

Estimation of Cs-137 emissions during wildfires and dust storm in Chernobyl Exclusion Zone in April 2020 using ensemble iterative source inversion method

Ivan V. Kovalets^{a,*}, Mykola Talerko^b, Roman Synkevych^a, Serhii Koval^a

^a Institute of Mathematical Machines and Systems Problems NAS of Ukraine, Prosp. Glushkova, 42, 03187, Kyiv, Ukraine

^b Institute for Safety Problems of Nuclear Power Plants, NAS of Ukraine, 12, Lysogirska St., Kyiv, 03028, Ukraine

HIGHLIGHTS

- Ensemble Iterative Source Inversion Method (EISIM) was developed.
- Uncertainties in source term parameters and meteorological data were taken into account.
- Emission inventories of ¹³⁷Cs resulting from wildfires in Chernobyl Exclusion Zone were estimated.
- Contribution of emissions resulting from dust storm was evaluated.

ARTICLE INFO

Keywords:

Inverse method
Wildfire
Dust storm
Chernobyl exclusion zone
FLEXPART
Source term estimation
Ensemble
Data assimilation

ABSTRACT

The emission inventories of cesium-137 resulting from the wildfires in the Chernobyl Exclusion Zone (3–24 April 2020) and from the dust storm (16–17 April 2020), which resuspended contaminated ash, were estimated using inverse modeling. The goal of this work was to take into account uncertainties of inexact known source term and meteorological input data for evaluation of emission inventories and their confidence limits, by developing Ensemble Iterative Source Inversion Method (EISIM). A set of source receptor matrices (SRMs) was calculated using FLEXPART atmospheric transport code with varying source term parameters (size distribution of emitted particles, height distribution of emissions) and meteorological input data. The covariance matrix of model errors was estimated by ensemble averaging of model results obtained after multiplication of the pre-calculated SRMs by the estimated emission inventories at the current iteration step. The emission inventories at each iteration step were evaluated for each of the ensemble members by solving the conventional variational source inversion problem. With iterations, the variance of model error was reduced by an order of magnitude. The estimated total emission of ¹³⁷Cs from wildfires was 448 GBq, close to the first guess estimation. By using emission inventories within the obtained confidence limits (from 39 to 1530 GBq), different combinations of source term parameters and input meteorological data, FLEXPART could fit observations with a correlation coefficient of more than 0.6 and a normalized mean squared error of less than 10. The obtained estimate of the total emission resulting from the wind resuspension during the dust storm was 27 GBq. The respective estimated confidence interval was from 3 to 93 GBq. By analyzing model error statistics, some of the source term parameters could be reliably evaluated. The fraction of the fine particles (0.25 μm) in total emissions $W_{sf(0.25)} \approx 0.1$ and the fraction of emission below a bottom height of convective plume $W_{h(1)} \approx 0.5$.

1. Introduction

Source term estimation (STE) methods are important tools used for evaluation of emission inventories following accidental releases of

hazardous pollutants in the atmosphere (Enting, 2002). The common framework to accomplish this task includes an atmospheric dispersion model (ADM) to evaluate concentrations of pollutant in places and at times of measurements and a minimization procedure to find an optimal

* Corresponding author..

E-mail addresses: ivkov084@gmail.com, ik@env.com.ua (I.V. Kovalets), ntalerko@gmail.com (M. Talerko), sunkevu4@gmail.com (R. Synkevych), sergejkoval167@gmail.com (S. Koval).

<https://doi.org/10.1016/j.atmosenv.2022.119305>

Received 30 May 2022; Received in revised form 22 July 2022; Accepted 26 July 2022

Available online 2 August 2022

1352-2310/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

estimate of emission inventories that minimizes deviation of simulated values from the respective measurements. However, different STE problems can have specific challenges that require special considerations for development of STE algorithms. For example, results of atmospheric transport can be sensitive to meteorological conditions, and input parameters of atmospheric transport models can be poorly known. So those details are to be accounted for by STE methods. Proper consideration of model uncertainties caused by errors in meteorological input data and source term parameters, namely size distribution of emitted particles and height distribution of the release, requires combining STE methodologies with ensemble data assimilation methods, such as Evensen (2009), Zhang et al. (2015), Zheng et al. (2007). One example of a problem that requires advanced STE methods is estimation of emission inventories resulting from wildfires in radioactively contaminated territories. Some of the complexities related to such events are listed below. Existing estimates of emission factor of radionuclide emissions during wildfires differ considerably, as reviewed by Ager et al. (2019). Even though the dynamics of wildfire areas is usually well known from the satellite data, the dynamics of the ‘burn factor’, a fraction of the burned fuel relative to its total amount on the territory of fire, is usually poorly known. This additionally complicates the STE task as mentioned by Kovalets and de With (2020). Uncertainties in size distribution of emitted radioactive aerosols and height distribution of the emission rates are significant.

All above-mentioned complexities were present in the simulation studies of a recent case of wildfire in the Chernobyl Exclusion Zone (ChEZ) in April 2020, which appeared to be the biggest wildfire in the history of ChEZ (Masson et al., 2021). Several studies of this extreme wildfire used different prior estimations of emission inventories, different assumptions regarding emission heights and size distribution of particles, etc (Masson et al., 2021; Evangeliou and Eckhardt, 2020; Talerko et al., 2021a; De Meutter et al., 2021; Baró et al., 2021; Protsak et al., 2020). An additional unique feature of this event was that on 16–17 April 2020 the wildfire in ChEZ was accompanied by a dust storm (Savenets et al., 2020; Talerko et al., 2021a, 2021b). The dust storm led to additional emissions of radionuclides, which were not taken into account in STE studies related to wildfires in ChEZ. Because of all mentioned complexities, significant deviations between different estimates of emissions exist, especially for the period of the dust storm (16–17 April) and for the last days of the wildfire (a more detailed review of previous estimates is presented in the text below). Therefore, this paper aims at taking into account uncertainties in the distribution of emitted particles, height distribution of releases, and meteorological input data in the STE process by developing an ensemble source inversion methodology and using it for quantification of emission inventories resulting from wildfires and dust storm in ChEZ in April 2020.

In the following, the paper first presents description of the wildfire in ChEZ in April 2020 and the relevant measurement data, together with a brief review of the previous STE studies (Section 2). The developed ensemble iterative source inversion method is then presented in Section 3. Section 4 presents results of the application of the developed STE method to the simulation of the wildfire event in ChEZ in April 2020.

2. Description of wildfires in ChEZ in April 2020

2.1. A brief review of the event

The biggest in the history of ChEZ wildfire started on April 3, 2020 (Talerko et al., 2021a). The intense period of the wildfire lasted more than two weeks and engulfed more than 40,000 ha of forests and grasslands in ChEZ (Talerko et al., 2021a). At the same time, wildfires happened outside ChEZ in the northern and northwestern parts of Ukraine, in Kyiv and Zhytomyr regions. On April 13, 2020, the wildfire was very close to the industrial site of Chernobyl NPP and spread on the highly contaminated territories of the ‘Red Forest’. Emissions of radionuclides during the wildfire lead to an increase in airborne

concentration of radionuclides in ChEZ by 1–4 orders of magnitude. In parts of Ukraine outside ChEZ, concentration of radionuclides also increased considerably. For example, the maximum detected daily average volume concentration of cesium-137 in Kyiv ($700 \mu\text{Bq}\cdot\text{m}^{-3}$, reported on 11 April) exceeded background values more than 100 times, albeit this value was still about 1000 times less than regulatory permissible $0.8 \text{ Bq}\cdot\text{m}^{-3}$, according to Ukrainian radiation safety regulations (NRBU, 1997). Radionuclides emitted by wildfires in ChEZ were also detected in many European countries (Masson et al., 2021). For example, weekly average values of the ^{137}Cs volume activity concentration detected in northeastern Germany and southeastern France were several times larger than the typical background in the respective areas (Masson et al., 2021).

In addition to the anthropogenic factor, extremely dry meteorological conditions in the preceding few months contributed to big wildfires, which were reviewed in a more detail by Talerko et al. (2021a). According to the last measurements in Chernobyl before 3 April, precipitation of moderate intensity (more than 3 mm/day) happened on March 10, 2020. In the preceding few weeks before wildfires, the dry sunny weather with air temperatures up to 20°C in the daytime alternated with frosts at night, when air temperatures dropped down to -7°C . The fire danger was characterized as ‘an extreme fire hazard’ according to Nesterov’s index, which is officially used in Ukraine for fire hazard classification (Eastaugh and Hasenauer, 2014). On 14 April, due to a cold front, there were rains in northwestern and north-central Ukraine, including ChEZ. This, together with the efforts of firefighters, led to a temporary stop of fires during the next 2 days. However, on 16 April, strong wind was observed in the whole northern Ukraine. Meteorological stations in Kyiv and Rivne detected wind gusts up to 19 and $20 \text{ m}\cdot\text{s}^{-1}$ respectively. This meteorological situation led to renewal of wildfires in ChEZ that lasted until approximately 23 April. The strong winds also led to powerful dust storms over a large part of Ukraine, including Rivne, Zhytomyr, and Kyiv regions (Talerko et al., 2021a, 2021b; Savenets et al., 2020). Due to the dust storm, daily average concentrations of $\text{PM}_{2.5}$, in particular in Kyiv, in the period of 16–18 April 2020 exceeded permissible values ($25 \mu\text{g}\cdot\text{m}^{-3}$) up to 6 times (Talerko et al., 2021a). On other days of April 2020, even when wildfires were strong and the wind was blowing from ChEZ to Kyiv, concentrations of $\text{PM}_{2.5}$ in Kyiv were much lower than during the dust storm. Under normal conditions, ChEZ is covered by vegetation, therefore, even high wind usually does not lead to intense wind resuspension of deposited radioactive materials. In contrast, wind lift of radioactivity from the conflagration territories can be large, especially during or immediately after wildfires, when newly formed ash is not yet mixed with the soil by mechanical mixing, wet removal, and chemical processes. Therefore, it seems that the intense resuspension of radionuclides on 16–17 April resulted from the compound extreme event of simultaneous wildfires and dust storm.

2.2. Previous STE studies

Estimation of emission inventories of radioactivity resulting from wildfires in ChEZ started during the wildfires (IRSN, 2020; Protsak et al., 2020). Later, more detailed studies appeared (Masson et al., 2021; Evangeliou and Eckhardt, 2020; Talerko et al., 2021a, De Meutter et al., 2021, Baró et al., 2021). Three main approaches to quantify emissions were used in different studies: a) estimating the area of fires with the help of satellite data and then using the emission factor to calculate emitted inventories (Protsak et al., 2020, 2022); b) calculating emissions by inverse simulation without a priori information on the emitted inventories during wildfires (Masson et al., 2021, Baró et al., 2021); c) combining approaches 1) and 2) when the estimates obtained by a) were adapted to better fit calculated concentrations to measurements (Evangeliou and Eckhardt, 2020; Talerko et al., 2021a; De Meutter et al., 2021). The difficulty of studies of type a) is that for the correct calculation of emissions by this method it is necessary to evaluate percentage

of biomass burnout, in addition to the area of wildfires (Kovalets and de With, 2020). The most rigorous methods of this type require even more data, namely specific activities of radionuclides in different compartments of the burnt biomass, together with the corresponding emission fractions (Protsak et al., 2022). Estimates by method b) are sensitive to the amount and accuracy of measurement data due to poor conditioning of the inverse problem (Enting, 2002). In most countries, levels of the observed ^{137}Cs volume concentrations were only 2–3 times higher than the average annual background values. The average daily background concentrations could fluctuate several times around the average values. Therefore, the usefulness of such measurements in solving the inverse problem is limited, because their “representative” error is relatively large. The most general way to estimate emissions is method c), but the use of inverse modeling methods requires adequate estimates of model errors and the first guess (prior) estimation of the solution. These model errors are the result of errors in the meteorological input data, physical assumptions of the model, assumptions about the size composition of particles, emission heights, and so on. Differences in these parameters lead to significant differences in estimates of emissions. For example, even estimates of the total emissions for the entire period of the fire, obtained in various studies, differ from 200 to 1800 GBq (Masson et al., 2021). Estimates of daily emissions can differ even more (Fig. 1).

