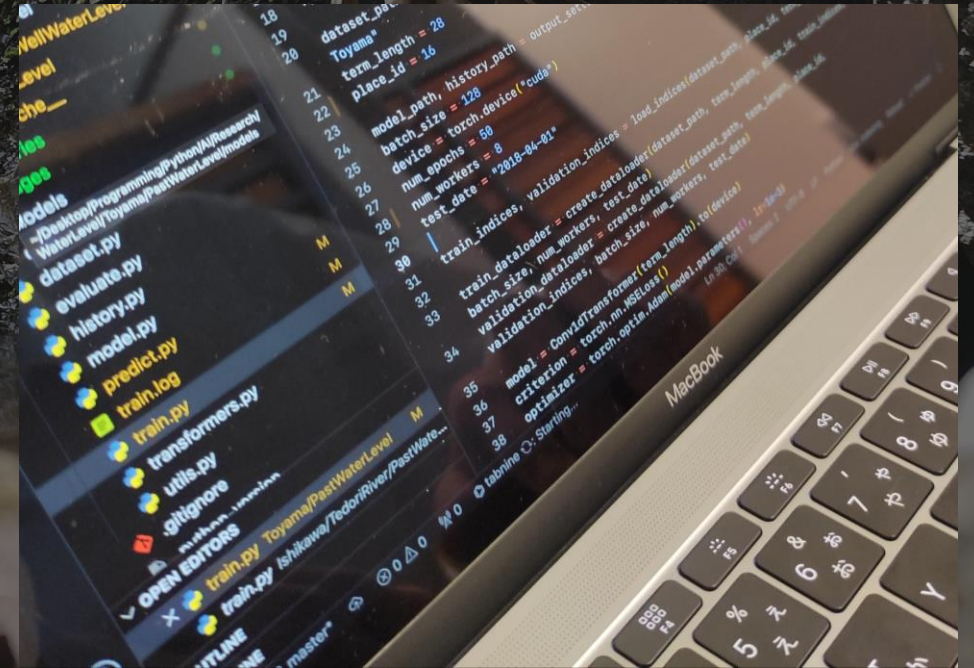


Development of the High Accuracy Groundwater Level Prediction Model Using Machine Learning



Kaijo Senior High School
Aoyama Kuya · Katsuyama Shouki

1-1. Background

Groundwater is an "invisible" water resource



Daily observations

- We've been researching Otomeyama park, Shinjuku City for 13 years
- We have strongly felt the need to preserve the water environment around us
(In Tokyo, 68 springs have disappeared in the past 10 years)
- Fluctuations in groundwater levels cause spring depletion, well depletion, land subsidence, and landslides
- We want to "visualize" groundwater and conserve it

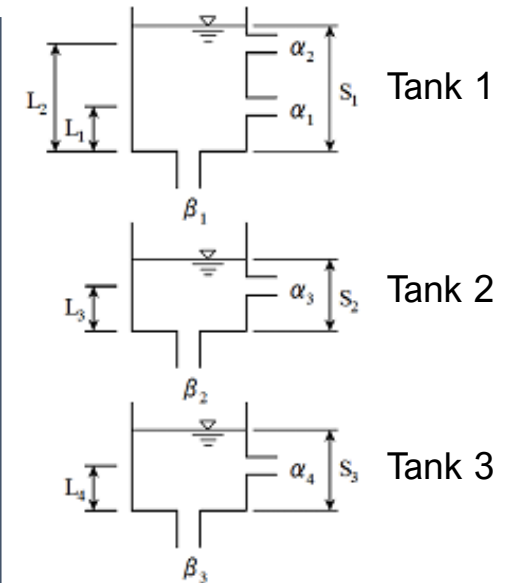
1-1. Background

Groundwater level prediction makes future groundwater "visible"

Tank Model

- Low versatility across regions
- Need more than 10 data
- Need one observation well per 2square kilometers

Place a heavy burden on local governments



In-line 3-stage tank model Method
(Japan Meteorological Agency)

1-2. Approach

1

Use only meteorological data and groundwater level data

2

Make the model applicable to different regions

Making Groundwater Level Prediction Easier

2-1. Model

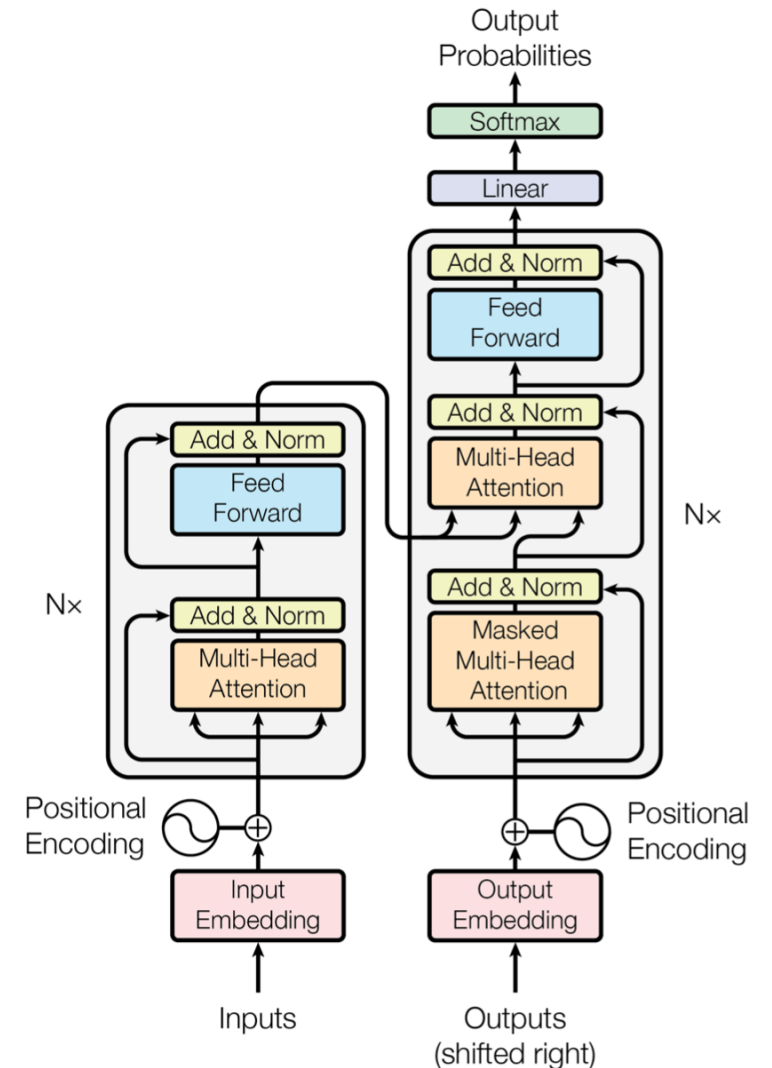
Transformer

A high-performance machine learning model for **time series data**

Proposed to use in **machine translation** or natural language processing

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin
“Attention is All You Need” 2017

