Comparison of Source-Location Algorithms for Atmospheric Samplers

Paul W. Eslinger, W. Steven Rosenthal, Ramesh S. Sarathi, Brian T. Schrom, Erin McCann

Author Affiliations:
Pacific Northwest National Laboratory
902 Battelle Blvd., Richland, Washington, 99354, USA

Abstract

Numerous algorithms have been developed to determine the source characteristics for an atmospheric radionuclide release, e.g., (Bieringer et al., 2017). This study compares three models that have been applied to the data collected by the International Monitoring System operated by the Comprehensive Nuclear-Test-Ban Treaty Organization Preparatory Commission to estimate source event parameters. Each model uses a different approach to estimate the parameters. A deterministic model uses a possible source region (PSR) approach (Ringbom et al., 2014) that is based on the correlation between predicted and measured sample values. A model (now called BAYEST) developed at Pacific Northwest National Laboratory uses a Bayesian formulation (Eslinger et al., 2019; Eslinger et al., 2020; Eslinger and Schrom, 2016). The FREAR model uses a different Bayesian formulation (De Meutter and Hoffman, 2020; De Meutter et al., 2021a; De Meutter et al., 2021b).

The performance of the three source-location models is evaluated with 100 synthetic release cases for the single xenon isotope, ^{133}Xe. The release cases resulted in detections in a fictitious network with 120 noble gas samplers. All three source-location models use the same sampling data.

The two Bayesian models yield more accurate location estimates than the deterministic PSR model, with FREAR having slightly better location performance than BAYEST. Samplers with collection periods of 3, 6, 8, 12, and 24-h were used. Results from BAYEST show that location accuracy improves with each reduction in sample collection length. The BAYEST model is slightly better for estimating the start time of the release. The PSR model has about the same spread in start times as the FREAR model, but the PSR results have a better average start time.

The Bayesian source-location algorithms give more accurate results than the PSR approach, and provide release magnitude estimates, while the base PSR model does not estimate the release magnitude. This investigation demonstrates that a reasonably dense sampling grid will sometimes yield poor location and time estimates regardless of the model. The poor estimates generally coincide with cases where there is a much larger distance between the release point and the first detecting sampler than the average sampler spacing.

Highlights

- Compare performance of three source-location algorithms using synthetic atmospheric samples
- Two Bayesian source-location models have quite similar results
- The Bayesian models give more accurate location estimates than a deterministic probable source region model
- Even a reasonably dense sampling grid sometimes yields poor location and release time estimates
- Shorter sample collection periods provide better location accuracy than longer periods

1. Introduction and background

The Comprehensive Nuclear-Test-Ban Treaty (1996) bans all nuclear explosions anywhere in the world. As of July 2023, 178 states have ratified the treaty, but more states must ratify it for the treaty to enter into force. The Preparatory Commission (CTBTO PrepCom, 2022) for the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) was also established in 1996 to build a global verification regime, the International Monitoring System (IMS) that will monitor treaty compliance when the Treaty enters into force. The IMS operates, or plans to operate, atmospheric radionuclide samplers at 80 locations around the

world. All 80 of the locations will have equipment to measure radioactive isotopes on airborne aerosols and half of them will also sample for noble gases, specifically ^{131m}Xe, ^{133}Xe, ^{133m}Xe, and ^{135}Xe.

This study uses synthetic release data to examine the empirical performance of three source-location algorithms used by analysts to make inferences from IMS atmospheric radioisotope samples. Of course, these three models are a subset of the many different published approaches (Bieringer et al., 2017; Fleming et al., 2012; Hutchinson et al., 2017; Rao, 2007; Redwood, 2011). The comparison considers only the isotope ^{133}Xe and estimates of the location (latitude and longitude), magnitude, and start time of the release. The first model (BAYEST) uses a Bayesian formulation (Eslinger et al., 2019; Eslinger et al., 2020; Eslinger and Schrom, 2016). The second model, Forensic Radionuclide Event Analysis and Reconstruction (FREAR) uses a different Bayesian formulation (De Meutter and Hoffman, 2020; De Meutter et al., 2021a; De Meutter et al., 2021b). The third model uses a possible source region (PSR) approach (Ringbom et al., 2014) that is based on the correlation between predicted and measured sample values.

2. Approach and models

Few controlled atmospheric release experiments are available, so the performance of the three source-location models is evaluated with 100 synthetic release cases. The analysis of each release case is considered independent of the other release cases, thus there is no overlap in the released plumes. The release cases resulted in detections in a hypothetical sampling network. All three models use the same sampling data in the analysis.

2.1. Release and sampling locations

A fairly dense sampling network is desired so that most releases will result in detections at more than one sampler. Choosing a constrained region will result in some releases exiting the region with no detections. The experiment design only considered releases that resulted in at least two detections.

A hypothetical network of 120 samplers was defined in the 48 contiguous states in the United States of America. The sampler locations are denoted by filled circles in Fig. 1. Then, 100 release locations were randomly selected inside the 48 contiguous states. The release locations are denoted by stars in Fig. 1. Finally, a release start time was selected randomly between January 15 and December 15, 2021, for each release event. Even though several of the release locations shown in Fig. 1 are close to each other, the releases occurred at different times and under different meteorological conditions. Each release occurred at ground level and lasted 3 h at a constant release rate. The total amount released for every event was 8.92×10^{12} Bq of ^{133}Xe. This amount is two to three orders of magnitude larger than the typical daily emission by a nuclear power plant, but some medical isotope production facilities emit up to two orders of magnitude more than this much ^{133}Xe in a day (Kalinowski, 2023). The exact magnitude of the release is immaterial for this study using synthetic releases, provided that it is large enough to produce detections at several samplers.

