Detection of nuclear testing from surface concentration measurements: Analysis of radioxenon from the February 2013 underground test in North Korea

May 21, 2017

R. J. Kurzeja, R. L Buckley, D. W. Werth, and S. R. Chiswell

Abstract:

A method is outlined and tested to detect and locate low level nuclear or chemical sources from time series of concentration measurements. The method uses a mesoscale atmospheric model to simulate concentration signatures from known or suspected sources at a receptor which are then regressed successively against segments of the measurement series to create time series of metrics that measure the goodness of fit between the signatures and the time series.

The method was applied to radioxenon data from the Comprehensive Test Ban Treaty (CTBT) collection site in Ussuriysk, Russia (RN58) after the Democratic People's Republic of Korea (North Korea) underground nuclear test on February 12, 2013 near Punggye. The metrics were found to be a good screening tool to locate data segments with a strong likelihood of origin from Punggye, especially when multiplied together to a determine the joint probability.

The metrics were used to find the probability that activity measured in February and April of 2013 originated from the underground test. A more detailed analysis of an RN58 data segment from April 3-5, 2013 was carried out for a grid of locations around Punggye and identified Punggye as the most likely point of origin.

The analysis of radioxenon isotopes at the RN58 CTBT site supports the conclusion that radioxenon was emitted from the test site continuously in April after the February 12 test and was detected intermittently at RN58, depending on the wind direction.

Extension of the method to higher resolution data sets, overlapping sources, and time-varying sources is discussed along with research that will refine the technique. We also discuss a path to evaluate uncertainty of the calculated probabilities.

1. Introduction

Radiation measurement networks have been installed to monitor normal and accidental atmospheric releases from nuclear reactors (e.g. SPEEDI , Chino et al., 1993), nuclear isotope production facilities (Wotawa et al., 2010) and also from nuclear incidents and tests, e.g., Le Petit et al. (2008), CTBT(2013). Nuclear materials production and weapons testing manifests itself in a range of environmental signatures, including seismic, infrasound, water, thermal, industrial, and atmospheric (Auer and Prior, 2014). Often activity is discovered by one means and then confirmed with other information. While monitoring of atmospheric signals from nuclear power plants and isotope production is straightforward because their locations are known and their source strengths can be estimated, the detection of undisclosed sources is hampered by ignorance of the time and location of the release.

The detection of underground tests is complicated because diffusion from the test location to the surface depends on the depth of the detonation, the geological properties of the area, and the level of control exercised during the test (Kalinowski, 2011). Dubasov (2010) noted that ~50% of Soviet underground testing experienced measurable releases of noble gases to the atmosphere. Carrigan and Sun (2014) discuss three phases of noble gas release from an underground test. Leaks to the surface may occur in the hours after a detonation driven by pressure in the detonation cavity. Emission in the days to weeks after the detonation may continue from multiphase convection due to cavity heat. Seepage from the detonation is possible months afterward aided by atmospheric pressure-induced pumping (Nilson et al., 1991).

The Comprehensive Test Ban Treaty (CTBT) verification program (Hoffman et al., 2000) includes seismic, hydro-acoustic, infrasound and radionuclide measurements to monitor underground nuclear testing. Measured particulates and gases include isotopes of xenon. However, because of the sparseness of the network, traces from specific upwind locations are episodic at each CTBT site with time gaps of weeks between events. Furthermore, often significant quantities of only one isotope are present which prohibits the use of isotopic ratios to confirm the origin.

Data from the CTBT network have been used to document underground tests conducted by the Democratic People's Republic of Korea (DPRK) in 2006, 2009 (Murphy et al., 2011, De Geer, 2012, Wright, 2013), and in 2013 (Ringbom et al. ,2014, hereafter referred to as Ring14). These tests were announced by the DPRK and confirmed by seismic data and atmospheric radionuclide signatures downwind of the test (Ring14). Since the tests were conducted underground, particles were probably not released directly into the atmosphere and confirmation of the tests was achieved with measurements of radioxenon isotopes which are not reactive and therefore more difficult to contain underground.

The CTBT atmospheric radioxenon analysis relies on notification or other indications that a test has occurred to direct attention to a particular time period for further analysis. In particular, Ring14 focused on events with enhanced radioxenon ^{131m}Xe activity on Apr 12/13, 2013 at RN58 (Ussuriysk, Russia) and in enhanced ^{133}Xe and ^{131m}Xe activity on Apr 7/8, 2013 at RN38 data (Takasaki, Japan). Further evidence was obtained from back trajectories from RN38 and RN58 to the suspected test location (Punggye, NK). The possibility of other radioxenon sources, e.g., nuclear power reactors and isotope production facilities, was considered.

The use of back trajectories from receptor locations is a common way to determine the origin of gases or particulates, Draxler (1987). Source-receptor models require back-trajectories from multiple receptors and continuous sampling to map sources on a regional or global grid, (e.g. Stohl et al. 1998, Becker et al., 2007, Siebert and Frank, 2004). These methods relate sources to receptors with a matrix of transport fields derived from global wind forecast models. Uncertainty in transport and measurements appears as an enlarged source region rather than as a reduced probability of origin from a particular location. In addition, these methods are unsuitable for searching data sets for weak signals or puff releases.

This paper presents a new method to detect and characterize weak atmospheric radionuclide signals imbedded in local and global background data. A source is inserted into an atmospheric model and the concentration time history, i.e., the signature, calculated at a receptor. The signature is important because it is a unique fingerprint of the source and can be used to (a),filter the observed time series to identify data segments most likely to be from the source, and (b), find the probability of origin from a particular location. An objective determination of probability with uncertainty is desirable for strong signals but essential for weak signals, which can either go unnoticed or given low priorit for lack of convincing evidence. Our new method differs from existing approaches in three important ways. First, observations at the receptor and model transport fields are used to determine the probability of origin from a source. Second, the method can be applied to a single receptor and an arbitrary source. Third, it can be extended to multiple isotopes and sources or linked to other observed phenomena associated with the upwind source.

We will demonstrate the method by investigating the DPRK test of February 12, 2013 with ^{133}Xe and ^{131m}Xe data presented by Ring14 . The RN58 site is a good choice for the analysis because it is close to the DPRK test facility and relatively unaffected by other radioxenon sources. The time period of April 2013 will be studied in detail because Ring14 found confirmatory evidence for the Feb 12 test in the RN58 and RN38 data on Apr 12/13 and Apr 8/9, respectively. After testing the method for these two events we search for other radioxenon events in the data series of likely origin from the Feb 12 test. The temporal resolution of the radioxenon data (12 hours) is too coarse for a detailed comparison with the model signature, however, gross features of the signature can be compared with the data.

Section 2 describes the analysis method while Section 3 describes the mesoscale and transport models used in the study. Sections 4-5 apply the method to radioxenon CTBT data at RN58. Sections 6 and 7 discuss the April 7/8 event at RN38 and introduce the arrival time metric. Sections 8 and 9 apply the method to other events in 2013 while Section 10 considers the possibility that Vladivostok is the source of the radioxenon. Section 11 examines the Apr 3/5 event at RN58 in more detail with a grid of sources around Punggye, while Sections 12 and 13 summarize the results.

2. The method

Fundamental to the analysis presented in this paper is the model-derived signature whose first function is to filter the data series to find events of interest. Fig. 1 shows a model signature along with measured ^{133}Xe activity from RN58 in Ussuriysk, Russia in April of 2013, to be discussed in greater detail in Section 4. The signature is the model-predicted activity at the receptor, RN58, for an assumed constant unit source at Punggye, believed to be the location of the underground test. (A variable source could be specified but is not justified by the available information). Null-valued wings have been added to each side of the model prediction to define its beginning and ending. Under ideal conditions the signature is assumed to uniquely identify a source from measurements at a receptor. However, in actual situations agreement between the model-derived signature and the observations is limited by model and data accuracy, data temporal resolution, the amount and accuracy of meteorological input data, and random variability in the model simulation. These conditions, together with a small source strength imply that the plume of interest may become indistinguishable from plumes from other sources.

Formally, we define a signature, $f_{i,N+M}$, to be the combination of a time sequence of N model-simulated non-zero concentrations at the receptor $[g_1, g_2, \ldots g_N]$, plus M/2 additional null values added before and after the non-zero values beginning at model point i, i.e.,

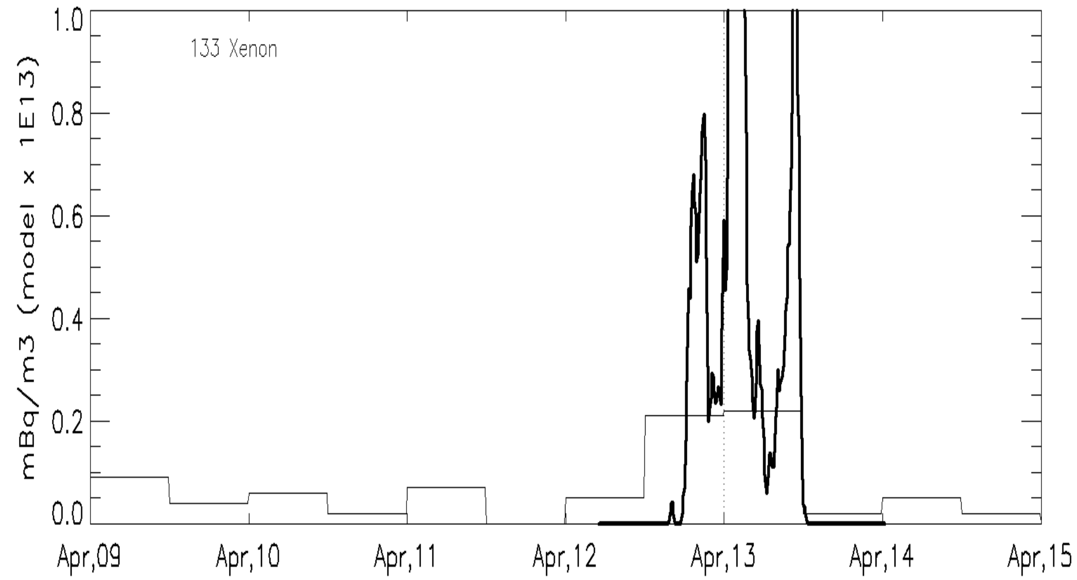

Fig. 1: ^{131}Xe acitivty (thin lines) and model signature (heavy line). The observed center of mass (*cm*) occurs at 00:00 UTC on April 13, while the modeled *cm* occurs at 02:24 UTC on April 13.

$f_{i,N+M} = [0, 0, \ldots, 0, g_1, g_2 \ldots g_N, 0, 0 \ldots 0]$.

Signature definition is important because it determines the search properties of the filter, but it can be somewhat arbitrary and an objective method for signature selection is desirable. In this study the non-zero edges of the model signature were chosen to be the times which encompass 90% of the plume's mass. These edges define the non-zero signature width (g_1 to g_N). The width of the wings was set at 0.7× the width of the non-zero signature. A value of 0.7 was selected by trial and error. A signature is shown in Fig. 1 for the Apr 12/13.

Complications arise when the simulated concentration is not approximately Gaussian in shape. For example, Fig. 2 shows model-predicted and measured activity at RN58 in Ussiriysk, Russia for April 2013. A double maximum can be seen on April 5-7 and a signature could be chosen that includes one or both maxima. A signature choice with both peaks might be overly restrictive and be very sensitive to small uncertainties in peak separation. On the other hand, a single-maximum signature would admit a greater number of events of interest. Neither choice is right or wrong but simply define different search patterns. Section 9 considers a case with alterative choices for the signature.

