(2) AI Technology in Nuclear Energy Field

Follow-up Training Course (FTC) Indonesia

On Reactor Engineering (RE)

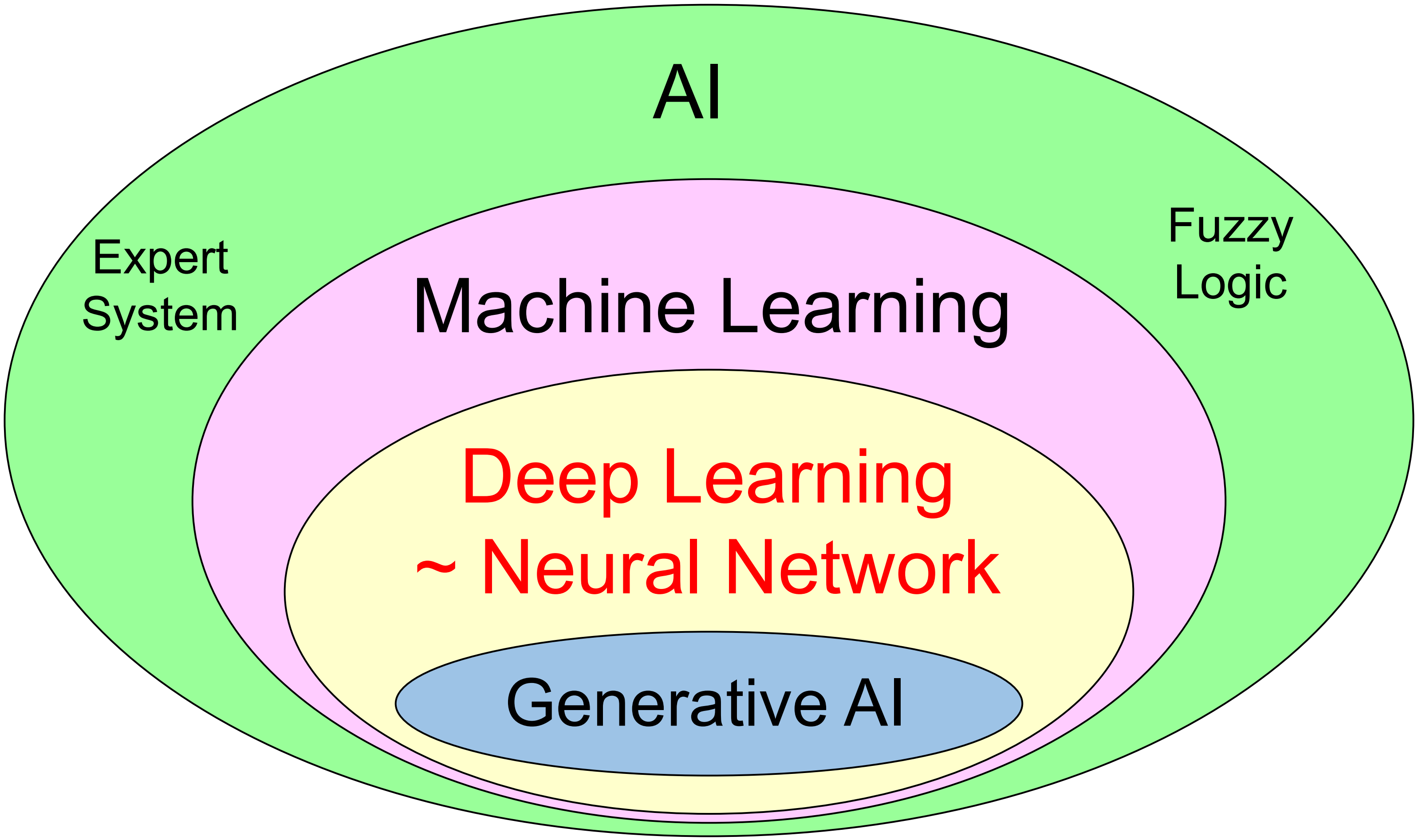
21 February 2025

Kunihiko NABESHIMA

E-mail : nabeshima.kunihiko@jaea.go.jp

Nuclear Human Resource Development Center
Tokai, Japan Atomic Energy Agency (JAEA)

Artificial Intelligence



History of Neural Network Model

● First Boom

1958~ Perceptron by Rosenblatt
2 layers (Input and Output)
Perceptron can only solve linearly separable problems

● Second Boom

1986~ Multi Layer Perceptron with Backpropagation Learning Algorithm by Rumelhart and Hinton

Problem of Overfitting and Vanishing Gradient

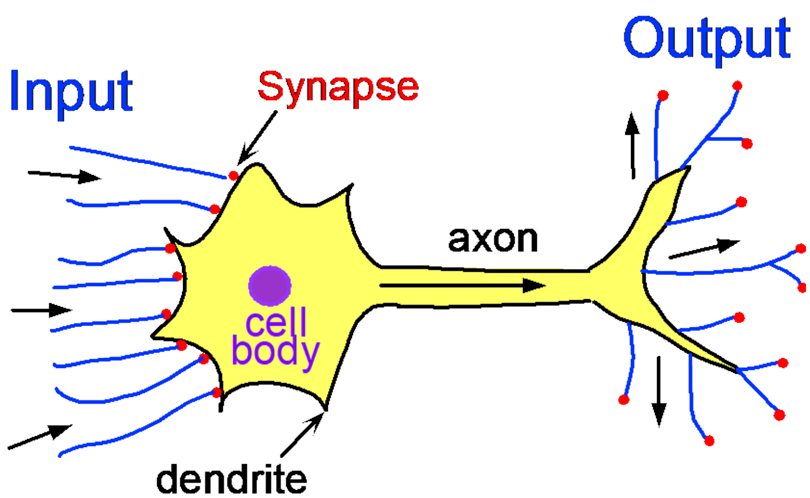
● Third and Fourth(?) Boom

2006~ Deep Learning (more than 4 layers) by Hinton, et al.
Dropout, ReLU (rectified linear unit)

2017~ Generative AI* (Transformer)
ChatDPT(OpenAI), Gemini(Google), Copilot(Microsoft)

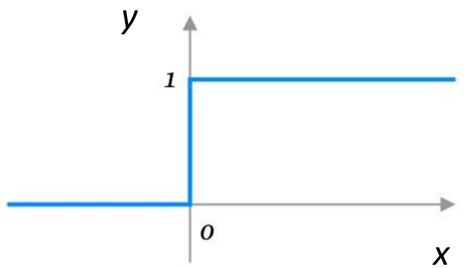
* AI requires a lot of electricity.

Neural Network Model



(a) Neuron Model

100~200 Billion Neurons in Brain

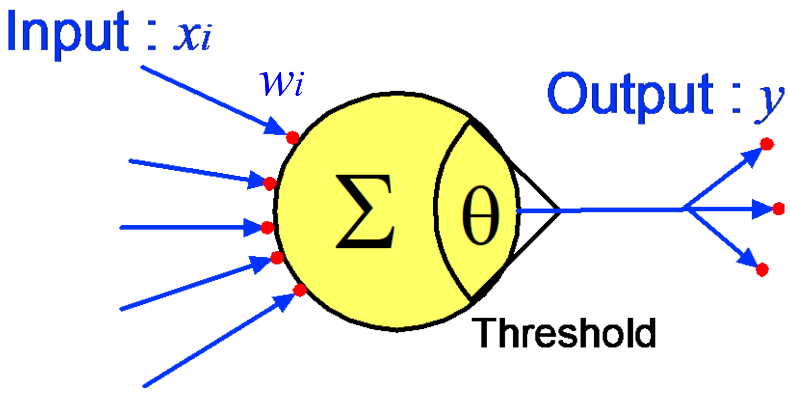


Sum of Input ≥ θ … Output=1
Sum of Input < θ … Output=0

θ : Threshold Level

Simple Behavior → Connection is important!

Neural Network Model



(b) Mathematical Model

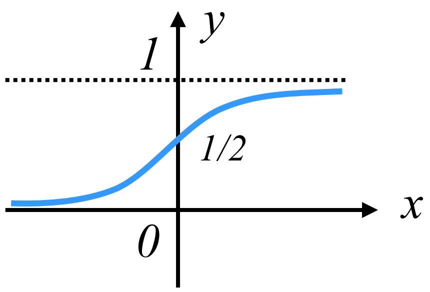
Sum of Input : $z = \sum_i w_i x_i$

w_i: $weight$

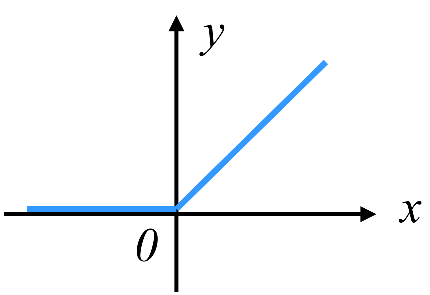
Output : $y = f(z - \theta)$

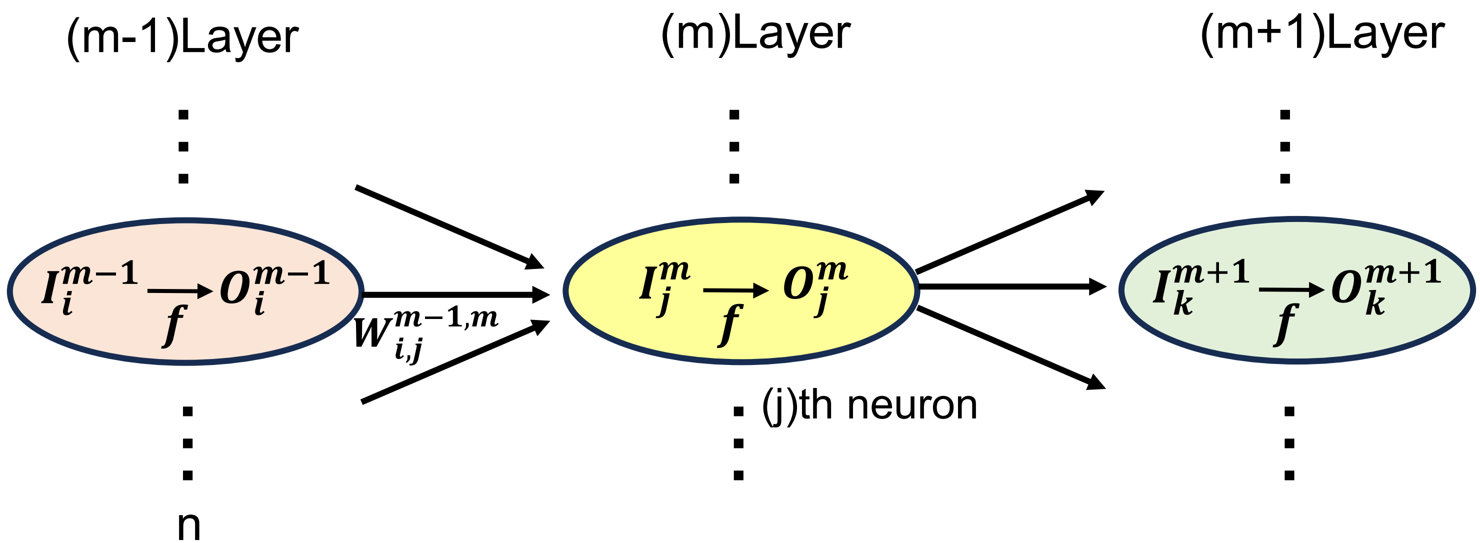
Activation Function f

Sigmoid $\quad \boldsymbol{f(x) = \frac{1}{1 + e^{-x}}}$



ReLU $\quad \boldsymbol{f(x) \begin{cases} 0 & (x < 0) \\ x & (x \geq 0) \end{cases}}$





$$O_j^m = f\left(I_i^m - \theta_j^m\right)$$

$$= f\left(\sum_{i=1}^{n} W_{i,j}^{m-1,m} \cdot O_i^{m-1} - \theta_j^m\right)$$

$\boldsymbol{I_j^m}$: Sum of Input of (m)Layer (j)th neuron

$\boldsymbol{O_j^m}$: Output of (m)Layer(j)th neuron

$\boldsymbol{W_{i,j}^{m-1,m}}$: Weight from (m-1)Layer (i)th neuron to (k)Layer (j)th neuron

$\boldsymbol{\theta_j^m}$: Threshold Level of (m)Layer (j)th neuron

$\boldsymbol{f}$: Activation Function

Learning: Optimization of weight (w) and threshold level (θ)

Backpropagation Learning Algorithm

Modification of weight and threshold level at (t) learning cycle

$$W_{i,j}^{m-1,m}(t) = W_{i,j}^{m-1,m}(t-1) + \Delta W_{i,j}^{m-1,m}(t)$$

$$\Delta W_{i,j}^{m-1,m}(t) = \boldsymbol{\eta} \cdot \delta_j^m \cdot O_i^{m-1} + \boldsymbol{\alpha} \cdot \Delta W_{i,j}^{m-1,m}(t-1)$$

$$\Delta\theta_j^m(t) = \eta \cdot \delta_j^m \cdot \theta_j^m + \alpha \cdot \Delta\theta_j^m(t-1)$$

η : Learning rate

δ_j^m : Error at (m)Layer (j)th neuron

α : Coefficient of Momentum term

Error at Output Layer (M)

$$\delta_j^M = (Y_j - O_j^M) \cdot \underline{f'(I_j^M)} = \underline{O_j^M \cdot (1 - O_j^M)} \cdot (Y_j - O_j^M)$$

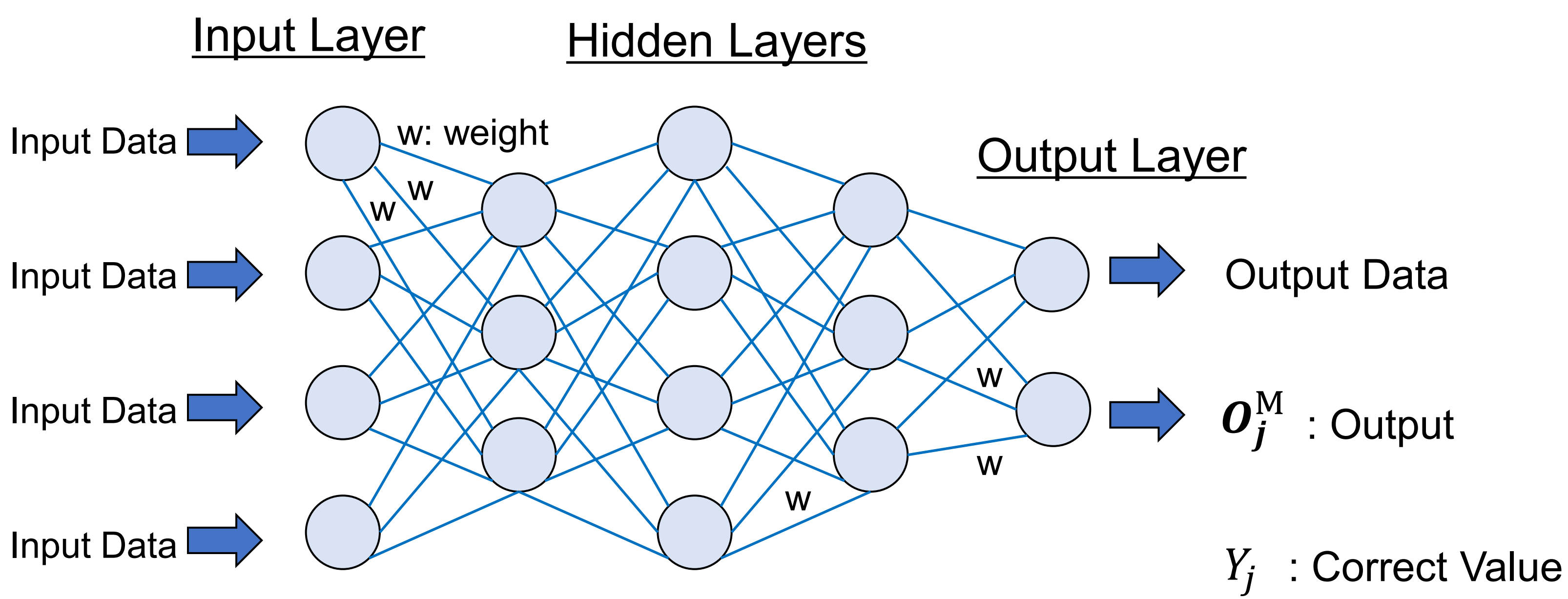
Sigmoid Function

$$\boldsymbol{f(x) = \frac{1}{1+e^{-x}}}$$

Error at Hidden Layer (m)

$$\delta_j^m = \underline{f'(I_j^m)} \cdot \sum_m W_{j,k}^{m,m+1} \cdot \delta_k^{m+1} = \underline{O_j^m \cdot (1 - O_j^m)} \cdot \sum_p W_{j,k}^{m,m+1} \cdot \delta_k^{m+1} \qquad (m = M-1, \ldots 2)$$

