ECWS-9 2025 Conference
The 9th International Electronic Conference on Water Sciences
11–14 November 2025 | Online
MDPI

Enhancing Basin-Scale Hydrological Insights in Greece by Integrating Machine Learning and Satellite Gravimetry.

Eleni A. Tzanou[1] *, Anastasia I. Triantafyllou[2], Elisavet G. Mamagiannou[2], Georgios S. Vergos[2]
[1]School of Surveying and Geoinformatics Engineering, Faculty of Engineering, International Hellenic University, Greece, GR-62124
[2]Laboratory of Gravity Field Research and Applications – GravLab, Department of Geodesy and Surveying, Aristotle University of Thessaloniki, Greece, GR-54124

INTRODUCTION

This study examines the potential of satellite gravimetry to monitor basin-scale hydrological variability across Greece by downscaling coarse-resolution terrestrial water storage anomalies derived from the Gravity Recovery and Climate Experiment (GRACE) and its successor mission, GRACE Follow-On. Monthly Liquid Water Equivalent (LWE) anomalies from mascon solutions ~1° resolution (~110 km) are refined to 0.1° (~10 km) using a supervised machine learning approach. The resulting high-resolution product is assessed through cross-comparisons with independent satellite datasets. By integrating satellite gravimetry, environmental indicators, and machine learning techniques, we propose a scalable framework for enhancing the spatial resolution of terrestrial water monitoring in data-scarce regions.

METHODOLOGY

The first step of the methodology refers to the downscaling the coarse spatial analysis of GRACE mascon's solutions by utilizing a boosting regression machine learning algorithm. As predictor variables we used time series of parameters derived from the monthly reanalysis of the ERA 5 satellite.

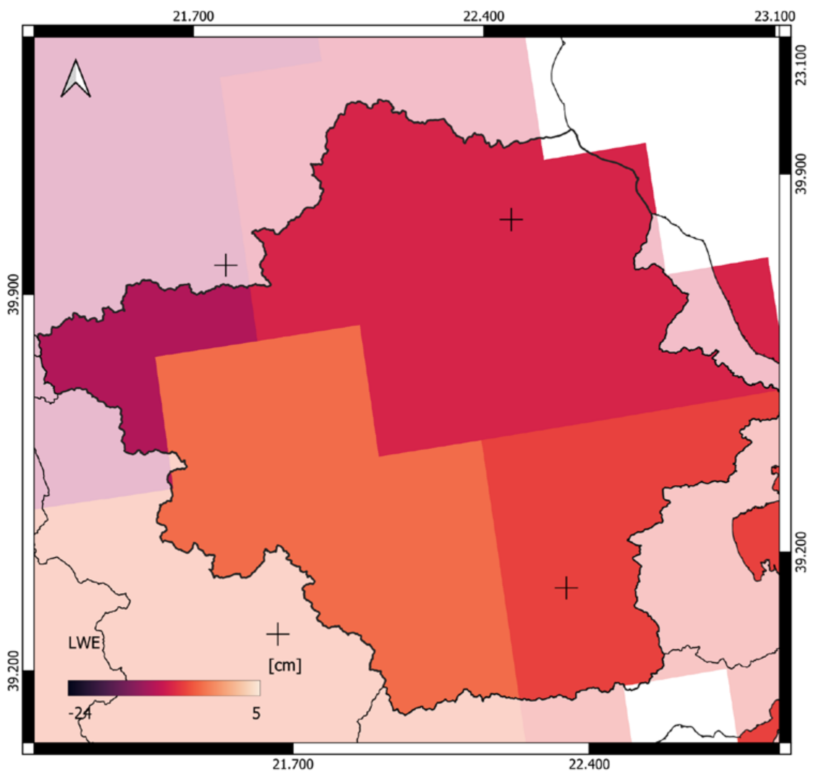
Predictor variables from ERA5 monthly reanalysis dataset	Units
Volumetric soil water layer 1(0 - 7cm)	[m^3 m^{-3}]
Volumetric soil water layer 2 (7 - 28cm)	[m^3 m^{-3}]
Volumetric soil water layer 3 (28 - 100cm)	[m^3 m^{-3}]
Volumetric soil water layer 4(100 - 298cm)	[m^3 m^{-3}]
Sub-surface runoff	[m]
Surface runoff	[m]
Evaporation	[m]
Total precipitation	[m]

