衛星影像用於水庫水質分析之研究

Research on the Use of Satellite Imagery for Reservoir Water Quality Analysis

逢甲大學營建及防災研究中心

組長	資料工程師	教授	助理教授
李承家	張智雄	雷祖強	何智超
Cheng-Jia Li	Chih-Hsiung Chang	Tsu-Chiang Lei	Chih-Chao Ho

經濟部水利署水利規劃分署

分署長	科長	正工程司	助理研究員
張廣智	葛武松	趙永楠	許家銓
Kuang-Chih Chang	Wu-Sung Ke	Yeong-Nan Chao	Chia-Chuan Hsu

摘　要

衛星影像具有大範圍與週期性之特性，常用於各地物辨識使用。臺灣所擁有之水資源相當有限，主要透過水庫做為水資源保存，其中水質監測是維護水域生態平衡和確保飲用水安全的關鍵活動，為確認水質變化狀況，過往環保署每月針對重要水庫依照區域進行點抽樣監測，評估各水庫水質狀態，但抽樣結果僅針對單點進行描述，無法以整體水庫範圍進行說明。故本研究嘗試結合衛星影像與環保署水質監測資料，採用深度學習模型依參數取得關聯，進行整體水庫範圍推估。以臺灣區域 13 座水庫進做為研究範圍，影像來源採用 MODIS 衛星影像，配合影像來源以 1 公里之網格依照水庫集水範圍進行分析單元建立，演算資料結合環境部水庫水質監測資訊、水庫監測資料等內容進行分析，針對水質優養化類別(葉綠素 a、透明度、總磷)，採用深度學習模式 LightGBM（Light Gradient Boosting Machine）進行模式訓練，由於各水庫範圍與狀態皆不同，本研究依照葉綠素 a 數值標準差區分為三群，以個別水庫 8 成訓練子集，2 成測試子集，於訓練子集加入該水庫最大最小值，隨機抽樣 50 次建模分析並進行模型參數篩選，以最佳成果建立模型。整體分析結果 Train R^2 分別為 0.81、0.71、0.46，Test R^2 為 0.66、0.5、0.17，整體以分群三成果最差，針對資料進行檢示，由於分群三水庫僅包含鳳山水庫，且長期處於優養化，葉綠素濃度約為其他水庫十倍，因此單獨進行模式訓練，但可用數據僅 71 筆，故其成效較差。綜合上述成果，深度學習水質推估測試約有 0.5 至 0.7 之精度，並可透過影像資訊推估水庫蓄水範圍內之各網格水質資訊，提供面化水質分析結果供水庫管理單位維管時之參考。

關鍵詞：影像分類，水質分析，LightGBM

Abstract

Satellite imagery, characterized by its extensive coverage and periodic observations, is frequently utilized for object recognition across various regions. Reservoir water quality monitoring is pivotal for maintaining aquatic ecological balance and ensuring the safety of drinking water. When applying satellite imagery for reservoir water quality analysis, the quality of the analysis is constrained by factors such as current climatic conditions, frequency of water quality monitoring, reservoir conditions, and seasonal variations. Deep learning models can autonomously derive correlations from various parameters to achieve more effective analytical outcomes.

Satellite imagery, characterized by its extensive coverage and periodic observations, is frequently utilized for object recognition across various regions. Taiwan, with its limited water resources, primarily relies on reservoirs for water storage. Monitoring water quality is a crucial activity for maintaining ecological balance in aquatic environments and ensuring the safety of drinking water. To monitor changes in water quality, the Environmental Protection Administration (EPA) has traditionally conducted monthly point-sampling surveys in key reservoirs to assess their water quality. However, the sampling results provide information limited to specific points and cannot represent the entire reservoir.

This study attempts to combine satellite imagery with EPA water quality monitoring data, utilizing deep learning models to establish correlations between parameters and estimate water quality across the entire reservoir. The study focuses on 13 reservoirs in Taiwan, using MODIS satellite imagery as the data source. The imagery is analyzed on a 1-kilometer grid, established according to the catchment areas of the reservoirs. The analysis incorporates environmental data, including water quality monitoring information from the Ministry of Environment and other reservoir monitoring data.

The study specifically addresses eutrophication-related water quality indicators such as chlorophyll-a, transparency, and total phosphorus. The Light Gradient Boosting Machine (LightGBM), a deep learning model, is employed for model training. Due to the varying conditions and ranges of the reservoirs, the study categorizes the reservoirs into three groups based on the standard deviation of chlorophyll-a values. For each reservoir, 80% of the data is used as a training subset, and 20% as a test subset. The training subset is supplemented with the reservoir's maximum and minimum values, and modeling is conducted by randomly sampling 50 times to select model parameters and achieve optimal results.

The overall analysis yields Train R^2 values of 0.81, 0.71, and 0.46, and Test R^2 values of 0.66, 0.5, and 0.17. The third group produced the poorest results, which included only Fengshan Reservoir, known for its long-term eutrophication with chlorophyll-a concentrations approximately ten times higher than other reservoirs. Due to the limited available data (only 71 records), the model's effectiveness for this group was relatively low. Overall, the deep learning-based water quality estimation achieved an accuracy ranging from 0.5 to 0.7. This approach allows for the estimation of water quality across the reservoir's water storage area at a grid level, providing spatially comprehensive water quality analysis results that can be used by reservoir management units for maintenance and management purposes.

Keywords: Image Classification, Water Quality Analysis, LightGBM