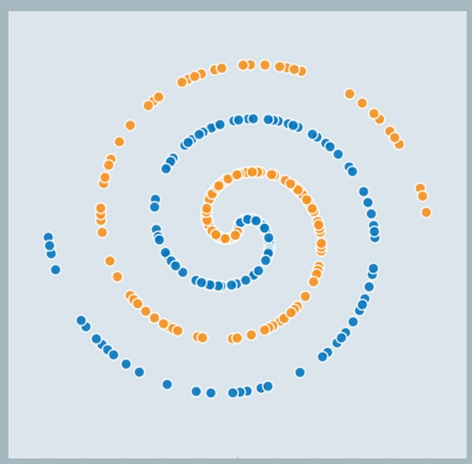
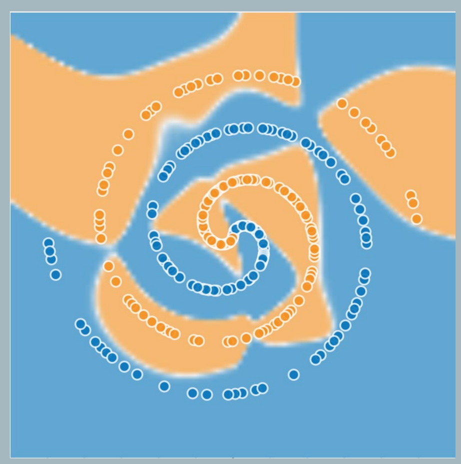
Estimating Soil Organic Carbon with Physics-Informed Neural Networks

James Sargeant

Supervisors: Prof Shyh Wei Teng, Prof Manoranjan Paul, Prof Manzur Murshed and Mr David Brennan

Problem and Motivation

- Accurate Soil Organic Carbon (SOC) readings are essential for farm management, carbon farming & market access
- Measurement is expensive at approx. $500 per sample
- Science-based estimation using soil and vegetation indices does not use all spectral information
- Estimation using neural networks only considers data, not any underlying science, as illustrated below:

Find blue/orange areas → Valid, but limited solution

Small improvements in SOC accuracy are worth millions of dollars across the farming community.

Proposed solution

Combine Spectral data and Indices in a Physics-Informed Neural Network (PINN):

- Soil and vegetation indices guide the network to an initial, scientifically-based solution
- Refine solution using all spectral data

Observed benefits of PNNs in other disciplines:

- Less data needed
- Solution valid outside data range
- Faster convergence to a solution

Science:

- For many soil properties, including SOC, reflectance at one wavelength is often linearly related to reflectance at another wavelength. Soil complexity degrades this relationship
- Soil & Vegetation Indices more resilient to roughness & moisture than raw satellite wavelength bands

Novelty: PINNS not previously used for SOC Estimation

Experimental design

Dataset:

- Finland and Sweden subset of the European Land Use/Cover Area frame Survey (LUCAS)
- 2,500 wavelengths resampled to 13 Sentinel-2 satellite bands
- 6,376 SOC measurements and spectra in dataset
- SOC: Range 0 – 560 g/kg, mean 110 g/kg

Single Index:

- Inverted Red Edge Chlorophyll Index (IRECI)
- IRECI good predictor of SOC (R^2 ~ 0.65 for dataset)