Backpropagation Learning Algorithm



Error at Output Layer (M)

$$\delta_j^M = (Y_j - O_j^M)\cdot \underline{f'(I_j^M)} = \underline{O_j^M\cdot(1-O_j^M)}\cdot(Y_j - O_j^M)$$

Sigmoid Function

$$\boldsymbol{f(x) = \frac{1}{1+e^{-x}}}$$

Error at Hidden Layer (m)

$$\delta_j^m = \underline{f'(I_j^m)}\cdot\sum_m W_{j,k}^{m,m+1}\cdot\delta_k^{m+1} = \underline{O_j^m\cdot(1-O_j^m)}\cdot\sum_p W_{j,k}^{m,m+1}\cdot\delta_k^{m+1} \qquad (m = M-1,\dots 2)$$

Recent trends in Deep Learning

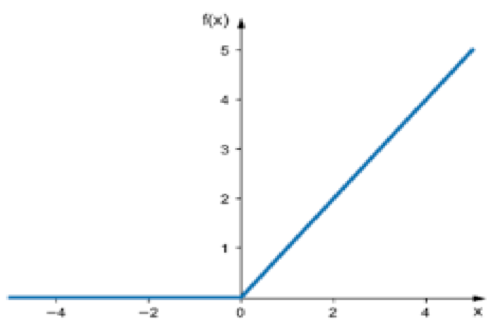
Disadvantages of Backpropagation Algorithm

Overfitting: Training data can be predicted correctly, but the error in test data is larger.

Vanishing Gradient : As the number of layers increases, early layers receive little or no updated weight information during backpropagation, and the learning speed slows down.

Solution

- Dropout:
  Some nodes are disabling during learning
- ReLU (Rectified Linear Unit) function is selected as activation function.
  f(x)=0, if x<0
  f(x)=x, if x≥0



Recent situation in Deep Learning

- Accelerating computer: Improving the performance of not only CPU but also GPU in particular.
- increased HDD capacity : Big data can be stored and processed.
- A lot of free Scientific Libraries and Tools by Python-Anaconda

□ Build a large network (a lot of inputs, outputs and layers)

□ Accelerate learning speed

□ Easy programing

□ Create a lot of simulation data for learning and testing

□ Store a lot of learning and testing data

Typical Deep Learning Networks

① **Convolutional neural network (CNN)** : A feature vector is extracted from an image in image processing. It is a dimension deleter and is good at image recognition.

② **Autoencoder** : It does not process complicated information as it is. It abstracts the data to reduce the amount of data, so large-scale learning is possible.

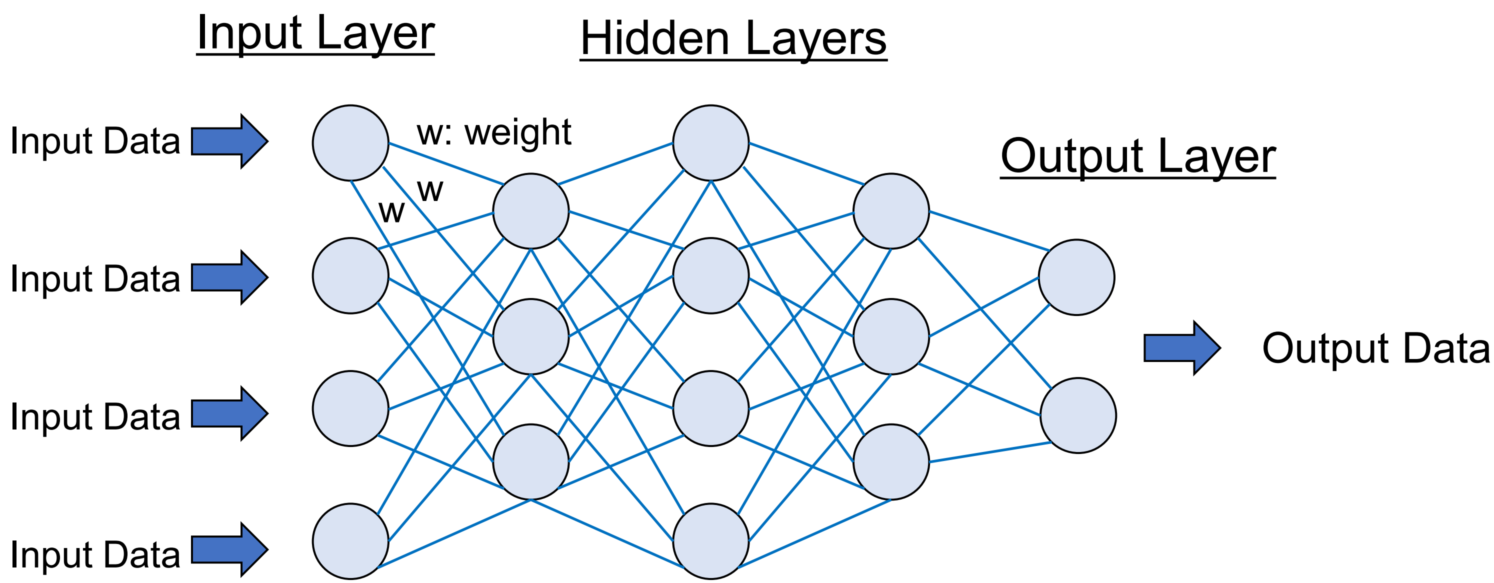
③ **LSTM (Long Short Term Memory)** : It is good for time series data and language processing. (a kind of RNN)

④ **GAN (Generative Adversarial Network)** : It is a hostile generation network of unsupervised learning that creates models (images) by reverse coupling of autoencoders. It consists of two networks, a generator and a classifier.

⑤ **Transformer** : It is useful for sequential data or natural language processing. Most of generative AI use this type.

Basic learning algorithm is back propagation

Deep Learning has made any modeling possible, but "Why these learning models work so well?"



Some Scientists

Elucidate learning model

- Structure of Brain
- Philosophy?

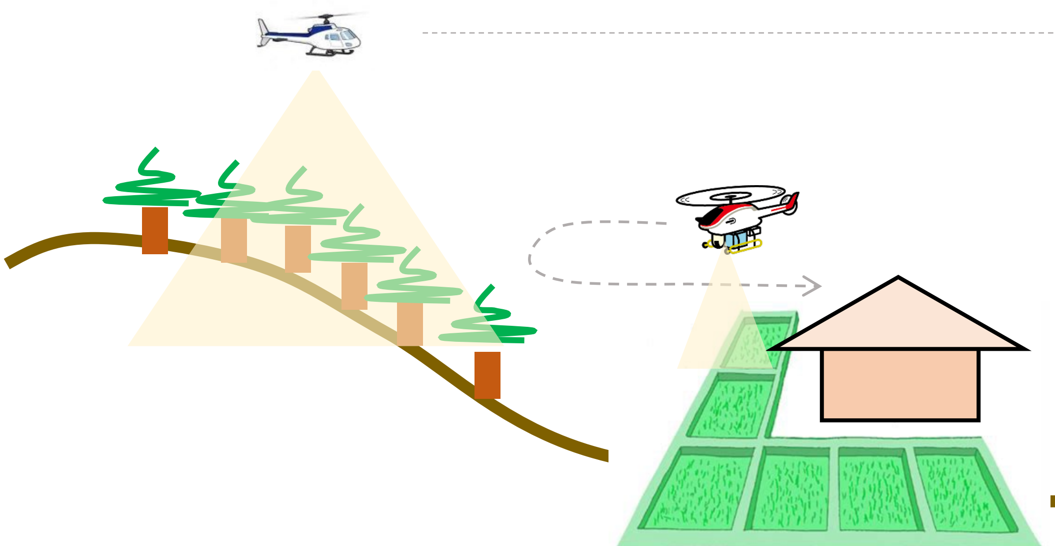
Other Scientists

Use deep learning as a tool

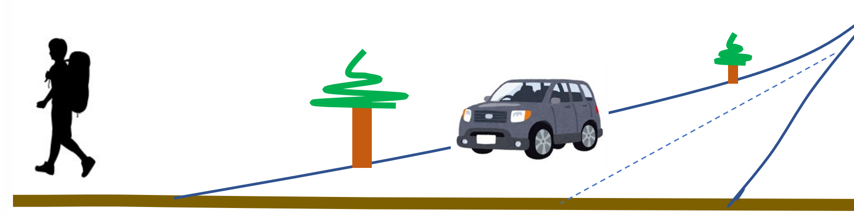
- **PyTorch** developed by FB (Meta)
- TensorFlow developed by Google
- MATLAB Deep Learning Toolbox

Application 1: Application of the artificial neural network (ANN) to the airborne radiation survey

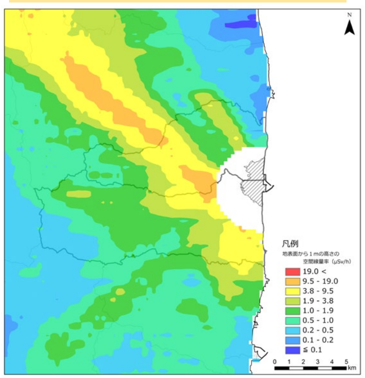
Monitoring methods of ambient dose equivalents



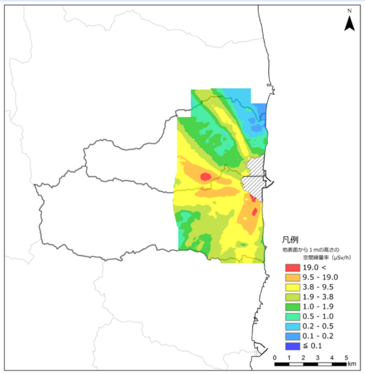
	Heli	UAV	Walk	Car
Mobility	◎	○	×	△
Accuracy	×	△	◎	○
Cost	△	○	○	○



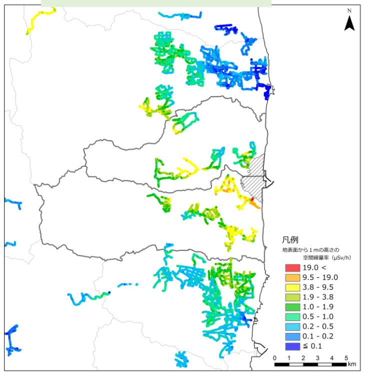
Manned helicopter



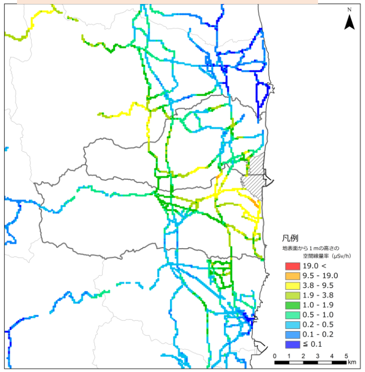
Unmanned Aerial Vehicle (UAV)



Walk survey

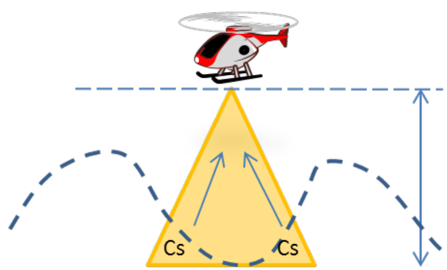


Carborne survey

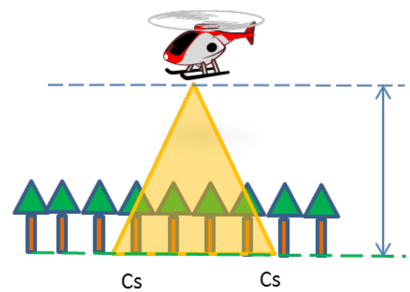


Dose rate mapping 1 m above the ground level around Fukushima Daiichi NPS by airborne radiation survey

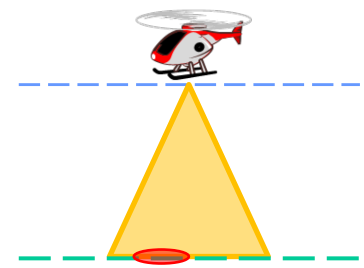
In the conventional method, the reenactment of air dose rate is not well at the area which deviates from the hypothetical conditions.



The unevenness of the terrain



Shielding effect (trees, etc.)



Local distribution of source

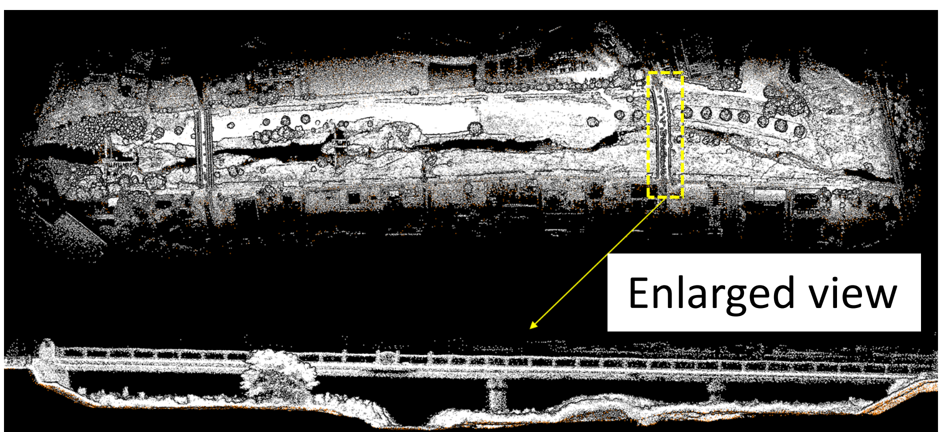
Inverse problem analysis

- long calculation time
- much time for creating parameters work

Walking survey
(Backpack survey)

Advances are being made in UAV-based ground surveying technology.

