Once all of the relevant factors had been accounted for and coded the data was exported from the GIS for statistical analysis. This analysis was conducted using R 2.8 (R Development Core Team 2006). R was selected primarily due to the fact it was freely available through the Internet and because of its ability to accommodate both spatial and non-spatial analysis.

At the block level, all data were screened and cleaned. From the A4 block harvest data, four records were removed due to the fact that the vines were replaced mid study and similarly seven were removed from the B22 data. During the screening phase, a further five records were considered to be outliers due to their unusually high values (more than 2 standard deviations from the mean) and were removed from the combined A4 and B22 data set. This combined data set was subsequently evaluated using the procedure of Bilodeau and Brenner (1999) and confirmed to be normal.

From the pruning data set a total of 36 records were removed, 29 associated with the A4 block and seven with the B22 block. Of these points, 11 were removed due to the vines removal as noted above. However, an additional 25 points were removed from the A4 block, as there was an issue with data collection during the 2003 field campaign. From the combined pruning data, a further 11 records were deemed to be outliers and were removed. As the combined data set were not normally distributed, a square root transformation was performed and the resulting transformed data confirmed to be normal, again using the procedure of Bilodeau and Brenner (1999).

3.3.1.3 DATA ANALYSIS

The usual method of accounting for variability associated with sampling dates is through the use orthogonal comparisons across dates; this methodology underpins

BACI designs and analysis (Osenberg & Schmitt 1996). In this case, however, such an analysis was not possible due to the prevalence of missing values (Table 3.7) and the lack of control blocks. In this situation a more 'robust' solution is obtained by fitting two versions of the same model and manually partitioning the data to account for all degrees of freedom encapsulated by the sampling date (Ellem Pers. Com.). Hence, data were aligned by date and coded as indicated in Table 3.8.

Table 3.7. Data completeness matrix. Data are arranged by Vintage, with X indicating missing data, with the relevant block indicated in parenthesis. Middle line indicates the timing of mining activities.

	2004	2005	2006	2007	2008
Canes	√	x (A4)	√	√	x (B22)
Weight	√	x (A4)	√	√	x (B22)
Yield	√	x (B22)	√	√	x (B22)
50 BW	√	x (B22)	√	√	x (B22)
Brix	√	x (B22)	√	√	x (B22)
pH	√	x (B22)	√	√	x (B22)
TA	√	x (B22)	√	√	x (B22)

Table 3.8. Partitioning harvest data by dates to calculate and account for degrees of freedom for the entire data set (2004-8).

Before			After					BA = 1 Degree Freedom (DF)
(04	04)	05	(06	06)	(07	07)	08	Vintage/BA = (1) + (1 + 1) = 3 DF
A4	B22	A4	A4	B22	A4	B22	A4	Dates/Vintage/BA = (1) + (1+1) = 3 DF

Thus, data that were collected before mining activities were separated from those collected post-mining. Because viticulture activities do not follow a strict calendar year and because pruning activities are designed to drive harvest in the following season, data were grouped according to the year in which harvest occurred. Hence, data associated with 2004 pruning are associated with the 2005 vintage. Because

harvest data for blocks A4 and B22 were always collected on separated days, collection date serves as both a proxy for both the vineyard blocks and the grape variety.

Two slightly different multivariate statistical models were used to analyse the harvest data, namely

$$\text{Response} = BA + (\text{Vintage}/BA) + (Date/\text{Vintage}/BA) + (BA : zone) \qquad \text{(Equation 2)}$$

and

$$\text{Response} = BA + (\text{Vintage}/BA) + (Date/\text{Vintage}/BA) + (property : elevation) \qquad \text{(Equation 3)}$$

The first model (Equation 2) specified an interaction term (BA:zone) to compare response variables differences within the zones Before and After mining progression. Similarly, the second mode (Equation 3) specified an interaction term (property:elevation) to compare response variables differences between blocks A4 and B22 that could be attributed to changed in elevation from LMS.

