



Spatial-temporal source term estimation using deep neural network prior and its application to Chernobyl wildfires

Antonie Brožová^{a,b,*}, Václav Šmídl^a, Ondřej Tichý^a, Nikolaos Evangeliou^c

^a Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodárenskou věží 4, Prague 18200, Czech Republic

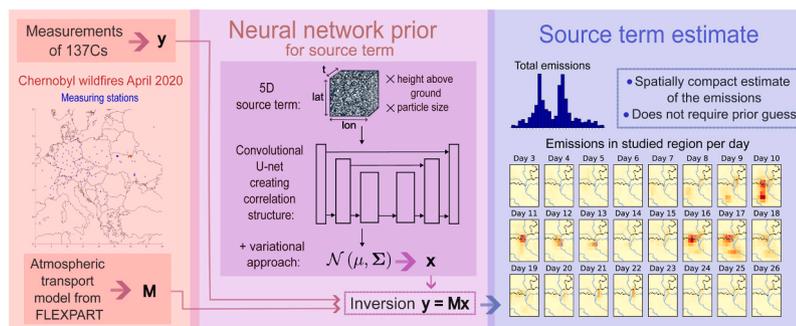
^b Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Trojanova 13, Prague 11200, Czech Republic

^c NILU, Department of Atmospheric & Climate Research (ATMOS), PO Box 100, Kjeller 2027, Norway

HIGHLIGHTS

- We propose deep neural network based prior for source term in atmospheric inversion.
- The neural network creates a parametric correlation structure.
- The number of parameters is smaller than the number of unknowns.
- Estimated emission of ¹³⁷Cs from Chernobyl wildfires in 2020 is comparable with SOTA.
- The source term is 5D: spatial-temporal with height above ground and particle size.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Atmospheric inversion
Spatial-temporal source
Deep image prior
Deep neural networks
Chernobyl wildfires

ABSTRACT

The source term of atmospheric emissions of hazardous materials is a crucial aspect of the analysis of unintended release. Motivated by wildfires of regions contaminated by radioactivity, the focus is placed on the case of airborne transmission of material from 5 dimensions: spatial location described by longitude and latitude in a given area with potentially many sources, time profiles, height above ground level, and the size of particles carrying the material. Since the atmospheric inverse problem is typically ill-posed and the number of measurements is usually too low to estimate the whole 5D tensor, some prior information is necessary. For the first time in this domain, a method based on deep image prior utilizing the structure of a deep neural network to regularize the inversion is proposed. The network is initialized randomly without the need to train it on any dataset first. In tandem with variational optimization, this approach not only introduces smoothness in the spatial estimate of the emissions but also reduces the number of unknowns by enforcing a prior covariance structure in the source term. The strengths of this method are demonstrated on the case of ¹³⁷Cs emissions during the Chernobyl wildfires in 2020.

* Corresponding author at: Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodárenskou věží 4, Prague 18200, Czech Republic.
E-mail address: brozoant@cvut.cz (A. Brožová).

<https://doi.org/10.1016/j.jhazmat.2025.137510>

Received 10 December 2024; Received in revised form 3 February 2025; Accepted 3 February 2025

Available online 5 February 2025

0304-3894/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

Estimation of a source term of unintended atmospheric emissions of a hazardous material is a crucial step for further evaluation of an incident. Its estimation is often formulated as an inverse problem of simulation of atmospheric transport using numerical models. The specifics of the inverse problem depend on the type of the release. A point-source release is a case, where the location of the release is known, which is the case of major accidental releases of radiation from Chernobyl and Fukushima Daiichi nuclear power plants [17,31]. Other minor releases of radiation, e.g. selenium-75 [9], ruthenium-106 [37,52], and iodine-131 [36], are also considered to be point-source emissions, however without the knowledge of the location. Such a case will be denoted as the source location since the task is to select a single location for all temporal steps. Examples of such techniques range from testing different locations in the spatial domain [10,61] to Bayesian techniques [12,13].

In more general cases, the emissions do not originate from a single location but from multiple locations simultaneously. Well-known examples are estimations of greenhouse gasses, methane [60,51,25], SF₆ [59], or black carbon [14,27]. All these studies are focused on spatial-temporal estimation from a limited number of ground measurements, which are not dense in the spatial domain. Therefore, the inverse problem becomes highly ill-posed and some form of regularization is needed. The most common regularization is some form of prior emissions knowledge. For example, Manning et al. [34] assumed the methane emissions to be constant over the analysis period, or the deviation of the prior (first guess) emissions from the optimized emissions are penalized directly within the loss function [38,40]. The standard source term determination methodology then minimizes the mismatch between measurement and the model regularized by the distance between the estimate and the prior (first guess) emissions [39]. Specification of the first guess is necessary to compensate for the lack of data but may cause significant bias in the result when inaccurate. This inaccuracy may be even more significant for inversion from deposition data [11,58] or for multi-species emissions [32,53] where distribution of species within the emissions can vary over time and its settings in the first guess is hard to adjust.

This paper aims to propose an inversion methodology for spatial-temporal sources that does not require a first guess. The inspiration is drawn from the image processing field since gridded elements of the source term can be seen as pixels and the emissions in each spatial location as pixel intensities. Moreover, it can be assumed that the emissions vary smoothly over the spatial locations and that they are sparse, which are two commonly used assumptions about images in image reconstruction problems. Similarly to atmospheric inversion, these problems are often formulated as inversion and, due to their ill-posedness, they are solved using a Bayesian framework with carefully chosen priors [1,56]. However, with the advancement of deep learning techniques, this approach has started to underperform [33]. The benefit of deep learning in this domain does not come from large data but, perhaps surprisingly, from the architectures of neural networks. This concept of deep learning regularization is known as the deep image prior (DIP) [57]. It is based on empirical observation, that convolutional networks like U-net by [45] are more likely to recover smooth images [47].

In this work, for the first time in literature, the DIP is applied to regularize the estimation of the spatial-temporal source term from concentration measurements. It builds on prior work in the area and can be seen as its extension. Specifically, it is demonstrated that DIP is, in essence, an estimation procedure of prior correlations. The prior correlations have been estimated in the source term determination context in temporal dimension [54] using the Variational Bayes methodology. However, the DIP allows correlations in multiple dimensions using the tools of deep learning. Estimation of the DIP within the variational framework has been considered in the image processing literature [26]. This approach was found to be superior to classical DIP due to simpler

tuning [5]. However, the number of observations in the image processing is much larger than that of concentrations available for a typical source-term determination. Therefore, it can substitute the first guess only to a certain limit. It would be clearly insufficient to estimate a large number of pixels from a small number of concentration measurements. Its suitability is expected in cases where a smaller number of sources contribute to the release simultaneously, making the problem spatial-temporal. The Chernobyl wildfires from April 2020 will be used as a demonstrating example.

The wildfires started on 3 April and lasted until approximately 23 April [43,49,30], with a span of a few degrees in latitude and longitude in the spatial domain. Due to the relatively long half-life, more than 30 years, of ¹³⁷Cs, the soil and vegetation are contaminated around Chernobyl after the nuclear power plant disaster, hence, ¹³⁷Cs emissions from the wildfires were measurable across almost the whole Europe [35]. Therefore, the ¹³⁷Cs emissions can be estimated from ambient concentration measurements, however, as shown by Talerko et al. [49], the origin of ¹³⁷Cs emissions is not a point source but a region with a radius of approximately 3 degrees. The source-receptor-sensitivities for this region are modeled using the atmospheric transport model FLEXPART 10.4 [41]. To our best knowledge, the method proposed in this paper is the first method aiming to estimate such a spatial-temporal distribution from radionuclide concentration measurements without predefined fire regions from the satellite data or ground inspection. All previous works such as Talerko et al. [30,49] use either predefined source regions or the first guess of ¹³⁷Cs emissions to regularize the inverse problem. The predefined regions by [30] will be used in this work to demonstrate that the spatial estimate obtained with deep neural network prior provides a meaningful estimate.

2. Review of inversion methods and regularization

2.1. Linear inverse problem formulation

The concept of the source-receptor-sensitivity (SRS) matrix [46] is adopted in this work, which means that the relation between each potential source and a measurement, given by its coordinates and measurement time interval, is modeled backward as a sensitivity to a unit release using an atmospheric transport model. The area of the potential emissions is covered by a grid of point sources and the sensitivity coefficient $m_{i,t,lon,lat}$ corresponding to measurement i in time-step t is obtained for each location on the grid described by its longitude, lon , and latitude, lat . By repetition of this procedure for each measurement and potential emissions interval, a 4D tensor $\tilde{\mathbf{M}}$ is constructed, which contains the SRS matrices $\tilde{\mathbf{M}}_{lon,lat}$ for each location on the grid. Denoting the measurement vector as \mathbf{y} and the emissions vector at location (lon, lat) as $\mathbf{x}_{lon,lat}$, the link between the measurements and reconstructions can be formulated as

$$\mathbf{y} = \sum_{lon,lat} \tilde{\mathbf{M}}_{lon,lat} \mathbf{x}_{lon,lat} + \mathbf{e}, \quad (1)$$

where \mathbf{e} is an error term of reconstruction. Hence, the vector \mathbf{y} is assumed to be explained by contributions from each element of the considered spatial domain. Moreover, the equation can be rewritten as

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}, \quad (2)$$

by simple rearranging all $\mathbf{x}_{lon,lat}$ into one array and $\tilde{\mathbf{M}}_{lon,lat}$ into one huge matrix. The goal is then to estimate the unknown emissions \mathbf{x} from the linear model (2).

