UTRECHT UNIVERSITY

Department of Information and Computing Science

Applied Data Science Master Thesis

A View Through Time: a Data Science Approach on the Spectral Alignment of Landsat 5, 7 and 8

First examiner:
Dr. Wiebe Nijland

Second examiner:
Dr. Mathieu Gravey

Candidate:
Jurrian van de Kraats

In cooperation with:
the Faculty of Geosciences Utrecht University and the Austrian Academy of Sciences - Institute for Interdisciplinary Mountain Research

July 28, 2023

Abstract

The Landsat satellite series provide geoscientist with over half a century with sensing data. Differences in sensor properties on these satellites create variations between sensing. A spectral alignment of the Landsat satellite series would be required to properly conduct longitudinal research on the development of the earth's surface. In this thesis four different models, two ordinary least square regressions, a polynomial regression and a single layer neural network regression, were compared in their performance on the transformation of Landsat 8 and Landsat 5 on Landsat 7 in a spatially and non-spatially dependent situation. Furthermore was tested if the azimuth and elevation formed a meaningful contribution to the transformation model. Here we show that in the transformation from Landsat 8 to Landsat 7 and Landsat 5 to Landsat 7 the single layer neural network regression, with the addition of the azimuth and elevation, has the lowest root mean squared error on spatially independent data. On spatially dependent data did the ordinary least squares regression without the azimuth and elevation as predictors contain the lowest root mean squared error. However, on local spatially dependent data did the single layer neural network contain the lowest root mean squared error as well. The difference in performance between a single neural network regression and complex linear regressions could shows that even in its simplest form, the neural network architecture forms a better fit for these transformations. Future research could study the implementation of backwards compatibility models, assessing the structural validity of the model and allowing more complex neural network regressions.

Contents

1. Introduction

The Landsat satellite series contains some of the longest temporal records of satellite captured land-surface observations as of to date. Since its first launch in 1972, eight different satellites of the series provide geoscientist with more than half a century of detailed data about earth surface and its development (Landsat Mission, 2022). However, were the variations between Landsat 8 and 9 seem negligible (Gross et al., 2022), can the same not be said about the differences between these satellites and the earlier iterations of Landsat (Mancino et al., 2020; Roy et al., 2016). Differences in the sensors and sensor quality (Flood, 2014), variations in sensor specificity (Flood, 2014; Roy et al., 2016), differences in data processing (Roy et al., 2016) and differing angles of reflectance caused by the satellite's line of orbit (Li et al., 2010; Roy et al., 2016) are possible causes for variations between sensing. A spectral alignment of the Landsat satellite series would be required to properly conduct longitudinal research on the development of the earths surface (Mancino et al., 2020).

In 2016 Roy et al. performed such an alignment of Landsat 7 and Landsat 8. By using a simple regression model they translated 75 to 80 percent of the variance from Landsat 7 to Landsat 8. However, if an identical model of Landsat 5 to 7 were to be created and chained to explain Landsat 8, the cumulative error would create a translated loss of almost 50 percent of the variance. More complex modelling or a model that translates multiple satellites could create a more accurate transformation.

To create such a model, it must be assumed that the data from both satellites convey the same information. Long temporal intervals could create inconsistencies by means of cloud-cover (Mancino et al., 2020), build-up or the passing of seasons (Roy et al., 2016). Furthermore, factors like the increase in CO_2 fertilisation create noticeable reflectance variations within spectral bands (Lim et al., 2004). These variations can be indistinguishable from the small variations between satellites (Huete et al., 2002), but should not be negated in the alignment process. Hence, any alignment method should entail the lowest possible variation in its temporal frame.

This assumption of low temporal variation requires overlap between the satellites in order to estimate their variations. With the long commission of Landsat 5, do three satellites, Landsat 5, 7 and 8, provide almost 40 years of spectral land-surface data, while also containing periods of overlap (Landsat Mission, 2022). A (nested) model that transforms both satellites to Landsat 8 could remove the cumulative error and align most of the available Landsat data. Unfortunately, there is no data published from the last six months of Landsat 5's commission. These six months were the only period of overlap between Landsat 5 and 8, making an accurate estimation between these satellite impossible. A similar result could in this case be obtained with a transformation from Landsat 5 and Landsat 8 to Landsat 7.

In previous research simple linear regression models were used to align the Landsat satellites (Flood, 2014; Ke et al., 2015; Roy et al., 2016). However, in other fields and applications more complex methods like polynomial regression (Storey et al., 2014; Tommaselli et al., 2015) or neural networks (Isa et al., 2021) were used to perform spectral sensor alignment. Furthermore, previous research has shown that adding variables that explain differences in reflectance, like the azimuth and elevation, could help in creating a more accurate alignment (Nagol et al., 2015; Zhang et al., 2015).

This thesis compared different modelling methods to find the transformation with the lowest root mean squared error from Landsat 5 and Landsat 8 to Landsat 7. It was important that in these models the structural validity is ensured (Wu et al., 2014). So called 'black box models' could base their transformation on other properties of the data, that future data might not contain. This made complex machine learning models not applicable for this study. The models used were an ordinary least squares regression, a polynomial regression and a single layer neural network regression. These models were compared on data with and without spatial dependencies. First the data acquisition and pre-processing will be described, followed by the model creation and their analysis, ending with the results and concluding remarks.

2. Data

2.1 Data Collection

In this thesis Google Earth Engine (Gorelick et al., 2017) was used to collect data from the Landsat Collection 2 Tier 2 datasets (courtesy of the U.S. Geological Survey). Six separate collections were conducted: two collections with random points in the Southern Cone (Figure 2.1), two full sample collections in the Northern Southern Cone and two full sample collections in parts of the Pacific North West. For each data collection were the clouds removed using the quality band and was the near infrared band saved. This was done with the assumption that the optimal model would not differ across bands. The following section will describe the collection method of each dataset.

The first dataset entailed the overlapping data from Landsat 7 and Landsat 8. A random sample of 60,000 30 by 30 meter pixels was created in the Southern Cone (Figure 2.2). For each Landsat 8 image of this area in the time frame of the 5^{th} of May 2015 until the 19^{th} of June 2015 were the Landsat 7 images collected with an overlap of one day or less. These images were reduced to one mean 'Landsat 7 value'. For each pixel with an Landsat 8 and Landsat 7 value, the azimuth and elevation were saved.

The Landsat 5 to Landsat 7 sample was the second dataset. A random sample of 200,000 30 by 30 meter pixels was created in the Southern Cone. This sample is greater than the Landsat 8 to Landsat 7 sample because days with overlap between Landsat 5 and 7 had a lot of cloud cover, reducing the overlapping pixels. For each Landsat 5 image in this area in the time frame of the 1^{st} of April 2000 until the 28^{th} of October 2000 were the Landsat 7 images collected with an overlap of six days or less. This interval is larger than the Landsat 8 to Landsat 7 interval because overlapping images were sparse. The Landsat 7 images were reduced to one mean 'Landsat 7 value' and for each overlapping pixel were the azimuth and elevation saved.

