

Source term estimation for anthropogenic radionuclides in the CTBT-context

FLEXPART Training 2026

Department Environmental Meteorology/Competence Unit CBRN-Incident Modeling

Christian Maurer

christian.maurer@geosphere.at

www.geosphere.at



1. Introduction

- The Comprehensive Nuclear Test-Ban Treaty (CTBT)
- The International Monitoring System (IMS)
- The Majak Reprocessing Plant / Kytschym Accident

2. The PSR method exemplified for a real case

- Elevated Ru-106 levels in autumn 2017
- Possible Source Region (PSR) analysis
- Hands-on – your turn
 - I. Setting-up the FLEXPART run
 - II. The COMMAND file in detail
 - III. Plotting binary output
 - IV. Python 3.8 packages for binary plotting
 - V. The concept and format of SRS files
 - VI. Create SRS files
 - VII. Poor man's inversion
 - VIII. The concept and format of PSR fields
 - IX. Create PSR files
 - X. Plot the PSR fields
 - XI. Python 3.8 packages for PSR plotting



3. The FREAR Tool

- FREAR overview
- Relating synthetic NDC Preparedness Exercise (NPE) RN measurements to a waveform event
- Cost-function approach
- Bayesian inference: Bayes theorem
- Bayesian inversion
- Refining the analysis: Fixing the source location



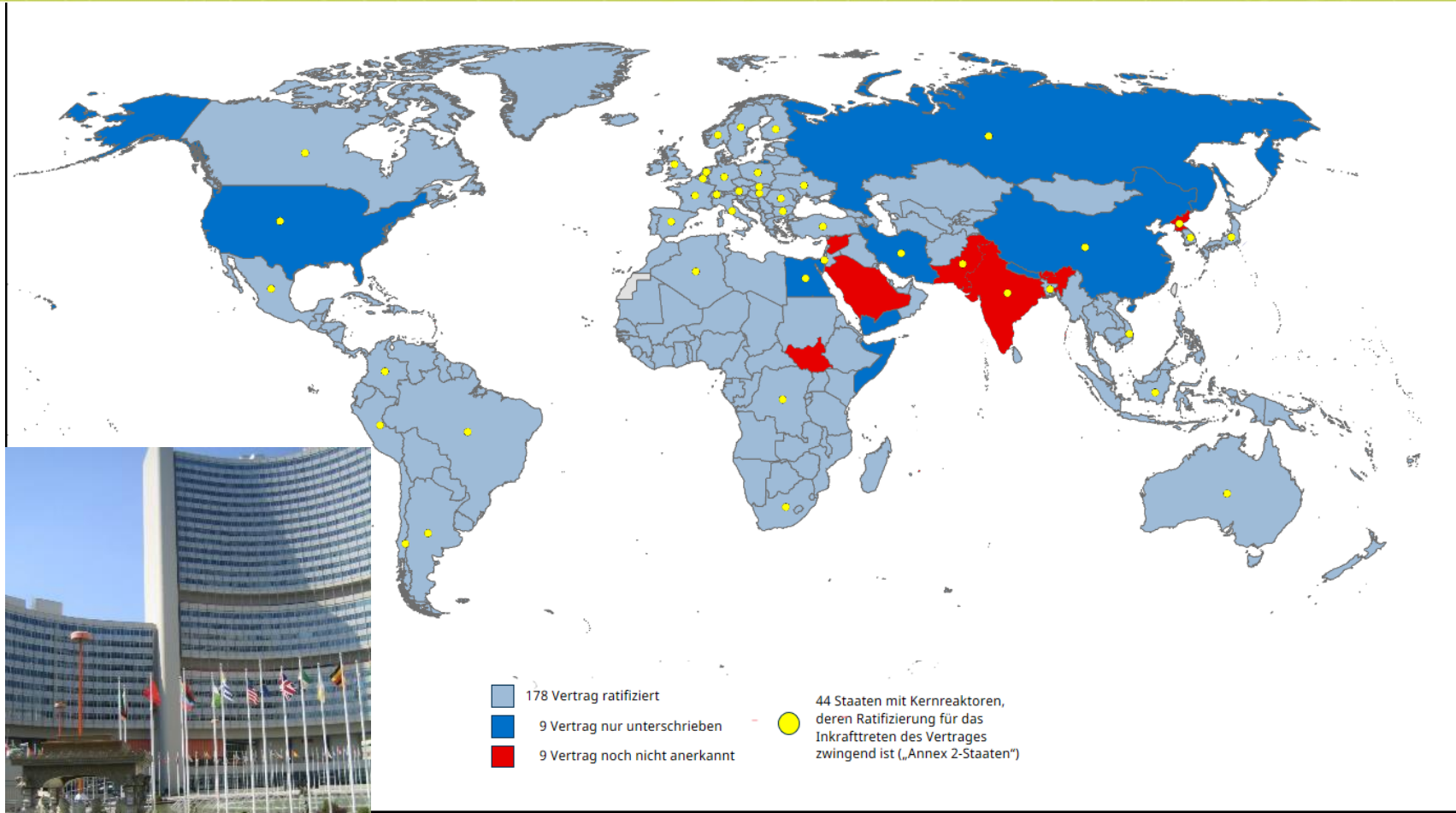
ARTICLE I - BASIC OBLIGATIONS

“Each State Party undertakes **not to carry out any nuclear weapon test explosion or any other nuclear explosion**, and to prohibit and prevent any such nuclear explosion at any place under its jurisdiction or control.”

The *Provisional Technical Secretariat* (PTS) was established in Vienna in 1997

Tasks of the PTS:

- Installation and maintenance of the *International Monitoring System* (IMS)
- Operating the *International Data Center* (IDC)
- Facilitating *On-Site Inspection* (OSI) – Technologies, Trainings



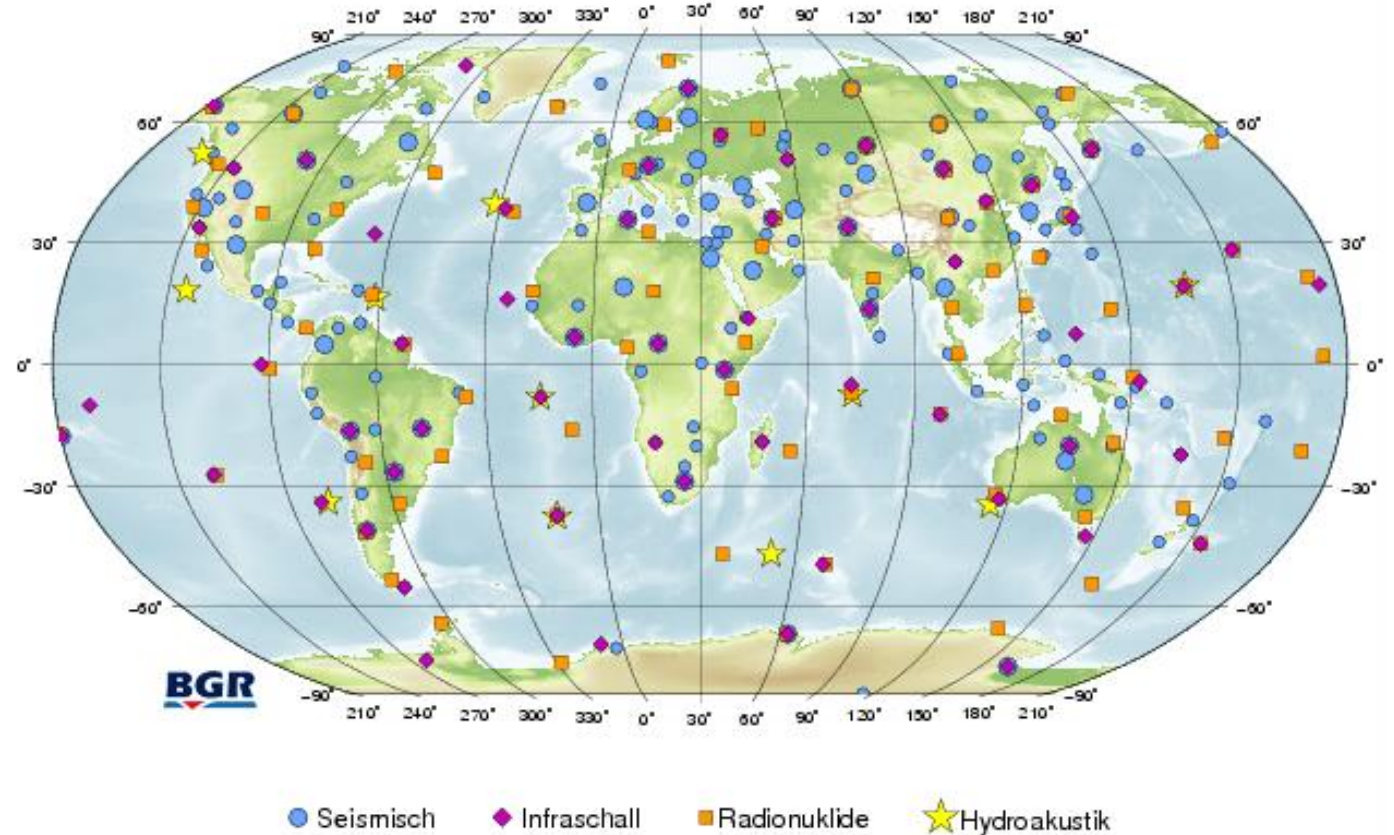
STATUS OF SIGNATURE AND RATIFICATION

196	187	178	9	18
Total States	Total Signed	Total Ratified	Not Signed	Not Ratified
Latest State Signatory: 🇸🇴 Somalia 📅 8-Sep-2023		Latest Ratifying State: 🇵🇬 Papua New Guinea 📅 13-Mar-2024		

Entry into force: As soon as 44 countries (Annex II states) which operate nuclear reactors have ratified the Treaty (not just signed) ...