Laser surveying



Photogrammetry

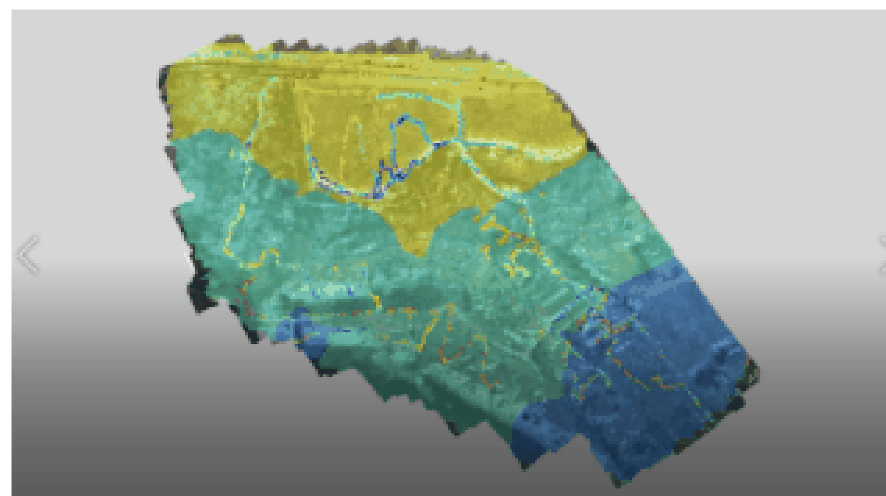
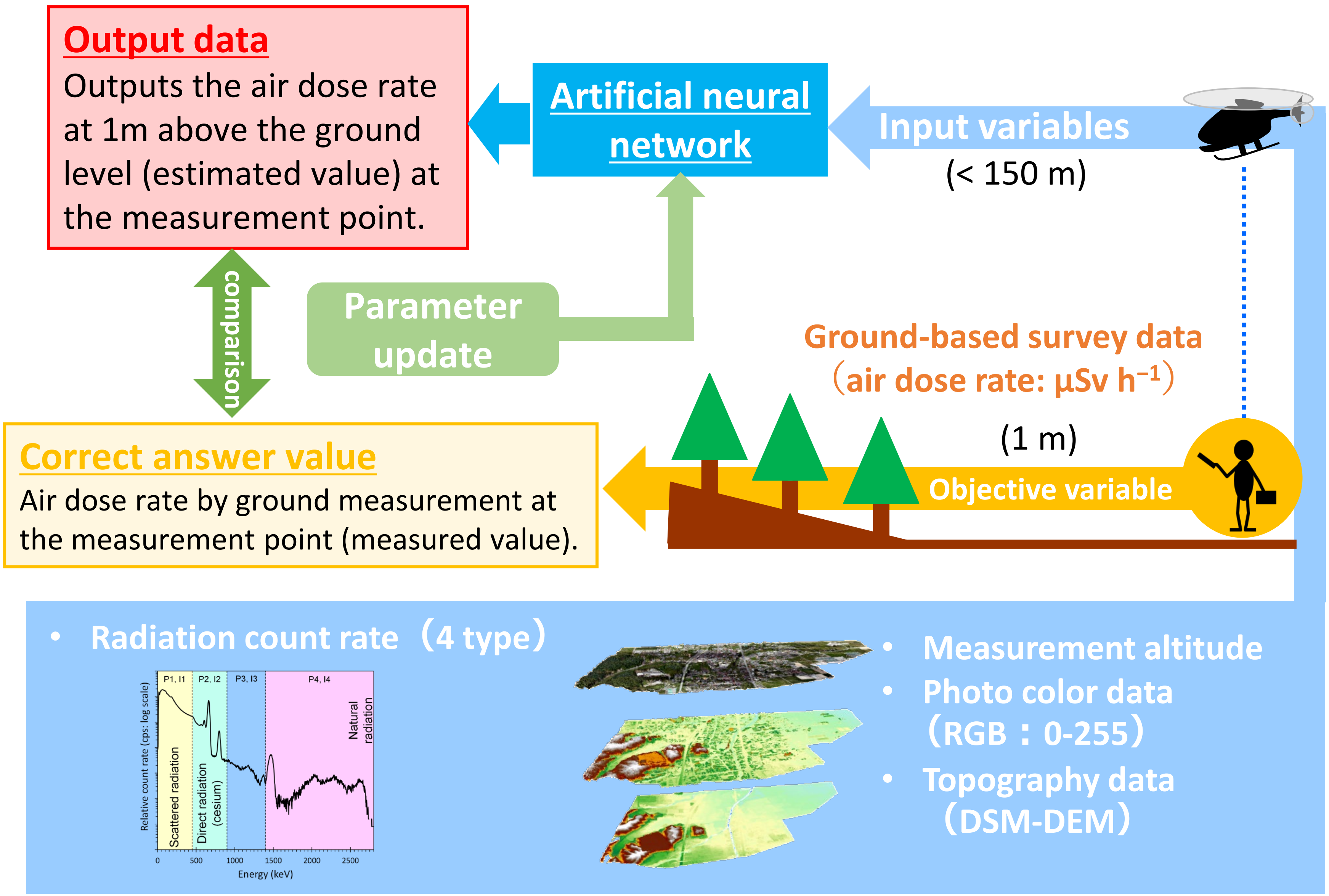


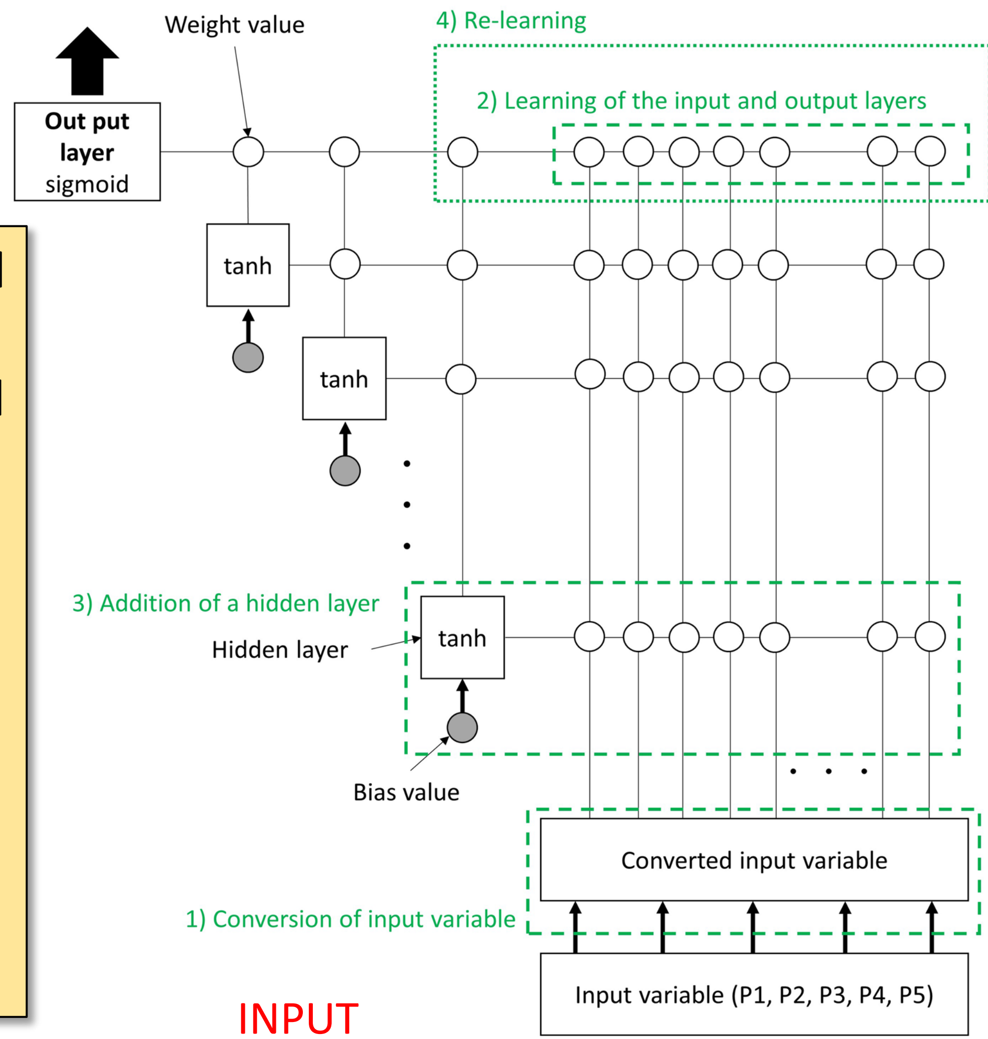
Image of Applying ANN



https://doi.org/10.1038/s41598-021-81546-4

Flow of the network construction

OUTPUT:
・Air dose rate at ground level



- The network is constructed by cascade correlation.
- The weight update method used is the adaptive sub-gradient method.
- The objective function is cross-entropy with added ridge regression.
- "NeuralWorks Predict" (commercially available software) is used for this application.

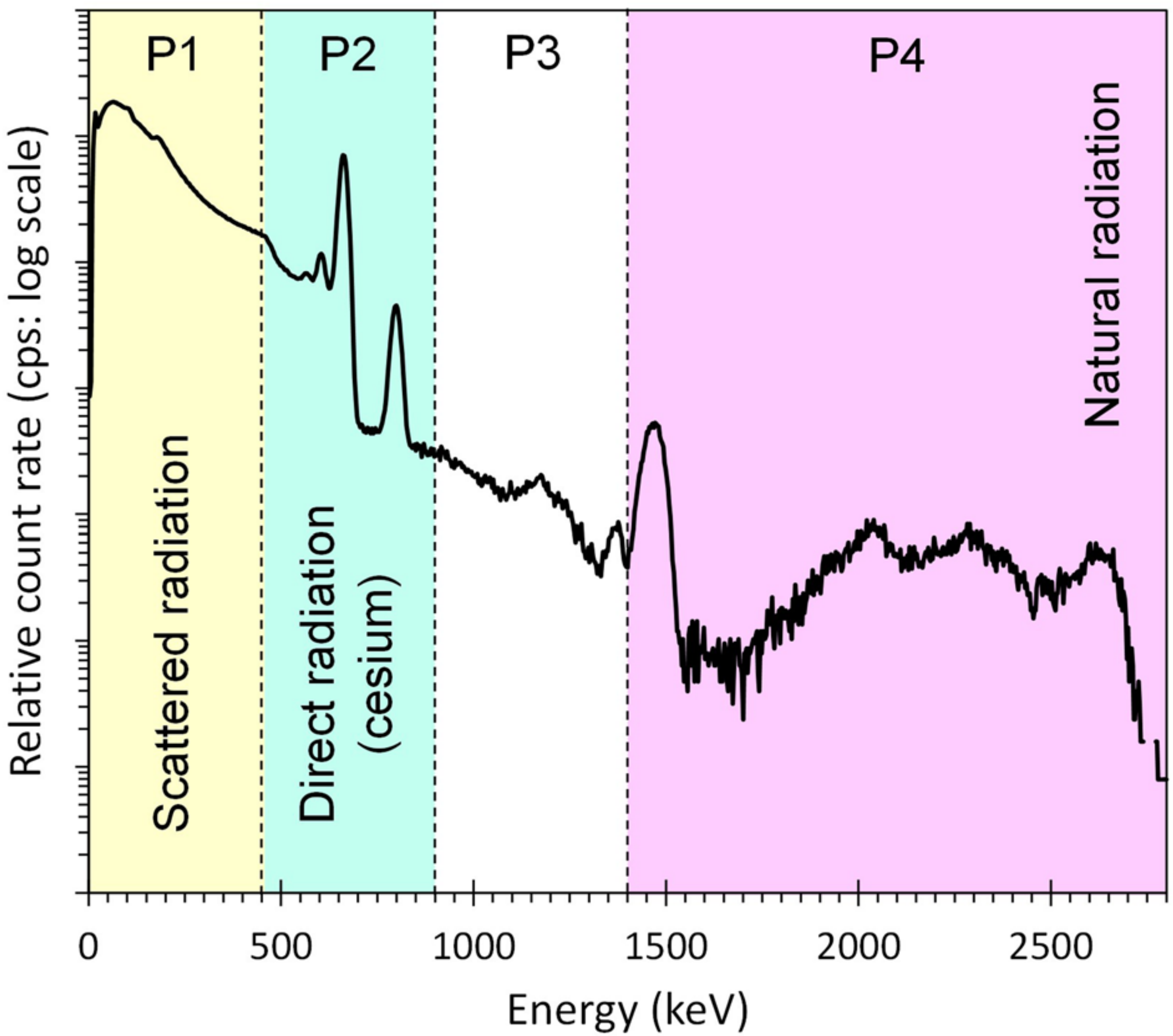
INPUT

P1-P4: Count rate of γ-ray spectral data divided by energy

P5: Distance from ground surface to helicopter position

https://doi.org/10.1038/s41598-021-81546-4

Typical gamma-ray spectrum of a UAV* spectrometer with a LaBr3 (Ce) sensor. (This spectrum was obtained at 50 m agl. around FDNPS.)



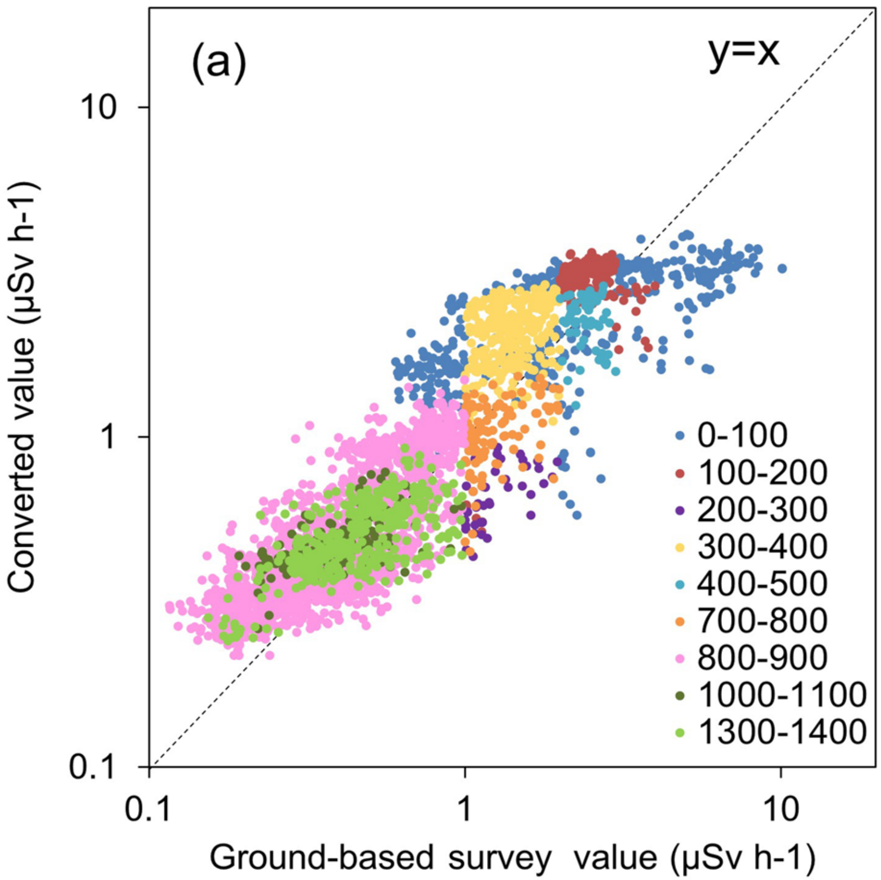
*UAV: Unmanned Aerial Vehicle (drone)