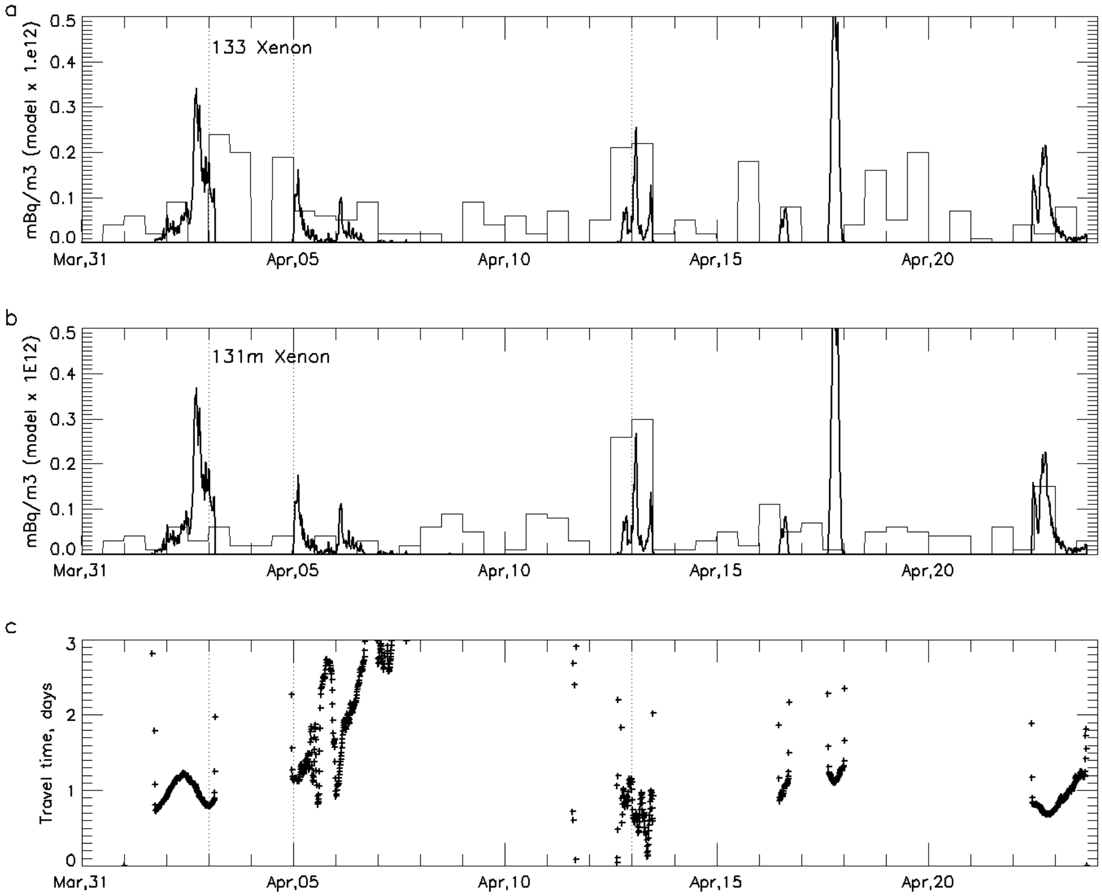

Fig 2. ^{133}Xe (a) and ^{131m}Xe (b) activity measured at RN58 from March 31-April 23, 2013. The thick line is the model simulated concentration. The lower panel (c) shows the model travel times from Punggye to RN58 in days. Vertical dotted lines are for time reference.

Figs. 1 illustrates the difference in resolution between the model simulation (10 minutes) and the CTBT data (12 hours) and explains why a detailed match of signature with data is not possible with the CTBT data. However, to illustrate the potential benefit of higher resolution data we have chosen to convert the 12-hr CTBT averages into 10-minute segments of constant value, rather than to average the model data over 12 hours.

The signature is used to create a time series of metrics by advancing the signature to each time, t, in the data series and then calculating the amplitude A(t) with a least-squares fit of the signature to the measurement segment, C_t, $C_{t+1,}$ $C_{t+2,}$ … C_{t+N+M}, i.e.,

$$A(t) \sim \min (\sum\{ \quad [C_t - A(t)x0]^2 + \ldots + [C_{t+M/2} - A(t)x0)]^2 + \ldots.$$

$$[C_{t+M/2+1} - A(t)xg_1]^2 + \ldots [C_{t+M/2+N} - A(t)xg_N]^2 + \ldots$$

$$[C_{t+M/2+N+1\,-}\, A(t)x0] + \ldots [C_{M+N} - A(t)x0]^2 \quad]$$

This process produces a time series of values defined with respect to the beginning point of the measurement segment. A more appropriate time stamp is with respect to the signature center of mass. Accordingly, we defined a new variable $A(t^*) = A(t+\Delta t)$, where Δt is the time from the beginning of the signature to the signature center of mass.

The goodness of fit between the signature and the data segment is characterized in terms of two amplitude metrics and two shape metrics. The amplitude metrics are the amplitude $m_A(t)$ and normalized amplitude $m_a(t)$. The shape metrics are the normalized error variance $m_v(t)$ and the Pearson correlation coefficient $m_p(t)$.

$m_a(t) = m_A(t)/(10 \times \text{median amplitude of the entire data set})$

$m_v(t) = 1.-2*\sum(f(t)-C(t))^2/[\sum C(t)^2+\sum f(t)^2]$, (sum over the signature length)

$m_p(t) = \sum(x_i - \underline{x})(y_i-\underline{y})/\surd\ [\sum(x_i - \underline{x})^2 \sum (y_i - \underline{y})^2]$

where averages are underlined variables.

The metrics are normalized to permit comparison between different scenarios. The factor of 10 in the normalized amplitude ensures that the largest metric for the RN38 data set is approximately equal to 1 and thus comparable to the other normalized metrics. The error variance is normalized by the average of the signature and the measurements over the signature length. The Pearson correlation coefficient is independent of amplitude and needs no normalization.

Weak signals by definition have low amplitude and it is desirable to also use metrics that don't depend on amplitude. Accordingly, we define amplitude and shape metrics as follows, which we use in the in the remainder of this paper.

amplitude metric = m_a,

shape metric = $m_v \times m_p$,

amplitude/shape metric = $m_a \times m_v \times m_p$

The method described here is a technique in pattern recognition called template matching (see e.g., Jain et al, 2000). Other methods are available to detect signals in a time series, for example, wavelet analysis (Torrence and Compo, 1998). Our approach is analogous to wavelet analysis but with a single basis function - the model-derived signature. The signature can be used to screen a data sets to find events of interest. The signature can also be used to estimate the probability of origin from a particular location. Probability and uncertainty estimates are essential to study weak signals, particularly if supplementary information is unavailable. Probability is also valuable to show the relative value of individual vs. multiple isotopes.

As will be discussed in Section 5, the continuous time series of metrics are converted to single values (metric events) by finding local maxima. Metric events can then be used for probability estimates. The metrics express the agreement between the source signature and the data segment and we assume they are proportional to the probability P of an exact match between the measured activity segment and a source's signature, given a release from that source, i.e, P = α x m, where m is the metric value and α is an empirical constant. For example, a Pearson correlation coefficient of 1.0 implies a 100% probability of an exact match of signature with a data segment ($\alpha = 1$), while $m_p = 0$ implies a zero probability of an exact match. An exact match between signature and data segment implies a 100% probability the data in the time segment originated from the assumed source.

While a single metric value does not appear to be a robust measure of a probability it should be noted data segments and model trajectories are averages during a trajectory of random turbulent events over many time scales. Time scales in atmospheric motion range from minutes or hours for convective motions to days for mesoscale events, to weeks for synoptic systems. Thus, model signature and data segments of 24 hours or more are averages of some but not all of the relevant turbulent scales.

The value of α could be determined empirically from a data base of observed scenarios. Furthermore, α may not be a constant but depend on variables such as the signal to noise ratio. In this study α is assumed to be 1.0. Also, as noted above, variability and uncertainty in the model and data may diminish the size of the metrics to a point where metrics from other sources are comparable in magnitude to the metrics from the proposed source. This will reduce the ability of the signature to uniquely identify the source.

To avoid these two problems we use the metrics to determine relative probability of origin by defining a time period of interest around an observed event, and then comparing the relative sizes of the metric of interest to the other metrics in the interval to find the relative probability of origin. Thus, the probability that a particular event originated from Punggey is proportional to the relative size of its metric compared to other metrics and inversely proportional to the number of events in the interval. For example, if there are few other events in the interval, with small values, then the probability originated from Punggye becomes more likely. This might be the case for air that had previously passed over the Pacific Ocean before passing over Punggye and then to RN58. On the other hand if an air mass first passes over industrialized regions in China, South Korea, or Japan before passing over Punggye the time interval of interest might contain many metrics with large values, which implies a lower probability that the event originated from Punggye.

The concept is formalized with Bayes rule. Consider a time period of interest, T, with L events in addition to the one suspected to be of Punggye origin. A nominal length of 7 days is selected for T because it is a typical time between return plumes from an upwind direction. Then we use Bayes rule as follows,

P_0 = probability of a source at Punggye, location $L_{0,}$

$P(m_0|L_0) = m_0$ = conditional probability of an exact match between the signature and the data segment , given a source at location L_0.

$P(L_0|m_0)$ = conditional probability of a source at Punggye L_0, given a metric of value m_0.

The probabilities for the L other observed events in the time interval of interest are defined analogously.

Then, according to Bayes rule the probability that a particular observed event with metric value of m_0 originated from Punggye (exact match of signature with data segment) is given by,

$$P(L_0|m_0) = \frac{P(m_0|L_0) \times P_0}{P(m_0|L_0) x P_0 + \{P(m_1|L_1) x P_1 + P(m_2|L_2) x P_2 + \ldots P(m_L|L_L) x P_L\} \quad x \quad 7\text{days}/T(\text{days})} \qquad \text{Eq 1}$$

The probabilities $P_{0 \ldots L}$ are termed prior probabilities and must be estimated. Some information is available. The probability of a radioxenon release from nuclear power reactors and radioisotope production facilities can be estimated from published reports, e.g., Kalinowski and Tuma, 2009, Riechmann and Kalinowski, 2008 and Wotawa et al, 2010. In addition, xenon releases from underground tests can be inferred from seismic data, public statements and satellite surveillance. Release from underground tests may also depend on fluctuations in the atmospheric pressure which draw radioxenon from underground fissures (Nilson et al., 1991). For simplicity the present analysis assumes that all prior probabilities equal 1.

The factor of 7 days/T(days) in the denominator converts the result to a standard 7 day interval. The probability $P(L_0|m_0)$ depends inversely on the number of events in the interval T which in turn depends on the length of the interval T. More (fewer) events in the interval will tend to reduce (increase) the probability and the factor adjusts this number to a 7-day period. Admittedly, the choice of the standard interval has an important effect on the calculated probability and 7 days may seem arbitrarily small. In fact, however, the decrease in the probability with interval length simply reflects the fact that as the test event is compared with more and more events (T increases) its uniqueness will diminish. A good analogy is a DNA identity test with a degraded DNA sample. The test DNA sample may be unique in a town with a small population but will lose its uniqueness when evaluated against a larger population, e.g., a state or country. The use of an arbitrary interval of 7 days is simply a way of standardizing the results so that individual events can be compared. A 7-day period is also the approximate plume return time at a receptor for a typical source location. Section 5 discusses how the event of interest was found for an actual case.

While the prior probabilities in Eq 1 have single values, the conditional probabilities are actually probability distributions. These distributions can be assumed to be Gaussian and defined by a mean and a standard deviation. In this study we consider the mean values only. Additional actual scenarios or model simulations are required to define the probability distributions, which is beyond the scope of the present work.

3. Mesoscale model description

Chemical signatures (time-dependent concentrations) were calculated at the RN58 CTBT monitoring station in Ussuriysk, Russia and at RN 38 (Takasaki, Japan) for February 2-21, and Apr 1-24 of 2013 with a mesoscale dynamic model and a Lagrangian particle transport model. The Regional Atmospheric

Modeling System (RAMS) (Pielke et al., 1992) produced a time-history of winds and turbulence on a 3-dimensional grid over Southeast Asia, which were then input to the Lagrangian model (HYSPLIT) to find the downwind concentration distribution. A constant unit source term from a Punggye release was assumed, but as discussed in Section 2, the inferred source strength will vary from event to event.

RAMS has been applied extensively to reproduce meteorological simulations on a range of scales (e.g. Cotton et al., 2002), and is ideal for this research. The modeling domain encompasses Japan and the Korean peninsula, centered over the Sea of Japan. A nested grid system centered at 39 °N and 133.5 °E was applied using a coarse grid (spanning 110 °E to 155 °E and 22 °N to 55 °N) with 30 km horizontal grid spacing, and a fine grid (spanning 119 °E to 148 °E and 28 °N to 50 °N) with 10 km horizontal grid spacing. The vertical grid uses 40 levels starting at 30 m for the lowest level, and increases 15% for each successive level until a maximum of 500 m is reached to a model top exceeding 12000 m. Model parameterizations used include the Mellor-Yamada planetary boundary layer scheme (Mellor and Yamada, 1982) to calculate turbulent diffusion, the Harrington radiation scheme (Harrington, 1997), and the LEAF-2 land surface scheme (Walko et al., 2000). Surface conditions are initialized using a 30 second topographic field to resolve many of the interactions between orography and airflow. Boundary conditions are supplied using the Global Forecast System (GFS, Environmental Modeling Center, 2003) with 0.5° horizontal resolution and at 3-hr intervals. Meteorological data for this study were generated at 10 minute intervals from 00 UTC, April 01 to 00 UTC, April 16, 2013.

The RAMS output (winds, turbulence, temperature, moisture) was subsequently used as input to the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT, Draxler and Hess, 1998) dispersion model. HYSPLIT has been used in numerous applications (e.g. Becker et al., 2007; Stunder et al., 2007) and is ideal for simulating the large-scale dispersion of emitted tracer. This Lagrangian model simulates the release of a large number of 'particles', recalculating at each time step the position of each one according to both a resolved wind field (from the RAMS model), as well as a dispersive term derived from the RAMS-simulated turbulent kinetic energy. Concentrations from HYSPLIT at three levels (0 – 50, 50 – 100, and 100 – 150 m) were generated for 10 minute averages assuming known nuclear power plants at Wolsong, Yonggwang, Ulchin, and Kori in South Korea, Ohi in Japan, as well as a medical isotope production (MIP) facility also located in Japan (near Toyko, Achim, 2013; Yonezawa, 2012). Although not currently operational, the Ohi facility was active during the period of interest (February to April 2013). In addition, a hypothetical seventh source is located in Punggye, North Korea, along with surrogate sources near Punggye to illustrate uncertainty associated with source location.

