National Aeronautics and Space Administration

Uncertainty-aware soil property mapping algorithm for Imaging Spectrometers

Principal Investigator: Nimrod Carmon (398); Co-Investigators: Dana Chadwick (329), David Schimel (329)

Program: FY22 R&TD Innovative Spontaneous Concepts

Objectives:

The objective of this work was to develop an uncertainty-aware partial least squares regression (PLSR) soil property mapping algorithm and to demonstrate its performance.

Background:

Soil, a crucial component of the Earth system, has a critical role in most natural and anthropogenic processes. Thus far, mapping soil conditions using imaging spectroscopy was done in a limited capacity for local regions using algorithms requiring fine-tuning and manual attention. In light of new orbital missions such as the surface biology and geology (SBG), there is a need for soil property mapping algorithms that are scalable, automated, and applicable globally. Because soil is a complex system with a multitude of physical, chemical, and biological elements affecting its spectral signature, spectral models to map soil properties are based on machine learning rather than physical modeling. These rely on large soil spectral libraries on which a prediction model is learned, to later apply the model to new samples for which only spectral measurements are available. While these models work well on laboratory-measured samples, applying them on remote-sensing reflectance maps is challenging due to sensor noise, retrieval errors, and mixed pixels. In contrast to laboratory measurements, remote-sensing reflectance estimates have large uncertainties which must be accounted for. If not, the estimated soil conditions maps would have large errors, reducing the scientific value of these measurements.

Approach and Results:

In this work, we developed a new capability of using reflectance uncertainties to develop uncertainty-aware soil property models. Our algorithm is based on traditional partial least squares regression (PLSR). Here, we enhanced the PLSR algorithm to be able to digest a posterior covariance matrix and weight predictor spectral bands appropriately (Algorithm 1), ensuring uncertain spectral channels are used less in the model. To evaluate the impact of this enhancement we've tested our new algorithm on the Valencia study site near Santa Clarita, California (Figure 1), with measurements from AVIRIS-NG. We calculated reflectance maps with associated uncertainties and mapped calcium carbonate ($CaCO_3$) using the traditional uncertainty-naïve, and our new uncertainty-aware approach (Figure 2). Our results indicate that the calcium carbonate maps calculated using our uncertainty-aware model are less coupled to the scene's topography compared to the uncertainty-naïve model. Moreover, while the uncertainty naïve model over-estimates the calcium carbonate maps beyond the training dataset's response range, the uncertainty aware model remains within the appropriate range of $0 \leq CaCO_3 \leq 90$.

Algorithm 1: Uncertainty-Aware Partial Least Squares Regression (UA-PLSR)

Data: $\mathbf{X}_{n*m}, \boldsymbol{y}_{n*1}, \boldsymbol{\rho}_{m*1}, \mathbf{S}_{m*m}$
Result: $\boldsymbol{b}, \mu(\mathbf{X}), \mu(\boldsymbol{y})$
Initialization
$\mathbf{X} \longleftarrow [\mathbf{X}^T - \mu(\mathbf{X}^T)]^T$ // per-sample mean center
$\mathbf{X}^* \longleftarrow (\mathbf{X} - \mu(\mathbf{X}))(\sigma(\mathbf{X}))^{-1}$ // per-variable mean center and scale
$\boldsymbol{y}^* \longleftarrow \boldsymbol{y} - \mu(\boldsymbol{y})$ // mean center
$\mathbf{X}_k \longleftarrow \mathbf{X}^*$
$\boldsymbol{y}_k \longleftarrow \boldsymbol{y}*$
for k *in* $ncomp$ **do**
  $\mathbf{L}\mathbf{L}^T = \boldsymbol{S}$ // Cholesky decomposition
  $\boldsymbol{C} = (\mathbf{X}_k\mathbf{L}^{-1})^T\boldsymbol{y}_k$ // leverage uncertainty
  $\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V} = SVD(\boldsymbol{C})$
  $\boldsymbol{x}_w^{(k)} = \boldsymbol{U}[:, 0], y_w^{(k)} = \boldsymbol{V}[0]$ // $\boldsymbol{x}_w^{(k)}$ is m*1, $y_w^{(k)}$ is 1*1
  $\boldsymbol{x}_s^{(k)} = \mathbf{X}_k\boldsymbol{x}_w^{(k)\,T}$ // $\boldsymbol{x}_s^{(k)}$ is n*1
  $\boldsymbol{y}_s^{(k)} = \boldsymbol{y}_k\boldsymbol{y}_w^{(k)}$ // $\boldsymbol{y}_s^{(k)}$ is n*1
  $\boldsymbol{x}_l^{(k)} = (\boldsymbol{X}_s^{(k)T}\boldsymbol{X}_s^{(k)})^{-1}(\boldsymbol{X}_s^{(k)T}\mathbf{X}_k)$ // $\boldsymbol{x}_l^{(k)}$ is 1*m
  $\mathbf{X}_{k+1} \longleftarrow \mathbf{X}_k - \boldsymbol{X}_s^{(k)}\boldsymbol{x}_l^{(k)}$ // deflate $\mathbf{X}_k$
  $y_l^{(k)} = (\boldsymbol{X}_s^{(k)\,T}\boldsymbol{X}_s^{(k)})^{-1}(\boldsymbol{X}_s^{(k)T}\boldsymbol{y}_k^{(k)})$ // $\boldsymbol{y}_l^{(k)}$ is 1*1
  $\boldsymbol{y}_{k+1} \longleftarrow \boldsymbol{y}_k - \boldsymbol{x}_s^{(k)}\boldsymbol{y}_l^{(k)}$ // deflate y
end
$\mathbf{W}_x = [\boldsymbol{x}_w^{(0)}, \boldsymbol{x}_w^{(1)}, ..., \boldsymbol{x}_w^{(ncomp-1)}]$ // W_x is m*ncomp
$\mathbf{L}_x = [\boldsymbol{x}_l^{(0)}; \boldsymbol{x}_l^{(1)}; ...; \boldsymbol{x}_l^{(ncomp-1)}]$ // L_x is ncomp*m
$\boldsymbol{l}_y = [y_l^{(0)}, y_l^{(1)}, ..., y_l^{(ncomp-1)}]$ // L_y is ncomp*1
$\boldsymbol{\beta} = \mathbf{W}_x * (\mathbf{L}_x\mathbf{W}_x)^{-1} * \boldsymbol{l}_y^T$ // B is m*1
$\hat{y}(\rho) = (\boldsymbol{\rho} - \mu(\mathbf{X}))^T * \boldsymbol{\beta} + \mu(\boldsymbol{y})$ // ρ is an observed reflectance

