



Source term inversion of short-lived nuclides in complex nuclear accidents based on machine learning using off-site gamma dose rate

Yongsheng Ling^{a,b,c}, Chengfeng Liu^a, Qing Shan^a, Daqian Hei^d, Xiaojun Zhang^e, Chao Shi^a, Wenbao Jia^a, Qi Yue^f, Jing Wang^{c,g,*}

^a Department of Nuclear Science and Technology, Nanjing University of Aeronautics and Astronautics, 211106 Nanjing, China

^b Collaborative Innovation Center of Radiation Medicine of Jiangsu Higher Education Institutions, 215021 Suzhou, China

^c Institute of Environmental Engineering, ETH Zürich, Zürich CH-8093, Switzerland

^d School of Nuclear Science and Technology, Lanzhou University, 730000 Lanzhou, China

^e Suzhou Guanrui Information Technology Co., Ltd, 215123 Suzhou, China

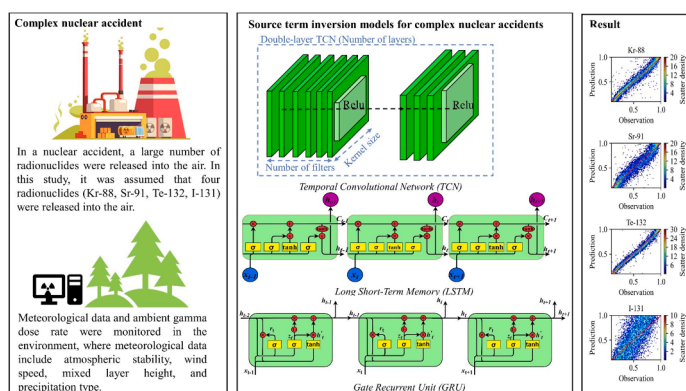
^f China Institute of Radiation Protection, 030006 Taiyuan, China

^g Lab of advanced analytical technologies, Empa, Dübendorf CH-8600, Switzerland

HIGHLIGHTS

- Machine learning was valuable for source term inversion in complex nuclear accidents.
- LSTM was preferable for feature extraction of time-series gamma dose rate.
- Ambient gamma dose rate was a key parameter for source term inversion.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Joao Pinto da Costa

Keywords:

Complex nuclear accident
Source-term inversion
Machine learning
Gamma dose rate

ABSTRACT

During nuclear accidents, large amounts of short-lived radionuclides are released into the environment, causing acute health hazards to local populations. Therefore, it is particularly important to obtain source-term information to assist nuclear emergency decision makers in determining emergency protective measures. However, it is extremely difficult to obtain reliable contaminant monitoring instrument readings to estimate the source term based on core conditions, release routes, and release conditions. Currently, a wide variety of source-term inversion methods are attracting increasing attention. In this study, the release rates of four typical short-lived nuclides (Kr-88, Sr-91, Te-132, I-131) in two complex nuclear accident scenarios were estimated using a machine-learning method. The results show that the best estimation performance is obtained with the long short-term memory network, and the mean absolute percentage errors for the release rates of the four nuclides at 10 h under the two nuclear accidents are 9.87% and 11.08%, 17.49% and 16.51%, 7.16% and 8.35%, and 38.83% and

* Correspondence to: ETH Zürich, Zürich CH-8093, Switzerland.

E-mail address: jing.wang@ifu.baug.ethz.ch (J. Wang).

<https://doi.org/10.1016/j.jhazmat.2023.133388>

Received 17 October 2023; Received in revised form 22 December 2023; Accepted 26 December 2023

Available online 28 December 2023

0304-3894/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

41.87%, respectively. Meanwhile, the mean absolute percentage errors for Te-132 (7.16% and 8.35%) were the lowest among all the estimated nuclides. In addition, stability analysis showed that the gamma dose rate was the key parameter affecting the estimation accuracy.

1. Introduction

On 11 March 2011, a strong earthquake and tsunami struck the coast of Japan, causing a serious accident at the Fukushima Daiichi Nuclear Power Plant. Large amounts of radionuclides were released into the environment [19]. These radionuclides include both short- and long-lived radionuclides, with short-lived radionuclides decaying rapidly during atmospheric dispersion and surface deposition [30] and posing a serious health hazard to local populations [38]. Therefore, it is particularly important to assess the source term (i.e. nuclide release rate) of short-lived radionuclides quickly and accurately to assist nuclear emergency decision-making authorities in developing emergency plans [21,34].

Generally, two methods are used for source term inversion: one is based on monitoring instrument data inside a nuclear power plant [2, 15]) and the other is based on environmental monitoring data outside the nuclear power plant [12,13,9]. However, in severe nuclear accidents, the monitoring instruments of power plants are often destroyed, making it impossible to obtain operational data from these plants [20]. Therefore, an increasing number of studies have focused on source-term inversion to assess the type and release rate of nuclides from environmental monitoring data outside nuclear power plants. Many studies have been conducted on source-term inversion based on meteorological parameters and air concentrations of radionuclides ([8,10,14,39,44, 45]). The concentration of radionuclides in the air was obtained by analysing samples from the environment, which represented the average value over a period. However, during nuclear accident emergencies, it is difficult to obtain radionuclide concentrations in real time, leading to difficulties in source-term inversion [36]. In contrast, the gamma dose rate, which can be measured in real time, may be more suitable for source-term inversion. Methods for source-term inversion based on the gamma dose rate include the source-receptor equation [36], Kalman filtering [24], least squares [43], and data assimilation [23]. However, the accuracy of these methods depends heavily on the quality and reliability of a priori information. Owing to the advantages of artificial neural networks with the ability of adaptive learning without setting a priori information, they are widely used in environmental radioactivity studies, such as radiation spectrum analysis [22,33], radionuclide concentration assessment [5,31], and dose distribution of radionuclides [35,41]. In our previous study, the estimation of the nuclide release rate for single nuclides [28] and multi-nuclides [25] and the classification of release classes [42] were achieved using back-propagation neural networks. To assess the release rate of multi-nuclides more accurately, researchers have used a time-serialised approach and applied recurrent neural networks [29] and temporal convolutional networks [26,27] with time-series analysis capabilities.

Although the above-mentioned machine learning-based inversion methods have obtained better results for estimating the release rate of multiple nuclides, they were conducted on a simplified accident scenario, characterised by unchanged meteorological parameters and constant release, which may lead to insufficient applicability of the methods. In real situations, changes in meteorological parameters and the two-stage release of nuclides may pose challenges and difficulties in source-term inversion. Changes in meteorological parameters may affect the transport and diffusion processes of nuclides, leading to nonlinear changes in the release patterns. In addition, the gap release of nuclides can lead to a coupling effect between the initial and subsequent release processes, complicating the source term inversion.

Two nuclear accident scenarios were investigated to study source-term inversion in complex nuclear accident scenarios. 1) Intermittent

changes in meteorological parameters during off-site environmental monitoring, because meteorological parameters have a significant impact on the diffusion and transport of radionuclides into the atmosphere. Intermittent changes in meteorological parameters (wind speed, wind direction, atmospheric stability, mixed layer height, and precipitation) were simulated to reflect the dynamic changes in a real nuclear accident. 2) Setting up a two-stage release of radionuclides because radioactive emissions during severe nuclear accidents may appear as a multi-stage release, such as the Fukushima accident [16]. In this case, a two-stage release of radionuclides was simulated. The second release started two hours after the end of the first release. These studies can simulate and assess the radionuclide release in complex nuclear accident scenarios. Our approach helps to explore the adaptability of machine-learning-based inversion methods in complex accident scenarios and to enhance the applicability in emergency scenarios that are closer to real scenarios.

2. Methodology

2.1. Nuclear accident data

In this study, two nuclear accident scenarios were simulated using the International Radiological Assessment System (InterRAS). InterRAS is a set of personal computer-based tools. InterRAS, as indicated by its name, is an international version of RASCAL [1,37] developed by U.S. NRC, which is introduced by IAEA for emergency response in eastern European countries. InterRAS contains tools to estimate the distance that urgent protective actions may be needed based on nuclear power plant conditions or release rates (ST-DOSE), to estimate early and long-term dose from field measurements of radionuclide concentrations (FM-DOSE), and to compute decay of radionuclides (DECAY). In this study, ST-DOSE is used to simulate the diffusion of radionuclides into the atmosphere and to calculate the gamma dose from the air to the ground. The source terms and auxiliary data were input into InterRAS to obtain the gamma dose rate at five downwind monitoring sites. A script written in python 3.7 was used to generate 30,000 sets of data on a Win7 system, of which 24,000 sets were used for training the neural network and the remaining 6000 sets were used for testing the neural network.

2.1.1. Source term

In this study, four short-lived nuclides, Kr-88, Sr-91, Te-132, and I-131, which are typical nuclear reactor fission products that can provide critical nuclear reactor accident information, were selected as mixed-source terms. Kr-88, Sr-91, and Te-132, as typical short-lived nuclides, decay rapidly after being released into the environment. Therefore, they can be used to confirm whether a nuclear accident has occurred and to assess its severity. I-131 can be released into the environment in large quantities during a nuclear accident and has serious public health effects; therefore, determining the release quantity of I-131 can help assess the post-accident radiological risk. The release rates of the four nuclides were referenced from the list of fission products from IAEA-TECDOC-955 [18], and the radioactive properties and release rates of the four fission products are listed in Table 1. The reactor inventory in the penultimate column of Table 1 was based on a light water reactor fission products inventory (per 1000 MWe) after 30 min from shutdown.

2.1.2. Auxiliary data

The auxiliary data included the nuclear accident release height and meteorological parameters. The meteorological parameters that affect the diffusion of nuclides in the atmosphere must be set in InterRAS. The

Table 1
Radioactive properties of the four short-lived radionuclides.

Radio-nuclides	Half-life	Air immersion dose rate conversion factor (Sv s^{-1} per Bq m^{-3})	Ground shine dose rate conversion factor (Sv s^{-1} per Bq m^{-2})	Reactor inventory Bq/ (1000 MW)	Order of magnitude of release rate (Bq/h)
Kr-88	2.8 h	1.18×10^{-15}	-	2.52×10^{18}	10^{18}
Sr-91	9.63 h	3.20×10^{-14}	6.66×10^{-16}	4.07×10^{18}	10^{18}
Te-132	3.3 d	5.50×10^{-14}	1.10×10^{-15}	4.44×10^{18}	10^{18}
I-131	8.1 d	1.82×10^{-14}	3.76×10^{-16}	3.15×10^{18}	10^{18}

Table 2
The value and physical meaning of auxiliary data during the nuclear accident.

Auxiliary data	Value range	Physical meaning
Release height	0 - 60 (m)	The release height affects the maximum extent of nuclides dispersion, and the wind speed varies with the height.
Wind speed	1 - 10 (m/s)	Wind speed directly affects the diffusion rate of radionuclides.
Wind direction	0 - 360	Wind direction affects the direction of movement of the radioactive plume.
Atmospheric stability	A - G	Pasquill atmospheric stability category. The tendency and degree to which an air mass returns to or leaves its original equilibrium position after a perturbation in the vertical direction. A-G indicates a state ranging from extremely unstable to extremely stable.
Mixed layer height	307 - 1030 (m)	Mixed layer height affects the diffusion of nuclides in the vertical direction.
Precipitation	None Light Rain Moderate Rain Heavy Rain	None Light Rain (rainfall rate < 25 mm/h) Moderate Rain (rainfall rate between 25 and 75 mm/h) Heavy Rain (rainfall rate > 75 mm/h)

Table 3
The variations of meteorological parameters during the nuclear accident.