For this study, a continuous unit release (1 Ci/hr) of material was used in all cases. Radioactive decay was determined from the isotopic half-life and the travel time from the source to the receptor. The travel time was found from simulations with no radioactive decay, and a second with a half-life of 9 hours.

4. CTBT data

Ring14 discussed the experimental techniques employed by the CTBT network and focused in particular on radioxenon data from the RN38 station in Takasaki, Japan (36.3N, 139.08E) and the RN58 station in

Ussuriysk, Russia (43.7N. 131.9E) after the DPRK underground test of February 12, 2013. The locations of the RN38 and RN58 sites along with the DPRK test site at Punggye (41.27N, 129.1E) are shown in Fig. 3. Nuclear power reactors operational at this time in South Korea and Japan are also indicated in the figure.

Ring14 concluded that three measurements at RN38, (beginning times of April 7, 19 UTC, April 8, 07 UTC, April 8, 19 UTC), and two measurements from RN58, (beginning times of April 12, 12 UTC and April 13,00 UTC) contained ^{133}Xe and ^{131m}Xe activity significantly above background. They found that the isotopic ratios were consistent with a fission explosion and attributed these high concentrations to the DPRK test of February 12. However, they noted that the ^{131m}Xe /^{133}Xe ratio at RN58 exceeded the ratio at RN38 by approximately a factor of 8, which is greater than their estimated maximum experimental error of 50%. They raised the possibility of different fractionation history to explain the different isotopic ratios inferred from the two data sets. Ring14 also linked the 5 samples to the DPRK test by means of back trajectories that passed over the underground test location at Punggye. Thus, Ring14's conclusion that radioxenon measured in April, 2013 at RN 38 and RN 58 originated from Punggye is based on the NPRK announcement, seismic data, anomalously large concentrations, the isotopic ratio of ^{131m}Xe /^{133}Xe and on back trajectories that associate the samples with the Punggye test site. Ring14 considered the possibility that the effluent may have originated from locations in northern China, but assigned it a low probability because no known radioxenon sources are located there.

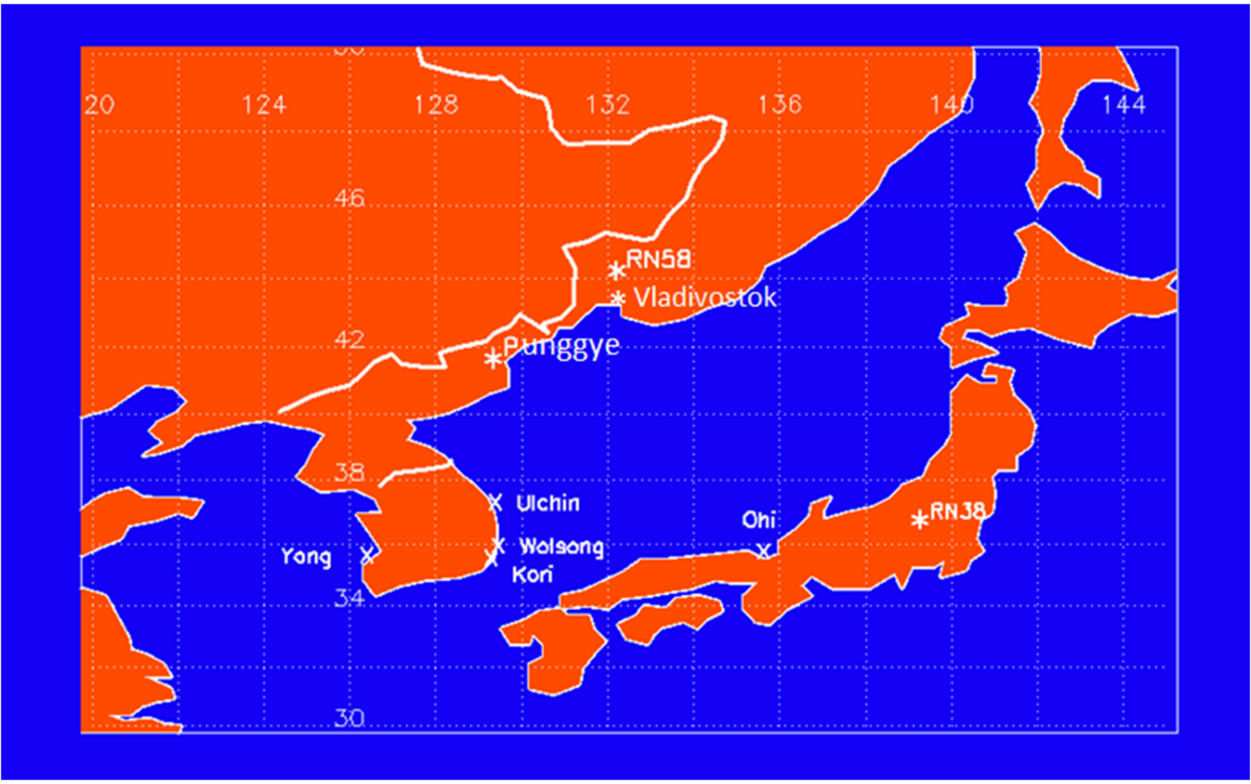

Fig. 3: Map showing the DPRK test site, CTBT monitoring locations RN38 and RN58, and commercial

nuclear power reactors in Japan and South Korea, and a potential radioxenon source at Vladivostok.

5. The April 12/13 event at RN58

This study focuses on radioxenon data measured at the Russian CTBT site RN58 which is ~400 km northeast of Punggye, in an area relatively unaffected by other radioxenon sources. Even though RN58 is close to the test site, its background radioxenon concentrations are ~1/2 those measured at RN38 in Japan. This implies that the data from RN58 can be analyzed without the need to consider overlap from other radioxenon sources as a first approximation. The ^{133}Xe and ^{131m}Xe activity from RN58 are shown in Fig. 2. The data were obtained from a scanned and digitized copy of Ring14's, Fig. 6 and some uncertainty in the amplitudes of the data points must be assumed. Since a continuous data record is necessary to match observations with model-derived signatures, missing RN58 values were replaced with randomly-chosen values from the two proceeding and two following data points. The global background radioxenon activity was subtracted from the data before processing. The global background is defined as the activity from distant sources (> 1000 km). It is characterized by low activities and long temporal periods. The background activity was selected to be the smallest 5-day running mean in the RN58 data for the period between February and May 2013.

Also shown in Fig. 2 is the model-simulated concentration with ten-minute resolution from a continuous unit release from Punggye. The model-predicted concentrations for the two isotopes differ slightly from one another because of differing half-lives of the two isotopes, i.e. 5.2 days for ^{133}Xe and 11.8 days for ^{131m}Xe. Seven distinct model signatures are seen in the simulation of ^{133}Xe. Four of these signatures correspond to enhanced measured activities, including the large activity seen on April 12-13, previously analyzed by Ring14, as well as events on April 3 and April 5-6. Model events on April 17 and 22 have no counterparts in the measured ^{133}Xe activities. The ^{131m}Xe simulation shows the same 7 events, but enhanced measured activities are seen only on April 12/13 and April 23. Low levels of ^{131m}Xe activity compared with ^{133}Xe are expected in early April and higher levels in late April because of its low initial concentration, and longer half-life, respectively.

The bottom panel of Fig. 2 shows the smoothed travel time from Punggye to RN58. The travel time was calculated from the ratio of the model predicted concentrations for an inert tracer to one with a 9-hour half-life. The travel time for the April 5-7 plume is much longer and more variable than for the other plumes, which suggests more uncertainty in the simulation.

The metric time series for the ^{133}Xe data set is shown in Fig. 4. The top panel shows the measured activity for the period Feb 5 to May 15 and model signatures for the period after April 2. Panel 4b shows the calculated amplitude while panels 4c to 4e show time series of metrics. We assume the metrics are independent and determine the composite metric, the joint probability, by multiplying them together (shown in Fig. 4f). The composite metric is more selective than the other metrics. For example, Fig 4f shows a reduction in the number of significant peaks comparable in size to the Apr 12/13 metric to approximately 10.

The composite metrics shown in Fig. 4f include the normalized amplitude (Fig. 4c) . As noted in Section 2 metrics obtained with the normalized amplitude are referred to as ‘amplitude/shape’ metrics. For the important case of weak signals the normalized amplitude will be small and subsequent sections also present results with the normalized amplitude excluded from the composite metric, i.e., ‘shape’ metrics.

As can be seen from the Fig. 4, the individual ^{133}Xe metrics are not particularly selective when viewed separately. In particular the error variance and Pearson correlation metrics are limited by the temporal resolution of the measurements. The advantages of better resolution will be discussed in Section 6 with respect to the RN38 data.

The analysis shown in Fig. 4 for ^{133}Xe was repeated but for ^{131m}Xe and the composite metric is shown in Fig. 5b along with the composite metric for ^{133}Xe (Fig 5a, identical to Fig 4f) . Fig. 5b shows a much larger composite metric on April 12-13 compared with other events in the period than found from ^{133}Xe activity. Since the two radioxenon isotope measurements are independent of each other, their metrics can be multiplied to produce a compound metric, which is shown in Fig. 5c. The compound metric for April 12-13 is seen to be ~10 times larger than the next largest metric. This result illustrates the better selection power of two isotopes compared with either isotope considered separately.

It is of interest to see how a selection of another signature affects the results. Accordingly, the analysis was repeated with the April 2/3 model signature instead of the April 12/13 signature. The results are shown in Fig. 6. The metrics in Figs. 5 and 6 are similar but note the relatively larger metrics obtained with the Apr 2/3 signature for the Apr 3 period compared with the Apr 12/13 signature. This can be interpreted as either evidence in support of a release from Punggye on Apr 2/3, or, as a result of an Apr 3 signature that is more typical of plumes reaching RN58 from various sources than the Apr 12/13 signature.

Measurement uncertainty was studied by repeating the analysis but with data sets modified to include measurement uncertainty (0.05 mBq/m3) based on Ring14’s Table 1. A random number generator was used to create three sequences of numbers with a normal distribution, and a standard deviation equal to 0.05 mBq/m^3 with a mean value of zero. These sequences were then added to the original data to produce three modified sets of radioxenon data that were analyzed as before. The calculated metrics were then compared with the original set. Relatively little effect was seen in the ^{133}Xe data because the data values generally exceed the uncertainty. The effect on the ^{131m}Xe data was greater and tended to smooth the metric time series without changing the general pattern.

The metric time series in Figs 4-6 are of limited value in statistical analysis, e.g., Eq. 1, where single-valued measures of agreement are needed. Single-valued metrics, i.e., event metrics, were extracted from the metric time series in a three-step process. First, visual inspection of the time series near an anomaly was used to define the domain of interest as described in the development of Eq. 1. The domain of interest includes the observed event and extends between the last and next predicted plumes at the receptor. The domain of interest is narrow enough to exclude other possible plumes from Punggye at the receptor. The domain of interest for the Apr 12/13 event is shown in Fig. 5 which shows that the number of events in the domain of interest varies between isotopes.

Event metrics are bounded by points where the time series becomes zero which occurs when the Pearson correlation coefficient becomes negative. The next step finds the time of the metric event from the concentration-weighted time-average between the zero points. The final step sets the value of the metric event to the value of the metric time series at the metric's time.

The above definition of the domain of interest is necessary in order to exclude all but one Punggye plume at the receptor. If, however, we assumed a puff release then a larger domain of fixed width, e.g., one month, could be chosen. The factor ' 7 days/T(days)' in Eq 1 converts the time period of interest to a standard 7-day period.