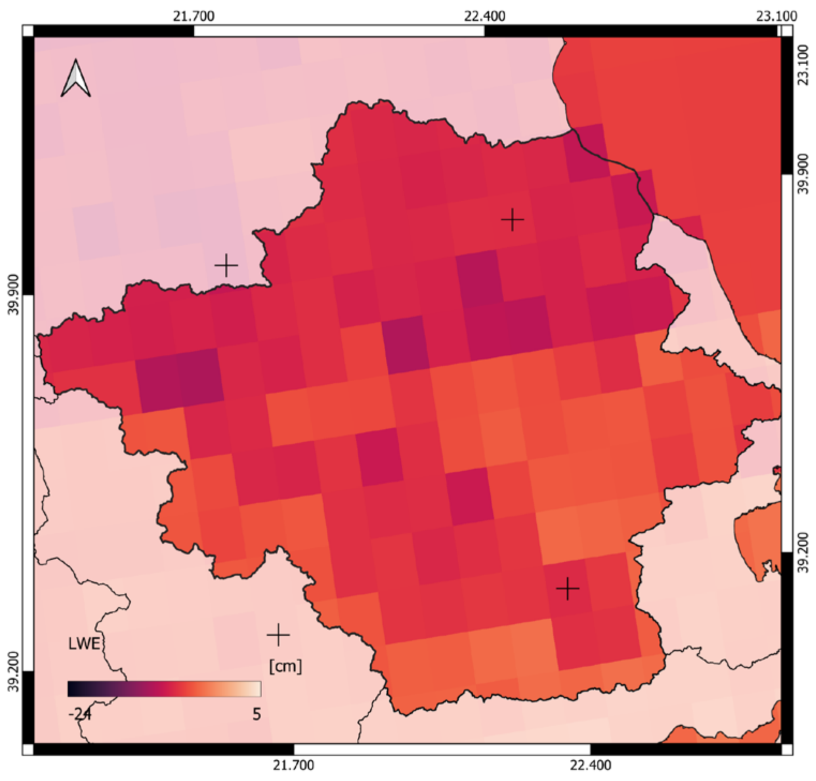
The ML technique that we used was based on a boosting function that adds up many small trees

$$y_t = f(x_t) = f_0 + \eta \sum_{m=1}^{M} T_m(x_t)$$

Where:

- f_0: a simple starting guess of LWE
- Each tree T_m is tiny and fixes leftover error from the previous trees.
- η (learning rate) keeps steps small and stable.





MAE	RMSE	R2	BIAS	CORR
2.4438	3.2448	0.8914	0.4608	0.9462

We evaluated 20 randomly chosen parameter combinations using a five-fold cross-validation setup with the negative MAE metric (i.e., minimizing MAE). The total grid contained 3^6 = 729 possible combinations, from which 20 were sampled at random for efficiency.

In the maps presented, which refers to October 2023 at Pinios Basin, the predicted LWE closely matches the real GRACE data, with a very high correlation (0.95) and $R^2 \approx 0.89$, meaning about 89% of the variance is explained; the low MAE (≈2.4 cm) and small positive bias (≈0.46 cm) suggest only minor systematic overestimation.

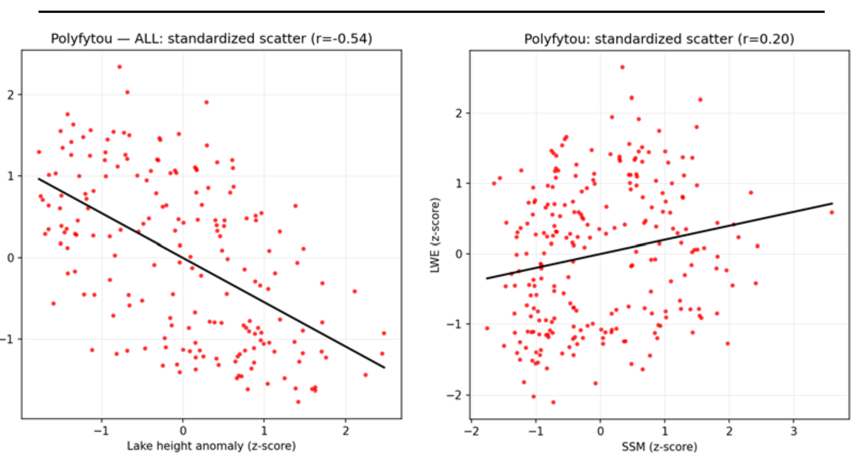
To further assess whether the predictions capture meaningful hydrological variability, we compared the results of the process of ML with data related to Lake Polifitou and Lake Kremasta, which are both important regulated reservoir. Specifically, for Lake Polifitou and Kremasta we obtained:

- Lake-level time series derived from satellite radar altimetry (Jason-1, Jason-2/OSTM, Jason-3, and Sentinel-6) from NASA GWM service.
- Soil moisture dataset derived from derived from Active Microwave Instrument - WindScat (AMI-WS) and ASCAT (Metop-A, Metop-B, Metop-C) from ESA's CDS service

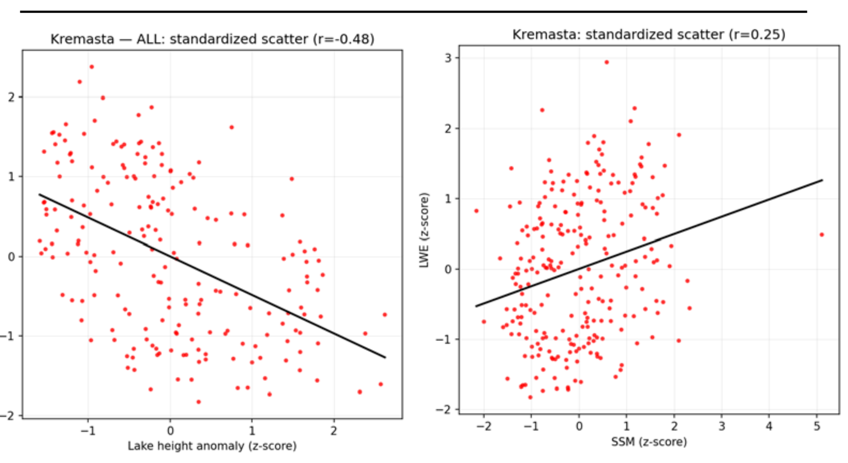
The comparison was performed through a number of statistical correlation measures (zero-lag Pearson correlation, Spearman rank correlation, Lagged correlation analysis).