None of the above publications studied the impact of these factors on estimates of emissions. An additional difficulty in the quantification of emissions during the event of April 2020 was combination of emissions from wildfires and emissions from the raised radioactive dust, in the areas where wildfires previously occurred. Except the work of Talerko et al. (2021b), dust storm, as a separate mechanism for the release of radioactive ash, was not considered. For some days at the end of the wildfires (20 April, 23 April), there are also significant discrepancies in emission estimates, probably due to insufficient measurements, as the atmospheric transport was mainly in the eastern direction, where there are very few measurements.

2.3. Measurement data used in this study

The largest and most representative dataset of radiological measurements during wildfires in ChEZ in April 2020 was published recently by Masson et al. (2021). This dataset contains about 1000 measurements of ^{137}Cs concentrations in air collected in more than 20 countries, including Ukraine. The measured concentrations are characterized by large variability from about 0.02 to 9000 $\mu\text{Bq}\cdot\text{m}^{-3}$. However, outside ChEZ, the maximum concentration was much smaller (700 $\mu\text{Bq}\cdot\text{m}^{-3}$). The available ‘no detection’ measurements also varied largely by the lower detectable limit - from 0.1 to 200 $\mu\text{Bq}\cdot\text{m}^{-3}$. The measurement interval of measurements varied from 1 to 30 days. For most of the observations, measurement error was reported. The median value of the

relative measurement error calculated from all such measurements was 16%.

For source term estimation, only those measurements were used for which the start and end dates were within the simulation period (from 3 April to April 27, 2020). Measurements within ChEZ were excluded because simplified source parameterization described below does not allow for representation of time histories of concentrations close to the areas of wildfires. In general, 454 concentration measurements were used for the STE study in this work. It is important to evaluate background concentrations of ^{137}Cs at different stations for the solution of the inverse problem. For many stations, Masson et al. (2021) presented measurements collected before and after the wildfires that could be used for that purpose. In this work, if for a station there existed measurements unaffected by the wildfires (i.e. with the final date before April 3, 2020 or with the start date after April 27, 2020), the background concentration was derived from such measurements. For the other stations, the background values were based on the literature review. If there was no relevant data in the published literature, the minimum measured concentrations during the period of measurements were used as background values. Details of the definition of background values are provided in the Supplement.

Although estimation of locations and areas of the wildfires, as well as the first guess of emissions in this study, are largely based on the previous study (Talerko et al., 2021a), for some dates (21–23 April) parameters of the wildfires had to be additionally evaluated. As in the previous studies, satellite-based data provided by NASA (2020) was used for that purpose. The data in the Atlas of radioactive contamination of Ukraine (<http://radatlas.isgeo.com.ua/>) was used for evaluation of contamination of the territories where wildfires took place. More details on the estimation of parameters of the releases are provided in the sections below.

3. Setup of the atmospheric dispersion model

3.1. Basic setup of FLEXPART ADM

FLEXPART is a Lagrangian atmospheric dispersion model which solves stochastic equations of particles’ movement in the atmosphere (Pisso et al., 2019; Stohl et al., 2005). The average wind components are obtained from the numerical weather prediction models, while stochastic components of the wind are internally calculated by solving the Langevin equation. To do so, a range of micrometeorological parameters, such as variances of wind fluctuations and others are parameterized using state-of-art theories of the planetary boundary layer. Wet and dry depositions are calculated by using known parameterizations and taking into account size distribution of particles, precipitation intensity, land use categories, etc. The available physical parameterizations and technical details of the modeling methodology are described in the works Pisso et al. (2019); Stohl et al. (1998, 2005); Stohl and Thomson (1999), and in the documentation available on the official FLEXPART web site: <https://www.flexpart.eu>.

In this work, FLEXPART version 10.4 was used (Pisso et al., 2019). This model can use meteorological data provided by operational versions of the numerical weather prediction system of ECMWF and the Global Forecasting System (GFS) of the US National Centre of Environmental Prediction (NCEP). In this work, we used final analysis data of GFS with 0.5 deg. spatial resolution and 6 h time resolution, as they are available from the link <https://www.ncei.noaa.gov/data/global-forecast-system/access/grid-004-0.5-degree/analysis>. For investigation of the sensitivity of the obtained source term estimations to errors in the meteorological fields, we used analysis data of the Global Ensemble Forecasting System (GEFS) also operated by NCEP (<https://www.ncei.noaa.gov/products/weather-climate-models/global-ensemble-forecast>, selection « GEFS, 21-Member, 003 (1°) | 01Jan2008–Present»). This system integrates 21 forecast realizations per analysis time. Output meteorological fields calculated in each realization are provided with 1

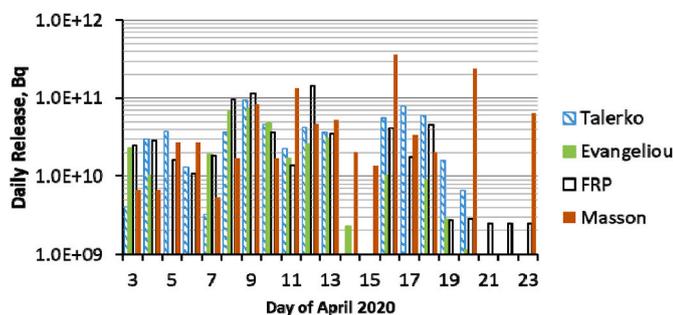


Fig. 1. Comparison of estimates of daily emissions of ^{137}Cs during wildfires in ChEZ from different works: ‘Masson’ – estimates by Masson et al. (2021), ‘Talerko’ – estimates by Talerko et al. (2021a), ‘Evangeliou’ – estimates by Evangeliou and Eckhardt (2020), averaged by Talerko et al. (2021a), ‘FRP’ – estimates in this work obtained by processing satellite data as described in section 3.2.

deg. spatial resolution and 6 h time resolution in GRIB2 files similar to the output data of GFS. However, additional conversion was necessary to enable FLEXPART to run on output files of GEFS. The converted files were created with the aid of the wgrib2 utility (<https://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/>). The meteorological fields read from GEFS files were written in the converted files in the same order as they were present in GFS files. At analysis times, GEFS did not provide precipitation intensities necessary for FLEXPART. Therefore, for each analysis time X, precipitation intensities were taken from a 6-hr forecast initialized at the previous analysis time, valid at X-6 hrs. The meteorological variables at pressure levels less than 10 mb were not available in GEFS files. Therefore, in the converted files, data from GFS final analysis files at pressure levels less than 10 mb was added.

The calculations were performed for the time period from April 3, 2020, 00 h to April 27, 2020, 00 h. The computational data was stored on the grid with 0.1 deg spatial resolution and 1-h time resolution. The spatial extent of the grid was defined to cover the measurements used in the STE procedure. The coordinates of the lower-left corner of the grid were: 0 deg. W, 37.8 deg. N and the number of nodes was 400 x 322. Physical options were defined to take into account effects of the subgrid variability of the planetary boundary layer (LSUBGRID = 1) and convection (LCONVECTION = 1). Parameters related to size distribution are described below.

3.2. Parameters of emissions from the wildfires

The first guess (prior) estimates of emission inventories of ^{137}Cs , together with coordinates and areas of emission sources, were taken in this work from Talerko et al. (2021a). Because estimates in Talerko et al. (2021a) are available only until 20 April, for estimation of emissions during 21–23 April, fire products from the Moderate Resolution Imaging Spectroradiometer (MODIS, https://firms.modaps.eosdis.nasa.gov/active_fire) were used. In particular, MODIS data contains fire radiative power (FRP, MWt) of each pixel. According to Talerko et al. (2021a), the emission inventory can be evaluated from the FRP data using formula

$$Q = C_r \cdot D \cdot FRE = C_r \cdot D \cdot \int_0^{\tau} FRP(t) dt. \quad (1)$$

Here $D[\text{Bq}\cdot\text{m}^{-2}]$ is the inventory of radionuclides in the regions

covered by wildfires, FRE is the total amount of energy released by wildfires during the considered time period, characterized by duration τ , and $C_r = 9.5 \cdot 10^{-4} \text{ m}^2 \cdot \text{MJ}^{-1}$ is coefficient evaluated by Talerko et al. (2021a). The emission inventories were evaluated from equation (1) for the whole time period of wildfires (Fig. 1, 'FRP'). As expected, for the period 3–20 April, the 'FRP' estimates are consistent with the estimates in Talerko et al. (2021a). The estimates obtained using formula (1) for the period 21–23 April were used as first guess estimates for the respective time period. Locations of centers of emission areas used in the simulation are shown in Fig. 2, while the details of the first guess estimation of the emission inventories, locations, and areas of the wildfires are provided in Supplement, Table S2.

Regarding parameters of size distribution of emitted aerosols, Evangelioiu and Eckhardt (2020) assumed that 20% of release consisted of particles with aerodynamic diameter $\mu_d = 0.25 \mu\text{m}$, 20% of release consisted of particles with $\mu_d = 8 \mu\text{m}$ and the rest were particles with $\mu_d = 16 \mu\text{m}$. The geometric standard deviation was set: $\sigma_g = 1.1$ and the same value was used in this study. In a laboratory study by Hao et al. (2018), it was estimated that from 0 to 12% of release was transported on particles with an aerodynamic diameter less than $2.5 \mu\text{m}$, while the rest was on particles larger than $10 \mu\text{m}$. As in the mentioned data, in our work, the release consisted of particles of 3 categories, with the values of μ_d : 0.25, 8, and $16 \mu\text{m}$. For the particles with $\mu_d = 0.25 \mu\text{m}$, two alternative values for their relative fraction were used: $W_{sf(0.25)} \in \{0.1, 0.2\}$, 10% as in Hao et al. (2018) or 20% as in Evangelioiu and Eckhardt (2020). For particles with $\mu_d = 8 \mu\text{m}$, two alternative values were also used for their relative fraction: $W_{sf(8)} \in \{0., 0.2\}$, 0% as in Hao et al. (2018) or 20% as in Evangelioiu and Eckhardt (2020). For each combination of $W_{sf(0.25)}$ and $W_{sf(8)}$, the rest was assumed to be particles with aerodynamic diameter $\mu_d = 16 \mu\text{m}$. Therefore, the set of used values for fractions of particles with different sizes could be formalized by the following formula:

$$\begin{aligned} W_{sf(0.25)} &\in \{0.1, 0.2\}, W_{sf(8)} \in \{0., 0.2\} \\ W_{sf(16)} &= 1 - (W_{sf(0.25)} + W_{sf(8)}). \end{aligned} \quad (2)$$

For setting parameters of height distribution, we used data from the Global Fire Assimilation System GFAS (Rémy et al., 2017; Di Giuseppe et al., 2018) available from Copernicus Atmosphere Monitoring Service (CAMS) Products (<https://apps.ecmwf.int/datasets/data/cams-gfas/>). The bottom and top heights of convective plumes formed by the

