Satellite-Based Estimation of Soil-geology Using Physics-Guided Machine Learning

Hrusikesha Pradhan, Kyra H. Adams, J.T. Reager, Kyongsik Yun, Benjamin D. Hamlington

Jet Propulsion Laboratory, California Institute of Technology

Jet Propulsion Laboratory
California Institute of Technology

NEURAL INFORMATION PROCESSING SYSTEMS

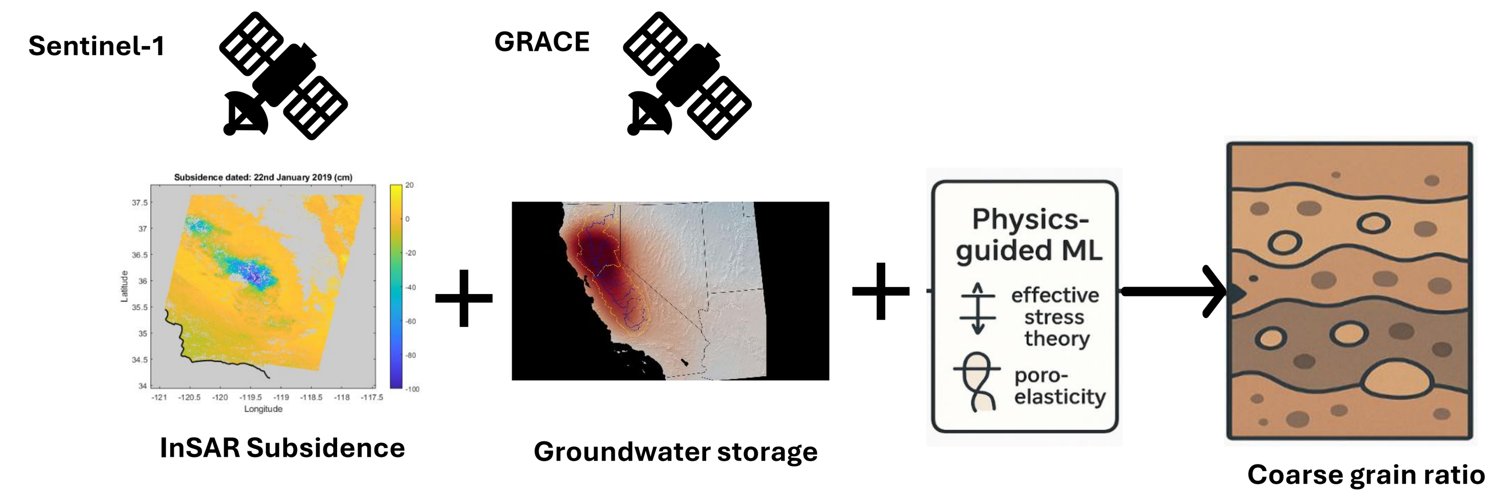
Background and Motivation

California's Central Valley supports a quarter of U.S. food production and relies heavily on groundwater, supplying ~20% of national demand. Intensive pumping has caused widespread land subsidence and loss of aquifer storage. The extent of this subsidence depends on the soil's coarse-grain ratio (CGR)—fine-grained layers compact more inelastically. Traditional CGR measurements from boreholes are sparse, costly and geographically limited.

Source- USGS

Contribution

- **Sky-to-subsurface approach**: Physics-guided ML framework integrating satellite observations with physics to infer subsurface geologic composition.
- **Data–physics integration: Sentinel-1 InSAR** subsidence and **GRACE-derived** groundwater-storage with **effective stress and poroelasticity theory.**
- **Scalable approach:** Maps soil characteristics, supporting improved groundwater management and climate adaptation in data-sparse regions.



