

データ駆動型アプローチによる表面塩化物イオン濃度に対する回帰モデルの構築

香川大学 賛助会員 岡崎百合子 正会員 岡崎慎一郎 港湾空港技術研究所 正会員 山路徹

1. 研究の背景と目的

港湾コンクリート構造物（以下、港湾 RC 構造物）における最も考慮すべき劣化要因は塩害となることから、港湾 RC 構造物の適切な維持管理には表面塩化物イオン濃度（以下、 C_0 ）を考慮した耐久性設計が肝要となる。港湾基準では、日本各地の港湾構造物での実測データに対し、朔望平均満潮位（以下、H.W.L.）からの距離をパラメータとして線形回帰した式（図-1）が用いられ、 C_0 の傾向を概ね捉えている一方、未だデータの全てに対して十分な回帰性能を有するものではなく、改善が求められている。本研究では、機械学習を援用し、港湾 RC 構造物の C_0 回帰モデルに対し、H.W.L.からの距離に加え、モデル性能向上に寄与する環境因子を探索した。

2. 学習に用いたデータ

今回の検討に用いた C_0 に関するデータは、港湾基準式の構築に用いられたものと同一であり、日本各地における 18 の港湾施設から収集されたものである（図-2）¹⁾。

C_0 に関するデータの頻度について、H.W.L.からの距離により 4 区分して図-3 に示す。166 のデータのうち、95%以上が 20kg/m^3 以下であり、それより高い C_0 のデータは稀である。H.W.L.からの距離別にデータを概観すると、白色で示された H.W.L.からの距離が長い領域（遠 H.W.L.領域）では C_0 の値は 5kg/m^3 以下に集中するのに対し、濃色で示された H.W.L.からの距離が短い領域（近 H.W.L.領域）では C_0 の値が $0\text{kg/m}^3 \sim 30\text{kg/m}^3$ と大きくばらついている。

このデータに対し、気温、降雨量、日照時間、有義波高といった環境データを加え、機械学習により C_0 回帰モデルを構築する。有義波高は国土交通省港湾局 NOWPHAS²⁾から、それ以外は気象庁データベースから、各施設に近接する観測所のデータを取得した。

3. 特徴量選択

機械学習モデルにおける説明変数は特徴量と称され、性能の高いモデル構築には、冗長的でない適切な特徴量選択が重要となる。このため、表-1 に示した環境データの中から、特徴量間の相関性を考慮しながらモデルに説明変数として加えるべき特徴量を選択できる逐次前方選択を実施した。

逐次前方選択は、最初のループとして、候補となる変数に対し、それぞれ 1 つずつを特徴量としたモデルを構築し、その性能比較により、最も性能の良いモデルで特徴量とした変数を採用する。次のループでは、前ループで採用した変数に、それ以外の変数を 1 つずつ追加したモデルを構築し、それらの性能比較により特徴量とする変数を決定し、特徴量を追加しても性能が向上しなくなるまで同様のループを繰り返す。モデルの性能比較では、5 分割交差検証を行い、平均平方二乗誤差（RMSE）を指標とした。なお、機械学習アルゴリズムには、本研究で対象とするデータ数が 166 と少なく、データの不均衡性が高いため、ロバスト性の高い Support Vector Regression を採用した。