agl: Above the ground level

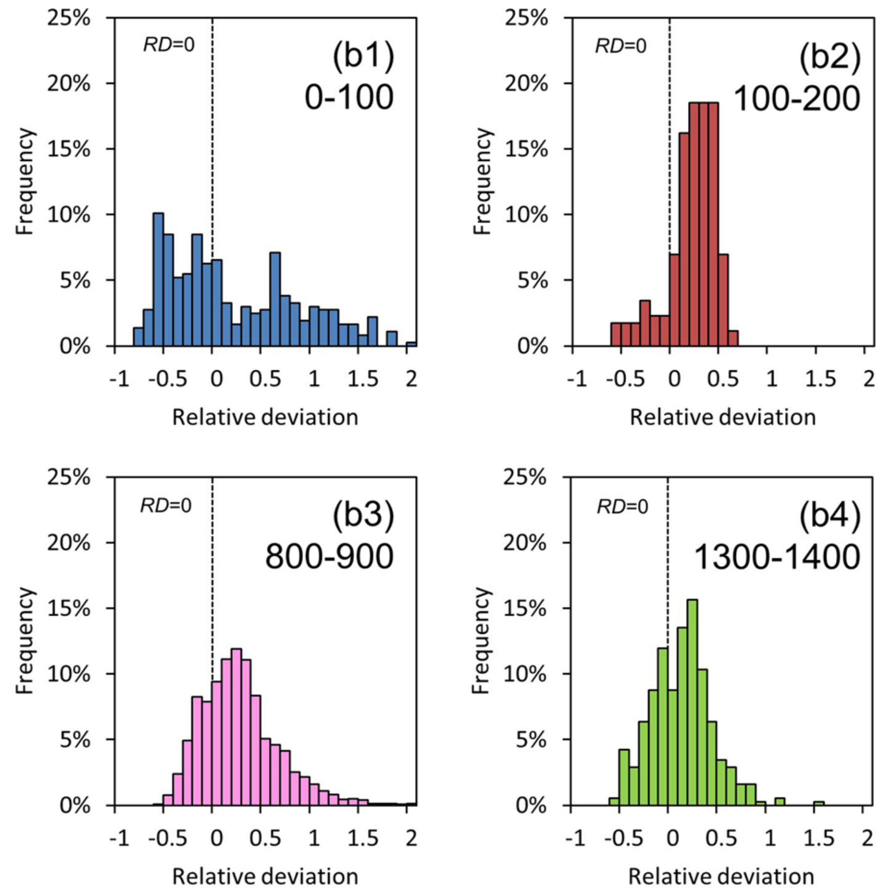
FDNPS: Fukushima Daiichi Nuclear Power Station

https://doi.org/10.1038/s41598-021-81546-4

Comparison of the ground-based survey values and the ANN conversion values for the number of training cases



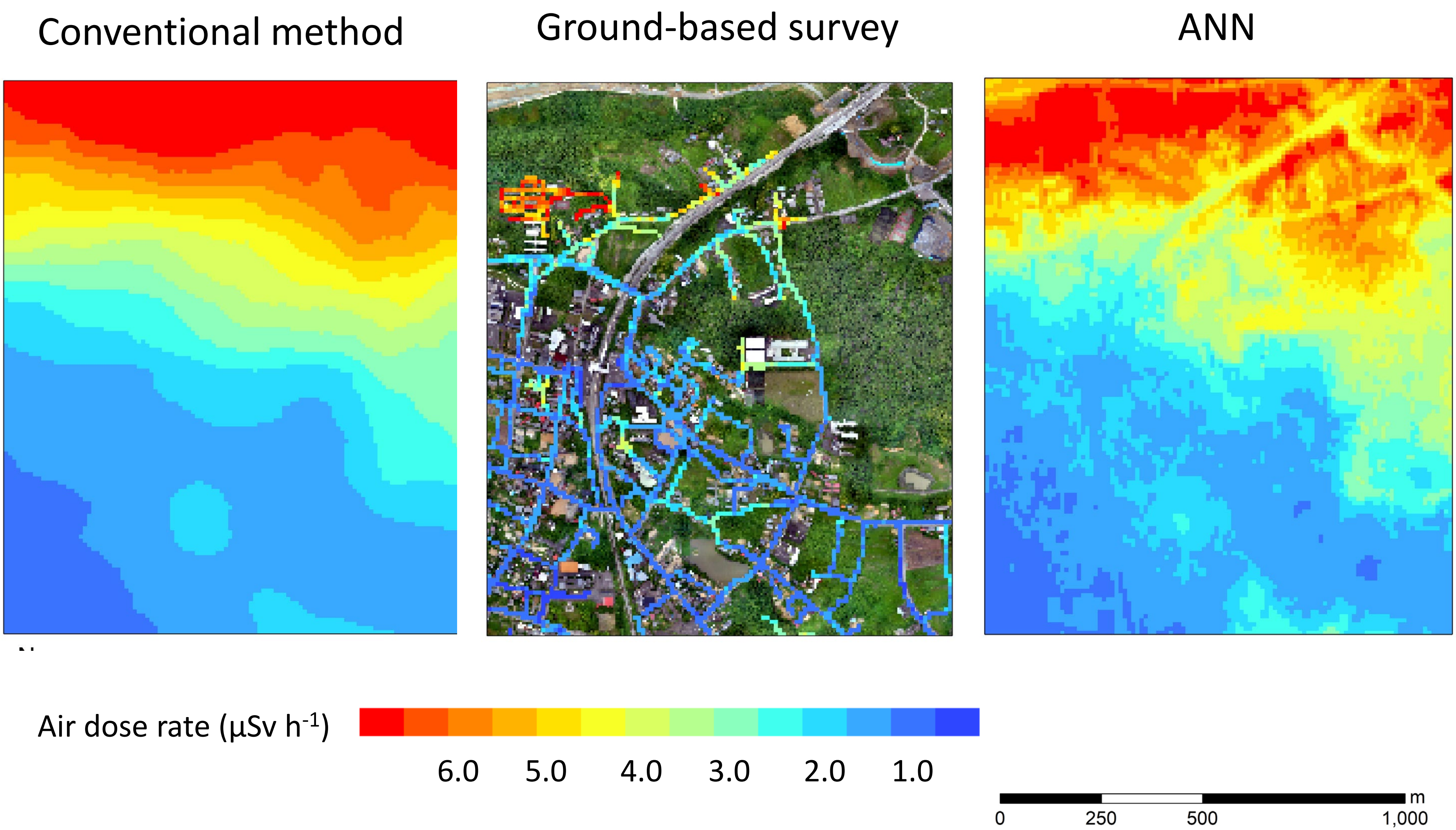
(a) scatter diagrams of the ground-based survey value and the ANN value



(b) histograms of the RD for each number of training cases

https://doi.org/10.1038/s41598-021-81546-4

Dose rate maps using machine learning

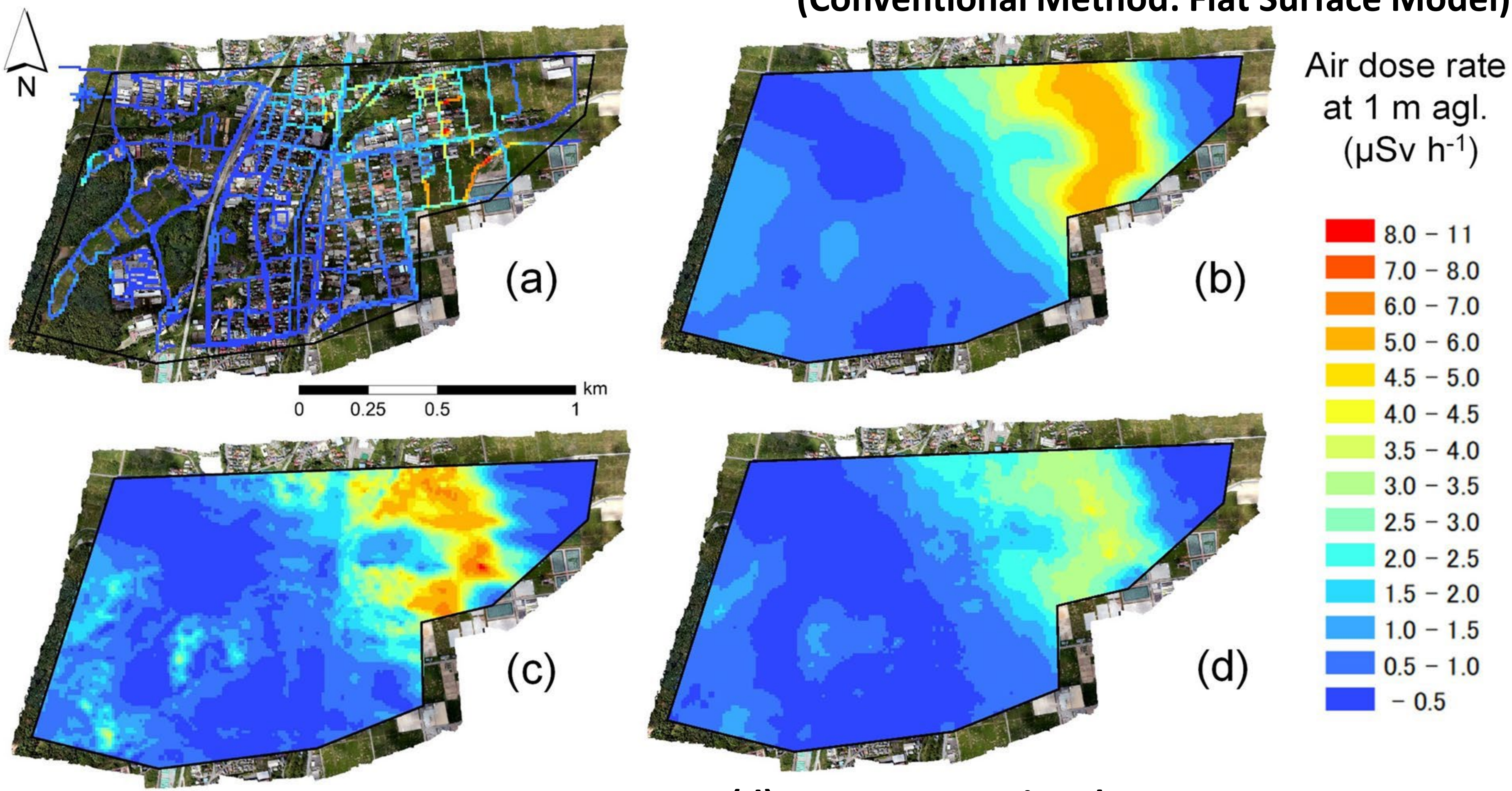


Air dose rate maps at 1 m above the ground level (agl.) by ArcGIS*

* https://www.esri.com/ja-jp/arcgis/about-arcgis/overview

(a) ground-based survey

(b) UAV-survey using the FSM (Conventional Method: Flat Surface Model)

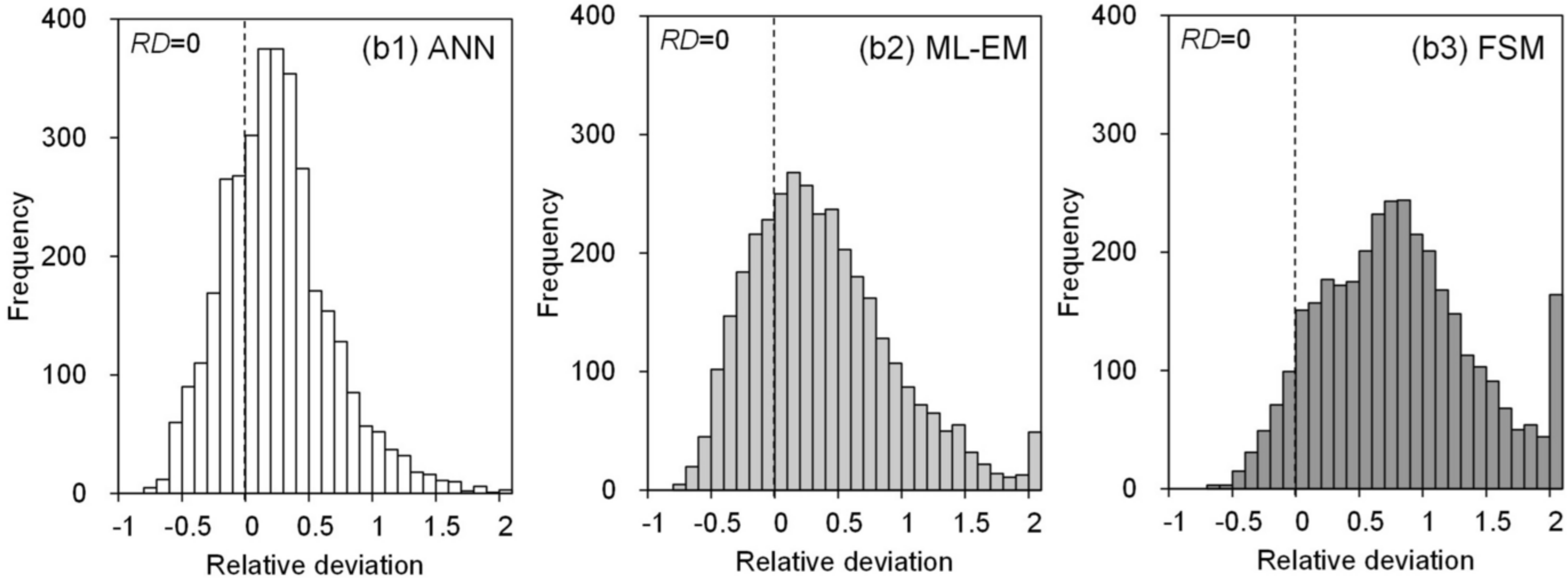


(d) UAV-survey using the ANN

(c) UAV-survey using the ML-EM (Inverse Problem Solution)

https://doi.org/10.1038/s41598-021-81546-4

Comparison of the ground-based survey values and the three types of converted values



$$Relative\ Deviation\ = \frac{Converted\ Value\ - Gland\_based\ Survey\ Value}{Gland\_based\ Survey\ Value}$$

https://www.nature.com/articles/s41598-021-81546-4

Application 2: Ultrasonic Testing in Fast Reactor

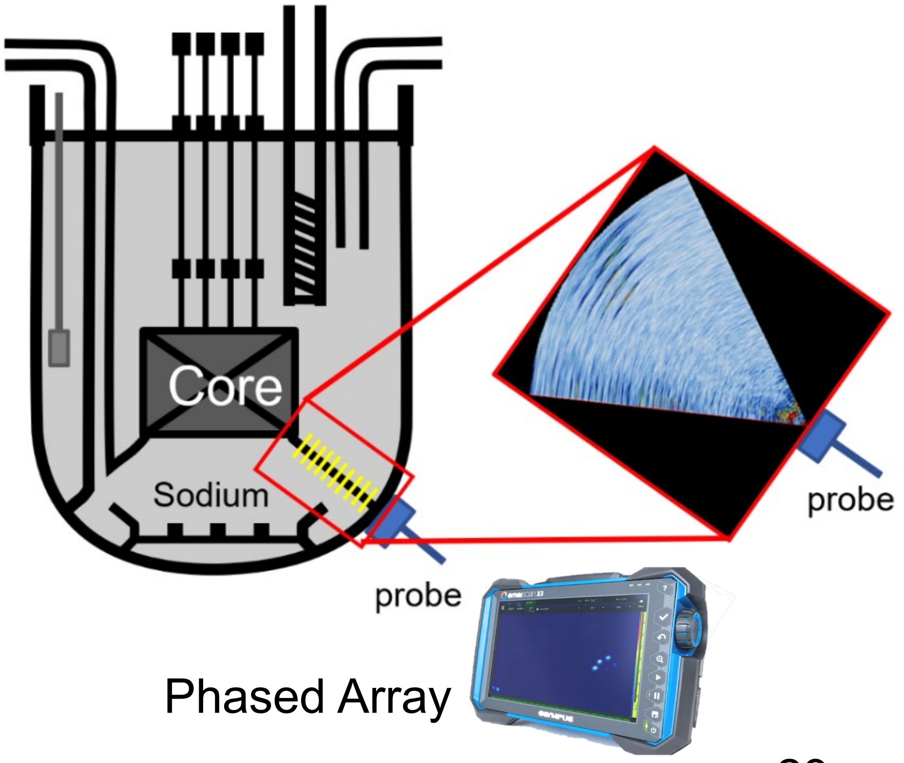
- Difficult to detect cracks of core support structure in fast reactor
  - Sodium is chemically active and invisible
  - It is difficult to extract sodium in the reactor vessel for inspection

