Evaluation of methodologies for forming a global view of radionuclide release events and their application to the data fusion pipeline at the CTBTO

Joshua Kunkle, Boxue Liu, Martin Kalinowski, Yuichi Kijima, Jolanta Kusmierczyk-Michulec, Mark Prior, Robin Schoemaker, Anne Tipka
Contact : joshua.kunkle@ctbto.org

wosmip

CTBTO PREPARATORY COMMISSION — COMPREHENSIVE NUCLEAR-TEST-BAN TREATY ORGANIZATION

Goal

Introduce refined input fields to the data fusion algorithm which are based on a global radionuclide (RN) event hypothesis. This should result in more accurate fusion results with reduced number of fused SHI events. Two methods have been investigated as prototypes for further studies.

The data fusion automatic pipeline

Currently SHI events are fused with the entire source-receptor-sensitivity (SRS) field of each RN station. In addition, a delayed release of 60 days is included.

The Problem

Data fusion does not provide a global view – fusion is applied for each RN Sample/SRS file and SHI event including all possible matches for 60 days. Too many of the data fusion matches are spurious and analysis is therefore overly cumbersome.

Concept

Estimate the probability of a release of in each *cell*, defined by the SRS granularity of 0.5 x 0.5 x 1 hour based on measurements at RN stations and ATM simulations. Groupings of high-probability cells can represent RN event hypotheses.

Motivated by Bayes Theorem, determine the probability of a release $P(R(lat, lon, t, I))$ of isotope *I* for a given grid cell at location (lat, lon, t). Here a flat prior is assumed.

Input Data

- SRS fields – backwards ATM simulations using Flexpart provide a dilution factor, $M(lat, lon, t)$, at each cell for each RN measurement, up to 14 days before the measurement time
- RN measurements – minimum detectable concentration (MDC) or measured activity concentration from Reviewed Radionuclide Report (RRR)
- Half-lives of relevant radionuclides

References

[1] P. Eslinger, H Miley, B. Schrom. Investigations of association among atmospheric radionuclide measurements. Journal of Environmental Radioactivity 241 (2022)
[2] H. Miley. The Development of an Association Algorithm for Radionuclide Measurements. SnT 2023 presentation, O3.6-364
[3] Machine Learning Enhanced Detection of Radionuclide Anomalies. SnT 2023 presentation O3.6-173
[4] P. De Meutter, Ian Hoffman. Bayesian Source Reconstruction of an Anomalous Selenium-75 Release at a Nuclear Research Institute. Journal of Environmental Radioactivity 218 (2020)

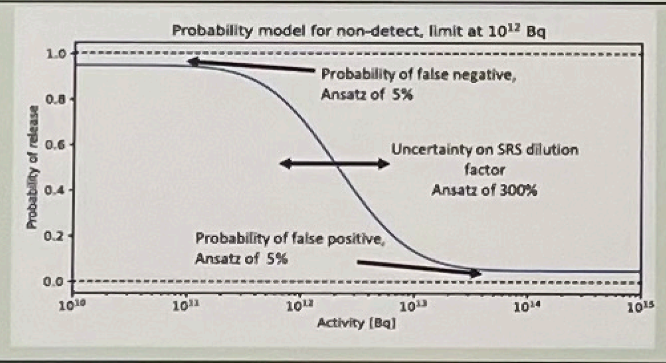
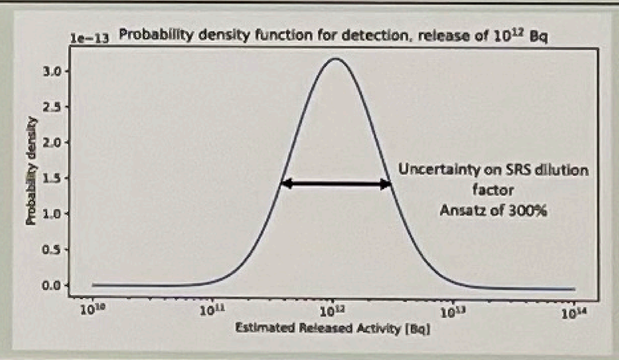
$$P(R(lat,lon,t,I)|C) = \frac{P(C\,|R(lat,lon,t,I))\,P(R(lat,lon,t,I))}{P(C)}$$

