A baseline for source reconstruction using the inverse atmospheric modelling tool *FREAR*

Pieter De Meutter

NERIS workshop, Dublin, 9 October 2023

Introduction to inverse modelling

Direct modelling

- Set of detections associated to a known release of radionuclides into the atmosphere
- Calculate the transport and dispersion of radionuclides in the atmosphere in space and time
- Assimilate observations to improve the results

Inverse modelling

- Set of detections associated to an unknown release of radionuclides into the atmosphere
- Use observations, atmospheric transport and dispersion modelling in a statistically coherent manner to determine the source parameters of an atmospheric release of radionuclides

source-receptor sensitivities obtained by ATM

source parameters (release location, timing, amount, …)

vector of observations

$$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{x} + \boldsymbol{\epsilon}$$

vector of combined model and observational errors

Forensic Radionuclide Event Analysis and Reconstruction - the “FREAR” code

- Initially developed with the purpose of CTBT verification
- Required input:

(1) source-receptor sensitivities ($\boldsymbol{M}$)

(2) observed airborne activity concentrations ($\boldsymbol{y}$)

(Can deal with both detections and instrumental non-detections; it takes into account the possibility for misses and false alarms)

- FREAR can solve the inverse modelling problem using two independent methods: a cost function optimization method and a Bayesian MCMC method
- Users can select the most appropriate source parameterization for a given problem (such as multiple release segments or a release from a fixed location), and can add their custom source parameterization if needed
- The Bayesian inference approach provides an estimate on the uncertainties in a natural way. Furthermore, an ensemble of atmospheric transport modelling can be used to better estimate model uncertainty
- Code written in R, available on GitLab under GPLv3

FREAR challenges and outlook

Some challenges:

- How to include different sources of observations (such as gamma dose rate measurements and deposition measurements)?
- FREAR performed well when applied to previous case studies; will it perform well when applied to the next case?
- …

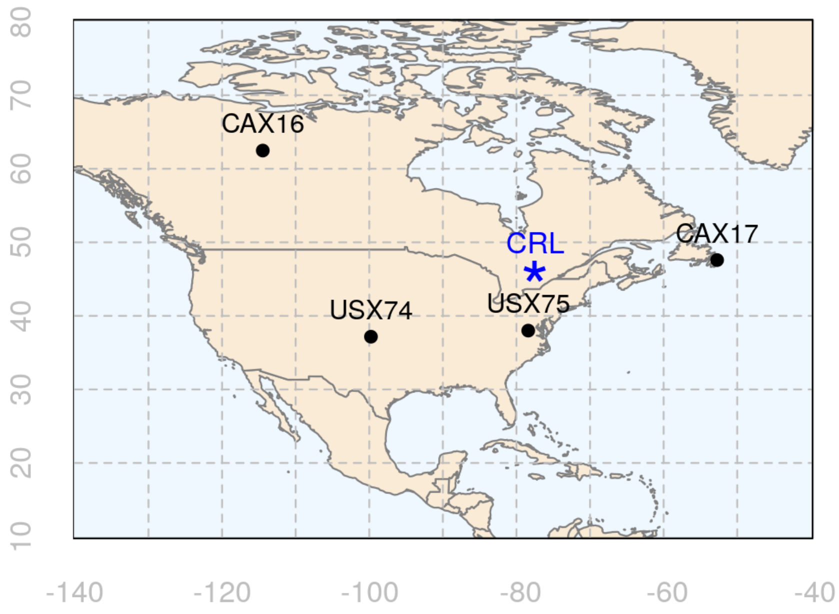
Outlook:

- Inclusion of gamma dose rate measurements → Pianoforte proposal IRENE
- Inclusion of deposition measurements → PhD student Stijn Van Leuven
- Apply FREAR over a set of test cases → this talk

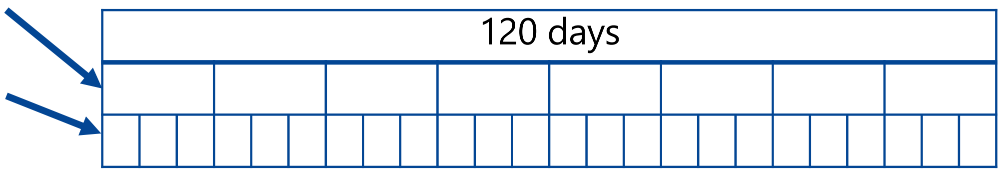
Purpose: to establish a baseline for source reconstruction to facilitate testing of data, methods and settings

Constructing a set of cases

- ^{133}Xe observations at four monitoring stations for the period 1 September 2014 – 30 December 2014 (120 d)
- Detections are linked with emissions from a (former) medical isotope production facility Chalk River Laboratories (CRL)
- Can we reconstruct the (known) source location of CRL?



