

# RNN による地下水位回帰モデルの構築と 入力データの影響度に関する考察

## CONSTRUCTION OF GROUNDWATER REGRESSION MODEL WITH RNN AND CONSIDERATION ON INFLUENCE OF INPUT DATA

笛田 泰成<sup>1</sup>・伊藤 真一<sup>2</sup>・酒匂 一成<sup>3</sup>・小田 和広<sup>4</sup>

Taisei FUETA, Shinichi ITO, Kazunari SAKO and Kazuhiro ODA

<sup>1</sup> 鹿児島大学大学院理工学研究科 (〒890-0065 鹿児島県鹿児島市郡元 1-21-24)

E-mail: k0249698@kadai.jp

<sup>2</sup> 鹿児島大学学術研究院工学域工学系 (〒890-0065 鹿児島県鹿児島市郡元 1-21-24)

E-mail: itou@oce.kagoshima-u.ac.jp

<sup>3</sup> 鹿児島大学学術研究院工学域工学系 (〒890-0065 鹿児島県鹿児島市郡元 1-21-24)

E-mail: sako@oce.kagoshima-u.ac.jp

<sup>4</sup> 大阪産業大学工学部都市創造工学科 (〒574-8530 大阪府大東市中垣内 3-1-1)

E-mail: oda@ce.osaka-sandai.ac.jp

**Key Words:** *ground water, machine learning, field measurement, recurrent neural network*

### 1. はじめに

集中豪雨の増加によって全国各地で斜面崩壊が多発しており、それに伴った人的被害も発生している。このような斜面崩壊に対しては適切な防災対策が必要であり、斜面崩壊の予知・予測の精度向上も急務となっている。降雨時の斜面崩壊の発生を予測するための手段の一つとして現地斜面のモニタリング<sup>1)</sup>がある。本手法では、降雨に伴って変化する物理量(土中水分量, 斜面変位など)を現地計測し、それらの計測データをリアルタイムに送信することで現時点の斜面の状態を遠隔で監視し, なおかつ, それらの計測データを蓄積することができる。近年では, これまで蓄積された現地計測データを有効活用して斜面崩壊の予測を試みる研究<sup>2)</sup>や土中水分量の計測データを活用した斜面の危険度評価に関する研究<sup>3)</sup>も多数行われている。その一方で, 膨大なデータを活用するための手段として AI 技術の導入も急速に進められており, 土木分野においても様々な適用事例<sup>4)</sup>が報告されているという現状もある。

筆者ら<sup>5)</sup>はこれまでに, AI 技術によって現地計測データを有効活用するための方法として, 体積含水率の現地計測データに基づく Recurrent Neural Network (以下 RNN)<sup>6)</sup>による回帰モデルの構築に関する研究を行ってきた。その結果, RNN によって構築された回帰モデルを用いると, 未学習の現地計測データを十分な精度で予測できることを明らかにした。しかし, 既往の研究<sup>5)</sup>で構築したモデルは降雨情報(半減期を変更した実効雨量)を与えた際に計測が行われている深度における体積含水率を出力するモデルであり, 斜面内の土中水分量に関する空間分布を把握することはできていなかった。降雨時の斜面崩壊に対する危険度評価を行うためには, 深さ方向の飽和度分布や地下水位の増減を予測できるモデルが望ましい。そこで, 本研究では, 地下水位の現地計測データに基づく RNN による地下水位回帰モデルの構築を試みる。このような, 降雨情報を入力してその時の地下水位を出力する RNN モデルを構築することができれば, 降雨量さえわかれば計測データがなくても地下水位を予測できるため, モニタ