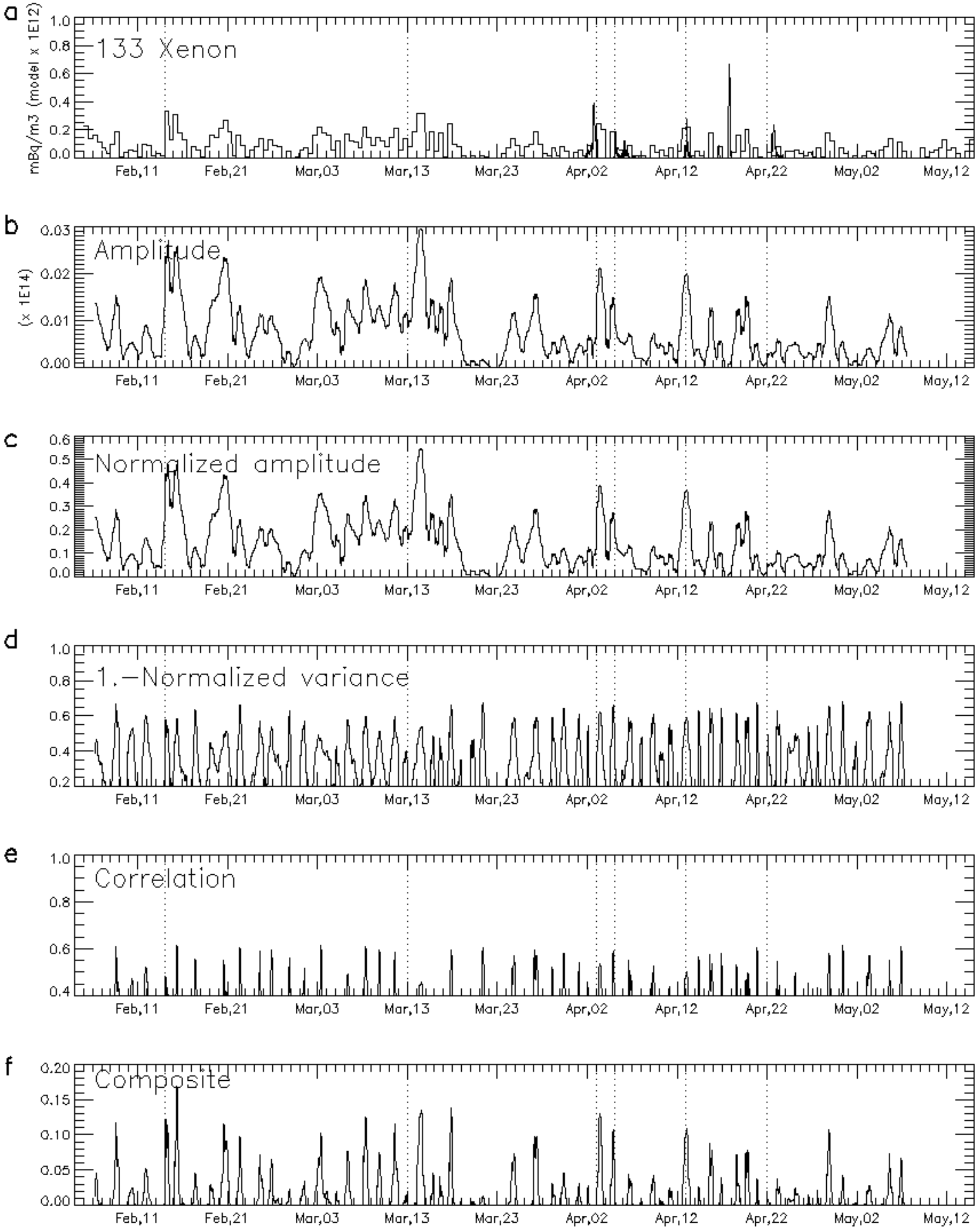

Fig. 4: ^{133}Xe data, model signature (April 12/13) and metrics. a, ^{133}Xe data and model signatures, b, amplitude metrics, c, normalized amplitude metrics, d, normalized variance metrics, e, Pearson correlation coefficient metrics, and f, composite shape metrics. The vertical dotted lines are for visual reference.

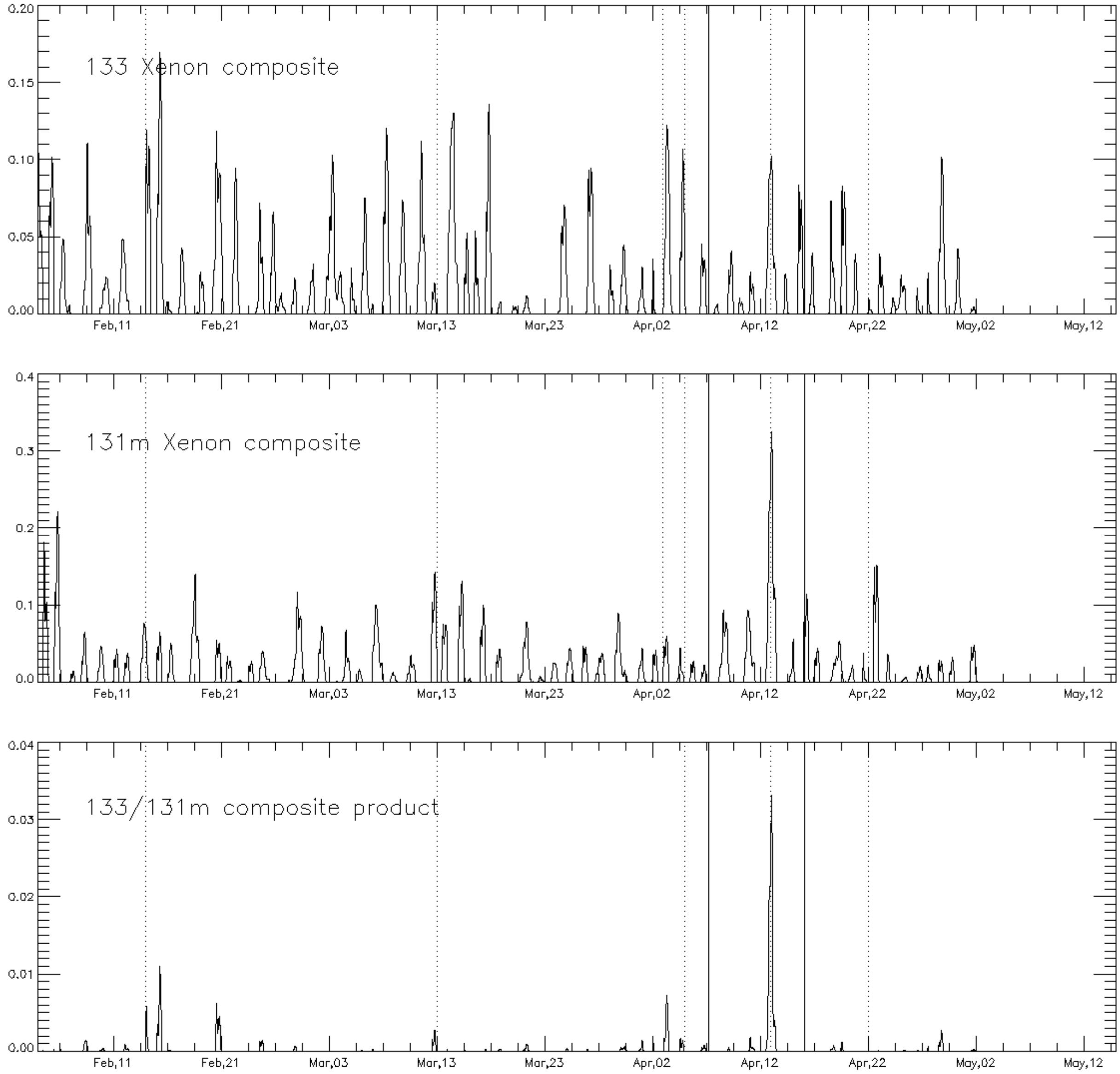

Fig. 5: Composite amplitude/shape metrics for ^{133}Xe (a), ^{131m}Xe (b) , and for their product (c). The Apr12/13 signature has been used. The vertical solid lines indicated the boundaries of the domain of interest around the Apr 12/13 event. The dotted lines are for visual reference.

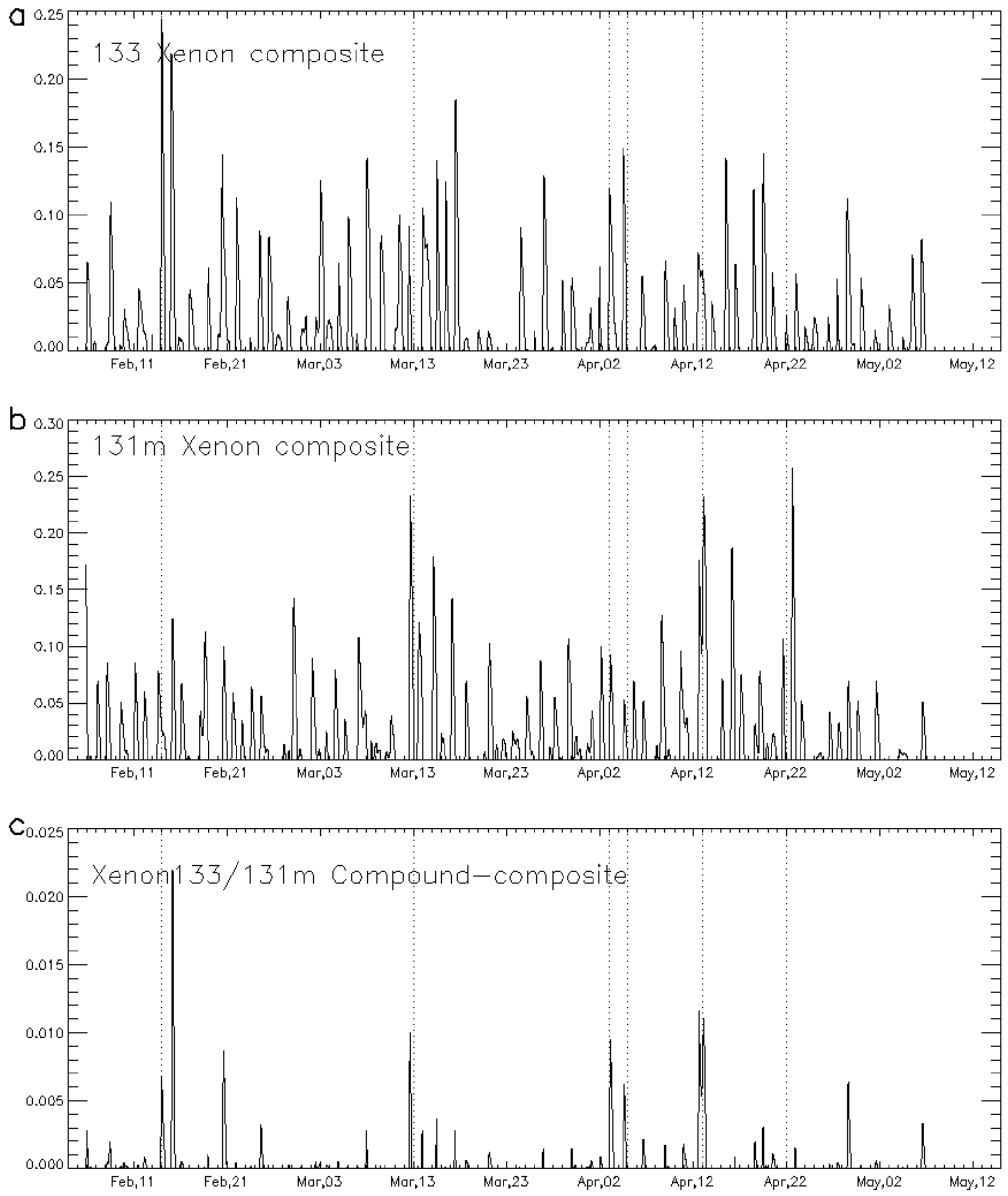

Fig. 6: As in Fig. 5, but with the Apr 2/3 signature.

Relative probabilities of origin can be found from the data in Figs. 5 with Eq 1. Table 1 shows metrics for 7 events between Apr 7 and 16 for the amplitude/shape metrics. Event 0 (Apr 12/13) best matches the model-predicted plume from Punggye and Table 1 shows the probability of Punggye origin for this plume (bottom row) obtained with metrics for ^{133}Xe, ^{131m}Xe, and ^{133}Xe/^{131m}Xe, namely, 44%, 58%, and 96%, respectively. The probability of Punggye origin for event 0 is higher when ^{131m}Xe activity is considered than for ^{133}Xe, and greater still when both isotopes are considered together. When both isotopes were considered only two events were found in the 9.2 day interval of interest.

Fig. 7a shows the same results presented in Table 1 (the first solid bar in each group of 4 bars). Fig 7a also shows probabilities when the shape metric is used (the first hatched bar in each group of 4 bars). In this case the probabilities of origin are 23%, 27% and 72%, respectively. These lower probabilities indicate the importance of the signal amplitude in source attribution but that a significant probability of origin is seen with shape metrics only. The format shown in Fig. 7 is a convenient for visualizing the results and will be used to describe events in the remainder of this paper.

It should be noted that each of the other 6 events in the Apr7-16 period has a non-zero probability of origin from Punggye, but these probabilities are 1/3 to 1/25 found for the 0'th plume, depending on the choice of isotope.

Event	**Prior Probability**	**133 Xe composite metric**	**131m Xe composite metric**	**133/131m Xe composite metric**
0	1.	0.098	0.306	0.028
1	1.	0.035	0.064	0.002
2	1.	0.003	0.093	
3	1.	0.013	0.040	
4	1.	0.023	0.092	
5	1.	0.052		
6	1.	0.037		
P(0|m$_0$)		**44%**	**58%**	**96%**

Table 1: Information for 7 events between Apr 7 and 16 at RN58 (see Fig.5). Event 0 (Apr 12/13) is the event that best matches the model prediction for a Punggye origin. The prior probabilities are in the second column and metric values in columns 3-5. The bottom row is the probability that event 0 originated from Punggye per one week period. The interval of interest is 9.2 days.

