Thermal sharpening using a linear regression model

Wagenmakers, Collin I.

Award date:
2018

Thermal Sharpening Using a Linear Regression Model

Bachelor End Project

Collin Wagenmakers

April 12, 2018

Abstract

Thermal infrared imagery acquired via remote sensing provide essential information in agriculture and water management. However, with present-day sensors high spatial resolution ($\sim 30m$) images are not regularly available. Sensors currently provide either low spatial resolution ($\sim 300m$) imagery at a high temporal frequency or high resolution imagery at a low temporal frequency. In this research a model to sharpen imagery to higher spatial resolutions is examined with data of an arable area in South Africa. The model uses multiple linear regression to predict the higher resolution imagery. The model assumes there is a unique and sufficiently linear relation between thermal and non-thermal imagery. The latter of which is available frequently and at high resolution. A low resolution VIIRS image is used for the prediction and this prediction is evaluated by a high resolution Landsat image of the same area. Both raw non-thermal data as well as transformed non-thermal data is used. The results show that the transformed data outperforms the raw data with mean squared errors (MSE) of $4.7135K^2$ and $6.7382K^2$ respectively. These results show that transformations of non-thermal data that increase the collinearity with the thermal data have the potential to improve the spatial resolution of thermal imagery. Further research is necessary to decrease the issues of collinearity between independent variables. Additionally, the model could be further developed by adding unique variables that have a strong linear correlation with the thermal data.

1 Preface

This research has been executed as a thesis project for the Applied Mathematics bachelor program of the TU/e. The assignment was commissioned by eLEAF, a company based in Wageningen, the Netherlands, which provides data and applications mainly on agriculture and water management. The aim of this research is to improve the algorithms used in the eLEAF applications. These applications include the support of sustainable water use, increase of food production and protection of environmental systems.

All satellite-based data used in this research was provided by eLEAF. The data represents reflectances of several ranges of the electromagnetic spectrum and land surface temperatures; all of the same area in South Africa.

Contents

2 List of abbreviations

The abbreviations in alphabetical order.

ET	Evapotranspiration
GVMI	Global vegetation moisture index
LSC	Land surface characteristic
LST	Land surface temperature
LWCI	Leaf water content index
MSI	Moisture stress index
NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index
NIR	Near-infrared
PSRI	Plant senescence reflectance index
RGB	Red-green-blue
SIMI	Shortwave infrared moisture index
SMC	Soil moisture content
SWCI	Surface water capacity index
SWIR	Shortwave infrared
TIR	Thermal-infrared
VIS	Visible
VSDI	Visible and shortwave infrared drought index

3 Introduction

3.1 Remote Sensing

Remote sensing is described as the acquisition of information about objects or phenomena at a non-intrusive distance from those objects or phenomena. The term most commonly refers to earth observations, either via satellite or aircraft. This paper will mostly focus on multi-spectral images taken with satellite-based sensor technologies. More specifically, sensors of the visible and near-infrared (VIS/NIR) spectrum, sensors of the shortwave infrared (SWIR) spectrum and sensors of the thermal-infrared (TIR) spectrum. Imagery from each spectrum is used to improve agricultural applications.

TIR remote sensing is the primary tool used for acquiring a land surface temperature (LST) of a geological surface area. LSTs provide valuable information for estimating soil moisture content (SMC) and evapotranspiration (ET). Up-to-date data on SMC and ET is of crucial importance in agricultural industries to enhance crop production and sustainable water management.

The issue, however, is that current sensors either lack the spatial resolution or temporal frequency to provide sufficiently useful information. Typically, satellites with a higher image resolution have a lower fly over rate and vice versa. As low resolution thermal data is available more regularly, naturally the question arises if high resolution LSTs can be accurately estimated using low resolution LSTs.

4 Theory

4.1 Land Surface Characteristics

To find so-called thermal sharpening methods, prior research has modeled relationships between other land surface characteristics (LSCs) and LSTs. Many of these LSCs are determined by non-TIR remote sensing. Imagery of non-TIR spectral ranges are generally available at higher resolution and higher temporal frequency. Consequently, accurate relationships between LSCs and LSTs are of invaluable importance. The next subsections will be dedicated to describing and defining the LSCs that are used in this research.

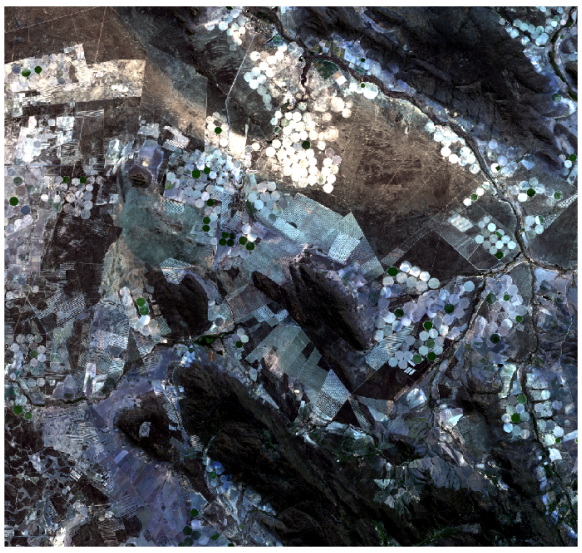


Figure 1: RGB image of area of research

Figure 1 represents the area in South Africa all data refers to in red-green-blue (RGB) colors. To get an indication of how the LSCs are correlated to the LST of this area, graphs of the provided data are added. These graphs can be seen in Figures 2-11. The data with which the LSCs are calculated can be found in Appendix A which contains the 13 spectral bands of satellite Sentinel-2.

Due to the way the concerning LSCs are defined, they all present a ratio of reflectance. As such, the LSCs discussed in this research are all dimensionless.

NDVI

A popular thermal sharpening method in recent years is TsHARP, a model that finds a relationship between the normalized difference vegetation index (NDVI) and the LST. The NDVI is an indicator that is used to analyze how much vegetation a surface area contains (Gandhi et al., 2015). This indicator is determined by reflectance from the VIS/NIR spectrum as follows:

$$NDVI = \frac{R_{NIR} - R_{RED}}{R_{NIR} + R_{RED}}. \tag{1}$$

Here, R_{NIR} and R_{RED} stand for the value of the NIR and red spectral reflectance respectively. As the spectral reflectance is a ratio of the reflected radiation over the incoming radiation, the values of NIR and R lie between 0 and 1. Consequently, the NDVI ranges between -1 and 1.

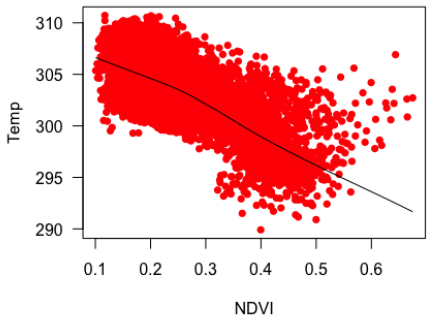


Figure 2: Temperature (in K) versus NDVI

Figure 2 shows that within the data a higher NDVI generally translates to a lower temperature. NDVI values around 0 (between -0.1 and 0.1) generally indicate a barren land surface. Higher positive values indicate thicker vegetation.

Albedo

Albedo is the ratio of reflected solar radiation out of the total incoming solar radiation (Coakley, 2003). Whereas most LSCs are defined by a spectral reflectances of a specific range of the electromagnetic spectrum, Albedo is defined by the reflectance of the entire electromagnetic spectrum. As such, Albedo is defined as a ratio with values between 0 and 1. An Albedo value of 0 indicates all light is absorbed, while a value of 1 indicates all light is reflected.

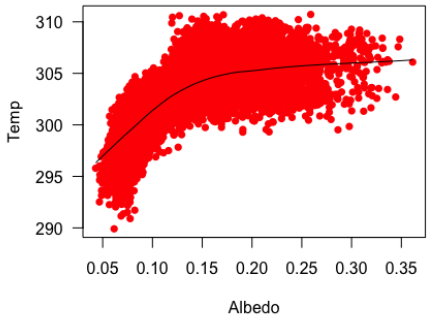


Figure 3: Temperature (in K) versus Albedo

Figure 3 captures a positive relation between Albedo and temperature of the provided data. As can be seen, the Albedo values range between 0.05 and 0.35, which indicate surface types like grass, forest and barren lands. Generally Albedo values between 0.5 and 0.9 indicate very high reflectance, only possible on surface types as ice or snow.

PSRI

The plant senescence reflectance index (PSRI) is used to detect stress or aging in plants (Merzlyak et al., 1999). Similarly to NDVI, this indicator is also determined by reflectance of the VIS/NIR spectrum. PSRI is determined as follows:

$$PSRI = \frac{R_{RED} - R_{BLUE}}{R_{VRE5}}. \tag{2}$$

R_{BLUE} and R_{VRE5} represent values of the blue and vegetation red edge (band 5) spectral reflectance respectively. As there are multiple wavelength ranges in the vegetation red edge spectrum, clarification is given that it concerns band 5 of Sentinel-2. The vegetation red edge spectral range is a specific range of the NIR spectrum for which healthy vegetation has high reflectance. Consequently, higher PSRI values indicate distress in the vegetation.

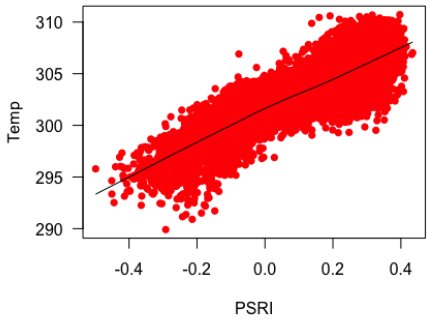


Figure 4: Temperature (in K) versus PSRI

As can be seen in Figure 4, higher temperatures correlate with higher PSRI in a rather linear relation. Healthy vegetation is well-hydrated, and thus generally more cool.

LWCI

The leaf water content index (LWCI) has similar applications as the PSRI. Both the LWCI and the PSRI are used to monitor vegetation health. More specifically, LWCI is used to monitor the relative water content of vegetation (Hunt Jr and Rock, 1989).

$$LWCI = \frac{-ln[1 - (R_{NNIR} - R_{SWIR11})]}{-ln[1 - (R_{NNIR} - 0.2)]}. \tag{3}$$

R_{NNIR} and R_{SWIR11} represent values of the narrow NIR and SWIR (band11) spectral reflectance respectively. Narrow NIR is a spectral range within the NIR spectrum used for classifying vegetation. SWIR reflectance is used for measuring SMC.

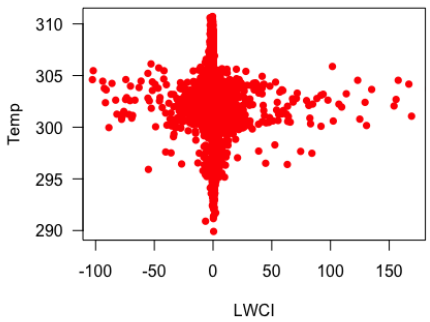


Figure 5: Temperature (in K) versus LWCI

Figure 5 shows that the LWCI for the extreme values in the temperature data set appears to approach zero. No clear relation with the LST is to be seen as of yet. Due to the logarithmic nature of the formula, the values of the LWCI have a very broad range.

MSI

The moisture stress index (MSI) is used, besides vegetation health monitoring, for fire hazard and drought analysis (Hunt Jr and Rock, 1989). The typical range of the MSI is between values of 0 and 3. Vegetation with an MSI value between 0.4 and 2 generally indicate green, healthy vegetation.

$$MSI = \frac{R_{SWIR11}}{R_{NNIR}}. \tag{4}$$

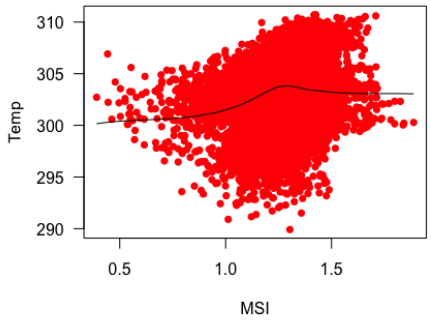


Figure 6: Temperature (in K) versus MSI

In Figure 6 can be seen that no clear relation between the LST and MSI is observable. It is noteworthy that all data falls in the green vegetation range of 0.4 to 2.

GVMI

The global vegetation moisture index (GVMI) is an indicator used to retrieve the water content of vegetation reliably. While the NDVI also provides information about the greenness of vegetation, the GVMI solely focuses on the quantity water in vegetation (Ceccato et al., 2002). A higher GVMI indicates higher water content in vegetation.

$$GVMI = \frac{(R_{NNIR} + 0.1) - (R_{SWIR11} + 0.02)}{(R_{NNIR} + 0.1) + (R_{SWIR11} + 0.02)}. \tag{5}$$

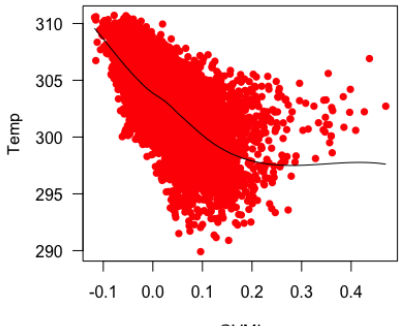


Figure 7: Temperature (in K) versus GVMI

As seen in Figure 7, there appears to be an inverse correlation between GVMI and temperature. This is unsurprising, as higher water content would logically relate to a lower temperature. However, from the data it is unclear if this inverse relation continues to be valid for GVMI values higher than 0.2.