リング斜面の拡充に対しても有効であると考えられる。

ニューラルネットワークなどによって構築されたモデルは、予測性能は高くても、モデルの内部構造がブラックボックスになる傾向が指摘されている。RNNによって構築された回帰モデルを用いて斜面の危険度評価を行うためには、そのモデルにおいてどのような入力データ(特徴量)が解析結果に影響を与えているのかを把握しておくことも重要である。近年、AIモデルが導き出した予測結果を説明するための方法として、説明可能なAI(Explainable AI, 以下XAI)<sup>7)</sup>が注目されている。XAIによる説明は、対象とする説明範囲の違いによって、大域説明と局所説明に大別される。本研究では構築したモデルに対して大域説明の方法であるPermutation importanceを適用して、入力データ(特徴量)の影響度についても考察することとする。本研究の目的は、地下水位の現地計測データに基づくRNNによる地下水位回帰モデルの構築と構築されたモデルの入力データの影響度について明らかにすることである。第2章では本研究で用いた現地計測データについて示す。第3章では、実際に降雨情報から地下水位を算出する回帰モデルを構築し、学習データに対する再現性やテストデータに対する汎化性能を評価する。第4章では、Permutation importanceによるモデルの説明を行い、地下水位の回帰モデルに対する入力データの影響度について考察する。

## 2. 現地計測データ

本研究の対象斜面<sup>8)</sup>は、風化した斑れい岩で構成されており、斜面長は14m程度、斜面勾配は45°程度である。当該斜面周辺では2014年の豪雨などによる11箇所の崩壊地が確認されている。この斜面周辺では、崩壊した土層の物理、浸透および強度特性を把握するための室内試験が行われている。粒度分布や透水係数などは既往の文献の試験結果<sup>8)</sup>を参照されたい。図-1は本研究で対象とした斜面の土層断面と観測機器の設置状況を示している。この斜面では、過去の崩壊により斜面下部のB-3地点とB-4地点の間に2m程度の滑落崖がある。合計13ヶ所で行った簡易貫入試験の結果を考慮して、土壌水分計付傾斜計、多段式地中変位計、地下水観測孔などを設置した。ここでは、体積含水率、地下水位、斜面変位、傾斜角度を1時間間隔で計測している。本研究では、地下水位の計測データに着目しており、図-1における斜面中部のW-2'で計測されたデータが、降雨に反応して地下水位が増減しておりデータ欠損も少なかったことから、W-2'の計測データを用いてこの後の分析を行うこととした。なお、降雨情報に関しては対象斜面近傍のアメダスの降雨データを用いた。図-2は本研究で用いた現地計測データを示している。同図の赤線は計測深度2.8mで計測された地下水位の経時変化を表している。計測期間は2017年5月19日から2018年9月21日までである。2017年10月