- Two sets of case studies:
  i. 8 cases with 15 d of observations
  ii. 24 cases with 5 d of observations

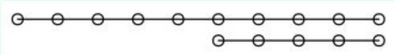


ATDM and FREAR setup

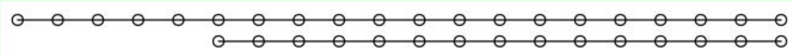
ECMWF input: 3-hourly at 0.5° x 0.5° for full NH

Flexpart output: daily at 0.5° x 0.5°, 0 – 100 m, full NH

5-day cases: 10 daily release segments



15-day cases: 20 daily release segments

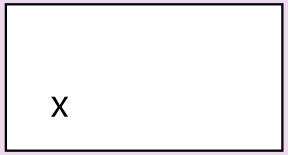


Use archived source-receptor sensitivities calculated by CTBTO (Flexpart + ECMWF)

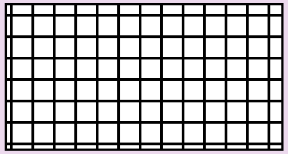
3-hourly output at 1° x 1°, 0 – 500 m

FREAR

Bayesian inference
- infer source term
- infer source location
- **y** = **M x**



Cost function
- optimize source term
- for each grid box
- **y** = **M x**

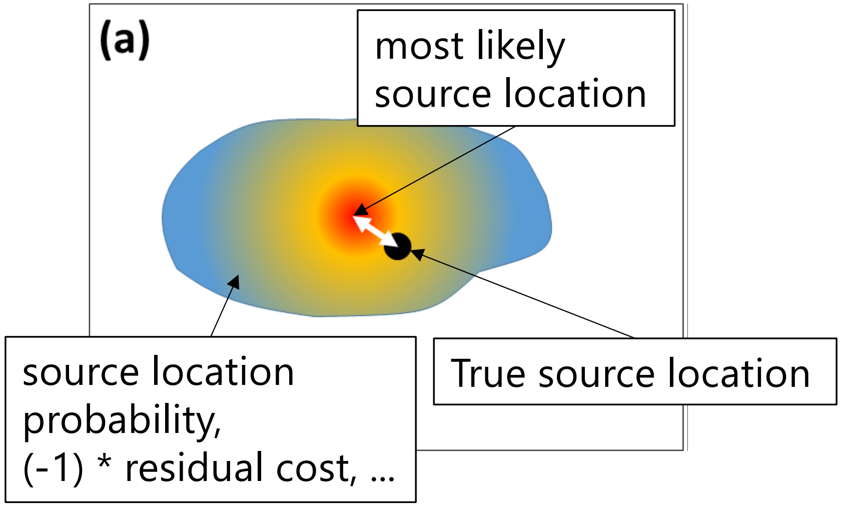


maximum-in-time PSR using Pearson / Spearman correlation
- for each grid box
- correlation between **y** and **M**

Accumulated-in-time FOR
- for each grid box
- area where **M[y>0]** > 0 for any time

Three verification metrics for source localisation

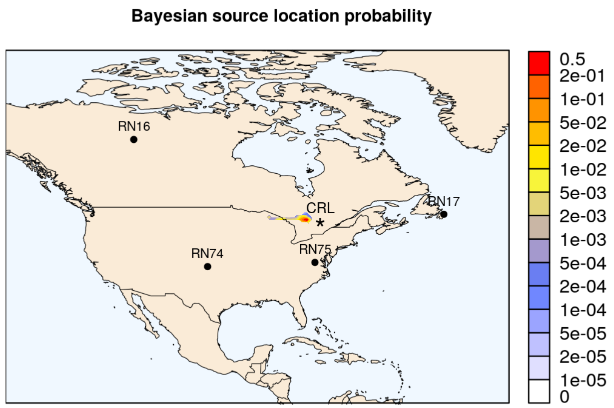
Distance [**0**, ∞] | Quantile at true location [0, **1**] | Fraction of domain excluded [0, **1**]