The International Monitoring System (IMS) was designed to detect underground, atmospheric and underwater nuclear explosions in the order of 1 kt TNT equivalent (corresponding to roughly a magnitude 4 earthquake).

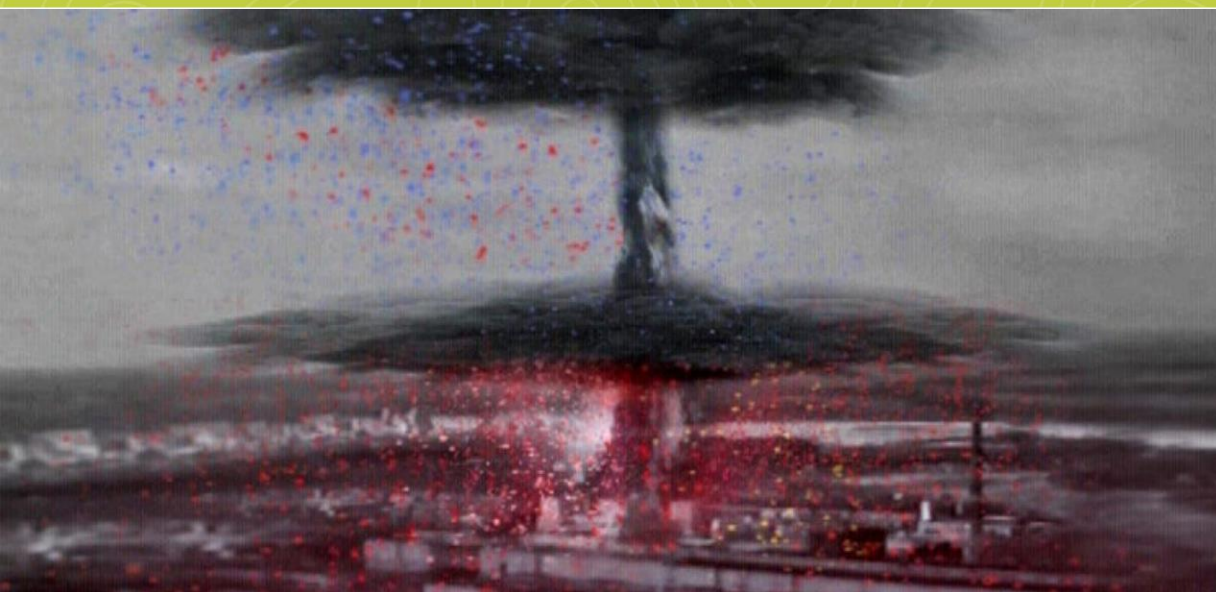
- **Four complementary verification technologies**
- > 80 % of the network currently installed
- **80 radionuclide stations** foreseen, whereof 40 will be equipped with noble gas measuring devices at Entry into Force
- **High value of the IMS in case of civil nuclear incidents and accidents!** -> see the Ruthenium case



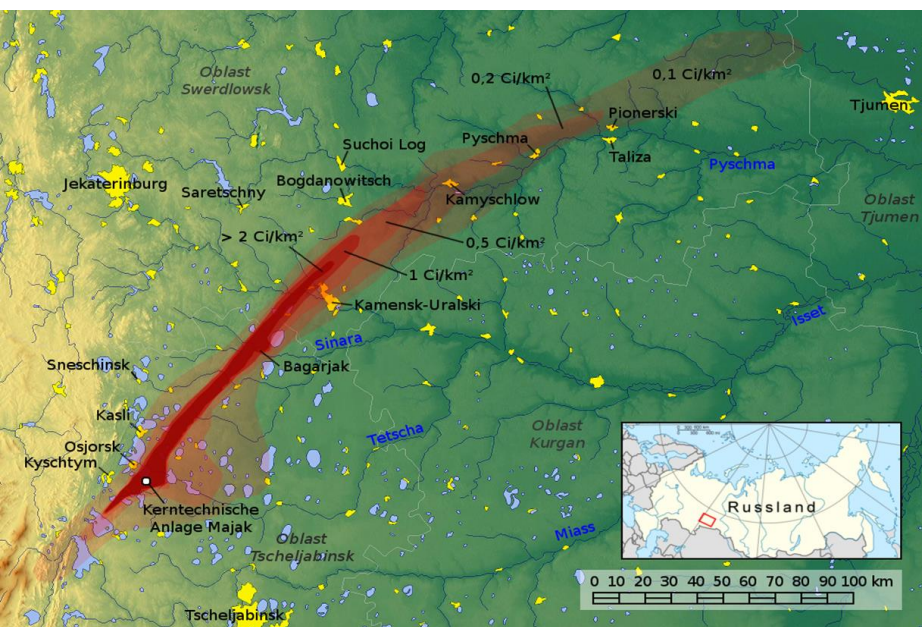
Hydroacoustic stations: 11 (11 certified) Seismic stations: 170 (> 150 certified)

Radionuclide stations: 80 (> 70 certified) Infrasound stations: 60 (53 certified)

The Majak Reprocessing Plant / Kytschym Accident



Apocalyptic scenes on September 29, 1957: “The radioactive cloud rose to a height of 1 km above the ground and glowed in vivid red colors. The birch leaves turned yellow instantly and fell to the ground. The forest soon looked bare...” (Eyewitness account)

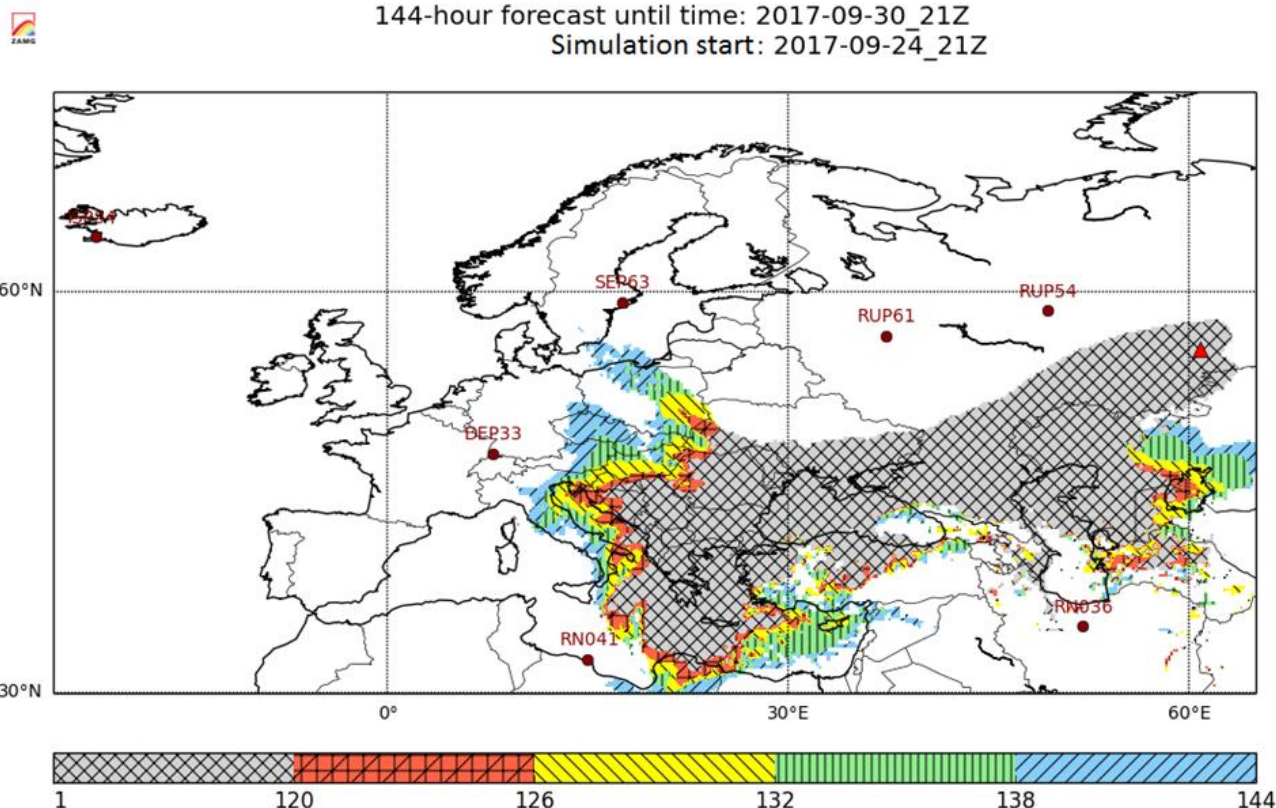


„Ural Trace“
Predominantly
Cesium-137
(HL: 30 years),
Strontium-90
(HL: 29 years)
and
Plutonium-239 (HL: 24.110 Jahre!!)

- Objective: **Production of weapons-grade plutonium for the construction of UDSSR's first atomic bomb**
- The cooling system for one of the tanks failed in 1956, and the nuclear waste solution dried out...
- A spark triggered a massive chemical (non-nuclear) explosion in 1957.
- 10,700 people evacuated; **received radiation doses up to 600 Millisieverts (mSv)**
- No impact on Western Europe; 90% of the released radioactivity remained confined to the plant site!



