

## 機械学習による宮古島地下ダム流域の硝酸態窒素濃度予測

Forecasting nitrate nitrogen concentration in Miyako subsurface dams using machine learning

武氣 祐璃子\*・○高野 愛\*\*・鈴木 麻里子\*\*\*・井上一 哉\*\*\*

Yuriko Takeki, Ai Takano, Mariko Suzuki and Kazuya Inoue

## 1. 緒論

地表水に乏しい沖縄県宮古島では、水資源として地下水を貯留する地下ダムが建設され、水の安定供給に貢献している。一方、地下ダムを保全管理する上で硝酸態窒素(以下、 $\text{NO}_3\text{-N}$ と記す)濃度の現状把握と変動予測は不可欠である。しかしながら、効率的に、かつ、迅速に地下ダムの $\text{NO}_3\text{-N}$ の濃度の変動を予測する手法は確立されていない。そこで本研究では、機械学習を用いて $\text{NO}_3\text{-N}$ 濃度の予測と濃度分布の可視化手法を構築する。

## 2. 機械学習の手順とアルゴリズムの選定

図1に示すように、宮古島は断層により砂川、仲原、福里の3流域に区分され、観測点K1~K13では2004年4月1日から60日間隔で水質を観測している。水質項目は $\text{NO}_3\text{-N}$ 、 $\text{Cl}^-$ 、pHなど13種類であり、例として図2に観測点K2の $\text{NO}_3\text{-N}$ 、 $\text{SO}_4^{2-}$ 、 $\text{Na}^+$ の経時変化を示す。

機械学習では12種の水質データに加えて、2004年4月1日からの経過日数の計13種を説明変数とし、目的変数である $\text{NO}_3\text{-N}$ 濃度を予測する。本研究では、2021年1月31日(以下、予測対象日と記す)の観測データを予測できる機械学習の手法を構築すべく、目的変数と説明変数ともに将来予測では未知であるため、まず、水質項目ごとに観測イベント間の変化量を求めた。図3に観測点K2における $\text{SO}_4^{2-}$ と $\text{Na}^+$ の変化量のヒストグラムを示す。次に、変化量を正規分布とみなしてLatin hypercube sampling<sup>1)</sup>にて各水質の変化量をランダム抽出した。最後に、予測対象日の1つ前の観測日である2020年11月24日の観測値に加算して説明変数とし、 $\text{NO}_3\text{-N}$ 濃度を予測した。抽出・加算・予測の過程を4000回繰り返し、得られた $\text{NO}_3\text{-N}$ 濃度のアンサンブル平均を当該観測点の予測値とした。

本研究では、予測に適した機械学習のアルゴリズムを選定すべく、Extra Trees (ET), Random Forest (RF), Support Vector Regression (SVR) の予測精度

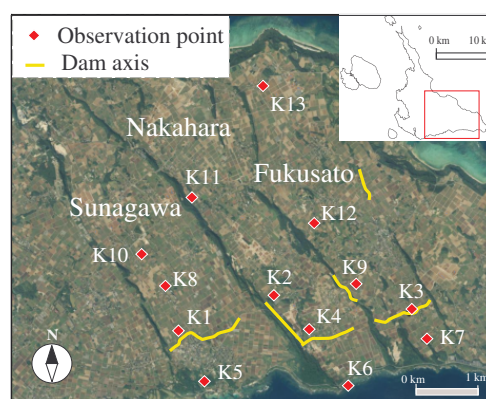


図1：宮古島の水質観測点  
(国土地理院のデータより作成)

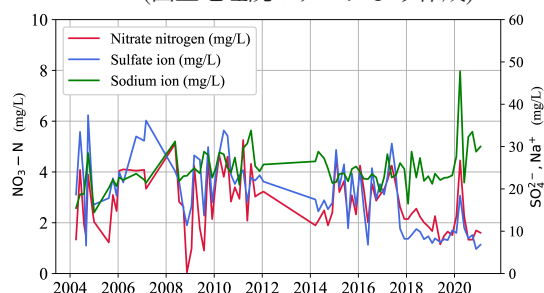


図2：観測点K2の水質項目の経時変化

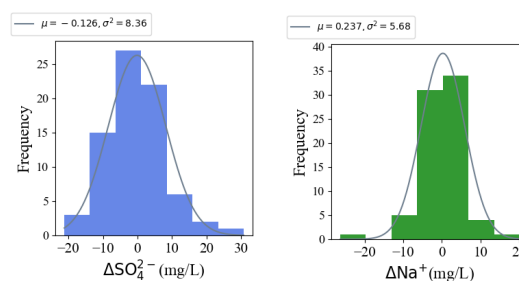


図3：観測点K2の水質の変化量と確率分布

\* 関西電力株式会社：Kansai Electric Power Co., Inc. \*\* 神戸大学農学部：Faculty of Agriculture, Kobe University

\*\*\* 神戸大学大学院農学研究科：Graduate School of Agricultural Science, Kobe University

キーワード：地下ダム、機械学習、宮古島、硝酸態窒素

を比較した。図4に各流域の代表点K1～K3の予測結果を示す。ETは高い精度を示しており、過学習しにくい点<sup>2)</sup>を踏まえて、NO<sub>3</sub>-N濃度予測のアルゴリズムに決定した。

