Characterising soil moisture in transport corridors using remote sensing

Andrew J Hardy

BSc (Hons) Geography
MSc Environmental Monitoring,
Modelling and Reconstruction

Thesis Submitted for the Degree of
Doctor of Philosophy

School of Civil Engineering and Geosciences
Newcastle University

November 2010

Abstract

This thesis assesses the ability of remote sensing techniques to characterise soil moisture in a transport corridor environment. Much of the world's transport networks are built on earthwork embankments or in cuttings. In the UK, many of these earthworks were constructed in the mid-19th Century and are susceptible to slope instability. Instability in transport corridors is often triggered by an increase in pore pressure, which is directly influenced by an increase in soil moisture. Although a number of studies have investigated the use of remote sensing techniques for estimating soil moisture, they have tended to be conducted under controlled conditions and few have considered their capacity for being operational. This study addresses this point by exploring the use of high spatial resolution digital elevation models (DEMs) and airborne hyperspectral imagery for characterising soil moisture in transport corridors.

A number of terrain (topographic wetness index (TWI), potential solar radiation, aspect) and spectral analysis (red edge position estimation, derivative stress ratios, continuum removal analysis, partial least squares (PLS) regression modelling, mapping biological indicator values) techniques were assessed using terrestrial systems over a test embankment, and airborne data for a transport corridor. The terrain analysis metrics TWI and potential solar radiation were found to be highly sensitive to the DEM spatial interpolation routine used, with a thin plate spline routine performing best in this study. This work also demonstrated that Ellenberg indicator values extended for the UK can be mapped successfully for transport corridor environments, providing potential for a number of different applications. Individually, the techniques were shown to be generally poor predictors of soil moisture. However, an integrated statistical model provided an improved characterisation of soil moisture with a coefficient of determination (R^2) of 0.67.

Analysis of the model results along with field observations revealed that soil moisture is highly variable over the transport corridor investigated. Soil moisture was shown to increase in a non linear fashion towards the toe of earthwork slopes, while contribution from surrounding fields often led to concentrations of moisture in cutting earthworks. Critically, while these patterns could be captured using the data investigated in this study, such spatial variability is rarely taken into account using analytical slope stability models, potentially raising important challenges in this respect.

Acknowledgements

First and foremost I would like to thank my supervisors Jon Mills and Stuart Barr for their guidance throughout my PhD. In addition, the support and advice from Pauline Miller, Martin Robertson, Meredith Williams and Mike Lim has been warmly received.

Thanks to Steve Hall (Network Rail) for providing access to the study area around Haltwhistle and Paul Hughes (Newcastle University) for access to the test embankment at Nafferton Farm.

Acknowledgement is given to the Engineering and Physical Sciences Research Council (EPSRC) for funding this study and to the data providers Network Mapping and the Geomatics Group at the Environment Agency for fielding my many queries.

Additional thanks to Iain Bramwell for collecting the terrestrial laser scanning data, Pauline Miller for processing the lidar data and Clare O'Reilly for training in identifying grass species.

Ultimately I would like to thank my family and my partner Marie for their continued support. This work is dedicated to them.

Table of contents

List of figures

List of tables

List of abbreviations and notations

AIC	Akaike Information Criterion
BIONICS	BIOlogical and eNgineering Impacts of Climate change on Slopes
CASI	Compact Airborne Spectrographic Imaging
DEM	Digital Elevation Model
EMR	ElectroMagnetic Radiation
FWHM	Full Width Half Maximum
GCP	Ground Control Point
GIS	Geographical Information System
GPS	Global Positioning System
GWR	Geographically Weighted Regression
ha	Hectares
IDW	Inverse Distance Weighted
LAI	Leaf Area Index
Lidar	LIght Detection And Ranging
ln	Natural logarithm
Lowess	LOcally WEighted Scatterplot Smoother
LSME	Land Surface Microwave Emission model
MHz	Mega Hertz
MNF	Minimum Noise Fraction
NDVI	Normalised Difference Vegetation Index
OLS	Ordinary Least Squares
PCA	Principal Components Analysis
PLS	Partial Least Squares
PPS	Precise Positioning System
PSF	Point Spread Function
RADAR	RAdio Detection And Ranging
REP	Red Edge Position
RMSE	Root Mean Square Error

RTK	Real Time Kinematic
SAR	Synthetic Aperture Radar
SCA	Specific Catchment Area
SM	Soil Moisture
TLS	Terrestrial Laser Scanning
TVDI	Temperature-Vegetation Dryness Index
TWI	Topographic Wetness Index
WH/m^2	Watt Hours per Metre Square
λ	Wavelength
nm	Nanometer
μm	Micrometer

1 Introduction

1.1 Instability in transport corridors

Much of the world's transport networks are located on cuttings and embankments which are all susceptible to instability (Lloyd et al., 2001). About 60% of UK rail networks are composed of embankment and cutting earthworks, equating to 16 000 kilometres (Manning et al., 2008). Many of these earthwork structures were constructed in the mid-19th century before the development of modern theories of soil mechanics and were therefore poorly compacted (Ridley et al., 2004).

Soil cuttings were crudely excavated by pick and shovel, or blasted using gunpowder, and embankments were constructed by tipping soil and rock into place using horse-drawn wagons. In the early part of the 20th century methods improved with steam-powered excavators replacing the pick and shovel, and steam locomotives replacing the horse and cart to tip material onto embankments (Figure 1.1) (Perry et al., 2003b). Despite these efforts, this offered little in the way of soil compaction that can be seen in today's modern earthwork construction. In addition, embankment slope angles were based on short-term angles of repose which in the modern age would be considered too steep (Perry et al., 2003b). As such, many of these structures failed during, or soon after, construction (Ridley et al., 2004; Perry et al., 2003b). Subsequent remediation seldom fully excavated and replaced these structures, leaving residual rupture surfaces such that the earthworks remain a significant hazard (Ridley et al., 2004; Perry et al., 2003b).

Historically, transport earthworks have not been treated as an asset to network operators and have received a low priority for maintenance and renewal (Perry et al., 2003b). As a result, engineering performance has suffered, increasing the risk of instability. More recently transport earthworks have received greater awareness within the UK transport industry, with embankments and cuttings comprising one third of total asset value for transport infrastructure (Ridley et al., 2004). Furthermore, it has been noted that transport earthworks may support additional assets, including road paving, lights, drainage, trackbed, or third party structures such as telecommunications masts (Perry et al., 2003b). These assets represent a significant commodity and

consequentially the cost of replacing these assets, due to earthwork instability, can far outweigh the cost of initially repairing the structure (Perry et al., 2003b).

Figure 1.1 Embankment construction in 1922 using a steam powered locomotive to tip material on the earthwork from Perry et al. (2003b).

As the UK's transport infrastructure ages, the risk of earthwork instability increases (Perry et al., 2003a; Phipps and McGinnity, 2001). Heavier loading and increasing passenger numbers have added pressure to the aging earthworks with detrimental effects (Perry et al., 2003b). In some areas of the world, notably South East Asia, rapid urban expansion has led to an increase in the pressure on available land (Kwak et al., 2006). In some cases has led to building construction upon transport earthworks, which has led to increasing concern over their stability (Zhang et al., 2005). Climate change is expected to further increase the risk of earthwork instability, with research suggesting that the UK climate is becoming more extreme, manifested by an increase in the magnitude of rainfall events (Ekström et al., 2005).

There are a number of examples in the UK where a combination of extreme rainfall events and aging earthworks have resulted in slope failures. On 2nd January

2003, a train carrying 105 passengers derailed on the Brighton to London line when it hit a landslide. The landslide was attributed to heavy rainfall that had persisted for several days before the event. Although local residents had previously expressed concern over the weakness of the cutting slopes, the infrastructure operator Network Rail had failed to reinforce the earthworks despite flagging up the area as vulnerable (Guardian, 2003).

This example of instability and others like it have immediate implications for passengers including extensive delays and risk to life but also an increase maintenance costs, risk of power failure, interruption to normal train operations and environmental degradation (Gyasi-Agyei et al., 2001). In light of this, the last two decades have seen an increase in the demand for improved reliability and safety on transport networks. This has resulted in a growing focus on appraisal and repair of embankments and cuttings, highlighted by an increase in spending. For example, over the period 1998 to 1999, the UK railway infrastructure operator Railtrack (now Network Rail) spent over £50 million was spent on earthwork maintenance (Perry et al., 2003b).

Despite the inherent problems with the UK's transport infrastructure, it is not economically or practically viable to replace earthwork structures (Ridley et al., 2004). In some cases remediation is undertaken solely to maintain transport operations and do not correct for permanent or ongoing deformations (Gyasi-Agyei et al., 2001). One of the reasons for this is that delays for even short periods of time, to allow for repair, can far outweigh remedial costs (Lloyd et al., 2001). In the past, unstable earthwork structures were met with a reactive response but there is a growing awareness to be proactive (Perry et al., 2003b).

The cost of unplanned, reactive repair to transport earthworks greatly exceeds the cost of proactive and continual maintenance (Perry et al., 2003b). For example, McGinnity et al. (1998) state that the cost of embankment remedial works fell from £3000-5000 per metre of track to £1000-2000 per metre when part of a proactive maintenance strategy. Similar examples can be found in highway maintenance where it is thought a proactive maintenance strategy for the M23 motorway in Surrey led to savings of about 80% compared with works carried out based on a reactive response (Perry et al., 2003b). It is therefore important to highlight areas prone to instability which can be achieved through risk-mapping.

Considered to be the most basic tool for reducing risk of instability, risk maps can provide vital information for administrators and policy makers (Yesilnacar and Suzen, 2006; Schulz, 2005). As a result, landslide hazard assessment has become a topic of major interest (Xie et al., 2003) with much of the research focussing on catastrophic deep-seated landslides. However, there is a growing body of research that studies shallow slope failures, typically 1-1.5 metres deep, which is the mode of failure that mainly effects transport earthworks (Fourie, 1996).

1.2 Instability and the role of soil moisture

The principal causes of earthwork instability include burrowing animals, vandalism, embankment age, construction type, change in earthwork materials, culvert deterioration, slope geometry, angle and height, moisture and vegetation shrink swell cycles, and the presence of water (Perry et al., 2003b). Many of these features can be clearly identified by experienced network inspectors, but some can be harder to spot. Specifically, this is the case where dense vegetation can mask instability processes at the earthwork surface, such as shallow slope failures. In the case of such failures, it is important to identify them early to minimise their detrimental effects (Perry et al., 2003b).

Shallow slope failures are one of the most common slope processes, and are found all over the world (Beguería, 2006). Such failures are generally associated with areas that experience prolonged rainfall, particularly in the sub-tropics, but there is growing evidence to suggest that compacted earthworks within the UK are becoming increasingly susceptible to such failures (Fourie, 1996). The trigger for most of these failures is a build up of pore water pressure (Beguería, 2006; Borga et al., 2002; Fukuoka, 1980), which is directly related to heavy or prolonged rainfall events and elevated soil moisture content (Collins and Znidarcic, 2004; Johnson and Sitar, 1990). This is illustrated by Figure 1.2 which shows the relationship between volumetric soil water content and pore water pressure (here denoted as matric suction), for different soil textures. As a result of this relationship, some failures have become annual events associated with wetter seasons (Baum et al., 2005; Lloyd et al., 2001).

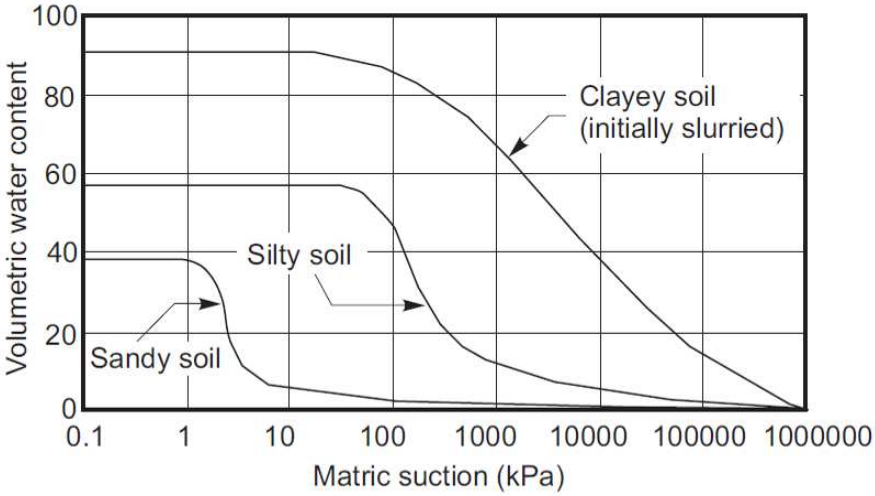


Figure 1.2 Diagram demonstrating the relationship between soil moisture and matric suction (pore water pressure) for different soil textures. From Fredlund and Xing (1994)

The intrinsic relationship between soil moisture and slope stability has led some researchers to suggest that areas of higher soil moisture content require less precipitation to trigger a landslide event (Manning et al., 2008; Fourie, 1996). This is particularly the case for railway embankments that have little or no drainage, allowing rainfall to percolate directly into the earthwork (Perry et al., 2003b). In fact, many earthwork problems are a result of poor or defective drainage, with just small improvements to drainage leading to significant improvements to slope stability (Perry et al., 2003b). This notion is reinforced by Ridley et al. (2004) who identified a strong correlation between earthwork failure events and reduced soil moisture deficit for railway embankment soils in the London area (Figure 1.3).

Slope instability in embankments can also be manifested through the cess heaving process (Perry et al., 2003b). This is where infiltration of water into an earthwork softens the clay foundations and is forced from beneath the track, therefore weakening the soil above. The effects of excess water infiltration can be exacerbated if water is shed from surrounding areas down cutting slopes (Perry et al., 2003b). An example of this was seen on 26th October 2008 when a night of heavy rainfall resulted in flood water spilling onto a cutting resulting in the failure of a ten-metre section with a further 20 metre section put at risk in Derbyshire, closing the Manchester to Sheffield

line indefinitely (MEN, 2008). Highway embankments are generally less prone to the effects of excessive water infiltration as they were constructed more recently than railway embankments with higher specifications, as well as a largely impermeable top layer. However, defective drainage, such as the example given in Derbyshire, can concentrate water onto a slope leading to instability (Perry et al., 2003b).

