Identification of atmospheric contamination source in an urban area by approximate bayesian computation methodology

NATIONAL CENTRE FOR NUCLEAR RESEARCH
ŚWIERK

Division of Nuclear Energy and Environmental Studies

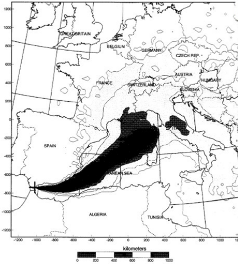
Piotr Kopka, Anna Wawrzynczak
National Centre for Nuclear Research

NERIS

Roskilde 05.04.2019

Motivation: historical events

- Accidental atmospheric releases of hazardous material:
- Release of ^{137}Cs in a steel mill in Algeciras, Spain, V.1998 r.
- Registered in Switzerland, France and Italy in June.

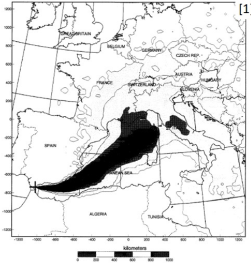


[1]

Figure 2. Seven-day average air concentration at 1200 UTC on 5 June 1998. Contours >10 (outermost or lightest), >100, >500, and >1000 uBq/m3 (innermost or darkest).

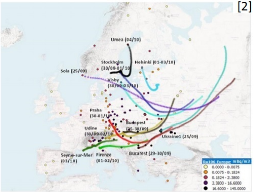
Motivation: historical events

- Accidental atmospheric releases of hazardous material:
- Release of ^{137}Cs in a steel mill in Algeciras, Spain, V.1998 r.
- Registered in Switzerland, France and Italy in June.
- ^{131}I - 2011,2012,2015,2017, ^{137}Cs-2015, ^{106}Ru - 2017.
- Source location was not known at the time when the first detections were reported.



[1]

Figure 2. Seven-day average air concentration at 1200 UTC on 5 June 1998. Contours >10 (outermost or lightest), >100, >500, and >1000 uBq/m3 (innermost or darkest).



[2]

Motivation: the urbanized areas

- Predicting the transport and dispersion of the contaminant becomes a critical problem for homeland defense


[3]

Motivation: the urbanized areas

- Predicting the transport and dispersion of the contaminant becomes a critical problem for homeland defense
- either **accidental** or **intentional**,
- either **toxic** or **radioactive**, etc

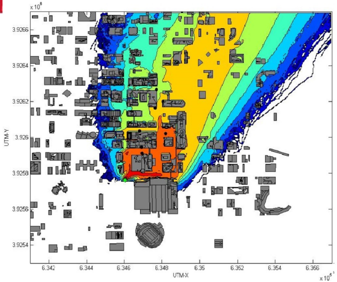
Motivation: the urbanized areas

- Predicting the transport and dispersion of the contaminant becomes a critical problem for homeland defense
- either **accidental** or **intentional**,
- either **toxic** or **radioactive**, etc
- The **most difficult** and requiring the **most resources** is the contamination in urbanized areas (micro-scale).

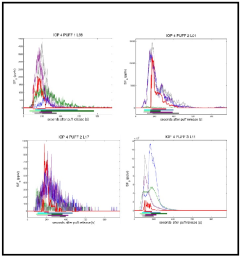
Motivation: the urbanized areas

- Predicting the transport and dispersion of the contaminant becomes a critical problem for homeland defense
- either **accidental** or **intentional**,
- either **toxic** or **radioactive**, etc
- The **most difficult** and requiring the **most resources** is the contamination in urbanized areas (micro-scale).
- Because, accurate modeling of atmospheric contaminant transportation in a dense urban area is not trivial.