Elevated Ru-106 levels in autumn 2017

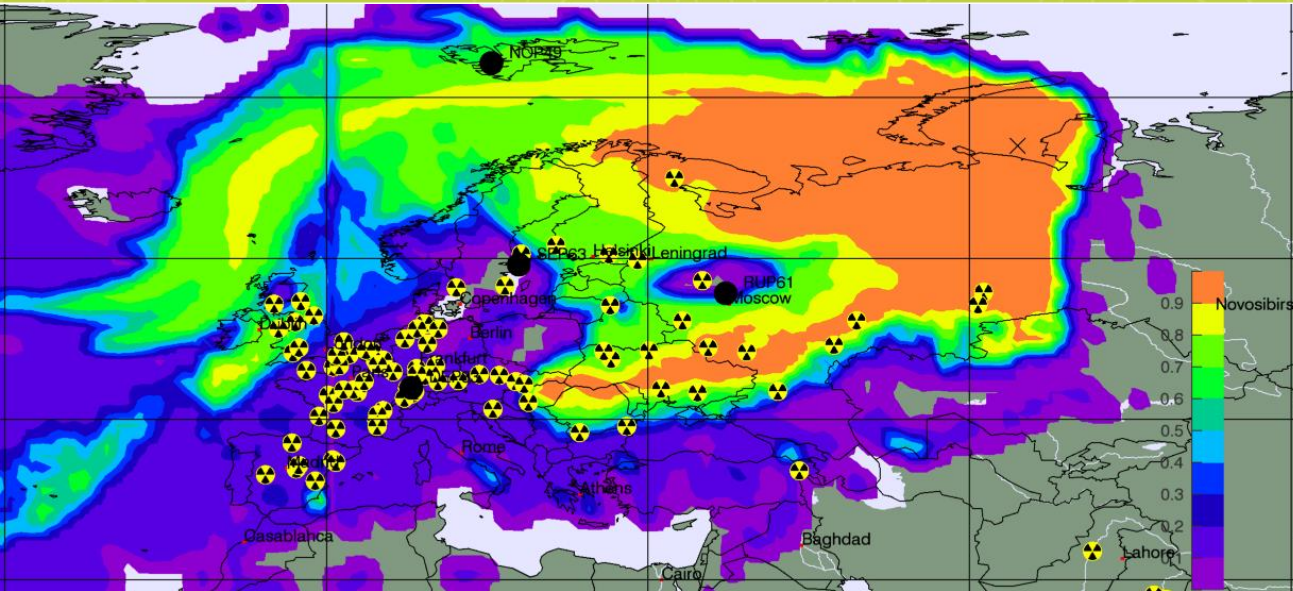


- **Unusually elevated Ru-106 (and Ru-103) measurements across Europe** (national and CTBTO station in Stockholm) in late September/early October 2017. Exceedance of 175 mBq/m^3 (the unit $\mu\text{Bq/m}^3$ is normally used) in Southeastern Europe.
- **No other nuclides** (e.g., Xenon, Barium, Lanthanum, Cesium, or Iodine) **were found** in the spectra: Atmospheric or underground **nuclear explosion as well as an accident in a nuclear power plant could thus be excluded** as the source of the Ruthenium.

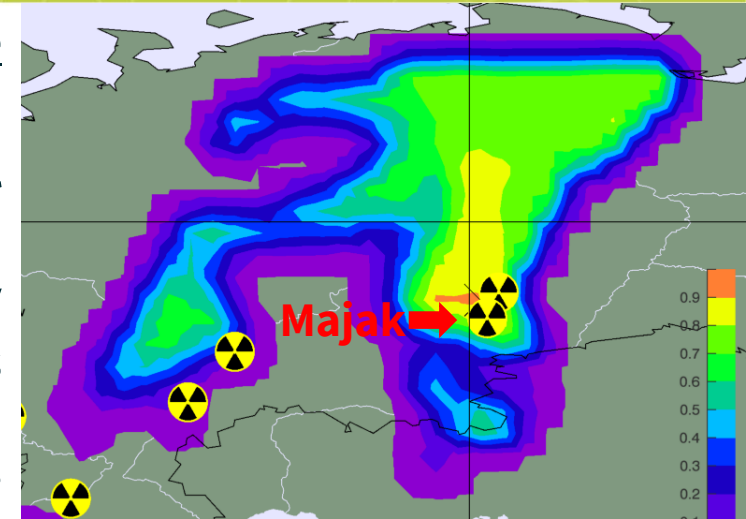
- According to the ratio of Ru-106 ($T_{1/2} = 374$ days) to Ru-103 ($T_{1/2} = 39$ days), the **radionuclides were generated 400 days before their release.**
- Conclusion: **Incident or accident at a nuclear fuel reprocessing plant or an isotope production facility**



Possible Source Region (PSR) analysis I



2) PSR field for time slot Sep., 25th, 15-18 UTC: **Geographical coincidence with Majak! Most likely release quantity is approximately 1 Petabecquerel (10^{15} Bq).**



1) First overview: Time-integrated PSR (20.- 30.9.2017). Field of highest integrated correlations is largely confined to Russia. Based on FLEXPART-9 and ECMWF 1.0° data. SRS fields from SEP63, DEP33, NOP49 and RUP61 as well as on the CTBTO/IDC software WEBGRAPE (*Web-connected graphics Engine*).

Conclusion: Majak is a plausible source of Ru!

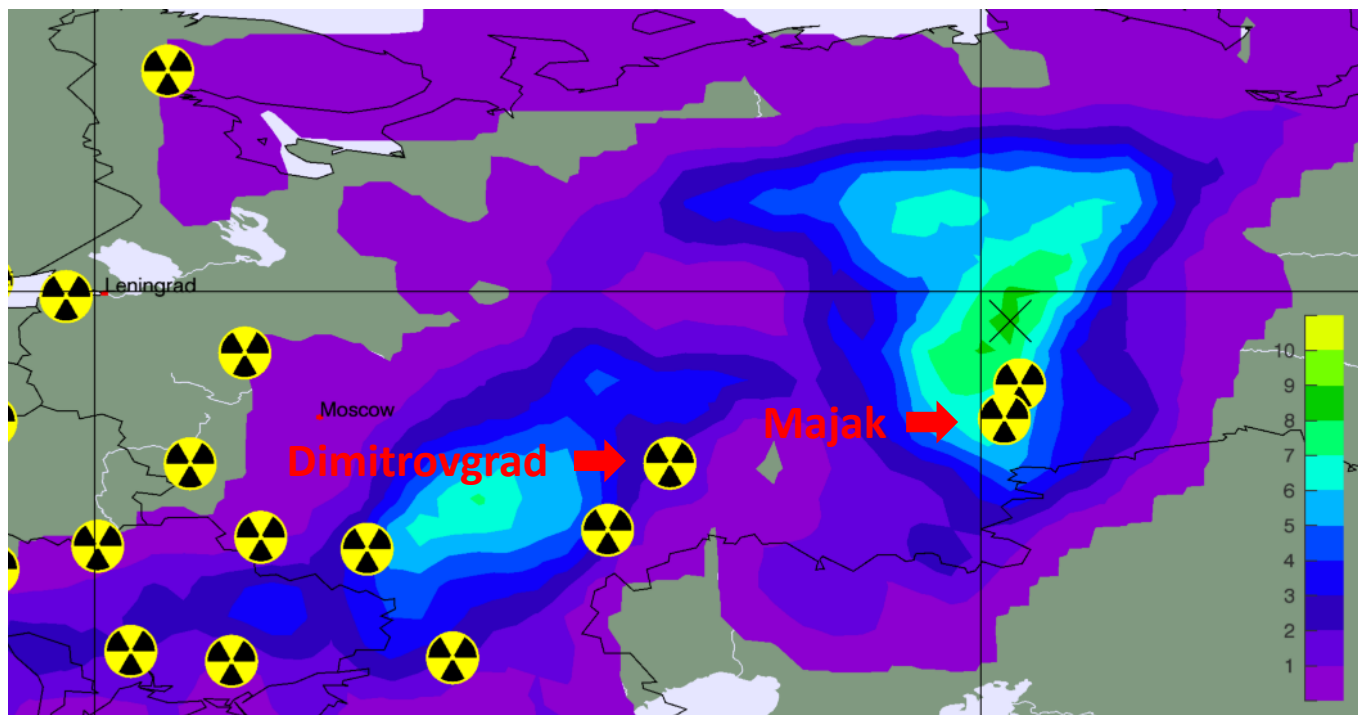
The concept behind PSR:

- For every grid cell and time step: Correlate FLEXPART source receptor sensitivities (SRS) and actual measurements. Determine *how well measurements fit a (hypothetical) release*.
- Assumption of ~ puff release (determined by the output time step).
- Also „non-detections“ can be helpful for constraining the source region.

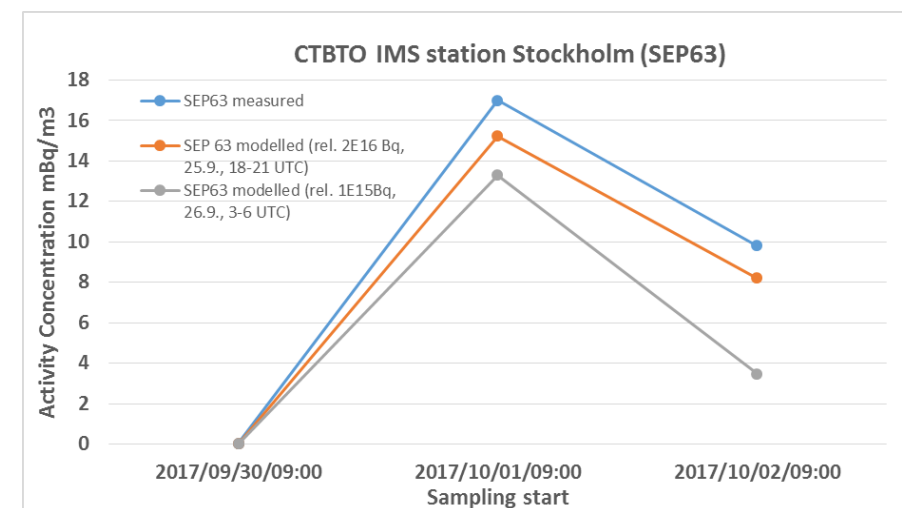


3) Multi-Model (Ensemble) overlap for time slot: Sept 25th 21:00 – Sept 26th 00:00 UTC

Overlap of PSR fields (in total 11) showing at least a correlation of 0.5



4) Testing of detailed source hypotheses based on fwd ATM and comparison to measurements



WMO Regional Specialized Meteorological Centers (RSMCs) supply CTBTO with results of dispersion calculations in case of exceptional radionuclide measurements (Level-5 Events). GeoSphere Austria is one of these designated centers! Aim: Ensemble Modeling.