NDWI

The normalized difference water index (NDWI) is used for detecting water bodies. Whereas the NDVI gives an indication of the amount of vegetation in an area, the NDWI gives an indication of how much water is in an area (Hardisky et al., 1983). NDWI values over 0.5 represent water bodies, while vegetation and soil have much smaller values, making it easy to distinguish the water bodies.

$$NDWI = \frac{R_{NNIR} - R_{SWIR11}}{R_{NNIR} + R_{SWIR11}}. \tag{6}$$

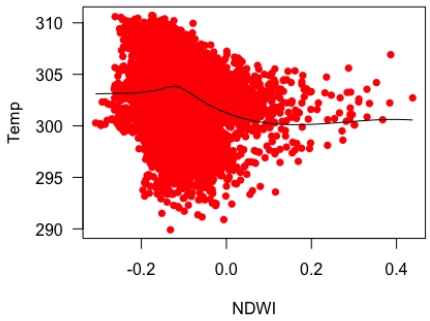


Figure 8: Temperature (in K) versus NDWI

No sufficiently large water bodies are detected in the data set. Consequently, the NDWI values are below 0.5 for the data set. In Figure 8, a clear relation between NDWI and the temperature cannot be directly determined.

VSDI

The visible and shortwave infrared drought index (VSDI) is used for monitoring moisture content for both vegetation and soil using SWIR and VIS spectral ranges. As stated by Zhang et al. (2013) "VSDI is theoretically based on the difference between moisture-sensitive bands (SWIR and red) and moisture reference band (blue)" (p. 4585). Generally, VSDI values between 0 and 1 indicate soil, vegetation or a combination, with higher values indicating higher moisture content. Values over 1 indicate bodies with high moisture content like water, snow or clouds. The formula for the VSDI is as follows:

$$VSDI = 1 - [(R_{SWIR11} - R_{BLUE}) + (R_{RED} - R_{BLUE})]. \tag{7}$$

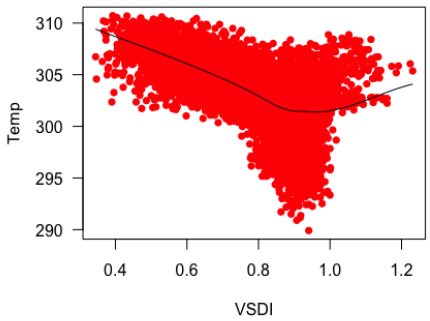


Figure 9: Temperature (in K) versus VSDI

No clear conclusion can be drawn from Figure 9. For the VSDI values between 0.4 and 1.0 it appears that the VSDI has a negative correlation with the temperature. However, it can be seen that there are VSDI values of over 1.0 that do not align with this conclusion. As the data from the NDWI concluded there are no large water bodies, therefore this is most probably an indication of clouds over the research area.

SWCI

The surface water capacity index (SWCI) is used for the monitoring of SMC. SWCI is used to monitor large areas where the vegetation coverage stays mainly consistent over time (Chen et al., 2009). Higher values indicate higher moisture content. The formula for SWCI is as follows:

$$SWCI = \frac{R_{SWIR11} - R_{SWIR12}}{R_{SWIR11} + R_{SWIR12}}. \tag{8}$$

The SWCI uses two different SWIR reflectance ranges. Whereas previously discussed LSCs used only band 11 of the Sentinel-2 SWIR range, the SWCI uses both band 11 and band 12.

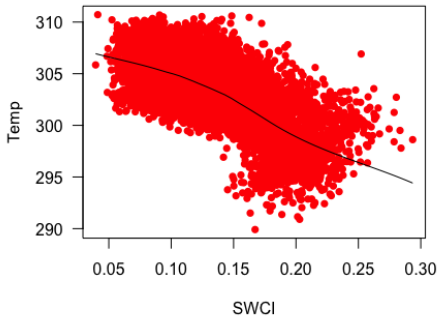


Figure 10: Temperature (in K) versus SWCI

The relation that can be seen in Figure 10 is in form very similar to the relation found for NDVI in Figure 2. Figure 2 and 10 indicate negative correlations with temperature, both of which are approximately linear relationships.

SIMI

The shortwave infrared moisture index (SIMI) makes use of reflectances in the SWIR spectrum to monitor SMC (Yao et al., 2011). Similarly to the SWCI, the SIMI uses two different ranges of the SWIR spectrum (band 11 and band 12). Unlike the SWCI, the SIMI has a negative correlation with SMC. The formula for the SIMI is as follows:

$$SIMI = \sqrt{\frac{R^2_{SWIR11} + R^2_{SWIR12}}{2}}. \tag{9}$$

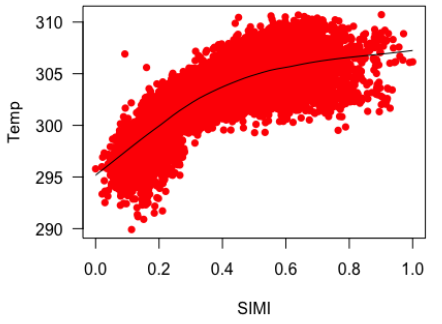


Figure 11: Temperature (in K) versus SIMI

Figure 11 shows a positive correlation between the SIMI and temperature. The relation appears to be of square root nature.

5 Method

5.1 Multiple Linear Regression

A statistical-based thermal sharpening tool, like TsHARP, is based on the assumption that there is a unique relationship between land surface characteristics (LSC) and thermal data (LST). For instance, the TsHARP algorithm uses NDVI to find such a relationship. By a similar approach, that relationship could be expanded by adding more LSCs to the equation.

Using linear regression analysis such a relationship can be found between LSCs and an available low resolution thermal image. Data of the LSCs should be transformed to the coarser resolution size of the LST, such that:

$$LST_{low} = f(LSC_{low}) + \epsilon. \tag{10}$$

Here, ϵ represents an error variable, which is assumed to be independent and identically distributed for all data. The function f is a depiction of the desired relationship.

For a single data point $\{y_i, X_1, X_2, ..., X_r\}$, with $y_i = LST_{low,i}$ as dependent variable and $X_1, X_2, ..., X_r$ as the r independent variables (or regressors) of the function f, a linear function model can be determined. By choosing a multiple linear regression model, the function for LSC then takes the form:

$$f(LSC) = \beta_0 1 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_r X_r = x^T \cdot \beta. \tag{11}$$

To the regressor vector $\bar{X}$ a constant is added. Then it follows that β is an $(r+1)$-dimensional parameter vector, with r the number of independent variables (excluding the constant) on which LST_{low} is dependent. Then x and β take the form:

$$x = \begin{pmatrix} 1 \\ X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_r \end{pmatrix},$$

such that:

$$y_i = x_i^T \cdot \beta + \epsilon_i. \tag{12}$$

An image consists of a data set of p rows and q columns. With a total of $n = pq$ data points of the form $\{y_i, X_{i,1}, X_{i,2}, ..., X_{i,r}\}_{i=1}^n$, it then follows that the model given in (3) can be expanded for the complete data set such that:

$$Y = X \cdot \beta + \epsilon, \tag{13}$$

with:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{pmatrix} = \begin{pmatrix} 1 & X_{1,1} & \cdots & X_{1,r} \\ 1 & X_{2,1} & \cdots & X_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n,1} & \cdots & X_{n,r} \end{pmatrix} \text{ and } \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}.$$

Now, a multiple linear regression analysis can be done to find a fitting relation between the dependent variable Y and the independent variables of regressor matrix X. As a result, the values for the so-called regression coefficients $\beta_1, ..., \beta_r$ of parameter vector β are estimated. Note that an independent variable $X_{i,k}$ can be a non-linear function of another variable $X_{i,l}$, but the model remains linear in parameter vector β.

Finally, using the found values of the regression coefficients, a higher resolution thermal image can be predicted, such that:

$$LST_{high} = f(LSC_{high}) + \epsilon. \tag{14}$$

Note that this method can obtain a resolution no better than the highest resolution the LSCs are available in. Typically, downscaling the LSC images to match the lower resolution of the LST is not an issue as will be discussed in the next section.

5.2 Bilinear Interpolation

A technique that is commonly used to resample imagery is bilinear interpolation. It can be used to both upscale and downscale the resolution of an image, though in this research it will only be used to downscale high resolution imagery. The basic principle of bilinear interpolation is that linear interpolation is performed twice in different directions. In Figure 12 an example is shown of linear interpolation.

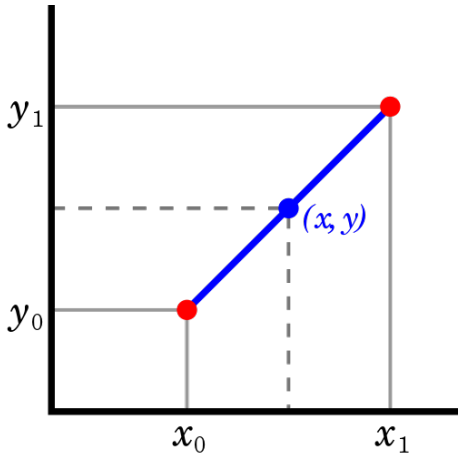


Figure 12: Example of linear interpolation (Berland, 2007)

Performing linear interpolation twice is a rather time consuming and devious approach. There is an alternative approach that renders a more elegant equation, namely:

$$f(x, y) = \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_3 xy. \tag{15}$$

Using linear algebra, the coefficients can be found by solving the following system:

$$\begin{pmatrix} 1 & x_1 & y_1 & x_1y_1 \\ 1 & x_1 & y_2 & x_1y_2 \\ 1 & x_2 & y_1 & x_2y_1 \\ 1 & x_2 & y_2 & x_2y_2 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} f(P_{11}) \\ f(P_{12}) \\ f(P_{21}) \\ f(P_{22}) \end{pmatrix}.$$

The error in bilinear interpolation is dependent on the interpolated function. In this case, the interpolated functions are the images representing the LSTs and LSCs. These images are of a highly erratic nature. Due to the unpredictable behavior of the LST and the LSCs, bilinear interpolation is an inadvisable approach for upscaling the resolution of these images. However, as stated previously, this interpolation method will be used for downscaling these images.

6 Implementation

6.1 Data

Using linear regression as explained previously, a relation will be found between the given LST of an area of land and data of different bands of spectral ranges in that same area of land. The LST image provided by eLEAF has been generated by the Visible Infrared Imaging Radiometer Suite (VIIRS), a commonly used thermal sensor. VIIRS provides eLEAF with low resolution LST images. The LST image that will be used for the linear regression can be seen in Figure 13.

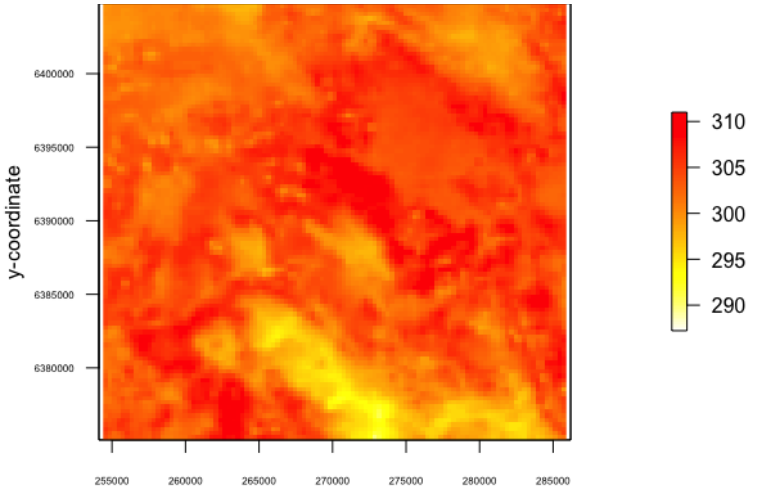


Figure 13: Low resolution (300m) image of LST (in K) generated by VIIRS

The pixels in Figure 13 represent a square surface area of $300m$ by $300m$. Hence, the spatial resolution of the image is 300m. In comparison, the much higher resolution images of the Landsat, another commonly used thermal sensor, have a spatial resolution of $30m$. Landsat provides eLEAF with high resolution LST images, although at a much lower frequency than VIIRS. In Figure 14 such a high resolution image of the same area is shown, taken by Landsat at approximately the same moment as the VIIRS image.

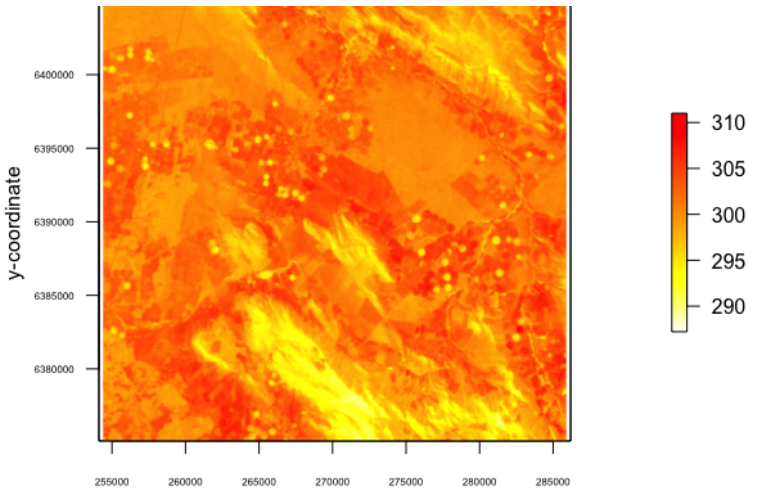


Figure 14: High resolution (30m) image of LST (in K) generated by Landsat

As can be seen in Figure 14, there are a number of circular yellow dots spread over the image. These are of high interest for eLEAF, as these dots represent agricultural areas. Also note how these dots are not visible in Figure 13. A model that can give reliable information on these areas would be useful. To be clear, the information given in Figure 14 will not be used to construct the model, but it will function as a reference.