2.1.1. Dimension of the source term

The simplest assumption on the source term is that the location is known, as well as the composition of the release. The source is then one-dimensional with the only unknown dimension denoting temporal

evolution of the source. In this contribution, however, the source term is considered to be five-dimensional: time-varying with three particle size fractions and possibly emitting simultaneously from different latitude, longitude, and height above ground level.

The matrix \mathbf{M} is often ill-conditioned even in simpler scenarios, implying an unreliable estimation via ordinary least squares (or ridge regression). The considered five-dimensional extension becomes even more sensitive and solid regularization has to be employed.

2.2. Standard linear source term estimation using first guess

The goal of the inversion is to minimize the reconstruction error J_{reco} , defined typically as a quadratic form

$$J_{\text{reco}}(\mathbf{x}) = (\mathbf{y} - \mathbf{M}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{M}\mathbf{x}), \quad (3)$$

with covariance matrix \mathbf{R} standing for tolerance of deviation between measurements and model reconstructions. However, the solution is often unstable and the first guess of emissions, \mathbf{x}_0 , is employed in majority of applications due to ill-conditionality of the inverse problem either of the point-source [15,55], or from the selected spatial-temporal domain [30,60,14,59]. The use of the first guess then can be seen as the form of the 3D-Var method [4] or regularization of the inverse problem [30]. Specifically a term measuring the distance of \mathbf{x} from \mathbf{x}_0 formulated as

$$J_{\text{prior}}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_0), \quad (4)$$

where \mathbf{B} is a covariance matrix which weights the penalization of the proximity of the estimated emissions from the first guess, is added to the reconstruction error J_{reco} forming a function

$$J(\mathbf{x}) = J_{\text{reco}}(\mathbf{x}) + J_{\text{prior}}(\mathbf{x}), \quad (5)$$

which, when minimized, leads to a solution of the inverse problem as close as possible to the first guess.

The matrix \mathbf{B} is often selected as a scalar coefficient [19], some form of a diagonal matrix, or as a weighted combination of the diagonal and differential operator [15,18] to favor smoothness of the solution. Determination of a reliable first guess, \mathbf{x}_0 , is a much more demanding task since it is often calculated or estimated based on very uncertain assumptions. Even obtaining a stable solution with a reasonable first guess may be a challenging task considering the ill-posedness of the inverse problem [55]. Determination of \mathbf{x}_0 becomes even more demanding considering the extra dimensions in terms of the unknown particle size distribution and unknown vertical profile of the emissions. Similar assumptions as were done by Kovalets et al. [30] can be made, e.g. the portion of the smallest fraction being between 10 % and 20 % [22, 16] or on the bottom and top heights of wildfire plumes [44]; however, the first guess will be strongly biased and can affect the results significantly.

One commonly used regularization, when first guess is not available, is Tikhonov regularization. In such a case, \mathbf{x}_0 is chosen to be zero vector, and matrix \mathbf{B} has the form $\lambda \mathbf{I}$, where λ needs to be estimated. Based on its value, the estimate is a compromise between zero array and direct solution of (3).

2.3. Bayesian inversion and variational bayes

The Bayesian interpretation of the standard inversion of the linear problem (5) is based on the correspondence of quadratic forms and the negative loglikelihood of a Gaussian distribution. The reconstruction error (3) and the prior (4) correspond to probabilistic models

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{M}\mathbf{x}, \mathbf{R}), \quad (6)$$

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}_0, \mathbf{B}). \quad (7)$$

Meaning of symbols \mathbf{x}_0 , \mathbf{B} , and \mathbf{R} is the same as in (5).

The negative loglikelihood is thus (5) which can be inverted using the Bayes rule to obtain:

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &= \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \\ \boldsymbol{\mu} &= \boldsymbol{\Sigma}(\mathbf{M}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{B}^{-1} \mathbf{x}_0), \\ \boldsymbol{\Sigma} &= (\mathbf{M}^T \mathbf{R}^{-1} \mathbf{M} + \mathbf{B}^{-1})^{-1}. \end{aligned} \quad (8)$$

Note that the choice of matrix \mathbf{B} is essential for regularization of the inversion in $\boldsymbol{\Sigma}$ and \mathbf{x}_0 steers the posterior mean.

However, the Bayesian formulation allows us to go beyond standard results, e.g. by estimating the covariance matrices \mathbf{B} . While the estimation of the full covariance matrix is problematic due to a low number of observations, the estimation of a bi-diagonal matrix \mathbf{D} such that $\mathbf{B}^{-1} = \mathbf{D}^T \mathbf{D}$ is possible. Bidiagonal matrix \mathbf{D} with fixed elements is commonly used in the standard approach to promote the smoothness of the solution [15]. Estimating the elements of matrix \mathbf{D} is possible using the Variational Bayes [54]. This approach is closely related to the proposed method for two reasons: i) the proposed prior can be seen as a generalization of the unknown correlations of neighboring elements of \mathbf{x}_0 (as will be further explained in Section 3.1), and ii) the Variational Bayes methodology is important to achieve reliable estimation of the DIP [5].

The Variational Bayes approach is a method of calculating approximate posterior distribution. It is based on defining a class of the approximate distributions, $q(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$, parameterized by $\boldsymbol{\theta}$ and solving minimization problem

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \operatorname{KL}(q(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}) || p(\mathbf{x}|\mathbf{y})) \quad (9)$$

$$= \underset{\boldsymbol{\theta}}{\operatorname{argmin}} - \mathbb{E}_q[\log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x}) - \log q(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})]. \quad (10)$$

where $\boldsymbol{\theta}^*$ is its optimal value, KL is the Kullback-Leibler divergence, and its alternative formulation (10) is known as the evidence lower bound [3]. While the former formulation is used with the conditional independence form of q , e.g. by Tichý et al. [54], the latter is used with gradient descent-based optimizations.

3. Inverse modeling using deep image prior

The proposed method is based on two approaches used in image processing, namely the deep image prior (DIP) [57] and Variational DIP [26]. Their relevance to the studied problem will be now briefly illustrated, their use will be explained and, finally, the proposed inversion algorithm will be derived.

3.1. Deep image prior

The DIP hypotheses stems from empirical observations in image reconstruction from noisy measurements (inverse modeling). Specifically, the quality of the reconstructed image is systematically better when the image is represented by a noise vector transformed by a deep neural network. The hypothesis states that the structure of the neural network (topology, layers, activation functions, etc.) imposes regularization of the inverse task. A common architecture in these tasks is the convolutional U-net [45]. While this hypothesis is supported by empirical studies, e.g., by Shi et al. [47], strict proof is not available. Therefore, our own intuition will be provided with a simple example.

3.1.1. Toy U-net example

Consider a two-layer U-net, $\mathbf{x} = \operatorname{Unet}_{\boldsymbol{\theta}}(\mathbf{z})$, with linear activation functions and only one channel with 2×2 kernels as displayed in Fig. 1. The input of the network, \mathbf{z} , is a realization of random noise (modified with one strong pixel for clarity). Parameters of the convolution kernels, $\boldsymbol{\theta}$, are initialized randomly but are estimated from the data. Therefore, the free parameters are displayed in color, each color in the convolution kernels denoting one parameter – four in total in this case. Each inner

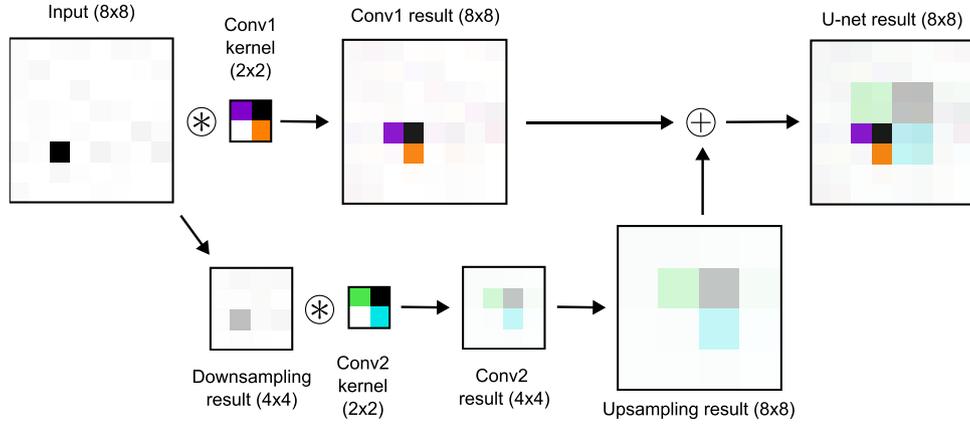


Fig. 1. Illustration of a two-layer U-net. On the top level, the input noise array is convolved with Conv1 kernel. The same input array is divided into a grid of 2×2 pixels, and it is downsampled by averaging each square of the grid. The downsampled input is convolved with Conv2 kernel and the upsampled by repeating every pixel once in both rows and columns. The upsampled result is then summed with the output of the first convolution, together forming the final output of the neural network.

layer starts with a downsampling operator and ends with an upsampling operator, thus effectively enlarging the image area on which the convolution operator acts.

The linear activations were chosen to allow for the analytical tractability of the representation. For a normally distributed input, $\mathbf{z} \sim \mathcal{N}(0, I)$, the toy U-net is a linear operator, say $\mathbf{x} = A\mathbf{z}$, hence the distribution of the output is $p(\mathbf{x}) = \mathcal{N}(0, A^T A)$. The convolution parameters thus impose covariance of the pixels of \mathbf{x} . The covariance matrix of a column-wise vectorization of image \mathbf{x} from Fig. 1 is displayed in Fig. 2. The colors in the covariance matrix correspond to the colors of the convolution parameters. The resulting multi-diagonal covariance structure is thus a generalization of the bi-diagonal structure of matrix D [54] mentioned in Section 2.3.