3.3.2 Vineyard Scale

3.3.2.1 Environmental Monitoring

Electromagnetic (EM) induction soil surveys were conducted over all the vineyards throughout the monitoring region during vine dormancy (June - July), post-harvest (February – March) or veraison (October – November) periods of each year. Apparent soil conductivity (EC_a) measurements were taken from every fourth vine row using an EM38 (Geonics, Ontario, Canada) sensor linked to the dGPS (Pro-XL, Trimble, Sunnyvale, California USA). Data from *in-situ* capacitance probes installed at multiple depths (data not presented) suggested that the water absorbed during

rainfall events typically moved below one metre within a week of any given event. When operated in vertical dipole mode, the EM38 typically samples to a depth of 1.2 metres (Lamb, Pers. Com.), though this depends on the soil type, amount of compaction, and other factors. Hence, after two weeks EM38 readings are no longer influenced by water absorbed during previous rainfall events at this site.

As vineyard trellising is known to distort EC_a measurements, all EM38 data were acquired using the protocol of Lamb *et al.* (2005). Figure 3.10 shows when the surveys were conducted in relation to the total monthly rain. Table 3.9 indicates the mining status of relevant blocks at the time EM38 surveys were conducted. The mining progression is also highlighted graphically in Figure 3.11.

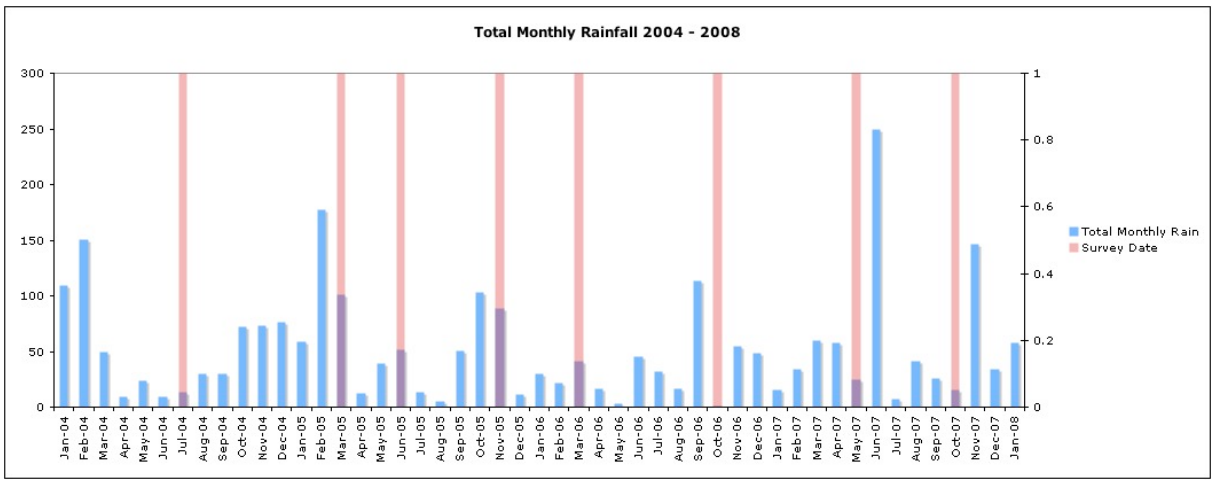

Figure 3.9. Monthly total rain for every year of the study. Redlines indicate months in which EM38 survey were conducted.

Table 3.9. Dates of vineyard scale field data collection, indicating the mining status of analysis blocks at the time of the survey

Block	5/7/04	10/3/05	15/6/05	7/11/05	15/3/06	31/10/06	7/5/07	22/10/07
A2	Before	Before	Before	Before	During	After	After	After
A3	Before	Before	During	During	After	After	After	After
A4	Before	Before	During	During	After	After	After	After
B16	Before	Before	Before	During	During	After	After	After
B17	Before	Before	During	After	After	After	After	After
B22	Before	Before	During	After	After	After	After	After

