

干拓地における排水管理への罰則付き回帰モデルの適用 Penalized Regression Model Application for Management of Drainage in Reclaimed Land

○長野 峻介*, 林 成也*, 藤原 洋一*, 高瀬 恵次*, 一恩 英二*

○CHONO Shunsuke*, HAYASHI Seiya*, FUJIHARA Yoichi*, TAKASE Keiji*, and ICHION Eiji*

1. はじめに

石川県小松市と加賀市の今江潟、柴山潟、木場潟は“加賀三湖”と呼ばれ、国営加賀三湖干拓事業（昭和27年～昭和44年）により今江潟の全面と柴山潟の2/3の面積は干拓され圃場整備がなされてきた。現在、加賀三湖干拓地および周辺地区では、集中管理施設によって用排水路の水位状況などを把握し一元的に灌漑排水管理を実施している。ただし、管轄する地区全体の状況を把握し、灌漑排水施設を的確に操作管理するには、管理者の経験を必用としている部分も大きい。そこで本研究では、加賀三湖干拓地の灌漑期における排水管理を解析するため、灌漑期における排水ポンプの操作とゲートの操作、用排水路の水位データの関係性を考察し、近年様々な分野で用いられてきた罰則付き回帰モデルを適用することによって、排水機場の排水量を予測し、管理者の経験に基づく排水機の操作管理のモデル化を行った。

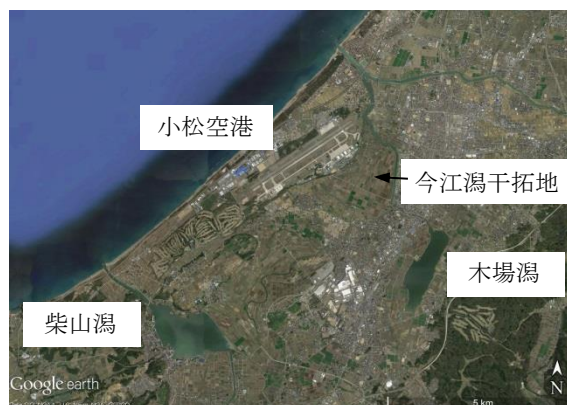


図1 加賀三湖干拓地
Fig. 1 Kagasanko Reclaimed Land

2. 罰則付き回帰モデルの適用

灌漑排水管理における排水機場の排水量の予測に、機械学習の手法の中で教師あり学習に分類される Lasso 回帰, Ridge 回帰, Elastic net といった罰則付き回帰モデルを用いて解析を行った。罰則付き回帰では、最小二乗法に罰則項（または正則化項）を追加することにより最適化を行う。最小二乗法による回帰では訓練データに最も適合した結果を返すが、過学習を引き起こす可能性がある。罰則項を加えることにより、過剰に訓練データに適合することを抑えて汎化性能を高めている。罰則項は L_1 罰則項と L_2 罰則項があり、Lasso 回帰では L_1 罰則項である係数の絶対値の和, Ridge 回帰では L_2 罰則項である係数の二乗和を用いて回帰係数を求め, Elastic net は L_1 罰則項と L_2 罰則項の両方を用いたものである。式 (1) において $q = 1$ とした場合は Lasso 回帰, $q = 2$ とした場合は Ridge 回帰となる。

$$\vec{b}^* = \operatorname{argmin}_{\vec{b}} (\mathbf{y} - \mathbf{X}\vec{b})^2 + \lambda \sum_{j=1}^M |W_j|^q \quad (1)$$

なお、加賀三湖干拓地における各地点の水位データ、ゲート、ポンプなどの操作データは2014年と2015年の灌漑期間4月1日から9月30日のデータを、回帰モデルの訓練データとして使用した。また、回帰モデルの解析にはプログラム言語 Python を用いた。Python のライブラリとして、数値計算に関しては NumPy, データフレームの扱いには Pandas, データの可視化には Matplotlib, 機械学習には Scikit-learn を使用した。

*石川県立大学生物資源環境学部 Ishikawa Prefectural Univ., Fac. of Bioresources and Environmental Sciences