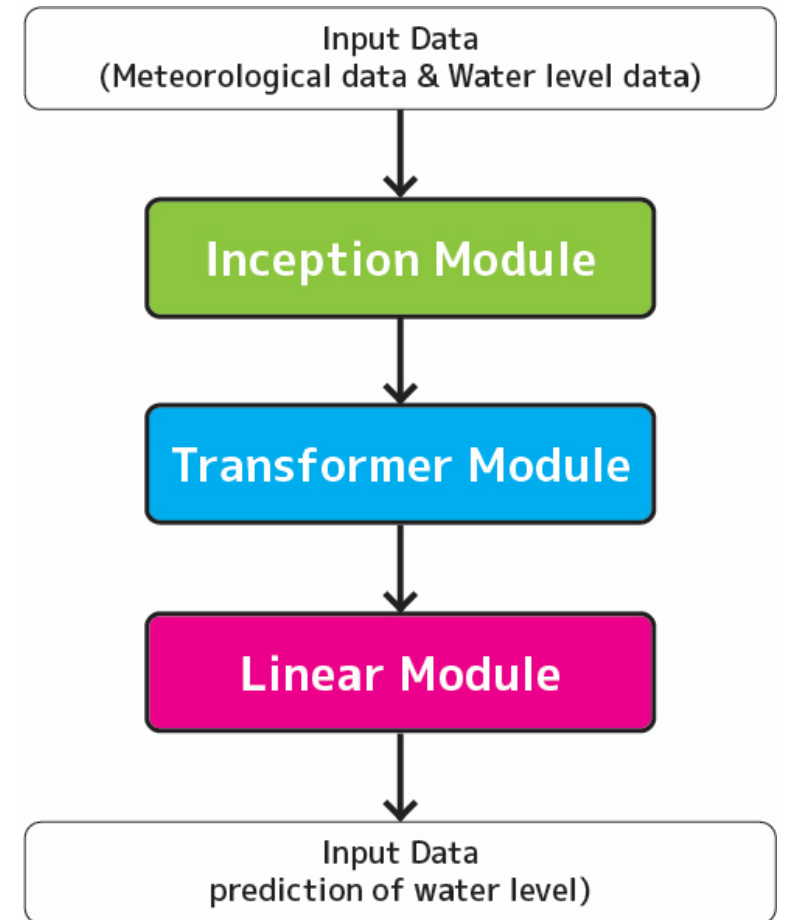


The Transformer-model architecture

2-2. Architecture

The model consists of three modules

1. Inception Module
2. Transformer Module
3. Linear Module

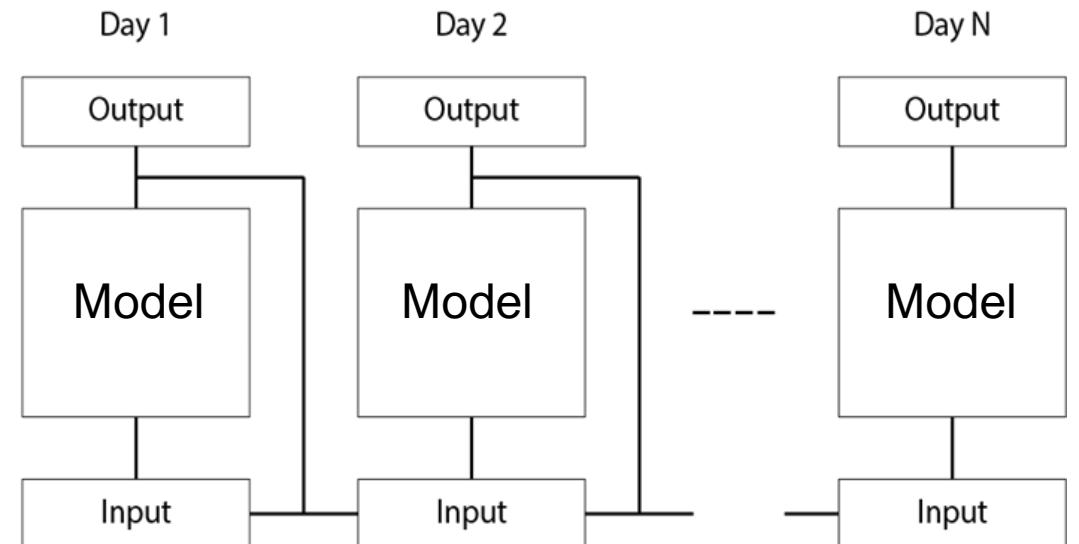


2-2. Architecture

Prediction Method

Passing the previous segment's output to the next segment's input

This model can predict without the latest ground water level data by predicting recurrently



2-3. Implementation

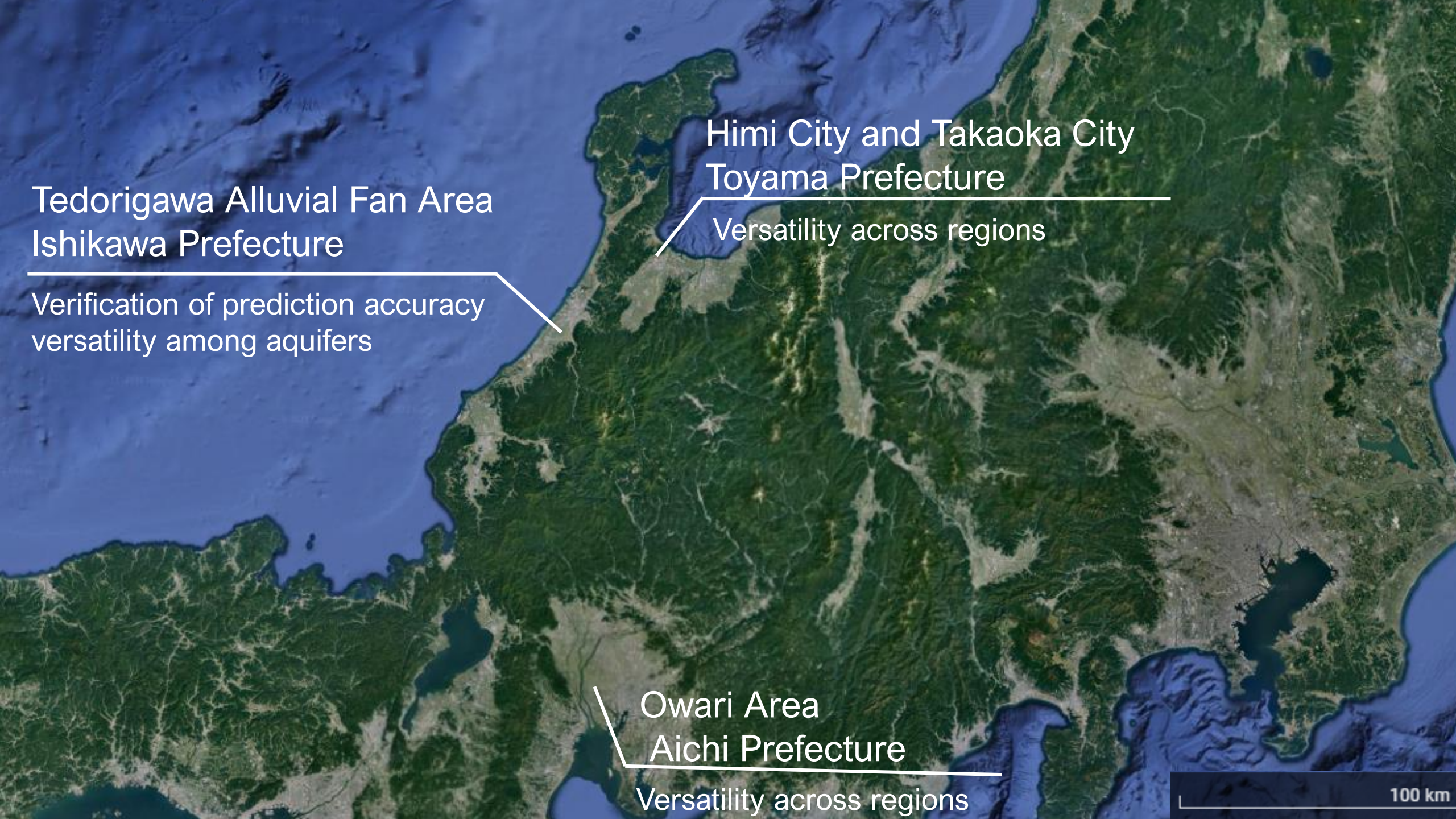
This model was implemented with Python, a programming language.

```
8     return softmax.bmm(value)
9
10 class AttentionHead(torch.nn.Module):
11     def __init__(self, dim_in: int, dim_k: int, dim_v: int):
12         super().__init__()
13         self.q = torch.nn.Linear(dim_in, dim_k)
14         self.k = torch.nn.Linear(dim_in, dim_k)
15         self.v = torch.nn.Linear(dim_in, dim_v)
16
17     def forward(self, query: Tensor, key: Tensor, value: Tensor) -> Tensor:
18         return scaled_dot_product_attention(self.q(query), self.k(key), self.v(value))
19
20 class MultiHeadAttention(torch.nn.Module):
21     def __init__(self, num_heads: int, dim_in: int, dim_k: int, dim_v: int):
22         super().__init__()
23         self.heads = torch.nn.ModuleList(
24             [AttentionHead(dim_in, dim_k, dim_v) for _ in range(num_heads)]
25         )
26         self.linear = torch.nn.Linear(num_heads * dim_v, dim_in)
27
28     def forward(self, query: Tensor, key: Tensor, value: Tensor) -> Tensor:
29         return self.linear(torch.cat([h(query, key, value) for h in self.heads], dim=-1))
30
31 def position_encoding(seq_len: int, dim_model: int, device: torch.device = torch.device("cuda")):
32     -> Tensor:
33     pos = torch.arange(seq_len, dtype=torch.float, device=device).reshape(1, -1, 1)
34     dim = torch.arange(dim_model, dtype=torch.float, device=device).reshape(1, 1, -1)
35     phase = pos / 1e4 ** (dim / dim_model)
36     return torch.where(dim.long() % 2 == 0, torch.sin(phase), torch.cos(phase))
37
38 def feed_forward(dim_input: int = 512, dim_feedforward: int = 2048) -> torch.nn.Module:
39     return torch.nn.Sequential(
40         torch.nn.Linear(dim_input, dim_feedforward),
41         torch.nn.ReLU(),
42         torch.nn.Linear(dim_feedforward, dim_input)
```

```
Epoch: 4
Training Model
0% [*****] 100% | ETA: 00:00:00
Total time elapsed: 00:00:04
Evaluating Model
0% [*****] 100% | ETA: 00:00:00
Total time elapsed: 00:00:01
Train Loss: 0.19856629803417677, Train R2 Score: 0.9230686720938687
Validation Loss: 0.13775256524483362, Validation R2 Score: 0.9433277536383271

Epoch: 5
Training Model
0% [*****] 100% | ETA: 00:00:00
Total time elapsed: 00:00:04
Evaluating Model
0% [*****] 100% | ETA: 00:00:00
Total time elapsed: 00:00:01
Train Loss: 0.15288805562150048, Train R2 Score: 0.9403594177064637
Validation Loss: 0.12273972765320823, Validation R2 Score: 0.9492265365188239

Epoch: 6
Training Model
0% [*****] 100% | ETA: 00:00:00
Total time elapsed: 00:00:04
Evaluating Model
```