The third dataset was a full sample of the overlapping data between Landsat

Figure 2.1: Map of the Southern Cone (source: Berglee, 2012)

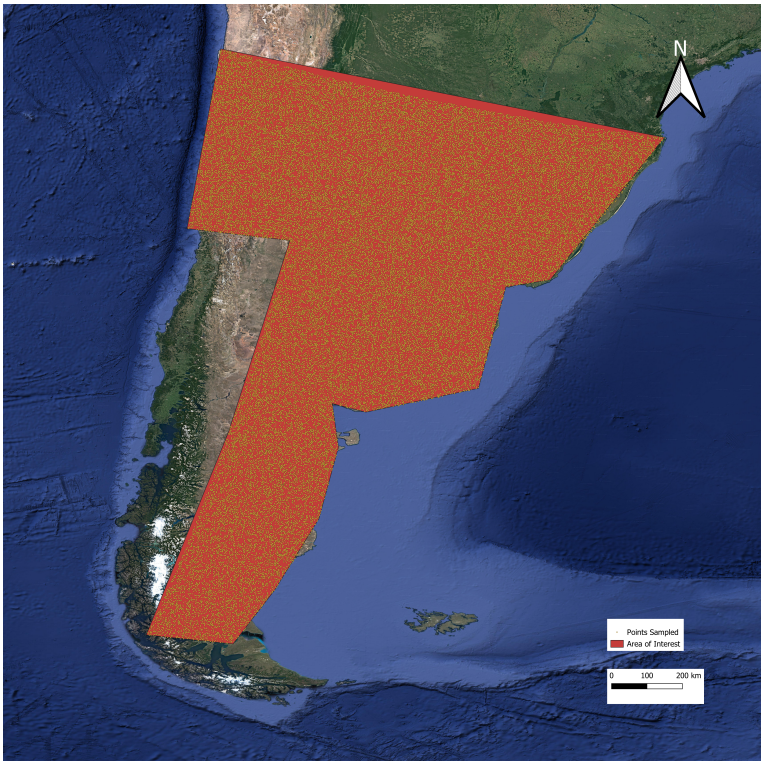

Figure 2.2: Area of interest with the sampled points
(Courtesy of Google Earth and the U.S. Geological Survey)

8 and Landsat 7 in the Northern Southern Cone. For each Landsat 8 image on the 1st of April 2015 were the Landsat 7 images with an overlap of one day or less collected. The Landsat 7 images were reduced to one mean 'Landsat 7 value' and for each overlapping pixel were the azimuth and elevation saved.

The fourth dataset followed the same process in the Pacific North West. The data was collected on the 6th of April 2015 with one day of overlap.

The fifth dataset was a full sample of the overlapping data between Landsat 5 and Landsat 7 in the Northern Southern Cone. For each Landsat 5 image from the 11th of December 2003 until the 1st of January 2004 were the Landsat 7 images with an overlap of six day or less collected and reduced to a mean Landsat 7 value. This data collection is wider than the Landsat 8 to Landsat 7 data collection because of a reduced amount of data due to cloud cover. Furthermore is overlapping data between Landsat 7 and Landsat 5 sparse, making it hard to find a substantial, diverse and overlapping area of interest. The azimuth and elevation of each pixel were saved.

The sixth dataset was collected in the Pacific North West from the 3rd of October 2000 until the 17th of October 2000 with an overlap of six days. The data was collected through the same process as the fifth dataset.

2.2 Data Properties

The Landsat 5 Thematic Mapper sensor is similar to the one used in Landsat 4. It was a newly added sensor beside the multi spectral sensor that could also be found in the first three Landsat satellites. The Thematic Mapper allowed sensing on a spatial resolution of 30 by 30 meters (USGS and EROS Center, Accessed 2023). Due to the unfortunate fail of Landsat 6, was the commission of Landsat 5 extended to more than 20 years.

The Landsat 7 Enhanced Thematic Mapper plus (ETM+) has similar spectral properties as the Thematic Mapper. Most differences come from the addition of new bands and its band processing (USGS and EROS Center, Accessed 2023). The ETM+ sensor forms the last sensor in the Thematic Mapper group.

The Landsat 8 Operational Land Imager (OLI) sensor has narrower spectral bands and improved geometry, calibration and signal to noise characteristics compared to the ETM+ sensor. Landsat 8 follows the same orbit as the decommissioned

Landsat 5 (Roy et al., 2016). Figure 2.3 shows the differences in captured wavelengths between the ETM+ and the OLI sensors.

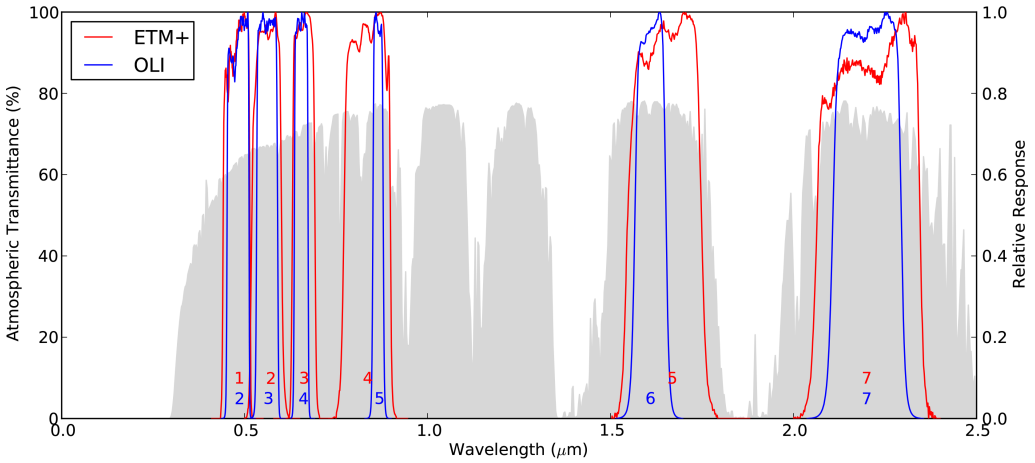

Figure 2.3: Differences between the ETM+ and OLI sensors (Source: Flood, 2014)

2.3 Data Processing

The data was processed in Python. The first and second dataset were split in intervals and the Landsat 8 and 5 data were plotted to the Landsat 7 data (Figure 2.4 and Figure 2.5). Intervals with no resemblance to a linear pattern were deleted with the assumption that this data conveyed different information (for example clouds that were not masked). Artificial outliers due to the data's sensing nature in the minimum and maximum values were deleted. The data were randomly split in a train (80%) and test set.