Meteorological parameters	Variation
Wind speed	The wind speed at the later time step is equal to the wind speed at the earlier time step minus 1 or plus 1 or unchanged.
Wind direction	The wind direction at the later time step is equal to the wind direction at the earlier time step minus 10 or plus 10 or unchanged.
Atmospheric stability	Atmospheric stability is determined based on the wind speed according to Pasquill's atmospheric stability classification.
Mixed layer height	When the mixed layer height at the earlier time step is greater than 669 m (average height), the value of the mixed layer height at the later time step is reduced by 50 m; when the value of mixed layer height at the earlier time step is less than or equal to 669 m, the value of mixed layer height at the later time step is increased by 50 m.
Precipitation	Changes are made in the following order: None - Light Rain - Moderate Rain - Heavy Rain - None

meteorological parameters included wind speed [4], wind direction, atmospheric stability [40], mixed-layer height [11], and precipitation. The ranges of the values and physical meanings of the auxiliary data are listed in Table 2.

2.1.3. Complex scenario 1: changing meteorological parameters

To study the inversion of the source term of the nuclear accident with

changing meteorological parameters, the meteorological parameters were changed at 2.5 and 6.5 h during the 10 h of the nuclear accident simulation. The initial auxiliary data were randomly generated within the range of values listed in Table 2 and the meteorological parameters were subsequently changed according to Table 3. Fig. 1 shows an example dataset. Five meteorological parameters were changed at hours 2.5 and 6.5, respectively. The release rates of the four radionuclides were randomly generated between 10^{18} Bq/h and 10^{19} Bq/h, but not including 10^{19} Bq/h. The four radionuclides were continuously released for 30 min at the 0th hour. Gamma dose rates were measured at 30-minute intervals over a 10-hour period at each of the five monitoring sites downwind, for a total of 20 gamma dose rate time steps. An example of this dataset is presented in Table 4.

2.1.4. Complex scenario 2: two-stage release of radionuclides

During the 10 h of the nuclear accident simulation, the four radionuclides were released for 30 min at the 0th hour and 2nd hour, with the same release rate. Five meteorological parameters were randomly generated based on Table 2 and kept constant. Gamma dose rates were measured in the same way as for complex scenario 1. An example of this dataset is presented in Table 5.

2.1.5. Simple scenario: constant meteorological parameters and release rates

In order to demonstrate the adaptation of machine learning for source inversion in complex nuclear accident scenarios, a simple nuclear accident scenario was set up as a comparison experiment. In the simple nuclear accident scenario, the simulation maintained constant meteorological parameters. Radionuclides were released only once, with the start time at the 0th hour and a duration of 0.5 h. With the exception of varying meteorological parameters and nuclide release patterns, all other data processing in the simple scenario were consistent with the two complex scenarios described above.

2.2. Source term inversion models for complex nuclear accidents

In this study, three models with time-series analysis capability were used to construct inversion models of complex nuclear accident source terms: the Temporal Convolutional Network (TCN), Long Short-Term Memory (LSTM), and Gate Recurrent Unit (GRU) models. The three models shared the same input data, including gamma dose rates at 20 time steps under five monitoring sites and six auxiliary data. Their common task was to assess the release rates of the four radionuclides. Since each model had different capabilities in feature extraction, this resulted in differences in the specific values of their output data.

In simulating nuclear accidents, the duration of release was set at only 0.5 h instead of 10 h. Consequently, this study was designed to assess the release rate of 0.5 h using the gamma dose rate at 20 time steps when the release duration is assumed to be known. Due to the special data format requirements of the machine learning models, the output data must have the same time dimension as the input data. Therefore, the release rates for the four nuclides were also extended to 20 time steps, but it does not mean that in the dose rate calculation process, the release duration is 10 h. There were differences in the output of the three nuclear accident scenarios. For the complex scenario 1 and simple scenario, the release rate of each nuclide remained constant over the 20 time steps. For the complex scenario 2, two-stage release rates were included in the output of the 20 time steps. The time steps from 0.5 h to 2 h include only the release rate of the first stage, while the time steps after 2.5 h include the release rates of the first and second stages. In addition, the released time of the nuclide and the gamma dose rate monitored time were simultaneous, and the duration of the release was set to 0.5 h, so that there was no need to calculate the start time and the duration of the release. In summary, the studies in this paper were primarily based on nuclide release rates during gamma dose rate monitoring, including one or more release periods.

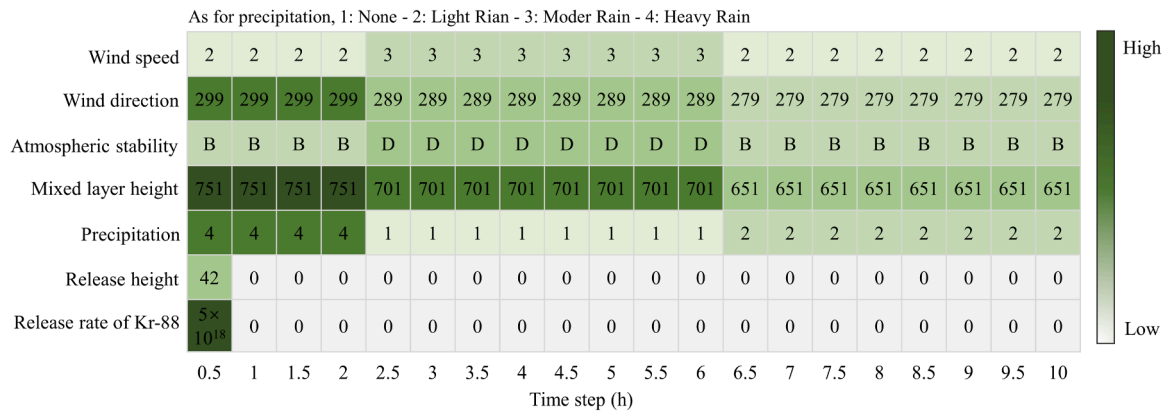


Fig. 1. An example dataset of meteorological parameters over time.

Table 4

An example dataset under changing meteorological parameters.

Source term		Time step (h)	Gamma dose rate (mSv/h)				
Nuclide	Release rate (Bq/h)		1 km	2 km	5 km	25 km	50 km
Kr-88	1.98×10^{18}	0.5	428	211	89.6	0.12	0
Sr-91	3.11×10^{18}	1	72	42	12.4	15.28	0
Te-132	4.95×10^{18}	1.5	90	47	15	2.6	3.15
I-131	3.57×10^{18}	2	100	60	17	1.8	1.65
		2.5	110	50	18	1.9	0.7
		3	110	70	19	2.1	0.6
		3.5	120	60	19	2.2	0.7
		4	120	70	20	2	0.7
		4.5	130	60	20	2	0.7
		5	90	70	20	3	0.7
		5.5	200	70	20	2	0.7
		6	100	70	30	2	0.8
		6.5	100	80	20	3	0.7
		7	200	70	20	2	0.8
		7.5	100	70	20	2	0.8
		8	100	70	20	3	0.3
		8.5	200	110	30	2	1
		9	100	0	20	3	1
		9.5	100	100	20	2	1
		10	200	100	20	2	0

2.2.1. TCN model

TCN [3] is a CNN-based sequence model that uses layer-wise convolutional operations to learn the features of temporal data. Specifically, a TCN comprises a series of convolutional layers, each of which contains convolutional kernels of different sizes to perform convolutional operations on the input sequence. Moreover, each convolutional layer is followed by an optional normalisation layer and an activation function, allowing the network to train and learn better.

In addition, the TCN employs residual blocks and a causal convolution to enhance the performance and stability of the network. The residual blocks allow the network to maintain deeper layers during training, thereby improving the expressiveness of the model and mitigating the gradient disappearance problem. In contrast, causal convolution allows the model to make assessment only from past data, avoiding information leakage and enhancing the robustness of the model.

As shown in Fig. 2, the inversion model of the nuclear accident source term based on TCN contains an input layer, two TCN layers, and a fully connected layer (FC). The input layer contained 24,000 input datasets, each containing the gamma dose rates for the five monitoring points at 20-time steps, where the gamma dose rates were normalised using Eq. (1).

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (1)$$

Table 5

An example dataset under two-stage release of radionuclides.

Source term		Time step (h)	Gamma dose rate (mSv/h)				
Nuclide	Release rate (Bq/h)		1 km	2 km	5 km	25 km	50 km
Kr-88	5.97×10^{18}	0.5	7800	2530	313	0	0
Sr-91	5.59×10^{18}	1	8500	2900	310	0	0
Te-132	9.89×10^{18}	1.5	10,000	3400	350	0	0
I-131	7.69×10^{18}	2	12,000	3600	410	0	0
Auxiliary data(Constant)		2.5	19,800	6530	713	2.64	0
		3	20,500	6900	810	10.73	0
Release height (m)	42	3.5	23,000	8400	750	4.09	0
Wind speed (m/s)	2	4	26,000	7600	910	0.09	0
Wind direction	299	4.5	24,000	9000	900	2.72	0.013
Atmospheric stability	B	5	32,000	8000	1000	10.82	0.307
Mixed layer height (m)	751	5.5	23,000	10,000	900	4.18	1.18
Precipitation	Heavy Rain	6	24,000	9000	1000	0.18	1.6
		6.5	32,000	10,000	1000	0.17	0.813
		7	30,000	9000	1000	0.2	0.707
		7.5	30,000	9000	1000	0.19	1.18
		8	20,000	10,000	1000	0.19	1.7
		8.5	40,000	10,000	1100	0.19	0.8
		9	20,000	10,000	1000	0.11	0.4
		9.5	40,000	9000	1000	0.2	0
		10	20,000	10,000	1000	0.2	0.1

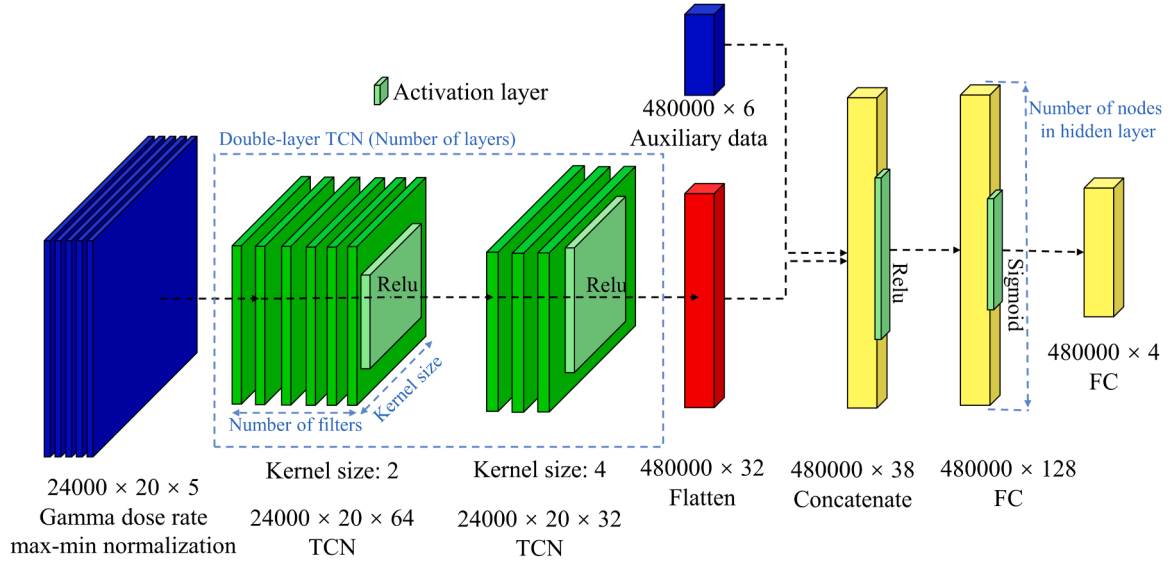


Fig. 2. Inversion model of nuclear accident source term based on TCN.

