Filling the gap: Estimation of soil composition using InSAR, groundwater depth, and precipitation data in California's Central Valley

Kyongsik Yun[1], Kyra H. Kim[1], Anshuman Pradhan[2], John Reager[1], Zhen Liu[1], Michael Turmon[3], Alexander Huyen[1], Thomas Lu[1], Venkat Chandrasekaran[2], and Andrew Stuart[2]

[1]NASA Jet Propulsion Laboratory
[2]California Institute of Technology
[3]Jet Propulsion Laboratory, California Institute of Technology

November 22, 2022

Abstract

California's Central Valley is responsible for $17 billion of annual agricultural output, producing 1/4 of the nation's food. However, land in the Central Valley is sinking at a rapid rate (as much as 20 cm per year) due to continued groundwater pumping. Land subsidence has a significant impact on infrastructure resilience and groundwater sustainability. It is important to understand subsidence and groundwater depletion in a consistent framework using improved models capable of simulating in-situ well observations and observed subsidence. Currently, groundwater well data is sparse and sampled irregularly, compromising our understanding of groundwater changes. Moreover, groundwater pumping data is a major missing piece of the puzzle. Limited data availability and spatial/temporal uncertainty in the available data have hampered understanding the complex dynamics of groundwater and subsidence. To address this limitation, we first integrated multimodal data including InSAR, groundwater, precipitation, and soil composition by interpolating data with the same spatial and temporal resolutions. We then identified regions with different temporal dynamics of land displacement, groundwater depth, and precipitation. Some areas (e.g., Helm) with coarser grain soil compositions exhibited potentially reversible land transformations (elastic land compaction). Finally, we fed the integrated data into the deep neural network of a gated recurrent unit-based sequence-to-sequence generation model. We found that the combination of InSAR, groundwater depth, and precipitation data had predictive power for soil composition using deep neural networks (correlation coefficient R=0.83, normalized Nash-Sutcliffe model efficiency NNSE=0.84). A random forest model was tested as baseline (R=0.65, NNSE=0.69). We also achieved significant accuracy with only 40% of the training data (NNSE=0.8), suggesting that the model can be generalized to other regions for indirect estimation of soil composition. Our results indicate that soil composition can be estimated using InSAR, groundwater depth and precipitation data. In-situ measurements of soil composition can be expensive and time consuming and may be impractical in some areas. The generalizability of the model sheds light on high spatial resolution soil composition estimation utilizing existing measurements.

Filling the gap: Estimation of soil composition using InSAR, groundwater depth, and precipitation data in California's Central Valley

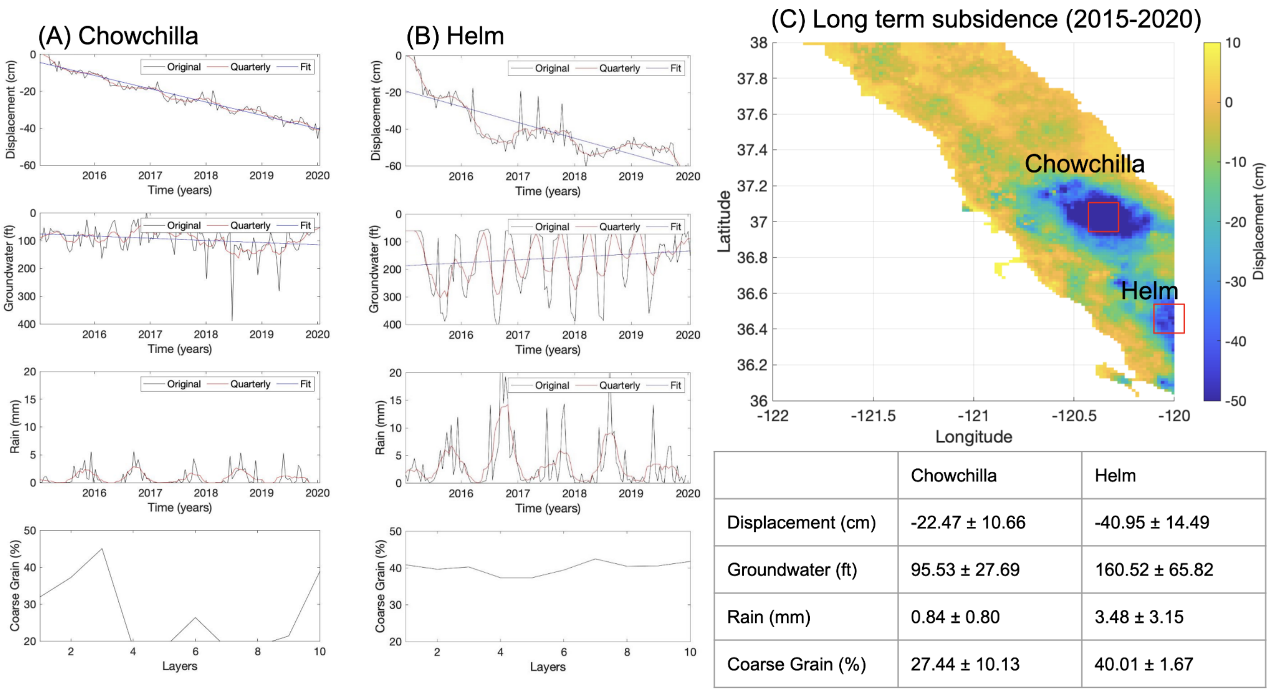
Kyongsik Yun, Kyra Kim, Anshuman Pradhan, John Reager, Zhen Liu, Michael Turmon, Alexander Huyen, Thomas Lu, Venkat Chandrasekaran, Andrew Stuart
Kyongsik.Yun@jpl.nasa.gov, Jet Propulsion Laboratory, California Institute of Technology