Forward modeling: concentration point estimation

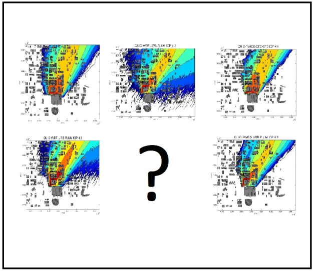


Forward

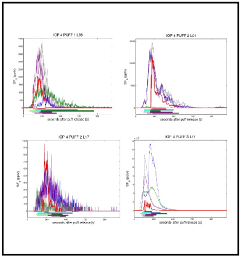


-
-
-

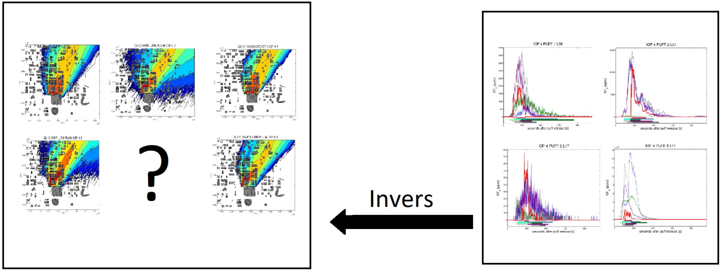
Inverse modeling: source terms estimation



Invers

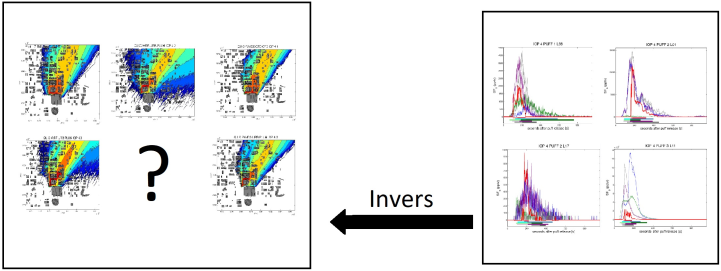


Inverse modeling: source terms estimation



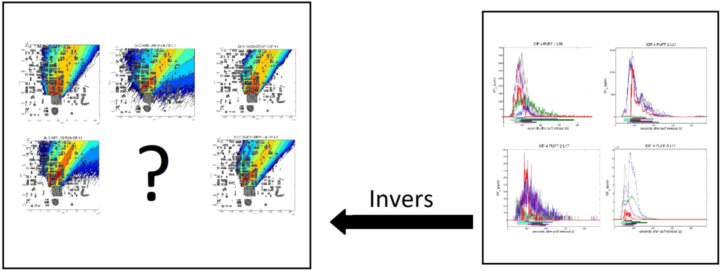
- *Questions*: How much material was released? When? Where? What are the potential consequences?

Inverse modeling: source terms estimation



- *Questions*: How much material was released? When? Where? What are the potential consequences?
- *Idea*: Build a model of pollutant transport in the atmosphere and compare estimated point concentrations with the measured data obtained from sensor networks.

Inverse modeling: source terms estimation



- *Questions*: How much material was released? When? Where? What are the potential consequences?
- *Idea*: Build a model of pollutant transport in the atmosphere and compare estimated point concentrations with the measured data obtained from sensor networks.
- *Problem*: Find the values of the pollutant transport model parameters, for which outcome will be the best „fitted" to the observational data.

Bayesian inference

- In the framework of Bayesian statistics all quantities are modeled as random variables with joint probability distributions.
- This randomness can be interpreted as parameter variability.
- It is reflected in the uncertainty of the true values.
- So, in practice we can are looking for the values of parameters which are the most probable - **Posterior Probability Distribution**.

Bayes theorem

$$\underbrace{\pi(\theta|d_{obs}, I)}_{posterior} = \frac{\overbrace{L(d_{obs}|\theta, I)}^{likelihood}\,\overbrace{\pi(\theta|I)}^{prior}}{\underbrace{\pi(d_{obs}|I)}_{evidence}} \propto \pi(d_{obs}|\theta, I)\pi(\theta|I) \qquad (1)$$

- θ represents possible model configurations e.g $\theta \equiv (x, y, q, \ldots)$, d_{obs} are observed data e.g $d_{obs} \equiv C_t^{Sj}, ..., C_t^{SN}$, e.g I background information (e.g. meteorological measurements)
- Probability $\pi(\theta|d_{obs}, I)$ of certain model configuration given observed measurements (d_{obs}) (also known as the posterior distribution)
- $L(d_{obs}|\theta, I)$ - the probability of the data d_{obs} conforming a given model configuration θ
- $\pi(\theta|I)$ - the possible model configurations before taking into account the measurements

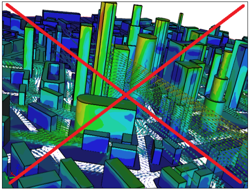
How to obtain posterior distribution of model parameters?