■ COMMAND FILE:

- Backward run: 25 September 2017 00 UTC to 3 October 2017 09 UTC
- Output every 3 hours 3-hourly averages
- Residence times output
- Binary output!
- No AGECLASS use and no fluxes output

■ OUTGRID FILE:

- Resolution 1.0 degree
- LLC: 0.0° N, 0.0° E
- 90x90 grid cells
- 1 output layer, 150 m a.g.l.

**TO DO: Prepare
option-files,
run FLEXPART
& plot results**

■ RELEASES FILE:

- Point source at CTBTO IMS station Stockholm (SEP63): 59.41° N, 17.95° E, 0-10 m a.g.l.
- 1st release:
 1. Release start: 30 September 2017, 09 UTC
 2. Release end: 1 October 2017, 09 UTC
 3. Total mass: 1 Bq of Ru-106 (species nr. 1)
 4. Particles released: 100,000
 5. COMMENT = "RELEASE_SEP63_1"
- 2nd and 3rd release:
 1. Release start: 1 and 2 October 2017, 09 UTC
 2. Release end 2 and 3 October 2017, 09 UTC
 3. Total mass, particles released and appropriate comments as for 1st release

Q1: Which of the first two samples (release ends on 1 and 2 October) could be related to the nuclear facilities of Dímítrovgrad or Majak in Russia at around Sep., 25th, 21 UTC, to 26th, 03 UTC? Check your plots.

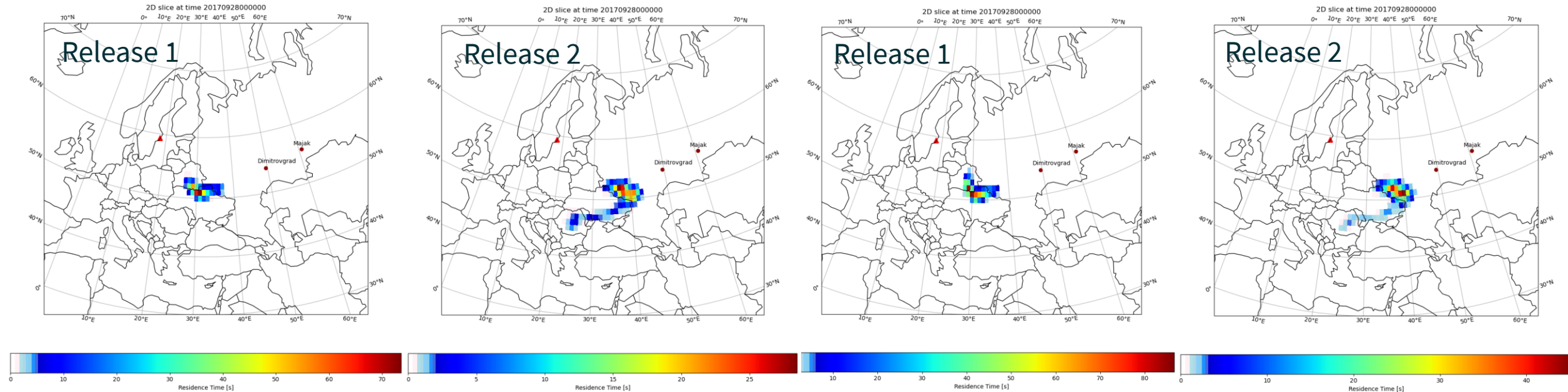


&COMMAND

LDIRECT= -1, ! Simulation direction in time ; 1 (forward) or -1 (backward)
IBDATE= 20170925, ! Start date of the simulation ; YYYYMMDD: YYYY=year, MM=month, DD=day
IBTIME= 000000, ! Start time of the simulation ; HHMISS: HH=hours, MI=min, SS=sec; UTC
IEDATE= 20171003, ! End date of the simulation ; same format as IBDATE
IETIME= 090000, ! End time of the simulation ; same format as IBTIME
LOUTSTEP= 10800, ! Interval of model output; average concentrations calculated every LOUTSTEP (s)
LOUTAVER= 10800, ! Interval of output averaging (s)
LOUTSAMPLE= 900, ! Interval of output sampling (s), higher stat. accuracy with shorter intervals
LSYNCTIME= 900, ! All processes are synchronized to this time interval (s)
IOUT= 1, ! Output type: [1]mass 2]pptv 3]1&2 4]plume 5]1&4, +8 for NetCDF output
LAGE SPECTRA= 0, ! Switch for calculation of age spectra (needs AGECLASSES);[0]off 1]on
OUTPUTFOREACHRELEASE= 1, ! Separate output fields for each location in the RELEASE file; [0]no 1]yes
IFLUX= 0, ! Output of mass fluxes through output grid box boundaries
IND_SOURCE= 1, ! Unit to be used at the source ; [1]mass 2]mass mixing ratio
IND_RECEPTOR= 1, ! Unit to be used at the receptor; [1]mass 2]mass mixing ratio 3]wet depo. 4]dry depo
LNETCDFOUT = 0, ! Switch to produce NetCDF output
NXSHIFT = 0, ! Shift of the global meteorological data by number of grid cells
/



Hands-on – your turn: Plotting binary output



ECMWF-ops

ECMWF-ERA5

- `python plot_FLEX_binary_V11.py ./output/ False 1 0 0 [1] alldates lcc False 0,65,30,75 mesh True` ([1] indicates selection of the 2nd release)
- Related python code is available in `./plotting_FLEXout/`
- **Create conda environment *environment.yml* or install python 3.8 packages listed on the next slide manually**



- ast
- sys
- os
- matplotlib
- imageio
- cartopy
- numpy
- struct
- traceback
- datetime
- pdb
- re
- time
- warnings



Hands-on – your turn: The concept and format of SRS files

17.950 59.410 20170930 09 20171001 09 0.10E+01 153 3 3 1.000 1.000 "Ru-106" -179.0000 0.0000 360 90

Source location longitude (degrees)

Source location latitude (degrees)

Collection start time

Collection stop time

Released mass (Bq)

Output simulation time length (= number of intervals)

Output time interval (hours)

Output time interval (hours, place holder)

Horizontal resolution in longitude direction (degrees)

Horizontal resolution in latitude direction (degrees)

Species Name

Lower left corner longitude of OUTGRID (degrees)

Lower left corner latitude of OUTGRID (degrees)

Number of grid cells in longitude direction

Number of grid cells in latitude direction

58.000 17.000 1 0.2222537E-05

59.000 17.000 1 0.1206449E+03

58.000 18.000 1 0.1381382E+02

.

.

Lower left corner latitude of respective grid box

Lower left corner longitude of respective grid box

Time step (positive if bwd)

Non-zero concentration

- **SRS file:** An ASCII file consisting of a header line with metadata information and four columns per line afterward with the location and time of non-zero FLEXPART concentrations for just one (!) vertical level
- Very intuitive way of depicting **FLEXPART output for one level and all time steps**, standard (post-processed) FLEXPART output format of CTBTO



- `make -f makefile_flex2srs_gcc_14.1` to compile `flexout2srsmnetcdf.f`. Required package: `netcdf`. Use **gcc 14.1**. Set **INCDIR** and **LIBDIR** accordingly!
 - Related subroutines are to be found in `./subroutines/`
 - `mkdir outputsrs, outputsrs/20171001/, outputsrs/20171002/` and `outputsrs/20171003/`
 - Create **CONTROL** file:
./output/
./summed_binary/ ! path for summed binary - NOT needed for current application
=====
- TO DO: Create ASCII-srs files (one level, all time steps) from binary FLEXPART output (one time step, all levels) for all 3 releases.**
- 1 ! Number of FLEXPART level to start at
 - 1 ! Number of FLEXPART level to stop at
 - 1.0 ! Multiplication factor to get s
 - fp ! Model name stamp in file name
 - f11 ! Version stamp in file name
 - 1 ! species to start with
 - 1 ! species to end with
 - 1 [2,3] ! selected release - VERY IMPORTANT for current application
 - 1900010100 ! selected start date-time
 - 2100010100 ! selected stop date-time
 - .false. ! indication, whether summed binary should be produced
 - 001 ! file suffix for summed binary - NOT needed for current application
 - part-sum ! summed species name for summed binary - NOT needed for current application
 - .false. ! indication, whether simple file name should be used
 - .false. ! indication, whether decay should be applied a posteriori
- `./flexout2srsmnetcdf.out` (Disregard output `CONTROL.full_levels` and `field_extension`.)
 - Gzip and move output `conc_RELEASE_SEP63_1_Ru-106.fp.2017100109.f10.1.srm`, `conc_RELEASE_SEP63_2_Ru-106.fp.2017100209.f10.1.srm` and `conc_RELEASE_SEP63_3_Ru-106.fp.2017100309.f10.1.srm` to `outputsrs/20171001/`, `outputsrs/20171002/` and `outputsrs/20171003/`. You can remove `*_integrated.srm` and `*depo*.srm` files. Not needed here.



$$m[\text{Bq/m}^3] = M[\text{s}] * s[\text{Bq/s} * \text{m}^3] \quad s[\text{Bq}] = m[\text{Bq/m}^3] / M[\text{s}] * 3 * 3.600[\text{s}] * 100,000[\text{m}] * 70,000[\text{m}] * 150[\text{m}]$$

Output time interval in hours Seconds of a hour 1.0° latitude in meters (ca.) 1.0° longitude at 50° latitude in meters (ca.) Vertical output layer extent

m...measurement

M...transfer coefficient matrix (aka source receptor sensitivity, footprint ...)