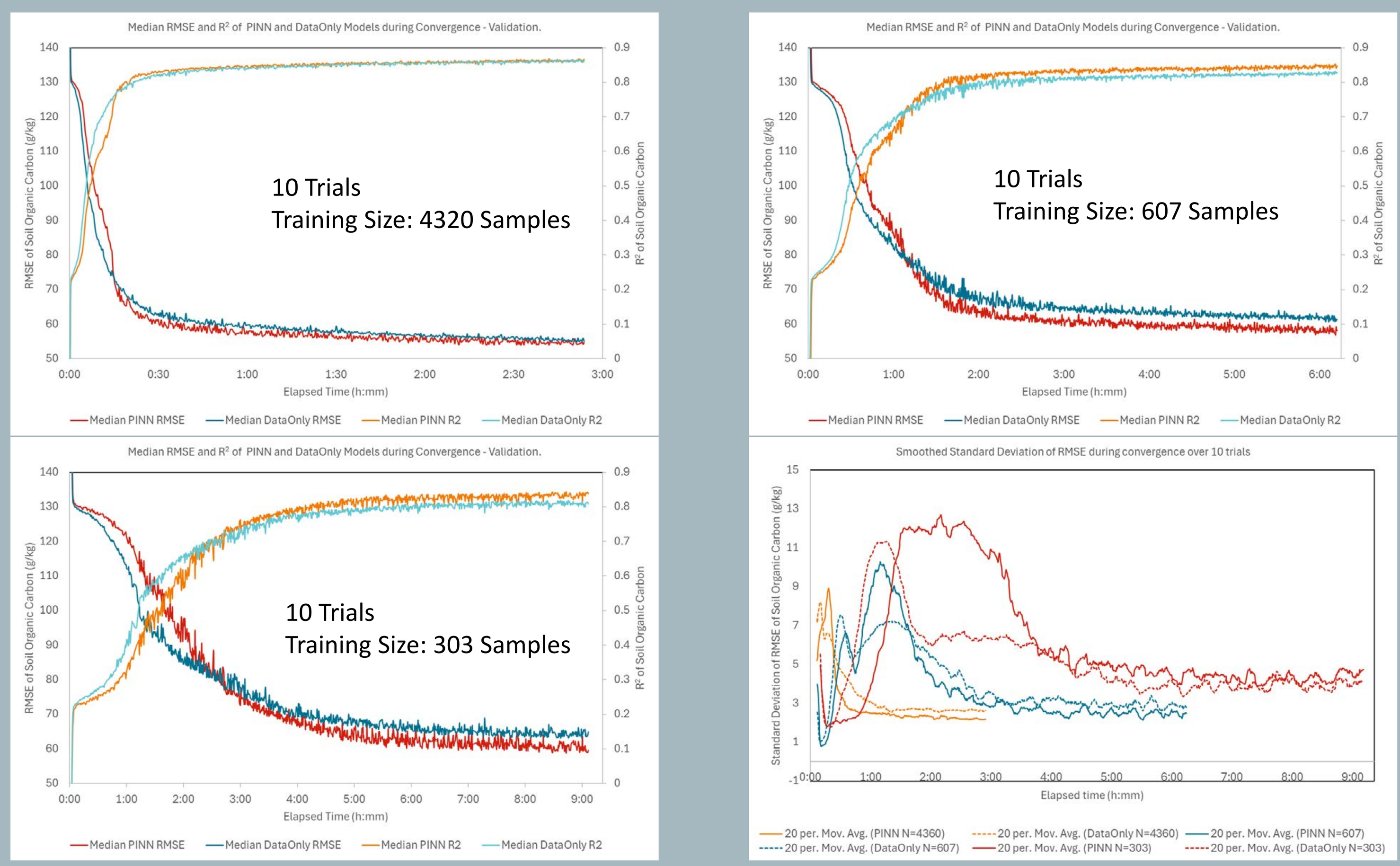
Neural Network Model:

- 2-stage Convolutional Neural Network similar to neural network model architecture found in literature
- Index incorporated in loss function, not model input

PINN version compared with Data Only version of same model using identical training and validation sets.

Preliminary Results

- PINN method accuracy better than Data Only method as sample size decreases
- PINN method more robust to noise than Data Only method for larger sample sizes



Future research includes the use of multiple indices and application of PINNs in Wimmera farms.

SOIL CRC
Federation University
Charles Sturt University
Deakin University
Wimmera CMA
Australian Government Department of Industry, Science and Resources
Cooperative Research Centres Program