➡ Ultrasonic Testing

- Difficult to detect from
  - Outside of Reactor Vessel
  - Long Distance

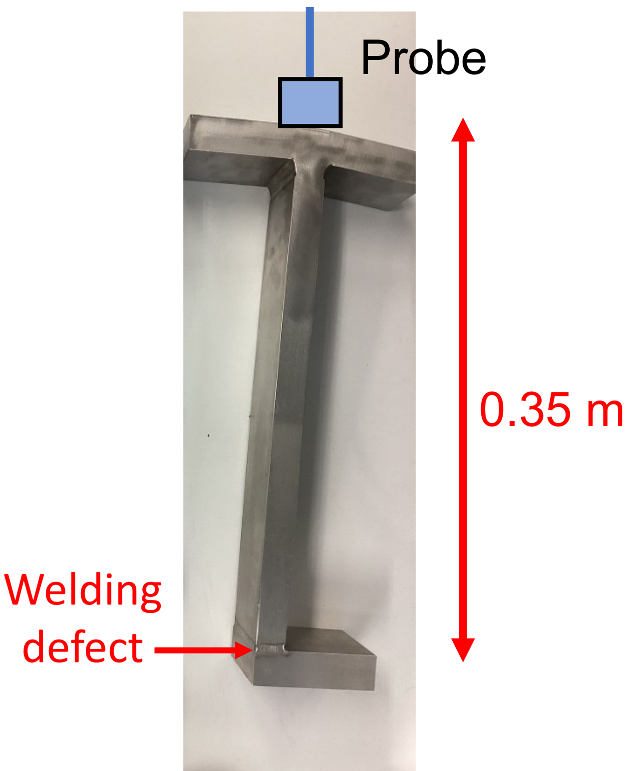
because of large “noise”

➡ Deep Learning (CNN: AlexNet)

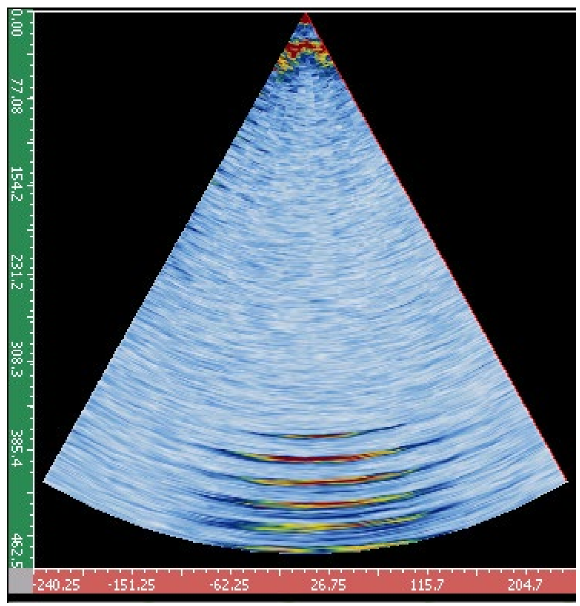


Images Data by Phased Array

- Actual length of core support structure is 1.1 m
  - Ultrasonic waves are attenuated
- We attempted to classify whether welding defects exist or not using small scale specimens (0.35 m)

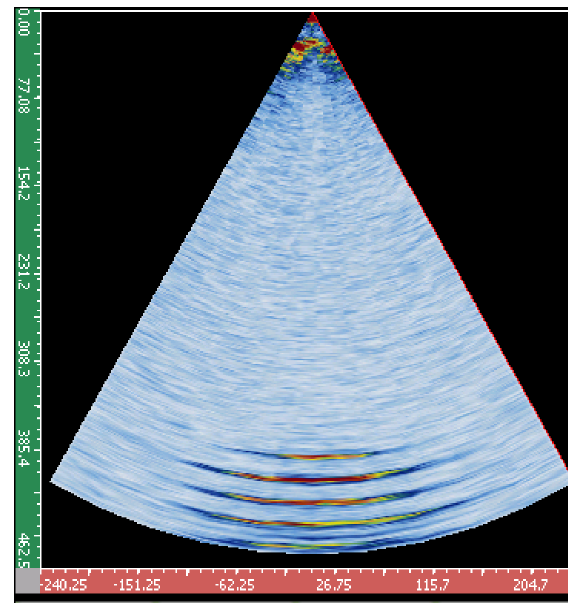


54 images



with welding defect

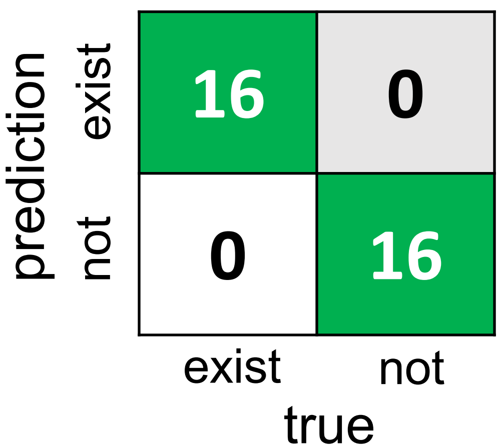
54 images

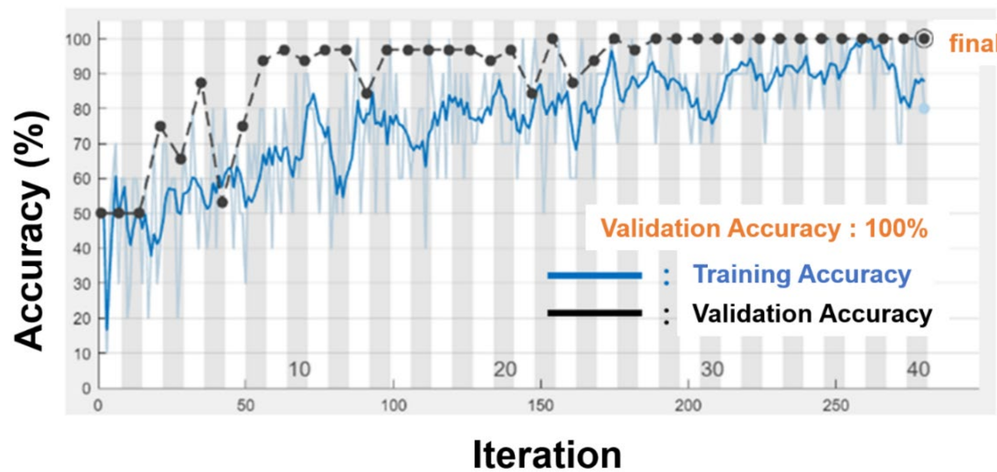


no welding defect

Transfer Learning for 0.35 m Specimens

- Prerequisite for transfer learning
  - 108 (54 images with and without welding defect) images are prepared
  - 70% (76 images) for training, 30% (32 images) for validating
  - initial learn rate→1.0×10^{-5}, epoch size→40, minibatch size→10
  - AlexNet is applied as pre-trained network





- There is no misclassification
- Criteria for classification is unknown → Explainable AI

Y. OTA, et al., "Explainable Machine Learning to Identify Flaws in Supporting Structures of Fast Reactor", NTHAS 13 (2024)

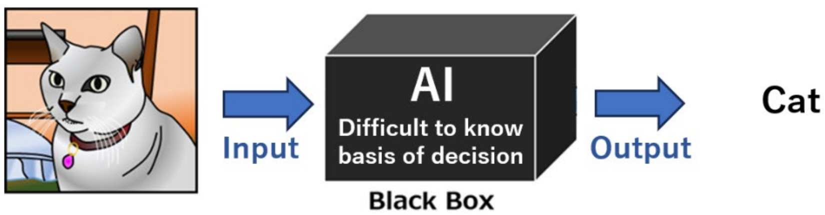
Transfer Learning, LIME Method

- Machine Learning enables us to determine whether welding defect exist in unknown data automatically
  - Transfer learning uses pre-train network (Alexnet etc.) and re-training for particular classification

◆ Disadvantage

➢ Unclearness of criteria

➡ Explainable AI



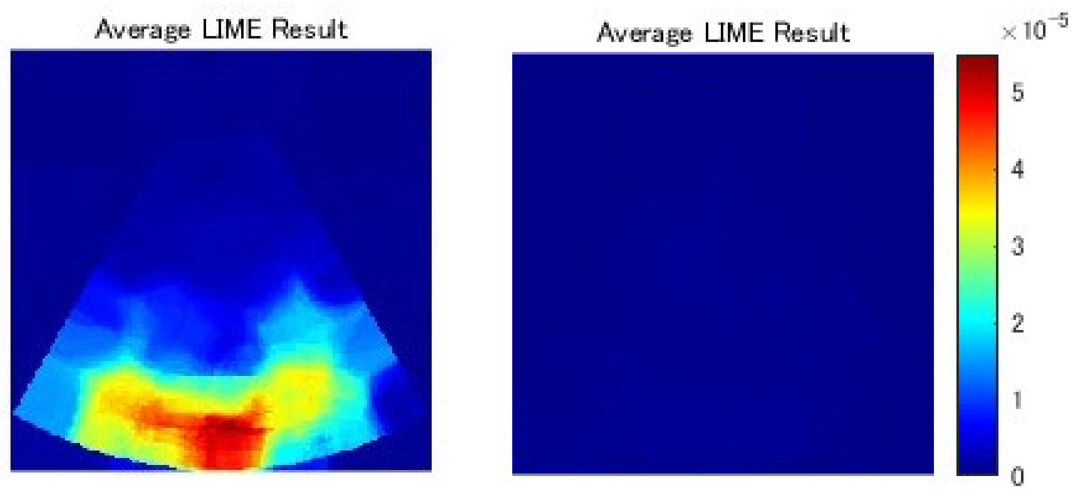
◆ LIME method enable us to visualize the criteria for classification by hiding part of images

Criteria of classification for 0.35 m specimens

- Prerequisite for LIME method
  - The number of hyperparameter → 50, Hide ratio → 40%
- LIME method is applied to phased array images individually → Overlap the results
- The bottom of the images are essential location for classification → Criteria is correct