The input for the independent variables in the model will be given by 13 different spectral bands. In Figure 15 all the spectral bands and their respective range on the electromagnetic spectrum can be seen. The 13 different spectral bands do not all have the same spatial resolution. As can be seen in Figure 15, there are three categories of spatial resolutions, namely resolutions of $10m$, $20m$ and $60m$.

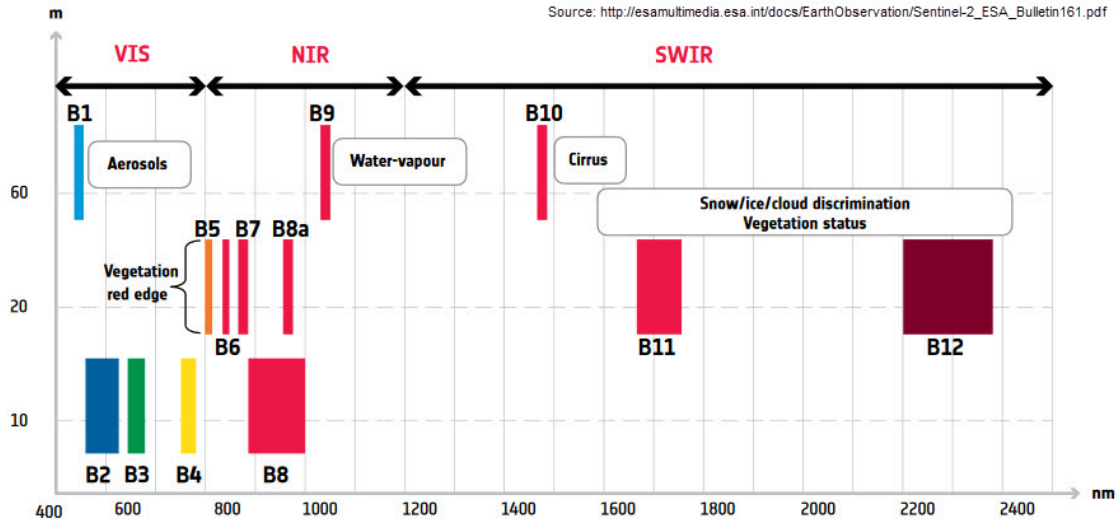
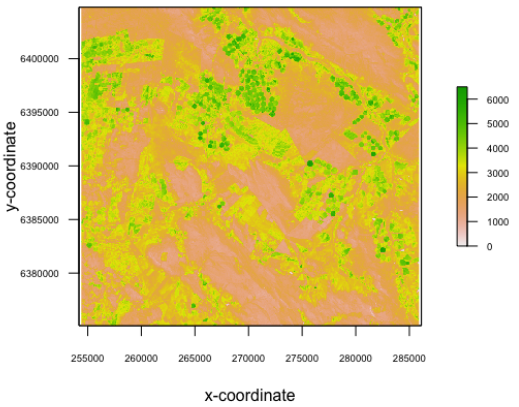


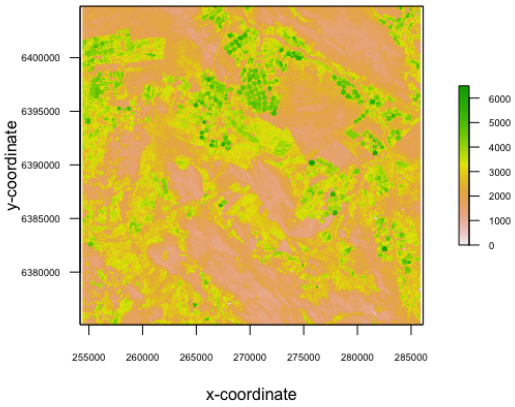
Figure 15: The 13 different spectral bands of Sentinel-2 that will be used as input (ESA, 2015)

eLEAF provided the imagery of all the 13 spectral bands in a spatial resolution of $20m$. So that means eLEAF uses a form of resampling to make all bands of identical resolution. These images represent the ratio between reflected and incoming radiation for a certain range in the electromagnetic spectrum. This ratio is given in prodecimille, so a pixel value of 1250 prodecimille corresponds with a reflection ratio of 0.125.

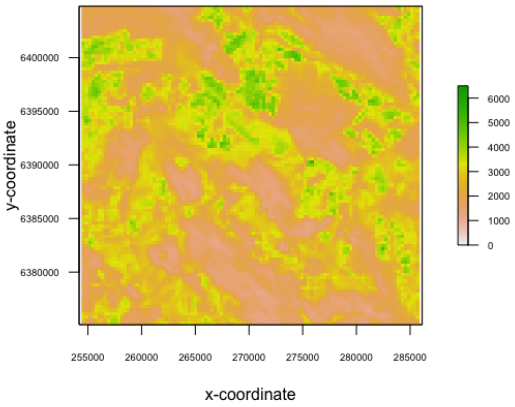
Bilinear interpolation will be used to resample the bands to the same resolution as the LST images. Consequently, a $300m$ resolution and a $30m$ resolution image is constructed for every spectral band. In Figure 16 an example can be seen of these resamples.



(a) Band 8a (resolution 20m)



(b) Band 8a (resolution 30m)



(c) Band 8a (resolution 300m)

Figure 16: Spectral reflectance (in prodecimille) of band 8a in different spatial resolutions

The spectral band images with a resolution of $300m$ will be used to perform the linear regression. Then, the spectral band images with a resolution of $30m$ will be used to predict the LST at that high resolution. Two models are used to make this prediction. These two models will be discussed in the next subsections.

6.2 Model 1

For the first model all the spectral bands are directly inserted in the model. As the computing time for finding a fitting model with 13 independent variables rapidly increases with each added degree, the variables are limited to at most the second degree. This also significantly reduces the number of the possible combinations in comparison with higher degrees and therefore the complexity of the formula. The resulting formula can be seen in (16). Note that the relation $*$ is defined as: $\alpha * \beta = \alpha + \alpha\beta + \beta$.

$$
\begin{aligned}
Temp \sim \quad & band1*(band2 + band3 + band4 + band5 + band6 + band7 + band8 \\
& \quad + band8a + band9 + band10 + band11 + band12) \\
& + band2*(band3 + band4 + band5 + band6 + band7 + band8 \\
& \quad + band8a + band9 + band10 + band11 + band12) \\
& + band3*(band4 + band5 + band6 + band7 + band8 + band8a \\
& \quad + band9 + band10 + band11 + band12) \\
& + band4*(band5 + band6 + band7 + band8 + band8a + band9 \\
& \quad + band10 + band11 + band12) \\
& + band5*(band6 + band7 + band8 + band8a + band9 + band10 \\
& \quad + band11 + band12) \\
& + band6*(band7 + band8 + band8a + band9 + band10 + band11 \\
& \quad + band12) \\
& + band7*(band8 + band8a + band9 + band10 + band11 + band12) \\
& + band8*(band8a + band9 + band10 + band11 + band12) \\
& + band8a * (band9 + band10 + band11 + band12) \\
& + band9*(band10 + band11 + band12) \\
& + band10 * (band11 + band12) \\
& + band11 * band12
\end{aligned}
\tag{16}
$$

After further studying the model, not all variables were found to be of sufficient significance. Using stepwise regression, a more fitting model can be found. Stepwise regression is a selection method used in statistics to find all variables that exceed a significance threshold. Using this for the formula described in (16), a new formula is found. This new formula can be seen in (17).

$$
\begin{aligned}
Temp \sim \quad & band1 + band2 + band3 + band4 + band5 + band6 \\
& + band7 + band8 + band8a + band9 + band10 + band11 + band12 \\
& + band1 \cdot band2 + band1 \cdot band3 + band1 \cdot band5 + band1 \cdot band6 + band1 \cdot band8 \\
& + band1 \cdot band8a + band1 \cdot band9 + band1 \cdot band10 + band1 \cdot band12 \\
& + band2 \cdot band3 + band2 \cdot band4 + band2 \cdot band5 + band2 \cdot band6 + band2 \cdot band7 \\
& + band2 \cdot band8 + band2 \cdot band8a + band2 \cdot band9 + band2 \cdot band10 \\
& + band2 \cdot band11 + band2 \cdot band12 + band3 \cdot band5 + band3 \cdot band6 \\
& + band3 \cdot band7 + band3 \cdot band8 + band3 \cdot band8a + band3 \cdot band9 \\
& + band3 \cdot band10 + band3 \cdot band11 + band3 \cdot band12 + band4 \cdot band5 \\
& + band4 \cdot band6 + band4 \cdot band7 + band4 \cdot band8a + band4 \cdot band9 \\
& + band4 \cdot band11 + band4 \cdot band12 + band5 \cdot band6 + band5 \cdot band7 \\
& + band5 \cdot band8a + band5 \cdot band10 + band5 \cdot band11 + band5 \cdot band12 \\
& + band6 \cdot band7 + band6 \cdot band8 + band6 \cdot band8a + band6 \cdot band9 \\
& + band6 \cdot band10 + band6 \cdot band11 + band6 \cdot band12 + band7 \cdot band8 \\
& + band7 \cdot band8a + band7 \cdot band9 + band7 \cdot band10 + band7 \cdot band11 \\
& + band7 \cdot band12 + band8 \cdot band8a + band8 \cdot band9 + band8 \cdot band10 \\
& + band8 \cdot band11 + band8 \cdot band12 + band8a \cdot band9 + band8a \cdot band11 \\
& + band8a \cdot band12 + band9 \cdot band11 + band9 \cdot band12 + band10 \cdot band11 \\
& + band10 \cdot band12 + band11 \cdot band12
\end{aligned} \tag{17}
$$

The found coefficients for the model are presented in Appendix B.

6.3 Model 2

For Model 2, the spectral bands will not be directly implemented in the formula. Instead they will be preprocessed to represent the LSCs mentioned in Section 4. An implication of this is that not all available spectral bands will be used, as not all spectral bands are used for the concerning LSCs. This formula can be seen in (18).

$$
\begin{aligned}
Temp \sim \quad & NDVI * (Albedo + PSRI + LWCI + MSI + GVMI \\
& \quad + NDWI + VSDI + SWCI + SIMI) \\
& + Albedo * (PSRI + LWCI + MSI + GVMI + NDWI \\
& \quad + VSDI + SWCI + SIMI) \\
& + PSRI * (LWCI + MSI + GVMI + NDWI + VSDI \\
& \quad + SWCI + SIMI) \\
& + LWCI * (MSI + GVMI + NDWI + VSDI + SWCI \\
& \quad + SIMI) \\
& + MSI * (GVMI + NDWI + VSDI + SWCI + SIMI) \\
& + GVMI * (NDWI + VSDI + SWCI + SIMI) \\
& + NDWI * (VSDI + SWCI + SIMI) \\
& + VSDI * (SWCI + SIMI) \\
& + SWCI * SIMI
\end{aligned} \tag{18}
$$

Similarly as with Model 1, stepwise regression is used to find the variables that exceed a significance threshold. The resulting formula can be found in (19)

$$
\begin{aligned}
Temp \sim \quad & NDVI + Albedo + PSRI + LWCI + MSI + GVMI \\
& + NDWI + VSDI + SWCI + SIMI + NDVI \cdot Albedo \\
& + NDVI \cdot PSRI + NDVI \cdot GVMI + NDVI \cdot NDWI + NDVI \cdot VSDI \\
& + NDVI \cdot SIMI + Albedo \cdot PSRI + Albedo \cdot MSI + Albedo \cdot NDWI \\
& + Albedo \cdot VSDI + Albedo \cdot SIMI + PSRI \cdot MSI + PSRI \cdot GVMI \\
& + PSRI \cdot NDWI + PSRI \cdot VSDI + PSRI \cdot SWCI + PSRI \cdot SIMI \\
& + LWCI \cdot MSI + LWCI \cdot GVMI + MSI \cdot GVMI + MSI \cdot VSDI \\
& + MSI \cdot SIMI + GVMI \cdot NDWI + GVMI \cdot VSDI + GVMI \cdot SWCI \\
& + GVMI \cdot SIMI + NDWI \cdot VSDI + NDWI \cdot SWCI + NDWI \cdot SIMI \\
& + VSDI \cdot SWCI + SWCI \cdot SIMI
\end{aligned}
\tag{19}
$$

The found coefficients for this model can also be found in Appendix B.