California's Central Valley is responsible for $17 billion of annual agricultural output, producing 1/4 of the nation's food. However, land in the Central Valley is sinking at a rapid rate (as much as 20 cm per year) due to continued groundwater pumping. Land subsidence has a significant impact on infrastructure resilience and groundwater sustainability. It is important to understand subsidence and groundwater depletion in a consistent framework using improved models capable of simulating in-situ well observations and observed subsidence. Currently, groundwater well data is sparse and sampled irregularly, compromising our understanding of groundwater changes. Moreover, groundwater pumping data is a major missing piece of the puzzle. Limited data availability and spatial/temporal uncertainty in the available data have hampered understanding the complex dynamics of groundwater and subsidence.

To address this limitation, we first integrated multimodal data including InSAR, groundwater, precipitation, and soil composition by interpolating data with the same spatial and temporal resolutions (every 2 weeks on a 1kmX1km grid). We then identified regions with different temporal dynamics of land displacement, groundwater depth, and precipitation (Figure 1). Some areas (e.g., Helm) with coarser grain soil compositions exhibited potentially reversible land transformations (elastic land compaction), which can inform government agencies of better management of groundwater use and water recharge for subsidence recovery.

We fed the integrated data into the deep neural network of a gated recurrent unit (GRU)-based sequence-to-sequence generation model. We found that the combination of InSAR, groundwater depth, and precipitation data had predictive power for soil composition using deep neural networks (correlation coefficient R=0.83, normalized Nash-Sutcliffe model efficiency NNSE=0.84). A random forest model was tested as baseline (R=0.65, NNSE=0.69) (Figure 2). We also achieved significant accuracy with only 40% of the training data (NNSE=0.8), suggesting that the model can be generalized to other regions for indirect estimation of soil composition.

Our results indicate that soil composition can be estimated using InSAR, groundwater depth and precipitation data. In-situ measurements of soil composition can be expensive and time consuming and may be impractical in some areas. The generalizability of the model sheds light on high spatial resolution soil composition estimation utilizing existing measurements.



	Chowchilla	Helm
Displacement (cm)	-22.47 ± 10.66	-40.95 ± 14.49
Groundwater (ft)	95.53 ± 27.69	160.52 ± 65.82
Rain (mm)	0.84 ± 0.80	3.48 ± 3.15
Coarse Grain (%)	27.44 ± 10.13	40.01 ± 1.67

Figure 1. Two representative regions of the Central Valley with significant subsidence with different characteristics. (A) Chowchilla has been shown to maintain monotonically decreasing land displacements, less fluctuating groundwater depth, relatively low precipitation, and high fine-grain ratio across the middle soil layers. (B) Helm, on the other hand, exhibited fluctuating land displacements, relatively large seasonal changes in groundwater depth, high precipitation, and a higher overall coarse-grain ratio across all soil layers. (C) A displacement map including Chowchilla and Helm (2015-2020).