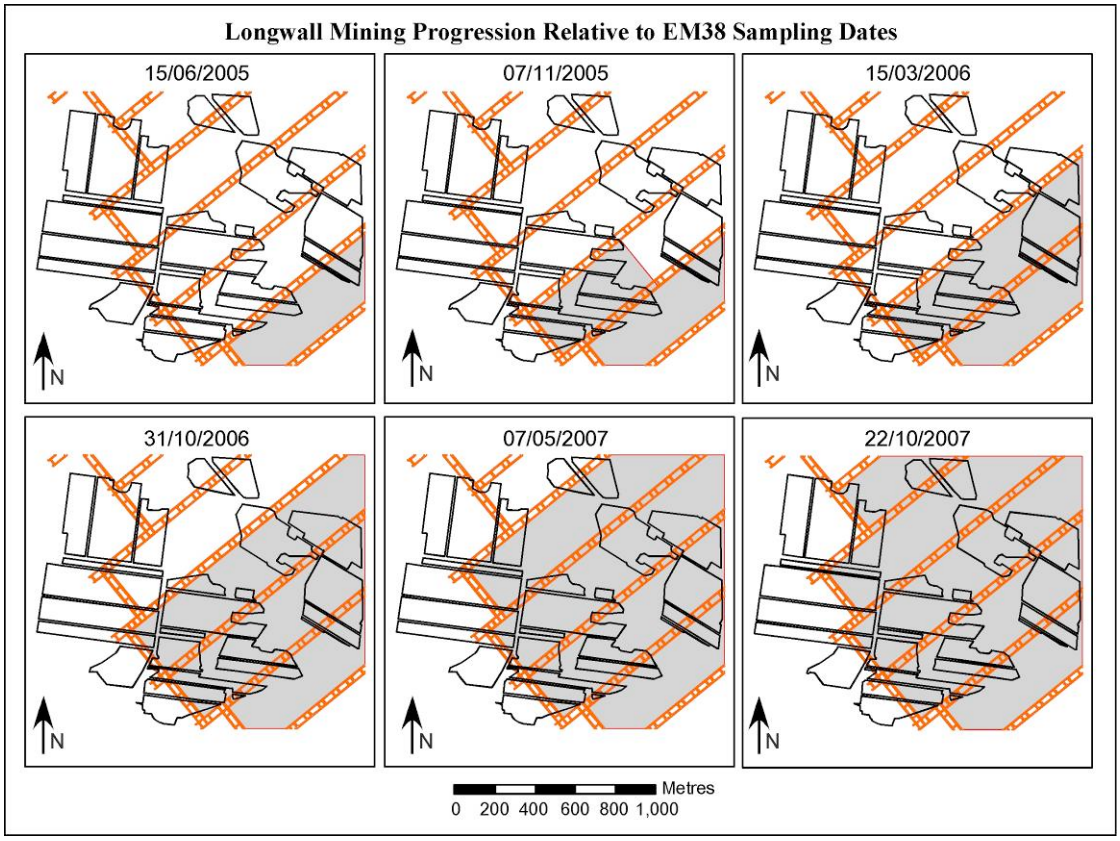

Figure 3.10. Progression of longwall mining through the vineyards relative to EM38 surveying. Grey indicates the extent of underground mining on the sampling date. Vineyards and longwalls as per Figure 3.2.

3.3.2.2 DATA PROCESSING AND SCREENING

The vineyard scale data were also screened for outliers. However, no data were removed from this data set; an examination of both 'unusually' high and low apparent

conductivity maps showed evidence of geographic patterns. The presence of these patterns suggests that these high and low values were indicative or real on-ground phenomena, and not merely the result of faulty sampling procedures. As such, treating these points as outliers would be inappropriate as it would have induced a geographic bias into the data sets.

However, the *buffer* function of the GIS was used to remove data that was within 5m of all block boundaries. This had two purposes. First, it helped minimized 'edge effects' induced by trellising systems as the EM38 moved in an out of vine rows. Secondly, it removed the influence of points repeated measurement on the compacted 'paths' between vineyards, which would have also skewed the results.