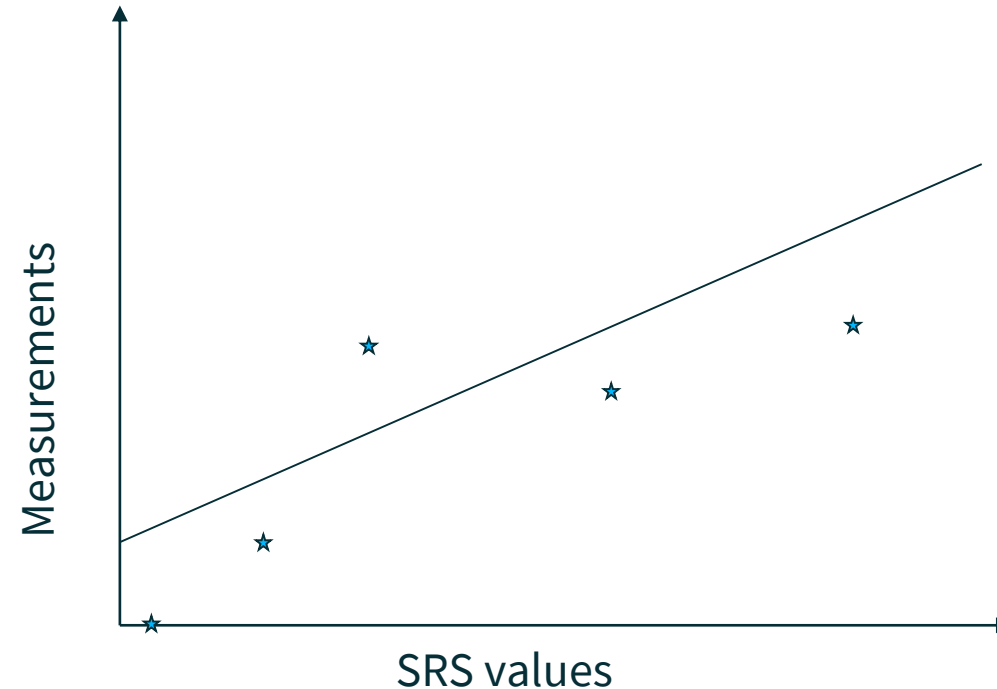
s...source

Q2: Assuming a release between Sep., 26th, 00 UTC, and 03 UTC, and a measurement of 1 mBq/m³ at SEP63 between 1 Oct., 9 UTC, and 2 Oct., 9 UTC, what is your source term estimate for Majak? Check the SRS file. **Hints:** 1) Majak is located at ~ 61°E and ~56° N. 2) Time steps are counted from collection stop backward. One time step equals three hours.



PSR file/result is a multi-sample product! Either based on:

- 1. the same station and different sampling times,**
- 2. different stations and the same sampling time or**
- 3. a combination thereof.**



Header is very similar to that of SRS files. However:

- **no station coordinates are given** in the header line as there has to be more than just one station (minimum two!) involved,
- a block after the header line lists the number and coordinates of involved stations.

```
-99.90 -99.90 20171003 09 20171003 09 0.10E+01 201 3 3 1.00 1.00 -179.0 0.0 360 89 "CORRELATION MAP"  
1  
17.95 59.41
```

Number and coordinates of all involved sampling stations.

```
51.00 14.00 16 0.6077475E-08 0.3692852E-08 0.4999775E-03
```

Lower left corner latitude of respective grid box, lower left corner longitude of respective grid box, time step, explained variance (R^2), slope of regression and offset of regression



Hands-on – your turn: Create PSR fields

- `make -f makefile_correl_gcc_14.1` to compile `locate_multi.f`. Use **gcc 14.1**.
- Related subroutines are to be found in `./subroutines/`
- *Hypothetical measurements* of 0.0, 1.0 and 0.5 mBq/m³ are assumed to be related to the three bwd fields.
- Create **CONTROL** file:

```
# SRS File archive
./outputsrs/
```

```
#
# Species measured
```

```
#
1          ! number of species
Ru-106  0.0    ! half-life [s]
```

```
#
# Scenario/Measurements
#
```

```
1E-18      ! background concentration, used only for FOI approach – NOT needed for current application
```

```
3          ! number of measurements
```

```
conc_RELEASE_SEP63 20171001 09 0.0E-3 ! station | collection stop | concentration above MDC (mBq -> Bq)
```

```
conc_RELEASE_SEP63 20171002 09 1.0E-3
```

```
conc_RELEASE_SEP63 20171003 09 0.5E-3
```

```
FALSE ! indication, whether srs-files are in CTBTO format (corner longitudes and latitudes of global field not given)
```

```
FALSE ! indication, whether time integration should be applied
```

```
10.75 ! days to start time integration before last collection stop
```

```
+++++
```

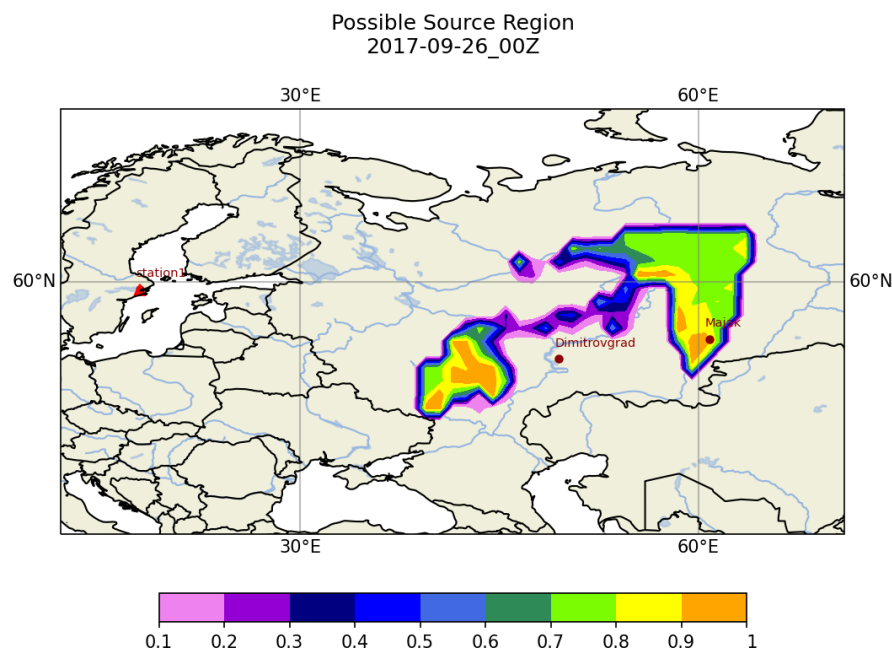
TO DO: Create PSR file based on 3 bwd FLEXPART SRS files and plot it.

- `./locate_multi`: Yields output `correlation.txt` Please ignore the source magnitude estimate given by `locate_multi` in the last but one line of the screen output! Using this information requires a non-standard FLEXPART output in bwd mode, i.e., concentration rather than residence time!

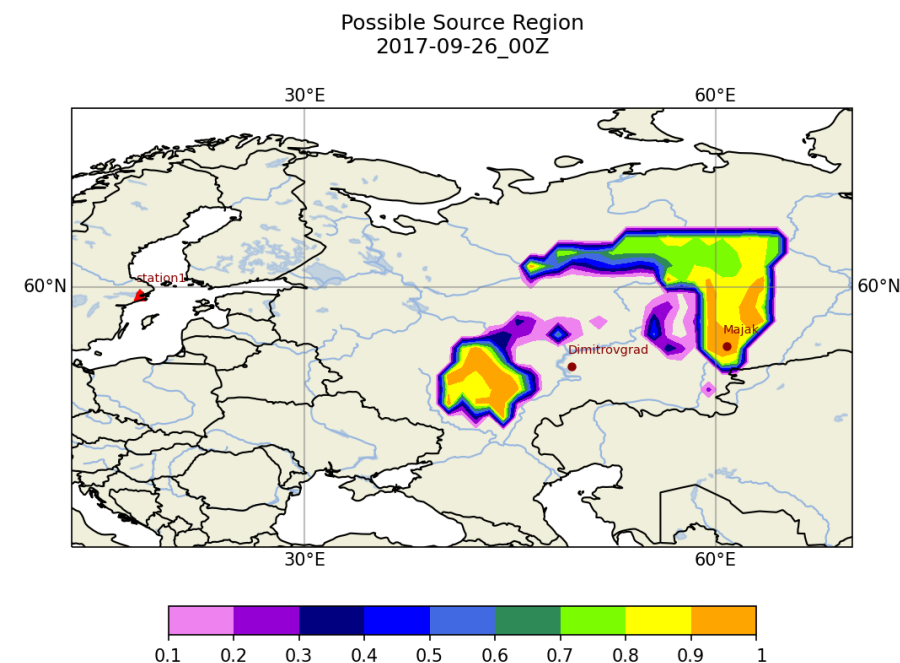


- Create conda environment *environment.yml* or install python 3.8 packages listed on the next slide manually
- Execute *python draw_corr_2D_plots.py*
- Related pyc-files in *./plotting_PSR/* (underlying python code is proprietary software of GeoSphere Austria)
- Generates intermediate output *correlation.txt.srm* (= the data) and *station_file* (= locations to plot)
- Plots: *corr__[yyyymmddhh].png*

ECMWF-ops:



ECMWF-ERA5:



To my best knowledge - package is very complex...:

- matplotlib
- gzip
- math
- os
- datetime
- sys
- pandas
- numpy
- logging
- time
- warnings
- mpl_toolkits
- scipy
- calendar
- re
- shutil
- cartopy
- subprocess
- imageio



FREAR (*Forensic Radionuclide Event Analysis and Reconstruction*) used for the first time at GeoSphere Austria in the context of the *NDC Preparedness Exercise* (NPE) 2024.