Note that the four free parameters of the convolution kernels define four degrees of freedom of the prior covariance matrix in this toy example. The full Unet used in DIP uses multi-channel convolutions and nonlinear activation functions, which further complicates the prior dependence of the elements. However, the number of parameters can still be much lower than the number of unknowns.

Since the training of the U-net uses only one sample of the input, the resulting output of the network is not a distribution but a numerical vector. Thus the U-net is essentially only a reparametrization of the unknown quantity. A single parameter influences many pixels on the

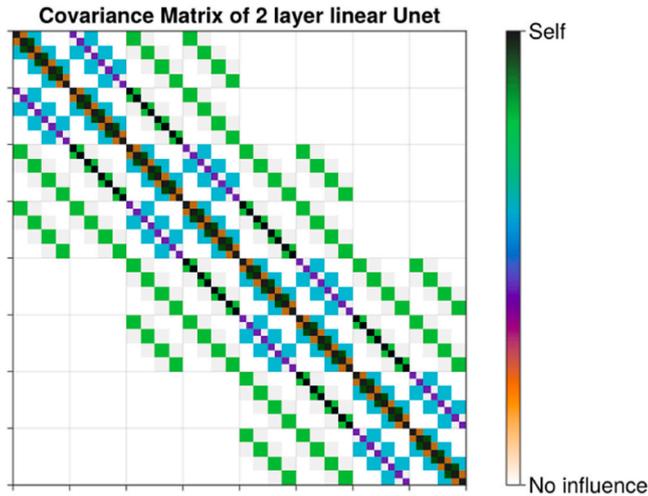


Fig. 2. Illustration of covariance structure created by two-layer U-net on random noise. The colors correspond to the colors (values) of the convolutional kernels Conv1 and Conv2 in Fig. 1.

output, thus decreasing its sensitivity to potential outliers. However, the prior term is not present in the loss function (DIP generally optimizes only J_{reco}). Despite its name, the original DIP is not a proper prior but a reparametrization. The regularization is achieved by careful tuning of the optimization method [5], which can be fragile. A much more stable extension of this approach is based on Variational Bayes approximation.

3.2. Variational DIP

Variational DIP [26] is an application of the DIP in a Bayesian framework using the Variational Bayes approximation (9). Specifically, the neural network of the U-net type is designed to have the output dimension twice the size of the unknown \mathbf{x} . The first part of the output is interpreted as the mean, $\boldsymbol{\mu}$, and the second part as the standard deviation, $\boldsymbol{\sigma}$, of the posterior distribution:

$$q(\mathbf{x}|\mathbf{y}) = \mathcal{N}(\boldsymbol{\mu}, \text{diag}(\boldsymbol{\sigma}^2)) \quad (11)$$

$$[\boldsymbol{\mu}, \boldsymbol{\sigma}] = \text{Unet}_\theta(\mathbf{z}), \quad (12)$$

where \mathbf{z} is a fixed sample of the same dimension as \mathbf{x} from uniform or standard normal distribution and θ denotes parameters of the U-net.

Evaluation of the lower bound (10) for likelihood (6) and (12) uses the parametrization trick [28] to approximate the expected value with respect to $q(\mathbf{x}|\mathbf{y})$, leading to the loss function

$$J_{\text{vdip}}(\boldsymbol{\mu}, \boldsymbol{\sigma}) = \frac{\omega}{2} (\mathbf{y} - \mathbf{M}\boldsymbol{\mu})^T (\mathbf{y} - \mathbf{M}\boldsymbol{\mu}) + \frac{\omega}{2} (-2\epsilon^T \text{diag}(\boldsymbol{\sigma})^T \mathbf{M}^T \mathbf{M} \boldsymbol{\mu} + \epsilon^T \text{diag}(\boldsymbol{\sigma})^T \mathbf{M}^T \mathbf{M} \text{diag}(\boldsymbol{\sigma}) \epsilon) - \sum_{i=1}^N \log \sigma_i \quad (13)$$

which was derived for non-informative prior $p(\mathbf{x}) \propto 1$. An appropriate prior for our task will be derived in the next section. The reparametrization trick can be used to approximate an expected value of an arbitrary function $g(\mathbf{x})$ with respect to q using a sample ϵ from the standard Gaussian distribution as $\mathbb{E}_q[g(\mathbf{x})] \approx g(\boldsymbol{\mu} + \text{diag}(\boldsymbol{\sigma})\epsilon)$ where ϵ is resampled in each iteration of the gradient descent optimization of (13).

Note that the first term of (13) is scaled J_{reco} from (3) for $\mathbf{x} = \boldsymbol{\mu}$. Since posterior mean $\boldsymbol{\mu}$ is the final estimate of \mathbf{x} , minimizing J_{vdip} is an alternative to minimizing J from (5) with the first guess term replaced by additional terms (regularization) arising from the DIP model.

3.3. Proposed inversion algorithm

In this Section, the general Variational DIP approach is specialized to estimate the spatio-temporal source term. This requires making specific

choices of the components of the probabilistic model (i.e. likelihood and the prior), as well as algorithmic choices of the implied optimization problem.

3.3.1. Likelihood model

The likelihood model is conventional as in (6), only the covariance matrix \mathbf{R} of the data is simplified to $\omega^{-1}\mathbb{I}$, where ω is a fixed precision parameter and \mathbb{I} is the identity matrix. This means that all measurement residues are i.i.d. distributed.

The source term, \mathbf{x} , is designed as a 5D tensor, where the first two axes correspond to location, the third axis to time, the fourth to height above ground level, and the fifth to particle sizes. The exact neural network Unet_θ utilized in the inversion is U-net with skip connections containing three levels of resolution, with downsampling via strided convolution with stride of 2 and upsampling performed by bilinear interpolation. The input \mathbf{z} is an array of the same size as the source term, so the number of levels is chosen so that the 3D array of size (longitude-steps) \times (latitude-steps) \times (time-steps) can be downsampled without reaching scalar on the lowest level. Instead of 2D convolution used in image processing to reconstruct 2D images, 3D convolution is used in the convolutional layer, effectively creating short-range dependence between neighboring locations and time steps by 3×3 convolutional filters. Height levels are treated as channels, with their number being doubled on lower level of the U-net. This means that the information about the emissions at each altitude is distributed among all others. Lastly, particle size fractions are treated as samples of a minibatch, so the emissions values with respect to fractions are not mixed by the convolutions. As a result, the number of parameters of the network is lower than the total number of unknown values in the 5D source term.

3.3.2. Prior model

Our method aims to regularize the problem without needing a first guess. However, non-informative prior distribution is insufficient to restrict the solution in the particle size fractions. Regularization can be achieved using a relatively vague prior that the ratio between two fractions at the same time and space is within a factor of 100, i.e. higher than 0.01 and lower than 100. While this assumption can be formalized to explicit $p(\mathbf{x})$ [53], the implied computational cost becomes prohibitive. Therefore, the expected value of the implied $\log p(\mathbf{x})$ is approximated by a masking approach.

Specifically, a vector of masking variables \mathbf{m} of the same dimension as $\boldsymbol{\mu}$ is introduced with entries $m_i = 0$ if the corresponding μ_i is within the prescribed bound and $m_i = 1$ otherwise. The loss function (13) is thus modified for the chosen prior as follows:

$$J_{\text{final}}(\boldsymbol{\mu}, \boldsymbol{\sigma}) = \frac{\omega}{2}(\mathbf{y} - \mathbf{M}\boldsymbol{\mu})^T(\mathbf{y} - \mathbf{M}\boldsymbol{\mu}) + \frac{\omega}{2}(-2\boldsymbol{\epsilon}^T \text{diag}(\boldsymbol{\sigma})^T \mathbf{M}^T \mathbf{M} \boldsymbol{\mu} + \boldsymbol{\epsilon}^T \text{diag}(\boldsymbol{\sigma})^T \mathbf{M}^T \mathbf{M} \text{diag}(\boldsymbol{\sigma}) \boldsymbol{\epsilon}) - \sum_{i=1}^N \log \sigma_i + \max\left(1, \sum_{i=1}^N m_i\right)^{-1} \sum_{i=1}^N (m_i r_i)^2, \quad (14)$$

where r_i denotes the difference between μ_i and the bounding value.

3.3.3. Proposed optimization strategy

The loss function J_{final} is minimized w.r.t. the parameters $\boldsymbol{\theta}$ through $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. This is performed by iterative minimization via gradient descent with the Adam optimizer [29] and the reparametrization variable $\boldsymbol{\epsilon}$ is sampled in every iteration. Finally, an estimate of the source term \mathbf{x} is the posterior mean $\boldsymbol{\mu}$, the output of the optimized network $\text{Unet}_\theta(\mathbf{z})$, where \mathbf{z} is fixed during the optimization.