3.3.2.3 DATA ANALYSIS

Both qualitative, visual inspections of the spatial datasets and quantitative statistical methods were used to identify potential problems within the data. The analysis of the EM38 data focused exclusively on blocks from vineyards A and B, as these were the first to be undermined. However, due to the 'staggered' progression of mining through the vineyards, not all blocks from these vineyards were analysed; the inclusion of the most recently mined blocks would have introduced a temporal bias into the analysis, as most of the data from the excluded blocks were pre-mining. Also, several blocks from vineyard B were excluded due to their proximity to the location at which mining commenced; the methods used to delineate the various zones is likely to be less accurate in these locations, due to the fact that fewer subsidence observations are collected there (Section 3.3.2.1). Therefore, the analysis of the vineyard and regional scale data focused on blocks A2, A3, A4, B16, B17, and B22. A reference 'control' block C (Figure 3.2) was also used.

Although the original study design did not have a 'control' block, a relevant area was found to be more than 200 metres away from the minimum subsidence contour of 20 millimetres. As such, it should be relatively free of influence from LMS effects. The inclusion of this region allowed for BACI- type comparisons for the EC_a (vineyard scale) data with the impacted areas within the vineyards.

Using the slope zones, the vineyard level data for blocks A2, A3, A4, B16, B17, B22 and C were analysed together to see whether or not the slope before, during or after mining had any significant impact on soil apparent conductivity. As the data were not normally distributed, a square root transformation was applied to the data prior to analysis.

Using R (R Development Core Team 2006), a linear regression model was applied according to

$$\sqrt{EC_a} = \text{Vineyard} + \text{Block} + \text{Year} + \text{Season} + \text{Aspect} + \text{Slope} + (\text{Slope} : \text{BDA}) + (\text{Zone} : \text{BDA})$$

(Equation 4).

The model included terms to account for differences between the vineyards and individual blocks, as well as terms to account for seasonal and yearly variation as well as terrain differences evident in the slope or the aspect. To highlight differences associated with the timing of mining, interaction terms were included to examine difference in slope zones before-during-after mining as well as a term to examine the mining zones before-during-after mining. The data for each block were then partitioned according the slope categories of Darmody *et al.* (1988), and a series of interaction diagrams were examined to see whether or not the various slope zones

exhibited trends consistent with a differential ponding of water throughout the zones (e.g. slope zone differences in EC_a).

The overall trends in apparent conductivity for undermined blocks was also compared to those of the control block. A series of BACI interaction diagrams (e.g. Figure 2.10) were used to examine difference between the trends in the control block and those in the mining zones before, during and after mining. Differences in the magnitudes from before versus after mining progression could be indicative of possible LMS impacts. Slope zone data were also examined using similar BACI comparisons.

3.3.3 Regional Scale

3.3.3.1 ENVIRONMENTAL MONITORING

Regional scale data were collected in the form of multispectral and panchromatic satellite images. With an individual image encapsulating 64 km^2, Digital Globe's QuickBird satellite platform was selected for its ability to capture the entire study region in a single pass. Images from the Standard 2A product were used. Both radiometric and geometric corrections were applied to these products (DigitalGlobe, 2006) The resulting images had a spatial resolution of 2.44 – 2.88 metres (multispectral) and 0.61 – 0.71 metres (panchromatic). These resolutions were considered capable of directly monitoring vine PAB and hence vine vigour. QuickBird images were obtained once a year in the spring. Image capture was designed to coincide with berry ripening (veraison) in late October to early November. The mining status of relevant blocks at the time of image capture is detailed in Table 3.10, and is presented graphically in Figure 3.12.

Table 3.10. Dates of the regional scale field data collection, indicating the mining status of relevant analysis blocks at time of QuickBird image capture.

Block	26/10/2003	03/11/2004	11/11/2005	14/11/2006	30/10/2007
A2	Before	Before	Before	After	After
A3	Before	Before	During	After	After
A4	Before	Before	During	After	After
B16	Before	Before	During	After	After
B17	Before	Before	After	After	After
B22	Before	Before	After	After	After