2.2. Atmospheric transport modeling

Atmospheric transport of the simulated release quantities was modeled using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Stein et al., 2015) maintained by the National Oceanic and Atmospheric Administration (NOAA). The transport runs were performed on the Linux version (revised in 2021) of HYSPLIT. The top of the atmospheric model domain was set to 10,000 m above ground level and the model used 45 vertical layers. Wet and dry deposition mechanisms were deactivated for the transport runs because xenon is a noble gas and its concentration in air typically does not depend on rainout or deposition processes (Grochala, 2007). Default model parameters (Draxler et al., 2020) were used for HYSPLIT.

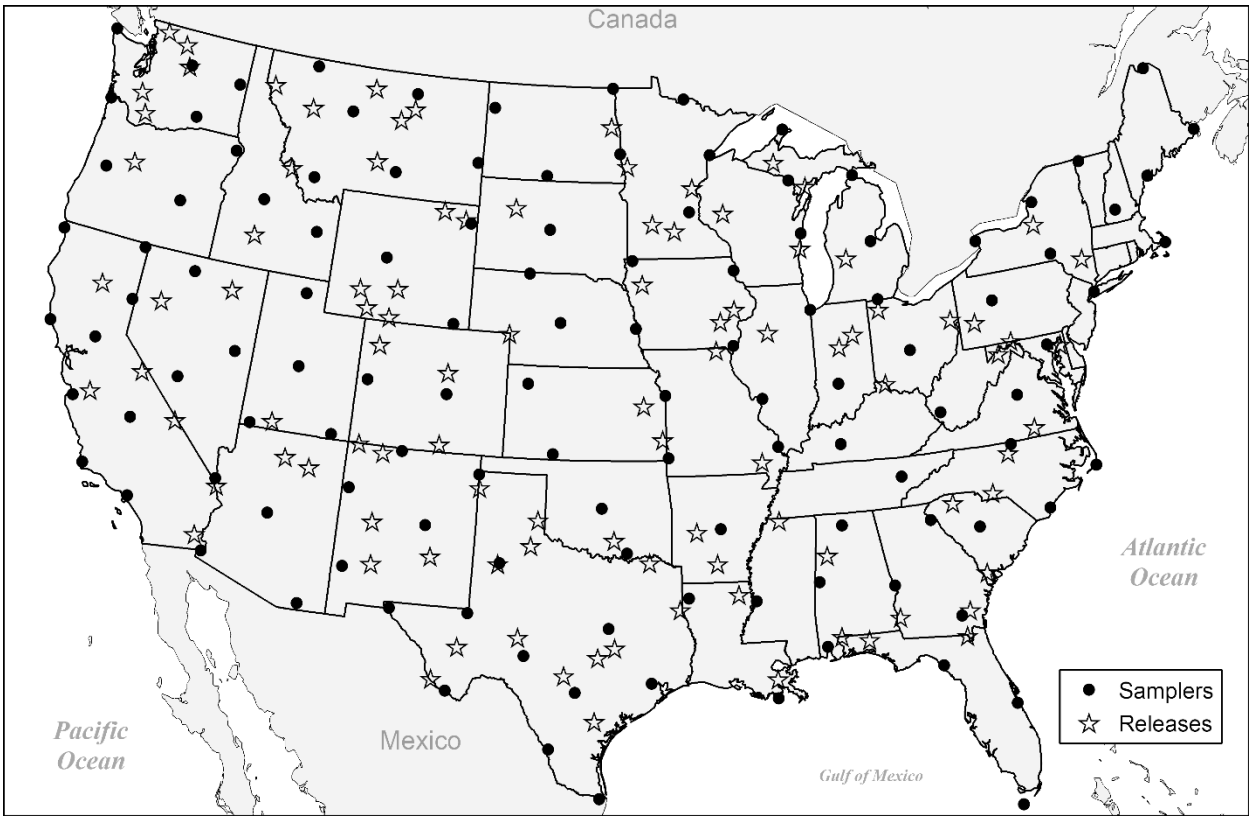

Fig. 1. Radionuclide sampler and release locations.

The transport model used archived global meteorological data on a 0.25-degree spatial grid and a 3-h time resolution (GFS, 2020). The 3-h releases were simulated using a ground-level point source and 5×10^6 particles. To support post-processing concentrations, atmospheric dilution factor data were output each hour on a 0.125° latitude-longitude grid for 144-h (6-d) after the start of the release. Dilution factors were averaged over the bottom 100 m of the atmospheric column. These data were then interpolated to specific sampler locations.

2.3. Sampler characteristics

Several designs of noble gas samplers are operational in the IMS. The Swedish-designed (SAUNA-II) system (Ringbom et al., 2003) uses a 12-h collection period. The next-generation Swedish-designed (SAUNA III) system (Ringbom et al., 2017), now in use at a few stations, uses a 6-h collection period. The currently-deployed French-designed SPALAX™ system (Fontaine et al., 2004) uses a 24-h collection period. The next-generation French system, (Topin et al., 2020) denoted by SPALAX NG has a variable collection time, but performance results have been published for an 8-h collection time. The American-designed Xenon International system uses a 6-h collection period (Brander et al., 2022). The currently-deployed Russian designed ARIX system (Dubasov et al., 2005) uses a 12-h collection period.

For this analysis, synthetic samples for every release were developed for samplers with 3, 6, 8, 12, and 24-h collection periods. No sampler in the IMS currently uses a 3-h collection period, but the analysis presented here uses samples developed for a 3-h collection period. This study focuses on comparing the different source-location algorithms, rather than sampler differences other than the length of the collection period; thus, each sample used a detection limit of 0.15 mBq/m^3.

2.4. Atmospheric transport for individual samples

All of the source location models use the effective atmospheric dilution factors between each sample and the suite of possible times and locations where the release might have occurred. This analysis used time-reversed atmospheric transport runs that start with a unit release over the sample collection period and go back in time to generate the dilution factors. Using this approach, only one atmospheric transport run is needed for each sample (Hourdin and Talagrand, 2006; Hourdin et al., 2006; Rao, 2007; Stohl et al., 2002). For forward-time atmospheric runs, it is necessary to run a separate model for every release time and location of a release that will be considered, although other methods can be utilized, e.g., (Rosenthal et al., 2022). However, it can be prohibitively expensive to do all the calculations.