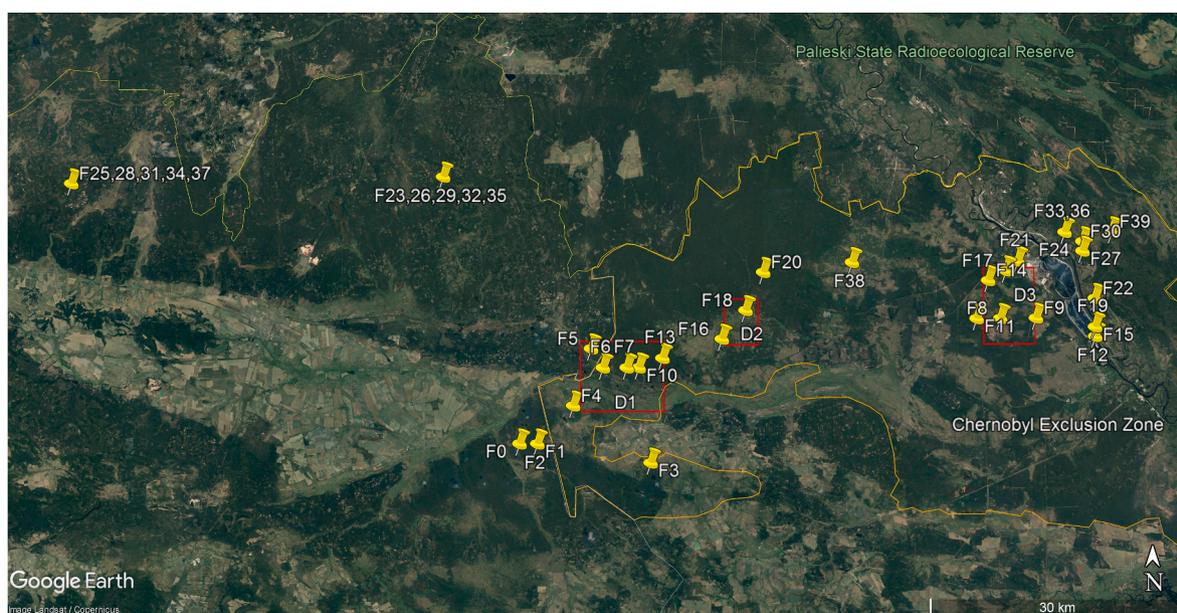


Fig. 2. Locations of sources F0-39 represent wildfires in the simulation study; sources D1-3 represent emissions due to wind resuspension during dust storms, they are shown together with their lateral boundaries (© Google Earth).

wildfires were estimated by GFAS and available at 0.1 deg. spatial resolution and 24 h time resolution (GFAS, 2022). Therefore, for each of the wildfire sources considered in this work (Supplement, Table S2), the top and upper heights of convective plumes z_{top} , z_{bot} were set equal to the corresponding values in the nearest node of GFAS (Fig. 3). In calculations, every wildfire source was split vertically into two parts: 1) bottom part, $0 < z \leq z_{bot}$ and 2) upper part, $z_{bot} < z \leq z_{top}$. The fraction of release in the bottom part of the source W_{h1} was estimated from 0.1 to 0.5 according to data presented by Sofiev et al. (2013). Hence in this work, we used 3 alternative values: $W_{h1} \in \{0.1, 0.25, 0.5\}$. The rest of the release was assumed from the upper part of the source: $W_{h2} = 1 - W_{h1}$.

3.3. Parameters of emissions from the dust storm

Talerko et al. (2021b) presented estimates of emissions of radioactive ash from the territories of the recent wildfires during the dust storm. The flux of radionuclides due to wind resuspension Q_d [$\text{Bq}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$] was parameterized by the formula $Q_d = q_s \cdot F_d$, where q_s [$\text{Bq}\cdot\text{kg}^{-1}$] was the specific activity of radioactive ash and F_d [$\text{kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$] was the mass flux of dust proportional to cubic power of friction velocity according to parameterization by Vautard et al. (2005). The estimate for specific activity q_s of radioactive ash was obtained by comparing formula (1) for radionuclides flux with known parameterization of mass flux F_w during wildfires (Ichoku and Kaufman, 2005):

$$F_w = C_e \cdot D \cdot F_{RE}. \quad (3)$$

By dividing Eq. (3) by (1) and using estimate of $C_e = 0.075 \text{ kg}\cdot\text{MJ}^{-1}$ from Ichoku and Kaufman (2005), an estimate for specific activity was obtained:

$$q_s = (C_r / C_e) D \approx 0.0137D. \quad (4)$$

Then flux of radionuclides due to wind resuspension during the dust storm was determined. The total estimated released inventory of ^{137}Cs during 16–17 April was 162 GBq, and about 90% of this activity was emitted on 16 April, when wind speed was the greatest.

In this work, first guess estimation of emissions created by dust storm was obtained by processing data presented by Talerko et al. (2021b). Twenty emission sources considered by Talerko et al. (2021b) were aggregated in this work into 3 sources (Fig. 2), emitting in total the same amount of radioactivity (162 GBq), of which 90% was emitted on 16 April (Supplement, Table S3).

Parameters of the size distribution of particles emitted due to wind resuspension were discussed by Talerko et al. (2021b), where it was assumed that the fraction of particles with $\mu_d = 1 \mu\text{m}$, $W_{sd(1)} = 0.4$, while

the rest of the radioactivity was transported by particles with $\mu_d = 10 \mu\text{m}$. However, based on measurement data published in the same work, an alternative value of $W_{sd(1)}$ could be proposed: $W_{sd(1)} = 0.1$. Finally, Wagenbrenner et al. (2013) studied dust emissions from the territories previously covered by the wildfires, and according to their data, the value of $W_{sd(1)} = 0.6$ could also be justified. Therefore, all three abovementioned values were used in this work as alternative options: $W_{sd(1)} \in \{0.1, 0.4, 0.6\}$. Similar to Talerko et al. (2021b), the rest of the radioactivity was assumed to come from particles with $\mu_d = 10 \mu\text{m}$: $W_{sd(10)} = 1 - W_{sd(1)}$.

4. Source term estimation method

4.1. STE as a minimization problem

STE can be considered as the data assimilation problem (Daley, 1991), in which minimization of the following cost function is performed (Enting, 2002):

$$J(\underline{q}) = \left(\underline{G}\underline{q} - \underline{y} \right)^T \left(\underline{F}^{-1} \right) \left(\underline{G}\underline{q} - \underline{y} \right) + \left(\underline{q} - \underline{q}_B \right)^T \left(\underline{B}^{-1} \right) \left(\underline{q} - \underline{q}_B \right). \quad (5)$$

Here, $\underline{y} \in R^{N_o}$ is the vector of measurements, N_o is the number of measurements, $\underline{q} \in R^{N_s}$ is the vector of unknown emission inventories from all N_s sources (Tables S2 and S3 in Supplement), $\underline{q}_B \in R^{N_s}$ is the vector of the corresponding first guess (prior) estimations of emission inventories, and $\underline{B} \in R^{N_s \times N_s}$ is the covariance matrix of errors in the first guess estimation. Source receptor matrix (SRM) $\underline{G} \in R^{N_o \times N_s}$ is calculated by the model, establishing connection between emission inventories and calculated concentrations in the points and at the times of measurements. Matrix $\underline{F} = \underline{M} + \underline{O}$, \underline{F} , \underline{M} , $\underline{O} \in R^{N_o \times N_o}$ is the sum of covariance matrices of model and observational errors \underline{M} and \underline{O} respectively.

As follows from Bayes' theorem (Enting et al., 2002), solution that minimizes cost function (5) corresponds to the maximum in posterior probability distribution function of the unknown vector \underline{q} , provided that probability distributions of errors of prior estimation \underline{q}_B , observations, and model errors are all Gaussian. In the context of this study, it is especially important to note that the covariance matrix of model errors \underline{M} represents errors, provided that vector \underline{q} is known exactly. Therefore, it represents errors introduced by different input parameters of the STE

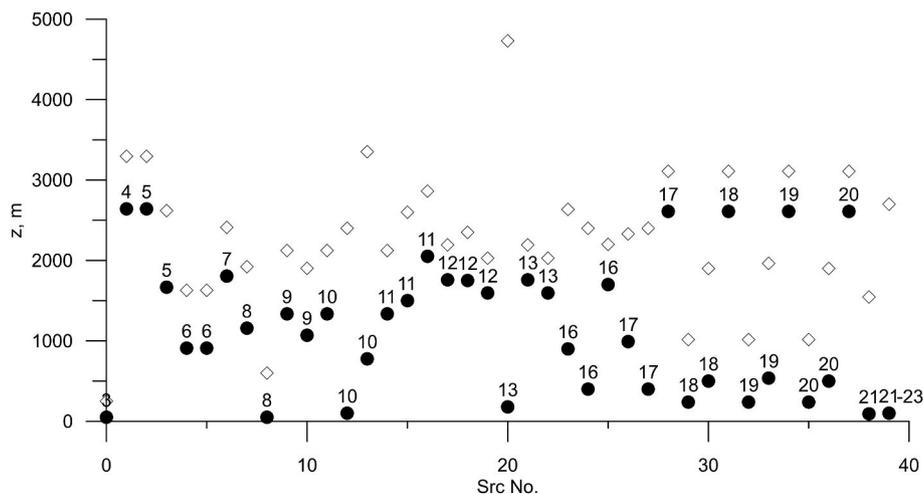


Fig. 3. Heights of the bottom, z_{bot} (black circles), and upper z_{top} (rhombus), boundaries of convective plumes for each of the wildfire sources considered in the study and derived from CAMS GFAS data. Data labels indicate days of April 2020, at which the corresponding sources were active.

problem, such as meteorological variables etc, but not of the unknown vector \underline{q} to be estimated by solving an inverse problem. Since in reality vector \underline{q} is not known, matrix \underline{M} is usually somehow parameterized. In this work, we propose to use an ensemble approach for calculation of the matrix \underline{M} (Evensen, 2009):

$$\begin{aligned} \underline{M}(\underline{q}) &= \frac{1}{N_e} \sum_{m=1}^{N_e} (\underline{c}_m - \langle \underline{c} \rangle) (\underline{c}_m - \langle \underline{c} \rangle)^T \\ \underline{c}_m &= \underline{G}_m \underline{q}, \quad 1 \leq m \leq N_e \\ \langle \underline{c} \rangle &= \frac{1}{N_e} \sum_{m=1}^{N_e} \underline{c}_m \end{aligned} \quad (6)$$

Here, column vectors \underline{c}_m are calculated with different model realizations obtained using different meteorological input datasets and different source term parameters, and N_e is the number of realizations (i. e. the size of the ensemble). Different model realizations are represented by the corresponding SRM \underline{G}_m , and matrix \underline{M} is calculated by ensemble averaging. As it is readily seen from equation (6), matrix \underline{M} and hence matrix \underline{F} , depend on the unknown vector \underline{q} : $\underline{F} = \underline{F}(\underline{q})$ and thus minimization problem (5) becomes a nonlinear regression:

$$J_m(\underline{q}_m) = \left(\underline{G}_m \underline{q}_m - \underline{y} \right)^T \left(\underline{F}(\underline{q}_m)^{-1} \right) \left(\underline{G}_m \underline{q}_m - \underline{y} \right) + \left(\underline{q}_m - \underline{q}_B \right)^T \left(\underline{B}^{-1} \right) \left(\underline{q}_m - \underline{q}_B \right) \xrightarrow{\underline{q}_m} \min \quad (7)$$

$1 \leq m \leq N_e$

Here, index m denotes that the minimization problem is solved separately for each SRM \underline{G}_m corresponding to a specific m -th model realization. In this work, we complemented minimization problem (7) with conditions of positive solution and zero systematic error. These conditions and the second regularization term in cost function (7) helped avoid overfitting:

$$\begin{aligned} \sum_{i=1}^{N_o} c_{m,i} &= \sum_{i=1}^{N_o} \sum_{j=1}^{N_s} g_{m,ij} q_{m,j} = \sum_{i=1}^{N_o} y_i \\ \underline{q}_m &> 0 \end{aligned} \quad (8)$$

$$J_m = \left(\underline{G}_m \underline{q}_m^{s+1} - \underline{y} \right)^T \left(\underline{F}(\underline{q}_m^s)^{-1} \right) \left(\underline{G}_m \underline{q}_m^{s+1} - \underline{y} \right) + \left(\underline{q}_m^{s+1} - \underline{q}_B \right)^T \left(\underline{B}^{-1} \right) \left(\underline{q}_m^{s+1} - \underline{q}_B \right) \xrightarrow{\underline{q}_m^{s+1}} \min \quad (9)$$

$1 \leq m \leq N_e, \quad 0 \leq s \leq N_{iter}$

Here $g_{m,ij}$ are elements of the matrix \underline{G}_m .

