



OPEN

# New method for visualizing the dose rate distribution around the Fukushima Daiichi Nuclear Power Plant using artificial neural networks

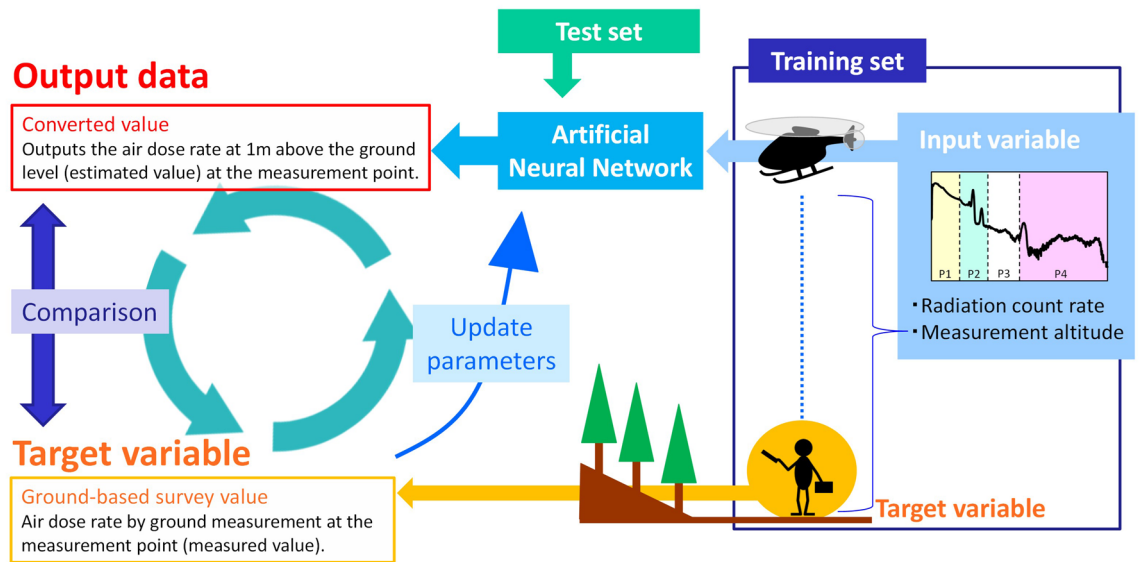
Miyuki Sasaki<sup>1✉</sup>, Yukihisa Sanada<sup>1</sup>, Estiner W. Katengeza<sup>2</sup> & Akio Yamamoto<sup>3</sup>

This study proposes a new method of visualizing the ambient dose rate distribution using artificial neural networks (ANNs) from airborne radiation monitoring results. The method was applied to the results of the airborne radiation monitoring which was conducted around the Fukushima Daiichi Nuclear Power Plant by an unmanned aerial vehicle. Much of the survey data obtained in the past were used as the training data for building a network. The number of training cases was related to the error between the ground and converted values by the ANN. The quantitative evaluation index (the root-mean-square error) between the ANN-converted value and the ground-based survey result converged at 200 training cases. This number of training case was considered a rough criterion of the required number of training cases. The reliability of the ANN method was evaluated by comparison with the ground-based survey data. The dose rate map created by the ANNs method reproduced ground-based survey results better than traditional methods.

Large quantities of radionuclides were released into the atmosphere after the Fukushima Daiichi Nuclear Power Plant (FDNPP) accident in March 2011<sup>1</sup>. Nine years after the accident, the ambient dose rate (air dose rate) has been decreased by radioactive decay, decontamination work, and weathering effect within the 80-km radius zones from the FDNPP<sup>2</sup>. As a quick and efficient survey method, manned helicopters and unmanned aerial vehicles (UAVs) have been developed to visualize the environmental distribution of the air dose rate in the airborne radiation survey (UAV-survey) after the FDNPP accident<sup>3,4</sup>. UAVs are effective tools for data collection over wide areas that are located around FDNPP because a person does not have to approach a dangerous place<sup>5,6</sup>. However, a UAV-survey is less accurate than a ground-based radiation survey using a handheld survey meter because it has a long distance from the ground surface source.

In a conventional approach, the count rates and the pulse height distributions collected by the UAV-survey are converted into air dose rates and deposition of radioactive cesium on the ground. In this approach, the ground is assumed to have a flat form, and this is called the flat source model (FSM)<sup>7</sup>. Improving UAV-survey accuracy requires accounting for the topographic features in radiation survey areas. Some reports suggested that the topographic effect was corrected by published big topographical data, such as the digital elevation model (DEM) and the digital surface model (DSM)<sup>8,9</sup>. A conversion method using the maximum-likelihood expectation-maximization (ML-EM) method was recently proposed to convert the value measured from the sky into an air dose rate of 1 m above the ground level (agl.)<sup>10,11</sup>. The conversion using the ML-EM method requires the creation of many types of attenuation parameters to improve the conversion accuracy. This method is effective for conversion in complex terrain areas, such as mountains or forests, which could not be reproduced by conventional FSM methods. However, the decision on the parameters needs many on-site measurement experiments or simulation work.

<sup>1</sup>Collaborative Laboratories for Advanced Decommissioning Science, Japan Atomic Energy Agency, 45-169 Sukakeba, Minamisoma, Fukushima 975-0036, Japan. <sup>2</sup>Department of Environment Systems, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan. <sup>3</sup>Graduate School of Engineering, Department of Applied Energy, Nagoya University, C2-4 Furouchou, Chikusaku, Nagoya 464-8603, Japan. ✉email: sasaki.miyuki@jaea.go.jp



**Figure 1.** Images of the application of the artificial neural network to the airborne radiation survey.

To solve this problem, we attempted herein to convert radiation measurement values using an artificial neural network (ANN). The ANN is applied to the study fields of medicine, physics, etc. as a kind of deep learning<sup>12,13</sup>. Simple and easy ANN software was recently published<sup>14,15</sup>. Moreover, the ANN method was applied to search for radiation sources for nuclear security<sup>16</sup>. Compared to the ML-EM method, the conversion by an ANN is expected to reduce both the workload of creating many parameters and the calculation time (Fig. 1). We have many data sets from the UAV-survey and the ground-based survey in the same locations because periodical surveys were conducted around the FDNPP as part of the national project<sup>4-6</sup>. These data sets were suitable for the training set in the ANN with published topographical data<sup>17</sup>. The air dose rate maps created by the UAV-survey are used for emergency monitoring data and determine the evacuation route of inhabitants; thus, the reliability of the created air dose rate map must be evaluated.