Thresholds:		
	Bayesian inference	source location probability = 0
	Cost function optimisation	residual cost > 2
	maximum-in-time PSR	maximum-in-time correlation < 0.1
	accumulated-in-time FOR	< 50% of SRS* overlapping

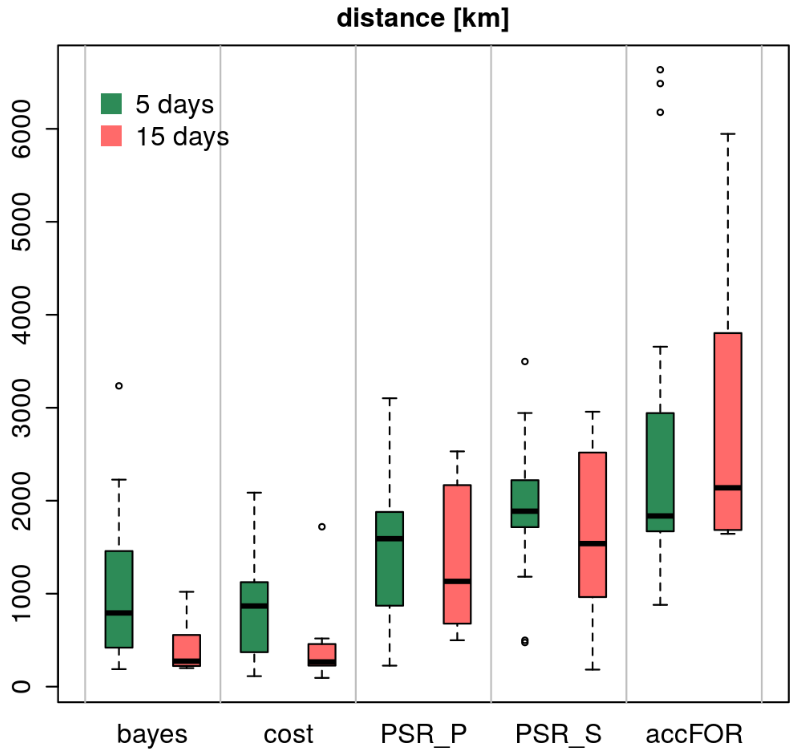
*: associated with detections

Results: example of different inverse modelling methods and verification metrics



Method	Distance [km]	Quantile	Excluded
bayes	199	0.000	0.998

Results: comparing different methods for 5 days and 15 days of observations (1/2)



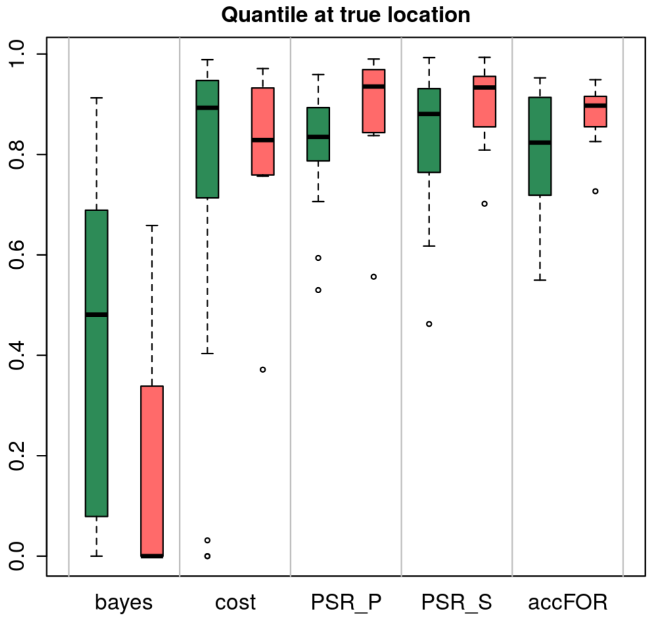
Comparing methods:

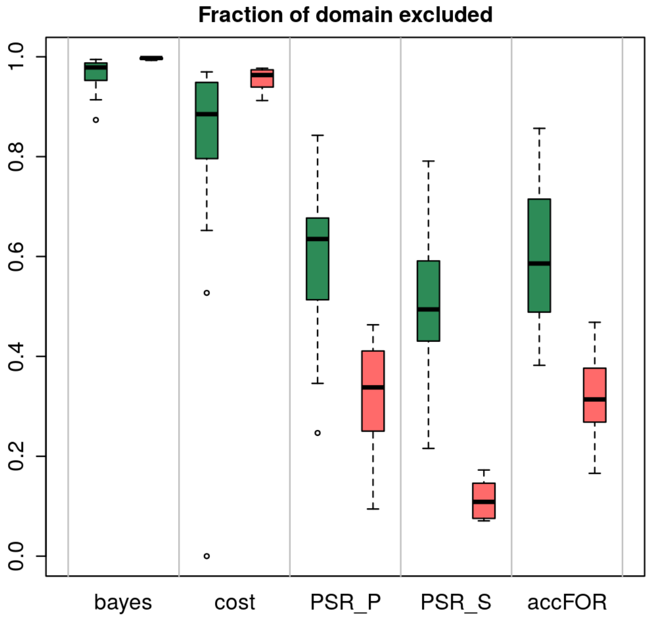
- *bayes* and *cost* are able to locate the source much better than other methods

Comparing 5 d vs 15 d observations:

- *bayes* and *cost* show a significant improvement (from 800 km to 270 km)
- only a modest or no improvement for other methods

Results: comparing different methods for 5 days and 15 days of observations (2/2)





Comparing methods:

- *bayes*: poor quantile score, other methods comparable
- *bayes* and *cost*: exclude large fraction of search domain

Comparing 5 d vs 15 d observations:

- *bayes* and *cost* show a deterioration in quantile score and an improvement in the fraction of domain excluded
- other methods show an improvement in the quantile score and a deterioration in the fraction of domain excluded

summed median scores					
	bayes	cost	corr (P)	corr (S)	FOR
5 days	1.43	1.69	1.41	1.34	1.37
15 days	1.00	1.79	1.20	1.04	1.21

When more information is available, the added value of more elaborate methods is higher compared to simpler methods

Results: Testing a new cost function (1/2)

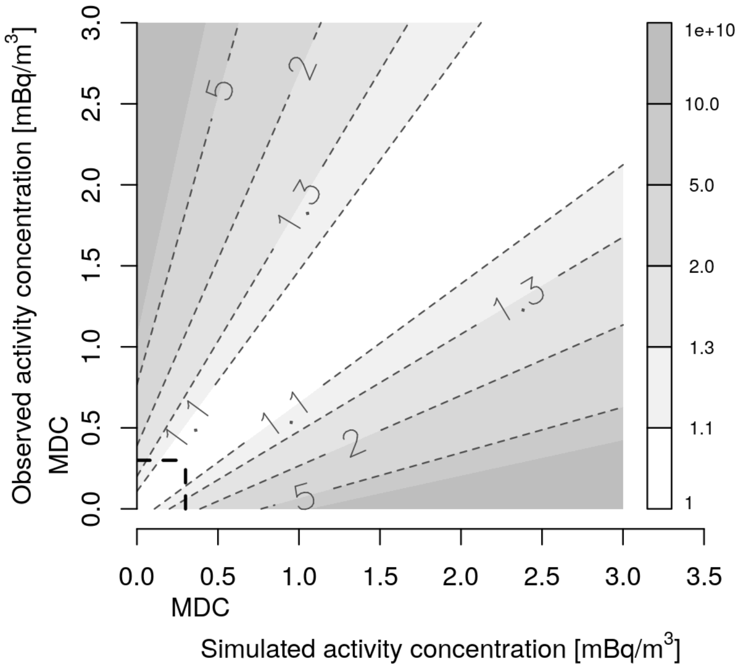
$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\left[\log(AC_{obs,i}+\mathrm{MDC}_i)-\log(AC_{sim,i}+\mathrm{MDC}_i)\right]^2\right)$$

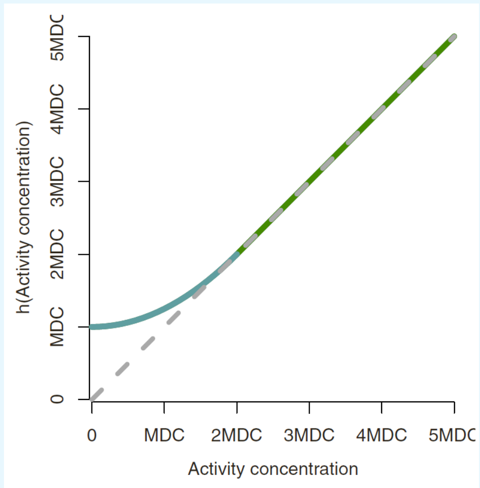
$$\exp\left(\frac{1}{n}\sum_{i=1}^{n}\left[\log(h(AC_{obs,i}))-\log(h(AC_{sim,i}))\right]^2\right)$$