- Approximate Bayesian Computation with the sequential extension [5].
- *Idea*: Accept θ as an approximate posterior draw if its associate data d is close enough to the observed data d_{obs}.
- d - expected concentration in sensors locations with source θ_i setup: $MODEL(\theta_i) \rightarrow d$
- $\rho(d, d_{obs})$ - chosen measure of discrepancy between d and d_{obs},
- ϵ - threshold value.
- So, parameters are a sample from $\pi(\theta|\rho(d, d_{obs}|I) < \epsilon)$

Requirements for dispersion model for STE in urban environment?

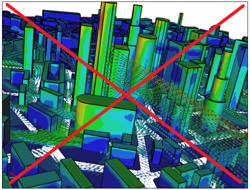
Requirements for dispersion model for STE in urban environment?

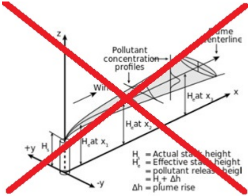
- Not to complicated and not required enormous computing power.



Requirements for dispersion model for STE in urban environment?

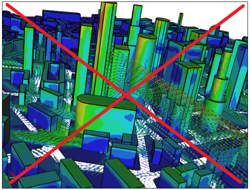
- Not to complicated and not required enormous computing power.
- But, should take into account necessary parameters such as wind field, the coefficients of turbulence, weather conditions, etc.

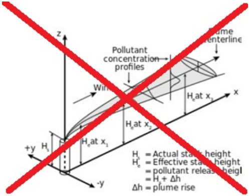


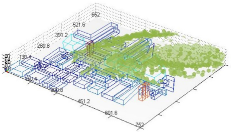


Requirements for dispersion model for STE in urban environment?

- Not to complicated and not required enormous computing power.
- But, should take into account necessary parameters such as wind field, the coefficients of turbulence, weather conditions, etc.
- A short (**QUIC**K!) computation time.

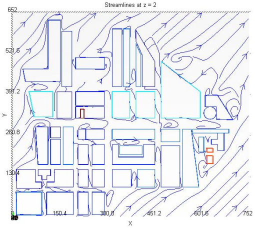






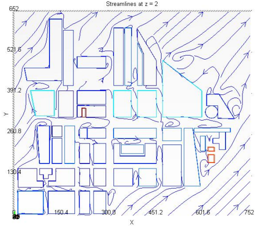
Quick Urban Industrial Complex (QUIC) Dispersion Modeling System Los Alamos

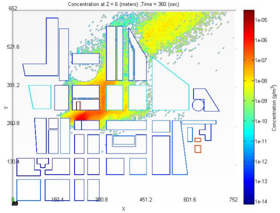
- QUIC-URB (originally developed by Rockle [6]) uses a 3D mass-consistent wind model to combine adequately resolved time-averaged wind fields around buildings



Quick Urban Industrial Complex (QUIC) Dispersion Modeling System Los Alamos

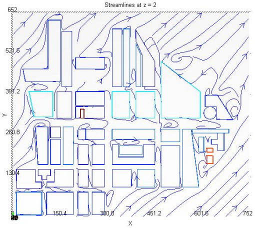
- QUIC-URB (originally developed by Rockle [6]) uses a 3D mass-consistent wind model to combine adequately resolved time-averaged wind fields around buildings
- QUIC-PLUME [7] is a Lagrangian particle model which describes gas dispersion by simulating the release of particles and moving them with an instantaneous wind composed of mean and turbulent components.