Outline:

- Developed and maintained by SCKCEN and Health Canada. Special thanks to Pieter de Meutter!
- Provides everything: From **FOR (SRS) overlap**, over **PSR** and **cost function analysis** (*deterministic approach, implicitly assumes normally distributed errors*) down to **Bayesian inference** (*stochastic approach*).
- **Assumes pollutant (radionuclide) originates from a single source at ground level.**
- **Source location may or may not be known.**
- Observation data must be provided with an uncertainty and a MDC (Minimum Detectable Concentration).
- A SRS (Source Receptor Sensitivity) field must be provided together with every observation.
- Continuous, rectangular or segmental releases can be inverted.

Bayesian source reconstruction of an anomalous Selenium-75 release at a nuclear research institute

Pieter De Meutter ^{a,b,c,*}, Ian Hoffman ^a

^a Radiation Protection Bureau, Health Canada, 775 Brookfield Road, Ottawa, Canada

^b Belgian Nuclear Research Institute, Boeretang 200, Mol, Belgium

^c Royal Meteorological Institute of Belgium, Ringlaan 3, Brussels, Belgium

Check for updates

ARTICLE INFO

Keywords:
Inverse modelling
Atmospheric dispersion modelling
Bayesian inference
Source reconstruction

ABSTRACT

Atmospheric transport and dispersion models are important tools in radiation protection as they help to estimate the impact of radionuclides released into the atmosphere. In particular, such models can be used in combination with radionuclide observations to estimate unknown source term parameters, or to improve source term estimates obtained through other methods. In this paper, a Bayesian inference system was used to determine the source term parameters and their corresponding credible intervals of a real-world anomalous ⁷⁵Se release at a nuclear facility in Belgium. Furthermore, a formulation is proposed that not only takes into account true detections, but also true instrumental non-detections, false alarms and real misses. The Bayesian inference system is able to correctly determine the known source location. The Bayesian inference is then refined by fixing the release location and by making stronger assumptions about the release period.

<https://www.sciencedirect.com/science/article/pii/S0265931X1930949X>

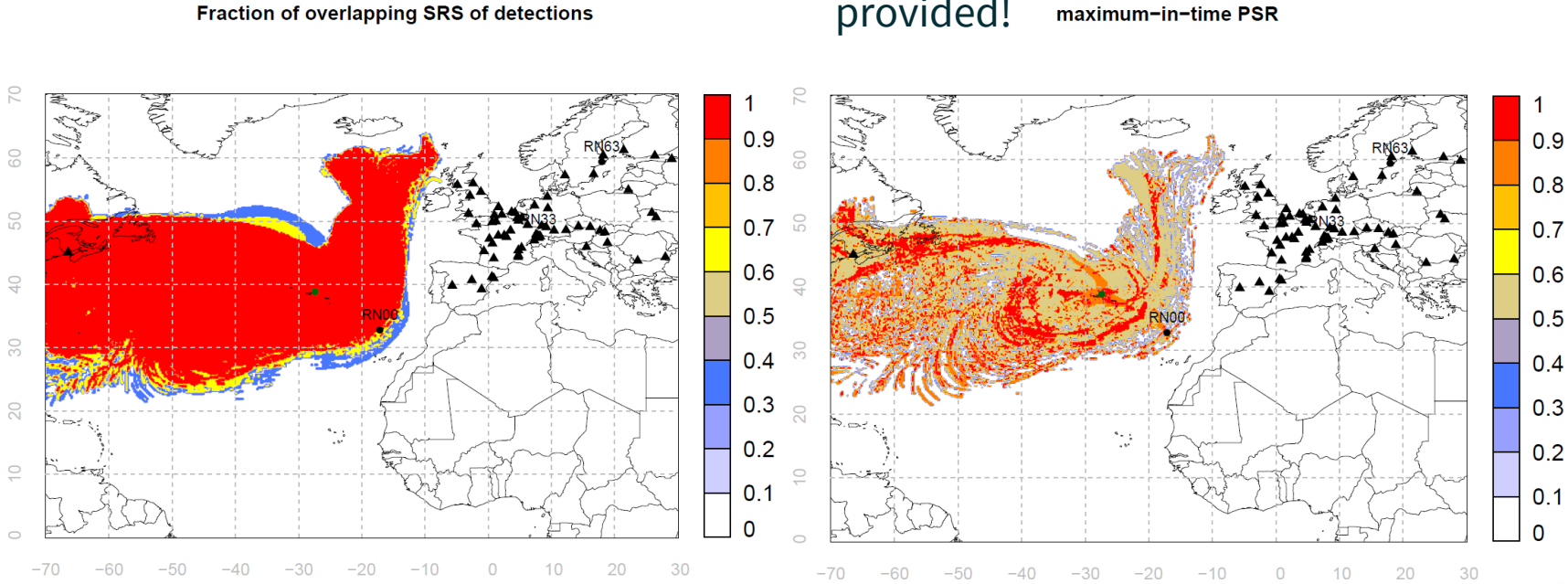
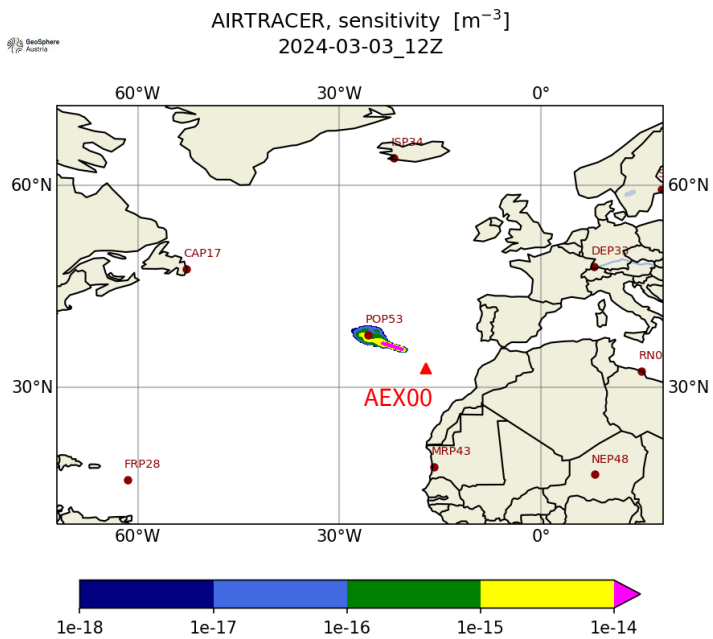


Relating synthetic NPE RN measurements to a waveform event

Station	Collection period	Waveform sens.	Measured Xe-133 /Xe-131m [mBq/m ³]
AEX00	2024030406-2024030418	25.2.-3.3.	2.8/0.6
AEX00	2024030418-2024030506	25.2.- 3.3.	0.2/-
AEX00	2024030506-2024030518	21.2.-3.3.	0.3/-

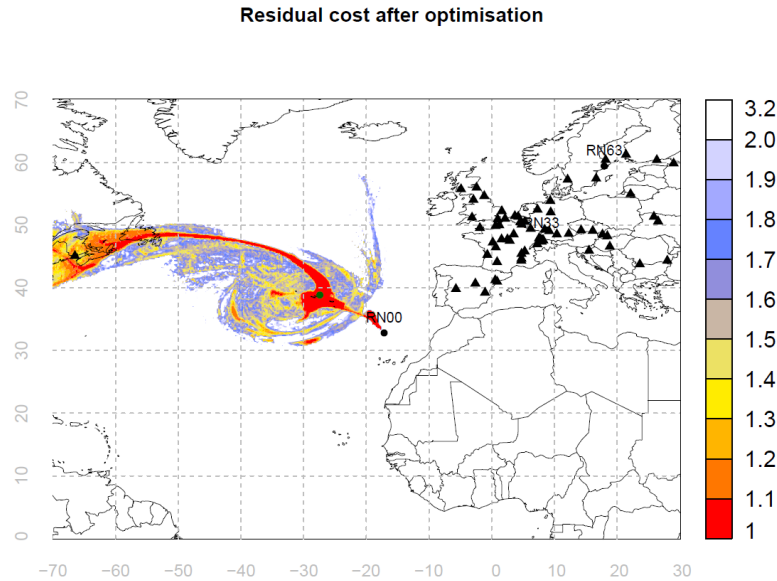
FOR overlap: Hypothesis of a common source supported.