The minimized loss function is high-dimensional and non-convex w.r.t. to the parameters $\boldsymbol{\theta}$ and the initialization of the neural network is

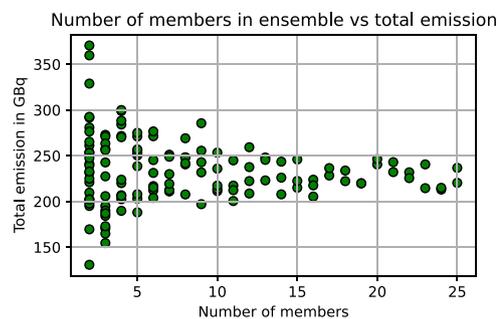


Fig. 3. Total estimated emission as a function of the number of ensemble members.

random (according to [23]), which, together with sampling in the reparametrization trick causes the optimization to arrive at different solutions based on the seed of a random sampler. In order to deal with this instability, two additional steps were added to the method: i) pretraining, and ii) ensembling.

Firstly, the neural network is pretrained to produce a constant array for both the mean and standard deviation of the posterior distribution. This is achieved by minimizing the logarithm of MSE between the output of the U-net and the constant targets. Although this does not ensure identical weights throughout the network, the output is the same at the start of the inversion process. Secondly, the classical ensembling approach [42] was found to overcome the problem with local minima well in our case. Specifically, the inversion is performed using K different sets of initial values (using K different seeds of the random sampler), and the final estimate of the source term is derived by averaging the source terms predicted from these runs.

A relatively low number of ensembles (e.g. $K = 15$) was found to be sufficient to obtain a reproducible result. This is demonstrated in Fig. 3 in terms of the average total estimated released activity for a varying number of ensemble members. Note that for independent runs, the estimated activity varies significantly and becomes more consistent with an increased number of ensemble members. Moreover, this approach allows us to quantify uncertainty in the final estimate of the source term.

The proposed algorithm will be called the Unet posterior, and its flowchart is displayed in Fig. 4. Implementation of the algorithm and used data are available at: <https://github.com/utia-Bayes/Unet-Chernobyl-Wildfires/>

4. Case description, data, and atmospheric transport modeling

4.1. Chernobyl fires in April 2020

The state of the art regarding the Chernobyl fires in April 2020 with a focus on studies considering the spatial-temporal ^{137}Cs will be now reviewed. Estimates using the assumption of point-source emissions and the bottom-up approach of Evangelidou and Eckhardt [16] yielded the total estimate of 341 GBq released during wildfires until 20 April. The top-down approach has been used by Baró et al. [2] for the period 4–13 April with an estimate of around 600 GBq. The whole period of wildfires

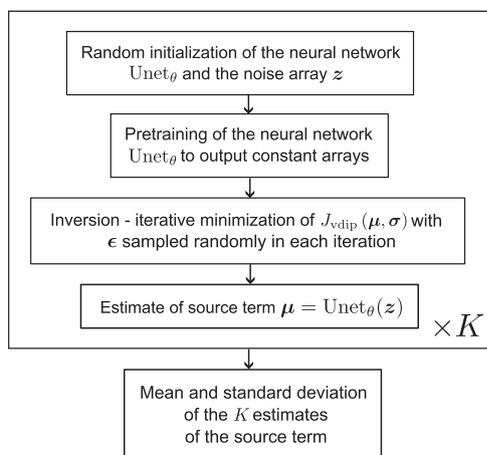


Fig. 4. Flowchart of the method Unet posterior.

has been considered by De Meutter et al. [8] with the result of 650 GBq within the uncertainty interval between 220 and 1810 GBq. These findings are in agreement with those of Masson et al. [35], who estimated that ^{137}Cs emissions were between 700 and 1200 GBq using the same dataset as used in this study.

The methods considering spatial-temporal emissions for this case are typically based on careful analysis of satellite data with locations of wildfires such as the Fire Information for Resource Management System (FIRMS). The spatial information can be subsequently combined with the landcover information and with a map of deposited ^{137}Cs caused by the Chernobyl accident and the emissions factors for radioactive particles. The first estimate of ^{137}Cs emissions was done by Protsak et al. [43] with an estimate of 690 GBq. This analysis was later updated by Talerko et al. [49] using a method based on the radionuclide inventory in the fire area. They provide a comprehensive spatial-temporal description of the spread of the wildfires. The most intense wildfires were observed in two periods, 3–13 April and 16–21 April. In the first period, wildfires broke out in the western part of the Chernobyl Exclusion Zone (CEZ) and spread to the northeast to the border with Belarus. The wildfires also broke out near Chernobyl Nuclear Power Plant (NPP) on 8 April and passed through the significantly contaminated area of the Red Forest, and, on 9 April, the area around the cooling pond of Chernobyl NPP had also been affected. Due to rain and the work of firefighters, the wildfires almost vanished on the next two days, 14 and 15 April. However, the wildfires flared up again on 16 April, also due to the dust storm combined with the strong wind [50]. This led to the spread of wildfires in the northern part of the CEZ, east of the Chernobyl NPP over the Pripyat river, and also the northwest of the CEZ in the Zhytomyr region. After 21 April, the intensity of wildfires decreased, although activity peak has been reported even after this date by several studies [35,8]. This bottom-up information on the locations of the wildfires was used by Kovalets et al. [30] as input information for inversion, where atmospheric concentration measurements and outputs of an atmospheric transport model were optimized, testing also hypothesis on different particle size fractions and emissions altitude, producing a total estimate of 469 GBq with uncertainty bounds from 36 to 1623 GBq.

4.2. ^{137}Cs measurements dataset

The most complete dataset of the ^{137}Cs concentration measurements of the April 2020 wildfires around Chernobyl [35] is used in this study. The complete dataset contains 1095 measurements from the period between January and June 2020, of which 858 have the start or end date within the period of wildfires, April 2020. The concentrations have relatively large interval of values ranging up to 180000 μBqm^{-3} . All measurements over 1000 μBqm^{-3} have been obtained inside the

Chernobyl Exclusion Zone (CEZ), while the maximum outside CEZ is approximately 700 μBqm^{-3} . The measurement locations include Ukraine and more than 20 other European countries, ensuring relatively good spatial coverage of the computational domain. The complete dataset and details can be found in [35].

4.3. Atmospheric transport modeling

To construct the SRSs, the Lagrangian particle dispersion model FLEXPART version 10.4 [41] is used in the backward “retroplume” mode. The computational particles are tracked backward in time using hourly ERA5 [24] assimilated meteorological analyses from the European Center for Medium-Range Weather Forecasts (ECMWF) with 137 vertical layers and a horizontal resolution of $0.5^\circ \times 0.5^\circ$. FLEXPART takes into account turbulence [6], unresolved mesoscale motions [48], and convection [20].

The output resolution was set to $0.5^\circ \times 0.5^\circ$, and the SRS coefficients are extracted for an area spanning between 48.5° and 54.0° N in latitude and between 27.5° and 33.0° E in longitude. Therefore, the emissions of ^{137}Cs will be estimated in the area of the size 11×11 with resolution 0.5° and with the Chernobyl NPP in the middle of the domain. Three different aerodynamic particle diameters are taken into account: 0.4 μm , 8 μm , and 16 μm , and 7 vertical layers corresponding to the height above ground level of 0–100, 100–500, 500–1000, 1000–1500, 1500–2000, 2000–2500, and 2500–3000 m.

Following the approach outlined by Kovalets et al. [30], the limitations of atmospheric transport modeling in simulating both short- and long-distance dispersion are addressed [41]. In particular, the model does not fully capture turbulent dispersion at short distances. Near the surface, the planetary boundary layer (PBL) is influenced by terrain, vegetation, and urban structures, creating complex flow patterns that are difficult to model accurately [7]. Additionally, at short distances, small errors in the source location, emission rate, or initial particle size distribution can lead to significant deviations in model predictions. Therefore, only those measurements taken outside the Chernobyl exclusion zone, which are all located at a considerable distance from the wildfires, are taken into account. After this filtering, the number of measurements considered in this study is reduced to 792.

5. Results

The emissions are estimated from 2 to 29 April, which creates a five-dimensional source term of size $11 \times 11 \times 28 \times 3 \times 7$ (latitude \times longitude \times time \times size fractions \times altitudes). For comparison, results obtained using the Tikhonov regularization with prior guess set to zeros, as explained in Section 2.2, are also presented. Since there is no ground-truth data of the emissions, the results are compared to a first guess used by Kovalets et al. in [30]. The spatio-temporal distribution of this first guess is based on satellite data provided by NASA. The associated emissions are calculated using the Atlas of radioactive contamination of Ukraine, which was used for evaluation of contamination of the territories where wildfires took place, see [49,30]. Kovalets et al. [30] assumed the end of the wildfires on 23 April, so the first guess was set to zero emissions after this date. Note that the first guess estimate was obtained from completely different data than our results and thus can be compared only to a certain extent.

5.1. Experimental setting

5.1.1. Tikhonov

The same constraints as in Unet posterior are used; that is, the source term contains only non-negative values, and ratios of size fractions are between 10^{-2} and 10^2 . It was implemented as a block-coordinate descent algorithm with each coordinate corresponding to one fraction and conjugate gradients were used to solve the constrained subproblem for each fraction in each step. The right trade-off between the

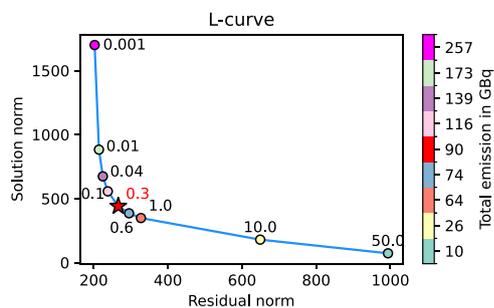


Fig. 5. L-curve of solutions with Tikhonov regularization for a set of different λ values. The red star denotes the chosen value. The color bar shows the total estimated emissions for the corresponding λ value.

regularization and reconstruction error can be found by plotting the norm of the solution against the norm of the reconstruction error for a set of values of λ . Such a curve is called the L-curve and the point with the highest curvature gives the optimal λ value [21]. The best λ parameter for this problem was chosen according to L-curve in Fig. 5 to be 0.3.