$R(lat, lon, t, I)$ = Release of isotope *I* at given location and time. Requires a minimum threshold

C = Configuration of linked RN detections & non-detections

Method 1 : Custom probability model

For each cell (lat, lon, t) and isotope *I*, match to RN measurements through SRS fields. The probability of a release is the combination of probabilities from all detects and non-detects

Non-Detects Nuclide **is not** quantified in RRR	Detects Nuclide **is** quantified in RRR
Places an **upper-limit** on the released activity within its SRS field	A release of the isotope **cannot be ruled out**
Calculate **upper-limit** at cell based on MDC of isotope *I* $A_0(lat,lon,t) = MDC\,\frac{e^{\lambda_I t}}{M(lat,lon,t)}$	Calculate **release** based on measured activity concentration, C_m, of isotope *I* $A_0(lat,lon,t) = C_m\,\frac{e^{\lambda_I t}}{M(lat,lon,t)}$
Represent release probability with an activation function, an adjusted CDF of log-normal distribution	Probability distribution of released activity modeled using log-normal distribution as an ansatz approximation
	

Method 2 : Statistical Method

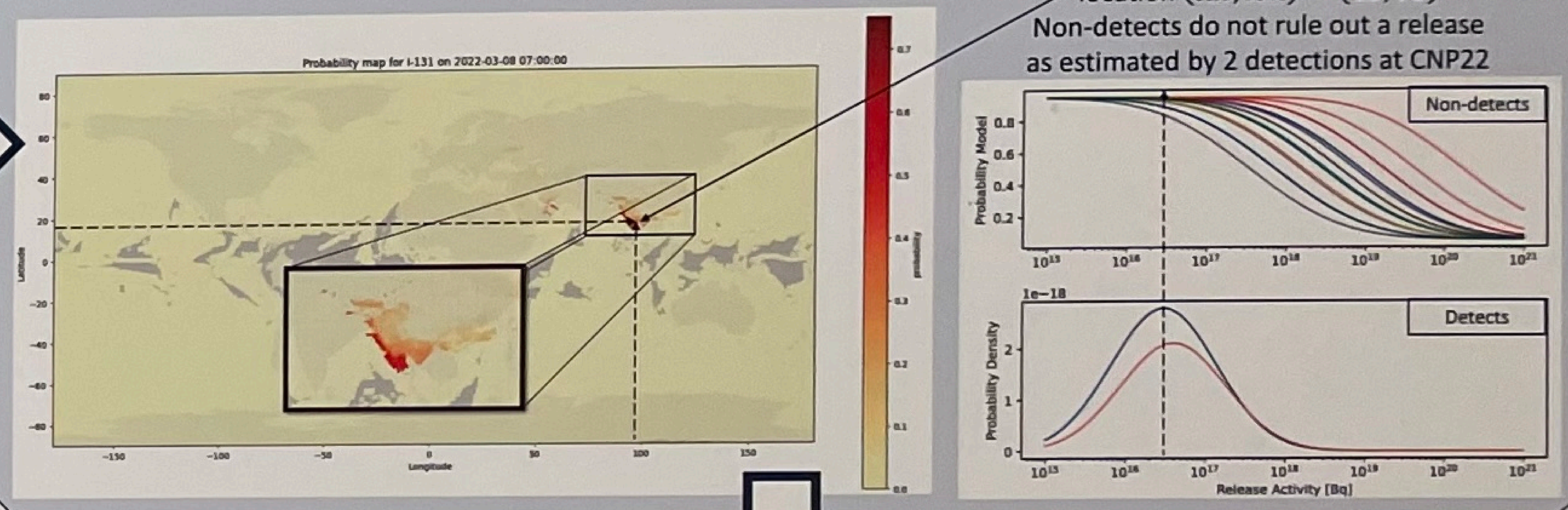
- This method is motivated from previous efforts based on station affinity [1,2]
- Each cell (lat, lon, t) is treated as a possible release of isotope *I*, modeled as a puff (within the 1-hour time block) of a fixed activity
- Match to RN stations through SRS fields and predict the measured activity concentration. Count all results that exceed the MDC as a detect

$$A_0 M(lat,lon,t)e^{-\lambda_I t} > MDC$$

- Accumulate counts for each station configuration, and time from detection (in hours)
- Aggregate counts over ~1 year of data to accumulate statistics
- For now, this works when the detection occurs for one or two RN measurements, otherwise the model becomes too complex