Spearman PSR: Inconclusive. Broad & patchy. No uncertainties provided! maximum-in-time PSR



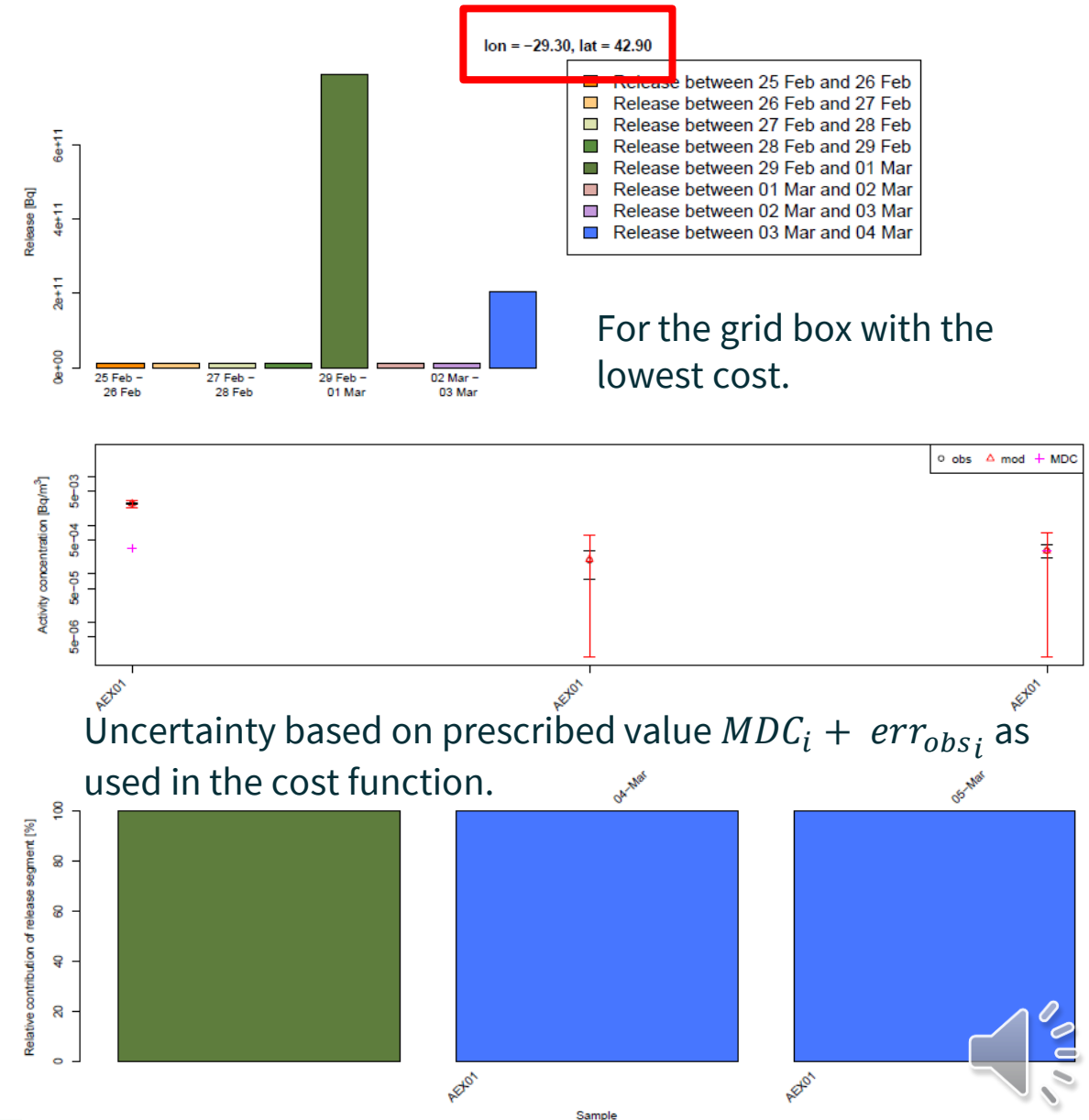
Green: Waveform event occurring on 14.1.2024
Integration period: 25.2.-3.3.





$$\frac{1}{N} \sum_{i=1}^N (\ln(obs_i) - \ln(mod_i))^2 + \log_{10}(1 + \sum_{j=1}^M \frac{par_j}{par_{max}})$$

$$mod_i[obs_i] < 2 * (MDC_i + err_{obs_i}) : mod_i[obs_i] = MDC_i + err_{obs_i} + mod_i[obs_i] * mod_i[obs_i] / (4 * (MDC_i + err_{obs_i}))$$
 is applied independently to every grid box of the domain. The results of a single optimization are (i) a set of optimized source parameters (with daily discretization j) and (ii) a resulting cost. Geometric mean of lower and upper plausible limit (set to $1E10$ and $1E15$ Bq) is taken as initial (daily) values.



$$p(\theta|c_{obs}) = \frac{p(c_{obs}|\theta) p(\theta)}{p(c_{obs})}$$

Product rule of probability calculus

$$\mathcal{L}(\theta) \equiv \mathcal{L}(\theta | \mathcal{D}, \varsigma) = \frac{1}{\prod_{i=1}^{N_m} \sqrt{2\pi} \sigma_i} \exp\left(-\frac{1}{2} \chi^2(\theta)\right)$$

Likelihood formulation

$$\chi^2(\theta) \equiv \sum_{i=1}^{N_m} \left(\frac{d_i - \tilde{C}_i(\theta)}{\sigma_i} \right)^2$$

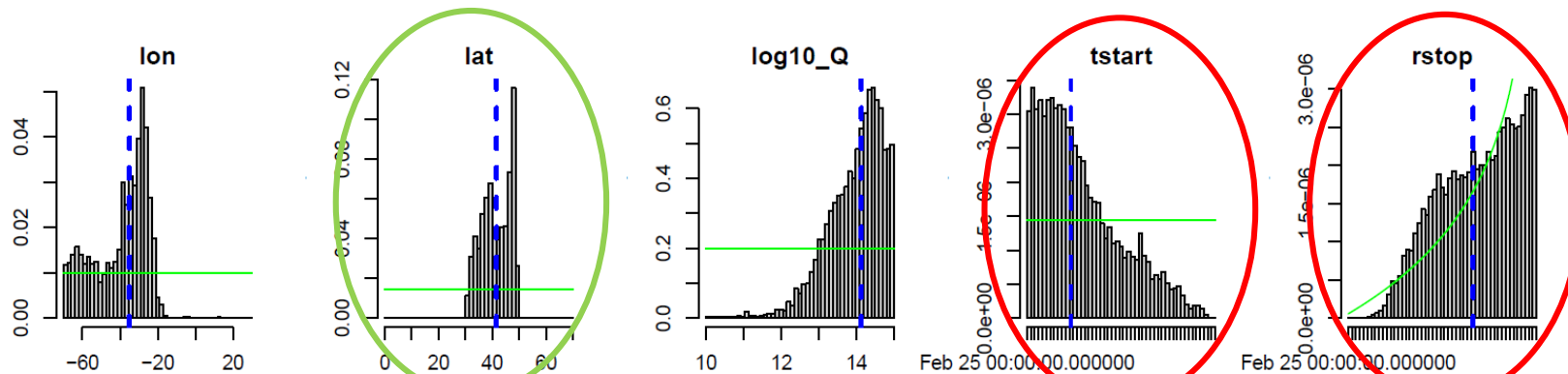
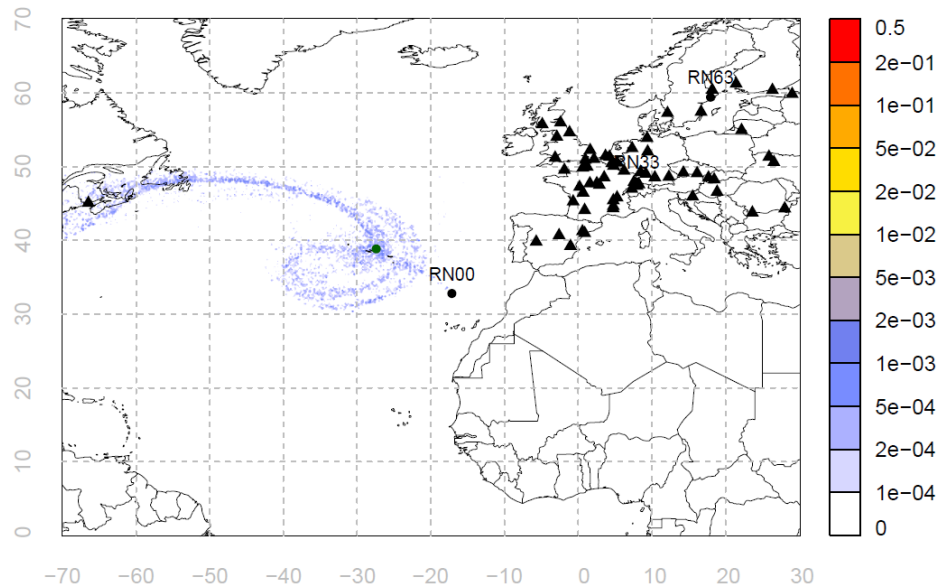
- $p(c_{obs})$: *Evidence* (normalization constant) given the prior information: Is there evidence in the concentration data for the existence of (a) source(s) in the domain investigated?
- $p(c_{obs}/\theta)$: *Likelihood*: Expresses how likely the observations are for a given source hypothesis. In practice this means comparing the observed and modelled activity concentrations.
- $p(\theta)$: *Uninformative, uniform a priori* is employed. The actual source location is supposed to be contained within a rectangular region with equal probability, the actual release time and the actual release strength occur within the intervals specified with equal probability too.
- $p(\theta/c_{obs})$: *Posterior probability density function*: Highly nonlinear, analytical evaluation of multi-dimensional integrals required for desired marginal posteriors and various statistics which is impossible ->

Markov Chain Monte Carlo (MCMC) algorithm is used to efficiently sample the posterior probability, respectively the high plausability regions of the source parameter space. *Multiple chains* are employed to check convergence.