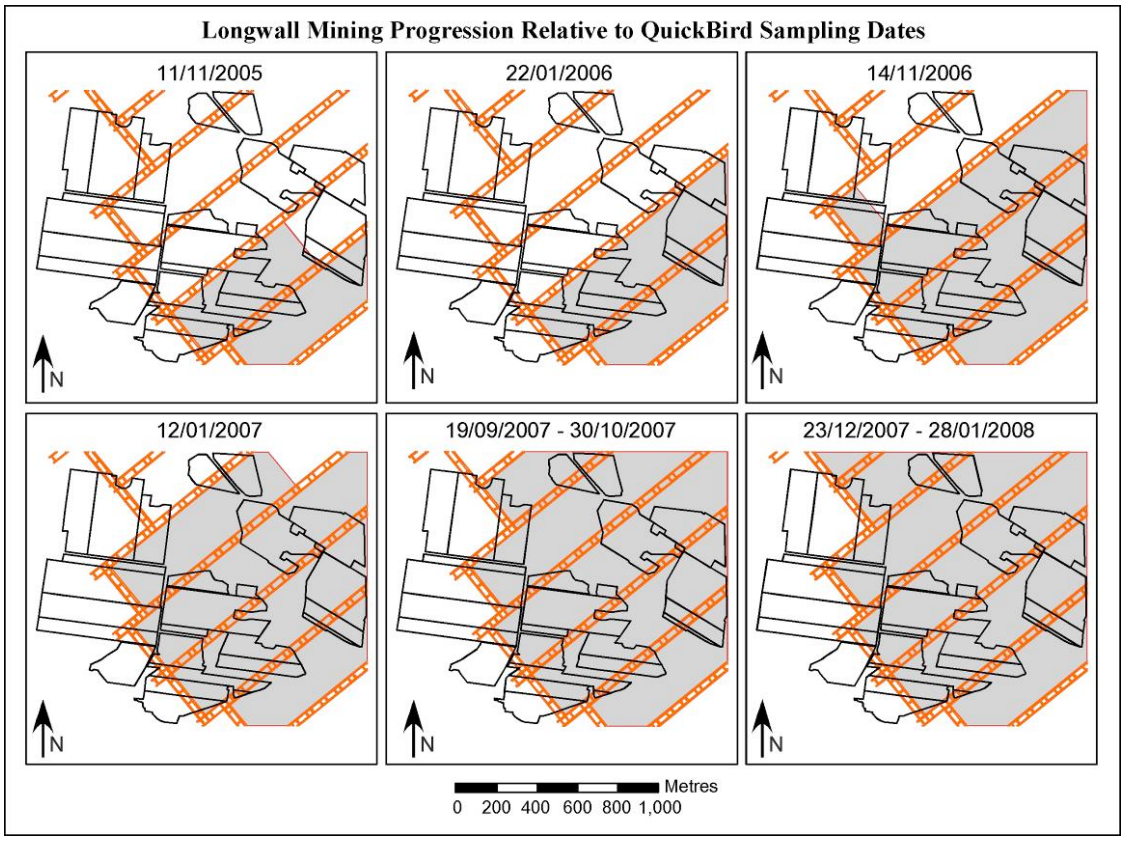

Figure 3.11. Progression of longwall mining through the vineyards relative to the capture of QuickBird imagery. Grey indicates the extent of underground mining on the sampling date. Vineyards and longwalls as per Figure 3.2.

3.3.3.2 DATA PROCESSIING AND SCREENING

All five of the QuickBird satellite images acquired during this study were orthorectified using Erdas Imagine 9.1 (Erdas 2007). These images were referenced to map coordinates using a high resolution, orthophotograph obtained by AAMHatch in 2004.

The spatial accuracy of the aerial photograph was independently tested by comparing the location of 30 well-distributed points within the aerial photograph with their real-world spatial locations as measured with the dGPS. In all cases, these points were less than one metre from their dGPS determined location.

In this process, a 5 metre DEM was used as the reference surface for the rectification process. This DEM was generated from 10 metre contour data for the region obtained from the NSW Department of Lands. A drainage enforce DEM was then generated using the *TopoRaster* function of ArcGIS's Spatial Analyst extension (ESRI 2007). From this 5 metre DEM new contours were generated and compared with the originals. Although this process appeared to generate some artefacts, particularly in the flatter areas of the terrain, overall the new contours were in good agreement with the original ones (Figure 3.10).