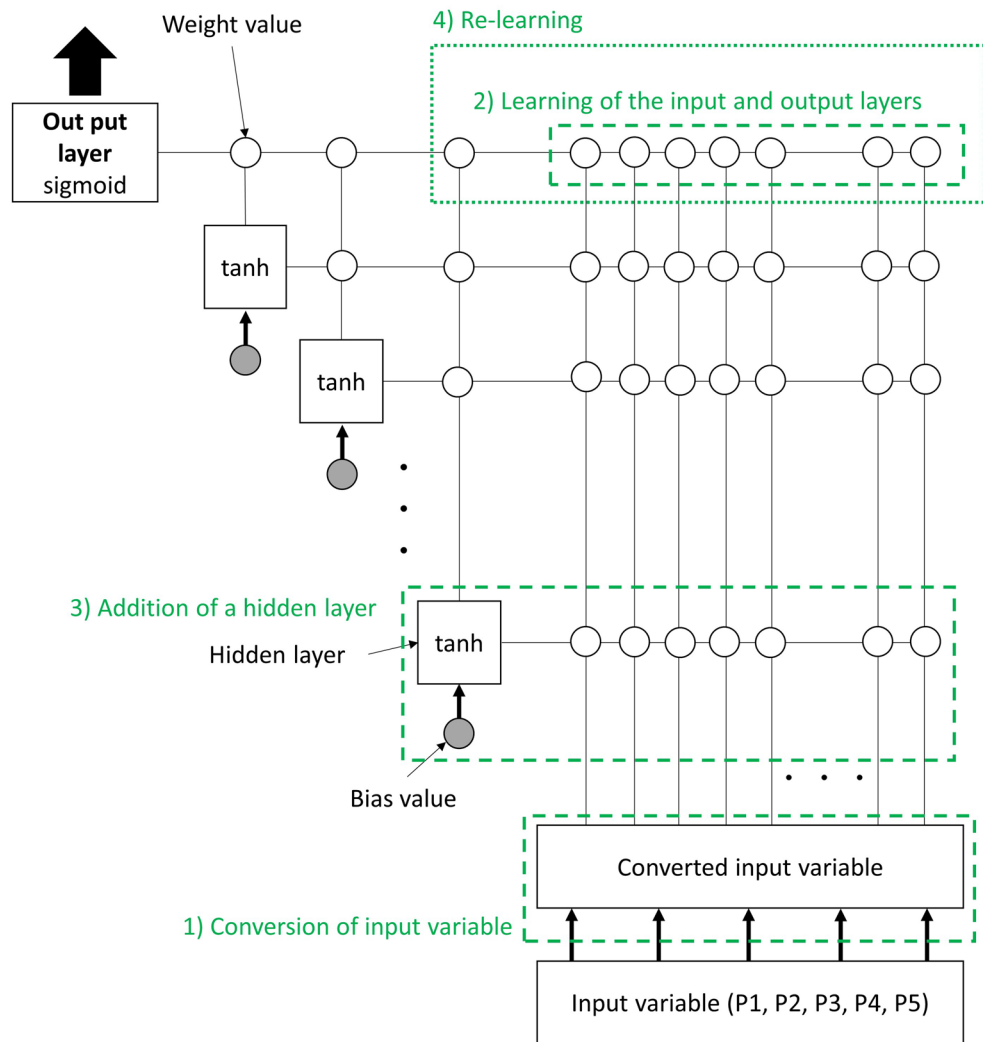
We developed herein a new conversion method using the ANN to improve the accuracy and the analysis time of the UAV-survey. This ANN method was optimized by evaluating the number of hidden layers constituting a network. In addition, we also evaluated the relationship between the number of data cases used for the training and the error between the ANN-converted UAV-survey and ground-based survey values. Finally, the ANN method validation was evaluated by comparison with the FSM and the ML-EM method.

## Materials and methods

**ANN method.** In this study, the ANN was constructed using NeuralWorks Predict (NeuralWare, Carnegie, USA), a software that uses cascade correlation to determine the optimal network structure with a simple graphical user interface<sup>18</sup>. This learning is classified as a regression problem of supervised learning. The cascade correlation is a constructive learning rule and a supervised learning algorithm that constructs a feed-forward network. Learning starts with a minimal network consisting only of the input and output layers. Minimizing the overall error of a network, new hidden layers are then added step by step. The cascade correlation builds a network in which one hidden layer has one neuron. Figure 2 depicts the ANN modeling flowchart. One training datum (set of input and target variables) is defined as a "training case." A dataset of multiple training cases is defined as a "training set."

The training set was randomly divided into two cases depending on the purpose of this method. In the ANN construction, 70% of all the training cases were randomly selected. The other training case was used to evaluate the calculation of the correlation coefficient with the target variable. The changes in the evaluation were observed each time the learning was repeated. Learning was stopped when the correlation coefficient stopped improving. The construction flow of the network is presented below and in Fig. 2:

- 1) Conversion of the input variable: The input variable is represented by some functions. The data with a high correlation between the input and target variables are normalized and added as the input data.
- 2) Learning of the input and output layers: Training is first implemented only on the input and output layers without the hidden layer. The output layer is a sigmoid function recommended from NeuralWorks Predict.
- 3) Addition of a hidden layer: If the index of the reliability (the correlation coefficient) of the output value in the training of former procedure 2) was not enough, one hidden layer would be added automatically. All hidden layers are tanh functions recommended from NeuralWorks Predict.
- 4) Re-learning: Learning will be performed again with the added hidden layer. The connection is fixed between the hidden and input layers, and the weight values are not updated. One hidden layer will be added again if the output value is not improved. When learning is judged to be sufficient, the hidden layer addition is stopped, as well as learning.