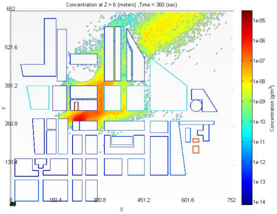




Quick Urban Industrial Complex (QUIC) Dispersion Modeling System Los Alamos

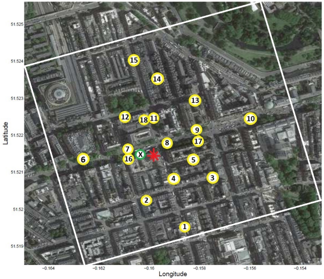
- QUIC-URB (originally developed by Rockle [6]) uses a 3D mass-consistent wind model to combine adequately resolved time-averaged wind fields around buildings
- QUIC-PLUME [7] is a Lagrangian particle model which describes gas dispersion by simulating the release of particles and moving them with an instantaneous wind composed of mean and turbulent components.
- Radioactive dispersion module is also supported





DAPPLE

- *Dispersion of Air Pollution and its Penetration into the Local Environment*
- *The building height* 10 *to* $64m$
- *Total mass emitted was* $323mg$ *of* $(PMCH, C7F14)$ *for 15 min*
- *10 samples taken over a 30 minute sampling period at 18 receptor*
- *The wind data sets take from rooftop Westminster City Council (WCC) (18 m)*
- *Funded by the Engineering for Health, Infrastructure and Environment Programme*



source position x=243.3m, y=282.8m, z=1.5m, mass q=323mg, duration and delay time l=900s, s = 0.0s

Final Bayesian STE framework setup

- Algorithm - ABC with Sequential Monte Carlo, Forward Model - QUIC , Data/Experiment: DAPPLE (28 Jun 07)
- Distance measure (Fractional Bias):

$$\rho(d^t, d^t_{obs}) = \frac{1}{18}\sum_{j=1}^{18}\left(\frac{1}{t}\sum_{i=1}^{t}\frac{|C_i^{Sj} - \hat{C}_i^{Sj}|}{C_i^{Sj} + \hat{C}_i^{Sj}}\right), \tag{2}$$

- Source parameters vector and prior definition:

$$\begin{aligned}
\pi(\theta^1) \equiv &(x, y) \sim U^{\Theta}([100, 600], [100, 500]) - (x = 243.3m, y = 282.8m)\\
&z \sim Gamma(3, 3) - z = 1.5m\\
&q \sim U(10, 500) - q = 323mg\\
&l \sim U(0, 1800) - l = 900s\\
&s \sim U(0, 180) - s = 0.0s
\end{aligned} \tag{3}$$

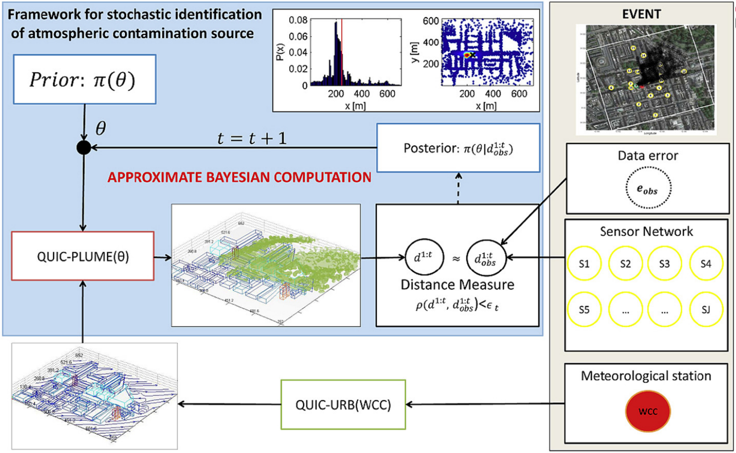


Figure 1: Framework for stochastic identification of atmospheric contamination source in an urban area

Results of reconstruction experiment

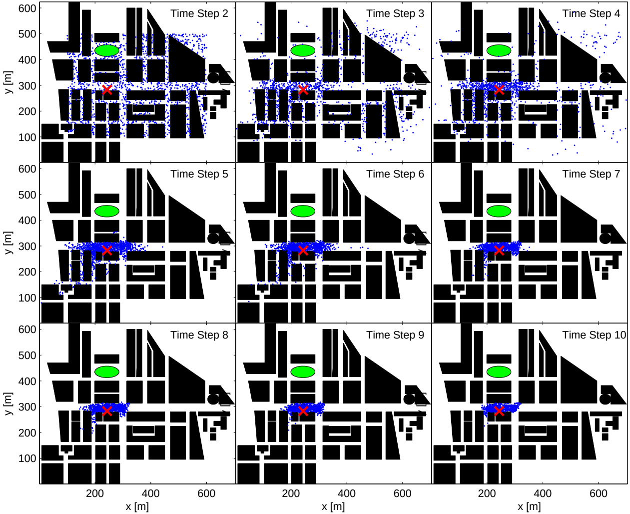


Figure 2: Scatter plot of all samples generate in the subsequent time steps $t = 2, 3, \ldots, 10$ in (x, y) space