$$h(x)=\begin{cases} f(x) & \text{if } x \leq 2\,\mathrm{MDC} \\ x & \text{else}\end{cases}$$

$$f(x)=\frac{x^2}{4\,\mathrm{MDC}}+\mathrm{MDC}.$$

Original cost

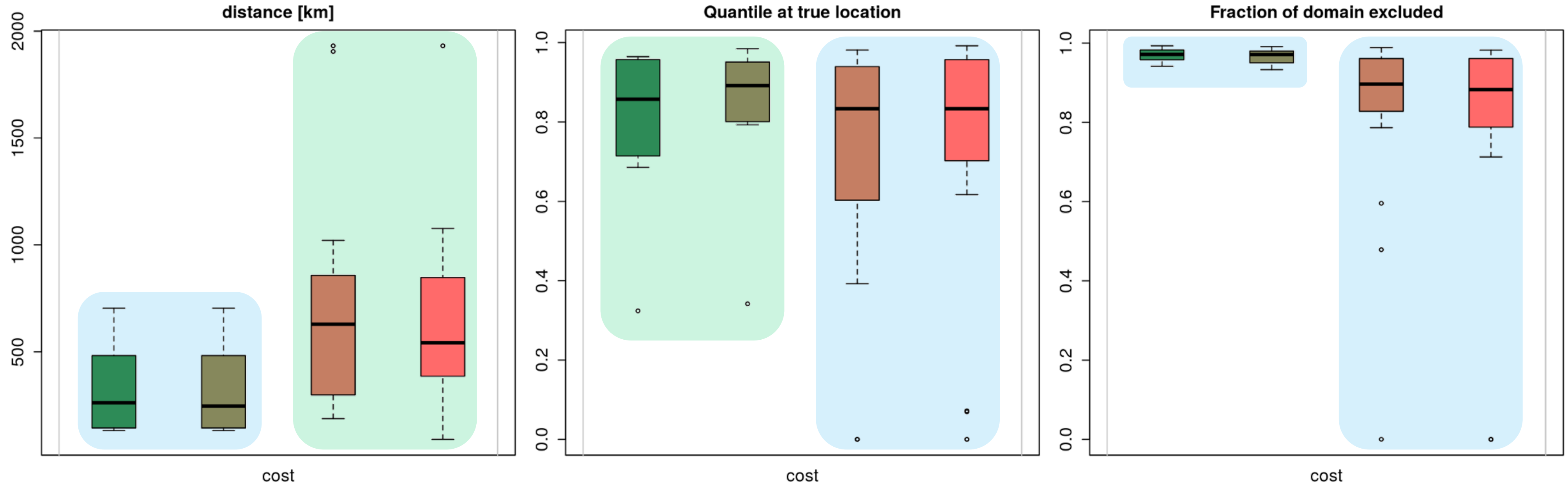




Results: Testing a new cost function (2/2)

- original cost function – 15 days
- new cost function – 15 days
- original cost function – 5 days
- new cost function – 5 days

We expect this new cost function to be slightly more sound. How does it perform on our set of test cases?
Effect of new cost function on scores is neutral / improvement



Summary and conclusions

Two sets of case studies have been defined for inverse modelling using ^{133}Xe observations associated to a (former) medical isotope production facility Chalk River Laboratories:

- 8 cases using 15 days of observations and 24 cases using 5 days of observations

These sets allow for:

- a comparison of data (observation selection, NWP input, ATM input, …)
- a comparison of methods (inverse modelling methods, source parameterizations, …)
- the testing of new or modified inverse modelling algorithms

Findings:

- Bayesian inference and cost function optimization are able to exclude a large fraction of the location search domain, contrary to simpler methods
- When more information is available, the added value of more elaborate methods is higher compared to simpler methods
- Bayesian inference underestimates uncertainties since the true source location sometimes falls outside the possible source region (note: other methods do not optimize for source location)

Thank you for your attention

- Questions?

- Acknowledgements:
  - Ian Hoffman (Health Canada)
  - Andy Delcloo (Royal Meteorological Institute of Belgium)
  - IMS station operators and CTBTO for making available the ^{133}Xe observations

SCK CEN
Belgian Nuclear Research Centre

Foundation of Public Utility

Registered Office: Avenue Herrmann-Debrouxlaan 40 – BE-1160 BRUSSELS
Operational Office: Boeretang 200 – BE-2400 MOL