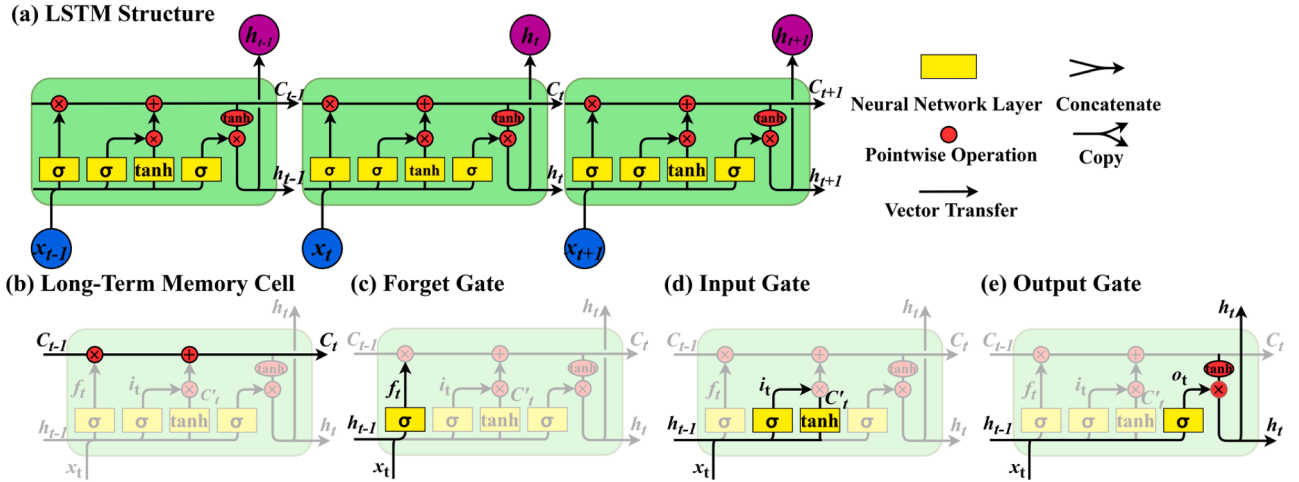


Fig. 3. Inversion model of nuclear accident source term based on LSTM.

where, y' is the normalized gamma dose rate, y is the raw gamma dose rate, y_{\min} is the minimum gamma dose rate, y_{\max} is the maximum gamma dose rate. For the normalization calculation process, the data were normalized for the 30,000 data sets, not only for a data set containing 20 time steps.

The model first extracts features through the first layer of the TCN (64 filters with a convolutional kernel size of 2) and then further extracts features through the second layer of the TCN (32 filters with a convolutional kernel size of 4). The extracted features were flattened, concatenated with six auxiliary data points, and fed into an FC layer containing 128 hidden nodes. Finally, the release rates of the four nuclides were assessed using an FC layer containing four hidden nodes.

2.2.2. LSTM model

The LSTM [17] is a neural network model commonly used for time-series data. This solves the problems of gradient disappearance and gradient explosion of traditional recurrent neural networks (RNN) and improves the memory capability of RNN.

The basic structure of LSTM is a series of gates, each of which is a nonlinear function that can be learned. The LSTM model has three gates: forget gate, input gate, output gate, and a large long-term memory cell.

The calculation process is illustrated in Fig. 3. Fig. 3(a) shows the LSTM structure at three time steps: x_t denotes the gamma dose rate at time step t , σ denotes the sigmoid activation function, \tanh denotes the hyperbolic tangent function, h_t denotes the hidden state at time step t , and C_t denotes the cell state at time step t .

Fig. 3(b) shows the cell state, which is calculated as follows:

$$C_t = f_t \times C_{t-1} + i_t \times C'_t \quad (2)$$

where, f_t is the forget gate, indicating which features of C_{t-1} are used to calculate C_t . i_t denotes which features of C'_t are used to update C_t .

Fig. 3(c) shows the forget gate, which is calculated as follows:

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f) \quad (3)$$

where W_f and b_f denote the weight and threshold of the forget gate, respectively. The values are mapped to the interval [0,1] by the sigmoid activation function, and if f_t is zero, the data are forgotten.

Fig. 3(d) shows the input gate, which is calculated as follows:

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \quad (4)$$

$$C'_t = \tanh(W_C \bullet [h_{t-1}, x_t] + b_C) \quad (5)$$

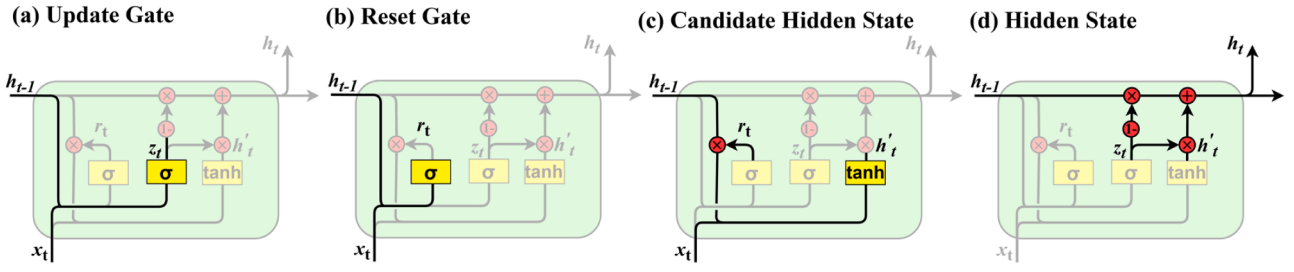


Fig. 4. Inversion model of nuclear accident source term based on GRU.

where W_i and b_i denote the weight and threshold of the input gate, respectively. i_t denotes the input data.

Fig. 3(e) shows the output gate, which is calculated as follows:

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \times \tanh(C_t) \quad (7)$$

where W_o and b_o denote the weight and threshold of the output gate, respectively. o_t denotes the output data.

2.2.3. GRU model

GRU [7] is similar to LSTM in that both are variants of the RNN. Compared with the LSTM model, the GRU model has only two gates, thus reducing the complexity of the model. A standard GRU unit has two gates: an update gate and a reset gate. These two gates determine the weight ratio between the input at the current moment and the output at the previous moment, thereby determining the output of the GRU at the current moment. The specific calculation process is shown in Fig. 4, with Fig. 4(a) showing the update gate calculation:

$$z_t = \sigma(W_z \bullet [h_{t-1}, x_t]) \quad (8)$$

Fig. 4(b) shows the reset gate, which is calculated as follows:

$$r_t = \sigma(W_r \bullet [h_{t-1}, x_t]) \quad (9)$$

Fig. 4(c) shows the candidate hidden state, which is calculated as follows:

$$h'_t = \tanh(W \bullet [r_t \times h_{t-1}, x_t]) \quad (10)$$

Fig. 4(d) shows the hidden state, which is calculated as follows:

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times h'_t \quad (11)$$

where z_t denotes the update gate, r_t is the reset state, h_t is the hidden state.

As shown in Eqs. (8)–(11), the GRU model first calculates the update gate and reset gate based on the hidden state at the previous moment and the input data at the current moment. Subsequently, the hidden state at the current moment is calculated based on the update and reset gates.

2.3. Evaluation metrics

Two commonly used evaluation metrics were used to fully evaluate the assessment performance of the model: the mean square error (MSE) and mean absolute percentage error (MAPE). The MSE was set as the loss function of the neural networks and was designed to assess the model's error throughout the optimization process, providing a quantitative measure of the model's fitting performance with the test dataset. MAPE was used to measure the estimative performance of a model in a test dataset. It shows the relative deviation of the assessed value from the observed value; a smaller value indicates that the model assesses the results more accurately. The formulas for calculating the two evaluation metrics are shown in Eqs. (12) and (13).

$$Loss = MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i')^2 \quad (12)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i'}{y_i'} \right| \quad (13)$$

where, n denotes 120,000 release rates that include 6000 test sets, each containing 20 time steps of release rates. y_i' denotes the real value of the normalized release rate, normalized by using Eq. (1). \hat{y}_i denotes the assessment of the normalized release rate.

2.4. Optimisation method

Optuna is an open-source Python software for hyperparametric automatic optimisation [32]. It uses a Bayesian optimisation algorithm and sample parallelisation to minimise the validation error of the neural network, thus improving the accuracy and convergence speed of the model. One advantage of Optuna is that it can quickly perform large-scale hyperparametric searches. In addition, it can be integrated with multiple deep learning frameworks such as TensorFlow, Keras, and PyTorch, which makes the model-tuning process more automated and efficient. In this study, the loss value of the test dataset was defined as the objective function when Optuna was used to optimise the hyperparameters.

2.5. Optimiser

The optimiser is an important part of a deep learning model that determines how the parameters are updated using a gradient descent algorithm. In deep learning, the model usually needs to optimise thousands, tens of thousands, or even more parameters, which directly affect its performance, generalisation ability, and accuracy. Therefore, it is crucial to select an appropriate optimiser that can effectively accelerate the training of the model and improve its performance.

In Keras [6], a deep learning framework, provides optimisers including SGD, RMSprop, Adam, Adadelta, Adagrad, Adamax, and Nadam. Each of the seven optimisers had different hyperparameters. To determine the optimal hyperparameter configuration for each optimiser, the Optuna optimisation method was used to optimise each of the seven optimisers, in which the search space of the hyperparameters was set based on the default values of Keras. The hyperparameters of the seven optimisers are listed in Table 6.

3. Experimental result

3.1. Neural network optimisation

To maximise the estimative performance of the neural network, some crucial parameters must be tuned during the training process of the neural network, including the optimiser and neural network structure.

Table 6

The hyperparameters of the seven optimisers.

Optimiser	Hyperparameter	Usage	Search space
SGD	Learning Rate	Control the step size of parameter update.	$(10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Momentum	Mitigate the fluctuation of gradient update direction.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
	Nesterov	Reduce the impact of gradient oscillation.	(True, False)
RMSprop	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Rho	Reduce the fluctuation of instantaneous gradient value.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
	Momentum	Mitigate the fluctuation of gradient update direction.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
Adam	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Beta1	Adapt to tasks with large gradient variations and improve the overall effectiveness of the model while maintaining gradient variations.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
	Beta2	Adapt to tasks with large gradient variations and high dimensionality, and improve the overall effectiveness of the model while maintaining gradient variations and high dimensionality.	(0.90, 0.92, 0.94, 0.96, 0.98, 1.0)
Adadelta	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Rho	Reduce the fluctuation of instantaneous gradient value.	(0.90, 0.92, 0.94, 0.96, 0.98, 1.0)
Adagrad	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Initial Accumulator Value	Initialise the learning rate and reduce the variance of parameter updates.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
Adamax	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Beta1	Adapt to tasks with large gradient variations and improve the overall effectiveness of the model while maintaining gradient variations.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
	Beta2	Adapt to tasks with large gradient variations and high dimensionality, and improve the overall effectiveness of the model while maintaining gradient variations and high dimensionality.	(0.90, 0.92, 0.94, 0.96, 0.98, 1.0)
Nadam	Learning Rate	Control the step size of parameter update.	$(10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1})$
	Beta1	Adapt to tasks with large gradient variations and improve the overall effectiveness of the model while maintaining gradient variations.	(0.0, 0.2, 0.4, 0.6, 0.8, 1.0)
	Beta2	Adapt to tasks with large gradient variations and high dimensionality, and improve the overall effectiveness of the model while maintaining gradient variations and high dimensionality.	(0.90, 0.92, 0.94, 0.96, 0.98, 1.0)

3.1.1. Optimiser

To save computational resources, a simple TCN model was constructed to search for the optimal hyperparameter configuration of the optimiser. The model structure was set up as a single-layer TCN containing six convolutional kernels of size 4 and a hidden layer containing 32 hidden nodes.