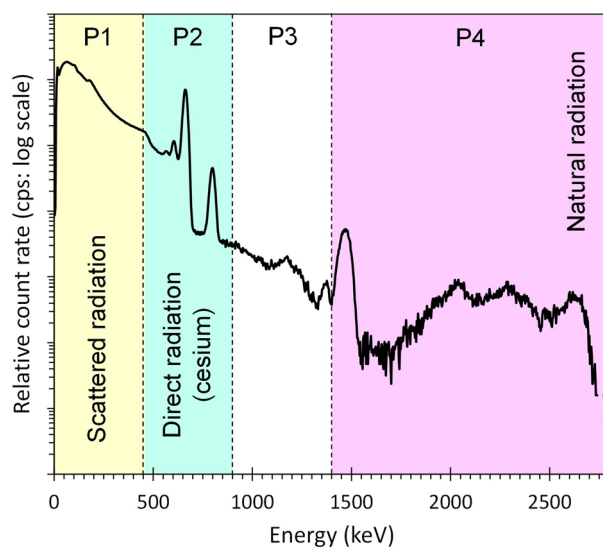
**Figure 2.** Flow of the network construction. The network was constructed by cascade correlation. The weight update method used was the adaptive subgradient method. The objective function was cross-entropy with added ridge regression.

The neural network was constructed with simple input variables and one hidden layer function for judging whether the UAV-survey data can be converted by the ANN. The output value of the network was compared with the target variable using the correlation coefficient ( $r$ ) which is defined as a criterion for sufficient learning. The learning was continued until no improvement of  $r$  was observed.  $r$  can be expressed as follows:

$$r = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum_i^n (x_i - \bar{x})^2)(\sum_i^n (y_i - \bar{y})^2)}} \tag{1}$$

where  $n$  is the total number of data;  $x$  is the converted value;  $y$  is the ground-based survey value (normalized target variable); and  $\bar{x}$  and  $\bar{y}$  are the averages for each value. The weight update method used was the adaptive subgradient method<sup>19</sup>. The objective function was cross-entropy with added ridge regression.

**Training set.** The radiation measurement data used for the ANN were acquired by the UAV (i.e., in this case, a type of unmanned helicopter from FAZER-R-G2, Yamaha Motor Co. Ltd., Shizuoka, Japan) around the FDNPP at the fiscal year of 2018 and 2019. The UAV was originally developed for spraying pesticides but was later adopted for radiation measurements. This UAV is operated manually during takeoff and landing and has a program operator for autonomous flight and an operator for the radiation detector. The UAV can conduct a programmed flight with the help of detailed self-localization using a real-time kinematic global positioning system. Its flight waypoints and altitude can be set. The detailed specifications of the UAV are given in a past research<sup>4,5</sup>. The dedicated radiation detector consisted of a LaBr<sub>3</sub> (Ce) scintillation detector (38 mm  $\phi$   $\times$  38 mm H  $\times$  three detectors) and detected gamma-rays in the energy range of 50–2800 keV (Japan Radiation Engineering Co. Ltd.,



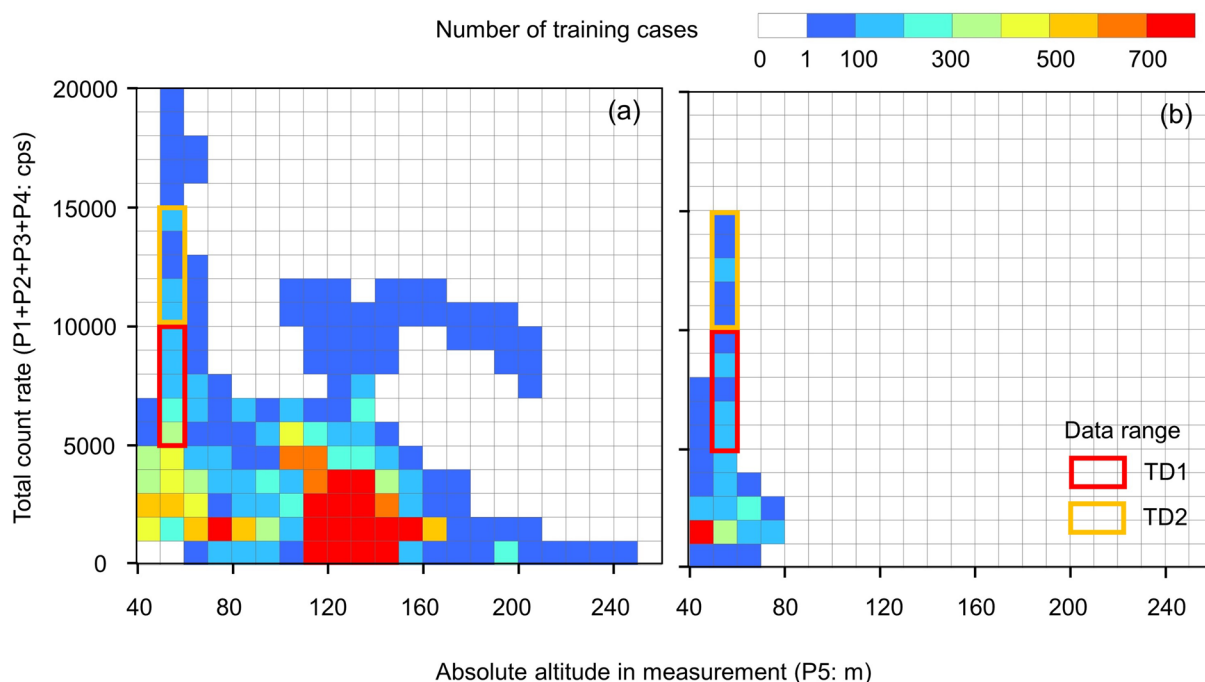
**Figure 3.** Typical gamma-ray spectrum of a UAV spectrometer with a LaBr<sub>3</sub> (Ce) sensor. This spectrum was obtained at 50 m agl. around the FDNPP.

