Modifying Priors to Improve Neural Network Predictions of Vegetation Biophysical Variables from Sentinel 2 Imagery

Hichem Lahiouel

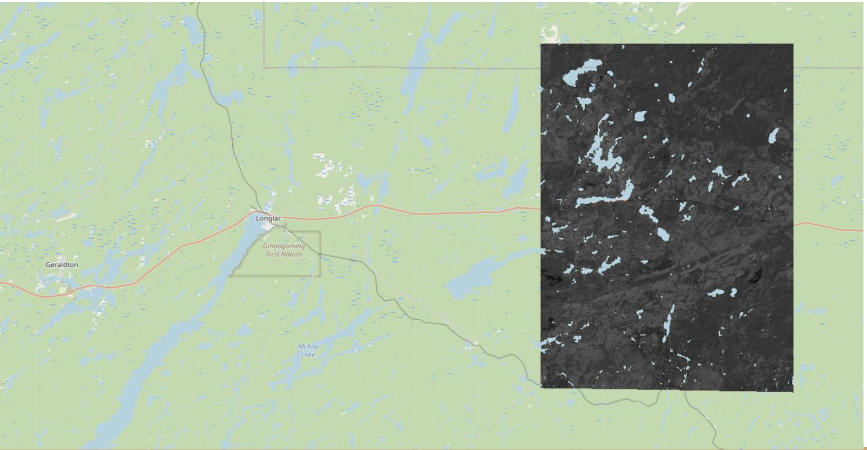
Vegetation Parameters

- Leaf Area Index (LAI)
  - Dimensionless quantity
  - Ranges from zero to fourteen
  - Total area of leaves per unit ground surface area
- Fraction of Cover (FCOVER)
  - Ranges from zero to one
  - Fraction of ground covered by green vegetation
- Fraction of Photosynthetically Active (Solar) Radiation (FAPAR)
  - Ranges from zero to one
  - Fraction of the solar radiation absorbed by live leaves for the photosynthesis activity

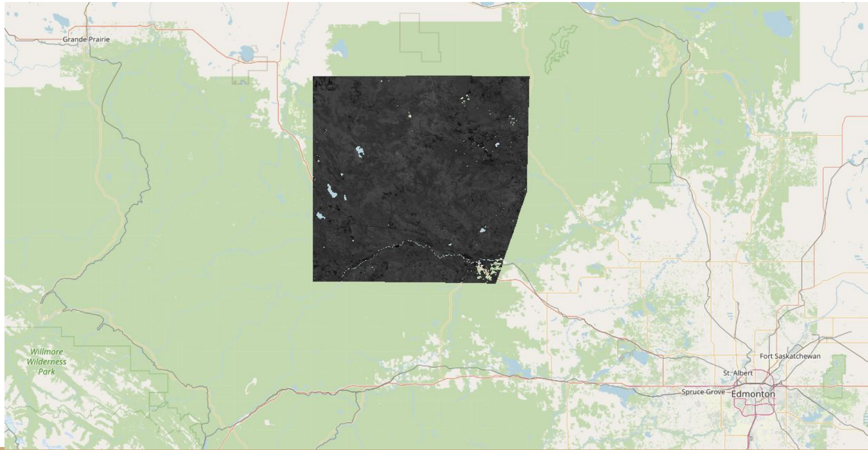
Retrieval Approach

- The Sentinel-2 data was retrieved from Google Earth Engine
- Three different regions in Canada were chosen for analysis
- These regions were chosen because they have a wide range of values for LAI, FAPAR, and FCOVER