importance

important area



with welding defect

no welding defect

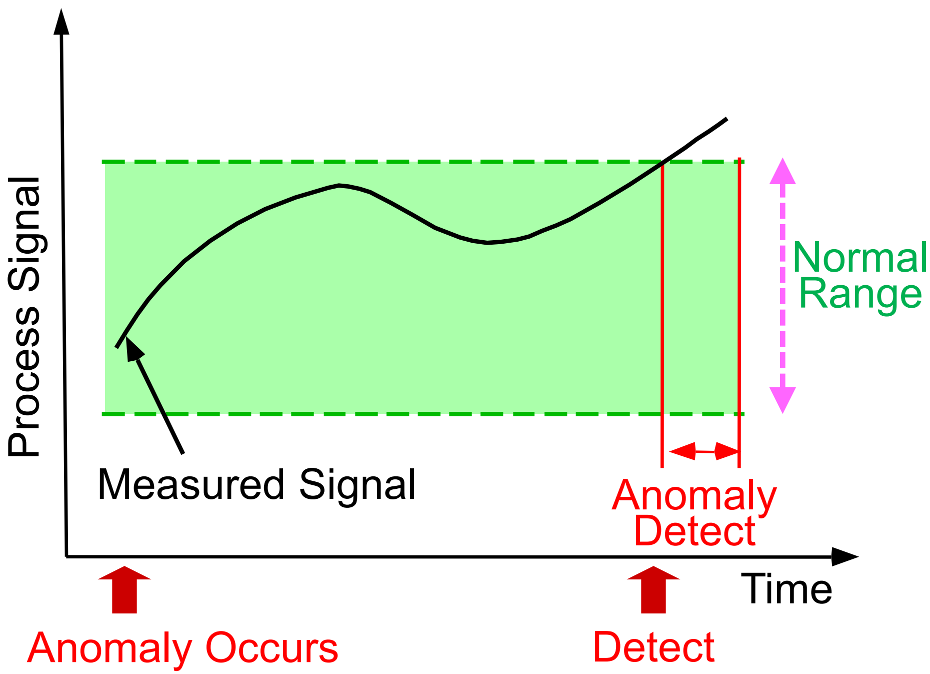
Application 3: Reactor monitoring

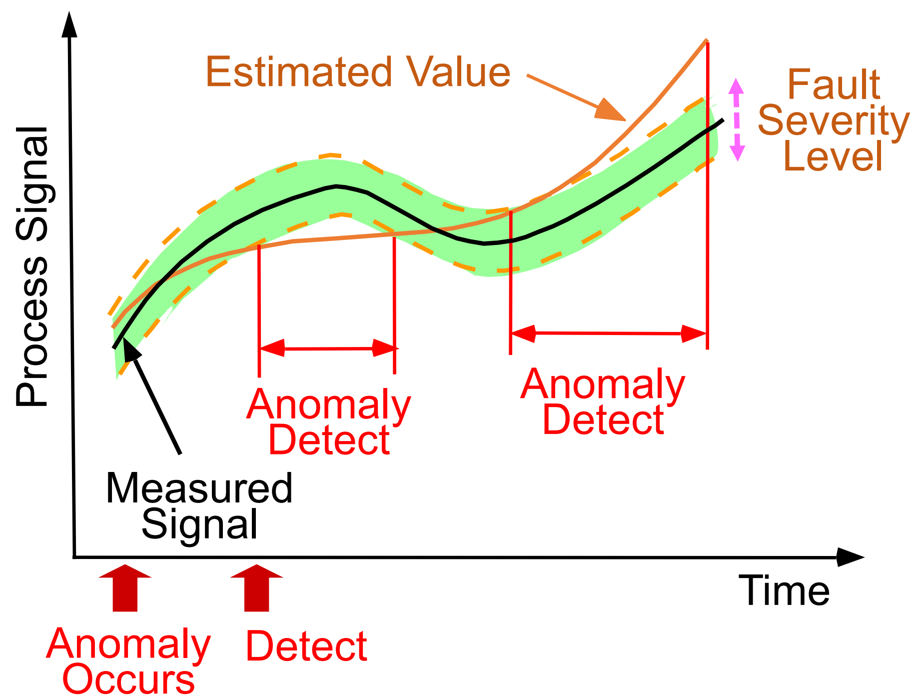
Conventional Alarm system

- Wide normal range
- Long time to detect anomalies

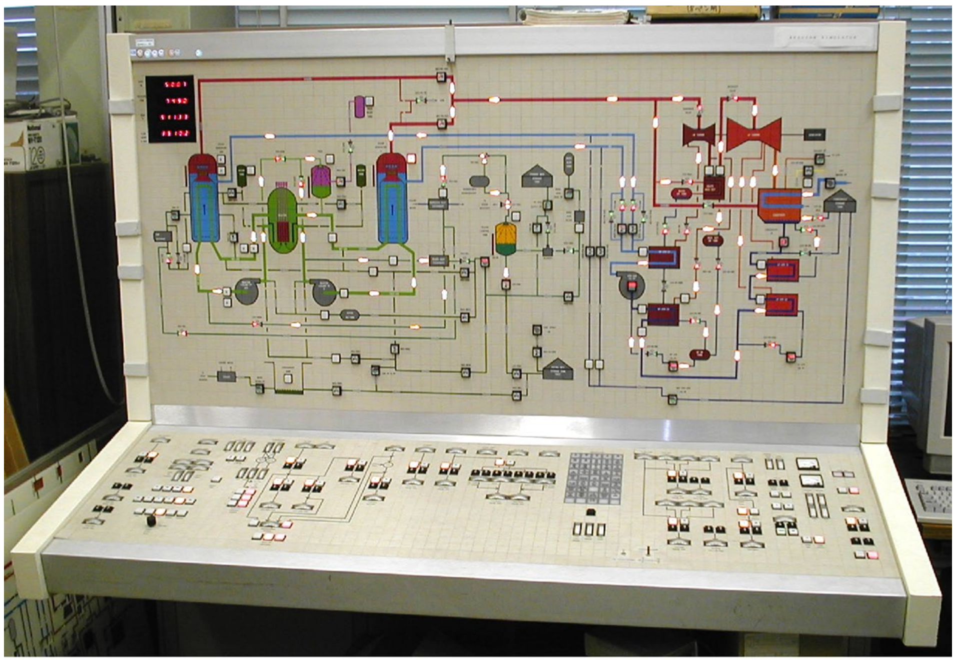
Model-Based by neural network

- Modeling of correlation among main process signals
- Monitoring difference between measured signal and estimated value





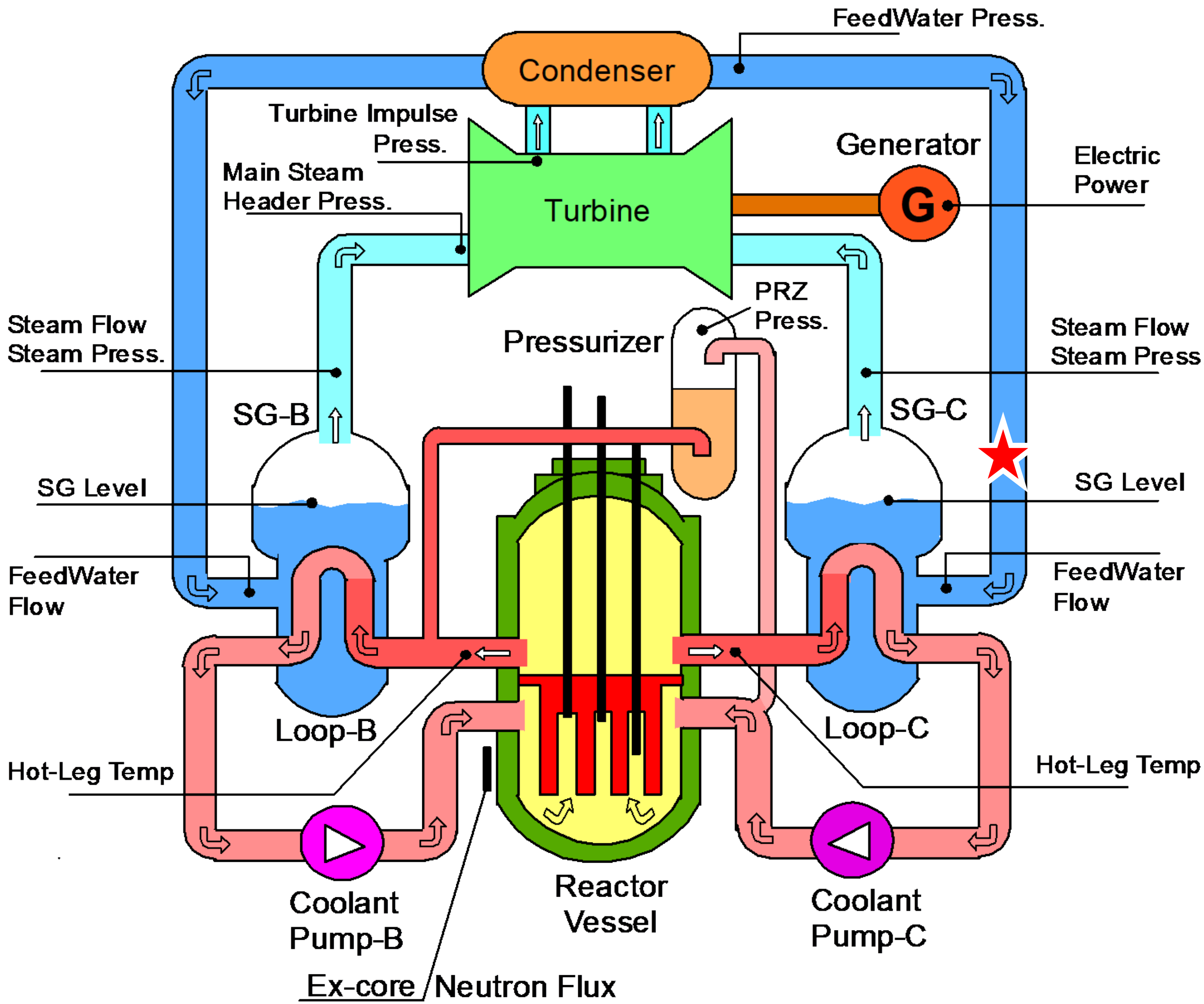
Case3-1: Feedforward and Recurrent Neural Network ~ Application to PWR simulator ~



PWR Plant Simulator (Surry-1 Model)

- SIMULATOR
  Surry-1 (USA), PWR
  822MWe, 3-Loops
- NETWORK TYPE
  Recurrent Neural Network with Adaptive Learning
- INPUT&OUTPUT
  Main Plant Signals: 22 Ch
- LEARNING DATA
  Normal Operation Data (Transient and Steady State Operation)

K. Nabeshima, et.al, "Nuclear Power Plant Monitoring with Recurrent Neural Network", J. Knowledge-based Intelligent Engineering System 4[4] (2000)



Schematic Representation of PWR Simulator (Surry-1)

* A loop is also modeled, but not on the display.

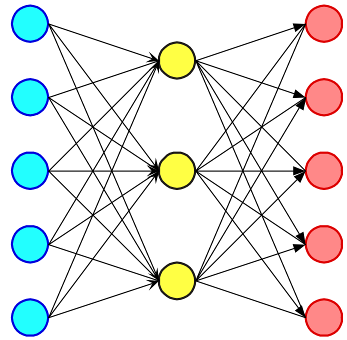
Monitoring signals (22 important signals)

Ch.	Signal	Maximum Error	Ch.	Signal	Maximum Error
1	Ex-core Neutron Flux -A	0.38848 [%]	12	Steam Flow (loop-C)	2.94088 [t/h]
2	Ex-core Neutron Flux -C	0.37494 [%]	13	Feedwater Flow (loop-B)	3.12966 [t/h]
3	Ex-core Neutron Flux -B	0.37883 [%]	14	Feedwater Flow (loop-C)	2.71875 [t/h]
4	Ex-core Neutron Flux -D	0.38532 [%]	15	Main Steam Header Pressure	0.09717 [kgf/cm^2]
5	Average Coolant Temp.	0.11759 [^oC]	16	Feedwater Pressure	0.07543 [kgf/cm^2]
6	Pressurizer Pressure	0.17125 [kgf/cm^2]	17	Hot-leg Temperature (loop-B)	0.10824 [^oC]
7	VCT (Vol. Cont. Tank) Level	0.38583 [%]	18	Hot-leg Temperature (loop-C)	0.19781 [^oC]
8	Turbine Impulse Pressure	0.13519 [kgf/cm^2]	19	Steam Pressure (loop-B)	0.07070 [kgf/cm^2]
9	Steam Generator Level (B)	0.09953 [%]	20	Steam Pressure (loop-C)	0.07122 [kgf/cm^2]
10	Steam Generator Level (C)	0.08940 [%]	21	Average Neutron Flux	0.58640 [%]
11	Steam Flow (loop-B)	2.82109 [t/h]	22	Generated Electric Power	2.31500 [MWe]

Anomaly Detection by Auto-associative Neural Network

Learning
Normal Operation

INPUT $X(t)$

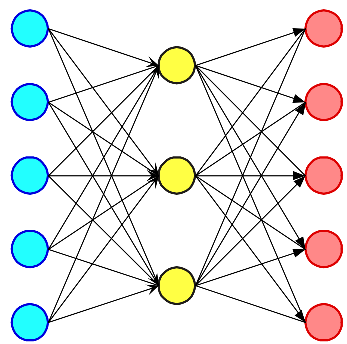


OUTPUT $X'(t) = X(t)$

Modeling of the correlation among plant signals

Testing
Normal Operation

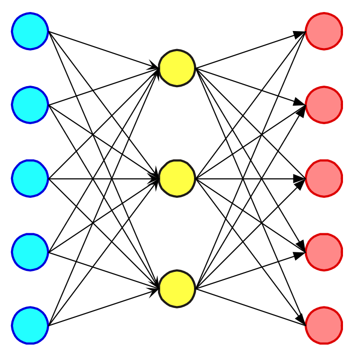
INPUT $X(t)$



OUTPUT $X'(t) \approx X(t)$

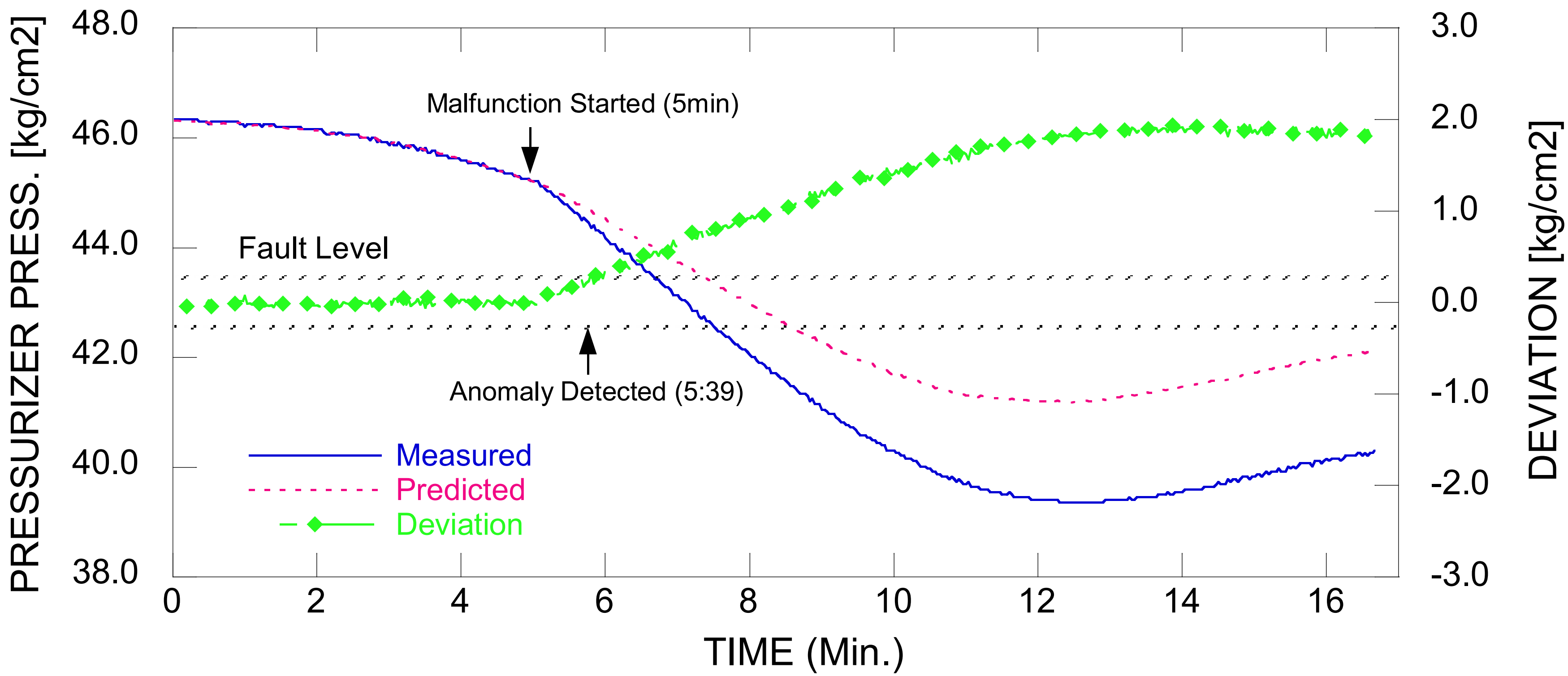
Abnormal

INPUT $X(t)$



OUTPUT $X'(t) \neq X(t)$

“Small Reactor Coolant System Leak” (56.7 l/min) during power decrease operation (Turbine: -2.0%/min)



Ch.6 (Pressurizer Pressure)

Fault Detection Channels for Leakage

No.	Malfunction	Detection Channel No.			Ch. No. Without Alarm	Conventional Alarm
		First	Second	Third		
1	Small Reactor Coolant System Leak (Large:56.7 *l*/min)	Ch.6 (0:18)	Ch.16 (2:26)	Ch.8,22 (3:32)	Ch.9	**No Alarm**
2	Small Reactor Coolant System Leak (Small:11.4 *l* /min)	Ch.8,22 (3:08)	Ch.2 (3:31)	Ch.4 (3:37)	Ch.9	**No Alarm**
3	Leakage of Atmospheric Steam Dump Valve (Large:5%)	Ch.11,12 (0:02)	Ch.13,14 (0:04)	Ch.2,4 (0:06)	Ch.5,9,10, 17,18	**No Alarm**
4	Leakage of Atmospheric Steam Dump Valve (Small:1%)	Ch.8,22 (7:47)	Ch.10 (8:11)	Ch.2,3,4 (8:13)	Ch.5,9,10, 17,18	**No Alarm**
5	Partial Loss of Feedwater (Large:90.7 ton/hr)	Ch.16 (0:04)	Ch.13,14 (0:10)	Ch.2,3 (0:32)	Ch.5,6,9,10, 13,17,18	**No Alarm**
6	Partial Loss of Feedwater (Small:9.07 ton/hr)	Ch.8,22 (3:53)	Ch.10 (4:15)	no	Ch.5,6,9,10, 13,17,18	**No Alarm**

- Anomaly Detection during steady state or transient operation is not difficult for Deep Learning.
- Anomaly Identification is difficult.