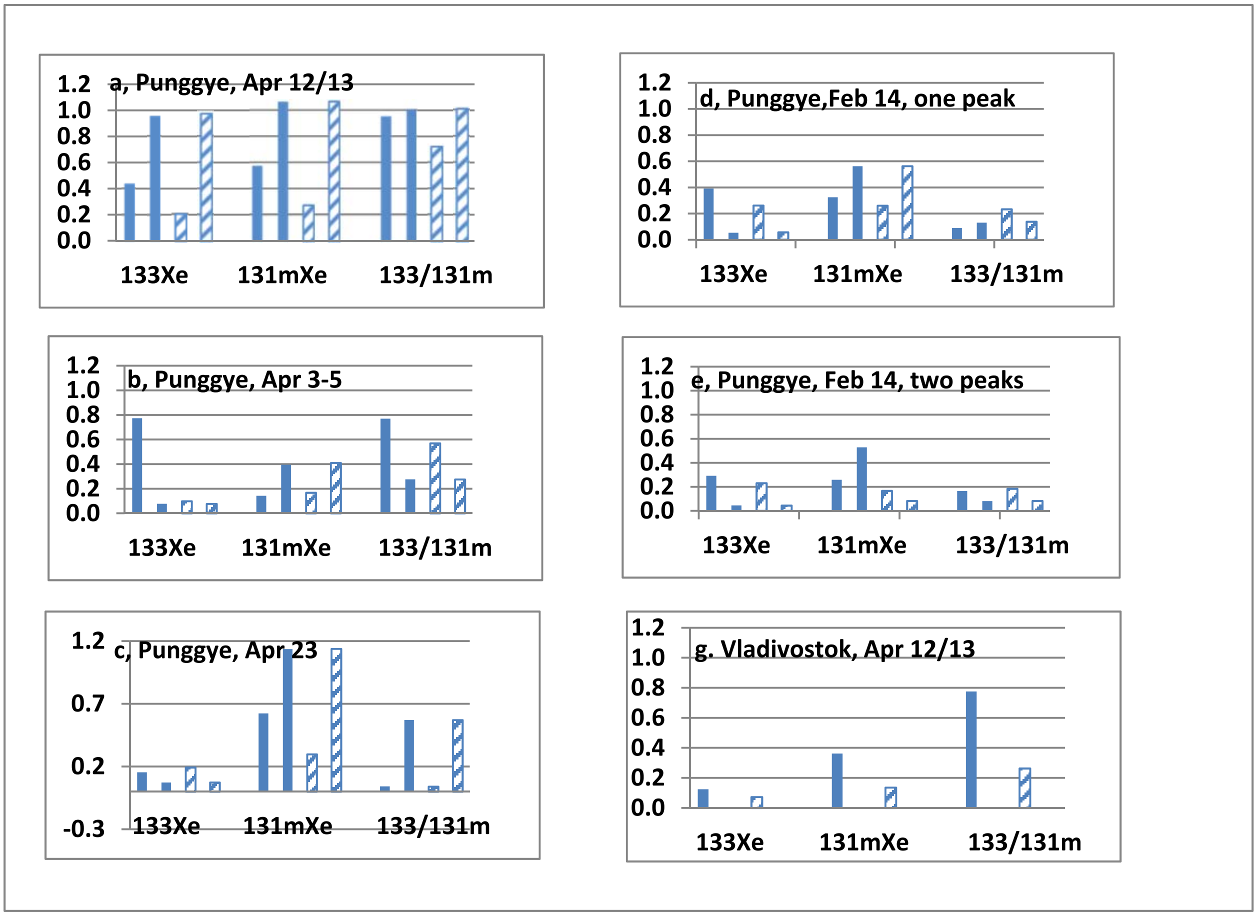

Fig. 7. Relative probability of Punggye origin for radioxenon events observed at RN58 in Ussuriysk, Russia. Figs. a-e were found with model signatures for a source at Punggye, NK. Fig g is for a model signatures from Vladivostok. Each group of four bars shows, respectively, the probabilities for the amp-shape metric (solid), the amp-shape arrival time metric (solid), the shape metric (hash) and the shape arrival time metric (hash).

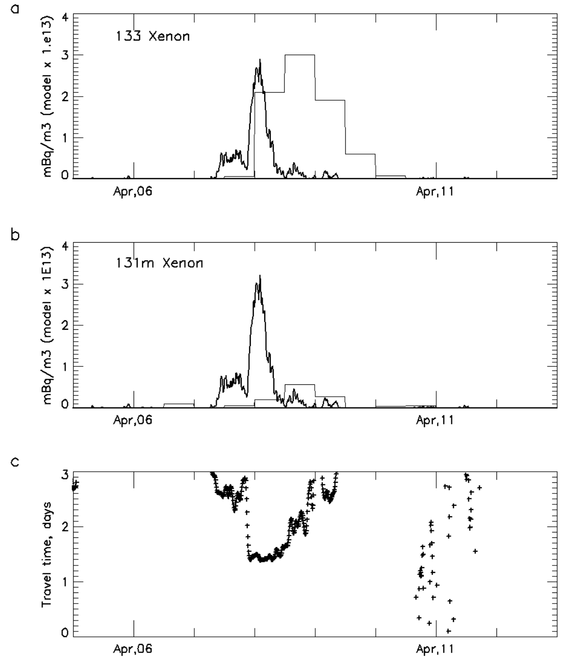

Fig. 8. The observed and modeled activities of ^{133}Xe (a), ^{131m}Xe (b) and the travel time to RN38 in Takasaki, Japan (c)

6. The April 8/10 event at RN38

We turn now to activity measured at the RN38 site in Japan, ~1200 km to the southeast of Punggye for additional insight into the method (see Fig. 3). As noted by Ring14, anomalously high ^{133}Xe and ^{131m}Xe activities were recorded at RN38 on April 8-10. Ring14 attributed this event to the Feb 12 underground test based on the isotopic ratio of ^{131m}Xe /^{133}Xe and atmospheric back trajectories. Although analysis of the entire time period is impossible without access to the complete data set, the measured activities for the period of April 4-12 are available from Ring14 (their Table 1 and Fig. 15). Figure 8 shows these data along with simulated time series from the RAMS/HYSPLIT simulation for the same period. The travel time varied between 1.3 and 2.7 days, with an average of 2.0 days (Fig 8c).

In contrast to the RN58 data, the signature and data at RN38 indicate a 2-day (four 12-hr averages) event rather than the 1-day (two 12-hr averages) event seen at RN58. The longer event permits a more detailed comparison between the observations and the model signature. The first point to note in Fig. 8 is the good agreement between the simulated and observed ^{133}Xe and ^{131m}Xe activities. Both measurements and signature show a central maximum value with lower shoulder values. However, the observed shoulder values are 1/3 to 2/3 the central value whereas the model shoulder are approximately 1/3 the center maximum. The longer signature at RN38 suggests that the shape metrics may be more important than at RN58. In fact, the Pearson Correlation coefficient was twice as large at RN38 as found for the Apr 12/13 event at RN58. This emphasizes the value of temporal resolution in the data series comparable to that of the model signature.

The travel time estimation (Fig. 2c, bottom) indicates that the first third of the plume may have been emitted approximately 1 day before the main body of the plume. The model plume center of mass is also 0.8 days early at RN38 compared with the observations. The longer signature length at RN38 is expected because of the longer travel time and greater lateral dispersion.

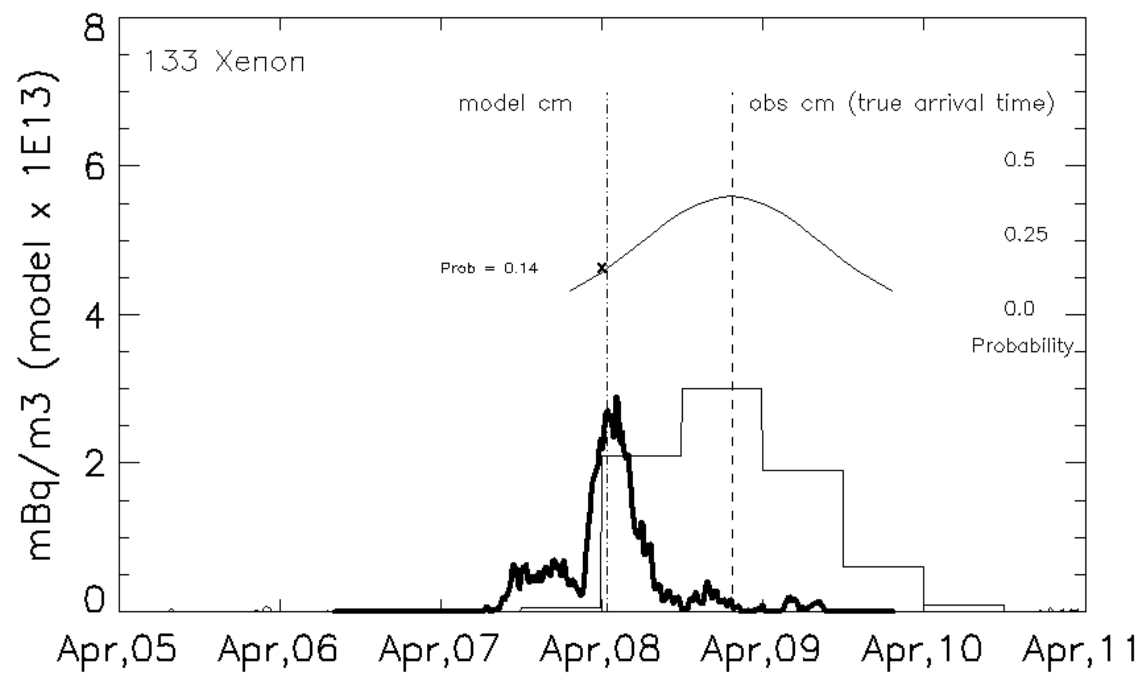

Fig. 9 The April 7-9 event at RN38 in Takasaki, Japan. The heavy and thin lines are the model simulation and data, respectively. The vertical dotted lines are the model and observed centers of mass. The thin curve (top) is a normal distribution of arrival times around the true arrival time (the observed center of mass).

7. Arrival time metric

The probability analysis presented in Section 5 does not include information on back trajectories, which are commonly used to link an observed event to an upwind location. This is mainly because time series filtering that includes back trajectories from each point in the time series is computationally expensive and difficult to interpret. In addition, back trajectory analysis typically defines a continuous spatial distribution of possible plume origins that extends from the receptor an infinite distance upwind, i.e., it does not limit the domain in the transport direction. Although the shape metrics vary with downwind distance and provide an approximate measure of the source's upwind distance, a more precise measure is desirable. It is also desirable to distinguish between plume trajectories and plume dispersion because of the different model processes involved. We now define a trajectory metric whose main purpose is to measure the upwind distance to the source that can be combined with the shape and amplitude metrics for improved source attribution

Trajectories are characterized by plume direction and speed. The plume center of mass is an appropriate measure of a plume's movement but require measurements at more than one location which is usually not practical. Accordingly, we focus only on the plume speed as measured by observations at the receptor.

For this purpose, we define the arrival time error as the difference at the receptor in arrival times of the model-predicted (or event) center of mass and the true center of mass. Fig. 9 illustrates this concept at RN38 on Apr 7/9. In this case, based on Ring14, we assume that the true arrival time is given by the center of mass of the observed event. Then the arrival time error of the model prediction is -0.8 days. The travel time error can be generalized by assuming that the arrival time error increases with travel time which suggests a definition of the normalized travel time as the ratio of arrival time error to the travel time. As shown in Fig. 9 the average travel time to RN38 is ~2.0 days which yields a normalized travel time error of-0.8 days/2.0 days = -0.4. The corresponding analysis of the RN58 plume (not shown) indicates an arrival time error of +0.1 days for the model compared with the observed event, which we also assume defines the true arrival. Then a travel time of 0.6 days (see Fig. 2) yields a normalized arrival time error of 0.1 days/0.6 days = +0.17. The average of this value and that found at RN38 yields an estimate of the standard deviation of the normalized arrival time error, σ_N, of,

$\sigma_N = (0.4+0.17)/2. = 0.28$

We next assume that the arrival time error obeys a normal probability distribution, i.e.,

$$P(\text{error}) = m_a = \alpha \times \exp(-((t-T_t)/\sigma_e)^2/2)/\sqrt{(2\pi)} \qquad \text{Eq 2}$$

Where σ_e is the effective standard deviation for the particular situation and α is a proportionality constant to be determined empirically. The effective standard deviation is found from σ_N and the travel time. For example, for the observed event at RN38 on April 7/8 the σ_e is,

$\sigma_e = 0.28 \times 2.0 \text{ days} = 0.56 \text{ days}.$

The above calculation is approximate because σ_N was obtained from only two events. More events are necessary to find reliable values of σ_N for a range of conditions. The true arrival time is also uncertain in general. The true arrival time at RN38 on Apr 8/9 is based isotopic ratios and on four 12-hr averages and is therefore known with reasonable accuracy. It is less accurately known for the Apr 12/13 event at RN58 because it is derived from only two data points. When the true arrival time cannot be found from an observed event, it must be estimated from other information. In some cases it can be found from other chemicals released concurrently from the source. It might also be found from another source near the source in question, e.g., from a power plant or city. Another option is to use a data base of simulated and measured tracer data to find typical arrival time errors for the model. If no supplementary data are available, the true arrival time can be estimated from the mean of an ensemble of model simulations after removal of the model bias. This option may be the only possibility for remote locations with no known sources.