Tedorigawa Alluvial Fan Area
Ishikawa Prefecture

Verification of prediction accuracy
versatility among aquifers

Himi City and Takaoka City
Toyama Prefecture

Versatility across regions

Owari Area
Aichi Prefecture

Versatility across regions

100 km

3-1. Data Used

Tank Model

- Groundwater level data
- Coefficient of transmissibility
- Coefficient of storage

- Amount of rainfall
- Amount of evapotranspiration
- Groundwater increment
- Penetration rate

- Industrial water
- Agricultural water
- Water for daily use
- Water for snow melt
- Water for fighting fires
- Water for environment

- Land use data

This Research

Provided directly by the local government

Using open data from the Japan Meteorological Agency

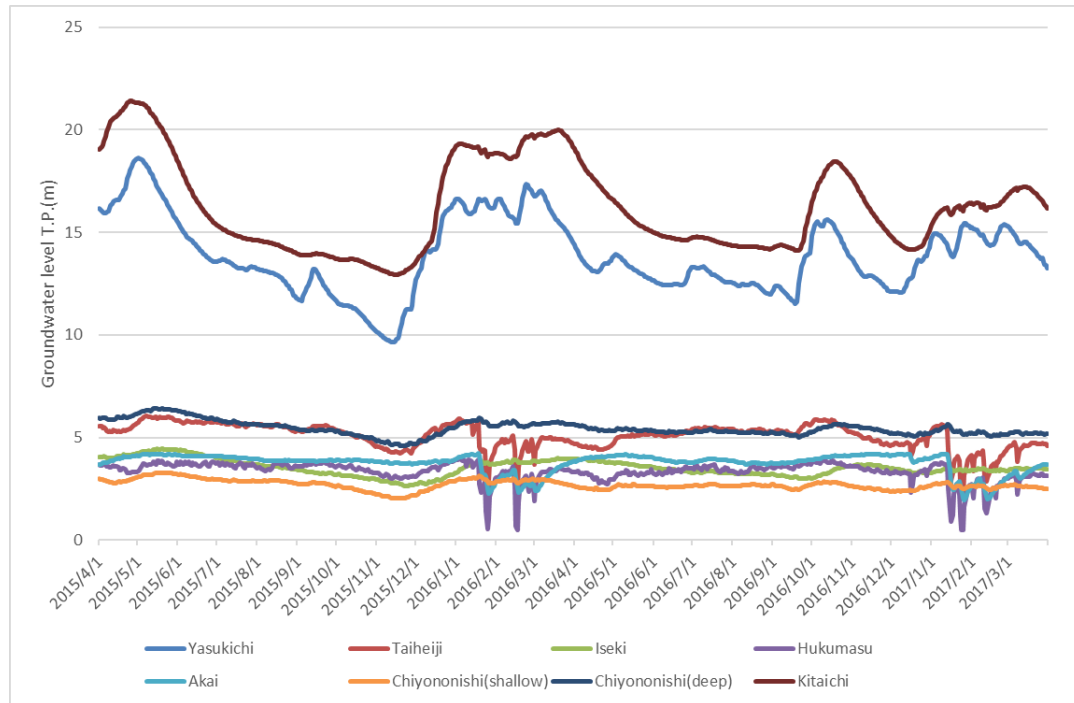
3-2. Dataset Tedorigawa Alluvial Fan Area

