

降雨時の地下水位発生予測に対する機械学習の適用性

鹿児島大学工学部 学生会員 ○馬場咲也子

鹿児島大学学術研究院理工学域工学系 正会員 伊藤真一

鹿児島大学学術研究院理工学域工学系 正会員 酒匂一成

1. はじめに

降雨時の斜面崩壊の発生には土中の水分状態が大きく影響する。その中でも、地下水位の発生は斜面崩壊の予知を考える上で重要な要因である。近年のセンシング技術の発達によって、地下水位などの土中の水分状態の現地計測が数多く普及している¹⁾。これにより、様々な斜面で現地計測データは膨大に蓄積されている。しかし、それらのデータを有効活用して、地下水位が発生する降雨量を客観的に求めた事例は多くない。一方で、近年のコンピューターの発達に伴って、様々な分野で機械学習が用いられている。本研究では、現地計測データを用いて地下水位が発生する降雨量を予測する問題に対する機械学習の適用性を検証する。

2. 現地計測データ

本研究では、京都市東山区の重要文化財後方の斜面で現地計測されたデータを用いて分析を行った。対象斜面では、雨量計、テンシオメータ、温度計などが設置され、降雨量、間隙水圧、地中温度がそれぞれ計測されている¹⁾。テンシオメータは5本設置されており、それぞれ深度20cm, 40cm, 60cm, 80cm, 100cmにおける間隙水圧を計測していた。本研究では、地下水位の発生について議論するために、深度100cmの計測データを用いた。図-1は10分間雨量と深度100cmで計測された間隙水圧の経時変化を示している。期間は2004年8月9日から2004年12月31日である。降雨に伴って間隙水圧は上昇し、降雨が終了すると間隙水圧は減少していることが分かる。

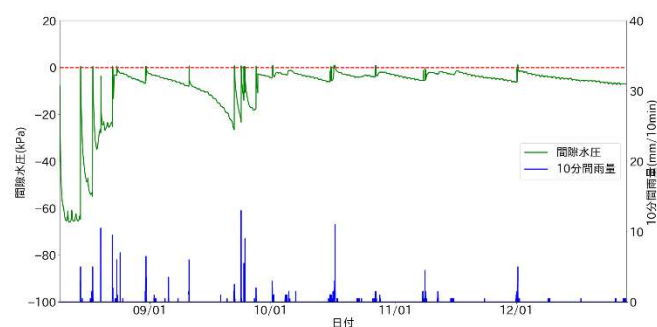


図-1 現地計測データ

3. 解析方法

現地の雨量計で計測された10分間雨量に基づいて、時間雨量と土壌雨量指数²⁾を算出した。この2つの降雨指標を入力ベクトル(説明変数)として、深度100cmにおける地下水位の有無(飽和・不飽和)を教師データ(目的変数)として機械学習に適用するデータを作成した。飽和・不飽和のクラス分類は、図-1における赤色の点線、つまり、間隙水圧が0.0kPa以上かどうかで判断した。図-2は適用データのスネーク曲線と飽和・不飽和のクラス分類結果を示している。全体的に、降雨指標の値が大きいほど飽和している傾向が確認できる。

機械学習のアルゴリズムとしては、ランダムフォレスト(Random Forest, 以下RFと呼ぶ)を用いた。解析ツールは、Pythonのscikit-learnを使用した³⁾。なお、機械学習には、モデル構築時に使用したデータに対して過剰に適合するモデルを構築してしまい、その結果、それ以外のデータへの汎化能力が著しく低下する過学習の可能性がある。本研究では、適用データをランダムに75%と25%に分割し、75%を学習データ、25%を検証データとして使用することにした。

4. 解析結果

第3章で作成した学習データを用いて、RFによる予測モデルを構築した。構築された予測モデルに対して、学習データと検証データをそれぞれ適用した結果、どちらも99%を超える精度が確認された。これは、現地計測データにおける不飽和のクラスが飽和のクラスと比較して異常に多かったためである。そこで、時間雨量と土壌雨量指数の軸上に機械学習の予測結果と実