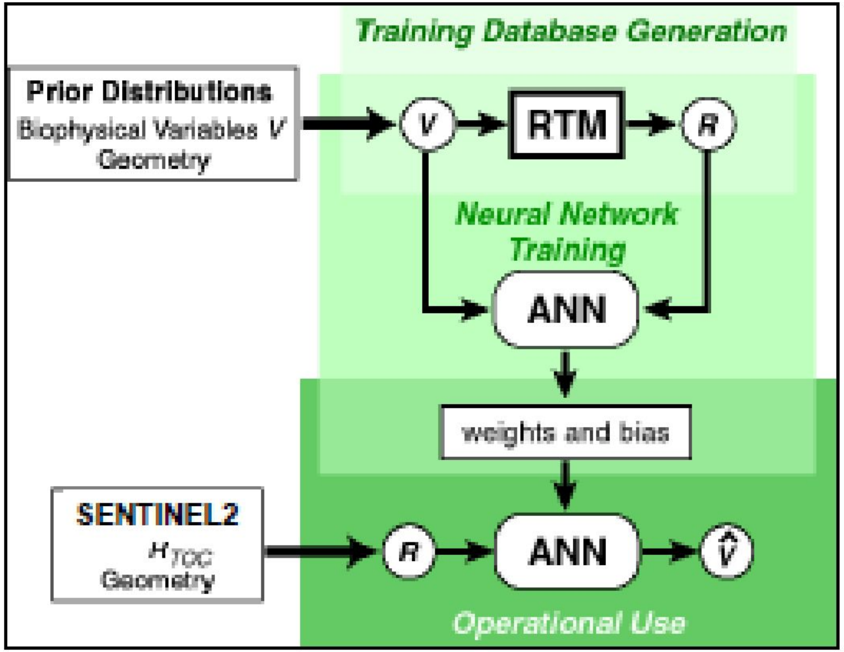
Geraldton, Ontario



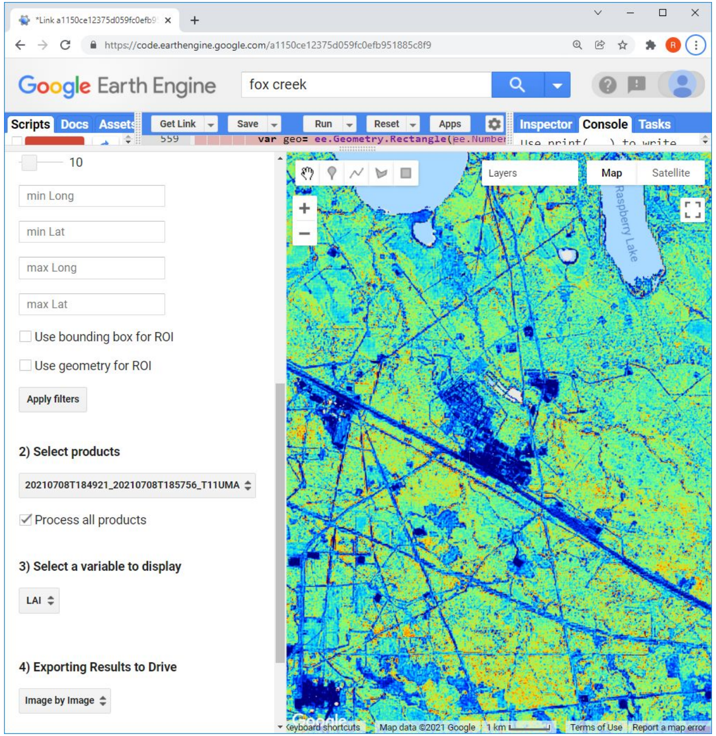
Fox Creek, Alberta



Process to get Calibration Data

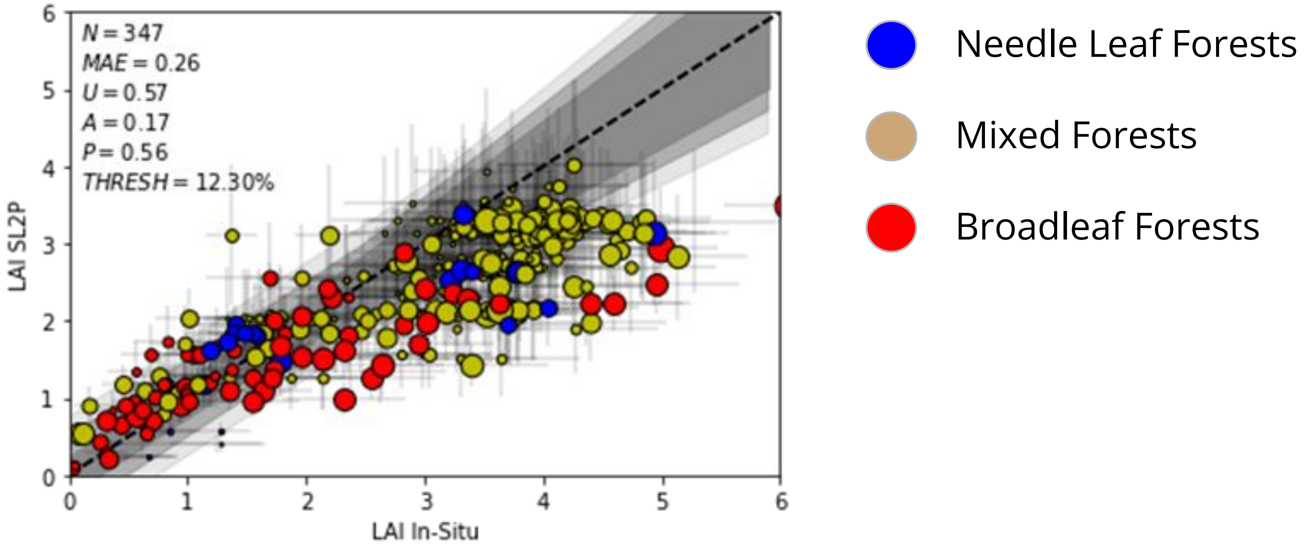


Process to get Images and Image Data