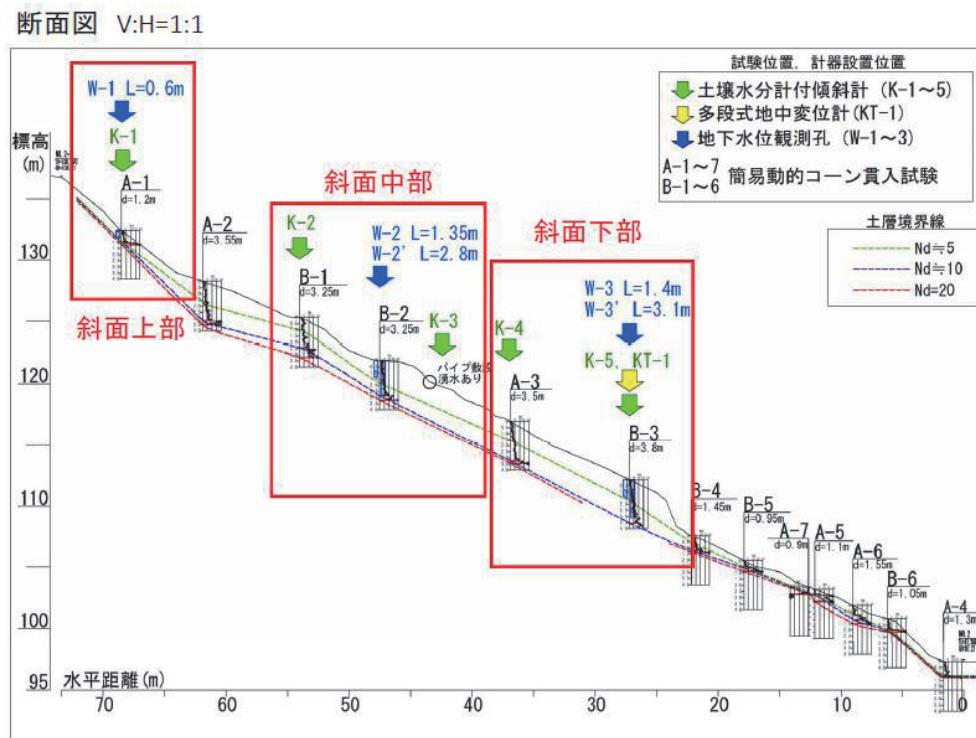


図-1 斜面の土層断面と観測機器の設置状況

22日の台風21号、2018年7月4日から7月8日頃までの2018年7月豪雨の影響で地下水位が地表面近くまで上昇していることがわかる。これらのデータをRNNによる回帰モデルを構築するために用いる学習データと構築したモデルの予測性能について評価するためのテストデータに分割した。図-3は本研究で用いる学習データとテストデータの経時変化をそれぞれ示している。2017年8月1日から12月31日までの現地計測データを学習データとし、2018年2月20日から7月21日までの現地計測データをテストデータとして用いた。

### 3. RNNによる地下水位回帰モデルの構築

本章では、地下水位の現地計測データに基づくRNNによる地下水位回帰モデルの構築を試みる。まず、本研究で用いた解析手法であるRNNについて説明する。そして、モデルを構築するための解析条件を示した後、対象斜面において現地計測されたデータ(降雨量と地下水位)を用いて回帰モデルを構築し、テストデータに対する汎化性能を確認する。

#### (1) リカレントニューラルネットワーク(RNN)

RNN<sup>6),9)</sup>は、ニューラルネットワークの構造に過去の情報を保持する隠れ層を定義することで、時間の概念を反映することができる。本研究では、RNNの中でもLong Short-Term Memory(以下LSTM)<sup>10)</sup>を用いて計算を行った。図-4はLSTMの概念図を示している。×印のノードは掛け合わせを意味している。

LSTMでは、長期の情報を保持するメモリセルと三つのゲート(入力ゲート、忘却ゲート、出力ゲート)がある。まず、メモリセルは次式で表される。

$$C_t = a_t + C_{t-1} \quad (1)$$

ここに、 $t$ は時間、入力 $x_t$ を活性化関数 $g_t$ で活性化した値を $a_t$ 、メモリセルの値は $C_t$ と表される。これにより過去の情報が保持され、長期の時間依存性を考慮することができる。

次に、ある入力に対しての信号があった場合にのみ機能する入力ゲート、出力ゲートに注目する。入力ゲートの値を $i_t$ 、出力ゲートの値を $o_t$ とする。

$$i_t = \sigma(W^i x_t + U^i h_{t-1} + b^i) \quad (2)$$

$$o_t = \sigma(W^o x_t + U^o h_{t-1} + b^o) \quad (3)$$

ここに、 $\sigma$ は活性化関数、 $h$ は隠れ層の値、 $W$ 、 $U$ は重み行列、 $b$ はバイアスベクトルである。これらのゲートを設けることで必要な時に過去のデータを活用することができ、効率的に長期に及ぶ学習を行うことが可能になる。

最後に忘却ゲートだが、メモリセル内の情報を書き換えるためのゲートである。時系列データを扱う場合、過去の情報を保存するだけでなく、忘れることも重要になる。忘却ゲートも入力ゲート、出力ゲートと同様に次式で表される。