Data type	Survey method	Parameter	Unit	Data range (Min–Max)		Description
				Training set (37,936 sets)	Test set (3442 sets)	
Input variable	UAV-based survey data					
		P1	cps	301–17,396	640–13,069	Count rate of 50–450 keV Scattered radiation
		P2	cps	35–1663	70–1350	Count rate of 450–900 keV Direct radiation from radiocesium
		P3	cps	14–27	16–26	Count rate of 900–1400 keV U and Th series natural radiation
		P4	cps	13–26	18–26	Count rate of 1400–2800 keV <sup>40</sup> K, U and Th series natural radiation
		GPS position data				
		P5	m (agl.)	44–244	43–75	Absolute altitude in measurement (Flight altitude)—(Surface altitude)
	Ground-based survey data					
Target variable		T1	μSv h <sup>-1</sup>	0.064–13	–	For training
Validation data		V1	μSv h <sup>-1</sup>	–	0.12–10	For validation

**Table 1.** List of training and test sets.

Ibaraki, Japan). Figure 3 depicts a typical gamma-ray spectrum. The UAV typically uses 50 m, 5 m s<sup>-1</sup> and 50 m agl. as the flight line space, flight speed and flight altitude, respectively. The typical flight altitude used to obtain the training set was 80 m, but the actual altitude above the ground ranged from approximately 40 m to 240 m to avoid buildings, mountains, steel towers, etc. A photographic survey using an unmanned helicopter was also conducted to construct a 3D orthophoto map for calculating the flight altitude above the ground or object surface. Each datum was averaged over a 10 m mesh.

Table 1 summarizes the input variables (training set). The count rate (i.e., P1, P2, P3, and P4) was obtained by dividing the gamma-ray spectrum into four regions (Fig. 3) to separate the following components: scattered radiation, P1; direct radiation from radiocesium, P2; and natural radiation, P3 and P4. P5 is the absolute altitude of the UAV, which is a major factor used to assess the air attenuation of gamma radiation. The absolute altitude was calculated by subtracting the surface altitude (DEM) from the flight altitude above sea level using the global navigation satellite system. The air dose rate at 1 m agl. (T1) was used as the target variable for the ANN and



**Figure 4.** Characteristics of the (a) training set (37,936 cases) and the (b) test set (3442 cases). The areas of TD1 and TD2 are the training cases presented in Fig. 6.

acquired using a handheld CsI (TI) (38 mm L × 38 mm W × 25 mm H) scintillation detector (Hot spot finder, Japan Shield Technical Research Co. Ltd., Fukushima, Japan). The training set (input and target variables) had 37,936 cases (Table 1 and Fig. 4a). Figure 4a illustrates a three-dimensional scatter diagram of the total count (P1 + P2 + P3 + P4), which is as an index of the radiation intensity, and the absolute altitude (P5) to enable us to understand the training set distribution. In this figure, the number of training cases on 120–150 m agl. of the absolute altitude in the measurement was relatively higher. In contrast, the condition of both high radiation intensity and high absolute altitude in the measurement was relatively small.

**Test set.** The test set different from the training set was used to evaluate the constructed ANN. The test set was obtained from Futaba-machi, which is located 3 km northwest from the FDNPP. This study area (1 km × 2 km) is still acknowledged as an evacuation area by the Japanese government. The flight line space, flight speed, and flight altitude of the UAV used were 50 m, 5 m s<sup>-1</sup>, and 50 m agl., respectively. For the output value validation, the ground-based survey data (V1) were obtained using a handheld survey meter at the same time as the UAV flight. All data were obtained for less than 8 h on August 19, 2018. Each datum was averaged over a 10 m mesh.

The test set had 3442 cases (Table 1). For comparison with the distribution of the training set, a three-dimensional scatter diagram of the total count (P1 + P2 + P3 + P4) and the absolute altitude (P5) was shown in Fig. 4b. The average flight altitude of the test set was 50 m. However, the actual altitude above the ground ranged from approximately 40 m to 80 m due to the avoidance of obstructions. Almost all test cases were understood to be within the range of the training set (Fig. 4a,b).

**Evaluation of the neural network condition.** The accuracy of the ANN method was expected to depend on the neural network condition (e.g., number of hidden layers and a distribution range of the training set). We conducted some verification tests to determine the effectiveness of the output validation by the neural network condition.

First, the output value validation was evaluated by limiting the number of hidden layers. NeuralWorks Predict automatically creates a hidden layer based on the correlation coefficient ( $r$ ) value. This software has the option for which the maximum layer is limited. Nine different types of networks (i.e., 0, 2, 4, 6, 8, 10, 12, 15, and 25 hidden layers) were constructed by the training set to optimize the number of hidden layers of a network. The test set was converted using each network. We calculated the root-mean-square error (RMSE) for nine types of network as follows using Eq. (2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - G_i)^2}{N}} \quad (2)$$

where  $N$  is the number of training cases;  $Y_i$  is the converted value (output value) of the UAV-survey data in mesh  $i$  ( $\mu\text{Sv h}^{-1}$ ; air dose rate at 1 m agl.); and  $G_i$  is the ground-based survey value (V1) in mesh  $i$  ( $\mu\text{Sv h}^{-1}$ ).

Second, we verified how many data cases can construct a network to converge the error between the ANN conversion value and the ground-based survey value using a limited training set. An ANN was constructed by

Data type	Total count rate* (cps)	P5 (m)	T1 ( $\mu\text{Sv h}^{-1}$ )	Number of cases
All training cases	370–19,099	43–244	0.064–13	41,378
TD1	0–5000	50–60	0.12–3.5	2469
TD2	5000–10,000	50–60	0.39–6.8	1544

**Table 2.** Limited training set for evaluating the neural network condition. \*Sum of P1, P2, P3, and P4.

randomly selecting data from a limited training set hereinafter referred to as “random training.” Two different kinds of data set were selected to evaluate the required number of training cases. The random training network was constructed by a limited training set (i.e., TD1 and TD2) (Table 2). The distribution of each random training set shown in Fig. 4 was selected by considering the radiation intensity or flight altitude of the test set. A random training dataset was extracted from the training (Fig. 4a) and test (Fig. 4b) sets. The number of training cases used for learning ranged from 50 to 1000 (50, 100, 200, 300, 400, 500, 750, and 1000). A random training network was constructed with the same settings. A total of 500 cases of data, other than those selected as the training set, were converted by the random training network to calculate the *RMSE* between the converted and ground-based survey values.