Using the SGD optimiser as an example, 20 searches were conducted during the hyperparameter search to ensure that each hyperparameter value in the search space was fetched. The search results were shown in Fig. 5(a) (results with large loss values for the individual test datasets were excluded to reduce bias). When the learning rate was 10^{-1} , momentum was 0.8, and Nesterov was True, the test dataset had the smallest loss value of 0.0360. In addition, when the learning rate was set to 10^{-4} , the loss values of the test datasets were all greater than 0.07, indicating that the learning rate had an important impact on the assessment performance of the model.

To avoid the influence of momentum and Nesterov on the experimental results and to determine the optimal learning rate, a univariate method was used to compare the loss values of the test dataset at different learning rates by selecting trials in which both momentum and Nesterov were equal in the search results. The comparison results were shown in Fig. 5(b) for four comparison trials. In the first comparison trial, the learning rate was ranked as 10^{-2} , 10^{-3} , and 10^{-4} when the momentum was 0.8, and Nesterov was False. In the second comparison trial, the learning rate was ranked as 10^{-1} , 10^{-2} when the momentum was 0.2, and Nesterov was True. In the third comparison trial, the learning rate was ranked as 10^{-1} , 10^{-3} and 10^{-4} when the momentum was 0.2, and Nesterov was False. In the fourth comparison trial, the learning rates were ranked as 10^{-2} and 10^{-4} when the momentum was zero and Nesterov was False. Combining the results of the four comparisons, the best learning rates were ranked as 10^{-1} , 10^{-2} , 10^{-3} , and 10^{-4} . Therefore, the optimal learning rate for the SGD was 10^{-1} .

As shown in Fig. 5(c), based on the above identified optimised learning rate of 10^{-1} , two trials were set up to study the optimal value of momentum by fixing Nesterov to False and True and changing momentum from 0 to 1.0. It could be observed that Nesterov had little effect on the model performance, and the best value of momentum was 0.8 when the optimiser was SGD. The other optimisers were used to search for hyperparameters in the same manner as in SGD, and the results were listed in Table 7. As shown in Table 7, the Nadam optimiser had the smallest loss value for the test dataset; therefore, it was selected as the best optimiser.

3.1.2. Structure optimisation of models under the complex scenario 1

To maximise the performance of neural networks, several network structure hyperparameters must be optimised, including the number of layers, number of filters per layer, kernel size for the TCN model, and number of layers and number of nodes per hidden layer for the LSTM and GRU models. Because the nuclear accident source term inversion model consists of each of these three models combined with a fully connected layer, the number of nodes in the hidden layer of the fully connected layer must also be optimised. The search spaces involved in these hyperparameters were listed in Table 8.

The optimisation results of the LSTM model obtained using the Optuna algorithm were shown in Figs. 6 and 7. For the single-layer LSTM model, the smallest loss value of 0.0242 could be achieved under the number of hidden nodes of 49, the number of hidden nodes in the fully connected layer of 128, and the learning rate of 10^{-3} . While, for the double-layer LSTM model, the smallest loss value of 0.0245 could be achieved under the number of hidden nodes of 27 and 61, the number of hidden nodes in the fully connected layer of 234, and the learning rate of 10^{-3} . In addition, Figs. 6–7 showed that the scatter points were roughly distributed around the optimal point. This indicated that the Optuna optimisation algorithm gradually approached the optimal point and concentrated most of the search resources around the optimal point, indicating high efficiency and stability. This stability could make the

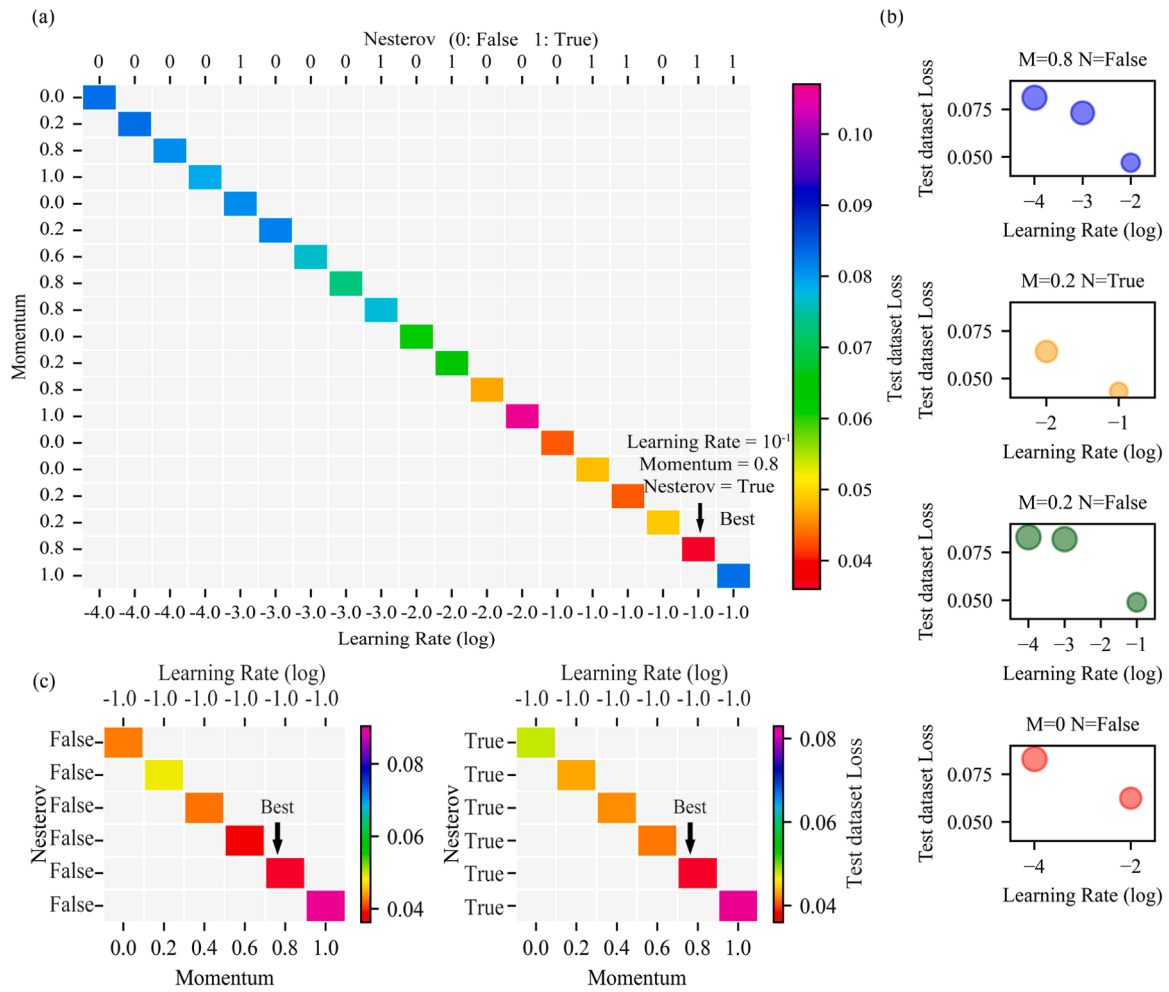


Fig. 5. The hyperparametric search results of SGD.

Table 7
Hyperparameter search results for seven optimisers.

Optimiser	Hyperparameter	Best value	Test dataset Loss
SGD	Learning Rate	10^{-1}	0.0360
	Momentum	0.8	
	Nesterov	True	
RMSprop	Learning Rate	10^{-3}	0.0353
	Rho	0.6	
	Momentum	0.4	
Adam	Learning Rate	10^{-3}	0.0350
	Beta1	0.8	
	Beta2	0.96	
Adadelta	Learning Rate	10^{-1}	0.0463
	Rho	0.92	
Adagrad	Learning Rate	10^{-1}	0.0377
	Initial Accumulator Value	0.2	
Adamax	Learning Rate	10^{-2}	0.0348
	Beta1	0.4	
	Beta2	0.98	
Nadam	Learning Rate	10^{-3}	0.0339
	Beta1	0.8	
	Beta2	0.98	

model reach the optimal point faster, which helped improve the accuracy and robustness of the model and provided a strong guarantee for large-scale applications.

The optimisation results of the three models were listed in Table 9. Among the three models, the LSTM and GRU models had similar test dataset loss values, both of which were significantly lower than those of

Table 8
Search space for network structure parameters.

Models	Hyperparameters	Search space
TCN	Number of layers	2
	Number of filters	(4 ~ 64)
	Kernel size	(2, 4, 8)
	Learning Rate	(10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1})
LSTM	Number of layers	2
	Number of nodes in hidden layer	(4 ~ 64)
GRU	Number of layers	2
	Number of nodes in hidden layer	(4 ~ 64)
FC	Number of nodes in hidden layer	(16 ~ 256)

the TCN model. This was because LSTM and GRU introduced a gating mechanism to control the flow of information, which could prevent the disappearance and explosion of gradients, thus outperforming TCN in processing long-sequence data. In addition, because the TCN used convolution for data feature extraction, it led to a significant increase in training time.

3.1.3. Structure optimisation of models under the complex scenario 2

The models for complex scenario 2 were optimised using the hyperparametric search space of complex scenario 1. The optimal structures of the three models were listed in Table 10. The double-layer LSTM model had the smallest loss value of the test dataset and was selected as the best model for complex scenario 2 in this study.

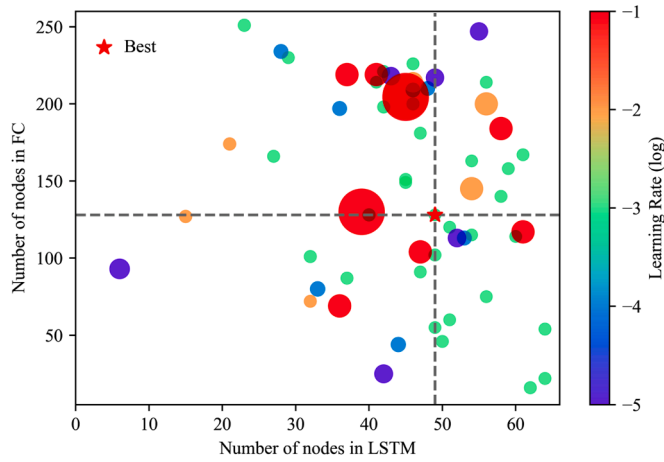


Fig. 6. Optimisation results for single-layer LSTM model. The size of the circle of the scatter represents the loss value of the test dataset.

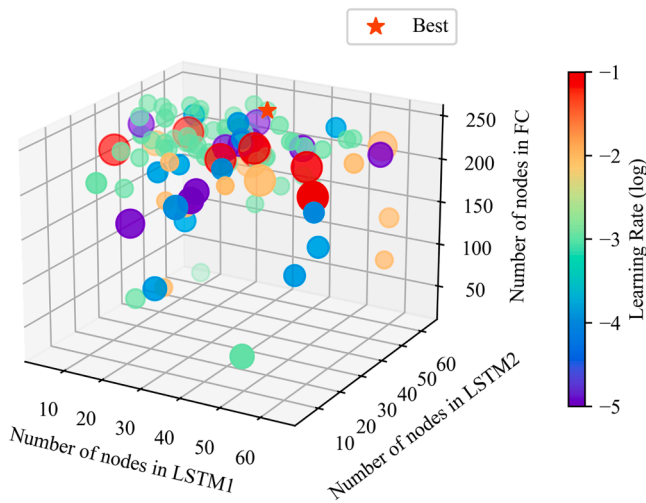


Fig. 7. Optimisation results for double-layer LSTM model.

3.1.4. Structure optimisation of models under the simple scenario

The optimization process of the neural network structure for the simple scenario was the same as for the two complex scenarios. The optimal structures of the three models were listed in Table 11. The single-layer LSTM model had the smallest loss value of the test dataset and was selected as the best model for the simple scenario in this study.