6.4 Compensating Operations

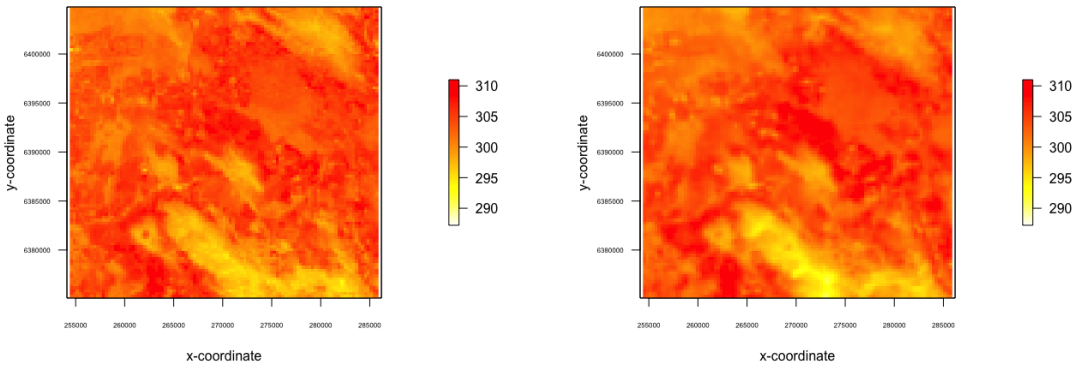
Now, using these models, two predictions will be performed with each. First the low resolution image will be reconstructed, then the high resolution image will be constructed. The results of these predictions will have outliers, as one would expect from a statistical model. As nonsensical outliers do not provide eLEAF with useful information (e.g. negative temperatures on the Kelvin scale), these outliers will be filtered out. Outliers in a prediction will be defined as elements that are respectively higher or lower than the maximum or minimum temperature of the low resolution image. These outliers will be replaced by the average of their neighbor elements. Simply leaving out the outliers would produce unwanted static on the resulting prediction. Therefore the choice is made to provide eLEAF with a reasonable indication of these areas, as it is essential to provide some statement on these areas. Implications of this will be discussed in the conclusion.

A final adjustment will also be added to the high resolution prediction. As the original low resolution and high resolution LST images by VIIRS and Landsat respectively were not taken at the same moment the temperatures differ slightly. The model will compensate that slight temperature deviation by adding the average difference of the prediction to all the values of the prediction.

7 Results

7.1 Prediction

Naturally, the question remains now to see how well the models can predict temperature based on the given input. Before making predictions on the high resolution temperatures, it makes sense to first try and reconstruct the image on which the model was based. First will be shown how both models can predict the low resolution image from which the relations are determined.



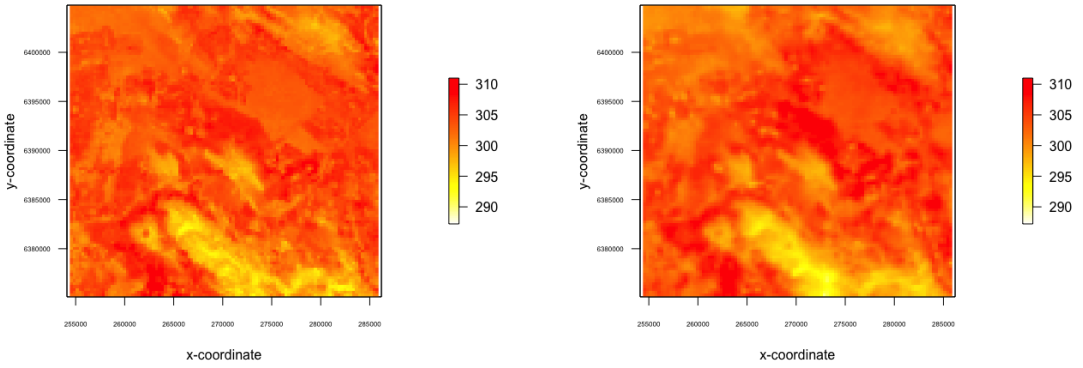
(a) Model 1 (b) VIIRS Observation

Figure 17: Low resolution (300m) comparison of Model 1 and VIIRS (in K)

In Figure 17, it can be seen how the first model, based purely on the direct input of all available spectral bands, can reproduce the low resolution image. Qualitatively, Figure 17a and Figure 17b appear to be similar images, but not exactly the same. It could be that the difference between the model and the reality further expand as the resolution gets higher. To get an indication on the error of the models the Mean Squared Error is computed. Its value is:

$$MSE_{Model1,low} = 1.4429.$$

Now the results of Model 2 will be examined. Similarly as with Model 1, first the reconstruction of the low resolution image is shown.



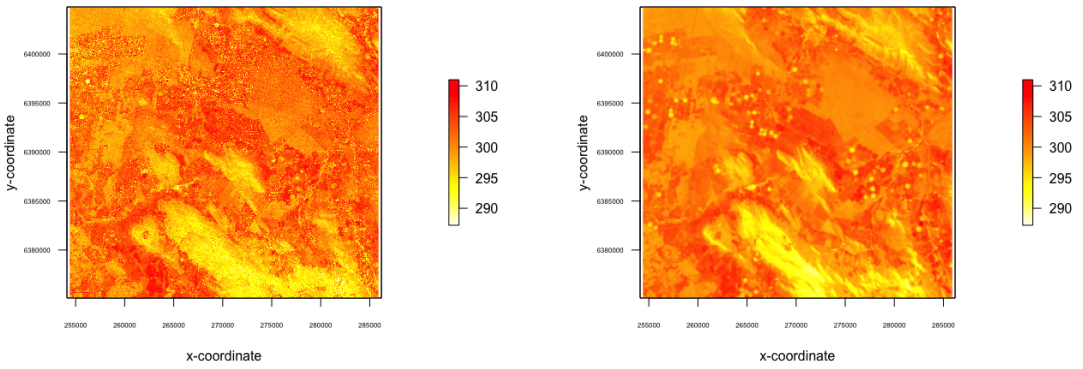
(a) Model 2 (b) VIIRS Observation

Figure 18: Low resolution (300m) comparison of Model 2 and VIIRS (in K)

In Figure 18, it can be seen how the second model, based on LSCs, reproduces the low resolution image. Comparing Figure 17a and Figure 18a it is clear that there are many similarities. However, there are also clearly some minor differences. From these first images it is not yet clear if Model 1 or Model 2 performs better in reproducing the low resolution image of the VIIRS satellite. Here, the Mean Squared Error is:

$MSE_{Model2,low} = 1.9120.$

For Model 2, the Mean Squared Error is higher than for Model 1. It could be that not including all spectral bands in this model results in a lower accuracy. Next, the higher resolution images will be presented.



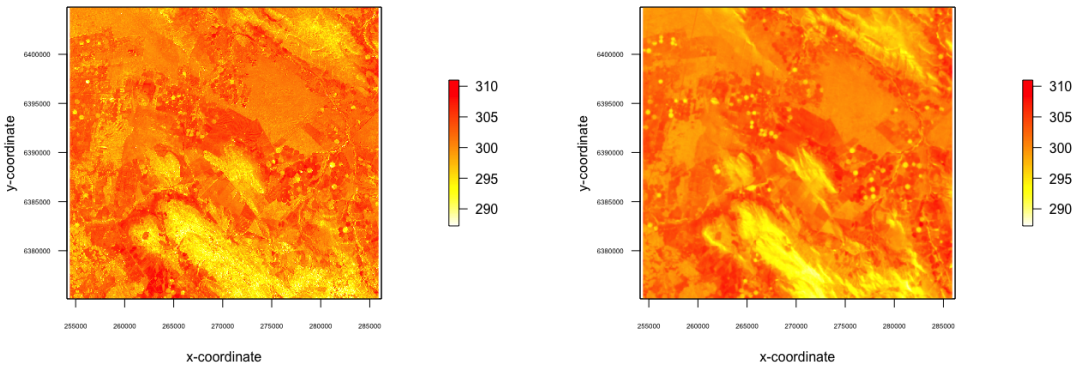
(a) Model 1 (b) Landsat Observation

Figure 19: High resolution (30m) comparison of Model 1 and Landsat (in K)

Even more interesting than reproducing the low resolution image is to predict the high resolution image. Figure 19 and Figure 20 show the comparison of the Landsat image and Model 1 and Model 2 respectively. In Figure 19 it can be seen that the model can produce roughly similar strokes of warm and cold areas as in the Landsat Image. However, there seems to be some noise in the form of small dots. Especially in cases where higher levels of detail are required, it would be difficult to make conclusions based on the image. The Mean Squared Error here is also substantially higher than in the lower resolution image:

$MSE_{Model1,high} = 6.7382.$

It is not surprising that the Mean Squared Error is higher for the high resolution image than for the low resolution image. Nonetheless, it is preferable to keep the Mean Squared Error lower rather than higher. Interesting is now to see how Model 1 performs in comparison to Model 2.



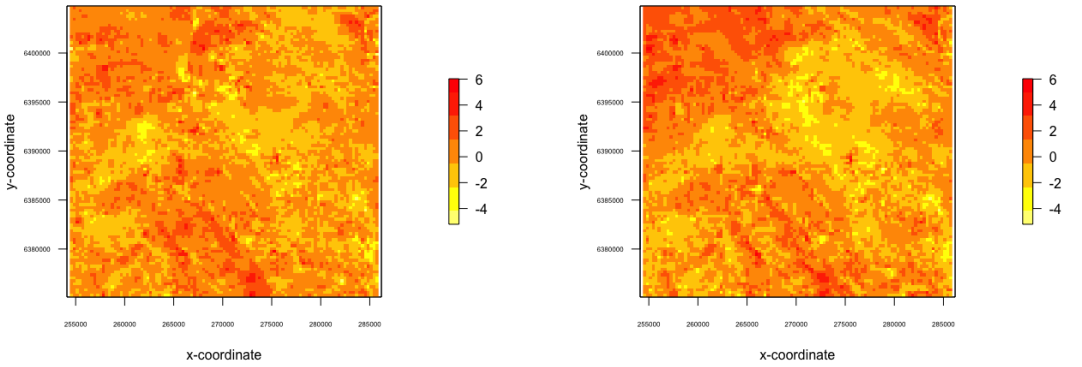
(a) Model 2 (b) Landsat Observation

Figure 20: High resolution (30m) comparison of Model 2 and Landsat (in K)

Most interesting to note in Figure 20 is that a high percentage of the low temperature dots from the Landsat image also appear in the Model 2 prediction. Also the edges of warm and cold areas appear to be sharper for Model 2 in comparison to Model 1. Thus, Model 2 seems to produce a more accurate image than Model 1 at this higher resolution. This is also reflected in the Mean Squared Error:

$MSE_{Model2,high} = 4.7135$.

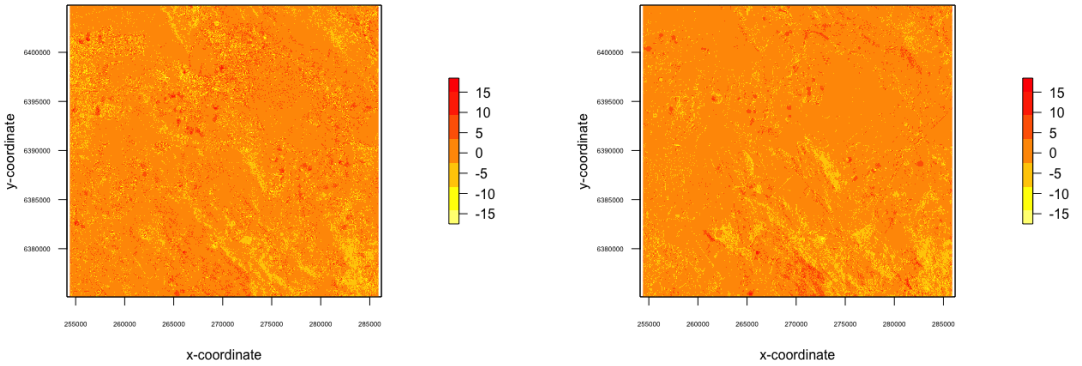
Surprisingly, the Model 2 Mean Squared Error performs worse for the low resolution. For the high resolution it performs significantly better.



(a) Difference of Model 1 (b) Difference of Model 2

Figure 21: Differences of the low resolution predictions and observations

Figure 21 shows a graph of the difference between the model and observation of the low resolution images. Darker red colors indicate the model overestimate the temperature, while lighter yellow colors indicate the model underestimate the temperature. Interesting to see is that the colors are not uniformly distributed over the image, but rather cluster in groups. The maximum difference in temperature from the observation is approximately 6K.



(a) Difference of Model 1 (b) Difference of Model 2

Figure 22: Differences of the low resolution predictions and observations

Similarly, the difference between the model and observation of the high resolution images can be seen in Figure 21. Here the maximum difference in temperature is approximately 18K, which is very significant. It is interesting to note that Figure 21a and 21b as well as Figure 22a and 22b show similar patterns. However, the patterns of Figure 21 and 22 do not appear to have similar patterns.

7.2 Diagnosis

To get better insights on how well the data fits a linear regression model, a look is taken at the diagnostics of both models. Consequently, further conclusions can be drawn on which model performs better.

Model 1

From the diagnostics of Model 1, presented in the three plots in Figures 23, 24 and 25, it can be concluded that the model is not perfect, but does sufficiently fit a linear regression model.