The third, fourth, fifth and sixth dataset were plotted to see if the data resembled a linear pattern (Figure 2.6). The wider the collection interval, the less linear the data became. This could be due to inaccuracies in the cloud masking and inaccuracies occurring during the collection interval. Artificial outliers due to the data's sensing nature in the minimum and maximum values were deleted. The data from the Northern Southern Cone were used as training data and the data from the Pacific North West was used as test data.

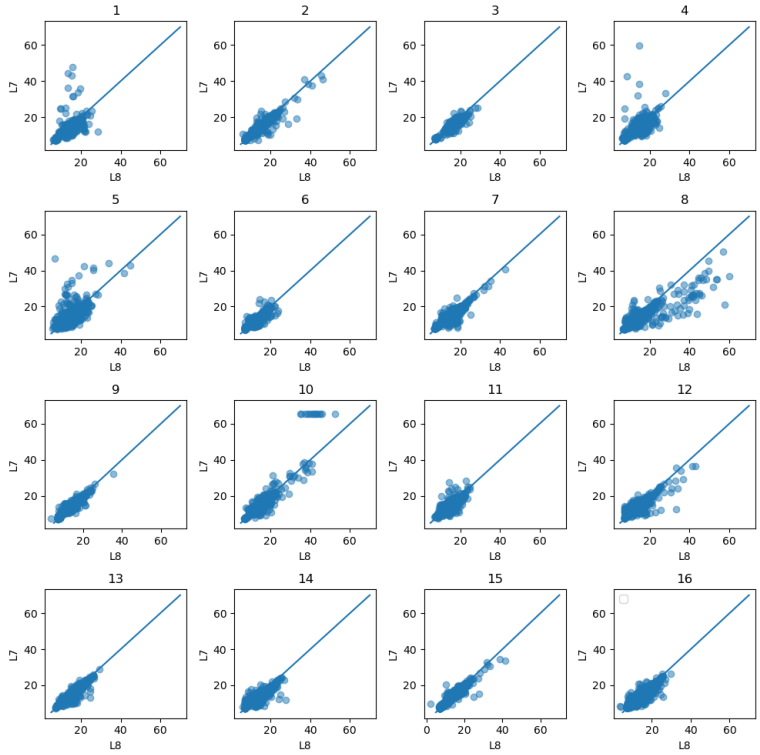

Figure 2.4: The random point sample of Landsat 8 based on download interval, before pre-processing, in thousands

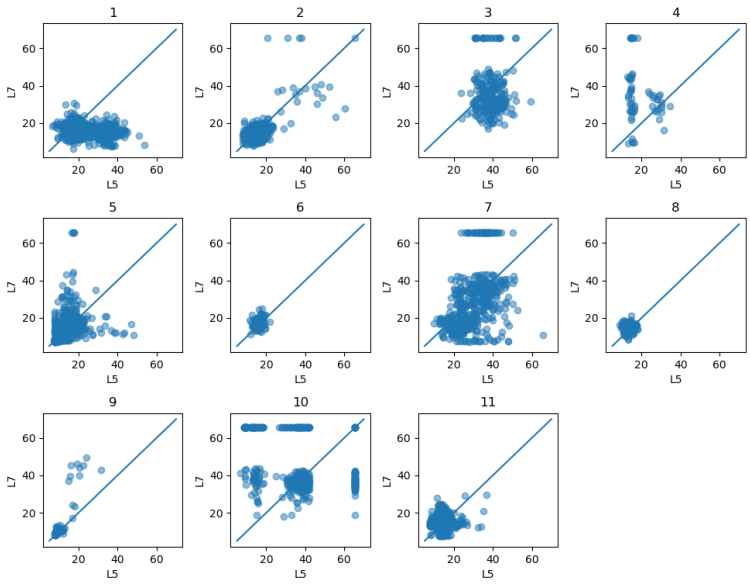

Figure 2.5: The random point sample of Landsat 5 based on download interval, before pre-processing, in thousands

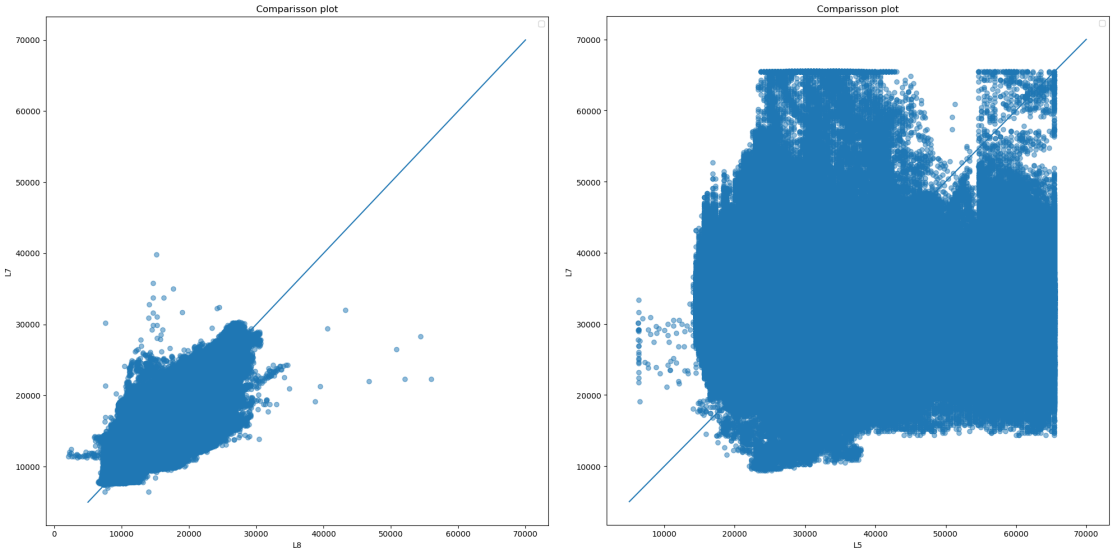

Figure 2.6: The full sample of Landsat 8 and Landsat 5 plotted to Landsat 7, before pre-processing

3. Method

For all four transformations conducted in this thesis, four models were tested. The first model is a benchmark ordinary least squares regression with only the reflectance data. The second model was an ordinary least squares regression with the reflectance data and the azimuth and elevation as predictors. The third model was a polynomial model with the same variables and the last model was a single layer neural network regression. Model calibration and comparison was assessed through the root mean squared error (RMSE; Figure 3.1). The following chapter will describe the modeling process of each of these models across the datasets.

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{\sigma_i}\right)^2} \tag{3.1}$$

3.1 Benchmark regressions

Two benchmark regressions were created to test model performance. The first benchmark regression was the same model used in Roy et al. (2016): a linear regression regressing the reflectance data of one satellite on the other. This model was created to compare the model performance of the other models used in this study to the state of the art. The model was created using python and the OLS function from the Statsmodels module (Seabold and Perktold, 2010).