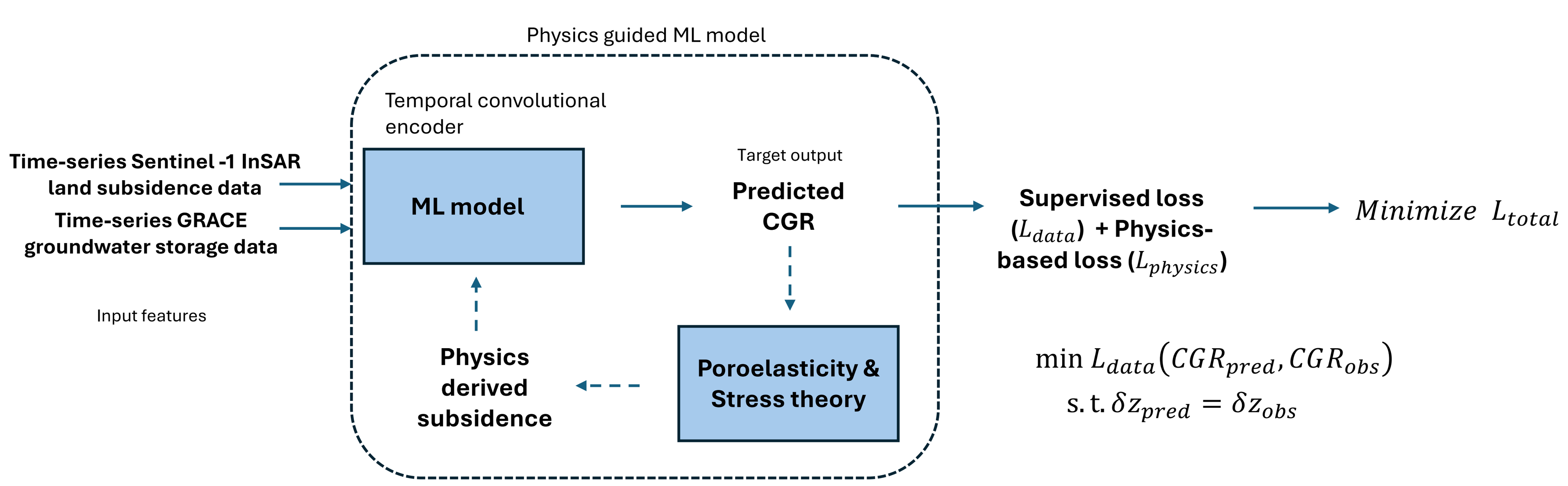
Climate Impact

- CGR helps quantify how **groundwater pumping drives land subsidence**, supporting better **groundwater management decisions**.
- CGR indicates how **aquifers respond to pumping**, helping identify **areas vulnerable to land subsidence**..
- By linking **satellite-based monitoring** with **physics-informed AI**, this work provides a scalable tool to assess how aquifers respond to changing climate and extraction pressures.
- Inferring **subsurface soil composition** from space enables identification of regions prone to **irreversible subsidence and water loss**, supporting adaptation strategies under increasing drought frequency.
- The framework **bridges Earth observation, hydrogeologic physics, and AI**, advancing **climate-resilient water management**.

Study Details

- Study focuses on **California's Central Valley**, a region experiencing severe land subsidence due to prolonged groundwater extraction.
- **Land Subsidence: Sentinel-1 InSAR** data (Tracks 42 & 144, 2014–2020) with **2 km × 2 km** resolution.
- **Groundwater Storage (GWS): GRACE-DA/GLDAS 2.2** data (daily, 0.25° resolution), interpolated to match the InSAR grid.
- **Soil Geology: USGS CVHM** geotexture dataset (~8,500 boreholes) providing coarse/fine-grain composition.
- **Temporal encoder**—implemented as a **Temporal Convolutional Network (TCN)**—with a **fully connected decoder** for spatial prediction.
- TCN captures groundwater–subsidence dynamics, while the combined **data-driven and physics-based loss** improves generalization.