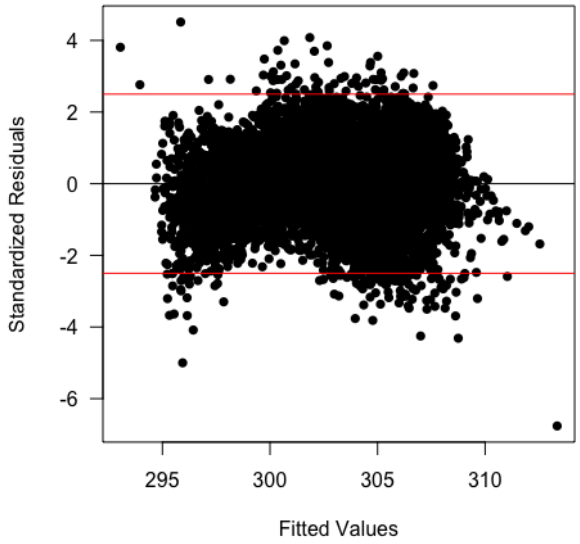


Figure 23: Standardized Residuals plot of Model 1

There are a bunch of outlying data points, as can be seen in Figure 23. Ideally, the points would be distributed in such a way that a horizontal band of points near the 0 line is formed. The red lines can be used as a reference to see how well this horizontal band is formed. All points outside this band formed by the red lines indicate a very high relative difference in the prediction versus the actual value. The band of points also shows a slight curve that gives cause for concern. There is one point with a value of approximately -6 which indicates a large outlier.

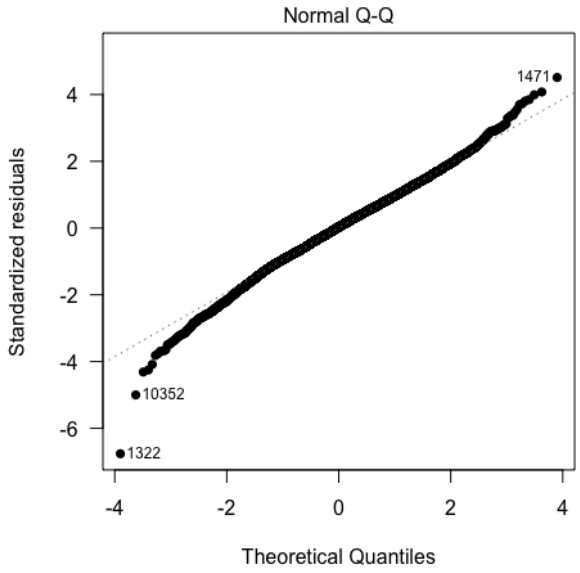


Figure 24: Normal Q-Q plot of Model 1

For the Normal Q-Q plot it is tested if the error components of the data are approximately

normally distributed. The data points should follow the diagonal dashed line and in Figure 24 the points sufficiently follow this line. Near the ends of the quantiles there is a curve, which again gives cause for concern.

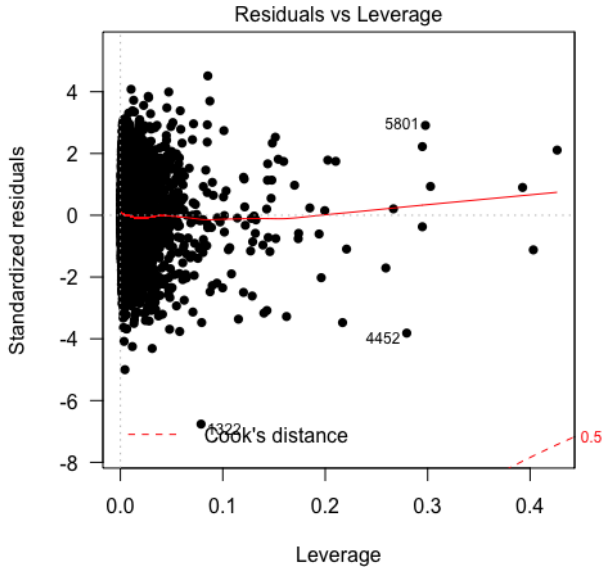


Figure 25: Cook's Distance plot of Model 1

Cook's Distance is a statistical tool used to check if there are data points that leverage the values of the fitted model. Figure 25 shows that all data falls within Cook's Distance. The conclusion that can be drawn is that the data points do not significantly leverage the model. However, some points do have a relatively high leverage value, even though they still are within Cook's Distance.

	Temp	b1	b2	b3	b4	b5	b6	b7	b8	b8a	b9	b10	b11	b12
Temp	1.00	0.56	0.58	0.63	0.71	0.70	0.65	0.64	0.64	0.64	0.63	0.76	0.77	0.75
b1	0.56	1.00	0.99	0.96	0.91	0.89	0.81	0.79	0.79	0.78	0.81	0.86	0.86	0.88
b2	0.58	0.99	1.00	0.98	0.94	0.92	0.85	0.82	0.82	0.82	0.84	0.88	0.89	0.91
b3	0.63	0.96	0.98	1.00	0.98	0.97	0.92	0.90	0.90	0.90	0.91	0.91	0.94	0.95
b4	0.71	0.91	0.94	0.98	1.00	0.99	0.94	0.92	0.92	0.91	0.92	0.93	0.97	0.98
b5	0.70	0.89	0.92	0.97	0.99	1.00	0.97	0.95	0.95	0.95	0.95	0.92	0.97	0.97
b6	0.65	0.81	0.85	0.92	0.94	0.97	1.00	1.00	1.00	1.00	0.99	0.86	0.91	0.90
b7	0.64	0.79	0.82	0.90	0.92	0.95	1.00	1.00	1.00	1.00	0.99	0.84	0.89	0.87
b8	0.64	0.79	0.82	0.90	0.92	0.95	1.00	1.00	1.00	1.00	0.99	0.84	0.89	0.87
b8a	0.64	0.78	0.82	0.90	0.91	0.95	1.00	1.00	1.00	1.00	0.99	0.84	0.89	0.87
b9	0.63	0.81	0.84	0.91	0.92	0.95	0.99	0.99	0.99	0.99	1.00	0.87	0.90	0.88
b10	0.76	0.86	0.88	0.91	0.93	0.92	0.86	0.84	0.84	0.84	0.87	1.00	0.95	0.94
b11	0.77	0.86	0.89	0.94	0.97	0.97	0.91	0.89	0.89	0.89	0.90	0.95	1.00	0.99
b12	0.75	0.88	0.91	0.95	0.98	0.97	0.90	0.87	0.87	0.87	0.88	0.94	0.99	1.00

Table 1: Correlations of all 13 bands with each other and the LST (Temp)

Table 1 shows that all bands are moderately to strongly correlated with the temperature. However, all bands are also strongly correlated with each other. The consequence is that the model could overfit the data. This means that in determining the coefficients there are more equations than unknowns. The result is that there are inconsistencies in the model. This is also a probable explanation for the noise that appears in the Model 1 prediction seen in Figure 19.

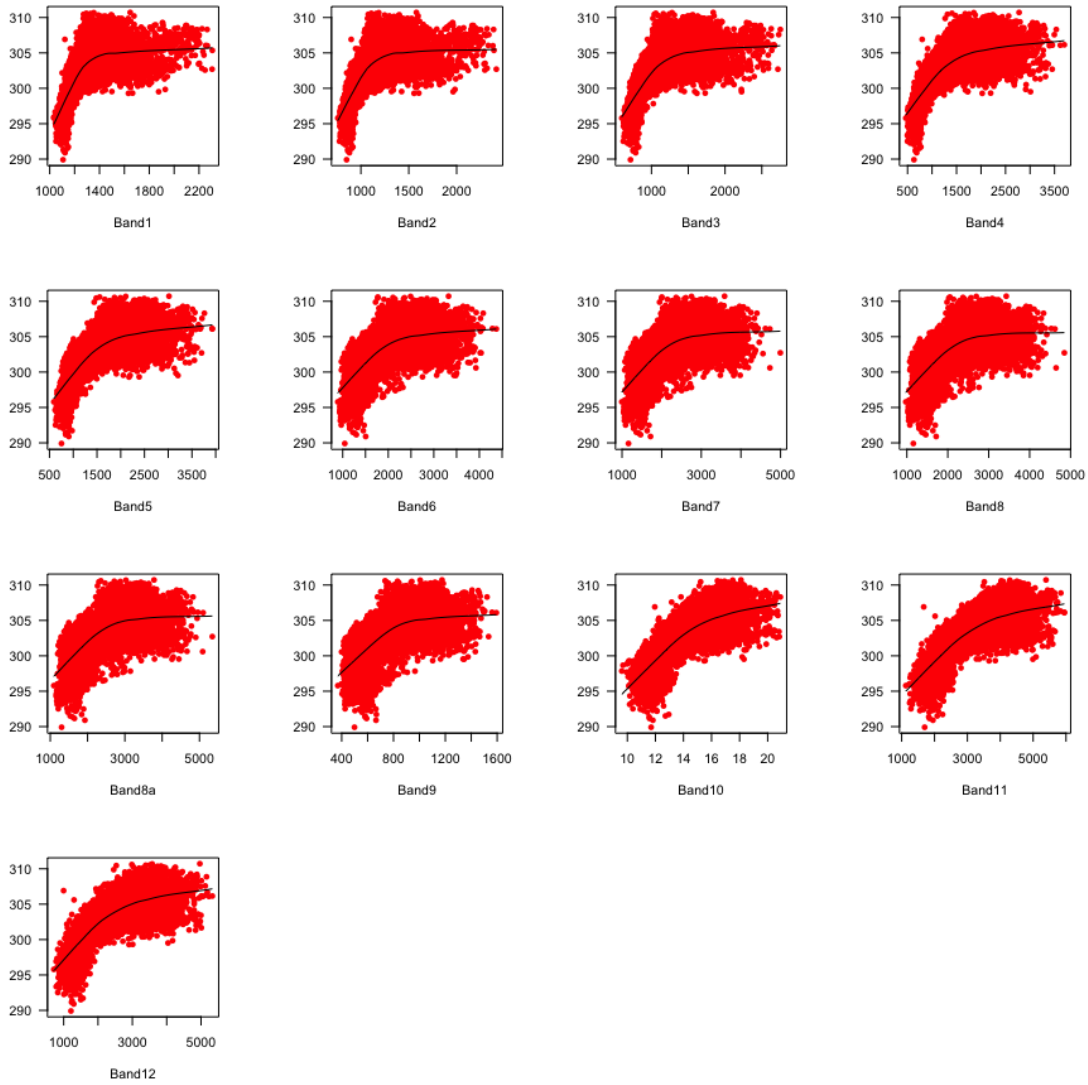


Figure 26: Scatterplots of variables of Model 1 versus LST

In Figure 26 can be seen that all bands have a very similar relation with the LST. Looking at the relations, it is not surprising that all bands have such a strong correlation with each other as seen in Table 1. None of the bands have a perfectly linear relation with the LST, but there seems to be a trend that the higher bands (especially, Band 10, 11 and 12) are more linear.

Model 2

For the diagnostics of Model 2, the three plots in Figures 27, 28 and 29 show that the data is still not perfect. However, it does further affirm that Model 2 performs better than the first model.

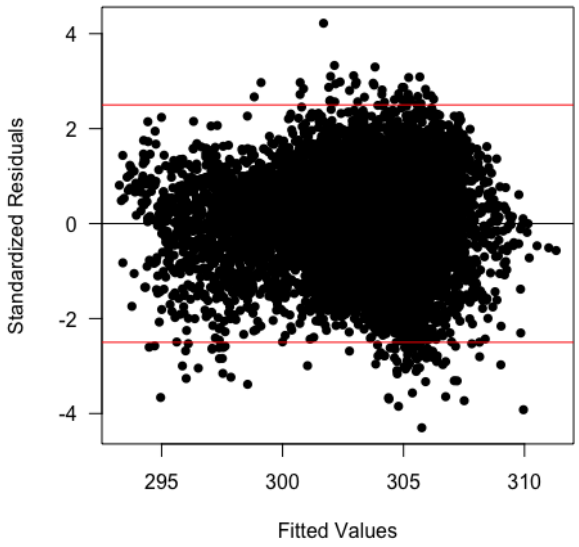


Figure 27: Standardized Residuals plot of Model 2

In Figure 27, the residuals form a slightly more uniform band than in Figure 23. There are still a lot of points with a standardized residual value outside the band formed by the red lines. Ideally, all points lie within the -2.5 and 2.5 horizontal band.

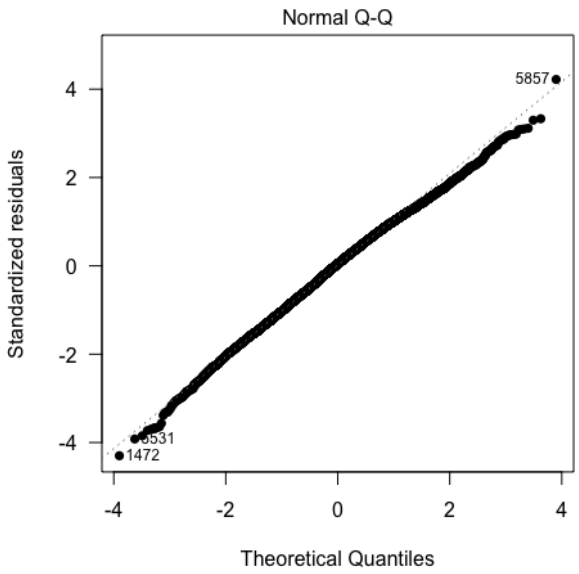


Figure 28: Normal Q-Q plot of Model 2

In Figure 28, the data points follow the diagonal dashed line very well. In comparison with Figure 24 there is a significant improvement. Also in this plot, tear the ends of the quantiles there is a slight curve, but it is less significant than can be seen in Figure 24.

