A multi-scale machine learning framework for interpretable groundwater level prediction

EGU26-2803
Session HS8.2.2

Sheng-Wei Wang[*], Wen-Chi Chen

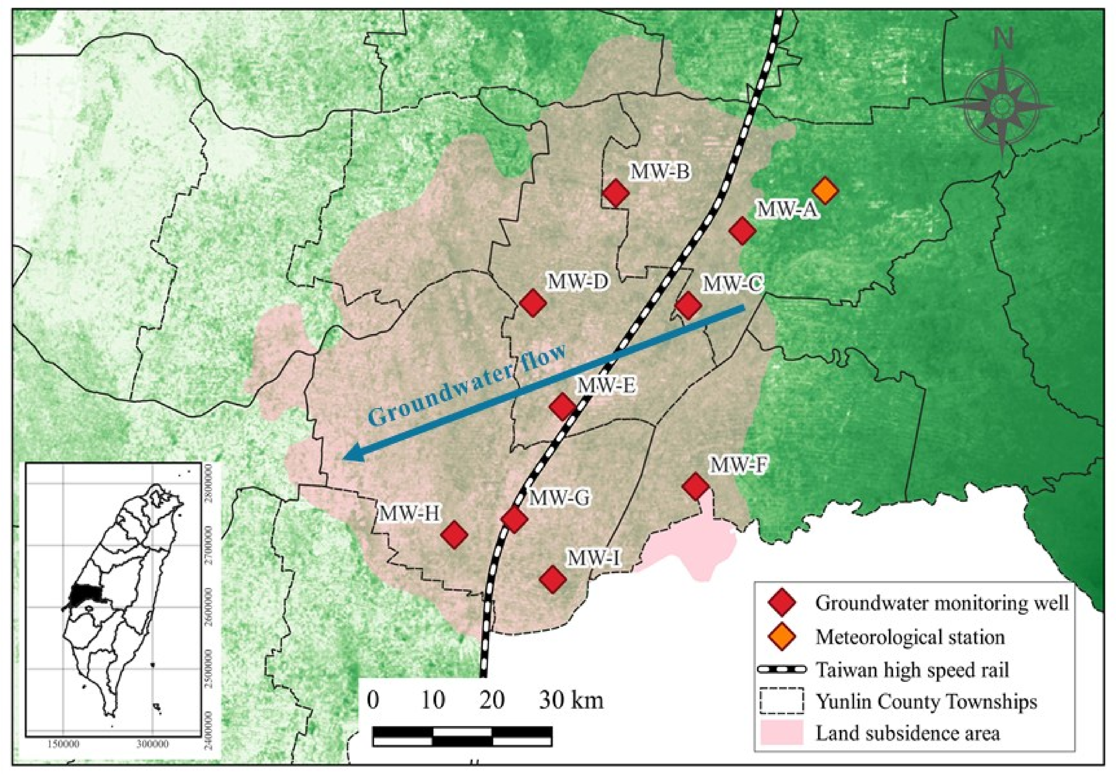
[*] Department of Civil Engineering, National Central University, Taiwan; Email: wangsw@mail.ncu.edu.tw

Introduction

- Groundwater level variations in agricultural aquifers are jointly controlled by short-term rainfall recharge and long-term pumping stress; however, these drivers are often recorded at different temporal scales, with groundwater level and rainfall available at daily resolution while pumping information is typically available only at monthly scale. This mismatch limits the ability of conventional machine learning models to provide both accurate predictions and meaningful hydrological interpretation.
- To address this issue, this study proposes a multi-scale machine learning framework that explicitly separates long-term groundwater trends from short-term event responses. By decomposing groundwater dynamics into monthly and daily components, the framework aims to improve predictive performance while preserving interpretability of recharge and pumping processes.
- The objectives of this study are to (1) evaluate whether pumping electricity can effectively represent long-term groundwater abstraction, (2) assess the performance of a two-stage multi-scale model for groundwater prediction, and (3) analyze groundwater response behavior using event-based diagnostics and interpretable machine learning techniques.

Materials and Methods

- The study area consists of an agricultural groundwater system with nine monitoring wells (MW-A to MW-I). Daily groundwater levels and rainfall data were collected, while monthly electricity consumption from surrounding pumping wells was used as a proxy for pumping intensity. Pumping electricity was further categorized by different usage types to reflect heterogeneous abstraction patterns across wells.
- A two-stage modeling framework was developed. In Stage A, a monthly-scale model was constructed using rainfall, pumping electricity, and autoregressive groundwater terms to represent long-term groundwater trends. The monthly predictions were then upsampled to daily resolution to provide a baseline groundwater state.
- In Stage B, daily residuals were defined as the difference between observed groundwater levels and the Stage A baseline. A daily-scale model was then built using residual lag (7 days), cumulative rainfall (7-day and 14-day), and antecedent dry-day indices to capture short-term recharge responses. The model was implemented using XGBoost with pseudo-Huber loss and a small nested time-series hyperparameter tuning scheme.



Distribution of the nine groundwater monitoring wells (MW-A to MW-I) and the meteorological station.