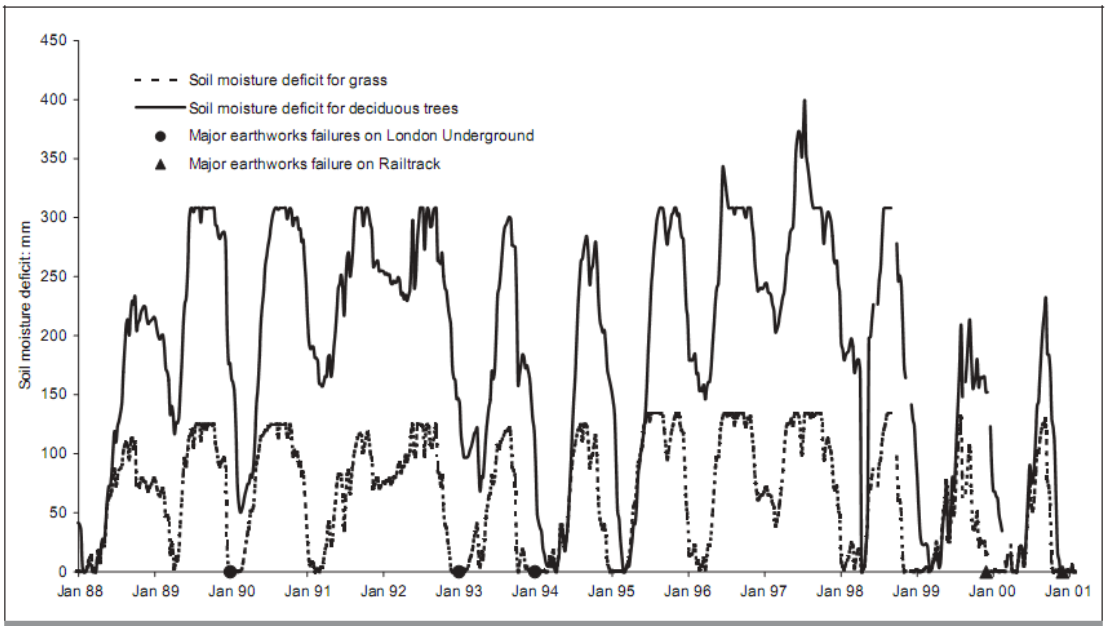


Figure 1.3 Variation of soil moisture deficit and major earthwork events for the London area for the period 1988-2001. From Ridley et al. (2004)

1.3 Measuring soil moisture

Network operators have a statutory obligation to conduct regular appraisals of earthwork stability (Perry et al., 2003b). As part of their appraisal, earthwork inspectors make subjective observations of soil moisture conditions based on descriptions of drainage type and conditions, and identifying potential sources of water ingress (Perry et al., 2003b). Additionally, they are advised to look for signs of failure by identifying seepage or ponding of water (Mickoviski and Van Beek, 2006). Similarly, inspectors try to identify potential future instability by looking for the presence of hydrophilic, water-loving vegetation, or changes in the vegetation, including rich dense vegetation that may indicate seepage (Perry et al., 2003b).

Although the collection of *in situ* soil moisture contents would be preferable, this process is time consuming and laborious. Traditional methods for measuring soil moisture involve collecting samples of soil, then measuring the weight of the sample before and after drying in an oven for 24 hours to calculate the gravimetric moisture content. Even with the advent of moisture impedance probes that have the ability to rapidly measure soil moisture in the field, both methods involve lengthy periods in the field by a number of inspectors. Furthermore, onsite inspection of embankment and cutting slopes carry obvious dangers, as well as the need for train drivers to reduce their speed in case of any impending danger. As a result, some authors have identified the use remote sensing to aid earthwork remediation as an important topic for future research (Perry et al., 2003b).

1.4 Potential for remote sensing

Although remote sensing techniques would never replace ground-based earthwork inspection, they have the potential to provide essential information on areas of known deterioration, as well as providing broad assessment of larger areas at the transport corridor scale (Perry et al., 2003b). As a result, aerial photographs are routinely used for making appraisals and providing historical information about an earthwork and its surrounding area (Perry et al., 2003b). However, more advanced remote sensing techniques have the potential to detect features not visible to the naked eye, including soil moisture (Perry et al., 2003b).

Studies have shown that surface soil moisture conditions are highly variable over time and space (Bárdossy and Lehmann, 1998). On site inspectors are expected to cover 5 km lengths of track a day when carrying out assessments of earthwork stability (Perry et al., 2003b) but this could result in misrepresentation of the soil moisture conditions as moisture patterns can change over periods of just hours. Furthermore, some earthwork structures, especially in the case of rural railway networks, can be difficult to access (Perry et al., 2003b), making on site assessments arduous and in some cases impossible. Remote sensing has the potential to provide a synoptic view over transport corridors, including those in a remote location where transport links can be vital links to the wider community (Lloyd et al., 2001).

Current methods for the remote inspection of transport corridor assets include video filming from a moving vehicle, radar, thermal imaging, low and high altitude aerial photography, and the use of satellite imagery (Perry et al., 2003b). It is thought that the use of such techniques is growing due to the improvements made to interpretive techniques, but authors note the need for further research in this area (Perry et al., 2003b).

The transport industry has identified the growing demand for digital data (Perry et al., 2003b). As part of this drive, network operators are increasingly making use of Geographical Information Systems (GIS) for the effective management of spatial data, where traditional paper based information systems are difficult to access and visualise (Perry et al., 2003b). Additionally, GIS can overlay different maps of interest (Perry et al., 2003b), as well as the ability to calculate risk of slope instability based on different spatial components (Dai et al., 2002). Modern remote sensing techniques often require processing and analysis within GIS software, providing digital data that can be integrated with other geographical information. GIS software therefore provides a platform for processing and storing soil moisture information as well as the ability to produce a meaningful map of soil moisture distribution for network operators.

1.5 Aims and objectives

The aim of this study is to explore the potential for characterising soil moisture in transport corridors using remotely sensed data. To achieve this aim the following objectives are addressed:

1. Identify and review existing techniques for characterising soil moisture from remotely sensed data and critically assess their operational capability in the context of transport corridor environments.
2. Carry out a proof of concept study for the techniques deemed applicable in objective 1, using ground based instruments at an earthwork scale and assess their potential for being scaled up to a transport corridor scale.
3. Based on the findings made in objectives 1 and 2, apply the techniques to a transport corridor using airborne remotely sensed datasets.
4. Design a method for integrating the most successful techniques into an overall model for characterising soil moisture in transport corridor environments.

1.6 Thesis structure

The main aim and key objectives, outlined in Section 1.5, are addressed in this thesis and are presented over the following six chapters. Chapter 2 presents a thorough review of current remote sensing techniques with the potential for characterising soil moisture. The techniques presented are critically assessed with reference to their ability to be operationally employed within a transport corridor environment.

Based on the findings of Chapter 2, a proof of concept study is carried out for several of the potentially most promising operational remote sensing methods. In this study, the techniques are evaluated over a test embankment under controlled conditions. This evaluation is carried out to gain a quantitative understanding of the relationships that exist between the distribution of soil moisture and measurements made using ground based remote sensing instruments. The techniques are also assessed for their potential to be scaled up to airborne sensors for a transport corridor.

The findings made in the embankment scale study are then used to set out the optimal methods for applying the techniques to airborne remotely sensed dataset for a real world transport corridor. This also includes techniques that could not be tested in the small scale study. The main product of the transport corridor scale work is an integrated model incorporating techniques that provide an improved characterisation of soil moisture distribution.

The methods for both the earthwork and transport corridor scale studies are reported in Chapters 3. The results of these methods are described in Chapters 4 and 5 for the earthwork and transport corridor scale studies respectively. These results are analysed and discussed in Chapter 6, with particular emphasis on a critical assessment of the results obtained and how they contribute to science. The key findings are concluded in Chapter 7, including suggestions for the direction of future research.

2 Review of techniques for characterising soil moisture using remote sensing

2.1 Introduction

Section 1.4 of this thesis outlined the potential for using remote sensing to monitor transport earthwork infrastructure. However, it also acknowledged that the transport industry has not fully embraced the potential of remotely sensed data, particularly for characterising soil moisture as an indicator of earthwork instability. This chapter provides a thorough critique of remote sensing techniques that have the potential to characterise soil moisture, with specific reference to their utility within transport corridor environments.

Remote sensing can help make observations about the Earth's surface in a number of different ways, according to the sensor that is being used, or the platform on which it is stationed. With regard to remote sensing of soil properties, some sensors can penetrate vegetation canopies and even the surface of the ground, to make direct measurements from the soil itself. Others can be used to make a measure of the Earth's surface, for instance elevation or topography, which can then be used to make indirect inferences about the soil below. Additional techniques involve making inferences from surrogate measurements, such as observations of vegetation characteristics. In light of this, the review is split into three main sections, based on whether the specific technique characterises soil moisture directly (Section 2.2), indirectly using terrain analysis (Section 2.3), or through surrogate measures using vegetation reflectance (Section 2.4).

2.2 Remote sensing of soil moisture

Microwave remote sensing is considered to be the most promising technique for the remote measurement of soil moisture over wide areas (Famiglietti et al., 1999). The advantage of using microwave sensors is that they can make observations at both day and night, and are relatively unaffected by the presence of cloud (Cracknell and Hayes, 2007). Furthermore, microwaves (with a spectral range of 1 mm to 1 m) can, under some circumstances, penetrate vegetation canopies and therefore obtain measurements relating directly to the soil surface. Microwave sensors can be divided into two distinct groups: passive sensors that detect microwave emissivity from the Earth's surface, and

active sensors that transmit a microwave signal to the ground and record the proportion of energy backscattered to the instrument (Campbell, 2006). This section will discuss the use of microwave remote sensing for measuring soil moisture using both passive and active sensors. In addition, this section discusses the use of optical sensors such as thermal imagers and sensors measuring spectral reflectance in the visible, near, and shortwave infrared. These sensors record reflected electromagnetic radiation which can be used to make inferences about Earth surface properties, such as soil moisture. As these inferences are made from a direct measure from the soil, i.e. reflected electromagnetic radiation, they are also considered in this section.

2.2.1 Passive microwave sensors

Passive microwave sensors measure emissions of microwave energy from the Earth's surface (Cracknell and Hayes, 2007). Water within a soil matrix increases the dielectric permittivity, this being the ability of a medium to conduct electrical energy (Sharkov, 2003). This means that a wetter soil has greater ability to absorb electromagnetic radiation which therefore lowers the amount of microwave emissivity (Schmugge et al., 2002). The ability of a material to conduct electrical energy is sometimes referred to as the dielectric constant. Water is known to have a much higher dielectric constant (approximately 80) compared to dry soils (< 5), which translates to a difference in emissivity of approximately 0.95 for dry soils and 0.6 for wet soils (Schmugge et al., 2002). The influence of water on the dielectric constant means that measures of microwave emissions can be used to measure soil moisture.

The measure of microwave emissivity from a soil is based on the scattering of electromagnetic radiation from the soil surface. In a soil, scattering at the surface is mainly a function of the dielectric constant and the roughness of the soil surface. In addition, shadowing caused by surface roughness and the relative view angle of sensor has also shown to have an influence on the degree of recorded emissivity (Kuria et al., 2007). The effects of shadowing and surface roughness are comparable, and in some cases greater, than the effect of soil moisture on the degree of backscatter (Kuria et al., 2007). For example, a rough soil surface has shown to typically reduce the microwave signal by 10% or 20%, and in some extreme cases the signal can be halved (Schmugge et al., 2002). Therefore, in order to make inferences of soil moisture content, the effects

of surface roughness must be accurately estimated (Shi et al., 2005). These estimates are based on empirical or physical surface emission models. Each have relative advantages and disadvantages in terms of complexity, ease of computation, effectiveness over vegetation, and transferability (for further details see Kuria et al. (2007), Crosson et al. (2005), Shi et al. (2005), and Wigneron et al. (1998)).

One of the main restricting factors on the use of passive microwave sensors is that the intensity of microwave emissivity from the Earth's surface is very low. This means that in order to record an adequate signal of microwave emissivity, the sensor must remain positioned over the area of interest for several days. However, components such as soil moisture are temporally variable and can change over a matter of hours (Famiglietti et al., 1998). This inevitably leads to a great deal of uncertainty in making characterisations of soil moisture using passive remote sensing systems. In addition, the weak intensity of emitted radiation from the Earth's surface means that a passive microwave instrument must have a large field of view, or footprint, which limits the potential spatial resolution (Cracknell and Hayes, 2007). For example, satellite based thermal-infrared sensors typically have a field of view three or four times smaller than passive microwave sensors (Cracknell and Hayes, 2007).

The low microwave emissivity from the Earth's surface also means that the signal can be perturbed by a number of different factors (Schmugge et al., 2002). As discussed earlier in this section, one of the main factors is surface roughness. However, the presence of vegetation can often have the most detrimental effect on the characterisation of soil moisture. Dense vegetation canopies can drastically reduce, and in some cases totally obscure, the microwave signal from the soil, as electromagnetic radiation is scattered by the surface of plant leaves and woody biomass, and absorbed by moisture in the vegetation (Notarnicola et al., 2006). As a result, many successful campaigns to map soil moisture using microwave sensors have been carried out over bare earth, or sparse vegetation canopies (Jackson et al., 1999; Burke et al., 1998; Wigneron et al., 1998; Jackson et al., 1995; Schmugge and Jackson, 1994). Schmugge et al. (2002) demonstrated this effect by explaining that measurements made over a mature corn crop can reduce the sensitivity of the emissivity signal by 25% compared to a bare soil.

Remote sensing in the microwave domain of the electromagnetic spectrum tends to centre on a small range of wavelengths (Table 2.1). Due to the influence of factors such as surface roughness and the presence of vegetation, the L band frequency is generally considered to be the optimal frequency for measuring soil moisture (Crosson et al., 2005). This is because it has a relatively long wavelength which has greater potential to penetrate vegetation canopies and is less susceptible to atmospheric attenuation (Wigneron et al., 2003; Kerr et al., 2001). Moreover, microwaves detected from a soil at this wavelength are derived from deeper within the soil, and are therefore seen to be a more accurate depiction of soil moisture content below the surface, or crust, of the soil (Njoku et al., 2003).

Band name	**Frequency (MHz)**	**Wavelength (cm)**
P	300 - 1000	30 - 100
L	1000 - 2000	15 - 30
S	2000 - 4000	7.5 - 15
C	4000 - 8000	3.75 - 7.5
X	8000 - 12000	2.5 - 3.75
K_u	12000 - 18000	1.667 - 2.5
K	18000 - 27000	1.111 - 1.667
K_a	27000 - 40000	0.75 - 1.111

Table 2.1 Table showing common microwave band names and relative frequency and wavelengths. From Mather (2004).

There have been a number of successful attempts to map soil moisture distribution with L band passive microwave sensors. Perhaps one of the most commonly cited examples include Jackson et al. (1999). This study was based around applying soil moisture retrieval algorithms, previously developed at a fine scale, to a regional scale. The algorithm requires information regarding soil temperature, vegetation type, vegetation water content, surface roughness, soil bulk density, and soil texture. The vegetation parameters were estimated from satellite imagery but the measures of soil properties were derived from field observations which not only limits operational potential but potentially limits its use outside the study area where soil properties are

likely to vary. The study found that the soil moisture retrieval algorithm could be scaled up to coarser resolution passive microwave data, with error levels in the region of 3%. Despite this success, the dominating vegetation types of the study area were pasture grassland, with relatively low vegetation canopies. As the algorithm relies on a great deal of parameterisation, in terms of vegetation, it is therefore likely to have a high degree of uncertainty when applied to other vegetation covers. This is of particular significance as the study was carried out to assess the potential for the algorithm to help inform global climate models, where undoubtedly these vegetation covers will be encountered.

