

機械学習および回帰分析併用による被圧帯水層の水源揚水井水位の将来予測

関西大学環境都市工学部 学生員 ○大西望央
 関西大学環境都市工学部 正会員 尹 禮分
 関西大学環境都市工学部 フェロー 楠見晴重

1. はじめに

京都盆地は良質な地下水が豊富であり、その賦存量は琵琶湖に匹敵すると推定されている。本研究の研究対象地域である京都盆地南部の南山城地域は桂川、宇治川、木津川の三川合流地域の南部に位置しており、良質な地下水が豊富に蓄えられている。この地域では、古くから上水道水としてだけでなく工業、農業用水としても地下水を利用していることから、地下水障害などを防ぐために地下水位の将来予測が必要となってくる。

本研究は機械学習を用いて、これらの地域の水源揚水井の運転水位の将来予測を行うことを検討した。

2. 研究対象地域

京都盆地の南部に位置する城陽市は地下水依存率が高く、上水道水源の85%を地下水に依存している。城陽市では13本の上水道水源揚水井を有している。

本研究では水源揚水井の運転水位と揚水量データを用いて機械学習を行い、地下水位予測モデルを作成、地下水位の将来予測を行うことを目的とする。城陽市の第1浄水場の2本、第2浄水場の4本、第3浄水場の7本、計13本の水源揚水井で予測を行った。

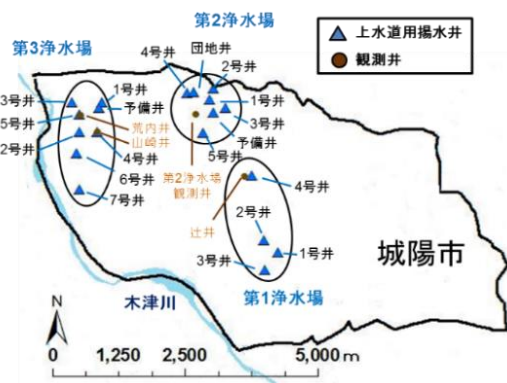


図.1 城陽市の揚水井と観測井の位置

3. SVR (Support Vector Regression)

SVRはSVMを回帰分析に適用したアルゴリズムであり、誤差の不感帯(ϵ チューブ)を設定することで、ノイズの影響を受けにくくするため、地下水位などの計測誤差の多いものの予測モデル構築に適した機械学習アルゴリズムである。また、カーネル関数を用いることで、非線形回帰分析にも応用が可能である。式(1)はSVRの回帰式である。カーネル関数はRBFカーネルを用いた。

$$f(x^{(j)}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x^{(i)}, x^{(j)}) + b \quad \dots (1)$$

x : 実測値 α : ラグランジュ関数

b : パラメータ

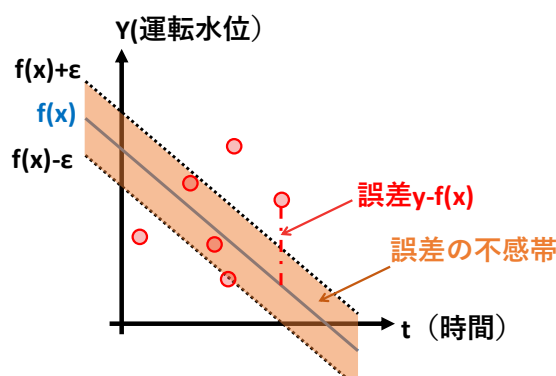


図.2 SVRの概要図

4. 線形回帰・SVR 併用モデルの作成

線形回帰分析には最小二乗法を用いた。これは各データと直線までの距離(誤差)の2乗和が最小になるような直線を求める方法である。本研究では、図.3に示すように、まず線形回帰モデルに観測値を入力する。出力された予測値と実測値との誤差をSVRモデルによって求め、線形回帰モデルによる予測値を補正、最終的な予測値として出力する。図.4のように特徴量には揚水井運転水位と揚水量午前0時、午後0時の12時間ごとのデータ7日分を用いた。