Fig. 9 illustrates the case when the true arrival time is known from the observations at RN38. As noted above, in this case σ_e=0.56 . The probability distribution of the arrival time error is shown in the top part of Fig. 9 and the probability of an arrival time error of 0.8 days for a model prediction is found to be 0.14. The corresponding probability for the RN58 plume is 0.33.

There are several ways to estimate the value of the empirical parameter α. We might assume that the probability of origin equals 100% when the arrival time for a plume from its location equals the true arrival time. Since the center point of a normal distribution has the value 0.4 in this case α equals 1.0/0.4 = 2.5. This may overestimate the probability since a plume from a distant or nearby source might also have an arrival time error of zero. Another possibility is to choose α empirically. For example, based on Ring14 we assume a 95% probability of origin from Punggye for the Apr 12/13 event at RN58 if both isotopes are considered. Then α is the constant that multiplies the probability of 0.33 to equal 0.95 or α =2.85, which we use for the rest of the calculations in this paper. Some of the probabilities shown in Fig 7 exceed 100% which reflect uncertainty in the estimate. The value of α can be adjusted based on additional data and also may prove to vary with the calculated probability.

We now can interpret the bar charts shown in Fig. 7. The first 4 bars in each figure indicate the probability of origin from Punggye respectively, with,

1. the amplitude/shape metric,
2. the arrival time error metric obtained with the amplitude/shape metric,
3. the shape metric,
4. the arrival time error metric obtained with the shape metric.

The probabilities for ^{131m}Xe and ^{133}Xe/^{131m}Xe are given by bars 5-8, and 9-12, respectively, in Fig. 7.

Fig. 7a shows probabilities of origin from Punggye of 0.42, 1.0, 0.2 and 1.05 for the four cases listed above for ^{133}Xe. The probabilities are larger for ^{131m}Xe and ^{133}Xe/^{131m}Xe. The arrival time error metrics for cases with and without the amplitude metric are also comparable. This is because the arrival time error is independent of amplitude unless the amplitude is comparable to the background, when the plume's center of mass becomes uncertain.

A net probability can be found by averaging the amplitude/shape and arrival time metrics. For example, the probability of origin from Punggye based only on ^{133}Xe when both the amplitude/shape metric and its corresponding arrival time metric are considered is (0.42+1.0)/2= 0.71.

It should be noted that the arrival time metrics were not found with Bayes rule but rather as absolute values that depend on the difference between arrival times of the model and data. This approach was chosen because the arrival time error is an absolute measure of agreement. Moreover, the determination of the arrival time errors requires the travel time, which is unknown for most events in a data series.

8. The Apr 3 and Apr 22 events at RN58

Elevated radioxenon activity was seen at RN58 on other occasions in April 2013 besides Apr 12/13 (see Fig. 2). High ^{133}Xe activity was observed on Apr 3/4, Apr 12/13, Apr 15, Apr 18/19, and Apr 20. High

^{131m}Xe activity was seen only on Apr 12/13 and Apr 22. Most of these peaks corresponded to maxima in model predictions except during Apr 15-Apr 20 when measured peaks had no model equivalent. A pronounced model peak on Apr 17 had no observed counterpart. Unusually large ^{131m}Xe activity was not seen on Apr 3/4, which is expected based on Ring14's projection that the ^{133m}Xe activity is ~1/10 that of ^{133}Xe at this time.

Some insight into the possibility that radioxenon was emitted from the underground test at other times in April, 2013 via atmospheric 'pumping' is obtained by comparing measured activity with the surface pressure. Fig. 10 shows the sea level pressure at Punggye for April, 2013. The vertical lines denote times one day before the beginning of the measured event which accounts for the approximate travel time. Apr 7 is the day of the lowest surface pressure (1002 mb), i.e., greatest removal of Xenon, and is also the approximate beginning of a plume that reached RN38 on April 8/9. The activity measured at RN38 was 2-10 times larger than measured at RN58 despite the greater transport distance to RN38 which implies a source strength 5-10 times greater than inferred for the Apr 12/13 event measured at RN58.

The pressure at Punggye preceding the Apr12/13 event is 1009 mb, which is ~ 7mb higher than on Apr 7, but then drops to 1002 mb on Apr 12-14 . The Apr 17 signature originated on about Apr 16, the end of a 4-day rise in atmospheric pressure which may explain the absence of significant observed activity. On the other hand, the pressure preceding the Apr 3 and Apr 22 ^{131m}Xe events was near 1015 mb which is ~ 11 mb higher than the minimum of 1002 mb observed on Apr 7 and implies lower emission and lower downwind activities. Therefore these two events provide a good test of the method at lower signal strengths.

The analysis described in Section 5 was repeated for the Apr 3 event and is summarized in Fig. 7b. A comparison of Figs 7a and 7b indicates smaller probabilities of Punggye origin for Apr 3 than for Apr 12/13. The probabilities for the combined isotopes are larger than values of individual isotopes and are in the range 0.3-0.8.

The smaller arrival time metrics for the Apr 3 event are explained by examination of Fig. 2 which shows that observed ^{133}Xe event on Apr 3 occurred approximately 12 hours earlier than the model maximum. The travel time is approximately one day which leads to a normalized arrival time error of 12 hours/24 hours = 0.5, which is larger than the found for the Apr 12/13 event at RN58 (=0.17) but comparable to the value found for the Apr 6/7 event at RN38 (=0.4). Thus, the arrival time metric for the Apr 3 event does not preclude a Punggye origin. Further analysis of the Apr 3 event is presented in Section 11.

An enhanced ^{131m}Xe activity was also seen on April 22. The method was applied to this event and Fig. 7c shows the resulting metrics. The ^{131m}Xe metrics are of greatest interest because it is expected to be much higher than that of ^{133}Xe because of its longer half-life (Ring14). The Apr 23 ^{131m}Xe metrics are comparable to those found on Apr 12/13 and we conclude that the Apr 22 event is of probable origin from Punggye.

The isotopic ratios ^{131m}Xe/^{133}Xe for the Apr 3 and Apr 22 events were approximately 0.1 and 6, respectively, which can be compared with Ring14's estimates of 0.05-0.2 and 0.4-2.0. However these estimates are uncertain because of the near background activities of ^{131m}Xe and ^{133}Xe, respectively, on the two days.

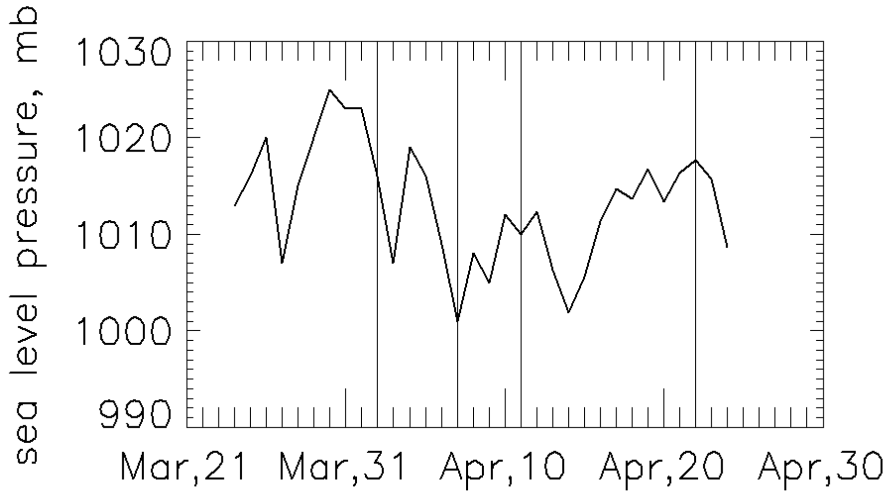

Fig. 10: Sea level atmospheric pressure at Punggye. The vertical lines denote the approximate release times 1 day before the observed radioxenon events on Apr 3, Apr 8, Apr 12 and Apr 23, at RN58, RN58, RN38, and RN58, respectively.

9. The Feb 14 event at RN58.

In previous sections we have examined events in April of 2013 for matches with model signatures. In this section we consider the event on February 14. As noted in the Section 1, radioxenon may reach the surface immediately after the detonation because of explosive pressure. Containment measures and geologic properties at the site will also affect movement of gases to the surface. The possibility of emission in the week after the test is suggested by a measured ^{133}Xe activity that was 2-3 times that measured on Apr 12/13 at RN58 (Figs. 4,11). No obvious anomalies were observed in the ^{131m}Xe activity in the week after the test.

To test the possibility that the February 14 events originated from the Feb 12 test, the analysis described in previous sections was repeated. Figure 11 shows the data and signatures for February, 2013. Two model peaks are seen on Feb 13 and two measured activity peaks on Feb 14/15. A smaller model peak is seen on Feb 17 that corresponds to a slight activity maximum.

Two signatures were considered. The first included only the second (larger) model peak on Feb 13 while the second included both model peaks. The results are summarized in Figs. 7d and 7e. The metrics are seen to be much lower (10-60%) than found for the April events. Only the ^{131m}Xe metrics for the one-peak signature are significant (30-60%). We conclude that the anomalously high ^{133}Xe activity seen on Feb 14/15 did not originate from the underground test on Feb 12.

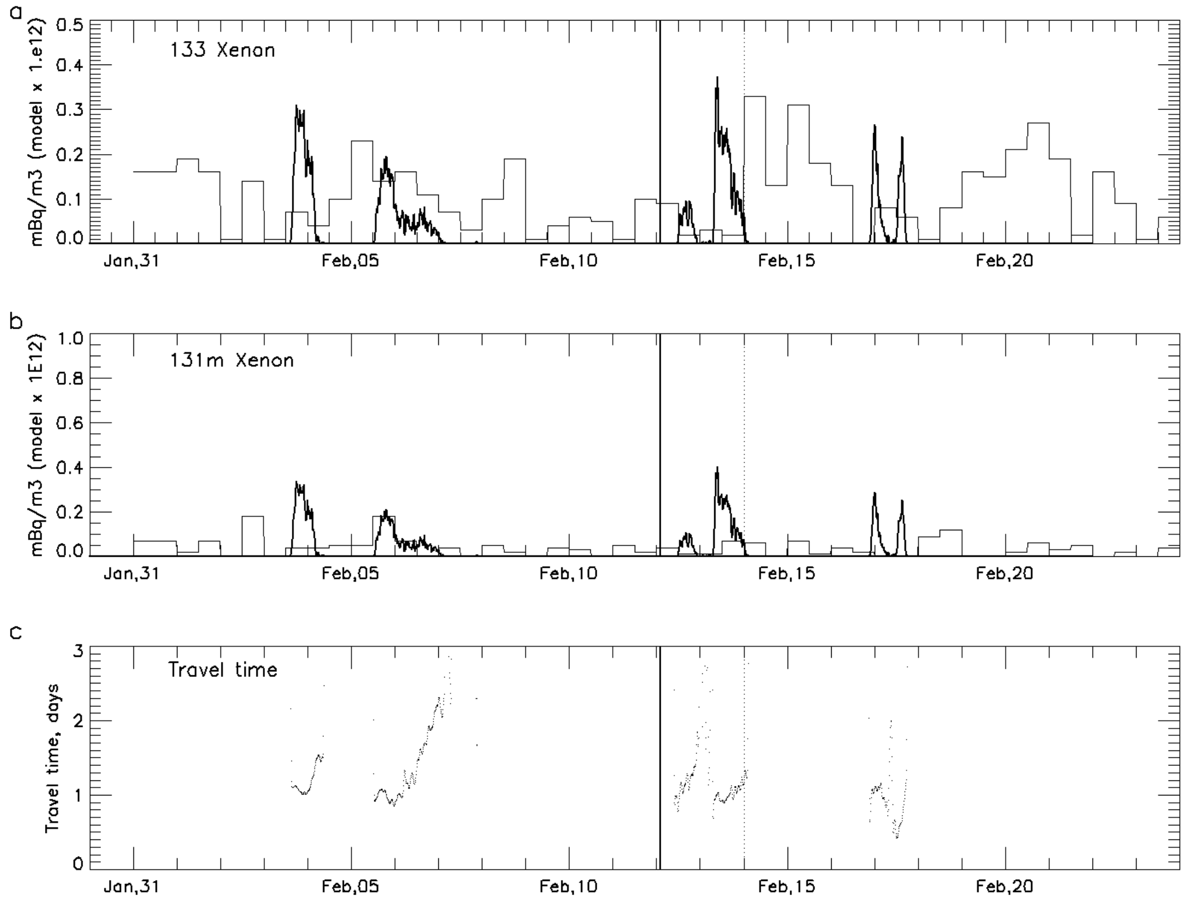

Fig 11: Observations and model simulations of 133Xe and ^{131m}Xe for February (a,b). Travel times are shown in (c). The heavy vertical line indicates the time of the underground test.

10. Possible Vladivostok origin of the Apr 12/13 event at RN58

The analysis in the preceding section has shown a high probability that the Apr 12/13 event at RN58 originated from Punggye. A relevant question is how the metrics from a Punggye release compare with those from a nearby location. A possible source of radioxenon is Vladivostok, Russia (Fig. 3), a refueling port for Russia's nuclear submarine fleet. Vladivostok (41.07N, 131.53E) is 100 km south of RN58 compared with Punggye which is 450 km southwest of RN58. Testing our method with a Vladivostok source is also a useful demonstration of how our method can distinguish between sources in the same general upwind direction, but at different upwind distances.