$$f_t = \sigma(W^f x_t + U^f h_{t-1} + b^f) \quad (4)$$

入力ゲートや忘却ゲートを導入することで、式(5)は次式のようになる。

$$C_t = i_t \circ a_t + f_t \circ C_{t-1} \quad (5)$$

ここに、 $\circ$ はアダマール積といい、ベクトルの要素積を表している。このようにLSTMではメモリセルと三つのゲートを導入することで、過去の情報の保

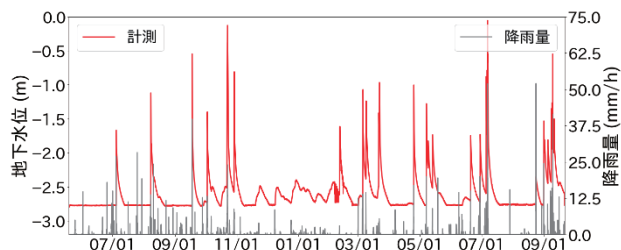
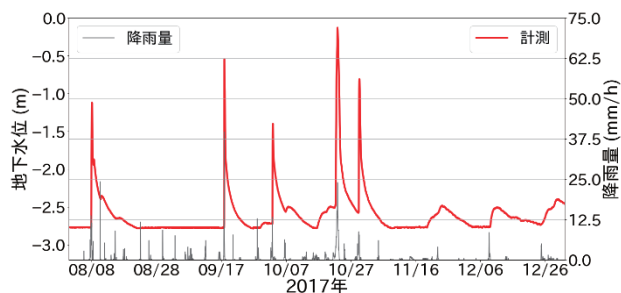
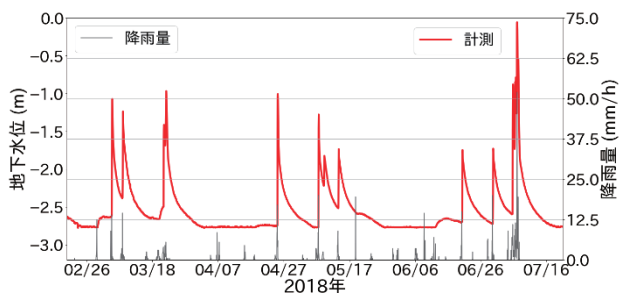


図-2 現地計測データ



(a) 学習データ



(b) テストデータ

図-3 学習データとテストデータの経時変化

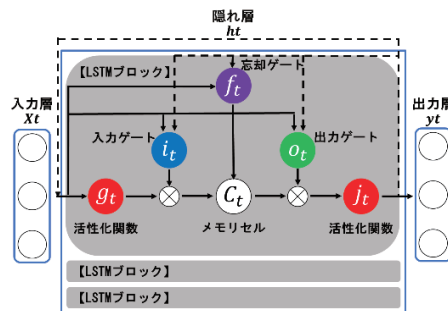


図-4 LSTMの概念図

存, ある入力に対しての信号があった場合に必要なタイミングでの過去の情報の使用, 書き換えを行っている。

## (2) 解析条件

降雨情報を入力して地下水位を出力する回帰モデルを構築する。本研究では, 既往の研究<sup>5)</sup>を参考にし, 入力データには半減期を変更した48通りの実効雨量を用いた。出力データには各時刻における地下水位の現地計測データを与えた。

実効雨量<sup>11)</sup>は以下の式を用いて算出した。

$$X_t = X_{t-1}e^{\beta} + R_t e^{\beta/2} \quad (6)$$

ここに,  $X_t$ は時刻 $t$ における実効雨量,  $R_t$ は時刻 $t-1$ から $t$ までの降雨量であり, 本研究では, 1時間雨量がこれに該当する。減少係数 $\beta$ は, 半減期 $M$ を用いて次式のように表される。

$$\beta = \ln(0.5)/M \quad (7)$$

実効雨量は流出孔が一つのタンクモデルと考えることができ, 半減期は流出孔の大きさを表している。そのため, 先行降雨の影響を考慮した長期的な降雨指標としては, 半減期が24h, 72hのような半減期の長い実効雨量が用いられ, より短期的な降雨強度の影響を評価するための降雨指標としては, 半減期が1.5hのような半減期の短い実効雨量が用いられることが多い。本研究では, 実効雨量の半減期を1時間, 2時間, …, 48時間と変更して48通りの実効雨量を算出した。図-5は対象斜面における実効雨量の経時変化の一例を示している。半減期が長いほど減少に時間がかかるため, 実効雨量は全体的に高い値となる。このように実効雨量は, 時間の経過に伴って値が滑らかに減少するため, 地下水を表現できる可能性があると考えられる。