The atmospheric transport runs for each sample used the HYSPLIT code and the same meteorological data set (GFS, 2020) as the runs used to generate the synthetic samples. However, these runs only used 10^6 particles. Even with the reduced number of particles, and ignoring most of the samples with no detections of ^{133}Xe, computing this suite of time-reversed atmospheric transport model runs took more than 2 months on a dedicated computer system with 168 cores. After the samplers with detections were identified, another eight samplers without detections were included. These eight were chosen to be the closest samplers (in distance) to the samplers with detections. The general thought was that these samplers without detections help define the outer boundaries of the plume resulting from the release.

2.5. Source-location models

The mathematical formulations of the source-term estimation models are extensive, so they are not replicated in this paper. The mathematical formulation for BAYEST is given in Section 2.3 of Eslinger et al. (2020). The underlying mathematics for FREAR is provided in Section 3 of De Meutter and Hoffman (2020). The PSR approach is briefly described in Section 5.1 of Ringbom et al. (2014).

Both BAYEST and FREAR are Bayesian models, so the user must specify a prior distribution on the model parameters (e.g., release time, location, magnitude, and duration). Uniform prior distributions were used for these parameters. Realistic limits on the uniform distributions were imposed. For example, the release location was constrained to a box containing the United States, and the release time was in the 36 hours before the time of the first detection.

All three source-location models used outputs from the same atmospheric transport model runs to collect atmospheric dilution factors in what is sometimes called a source-receptor matrix. The PSR approach uses a correlation approach without tunable parameters. However, both BAYEST and FREAR need the uncertainty in the sampled concentrations. The uncertainty models are different, but the authors attempted to use consistent uncertainty levels. For example, the sample data uncertainty for both models used the maximum of the sample minimum detectable concentration and 15% of the sample concentration. These values are in the range of uncertainty for current sampling techniques.

3. Results

This section contains summary comparisons of the location, time, and magnitude accuracy of each of the source-location algorithms.

3.1. Release location results

The first location result uses the distance (in km) between the release location and the best single location predicted by the model. The Bayesian formulation in the BAYEST and FREAR models calculate the 5-dimensional posterior probability distribution for the source parameters (latitude, longitude, magnitude, start time, stop time or duration). However, each model outputs summary information on the marginal distributions for individual parameters rather than the entire distribution. Thus, comparisons use different statistics. For FREAR, the best location is chosen using the mean value of the posterior latitude

and longitude probability distributions. For BAYEST, the best location is chosen as the location of the (latitude, longitude, start time, release length, magnitude) combination that results in the largest likelihood ratio (highest posterior probability). For PSR, the best location is chosen as the location that results in the highest correlation between the measured and modeled samples. The BAYEST and PSR models do a grid search in their solutions, where the grid uses 0.05° increments (approximately 4 km in the east-west direction and 5 km in the north-south direction). The FREAR model treats location as continuous in latitude and longitude.

The set of offset distances for each model is shown in Fig. 2. Each of the 100 values for the individual models are the average of the results using 3, 6, 8, 12, and 24-h collection periods. A perfect model would show as a vertical line at an offset of zero km. The FREAR model is slightly better than the BAYEST model, but both Bayesian models are substantially better than the PSR model. For example, only 40% of the PSR cases have a location offset of 100 km or less. By comparison, 66% of the BAYEST and 74% of the FREAR cases have a location offset of 100 km or less.

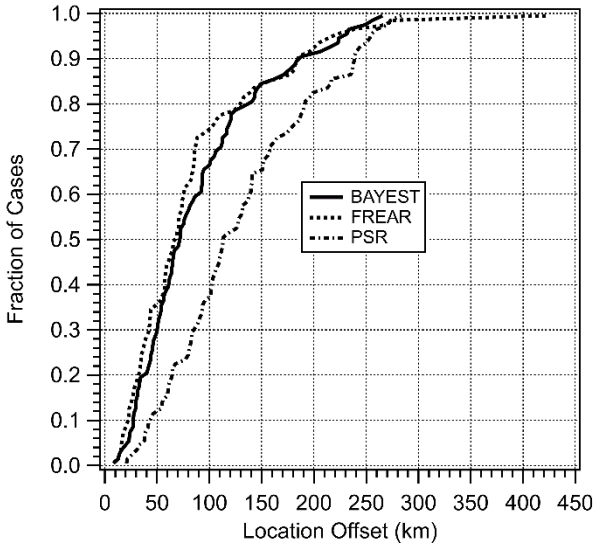

Fig. 2. Comparison of the distance (in km) between the modeled release locations and the predicted release locations for the three source-location models. Each of the 100 values for the individual models are the average of the results using 3, 6, 8, 12, and 24-h collection periods. A perfect model would show as a vertical line at an offset of 0 km.

The set of offset distances between the modeled release location and the predicted release location for the BAYEST and FREAR models is shown in Fig. 3 for each of the 3, 6, 8, 12, and 24-h collection periods. The average curves shown in Fig. 2 indicate that the BAYEST and FREAR models have very similar average performance. However, the BAYEST model shows consistently better location performance for shorter collection periods than for longer collection periods. The FREAR model shows similar performance for predicted locations within 100 km of the release location, but consistently poorer performance than BAYEST outside of 100 km when the collection period is less than 8 h. The reason for this somewhat unexpected difference is not known.