As previously mentioned, the L band frequency is considered to be the most optimal for the detection of soil moisture using passive microwave systems. However, the long nature of L band wavelengths requires a long aperture on the sensing instrument. This has long seen to be a limiting factor on the use of microwave radiometry due to the challenges facing satellite engineers (Cracknell and Hayes, 2007; Jackson et al., 1999). For instance, in order to provide an adequate footprint of less than 60 km, a satellite sensor must have an antenna length of above 4 m (Kerr et al., 2001) which would lead to a payload beyond the capabilities of present satellite systems (Drinkwater et al., 2009). Developments in antenna design have recently made this possible through synthetic aperture whereby a large antenna is simulated from a number of small antennas distributed over a Y-shaped frame (see Kerr et al. (2001) and Drinkwater et al. (2009) for a detailed description). This theory has been materialised through the European Space Agency's launch of the Soil Moisture and Ocean Salinity (SMOS) satellite in November 2009. Despite these advances, SMOS provides measures of microwave emissivity at a maximum ground resolution of 50 km (Kerr et al., 2001). This characteristic essentially rules out the use of passive microwave sensors for monitoring transport corridors, as earthworks are much smaller than the footprint that can be achieved, i.e. in the region of metres rather than kilometres.

Sensors operating at higher frequencies, such as bands S, C and X (see Table 2.1 for frequencies and corresponding wavelengths), have also shown to have potential for the retrieval of soil moisture (Crosson et al., 2005; Njoku et al., 2003). The advantage of these frequencies is that due to their relatively small wavelengths, they require smaller antennas, and as a result, have the potential to provide finer spatial scale footprints.

Example systems include the Advanced Microwave Scanning Radiometer for Nasa's Earth Observation System (AMSR-E). This system measures microwave emissivity in the C and X band frequencies, with an average spatial resolution of 56 km and 38 km respectively. Sahoo et al. (2008) compared observed soil moisture measurements with estimates of soil moisture from AMSR-E, producing correlation coefficients up to 0.81. Despite a high correlation, these estimates could only be made using a heavily parameterised Land Surface Microwave Emission Model (LSMEM). This model is based on radiative transfer theory, which describes the way in which electromagnetic radiation interacts with the Earth's surface. The microwave emissivity observed by the sensor is taken to be a contribution of signals from the soil, vegetation, water bodies, snow cover and atmosphere. In the case of Sahoo et al. (2008) and Njoku et al. (2003), a number of assumptions are made: the influence of the atmosphere is assumed to be constant, there is assumed to be no difference between canopy and surface temperature, and multiple scattering at the Earth's surface is ignored (see Sahoo et al. (2008) and Njoku et al. (2003) for more details).

Although studies have shown that reasonable estimates of soil moisture can be made using C and X band microwave radiometers, they tend to be limited to relatively sparse vegetation canopies, despite the best efforts of radiative transfer models. This is because the smaller wavelengths associated with bands C and X are highly perturbed by the presence of vegetation, as well as being more vulnerable to influences from atmospheric attenuation and the effects of surface roughness, compared to measures of emissivity at the L band frequency (Kerr et al., 2001). This could have serious implications for mapping soil moisture in transport corridors where earthworks tend to be covered in dense vegetation, including tall shrubs and deciduous trees, to aid stability.

This section has demonstrated a trade off that exists when using passive microwave sensors to estimate soil moisture. High frequency microwave bands (X and C) have the potential to be sensed at a relatively fine spatial resolution; whereas low frequency bands (L band) tend to be relatively unperturbed by influences from the atmosphere, surface roughness and the presence of vegetation. Despite high correlations with observed soil moisture using both high and low frequency sensors, their limitations give them no potential for use in characterising soil moisture in

transport corridor earthworks, where high spatial resolution information is required, often over dense vegetation canopies.

2.2.2 Active microwave sensors

Active microwave sensors are radar (radio detection and ranging) systems which form an image by detecting the reflectance of a microwave signal transmitted by the sensor. Similar to the passive variety, acquiring imagery using these sensors has the advantage of being relatively unaffected by time of day, or cloud cover (depending on atmospheric conditions relating to moisture and aerosols). They also function over similar frequencies and corresponding wavelengths (Table 2.1).

Another trait shared with passive microwave systems is that active microwave sensors using real apertures are restricted to coarse spatial resolutions. As discussed in Section 2.2.1, this is due to the length of the antenna needed to emit and receive (in the case of active sensors) microwave energy. As a result, real aperture systems tend to be limited to large scale observations, such as observations of the Earth's geoid, and various meteorological applications (e.g. wind direction and speed) (Cracknell and Hayes, 2007). Conversely, most applications associated with earth surface properties, including soil moisture, tend to be carried out using Synthetic Aperture Radar (SAR) systems.

Section 2.2.1 alluded to the development of synthetic apertures which are used in passive microwave remote sensing to simulate a much larger antenna, thereby reducing the potential size of the footprint, for example, the Soil Moisture and Ocean Salinity (SMOS) satellite (Kerr et al., 2001). SAR uses similar technology to not only receive microwave signals, but also to emit microwave energy. This technology requires sophisticated system designs, as well as complex processing procedures. In basic terms, the SAR sensor emits a series of microwave pulses towards the Earth's surface as the platform, satellite or aircraft, flies over the area of interest (Ulaby et al., 1996). This results in a number of overlapping footprints running along the direction of the flight path (Figure 2.1). The system uses precise measures of the amplitude and phase of the return signal to calculate the position of the sensed object relative to the antenna. The Doppler effect is exploited in order to detect whether the target is behind or ahead of the sensor. Cracknell and Hayes (2007) and Woodhouse (2006) provide more detailed

descriptions of the principles behind the SAR system but the most important feature concerning this review is the fact that better spatial resolution can be achieved compared to real aperture systems.

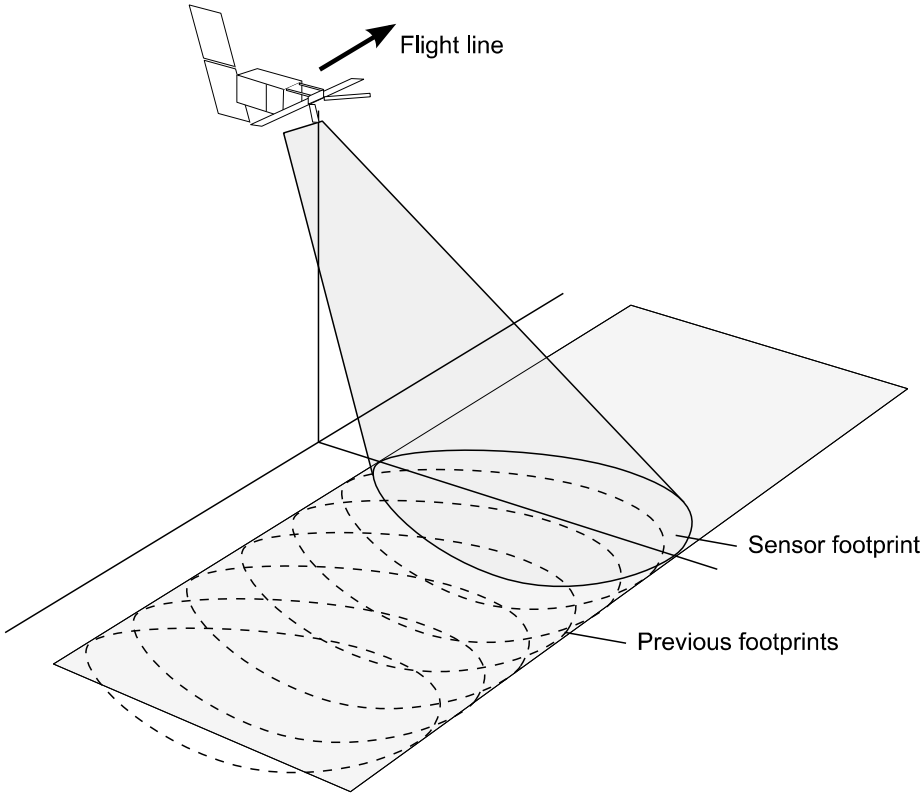


Figure 2.1 Illustration showing the geometry of the Synthetic Aperture Radar (SAR) system. Large footprints are designed to overlap and are combined to synthesise a larger antenna. From Woodhouse (2006).

Microwave energy emitted and received by a radar system can be horizontally or vertically polarised. Many systems have the ability to be configured to specify the way in which the system emits and receives a microwave signal. A system that both emits and receives a horizontally polarised signal is referred to as ‘like polarised’. Conversely, a system that emits horizontally polarised signal and received only vertically polarised returns is referred to as a ‘cross polarised’ system. By comparing a ‘like polarised’ image and a ‘cross polarised’ image for the same area, features that influence the polarisation of the signal can be identified. This type of analysis is known as depolarisation.

Depolarisation is related to the physical characteristics and the dielectric properties of the target. As discussed at the start of Section 2.2.1, the amount of water in a soil matrix can have a strong influence on the dielectric constant. In terms of active microwave imagery, higher amplitude backscatter tends to be found over wetter soils (Figure 2.2). Using these principles, it is possible to map soil moisture distribution using SAR imagery. The ENVISAT (Environmental Satellite) C-band ASAR (Advanced Synthetic Aperture Radar) is an example of a system that has the ability to provide imagery with different polarisations. Baghdadi et al. (2006) compared observed soil moisture measurements with microwave backscatter from ASAR using a number of different polarisations and incident angles. Although high correlations were found (RMSE of around 6%), the study found that the use of cross polarised imagery does not help to improve estimations. In fact, the most significant factor was the incidence angle, whereby, higher errors (RMSE between 9.1% and 9.6%) were found with low incidence angles (~40°). This finding is related to Figure 2.3 which demonstrates reduced sensitivity for low incidence angles, due to reduced scattering in the direction of the sensor. The use of cross polarised images for characterising soil moisture also has potential limitations as factors, such as surface roughness, can also lead to depolarisation (Baghdadi et al., 2007).

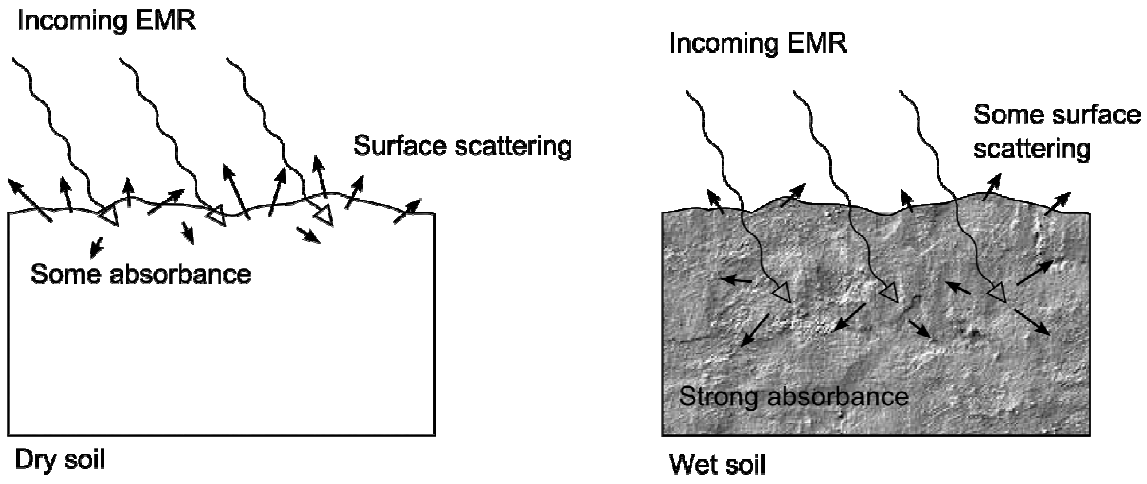


Figure 2.2 Diagram of soil profiles showing the interaction of electromagnetic radiation with dry and wet soils. Dry soils tend to reflect microwave energy whereas wet soils tend to absorb energy due to its higher dielectric constant.

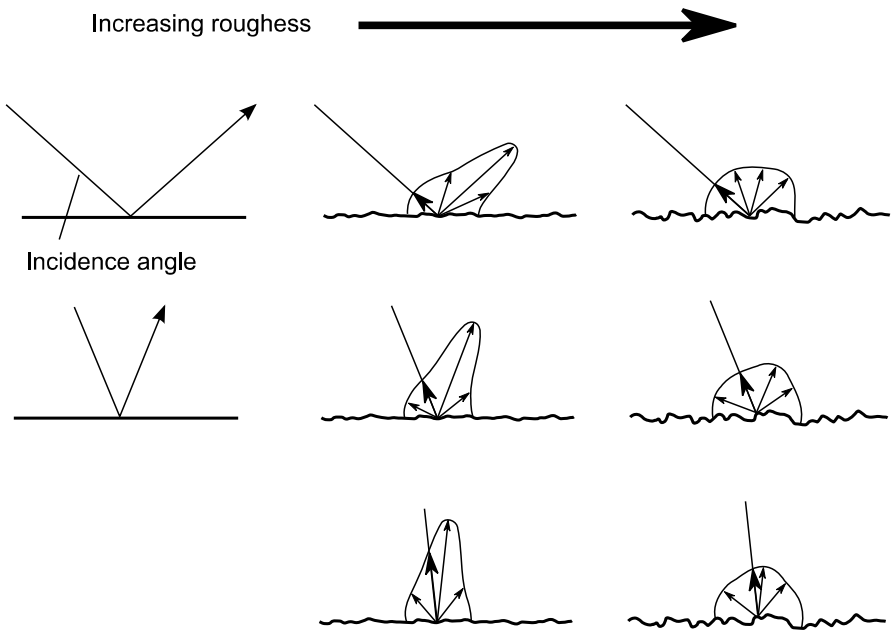


Figure 2.3 Diagram showing the effects of surface roughness and incidence angle on scattering. Smooth surfaces show only specular reflection. Rough surface show increasing diffuse scattering. For the roughest surface, reflectance is almost entirely diffuse and is not influenced greatly by incidence angle. Modified from Woodhouse (2006).

The example of Baghdadi et al. (2006) improved estimations of soil moisture (RMSE of 3.0%) by exploiting the relationship between incidence angle and surface roughness. Again, Baghdadi et al. (2007) found that SAR imagery acquired at different incidence angles could be used to make vastly improved estimations. This is because the ratio of backscatter coefficients can reduce the effects of other soil surface components, namely, surface roughness. It is possible to capture two images with different incident angles using satellite sensors such as the Canadian C-band RADARSAT (Radar Satellite). Despite this, the time taken for the second image to be collected (a matter of days) means that soil moisture conditions are likely to have changed (Baghdadi et al., 2007; Quesney et al., 2000). This, therefore, reduces the likelihood of capturing a reliable characterisation of soil moisture, particularly over temperate areas where the temporal variability of soil moisture is high (Famiglietti et al., 1998). In addition, as discussed in Section 2.2.1, the relatively short wavelength of

C band microwave energy (Table 2.1) means that it is strongly susceptible to scattering by vegetation canopies (Figure 2.4). As a result, the examples of Baghdadi et al. (2006; 2007) have been restricted to low vegetation canopies or bare soils. Therefore, these methods hold little relevance for vegetated earthwork embankments.