Methodology

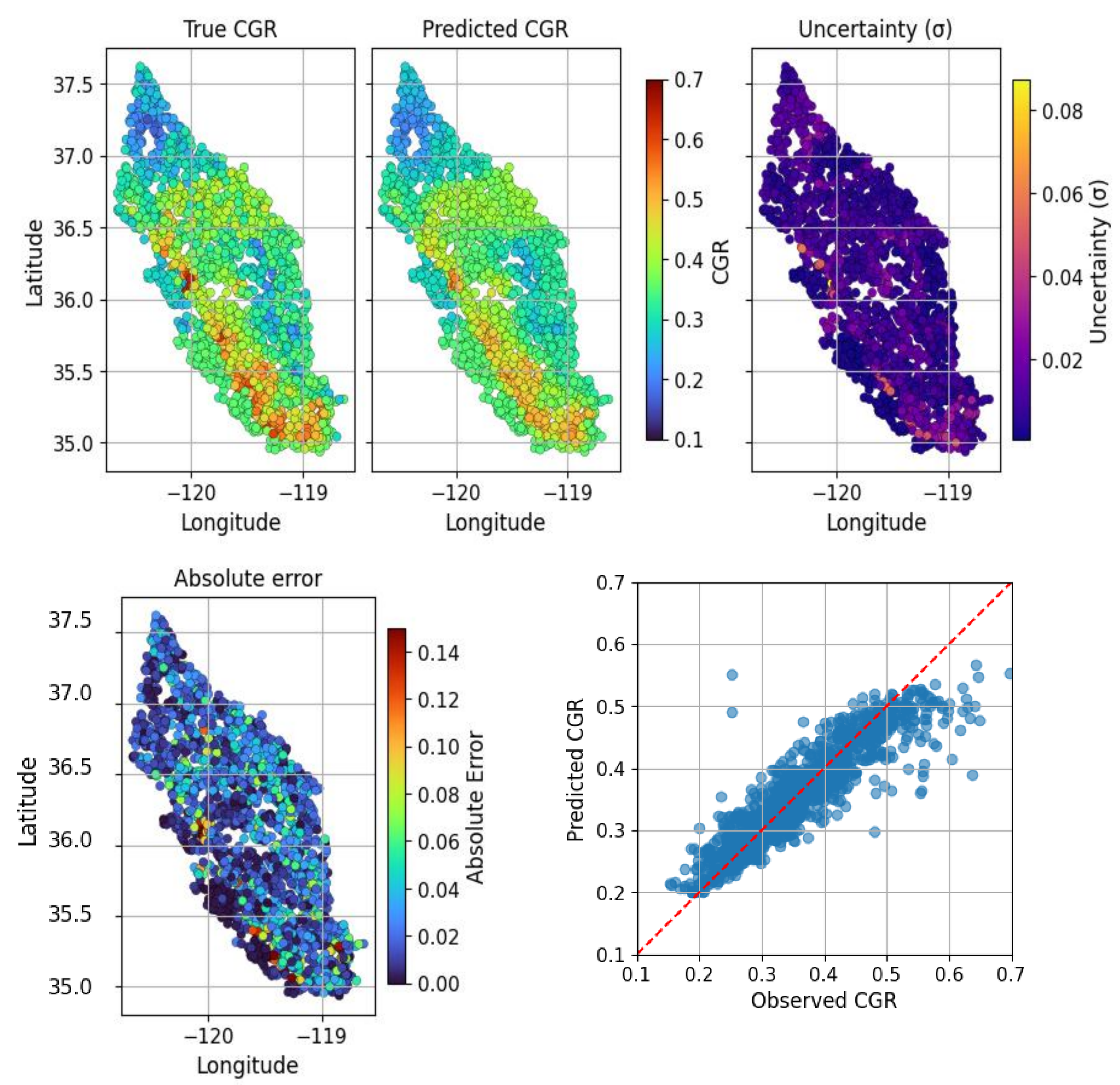


$$S_{sk} = S_{sk,fine}(1 - CGR_{pred}) + S_{sk,coarse}CGR_{pred}, \qquad \delta z_{pred} = -\frac{S_{sk}}{S_s}\delta d_{gws}$$

$$L_{total} = \lambda_d L_{data}(CGR_{pred}, CGR_{obs}) + \lambda_p L_{physics}(\delta z_{pred}, \delta z_{obs})$$

Results

Case A



- **Case A**: Physics-informed ML model, integrating poroelasticity theory with GRACE and Sentinel-1, achieves strong performance on unseen test data (Correlation: 0.90, NSE: 0.81, Mean Uncert.: 0.011).

Results Continued…

- **Case B**: To evaluate generalization to new geographic regions, location was excluded. The model still performs well (Correlation: 0.88, NSE: 0.76, Mean Uncert.: 0.014), demonstrating improved spatial robustness.
- **Case C**: Excluding physics improves numerical fit (Correlation: 0.92, NSE: 0.83, Mean Uncert.: 0.0235), model tends to overfit noise and lacks ability to leverage groundwater information and is less physically consistent.

Discussion and Future Work

- This study demonstrates that CGR can be effectively predicted remotely by leveraging satellite based temporal measurements, and by exploiting the interdependency of subsidence and GWS variations.
- Embedding physics (GWS and poroelasticity) ensures causality, physical consistency, reduces prediction uncertainty and extrapolation risks.
- We plan to expand model training to include additional geographic regions and longer multi-year time series, enabling improved spatial generalization and better capture of wet/dry cycles and seasonal aquifer behavior.
- Integrating higher temporal resolution InSAR data from NASA–ISRO's NISAR mission to improve subsidence characterization and test global scalability.

Acknowledgement

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA) (80NM0018D0004).

LinkedIn