## (3) 解析結果

図-3に示す学習データを用いて地下水位の回帰モデルを構築した。そして, 構築されたモデルに対して学習データを与えて構築されたモデルの再現性を確認した。図-6は学習データに対する再現解析結果を示している。同図より, 構築したモデルは降雨情報を用いて地下水位の現地計測データを良好に再現できていることが確認できる。また, 2017年10月22日を中心として台風21号の激しい降雨時の急激な地下水位の上昇も十分に再現ができていることがわかる。次に, 構築されたモデルに対して未学習のテストデータを適用してモデルの汎化性能を評価した。図-7はテストデータに対する予測解析結果を示しており, 図-8は2018年7月豪雨の予測解析結果

を示している。テストデータはモデル構築時には用いていないにも関わらず, 解析結果は計測データを概ね再現できている。また, 2018年7月豪雨の影響により地下水位が地表面近くまで上昇している計測データに対してもある程度再現ができていることから, 構築されたモデルは十分な汎化性能を有しているモデルであると判断した。以上の結果から, 半減期 $M$ を変更した実効雨量を多数計算し, それらの計算結果をRNNに学習させることで, 地下水位の現地計測データを良好に再現ができる回帰モデルの構築が可能であることが明らかになった。また, 本研究ではRNNによるモデル化を目的としているため

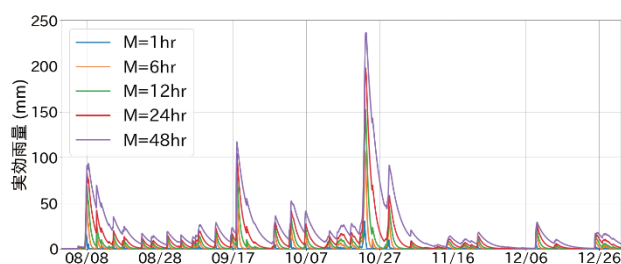


図-5 実効雨量の経時変化の一例

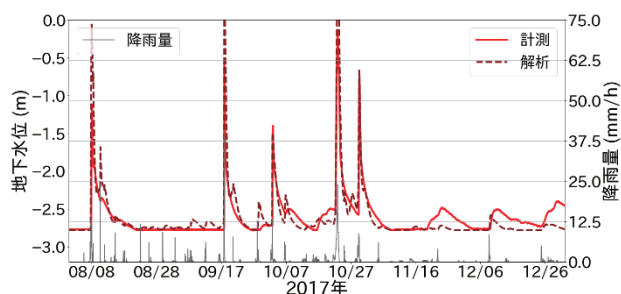


図-6 学習データに対する再現解析結果

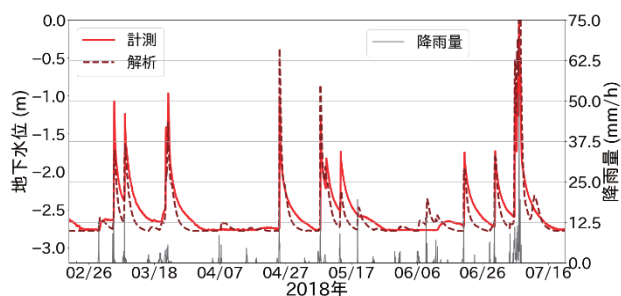


図-7 テストデータに対する予測解析結果

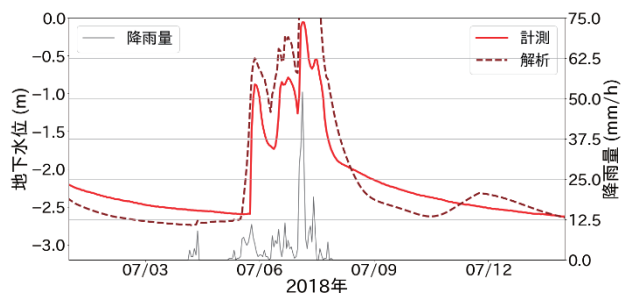


図-8 2018年7月豪雨の予測解析結果

W-2'地点に議論を絞ったが、最終的に斜面全体の安定計算などを行う際は計測地点ごとに地下水位の回帰モデルを構築し、それらの解析結果から安定性を評価することが望ましい。