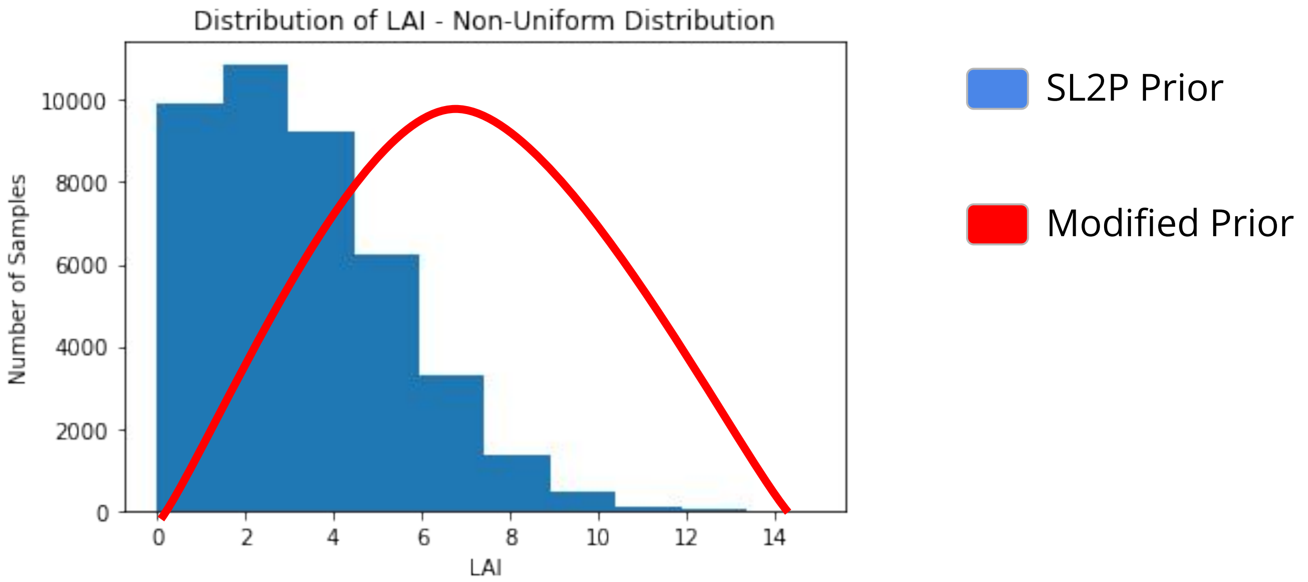
Research Question

- How can we reduce the bias towards low values of LAI and improve estimation of high values of LAI?
    - More specifically, can we improve our estimates of SL2P by modifying prior distributions given local information?