3.2. Source term inversion for three nuclear accident scenarios

Based on the optimised results of the neural network, the best-saving model was imported to assess the release rates of the four nuclides. To compare the advantages and disadvantages of the three models accurately, their MAPE on the test dataset were first compared, as shown in Fig. 8. The trends of the three models were the same for the three nuclear accident scenarios, and the MAPE of the four nuclides gradually decreased with increasing time steps, among which Kr-88 and Te-132 decreased sharply and remained stable after the 4th hour. Sr-91 levelled off after the 7th hour, while I-131 remained on a rapid downward trend for 10 h. From the value of MAPE, the inversion accuracies for the four nuclides were ranked as Te-132, Kr-88, Sr-91, and I-131. Among the three models, the TCN model had the fastest MAPE decline with the time step for the four nuclides, and the LSTM model had the highest assessment accuracy for the four nuclides. Comparing the two complex nuclear accident scenarios, the MAPE values of Kr-88, Sr-91, and Te-132

Table 9

Optimisation results of network structure parameters under the complex scenario 1.

Models	Hyperparameters	Best value	Test dataset Loss	Training time (s)
single-layer TCN	Learning rate	10^{-4}	0.0269	11,711
	Number of filters	52		
	Kernel size	2		
	Number of nodes in hidden layer of FC	142		
double-layer TCN	Learning rate	10^{-4}	0.0282	18,015
	Number of filters of TCN1	24		
	Number of filters of TCN2	24		
	Kernel size of TCN1	4		
	Kernel size of TCN2	4		
	Number of nodes in hidden layer of FC	94		
single-layer LSTM	Learning rate	10^{-3}	0.0242	5614
	Number of nodes in hidden layer of LSTM	49		
	Number of nodes in hidden layer of FC	128		
double-layer LSTM	Learning rate	10^{-3}	0.0245	8923
	Number of nodes in hidden layer of LSTM1	27		
	Number of nodes in hidden layer of LSTM2	61		
	Number of nodes in hidden layer of FC	234		
single-layer GRU	Learning rate	10^{-3}	0.0247	5312
	Number of nodes in hidden layer of GRU	39		
	Number of nodes in hidden layer of FC	208		
double-layer GRU	Learning rate	10^{-3}	0.0251	8420
	Number of nodes in hidden layer of GRU1	64		
	Number of nodes in hidden layer of GRU2	44		
	Number of nodes in hidden layer of FC	224		

Table 10

Optimal structure of three models for the complex scenario 2.

Model	Hyperparameter	Best value	Test dataset Loss	Training time (s)
single-layer TCN	Learning rate	10^{-4}	0.0195	7312
	Number of filters	64		
	Kernel size	8		
	Number of nodes in hidden layer of FC	152		
double-layer LSTM	Learning rate	10^{-3}	0.0182	7212
	Number of nodes in hidden layer of LSTM1	46		
	Number of nodes in hidden layer of LSTM2	48		
	Number of nodes in hidden layer of FC	173		
double-layer GRU	Learning rate	10^{-3}	0.0187	7211
	Number of nodes in hidden layer of GRU1	62		
	Number of nodes in hidden layer of GRU2	39		
	Number of nodes in hidden layer of FC	92		

were slightly lower than those of complex scenario 2, and I-131 was significantly lower than that of complex scenario 2 in complex scenario 1. The MAPEs for the four nuclides under the two complex nuclear accident scenarios were different from those of the simple nuclear accident scenario, but the difference was not significant.

Table 11
Optimal structure of the three models for the simple scenario.

Model	Hyperparameter	Best value	Test dataset Loss	Training time (s)
single-layer TCN	Learning rate	10^{-4}	0.0254	11,811
	Number of filters	64		
	Kernel size	2		
	Number of nodes in hidden layer of FC	138		
single-layer LSTM	Learning rate	10^{-3}	0.0231	5714
	Number of nodes in hidden layer of LSTM1	64		
	Number of nodes in hidden layer of FC	210		
	Learning rate	10^{-3}	0.0238	5614
single-layer GRU	Number of nodes in hidden layer of GRU1	42		
	Number of nodes in hidden layer of FC	200		
	Number of nodes in hidden layer of FC	200		

As shown in Table 12, there was a slight increase in MAPE for all four nuclides under the two complex scenarios compared with the results from the simple accident scenario with constant meteorological conditions and a single homogenous release of nuclides. The largest increase

in MAPE (5.86%) was from the inversion of I-131 under complex accident scenario 2. This suggested that source-term inversion could be effectively performed using the LSTM model under both the simple and complex accident scenarios described in this study.

It was worth noting from Tables 9 and 10 that during the optimisation of the neural network, the loss values of the test dataset of the LSTM model under complex scenario 1 were larger than those of complex scenario 2. However, Fig. 8 and Table 12 showed that the MAPE under complex scenario 1 was lower than that under complex scenario 2. This

Table 12
MAPE of four nuclides in three nuclear accident scenarios based on LSTM model.

Nuclear accident scenario	Kr-88	Sr-91	Te-132	I-131
Simple scenario (Constant meteorological parameters and release rates)	9.17%	15.22%	5.80%	36.01%
Complex scenario 1 (Changing meteorological parameters and constant release rates)	9.87%	17.49%	7.16%	38.83%
Complex scenario 2 (Constant meteorological parameters and changing release rates)	11.08%	16.51%	8.35%	41.87%

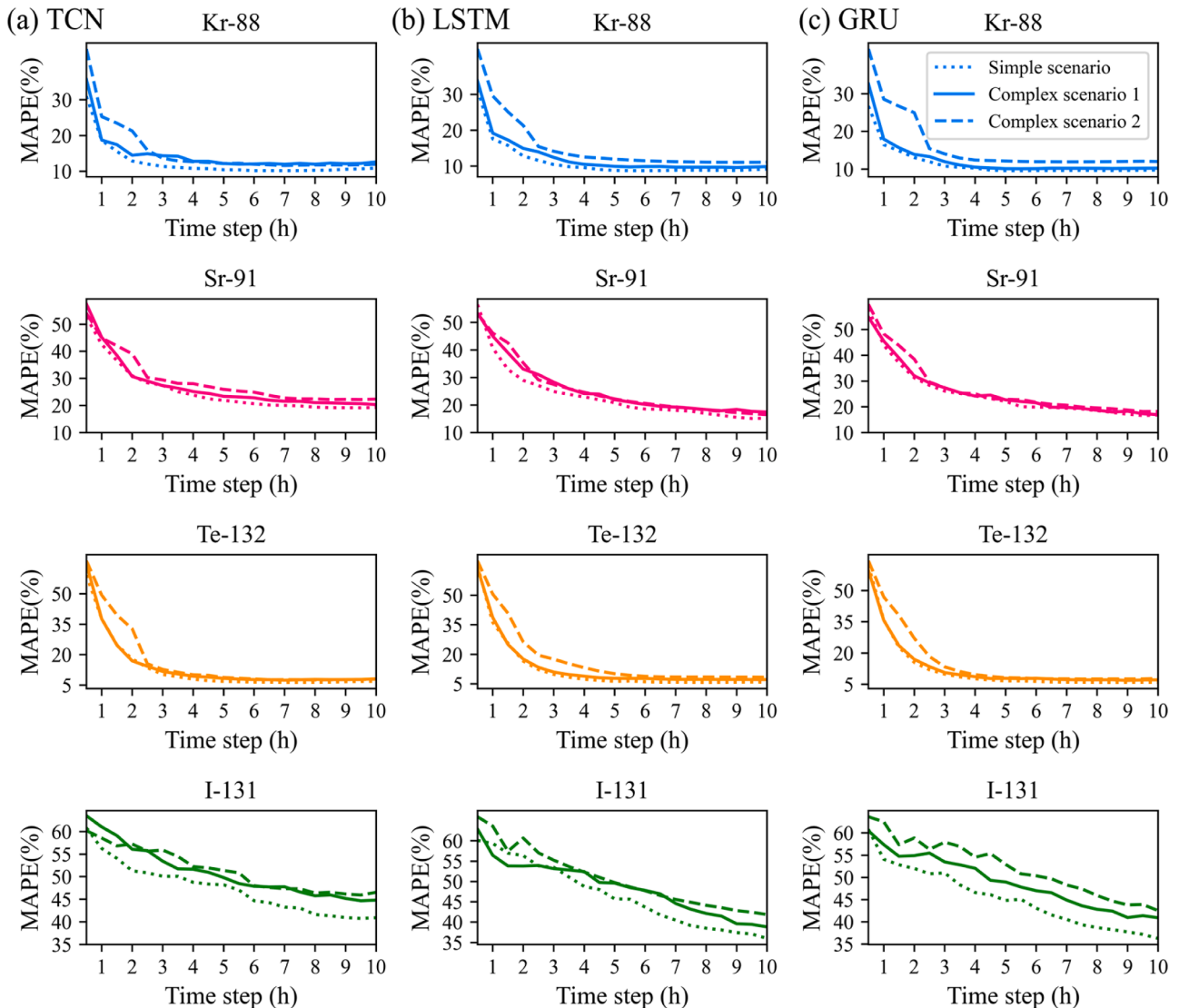


Fig. 8. MAPE of the three models on the test dataset.

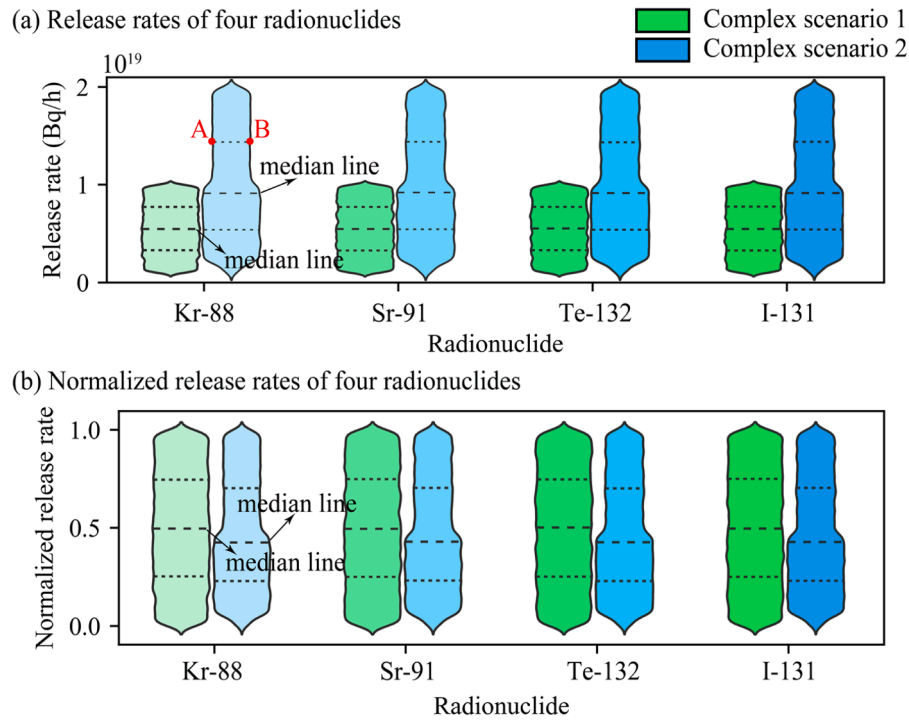
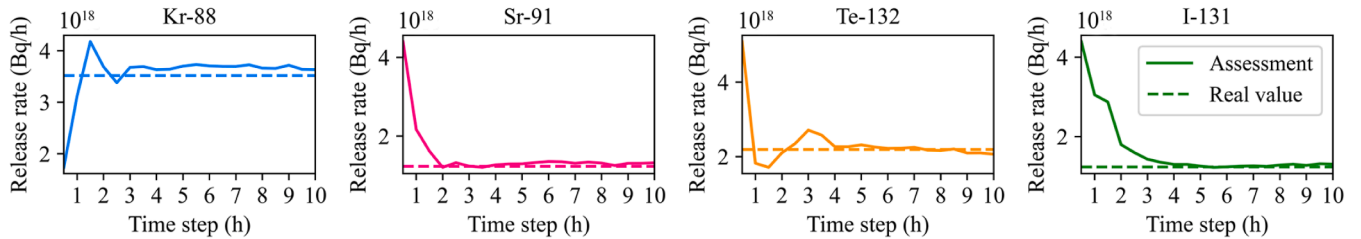


Fig. 9. Distribution of radionuclide release rate under two nuclear accident scenarios.