**ML-EM method.** The detailed procedure of the ML-EM method is given in the paper of Sasaki et al. (2019)<sup>10,11</sup>. The ML-EM method is a type of inverse estimation method that employs radiation information from various directions. The ML-EM method is theoretically expressed as follows by Eq. (3):

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^D C_{ij}} \sum_{i=1}^D \frac{y_i C_{ij}}{\sum_{j=1}^B C_{ij} \lambda_j^{(k)}} \quad (3)$$

where,  $k$  is the number of calculation iterations;  $j$  is the calculation position on the ground;  $B$  is the total number of ground-based calculation positions;  $i$  is the detection position; and  $D$  is the total number of detection positions.  $\lambda$  is the calculated intensity value on the ground.  $y_i$  is the measured count rate which is corrected by the detector response of radiocesium point source.  $C_{ij}$  is the parameter containing the attenuation coefficient.  $C_{ij}$  accounts for the attenuation by air, soil, and forest and is expressed as follows:

$$C_{ij} = F_x(x) \cdot F_\theta(\theta) \cdot F_h(h) \quad (4)$$

$$F_x(x) = x^{-\alpha_1} (x \leq 200) \quad (5)$$

$$F_x(x) = \frac{200^{-\alpha_1} \exp(-\alpha_2 x)}{\exp(-200\alpha_2)} (200 < x) \quad (6)$$

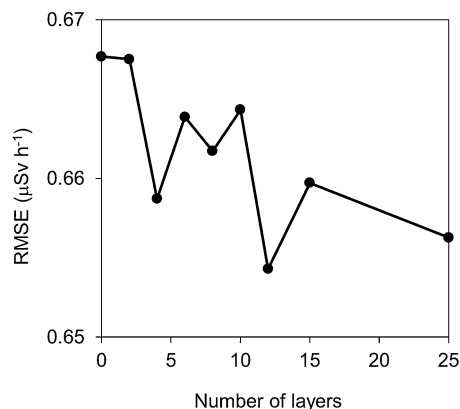
$$F_\theta(\theta) = 1 + A \exp[-((\theta - \gamma)/\sigma)^2] \quad (7)$$

$$F_h(h) = \exp(-\phi h) \quad (8)$$

$F_x$  is the attenuation rate due to air  $F_x$  and is calculated from the distance  $x$  (m) between  $i$  and  $j$  by using the particle and heavy ion transport code system (PHITS) developed by the Japan Atomic Energy Agency<sup>20</sup>.  $F_x$  was obtained from the decrease in the number of incident photons according to the distance from the point source of  $^{137}\text{Cs}$ . The calculation result yielded 1.97 and 0.010 as parameters  $\alpha_1$  and  $\alpha_2$ , respectively.  $F_\theta$  is the attenuation rate due to the effect of the soil scattering calculated from the angle  $\theta$  (deg) formed by  $i$ – $j$  and the soil surface. Parameters  $A$ ,  $\gamma$  and  $\sigma$  corresponded to the values of  $-0.727$ ,  $-35.2$ , and  $34.5$ , respectively.  $F_\theta$  was evaluated using the PHITS from the decrease in the number of incident photons when the angle between the soil surface and  $i$ – $j$  was changed upon centering on the point source of  $^{137}\text{Cs}$ .  $F_h$  is the attenuation rate due to forest shielding calculated from the distance  $h$  (m) that passes through the forest area between  $i$  and  $j$ .  $h$  is calculated using the DEM and the DSM. Accordingly, 0.061 was used for parameter  $\phi$ . Our previous research on the relationship between the air dose rate by the ground-based survey and the UAV-survey in the forest area<sup>11</sup> showed that  $F_h$  depended on the tree height.

**FSM method.** The count rate obtained by a dedicated UAV-survey detector ( $C_{all}$ : 50–2800 keV) was converted into the air dose rate ( $Y$ ) using a conversion factor ( $CD$ :  $0.00028 \mu\text{Sv h}^{-1} \text{ cps}^{-1}$ ) and an attenuation factor ( $AF$ :  $0.0061 \text{ m}^{-1}$ ). The  $CD$  was obtained in advance by comparing the count rate in the calibration area at altitude  $H_{std}$  (m) with that measured using a survey meter on the ground. The  $AF$  was calculated from the data obtained by varying the flight altitude on the ground from 20 to 150 m. Equation (9) shows the conversion expression in mesh  $i$ .

$$Y_i = (C_{all} - C_{BG}) \cdot CD \cdot \exp[-AF(H_{std} - H_{mi})] \quad (9)$$



**Figure 5.** RMSE of the ground-based survey value (V1) and the ANN conversion values when the number of hidden layers increases. This is the result of the test set converted to the ANN constructed by the training set.

where  $C_{BG}$  represents the background counts, including the cosmic ray and the natural radionuclides in the detector crystal, and  $H_m$  (m) is the flight altitude above the ground.

**Mapping.** Mapping was performed by supplementing unmeasured areas via the interpolation of the measured results. Various methods (e.g., kriging and spline approaches) were proposed for the interpolation, but the kriging method, which linearly assigns weights to the values of the measurement points and in inverse proportion to the distance, was applied herein to the UAV-survey data. The kriging method is easy to use when analyzing a large amount of data because its parameter setting is simple<sup>7</sup>. The interpolation processing was conducted using ArcGIS software (Environmental Systems Research Institute Inc., California, USA). The spatial resolution of the resulting contour map for the air dose rate was 10 m × 10 m.

**Validation method.** The ground-based survey data ( $G$ ) were compared with the air dose rate calculated using the UAV-survey ( $Y$ ). Both data were compared by visualizing the unevenness using a scatter diagram. The relative deviation ( $RD$ ) of each measurement cell was calculated as follows to evaluate the accuracy of the scheme used in this study:

$$RD_i = (Y_i - G_i)/G_i \quad (10)$$

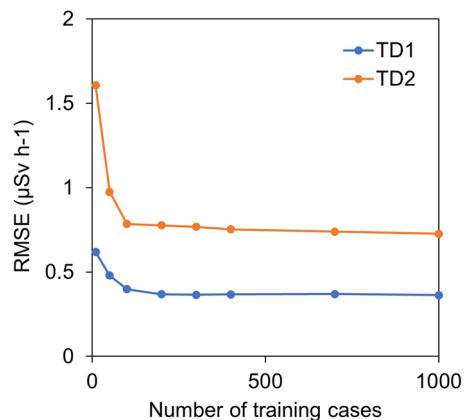
The calculated  $RD$ s were used to evaluate the total error and the statistical uncertainty shown as a histogram of frequency.