The second benchmark regression was an ordinary least squares regression with the azimuth and elevation as a predictor. This model was formed to assess if adding these predictors would add substantial value to the regression model. This was determined by the change in the RMSE and if this coincided with a decreasing Akaike Information Criteria (AIC) to account for model complexity (James et al., 2013). The model was fitted using the OLS function from the Statsmodels module (Seabold and Perktold, 2010). Both models were trained on the train dataset and evaluated on the test dataset.

3.2 Polynomial regression

A polynomial regression is a variation of the linear regression model where the linear function is replaced with an polynomial version (Figure 3.2). This allows the model to account for non-linear relationships while still allowing an ordinary least squares regression fit (James et al., 2013).

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \ldots + \beta_n x_i^n + \varepsilon_i \tag{3.2}$$

Ridge regression was used to determine the ideal order of polynomials for each transformation. Ridge regression is the process of step-wise adding variables to a model and using an error metric to determine the best model fit (James et al., 2013). The polynomial transformation function of Scikit-Learn (Pedregosa et al., 2011) was used to create the polynomial structure. The regression model was created using the OLS function from the Statsmodels module (Seabold and Perktold, 2010). For each polynomial, a model was fitted on the train-dataset. The final model was the model with the lowest RMSE on the test-dataset.

3.3 Single Layer Neural Network Regression

A single layer neural network regression contains one neuron activated by a linear regression through the weights of the input variables. In other words is the independent variable transformed through the weights to become the dependent variable. In essence does this not differ from a conventional linear regression without an intercept. However, where the ordinary least squares regression determines the predictor values through the overall lowest error (James et al., 2013), does the single layer neural network regression determine the predictors based on the lowest error value for each tuple (Martín Abadi et al., 2015). This optimisation process is conducted for each tuple of the dataset and re-iterated until the validation error converges.

The Tensorflow Keras Python modules (Martín Abadi et al., 2015) were used to fit the single layer neural network regression. The model normalised the data and split the data in a train (80%) an validation set for each iteration (epoch). The optimum amount of epochs was determined by visualising the validation error and

using the elbow method (James et al., 2013) to determine the epoch of convergence. The model eventual performance was determined on the test dataset.

4. Results

Sixteen models were fitted. Figure 4.1 shows the results of the analysis. These results are explained in the following chapter.

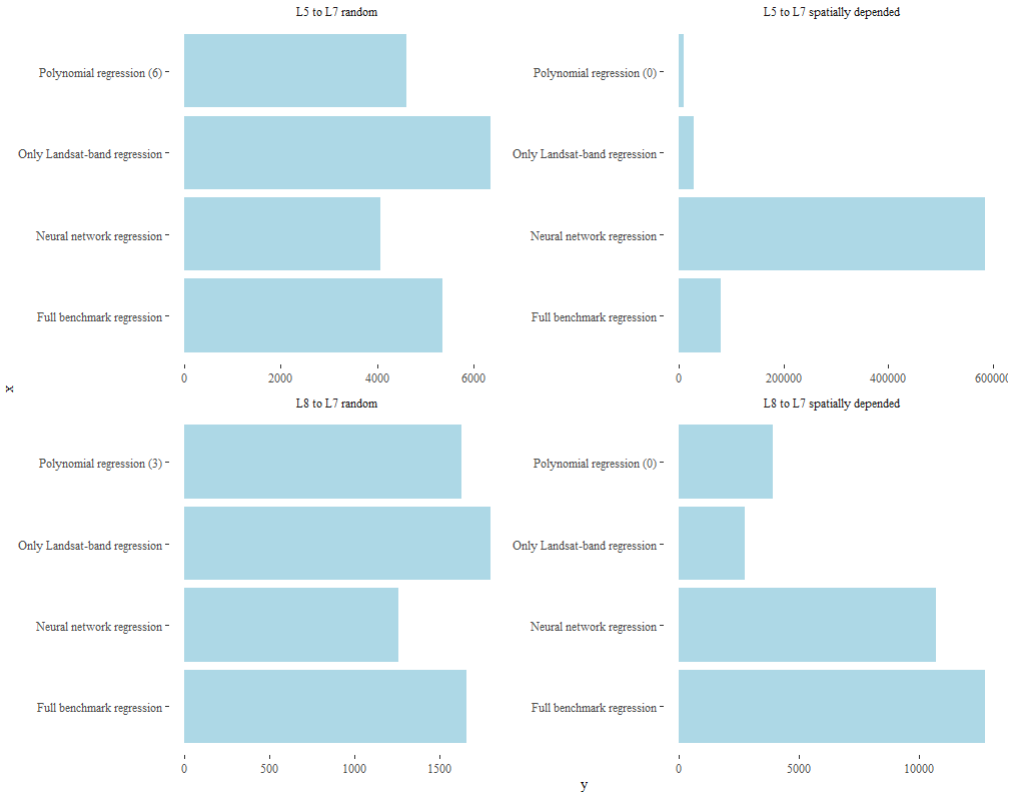

Figure 4.1: The Analysis Results

4.1 Spatially independent analysis

4.1.1 Landsat 8 to Landsat 7

The benchmark regression with only the reflectance data of Landsat 7 regressed on Landsat 8 had a RMSE of 1802 with an AIC of 312600. The regression with the added azimuth and elevation had a RMSE of 1659 with an AIC of 309000. This lower error and lower AIC makes the second model a better model fit.

Ridge regression showed that 3 is the optimum amount of polynomials for this

dataset (Figure 4.2). The RMSE of the polynomial model is 1633. The single layer neural network had a RMSE of 1264 with an epoch of 500 (Figure 4.3).