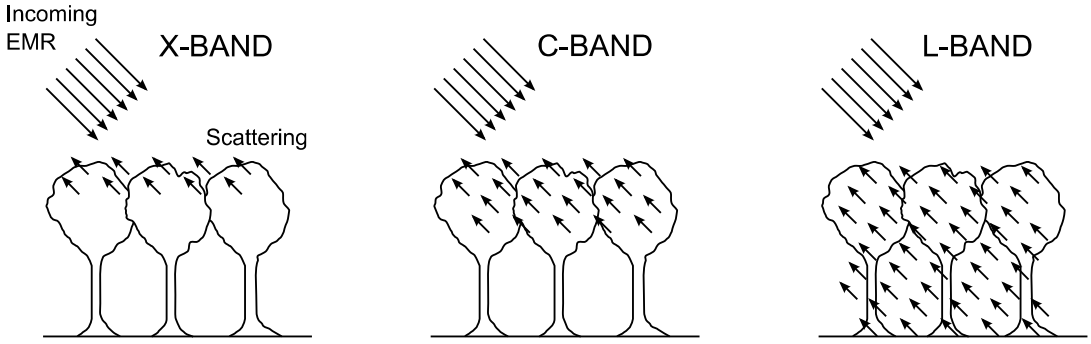


Figure 2.4 Diagram showing the interaction of electromagnetic radiation with vegetation canopies at different microwave frequencies. X-band is the highest frequency with a relatively short wavelength resulting in a high degree of scattering. L-band is the lowest frequency with a relatively long wavelength which penetrates the vegetation canopy and scatters from the ground. Corresponding frequencies and wavelengths can be found in Table 2.1. Adapted from Rao (1992).

Attempts have been made to make estimations of soil moisture using C band SAR imagery that are free from the influence of surface roughness and vegetation. Quesney et al. (2000) developed such a method using the European Space Agency's ERS (European Remote Sensing) SAR, that operates in the C band frequency at an incidence angle of >10°. Their method involved deriving soil moisture over specific calibration targets, that were corrected for the affects of vegetation and surface roughness. This allowed for the development of a soil moisture index that could then be applied to rest of the study area, in this case, entire watersheds. Despite showing high correlation coefficients with observed soil moisture (R^2 of 0.89), this method involved a high degree of parameterisation to correct for the effect of surface roughness, but especially, vegetation (Le Hegarat-Mascle et al., 2002). The effect of vegetation on the microwave backscatter was also modelled, but in a more comprehensive manner. They made use of a radiative transfer model based on the principals described in Karam et al.

(1992). This model requires information regarding vegetation height and plant density. Quesney et al. (2000) explained that these parameters can be derived using satellite vegetation reflectance data (in the form of established near infrared/red ratios from Landsat imagery as discussed later in Section 2.4.1) and knowledge of the plant type, in this case knowledge of the agricultural crop. Due to changeable nature of these parameters, this information must be continually updated throughout the vegetation growth cycle. This requires multiple satellite images and ground truth measurements, which can hinder the use of these methods in an operational context.

In a follow up to the method presented by Quesney et al. (2000), Le Hegarat-Mascle et al. (2002) developed these techniques in an attempt to eliminate the restrictions enforced by the reliance on continually updated information regarding vegetation. The previously developed method by Quesney et al. (2000) was validated over watersheds with varying soil types, composition, and vegetation cover. Rather than continually updating with ground truth data, the change in vegetation parameters were estimated using a crop growth model. Their results indicated reasonable correlation coefficients (R^2 0.63 – 0.85); however, the authors acknowledge that their methods cannot be applied to dense vegetation covers, including forests, ungrazed pasture and developed crops (Le Hegarat-Mascle et al., 2002). This limitation was attributed to the nature of C band microwave energy which, as discussed in Section 2.2.1, cannot penetrate dense vegetation canopies. Again, this calls into question the potential for using such techniques over densely vegetated earthworks. Perhaps one suggestion would be to apply these methods during the winter where the presence of leaves is limited, thereby reducing the effects of canopy scattering. However, the methods presented by Le Hegarat-Mascle et al. (2002) rely on the estimation of vegetation parameters from optical satellites which require leaf-on conditions during spring or summer months.

Recent attempts have been made to improve the retrieval of soil moisture over vegetated areas. Pierdicca et al. (2010) used an integration between airborne C band microwave imagery, from AirSAR, and information about vegetation cover from Landsat satellite imagery (further details regarding the use of satellite imagery for deriving information about vegetation can be found in Section 2.4.1). Initially, SAR backscatter data is collected for bare soils, i.e. before vegetation growth. This is then

used to estimate the change in backscatter for different times of the year, with the assumption that there is a relationship between soil roughness conditions and time of year. Similar to Le Hegarat-Mascle et al. (2002), vegetation conditions are quantified using Landsat imagery. Comparisons between estimated and observed soil moisture showed reasonable correlation coefficients (R^2 of 0.72). Although this is not as high as other studies mentioned in this Section, the study site included the presence of well developed vegetation cover. In addition, these estimates were made over a relatively fine spatial resolution, between 6.6 - 9.2 m, as they were made using a relatively high frequency microwave (C band) sensor, on board an airborne platform.

Both Le Hegarat-Mascle et al. (2002) and Pierdicca et al. (2010) have demonstrated the potential for using C band SAR to estimate soil moisture over vegetated areas. However, the models used to describe the effects of vegetation canopies are highly complex, using a number of parameters to characterise the vegetation at hand. This is therefore likely to limit the use of their soil moisture retrieval algorithms outside the study area. Furthermore, despite the relatively high spatial resolution at which these estimates are made, they remain too coarse to provide a characterisation of soil moisture distribution over transport corridor earthworks with a typical width of 5-20 m. The launch of TerraSAR-X, with a possible resolution of 1 m (Werninghaus et al., 2002), provides an exciting possibility for future research (Baghdadi et al., 2007), but currently there are no reported attempts to map soil moisture using this system. However, as this system operates at the high frequency X band, the soil moisture signal is likely to be highly vulnerable to the effects of vegetation and surface roughness.

This section has shown the potential for retrieving soil moisture from active microwave sensors. Despite excellent correlations with observed soil moisture, there are a number of inherent limitations that restrict the use of such systems. Many of these limitations are the result of a trade off between achievable spatial resolution and the vulnerability to the effects of soil surface roughness and the presence of vegetation. In general terms, although shorter frequency microwave energy (longer wavelength) is less susceptible to the influence of vegetation and surface roughness, the long aperture required for these systems limits the spatial resolution of the resulting imagery. Although recent advances have strived to improve the estimation of soil moisture using

high frequency microwave sensors (fine spatial resolution imagery) over vegetated areas, subsequent retrieval algorithms are highly parameterised and likely to be limited to their use outside the training area. Nevertheless, even the highest spatial resolution microwave sensors are insufficient for providing a detailed characterisation of soil moisture over a transport corridor earthwork.

2.2.3 Thermal sensors

The use of thermal imagery to map soil moisture is perhaps the only remote sensing method that has been considered in transport engineering (Perry et al., 2003b). Similar to passive microwave sensors, thermal sensors detect emissivity from the Earth's surface, but at much shorter wavelengths, typically 3.5 - 5.5 μm and 8 – 14 μm. Such sensors have been applied to a range of different applications, such as the detection of natural geothermal anomalies, and monitoring heat loss from building roof surfaces. Thermal channels have been incorporated into a number of satellite platforms, including ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer), MODIS (Moderate Resolution Imaging Spectroradiometer), AVHRR (Advanced Very High Resolution Radiometer) and Landsat ETM+ (Enhanced Thematic Mapper Plus), providing thermal imagery at a range of spatial resolutions (Landsat has the highest spatial resolution at 60 m). However, in terms of monitoring transport corridor earthworks, the potential of using thermal imagery lies in the development of relatively lightweight sensors. Due to the reduced payload, these sensors can be installed on light aircraft, providing the opportunity to obtain high spatial resolution imagery (Minacapilli et al., 2009).

The physical principles behind retrieving soil moisture from thermal imagery bears similarities with estimations made using microwave sensors. As described in Section 2.2.1, this relates to the fact that wetter soils have a higher dielectric constant, meaning they have greater ability to conduct electromagnetic radiation (see Figure 2.1). In terms of thermal radiation, the higher dielectric constant of a wet soil means that it has a greater ability to retain thermal energy. As a result, soil moisture is a major control on the daily range in soil temperature, where wetter soils exhibit a lower range in temperature (Verstraeten et al., 2006). Using this principal, soil moisture can be

characterised through thermal inertia mapping, which quantifies the diurnal change in temperature for a specific target.

Rather than making direct measures of thermal inertia, remote sensing techniques estimate the contributing components, such as day and night time surface temperatures (see Minacapilli et al. (2009) for a breakdown of the formulas used). The estimate of thermal inertia is used to invert an equation that describes the relationship between the thermal properties of a soil (heat capacity and thermal conductivity) and the soil water content. However, the relationship between the thermal conductivity of a soil and the soil moisture content is complex due to the effects of soil composition (Minacapilli et al., 2009). Similar to the complications associated with microwave signals from variations in surface roughness (see Sections 2.2.1 and 2.2.2), the relationship between soil moisture and thermal conductivity is modelled using a range of approaches.

Physically based models, used to model the effect of thermal conductivity on soil moisture content, tend to be highly parameterised and are therefore unlikely to be useful outside the training area (Minacapilli et al., 2009; Lu et al., 2007). As a result, empirically based models have been developed that use set parameters to express different soil textures. Despite this generalisation, studies have found a good correspondence between observed and estimated soil thermal conductivity under laboratory conditions (Lu et al., 2007). Minacapilli et al. (2009) used this kind of modelling in a thermal inertia mapping campaign to characterise soil moisture demonstrating correlation coefficients up to 0.75. As well as producing reasonable correlations, these estimates were made using Airborne Thematic Mapper (ATM) with a spatial resolution of 4 m, which would be adequate to carry out monitoring campaigns over transport corridor earthworks. However, this study was applied to bare soils only and therefore has no bearing on the estimation of soil moisture over vegetated areas. In addition, soil wetness was artificially induced using irrigation technology, which is likely to produce more distinct differences in soil moisture values compared to a naturally occurring distribution of soil moisture.

The main limitations, associated with using thermal imagery to characterise soil moisture, are inherent to all sources of optical remotely sensed data. Notably, these limitations include vulnerability to atmospheric attenuation and the inability to penetrate cloud cover and vegetation canopies (Kerr, 2007). Although the presence of vegetation

remains a major limiting factor, attempts have been made to estimate soil moisture for areas with agricultural crops. Crow et al. (2008) explained that through accurate interpretation of the thermal signal from vegetation canopies, it is possible to detect differences in surface temperature, due to changes in evapotranspiration, which can be used to infer soil moisture conditions. This was achieved through the assimilation of two soil-vegetation atmosphere transfer modelling approaches (see Crow et al. (2008) for full details of the methods and models used). The first modelling approach uses thermal remote sensing observations to predict the temperature from the vegetation canopy and the soil surface. The second is a water balance model that uses observed rainfall data and predicted evapotranspiration to determine soil moisture within the vegetation root zone. The results of this work showed good correspondence with observed soil moisture, with a correlation coefficient of 0.84. However, the strength of the correlations were found to be strongly dependent on the accuracy of vegetation input parameters into the aforementioned models (Crow et al., 2008). This involved extensive and continued field work, which limits the potential for such a technique to become operational. In addition, the study site used by Crow et al. (2008) was located over agricultural fields with a land cover of homogenous corn crop. It is likely that a transport corridor earthwork, with heterogeneous vegetation cover, is likely to present a much more challenging environment to model.

Another approach to deriving soil moisture over vegetated areas is referred to as the ‘triangle method’ (Vicente-Serrano et al., 2004; Sandholt et al., 2002), which exploits the relationship between soil moisture, surface temperature and fractional vegetation cover. In terms of remotely sensed data, surface temperature can be determined using a thermal sensor, and vegetation cover can be estimated using reflectance measurements in the red and near infrared regions of the electromagnetic spectrum. Specifically, a normalised difference vegetation index (NDVI) is used to represent vegetation cover, which uses reflectance from the red and infrared regions of the spectrum, that are sensitive to plant leaves (Bannari et al., 1995) (discussed in greater detail in Section 2.4.1). Analysis of the relationship between derived surface temperature and NDVI, referred to as the temperature-vegetation dryness index (TVDI), is used to provide information on soil moisture conditions by estimating evapotranspiration rates (see Sandholt et al. (2002) for full details). This relationship is

summarised in Figure 2.5 with high TVDI values relating to limited water availability, and low values corresponding to maximum evapotranspiration due to high soil moisture contents (Vicente-Serrano et al., 2004; Sandholt et al., 2002). This method has advantages over other techniques, such as those described by Crow et al. (2008), as estimations of TVDI rely solely on remotely sensed data, therefore increasing the potential for being used operationally (Sandholt et al., 2002).

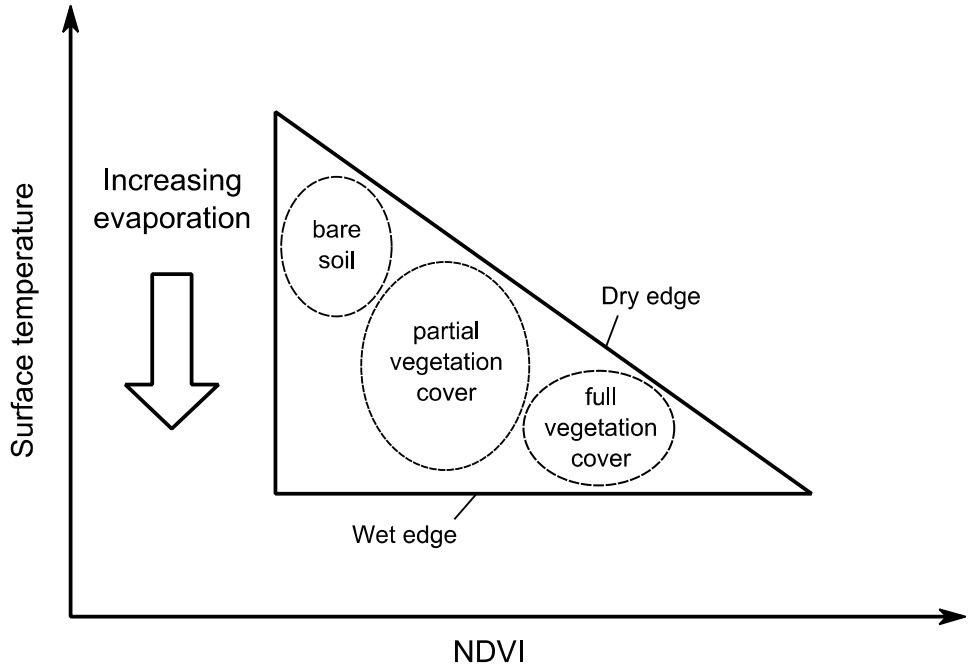


Figure 2.5 Simplification of the triangle method that defines the relationship between surface temperature, vegetation cover (represented by NDVI) and evaporation, which is used in the calculation of TVDI to infer soil moisture conditions. Modified from Sandholt et al. (2002).