RESULTS

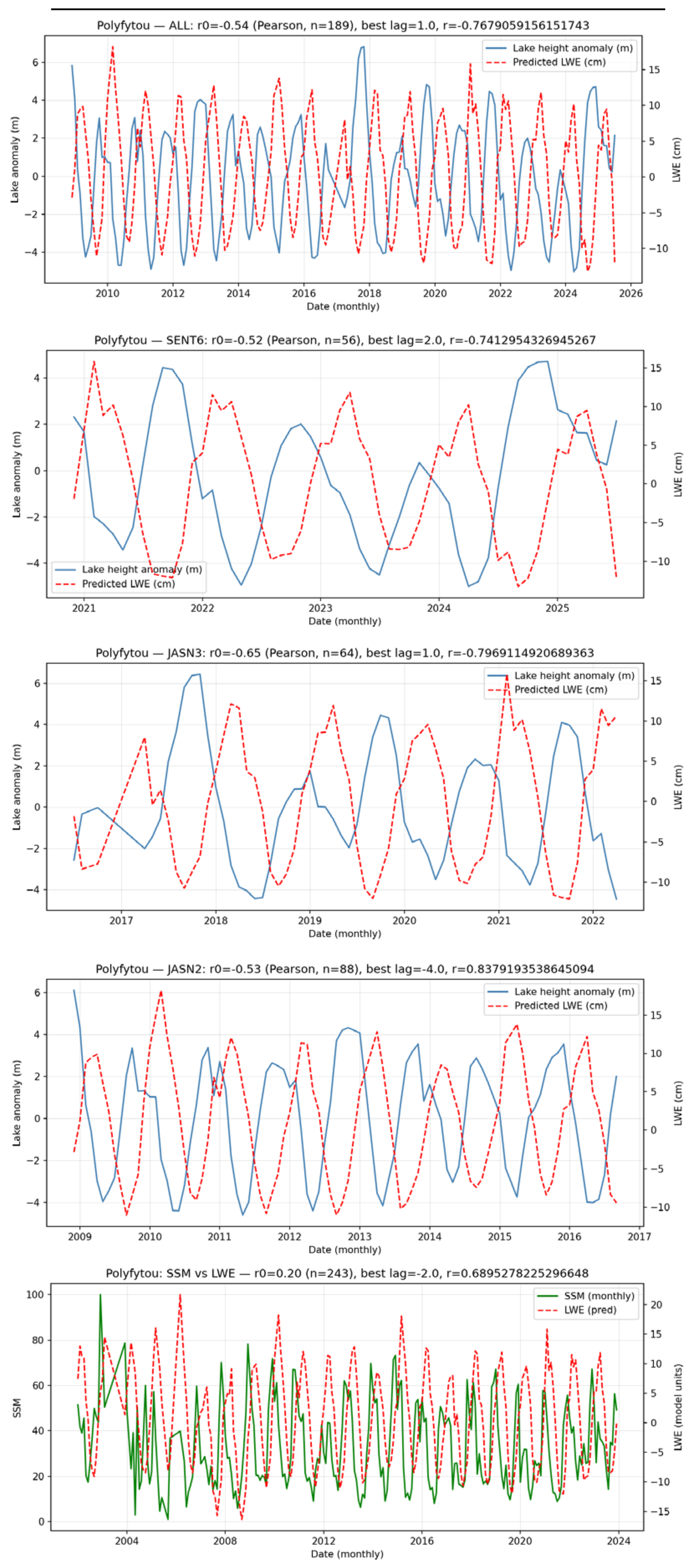
Lake Polifitou



Lake Kremasta

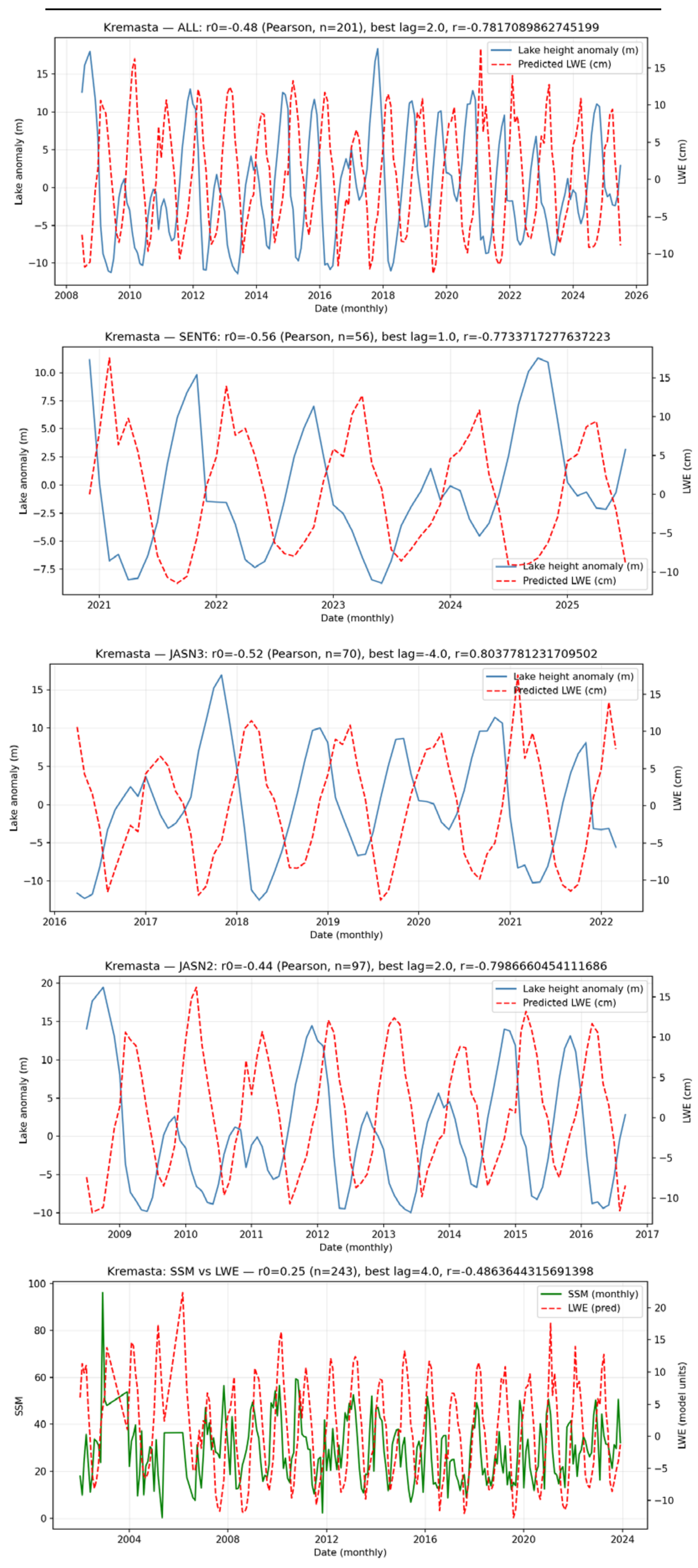


Lake Polifitou



Mission	Pearson	Spearman	Bestlag	Bestlag_r
ALL	-0.54496	-0.54216	1	-0.76791
JASN2	-0.53475	-0.53126	-4	0.83792
JASN3	-0.64719	-0.64712	1	-0.79691
SENT6	-0.51636	-0.51761	2	-0.74130

Lake Kremasta



Mission	Pearson	Spearman	Bestlag	Bestlag_r
ALL	-0.48484	-0.49598	2	-0.78171
JASN2	-0.4386	-0.46538	2	-0.79867
JASN3	-0.52411	-0.53383	-4	0.80378
SENT6	-0.56131	-0.55126	1	-0.77337

The correlation over Lake Polifitou demonstrates a good hydrological agreement across most satellite missions. For the aggregate "ALL" altimetry record, the zero-lag correlation reaches *r = 0.545* (Spearman = 0.542), while mission-specific analyses show even stronger relationships, notably for JASON-3 (r_0 = 0.65) and Sentinel-6 (r_0 = 0.52). These missions also exhibit best correlations at short positive lags (+1 to +2 months), suggesting that modeled storage anomalies slightly lead observed lake height changes, which is hydrologically logical, since GRACE senses basin-wide mass gain before that water accumulates at the reservoir surface. The soil-moisture comparison at the same point reinforces this interpretation. Although the zero-lag correlation is modest (r = 0.20), the maximum correlation increases sharply (r ≈ 0.69) when the soil-moisture series is shifted 2 months earlier (lag = −2).