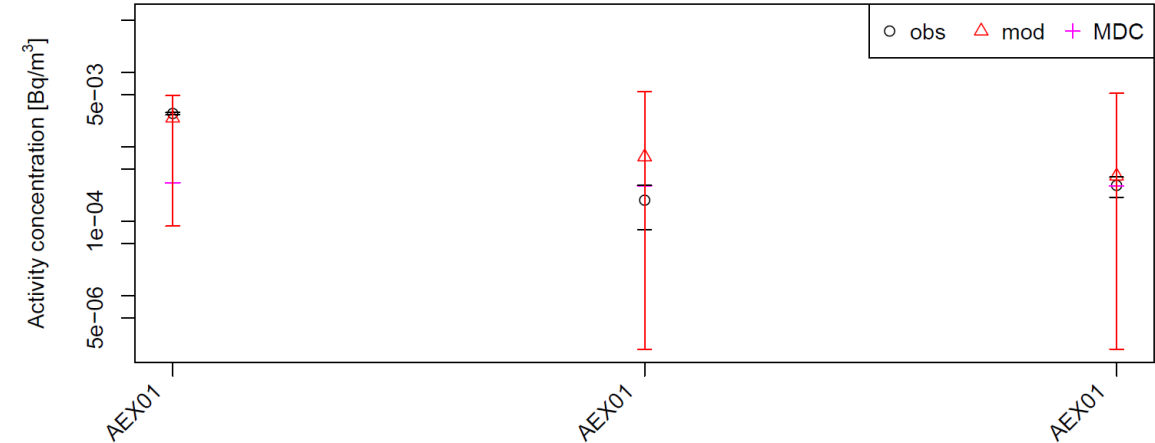
Counting the number of times a sample of the marginal bivariate longitude-latitude posterior distribution was present in a certain grid box divided by the total number of samples

Bayesian source location probability



Full pdf for five parameters.

Modelled concentrations using full chain



Poorly constrained parameters
tend toward the a priori.

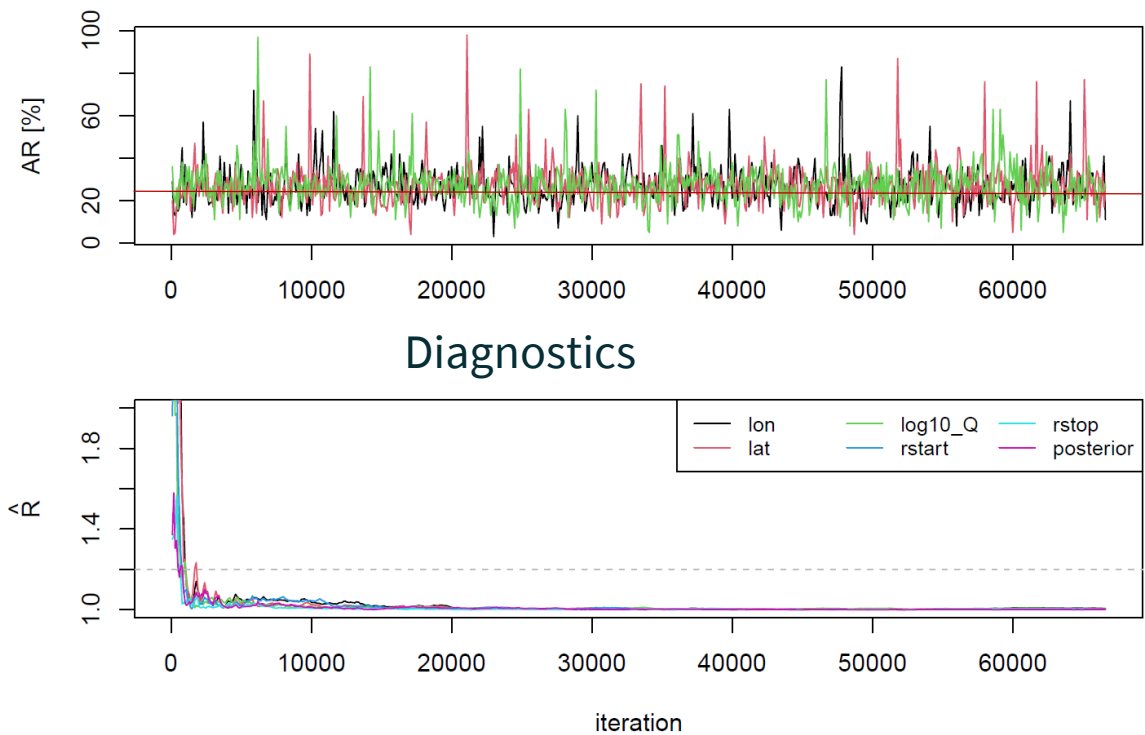
srsrelError=0.5

Modelled activity uncertainties obtained considering the full **MCMC chain**: For each sample of source parameters the corresponding activity concentrations are calculated for the 2.5 to 97.5% credible interval.



	lon	lat	log10_Q	rstart	rstop
0.025	-67.85	31.50	12.25	2024-02-25 02:06	2024-02-27 04:01
0.5	-35.11	41.48	14.13	2024-02-26 20:47	2024-03-01 16:33
0.975	-21.12	48.93	14.95	2024-03-02 08:37	2024-03-03 21:35
mean	-39.16	41.32	13.97	2024-02-27 06:53	2024-03-01 11:50

Actual location: -27.41 °E and 38.84°N . A considerable improvement compared to PSR fields....!

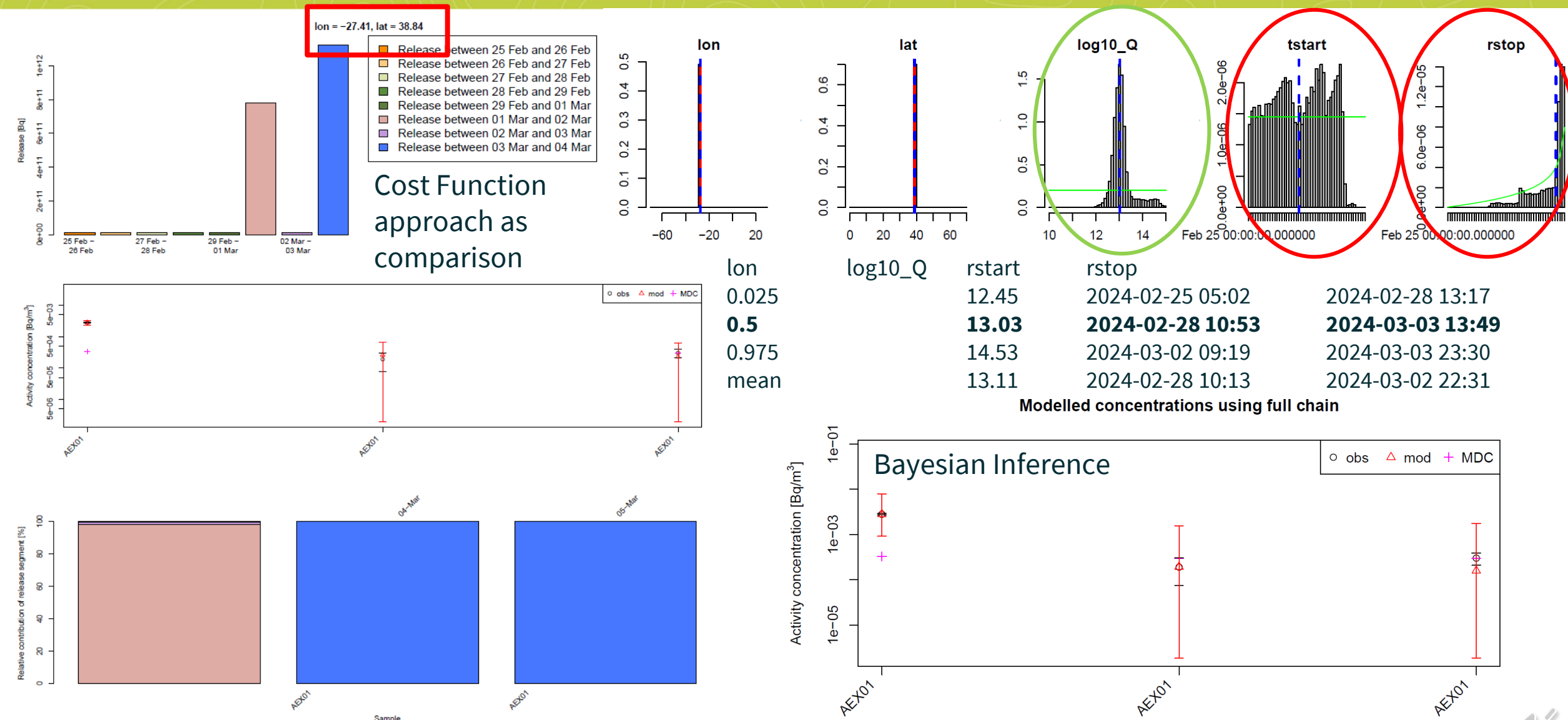


200000 iterations in total -> ~ 70000 per chain.

- Overall *acceptance rate (AR)* for the three Markov chains are above the optimum asymptotic acceptance rate of 23.4% -> *Sampling was efficient*. Acceptance rates – calculated over several 100s of iterations - vary during the sampling since these are dependent on the position of the source parameters relative to the posterior distribution. If rates are too low, reducing the perturbation step when generating a new candidate point for the source parameters might help.
- *Gelman and Rubin reduction coefficient (R)* around 1 -> *Algorithm converged*. Variance for source parameters within chains is comparable to variance between the 3 chains.



Refining the analysis: Fixing the source location



Agreement between cost function and Bayesian inference on an accumulated release of ~1E13 Bq between Feb., 28th, and March, 4th. Time continues to be poorly constrained - 12 hourly sampling intervals likely too coarse.

THANK YOU

Environmental Department/CBRN-Incident Modeling
christian.maurer@geosphere.at