A preliminary comparison of the observed activity with the model prediction explains why the Apr 12/13 event is probably not of Vladivostok origin. Fig. 12 shows the radioxenon data at RN58 on Apr 12/14 and the model signature for a release from Vladivostok which can be compared with the Punggye plume (Fig. 2). As can be seen the Vladivostok model event is 12 hours long instead of 24 hours long and arrived 0.25 days late instead of 0.1 days late, as found for the Punggye plume. The travel time to RN58 from Vladivostok is only ~0.15 days instead of 0.6 days as found for the Punggye plume, which yields a normalized arrival time error approximately 10 time larger than for the Punggye plume, i.e., 0.25/0.15 = 1.7 vs 0.16 found for the Punggye plume. This large arrival time error implies that the corresponding metrics will likely rule out Vladivostok as the source, regardles other metrics.

To test this hypothesis the metrics for the Vladivostok plume were calculated and are shown in Fig. 7g. They amplitude/shape metrics and and shape metrics are generally lower than found for the Punggye plume while the arrival time metrics are approximately zero for the reason given above. We conclude that the Apr 12/13 events seen at RN58 are not from a Vladivostok release.

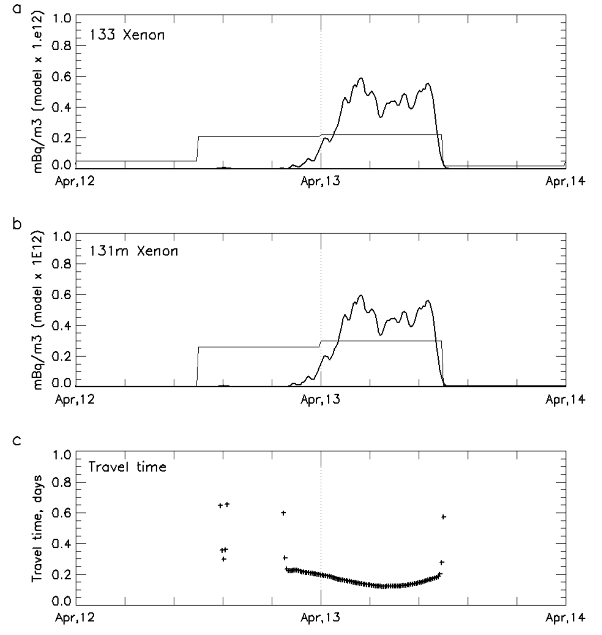

Fig 12: Measured and simulated radioxenon activity at RN58 for a release from Vladivostok, Russia.

11. The April 3 event: multiple release scenarios

As discussed in Section 8, the metrics for the Apr 3 event were lower than for the Apr 12/13 event and the ^{131m}Xe activity was near the background level. However, the relatively high metrics for the two isotopes considered together suggest a possibility that the Apr 3 events may have originated from the underground test. It is instructive to repeat the analysis for locations around Punggye under the assumption that the location with the largest metrics is the most likely origin. This approach depends on the relative sizes of the metrics rather than their absolute values. The grid of selected release locations around Punggye is listed in Table 3 and shown in Fig. 13. The distances of the release locations from Punggye vary from between 50 to 300 km.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
-	½ E	½ NE	½ N	½ S	½SW	½ W	1 E	1 NE	1 N	1 S	1SW	1 W	3S	3SW	3W

Table 3: Displacement in degrees latitude and longitude from Punggye for the 15 simulations. Note, ½ NE denotes ½ degree north and ½ degree east of Punggye. The first entry denotes a source at Punggye.

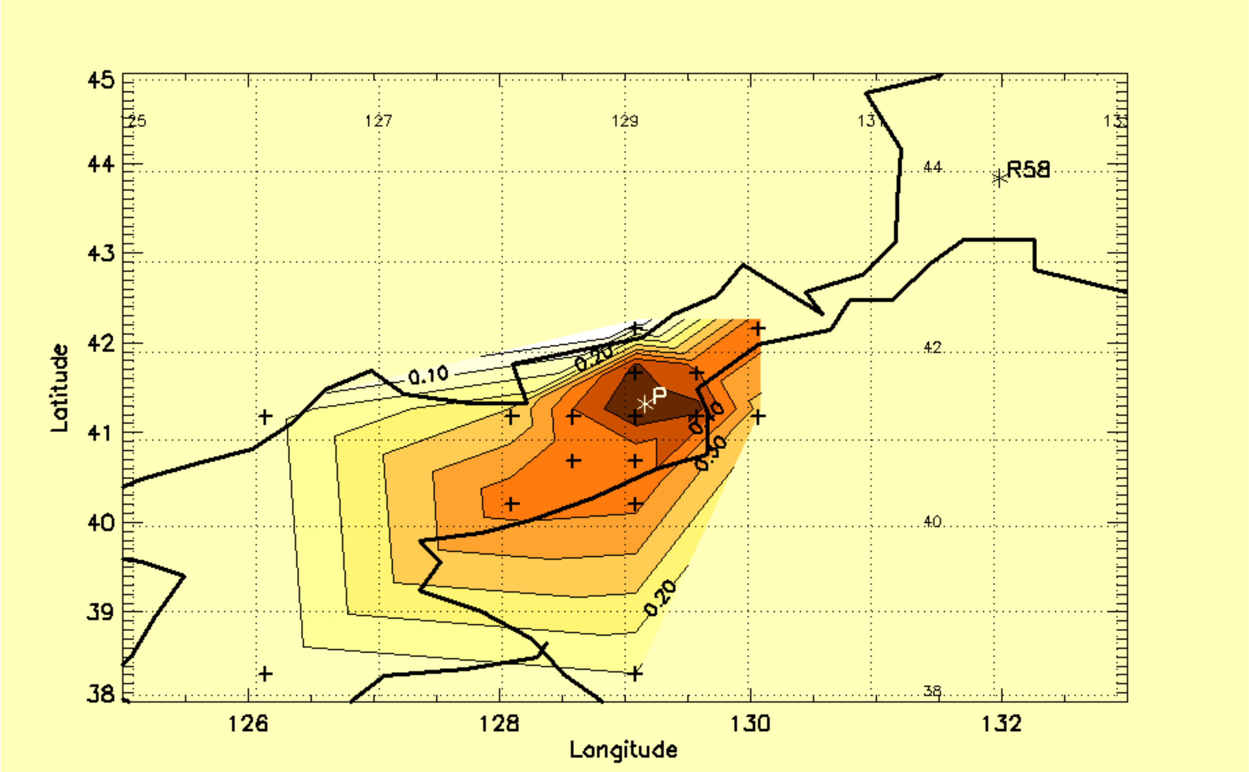

Fig. 13: Probability of origin of the April 3 event from Punggye vs a neighboring location. Punggye and RN58 are indicated with 'P' and 'RN58'. Crosses denote alternate release locations.

The ^{133}Xe metrics were calculated for each of the 16 release locations and the probabilities of origin from Punggye were calculated with Eq 1 by pairing the Punggye metrics with each of the 15 neighboring release locations. The results are shown in Figs. 13 and 14. Fig. 13 shows the relative probabilities of origin for each location versus the Punggye probability obtained from the ^{133}Xe activity. In principle, the probability can range from 0.0 to 1.0 with a value at Punggye exactly equal to 0.5 because the probability of origin from two infinitely close points must be the same. A location with a value greater (smaller) than 0.5 is one with greater (lower) probability of origin than from Punggye.

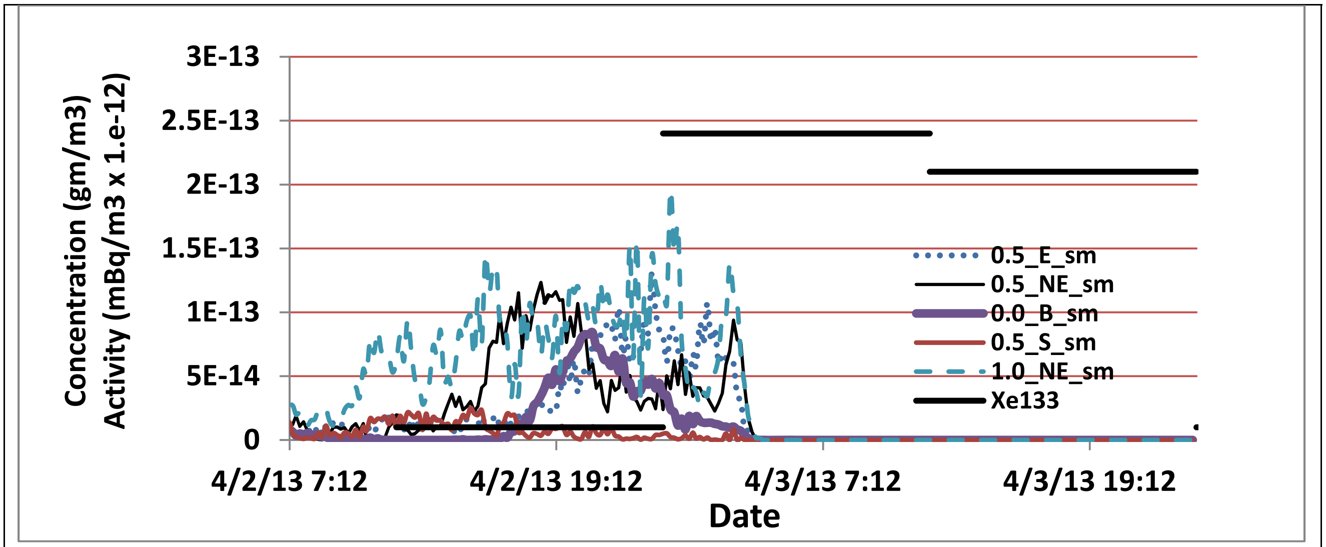

Fig. 14: Model simulations of the ^{133}Xe concentration at RN58 from releases at Punggye (0.0_B_sm), and from locations ½ degree east, northeast, and south of Punggye, and a location 1 degree northeast of Punggye. The solid bars denote the measured ^{133}Xe concentrations in mBq/m3 (×1e-12).

As seen in Fig. 13, Punggye was found to be the most likely origin for the Apr 3 plume. The reasons for this are shown in Fig. 14 which shows 5 plumes from locations around Punggye as well as the Punggye plume. As can be seen most plumes are quite different from the Punggye plume with a poor match with the observed activity, which explains why their metrics are lower than the Punggye metrics.

It may seem that metric values and the implied probabilities must invariably decrease with increasing source distance normal to a line from Punggye to the receptor. This belief may be based on the concept of a straight line Gaussian plume in which the plume centerline coincides with the line between source and receptor. However, in actual situations plumes almost invariably follow curved paths to the receptor and tend to be advected across the receptor at an angle with respect to source-receptor line. Thus, the locus of maximum source probability will in general follow a curved path from source to receptor and the most likely source will usually not be on the straight line from source to receptor.

The multiple source analysis presented in this section is more computationally expensive that the event analysis discussed in Sections 5-9, but it is more persuasive because each plume is described exactly by the model which leads to more precise metrics. On the other hand, systematic errors in the model or the data will lead to biases in the most probable source location. Most obvious is that a bias in the speed of the model transport winds will lead to a source location closer or further from the receptor.

12. Discussion

This study has described a new method to associate segments in a concentration data series with an upwind source through metrics that measure the degree of agreement between a data segment and a model signature - the predicted concentration at the receptor. A major motivation was the belief that more information could be extracted from a data series if it was compared objectively to a specific signature of interest. Even though a unique model signature is only expected under ideal conditions, i.e., a perfect model with correct initial and boundary conditions, no turbulent variability, and high temporal resolution in the concentration measurements, nevertheless the concept of signature uniqueness is the basis for our assumption that metrics that measure agreement between the signature and a data segment are proportional to the probability of origin from the signature location. Thus, the signature is a degraded 'fingerprint' of the source which must be compared with signals from other sources to determine uniqueness.

Four metrics were defined to quantify agreement between the model signature and data segments in the time series. These are the amplitude, variance, Pearson correlation coefficient and the plume arrival time error. The arrival time error metric converts the complicated information from a back trajectory into a single metric that can be applied to a time series.

Some of the metrics (variance, Pearson correlation coefficient) depend on temporal resolution of the activity measurements. Plumes from close sources are narrow and sharply defined and require high temporal resolution measurements. The 12-hour resolution in the CTBT data limited the signature length at RN58 (Ussuriysk, Russia, 400 km from Punggye) to 2 data points (24 hours) and to 4 data points (48 hours) at RN38 (Takasaki, Japan , 1200 km from Punggye). For this reason the metrics were better defined at RN38 than at RN58.

The problem of inferring source properties from downwind measurements has been addressed in many previous studies, see for example Hutchinson et al., 2017. One common approach is the source-receptor method which relates sources to receptor points with a matrix derived from back trajectories, e.g. Siebert and Frank (2004). Typically this approach requires multiple measurements of high signal to background ratio and yields areas of probable origin rather than the probability of origin from a specific source. A single receptor leads to a continuous upwind domain from the receptor. Ring14 added precision to the process by considering the overlap region derived from a sequence of source areas derived from backward trajectories. This process reduced the size of the probable source area. They also considered the overlap of source areas derived from the RN58 and RN38, which is an abbreviated form the source-receptor matrix method. These methods differ from the method described here because the time segments are based on inspection and are equally-weighted. In addition, the results and uncertainty will be given in terms of a source area rather than as a probability of origin.