5.1.2. Unet posterior

Since the input to the neural network is random and the parameters of the networks are randomly initialized, averaged result from $K = 15$ runs of the algorithm with different seeds is reported. The code was implemented in Julia and executed on NVIDIA GeForce RTX 2080 Ti. Note that no other training dataset was used to estimate the source term and all measurements are used in every step of the inversion. The parameter ω was set to 10^{-2} . The total amount of parameters in the U-net is 6314, which, compared to the unknown source term, is $10\times$ lower number of unknowns.

5.2. Correspondence between reconstructions and the data

Fig. 6 shows that both Tikhonov regularization and Unet posterior fit

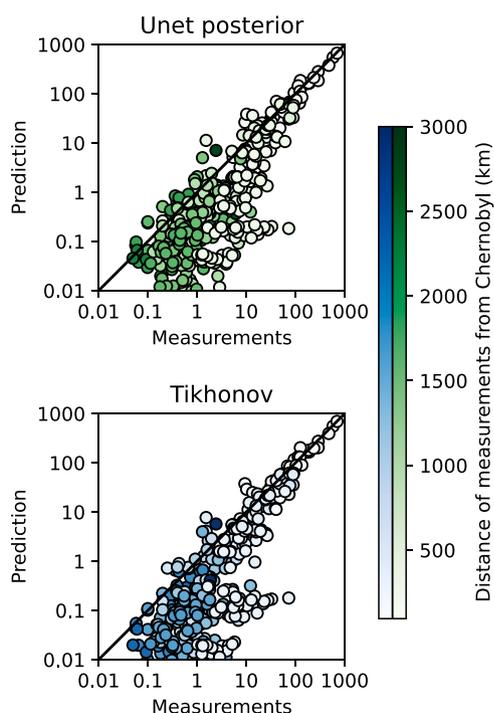


Fig. 6. Scatterplot of data and predicted values. The color denotes the distance from Chernobyl NPP.

Table 1

Comparison of the total estimated release. LB and UB stand for lower bound and upper bound on total emissions, r stands for Pearson correlation coefficient. LB, Total, and UB are in GBq.

	LB	Total	UB	r
Unet posterior	189	225	377	0.974
Tikhonov	-	90	-	0.979
[16]	-	341	-	0.58
[43]	-	690	-	-
[49]	-	574	-	-
[35]	700	950	1200	-
[30]	36	469	1623	0.85

the data well. The highest measurements that are closest to the Chernobyl NNP are fitted the best and deviations appear further away from this area.

Table 1 shows a quantitative comparison of the algorithms. From the Pearson correlation coefficient r , it can be seen that both models reflect the data very well. Tikhonov regularization reaches a slightly better result than Unet posterior, which has lower freedom in estimation due to a ten times lower number of parameters. Nevertheless, compared to results reported by Evangelidou and Eckhardt [16] and Kovalets et al. [30], the two methods perform similarly and achieve a better fit with the data than previous methods.

5.3. Spatial estimate

While the solutions returned by both Unet posterior and Tikhonov correspond well to the data, this does not necessarily mean that their estimates are equally accurate. The ill-posedness of the inversion may allow for very different solutions depending on their priors, making a good fit with the data only a necessary condition. Therefore, the spatial-temporal estimates of the source term will be compared with the first guess created by Kovalets, which is based on different data.

Fig. 7 shows the estimated activity in time and space together with the Kovalets first guess. Both Unet posterior and Tikhonov achieved solution that has a similar time profile as the first guess. Both methods struggle to estimate emission peaks between 3 and 7 April which may be caused by a large number of measurements contributing to the emissions. The number of related non-zero SRS coefficients is higher in this time period than in the rest of the month, making the choice of the correct spatial location of the emissions more difficult. Between 8 and 20 April, both estimates are closer to the Kovalets first guess in time and the peaks are also located similarly. Nevertheless, looking at 10 April, both methods deviate significantly from the almost point emission first guess by Kovalets. Moreover, from 16 to 19 April, some emissions are estimated southwest of the Chernobyl NPP, which is not in the first guess. On the other hand, on 16 April, the emissions spread more to the west (compared to the beginning of April) due to the dust storm, and Unet posterior shows this behavior as well. Both models estimated lower emissions at the end of the month compared to earlier days, aligning with Kovalets first guess. However, a deviation from the first guess emerges on April 20, peaking on April 22, which is estimated by both Unet posterior and Tikhonov. These emissions largely contribute to a high measurement of $290 \mu\text{Bqm}^{-3}$ measured between 20 and 23 April by a station in Ukraine. Therefore, if the high emissions in this area are incorrect, it is likely caused by an inaccuracy in an atmospheric model. The following two paragraphs will discuss more closely deviations specific to each of the two methods.

Unet posterior differs from the Kovalets first guess primarily due to a greater dispersion of the emissions, which may be caused by the spatial convolutions used in U-net. Apart from that, significant emissions are convoluted in two distinct regions that were not predicted by the first guess. In order to get a better understanding of these, Fig. 8 shows what emissions were averaged in the ensemble of Unet posterior on 10 and 17 April. On 10 April, strong emissions are estimated not only in the

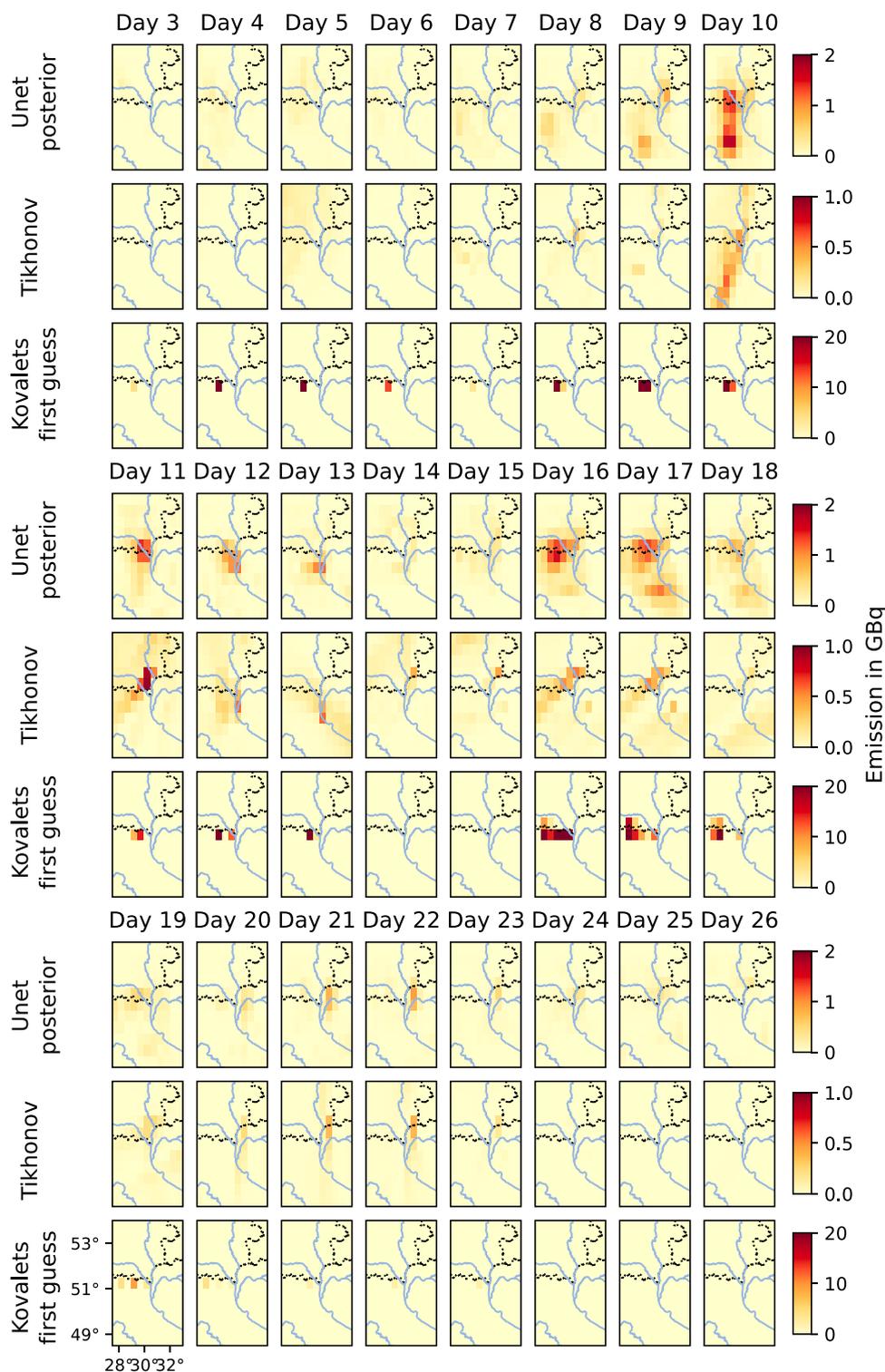


Fig. 7. Comparison of estimation of spatial-temporal emissions by Unet posterior, Tikhonov regularization and first guess used by Kovalets et al. [30] for the studied region: area around Chernobyl, Ukraine. Each map shows emissions on one day, summed over particle sizes and altitudes. Estimates on 2 and 27–29 April are cropped, they do not contain any significant emissions.