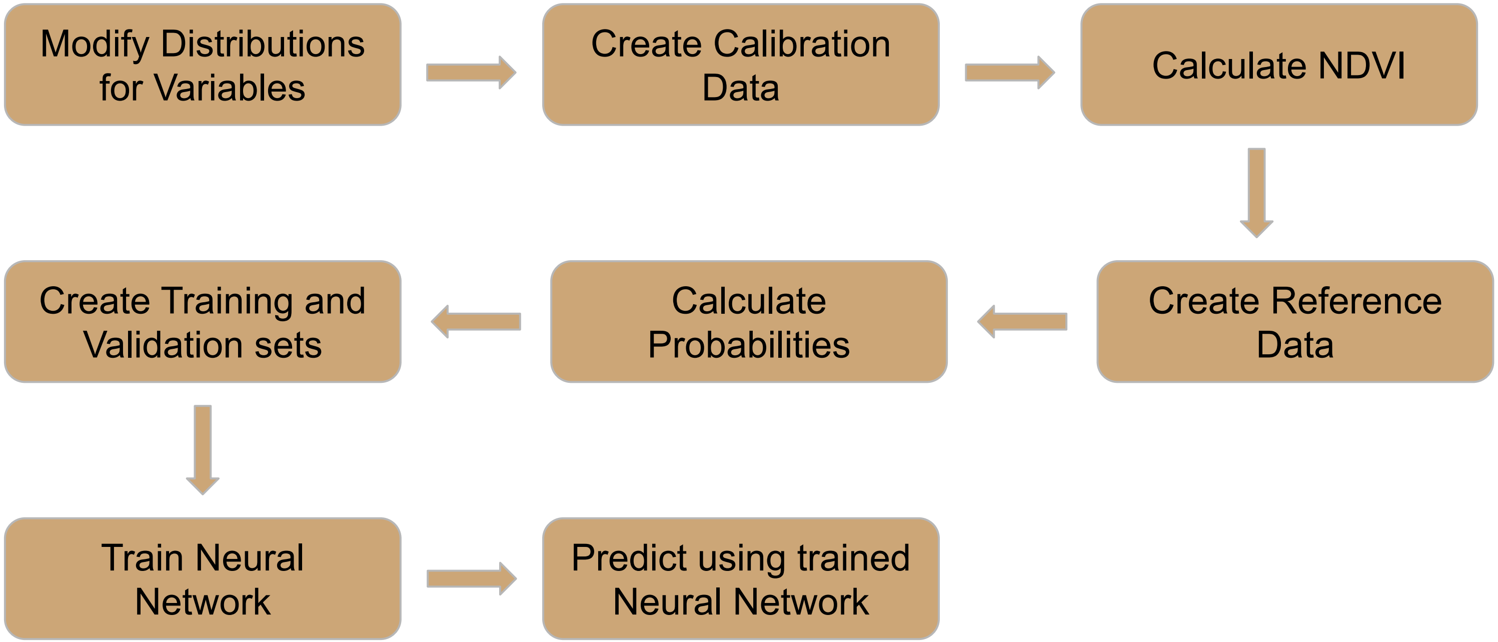


Modifying Prior Distributions

- The distributions for the different variables are determined in an Excel file
- One experiment I did was to compare the impact of modifying the prior distribution of LAI such that the LAI values varied from 1 to 10 (represented local distributions)



Workflow Chart



Calibration Database

- Contains spectral and angle data, plus data for the Biophysical vegetation variables themselves
- Simulations using PROSAIL Radiative Transfer Models with Default Priors