Using the triangle method and the calculation of TDVI values, studies have found reasonable correlations with soil moisture values estimated using hydrological models. For example, Sandholt et al. (2002) found correlation coefficients up to 0.81 using data from the AVHRR sensor. They also found that their estimates were insensitive to surface cover indicating stability of the developed technique. However, it should be noted that this study was conducted for a semi arid region of West Africa where vegetation tends to be sparse (Sandholt et al., 2002). Vicente-Serrano et al. (2004) used

similar techniques to Sandholt et al. (2002) for estimating soil moisture over an area in north east Spain, using data from both AVHRR and Landsat satellites. Although the study site had greater vegetation cover, comparisons with soil moisture derived from hydrological models showed relatively poor results, up to an R^2 of 0.42. This suggests that the triangle method is better suited to areas with sparse vegetation cover and may therefore not be applicable for UK transport corridors. Perhaps more crucial is the coarse spatial resolution of the AVHRR sensor (1.1 km) which is not adequate for monitoring transport corridor earthworks.

2.2.4 Visible, near and shortwave infrared

Thermal imagery focuses on an area of the electromagnetic spectrum with relatively long wavelengths. Efforts have also been made to study the effects of soil moisture on reflectance in the visible, near and shortwave infrared region of the spectrum, ranging from 400 nm to 2500 nm. The influence of soil moisture content on the reflectance of a soil in this spectral domain follows similar lines to that of thermal and microwave reflectance or emissivity. Specifically, an increase in water in a soil matrix increases the amount of electromagnetic energy that is absorbed by the soil (Moran et al., 2004). This results in the familiar darkening of a soil after wetting (Lobell and Asner, 2002) which forms the basis of estimations of soil moisture using reflectance in the visible, near and shortwave infrared.

Lobell and Asner (2002) gathered reflectance measurements in this spectral domain for four different soil types with varying degrees of wetness under laboratory conditions. The relationship between soil moisture and reflectance was expressed as the best fit coefficients for an exponential model. Reflectance at the wavelengths 600 nm, 1200 nm and 2200 nm, were shown to have a particularly strong exponential relationship with observed soil moisture, although actual correlation coefficients were not reported. Their findings also showed that reflectance measurements in the visible domain were only sensitive to soil moisture contents up to 20%, whereas reflectance in longer wavelengths, such as the near and shortwave infrared, continued to respond to soil moisture content value up to 50%.

Weidong et al. (2002) also examined the influence of soil moisture on reflectance measurements in the visible, near and shortwave infrared regions of the spectrum.

Reflectance in this domain is strongly affected by soil type and texture (Moran et al., 2004; Muller and Decamps, 2001) and therefore Weidong et al. (2002) analysed the effects over 10 different soil types. They found very high correlations, R^2 above 0.88, when reflectance measurements were compared to soil moisture values for individual soils. High correlations, R^2 between 0.7 and 0.85, were also found when comparisons were made using the whole set of soils, showing promise for developing a universal model. However, this was only applicable to soils with relatively low soil moisture content as poorer correlations were found over wetter soils (R^2 less than 0.5). This is due to the non linear relationship between soil moisture and reflectance over this spectral domain. Specifically, under low soil moisture conditions, reflection tended to decrease with an increase in soil moisture, whereas, at higher soil moisture levels, reflectance tended to increase with an increase in soil moisture (Weidong et al., 2002).

Despite the complexities illustrated by the latter study, the work reported by Lobell and Asner (2002) and Weidong et al. (2002) demonstrated the potential for using reflectance measurements in the near and shortwave infrared to infer soil moisture content. However, this work was limited to a laboratory setting where factors such as soil type, surface roughness and texture could be controlled. Lihua et al. (2005) extended this work by estimating soil moisture using visible imagery gathered from an unmanned aerial vehicle and hyperspectral data in the field using a spectroradiometer. Despite Lobell and Asner (2002) previously showing the visible part of the spectrum to be insensitive to soil moisture content value above 20%, Lihua et al. (2005) found strong correlations with observed soil moisture, up to an R^2 of 0.9.

One the advantages of taking remotely sensed measurements of reflectance in the visible, near and shortwave infrared, is the potential for acquiring very high spatial resolution imagery (0.5 m with some commercial satellite sensors) (Lobell and Asner, 2002). There are a number of sensors with the capability to record reflectance in the visible, near and shortwave infrared. Those with a fine spatial and spectral resolution are typically hyperspectral sensors onboard aircraft, measuring spectral reflectance over numerous bands. Ben-Dor et al. (2002) used the DAIS-7915 (Digital Airborne Imaging Spectrometer) scanner with 79 channels across the spectral region 400 nm to 1400 nm at a spatial resolution of 8 m to map soil properties over farmland in Northern Israel. They compared reflectance measurements against different soil properties, including

soil moisture content, using multiple linear regression. Correlations between observed soil moisture and estimates made using the airborne remotely sensed data were not as good as those derived from laboratory experiments, with R^2 values of 0.65 and 0.85 respectively. Nevertheless, this study demonstrated that a reasonable characterisation of soil moisture can be made, at a high spatial resolution, using airborne sensors

The previously mentioned studies have all shown high correlations with observed soil moisture. However, these estimates were made over bare soil as reflectance measurements made in the visible, near and shortwave infrared are highly perturbed by the presence of vegetation (Moran et al., 2004; Muller and Decamps, 2001). This is because the relatively short wavelength of electromagnetic energy within this spectral domain is absorbed and scattered by leaves and branches. Therefore, techniques using this spectral domain to make direct inferences of soil moisture would have little use over a vegetated transport corridor earthworks.

2.3 Characterising soil moisture using terrain analysis

Rather than detecting soil moisture directly from the soil surface, research has shown potential for characterising soil moisture conditions through indirect methods. Specifically, topography has been shown to have a strong influence on the processes that control soil moisture distribution, including surface and subsurface flows, and evapotranspiration (Tenenbaum et al., 2006; Bárdossy and Lehmann, 1998). Researchers have exploited this relationship through the development of terrain analysis routines such as wetness indices and calculations of potential evaporation. These calculations are applied to digital elevation models (DEMs) which are generated using a number of different techniques, which are discussed in Section 2.3.3.

2.3.1 Wetness indices

Terrain analysis using DEMs is a well established practise in hydrology (Sørensen et al., 2006). Perhaps one of the most common examples of this is through the use of a topographic wetness index (TWI). The theory behind this index was first presented by Beven and Kirkby (1979), and can be expressed by Equation 2.1.

$$TWI = \ln\left(\frac{a}{\tan\beta}\right)$$

Equation 2.1

Where a is the local upslope area draining through a certain point per unit contour length, and β is the local surface slope (Sørensen et al., 2006). The tangent curvature of the slope provides a measure of flow convergence or divergence, and the upslope area is a measure of the potential area that can contribute flow through a unit contour length (Western et al., 1999a). The specific calculations behind the TWI are detailed in Section 3.2.1.1.

To put the TWI calculation into context, take the scenario in Figure 2.6. The input DEM shows a simplified version of a catchment area, with steep sides and a flat base (step 1). The upslope contributing area in the TWI calculation (Equation 2.1) indicates that flow directions travel from the steep sides of the catchment, towards the flat base (step 2). As the flat base is located towards the end of the flow direction paths, it has a higher degree of contribution from further up the catchment, compared to the valley sides that have little or no other cells flowing into them (step 3). This illustrates the estimation of accumulation based on upslope contributions. The curvature of the slope explains that moisture is more likely to be retained on flatter slopes, as opposed to steep slopes for which moisture is likely to flow downslope. The combination of contributing area and the curvature of the slope means that flat areas which receive contribution from upslope areas have higher TWI values than steep areas with little or no upslope contribution. Areas with high values are known as areas of topographic convergence, which is depicted in Figure 2.6 (step 3) where the base of the catchment has a higher TWI value (red grid squares) than the catchment sides (white grid squares). However, it should be acknowledged that values are expressions of subsurface flow and susceptibility to wetness rather than direct predictions of soil moisture content (Tenenbaum et al., 2006; Beven, 1997). As a result, the calculation of TWI does not take into account other controls on the distribution of soil moisture, such as soil characteristics (Schmidt and Persson, 2003).

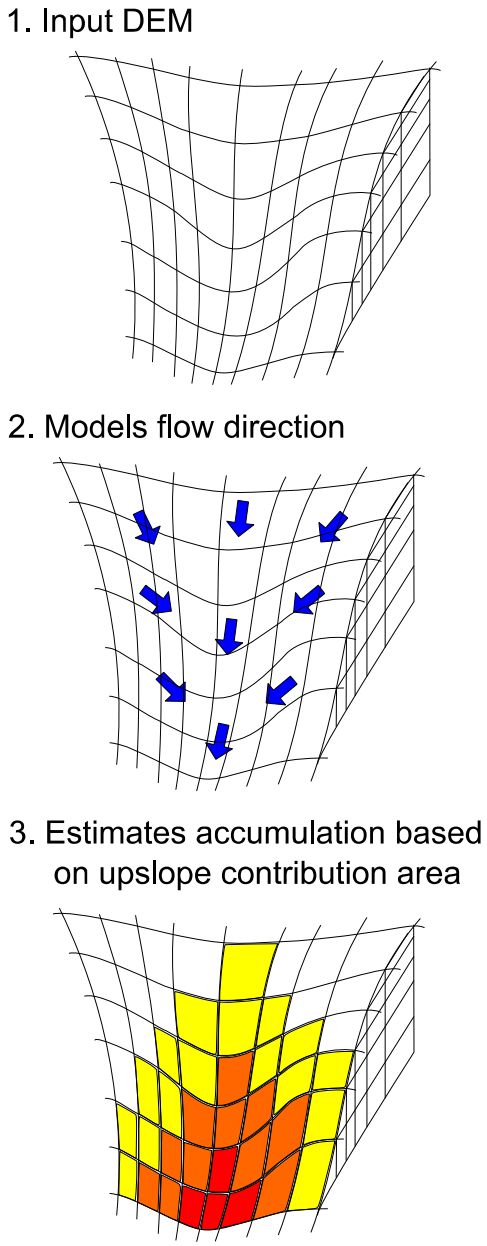


Figure 2.6 Illustration showing key components in the calculation of a TWI, including the calculation of flow direction and upslope contributing area. Red areas refer to high TWI values, corresponding to wetter areas.

A common application for TWIs is in slope stability studies (Baum et al., 2005; Gritzner et al., 2001; Pack et al., 1998). This is based on the observation that areas with higher soil moisture content tend to occur in areas of topographic convergence, as illustrated in Figures 2.6 and 2.7, which is commonly associated with slope instability

(Pack et al., 1998). This coincidence is considered so intrinsic that entire stability models have been explicitly based around TWI calculations, such as the SINMAP (Stability Index Mapping) model developed by Pack et al. (1998). However, these studies make no attempt to draw comparisons with observed soil moisture measurements. Despite this, there is a growing body of work that specifically aims to assess the quality of TWIs with observed soil moisture measurements.

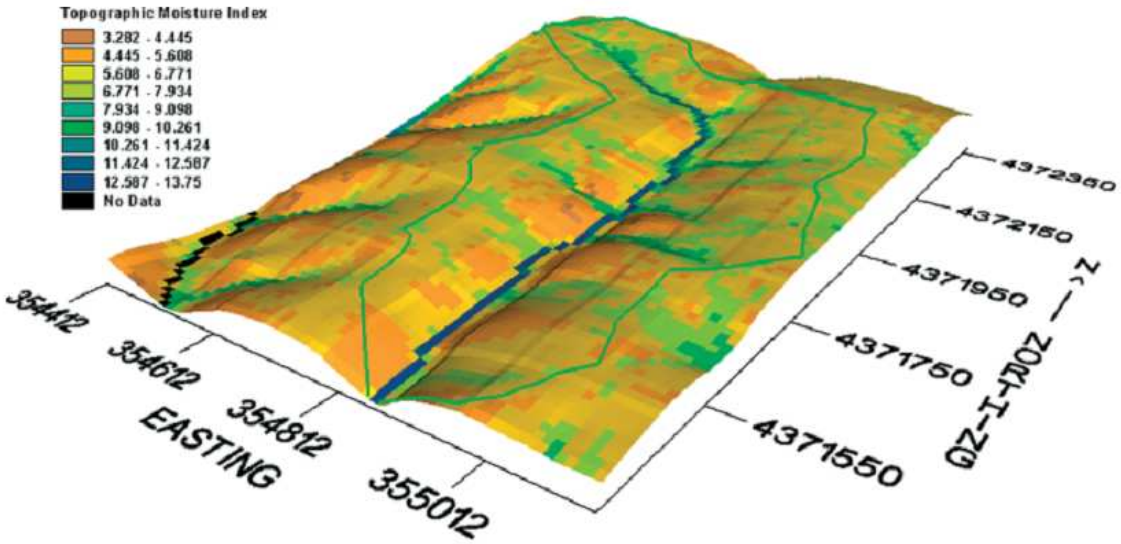


Figure 2.7 Example TWI calculated from a 11.25 m DEM from a forested catchment near Baltimore, USA, from Tenenbaum et al. (2006). Notable patterns include increased wetness (blue) in an area of convergence running down the centre of the scene and lower wetness values (dark orange) for the steep slopes surrounding this area of convergence.

Schmidt and Persson (2003) showed a reasonable comparison between surface soil moisture measurements and a TWI generated from a 3 m DEM for varying relief in Central Sweden and North East Germany. Best results were found over undulating, hummocky terrain with an R^2 of 0.64. However, poor results were found in areas of low relief due to the lack of topographically converging areas. As Schmidt and Persson (2003) explained, the concept behind the TWI is only valid for areas with a high degree of lateral water movement, which is controlled by relative relief. This is supported by Barling et al. (1994), who explained that soil moisture distribution is not sufficiently controlled by topography in areas that exhibit an overall shallow slope angle (< 6˚).

Whereas transport corridors as a whole may exhibit both shallow sloping low relief and steep sloping hummocky terrain, slopes on embankment and cutting earthworks generally tend to be more than 6˚, therefore showing potential for using a TWI to characterise soil moisture in such an environment.

Western et al. (1999a) compared a TWI, calculated using a 5 m DEM for an undulating catchment near Melbourne, Australia, against a detailed survey of soil moisture measurements made throughout the year. Correlation coefficients varied greatly throughout the measurement period, ranging from no correlation, to an R^2 of 0.42 found during the wet season. Interestingly, slight improvements to correlations were obtained when soil moisture was compared to the natural logarithm of the contributing area, i.e. taking the tangent of the slope (β) out of the TWI calculation in Equation 2.1, with a correlation coefficient of up to 0.5. This correlation was only achieved during wet conditions suggesting that contributing area becomes the dominant control on soil moisture distribution (Western et al., 1999a). This implies that during time of increased rainfall, lateral movement of water through a terrain becomes more dominant than the control of slope curvature, which may be more applicable to long term controls on soil moisture distribution as wetness concentrates in areas of convergence. Better results were achieved when the TWI was coupled with a calculation of potential solar radiation, which is discussed in Section 2.3.2.