## Results and discussion

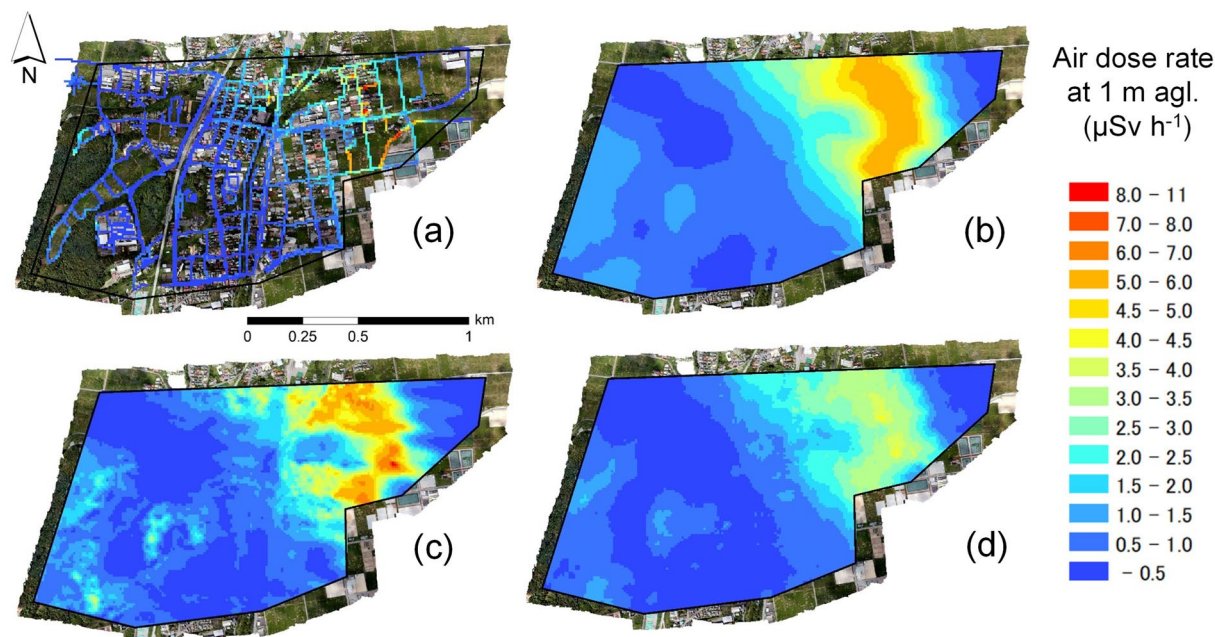
**Network construction.** Figure 5 shows the RMSE of the ground-based survey value (T1) and the output values under the condition of hidden layers in the network. The RMSE value tended to decrease with the layer increase, although some variation exists. This result suggests that the conversion accuracy improved as the number of hidden layers increased. In the case herein, a network with 12 hidden layers was most suitable for conversion from the UAV-survey data to the air dose rate on the ground. From this result, the network shown in Fig. 2 was constructed with 12 hidden layers for the test set calculation.

**Number of training cases.** Two types (i.e., TD1 and TD2) of the random training set were created to evaluate the relationship between the number of training cases and the error between the output value by the ANN and the ground-based survey value. The RMSE was calculated by comparing the air dose rate with the ground-based survey data that differed from the random training set as an evaluation index. Figure 6 shows the RMSE variation with an increasing number of random training cases. Both RMSEs converged when the number of training cases was approximately 200 or more in random training. The RMSE with two types of different training data sets saturated at 200; thus, this number of training case was considered as a rough criterion for the required number of training data set. In addition, the RMSE values were saturated with a different value, which was expected to affect the statistics error depending on the radiation intensity.

**Validation of the ANN method.** Figure 7a shows a 10 m mesh contour map of the air dose rate at 1 m agl. of the ground-based survey (V1). Figure 7b–d illustrate the results of converting the measurement values in sky to air dose rates at 1 m agl. using the FSM and those when using the ML-EM and the output value of using the ANN, respectively. This study area is a known path of the radioactive plume at the time of the accident<sup>6</sup>. The east side used as a residential region of this study area had a relatively high air dose rate of 2.0–10  $\mu\text{Sv h}^{-1}$ , while the forest region in the western area had a relatively low air dose rate of 0.5–3.0  $\mu\text{Sv h}^{-1}$ . The incline of the air dose rate for this study area was large compared with those from the other regions surrounding the FDNPP.



**Figure 6.** Relationship between the number of training cases and the error. This is the *RMSE* of the ground-based survey value and the ANN conversion values when the number of training cases increases in a random training network. TD1 and TD2 were constructed with the dataset shown in Table 2.

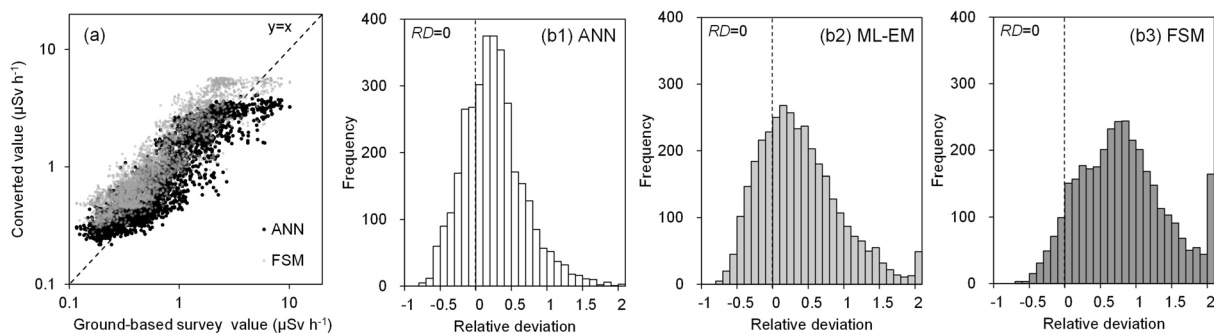


**Figure 7.** Air dose rate maps at 1 m agl. by (a) ground-based survey, (b) UAV-survey using the FSM, (c) UAV-survey using the ML-EM, and (d) UAV-survey using the ANN. This map was created using ArcGIS 10.5. <https://www.esri.com/ja-jp/arcgis/about-arcgis/overview>.