Results of reconstruction experiment

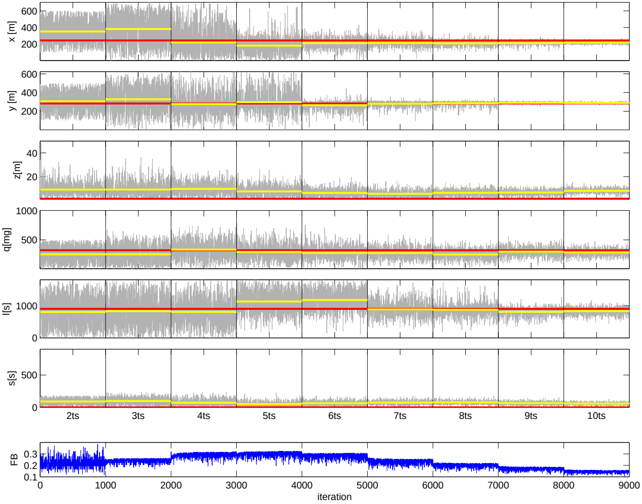


Figure 3: The trace plots for all searched parameters $\theta \equiv (x, y, z, q, l, s)$ in all time steps.

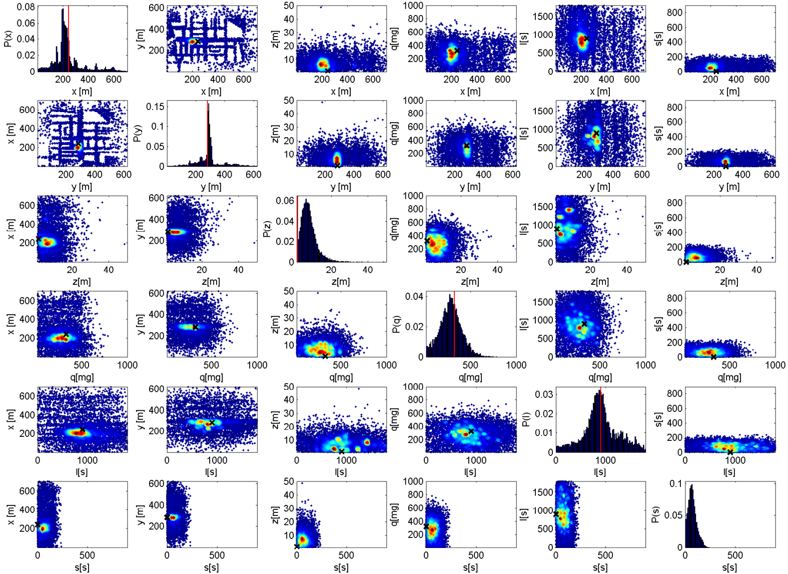


Figure 4: Bivariate and marginal posterior distributions for all parameters

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI_{LB}^{90\%}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI_{UB}^{90\%}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI^{90\%}_{LB}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI^{90\%}_{UB}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

- Posterior distributions strongly support the correct location of the source (x, y)

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI_{LB}^{90\%}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI_{UB}^{90\%}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

- Posterior distributions strongly support the correct location of the source (x, y)
- The difference between the estimated z value and target source height equals $\approx 3.5m$.

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI_{LB}^{90\%}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI_{UB}^{90\%}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

- Posterior distributions strongly support the correct location of the source (x, y)
- The difference between the estimated z value and target source height equals $\approx 3.5m$.
- The released mass is underestimated about $\sim 50mg$.