Tenenbaum et al. (2006) showed a similar spread in correlation coefficients when they compared soil moisture measurements to TWI calculations made using DEMs with a range of spatial resolutions (0.5 m to 30 m), over a forested area with undulating terrain, and a suburban area with relatively low relief in Baltimore, USA. The strongest correlations (R^2 of 0.8) were found using a 11.25 m DEM, when compared to soil moisture measurements made in wetter conditions. Moderate correlations, R^2 up to 0.56, were found using a TWI derived from a 0.5 m DEM. This shows promise for characterising soil moisture at a fine scale. However, to achieve these results, the wetness index was averaged over an 11 x 11 kernel producing a conceptual grid with a 5.5 m resolution. This approach was adopted so the TWI grid values were comparable to the observed the soil moisture measurements. In doing so, this would suppress any fine scale variation predicted by the calculation which could be

crucial in precision engineering over transport earthworks where important hydrological features, such as drains, are typically narrow (width of approximately 1 m).

The degree of success in using wetness indices to predict soil moisture distribution has varied greatly. It is generally held that higher correlations with observed soil moisture can be found when calculations are applied to coarser scale DEMs (Tenenbaum et al., 2006). However, this is highly dependent on the complexity of the terrain and environment being modelled. For example, Gritzner et al. (2001) found only partial success when using a TWI as part of an assessment into landslide potential in a 875 km^2 catchment in Idaho, USA, as the coarse 30m DEM failed to identify features such as roads. This leads to the suggestion that the success of the TWI calculation is dependent on the ability of the DEM to identify controlling hydrological features (Tenenbaum et al., 2006; Gritzner et al., 2001). Based on these findings, if a TWI was used to characterise soil moisture in a transport corridor, a trade off would have to be made between sufficient spatial resolution, with the ability to represent all major hydrological features, and the size of the dataset to allow for reasonable processing time.

One criticism of TWIs is that they assume the entire upslope contributing area flows to a specific point by subsurface flow, which in some environments can be a very slow process (Wilson and Gallant, 2000; Western et al., 1999a). This can be a particular disadvantage in applications such as pollutant movement over a catchment. Here, the steady-state estimations made by TWIs do not take into account the time since the rainfall event which can affect surface and subsurface flow patterns (Kim and Jung, 2003). This has led to the development of quasi and fully dynamic wetness indices that incorporate a temporal dimension, taking into account hydraulic functions such as discharge rate, conductivity and gravitational acceleration (Kim and Jung, 2003; Barling et al., 1994). However, these models require knowledge of soil properties, such as grain size composition and effective porosity, which can be cumbersome to measure accurately in the field, and would likely detract from the original advantage of using a TWI to characterise soil moisture distribution with less effort than using conventional ground-based methods.

2.3.2 Potential solar radiation

As discussed earlier in Section 2.3.1, wetness indices tend to work better in wetter overall conditions, where soil moisture distribution is more strongly controlled by topographic convergence (Western et al., 1999a; Grayson et al., 1997). However, during dry periods, soil moisture distribution has been shown to be a function of aspect and resulting evapotranspiration (Tenenbaum et al., 2006; Western et al., 1999a; Isard, 1986). Potential evapotranspiration can be characterised by a calculation of potential solar radiation (Western et al., 1999a).

The calculation of potential solar radiation uses a DEM to model the way in which solar radiation reaches the Earth's surface from the Sun. Solar radiation arrives at the Earth's surface in three different ways: (1) direct radiation, as the name suggests, is received via a direct line from the sun, and is the largest contributor to total radiation; (2) diffuse radiation reaches the Earth's surface by way of scattering in the atmosphere; (3) reflected radiation is scattered by features on the Earth's surface. Generally, the latter route is only significant where there is large snow cover and is therefore not considered in the calculation of potential solar radiation. Although spatial variation in humidity would also have an effect on potential evaporation, it is considered insignificant at a small catchment scale (approximately 500 m^2) due to mixing in the atmospheric boundary layer (Western et al., 1999a), and is also excluded from the calculation.

Fu and Rich (2002) modelled the principles of solar radiation to calculate potential solar radiation. This calculation begins by generating a hemispherical viewshed for each location in the input DEM. This viewshed is a representation of the area of sky that is visible to that point based on the presence of any obstructions (Figure 2.8). It is therefore important to have a DEM that sufficiently represents topography, as coarse DEMs may produce under, or overestimations of potential solar radiation for any given grid cell. The hemispherical viewshed is used in conjunction with information about the position of the sun and sky direction to calculate total radiation. A sunmap is used to represent the position of the sun by taking into account the time of the day, time of year and latitude of the area of interest (see Figure 2.9 for a graphical representation). As a result, potential solar radiation can be calculated for any length of time from specific days, to entire years. Sky direction is represented by a skymap that produces a

hemispherical view of the sky, with different areas of the map defined by the zenith and azimuth direction (Figure 2.10). The different areas on the sky map have a unique value that is used to quantify the contribution of diffuse radiation from atmospheric scattering. These contributions are generalisations of a complicated process which is normally quantified using complex atmospheric contribution models. Total radiation is calculated by overlaying the viewshed with the sunmap and skymap. An example of this can be found in Figure 2.11 which shows a map of surface temperature predicted from the potential solar radiation calculation outlined by Fu and Rich (2002).

The potential solar radiation calculation has been used in studies to achieve improved estimations of soil moisture particularly in dry conditions (Western et al., 1999a). However, the calculation is rarely used to delineate soil moisture on its own. Greater success has been found using a combination of potential solar radiation and TWI (Western et al., 1999a; Western et al., 1999b). Specifically, Western et al. (1999a) found that a combination of potential solar radiation and TWI produced an R^2 of 0.61 with observed soil moisture during wet conditions, an improvement by a factor of nearly 20% compared to using the TWI on its own.

Advances in the collection of remotely sensed topographic data and terrain analysis using GIS has led to the development of a number of hydrological models (Du et al., 2007; Boughton, 2005). The hydrological model TOPOG uses terrain analysis to predict hydrological characteristics including soil moisture distribution and catchment erosion (Huang et al., 2005). This model incorporates information from both wetness and solar radiation calculations from a DEM in order to describe the way in which water moves through a landscape. Soil moisture estimates made using such models have been successfully applied to the assessment of slope stability (Ray and Jacobs, 2007) but few reported attempts have been made to compare soil moisture distributions with observed soil moisture measurements. One of the criticisms of these kind of models is that they require a large number of parameters (Du et al., 2007), including soil depth, hydraulic conductivity and vegetation cover, all of which are spatially variable (Huang et al., 2005). As a result, models are likely to be tailored to specific study areas and have limited use in an operational context. In addition, models such as TOPOG require a great deal of terrain data and are therefore only considered to be applicable to small catchment scale areas (less than 10 km^2) (Boughton, 2005; Huang et al., 2005).

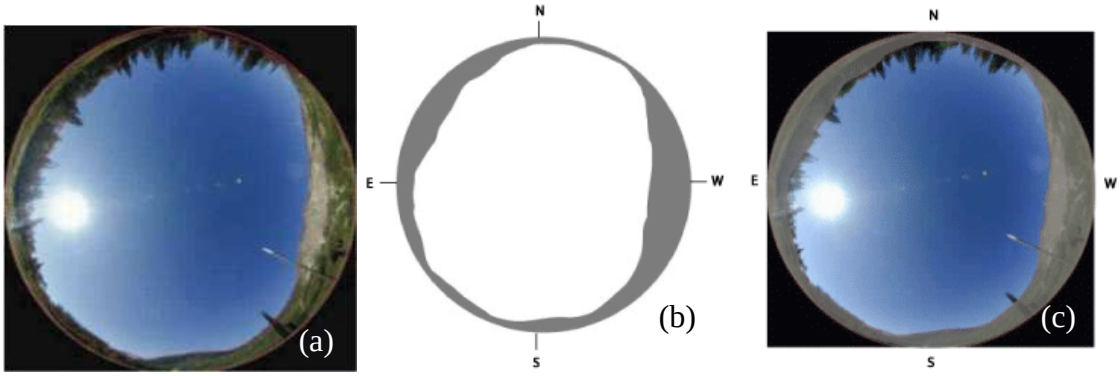


Figure 2.8 Illustration of a viewshed calculated for a specific cell. Image (a) shows the hemispherical view from a particular DEM grid square. Image (b) shows areas in grey that are considered to be obstructions. Image (c) shows the same hemispherical view as (a) but with the area of sky that is not visible due to obstructions masked out in grey. Images taken from ESRI's ArcGIS help documents (ESRI, 2007a).

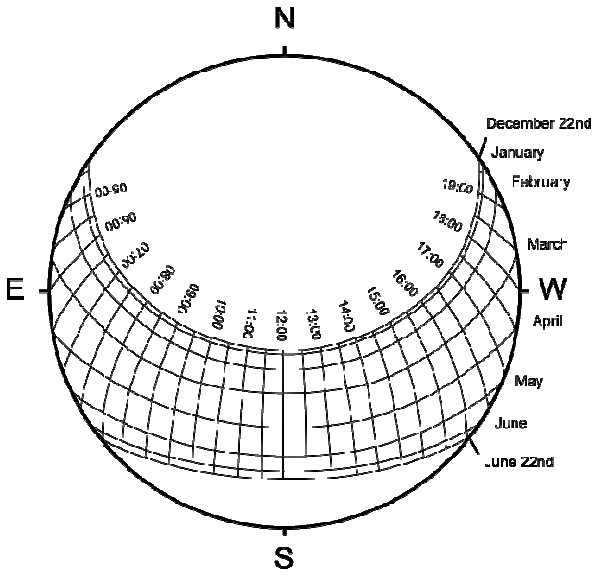


Figure 2.9 Schematic representation of the sunmap that quantifies direct solar radiation from time of year, time of year and latitude of the area of interest. This particular sunmap is for a latitude of 45° N, calculated for the period December 21st to June 21st. Each square represents half hour intervals throughout the day and monthly intervals through the year. Adapted from ESRI's ArcGIS help documents (ESRI, 2007a).

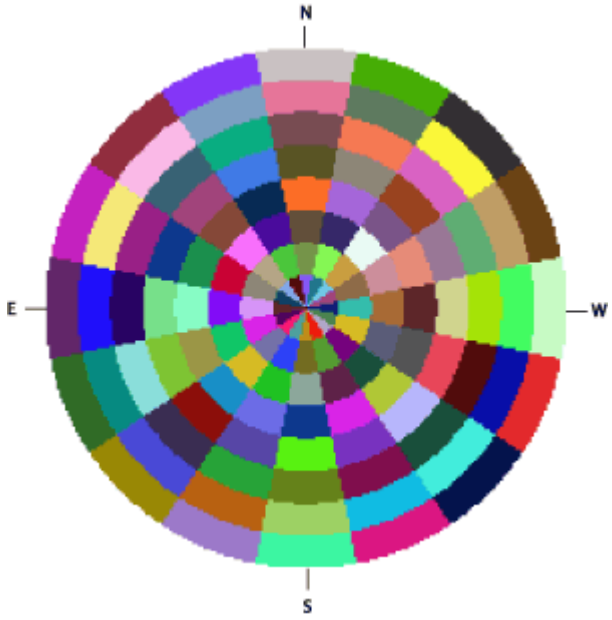


Figure 2.10 Schematic representation of the sky map that is used to represent diffuse radiation from zenith and azimuth directions. Each box represents a portion of the sky with a unique calculation of diffuse solar radiation. From ESRI's ArcGIS help documents (ESRI, 2007a).

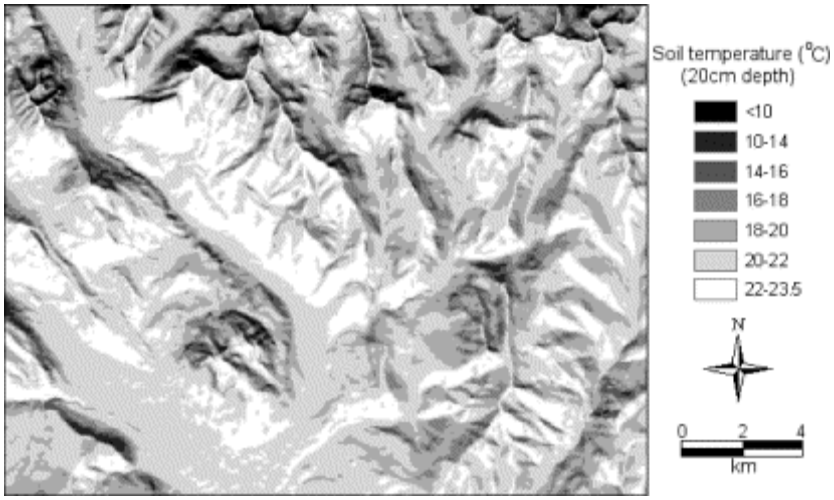


Figure 2.11 Map of soil temperature predicted using linear regression between observed temperature measurements and potential solar radiation calculated using a 30 m DEM, from Fu and Rich (2002). Noticeable patterns include a reduction in predicted surface temperature on north facing slopes.

2.3.3 Issues relating to input DEM

Sections 2.3.1 and 2.3.2 relate to indirect methods for characterising soil moisture using DEMs. However, there are a number of issues relating to the input DEM that can affect these characterisations, particularly with the calculation of TWIs. Gritzner et al. (2001) highlighted the need for a DEM with a resolution that sufficiently represents the hydrological features within the area of interest. Naturally, this varies with the complexity of the area at hand. Tenenbaum et al. (2006) found that urban areas require much finer scale DEMs to produce an accurate TWI (< 1 m) compared to a forested area with low relief (5-10 m). Transport corridors often run through urban areas and would therefore require a fine scale DEM. In addition, transport infrastructure such as bridges, embankments and cuttings can be important hydrological features in themselves (Tenenbaum et al., 2006; Gritzner et al., 2001) and therefore need to be represented in the input DEM. As a result, it may be advantageous to consider the use of digital surface models (DSMs). Whereas DEMs are a depiction of terrain, DSMs include other features, such as buildings, roads and vegetation, that are normally filtered out in routines for DEM generation. However, features such as roads and buildings are relatively small and may not be detected in coarser scale terrain data. For example, Gritzner et al. (2001) found inaccuracies using a 30 m DEM to delineate soil moisture as input to an assessment of landslide risk due to subsequent terrain analyses not taking into account features such as roads.

Issues concerning DEM resolution relate to the methods used for capturing elevation data. Data gathered using portable RTK (Real Time Kinetic) GPS tend to be cheap (providing the equipment is available) and provide accurate measurements, but they are labour intensive over large areas, and the need for a constant radio link has proved to be problematic in undulating terrain (Schmidt and Persson, 2003). Spot elevation values can be derived over wide areas from aerial photography and photogrammetric techniques (Schmidt and Persson, 2003). However, some authors have noted that photogrammetric techniques may not provide adequate resolution to represent hydrological features, particularly in complex urban environments, or where there are dense vegetation canopies (Tenenbaum et al., 2006; Haugerud and Harding, 2001) and would therefore be problematic in a transport corridor environment. For this reason, DEMs derived from airborne laser scanning (ALS) tend to be favoured as they

consistently achieve higher point density returns (Schmidt and Persson, 2003) and have the ability to provide reasonable ground returns over vegetated areas (Schulz, 2005). In addition, terrain data derived from ALS has shown to be useful for depicting fine scale features such as roads and buildings (Hollaus et al., 2005), which were identified earlier in this section as having an influence on the accuracy of terrain analysis calculations, including TWI.