Training data

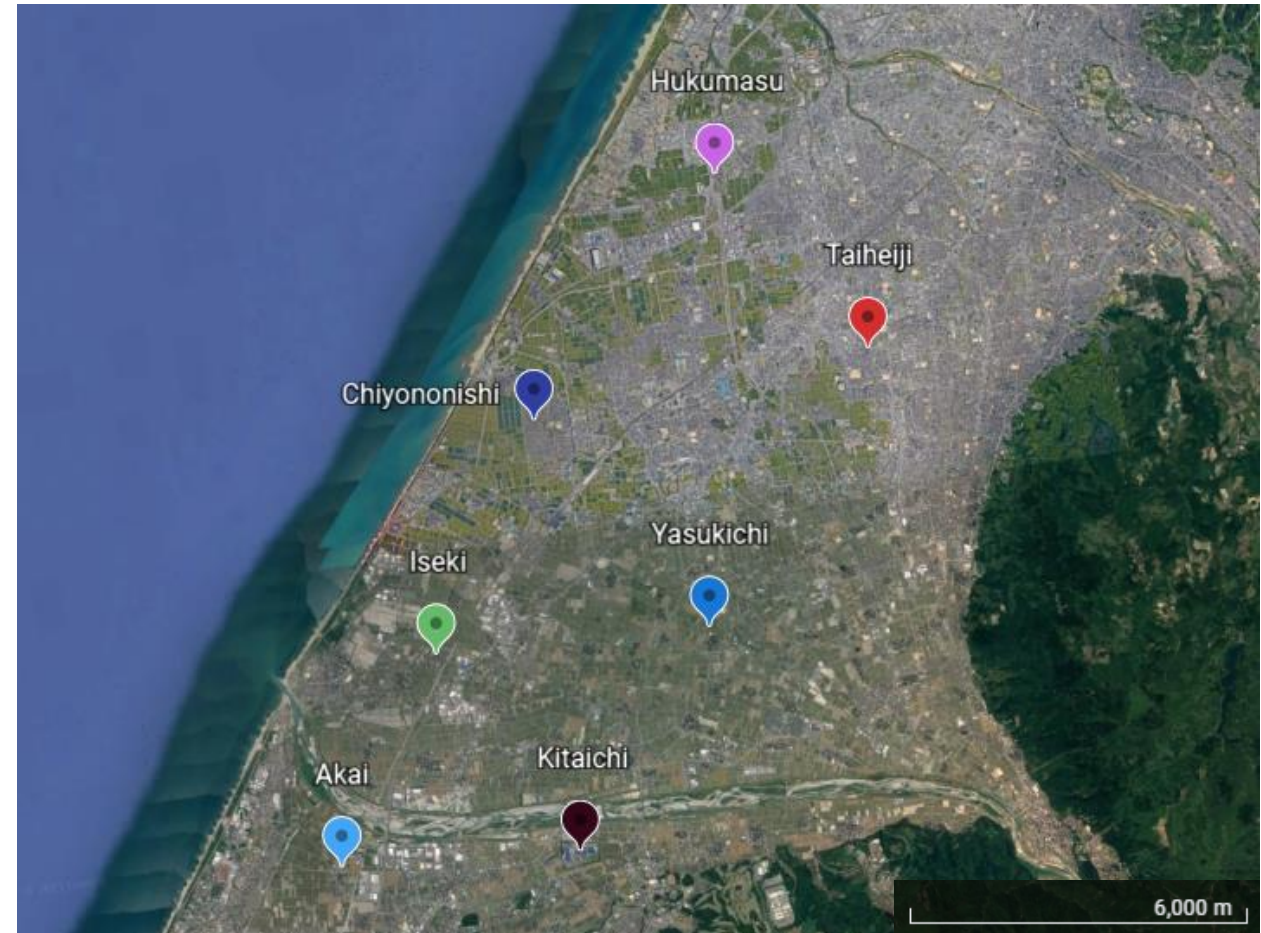
1974/04/01~2015/03/31(about 40 years)

Validation data

2015/04/01~2017/03/31



Groundwater level fluctuation graph



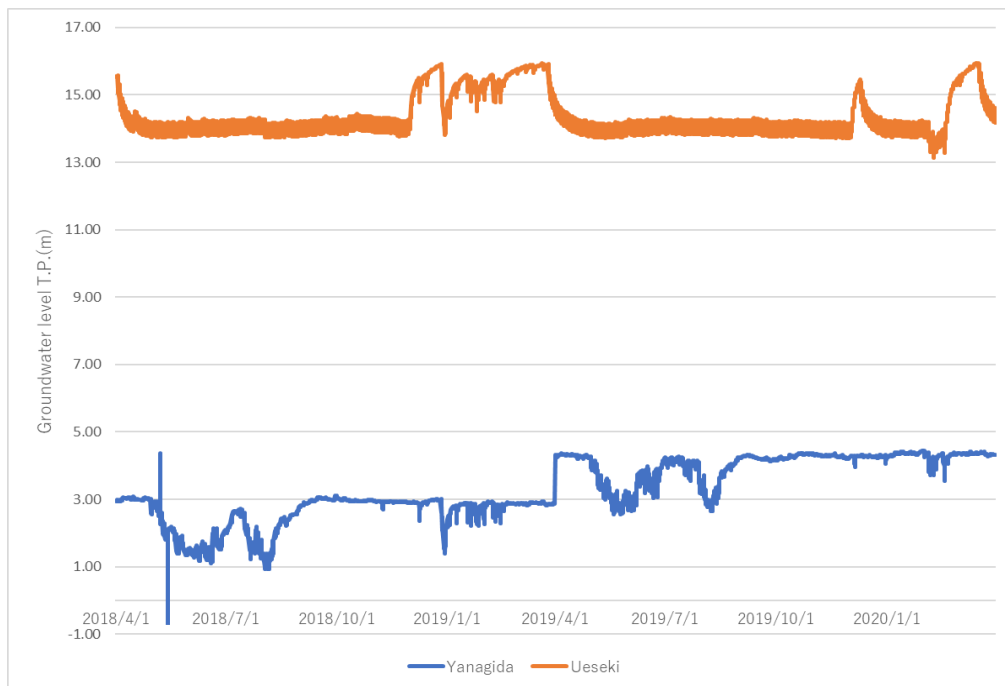
3-2. Dataset Himi city and Takaoka city

Training data

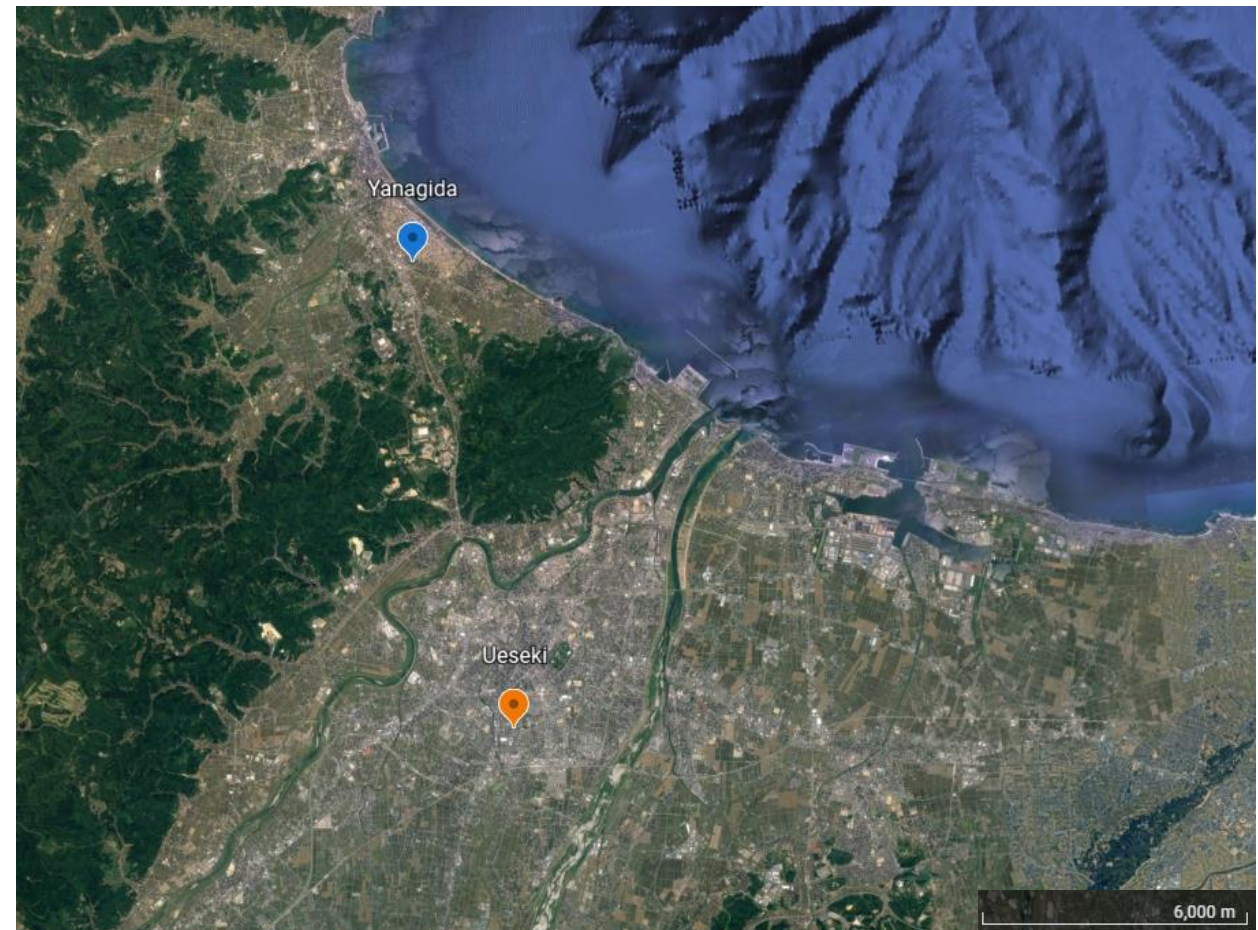
2007/04/01 ~ 2018/03/31 (about 10 years)

