 24 h, whereas RN38 is equipped with the Sauna system<sup>32</sup> and measures radioxenon activity concentrations every 12 h. Finally, RN58 is equipped with the Arix system and measures radioxenon activity concentrations every 12 h<sup>33</sup>.

**ATM model and weather data.** Two sets of meteorological data were used. First, deterministic weather data of the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasts (ECMWF) (4 analyses per day were combined with forecasts having a 3 h lead time, so that meteorological data were available every 3 h). Second, unique meteorological data created by running the current latest version (cy43r1) of ECMWF's Ensemble Data Assimilation (EDA) system<sup>34</sup> for the period March–April 2013. This ensemble consisted of 26 independent lower-resolution 4D-Var assimilations, of which 25 use perturbed observations, sea-surface temperatures and model physics. More information on the type of perturbations can be found in another study<sup>34</sup>. The EDA cycles its own background error and covariance estimates, therefore it can be seen as a variational implementation of a perturbed observation Ensemble Kalman Filter. For each perturbed EDA member, the perturbations with respect to the ensemble mean were calculated. These perturbations were then added and subtracted from the unperturbed member, so that 50 perturbed members were obtained. Meteorological data were available every three hours, having horizontal grid spacings of 0.5° and having 137 non-uniform vertical levels with the top level at 0.01 hPa.

The Lagrangian particle model Flexpart<sup>35–37</sup> version 9.02 was used for the atmospheric transport and dispersion calculations. Flexpart was used in backward mode<sup>38</sup>, giving source-receptor-sensitivities as output. Radioactive decay was taken into account for both  $^{133}\text{Xe}$  and  $^{131m}\text{Xe}$ . The ECMWF Mars extraction software for Flexpart available at the Flexpart website<sup>39</sup> was used in modified form.

**Inverse modelling method.** Consider a vector of observed concentrations  $\mathbf{y}$ . Inverse modelling involves finding a source term  $\mathbf{x}(x, y, z, t)$  so that the following relation is true:

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \boldsymbol{\varepsilon} \quad (2)$$



Here,  $\mathbf{M}$  is the source-receptor-sensitivity matrix<sup>38,40</sup> and  $\varepsilon$  is the combined observation and model error. Forward modelling is often used to calculate  $\mathbf{M}\mathbf{x}$  after selecting an initial guess source term  $\mathbf{x}$ . Via an iterative process, the source term is then refined. However, with Flexpart, it is possible (and, if the source location is not known, more efficient) to perform a backward calculation for each observation used in the inverse modelling; the result is the source-receptor-sensitivity matrix  $\mathbf{M}$ , thereby avoiding the need to rerun the atmospheric transport model during the optimisation: only the source term  $\mathbf{x}(x, y, z, t)$  must be varied until Eq. 2 holds.

An exact match between simulations and observations is not possible since both the source-receptor-sensitivity matrix and the observations contain uncertainties. Instead, the disagreement between the observed activity concentrations  $\mathbf{y}$  and the simulated activity concentrations  $\mathbf{M}\mathbf{x}$  should be minimised while taking into account the uncertainty in the observations and the model. The disagreement is quantified by a cost function. Two cost functions were used in this study. A previous study<sup>41</sup> combined the mean square error (*mse*) and the correlation (*cor*) into a single likelihood weight. It was argued that *mse* is mainly sensitive to changes in source amount, while temporal *cor* is mainly sensitive to changes in arrival time and duration. We followed this reasoning for constructing the following cost function (note the normalisation of the *mse*):

$$\text{cost function}(\mathbf{x}) = \sum_{\text{samples}} (AC_{\text{obs}} - AC_{\text{sim}})^2 \cdot \left( \sum_{\text{samples}} (AC_{\text{obs}})^2 \right)^{-1} + 1 - \text{cor}(AC_{\text{obs}}, AC_{\text{sim}}) \quad (3)$$

where  $AC_{\text{obs}}$  and  $AC_{\text{sim}}$  are the observed and simulated activity concentration. An alternative cost function used in this study is the geometric variance (Eq. 4). This cost function was found to give the best results when applying inverse modelling to a large domain and noisy measurements<sup>42</sup>. Here, a parameter  $\alpha$  has been added representing the minimum detectable concentration, such that non-detections could be used in the inverse modelling:

$$\text{cost function}(\mathbf{x}) = \exp \left( \frac{1}{n} \sum_{\text{samples}} (\log(AC_{\text{obs}} + \alpha) - \log(AC_{\text{sim}} + \alpha))^2 \right) \quad (4)$$

with  $n = 57$  the number of observations and  $\log$  the natural logarithm. We assumed that the signal of interest originated from a single grid box source. Therefore, for each grid box in the lowest model level, we performed the inverse modelling and obtained a cost function value and an associated optimal source term. It was assumed that the most likely source locations are those grid boxes with the lowest cost function values.