Probability model Results

The probability of each cell is determined by convoluting the detect PDF with the non-detect probabilities. In the case of multiple detects, the result using the detect PDF having the highest probability is selected The model provides hour-by-hour probability maps for each isotope



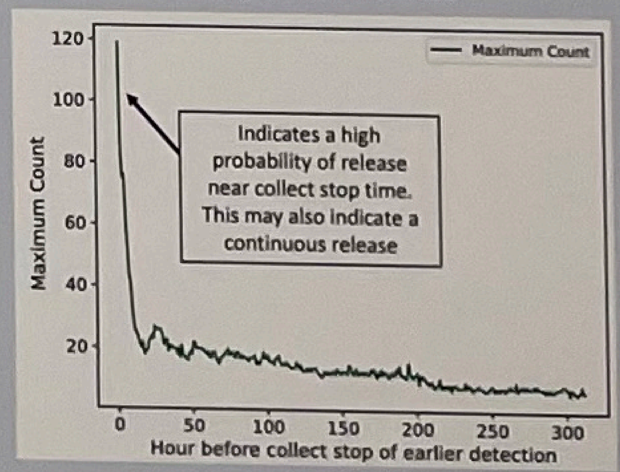
Comparison and Conclusions

- Both models can provide release localization, with varying results. The highest probability areas of the two methods do not agree, but there is some overlap
- The highest probability area of Method 1 lies in the ocean. This may indicate a need for tuning or to reconsider the prior distribution
- Method 2 tends to provide smoother probability regions while those for Method 1 are more irregular
- Method 2 currently has no limit to model complexity with larger numbers of detecting stations, this issue needs to be addressed, possibly using machine learning
- Generally, the models require further testing and tuning
- These studies will be used as a starting point to investigate other methods such as in [3,4]

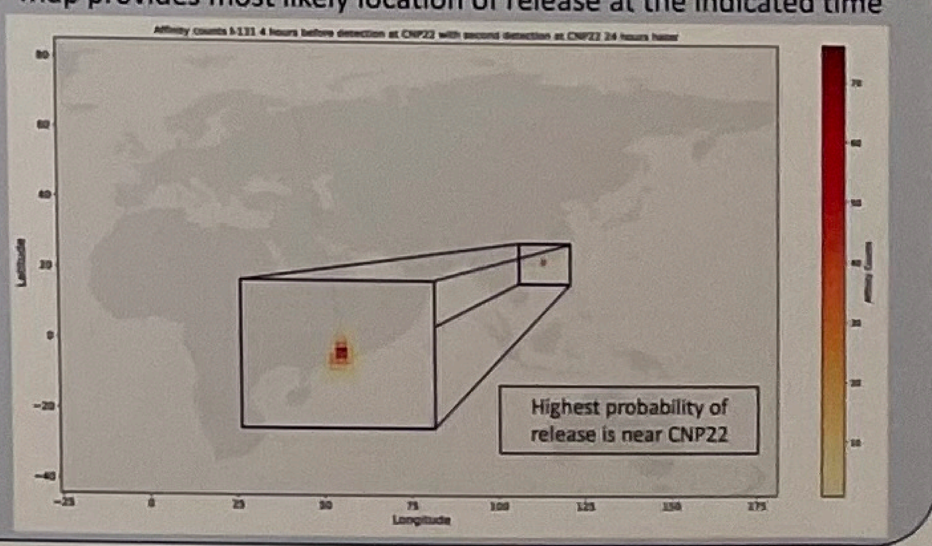
Statistical Method Results

Consider a case of two detections of I-131 at CNP22 1 day apart. This is comparable to the result from Method 1. The model provides a probability over time and space.

Maximum count vs time gives an estimate of the most likely time of release



Map provides most likely location of release at the indicated time



PUTTING AN END TO NUCLEAR EXPLOSIONS