After solving the nonlinear regression problems (7),(8) for every index m , the ensemble of the emission inventories was obtained. Confidence intervals of the estimated emission inventories were assessed by processing the obtained ensemble of estimates. For every source, the 2.5-th and 97.5-th percentiles of the values in the vector of the obtained emission estimates were calculated, yielding respectively lower and upper confidence limits of the respective emissions.

4.2. Covariance matrices

As it was mentioned above, the covariance matrix of model errors is calculated by processing ensemble of model runs using equation (6). This ensemble consists of model runs performed using data of each of the 21 members of the GEFS meteorological system (section 3.1) and using the final analysis dataset of the GFS system (i.e. in total $N_M = 22$ meteorological datasets). Runs with each meteorological dataset are combined with each of the combinations of parameters $W_{sf(0.25)}$, $W_{sf(8)}$, $W_{h(1)}$, $W_{sd(1)}$, discussed in sections 3.2-3.3. In total, there are $N_p = 2 \times 2 \times 3 \times 3 = 36$ combinations of those parameters and, therefore, the number of ensemble members $N_e = N_M \times N_p = 22 \times 36 = 792$.

The covariance matrix of measurement errors \underline{Q} is set to be diagonal, and its diagonal elements (variances) are equal to a sum of squared instrumental error and background value θ for the corresponding station: $\sigma_{oi}^2 = \sigma_{di}^2 + \theta_i^2$. In the cases where instrumental error was not reported, it was set to 16% of the corresponding measurement value (here the median value of all reported relative measurement errors mentioned in section 2.3 was used). In cases when the measurement value was less than the lower detectable limit (LDL), the instrumental error σ_d was set to $\sigma_d = LDL$, while the measurement value was set to the background value at the corresponding station.

The covariance matrix of the errors in the first guess approximation \underline{B} was also set to be diagonal. In operational air quality prediction systems that are used for solving STE problems, it is convenient to set corresponding dispersion parameters proportional to the values of first guess emission rates with the coefficient of proportionality ~ 1 (Winiarek et al., 2011; Kovalets et al., 2019). Therefore, in this study, the diagonal elements of matrix \underline{B} were set accordingly: $\sigma_{Bi}^2 = q_{Bi}^2$.

4.3. Minimization procedure

The nonlinear regression problem (7)-(8) was solved in this work by iteration procedure:

where s was the iteration number and the total number of iterations was N_{iter} . At each iteration step s , the covariance matrix \underline{F} was calculated using solution obtained at the previous step, while at the zeroth iteration step \underline{q}_m^0 it was set to $\underline{q}_m^0 = \underline{q}_B$, and thus:

$$\underline{F}(\underline{q}_m^0) = \underline{M}^0 + \underline{Q} \quad (10)$$

Here, \underline{M}^0 is defined as $\underline{M}^0 = \underline{M}(\underline{q}_B)$.

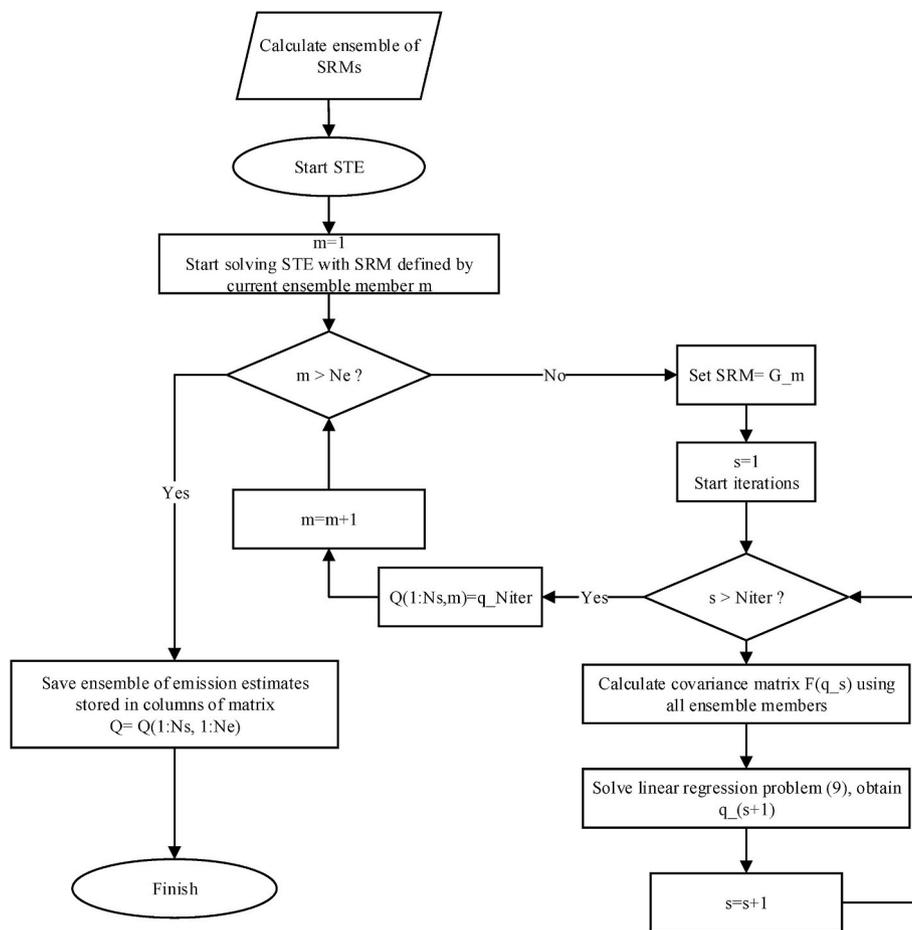


Fig. 4. Flowchart of the ensemble iterative source inversion method (EISIM).

Then, at each iteration, linear regression problem (9) with restrictions in the form of relationships (8) was solved for the unknown \underline{q}_m^{s+1} . The method of solving linear regression problem is described in Supplement, section S3. The flowchart of the described source inversion method is shown in Fig. 4. Preliminary testing of the algorithm had shown that usually 50 iterations were sufficient for convergence (Fig. 5a). Therefore, in the next applications, the number of iterations was set to $N_{iter} = 50$. The resulting solution obtained after the end of the iteration cycle was further denoted as $\underline{q}_m = \underline{q}_m^{Niter}$.

As it follows from Fig. 5b, the norm of the main diagonal of the model error covariance matrix \underline{M} estimated by ensemble averaging using equation (6) not only converges to a stationary value, but also decreases by a factor of 10. Therefore, the algorithm can find a solution of the STE problem that reduces the variance of differences between simulated results induced by variable meteorological fields and parameters of the source term.

To our knowledge, such an iterative STE method based on the processing of the ensemble of runs has not been used previously. In the studies Zhang et al. (2014), Zhang et al. (2015), covariance matrix of model errors was also defined by iterations to improve the overall convergence of the ensemble Kalman filtering method (EnKF). Despite the usage of an ensemble of model realizations, our method cannot be considered EnKF, because time is not explicitly present in its formulation. Therefore, we simply denote it as Ensemble Iterative Source Inversion Method (EISIM).

4.4. Computational details

Parameter estimation problems in which the uncertainty of the

model is included in the methodology are usually time-consuming. In our approach, the main computational efforts were spent on calculations of the elements of source receptor matrices by running the FLEXPART model. The total number of FLEXPART runs for calculation of SRMs related to wildfires was $N_F = 3 \times N_S \times N_M = 3 \times 40 \times 22 = 2640$. Each FLEXPART run was performed for one of the 40-sources described in section 3.1 (split vertically into two sub-sources), for one of three size fractions of emitted aerosols (0.25, 8, or 16 μm), and one of the 22 meteorological ensemble members. Simulations were performed with reference ('unit') release rate and thus calculations of elements of SRMs for particular values of size fractions of emitted aerosols and fractions of emissions from different heights could be easily recalculated. The number of particles used to represent a single source, lasting 24 h, was set to $2 \cdot 10^5$. Calculations were performed on a personal computer with 72 computing cores Intel(R) Xeon(R) Gold 6240 CPU @ 2.60 GHz and on a 48-computing-core Intel Xeon Processor (Cascadelake) CPU @ 2.095 GHz of the Cloud Computing Platform of the Ukrainian National Grid Infrastructure. The parallelization of computations was implemented by distributing separate FLEXPART runs on different computational cores. The time needed for a single FLEXPART run varied from 3 to 25 h depending on the start time of the source and the size of the particles (for coarser particles calculations were faster due to rapid deposition of particles). The total time needed to execute all necessary FLEXPART runs related to wildfires on the computational infrastructure described above was 14 days. The time needed to execute FLEXPART runs representing dust storm required 1 additional day because emissions created by dust storm were represented by only two additional sources. The FLEXPART output files were processed before performing the minimization procedure. Processing included extraction of output concentrations, their spatial interpolation to the points of measurements, time

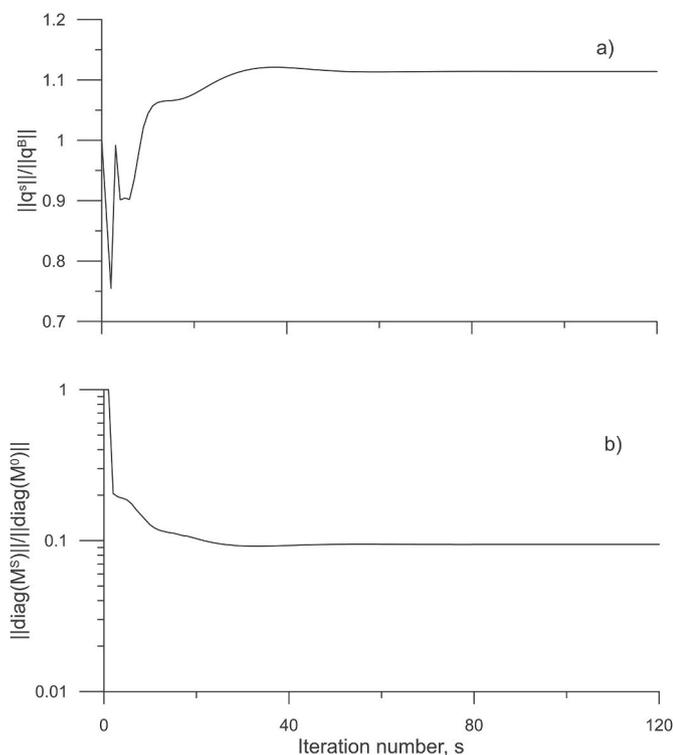


Fig. 5. Example of the iteration number dependence of (a) non-dimensional norm of the solution $\|q_m^s\|/\|q_B\|$ and (b) non-dimensional norm of the main diagonal of the model-error covariance matrix $\|diag(\underline{M})\|/\|diag(\underline{M}^0)\|$; the example is shown for model realization $m = 1$.

integration following measurements time intervals, and saving the resulting arrays in intermediate files. The processed FLEXPART data was then used in minimization procedure (9) that was performed on a Windows machine with processor Intel (R) Core (TM) i5-8400 CPU @ 2.8 GHz, 6 Cores. The minimization problem (9) for a single ensemble member was solved within about 5 min. Performing runs in parallel using 6 computational cores, it took about 12 h to solve the minimization problem (9) for all 792 ensemble members.

From the above it is clear that the dependence of computational

complexity of the algorithm on the number of parameters depends first of all on the number of additional FLEXPART runs. For example, adding a new size fraction increases the computing time by a factor of 4/3, i.e. by 33%. However, preserving the number of size fractions, the number of the configurations of their relative weights $W_{sf(0.25)}$, $W_{sf(8)}$, $W_{sf(16)}$ can be increased substantially without a significant increase in computational time of the algorithm, because it will not require additional FLEXPART runs, only an increase in the number of cycles in minimization procedure (9).