#### 4. 構築されたモデルの入力データの影響度

ここでは、前章で構築した回帰モデルの入力データの影響度について確認することで、対象斜面の地下水位挙動に影響を与える実効雨量の半減期  $M$  について議論する。具体的には、Permutation importance を用いて、どのような入力データが解析結果に影響を与えているのかについて分析する。

##### (1) Permutation Importance

機械学習によって構築されたモデルにおける各特徴量の影響度を把握することは、モデルを理解するために重要である。図-9 は特徴量 A の影響度を確認する場合の Permutation importance の概念図を示している。Permutation importance では、確認したい特徴量の値をランダムに並べ替えて構築されたモデルに与える。これにより、正解値と予測値の間の誤差が

特徴量A	特徴量B	特徴量C	...	正解値	予測値
A1	B1	C1	...	0.11	0.12
A2	B2	C2	...	0.18	0.17
A3	B3	C3	...	0.63	0.65
A4	B4	C4	...	0.77	0.79
⋮	⋮	⋮	⋮	⋮	⋮

特徴量Aのみをランダムに並び替え  正解値と予測値の間の誤差が増加

特徴量A	特徴量B	特徴量C	...	正解値	予測値
A38	B1	C1	...	0.11	0.56
A62	B2	C2	...	0.18	0.47
A11	B3	C3	...	0.63	0.22
A247	B4	C4	...	0.77	0.19
⋮	⋮	⋮	⋮	⋮	⋮

図-9 Permutation importance の概念図

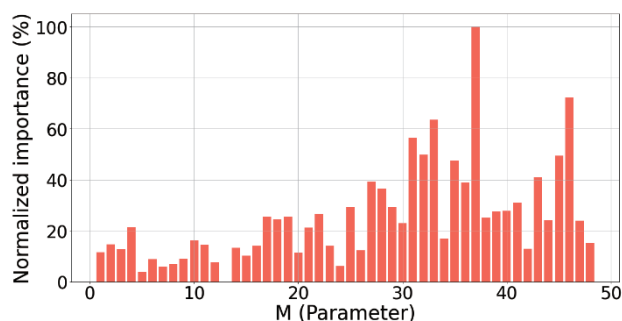


図-10 Permutation importance によって得られた各特徴量の影響度

増加することが予想されるが、この誤差の増加量が大きい特徴量は、モデルの予測がその特徴量に依存していると考えられる。このような特徴量の値のランダムな並べ替えを全ての特徴量に対して行い、それぞれの誤差の増加量を比較することで重要な特徴量を判断する。Permutation importance は並べ替えのランダム性に依存することや同一の傾向を持つ特徴量があると影響度が低く算出されることがあるなどの留意点<sup>7)</sup>もあるが、モデルの再学習の必要がないため高速で各特徴量の影響度を算出することができ、シンプルかつ効果的な方法である。

##### (2) Permutation importance による構築されたモデルの説明

構築されたモデルの大域説明(どの特徴量が効いているかの説明)を試みた。図-10 は Permutation importance によって得られた各特徴量の影響度を示している。なお、この影響度は、最も影響が大きかった実効雨量(半減期:  $M=37$  時間)の影響度が 100% になるように正規化している。これらの結果から、本研究で構築した地下水位の回帰モデルは半減期が長い実効雨量の影響が相対的に大きく、半減期の短い実効雨量の影響は小さいことが確認できる。これは、対象斜面で計測された地下水位の経時変化と半減期が長い実効雨量の経時変化が対応しているためであると考えられる。このように構築されたモデルに対して影響の大きい入力データを把握しておくことで、その斜面における地下水挙動の特徴も把握することができる。このようなモデルの中身の説明は、防災分野に機械学習を導入する上で重要な観点であると考えられる。

#### 5. おわりに

本研究では、地下水位の現地計測データに基づく RNN による地下水位回帰モデルの構築と構築されたモデルの入力データの影響度について明らかにすることを目的として、まず、降雨情報を入力してその際の地下水位を出力する回帰モデルを構築した。そして、構築されたモデルに対してモデルの説明をする手法の一つである Permutation importance を適用して、その結果について考察した。得られた知見を以下に示す。