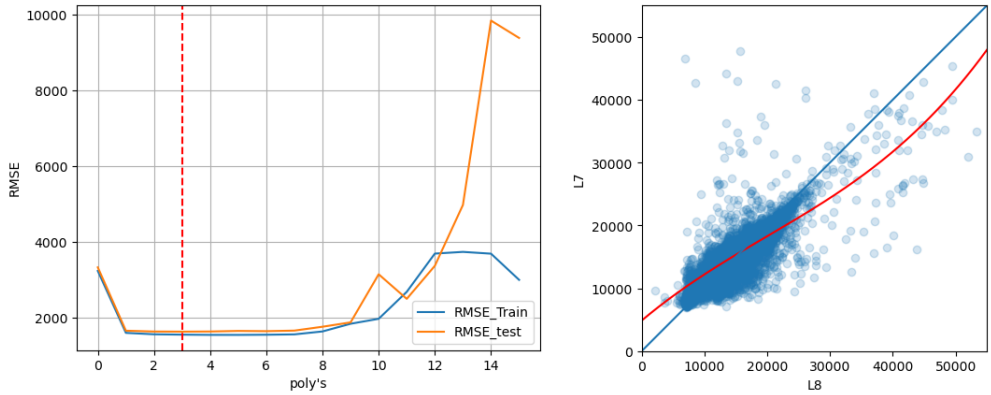

Figure 4.2: Results Polynomial Regression

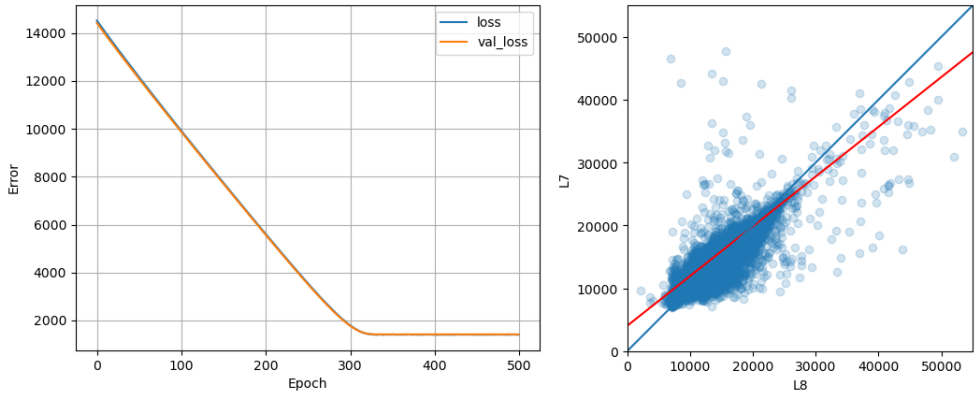

Figure 4.3: Results Single Layer Neural Network Regression

4.1.2 Landsat 5 to Landsat 7

The RMSE of the benchmark regression with only Landsat 7 regressed on Landsat 5 was 6348. The regression with the inclusion of the azimuth and elevation was 5344. The AIC of the second model was lower (131,200 compared to 129,000). This makes the model with all the variables a better model fit.

Ridge regression showed that 6 polynomials forms the optimum structure for this dataset (Figure 4.4). This results in a model with a RMSE of 4609. The single layer neural network regression had a RMSE of 4058 on a epoch of 1500 (Figure 4.5).

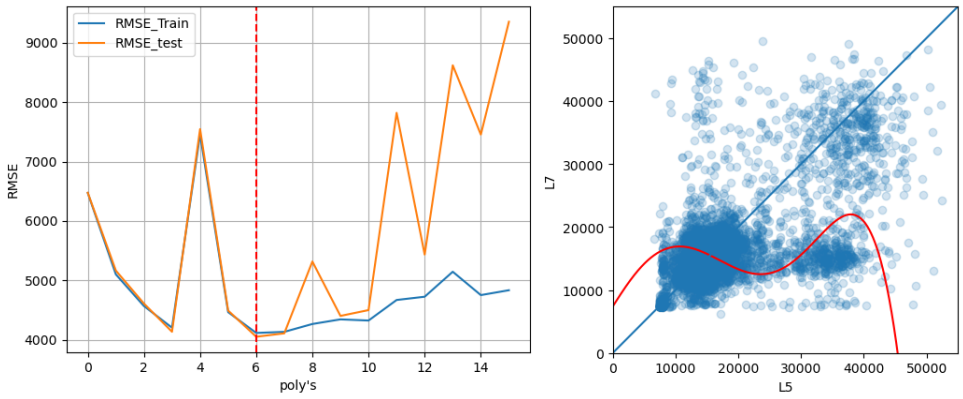

Figure 4.4: Results Polynomial Regression

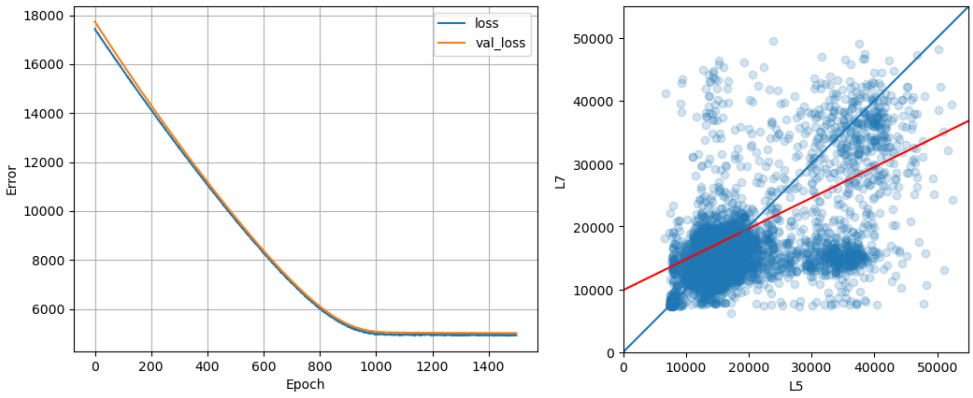

Figure 4.5: Results Single Layer Neural Network Regression

4.2 Spatially dependent analysis

4.2.1 Landsat 8 to Landsat 7

The benchmark regression with only the reflectance data from Landsat 8 and Landsat 7 had a RMSE of 2763. The regression with added azimuth and elevation had a RMSE of 12752. The AIC of the first model was 199,700,000 the AIC of the second model was 198,300,000 making the model with only the reflectance data a better model fit.

Ridge regression showed that 0 polynomials fitted the test data best (Figure 4.6). This model had a RMSE of 3940. The neural network model had a RMSE of 107,290 on 3 epochs, because the error immediately converged (Figure 4.6).

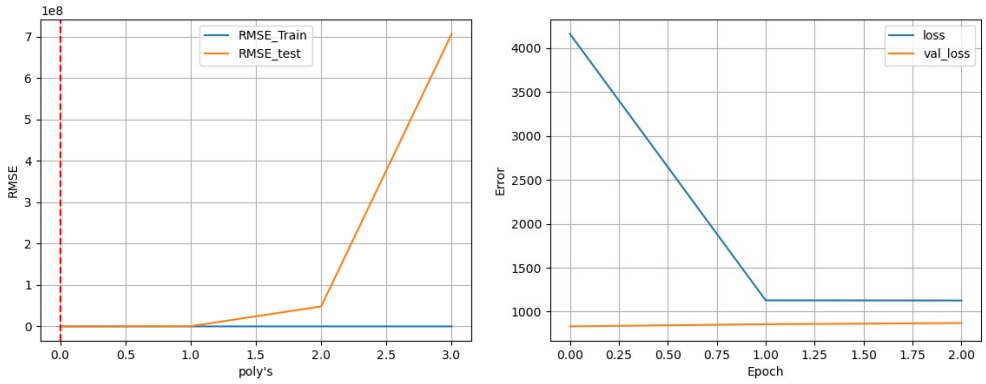

Figure 4.6: The best polynomial and epoch

4.2.2 Landsat 5 to Landsat 7

The RMSE on the benchmark regression with only the reflectance data regressing Landsat 7 on Landsat 5 was 28,244. The RMSE on the model with the azimuth and elevation was 79,996. The AIC on the first model was 53,500,000. The AIC on the second model was 50,160,000. However, the clear difference in the error makes the first model a better fit.