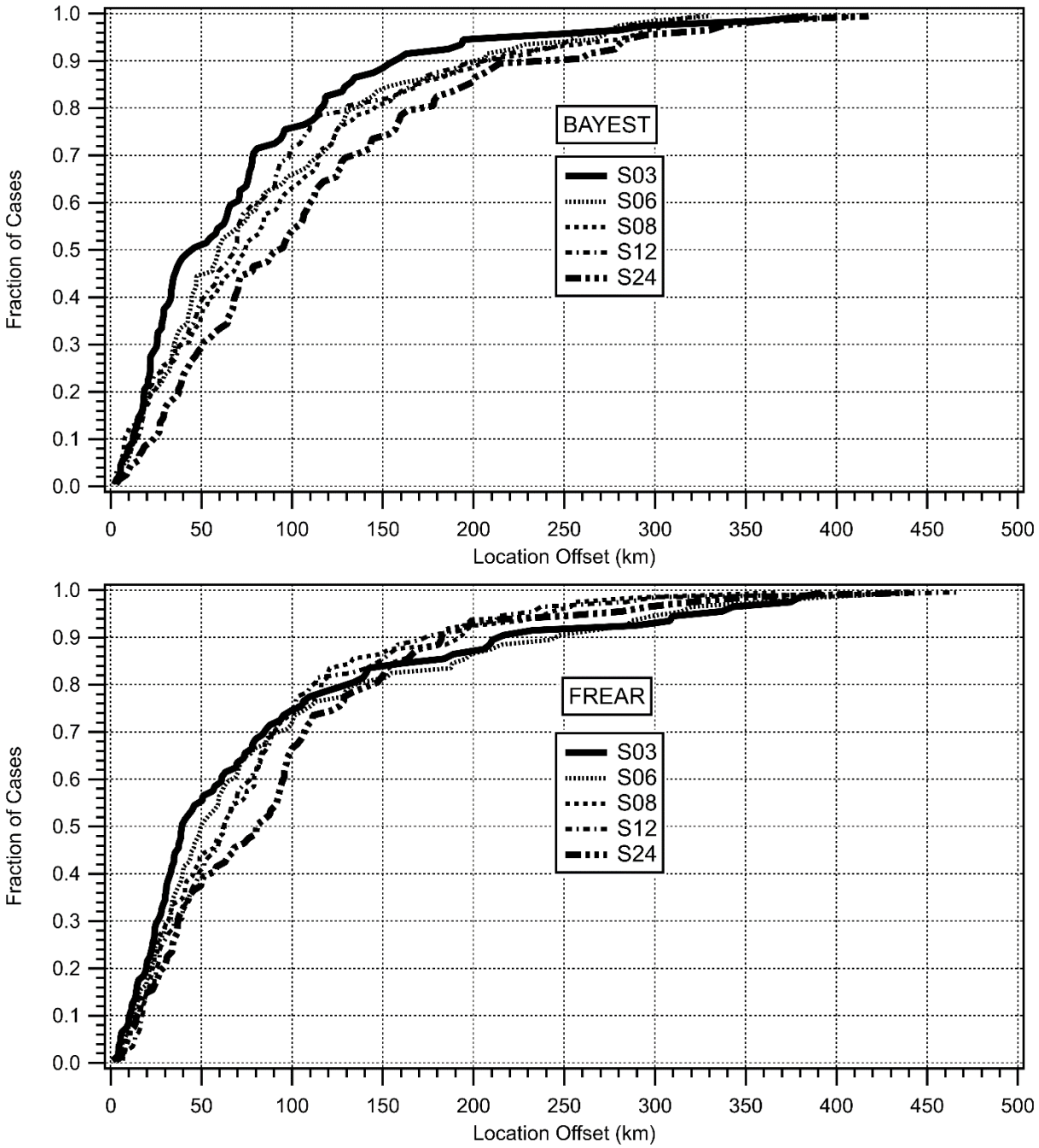

Fig. 3. Comparison of the distance (in km) between the modeled release locations and the predicted release locations for the BAYEST (top pane) and FREAR (bottom pane) source-location models for the sampling data using 3, 6, 8, 12, and 24-h collection periods. A perfect model would show as a vertical line at an offset of 0 km.

The data in Fig. 2 and Fig. 3 are based on point estimates of the release location. The Bayesian models calculate a probability distribution for the release location. Fig. 4 shows the posterior probability that the predicted release location is within 100 km of the modeled release location for the BAYEST source-location model for the sampling data using 3, 6, 8, 12, and 24-h collection periods. It is desirable to have a high probability that the predicted release location is within 100 km of the modeled release location; thus, the best possible performance for this metric would be a vertical line on the right side of the plot probability

1.0. This plot shows a consistent degradation in the posterior probability estimates as collection periods lengthen from 3-h to 24-h. If each curve is summarized by an average value, for 3, 6, 8, 12, and 24-h collections, respectively, the average posterior probabilities the release is within 100 km of the true release point are 0.602, 0.533, 0.458, 0.399, and 0.275.

There is evidence that shorter collection time periods, such as 6, 8, or 12- h, rather than 24-h, improve network performance (Eslinger and Miley, 2022; Eslinger et al., 2023) in detecting that a release has occurred. The current work expands on that body of evidence using the source-location estimation codes rather than simply demonstrating that shorter collection periods should result in more samples with detectable concentrations.

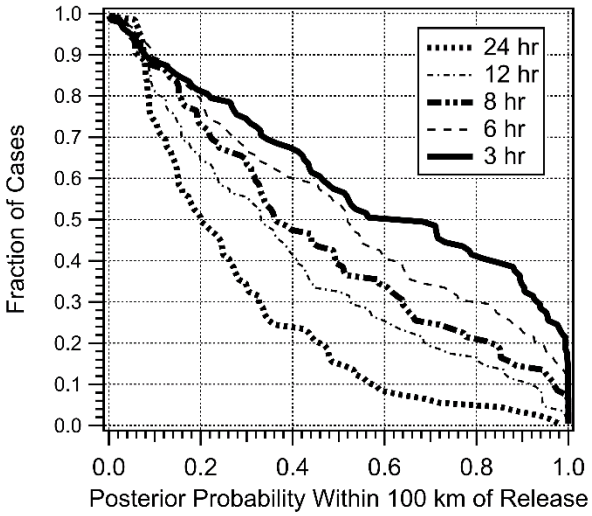

Fig. 4. Posterior probability that the predicted release location is within 100 km of the modeled release location for the BAYEST source-location model for the sampling data using 3, 6, 8, 12, and 24-h collection periods. The best possible performance for this metric would be a vertical line on the right axis at probability 1.0.