Chernobyl area but also around longitude of 29.75° E and latitude of 49.75° N. As the figure shows, the estimate in this area is higher than around Chernobyl due to three runs of the fifteen that were averaged to get the estimate, so while the emissions are estimated high in this location, there is higher uncertainty than in the area around Chernobyl. Furthermore, this part of ^{137}Cs was estimated to be emitted at higher

levels than the one around Chernobyl (in Fig. 11 can be seen that the portion of particles in height of 2500–3000 m is higher). The source-receptor sensitivity is lower in higher altitudes, which allows for higher values of emissions and higher uncertainty. Another significant difference appears around 17 April in an area around a longitude of 29.75° E and a latitude of 49.75° N, where, again, the higher emissions

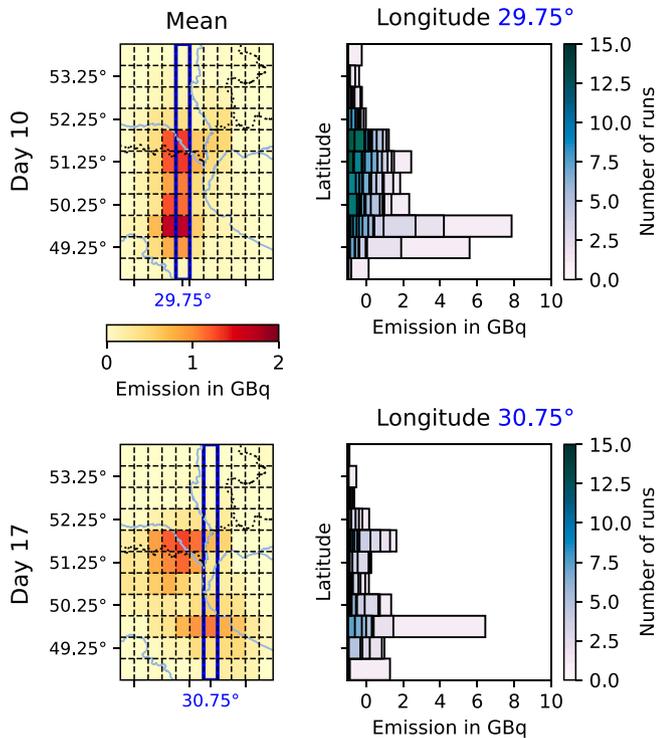


Fig. 8. Uncertainty in the estimate of the emissions for day 10 and day 17. The maps on the left side of the figure show estimate of emissions obtained by Unet posterior - that is by averaging the results of 15 runs of the algorithm. The barplot on the right side shows emissions estimated by each run of the algorithm for locations highlighted by the blue rectangle in the maps (longitude 29.75°, resp. 30.75° E and all latitudes considered in the source term). Each bar location corresponds to one latitude and the overlapping bars show emissions achieved by each run.

are caused mainly by two runs of the fifteen. On the other hand, most runs estimate activity stronger than around Chernobyl. In this case, the high emissions on 16 April are mostly estimated in heights lower than 2000 m and move higher on 17 and 18 April, as Fig. 11 shows.

The specific deviations of the Tikhonov estimate are mostly spatial artifacts. Particularly, the whole upper left corner of the considered area exhibits unexpected activity on 5 April. From 11 to 13 April, there is a blurring of the central emissions with a different direction every day, followed by a cloud of emissions with the strongest activity in the upper left corner of the map on 14 and 15 April. Since the area of estimation was specifically chosen so that the whole area of the Chernobyl wildfires is contained inside, the activity certainly should not be peaking in a corner pixel. From 16 to 19 April, there is another cloud, this time closer to the lower right corner, which is clearly separated from the main emissions around CEZ.

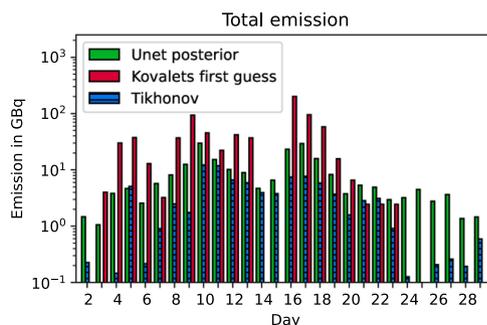


Fig. 9. Total emissions per day.

Overall, the spatial estimate provided by Unet posterior is more compact than the one estimated with Tikhonov regularization. There are fewer deviations from Kovalets first guess, and it also offers a way to assess the uncertainty of the estimate through variation of the ensemble. Nevertheless, the deviations also appear in the same area when Tikhonov regularization is used, which suggests that there may be an inaccuracy in the atmospheric model.

5.4. Total emissions

The total emissions estimated by Unet posterior is 225 GBq. Table 1 also reports the lower and upper bound on the estimate, which were chosen as the minimum and maximum total emission of the individual runs of the ensemble.

Fig. 9 depicts the time profiles and shows that the Unet posterior estimate follows the same time pattern as the Kovalets first guess. Since the first guess of Kovalets et al. [30] is based on different data, it is hard to decide which estimate of total emissions is closer to reality. The estimate of Kovalets et al. [30] obtained after performing inversion with ¹³⁷Cs dataset by Masson et al. [35] is reported in Table 1 and it can be seen that the estimate of Unet posterior fits in their confidence interval. Higher activity estimated in time periods where there should be no fires according to Kovalets first guess could be attributed to convolution Unet posterior uses in the time axis. The learned convolution kernel cannot capture both smoothness in times of high activity and sharp changes in emissions on 3–4, 13–14, 15–16, and 23–24 April.

The Tikhonov estimate follows the same profile, too, but its total emissions are even lower. This is caused by the regularization term, which pushes the estimate towards zero and, as the L-curve shows (Fig. 5), the total emissions depend strongly on the λ value.

Fig. 10 compares emissions in individual elements of the source term, showing that while the Unet posterior achieves higher values, it also estimates a higher portion of the elements to be closer to zero than Tikhonov with optimal λ value ($\lambda=0.3$). On the other hand, choosing λ so that the total emissions are approximately the same to the one predicted by U-net posterior ($\lambda=0.001$, total emissions=257 GBq), pushes the maximum value to be roughly the same as the one for U-net posterior. Nevertheless, the distribution of values shows that there would be approximately the same amount of elements with activity from 10^4 to 10^8 GBq, which would blur spatial estimate and even more strengthen the artifacts that can be seen in Fig. 7. This clearly illustrates that the Unet posterior provides some prior information useful for recovering which areas contain some emissions and which do not and helps to find compact and meaningful spatial emissions.

5.5. Particle sizes and vertical layers

Fig. 11 illustrates the emissions in each height level w.r.t. each particle size. The emissions of the three fractions estimated by the Unet

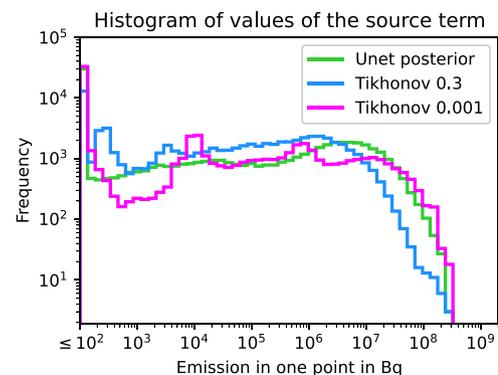


Fig. 10. Histogram of values in the 5D source term estimate.

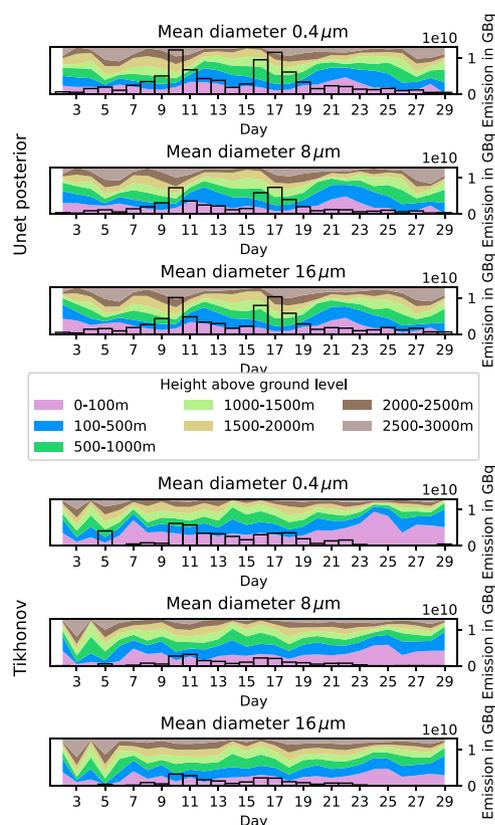


Fig. 11. Total emissions per day per fraction and height above ground level.

posterior have very similar time profile as well as their ratios in all height levels. In total, 41 % of the total emissions are made up by the smallest particles with diameter $0.4 \mu\text{m}$, 25 % by particles of diameter $8 \mu\text{m}$, and 34 % by particles with diameter $16 \mu\text{m}$. The emissions in vertical layers in the corresponding order from 0 to 100 m to 2500–3000 m make up 15 %, 16 %, 16 %, 14 %, 14 %, 12 %, and 13 % of the total emissions, while, as previously noted, the emissions at higher levels have higher uncertainty.