Ridge regression showed that 0 polynomials was optimum fit for this data (Figure 4.7). The RMSE for the polynomial model was 10,580. The single layer neural network regression had a RMSE of 584,799 with an epoch of 10 (Figure 4.7).

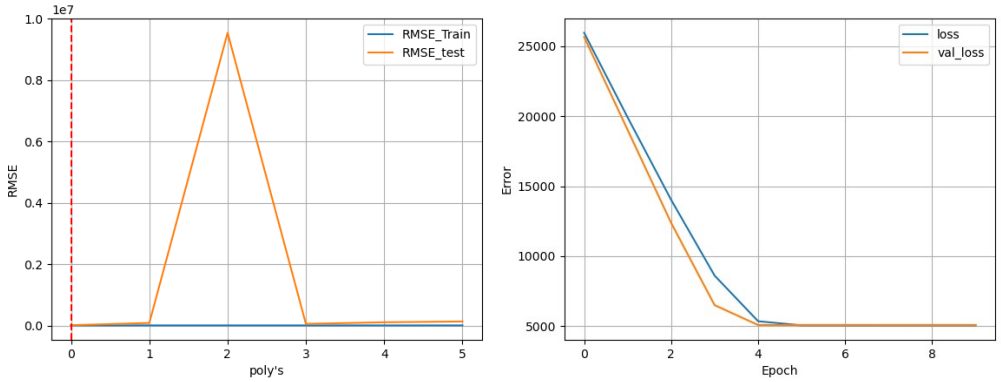

Figure 4.7: The best polynomial and epoch

5. Discussion and Conclusion

5.1 Discussion

Model performance on the spatially dependent models could have been distorted due to over-fitting. The spatially dependent models used the Northern Southern Cone as training data. This small area of interest also has a small variation in its azimuth and elevation. The model could create nonsensical results because it is tested on an azimuth and elevation it has not seen before, because it is from another hemisphere. This assumption of overfitting on the azimuth and elevation is best shown in the second order polynomial model of the Landsat 5 to Landsat 7 transformation. This model is the only spatially depended model with a large, integer predictor value (20.63) on the product of the azimuth and elevation. This results in a large outlier in the error (Figure 4.7) when tested on spatial data from another hemisphere.

The validation data used to train the neural network showed a RMSE of 870 on the Landsat 8 to Landsat 7 transformation and a RMSE of 5073 on the Landsat 5 to Landsat 7 transformation. These are also the lowest validation or train errors in the spatially dependent models. This makes the neural network also better on 'local spatially dependent data'.

Cloud cover had a big impact on the data availability in the spatially independent models and the tidiness of the data collected. The spatially dependent model showed that spatial dependency does not change the model performance in local areas. Future research is needed to determine if the addition of the azimuth and elevation would in this case have a strong influence on the model performance. Furthermore, could future studies use spatially dependent data for the transformation model to combat problems in cloud cover.

A model containing data from all over the world would allow for better a all round application. However, gathering and processing this amount of data to fit a good model would be computationally very heavy and time consuming. In this case would it be more efficient to create smaller transformation models from

different locations on earth.

Furthermore does the single layer neural network regression outperform complex regression methods. This is also the case in less clean data, like the data used in the models transforming Landsat 5 in to Landsat 7, were the data does not resemble a linear pattern. The addition of an extra layer could be a very easy way to make the neural network model even more accurate and account better for some non-linear patterns. However, the addition of extra layers makes the model hard to explain and compromises the structural validity. Future research in the creation of recalibration models, for example a Landsat transformation specific version of the SRM (Lee et al., 2019), would ensure the structural validity and allow for the usage of more complex neural networks.

5.2 Conclusion

In this thesis four different models were compared in their performance on the transformation of Landsat 8 and Landsat 5 on Landsat 7 in a spatially and non-spatially dependent situation. Furthermore was tested if the azimuth and elevation formed a meaningful contribution to the transformation model.

In the non-spatially dependent models does the single layer neural network regression perform best. This model had a lower RMSE by a considerable margin in the both transformations. The polynomial models were the second best fit. There was a difference in the polynomial structure between Landsat 8 and Landsat 5. Furthermore do the benchmark regressions show that the addition of the azimuth and elevation creates a better model fit.

In the spatially dependent models does the benchmark regression with only the Landsat reflectance data perform best. The polynomial models both had the best fit with zero polynomials, effectively giving a mean value. The single layer neural network regression had the highest RMSE on both satellites. However, on local spatially depended data has the single layer neural network regression the lowest RMSE.

Acknowledgements

This thesis would not be possible without the supervision and helping hand of Dr. Mathieu Gravey.

I am also thankful to Dr. Wiebe Nijland for providing important feedback during my defence

Many thanks to Quirine Donckers for helping me when I struggled coding.

Thanks should also go to Lisa Brandenburg for helping me for when I was stuck in the writing process.

Bibliography

Berglee, R. (2012). World regional geography: People, places and globalization.

Flood, N. (2014). Continuity of reflectance data between landsat-7 etm+ and landsat-8 oli, for both top-of-atmosphere and surface reflectance: A study in the australian landscape. *Remote Sensing*, *6*(9), 7952–7970.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2017.06.031

Gross, G., Helder, D., Begeman, C., Leigh, L., Kaewmanee, M., & Shah, R. (2022). Initial cross-calibration of landsat 8 and landsat 9 using the simultaneous underfly event. *Remote Sensing*, *14*(10), 2418.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote sensing of environment*, *83*(1-2), 195–213.

Isa, S. M., Suharjito, Kusuma, G. P., & Cenggoro, T. W. (2021). Supervised conversion from landsat-8 images to sentinel-2 images with deep learning. *European Journal of Remote Sensing*, *54*(1), 182–208.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in r, second edition*. Springer. https://faculty.marshall.usc.edu/gareth-james/ISL/

Ke, Y., Im, J., Lee, J., Gong, H., & Ryu, Y. (2015). Characteristics of landsat 8 oli-derived ndvi by comparison with multiple satellite sensors and in-situ observations. *Remote sensing of environment*, *164*, 298–313.

Landsat Mission, U. S. G. S. (2022). About. https://www.usgs.gov/landsat-missions/about

Lee, H., Kim, H.-E., & Nam, H. (2019). Srm: A style-based recalibration module for convolutional neural networks. *Proceedings of the IEEE/CVF International conference on computer vision*, 1854–1862.

Li, F., Jupp, D. L., Reddy, S., Lymburner, L., Mueller, N., Tan, P., & Islam, A. (2010). An evaluation of the use of atmospheric and brdf correction to standardize landsat data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(3), 257–270.

Lim, C., Kafatos, M., & Megonigal, P. (2004). Correlation between atmospheric co2 concentration and vegetation greenness in north america: Co2 fertilization effect. *Climate Research*, *28*(1), 11–22.