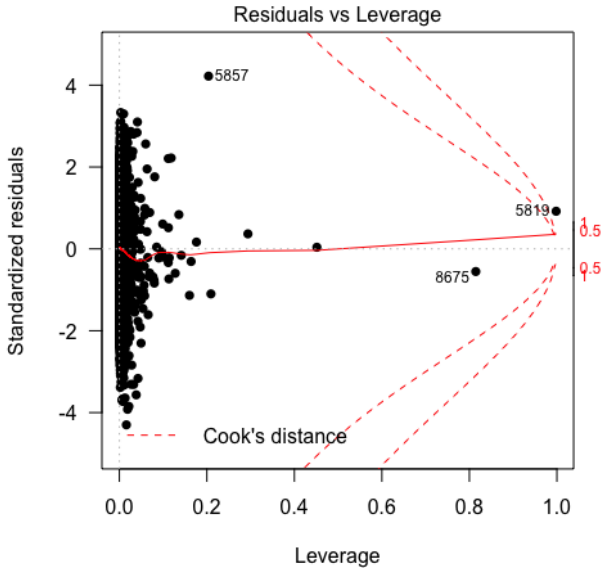


Figure 29: Cook's Distance plot of Model 2

Figure 29 shows that there is a point that falls outside Cook's Distance. This point and the other points with a relatively high leverage could be further investigated to find a cause for their behavior. In comparison with 25 there are less points in

	Temp	NDVI	Albedo	PSRI	LWCI	MSI	GVMI	NDWI	VSDI	SWCI	SIMI
Temp	1.00	-0.70	0.68	0.83	-0.00	0.18	-0.59	-0.18	-0.79	-0.67	0.76
NDVI	-0.70	1.00	-0.71	-0.76	-0.01	-0.53	0.87	0.55	0.75	0.90	-0.80
Albedo	0.68	-0.71	1.00	0.87	-0.01	-0.13	-0.42	0.10	-0.93	-0.81	0.96
PSRI	0.83	-0.76	0.87	1.00	-0.01	0.01	-0.56	-0.04	-0.94	-0.80	0.91
LWCI	-0.00	-0.01	-0.01	-0.01	1.00	0.02	-0.01	-0.02	0.01	0.00	-0.01
MSI	0.18	-0.53	-0.13	0.01	0.02	1.00	-0.81	-0.99	-0.04	-0.32	0.09
GVMI	-0.59	0.87	-0.42	-0.56	-0.01	-0.81	1.00	0.83	0.58	0.74	-0.62
NDWI	-0.18	0.55	0.10	-0.04	-0.02	-0.99	0.83	1.00	0.07	0.35	-0.11
VSDI	-0.79	0.75	-0.93	-0.94	0.01	-0.04	0.58	0.07	1.00	0.82	-0.98
SWCI	-0.67	0.90	-0.81	-0.80	0.00	-0.32	0.74	0.35	0.82	1.00	-0.87
SIMI	0.76	-0.80	0.96	0.91	-0.01	0.09	-0.62	-0.11	-0.98	-0.87	1.00

Table 2: Correlations of all 10 LSCs with each other and the LST (Temp)

Table 2 is very different from Table 1. Whereas all bands are strongly positively correlated with each other, here it can that the LSCs have more unique correlations. Immediately can be seen that MSI and NDWI are very weakly correlated to the temperature and LWCI is not correlated at all. The LSCs that are strongly correlated to temperature, either positively or negatively, are also strongly correlated to each other. Most of these LSCs are used for monitoring SMC, so it is not surprising that they are correlated to each other. The mutual correlations in 2 are not as high as in 1 though, so that explains why there is less noise in the Model 2 prediction as seen in 20.

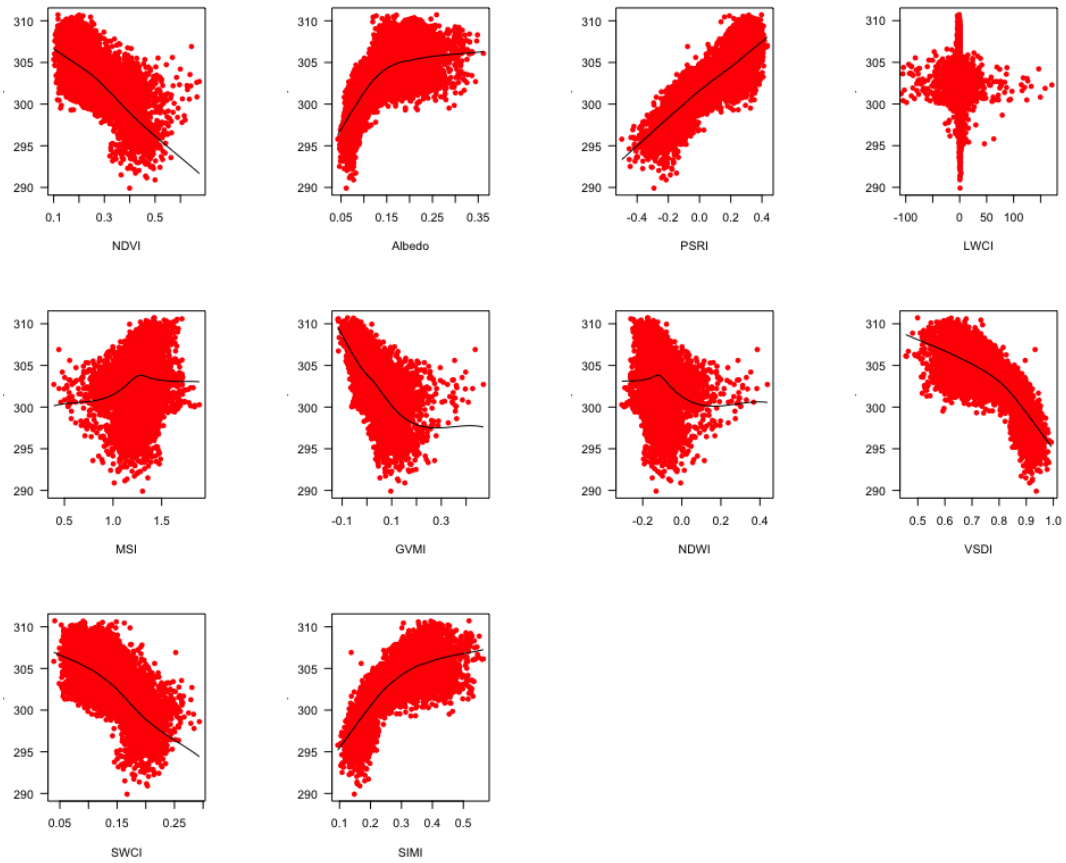


Figure 30: Scatterplots of variables of Model 2 versus LST

Unlike Figure 26, in Figure 30 can be seen that there are multiple unique relations with the LST. PSRI stands out, because of its linear relation with the LST. Unsurprisingly, PSRI also has the highest correlation with the LST as seen in Table 2. NDVI, SWCI and VSDI also have a moderately linear relationship with the LST.

8 Conclusion

In this conclusion, first a few shortcomings of the model will be examined. Then, also based on these shortcomings, recommendations for further research will be given.

8.1 Shortcomings

The reference material

The high resolution image provided by Landsat was used as reference material for the quality of the model. The goal was to provide eLEAF with a high resolution image of the LST, making use of the low resolution image of the LST. In the model, values for the prediction are shifted by the average difference between the prediction and the original high resolution image. In practice this would not be possible, as the high resolution would normally not be available. However, as previously mentioned, the images were not taken at exactly the same moments, but approximately 30 minutes apart. The extra solar exposure seems a good explanation for the average $2.44K$ difference in temperature between the two images.

It is unclear how the time in between the taking of the images has affected the LST. Logically, more solar exposure would result in higher temperatures, but it is unknown how exactly this affects every different surface type. As a consequence the high resolution image by Landsat might differ in unexpected ways from the low resolution image by VIIRS. This could be a reason why the Mean Squared Error for the high resolution images is higher than for the low resolution images.

Reliability of the input

In the original images of the spectral bands there are a few data points that have extremely high values. This could be due to a fault in the sensor, or a rare surface type. Whatever the cause, these points could have an influence on the model, though it is unclear how much they would affect it. After resampling with bilinear interpolation, most of these points seem to have been filtered out. However, there are some, although very few, data points on the edge of the image that do not get filtered out in the resampling to a spatial resolution of $30m$.

The question is how much these points influence the model. It is clear that these points cause outliers in the prediction. It is unclear what their effect is on determining the coefficients of the model. As previously stated, these data points seem to get flattened out in the resampling to the low spatial resolution of $300m$. However, the outliers could still influence the values of the resampled images.

Approach to outliers in the prediction

As discussed in section 3.4, outliers in the prediction are replaced by the average of their neighboring data points. However, this can be a very inaccurate approach, especially at lower resolutions. Due to the non-linear behavior of temperature, a singular peak in temperature could occur in the concerning data point that the neighboring data points do not capture. The lower the resolution, the higher the chance of such a peak occurring. Additionally, such peaks could very well be the cause of outliers in the model in the first place. Consequently, using this approach on a cluster of outliers would even further decrease the accuracy of the prediction.

In reality, LST is represented by a continuous 2D-grid instead of a discrete 2D-grid of pixels used in this research. So if the distribution of outliers on a discrete grid is sufficiently uniform and the grid is at a sufficiently high resolution, the approach of averaging neighbors is effective. However, for this research the expertise on the behavior of LST is too limited to draw a concrete conclusion of the effectiveness of this approach.

8.2 Recommendations

Testing on more data

The most urgent suggestion for new research is to test this model on different datasets. There is a wide variety of options to research.

Firstly, the same area grid should be tested at different times. Further research might compare if the LSCs still have the same relation with temperature at various times of the day. Additionally, even broader time gaps should be compared to observe the consequences of seasonal shifts. Secondly, it is advised the model is tested on different areas of interest. On the one hand, it should be researched if the models of this research are indeed effective in other similar landscapes. On the other hand, further research should also consider dissimilar landscapes. Different surface types could influence the relation of the LST with the LSCs. If so, additional research would answer how large the differences are.

If possible, the ultimate goal should be to find a universal model that accurately predicts the LST for any surface type at any time of day.

Filtering out outliers of the input

A suggestion for further research is that the input data is filtered so that outliers do not get taken into account in the model.

The input data has outliers as previously stated. In the results presented in this research, these outliers are used in the fitting of the model. As a result, these outliers could lever the coefficients to their favor. As seen in Figure 25 and 29, there are data points in both models that potentially influence the model. Further research on these data points can determine if the data points are indeed erroneous outliers that can be filtered out of the input data.

Testing different models

In both Model 1 and Model 2 a linear regression model is formed with stepwise selection. In further research additional statistical models could considered. A few options are discussed below.

Box-Cox transformation
Looking at Figure 26, the relations all are similar to a square root relation. Consequentially, a Box-Cox transformation quickly comes to mind, with the goal to normalize the non-linear relation between the spectral bands and the LST. Such a power transformation takes the form:

$$Y^{\lambda} = X \cdot \beta + \epsilon, \tag{20}$$

with $\lambda \in \mathbb{R}$. Preferably for Box-Cox transformations, λ takes a value between -2 and 2. However, in the case of Model 1, an optimal λ takes a relatively high value of around 6. The consequence is that such a high exponent increases the gap between outliers immensely. In this data set, as seen in Figure 23 there are a lot of outliers. As such, this is not a recommended statistical model for the data set.

Nonlinear regression
As shown in the Figures 26 and 30, many of the used independent variables do not have a linear relation with the dependent variable: the LST. Consequently, a nonlinear combination of the independent variables could be a better fitting model for the observational data.

Ridge regression & LASSO
An issue with the current model is that there is a high collinearity between the independent variables. Variable selection via stepwise regression, which is used in this research, can be inadequate at taking the effects of strong correlations into account. There are other methods that are better suited for dealing with collinearity in linear regression models. High collinearity can lead to an ill-posed set of equations - a non-unique solution. Standard linear regression as used in this research then leads to overfitting. Ridge regression uses Tikhonov regularization to prevent overfitting.

Least absolute shrinkage and selection operator (LASSO) uses both regularization and variable selection to prevent overfitting.

The gain of applying either ridge regression or LASSO is that the prediction produces less noise. In practice, this means clearer images. Thus, it is recommended to further research the potential of applying ridges regression or LASSO.

Finding more suitable variables

As shown in Table 1 and 2, there is high collinearity between all the independent variables that also have a strong correlation with the dependent variable. The consequence of this is overfitting, which produces noise in the prediction. As discussed earlier, ridge regression or LASSO could provide a solution for overfitting. However, as seen in Figure 22 there are clusters of large differences between the model and the observation that indicate the model missed low temperature agricultural areas. Although there are slight differences for the two models, it appears that some of these errors appear in the same clusters of data. Also, these clusters involve relatively large sets of data, which could be an indication that the error in the prediction is not due to overfitting. In stead, it could indicate that an essential independent variable for accurate predictions is missing in the model.

To improve the current model, a more suitable set of variables could be used for the linear regression. It is recommended to research if there are other variables that have a strong correlation with the LST, while having little collinearity with the other independent variables. Such variables have the potential to greatly increase the accuracy of the model.

9 Bibliography

Berland (2007). Linear Interpolation. https://commons.wikimedia.org/wiki/File:LinearInterpolation.svg. Accessed: 2018-04-11.

Ceccato, P., Gobron, N., Flasse, S., Pinty, B., and Tarantola, S. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data: Part 2. Validation and applications. *Remote Sensing of Environment*, 82(2-3):198–207.