キーワード 地下水、水源揚水井、京都盆地、機械学習

連絡先 〒 564-8680 大阪府吹田市山手町3丁目3番35号 関西大学 TEL 06-6368-0837



図.3 線形回帰・SVR 併用モデルの概要図

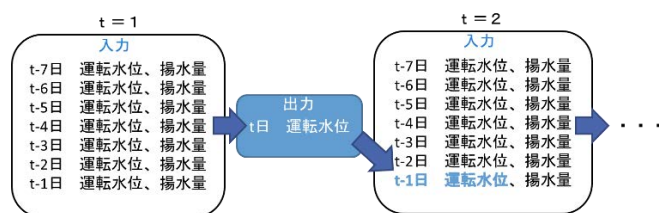


図.4 データの予測方法

5. 水位予測結果

図.5は解析対象とした第3浄水場1号井の土質柱状図である。深度300m、適正揚水量は4000m³/dayである。図.6は第3浄水場1号井の運転水位と揚水量の全実測データを表している。図.7は実測水位と線形回帰・SVR併用モデルによって予測した運転水位を表している。1999年3月31日～2022年1月1日までの7489日間を学習期間、2022年1月1日～2022年1月30日までの30日間を検証期間とした。

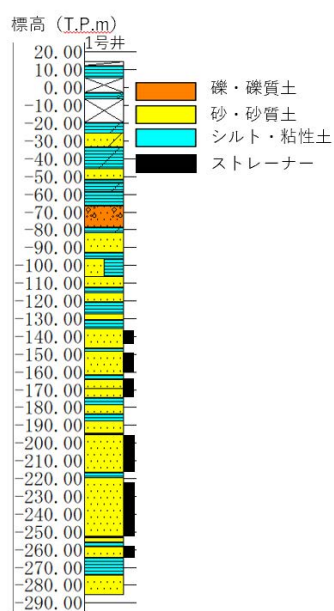


図.5 第3浄水場1号井の土質柱状図

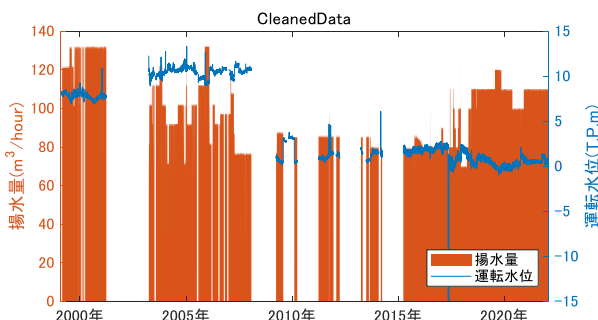


図.6 第3浄水場1号井の全実測データ

揚水量を用いず、運転水位のみを用いた予測モデルとの精度比較を行った。精度比較には式(2)～(4)示すRMSEとMAE、MASEの3種類を用いた。精度の比較結果を図.8に表す。地下水位のみを用いて予測を行うモデルよりも、地下水位と揚水量の両方を用いて予測を行うモデルの方がRMSE、MAE、MASEのすべてに置いて値が小さくなった。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad \dots (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \max(|\hat{y}_i - y_i| - \sigma, 0) \quad \dots (3)$$

$$MASE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \right| \quad \dots (4)$$

$\hat{y}_i$: 予測値 y_i : 観測値

n : 予測数 σ : 標準偏差

T : サンプルデータ数

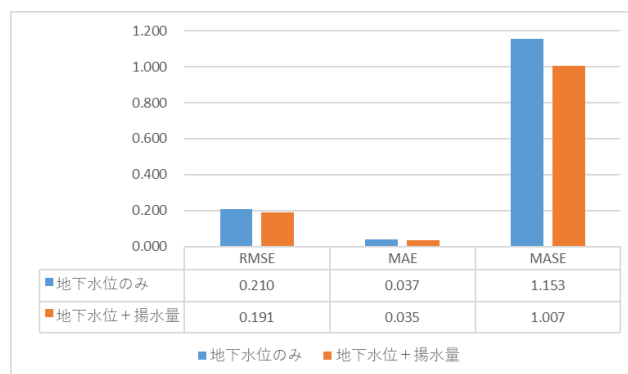


図.8 第3浄水場1号井予測結果の精度

6. まとめ

水源揚水井の地下水位将来予測において、地下水位のみよりも地下水位と揚水量、の2種類の特徴量を用いることでより精度の高いモデルの構築が可能であることが認められた。今後は、過学習を防ぐためにデータクリーニングを行いモデル改善し、水源揚水井の維持管理のために適切な揚水量の推定などを行う予定である。

参考文献

竹内一郎、他：機械学習プロフェッショナルシリーズ サポートベクトルマシン、講談社、p42-61、2015。