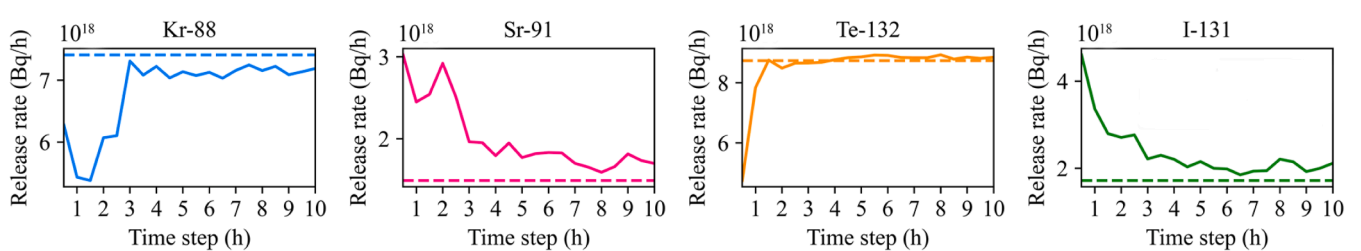
indicated that the accuracy of the source-term inversion under complex scenario 2 was worse. In complex scenario 2, the four radionuclides were continuously released for 30 min at the 2nd hour, indirectly reducing

the release rate for the first 4 time steps. Therefore, normalizing the radionuclide release rate decreased the overall radionuclide release rate over a 10 h period.

(a) Simple scenario



(b) Complex scenario 1



(c) Complex scenario 2

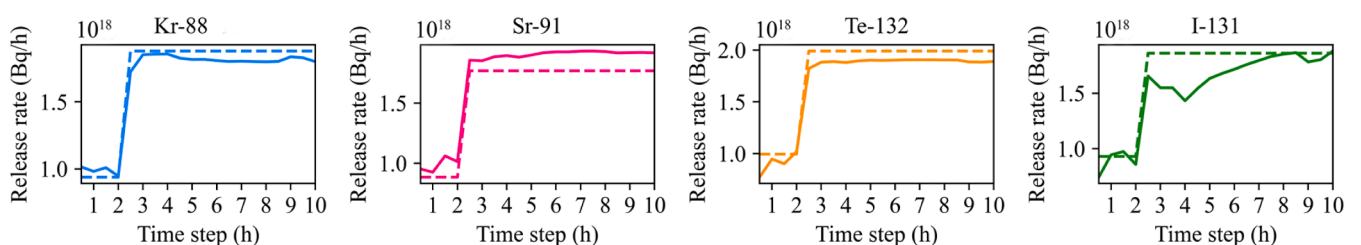


Fig. 10. Evolution of the assessed values over time for a set of test dataset under the three nuclear accident scenarios.

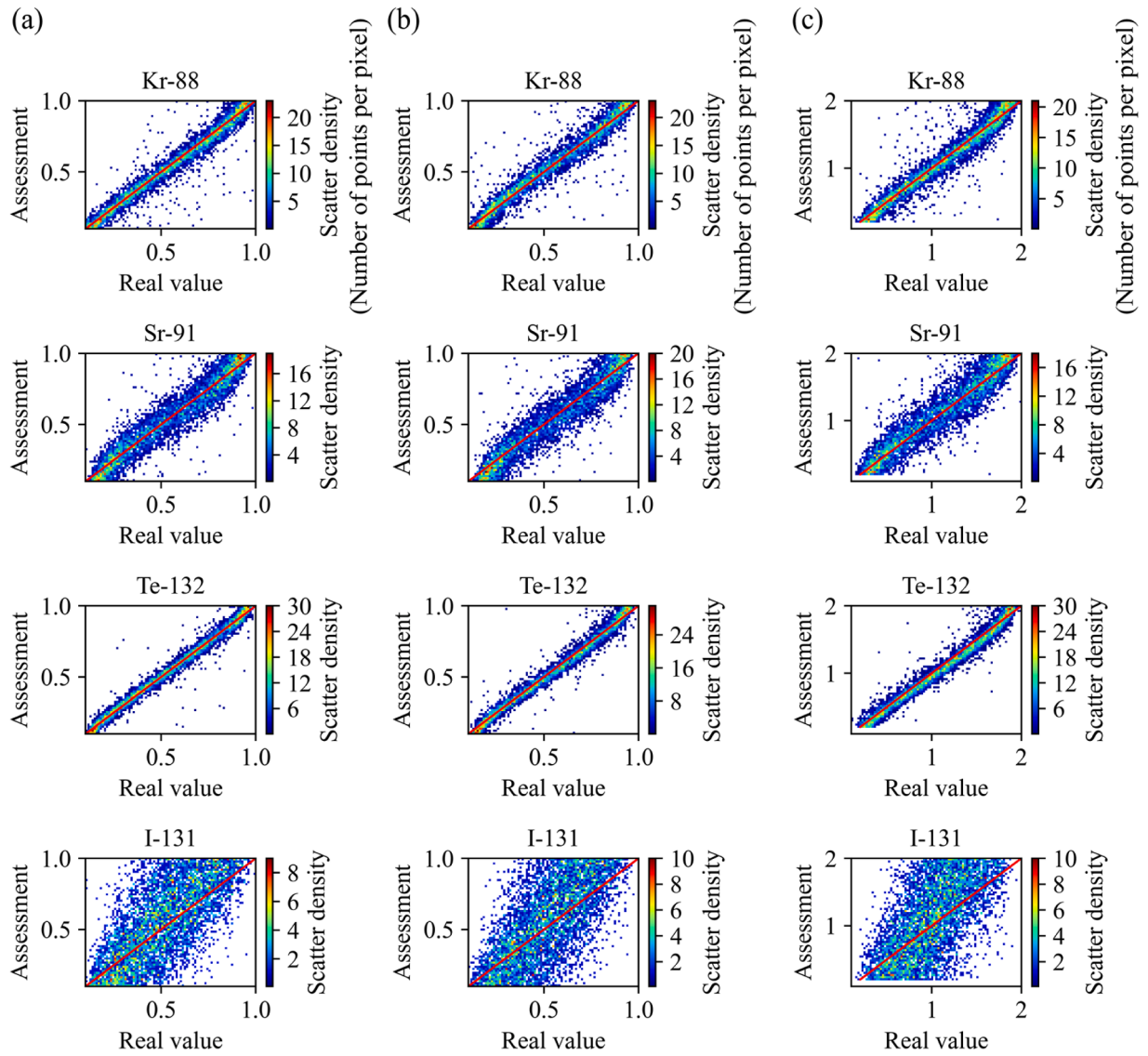


Fig. 11. Scatter plot of assessed values for the last time step on the test dataset for the three nuclear accident scenarios. (a) Simple scenario; (b) Complex scenario 1; (c) Complex scenario 2. The scatter represented the assessed values, and the red line indicated that the assessed values overlap with the real values.

Fig. 9 showed violin plots of 600,000 release rates distributions for four radionuclides, comprising 30,000 sets of samples. Each set encompassed 20 time steps release rate data. The horizontal width of the violin plot (e.g., the distance between points A and B in Fig. 9(a)) represented the probability density of the release rate, and the vertical height indicated the magnitude of the corresponding release rate.

As shown in Fig. 9(a), in complex scenario 1, the release rate exhibited a uniform distribution spanning from 1×10^{18} Bq/h and 1×10^{19} Bq/h. For complex scenario 2, the release rates for the last 16 time steps were twice as high as those in the first 4 time steps within each set of release rate data, accounting for 80% of the total 20 time steps. Within each of the 30,000 sets of release rate data, the first 4 time steps showed a uniform distribution ranging from 1×10^{18} Bq/h to 1×10^{19} Bq/h. It was notable that only when the release rates for the first 4 time steps were between 0.5×10^{19} Bq/h and 1×10^{19} Bq/h, the release rates for the last 16 time steps were distributed between 1×10^{19} Bq/h and 2×10^{19} Bq/h. This distribution accounted for 50% of the 30,000 sets of release rate data. Consequently, among the 600,000 release rates, only 40% were observed between 1×10^{19} Bq/h and 2×10^{19} Bq/h, leading to the narrower width of the upper half of the violin plot compared to the lower half.

The dashed line at the middle of the violin plot represented the median release rate, revealing significantly higher release rates in complex scenario 2 when compared to complex scenario 1. As shown in Fig. 9(b), after normalizing the release rate using Eq. 1, the normalized release rates under complex scenario 2 were significantly lower than those under complex scenario 1 according to the median line of the violin plot. Therefore, the test dataset under complex scenario 2 showed smaller test dataset loss values during model training when MSE was used as the loss function.

To analyse the variation in the assessed values over time, a set of data was selected from each of the 6000-test dataset under the three nuclear accident scenarios, as shown in Fig. 10. Fig. 10(a) indicated that the assessed values of the four nuclides gradually converged to the real values under the simple scenario. Fig. 10(b) showed that, for complex scenario 1, the assessed value of Te-132 was already very close to the real value at the 2nd hour and stabilised near the real value at the subsequent time step. The assessed value for Kr-88 also began to stabilise at approximately the real value at the 3rd hour. For Sr-91 and I-131, the assessed values gradually converged to the real values during 10 h. As shown in Fig. 10(c), for complex scenario 2, the real values of the four nuclides increased rapidly in the 2nd hour, indicating that the two-stage

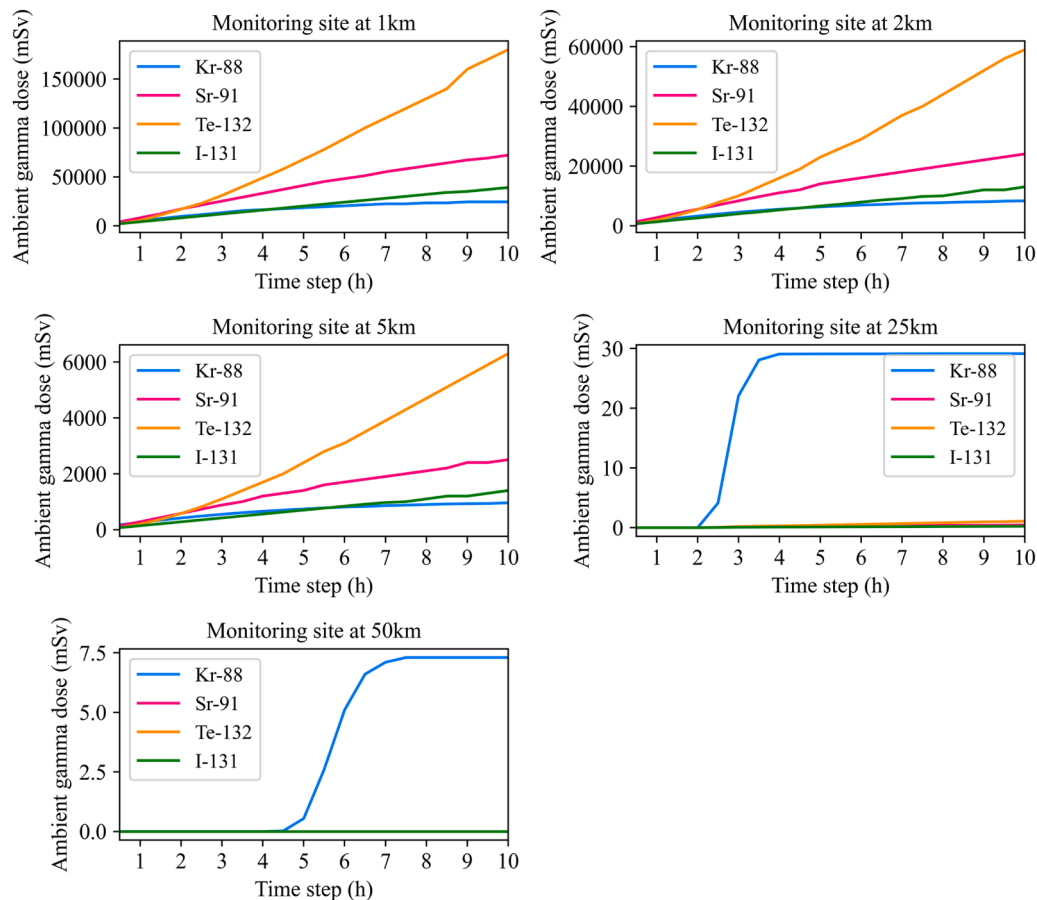


Fig. 12. Ambient gamma dose contributed by the four nuclides at the five monitoring sites downwind during 10 h.

release of the four nuclides began in the 2nd hour. Te-132, Kr-88, and Sr-91 all started to stabilise near the real values at the 2nd hour, and I-131 gradually converged to the real values during 10 h.