Case 3-2: Feedforward Neural Network with Adaptive Learning ~ Real-time Application to PWR Plant ~

Borssele NPP (The Netherlands)

- Electric Power : 470 MWe
- Coolant Loop : 2 Loop
- Steam Generator : 2

Conventional Method:
Off-line Fault Detection by Noise Analysis

K. Nabeshima, et.al, “Real-Time Power plant Monitoring with Neural Network”, J. Nucl. Sci. Technol. 35[2] (1998)

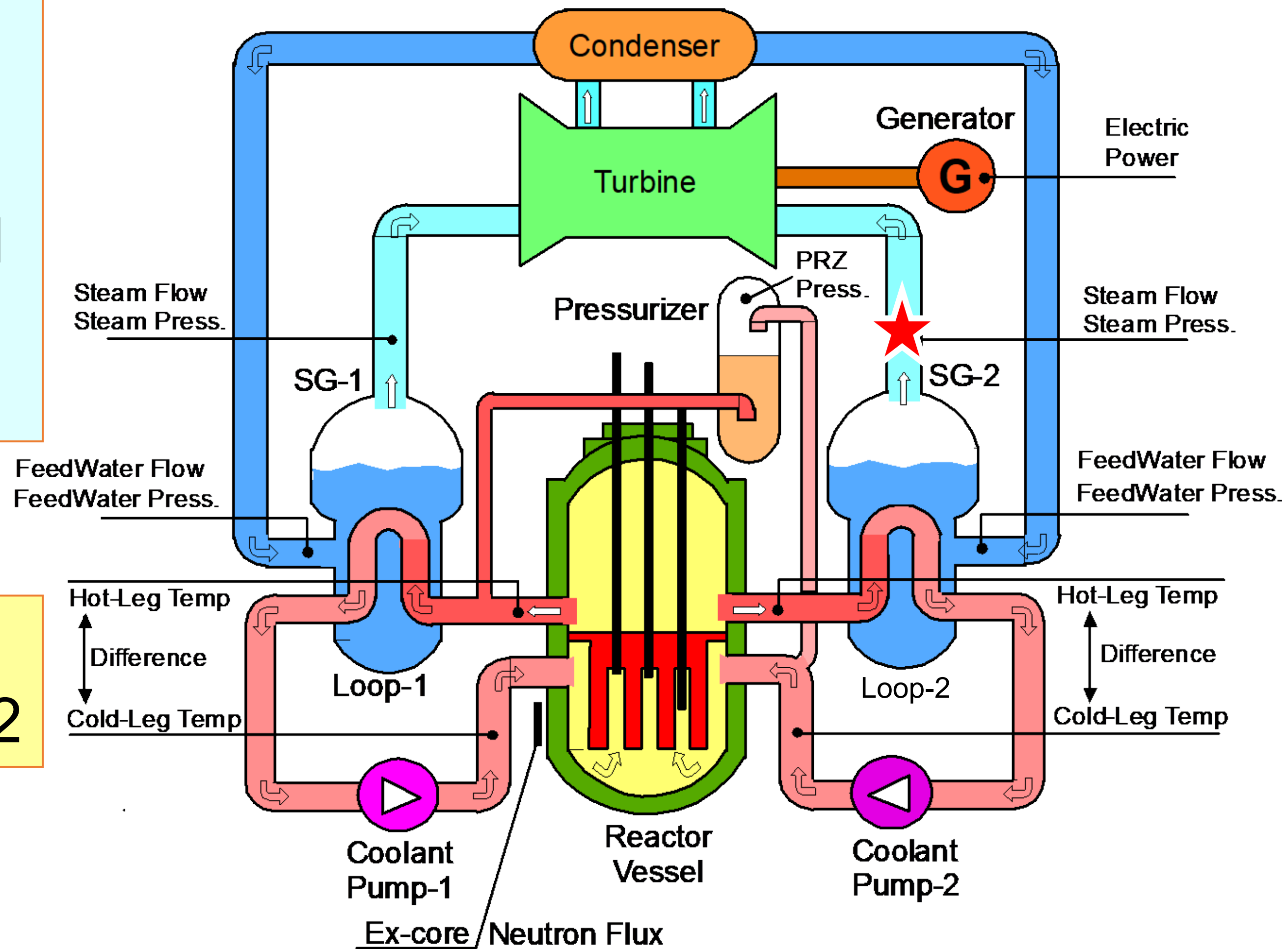
Reverse Phase Rinsing Operation

<u>Rinsing Operation*</u>
<u>at Last Shutdown</u>

Dynamics of Loop-1 and Loop-2 was reversed

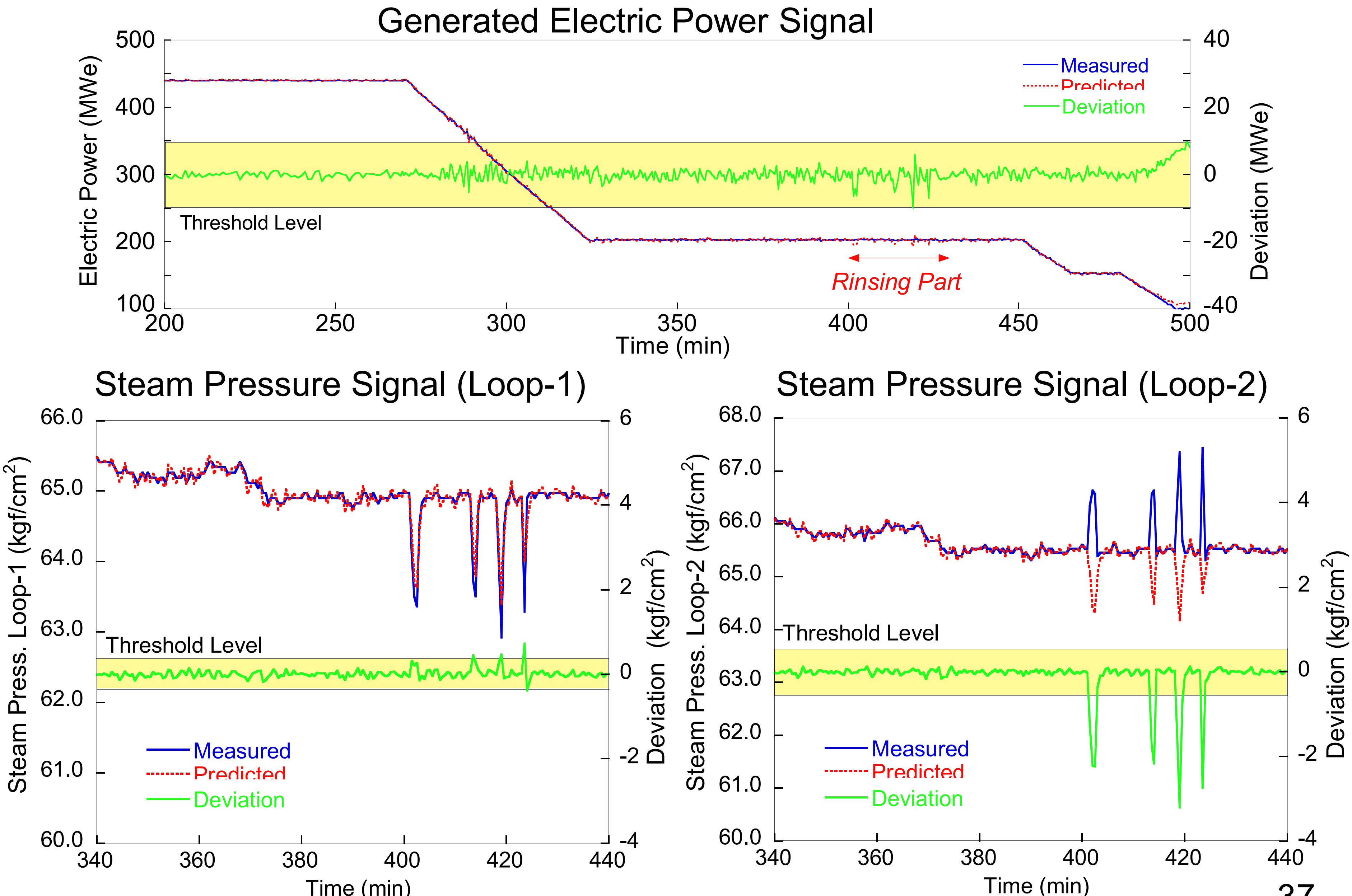
Anomaly Detection at Secondary Loop-2

* Rinsing :
Operation mode for cleaning Condenser

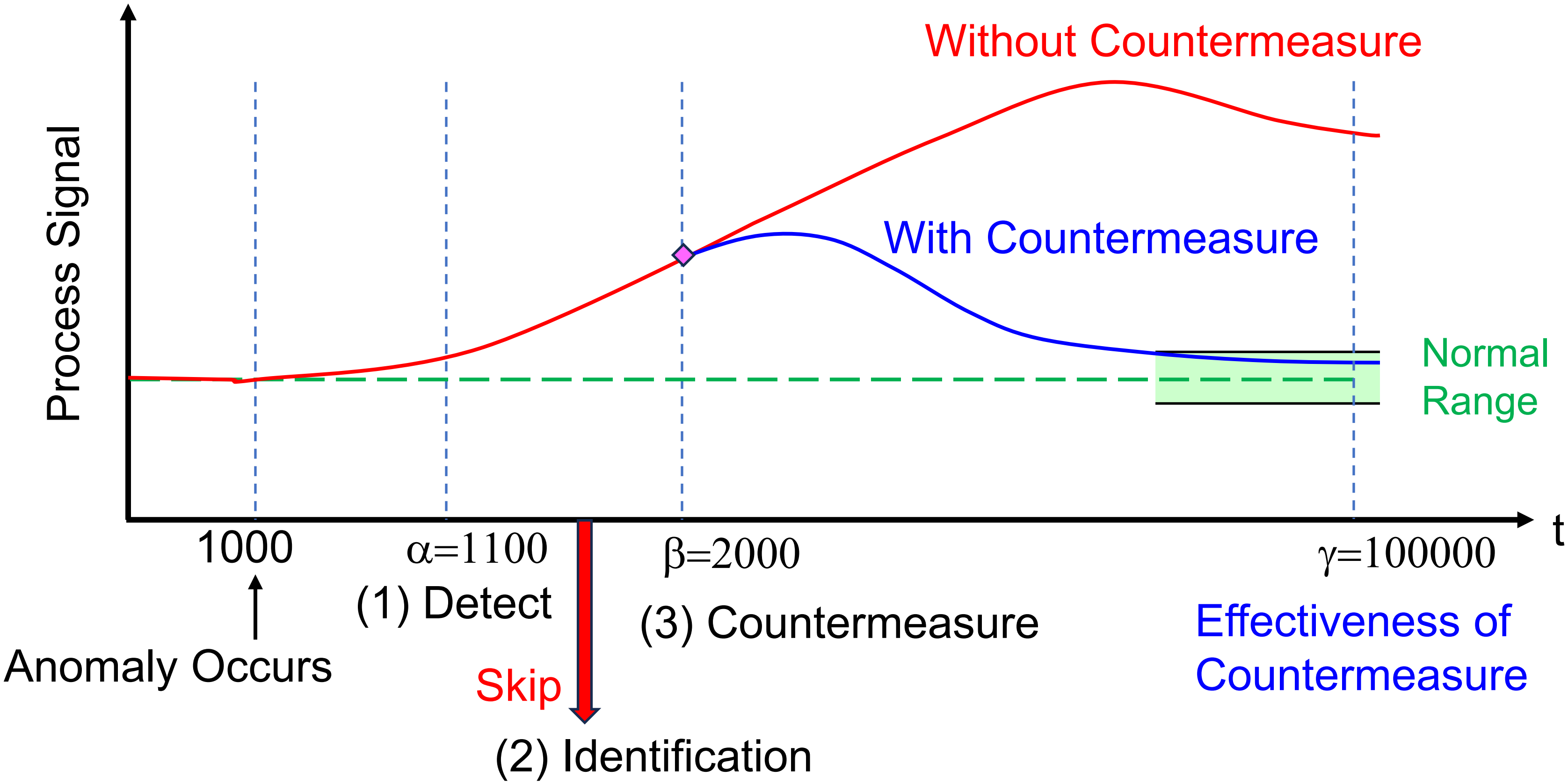


Overview and Signals of Borssele NPP

Case 2 : Real-time Anomaly Detection During Last Shutdown



Case 3-3: Reactor Monitoring for Operator Support ~ Application to HTGR ~



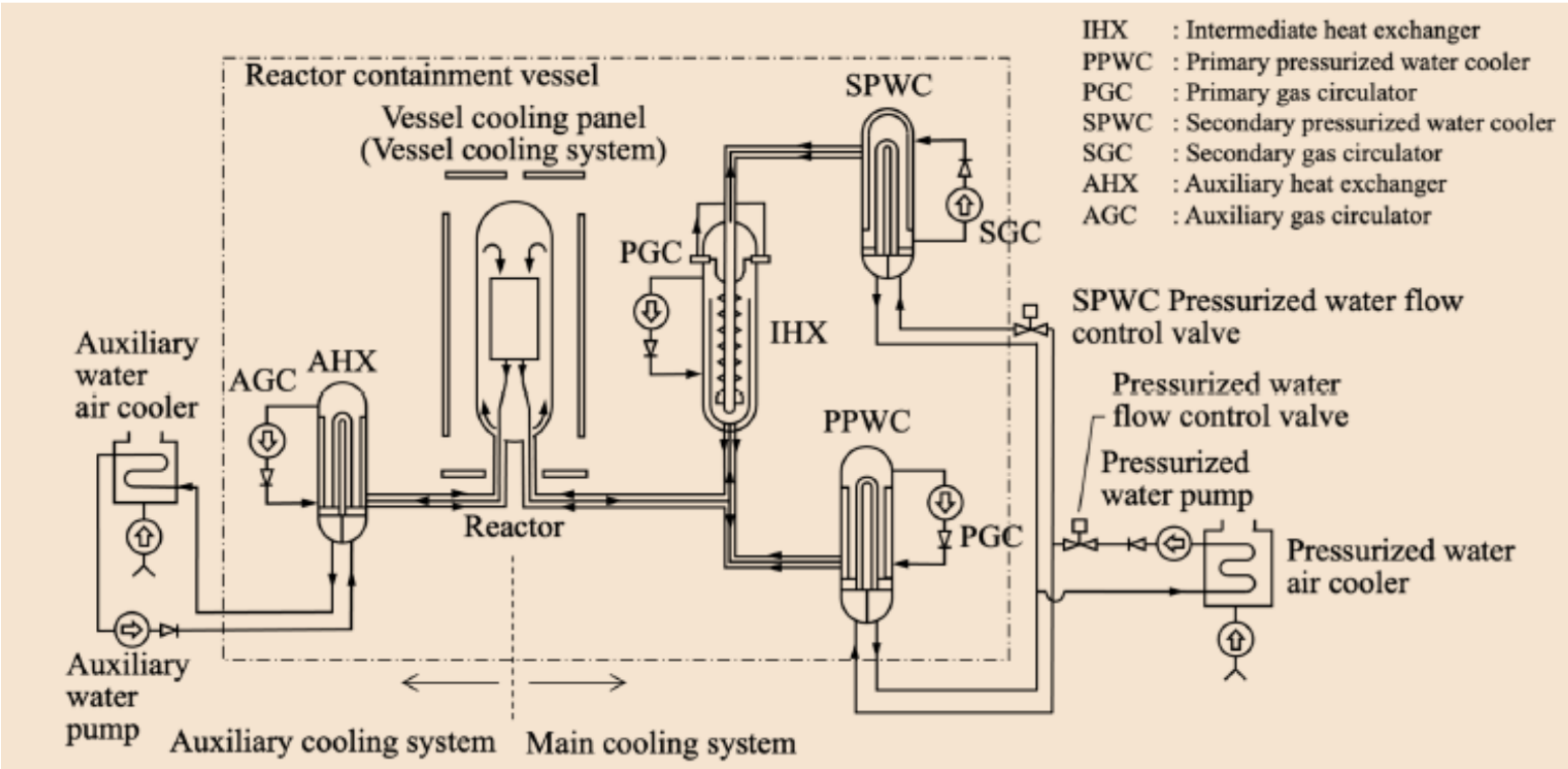
① Anomaly Detection by Deep Learning
② Identification (Place, Cause) → SKIP
③ Countermeasure Selected by Reinforcement Learning