Fig. 3 shows that between 50% and 75% of the release cases result in a location estimate that is within 100 km of the modeled release location, depending on the choice of sampler collection period length and the source location model. As shown in Fig. 1, this study uses 120 samplers in the continental United States. The sampling network has a minimum spacing of about 200 km, and an average distance to the three closest neighbors of about 280 km. As noted in the introduction, the IMS uses 80 radionuclide samplers over the entire globe. However, the IMS was initially designed to have a high probability of detecting releases from a 1 kt nuclear explosion (IMS Expert Group, 1995) rather than attempting to use the radionuclide samples to locate the release point. The releases in this study were about 3 orders of magnitude smaller than the initial ^{133}Xe release value considered for the design of the IMS, and every release was detected multiple times in the fictional sampling network.

In general, the forward-time and reversed-time atmospheric transport runs do not yield identical dilution factors (Eslinger and Schrom, 2019) but they are close in many cases. In other cases, they may differ significantly, especially over mountainous terrain. Thus, one possible reason for poor location performance is the difference in the atmospheric transport runs that produced the synthetic sampling data (which used the time-forward mode) and the runs used to develop dilution factors for the source-location models (which used the time-reversed mode).

Another major reason the predicted release location is sometimes poorly estimated is the spatial relationship of the air-borne release plume and the sampling stations. One such case is illustrated in Fig. 5 using 6-h sampling data. The maximum value of the release plume skirts the four closest sampling locations, results in single samples near the detection limit at two other sampling locations, and the first “strong”

sample concentration is obtained about 1100 km from the release location. A comparison for the location discrepancies for each of the models is provided in Table 1 for this release case. FREAR does a little better for this case, but the results are highly variable.

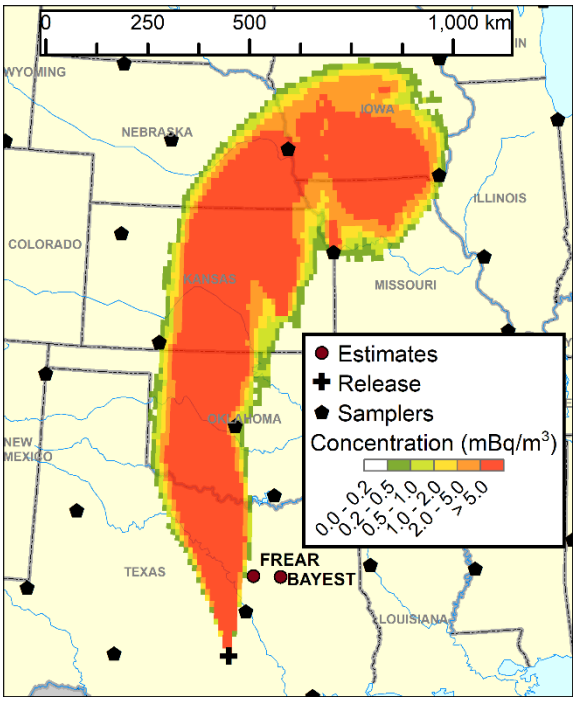

Fig. 5. Trace of the maximum surface-level concentration of ^{133}Xe for plume movement for 30 hours after the release. The plume skirts the closest samplers and registers the first strong concentration in a sampler 1100 km from the release point. The BAYEST and FREAR models pick a release point on the order of 200 km in error using 6-h sampling data.

Table 1
Delta location (km) between release point and estimated release point for one release case where the released plume skirts the closest samplers.

Sampling Period (h)	BAYEST	PSR	FREAR
3	280	317	32
6	228	228	199
8	74	74	83
12	165	165	233
24	278	278	147

3.2. Release time results

An event time metric uses the time difference (h) between the release start time and the best release start time predicted by the model. The codes output different summary statistics, so slightly different measures of the release time are used. For FREAR, the best time is chosen using the mean value of the posterior release start time probability distribution. For BAYEST, the best time is chosen as the time associated with the largest likelihood ratio (highest posterior probability). For PSR, the best time is chosen as the release start time that results in the highest correlation between the measured and modeled samples, using the joint location-time solution. The set of release start time discrepancies for each model is shown

in Fig. 6. Each of the 100 values for the individual models are the average of the results using 3, 6, 8, 12, and 24-h collection periods.

A time offset of 0 is a perfect match. Negative times in Fig. 6 indicate a release start time that is before the release event start time. Using this metric, the BAYEST model performs the best, with only 3 cases out of 100 with average start times more than 12 hours in error, compared to 9 cases for the PSR model and 15 cases for the FREAR model. The PSR model has about the same spread in delta times as the FREAR model, but the PSR results have an average value of -0.1 h while FREAR has an average value of -5.4 h. The BAYEST model has an average delta time of -0.6 h. Part of the different between FREAR and the other models may be due to comparing different statistics (mean value versus value from highest posterior probability).

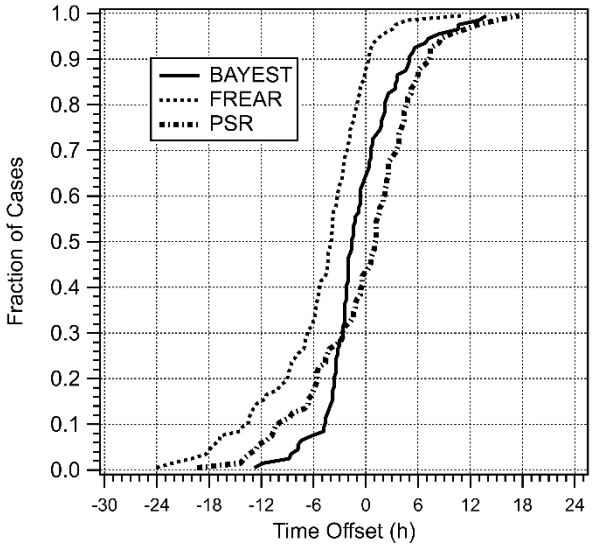

Fig. 6. Comparison of the time offset (h) between the modeled release start times and the predicted release start times for the three source-location models. Each of the 100 values for the individual models are the average of the results using 3, 6, 8, 12, and 24-h collection periods. A time offset of 0 is a perfect match.