The release is assumed to took place between 20 March 2013 and 15 April 2013, with release intervals of two days. The latter is sufficient to resolve the synoptic scale signal<sup>43</sup> and helps to regularise the problem. Although such a long release window complicates the inversion due to longer computation times and more unknowns in the inversion, it allows to identify remote sources.

The optimisation was done using the routine *nlinmb* from the R statistical software<sup>44</sup> that uses a quasi-Newton method for the minimisation with bounds on the release term ( $Q_{\text{min}} = 10^9$  Bq/day and  $Q_{\text{max}} = 10^{13}$  Bq/day were used unless otherwise mentioned).

**Data availability.** The data generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## References

1. ENSDF. ENSDF National Nuclear Data Center, Evaluated Nuclear Structure Data File. <http://www.nndc.bnl.gov>, Last accessed: 2018-02-12 (2016).
2. CTBTO. CTBTO Preparatory Commission. <http://www.ctbto.org>. Accessed: 2018-01-09 (2017).
3. Stocki, T. *et al.* Automated radionuclide monitoring for the Comprehensive Nuclear-Test-Ban Treaty in two distinctive locations: Ottawa and Tahiti. *Journal of Environmental Radioactivity* **80**, 305–326 (2005).
4. Hutchinson, M., Oh, H. & Chen, W.-H. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Information Fusion* **36**, 130–148 (2017).
5. Stohl, A. *et al.* 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* **12**, 2313–2343 (2012).
6. Winiarek, V., Bocquet, M., Saunier, O. & Mathieu, A. 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. *Journal of Geophysical Research: Atmospheres* **117** (2012).
7. Saunier, O. *et al.* An inverse modeling method to assess the source term of the Fukushima nuclear power plant accident using gamma dose rate observations. *Atmospheric Chemistry and Physics* **13**, 11403–11421 (2013).
8. Chai, T., Draxler, R. & Stein, A. Source term estimation using air concentration measurements and a Lagrangian dispersion model—experiments with pseudo and real cesium-137 observations from the Fukushima nuclear accident. *Atmospheric Environment* **106**, 241–251 (2015).
9. Stohl, A. *et al.* Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption. *Atmospheric Chemistry and Physics* **11**, 4333–4351 (2011).
10. Chazette, P. *et al.* Eyjafjallajökull ash concentrations derived from both lidar and modeling. *Journal of Geophysical Research: Atmospheres* **117** (2012).
11. Tichý, O. *et al.* Bayesian inverse modeling and source location of an unintended 131I release in Europe in the fall of 2011. *Atmospheric Chemistry and Physics* **17**, 12677–12696 (2017).
12. Saey, P. R. *et al.* A long distance measurement of radionuclide in Yellowknife, Canada, in late October 2006. *Geophysical Research Letters* **34** (2007).
13. Ringbom, A. *et al.* Measurements of radionuclide in ground level air in South Korea following the claimed nuclear test in North Korea on October 9, 2006. *Journal of Radioanalytical and Nuclear Chemistry* **282**, 773 (2009).
14. Ringbom, A. *et al.* Radionuclide detections in the CTBT international monitoring system likely related to the announced nuclear test in North Korea on February 12, 2013. *Journal of Environmental Radioactivity* **128**, 47–63 (2014).