キーワード：排水管理, 計画手法, 排水施設

3. 解析結果

解析結果から各モデルの RMSE 値（平均二乗平方根誤差）と R^2 （決定係数）を表 1 に示した。表 1 の training は訓練データ, 10 - foldCV は交差検定（10 分割）を示し、交差検定での RMSE 値が小さいほど過学習が抑えられ正則化していることになる。排水機場の排水量の観測値と ElasticNet モデルによる予測値との比較を図 2 に示し、2014 年の観測値と予測値の時系列変化とそれらの差を図 3 に示した。

表 1 において、最小二乗（OLS）モデルと比較して、Elastic net モデルは、訓練データに対する RMSE（training）は大きいですが、交差検定での RMSE（10-foldCV）が小さくなっている。Elastic net モデルでは正則化により未知のデータに対処する予測能力が高く、実際の操作管理に適用する際において ElasticNet のような罰則項を取り入れた手法が有用であることがわかる。図 2 において、排水量の観測値が縦に重なっている箇所がある。これは、排水量を予測した排水機場には、排水機が 4 台設置されており、どの排水機を稼働させたかによってこの傾向が表れている。図 3 において、観測値と予測値の差は 4 月や 9 月が大きくなっている。この期間は、降水量や灌漑量が少なく排水機の稼働頻度が少ないために訓練データも少なく、灌漑期の中でも予測精度が低下している。

4. おわりに

本研究では加賀三湖干拓地の灌漑期における排水管理に罰則付き回帰モデルを適用し排水機場の排水量を予測した。最小二乗モデルと比較して、交差検定での RMSE が小さくなった ElasticNet のように罰則付き回帰モデルの有用性が確かめられた。今後の課題として、より回帰精度を上げるため訓練データを増やすこととそのデータ項目の選定、パターン解析の検討などが挙げられる。

表 1 各モデルの R^2 と RMSE
Table 1 R^2 and RMSE

Model	R^2 (training)	RMSE (training)	RMSE (10-foldCV)
Ridge	0.8589	1270.5	1359.8
Lasso	0.8528	1271.4	1343.2
ElasticNet	0.8583	1271.4	1343.2
OLS	0.8593	1268.1	1367.7

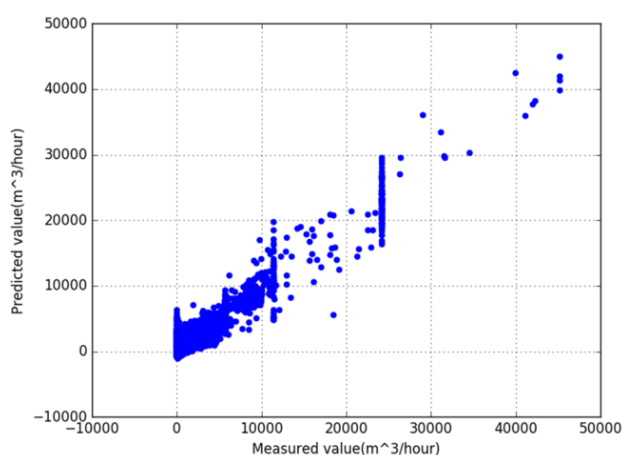


図 2 排水量の観測値と予測値

Fig. 2 Measured Value and Predicted Value of Drainage

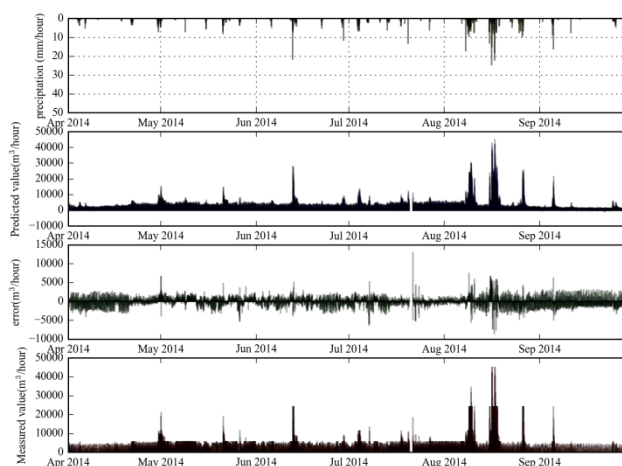


図 3 2014 年の観測値と Elastic Net による予測値

Fig. 3 Measured Value and Predicted Value of Elastic Net in 2014