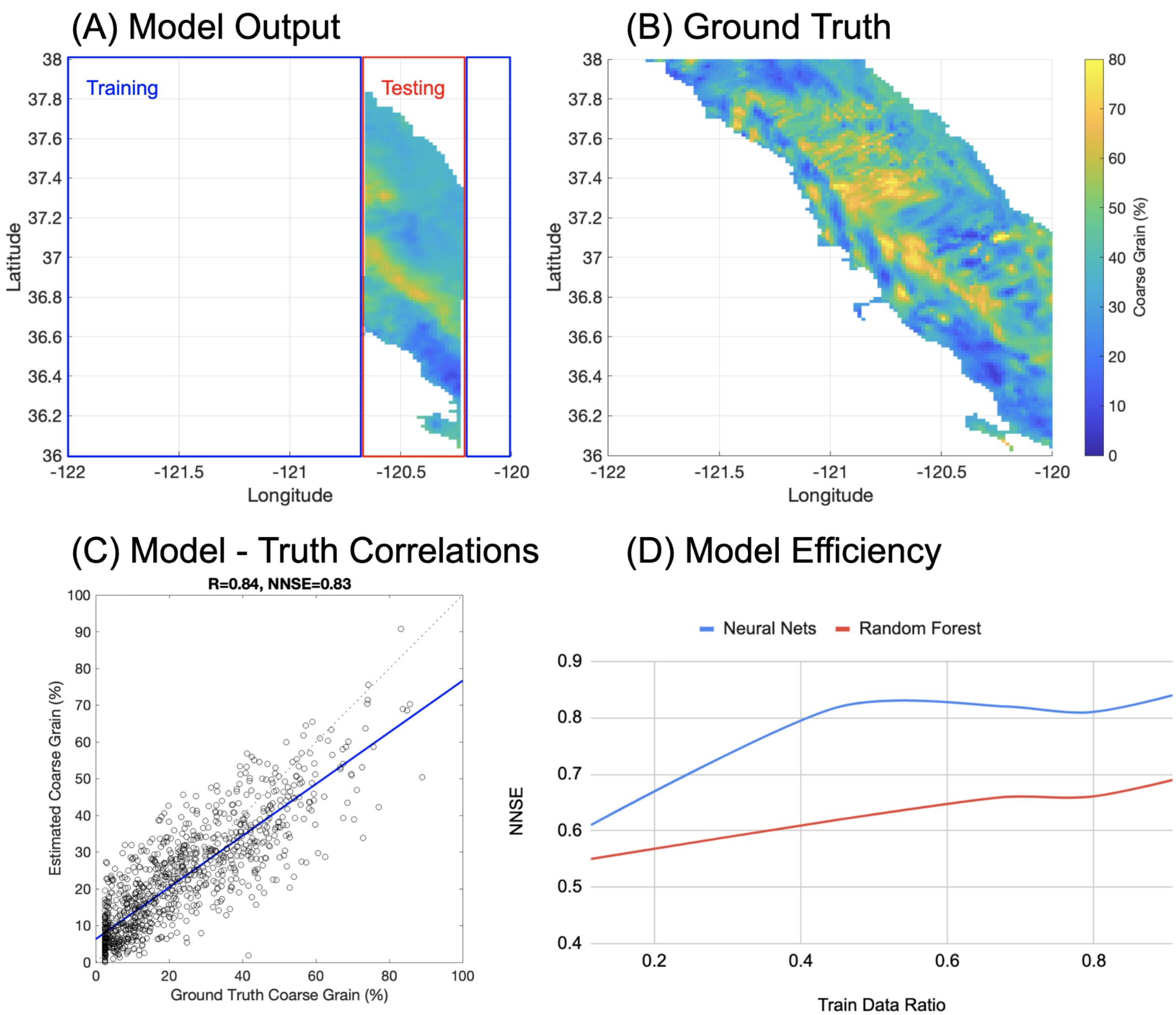


Figure 2. Deep neural network model of soil composition estimation. (A) Example training and testing areas randomly selected for validation. (B) Ground truth coarse-grained ratio of Central Valley in soil layer 1. (C) Correlation plot between ground-truth coarse-grain and the estimated coarse-grain ratios (correlation coefficient R=0.84, normalized Nash-Sutcliffe model efficiency NNSE=0.83). (D) NNSE of neural networks and random forest models over various training data ratios (0.1-0.9).