Validation data

2018/04/01 ~ 2020/03/31



Groundwater level fluctuation graph



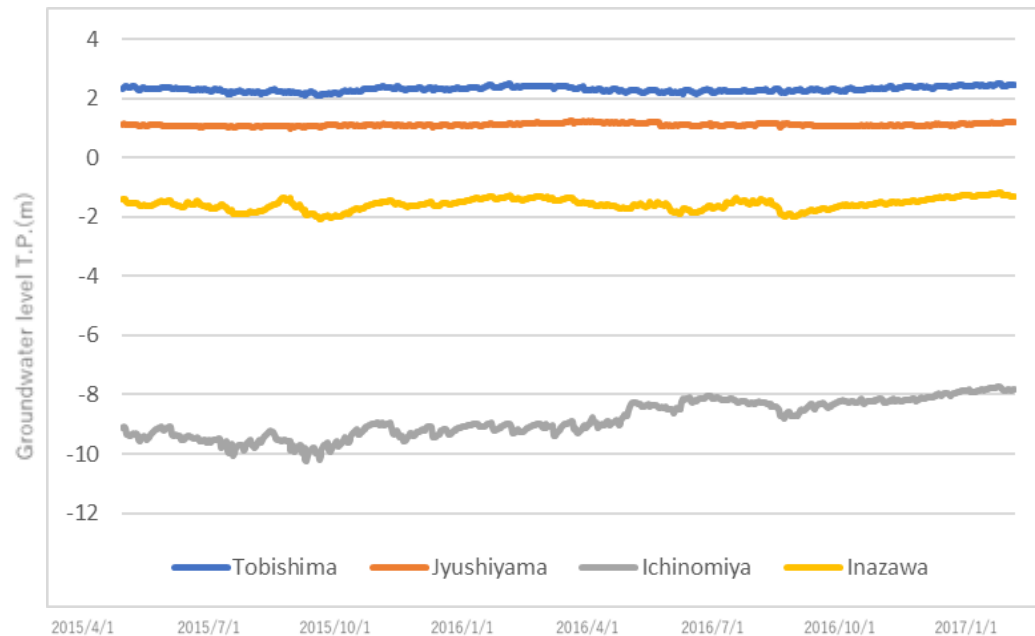
3-2. Dataset Owari Area, Aichi Prefecture

Training data

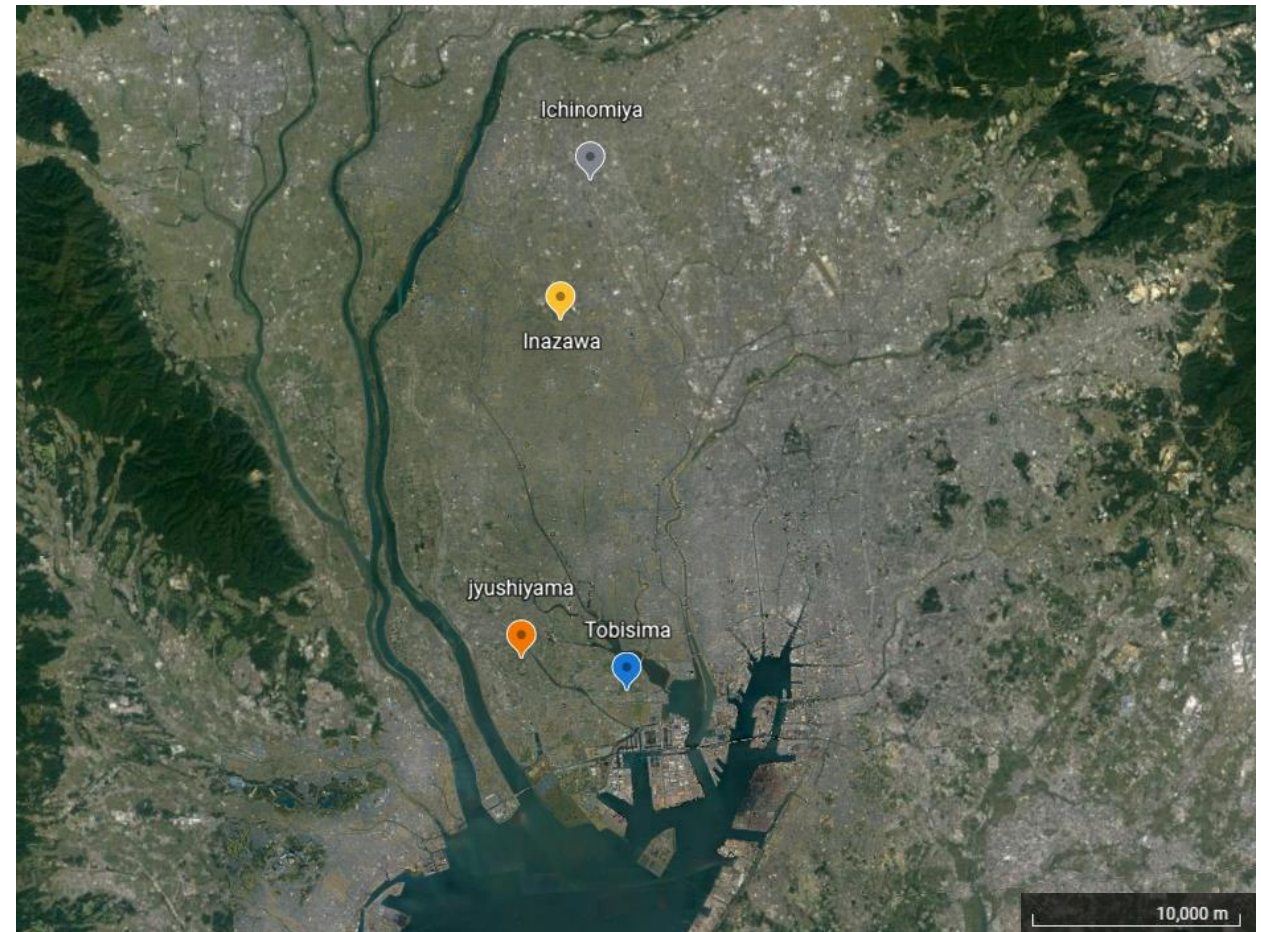
1974/04/01~2015/03/31(about 40 years)