The altimetry–LWE correlations for Lake Kremasta reveal a similar but slightly weaker pattern. Across all missions combined, the zero-lag Pearson correlation is 0.48, increasing to 0.56 for Sentinel-6 and 0.52 for JASON-3. Best correlations are typically found at short positive lags (1–2 months), with values up to r ≈ 0.78–0.80, suggesting that LWE variations precede the lake-level response by roughly one to two months—this is what it was observed at lake Polifitou too. For the Kremasta site, the soil-moisture relationship shows low zero-lag correlation (r = 0.25) and a negative best-lag correlation (r = −0.49 at +4 months), implying a delayed coupling between shallow soil moisture and total storage.

CONCLUSION

The downscaled GRACE LWE product demonstrates a strong ability to reproduce basin-scale hydrological variability. Comparison with altimetry over the lakes yielded correlation coefficients between 0.51 and 0.65 at zero lag, improving to 0.7–0.8 when accounting for a 1–4-month lag, indicating that are physically consistent with observed lake-level changes. The observed tendency for LWE to lead altimetry reflects realistic catchment behavior, where basin-wide water mass increases (soil and groundwater recharge) precede surface storage responses in the reservoirs. Correlations with surface soil moisture (SSM) were lower at zero lag (≈0.2) but strengthened markedly (to ≈0.5–0.7) when accounting for the faster response of surface moisture to precipitation, confirming that SSM leads LWE by 1–2 months—a hydrologically coherent sequence of rainfall infiltration and subsurface recharge. Together with the direct comparison to the official GRACE dataset ($R^2 \approx 0.89$, Corr ≈ 0.95), these results demonstrate that the ML downscaling approach can capture both rapid surface responses and slower subsurface dynamics.