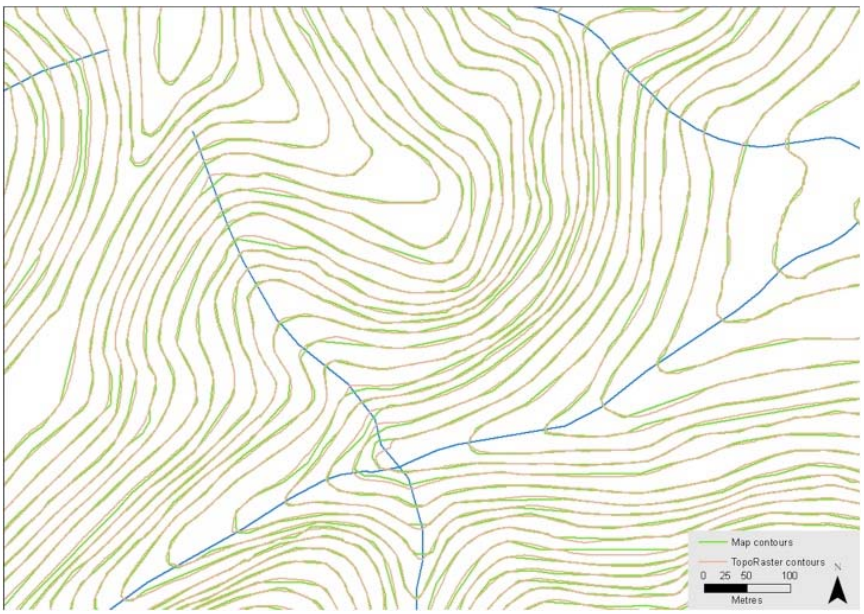

Figure 3.12. Comparison of original contours with contours from the TopoRaster generated 5 m DEM. Image courtesy of Ross Jenkins

The images were subsequently orthorectified against the 2004 orthophotograph using the five metre TopoRaster generated DEM. For each image, a minimum of 20 well-distributed ground control points (GCP) were used to spatially locate the satellite

images. As the root mean square (RMS) error of each GCP was always observed to be less than one, which meant that the every feature evident within the image should be within one pixel of its location on the orthophotograph. Hence, for the panchromatic image, every feature should be within 0.7 metres of its corresponding location, and every feature evident in the multispectral image should be within 2.4 metres. Because the panchromatic and multispectral images were captured simultaneously, the same GCP points were used to rectify both images, effectively rendering the spatial accuracy of the multispectral image comparable to that of the panchromatic image (Table 3.11 – See Appendix 1 for a more details on the rectification process).

The accuracy of the rectification process was independently verified using a minimum of 20 well-distributed check points (CP). Like the GCPs, the RMS error of each of the CP was less than one, indicating that each verification feature was within 0.7 metres (panchromatic) or 2.4 metres (multispectral) of its corresponding location on the reference orthophoto. Hence, accounting for GPS errors each pixel on the images should be within 1.7 metres (panchromatic) or 3.4 metres (multispectral) of its actual on-ground position.

These orthorectified satellite images were subsequently converted to a series of normalised difference vegetation index (NDVI) maps based on the per-pixel calculation using

$$NDVI = \frac{NIR - R}{NIR + R}$$ (Rouse *et al.* 1974) (Equation 4)

where NIR and R are, respectively, reflected electromagnetic radiation in the near infrared and red wavebands. The NDVI ratio was selected, as it is a standard ration

that is widely known. Furthermore, the ratio reduces distortions from several potential sources of noise such as ‘sun illumination differences, cloud shadows, some atmospheric attenuation, and some topographic variations’ (Jensen, 2005 p314). Although the NDVI can be impacted by other forms of noise (e.g. additive effects of atmospheric path radiance – Huete *et al*., 2002), it is reasonably resistant to man atmospheric effects. Whilst NDVI values of thin and broken vegetation have been shown to be significantly impacted by atmospheric effects (as much as 50% - Jensen *et al.,* 2002), it is not clear that these factors are relevant to viticultural settings.