High Temperature Engineering Test Reactor (HTTR) in JAEA

Thermal Power:	30 MW
Coolant:	Helium Gas
Outlet Temp. :	950 ^{o}C
Inlet Temp. :	395 ^{o}C
Core Material:	Graphite
Pressure:	4MPa

1998 : First criticality
2001 : Full power operation
2010 : 50 days continuous 950^{o}C Operation
2010 : Loss of core flow test at 9MW
(Great East Japan Earthquake : 2011)
2021 : Restart
2022 : Loss of core cooling test at 9MW
2024 : Loss of core cooling test at 30MW

Reactor Cooling System of HTTR in JAEA



Abnormal situation data is created by ACCORD:

Plant dynamic analysis code for high temperature gas-cooled reactors

Name of disturbance and their value range

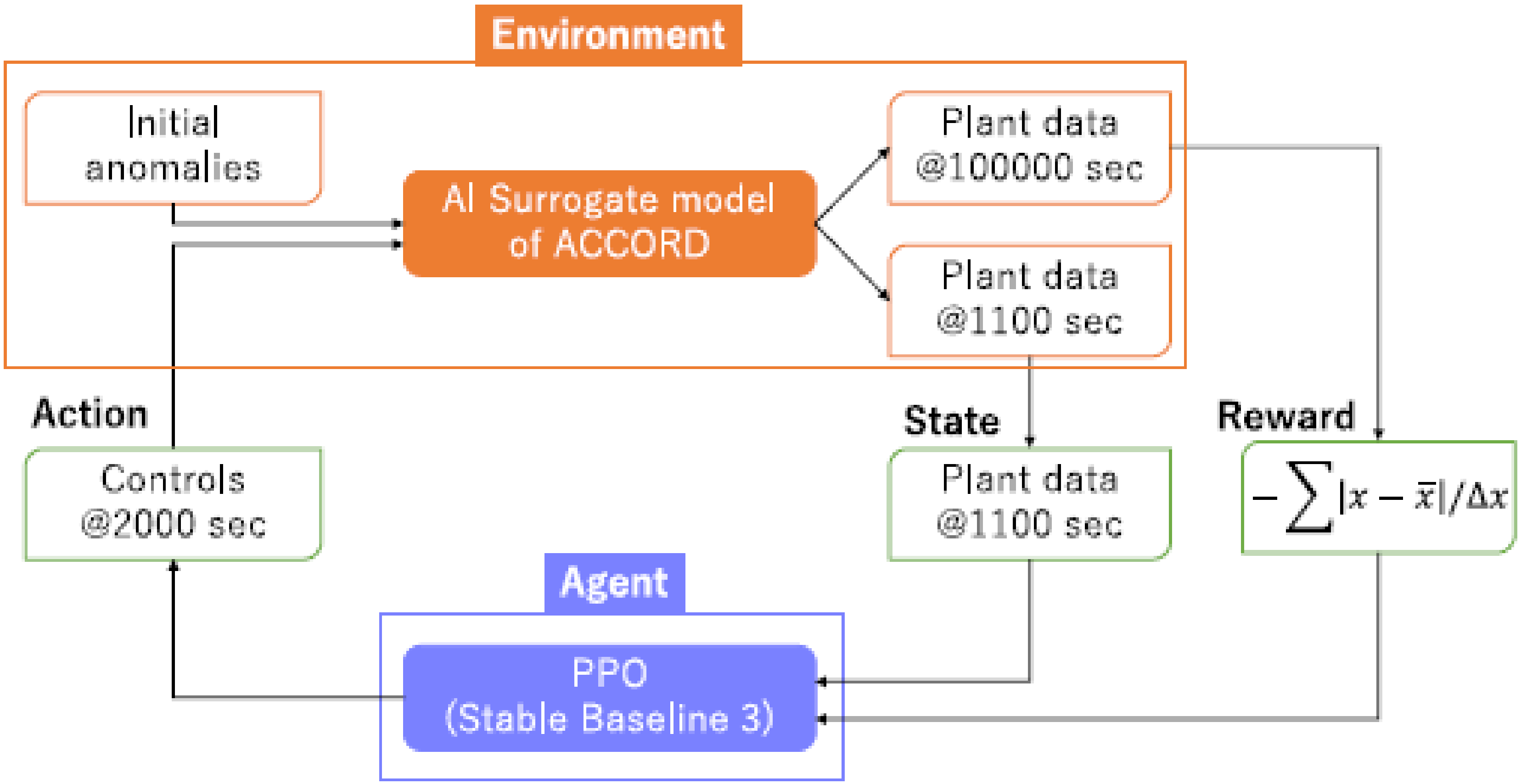
Name	Description	Unit	Steady Value	Range (%)[a]
FLW4(1)	AC air flow rate (pressurized water system)	kg/s	605	30 to 120
TMP4(1)	AC air inlet temperature (pressurized water system)	°C	33	10 to 150
DBPSCR	Reactor core bypass flow ratio	—	0.098	81 to 121
VLVBPS	AC bypass value opening rate	%	58	20 to 170
VLVACL	AC flow control value opening rate	%	42	40 to 230
GCI	G/C rotation speed for IHX	rpm	8080	80 to 110
GC1P1	Rotation speed of primary gas circulator	rpm	6975	41 to 120
GC2P	Rotation speed of secondary gas circulator	rpm	7100	40 to 120
WPMP1	Rotation speed of pressurized water pump	rpm	2970	10 to 120
VVPWC1	PPWC pressurized water flow control valve opening rate	%	40	10 to 250
VVPWC2	SPWC pressurized water flow control valve opening rate	%	100	10 to 100

*AC = air cooler; G/C = gas circulator; IHX = intermediate heat exchanger; PPWC = primary pressurized water cooler; SPWC = secondary pressurized water cooler.
[a]Steady value: 100%.

Components used for countermeasures

Name	Description (Physical quantity [Unit])	Range* (%)
GC2P	Secondary gas circulator (Rotation speed [rpm])	90 – 110
WPMP1	Pressurized water pump (Rotation speed [rpm])	90 – 110
VVPWC1	Pressurized water flow control valve (Openness [Opening degree])	25 – 250
VVPWC2	SPWC Pressurized water flow control valve (Openness [Opening degree])	10 – 100
GCA1	Auxiliary gas circulator: No. 1** (Rotation speed [rpm])	0 – 110
GCA2	Auxiliary gas circulator: No. 2** (Rotation speed [rpm])	0 – 110
WPMP2	Auxiliary water pump: No. 1 (Rotation speed [rpm])	90 – 110
WPMP2B	Auxiliary water pump: No. 2 (Rotation speed [rpm])	90 – 110

Schematic diagram of Counter-Measure Proposal Module (CMPM)



S. Takaya, A. Seki, M. Yoshikawa, N. Sasaki and X. Yan, “Proposal of a novel AI-based plant operator support system for the safety of nuclear power plants”, Mechanical Engineering Journal, Vol.11, No.2 (2024)

AI surrogate models of ACCORD:

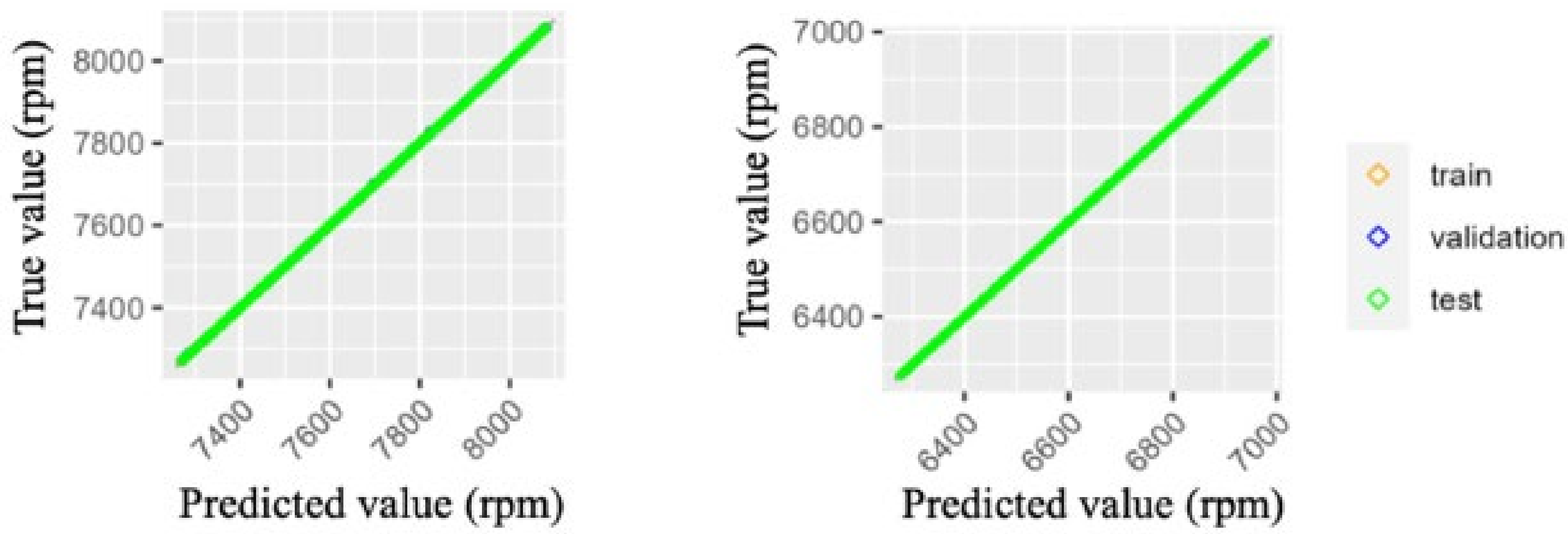
- Feedforward fully-connected neural network
- Hidden layers: 3
- Number of nodes in each layer: 100
- Activation function: leaky ReLU
- Training Data: 3200 cases

CMPM (Counter-Measure Proposal Module):

- Deep reinforcement learning algorithms: Proximal policy optimization (PPO)
- Reward function: $-\sum_{i=1}^{2} \frac{|\bar{x}_i - x_i|}{\Delta x_i}$

$\bar{x}_i$: value during normal operation
i=1: reactor power
i=2: reactor outlet temperature

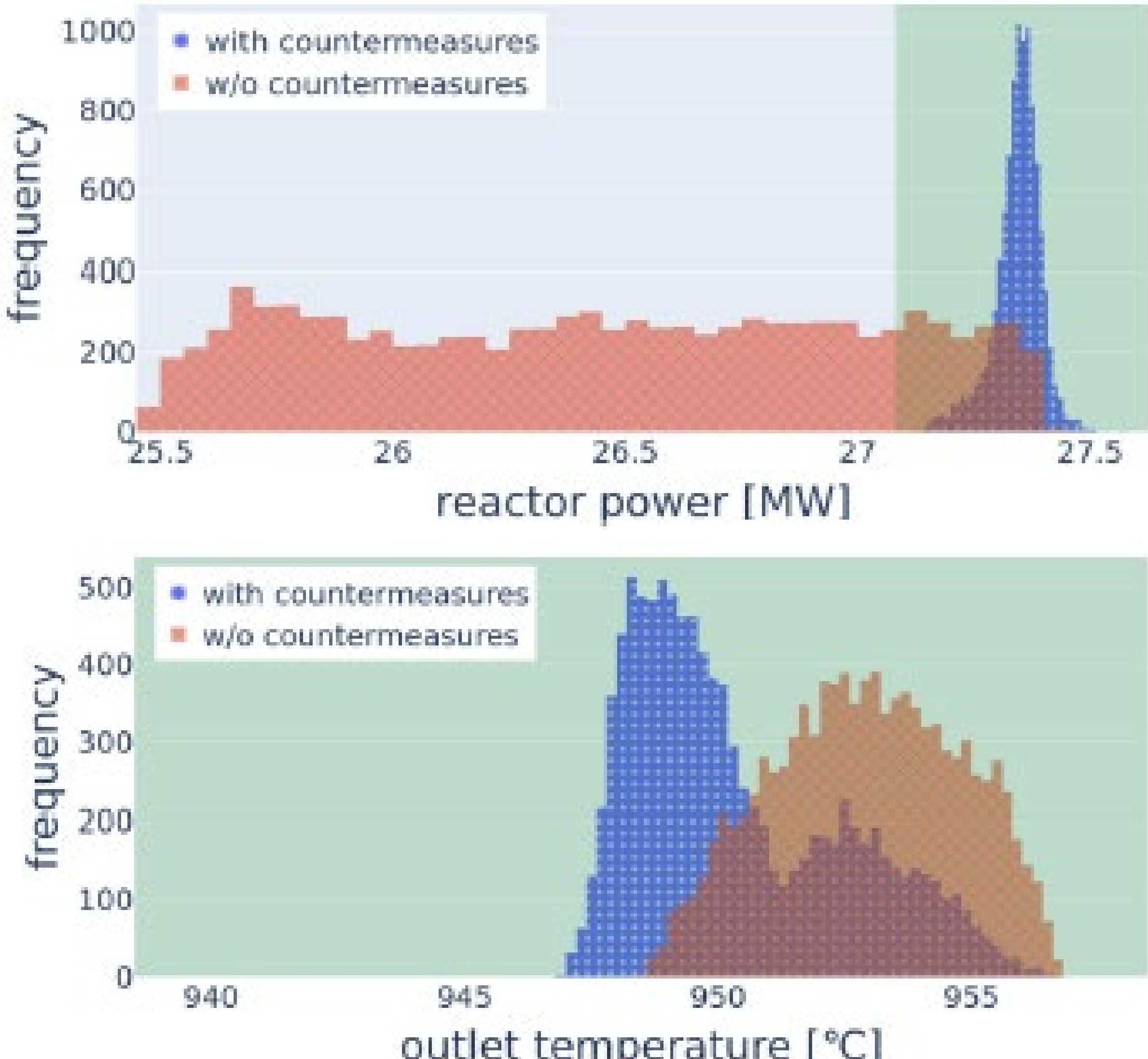
Comparisons between true and predicted rotation speed of primary gas circulators



(a) IHX(GCI)

(b) PPWC (GC1P1)

Comparison of distributions of plant parameters with and without countermeasures (1000 cases)



What is most important for Deep Learning?

- Selection of important input signals for appropriate modeling
- Preprocessing of input signals

  (ex. stochastic parameters)
- Type of network
- Number of layers
- Number of unit in layers
- Learning parameters

Thank you for your attention!