- Model performance was evaluated using five-fold time-series cross-validation and out-of-fold predictions. Interpretability and process understanding were assessed using SHAP analysis, event-aligned composite response analysis, and lag-to-peak diagnostics, with comparisons among observed data, no-pumping simulations, and pumping-included simulations.

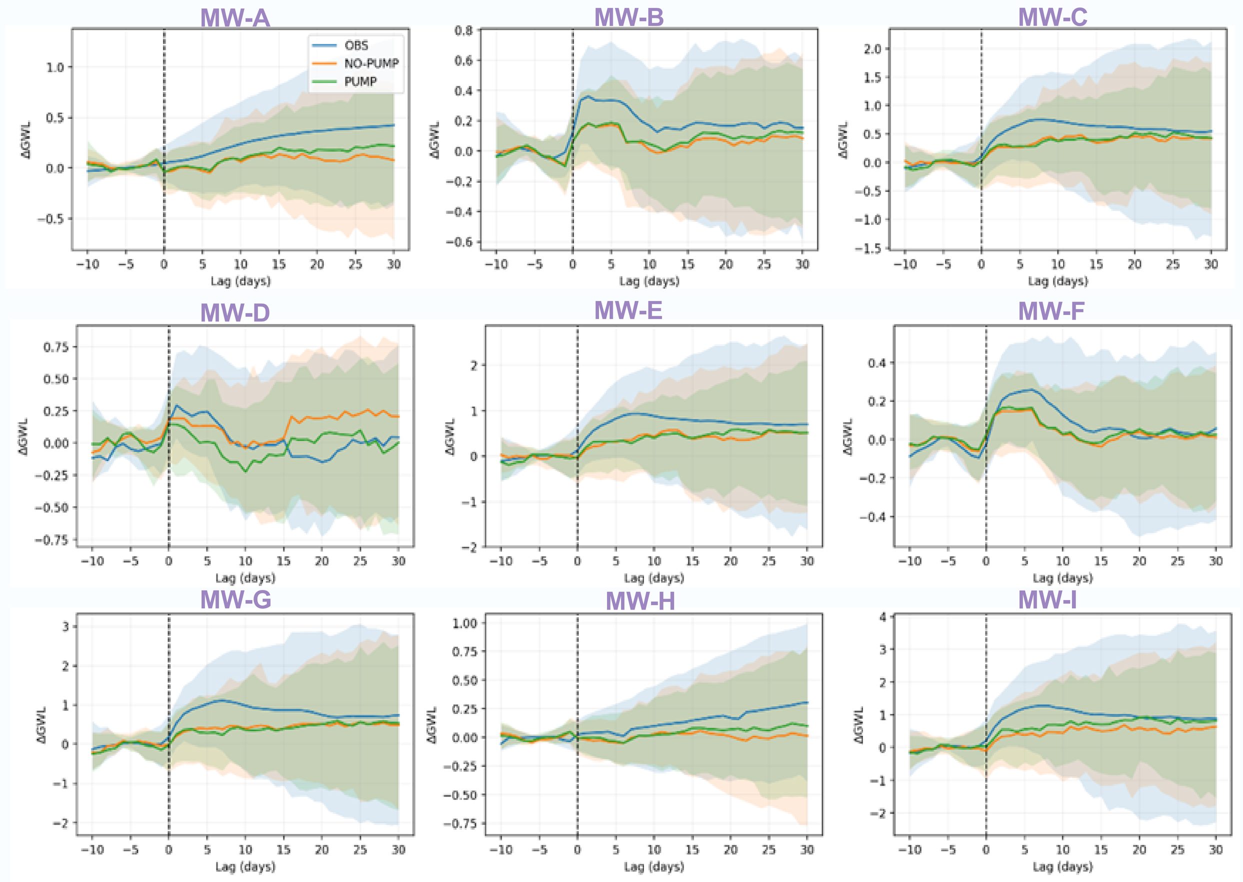
Results

- At the daily scale, incorporating pumping information improved prediction performance in 8 out of 9 wells, indicating that pumping information enhances the representation of groundwater variability primarily by improving the baseline trend.
- At the monthly scale, the pumping-included model also outperformed the no-pumping model in 8 out of 9 wells. These results demonstrate that pumping electricity is an effective proxy for long-term groundwater abstraction.