Validation data

2015/04/01~2017/03/31

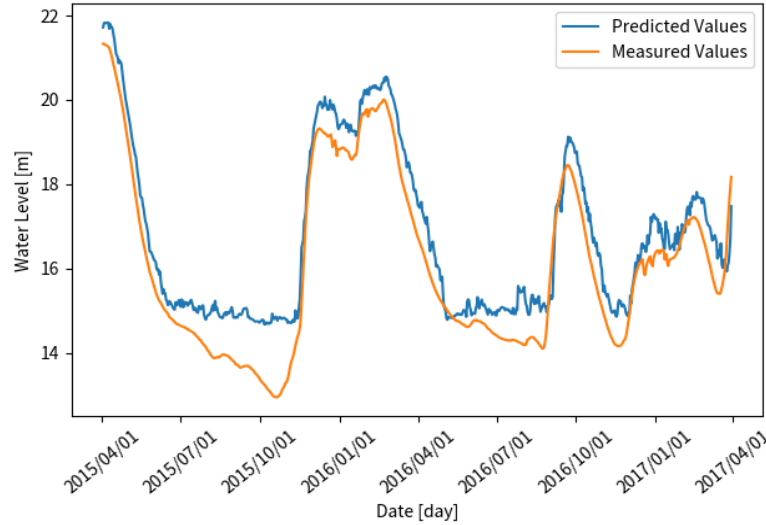


Groundwater level fluctuation graph

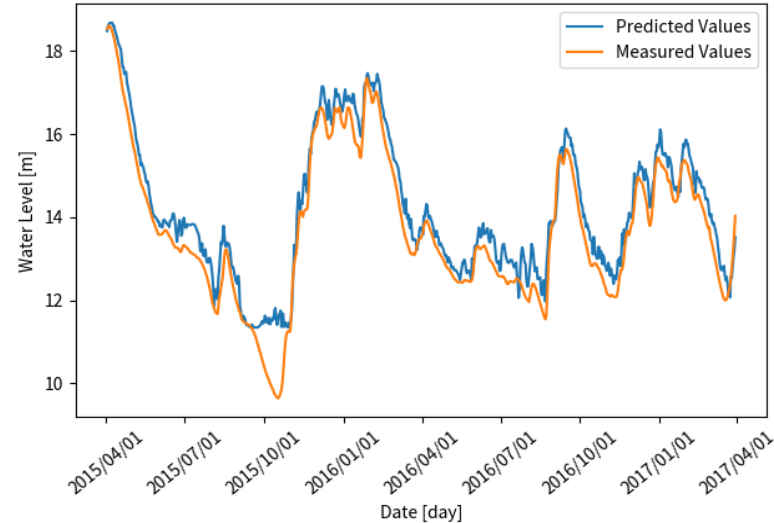


4-1. Results 1 Tedorigawa Alluvial Fan Area

Kitaichi, 2nd aquifer



Yasukichi, 2nd aquifer



Model	Yasukichi	Kitaichi
Chono et al.(2017)	0.0063	0.0032
This research	0.0035	0.0019

MSE score for each location

$$MSE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

y: measured values, \hat{y} : output values *N*: the number of data

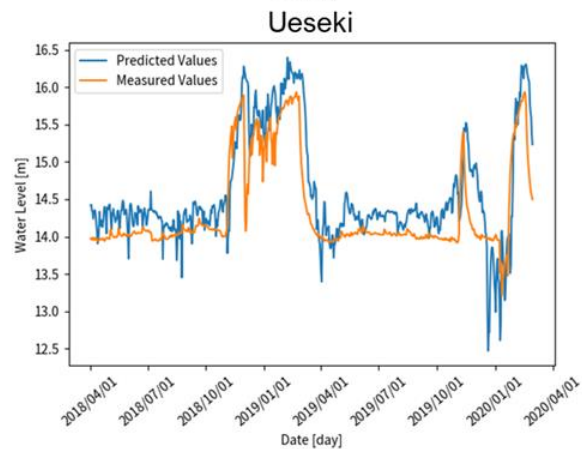
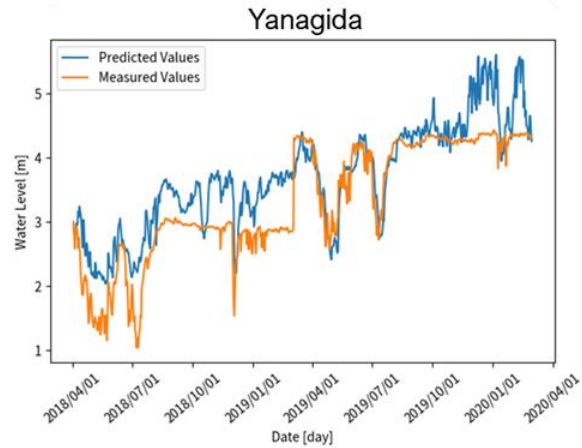
The lower the MSE score, the higher the accuracy

⇒ This model showed a high accuracy

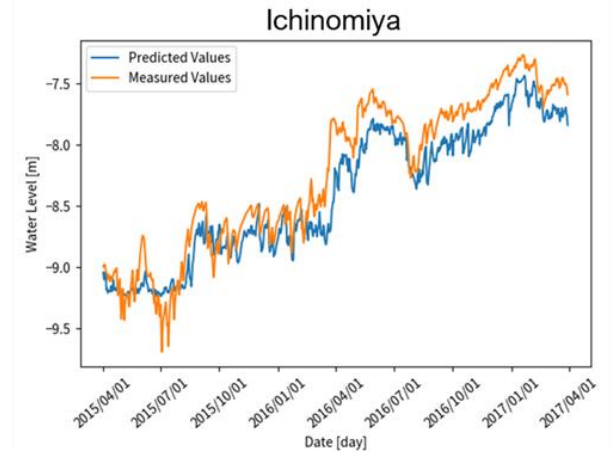
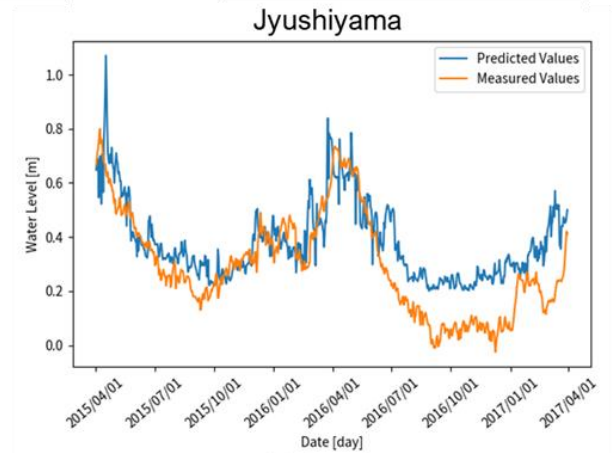
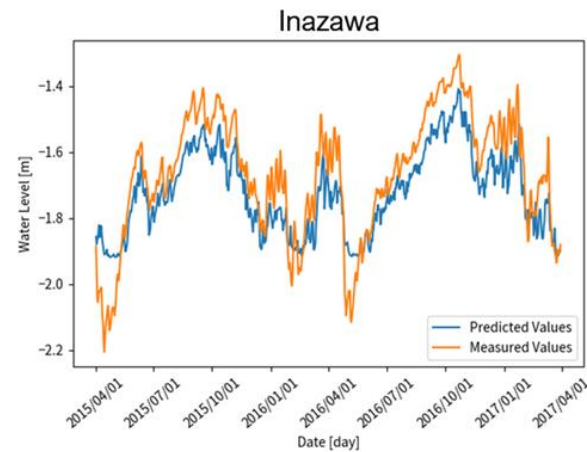
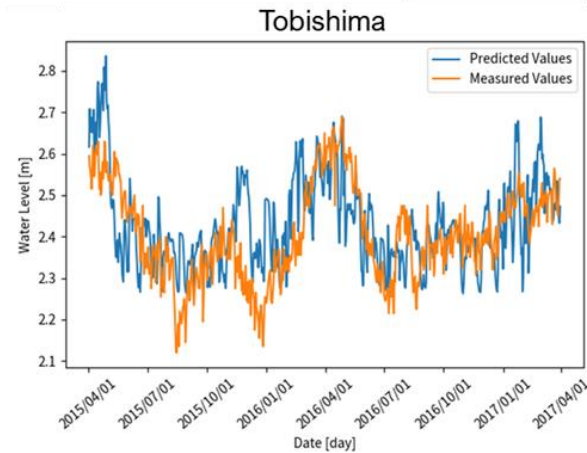
4-1. Results 1 Toyama and Aichi prefecture