図-4 に逐次前方選択を行った際の各ループにおける最小 RMSE の推移を示す。最初のループでは「H.W.L.か

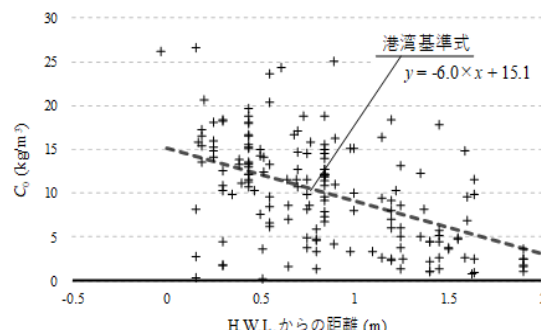


図-1 港湾基準式

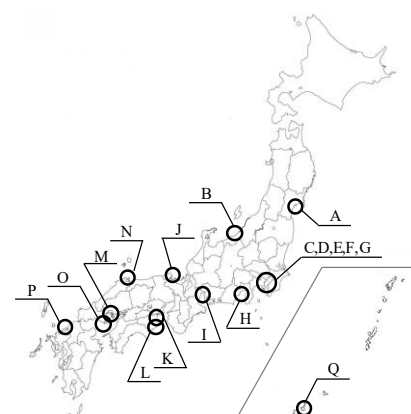


図-2 データを収集した港湾施設位置

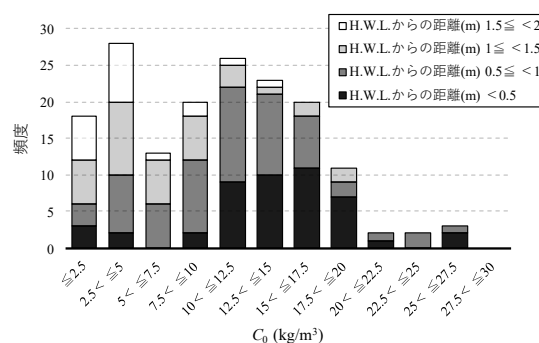


図-3 データの頻度

らの距離」が最も小さい RMSE となる 5.20 を示し、港湾基準式に用いられた本変数の有用性を再確認した。2 ループ目の「有義波高」追加時に、RMSE は大幅に低下し、続いて、「風速」、「降雨量最大値」を追加するごとに RMSE は若干低下し、それ以降は特徴量の追加に伴って RMSE が増加した。この結果より、今回対象としたデータから C_0 の回帰モデルを構築する上で説明性の高い環境因子として、上述の 3 変数を選択した。

4. 環境因子がモデル性能に与える影響

H.W.L.からの距離に加え、3 つの環境因子を特徴量とした機械学習モデルを構築し、既往の港湾基準式と比較することで、環境因子がモデル性能に与える影響を調査した。図-5 に、モデル回帰値と実測値の比較、および回帰誤差のヒストグラムを、港湾基準式と本研究での構築モデルについてそれぞれ示す。回帰値と実測値の比較より、港湾基準式では実測値の増加に対する回帰値の増加不足が示されているが、本モデルではその傾向が緩和され、決定係数 R^2 も港湾基準式の 0.28 から 0.54 に改善し、実測値と回帰値の相関性が高くなった。 C_0 の実測が 5 kg/m^3 以下の低 C_0 領域で特に大きな改善が確認される一方、 20 kg/m^3 以上の高 C_0 領域では実測と回帰値の乖離が大きく残存する。この理由として、図-3 に示したように高 C_0 領域のデータ数の少なさに起因する学習不足、また、表-1 に示す以外の特徴量が必要である可能性等が挙げられるが、今後の課題としたい。また、回帰誤差のヒストグラムより、環境因子追加により特に近 H.W.L.領域において改善していることが分かり、本モデルで採用した環境因子は、近 H.W.L.領域への影響が大きいことが示唆された。