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI_{LB}^{90\%}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI_{UB}^{90\%}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

- Posterior distributions strongly support the correct location of the source (x, y)
- The difference between the estimated z value and target source height equals $\approx 3.5m$.
- The released mass is underestimated about $\sim 50mg$.
- The most probable duration of the release was estimated almost perfectly.

Results of reconstruction experiment

Parameters	$x[m]$	$y[m]$	$z[m]$	$q[mg]$	$l[s]$	$s[s]$
θ^*	243.3	282.8	1.5	323.0	900.0	0.0
θ^{MAP}	203.5	291.8	5.4	265.7	881.0	71.0
$mean(\theta)$	243.9	290.9	7.6	280.0	903.3	72.4
$std(\theta)$	124.6	84.3	4.1	118.7	352.3	41.8
$CI_{LB}^{90\%}(\theta)$	35.0	150.0	1.1	59.1	359.7	0.7
$CI_{UB}^{90\%}(\theta)$	466.2	461.6	12.8	461.3	1624.0	130.9

Table 1: Summary of the essential statistics obtained from the results of inversion procedure.

- Posterior distributions strongly support the correct location of the source (x, y)
- The difference between the estimated z value and target source height equals $\approx 3.5m$.
- The released mass is underestimated about $\sim 50mg$.
- The most probable duration of the release was estimated almost perfectly.
- One minute delay can be seen as important.

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.
- Bayesian inference (ABC) with fast-running QUIC-PLUME dispersion model has been successfully validated against real DAPPLE experiment.

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.
- Bayesian inference (ABC) with fast-running QUIC-PLUME dispersion model has been successfully validated against real DAPPLE experiment.
- This lead to estimation of six atmospheric contamination source parameters i.e.: contamination source position (x, y, z) in city environment, mass of release (q), start time of release (s) and its duration (l).

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.
- Bayesian inference (ABC) with fast-running QUIC-PLUME dispersion model has been successfully validated against real DAPPLE experiment.
- This lead to estimation of six atmospheric contamination source parameters i.e.: contamination source position (x, y, z) in city environment, mass of release (q), start time of release (s) and its duration (l).
- Clear interpretation for decision makers.

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.
- Bayesian inference (ABC) with fast-running QUIC-PLUME dispersion model has been successfully validated against real DAPPLE experiment.
- This lead to estimation of six atmospheric contamination source parameters i.e.: contamination source position (x, y, z) in city environment, mass of release (q), start time of release (s) and its duration (l).
- Clear interpretation for decision makers.
- Can be used for a radiological event.

Conclusions

- The framework for stochastic event reconstruction of the atmospheric contaminant in an urban environment has been presented.
- Bayesian inference (ABC) with fast-running QUIC-PLUME dispersion model has been successfully validated against real DAPPLE experiment.
- This lead to estimation of six atmospheric contamination source parameters i.e.: contamination source position (x, y, z) in city environment, mass of release (q), start time of release (s) and its duration (l).
- Clear interpretation for decision makers.
- Can be used for a radiological event.
- Published: *Framework for stochastic identification of atmospheric contamination source in an urban area P Kopka, A Wawrzynczak Atmospheric Environment 195, 63-77*

THANK YOU
FOR YOUR ATTENTION

Bibliography

1 Aluzzi, F. J., at al. (1999). Comparison of gridded versus observation data to initialize ARAC dispersion models for the Algeciras, Spain steel mill CS-137 release. Lawrence Livermore National Lab., CA (US).

2 Report on the IRSN's investigations following the widespread detection of 106Ru in Europe early October 2017

3 https://www.uvu.edu/esa/jackrabbit/docs/jrii17/finalreports/uvu-jack-rabbit-final-report-2017.pdf

4 http://www.dapple.org.uk

5 F. V. Bonassi, M. West, et al., Sequential Monte Carlo with Adaptive Weights for Approximate Bayesian Computation, Bayesian Analysis 10 (1) (2015) 171-187.

6 R. Rockle, Bestimmung der Stromungsverhaltnisse im Bereich komplexer Bebauungsstrukturen, na, 1990.

7 M. D. Williams, M. J. Brown, B. Singh, D. Boswell, Quic-plume theory guide, Los Alamos National Laboratory (2004) 43.

Roskilde 05.04.2019