The mean and standard deviation of the time offset (h) between the modeled release start times and the predicted release start times for the three source-location models are provided in Table 2 for different sample collection lengths. On average, BAYEST has a smaller offset time and lower standard deviation, but different models have better performance in different sampling collection times. The smallest bias on the release time estimate occurs for the medium length sampling periods, while naively, one would expect the smallest bias for the shorter sampling times. More work is needed to advance a reason for this result.

Table 2
Means and standard deviations of the time offset (h) between the modeled release start times and the predicted release start times for the three source-location models. The statistics for each collection period length are based on 100 release events.

	BAYEST		**PSR**		**FREAR**	
Collection	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
3-h	-1.38	4.97	1.44	10.34	-5.94	8.32
6-h	-1.21	5.92	1.04	9.88	-5.02	7.83
8-h	-0.38	6.48	0.04	9.24	-4.86	7.21
12-h	-0.12	6.05	-0.32	9.65	-4.72	6.51
24-h	-0.41	8.89	-2.69	12.31	-6.39	6.88
Average	-0.57	4.65	0.25	7.01	-4.87	6.39

3.3. Release magnitude results

Fig. 7 shows a comparison of the estimated release magnitudes for the three source-location models. The BAYEST and FREAR models develop the posterior probability distribution for the release magnitude but output different summary statistics, so slightly different measures of the release magnitude are also used. For FREAR, the best magnitude is chosen using the mean value of the posterior magnitude probability distribution. For BAYEST, the best magnitude is chosen as the magnitude associated with the largest likelihood ratio (highest posterior probability). The PSR model uses a correlation approach where the magnitude cancels out, so the estimation process does not produce a magnitude. For purposes of this analysis, a magnitude is assigned to the results from PSR using a two-step approach. First, the location and time that gives the highest correlation is selected. Then, the magnitude is selected so that the average value of the predicted sample values is the same as the average value of the measured sample values.

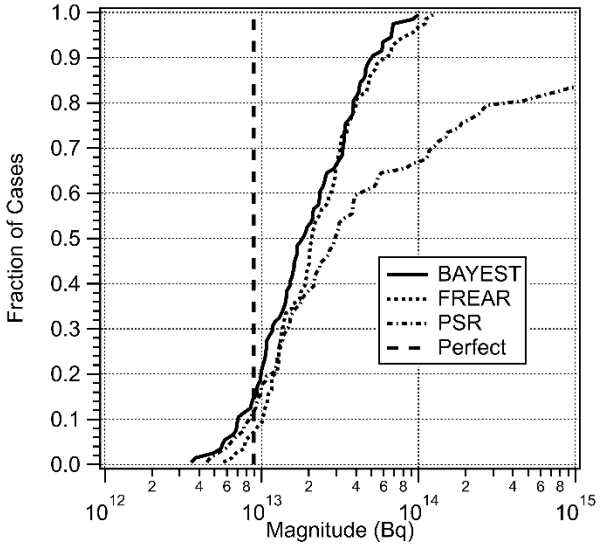

Fig. 7. Comparison of the estimated release magnitudes of ^{133}Xe (Bq) for the three source-location models. Each of the 100 values for the individual models are the average of the results using 3, 6, 8, 12, and 24-h collection periods. The released amount was 8.92×10^{12} Bq.

The estimated magnitudes for the BAYEST and FREAR models are very similar. Both generally appear to be too high compared to the release amount of 8.92×10^{12} Bq. However, about 95% of the cases are within an order of magnitude of the release amount, shown by the vertical dashed line in Fig.7 The results from the modified PSR method are reasonable for some cases, though 34% of the magnitudes are more than an order of magnitude too large, and 6% of the cases are more than 3 orders of magnitude too large.

3.4. High-probability (plausible) intervals

The BAYEST model can output plausible intervals that are designed to contain the actual release start time and magnitude with a high level of certainty. This section discusses the accuracy of the uncertainty interval rather than the accuracy of the time and magnitude estimates. Specifically, these runs were set to produce 95% plausible intervals. For magnitude, 85% of the 500 plausible intervals included the release magnitude. This suggests that the 95% plausible intervals for magnitude produced by the code are slightly too narrow because wider intervals would include the release magnitude more often. For the release start time, 92% of the 500 plausible intervals included the release start time, so the 95% plausible intervals for the start time are approximately correct. The BAYEST code doesn't currently have a function to produce

an output that is comparable to a simple plausible interval on location, although it can calculate the probability that the estimated release location is within a specified distance from a known release location.

The FREAR model outputs the 2.5% and 97.5% levels of the parameter distributions, which is equivalent to a 95% plausible interval. For magnitude, 63% of the intervals for magnitude included the actual release magnitude. For the release start time, 61% of the intervals included the release start time. For location, 53% of the implied plausible intervals for location included the release location. Thus, the plausible intervals for location, time, and magnitude seem to be too narrow.

4. Discussion and conclusions

A large numerical experiment with 100 synthetic release cases and a fictitious network of 120 samplers distributed throughout the United States with an average minimum spacing of 280 km was used to compare the performance of three source-location models. The three algorithms are BAYEST, FREAR, and PSR. Each of the 100 release cases were evaluated using samples with 3, 6, 8, 12, and 24-h collection periods.