Predict to a certain extent of accuracy

Himi city • Takaoka city
Toyama Prefecture



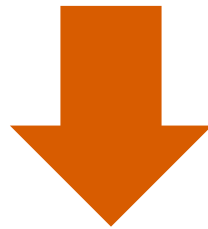
Owari Area
Aichi prefecture



4-1. Results 1

Types of versatility

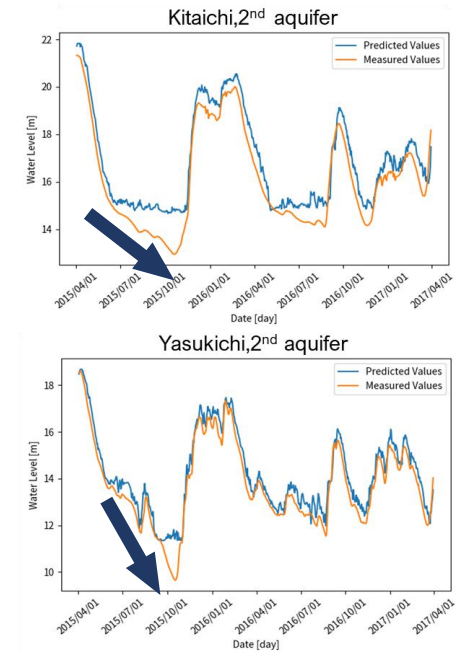
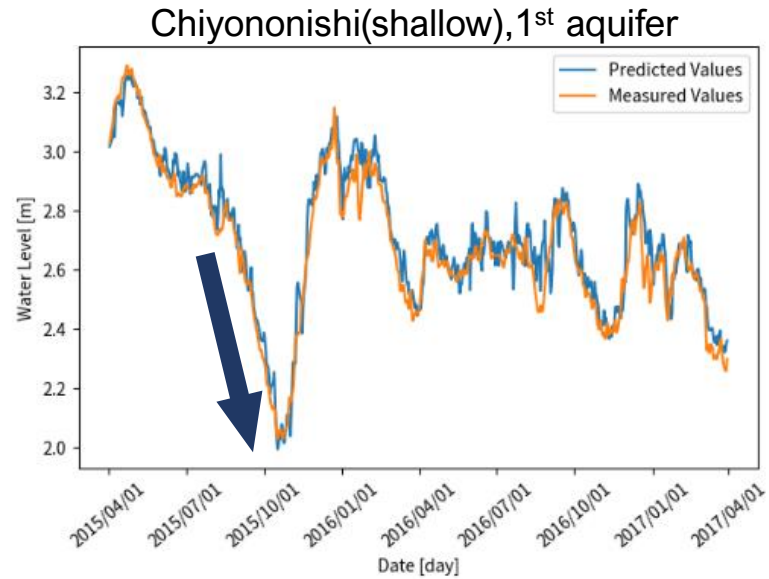
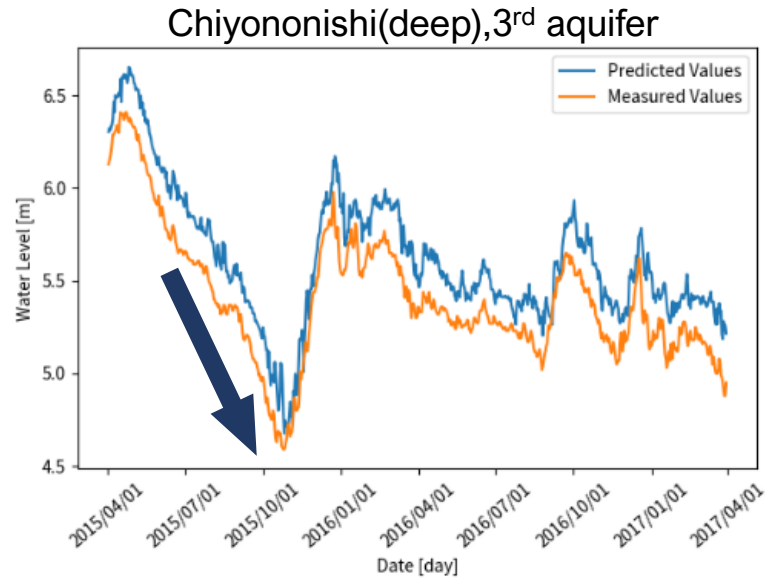
- Between **Wells**
- Between **Aquifers**
- Between **Regions**



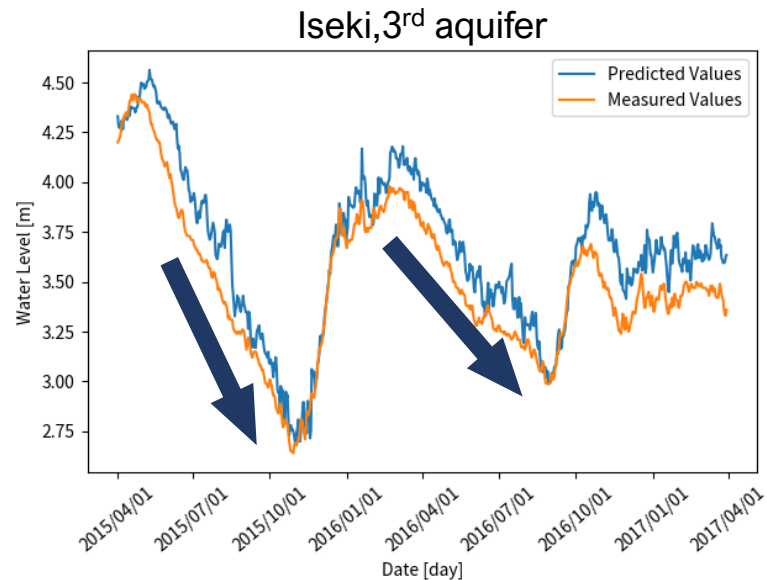
High versatility was confirmed
by using the same model

4-2. Results 2

Periods of unusual variation



←
landslide

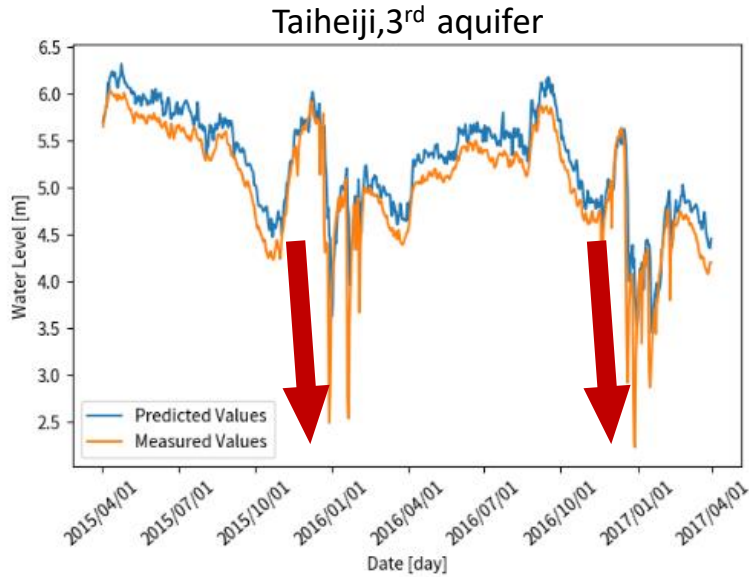


Landslide that occurred in 2015(Forestry Agency)