Chen, H., Zhang, H., Shen, S.-h., Yu, W., and Zou, C. (2009). Construction and validation of a new model for cropland soil moisture index based on MODIS data. 7454(August 2009):745418–1–745418–8.

Coakley, J. (2003). Reflectance and Albedo, Surface. *Encyclopedia of Atmospheric Sciences*, pages 1914–1923.

ESA (2015). The 13 spectral bands of Sentinel-2 as illustrated by the European Space Agency. http://esamultimedia.esa.int/docs/EarthObservation/Sentinel-2_ESA_Bulletin161.pdf. Accessed: 2018-04-11.

Gandhi, G. M., Parthiban, S., Thummalu, N., and Christy, A. (2015). NDVI: Vegetation Change Detection Using Remote Sensing and Gis - A Case Study of Vellore District. *Procedia Computer Science*, 57:1199–1210.

Hardisky, M. A., Klemas, V., and Smart, R. M. (1983). The influence of Soil Salinity, Growth Form, and Leaf Moisture on the Spectral Radiance of Spartina alterniflora Canopies. *Photogrammetric Engineering and Remote Sensing*, 49(1):77–83.

Hunt Jr, E. R. and Rock, B. N. (1989). Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing of Environment*, 30(1):43–54.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of leaf senescence and fruit ripening. *Physiol. Plant*, 106:135.

Yao, Y., Qin, Q., Zhao, S., and W., Y. (2011). Retrieval of soil moisture based on MODIS shortwave infrared spectral feature. *International Journal of Disaster Risk Science*, 4(30):9–14.

Zhang, N., Hong, Y., Qin, Q., and Liu, L. (2013). VSDI: a visible and shortwave infrared drought index for monitoring soil and vegetation moisture based on optical remote sensing. *International Journal of Remote Sensing*, 34(13):4585–4609.

10 Appendix

10.1 A: Input

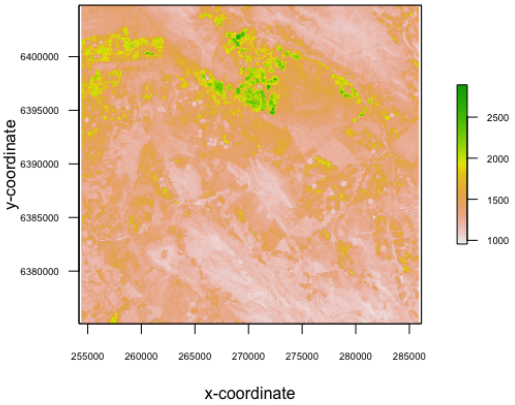


Figure 31: Spectral reflectance (in prodecimille) of band 1

Band 1 is referred to as the coastal aerosol band and has a central wavelength of 0.443 μm. The values can be seen in Figure 31.

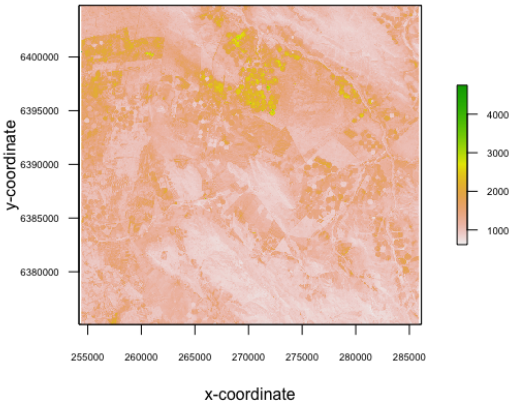


Figure 32: Spectral reflectance (in prodecimille) of band 2

Band 2 is referred to as the blue band and has a central wavelength of 0.490 μm. The values can be seen in Figure 32.

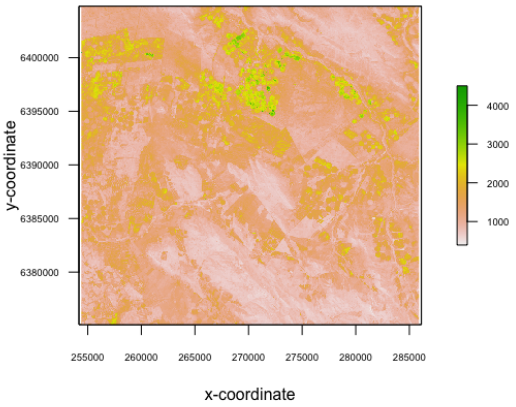


Figure 33: Spectral reflectance (in prodecimille) of band 3

Band 3 is referred to as the green band and has a central wavelength of 0.560 μm. The values can be seen in Figure 33.

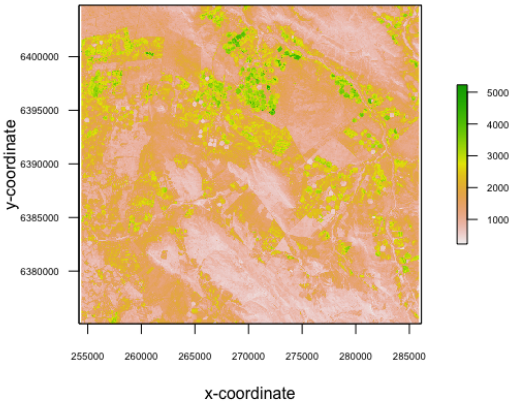
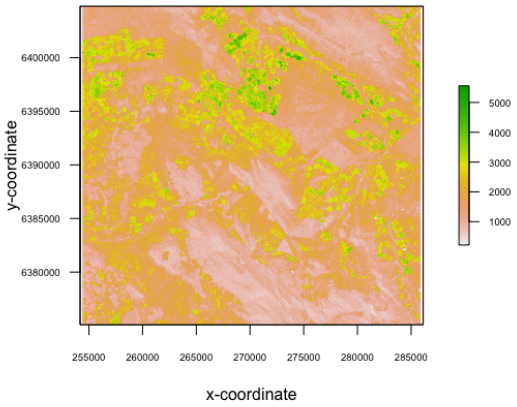
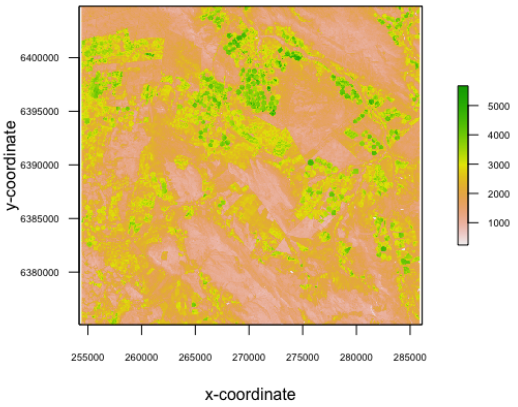


Figure 34: Spectral reflectance (in prodecimille) of band 4

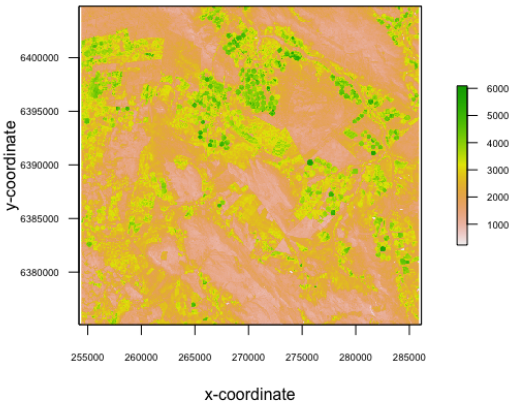
Band 4 is referred to as the red band and has a central wavelength of 0.665 μm. The values can be seen in Figure 34.



(a) Band 5



(b) Band 6



(c) Band 7

Figure 35: Spectral reflectance (in prodecimille) of band 5, 6 and 7

Band 5, 6 and 7 are referred to as the vegetation red edge bands and have a central wavelength of 0.705 μm, 0.740 μm and 0.783 μm respectively. The values can be seen in Figure 35.

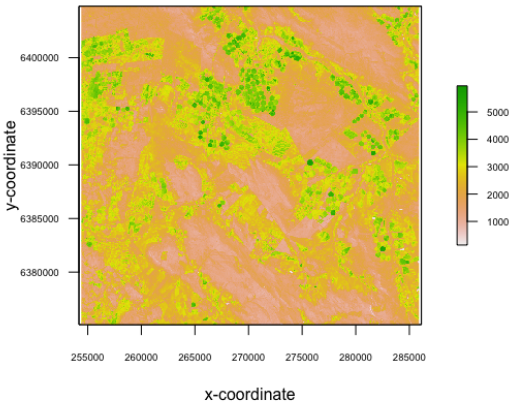


Figure 36: Spectral reflectance (in prodecimille) of band 8

Band 8 is referred to as the NIR band and has a central wavelength of 0.842 μm. The values can be seen in Figure 36.

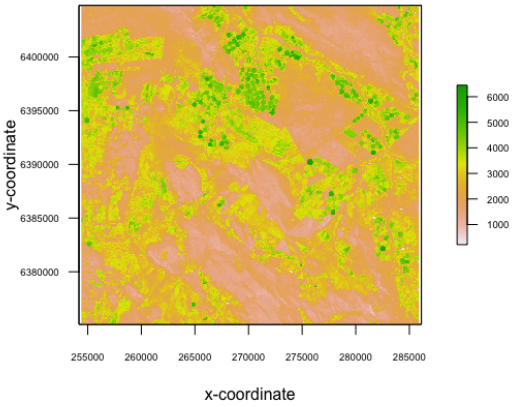


Figure 37: Spectral reflectance (in prodecimille) of band 8a

Band 8a is referred to as the narrow NIR band and has a central wavelength of 0.865 μm. The values can be seen in Figure 37.

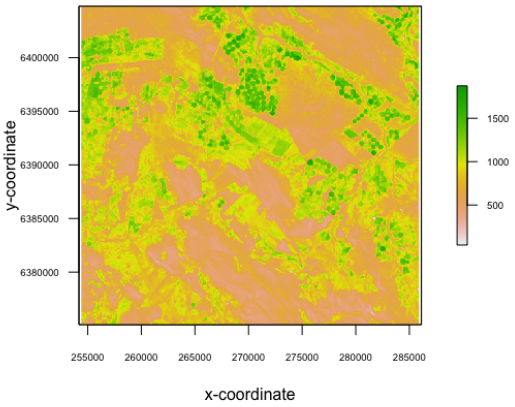


Figure 38: Spectral reflectance (in prodecimille) of band 9

Band 9 is referred to as the water vapour band and has a central wavelength of 0.945 μm. The values can be seen in Figure 38.

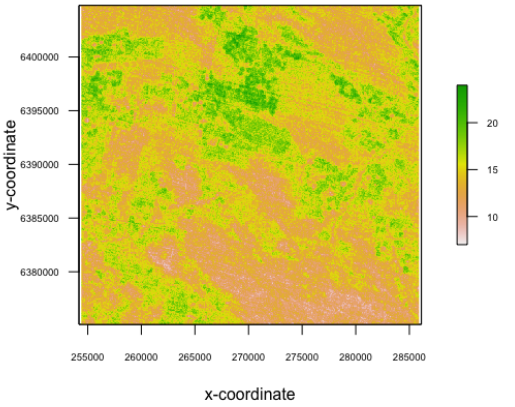
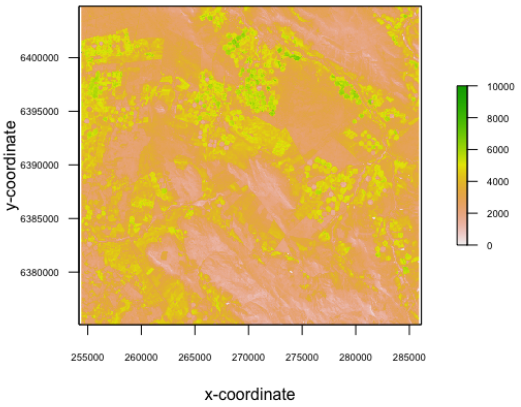
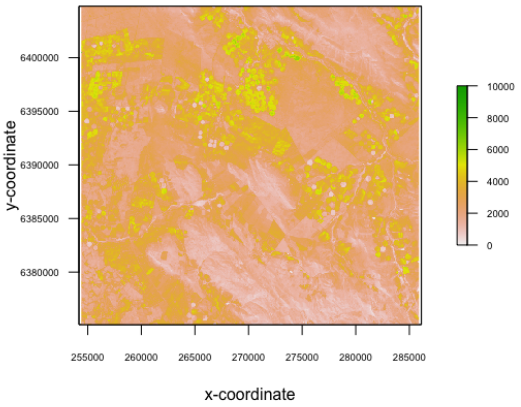


Figure 39: Spectral reflectance (in prodecimille) of band 10

Band 10 is referred to as the SWIR-Cirrus band and has a central wavelength of 1.375 μm. The values can be seen in Figure 39.



(a) Spectral reflectance (in prodecimille) of band 11



(b) Band 12

Figure 40: Spectral reflectance (in prodecimille) of band 11 and 12

Band 11 and 12 are referred to as the SWIR bands and have a central wavelength of 1.610 μm and 2.190 μm respectively. The values can be seen in Figure 40.