5. Results of calculations

Ensemble of solutions of STE problem for the case of wildfires in ChEZ was obtained after application of procedure (9). The total ensemble-average daily emission inventories of ^{137}Cs during the period of wildfires are shown in Fig. 6. Emissions for 21–23 April presented in Fig. 6 were estimated for the 3-day period and then daily emissions were recalculated assuming a constant emission rate. As it can be noted from Fig. 6, almost each day, the first guess emissions from wildfires (Fig. 6a) fall within the 95% confidence intervals of the obtained solutions. In contrast, estimate of the first guess emission due to wind resuspension on the first day of the dust storm (16 April) does not fall in the confidence interval of the respective estimated emissions. The average value of the estimated emission due to wind resuspension on 16 April is 23 GBq, which is about 6 times less than the respective first guess estimation.

Evaluation of individual solutions from the ensemble can be performed using independent metrics, such as normalized mean squared errors (NMSEs) and correlation coefficients r of the respective simulated concentrations of ^{137}Cs , by comparing to the measurements:

$$NMSE_m = \frac{(1/N_o) \sum_{i=1}^{N_o} (c_{m,i} - y_i)^2}{\langle c_m \rangle \langle y \rangle}$$

$$r_m = \frac{\sum_{i=1}^{N_o} (c_{m,i} - \langle c_m \rangle)(y_i - \langle y \rangle)}{\sqrt{\sum_{i=1}^{N_o} (c_{m,i} - \langle c_m \rangle)^2} \sqrt{\sum_{i=1}^{N_o} (y_i - \langle y \rangle)^2}} \quad (11)$$

$$\langle c_m \rangle = \frac{1}{N_o} \sum_{i=1}^{N_o} c_{m,i}; \quad \langle y \rangle = \frac{1}{N_o} \sum_{i=1}^{N_o} y_i$$

Here, as in equation (9), index m is used to denote model realization for which evaluation of the error statistics is performed, y represents

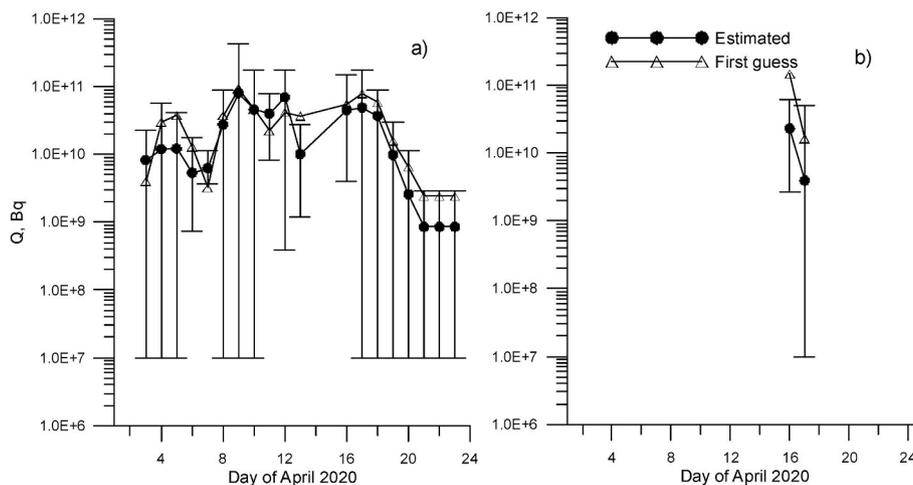


Fig. 6. Daily emission inventories of ^{137}Cs during the period of wildfires, obtained by averaging the ensemble of solutions (black circles) together with confidence intervals; first guess estimations for the same dates are shown by triangles; a) emissions resulting from wildfires; b) emissions created by wind resuspension during the dust storm on 16–17 April 2020.

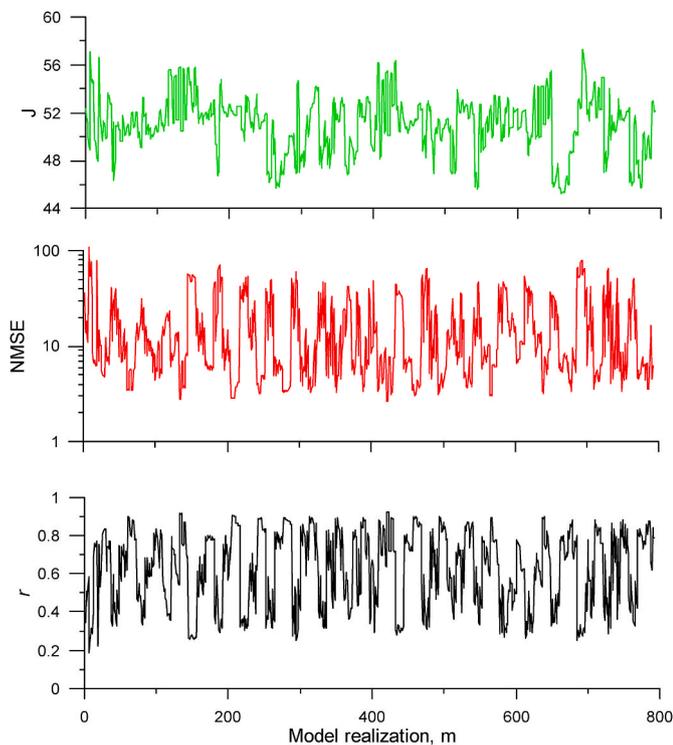


Fig. 7. Optimized cost functions J defined in (7), correlation coefficients r , and NMSEs (11) of model results as compared to measurements of ^{137}Cs concentrations; statistical indicators obtained in simulations with the emission rates resulting from the solution of STE problem for the respective model realization m .

observations, c represents respective concentration values calculated by model and the averaging is performed over the total number of measurements N_o .

Fig. 7 presents optimized cost functions J_m (7) obtained when solving STE problem (9) for the respective ensemble member m , together with the NMSE_m and r_m obtained by using the corresponding solution q_m . The statistical dependence of metrics NMSE and J is weak and their mutual correlation coefficient equals 0.2. The magnitude of the mutual correlation of r and J is even smaller and equals -0.08 . From the results presented in Fig. 7, it can also be noted that the degree of variation of J is

much smaller than that for NMSE and r . While J varies from 46 to 58, i.e. by about 20% from the average value, the NMSE values span more than an order of magnitude (from $\text{NMSE} \approx 2$ to $\text{NMSE} \approx 100$), and the values of r change from $r \approx 0.2$ to $r \approx 0.9$, i.e. by a factor of 4.5. It can be concluded that NMSE and r can be used for independent evaluation of the obtained solutions even though they are obtained by comparing model results to the same measurements that are used for solving the STE problem.

The median values of NMSE and r presented in Fig. 7 are $\text{NMSE} = 9.8$ and $r = 0.67$, which are quite good taking into account significant uncertainties in release estimation and spatial scale of the considered dispersion problem (~ 1000 km). In particular, the abovementioned value of NMSE is within the reported range of errors of different atmospheric models applied for the conditions of the European Tracer Experiment ETEX with an exactly known source term (Bellasio et al., 2012).

For a sensitivity test, we selected from the full ensemble a subset of the highest quality solutions that satisfy the following criteria: $\text{NMSE} \leq 10$, $r \geq 0.6$. This subset consisted of $N_{e2} = 406$ members. The ensemble-average emissions obtained by averaging over the selected subset (Estimate 2) were then compared against the emissions averaged over the full ensemble (Estimate 1) in Fig. 8. As it can be seen from Fig. 8, emissions from wildfires in Estimate 2 were very close to emissions in Estimate 1, confirming robustness of the obtained estimates. Emissions due to wind resuspension for 16 April in Estimate 2 were also close to the respective values in Estimate 1. For 17 April, emissions in Estimate 2 were reduced by a factor of about 2 compared to Estimate 1. This sensitivity was the result of two main factors: 1) emissions due to wind resuspension were largely reduced on 17 April due to a decline in wind speed and became poorly distinguishable on the background of the emissions from the wildfires; 2) relatively small number of measurements were affected by the dust storm because of the prevailing eastward direction of the wind.

The total emissions according to Estimates 1 and 2 and their

Table 1

Total inventories of emissions of ^{137}Cs from wildfires (3–24 April 2020) and from wind resuspension during the dust storm (16–17 April 2020); respective confidence limits are in brackets.

Source	First guess, GBq	Estimate 1, GBq	Estimate 2, GBq
Wildfires	586	461 (19 ≤ Q ≤ 1590)	448 (39 ≤ Q ≤ 1530)
Dust Storm	162	27 (3 ≤ Q ≤ 112)	21 (3 ≤ Q ≤ 93)

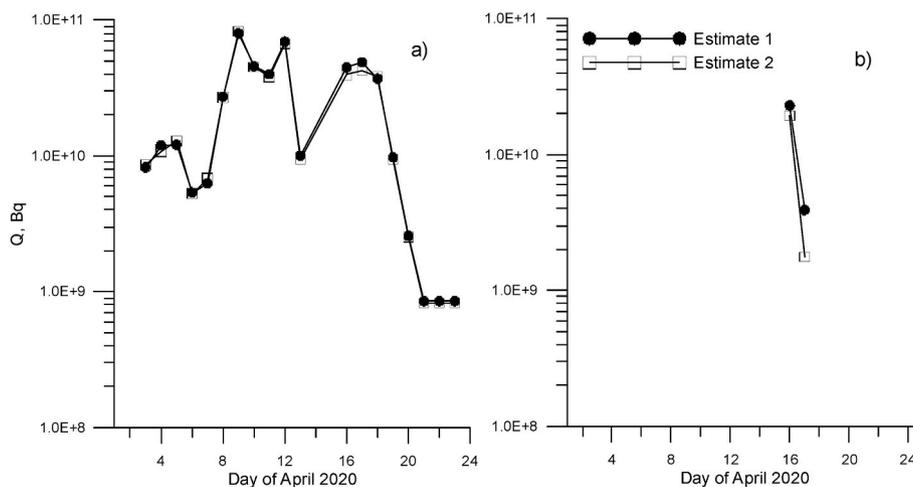


Fig. 8. Daily emission inventories of ^{137}Cs during the period of wildfires, obtained by averaging the full ensemble of solutions (black circles, Estimate 1) and obtained by averaging those estimates for which the resulting model errors satisfied the following criteria: $\text{NMSE} \leq 10$, $r \geq 0.6$ (open squares, Estimate 2); a) emissions formed by wildfires; b) emissions created by wind resuspension during the dust storm on 16–17 April 2020.

Table 2

Average variations of total emissions δQ_φ caused by changing the input meteorological data ('Meteo'), parameters of height distribution of the release ('Wh'), parameters describing the size distribution of particles emitted by wildfires ('Wsf'), and parameters describing the size distribution of particles resulting from wind resuspension ('Wsd'); parameter $\lambda_\varphi = \delta Q_\varphi / \Sigma \delta Q_\varphi$ is the relative value of δQ_φ summed over all variations.

Parameter	Meteo	Wh	Wsf	Wsd
δQ_φ , GBq	404	216	134	74.7
$\delta Q_\varphi / \Sigma \delta Q_\varphi$	0.49	0.26	0.16	0.09

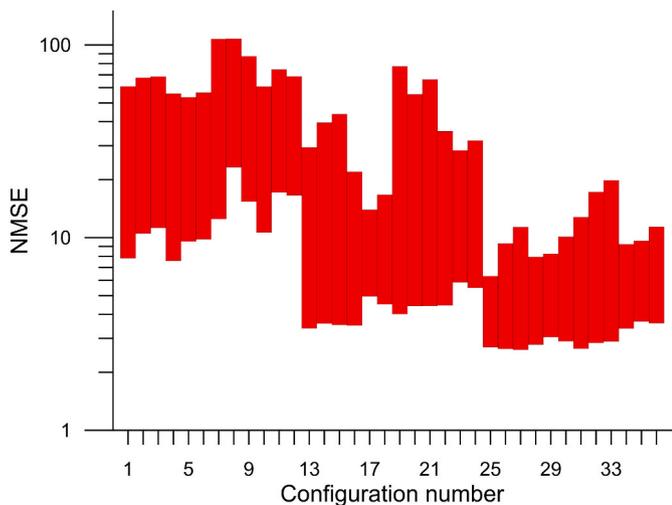


Fig. 9. The range of NMSEs obtained with the emission rates resulting from solving the STE problem, depending on the configuration of source term parameters; for each configuration, the range of NMSEs is created by different meteorological input data sets.

respective confidence intervals are provided in Table 1. The difference between the total emissions and the respective confidence intervals in Estimates 1 and 2 is small, confirming robustness of the obtained results.