Prediction with
high accuracy

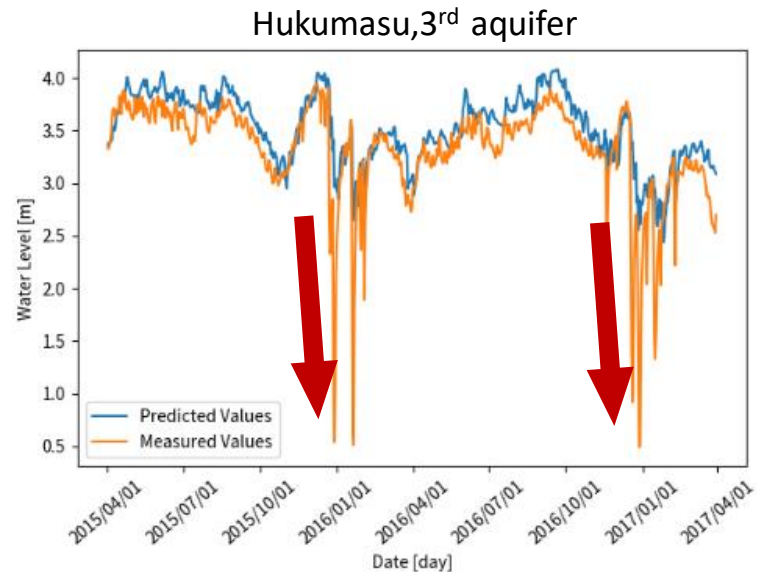
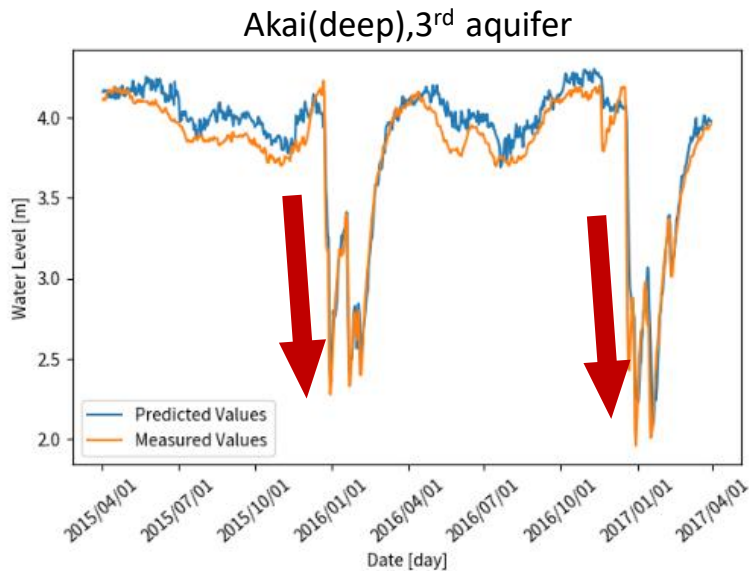
4-2. Results 2

Periods of unusual variation



←
Snow removal
equipment

Half the accuracy of normal

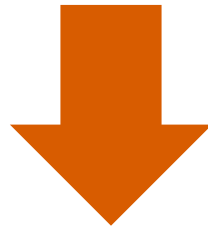


Snow Removal Equipment Using Groundwater (Hukui Shinbun Online)

4-2. Results 2 Tedorigawa Alluvial Fan Area

Periods of unusual variation

- Lowering of the groundwater level due to **landslides**
- Lowering of groundwater level due to operation of snow **removal equipment**



Some prediction is possible

5. Conclusion

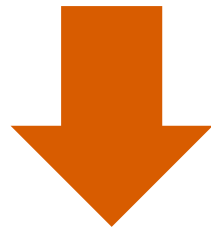
- 1 Only used meteorological data and groundwater level data
- 2 Showed a high accuracy
- 3 *Versatility across wells, aquifers and regions*
- 4 Making Groundwater Level Prediction Easier

It will lead to the protection of the water environment around all people

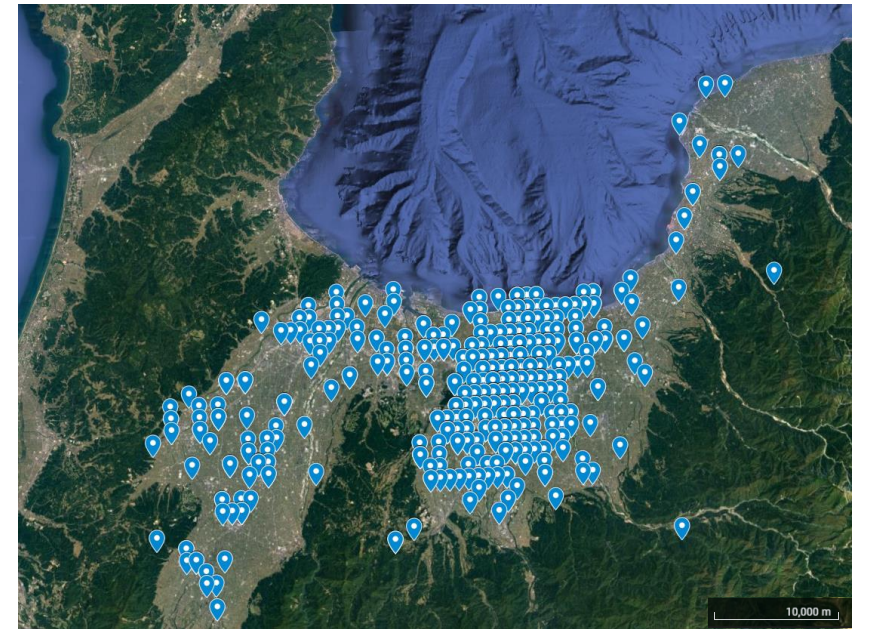
6. Prospect

Use of groundwater data ledgers

- Over 60,000 boring data available
- Model with a larger number of points



Prediction will be possible even at points where there are no observation wells



Points with drilling data (Toyama Prefecture)

7. Acknowledgement

We would like to thank the following people for their various supports

- Mr. Yamada Naoki, Teacher, Kaijo Junior and Senior High School
- Members of Environmental Conservation Division,
Living Environment and Culture Department, Toyama Prefecture
- Members of Environmental Policy Division,
Living Environment Department, Ishikawa Prefecture
- Members of Water and Air Environment Division,
Policy Department, Environment Bureau, Aichi Prefecture
- Mr. Kokubun Kuniki (Technical Support Section,
Tokyo Metropolitan Government Civil Engineering Technical Support
and Human Resource Development Center)
- Mr. Nakayama Toshio (Technical Support Section,
Tokyo Metropolitan Government Civil Engineering Technical Support
and Human Resource Development Center)



8. References

- (1) 農林水産省農村振興局農村政策部農村環境課
『農業地域における持続的な地下水利用の手引き』 2018,p.p.2-5
- (2) 環境省水・大気環境局土壌環境課地下水・地盤環境室
『地下水保全ガイドライン(第二版)～地下水保全と持続可能な地下水利用のために～』 2021,p.p.33～91
- (3) 長野 峻介,野村 和哉,藤原 洋一,田中健二,高瀬 恵次,一恩 英二
『ランダムフォレストを用いた手取川扇状地における地下水位変動解析』 2017
- (5) Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin
『Attention Is All You Need』 2017
- (6) Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich 『Going deeper with convolutions』 2014
- (7) Sergey Ioffe, Christian Szegedy
『Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift』 2015
- (8) Sepp Hochreiter, Jurgen Schmidhuber 『LONG SHORT-TERM MEMOR』 1997
- (9) Diederik P. Kingma, Jimmy Ba 『Adam: A Method for Stochastic Optimization』 2014
- (10) 名古屋通商産業局総務部開発業務課
『石川県手取川・犀川下流地域地下水利用適正化調査報告書』 1974,p.p.40～43
- (11) 林野庁近畿中国森林管理局石川森林管理署
『平成29年度における手取川上流崩壊地対策について』 2017,p.p.1～6
- (12) 柳井清治 『近年の土砂災害シリーズ 2015年5月に白川源流で発生した地すべりと濁水の発生』 2017,p.87