15. Becker, A. *et al.* Global backtracking of anthropogenic radionuclides by means of a receptor oriented ensemble dispersion modelling system in support of Nuclear-Test-Ban Treaty verification. *Atmospheric Environment* **41**, 4520–4534 (2007).
16. Eslinger, P. W. & Schrom, B. T. Multi-detection events, probability density functions, and reduced location area. *Journal of Radioanalytical and Nuclear Chemistry* **307**, 1599–1605 (2016).
17. Yee, E. Automated computational inference engine for Bayesian source reconstruction: Application to some detections/non-detections made in the CTBT International Monitoring System. *Applied Mathematical Sciences* **11**, 1581–1618 (2017).
18. Hofman, R. & Seibert, P. Analysis of the April 2013 radioxenon detections based on formal inverse modeling (presented at the ATM Workshop in Stockholm, 2014).
19. De Meutter, P., Camps, J., Delcloo, A. & Termonia, P. Assessment of the announced North Korean nuclear test using long-range atmospheric transport and dispersion modelling. *Scientific Reports* **7**, 8762 (2017).
20. Issartel, J.-P. & Bayerel, J. Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmospheric Chemistry and Physics* **3**, 475–486 (2003).
21. Achim, P. *et al.* Characterization of Xe-133 global atmospheric background: Implications for the International Monitoring System of the Comprehensive Nuclear-Test-Ban Treaty. *Journal of Geophysical Research: Atmospheres* **121**, 4951–4966 (2016).
22. Generoso, S. *et al.* Seasonal variability of Xe-133 global atmospheric background: characterization and implications for the International Monitoring System of the Comprehensive nuclear-Test-Ban Treaty. *Journal of Geophysical Research: Atmospheres* (2018).
23. England, T. & Rider, B. *Evaluation and compilation of fission product yields 1993*. (Tech. Rep., Los Alamos National Lab., NM (United States), 1995).
24. Kalinowski, M. B. *et al.* Discrimination of nuclear explosions against civilian sources based on atmospheric xenon isotopic activity ratios. *Pure and Applied Geophysics* **167**, 517–539 (2010).
25. Ringbom, A. *et al.* The SAUNA III project (presented at the CTBT: Science and Technology conference in Vienna, 2017).
26. Topin, S. *et al.* SPALAX-new generation: Deployment, operation and performances (presented at the CTBT: Science and Technology conference in Vienna, 2017).
27. Eslinger, P. W. *et al.* International challenge to predict the impact of radioxenon releases from medical isotope production on a Comprehensive Nuclear-Test-Ban Treaty sampling station. *Journal of Environmental Radioactivity* **157**, 41–51 (2016).
28. Maurer, C. *et al.* 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* (2018).
29. Gueibe, C., Camps, J. & Van Der Meer, K. Xenon mitigation project—phase I: Adsorption materials. Tech. Rep., SCK-CEN-BLG-1099, ISSN 1379-2407 (2014).
30. Fontaine, J.-P., Pointurier, F., Blanchard, X. & Taffary, T. Atmospheric xenon radioactive isotope monitoring. *Journal of Environmental Radioactivity* **72**, 129–135 (2004).
31. Topin, S. *et al.* Spalax new generation: New process design for a more efficient xenon production system for the CTBT noble gas network. *Journal of Environmental Radioactivity* **149**, 43–50 (2015).
32. Ringbom, A., Larson, T., Axelsson, A., Elmgren, K. & Johansson, C. Sauna: a system for automatic sampling, processing, and analysis of radioactive xenon. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **508**, 542–553 (2003).
33. Dubasov, Y. V. *et al.* The automatic facility for measuring concentrations of radioactive xenon isotopes in the atmosphere. *Instruments and Experimental Techniques* **48**, 373–379 (2005).
34. Bonavita, M., Hólm, E., Isaksen, I. & Fisher, M. The evolution of the ECMWF hybrid data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **142**, 287–303 (2016).
35. Stohl, A., Hittenberger, M. & Wotawa, G. Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. *Atmospheric Environment* **32**, 4245–4264 (1998).
36. Stohl, A. & Thomson, D. J. A density correction for Lagrangian particle dispersion models. *Boundary-Layer Meteorology* **90**, 155–167 (1999).
37. Stohl, A., Forster, C., Frank, A., Seibert, P. & Wotawa, G. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmospheric Chemistry and Physics* **5**, 2461–2474 (2005).
38. Seibert, P. & Frank, A. Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. *Atmospheric Chemistry and Physics* **4**, 51–63 (2004).
39. Flexpart. Official Flexpart website. <http://www.flexpart.eu>. Last accessed: 2018-02-12 (2016).
40. Wotawa, G. *et al.* Atmospheric transport modelling in support of CTBT verification—overview and basic concepts. *Atmospheric Environment* **37**, 2529–2537 (2003).
41. Lucas, D. D., Simpson, M., Cameron-Smith, P. & Baskett, R. L. Bayesian inverse modeling of the atmospheric transport and emissions of a controlled tracer release from a nuclear power plant. *Atmospheric Chemistry and Physics* **17**, 13521–13543 (2017).
42. Cervone, G. & Franzese, P. Monte Carlo source detection of atmospheric emissions and error functions analysis. *Computers & Geosciences* **36**, 902–909 (2010).
43. De Meutter, P., Camps, J., Delcloo, A., Deconinck, B. & Termonia, P. Time resolution requirements for civilian radioxenon emission data for the CTBT verification regime. *Journal of Environmental Radioactivity* **182**, 117–127 (2018).
44. Richard, A. & Becker, O. S. version by Ray Brownrigg. Enhancements by Thomas P Minka, A. R. W. R. & Deckmyn., A. *maps: Draw Geographical Maps*. <https://CRAN.R-project.org/package=maps>. R package version 3.2.0 (2017).
45. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/> (2017).

## Acknowledgements

The authors would like to thank Dominique Lucas (ECMWF) and Massimo Bonavita (ECMWF) for their help with running the latest version of the EDA system at ECMWF, Martin Leutbecher (ECMWF) for interesting discussions on the use of the EDA system, and the CTBTO Preparatory Commission for making available the IMS radioxenon data. The two anonymous reviewers are thanked for their constructive remarks. One of the authors (P.D.M.) acknowledges funding from Engie under contract number JUR2015-28-00.

## Author Contributions

P.D.M. and A.D. prepared the meteorological data and performed the Flexpart simulations; P.D.M. and J.C. analysed the output data; all authors contributed to the manuscript; P.T. supervised the research.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-28403-z>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018