Mancino, G., Ferrara, A., Padula, A., & Nolè, A. (2020). Cross-comparison between landsat 8 (oli) and landsat 7 (etm+) derived vegetation indices in a mediterranean environment. *Remote Sensing*, *12*(2), 291.

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Jia, Y., Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, … Xiaoqiang Zheng. (2015). TensorFlow: Large-scale machine learning on heterogeneous

systems [Software available from tensorflow.org]. https://www.tensorflow.org/

Nagol, J. R., Sexton, J. O., Kim, D.-H., Anand, A., Morton, D., Vermote, E., & Townshend, J. R. (2015). Bidirectional effects in landsat reflectance estimates: Is there a problem to solve? [Global Land Cover Mapping and Monitoring]. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 129–135. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2014.09.006

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Roy, D. P., Kovalskyy, V., Zhang, H., Vermote, E. F., Yan, L., Kumar, S., & Egorov, A. (2016). Characterization of landsat-7 to landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote sensing of Environment*, *185*, 57–70.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 operational land imager on-orbit geometric calibration and performance. *Remote sensing*, *6*(11), 11127–11152.

Tommaselli, A., Oliveira, R. A., Nagai, L. Y., Imai, N. N., Miyoshi, G. T., Honkavaara, E., & Hakala, T. (2015). Assessment of bands coregistration of a light-weight spectral frame camera for uav. *GeoUAV-ISPRS Geospat. Week*, *192*.

USGS and EROS Center. (Accessed 2023). USGS EROS Archive - Landsat Archives: Landsat 4-5 Thematic Mapper (TM) Level 1 Data.

Wu, W., Dandy, G. C., & Maier, H. R. (2014). Protocol for developing ann models and its application to the assessment of the quality of the ann model development process in drinking water quality modelling. *Environmental Modelling & Software*, *54*, 108–127.

Zhang, H. K., Roy, D. P., & Kovalskyy, V. (2015). Optimal solar geometry definition for global long-term landsat time-series bidirectional reflectance normalization. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(3), 1410–1418.

Appendices

A. Code

For this thesis Google Earth Engine and Python were used. The following GitHub link will have all the code files and results notebook. In the REAMDE are the links to the Google Earth enigne code

https://github.com/JJkriskras/ADS_thesis_code

B. Model results

A more detailed version of the results can be accessed in the notebook on GitHub (see Appendix A)

B.1 Spatially independent analysis

B.1.1 benchmark regression L8 to L7 with only Landsat data

Dep. Variable:	L7	**R-squared (uncentered):**	0.985
Model:	OLS	**Adj. R-squared (uncentered):**	0.985
Method:	Least Squares	**F-statistic:**	1.159e+06
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	17:50:01	**Log-Likelihood:**	-1.5630e+05
No. Observations:	17558	**AIC:**	3.126e+05
Df Residuals:	17557	**BIC:**	3.126e+05
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L8	0.9668	0.001	1076.425	0.000	0.965	0.969

Omnibus:	8462.274	**Durbin-Watson:**	1.979
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	4203100.077
Skew:	0.975	**Prob(JB):**	0.00
Kurtosis:	78.772	**Cond. No.**	1.00

B.1.2 benchmark regression L8 to L7 with azimuth and elevation

Dep. Variable:	L7	**R-squared (uncentered):**	0.988
Model:	OLS	**Adj. R-squared (uncentered):**	0.988
Method:	Least Squares	**F-statistic:**	4.751e+05
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	17:50:01	**Log-Likelihood:**	-1.5450e+05
No. Observations:	17558	**AIC:**	3.090e+05
Df Residuals:	17555	**BIC:**	3.090e+05
Df Model:	3		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L8	0.7583	0.003	222.832	0.000	0.752	0.765
SA	71.2263	2.110	33.763	0.000	67.091	75.361
SE	26.5786	2.219	11.976	0.000	22.229	30.929

Omnibus:	15268.512	**Durbin-Watson:**	2.002
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	4947082.759
Skew:	3.314	**Prob(JB):**	0.00
Kurtosis:	84.965	**Cond. No.**	3.51e+03

B.1.3 best polynomial regression L8 to L7

poly = 0, RMSE train = 3237.480984150923, RMSE test = 3331.4354458665084

poly = 1, RMSE train = 1600.9305804834112, RMSE test = 1657.9085971432862

poly = 2, RMSE train = 1563.5277874805377, RMSE test = 1636.344130424353

poly = 3, RMSE train = 1554.0037265496317, RMSE test = 1633.1000066800993

poly = 4, RMSE train = 1547.821140292814, RMSE test = 1637.1281387208548

poly = 5, RMSE train = 1547.36771455646, RMSE test = 1652.1908289678972

poly = 6, RMSE train = 1551.6946317611937, RMSE test = 1645.9695703221666

poly = 7, RMSE train = 1561.3355413176848, RMSE test = 1660.7800144719188

poly = 8, RMSE train = 1638.5621952590307, RMSE test = 1762.816762511343

poly = 9, RMSE train = 1841.6285096008864, RMSE test = 1877.9583481096445

poly = 10, RMSE train = 1972.5607309279496, RMSE test = 3148.721902945168

poly = 11, RMSE train = 2685.373428730538, RMSE test = 2499.5832890474685

poly = 12, RMSE train = 3695.5919585157535, RMSE test = 3372.4646471983174

poly = 13, RMSE train = 3740.3125318683065, RMSE test = 4980.696438519493

poly = 14, RMSE train = 3693.3305443966797, RMSE test = 9850.908343258436

poly = 15, RMSE train = 3001.23650867341, RMSE test = 9392.384681776468

B.1.4 single layer neural network regression L8 to L7

Kernel results:

array([[2872.5686], [-66.48658], [229.00797]])

B.1.5 benchmark regression L5 to L7 with only Landsat data

Dep. Variable:	L7	**R-squared (uncentered):**	0.874
Model:	OLS	**Adj. R-squared (uncentered):**	0.874
Method:	Least Squares	**F-statistic:**	4.492e+04
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	18:21:05	**Log-Likelihood:**	-65591.
No. Observations:	6458	**AIC:**	1.312e+05
Df Residuals:	6457	**BIC:**	1.312e+05
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L5	0.8459	0.004	211.945	0.000	0.838	0.854

Omnibus:	1038.953	**Durbin-Watson:**	1.873
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	9444.846
Skew:	-0.494	**Prob(JB):**	0.00
Kurtosis:	8.842	**Cond. No.**	1.00

B.1.6 benchmark regression L8 to L7 with azimuth and elevation

Dep. Variable:	L7	**R-squared (uncentered):**	0.911
Model:	OLS	**Adj. R-squared (uncentered):**	0.911
Method:	Least Squares	**F-statistic:**	2.198e+04
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	18:21:05	**Log-Likelihood:**	-64482.
No. Observations:	6458	**AIC:**	1.290e+05
Df Residuals:	6455	**BIC:**	1.290e+05
Df Model:	3		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L5	0.5419	0.008	65.510	0.000	0.526	0.558
SA	63.5509	16.629	3.822	0.000	30.953	96.149
SE	85.8864	19.698	4.360	0.000	47.271	124.502