The BAYEST and FREAR source-location algorithms had very similar performance, with FREAR slightly better on the location estimates while BAYEST was slightly better on the release start time estimates. The PSR model was not as good as either Bayesian model on location estimates but was in the middle of the three models on release start time estimates. The two Bayesian models had very similar performance for the release magnitude estimates. The PSR model does not directly provide a release magnitude estimate, and a modification matching the average sample concentrations gave poor results for more than half of the release cases. Different model formulations might provide better overall magnitude estimates for a PSR-like model.

Other simulation studies have indicated that collection time periods shorter than 24-h improve the IMS network performance (Eslinger and Miley, 2022; Eslinger et al., 2023) for detecting the occurrence of releases. This study further shows, in Fig. 4, that shorter collection periods provide more accurate release location estimates.

As shown in Fig. 2 and Fig. 6, many release cases had large location or time discrepancies when compared to the simulated release events. Examination of several cases with the poorest performance all showed wind patterns where the release plume bypassed the closest samplers and moved a large distance before it encountered a sampling location.

Overall, the two Bayesian models provided better solutions than the PSR model. Even though the PSR model was competitive on the release start time analysis, as shown in Fig. 6, it is not as good on the release location estimates. For example, 50% of the FREAR cases were within 68 km of the release location. Similarly, 50% of the BAYEST cases were within cases 72 km, while 50% of the PSR results were within 113 km of the release location.

For this study, the PSR model was the easiest of the three methods to implement and run. Unfortunately, the PSR model doesn't have an integrated release magnitude algorithm. In situations where a release magnitude estimate is desired, one could employ a Bayesian model, or use the PSR results on location and time as starting points for a more complex release estimation model.

Even a reasonably dense sampling grid will sometimes yield poor location and time estimates and the poor estimates are not unique to a particular model. The poor estimates generally coincide with cases where there is a much larger distance between the release point and the first detecting sampler than the average sampler spacing. This is a significant source of error in each method, with approximately 30% of the cases for BAYEST and FREAR having a location error of 100 km or more, while the PSR method has a 30% chance of a location error of about 160 km or more. While this simulation study found consistent performance between the FREAR and BAYEST algorithms, results may vary for other geographies, climates, and sampler densities.

Acknowledgments

This Low Yield Nuclear Monitoring (LYNM) research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, NNSS, PNNL, and SNL. The authors gratefully acknowledge the NOAA Air Resources Laboratory for provision of the HYSPLIT transport and dispersion model used in this publication and the associated meteorological data.

References

Bieringer, P.E., Young, G.S., Rodriguez, L.M., Annunzio, A.J., Vandenberghe, F., Haupt, S.E., 2017. Paradigms and commonalities in atmospheric source term estimation methods. Atmos. Environ. 156, 102-112. doi:10.1016/j.atmosenv.2017.02.011

Brander, S., Baur, S., Krais, R., Ross, J.O., Orr, A., Sayne, R., Howard, M., Mayer, M., Panisko, M., Hayes, J.C., Bollhöfer, A., 2022. Phase II testing of Xenon International on Mount Schauinsland, Germany. J. Environ. Radioact. 255, 107034. doi:10.1016/j.jenvrad.2022.107034

Comprehensive Nuclear-Test-Ban Treaty. 1996. Text of the Comprehensive Nuclear-Test-Ban Treaty. United Nations Office for Disarmament Affairs (UNODA), Status of Multilateral Arms Regulation and Disarmament Agreements, CTBT. Accessed on September 20, 2012, at http://www.ctbto.org/the-treaty/treaty-text/

CTBTO PrepCom. 2022. Comprehensive Nuclear-Test-Ban-Treaty Organization (CTBTO) Preparatory Commission Web Page. Accessed on July 22, 2022, at http://www.ctbto.org/

De Meutter, P., Hoffman, I., 2020. Bayesian source reconstruction of an anomalous Selenium-75 release at a nuclear research institute. J. Environ. Radioact. 218, 106225. doi:10.1016/j.jenvrad.2020.106225

De Meutter, P., Hoffman, I., Ungar, K., 2021a. How can we determine the origin of radionuclide observations? Presenting the Bayesian source reconstruction algorithm "FREAR", CTBT: Science and Technology Conference 2021, Viernna, Austria, https://conferences.ctbto.org/event/7/contributions/992/attachments/58/146/P2.4-373_DeMeutter.pdf

De Meutter, P., Hoffman, I., Ungar, K., 2021b. On the model uncertainties in Bayesian source reconstruction using an ensemble of weather predictions, the emission inverse modelling system FREAR v1.0, and the Lagrangian transport and dispersion model Flexpart v9.0.2. Geosci. Model Dev. 14(3), 1237-1252. doi:10.5194/gmd-14-1237-2021

Draxler, R.R., Stunder, B., Rolph, G., Stein, A., Taylor, A., 2020. HYSPLIT4 User's Guide, Air Resources Laboratory, National Oceanic and Atmospheric Administration (NOAA), Silver Spring, Maryland, https://www.arl.noaa.gov/documents/reports/hysplit_user_guide.pdf

Dubasov, Y.V., Popov, Y.S., Prelovskii, V.V., Donets, A.Y., Kazarinov, N.M., Mishurinskii, V.V., Popov, V.Y., Rykov, Y.M., Skirda, N.V., 2005. The АРИКС-01 Automatic Facility for Measuring Concentrations of Radioactive Xenon Isotopes in the Atmosphere. Instrum Exp Tech 48(3), 373-379. doi:10.1007/s10786-005-0065-3

Eslinger, P.W., Lowrey, J.D., Miley, H.S., Rosenthal, W.S., Schrom, B.T., 2019. Source term estimation using multiple xenon isotopes in atmospheric samples. J. Environ. Radioact. 204, 111-116. doi:10.1016/j.jenvrad.2019.04.004

Eslinger, P.W., Lowrey, J.D., Miley, H.S., Rosenthal, W.S., Schrom, B.T., 2020. Source type estimation using noble gas samples. J. Environ. Radioact. 225, 106439. doi:10.1016/j.jenvrad.2020.106439