Figure 8 shows a comparison of the ground-based survey values (V1: Fig. 7a) and the three types of converted values (i.e., ANN, ML-EM, and FSM). The panel is a scatter diagram of the ANN and FSM values and the ground-based survey values in the same location. Three types of UAV-survey conversion values were correlated with the ground-based survey. The FSM tended to be relatively higher than the ground-based survey. The test set area in Fig. 7 was decontaminated in some areas after the accident; hence, the air dose rate at the decontamination area was locally decreased compared with the surroundings. In this case, the UAV-survey accuracy was expected to be affected by the changing field of view of the radiation detector<sup>21</sup>. On the contrary, the ML-EM method used the surrounding topographical and multiple measurement information. Therefore, the bias expressed as the *RMSE* or *RD* was better than the FSM. The ANN bias was reduced in the learning process to improve the error between the ground-based survey and output values. Figures 8(b1–3) present the *RD* histograms for quantitatively evaluating this tendency. Table 3 shows the *RMSE* with the ground-based survey and the interquartile ranges of the *RD* (i.e., 25%, 50% and 75%). The 50% interquartile range of the ANN (0.22) was near zero compared with the other two. In addition, the ANN had the smallest *RMSE* value, indicating that it was best in reproducing the ground-based survey value.

For a detailed evaluation of the ANN accuracy, the correlation with the ANN result and the number of training cases were evaluated and shown in Fig. 9. Figures 9(b1–4) depict examples for the correlation with the

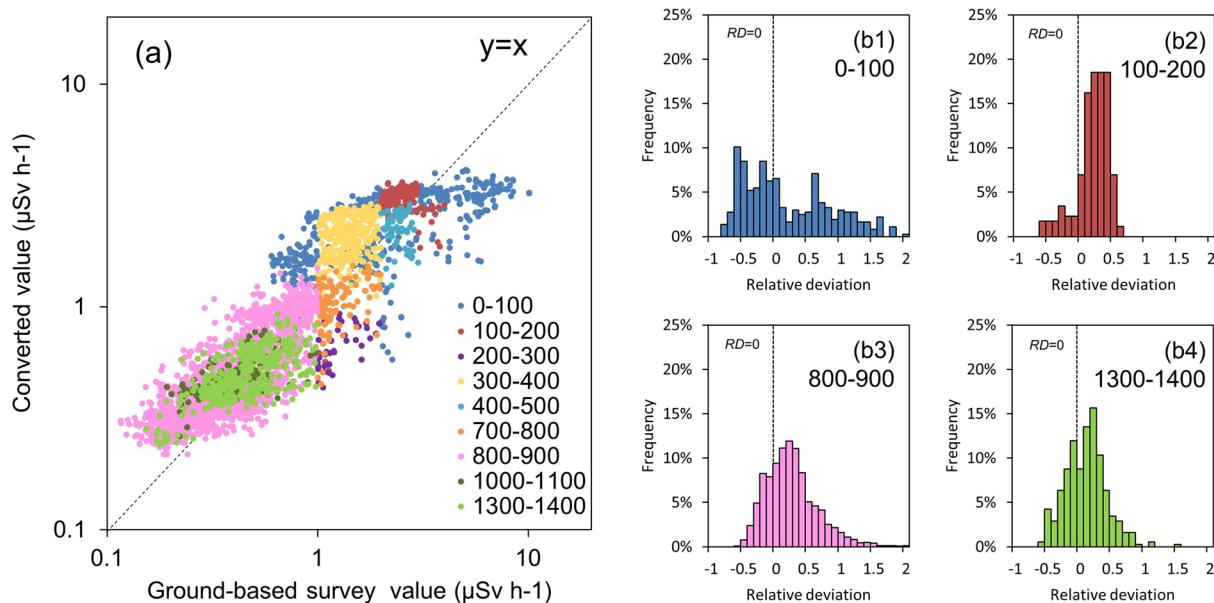




**Figure 8.** Comparison of the ground-based survey values and the three types of converted values (i.e., ANN, ML-EM, and FSM): (a) scatter diagrams of the ground-based survey value and the ANN and FSM values in the same location; and (b) histograms of the RDs (i.e., ANN, ML-EM, and FSM).

Parameters		ANN	ML-EM	FSM
Metrics	Range (%)			
RMSE		0.66	0.85	1.00
Interquartile range of RD	25	-0.04	-0.04	0.37
	50	0.22	0.30	0.78
	75	0.47	0.70	1.18

**Table 3.** Evaluation result of the converted value of the UAV-survey data in the FSM, ML-EM, and ANN. All data numbers are 3442.



**Figure 9.** Comparison of the ground-based survey values and the ANN conversion values for the number of training cases in the ANN of Fig. 7d: (a) scatter diagrams of the ground-based survey value and the ANN value; and (b) histograms of the RD for each number of training cases.

number of training cases and the RD of the ANN and ground-based values. The number of training cases was classified and counted every 10 m at altitude, every 5000 cps of the total count (P1 + P2 + P3 + P4), and every 1  $\mu\text{Sv h}^{-1}$  of the ground-based survey value (T1). The histograms show a few changes in the RD and RMSE with the change of the number of training cases. Table 4 presents the RMSE between the ground-based survey value and the ANN value, and the interquartile ranges of 25%, 50%, and 75% to quantitatively evaluate this tendency. The RMSE decreased as the number of training cases increased. As a result, the training case number which is

Number of training cases		0–100	100–200	200–300	300–400	400–500	700–800	800–900	1000–1100	1300–1400	All data
RMSE		1.72	0.78	0.65	0.77	0.52	0.35	0.18	0.13	0.15	0.66
Inter-quartile range	25%	– 0.36	0.12	– 0.53	0.28	– 0.18	– 0.30	0.01	0.08	– 0.09	– 0.04
	50%	0.02	0.28	– 0.48	0.44	– 0.01	– 0.15	0.24	0.22	0.16	0.22
	75%	0.70	0.41	– 0.41	0.69	0.07	– 0.03	0.49	0.43	0.32	0.47
Data number		366	173	39	302	60	109	1849	167	377	3442

**Table 4.** Evaluation result of the ANN value for each number of training cases. All data numbers are 3442.

necessary to calculate reliable converted value was 200 as well as the simulated verification as shown in Fig. 6. The tendency of the accuracy with the number of training cases was clarified from this application to the field data.