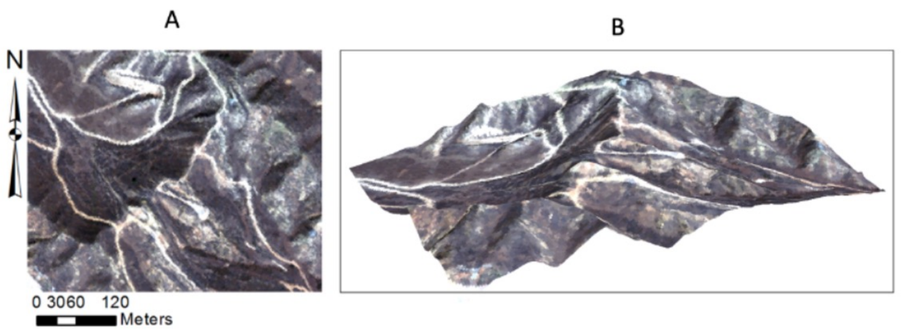

Figure 1: The Valencia study site near Santa Clarita, California. (A) A true color RGB image of the study site. (B) a true color RGB overlayed on a digital elevation map to show the rugged topography

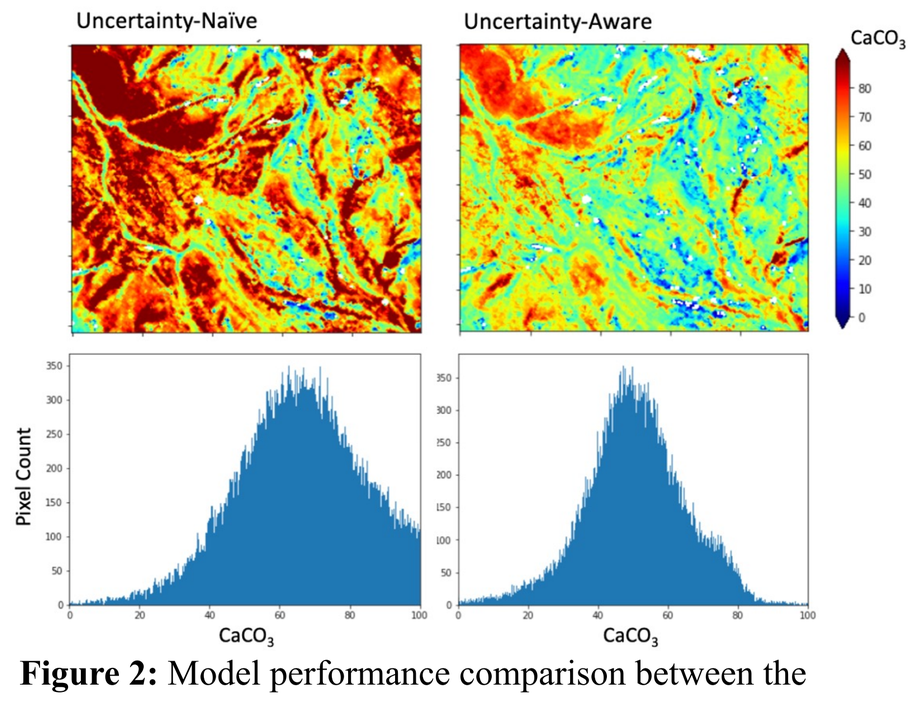

Figure 2: Model performance comparison between the uncertainty naive (left column) and uncertainty aware (right column) PLSR models for calcium carbonate. Upper panels show a spatial map of the calcium carbonate estimates, while bottom panels show the value distributions of these estimates.

Significance/Benefits to JPL and NASA:

Uncertainty in reflectance maps from remote-sensing imaging spectrometers is an unavoidable reality. While airborne campaigns can afford to wait for optimal atmospheric conditions, reducing this uncertainty, orbital instruments cannot. Hence, estimating and leveraging reflectance uncertainty would ensure the Biogeophysical maps calculated based on these reflectance maps are accurate. In this work, we've demonstrated a new capacity that enables leveraging uncertainty estimates within model development and showed initial results suggesting improved model performance. While developed and evaluated for soil property maps, this approach is seamlessly applicable for any Biogeophysical mapping using PLSR models.

This new capacity would be useful for orbital imaging spectroscopy missions such as SBG which would encounter challenging atmospheric conditions that reduce the reflectance quality. With the approach developed in this work, these reflectance uncertainties can be systematically leveraged to result in more accurate surface conditions maps that are less biased by uncertainties originating from instrument noise and/or challenging atmospheric conditions.

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

Clearance Number: CL#22-5213
Poster Number: RPC-116

PI/Task Mgr. Contact Information:
Email: nimrod.carmon@jpl.nasa.gov