The use of a model signature differs from other techniques that associate a signal with a particular source by means of a unique trait of the source, such as the isotopic ratio, because it can be applied to a single isotope, or to multiple isotopes for enhanced selectivity. It also differs from other plume tracking approaches because it does not require multiple receptors and a source long enough in time to enable detection at more than one receptor.

The method was demonstrated with the radioxenon data collected at RN58 and RN38 after the February 2013 underground nuclear test at Punggye, North Korea. The method confirmed previous results that the radioxenon data on April 12-13 at RN58 and April 7-8 at RN38 originated at the North Korean test at Punggye, The approach was applied to other events in April 2013 and identified an event on April 3 and a second on April 20 that are of likely Punggye origin. We also studied events in Feb 13/14 after the Feb 12 test. Probabilities that this event originated from Punggye, while greater than zero, were significantly lower than for the April events, despite high measured ^{133}Xe activity.

An alternative use of the technique compared the metrics for points around Punggye to those from Punggye for the April3/5 event. Metrics from model simulations for 15 locations around Punggye were calculated and evaluated with Bayes rule. The largest metrics, greatest probability, were found for a Punggye release. These results are significant because minor flaws in the model should have the same general effect on all the metrics and thus not alter the relative sizes of the metrics.

The test of the April 3 event is similar to inverse methods that find a source location or strength by minimizing a cost function, such as the variance between the predicted and observed concentration at one or more receptors , e.g., Hutchinson et al, 2017. For example, our metrics could be combined into a single cost function for this purpose. However, this approach does not yield probabilities of origin or their uncertainty.

The use of Bayes rule for the detection of undisclosed sites requires estimates of the prior probabilities which may be highly uncertain. For this reason a better strategy is to assign only one probability to the unknown plumes and one to the event in question. The uncertainty in the prior probability emphasizes the importance of accurate likelihood probabilities. The greatest chance for a false positive with our method is when a signal from a distant or nearby source happen to produce a plume whose center of mass coincides exactly with a model-predicted center of mass at the receptor. This would yield a large arrival time error metric and a high (false) probability of origin. In this case however, smaller shape metrics should result that will allow differentiation between the plumes. This example emphasizes the value of examining the shape and arrival time metrics separately.

Limitations to the study should be kept in mind. It must be emphasized that our results are preliminary since only a few events have been analyzed. A complete analysis of uncertainty with additional observations and model simulations is required to establish the method for a variety of situations.

For convenience, we assumed that emissions were steady and continuous during the signature. This is not a serious limitation because the assumption applies to each event individually, i.e., the emission rate is constant during an event but different from event to event. In additions, temporal variation in the source could be included in the analysis if known.

A second limitation is that the method tests hypothetical sources and is not efficient for finding the most likely origin of the source, which requires the creation and testing of many signatures. In practice, however, the evaluation of a set of sources is possible and could be incorporated into a more complicated software algorithm.

However, the most important limitation of this study is the lack of a data base of metric values for a variety of scenarios. The availability of such a data base would permit a direct conversion of metric values to probability of origin.   Equally useful would be model experiments with an ensemble of model simulations that encompass the known uncertainties to determine the mean values and variances of the metrics.   This would allow us to determine the uncertainty of the probabilities used in Bayes rule.

Ensemble studies would also elucidate the role of background data on the metric values. In general, increased background concentrations will increase the uncertainty in the metrics and reduce the size of some metrics. It should be recognized that signals commonly lumped under the category of ’background’ may simply be from unknown sources, which suggests that the true may be much lower than currently assumed.

13. Conclusions

This study has shown that much additional information about upwind sources can be extracted from a time series of measurements at a receptor through the use of a model-derived prediction (signature) at the receptor. The signature was used to develop two significant products. The first is a filter that isolates segments of data similar to the signature. This process replaces subjective search of a data series for anomalous events. The second is the use of the signature to define metrics which measure the agreement between an event’s amplitude, shape and arrival time with that of the signature, from which the probability of origin for the event can be found.

The technique has several advantages over other backtracking efforts. First, since the method does not involve triangulation, only one receptor is needed and the source need not be of any duration. Second, the technique can be extended to multiple isotopes and arbitrary measurement strategies. Third, probability of origin and its uncertainty are calculated which is critical to decide how much credibility should be attached to weak signals.

The technique can also be applied to arbitrary sampling strategies. For example, it is straightforward to calculate the model signature for a moving, e.g., aircraft, sample with an arbitrary flight path. The value of the moving sample can be found by comparing the metrics with metrics from a stationary surface sample. The value of high resolution data can be determined by comparing metrics from high resolution data with results with averaged data to decide whether improved instruments are cost-effective. Enhanced resolution can also be achieved with overlapping samples. For example, increased resolution can be achieved with 12-hour samples overlapping in 6 hour intervals.

The technique has considerable potential. One is guidance for improvements to transport models and quantification of model errors. For example, understanding the systematic error in the model-derived transport speed as a function of receptor distance will permit a simple correction to the arrival time error. Better wind and turbulence data will reduce the signature uncertainty.   The value of high resolution measurements at the receptor has also been shown

References.

Achim, P., Morin, M., Gross, P., Le Petit, G., Monfort, M., 2013. Highlighting radioxenon local sources by studying the seasonal variability of worldwide atmospheric background. Application to JPX38-Takasaki detections. CTBTO Science and Technology Conference.

Auer, M. and M. Prior, 2014. A New Era of Nuclear Test Verification. *Physics Today*, Sept. 2014.

Becker, A., Wotawa, G., De Geer, L.-E., Seibert, P., Draxler, R.R., Sloan, C., D'Amours, R., Hort, M., Glaab, H., Heinrich, P., Grillon, Y., Shershakov, V., Katayama, K., Zhang, Y., Stewart, P., Hirtl, M., Jean, M., Chen, P., 2007: Global backtracking of anthropogenic radionuclides by means of a receptor oriented ensemble dispersion modelling system in support of nuclear-test-ban treaty verification. *Atmos. Environ.*.,41, 4520-4534.

Carrigan, C. and Y. Sun, 2012. Detection of noble gas radionuclides from an underground nuclear explosion during a CTBT on-site inspection. *Pure and Appl. Geophys.*, August. DOI 10.1007/s00024-012-0563-8

Chino, M., H. Ishikawa, H. Yamazawa, 1993. SPEEDI and WSPEEDI: Japanese Emergency Response Systems to Predict Radiological Impacts in Local and Workplace Areas due to a Nuclear Accident. *Radiation Protection Dosimetry*, 50 (2-4), 145–152.

Cotton. W.R., R. A. Pielke, R. L. Walko, G.E. Liston, C. J. Tremback, H. Jiang, R. L . McAnelly, J. Y. Harrington, M.E. Nichols, G.G. Carrio and J.P. Mcfaddent, 2002. RAMS 2001: Current status and future directions*. Meteorol. Atmos. Phys.,* DOI 10.1007/s00703-001-0584-9.

CTBTO, 2013. CTBTO Preparatory Commission Web Page. http://www.ctbto.org/

De Geer, Lars-Erik, 2012. Radionuclide Evidence for Low-Yield Nuclear Testing in North Korea in April/May 2010. *Science and Global Security*, 20 1-29. DOI: 10.1080/08929882.2012.652558.

Draxler, R. 1987. Sensitivity of a trajectory model to the spatial and temporal resolution of the meteorological data during CAPTEX*, J. Climate and Applied Meteor*., 26, 577-588.

Draxler, R.R., and G.D. Hess, 1998. An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition*. Aust. Meteor. Mag*., 47, 295-308.

Dubasov, Y., 2010.Underground nuclear explosions and release of radioactive noble gases, *Pure Appl. Geophys*., 167, 455–461.

Environmental Modeling Center, 2003: The GFS Atmospheric Model. NOAA/NCEP/Environmental Modeling Center Office Note 442, 14 pp. [Available online at http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html.]

Harrington, J. Y., 1997.The effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus, PhD Dissertation., Atmospheric Science Paper No 637, Colorado State University, Department of Atmospheric Science, Fort Collins, CO 80523, USA, 289 pp..

Hoffmann, W., R. Kebeasy, and P. Firbas, P. ,2000), Introduction to the verification regime of the Comprehensive Nuclear-Test-Ban Treaty, Phys. *Earth Planet. Inter*. 113, 5–9.

Hutchinson, M. O., Hyondong, W-H Chen, 2017. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensor, *Information Fusion*, 36, 130-148.

Jain, A. K., R. P. Q. Duin and J. Mao, 2000. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and machine Intelligence*, 22, January 2000, p4-37.

Kalinowski, M. and M. Tuma, 2009. Global radioxenon emission inventory based on nuclear power reactor reports. *J. Environ. Radioactivity*, 100, 58-70.

Kalinowski, M. B., 2011: Characterisation of prompt and delayed atmospheric radioactivity releases from underground nuclear tests at Nevada as a function of release time. *J. of Environ. Radioactivity*, 102, 824-836.

Le Petit, Gilbert, P. Armand, G. Brachet, T. Taffary, J-P Fontaine, P. Achim, X. Blanchard, J-C. Piwowarczyk, F. Pointurier, 2008. Contribution to the development of atmospheric radioxenon monitoring. *J of Radioanalytic and Nuclear Chemistry*. 276/2, p391-398.

Mellor, G. L. and Yamada, T., 1982. Development of a turbulence closure model for geophysical fluid problems, *Rev. Geophys. Space Phy.,* 20, 851–875.

Murphy, J.R., Stevens, J.L., Kohl, B.S., Benett, T.J., Barker, B.W., 2014. Supplemental analysis of the seismic characteristics of the 2006 and 2009 North Korenuclear tests. In: *Proceeedings of the 2011 Monitoring Research Review:Ground-Based Nuclear Explosion Monitoring Technologies*. NNSA, Tucson,Arizona, pp 513-523.

Nilson, R.H., Peterson, E.W., Lie, K.H., Burkhard, N.R., and Hearst J.R., 1991. Atmospheric pumping: a mechanism causing vertical transport of contaminant gases through fractures permeable media, *J. Geophys. Res.*, 96(B13), 21933–21948, 1991.

Pielke, R. A., Cotton, W. R., Walko, R. L., Tremback, C. J., Lyons, W. A., Grasso, L. D., Nicholls, M. E., Moran, M. D., Wesley, D. A., Lee, T. J., and Copeland, J. H., 1992. A Comprehensive Meteorological Modeling System – Rams, *Meteorol. Atmos. Phys*., 49, 69–91.

Riechmann, B. and Kalinowski, M., 2008. Implications of Xe emissions from medical isotope production, *Geophys. Res. Abstracts*, Vol. 10, EGU2008-A-07898.

Ringbom, A., et al,, 2014. Radioxenon detections in the CTBT international monitoring system likely related to the announced nuclear test in North Koread on February 12 2013. *J. of Environ. Radioactivity,* 128,47-63.

Seibert, P. and A. Frank, 2004. Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, *Atmos. Chem. Phys*., 4, 51–63, 2004.

Stohl, A., M. Hittenberger, and G. Wotawa, 1998. Validation of the Lagrangian particle dispersion model Flexpart against large-scale tracer experiment data, *Atmos. Environ*. 32(24), 4245–4264.

Stunder, Barbara J. B., J. Heffter, R. Draxler, 2007. Airborne Volcanic Ash Forecast Area Reliability. *Weather Forecasting*, 22, 1132–1139.

Torrence, C., and G. P. Compo, 1998. A Practical Guide to Wavelet Analysis. *Bull. of the American Meteorol. Soc.*, 79 (1), 61-78.

Walko, R. L., and L. E. Band, J. Baron, T. G. F. Kittel, R. Lammers, Tsengdar J. Lee, D. Ojima, R. A. Pielke Sr., C. Taylor, C. Tague, C. J. Tremback, and P. L. Vidale, 2000. Coupled Atmosphere–Biophysics–Hydrology Models for Environmental Modeling, *J. Appl. Meteor*., 39, 931–944.

Wotawa, G,, A. Becker, M. Kalinowski, P. Saey, M. Tuma, M. Zähringer, 2010. Computation and analysis of the global distribution of the radioxenon isotope 133Xe based on emissions from nuclear power plants and radioisotope production facilities and its relevance for the verification of the nuclear-test-ban treaty, *Pure and Applied Geophysics*, 167 (4-5), 541-557.

Wright, Christopher, 2013. Low-Yield Nuclear Testing in North Korea in May 2010: Assessing the Evidence with Atmospheric Models and Xenon Activity Calculations*. Science and Global Security*, 21 3-52. DOI: 10.1080/08929882.2013.754310.

Yonezawa C., 2012. Investigation of possible Radioxenon Sources in Japan. INGE Workshop, Ibaraki, Japan, 5-9 November .*Atmos. Environ*, 2007, 41, 4520-4534.