The scatter plot (Fig. 11) showed the overall assessment of the four nuclides in the last time step based on the LSTM model, which helped to analyze the deviation density distribution between the assessed and real values. It could be real from Fig. 11 that there was no significant difference in the scatter density distributions between the simple and complex scenarios. For Kr-88, Sr-91, and Te-132, the scattered points had a narrow distribution width and were heavily concentrated around the ideal agreement line, indicating high reliability. For I-131, although the distribution of the scattered points was broad, there was still a decent number of scattered points clustered around the ideal agreement line.

From Figs. 8 and 11, the assessment accuracies of the four nuclides were ranked as Te-132, Kr-88, Sr-91, and I-131, regardless of the source term inversion model or nuclear accident scenario. To analyse the reasons for this phenomenon, an analysis of the dose produced by a single nuclide over a 10 h period was performed using the auxiliary data in Table 5, with a fixed release rate of 1×10^{19} Bq/h for each nuclide. The ambient gamma doses for the four nuclides at the five downwind monitoring sites were shown in Fig. 12. Among the five monitoring points, the gamma dose rates at the first three points had the largest proportion, thus significantly impacting the accuracy of source term inversion. Among the three solid particle radionuclides (Te-132, Sr-91, and I-131), Te-132 contributed the highest gamma dose during the 10-hour nuclear accident simulation period, displaying a significant difference in gamma dose contribution compared to the others. Therefore, during the source term inversion process, Te-132 could be effectively identified by the three machine learning models, showing the highest

accuracy in source term inversion.

Even though the gamma dose contribution of I-131 was not significantly different from Sr-91, due to the much longer half-life of I-131 (3.3 days) compared to Sr-91 (9.63 h), the quantity of I-131 experienced minimal changes within the 10-hour simulation of a nuclear accident. This resulted in a 21.34% higher MAPE for I-131 compared to Sr-91 in complex scenario 1. As for the only gaseous radionuclide in the study, Kr-88, despite having the lowest gamma dose contribution at the first three monitoring points, its sufficiently short half-life (2.8 h) ensured significant change in its quantity within the 10-hour period. This change was effectively captured by the three machine learning models, leading to higher accuracy in source term inversion.

In summary, the gamma dose contribution and half-life of radionuclides appear to be the primary factors determining the accuracy of source term inversion. Nuclides with higher gamma dose contributions and shorter half-lives show higher accuracy in the source term inversion process, and these two factors together affect the accuracy of the source term inversion.

3.3. Stability analysis

In real applications, the data are subject to certain errors during the measurement process. It is critical to verify the stability of the LSTM model and to identify the key parameters affecting its accuracy. Simulation experiments with artificially added errors were conducted on the raw data under complex scenario 1, including the gamma dose rate, release height, wind speed, wind direction, atmospheric stability, mixed layer height, and precipitation, to further analyse the model performance after adding errors. From these experiments, it was possible to gain a deeper understanding of the stability of the model and assess its

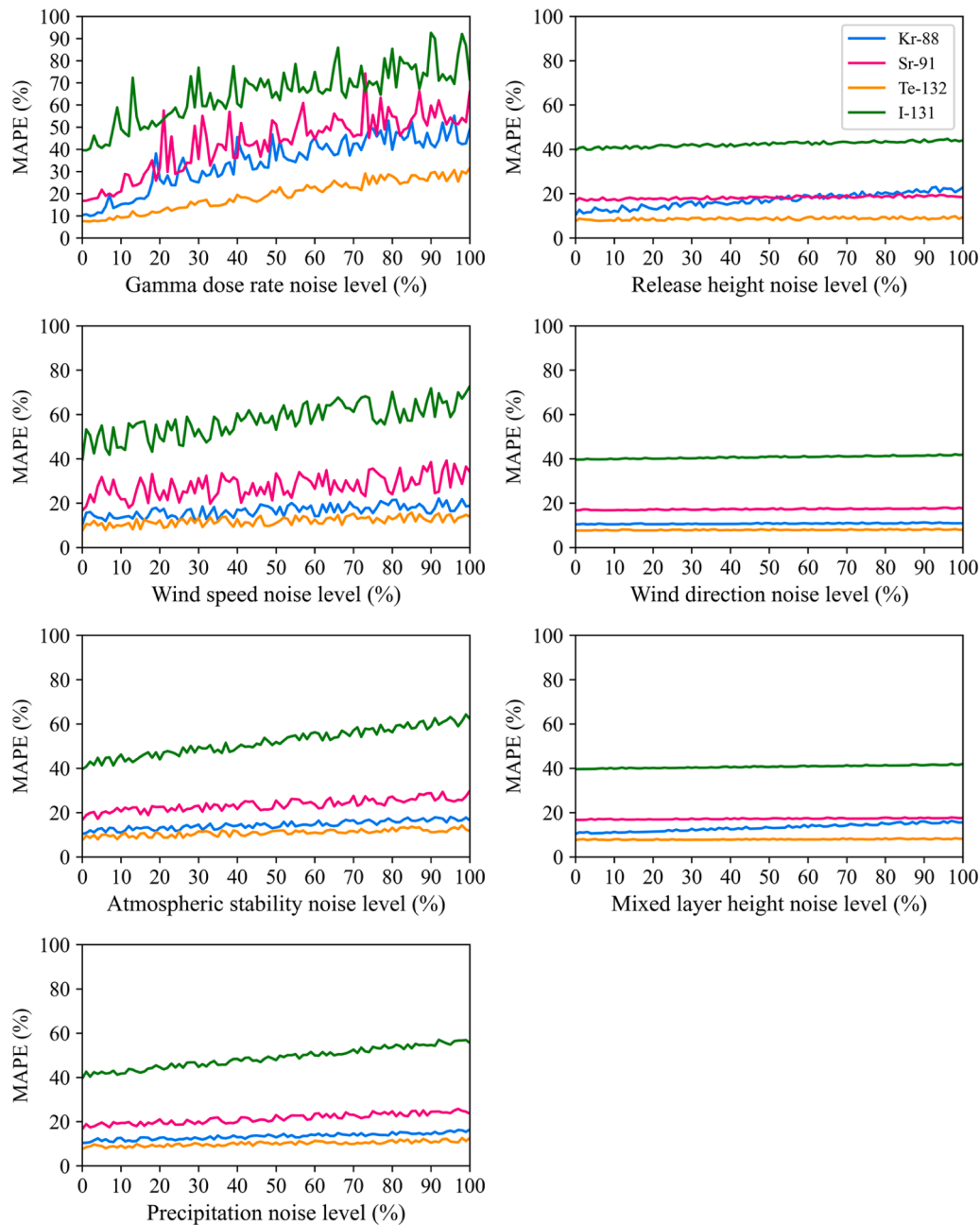


Fig. 13. Assessment performance at the last time step after adding noise to the test dataset.

reliability in practical applications.

As shown in Fig. 13, noise was added to the data for the seven inputs in complex scenario 1. The gamma dose rate was the most sensitive to noise, and the MAPE of the four nuclides gradually increased in the last time step as the noise level increased. After adding 100% noise, the MAPE of Kr-88, Sr-91, Te-132, and I-131 increased to 49.04%, 65.94%, 31.57%, and 71.62%, respectively, indicating that the gamma dose rate significantly affected the estimative performance of LSTM. Although the assessment performance of the four nuclides deteriorated after being affected by the strong noise, Te-132 still exhibited good assessment results, indicating its high stability in noisy environments. The release height had a significant effect on the gaseous nuclide Kr-88, with the MAPE increasing by 11.95% when the noise level reached 100%, whereas this effect was not significant for the other radionuclides. Wind speed has been identified as an important factor influencing the diffusion of radionuclides. As the noise level increases, the MAPE of the four

radionuclides gradually increase. Because the five monitoring points in this study were located downwind, the impact of the wind direction on the estimated source term was deemed insignificant. Atmospheric stability had an acceptable effect on the estimated release rates of the four radionuclides. The mixed-layer height had a minor effect on the estimation of the source term. Furthermore, precipitation was identified as a factor causing the wet deposition of radionuclides. With increasing noise levels, the MAPE of the four nuclides gradually increased.

4. Conclusions

In this study, a methodology was developed for source term inversion based on the gamma dose rate under two complex nuclear accidents (changing meteorological parameters and two-stage release of radionuclides). This method adopted three neural network models (TCN, LSTM, and GRU) to estimate the release rates of four radionuclides (Kr-

88, Sr-91, Te-132, and I-131). Through Optuna optimisation, the Nadam optimiser and LSTM model were selected as the best optimiser and source term inversion model, respectively, for two complex nuclear accidents. The best models for the two complex nuclear accidents were the single-layer LSTM (where the learning rate was 10^{-3} and the number of nodes in the hidden layer of the LSTM was 49, and the number of nodes in the hidden layer of the FC was 128) and the double-layer LSTM (where the learning rate was 10^{-3} , the number of hidden layer nodes for the first LSTM was 46, the number of hidden layer nodes for the second LSTM was 48, and the number of hidden layer nodes for the FC layer was 173). The results of the model estimation showed that for the two types of complex nuclear accidents, the MAPE of the release rate for the four radionuclides gradually decreased as the time-series gamma dose rate was the input, in which the MAPE of Te-132 in the last time step was 7.16% and 8.35%, respectively, indicating that the LSTM model has a high accuracy. Stability analysis showed that the gamma dose rate had a crucial influence on the source term inversion, and the wind speed, atmospheric stability, and precipitation also influenced the model assessment, which should be maintained as accurately as possible. Based on these results, the proposed method provided a framework for source-term inversion in complex nuclear accident scenarios.

It can be concluded that the models have some limitations in multi-nuclide source term inversion, mainly due to the lack of sufficient information on the nuclides with low gamma dose rate contributions, which leads to significant discrepancies between the assessed and real values for these nuclides.

Environmental implication

Kr-88, Sr-91, Te-132, and I-131 are the main short-lived nuclides released into the atmosphere during nuclear accidents, and their source terms are necessary to study the consequences and environmental behavior of short-lived nuclides. However, variations of meteorological parameter and two-stage release of radionuclides can increase the difficulty of source term inversion. We propose three machine learning models for source term inversion and obtain a good accuracy.

Data Availability

Data will be made available on request.

CRediT authorship contribution statement

Liu Chengfeng: Writing – original draft, Software, Methodology, Conceptualization. **Ling Yongsheng:** Project administration, Formal analysis. **Wang Jing:** Writing – review & editing, Supervision. **Yue Qi:** Software, Investigation. **Jia Wenbao:** Supervision, Investigation. **Shi Chao:** Investigation. **Zhang Xiaojun:** Resources. **Hei Daqian:** Funding acquisition. **Shan Qing:** Validation, Software.