表-1 特徴量候補とした環境変数

H.W.L.からの距離	気温	平均
有義波高		最大
年降雨量	平均	風速
	最大	日照時間
	最小	平均
	回数	最大
		最小

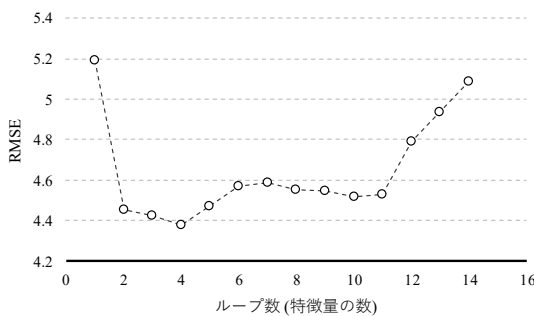


図-4 逐次前方選択結果

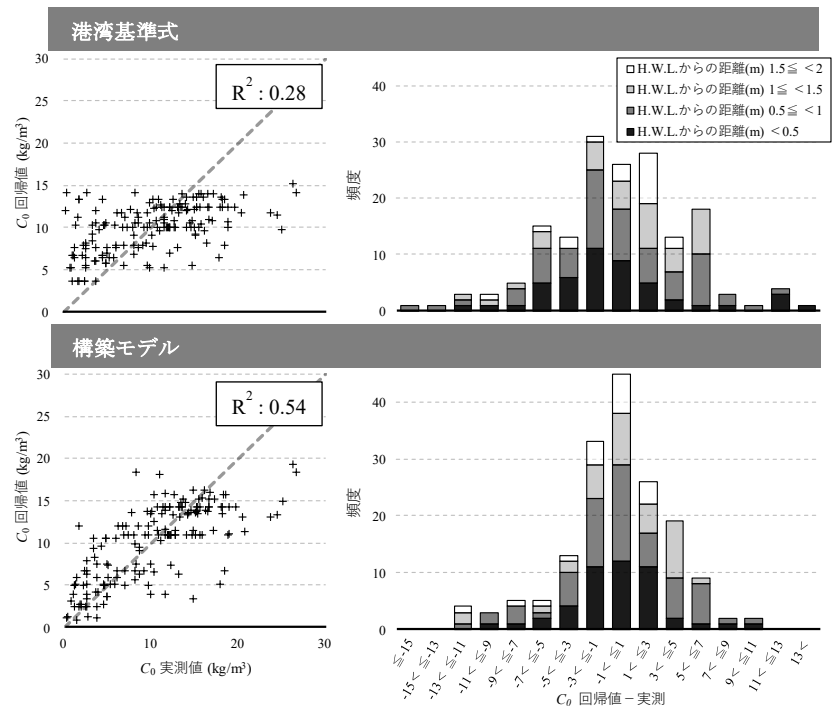


図-5 回帰値と実測値の比較および回帰誤差ヒストグラム

最後に、環境因子が、港湾基準式で示されているような H.W.L.からの距離と C_0 に与える影響を把握するため、3 つの環境因子のうち、最も影響の大きい「有義波高」に焦点をあて、パラメトリックスタディを行った結果を図-6 に示す。有義波高 1.5m においては港湾基準式に類似し、有義波高が小さくなるにつれ、 C_0 は低くなり、その影響度は、遠 H.W.L.領域では小さく、近 H.W.L.領域では大きくなる。これは、海面付近の近 H.W.L.領域では、H.W.L.からの距離が同じであっても、有義波高の大きさにより施設表面に付着する塩化物イオン量が大きく異なるためと解釈でき、本モデルではそうした事象を適切に学習したことが示されている。

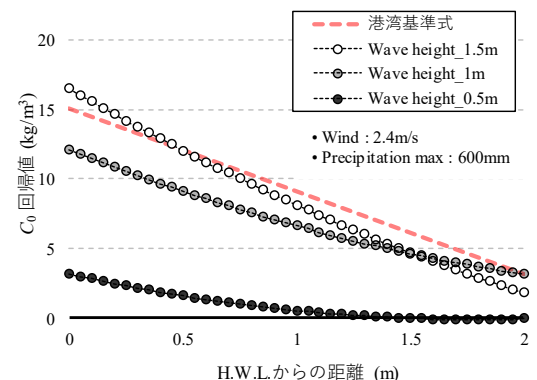


図-6 有義波高をパラメータとした C_0 回帰値

参考文献：1)山路ら：土木学会論文集 E, 64(2),335-347, 2) <https://www.mlit.go.jp/kowan/nowphas/>