Eslinger, P.W., Miley, H.S., 2022. Projected network performance for next-generation xenon monitoring systems. J. Environ. Radioact. 251-252, 106976. doi:10.1016/j.jenvrad.2022.106976

Eslinger, P.W., Miley, H.S., Burnett, J.L., Lidey, L.S., Mendez, J.M., Schrom, B.T., Sharma, M.K., 2023. Projected network performance for next generation aerosol monitoring systems. J. Environ. Radioact. 257, 107088. doi:10.1016/j.jenvrad.2022.107088

Eslinger, P.W., Schrom, B.T., 2016. Multi-detection events, probability density functions, and reduced location area. J Radioanal Nucl Chem 307(3), 1599-1605. doi:10.1007/s10967-015-4339-3

Eslinger, P.W., Schrom, B.T., 2019. Utility of atmospheric transport runs done backwards in time for source term estimation. J. Environ. Radioact. 203, 98-106. doi:10.1016/j.jenvrad.2019.03.006

Fleming, Z.L., Monks, P.S., Manning, A.J., 2012. Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition. Atmospheric Research 104(Supplement C), 1-39. doi:10.1016/j.atmosres.2011.09.009

Fontaine, J.P., Pointurier, F., Blanchard, X., Taffary, T., 2004. Atmospheric xenon radioactive isotope monitoring. J. Environ. Radioact. 72(1–2), 129-135. doi:10.1016/S0265-931X(03)00194-2

GFS. 2020. Concatenated short-term NCEP Global Forecast System (GFS) forecast model archive. Air Resources Laboratory, National Oceanic and Atmospheric Administration. Accessed on January 10, 2020, at ftp://arlftp.arlhq.noaa.gov/archives/gfs0p25/

Grochala, W., 2007. Atypical compounds of gases, which have been called 'noble'. Chemical Society Reviews 36(10), 1632-1655. doi:10.1039/B702109G

Hourdin, F., Talagrand, O., 2006. Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parametrization of subgrid-scale transport. Q. J. Roy. Meteor. Soc. 132(615), 567-583. doi:10.1256/qj.03.198.A

Hourdin, F., Talagrand, O., Idelkadi, A., 2006. Eulerian backtracking of atmospheric tracers. II: Numerical aspects. Q. J. Roy. Meteor. Soc. 132(615), 585-603. doi:10.1256/qj.03.198.B

Hutchinson, M., Oh, H., Chen, W.-H., 2017. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. Information Fusion 36, 130-148. doi:10.1016/j.inffus.2016.11.010

IMS Expert Group, 1995. International Monitoring System Expert Group Report based on Technical Discussions held from 6 February to 3 March 1995, Conference on Disarmament. United Nations, Geneva, Switzerland, p. 97, https://digitallibrary.un.org/record/205422/files/CD_NTB_WP.224-EN.pdf

Kalinowski, M.B., 2023. Global emission inventory of 131mXe, 133Xe, 133mXe, and 135Xe from all kinds of nuclear facilities for the reference year 2014. J. Environ. Radioact. 261, 107121. doi:10.1016/j.jenvrad.2023.107121

Rao, K.S., 2007. Source estimation methods for atmospheric dispersion. Atmos. Environ. 41(33), 6964-6973. doi:10.1016/j.atmosenv.2007.04.064

Redwood, M., 2011. Source Term Estimation and Event Reconstruction: A Survey, Dstl/CR51790, Atmospheric Dispersion Modelling Liaison Committee Report: ADMLC-R6-2011-1, Health Protection Agency, Chilton, Didcot, UK, pp. 96-151, https://admlc.files.wordpress.com/2014/05/admlc-r6.pdf

Ringbom, A., Aldener, M., Axelsson, A., Fritioff, T., Kastlander, J., Mortsell, A., Olsson, H., 2017. Analysis of data from an intercomparison between a Sauna II and a Sauna III system, Science and Technology 2017 Conference, Vienna, Austria, https://ctnw.ctbto.org/DMZ/abstract/21838

Ringbom, A., Axelsson, A., Aldener, M., Auer, M., Bowyer, T.W., Fritioff, T., Hoffman, I., Khrustalev, K., Nikkinen, M., Popov, V., Popov, Y., Ungar, K., Wotawa, G., 2014. Radioxenon detections in the CTBT international monitoring system likely related to the announced nuclear test in North Korea on February 12, 2013. J. Environ. Radioact. 128(0), 47-63. doi:10.1016/j.jenvrad.2013.10.027

Ringbom, A., Larson, T., Axelsson, A., Elmgren, K., Johansson, C., 2003. SAUNA—a system for automatic sampling, processing, and analysis of radioactive xenon. Nucl Instrum Meth A 508(3), 542-553. doi:10.1016/s0168-9002(03)01657-7

Rosenthal, W.S., Eslinger, P.W., Schrom, B.T., Miley, H.S., Baxter, D.J., Fast, J.D., 2022. Enabling probabilistic retrospective transport modeling for accurate source detection. J. Environ. Radioact. 247, 106849. doi:10.1016/j.jenvrad.2022.106849

Stein, A.F., Draxler, R.R., Rolph, G.D., Stunder, B.J.B., Cohen, M.D., Ngan, F., 2015. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. Bulletin of the American Meteorological Society 96(12), 2059-2077. doi:10.1175/BAMS-D-14-00110.1

Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., Seibert, P., 2002. A replacement for simple back trajectory calculations in the interpretation of atmospheric trace substance measurements. Atmos. Environ. 36(29), 4635-4648. doi:10.1016/S1352-2310(02)00416-8

Topin, S., Gross, P., Achim, P., Generoso, S., Cagniant, A., Delaune, O., Morin, M., Philippe, T., Fontaine, J.-P., Moulin, C., Douysset, G., Le Petit, G., 2020. 6 months of radioxenon detection in western Europe with the SPALAX-New generation system - Part1: Metrological capabilities. J. Environ. Radioact. 225, 106442. doi:10.1016/j.jenvrad.2020.106442