ALS systems are typically mounted on an aircraft allowing timely acquisition of high density elevation data over large areas. Typically this can produce high spatial resolution DEMs of less than 1 m, with errors better than 25 cm, depending on a number of conditions including vegetation cover (Hollaus et al., 2005). DEMs derived in this way have been successfully used to generate wetness parameters for a number of slope stability studies (Baum et al., 2005; Schulz, 2005; Borga et al., 2002). Additionally, ALS derived DEMs have been used in a number of soil moisture distribution studies such as Schmidt and Persson (2003) and Tenenbaum et al. (2006), but currently there have been no attempts to apply these methods for characterising soil moisture in a transport corridor environment.

Despite the notable advantages of using ALS sensors for generating DEMs, the ground surface signal can be strongly perturbed by the presence of vegetation (Reutebuch et al., 2003). To overcome this, filtering algorithms have been developed to maximise the signal from the ground by eliminating the signal from the vegetation canopy and other obstructions (Haugerud and Harding, 2001). This process has provided excellent results, as well as being fully automated (Haugerud and Harding, 2001). As a result, the filtering procedure has become a common practise in DEM processing (Lim et al., 2007) and undergoes continual development. For a more comprehensive review of filtering algorithms see Sithole and Vosselman (2004).

SAR sensors can also be used to derive elevation values and build high spatial resolution DEMs using SAR interferometry (also known as InSAR) where the difference in the signal from two separate SAR antennae are used to construct a map of surface elevation (Mather, 2004; Rosen et al., 2000). However, studies have shown that the signal can be perturbed by different land covers, particularly vegetation (Rosen et al., 2000) and as a result is not as accurate a ALS systems for generating a DEM (Chang et al., 2004). Based on these findings, a DEM derived from ALS would be most

applicable in a transport corridor environment due to the presence of dense vegetation and the need for a fine spatial resolution.

2.4 Remote sensing of vegetation as a surrogate for characterising soil moisture

Section 2.2 discusses remote sensing techniques that infer soil moisture directly from the Earth's surface, whereas Section 2.3 refers to techniques that make an indirect characterisation of soil moisture from DEMs derived from remotely sensed data. This section centres on the characterisation of soil moisture through surrogate measurements. Specifically, the techniques discussed involve measures of vegetation spectral reflectance that can be used to make inferences about the soil below. This discussion will show that these techniques are potentially important for monitoring transport corridors as earthworks tend to be covered in dense vegetation, which can often hinder the use of the remote sensing techniques discussed in Sections 2.2 and 2.3.

Plants get their water from soil and therefore changes in soil moisture are felt by the plant (Porporato et al., 2001). The response of vegetation to changes in soil moisture is termed ecohydrology (Rodriguez-Iturbe et al., 2001) and can be categorised into two main responses: (1) an increase in water availability leads to an increase in nutrient availability, which results in an increase in growth or biomass (Goldberg and Miller, 1990); (2) an excess of water, leading to waterlogging, causing the depletion of oxygen in the soil, results in plant stress, highlighted by reduced chlorophyll production and overall biomass (Smith et al., 2004a; Trought and Drew, 1980).

To some extent this theory is put into practice by transport corridor inspectors as they walk railway routes and attempt to identify areas of increased biomass that could be indicative of an increase in soil moisture content (Perry et al., 2003b). However, there is no reference in the literature, or from engineering professionals, to any attempts made by inspectors to identify stressed vegetation as an indicator of waterlogged areas. In addition, there are no documented studies that have dealt with the response of vegetation to an increase in soil moisture on earthwork slopes. This is a potentially important area of research as vegetation spectral reflectance in the visible and near infrared domains has been shown to be able to identify stressed vegetation (Kempeneers et al., 2005; Li et al., 2005; Smith et al., 2005b; Adams et al., 1999; Carter, 1993), and

therefore has the potential to be scaled up to airborne sensors for application over wide areas (Kempeneers et al., 2005; Li et al., 2005).

2.4.1 Remote sensing of vegetation change

Since the launch of the first satellites with the capability to monitor the Earth's resources in the 1970s, researchers have developed methods to establish relationships between spectral reflectance and vegetation parameters (Ustin et al., 1999; Bannari et al., 1995; Rouse et al., 1974; Rouse et al., 1973). To understand the relationship between spectral reflectance and vegetation we must first consider the properties of plant leaves.

Figure 2.12 shows a cross section of a typical leaf. Towards the upper part of the leaf, the palisade layer is made up of elongated tissue that contains chloroplasts. These are cells that include lens shaped chlorophyll pigments that facilitate photosynthesis by absorbing sunlight. However, chlorophyll does not absorb sunlight equally. Light in the red and blue regions of the spectrum is absorbed by the pigments and green light is reflected, giving leaves their distinctive green colour. This top layer of the leaf, including the cuticle and upper epidermis, is almost invisible to infrared light. This passes through to the internal spongy mesophyll tissue where it is strongly scattered or reflected (Campbell, 2006).

Attempts to analyse vegetation using broad band satellite imagery has involved the development of vegetation indices, using the red and infrared regions of the spectrum, which have a particularly defined response to plant leaves (Bannari et al., 1995). The most common of such indices is the Normalised Difference Vegetation Index (NDVI), which was briefly introduced in Section 2.2.3. This was developed by Rouse et al. (1973) and remains extensively used today for many multispectral operations, using sensors with just a few spectral bands, such as Landsat with seven spectral channels. However, such indices were developed for broad band multispectral sensors which hold only limited information about vegetation targets (Cho and Skidmore, 2006; Thenkabail et al., 2000). The development of hyperspectral sensors, with near continuous reflectance values over the spectrum (of a given range) have increased the potential amount of information that can be derived about vegetation from their spectral reflectance (Thenkabail et al., 2000).

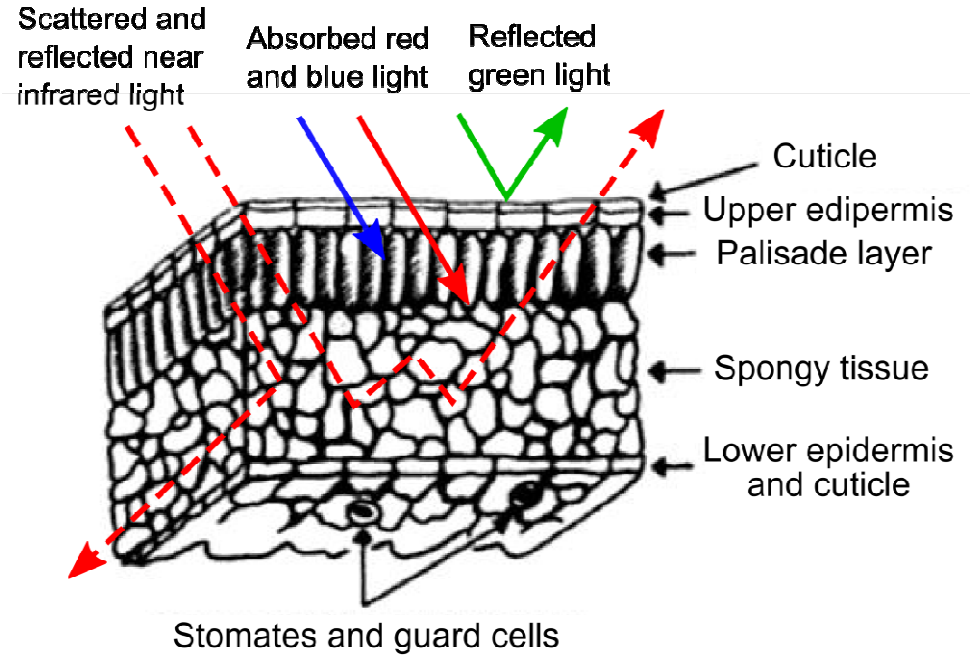


Figure 2.12 Cross section of a typical leaf showing interaction of leaf structure with visible and near infrared light. Adapted from Campbell (2006).

Features within a continuous spectral reflectance curve of a plant can reveal a number of characteristics, such as biomass, leaf area, senescence, stress and species type (Mather, 2004). More specifically, the red edge region of the electromagnetic spectrum is extremely sensitive to vegetation and as a result is possibly the most studied spectral feature (Baranoski and Rokne, 2005). The red edge (Figure 2.13) is defined by an abrupt change in reflectance from 680 nm (strong absorption) to 740 nm (strong reflectance). This occurs due to a combination of absorbance by leaf chlorophyll pigments and reflectance by internal scattering within the leaf (Baranoski and Rokne, 2005; Dawson and Curran, 1998).

A number of techniques have been developed to exploit the red edge feature. The following sub-section discusses those techniques with relation to changes in vegetation due to changes in soil moisture conditions.

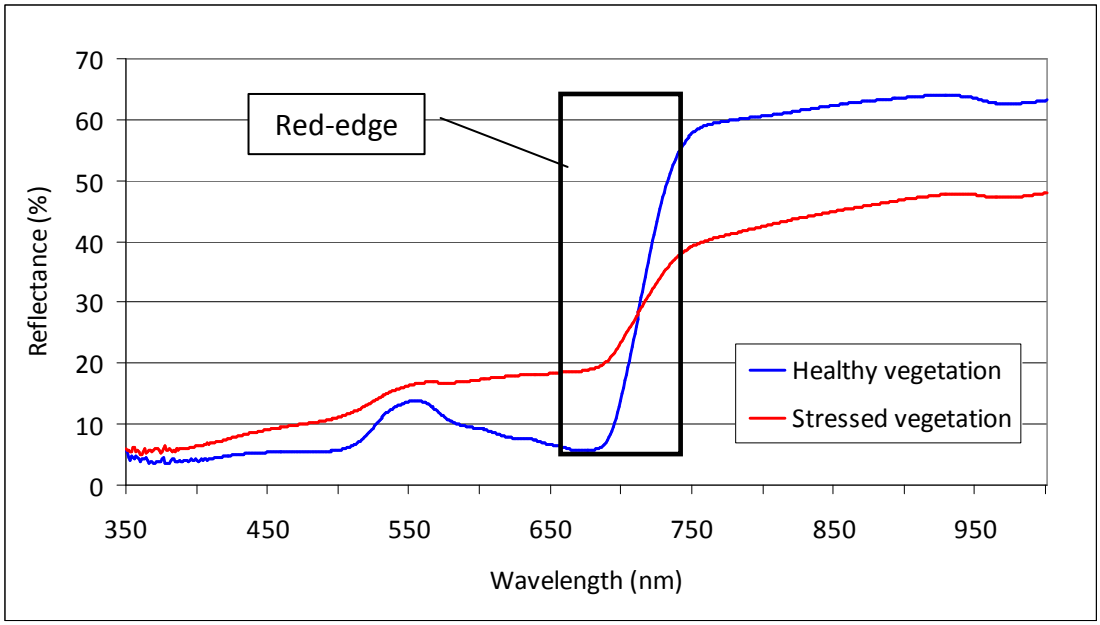


Figure 2.13 Example of healthy and stressed vegetation spectra measured over grass canopies using field spectroscopy with highlighted red edge feature.

2.4.1.1 Remote sensing of the red edge feature

Plant tissue under stress increases its heat production in order to dissipate excess energy, which can lead to a reduction in chlorophyll production (Zarco-Tejada et al., 2003). This propagates itself in the leaf reflectance through a shift in the red edge towards longer wavelengths (Cho and Skidmore, 2006; Noomen et al., 2006; Li et al., 2005; Smith et al., 2005b; Smith et al., 2004a; Merton, 1998). Authors have identified this shift by examining the red edge position (REP). For example, Smith et al. (2005b) found a shift in REP of up to 30 nm towards shorter wavelengths (from 718 nm to 688 nm) when plants were subjected to a range of different environmental stresses over a period of four weeks. In terms of soil moisture, an increase in soil moisture can lead to oxygen displacement in the soil, therefore reducing the availability of oxygen to a plant leading to stress (Smith et al., 2004a). Although many of the following studies do not explicitly address the impact of changes in soil moisture on the red edge feature, they all refer to different sources of plant stress.

The position of the red edge can be estimated using a number of different techniques. Perhaps the most basic technique for defining the REP uses linear interpolation to define a straight line between the point of minimum and maximum

reflectance, typically 670 nm and 780 nm respectively (as seen in Figure 2.13) (Dawson and Curran, 1998). The REP is subsequently derived from the slope of the line. This technique has been extended by employing an inverted Gaussian technique, that uses a least-squares approach to fit a normal curve to the red edge (Cutler and Curran, 1996). The REP is then defined as the midpoint between the rising limb of this curve. Although these two techniques are relatively simple, they require predefined points for the minimum and maximum points of reflectance which may alter between species and stage of senescence (Dawson and Curran, 1998). Furthermore, they do not take into account the possible influence of the canopy, including the effects of shadowing and soil background (Dawson and Curran, 1998).

More advanced techniques for estimating the position of the red edge largely focus on exploiting the derivative of the spectra. With the advent of hyperspectral sensors, many early studies concentrated on optimising simple band ratio calculations that were previously developed for multispectral sensors (Tsai and Philpot, 1998). Such techniques treated each band as an independent variable and ignored any band to band relationship (Tsai and Philpot, 1998). Derivative analysis overcame this by taking into account the surrounding bands and is therefore an expression of the rate of change in reflectance from one band to the next (Equation 2.2) (Schowengerdt, 2007; Smith et al., 2004a).

$$\rho' = \frac{\rho_{+1} - \rho_{-1}}{\lambda_{+1} - \lambda_{-1}}$$

Equation 2.2

Where, ρ' is the first derivative of the spectrum, ρ is the reflectance value and λ is the corresponding wavelength.

Although reflectance of vegetation in the red edge region of the spectrum is largely a function of leaf properties, there are also contributions from shadow, soil background and vegetation underneath the target (Dawson and Curran, 1998). The advantage of derivative analysis is that it maximises the signal from the vegetation target and suppresses background features (Dawson and Curran, 1998; Tsai and Philpot, 1998; Filella and Penuelas, 1994). As the derivative of the spectrum is concerned with change in reflectance, rather than absolute reflectance, it is also relatively insensitive to the effects of atmospheric scattering and illumination intensity (Dawson and Curran,

1998; Tsai and Philpot, 1998). This attribute makes analysis of the derivative of the spectra particularly useful for being applied to spectral reflectance data gathered from an airborne platform where variations in cloud cover and sun angle can change absolute reflectance values (Tsai and Philpot, 1998). Despite this, the derivative spectrum is known to be sensitive to noise (Tsai and Philpot, 1998). As a result, some authors have chosen to smooth spectra prior to calculating the derivative. Smith et al. (2004a) found that a five point weighted average provided sufficient smoothing without losing any fine scale spectral detail and is easy to calculate compared to other commonly used smoothing algorithms including the Savitzky-Golay least squares fitting procedure (Tsai and Philpot, 1998).