In order to evaluate the impact of changes in different input parameters upon variations of estimated emission inventories, we calculated average variations of total emissions δQ_φ caused by changes of the respective group of parameters φ with other parameters fixed. Namely, the values of δQ_{wh} , δQ_{wsf} , δQ_{wsd} , and δQ_{Meteo} were evaluated. Here, the value δQ_{wh} accounts for the impact of parameters responsible for the height distribution of the release ($W_{h(1)}$, $W_{h(2)}$); δQ_{wsf} characterizes the impact of the size distribution of particles emitted from wildfires (parameters $W_{sf(0.25)}$, $W_{sf(8)}$, $W_{sf(16)}$); δQ_{wsd} characterizes the impact of the size distribution of particles emitted by wind resuspension (parameters $W_{sd(1)}$, $W_{sd(10)}$); δQ_{Meteo} accounts for the impact of the meteorological input data. As is seen from the data presented in Table 2, changes in meteorological input data have the most significant impact on the variability of the total emission inventory. The corresponding relative value of emission variability $\lambda_{Meteo} = \delta Q_{Meteo} / \Sigma \delta Q_\varphi = 0.49$. Other parameters cause significantly smaller variations in estimated emissions: $\lambda_{wh} = 0.26$ for parameters describing the height distribution of the release, $\lambda_{wsf} = 0.16$ for parameters related to the size distribution of particles emitted from the wildfires, and $\lambda_{wsd} = 0.09$ for parameters related to the size distribution of particles emitted by wind resuspension.

As was mentioned in section 4.2, for each of the 22 input meteorological datasets, 36 respective configurations of source term parameters, describing the size distribution of emitted particles and height distributions of wildfire emissions were applied. The obtained values of NMSE may be used to identify the most suitable parameter configurations. To do that, we present in Fig. 9 the range of obtained NMSEs for each of the source term parameter configurations. The variability of NMSE for a

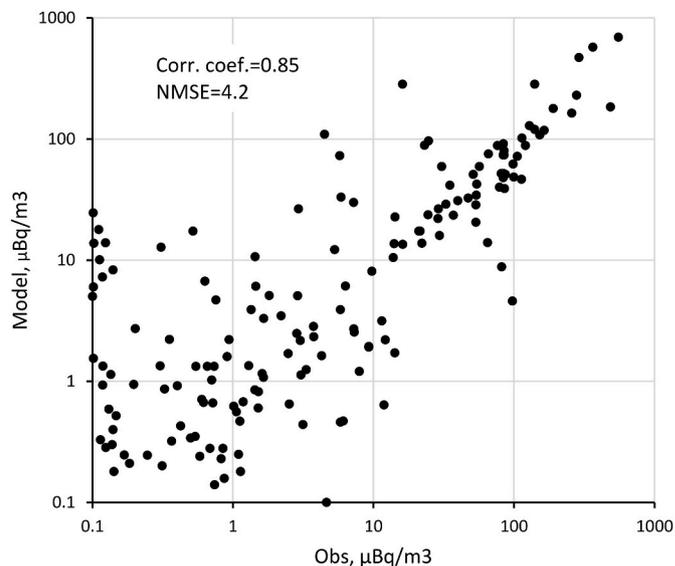


Fig. 10. Calculated concentrations in Ensemble run vs. observations; values less than $0.1 \mu\text{Bq}\cdot\text{m}^{-3}$ are clipped.

given configuration is created by different meteorological input data. The values of parameters for each configuration are provided in Supplement (Table S4). As can be noted from Fig. 9, the narrowest NMSE bands, together with the smallest minimum values of the normalized error, are found for the parameter configurations from 25 to 30. This range of configurations corresponds to the following fixed values of source term parameters: $W_{h(1)} = 0.5$, $W_{sf(0.25)} = 0.1$, varying other parameters according to the description in section 3.3. Therefore, it can be concluded that parameters $W_{h(1)}$, and $W_{sf(0.25)}$ have dominating influence on the solution of the STE problem, and the abovementioned values $W_{h(1)} = 0.5$, $W_{sf(0.25)} = 0.1$ can be considered most reliable estimates for those parameters out of the proposed values.

Reconstruction of the space-time distributions of near-surface concentrations during wildfires can be provided by ensemble averaging the calculated concentration fields in different ensemble realizations:

$$c^e(x, y, t) = \frac{1}{N_{e2}} \sum_{m=1}^{N_{e2}} c_m(x, y, t) \quad (12)$$

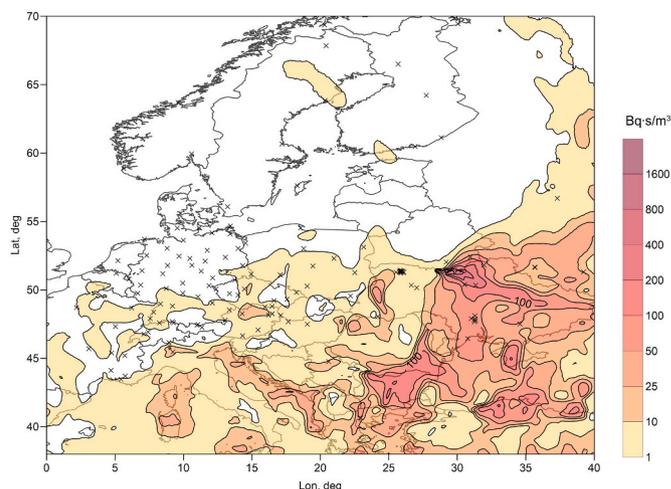


Fig. 11. Time integrated concentration of ^{137}Cs calculated in Ensemble run for the whole simulation period (April 03, 2020, 00 UTC - April 27, 2020, 00 UTC); locations of measurement stations used for solving STE problem are shown by symbols; isolines are shown for values 1, 10, 25, 50, 100, 200, 400, 800, 1600 $\text{Bq}\cdot\text{s}\cdot\text{m}^{-3}$.

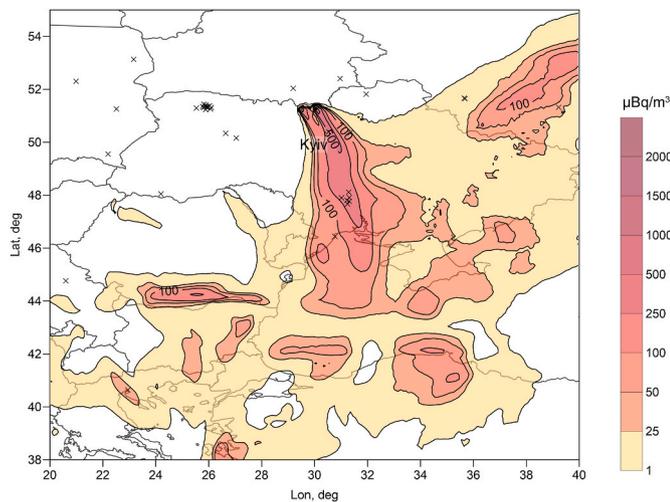


Fig. 12. The simulated distribution of ground-level concentration of ^{137}Cs averaged for 24-hr period (April 10, 2020, 06 UTC - April 11, 2020, 06 UTC) obtained in the Ensemble run; locations of measurement stations used for solving the STE problem are shown by symbols; isolines are shown for values 1, 25, 50, 100, 250, 500, 1500, 2500 $\mu\text{Bq}\cdot\text{m}^{-3}$.

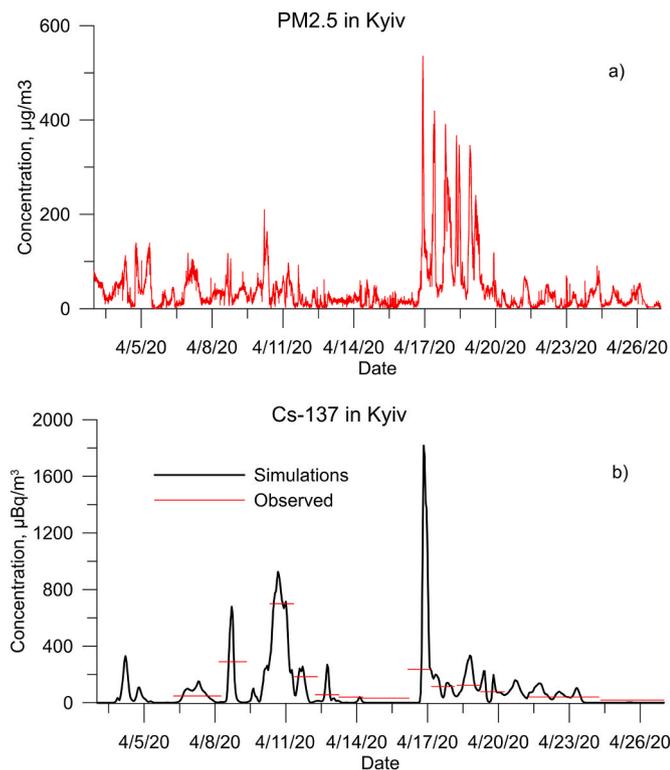


Fig. 13. a) Time series of PM2.5 concentrations in Kyiv according to measurement data from www.saveecobot.com; b) time series of observed (Masson et al., 2021) and simulated in Ensemble run concentrations of ^{137}Cs in Kyiv.

Here, only those ensemble members were selected that provided the highest quality solutions as described above ($\text{NMSE} \leq 10$, $r \geq 0.6$). The concentration field $c^e(x, y, t)$ calculated in this way will be henceforth denoted as ‘Ensemble run’ or ‘Ensemble estimation’. The scatter plot of the calculated concentrations in the Ensemble run vs. observed values is provided in Fig. 10. The resulting error indicators of the concentrations evaluated by formula (12) are $\text{NMSE} = 4.2$ for normalized error and $r = 0.85$ for the correlation coefficient. From Fig. 10 it can be noted that the scatter is reduced as the observed amounts get larger. This is explained

by the fact that as the observed concentrations become closer to the background, the relative impact of the respective background concentrations on elements of the observation error covariance matrix \underline{Q} becomes comparable to or greater than the impact of instrumental errors as described in section 4.2. Hence, larger relative deviations are allowed by the minimization algorithm for smaller concentrations.

Fig. 11 presents time-integrated concentrations of ^{137}Cs obtained in the Ensemble run for the whole simulation period. As it can be seen from this figure, the dominant direction of atmospheric transport was to the east, while the majority of measurements were in the western and central parts of Europe. The lack of measurements in the dominant direction of plume dispersion is one of the difficulties in solving the STE problem for this case.

Fig. 12 presents concentrations of ^{137}Cs simulated in the Ensemble run for 10 April when the maximum concentration was observed in Kyiv during the period of wildfires ($700 \mu\text{Bq}\cdot\text{m}^{-3}$). The time series of observed and simulated ^{137}Cs concentrations in Kyiv are shown in Fig. 13. The simulated concentration for that day ($547 \mu\text{Bq}\cdot\text{m}^{-3}$) is close to the observed. Fig. 13 also shows observed concentrations of particulate matter PM2.5 in Kyiv from April 3 to April 27, 2020. As is seen from the figure, on 10 April increased levels of PM2.5 were also observed in Kyiv, up to $200 \mu\text{g}\cdot\text{m}^{-3}$. The coincidence of peaks in ^{137}Cs and PM2.5 concentrations indicate that the particulate matter observed on 10 April also originated from the wildfires in ChEZ.