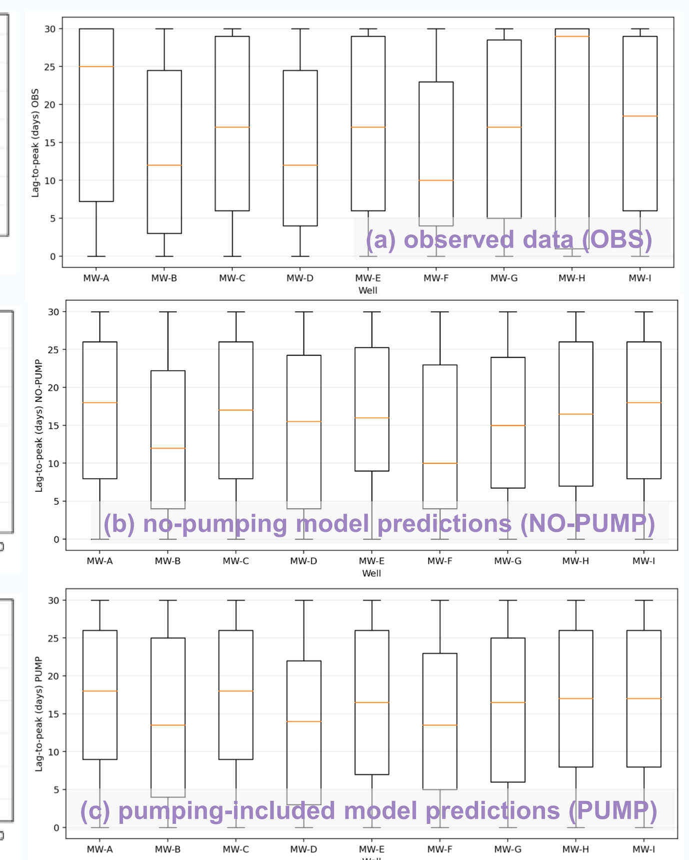
Scale	Scenario	RMSE	MAE	R^2
Monthly	No-pump	1.983	1.463	0.263
Monthly	Pump	1.639	1.189	0.420
Daily	No-pump	1.872	1.302	0.302
Daily	Pump	1.538	1.098	0.381

Discussion

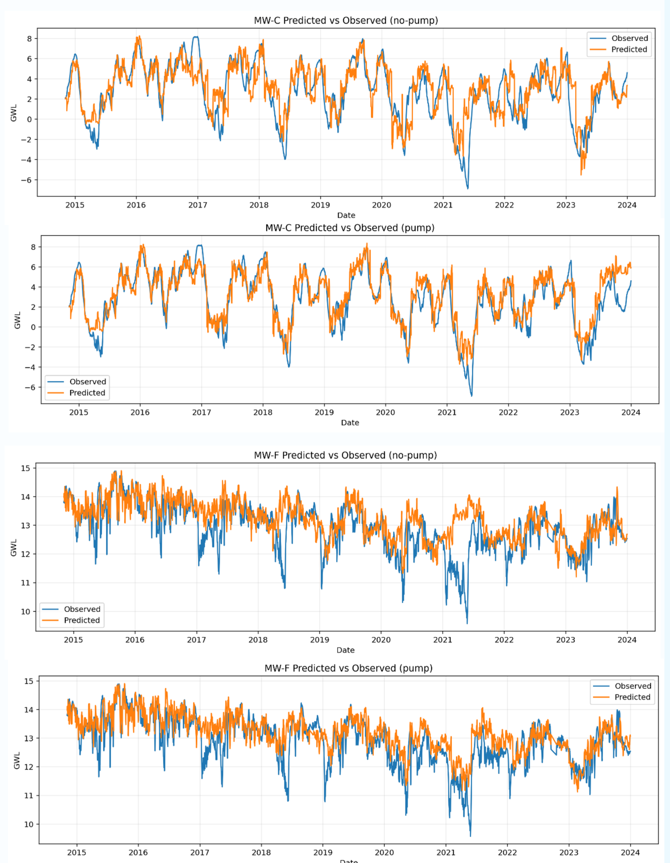
- Event-aligned composite analysis reveals substantial heterogeneity in groundwater response across wells, with varying magnitudes and timing of post-rainfall responses. While pumping improves the long-term trend representation, it does not significantly alter event-scale response timing, suggesting that short-term dynamics are primarily controlled by hydrogeological conditions.
- Lag-to-peak analysis further confirms that groundwater response times differ significantly among wells and remain relatively consistent across scenarios. This indicates that aquifer properties and hydraulic connectivity dominate event-scale behavior, whereas pumping mainly affects background groundwater levels.
- Predicted-versus-observed time series demonstrate that the multi-scale framework successfully captures both long-term trends and short-term fluctuations. The pseudo-Huber loss function provides marginal but consistent improvements in robustness, particularly in wells with more variable residual behavior.



Event-aligned groundwater responses for nine wells (MW-A–MW-I) under OBS, NO-PUMP, and PUMP scenarios. ΔGWL is normalized to pre-event conditions (day 0 = rainfall onset); lines show means and shading indicates interquartile ranges.



Lag-to-peak groundwater response distributions for all wells, defined as days from rainfall onset to peak ΔGWL.



Comparison of observed and simulated groundwater levels for MW-C and MW-F under no-pump and pump scenarios.

Conclusion

- Future work will refine pumping representation by incorporating spatial weighting, well distribution, and seasonal abstraction patterns to improve model performance and interpretability.
- Additional analysis will classify wells into pumping-dominated, recharge-dominated, and mixed-control types to better understand the drivers of groundwater variability across the study area.
- The framework will be extended to scenario-based applications including drought analysis and groundwater management, to support sustainable water resources under climate and anthropogenic change.