Table 3.11. Satellite imagery detail and orthorectification accuracy.

Date	Image Type	Off-Nadir (°)	Point Type	Pixel Size (m)	Mean RMS	Mean Error (m)
26-10-2003	Pan	1.6	Control	.6	.587	0.35
			Check	.6	.479	0.29
	MS	1.6	Control	2.4	.147	0.35
			Check	2.4	.129	0.31
06-01-2004	Pan	4.7	Control	.6	.583	0.35
			Check	.6	.644	0.39
	MS	4.7	Control	2.4	.146	0.35
			Check	2.4	.161	0.39
03-11-2005	Pan	18.3	Control	.6	.515	0.31
			Check	.6	.656	0.39
	MS	18.3	Control	2.4	.124	0.30
			Check	2.4	.171	0.41
04-03-2005	Pan	2.5	Control	.6	.495	0.30
			Check	.6	.541	0.32
	MS	2.5	Control	2.4	.129	0.31
			Check	2.4	.130	0.31
11-11-2005	Pan	23.6	Control	.6	.531	0.32
			Check	.6	.687	0.41
	MS	23.6	Control	2.4	.130	0.31
			Check	2.4	.170	0.41
22-01-2006	Pan	19.1	Control	.6	.585	0.35
			Check	.6	.571	0.34
	MS	19.1	Control	2.4	.150	0.36
			Check	2.4	.140	0.34
14-11-2006	Pan	21.0	Control	.6	.586	0.35
			Check	.6	.700	0.42
	MS	21.0	Control	2.4	.155	0.37
			Check	2.4	.177	0.42
12-01-2007	Pan	23.6	Control	.6	.655	0.39
			Check	.6	.633	0.38
	MS	23.6	Control	2.4	.335	0.80
			Check	2.4	.276	0.66
19-09-2007	Pan	22.8	Control	.6	.643	0.39
			Check	.6	.661	0.40
	MS	22.8	Control	2.4	.158	0.38
			Check	2.4	.167	0.40
30-10-2007	Pan	22.6	Control	.6	.495	0.30
			Check	.6	.543	0.33
	MS	22.6	Control	2.4	.126	0.30
			Check	2.4	.136	0.33
23-12-2007	Pan	2.5	Control	.6	.603	0.36
			Check	.6	.670	0.40
	MS	2.5	Control	2.4	.154	0.37
			Check	2.4	.172	0.41
28-01-2008	Pan	3.7	Control	.6	.423	0.25
			Check	.6	.608	0.36
	MS	3.7	Control	2.4	.103	0.25
			Check	2.4	.159	0.38

The regional scale data were also screened for outliers. However, no data were removed from this set; again, an examination of both 'unusually' high and low apparent NDVI maps showed evidence of geographic patterns. The presence of these patterns suggests that these high and low values were indicative or real on-ground phenomena, and not merely the result of faulty sampling procedures. As such, treating these data as outliers would be inappropriate as it would have induced a geographic bias into the data sets.

3.3.3.3 DATA ANALYSIS

As with the EM38 data, both qualitative visual inspections of the spatial datasets and quantitative statistical methods were used to identify problems discernable within the data. The regional scale NDVI data were also analysed in R (R Development Core Team 2006) using a linear regression model

$$\sqrt{\text{NDVI}} = \text{Vineyard} + \text{Block} + \text{Vintage} + \text{Aspect} + (\text{Slope} : \text{BDA}) + (\text{Zone} : \text{BDA}) \quad \text{(Equation 5)}$$

to account for differences between the vineyards and between the blocks. Yearly differences were also accounted for, as well as differences relating to aspect. As with the vineyard scale data, to highlight differences associated with the timing of mining, interaction terms were included to examine difference in slope zones before-during-after mining as well as a term to examine the mining zones before-during-after mining.

As with the EM38 data, a series BACI of interaction diagrams were used to examine difference between the trends in the control block and those in the mining zones before, during and after mining. Differences in the magnitudes from before versus after mining progression could be indicative of possible LMS impacts. As with the

EM38 data, comparisons slope zone comparisons were also made before and after mining.