Sentinel-2 Spectral Bands

Geometry

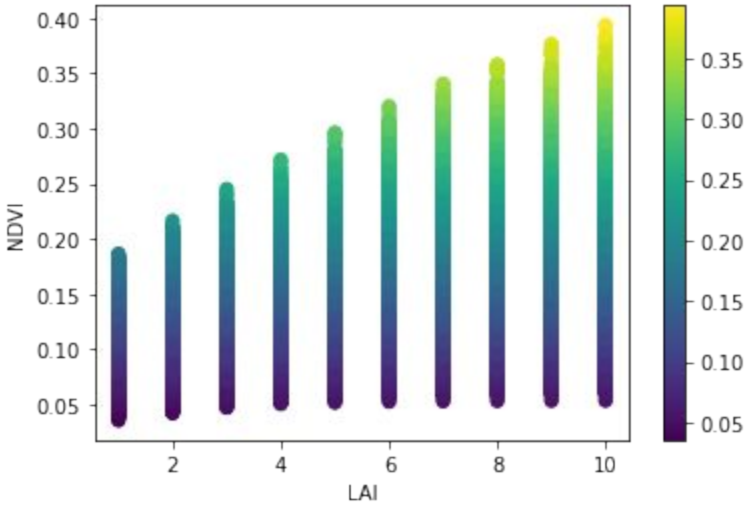
Variables

	B0	B1	B2	B3	B4	B5	B6	B7	A1	A2	A3	LAI	FAPAR	FCOVER
0	0.1312746755	0.0743505601	0.1729546392	0.2952160738	0.3209243927	0.3303722287	0.1873586099	0.1113007554	0.9638065534	0.5275411500	-0.5752876765	1	0.4391203617	0.4803244610
1	0.1019407701	0.0868185076	0.1599851769	0.3680334831	0.4389734403	0.4759755142	0.3043390179	0.1996883219	0.9995852344	0.5284670738	-0.4777431279	1	0.5296906764	0.5338211214
2	0.0577228876	0.0543735427	0.0726863963	0.1146353863	0.1262129829	0.1300015294	0.1048168652	0.0853760008	0.9807097061	0.6843199822	0.6909005295	1	0.2406051357	0.2324486617
3	0.0608335819	0.0423475494	0.0917885392	0.1738484853	0.1980468499	0.2153382831	0.1396374825	0.0906137034	0.9865515976	0.7869673865	0.8590635747	1	0.3515168619	0.3732491055
4	0.0789422177	0.1359804109	0.1677317262	0.2554469658	0.2962514793	0.3108466204	0.3167107908	0.2088934617	0.9999190888	0.8237423248	-0.9085753290	1	0.2890892994	0.2565897729
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
122875	0.0642385315	0.0276301255	0.0866383370	0.3106952678	0.4255780128	0.4466691326	0.1173674161	0.0414608341	0.9641458993	0.8501574022	-0.9266152189	10	0.9454688738	0.9530549586
122876	0.0913131658	0.0504649365	0.1222839075	0.3081750760	0.4047671458	0.4293059492	0.1725548796	0.1005631041	0.9616583754	0.8710934328	0.9265827175	10	0.8085248415	0.8034524935
122877	0.0621419307	0.0380236717	0.0921177459	0.2614424931	0.3577263614	0.3626203975	0.1294881968	0.0624406803	0.9757927762	0.8637999983	-0.9813994186	10	0.8185069725	0.8122050332
122878	0.0514800307	0.0288523379	0.0690024543	0.4158467721	0.6343113015	0.6354447032	0.1810943898	0.0474550174	0.9606931097	0.8268533675	-0.8694016324	10	0.9810850825	0.9992241257
122879	0.0668233773	0.0250243879	0.0907567987	0.4369922465	0.6772450624	0.6972009957	0.2347272437	0.0863371443	0.9591264867	0.8382274319	0.9967558231	10	0.9762029807	0.9928982883

122880 rows × 14 columns

Reference Database

- Normally we would choose a site and extract data from the site
- Contains the same kind of data as the Calibration data
- Simulated field sampling by extracting 200 samples from Calibration data
    - Samples were chosen by choosing the lowest 20% of NDVI