As is seen from Fig. 13, especially high concentrations of PM2.5 were observed in Kyiv during the dust storm that started on April 16, 2020. The start of the rapid increase of ^{137}Cs concentration in Kyiv was simulated in the Ensemble run practically at the same time (16 April,

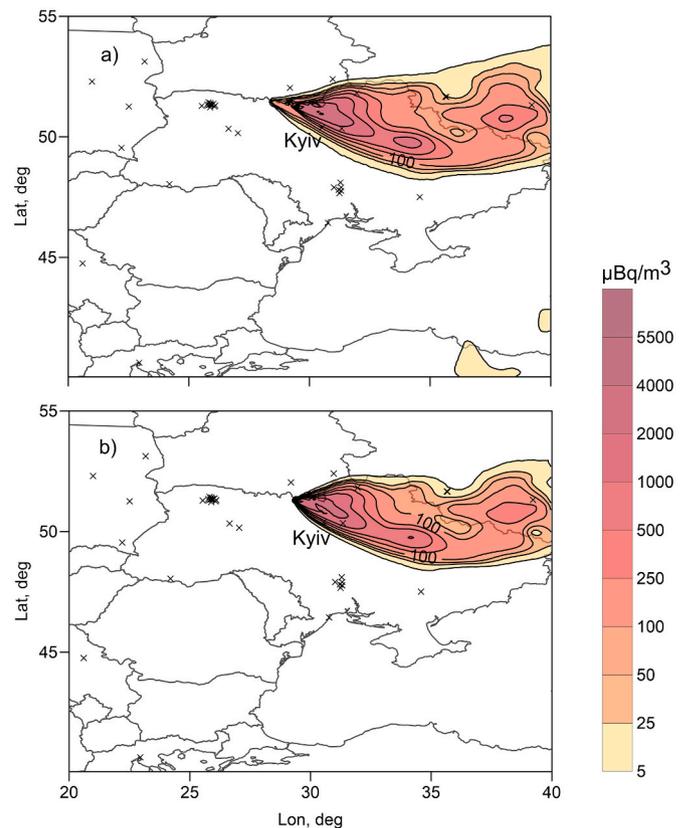


Fig. 14. The simulated ground-level distribution of concentration of ^{137}Cs in the Ensemble run averaged for the 24 h period (April 16, 2020, 06 UTC - April 17, 2020, 06 UTC); a) concentration created by all sources; b) concentration created only by wind resuspension during the dust storm; locations of measurement stations used for solving STE problem are shown by symbols; isolines are shown for values 1, 25, 50, 100, 250, 500, 1500, 2500 $\mu\text{Bq}\cdot\text{m}^{-3}$.

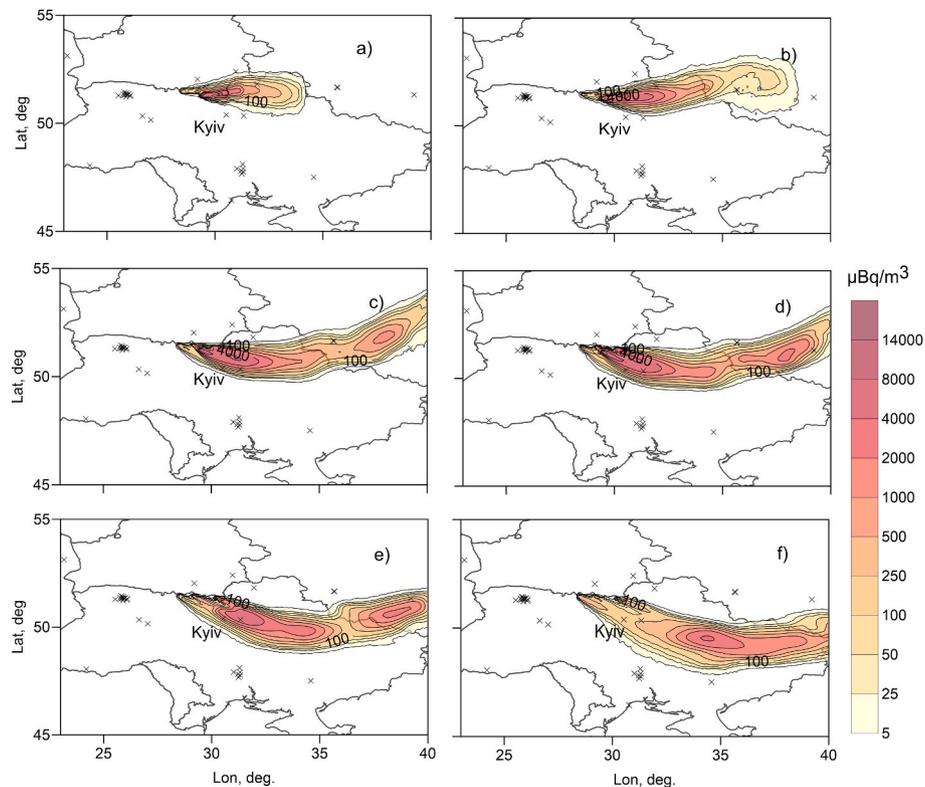


Fig. 15. Concentration distributions of ^{137}Cs near ground simulated in Ensemble run for different times: April, 16th, 06:00 (a), 12:00 (b), 18:00 (c), 20:00 (d), 24:00 (e), April, 17th, 06:00 (f); isolines are shown for values 5, 25, 50, 100, 250, 500, 1000, 2000, 4000, 8000, 14,000 $\mu\text{Bq}\cdot\text{m}^{-3}$.

18:00 UTC) when the increase of $\text{PM}_{2.5}$ was observed. However, this sharp increase in ^{137}Cs concentrations dropped down according to Ensemble run on the morning of 17 April. The daily average concentration of ^{137}Cs in Kyiv was somewhat overestimated in the Ensemble run: $357 \mu\text{Bq}\cdot\text{m}^{-3}$ vs observed $236 \mu\text{Bq}\cdot\text{m}^{-3}$ on 16 April; $197 \mu\text{Bq}\cdot\text{m}^{-3}$ vs observed $114 \mu\text{Bq}\cdot\text{m}^{-3}$ on 17 April; $156 \mu\text{Bq}\cdot\text{m}^{-3}$ vs observed $124 \mu\text{Bq}\cdot\text{m}^{-3}$ on 18 April.

If concentrations of ^{137}Cs and $\text{PM}_{2.5}$ in Kyiv are compared between two time periods, 10 April and 16–18 April, one can see that while on the one hand, $\text{PM}_{2.5}$ concentrations were considerably higher during the dust storm that started on 16 April, on the other hand, concentrations of ^{137}Cs were higher on 10 April. Therefore, it can be concluded that particulate matter measured in Kyiv during the dust storm originated mostly outside ChEZ. This is also confirmed by Fig. 14, which presents spatial distribution of near-surface concentration of ^{137}Cs averaged for the period April 16, 2020, 06 UTC–April 17, 2020, 06 UTC calculated in the Ensemble run. It can be seen that Kyiv is located on the rear of the plume and the maximum concentrations of cesium created by wildfires in ChEZ do not cover Kyiv.

Instantaneous concentrations for different times from April 16 to April 17, 2020 are presented in Fig. 15. It can be seen that during most of 16 April, the plume did not reach Kyiv, but by 18:00 on 16 April, it had turned towards Kyiv. This happened at the time when a rapid increase in $\text{PM}_{2.5}$ concentrations was observed in Kyiv (as shown in Fig. 13). During the remaining time, the plume was transported towards Kyiv, but the intensity of emissions decreased due to a decrease in wind speed and therefore the concentration of ^{137}Cs in Kyiv was reduced. At the same time, as shown in Fig. 13, the concentrations of $\text{PM}_{2.5}$ remained high even on 19 April. This is explained by the fact that the dust storm covered practically the whole of northern Ukraine (Savenets et al., 2020), therefore, high levels of $\text{PM}_{2.5}$ were observed in Kyiv even after 17 April when wind speeds decreased.

Together with the concentrations of ^{137}Cs created by all sources, Fig. 14 also presents concentrations created only by wind resuspension

during the dust storm, as simulated in the Ensemble run. It is interesting to note that despite the fact that according to Fig. 8 emission due to wind resuspension on 16 April was about 2 times less than emissions due to the wildfires, the surface concentrations at distances up to a few hundred km from ChEZ were mainly created by emissions resulting from wind resuspension. This is explained by the fact that while emission from wind resuspension is created close to the Earth's surface, emissions from wildfires were height distributed with the lower bottom of the convective plumes on 16 April at heights more than 500 m, while top heights reached 2500 m (Fig. 3).

6. Discussion and conclusions

In this work, we estimated emission inventories of ^{137}Cs during wildfires in Chernobyl Exclusion Zone (3–24 April 2020) accompanied by dust storm (16–17 April 2020). Uncertainties in meteorological input data, size distribution of the emitted particles, and height distribution of the release were taken into account using the developed Ensemble Iterative Source Inversion Method (EISIM). In this method, a set of source receptor matrices is calculated using FLEXPART atmospheric transport code with varying source term parameters and input meteorological data. The covariance matrix of model errors is estimated by ensemble averaging model results obtained by multiplication of the pre-calculated SRMs by the estimates of emission inventories at the current iteration step. The emission inventories at the next iteration step are evaluated for each of the ensemble members by solving the conventional problem of minimizing a quadratic cost function that characterizes differences between calculated results and observations. Iterations are repeated until a stable solution is obtained for each of the ensemble members. During the course of the iterations, the variance of model error was reduced by a factor of 10 compared to the initial value. The result of the algorithm application is the ensemble of estimates of emission inventories.

The previously published emission inventories resulting from

wildfires (Talerko et al., 2021a) and dust storm (Talerko et al., 2021b) were used as first guess estimations. Simulations of cesium concentrations in the air were performed for the period from April 3 to April 27, 2020. The meteorological input data from 21 realizations of the Global Ensemble Forecasting System (GEFS) operated by NCEP together with GFS final analysis data was used to represent uncertainties in model results caused by errors in meteorological data. The bottom and top heights of convective plumes were evaluated using data from the Global Fire Assimilation System. The fraction of emissions below the convective plume bottom height varied between different members of ensemble estimates from 0.1 to 0.5, while the rest was emitted between the plume bottom and top heights. The size distribution of particles emitted due to the wildfires was defined based on the literature review. It varied between different ensemble members from 0.1 to 0.2 for fine particles (0.25 μm), and from 0 to 0.2 for particles of the 8 μm size. The rest of the release was assumed to be coarse particles (16 μm). A similar approach was used to represent size distribution of particles emitted due to wind resuspension during the dust storm: the fraction of fine particles (1 μm) varied from 0.1 to 0.6, while the rest of the release was assumed to consist of particles of 10 μm size.

The estimated total emission of ^{137}Cs from wildfires was close to the first guess estimations: 448 GBq vs. 586 GBq. However, the confidence interval of the emitted inventory was large: from 39 to 1530 GBq. By using emission inventories within the confidence interval and different combinations of source term parameters and input meteorological data, FLEXPART could fit observations with zero systematic error, correlation coefficient more than 0.6, and normalized mean squared error less than 10. By analyzing model error statistics, some of the source term parameters could be reliably evaluated: the fraction of the fine particles (0.25 μm) in the total emissions was evaluated to be ≈ 0.1 and the fraction of emission below a bottom height of the convective plume was evaluated to be ≈ 0.5 .

The obtained estimate of the total emission resulting from wind resuspension during the dust storm was 27 GBq. It is by a factor of 6 smaller than the respective first guess emission (162 GBq), which did not fall into the estimated confidence interval (from 3 to 93 GBq). The difficulty in obtaining a robust estimate of emissions resulting from wind resuspension is due to the following: 1) emissions were largely reduced on 17 April because of a decline in wind speed, as they became poorly distinguishable from the background of emissions from the wildfires; 2) relatively small number of measurements were affected by the dust storm, because the prevailing direction of the wind was eastward.