The Tikhonov estimated amounts of 53 % to particles of diameter $0.4 \mu\text{m}$, 23 % to $8 \mu\text{m}$, and 24 % to $16 \mu\text{m}$. The distribution among vertical layers is 21 %, 19 %, 16 %, 14 %, 12 %, 12 %, and 6 % going from 0 to 100 m to 2500–3000 m.

Both models assume that the highest amount of ^{137}Cs is carried by the smallest particles and that, overall, the amount decreases with growing height above ground level.

6. Conclusion

This paper introduced a novel inversion method called Unet posterior for spatial-temporal source term estimation using a deep learning concept known as the deep image prior. The use of convolutional neural network allowed us to represent the source as a 5D tensor containing not only spatial location and time, but also particle size and height above ground level. Combining this approach with the variational Bayes methodology leads to a method that does not require the first guess to find a compact and reasonable estimate of location of sources in a chosen area. It was applied to the problem of spatial-temporal emissions of ^{137}Cs from the Chernobyl wildfires in 2020. Compared to common Tikhonov regularization, it provides a cleaner estimate of the source without compromising the estimation of the total emissions (Unet posterior estimated 225 GBq of ^{137}Cs). Both methods located the main emissions quite well according to the first guess used by [30], and deviations appeared in similar areas, which may suggest inaccuracies in the

atmospheric model.

Environmental Implications

The source term estimation of atmospheric release is crucial for further analysis and appropriate countermeasures. A demanding situation is considered, where the material is emitted from potentially many places from the domain. This is the case of wildfires in a radiation-contaminated area as happened in 2020 in Chernobyl area. A completely new prior, which is based on deep neural networks reducing the number of unknowns, is proposed for this task. This prior does not require the first guess and provides more accurate estimates than existing methods. This novel approach can be applied to estimation of other spatial-temporal sources of hazardous materials.

CRedit authorship contribution statement

Antonie Brožová: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Conceptualization. **Václav Šmídl:** Writing – original draft, Visualization, Supervision, Methodology, Conceptualization. **Ondřej Tichý:** Writing – original draft, Supervision, Methodology, Funding acquisition, Data curation. **Nikolaos Evangelios:** Investigation, Funding acquisition, Data curation.

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

Code, data, and results are shared on github, there is a link in the manuscript.

Acknowledgments

This research has been supported by the Czech Science Foundation, grant no. GA24-10400S. Antonie Brožová was also supported by MEYS grant SGS24/141/OHK4/3T/14. FLEXPART model simulations are cross-atmospheric research infrastructure services provided by ATMO-ACCESS (EU grant agreement No 101008004). Nikolaos Evangelios was funded by the same EU grant. The computations were performed on resources provided by Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway.

References

- [1] Babacan, S.D., Molina, R., Do, M.N., Katsaggelos, A.K., 2012. Bayesian blind deconvolution with general sparse image priors. *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7–13, 2012, Proceedings, Part VI 12*. Springer, pp. 341–355.
- [2] 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, 467.
- [3] Bishop, C., 2006. *Pattern recognition and machine learning*. Springer.
- [4] Bocquet, M., Vanderbeeken, P., Farchi, A., Dumont Le Brazidec, J., Roustan, Y., 2024. Bridging classical data assimilation and optimal transport: the 3d-var case. *Nonlinear Process Geophys* 31, 335–357.
- [5] Brozova, A., Smidl, V., 2024. Avoiding undesirable solutions of deep blind image deconvolution. In: *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2024)*, SCITEPRESS.559–566.
- [6] Cassiani, M., Stohl, A., Brioude, J., 2015. Lagrangian stochastic modelling of dispersion in the convective boundary layer with skewed turbulence conditions and a vertical density gradient: formulation and implementation in the FLEXPART model. *Bound-Layer Meteorol* 154, 367–390.
- [7] Cassiani, M., Stohl, A., Oľivić, D., Seland, Ø., Bethke, I., Pisso, I., Iversen, T., 2016. The offline lagrangian particle model flexpart–noresm/cam (v1): model description and comparisons with the online noresm transport scheme and with the reference flexpart model. *Geosci Model Dev* 9, 4029–4048.