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.

Acknowledgments

This study was supported by the Fundamental Research Funds for the Central Universities (Grant No. NS2022056) and by a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors thank Editage (www.editage.cn)

for English language editing. The authors thank the editors and anonymous reviewers for their constructive comments and suggestions that helped improve the paper.

References

- [1] Athey, G.F., Sjoreen, A.L., and McKenna, T.J. 1994. RASCAL Version 2.1 workbook. Volume 2, Revision 2. United States. <https://doi.org/10.2172/10105537>.
- [2] Ayo-Imoru, R.M., Cilliers, A.C., 2018. Continuous machine learning for abnormality identification to aid condition-based maintenance in nuclear power plant. *Ann Nucl Energy* 118, 61–70. <https://doi.org/10.1016/j.anucene.2018.04.002>.
- [3] Bai, S., Kolter, J.Z., Koltun, V., 2018. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. <https://doi.org/10.48550/arXiv.1803.01271>.
- [4] Bao, B., 2011. Study of Coastal wind Characteristics in China. Ocean University of China. (<https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2012&fileame=1011230923.nh>).
- [5] Barbosa, S., Huisman, J.A., Azevedo, E.B., 2018. Meteorological and soil surface effects in gamma radiation time series - Implications for assessment of earthquake precursors. *J Environ Radioact* 195, 72–78. <https://doi.org/10.1016/j.jenvrad.2018.09.022>.
- [6] Chollet, F., 2015. Keras. GitHub. (<https://github.com/fchollet/keras>).
- [7] Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. <https://doi.org/10.48550/arXiv.1412.3555>.
- [8] Davoine, X., Bocquet, M., 2007. Inverse modelling-based reconstruction of the Chernobyl source term available for long-range transport. *Atmos Chem Phys* 7, 1549–1564. <https://doi.org/10.5194/acp-7-1549-2007>.
- [9] Dong, X., Fang, S., Zhuang, S., Xu, Y., Zhao, Y., Sheng, L., 2023. Objective inversion of the continuous atmospheric ¹³⁷Cs release following the Fukushima accident. *J Hazard Mater* 447, 130786. <https://doi.org/10.1016/j.jhazmat.2023.130786>.
- [10] Evangelou, 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. <https://doi.org/10.5194/acp-17-8805-2017>.
- [11] Fan, S., Zhu, W., Wang, A., Guo, L., Dong, J., 2005. Study on the meteorological characteristics of the boundary layer in the Pearl River Delta. *Acta Sci Nat Univ sunyatseni*. (https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C44YLTlOAiTRKgcHrJ08w1e7F1IFNsBV5UthMzOht_dKJ6CT0BCIP_BaVFz8WXRclFs9p_TwWIWEXfD_lcxgRJ8&uniplatform=NZKPT).
- [12] 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. <https://doi.org/10.1016/j.jhazmat.2022.129806>.
- [13] Fang, S., Dong, X., Zhuang, S., Tian, Z., Zhao, Y., Liu, Yun, Liu, Yuanyuan, Sheng, L., 2023. Inversion of ¹³⁷Cs emissions following the Fukushima accident with adaptive release recovery for temporal absences of observations. *Environ Pollut* 317, 120814. <https://doi.org/10.1016/j.envpol.2022.120814>.
- [14] 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>.
- [15] Hashemian, H.M., 2011. On-line monitoring applications in nuclear power plants. *Prog Nucl Energy* 53, 167–181. <https://doi.org/10.1016/j.pnucene.2010.08.003>.
- [16] Hirano, M., Yonomoto, T., Ishigaki, M., Watanabe, N., Maruyama, Y., Sibamoto, Y., Watanabe, T., Moriyama, K., 2012. Insights from review and analysis of the Fukushima Dai-ichi accident. *J Nucl Sci Technol* 49, 1–17. <https://doi.org/10.1080/18811248.2011.636538>.
- [17] Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput* 9, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [18] IAEA, 1997. 955: Generic Assessment Procedures for Determining Protective Actions during a Reactor Accident. IAEA, Vienna.
- [19] Katata, G., Chino, M., Kobayashi, T., Terada, H., Ota, M., Nagai, H., Kajino, M., Draxler, R., Hort, M.C., Malo, A., Torii, T., Sanada, Y., 2015. Detailed source term estimation of the atmospheric release for the Fukushima Daiichi Nuclear Power Station accident by coupling simulations of an atmospheric dispersion model with an improved deposition scheme and oceanic dispersion model. *Atmos Chem Phys* 15, 1029–1070. <https://doi.org/10.5194/acp-15-1029-2015>.
- [20] Katata, G., Ota, M., Terada, H., Chino, M., Nagai, H., 2012. Atmospheric discharge and dispersion of radionuclides during the Fukushima Dai-ichi Nuclear Power Plant accident. Part I: Source term estimation and local-scale atmospheric dispersion in early phase of the accident. *J Environ Radioact* 109, 103–113. <https://doi.org/10.1016/j.jenvrad.2012.02.006>.
- [21] Klampanos, I.A., Davvetas, A., Andronopoulos, S., Pappas, C., Ikonopoulos, A., Karkalatsis, V., 2018. Autoencoder-driven weather clustering for source estimation during nuclear events. *Environ Model Softw* 102, 84–93. <https://doi.org/10.1016/j.envsoft.2018.01.014>.
- [22] Koo, B.T., Lee, H.C., Bae, K., Kim, Y., Jung, J., Park, C.S., Kim, H.-S., Min, C.H., 2021. Development of a radionuclide identification algorithm based on a convolutional neural network for radiation portal monitoring system. *Radiat Phys Chem* 180, 109300. <https://doi.org/10.1016/j.radphyschem.2020.109300>.

- [23] Kovalets, I., Andronopoulos, S., Hofman, R., Seibert, P., Ievdin, I., 2016. Advanced method for source term estimation and status of its integration in JRODOS. *Radioprotection* 51, S121–S123. <https://doi.org/10.1051/radiopro/2016046>.
- [24] 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. <https://doi.org/10.1016/j.jhazmat.2019.120770>.
- [25] Ling, Y., Chaojun, C., Wenyu, H., Hei, Daqian, Hei, Da.Qian, Qing, S., Jia, W., 2019. A new method for nuclear accident source term inversion based on GA-BPNN algorithm. *Neural Netw World* 29, 71–82. <https://doi.org/10.14311/nnw.2019.29.006>.
- [26] Ling, Y., Huang, T., Yue, Q., Shan, Q., Hei, D., Zhang, X., Shi, C., Jia, W., 2022. Improving the estimation accuracy of multi-nuclide source term estimation method for severe nuclear accidents using temporal convolutional network optimized by Bayesian optimization and hyperband. *J Environ Radioact* 242, 106787. <https://doi.org/10.1016/j.jenvrad.2021.106787>.
- [27] Ling, Y., Liu, C., Shan, Q., Hei, D., Zhang, X., Shi, C., Jia, W., Wang, J., 2023. Inversion Method for Multiple Nuclide Source Terms in Nuclear Accidents Based on Deep Learning Fusion Model. *Atmosphere*. <https://doi.org/10.3390/atmos14010148>.
- [28] Ling, Y., Yue, Q., Chai, C., Shan, Q., Hei, D., Jia, W., 2020. Nuclear accident source term estimation using Kernel Principal Component Analysis, Particle Swarm Optimization, and Backpropagation Neural Networks. *Ann Nucl Energy* 136, 107031. <https://doi.org/10.1016/j.anucene.2019.107031>.
- [29] Ling, Y., Yue, Q., Huang, T., Shan, Q., Hei, D., Zhang, X., Jia, W., 2021. Multi-nuclide source term estimation method for severe nuclear accidents from sequential gamma dose rate based on a recurrent neural network. *J Hazard Mater* 414, 125546. <https://doi.org/10.1016/j.jhazmat.2021.125546>.
- [30] Lodde, G.M., Good, B.A., Surgeoner, D.M., 1989. Assessment and control of the Three Mile Island Unit 2 reactor building atmosphere. *Nucl Technol*. <https://doi.org/10.13182/NT89-A27750>.
- [31] Luo, H., Karki, B.B., Ghosh, D.B., Bao, H., 2021. Deep neural network potentials for diffusional lithium isotope fractionation in silicate melts. *Geochim Et Cosmochim Acta* 303, 38–50. <https://doi.org/10.1016/j.gca.2021.03.031>.
- [32] Optuna. Available online: (<https://optuna.readthedocs.io/en/stable/>) (accessed on 19 February 2023).
- [33] Qi, S., Wang, S., Chen, Y., Zhang, K., Ai, X., Li, J., Fan, H., Zhao, H., 2022. Radionuclide identification method for NaI low-count gamma-ray spectra using artificial neural network. *Nucl Eng Technol* 54, 269–274. <https://doi.org/10.1016/j.net.2021.07.025>.
- [34] Raskob, W., Gering, F., Lochard, J., Nisbet, A., Starostova, V., Tomic, B., 2011. Nuclear and Radiological Emergency Management and Rehabilitation Strategies: Towards a EU approach for decision support tools. *Nucl Eng Des*, Seven–Eur Comm Conf Euratom Res Train React Syst (Fission Saf 2009) 241, 3395–3402. <https://doi.org/10.1016/j.nucengdes.2011.03.008>.
- [35] Sasaki, M., Sanada, Y., Katengeza, E.W., Yamamoto, A., 2021. New method for visualizing the dose rate distribution around the Fukushima Daiichi nuclear power plant using artificial neural networks. *Sci Rep* 11, 1857. <https://doi.org/10.1038/s41598-021-81546-4>.
- [36] Saunier, O., Mathieu, A., Didier, D., Tombette, M., Quélo, D., Quélo, D., Winiarek, V., Bocquet, M., 2013. An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations. *Atmos Chem Phys* 13, 11403–11421. <https://doi.org/10.5194/acp-13-11403-2013>.
- [37] Sjoreen, A., Athey, G., Ramsdell, J., McKenna, T., 1994. RASCAL Version 2.1, User's Guide. Nuclear Regulatory Commission.
- [38] Ten Hoeve, J.E., Jacobson, M.Z., 2012. Worldwide health effects of the Fukushima Daiichi nuclear accident. *Energy Environ Sci* 5, 8743–8757. <https://doi.org/10.1039/C2EE22019A>.
- [39] Tichý, O., Smidl, V., Hofman, R., Hofman, Radek, Šindelářová, K., Hýža, M., Stohl, A., 2017. Bayesian inverse modeling and source location of an unintended ¹³¹I release in Europe in the fall of 2011. *Atmos Chem Phys* 17, 12677–12696. <https://doi.org/10.5194/acp-17-12677-2017>.
- [40] Wang, X., Kang, L., 2003. *Nuclear Accident Consequence Evaluation Methods and their New Developments*. China Atomic Energy Press, Beijing.
- [41] Yeşilkanat, C.M., Kobya, Y., 2015. Determination and mapping the spatial distribution of radioactivity of natural spring water in the Eastern Black Sea Region by using artificial neural network method. *Environ Monit Assess* 187. <https://doi.org/10.1007/s10661-015-4811-0>.
- [42] Yue, Q., Jia, W., Huang, T., Shan, Q., Hei, D., Zhang, X., Ling, Y., 2020. Method to determine nuclear accident release category via environmental monitoring data based on a neural network. *Nucl Eng Des* 367, 110789. <https://doi.org/10.1016/j.nucengdes.2020.110789>.
- [43] Zhang, X., Raskob, W., Landman, C., Trybushnyi, D., Li, Y., 2017. Sequential multi-nuclide emission rate estimation method based on gamma dose rate measurement for nuclear emergency management. *J Hazard Mater* 325, 288–300. <https://doi.org/10.1016/j.jhazmat.2016.10.072>.
- [44] 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>.
- [45] 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>.