By analyzing the impact of changes of different input parameters on estimated emission inventories, it was shown that meteorological input data has the most significant impact on the variability of the total emission inventory. The corresponding relative value of emission variability is caused by different meteorological data: $\lambda_{\text{Meteo}} = 0.49$. Changes in other parameters cause significantly smaller variations in estimated emissions: $\lambda_{\text{Wh}} = 0.26$ for parameters describing the height distribution of the release, $\lambda_{\text{Wsf}} = 0.16$ for parameters related to the size distribution of emitted particles from the wildfires, and $\lambda_{\text{Wsd}} = 0.09$ for parameters related to size distribution of particles emitted by wind resuspension.

The obtained results demonstrate the need for a more in-depth future research regarding combination of wildfires and dust storms, especially for the cases when wildfires happen at radioactively contaminated territories and the contaminated ash is then re-emitted by wind resuspension during a dust storm.

CRedit authorship contribution statement

Ivan V. Kovalets: Conceptualization, Methodology, Investigation, Software, Writing – original draft. **Mykola Talerko:** Conceptualization, Investigation, Methodology, Writing – review & editing. **Roman Synkevych:** Methodology, Investigation, Software, Visualization, Writing – review & editing. **Serhii Koval:** Software, Data curation, Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by a grant from the National Research Foundation of Ukraine No. (2020).02/0048. The computational resources of the Cloud Computing Platform of the Ukrainian National Grid Infrastructure operated by Bogolyubov Institute for Theoretical Physics were used in this study. The authors acknowledge the use of data from FIRMS by NASA's Earth Science Data and Information System (ESDIS), data of Global Fire Assimilation System from Copernicus Atmosphere Monitoring Service Products (CAM5) and data of Global Ensemble Forecasting System (GEFS) operated by NCEP. We thank two anonymous reviewers and Oleksiy Danylenko for their valuable comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atmosenv.2022.119305>.

References

- Ager, A.A., Lasko, R., Myroniuk, V., Zibitsev, S., Day, M.A., Usenya, V., Bogomolov, V., Kovalets, I., Evers, C., 2019. The wildfire problem in areas contaminated by the Chernobyl disaster. *Sci. Total Environ.* 696, 133954 <https://doi.org/10.1016/j.scitotenv.2019.133954>.
- Baró, R., Maurer, C., Brioude, J., Arnold, D., Hirtl, M., 2021. The environmental effects of the April 2020 wildfires and the Cs-137 Re-suspension in the chernobyl exclusion zone: a multi-hazard threat. *Atmosphere* 12 (4), 467. <https://doi.org/10.3390/atmos12040467>.
- Bellasio, R., Scarpato, S., Bianconi, R., Zeppa, P., 2012. APOLLO2, a new long range Lagrangian particle dispersion model and its evaluation against the first ETEX tracer release. *Atmos. Environ.* 57, 244–256. <https://doi.org/10.1016/j.atmosenv.2012.04.017>.
- Daley, R., 1991. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, 1991.
- De Meutter, P., Gueibe, C., Tomas, J., den Outer, P., Apituley, A., Bruggeman, M., Camps, J., Delcloo, A., Knetsch, G.J., Roobol, L., Verheyen, L., 2021. The assessment of the April 2020 chernobyl wildfires and their impact on Cs-137 levels in Belgium and The Netherlands. *J. Environ. Radioact.* 237, 106688 <https://doi.org/10.1016/j.jenvrad.2021.106688>.
- Di Giuseppe, F., Rémy, S., Pappenberger, G., Wetterhall, F., 2018. Combining fire radiative power observations with the fire weather index improves the estimation of fire emissions. *Atmos. Chem. Phys. Discuss.* 18, 5359–5370. <https://doi.org/10.5194/acp-2017-790>.
- Eastaugh, C.S., Hasenauer, H., 2014. Deriving forest fire ignition risk with biogeochemical process modeling. *Environ. Model. Software* 55, 132–142. <https://doi.org/10.1016/j.envsoft.2014.01.018>.
- Enting, I.G., 2002. *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, Cambridge, 2002.
- Evangelou, N., Eckhardt, S., 2020. Uncovering transport, deposition and impact of radionuclides released after the early spring 2020 wildfires in the Chernobyl Exclusion Zone. *Sci. Rep.* 10, 10655 <https://doi.org/10.1038/s41598-020-67620-3>.
- Evensen, G., 2009. *Data Assimilation the Ensemble Kalman Filter*. Springer-Verlag Berlin Heidelberg, 2009.
- GFAS, 2022. ECMWF: ECMWF CAMS global fire assimilation system. Available online: <https://apps.ecmwf.int/datasets/data/cams-gfas/>. (Accessed 17 May 2022).
- Hao, W.M., Baker, S., Lincoln, E., Hudson, S., Lee, S.D., Lemieux, P., 2018. Cesium emissions from laboratory fires. *J. Air Waste Manag. Assoc.* 68 (11), 1211–1223. <https://doi.org/10.1080/10962247.2018.1493001>.
- Ichoku, C., Kaufman, J.Y., 2005. A method to derive smoke emission rates from MODIS fire radiative energy measurements. *IEEE T. Geosci. Remote.* 43 (11), 2636–2649. <https://doi.org/10.1109/TGRS.2005.857328>.
- IRSN, 2020. Fires in Ukraine in the exclusion zone around the Chernobyl power plant: latest measurement results and assessment of environmental and health consequences. https://www.irsna.fr/EN/newsroom/News/Documents/IRSN_Information-Report-Fires-in-Ukraine-in-the-Exclusion-Zone-around-chernobyl-NPP_050520_20.pdf. (Accessed 17 May 2022).

- Kovalets, I., de With, G., 2020. Wildfires in the Chernobyl exclusion zone – a summary of the 2020 event. *Dutch Journal for Radiation Protection* 11 (1), 31–36.
- Kovalets, I., Andronopoulos, S., Hofman, R., Seibert, P., Ievdin, I., Pylypenko, O., 2019. Advanced source inversion module of the JRODOS system. In: Agarwal, R., Agarwal, A., Gupta, T., Sharma, N. (Eds.), *Pollutants from Energy Sources. Energy, Environment, and Sustainability*. Springer, Singapore, pp. 149–186. https://doi.org/10.1007/978-981-13-3281-4_10.
- Masson, O., Romanenko, O., Saunier, O., Kirieiev, S., Protsak, V., Laptev, G., Voitsekhoivych, O., Durand, V., Coppin, F., et al., 2021. Europe-wide atmospheric radionuclide dispersion by unprecedented wildfires in the chernobyl exclusion zone, April 2020. *Environ. Sci. Technol.* 55 (20), 13834–13848. <https://doi.org/10.1021/acs.est.1c03314>.
- NASA, 2020. Fire Information for Resource Management System (FIRMS). <https://firms.modaps.eosdis.nasa.gov/>. (Accessed 18 May 2022).
- NRBU, 1997. *Norms of Radiation Safety of Ukraine (NRBU-97)*, 1997. Ministry of Health of Ukraine, Kyiv (in Ukrainian).
- Pisso, I., Sollum, E., Grythe, H., Kristiansen, N.I., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R.L., et al., 2019. The Lagrangian particle dispersion model FLEXPART version 10.4. *Geosci. Model Dev. (GMD)* 12, 4955–4997. <https://doi.org/10.5194/gmd-12-4955-2019>.
- Protsak, V., Voitsekhoivych, O., Laptev, G., 2020. Estimation of Radioactive Source Term Dynamics for Atmospheric Transport during Wildfires in Chernobyl Zone in Spring 2020. *Ukrainian Hydrometeorological Institute*. URL: <https://uhmi.org.ua/msg/fire2020/analytical.pdf> (date of access. (Accessed 17 May 2022)).
- Protsak, V., Laptev, G., Voitsekhoivych, O., Hinchuk, T., Korychenskyi, K., 2022. Methodology for estimating the emission of radionuclides into the atmosphere from wildfires in the Chernobyl Exclusion Zone. In: EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-11620. <https://doi.org/10.5194/egusphere-egu22-11620>.
- Rémy, S., Veira, A., Paugam, R., Sofiev, M., Kaiser, J.W., Marengo, F., Burton, S.P., Benedetti, A., Engelen, R.J., Ferrare, R., Hair, J.W., 2017. Two global data sets of daily fire emission injection heights since 2003. *Atmos. Chem. Phys.* 17, 2921–2942. <https://doi.org/10.5194/acp-17-2921-2017>.
- Savenets, M., Osadchyi, V., Oreshchenko, A., Pysarenko, L., 2020. Air quality changes in Ukraine during the April 2020 wildfire event. *Geographica Pannonica* 24 (4), 271–284. <https://doi.org/10.5937/gp24-27436>.
- Sofiev, M., Vankevich, R., Ermakova, T., Hakkarainen, J., 2013. Global mapping of maximum emission heights and resulting vertical profiles of wildfire emissions. *Atmos. Chem. Phys.* 13, 7039–7052. <https://doi.org/10.5194/acp-13-7039-2013>, 2013.
- Stohl, A., Thomson, D.J., 1999. A density correction for Lagrangian particle dispersion models. *Bound.-Layer Met.* 90, 155–167.
- Stohl, A., Hittenberger, M., Wotawa, G., 1998. Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiments. *Atmos. Environ.* 32, 4245–4264.
- Stohl, A., Forster, C., Frank, A., Seibert, P., Wotawa, G., 2005. Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* 5, 2461–2474. <https://doi.org/10.5194/acp-5-2461-2005>, 2005.
- Talerko, M., Kovalets, I., Lev, T., Igarashi, Y., Romanenko, O., 2021a. Simulation study of the radionuclide atmospheric transport after wildland fires in the Chernobyl Exclusion Zone in April 2020. *Atmos. Pollut. Res.* 12 (3), 193–204. <https://doi.org/10.1016/j.apr.2021.01.010>, 2021.
- Talerko, M.M., Lev, T.D., Kashpur, V.O., 2021b. Estimation of the contribution of dust storm on April 16, 2020 to radioactive contamination of the atmosphere during forest fires in the exclusion zone. *Nuclear Power and the Environment* 1 (20), 81–95. <https://doi.org/10.31717/2311-8253.21.1.7> (in Ukrainian).
- Vautard, R., Bessagnet, B., Chin, M., Menut, L., 2005. On the contribution of natural Aeolian sources to particulate matter concentrations in Europe: testing hypotheses with a modelling approach. *Atmos. Environ.* 39 (18), 3291–3303. <https://doi.org/10.1016/j.atmosenv.2005.01.051>.
- Wagenbrenner, N.S., Germino, M.J., Lamb, B.K., Robichaud, P.R., Foltz, R.B., 2013. Wind erosion from a sagebrush steppe burned by wildfire: measurements of PM10 and total horizontal sediment flux. *Aeolian Research* 10, 25–36. <https://doi.org/10.1016/j.aeolia.2012.10.003>.
- Winiarek, V., Vira, J., Bocquet, M., Sofiev, M., Saunier, O., 2011. Towards the operational estimation of a radiological plume using data assimilation after a radiological accidental atmospheric release. *Atmos. Environ.* 45 (17), 2944–2955. <https://doi.org/10.1016/j.atmosenv.2010.12.025>.
- Zhang, X.L., Su, G.F., Yuan, H.Y., Chen, J.G., Huang, Q.Y., 2014. Modified ensemble Kalman filter for nuclear accident atmospheric dispersion: prediction improved and source estimated. *J. Hazard Mater.* 280, 143–155. <https://doi.org/10.1016/j.jhazmat.2014.07.064>.
- Zhang, X.L., Su, G.F., Chen, J.G., Raskob, W., Yuan, H.Y., Huang, Q.Y., 2015. Iterative ensemble Kalman filter for atmospheric dispersion in nuclear accidents: an application to Kincaid tracer experiment. *J. Hazard Mater.* 297, 329–339. <https://doi.org/10.1016/j.jhazmat.2015.05.035>.
- Zheng, Q., Leung, J.K.C., Lee, B.Y., Lam, H.Y., 2007. Data assimilation in the atmospheric dispersion model for nuclear accident assessments. *Atmos. Environ.* 41 (11), 2438–2446. <https://doi.org/10.1016/j.atmosenv.2006.05.076>.