Omnibus:	1989.302	**Durbin-Watson:**	1.994
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	13152.433
Skew:	1.308	**Prob(JB):**	0.00
Kurtosis:	9.483	**Cond. No.**	7.64e+03

B.1.7 best polynomial regression L5 to L7

poly = 0, RMSE train = 6472.485985373205, RMSE test = 6476.321181115989

poly = 1, RMSE train = 5102.753696523195, RMSE test = 5166.521051078901

poly = 2, RMSE train = 4570.3049611139695, RMSE test = 4612.000178619161

poly = 3, RMSE train = 4206.309242114672, RMSE test = 4137.493351926059

poly = 4, RMSE train = 7458.176129714716, RMSE test = 7545.894383066135

poly = 5, RMSE train = 4466.963592013505, RMSE test = 4493.795302143866

poly = 6, RMSE train = 4117.612227068417, RMSE test = 4053.5183660590537

poly = 7, RMSE train = 4135.960342588092, RMSE test = 4111.187545786148

poly = 8, RMSE train = 4268.0712618432135, RMSE test = 5318.975807655074

poly = 9, RMSE train = 4345.930615677909, RMSE test = 4403.9876287140705

poly = 10, RMSE train = 4326.280910780393, RMSE test = 4501.967049811148

poly = 11, RMSE train = 4671.924767299623, RMSE test = 7819.978267875041

poly = 12, RMSE train = 4726.83510332011, RMSE test = 5435.319420807169

poly = 13, RMSE train = 5147.17565435278, RMSE test = 8619.660125911589

poly = 14, RMSE train = 4754.193130209675, RMSE test = 7454.457743950646

poly = 15, RMSE train = 4836.273218452313, RMSE test = 9352.037380962503

B.1.8 single layer neural network regression L5 to L7

Kernel results:

array([[4037.1533], [-1347.053],[1815.966]])

B.2 Spatially dependent results

B.2.1 Benchmark regression with only Landsat data L8 to L7

Dep. Variable:	L7	**R-squared (uncentered):**	0.993
Model:	OLS	**Adj. R-squared (uncentered):**	0.993
Method:	Least Squares	**F-statistic:**	1.764e+09
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	19:03:44	**Log-Likelihood:**	-9.9848e+07
No. Observations:	11728517	**AIC:**	1.997e+08
Df Residuals:	11728516	**BIC:**	1.997e+08
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L8	0.9537	2.27e-05	4.2e+04	0.000	0.954	0.954

Omnibus:	1575425.208	**Durbin-Watson:**	0.127
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	9703226.575
Skew:	-0.494	**Prob(JB):**	0.00
Kurtosis:	7.345	**Cond. No.**	1.00

B.2.2 Benchmark regression with azimuth and elevation data L8 to L7

Dep. Variable:	L7	**R-squared (uncentered):**	0.994
Model:	OLS	**Adj. R-squared (uncentered):**	0.994
Method:	Least Squares	**F-statistic:**	6.609e+08
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	19:03:46	**Log-Likelihood:**	-9.9166e+07
No. Observations:	11728517	**AIC:**	1.983e+08
Df Residuals:	11728514	**BIC:**	1.983e+08
Df Model:	3		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L8	0.7971	0.000	6029.355	0.000	0.797	0.797
SA	119.2030	0.256	465.966	0.000	118.702	119.704
SE	-54.0901	0.277	-194.988	0.000	-54.634	-53.546

Omnibus:	1227493.940	**Durbin-Watson:**	0.104
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	7082481.204
Skew:	-0.342	**Prob(JB):**	0.00
Kurtosis:	6.745	**Cond. No.**	1.75e+04

B.2.3 best polynomial regression L8 to L7

poly = 0, RMSE train = 2342.552811029791, RMSE test = 3939.5960603785575
poly = 1, RMSE train = 1134.9192748910957, RMSE test = 474687.19361656654
poly = 2, RMSE train = 1108.2462793100044, RMSE test = 48217086.92868242
poly = 3, RMSE train = 1106.7205788835881, RMSE test = 707345631.8799596

B.2.4 single layer neural network regression L8 to L7

Kernel results:

array([[2037.2625], [-172.52896], [8.32545]])

B.2.5 Benchmark regression with azimuth and elevation data L5 to L7

Dep. Variable:	L7	**R-squared (uncentered):**	0.871
Model:	OLS	**Adj. R-squared (uncentered):**	0.871
Method:	Least Squares	**F-statistic:**	1.684e+07
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	19:10:26	**Log-Likelihood:**	-2.6748e+07
No. Observations:	2501170	**AIC:**	5.350e+07
Df Residuals:	2501169	**BIC:**	5.350e+07
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L5	0.7541	0.000	4103.088	0.000	0.754	0.754

Omnibus:	3476.917	**Durbin-Watson:**	0.017
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	3835.314
Skew:	0.055	**Prob(JB):**	0.00
Kurtosis:	3.157	**Cond. No.**	1.00

B.2.6 Benchmark regression with azimuth and elevation data L5 to L7

Dep. Variable:	L7	**R-squared:**	0.132
Model:	OLS	**Adj. R-squared:**	0.132
Method:	Least Squares	**F-statistic:**	1.901e+05
Date:	Fri, 28 Jul 2023	**Prob (F-statistic):**	0.00
Time:	19:10:27	**Log-Likelihood:**	-2.5078e+07
No. Observations:	2501170	**AIC:**	5.016e+07
Df Residuals:	2501167	**BIC:**	5.016e+07
Df Model:	2		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
L5	-0.0369	0.000	-117.477	0.000	-0.038	-0.036
SA	306.3343	0.212	1445.174	0.000	305.919	306.750
SE	-1136.6851	1.877	-605.662	0.000	-1140.363	-1133.007

Omnibus:	432219.596	**Durbin-Watson:**	0.027
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1205545.479
Skew:	0.930	**Prob(JB):**	0.00
Kurtosis:	5.848	**Cond. No.**	2.00e+04

B.2.7 Best polynomial model L5 to L7

poly = 0, RMSE train = 5873.802242045138, RMSE test = 10579.748674604894
poly = 1, RMSE train = 5472.494897272348, RMSE test = 84408.315174281
poly = 2, RMSE train = 5227.759611448572, RMSE test = 9543521.339881321
poly = 3, RMSE train = 5204.106205045842, RMSE test = 54356.65510076382
poly = 4, RMSE train = 5201.477334748991, RMSE test = 106392.56186232778
poly = 5, RMSE train = 5201.417321617655, RMSE test = 132444.1627951805

B.2.8 Single Layer Neural Network L5 to L7

Kernel results:

array(array([[-335.43805],[1197.2877],[-813.3402]])