### 3. 宮古島地下ダム流域の硝酸性窒素濃度マップの提示

ダム湖の保全管理に向けて、点情報を面情報に拡張すべく、各観測点のアンサンブル予測結果に基づき、逆距離加重法にて流域ごとに濃度分布を可視化した。予測対象日の実際の観測値から作成した濃度マップを図5に示す。仲原流域におけるNO<sub>3</sub>-N濃度は砂川、福里流域と比較すると低い傾向にある。仲原流域では地下ダムが建設中であり、止水壁を完全に締め切っていないため、南東方向の地下水流れに従って海へNO<sub>3</sub>-Nは流出していると推測される。また、いずれの流域においても南側のNO<sub>3</sub>-N濃度は高い傾向にあることから、止水壁付近ではNO<sub>3</sub>-Nが滞留していると推測される。

構築した手法の予測性能を評価するため、予測対象日の68日前(2020年11月24日)までの水質変動の履歴からNO<sub>3</sub>-N濃度を予測し、作成した濃度マップを図6(a)に示す。精度を評価するため、図6(b)には予測値から観測値

を差し引いた差分マップを作成した。予測値と観測値の誤差の範囲は最大でも±1 mg/lであり、良好な結果と言える。図6(c)は予測対象日の186日前(2020年7月29日)までの水質変動の履歴からNO<sub>3</sub>-N濃度を予測し、作成した濃度マップである。図6(d)に示す差分マップと図6(b)を比較すると、予測対象日までの日数が増えると、予測値と観測値のずれは大きくなる一方、±2 mg/lの範囲に誤差は収まっている。現状のNO<sub>3</sub>-N濃度は3～4 mg/lで推移していることを考慮すると、予測として十分と判断される。

### 4. 結論

本研究では、機械学習を用いて宮古島地下ダム流域のNO<sub>3</sub>-N濃度の予測と濃度分布の可視化手法を構築した。68日後、186日後までの予測の誤差はそれぞれ±1 mg/l、±2 mg/lの範囲に収まり、精度は良好と判断される。

参考文献：1) Huntington, D.E., Lyrntzis, C.S.: Improvements to and limitations of Latin hypercube sampling, *Prob. Eng. Mech.*, 13(4), pp.245-253, 1998. 2) Geurts, P., Ernst, D. and Wehenkel, L.: Extremely randomized trees, *Machine Learning*, 63, pp.3-42, 2006.

謝辞：本研究は、(一社)土地改良建設協会「2021年度国営事業地区等フィールド調査学生支援事業」による支援を受けました。

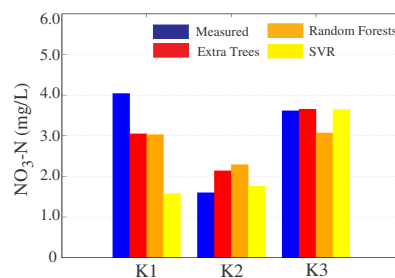


図4：観測点K1～K3での精度比較

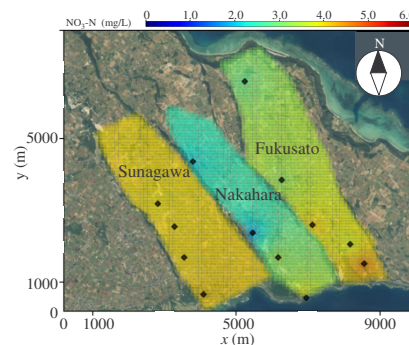


図5：2021/1/31の観測値の濃度マップ

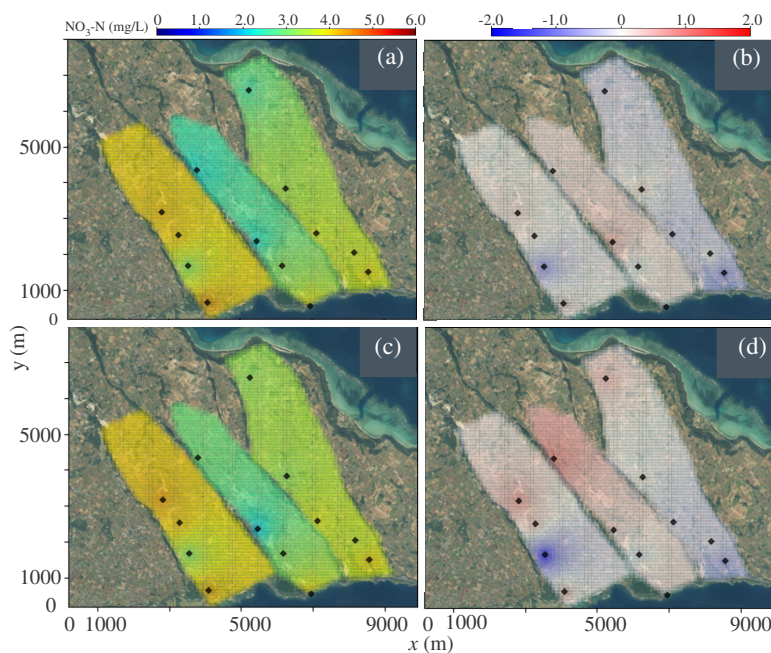


図6：(a) 68日前までのデータに基づく予測、(b) (a)と観測値の差分、(c) 186日前までのデータに基づく予測、(d) (c)と観測値の差分