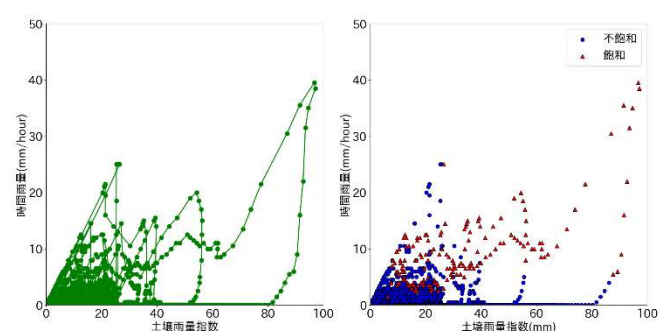


図-2 スネーク曲線と飽和・不飽和のクラス分類結果

際の計測データの値をプロットした。図-3 は学習データに対する識別結果を、図-4 は検証データに対する識別結果を示している。図中の赤色の領域は RF によって飽和と予測された降雨の組み合わせ(時間雨量と土壌雨量指数)であり、青色の領域は不飽和と予測されたそれを表している。このため、図-3 と図-4 における領域の色分けは同一である。次に、図-3 と図-4 における橙色の三角は現地計測データによって飽和していた降雨の組み合わせであり、青色の丸は不飽和状態のそれを表している。これらの値は学習データと検証データで異なる。図-3 と図-4 の結果から、RF によって構築された予測モデルは、学習データの傾向を精度良く再現できているだけでなく、検証データに対する汎化能力も高いことが分かる。図-5 は RF による予測結果とそれに基づく基準線を示している。赤色の横線の値は時間雨量 4.0(mm/h) であり、赤色の縦線の値は土壌雨量指数 26.3 である。まず、時間雨量が 4.0(mm/h) を下回ると、ほとんどの領域で不飽和と予測されている。この傾向は土壌雨量指数の値が大きい(長期的に長雨が降っていた)場合も同様である。したがって、対象斜面における地下水位の発生には時間雨量の影響が大きく、その基準値は 4.0(mm/h) であるといえる。次に、赤色の縦線に注目すると、土壌雨量指数 26.3 以上、かつ、時間雨量 4.0(mm/h) 以上である場合は、ほとんどの領域で地下水位が発生する(飽和)と予測されている。つまり、雨が継続的に降り続いて土中に多くの水が保水されている状況で、時間雨量 4.0(mm/h) 以上の雨が降った場合は、深度 100cm において地下水位が発生している可能性は極めて高いと考えられる。

以上の結果から、現地計測データを用いて地下水位が発生する降雨量を予測するという問題に対して機械学習は十分に適用可能であるといえる。このように、降雨基準を予め把握しておくことができれば、それ以上の現地計測を行わなくても、地下水位の発生を降雨量によって予測できると考えられる。

5. まとめ

本研究では、現地計測データを用いて地下水位が発生する降雨量を予測する問題に対する機械学習の適用性について検証した。得られた知見を以下に示す。

(1) 構築された予測モデルは、学習データの傾向を再現できるだけでなく、検証データに対する汎化能力も高いことが分かった。

- (2) 対象斜面における地下水位の発生には時間雨量の影響が大きく、その基準値は 4.0(mm/h) であることが分かった。
- (3) 対象斜面では、土壌雨量指数 26.3 以上、かつ、時間雨量 4.0(mm/h) 以上である場合に地下水位が発生する可能性が高いことが分かった。
- (4) 現地計測データを用いて地下水位が発生する降雨量を予測するという問題に対して機械学習は十分に適用可能であることが分かった。

参考文献：1) 酒匂一成，深川良一，岩崎賢一，里見知昭，安川郁夫：降雨時の斜面災害防止のための重要文化財周辺斜面における現地モニタリング，地盤工学ジャーナル，Vol.1, No.3, pp.57-69, 2006. 2) 岡田憲治，牧原康隆，新保明彦，永田和彦，国次雅司，斉藤清：土壌雨量指数，日本気象学会機関紙「天気」，Vol.48, No.5, pp.59-66, 2001. 3) Andreas C. Muller, Sarah Guido: Pythonではじめる機械学習—scikit-learnで学ぶ特徴量エンジニアリングと機械学習の基礎，オライリージャパン，2017.

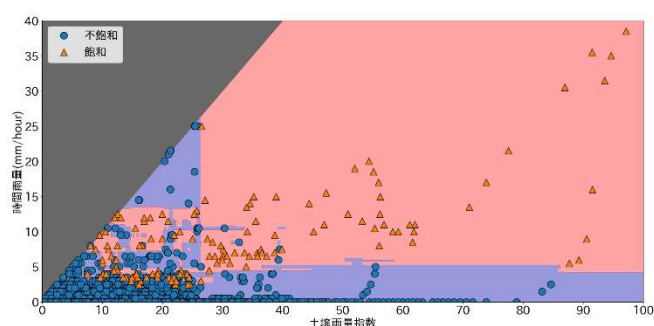


図-3 学習データに対する識別結果

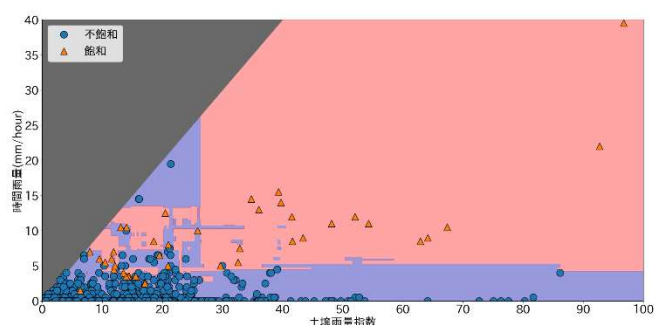


図-4 検証データに対する識別結果

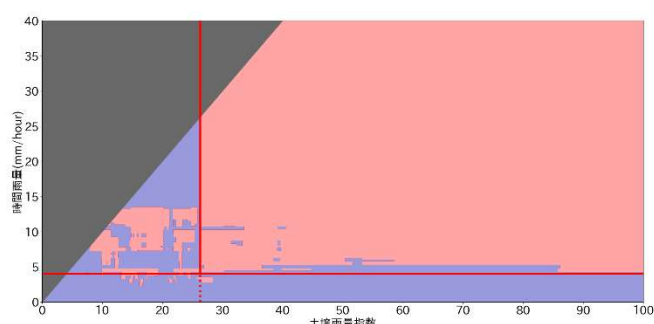


図-5 RFによる予測結果とそれに基づく基準線