- [8] De Meutter, P., Gueibe, C., Tomas, J., denOuter, P., Apituley, A., Bruggeman, M., Camps, J., Delcloo, A., Knetsch, G.J., Roobol, L., et al., 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.
- [9] De Meutter, P., Hoffman, I., 2020. Bayesian source reconstruction of an anomalous Selenium-75 release at a nuclear research institute. *J Environ Radioact* 218, 106225.
- [10] De Meutter, P., Hoffman, I., Ungar, K., 2021. On the model uncertainties in Bayesian source reconstruction using an ensemble of weather predictions, the emission inverse modelling system FREAM v1. 0, and the Lagrangian transport and dispersion model Flexpart v9.0.2. *Geosci Model Dev* 14, 1237–1252.
- [11] Dong, X., Fang, S., Zhuang, S., Xu, Y., 2024. Inversion of critical atmospheric ¹³⁷Cs emissions following the Fukushima accident by resolving temporal formation from total deposition data. *Environmental Science & Technology*.
- [12] DumontLeBrazidec, J., Bocquet, M., Saunier, O., Roustan, Y., 2020. MCMC methods applied to the reconstruction of the autumn 2017 Ruthenium 106 atmospheric contamination source. *Atmos Environ: X* 6, 100071. <https://doi.org/10.1016/j.aeoa.2020.100071>.
- [13] Dumont Le Brazidec, J., Bocquet, M., Saunier, O., Roustan, Y., 2021. Quantification of uncertainties in the assessment of an atmospheric release source applied to the autumn 2017 ¹⁰⁶Ru event. *Atmos Chem Phys* 21, 13247–13267. <https://doi.org/10.5194/acp-21-13247-2021>.
- [14] Eckhardt, S., Pisso, I., Evangeliou, N., Zwaafink, C.G., Plach, A., McConnell, J.R., Sigl, M., Ruppel, M., Zdanowicz, C., Lim, S., et al., 2023. Revised historical Northern Hemisphere black carbon emissions based on inverse modeling of ice core records. *Nat Commun* 14, 271.
- [15] Eckhardt, S., Prata, A., Seibert, P., Stebel, K., Stohl, A., 2008. Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling. *Atmos Chem Phys* 8, 3881–3897.
- [16] Evangeliou, 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.
- [17] Evangeliou, N., Hamburger, T., Cozic, A., Balkanski, Y., Stohl, A., 2017. Inverse modeling of the Chernobyl source term using atmospheric concentration and deposition measurements. *Atmos Chem Phys* 17, 8805–8824.
- [18] Fang, S., Dong, X., Zhuang, S., Tian, Z., Chai, T., Xu, Y., Zhao, Y., Sheng, L., Ye, X., Xiong, W., 2022. Oscillation-free source term inversion of atmospheric radionuclide releases with joint model bias corrections and non-smooth competing priors. *J Hazard Mater* 440, 129806.
- [19] Fang, S., Zhuang, S., Li, X., Li, H., 2021. Automated release rate inversion and plume bias correction for atmospheric radionuclide leaks: a robust and general remediation to imperfect radionuclide transport modeling. *Sci Total Environ* 754, 142140. <https://doi.org/10.1016/j.scitotenv.2020.142140>.
- [20] Forster, C., Stohl, A., Seibert, P., 2007. Parameterization of convective transport in a Lagrangian particle dispersion model and its evaluation. *J Appl Meteorol Climatol* 46, 403–422.
- [21] Hansen, P., 1992. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev* 34, 561–580.
- [22] Hao, W., Baker, S., Lincoln, E., Hudson, S., Lee, S., Lemieux, P., 2018. Cesium emissions from laboratory fires. *J Air Waste Manag Assoc* 68, 1211–1223.
- [23] He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: 2015 IEEE International Conference on Computer Vision (ICCV), 1026–1034. [10.1109/ICCV.2015.123](https://doi.org/10.1109/ICCV.2015.123).
- [24] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The ERA5 global reanalysis. *Q J R Meteorol Soc* 146, 1999–2049.
- [25] Houweling, S., Bergamaschi, P., Chevallier, F., Heimann, M., Kaminski, T., Krol, M., Michalak, A., Patra, P., 2017. Global inverse modeling of CH₄ sources and sinks: an overview of methods. *Atmos Chem Phys* 17, 235–256.
- [26] Huo, D., Masoumzadeh, A., Kushol, R., Yang, Y.H., 2023. Blind image deconvolution using variational deep image prior. *IEEE Trans Pattern Anal Mach Intell* 45, 11472–11483.
- [27] Jia, M., Evangeliou, N., Eckhardt, S., Huang, X., Gao, J., Ding, A., Stohl, A., 2021. Black carbon emission reduction due to COVID-19 lockdown in China. *Geophys Res Lett* 48, e2021GL093243.
- [28] Kingma, D.P., 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
- [29] Kingma, D.P., Ba, J., 2015. Adam: A method for stochastic optimization. In: Bengio, Y., LeCun, Y. (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings. (<http://arxiv.org/abs/1412.6980>).
- [30] Kovalets, I., Talerko, M., Synkevych, R., Koval, S., 2022. Estimation of cs-137 emissions during wildfires and dust storm in chernobyl exclusion zone in april 2020 using ensemble iterative source inversion method. *Atmos Environ* 288, 119305.
- [31] Li, X., Sun, S., Hu, X., Huang, H., Li, H., Morino, Y., Wang, S., Yang, X., Shi, J., Fang, S., 2019. Source inversion of both long- and short-lived radionuclide releases from the Fukushima Daiichi nuclear accident using on-site gamma dose rates. *J Hazard Mater* 379, 120770.
- [32] Ling, Y., Liu, C., Shan, Q., Hei, D., Zhang, X., Shi, C., Jia, W., Yue, Q., Wang, J., 2024. Source term inversion of short-lived nuclides in complex nuclear accidents based on machine learning using off-site gamma dose rate. *J Hazard Mater* 465, 133388.
- [33] López-Tapia, S., Molina, R., Katsaggelos, A.K., 2021. Deep learning approaches to inverse problems in imaging: Past, present and future. *Digit Signal Process* 119, 103285.
- [34] Manning, A., O'Doherty, S., Jones, A., Simmonds, P., Derwent, R., 2011. Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach. *J Geophys Res: Atmos* 116.
- [35] Masson, O., Romanenko, O., Saunier, O., Kirieiev, S., Protsak, V., Laptev, G., Voitsekhovich, O., Durand, V., Coppin, F., Steinhäuser, G., et al., 2021. Europe-wide atmospheric radionuclide dispersion by unprecedented wildfires in the chernobyl exclusion zone, April 2020. *Environ Sci Technol* 55, 13834–13848.
- [36] Masson, O., Steinhäuser, G., Wershofen, H., Mieltski, J.W., Fischer, H.W., Pourcelot, L., Saunier, O., Bieringer, J., Steinkopff, T., Hýža, M., Möller, B., Nikolic, J., Dalaka, E., Dalheimer, A., de Vismes-Ott, A., Eleftheriadis, K., Forte, M., GascoLeonarte, C., Gorzkiewicz, K., Homoki, Z., Isajenko, K., Karhunen, T., Katzberger, C., Kierepko, R., KovendinéKónyi, J., Malá, H., Nikolic, J., Povinec, P.P., Rajacic, M., Ringer, W., Rulík, P., Rusconi, R., Sáfány, G., Sykora, I., Todorovic, D., Tschiersch, J., Ungar, K., Zorko, B., 2018. Potential source apportionment and meteorological conditions involved in airborne ¹³¹I detections in January/February 2017 in Europe. *Environ Sci Technol* 52, 8488–8500.
- [37] Masson, O., Steinhäuser, G., Zok, D., Saunier, O., Angelov, H., Babić, D., Bečková, V., Bieringer, J., Bruggeman, M., Burbidge, C.I., Conil, S., Dalheimer, A., De Geer, L.E., de Vismes Ott, A., Eleftheriadis, K., Estier, S., Fischer, H., Garavaglia, M.G., GascoLeonarte, C., Gorzkiewicz, K., Hainz, D., Hoffman, I., Hýža, M., Isajenko, K., Karhunen, T., Kastlander, J., Katzberger, C., Kierepko, R., Knetsch, G.J., KövendinéKónyi, J., Lecomte, M., Mieltski, J.W., Min, P., Möller, B., Nielsen, S.P., Nikolic, J., Nikolovska, L., Penev, I., Petrinc, B., Povinec, P.P., Querfeld, R., Raimondi, O., Ransby, D., Ringer, W., Romanenko, O., Rusconi, R., Saey, P.R.J., Samsonov, V., Šilobritiene, B., Simion, E., Söderström, C., Šostarić, M., Steinkopff, T., Steinmann, P., Sýkora, I., Tabachny, I., Todorovic, D., Tomankiewicz, E., Tschiersch, J., Tsihranski, R., Tzortzis, M., Ungar, K., Vidic, A., Weller, A., Wershofen, H., Zagvyai, P., Zaleska, T., Zapata García, D., Zorko, B., 2019. Airborne concentrations and chemical considerations of radioactive ruthenium from an undeclared major nuclear release in 2017. *Proc Natl Acad Sci* 116, 16750–16759.
- [38] Meirink, J., Bergamaschi, P., Krol, M., 2008. Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion. *Atmos Chem Phys* 8, 6341–6353.
- [39] Melinc, B., Zaplotnik, Ž., 2024. 3d-var data assimilation using a variational autoencoder. *Q J R Meteorol Soc* 150, 2273–2295.
- [40] Pison, I., Bousquet, P., Chevallier, F., Szopa, S., Hauglustaine, D., 2009. Multi-species inversion of CH₄, CO and H₂ emissions from surface measurements. *Atmos Chem Phys* 9, 5281–5297.
- [41] Pisso, I., Sollum, E., Grythe, H., Kristiansen, N.I., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R.L., Groot Zwaafink, C.D., Evangeliou, N., Sodemann, H., Haimberger, L., Henne, S., Brunner, D., Burkhardt, J.F., Fouilloux, A., Brioude, J., Philipp, A., Seibert, P., Stohl, A., 2019. The Lagrangian particle dispersion model FLEXPART version 10.4. *Geosci Model Dev* 12, 4955–4997. <https://doi.org/10.5194/gmd-12-4955-2019>. (<https://gmd.copernicus.org/article/12/4955/2019/>).
- [42] Polikar, R., 2012. Ensemble learning. *Ensemble Mach Learn: Methods Appl* 1–34.
- [43] Protsak, V., Voitsekhovich, O., Laptev, G., 2020. Estimation of radioactive source term dynamics for atmospheric transport during wildfires in Chernobyl zone in Spring 2020. *Ukrainian hydrometeorological institute (in Ukrainian)*.
- [44] Rémy, S., Veira, A., Paugam, R., Sofiev, M., Kaiser, J., Marenco, F., Burton, S., Benedetti, A., Engelen, R., Ferrare, R., et al., 2017. Two global data sets of daily fire emission injection heights since 2003. *Atmos Chem Phys* 17, 2921–2942.
- [45] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18. Springer, pp. 234–241.
- [46] Seibert, P., Frank, A., 2004. Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. *Atmos Chem Phys* 4, 51–63.
- [47] Shi, Z., Mettes, P., Maji, S., Snoek, C.G., 2022. On measuring and controlling the spectral bias of the deep image prior. *Int J Comput Vis* 130, 885–908.
- [48] 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.
- [49] Talerko, M., Kovalets, I., Lev, T., Igarashi, Y., Romanenko, O., 2021. Simulation study of radionuclide atmospheric transport after wildland fires in the Chernobyl exclusion zone in April 2020. *Atmos Pollut Res* 12, 193–204.
- [50] Talerko, M., Lev, T., Kashpur, V., 2021. Estimation of the contribution of dust storm on April 16, 2020 to radioactive contamination of the atmosphere during forest fires in the Exclusion zone. *Yad Energ to Dovkyllyia (Ukrainian)* 81–95.
- [51] Thompson, R., Sasakawa, M., Machida, T., Aalto, T., Worthly, D., Lavric, J., LundMyhre, C., Stohl, A., 2017. Methane fluxes in the high northern latitudes for 2005–2013 estimated using a bayesian atmospheric inversion. *Atmos Chem Phys* 17, 3553–3572.
- [52] Tichý, O., Šmídl, V., Evangeliou, N., 2022. Source term determination with elastic plume bias correction. *J Hazard Mater* 425, 127776. <https://doi.org/10.1016/j.jhazmat.2021.127776>.
- [53] Tichý, O., Šmídl, V., Hofman, R., Evangeliou, N., 2018. Source term estimation of multi-specie atmospheric release of radiation from gamma dose rates. *Q J R Meteorol Soc* 144, 2781–2797. <https://doi.org/10.1002/qj.3403>.

- [54] Tichý, O., Šmídl, V., Hofman, R., Stohl, A., 2016. LS-APC v1.0: a tuning-free method for the linear inverse problem and its application to source-term determination. *Geosci Model Dev* 9, 4297–4311.
- [55] Tichý, O., Ulrych, L., Šmídl, V., Evangeliou, N., Stohl, A., 2020. On the tuning of atmospheric inverse methods: comparisons with the European Tracer Experiment (ETEX) and Chernobyl datasets using the atmospheric transport model FLEXPART. *Geosci Model Dev* 13, 5917–5934. <https://doi.org/10.5194/gmd-13-5917-2020>.
- [56] Tzikas, D.G., Likas, A.C., Galatsanos, N.P., 2009. Variational bayesian sparse kernel-based blind image deconvolution with student's-t priors. *IEEE Trans Image Process* 18, 753–764.
- [57] Ulyanov, D., Vedaldi, A., Lempitsky, V., 2018. Deep image prior, In: Proceedings of the IEEE conference on computer vision and pattern recognition, 9446–9454.
- [58] Van Leuven, S., De Meutter, P., Camps, J., Termonia, P., Delcloo, A., 2025. Source reconstruction via deposition measurements of an undeclared radiological atmospheric release. *EGUsphere* 2025, 1–29.
- [59] Vojta, M., Plach, A., Thompson, R., Stohl, A., 2022. A comprehensive evaluation of the use of Lagrangian particle dispersion models for inverse modeling of greenhouse gas emissions. *Geosci Model Dev* 15, 8295–8323.
- [60] Wittig, S., Berchet, A., Pison, I., Saunois, M., Thanwerdas, J., Martinez, A., Paris, J. D., Machida, T., Sasakawa, M., Worthy, D.E., et al., 2023. Estimating methane emissions in the Arctic nations using surface observations from 2008 to 2019. *Atmos Chem Phys* 23, 6457–6485.
- [61] Xu, Y., Fang, S., Dong, X., Zhuang, S., 2024. A spatiotemporally separated framework for reconstructing the sources of atmospheric radionuclide releases. *Geosci Model Dev* 17, 4961–4982.