- (1) 対象斜面における地下水位の現地計測データを用いて RNN による回帰モデルを構築した。構築されたモデルは学習に用いたデータだけでなく、未学習のテストデータを与えた際も地

下水位の現地計測データを概ね再現できることがわかった。このことから、モデル構築のためのアルゴリズムとして RNN を用いて、入力データとしては半減期を変更した多数の実効雨量の計算値を与える方法は、地下水位回帰モデルを構築するための方法として有効であると考えられる。

- (2) **Permutation importance** によってモデルの構築に影響の大きい入力データについて考察した。その結果、本研究の対象斜面では、半減期の大きい実効雨量が入力データとして大きく影響していることがわかった。このような構築されたモデルに対して影響の大きい入力データを把握することは、その斜面における地下水挙動の特徴の把握に繋がるため、防災分野に機械学習を導入する上でも重要な観点であると考えられる。

以上の結果から、本研究の目的であった降雨情報からその時の地下水位を算出する RNN による回帰モデルに対しては入力データとしては実効雨量を用いる方法が適切であることが明らかになった。また、構築されたモデルに対して **Permutation importance** を適用した場合については、深層学習などの内部構造が複雑な AI が導き出した予測結果を説明するための方法として有効であることが明らかになった。

#### 参考文献

- 1) 酒匂一成, 横田祐介, 里見知昭, 檀上徹, 深川良一: 無線センサネットワークを利用した斜面内の負の間隙水圧の長期多点計測システム, 土木学会論文集 C(地圏工学), Vol.74, No.2, pp.141-163, 2018.
- 2) 笹原克夫, 酒井直樹: 地表面変位と地下水位のモニタリングに基づく斜面崩壊発生予測, 地盤工学ジャーナル, Vol.9, No.4, pp.671-685, 2014.
- 3) 檀上徹, 酒匂一成, 深川良一, 酒井直樹, 岩佐直人, Nghiem Minh Quang: 降雨量, 不飽和浸透挙動, 変位量の観測結果に基づく降雨時表層すべり型崩壊過程の検証, 土木学会 論文集 C (地圏工学) Vol.68, No.3, pp.508-525, 2012.
- 4) 古田均, 野村泰稔, 広兼道幸, 一言正之, 小田和広, 秋山孝正, 宇津木慎司: AI のインフラ分野への応用, 電気書院, 2019.
- 5) 伊藤真一, 小田和広, 小泉圭吾, 酒匂一成: 体積含水率の現地計測データの予測に対するリカレントニューラルネットワークの適用性, AI・データサイエンス論文集, Vol. 1, No.1, pp.445-452, 2020.
- 6) Chollet, F.: Python と Keras によるディープラーニング, マイナビ出版, 2019.
- 7) 大坪直樹, 中江俊博, 深沢祐太, 豊岡祥, 坂元哲平, 佐藤誠, 五十嵐健太, 市原大暉, 堀内新吾: XAI(説明可能な AI)そのとき人工知能はどう考えたか?, リックテレコム, 2021.
- 8) 小田和広, 矢野晴彦, 鏡原聖史, 深川良一, 片岡沙都紀, 小泉圭吾, 小山倫史, 笹原克夫, 辻野裕之, 藤本将光, 伊藤真一, 大段恵司: 京都府綾部市安国寺裏斜面におけるモニタリング計測器配置の検討, Proceedings of the Kansai Geo-Symposium 2017, pp.212-217, 2017.
- 9) 巢鴨悠輔: 詳細ディープラーニング第2版, マイナビ出版, 2019.
- 10) Hochreiter, S. and Schmidhuber. J.: Long short-term memory, *Neural Computation*, Vol.9. No.8, pp.1735-1780, 1997.
- 11) 小杉賢一朗, 藤本将光, 山川陽祐, 正岡直也, 糸数哲, 水山高久, 木下篤彦: 山体基岩内部の地下水位変動を解析するための実効雨量に基づく関数モデル, 砂防学会誌, Vol.66, No.4, pp.21-32, 2013.

(2022. 5. 30 受付)