## Conclusion

In this study, we attempted to establish a new conversion method of the UAV-survey using published ANN software. The big data of the UAV-survey that we acquired after the FDNPP accident was used as the ANN training set. Constructing an ANN using the training set around the FDNPP resulted in an optimized hidden layer with 12 layers. The number of training cases was related to the error between the ground value and the ANN-converted value. The evaluation index between the ANN-converted value and the ground-based survey converged at 200 training cases; hence, this number was considered as a rough criterion of the required minimum number. It is thought that the number of training cases can be used as an index to evaluate the reliability of the ANN-converted value. The dose rate map made by the ANN method reproduced ground-based survey results much better than the traditional methods. However, this study used only the basic input variables in the radiation measurement and did not use terrain and photographic data as the input variables. The conversion accuracy is expected to be improved when these input variables are added. Future research will focus on the application of ANN conversion using terrain and photographic data and evaluate the effect of each input variable on conversion.

## Data availability

The DEM data set we used in the study can be found on the web site of Geospatial Information Authority of Japan<sup>17</sup>. The UAV-survey data set we used in the study can be found on web site of Nuclear Regulation Authority<sup>22</sup>. All data and the code that support the results within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Received: 31 July 2020; Accepted: 3 December 2020

Published online: 20 January 2021

## References

- Katata, G. *et al.* Detailed source term estimation of the atmospheric release for the Fukushima Dai-ichi Nuclear Power Station accident by coupling simulations of atmospheric dispersion model with an improved deposition scheme and oceanic dispersion model. *Atoms. Chem. Phys.* **15**, 1029–1070 (2015).
- Saito, K. *et al.* Summary of temporal changes in air dose rates and radionuclide deposition densities in the 80 km zone over five years after the Fukushima Nuclear Power Plant accident. *J. Environ. Radioact.* **210**, 105878 (2019).
- Sanada, Y., Sugita, T., Nishizawa, Y., Kondo, A. & Torii, T. The aerial radiation monitoring in Japan after the Fukushima Dai-ichi nuclear power plant accident. *Prog. Nucl. Sci. Technol.* **4**, 76–80 (2014).
- Sanada, Y., Urabe, Y., Sasaki, M., Ochi, K. & Torii, T. Evaluation of ecological half-life of dose rate based on airborne radiation monitoring following the Fukushima Dai-ichi nuclear power plant accident. *J. Environ. Radioact.* **192**, 417–425 (2018).
- Sanada, Y., Orita, T. & Torii, T. Temporal variation of dose rate distribution around the Fukushima Daiichi nuclear power station using unmanned helicopter. *Appl. Radiat. Isot.* **118**, 308–316 (2016).
- Sanada, Y. & Torii, T. Aerial radiation monitoring around the Fukushima Dai-ichi nuclear power plant using an unmanned helicopter. *J. Environ. Radioact.* **139**, 294–299 (2015).
- IAEA, Guidelines for radioelement mapping using gamma ray spectrometry data. IAEA-TECDOC-1363 (2003).
- Ishizaki, A., Sanada, Y., Ishida, M. & Munakata, M. Application of topographical source model for air dose rates conversions in aerial radiation monitoring. *J. Environ. Radioact.* **180**, 82–89 (2017).
- Tyler, A. N., Sanderson, D. C. W., Scott, E. M. & Allyson, J. D. Accounting for spatial variability and fields of view in environmental gamma ray spectrometry. *J. Environ. Radioact.* **33**, 213–235 (1996).
- Sasaki, M., Ishizaki, A. & Sanada, Y. Development of analysis method for airborne radiation monitoring using the inverse problem solutions. *Prog. Nucl. Sci. Tech.* **6**, 63–67 (2019).
- Sasaki, M., Sanada, Y. & Yamamoto, A. Application of the forest shielding factor to the maximum-likelihood expectation maximization method for airborne radiation monitoring. *Radiat. Protect. Dosim.* **184**, 400–404 (2019).
- Hinton, G. E., Osindero, S. & Ten, Y. W. A fast learning algorithm for deep belief nets. *Neural Comput.* **18**, 1527–1554 (2006).
- Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **10**, 257–273 (2017).
- Watanabe, K., Matsushita, Y., Kobayashi, I. & Kuroda, N. Artificial neural network modeling for predicting final moisture content of individual Sugi (*Cryptomeria japonica*) samples during air-drying. *J. Wood. Sci.* **59**, 112–118 (2013).
- Walczak, S. & Velanovich, V. Prediction of perioperative transfusions using an artificial neural network. *PLoS ONE* **15**, e0229450 (2020).
- Kangas, L. J., Keller, P. E., Siciliano, E. R., Kouzes, R. T. & Ely, J. H. The use of artificial neural networks in PVT-based radiation portal monitors. *Nucl. Inst. Meth. Phys. Res. A.* **587**, 398–412 (2008).
- GSI (Geospatial information authority of Japan), Maps & Geospatial Information. [http://www.gsi.go.jp/ENGLISH/page\\_e30031.html](http://www.gsi.go.jp/ENGLISH/page_e30031.html) (Accessed July 1, 2020)
- Fahlman, S. E. & Lebiere, C. The cascade-correlation learning architecture. *Adv. Neural. Inf. Process. Syst.* **2**, 524–532 (1990).

19. Duchi, J., Hazan, E. & Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **12**, 2121–2159 (2011).
20. Sato, T. *et al.* Particle and heavy ion transport code system, PHITS, Version 252. *J. Nucl. Sci. Technol.* **50**(9), 913–923 (2013).
21. Malins, A., Okumura, M., Machida, M. & Saito, K. Topographic effects on ambient dose equivalent rates from radiocesium fallout. In: Proceedings of Joint International Conference on Mathematics and Computation, Supercomputing in Nuclear Applications and the Monte Carlo Method (M&C + SNA + MC 2015) (2015).
22. Nuclear Regulation Authority, Monitoring information of environmental radioactivity level. <https://radioactivity.nsr.go.jp/en/> (Accessed July1, 2020).

### Author contributions

M.S. performed statistical analysis and manuscript preparation. M.S., Y.S., E.W.K and A.Y. discussed the results. M.S. and Y.S. carried out the monitoring. Every authors have given approval to the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to M.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021