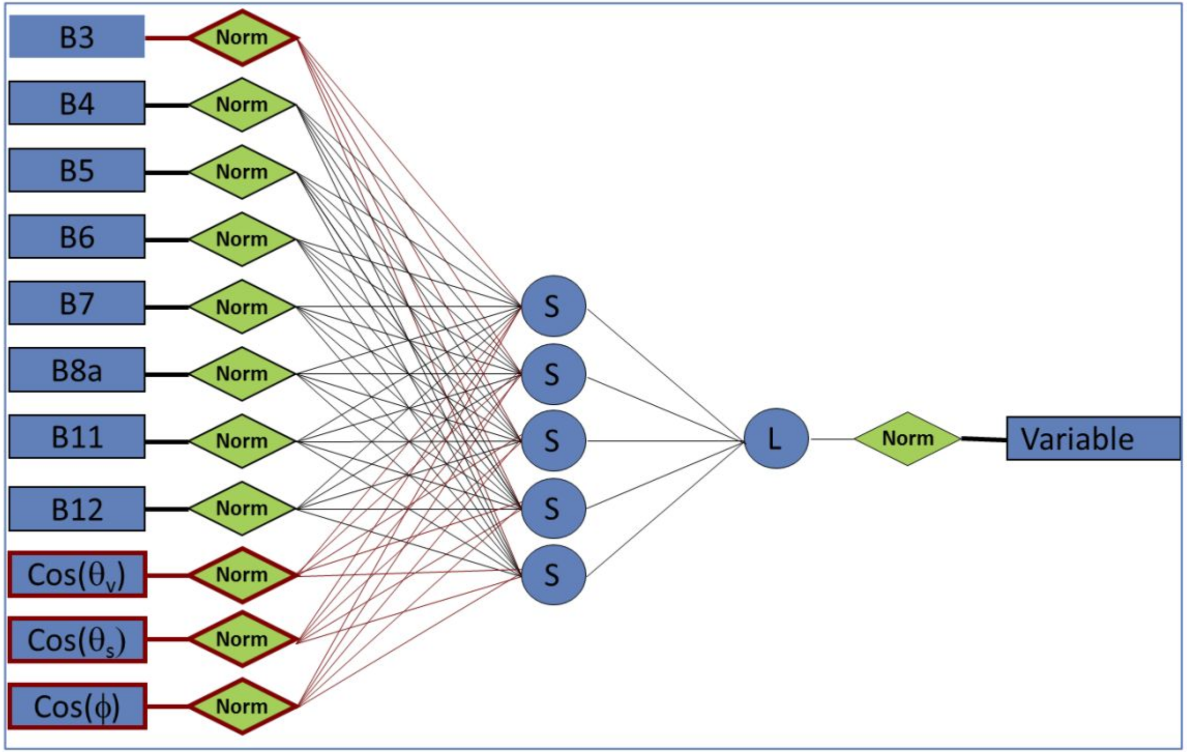


Fraction of Samples for given LAI in initial Calibration

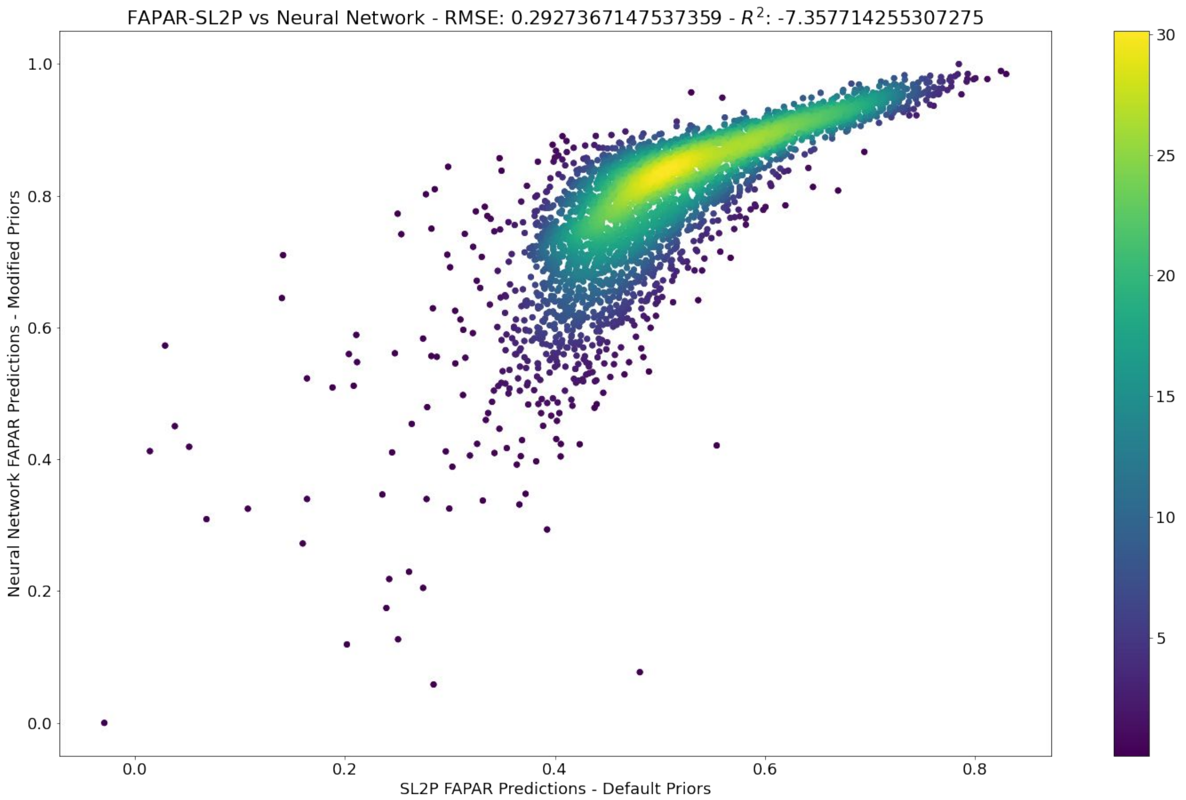
Tensorflow Neural Network

- The neural network was created in Tensorflow
- A sequential model was used with three layers
  - The first layer uses a Sigmoid activation function
  - The second and third layer use a Rectified Linear Unit for the activation function
- Tensorflow allows you to use training and validation weights for training
- In addition, Tensorflow has a feature that allows you to monitor important metrics such as Mean Squared Error and Mean Absolute Error
- Tensorflow also has a feature that will stop training when the loss function stops decreasing for a certain amount of epochs

Neural Network Structure



Sample Results for FAPAR Variable (Geraldton Region)



Number of Samples

Conclusions

- Modifying prior distributions impacts predictions of variables and may reduce bias

- It would be interesting to see how to incorporate the uncertainties in the data into the process and see how they affect the final results
- Further research could be done on how to improve the Neural Networks
- The results need to be validated with in-situ data

- It was really interesting to become more familiar with different tools in Python that are used for Data Science
  - For example, I gained experience with Pandas, Tensorflow, and Sci-Kit Learn

- Added data for subset ids, NDVI, and probabilities
  - NDVI
    - Normalized Difference Vegetation Index

	B0	B1	B2	B3	B4	B5	B6	B7	A1	A2	A3	LAI	FAPAR	FCOVER	NDVI	subset_id	prob
0	0.1312746755	0.0743505601	0.1729546392	0.2952160738	0.3209243927	0.3303722287	0.1873586099	0.1113007554	0.9638065534	0.5275411500	-0.5752876765	1	0.4391203617	0.4803244610	0.0561969505	0	0.9757562971