10.2 B: Model Values

	Coefficients
(Intercept)	2.926E+02
band1	-4.405E+02
band2	9.632E+02
band3	-6.603E+02
band4	-1.009E+03
band5	2.457E+03
band6	-3.864E+03
band7	4.869E+03
band8	-9.754E+01
band8a	-1.780E+03
band9	-4.549E+02
band10	-1.830E+04
band11	1.691E+02
band12	-2.617E+02
band1:band2	-2.136E+03
band1:band3	7.002E+03
band1:band5	-7.741E+03
band1:band6	-1.673E+04
band1:band8	2.799E+04
band1:band8a	-1.911E+04
band1:band9	1.761E+04
band1:band10	6.916E+05
band1:band12	3.542E+03
band2:band3	-6.908E+03
band2:band4	1.021E+04
band2:band5	-1.893E+04
band2:band6	9.026E+04
band2:band7	-6.145E+04
band2:band8	-5.033E+04
band2:band8a	4.599E+04
band2:band9	-1.704E+04
band2:band10	-8.557E+05
band2:band11	3.895E+03
band2:band12	-7.018E+03
band3:band5	1.709E+04
band3:band6	-5.550E+04
band3:band7	3.921E+04
band3:band8	2.433E+04
band3:band8a	-1.831E+04
band3:band9	-7.214E+03
band3:band10	4.055E+05

Table 3: Coefficients of Model 1 (part 1)

	Coefficients
band3:band11	-6.698E+03
band3:band12	4.921E+03
band4:band5	-6.640E+03
band4:band6	-7.272E+03
band4:band7	1.159E+04
band4:band8a	-9.207E+03
band4:band9	1.270E+04
band4:band11	1.004E+04
band4:band12	-7.823E+03
band5:band6	1.794E+04
band5:band7	-1.658E+04
band5:band8a	2.930E+03
band5:band10	-4.254E+05
band5:band11	-9.530E+03
band5:band12	1.073E+04
band6:band7	2.785E+03
band6:band8	-9.765E+03
band6:band8a	1.225E+04
band6:band9	-2.441E+04
band6:band10	9.805E+05
band6:band11	6.341E+03
band6:band12	-1.081E+04
band7:band8	8.143E+03
band7:band8a	-1.549E+04
band7:band9	2.467E+04
band7:band10	-1.211E+06
band7:band11	-8.582E+03
band7:band12	1.076E+04
band8:band8a	3.158E+03
band8:band9	-2.098E+04
band8:band10	5.514E+05
band8:band11	-3.377E+03
band8:band12	2.742E+03
band8a:band9	1.494E+04
band8a:band11	8.749E+03
band8a:band12	-6.850E+03
band9:band11	-3.932E+03
band9:band12	2.466E+03
band10:band11	-5.809E+04
band10:band12	8.476E+04
band11:band12	-2.702E+02

Table 4: Coefficients of Model 1 (part 2)

In Tables 3 and 4 are the coefficients determined for Model 1. Generally, higher values indicate a larger weight in the model.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
band1	1	30625.8	30625.8	20967.2	0.00E+00
band2	1	3031.3	3031.3	2075.3	0.00E+00
band3	1	11781.2	11781.2	8065.7	0.00E+00
band4	1	16312.9	16312.9	11168.3	0.00E+00
band5	1	11.4	11.4	7.8	5.21E-03
band6	1	546.5	546.5	374.1	6.59E-82
band7	1	3060.4	3060.4	2095.2	0.00E+00
band8	1	60.4	60.4	41.4	1.32E-10
band8a	1	185.6	185.6	127.1	2.64E-29
band9	1	1518.5	1518.5	1039.6	2.12E-217
band10	1	2958.3	2958.3	2025.3	0.00E+00
band11	1	973.2	973.2	666.3	2.00E-142
band12	1	748.6	748.6	512.5	8.76E-111
band1:band2	1	2060.2	2060.2	1410.5	1.94E-289
band1:band3	1	690.7	690.7	472.9	1.48E-102
band1:band5	1	2313.8	2313.8	1584.1	2.22E-322
band1:band6	1	30.7	30.7	21.0	4.57E-06
band1:band8	1	325.4	325.4	222.8	7.32E-50
band1:band8a	1	386.9	386.9	264.9	7.95E-59
band1:band9	1	240.0	240.0	164.3	2.49E-37
band1:band10	1	67.8	67.8	46.4	1.02E-11
band1:band12	1	101.4	101.4	69.4	8.98E-17
band2:band3	1	37.1	37.1	25.4	4.78E-07
band2:band4	1	66.7	66.7	45.7	1.46E-11
band2:band5	1	23.3	23.3	15.9	6.59E-05
band2:band6	1	138.5	138.5	94.8	2.58E-22
band2:band7	1	566.6	566.6	387.9	8.32E-85
band2:band8	1	25.6	25.6	17.5	2.90E-05
band2:band8a	1	40.8	40.8	28.0	1.27E-07
band2:band9	1	72.3	72.3	49.5	2.08E-12
band2:band10	1	24.2	24.2	16.6	4.74E-05
band2:band11	1	21.7	21.7	14.9	1.16E-04
band2:band12	1	48.1	48.1	32.9	9.89E-09
band3:band5	1	212.5	212.5	145.5	2.85E-33
band3:band6	1	75.7	75.7	51.8	6.55E-13
band3:band7	1	9.5	9.5	6.5	1.06E-02
band3:band8	1	93.6	93.6	64.1	1.31E-15
band3:band8a	1	23.8	23.8	16.3	5.55E-05
band3:band9	1	288.1	288.1	197.2	2.15E-44
band3:band10	1	5.0	5.0	3.4	6.35E-02

Table 5: Analysis of variance of Model 1 (part 1)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
band3:band11	1	2.4	2.4	1.6	1.99E-01
band3:band12	1	384.1	384.1	262.9	2.07E-58
band4:band5	1	186.5	186.5	127.7	1.98E-29
band4:band6	1	301.7	301.7	206.5	2.19E-46
band4:band7	1	58.8	58.8	40.3	2.29E-10
band4:band8a	1	26.6	26.6	18.2	2.02E-05
band4:band9	1	184.0	184.0	126.0	4.58E-29
band4:band11	1	251.4	251.4	172.1	5.31E-39
band4:band12	1	1.0	1.0	0.7	4.09E-01
band5:band6	1	102.8	102.8	70.4	5.50E-17
band5:band7	1	30.7	30.7	21.0	4.67E-06
band5:band8a	1	94.5	94.5	64.7	9.54E-16
band5:band10	1	56.3	56.3	38.6	5.50E-10
band5:band11	1	17.2	17.2	11.8	5.92E-04
band5:band12	1	25.5	25.5	17.4	2.99E-05
band6:band7	1	10.9	10.9	7.5	6.20E-03
band6:band8	1	39.5	39.5	27.1	2.01E-07
band6:band8a	1	118.9	118.9	81.4	2.13E-19
band6:band9	1	107.5	107.5	73.6	1.09E-17
band6:band10	1	18.3	18.3	12.5	4.03E-04
band6:band11	1	7.9	7.9	5.4	1.99E-02
band6:band12	1	43.3	43.3	29.7	5.27E-08
band7:band8	1	7.1	7.1	4.9	2.71E-02
band7:band8a	1	16.7	16.7	11.4	7.20E-04
band7:band9	1	106.8	106.8	73.1	1.40E-17
band7:band10	1	10.0	10.0	6.9	8.73E-03
band7:band11	1	212.7	212.7	145.6	2.62E-33
band7:band12	1	5.8	5.8	3.9	4.69E-02
band8:band8a	1	57.7	57.7	39.5	3.44E-10
band8:band9	1	40.0	40.0	27.4	1.73E-07
band8:band10	1	17.7	17.7	12.1	5.03E-04
band8:band11	1	0.0	0.0	0.0	9.58E-01
band8:band12	1	11.1	11.1	7.6	5.83E-03
band8a:band9	1	94.9	94.9	65.0	8.50E-16
band8a:band11	1	47.2	47.2	32.3	1.33E-08
band8a:band12	1	74.3	74.3	50.8	1.07E-12
band9:band11	1	41.1	41.1	28.1	1.16E-07
band9:band12	1	37.7	37.7	25.8	3.83E-07
band10:band11	1	14.3	14.3	9.8	1.75E-03
band10:band12	1	2.7	2.7	1.8	1.76E-01
band11:band12	1	48.0	48.0	32.9	1.00E-08
Residuals	10313	15063.7	1.5		

Table 6: Analysis of variance of Model 1 (part 2)

The analysis of variance (ANOVA) of Model 1 as can be seen in Tables 5 and 6 is used as a tool to determine the significance of a variable in the model. Most importantly, Pr(>F) values of less than 0.05 typically indicate a significant relation between the independent and the dependent variable.

	Coefficients
(Intercept)	1.387E+02
NDVI	-4.356E+02
Albedo	-1.892E+03
PSRI	7.245E+01
LWCI	9.292E-02
MSI	1.088E+03
GVMI	-4.506E+03
NDWI	6.510E+03
VSDI	-1.875E+01
SWCI	-4.947E+02
SIMI	3.573E+02
NDVI:Albedo	-1.688E+02
NDVI:PSRI	2.775E+02
NDVI:GVMI	5.686E+02
NDVI:NDWI	-3.290E+02
NDVI:VSDI	3.478E+02
NDVI:SIMI	2.822E+02
Albedo:PSRI	4.460E+02
Albedo:MSI	4.200E+02
Albedo:NDWI	1.227E+03
Albedo:VSDI	1.360E+03
Albedo:SIMI	1.280E+03
PSRI:MSI	-7.317E+01
PSRI:GVMI	-3.870E+02
PSRI:NDWI	-1.332E+02
PSRI:VSDI	8.508E+01
PSRI:SWCI	-3.103E+02
PSRI:SIMI	-4.834E+02
LWCI:MSI	-6.626E-02
LWCI:GVMI	-2.149E-01
MSI:GVMI	1.521E+02
MSI:VSDI	-9.051E+02
MSI:SIMI	-1.074E+03
GVMI:NDWI	-2.128E+02
GVMI:VSDI	4.207E+03
GVMI:SWCI	-1.630E+03
GVMI:SIMI	4.325E+03
NDWI:VSDI	-5.833E+03
NDWI:SWCI	8.650E+02
NDWI:SIMI	-6.565E+03
VSDI:SWCI	8.006E+02
SWCI:SIMI	3.286E+02

Table 7: Coefficients of Model 2

In Table 7 are the coefficients determined for Model 2. Generally, higher values indicate a larger weight in the model.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
NDVI	1	48062.5	48062.5	25035.6	0.00E+00
Albedo	1	6425.4	6425.4	3347.0	0.00E+00
PSRI	1	15533.9	15533.9	8091.5	0.00E+00
LWCI	1	0.1	0.1	0.0	8.56E-01
MSI	1	852.1	852.1	443.9	1.64E-96
GVMI	1	344.1	344.1	179.3	1.53E-40
NDWI	1	881.6	881.6	459.2	1.01E-99
VSDI	1	723.8	723.8	377.0	1.60E-82
SWCI	1	212.2	212.2	110.5	1.01E-25
SIMI	1	384.1	384.1	200.1	5.27E-45
NDVI:Albedo	1	149.1	149.1	77.7	1.40E-18
NDVI:PSRI	1	67.4	67.4	35.1	3.24E-09
NDVI:GVMI	1	271.3	271.3	141.3	2.22E-32
NDVI:NDWI	1	51.3	51.3	26.7	2.36E-07
NDVI:VSDI	1	19.4	19.4	10.1	1.46E-03
NDVI:SIMI	1	535.5	535.5	278.9	8.22E-62
Albedo:PSRI	1	68.5	68.5	35.7	2.43E-09
Albedo:MSI	1	3.8	3.8	2.0	1.61E-01
Albedo:NDWI	1	0.3	0.3	0.2	6.77E-01
Albedo:VSDI	1	201.1	201.1	104.7	1.83E-24
Albedo:SIMI	1	524.5	524.5	273.2	1.36E-60
PSRI:MSI	1	526.8	526.8	274.4	7.54E-61
PSRI:GVMI	1	81.0	81.0	42.2	8.75E-11
PSRI:NDWI	1	38.5	38.5	20.1	7.59E-06
PSRI:VSDI	1	0.1	0.1	0.1	7.87E-01
PSRI:SWCI	1	833.2	833.2	434.0	1.87E-94
PSRI:SIMI	1	0.0	0.0	0.0	9.08E-01
LWCI:MSI	1	0.1	0.1	0.1	8.18E-01
LWCI:GVMI	1	21.6	21.6	11.3	7.99E-04
MSI:GVMI	1	0.4	0.4	0.2	6.38E-01
MSI:VSDI	1	20.1	20.1	10.5	1.22E-03
MSI:SIMI	1	62.6	62.6	32.6	1.17E-08
GVMI:NDWI	1	60.6	60.6	31.6	1.97E-08
GVMI:VSDI	1	27.5	27.5	14.3	1.56E-04
GVMI:SWCI	1	190.6	190.6	99.3	2.79E-23
GVMI:SIMI	1	10.6	10.6	5.5	1.90E-02
NDWI:VSDI	1	391.1	391.1	203.7	8.82E-46
NDWI:SWCI	1	42.5	42.5	22.2	2.54E-06
NDWI:SIMI	1	95.9	95.9	50.0	1.66E-12
VSDI:SWCI	1	184.1	184.1	95.9	1.53E-22
SWCI:SIMI	1	9.1	9.1	4.8	2.93E-02
Residuals	10353	19875.4	1.9		

Table 8: Analysis of variance of Model 2

The analysis of variance (ANOVA) of Model 2 as can be seen in Table 8 is used as a tool to determine the significance of a variable in the model. Most importantly, Pr(>F) values of less than 0.05 typically indicate a significant relation between the independent and the dependent variable.