1	0.1019407701	0.0868185076	0.1599851769	0.3680334831	0.4389734403	0.4759755142	0.3043390179	0.1996883219	0.9995852344	0.5284670738	-0.4777431279	1	0.5296906764	0.5338211214	0.1278920384	1	0.9398261617
2	0.0577228876	0.0543735427	0.0726863963	0.1146353863	0.1262129829	0.1300015294	0.1048168652	0.0853760008	0.9807097061	0.6843199822	0.6909005295	1	0.2406051357	0.2324486617	0.0628120376	2	0.9940720495
3	0.0608335819	0.0423475494	0.0917885392	0.1738484853	0.1980468499	0.2153382831	0.1396374825	0.0906137034	0.9865515976	0.7869673865	0.8590635747	1	0.3515168619	0.3732491055	0.1066063936	3	0.9672943150
4	0.0789422177	0.1359804109	0.1677317262	0.2554469658	0.2962514793	0.3108466204	0.3167107908	0.2088934617	0.9999190888	0.8237423248	-0.9085753290	1	0.2890892994	0.2565897729	0.0978285043	4	0.9668794042
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
122675	0.0642385315	0.0276301255	0.0866383370	0.3106952678	0.4255780128	0.4466691326	0.1173674161	0.0414608341	0.9641458993	0.8501574022	-0.9266152189	10	0.9454688738	0.9530549586	0.1795355904	12283	0.9454364076
122676	0.0913131658	0.0504649365	0.1222839075	0.3081750760	0.4047671458	0.4293059492	0.1725548796	0.1005631041	0.9616583754	0.8710934328	0.9265827175	10	0.8085248415	0.8034524935	0.1642494776	12284	0.9448967476
122677	0.0621419307	0.0380236717	0.0921177459	0.2614424931	0.3577263614	0.3626203975	0.1294881968	0.0624406803	0.9757927762	0.8637999983	-0.9813994186	10	0.8185069725	0.8122050332	0.1621277372	12285	0.9591919288
122678	0.0514800307	0.0288523379	0.0690024543	0.4158467721	0.6343113015	0.6354447032	0.1810943898	0.0474550174	0.9606931097	0.8268533675	-0.8694016324	10	0.9810850825	0.9992241257	0.2088839645	12286	0.9188533966
122679	0.0668233773	0.0250243879	0.0907567987	0.4369922465	0.6772450624	0.6972009957	0.2347272437	0.0863371443	0.9591264867	0.8382274319	0.9967558231	10	0.9762029807	0.9928982883	0.2294218829	12287	0.9076213421

122680 rows × 17 columns