Studies have successfully used derivative spectra to improve estimations of vegetation parameters, compared to raw spectra (Malthus and Madeira, 1993). In particular, authors have noted the significance of two peaks in the first derivative of plant spectra (Figure 2.14) (Cho and Skidmore, 2006; Smith et al., 2004b; Zarco-Tejada et al., 2003). Zarco-Tejada et al. (2003) explained that this double peak feature is entirely the result of reflectance from chlorophyll pigments in the leaf. Extensive work by Smith et al. (2005a; 2005b; 2004a; 2004b) extended upon this theory by calculating a ratio of derivative reflectance values at 702 nm and 725 nm that correspond to the double peak feature. They found that this ratio was a strong indicator of stress in plants. However, the specific wavelengths where the double peaks occur may vary from plant to plant, as highlighted by the different peaks found by Smith et al. (2004b), Zarco-Tejada et al. (2003), Cho and Skidmore (2006) and Clevers et al. (2004). As a result, a predefined stress ratio may not be applicable to all species types and may have to be tailored to site specific vegetation types (Kempeneers et al., 2005). The stress ratio has not been tested for embankment vegetation in response to an increase in soil moisture.

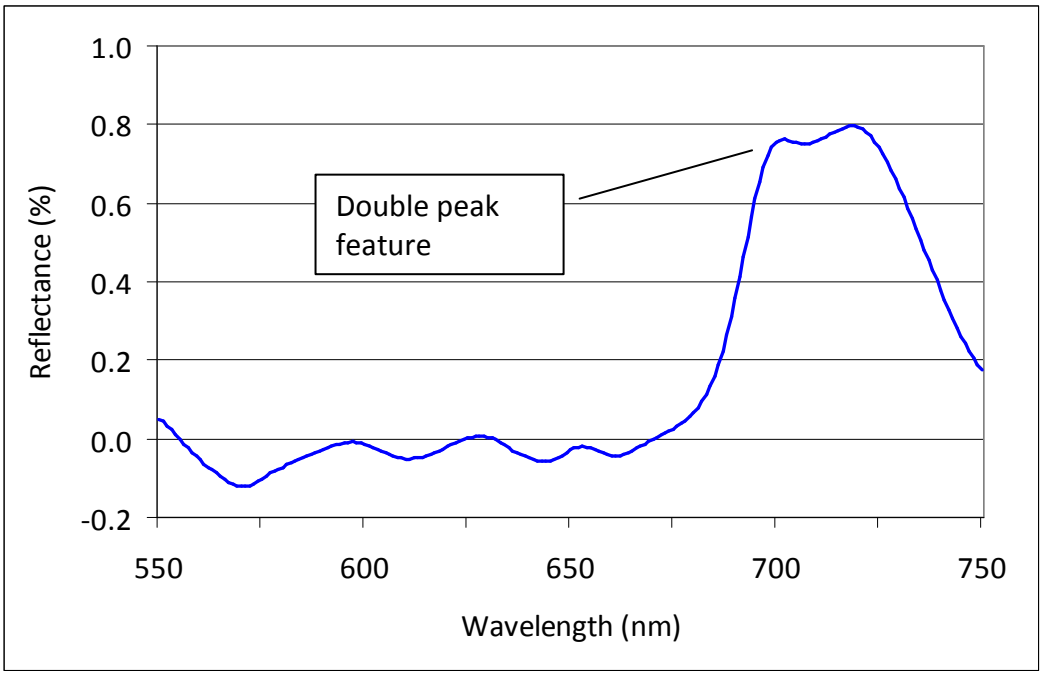


Figure 2.14 Example of the double peak feature at the maximum point of inflection in derivative spectra for the canopy of healthy grass measured using field spectroscopy.

Dawson and Curran (1998) presented a technique that has the benefits of being relatively simple to calculate, sufficiently flexible to allow for different points of minimum/maximum reflectance, and require no *a priori* knowledge about the spectrum being analysed. Their technique, known as the Lagrangian interpolation technique, uses the derivative of the spectra (ρ') (calculated using Equation 2.2). The calculation is described in Equation 2.3.

$$\rho'_{\lambda} = \frac{(\lambda-\lambda_i)(\lambda-\lambda_{i+1})}{(\lambda_{i-1}-\lambda_i)(\lambda_{i-1}-\lambda_{i+1})}\rho'_{\lambda(i-1)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})}{(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})}\rho'_{\lambda(i)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_i)}{(\lambda_{i+1}-\lambda_{i-1})(\lambda_{i+1}-\lambda_i)}\rho'_{\lambda(i+1)}$$

Equation 2.3

Where, ρ' represents the first derivative at a given wavelength, λ_i is the band with the maximum first derivative value, with λ_{i-1} and λ_{i+1} representing the two bands either side. $\rho'_{\lambda(i)}$, $\rho'_{\lambda(i-1))}$ and $\rho'_{\lambda(i+1)}$ represent the first derivative values which can be determined for any band position on a theoretical red-edge. In the case of Dawson and Curran

(1998), the wavelengths 695 nm, 700 nm and 705 nm were used (Figure 2.15). REP can then be calculated using Equations 2.4 and 2.5.

$$REP = \frac{A(\lambda_i + \lambda_{i+1}) + B(\lambda_{i-1} + \lambda_{i+1}) + C(\lambda_{i-1} + \lambda_i)}{2(A + B + C)}$$

Equation 2.4

Where,

$$A = \frac{\rho'_{\lambda(i-1)}}{(\lambda_{i-1} - \lambda_i)(\lambda_{i-1} - \lambda_{i+1})}, B = \frac{\rho'_{\lambda(i)}}{(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})}, \text{and}$$

$$C = \frac{\rho'_{\lambda(i+1)}}{(\lambda_{i+1} - \lambda_{i-1})(\lambda_{i+1} - \lambda_i)},$$

Equation 2.5

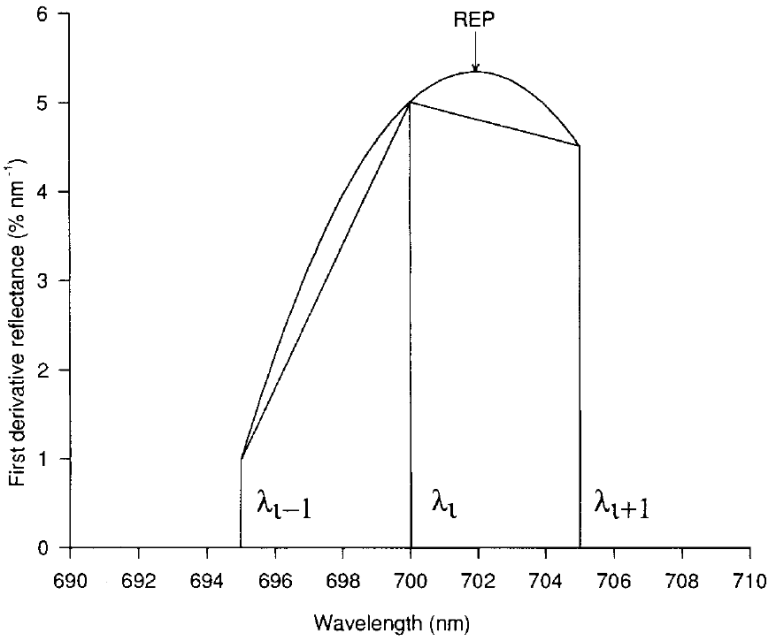


Figure 2.15 Example of derivative spectra for vegetation with the point of maximum inflection (λι) and the two points either side (λι-1 and λι+1) for use in calculating Lagrangian REP. From Dawson and Curran (1998).

Dawson and Curran (1998) demonstrated that the Lagrangian technique is comparable to the more complex Gaussian technique, making it better suited to large remotely sensed datasets due to reduced processing demands. This is supported by studies such as Li et al. (2005) who successfully used the Lagrangian technique to identify a shift in REP over two epochs of AVIRIS data, as a response to oil induced vegetation stress. They compared their results to REP estimated using polynomial fitting and found that the latter was less consistent in detecting a shift in the red edge over the whole image, compared to the Lagrangian technique. However, they also found that, unlike the polynomial technique, the Lagrangian method was sensitive to the band combination, meaning that it may provide different estimates of REP when applied to reflectance measurements from different sensors, due to different band combinations which are likely to be sensitive to slightly different wavelengths. Based on these findings they concluded that it was difficult to state which of the two methods should be favoured, but stated that the polynomial method may be preferred due its insensitivity to band combinations (Li et al., 2005).

A simplified calculation of REP can be found based on fitting a linear model to the red edge but assuming a straight line between reflectance at 700 nm and 740 nm. The reflectance of the red edge is then estimated to be half way between the absorption feature at about 670 nm and the reflectance feature at about 780 nm (Clevers et al., 2004). REP can then be defined by a linear interpolation between 700 and 740 nm using Equation 2.6.

$$REP = 700 + 40\left(\frac{\{(R_{670} + R_{780})/2 - R_{700}\}}{R_{740} - R_{700}}\right)$$

Equation 2.6

Where, R_{670}, R_{700}, R_{740} and R_{780} represent reflectance values at the wavelengths 670, 700, 740 and 780 nm respectively. The constants 700 and 40 are used to represent spectral range from 700 nm to 740 nm. Clevers et al. (2004) compared this technique with other methods, including 3rd and 6th order polynomial fitting, inverted Gaussian, and maximum of the first derivative, for estimating REP to detect vegetation stress caused by heavy metal contamination. They found that the maximum first derivative provided the only significant correlation, with R^2 values up to 0.64.

Despite the success of derivative analysis, some authors have noted that predictor models, based on multiple linear regression of derivative spectra, can be inconsistent when applied to different scales or sensors (Huang et al., 2004; Kokaly and Clark, 1999). To overcome this, Kokaly and Clark (1999) developed a technique known as continuum removal analysis which isolated the spectral feature of interest, in this case the red edge. The line of continuum is defined as a hull over an absorption feature, which is then removed by dividing the reflectance value at each wavelength along the line by the reflectance value of absorption feature (see Figures 2.16 and 2.17) (Mutanga et al., 2005; Curran et al., 2001). When predicting nitrogen content they found an improvement in the correlation coefficient from 0.65 to 0.85, when compared with derivative analysis (Kokaly and Clark, 1999). Noomen et al. (2006) studied the effects of plant stress induced by gas leakage through continuum removal analysis and found an increase in reflectance in the red region of 550 nm to 750 nm, that defines the red edge absorption feature (Figure 2.17). Mutanga et al. (2005) also used continuum removal analysis to relate spectral reflectance in the same region with plant nutrients using stepwise regression achieving high correlation coefficients (up to 0.77). Huang et al. (2004) found that improvements can be made over stepwise linear regression by using a partial least squares (PLS) approach.

The examples of Mutanga et al. (2005) and Kokaly and Clark (1999) demonstrated the potential for using continuum removal analysis to estimate vegetation parameters. Currently there have been no reported attempts to study the relationship between soil moisture and the continuum removed spectra of plant leaves. Some studies have used continuum removed spectra in the shortwave infrared over bare soils to estimate moisture content (Whiting et al., 2004). Similar to the studies cited in Section 2.2.4, excellent correlations with observed soil moisture can be obtained (R^2 up to 0.91); however, their methods would have little application over vegetated earthworks.

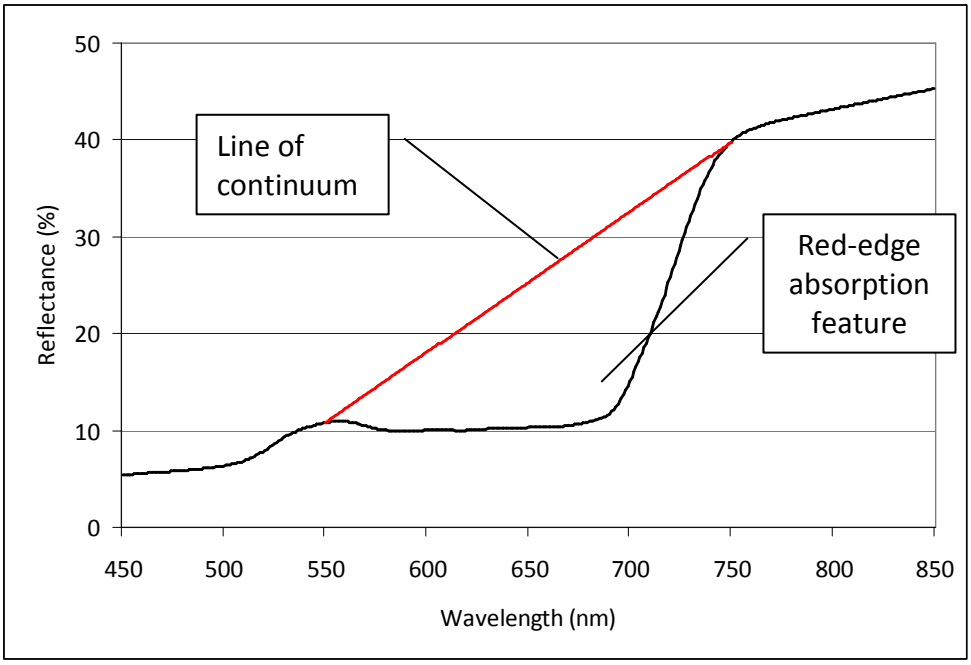


Figure 2.16 Red edge absorption feature with line of continuum for canopy reflectance from healthy grass measured using field spectroscopy.

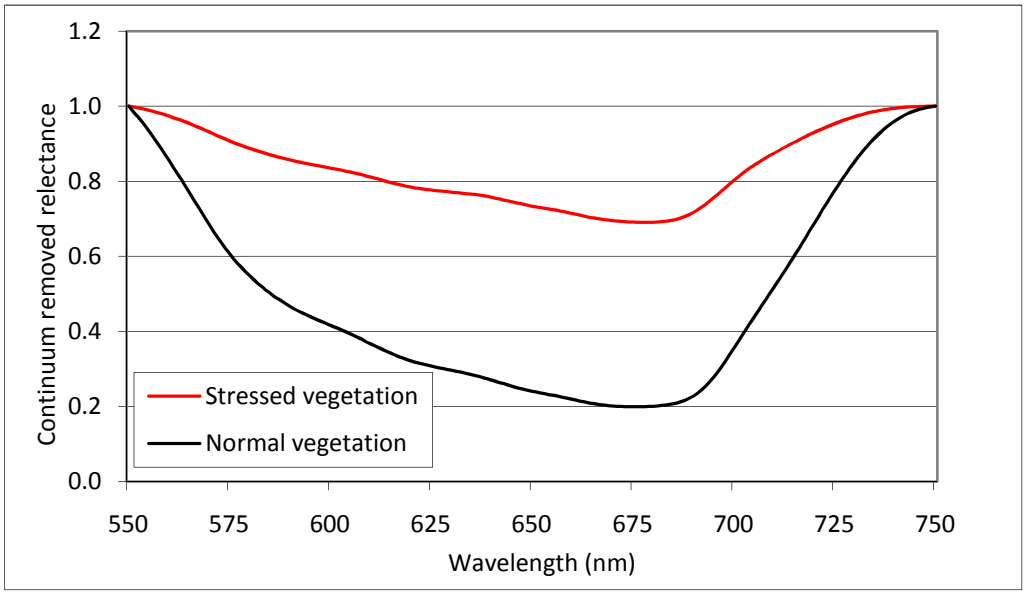


Figure 2.17 Continuum removed spectra for normal and stressed vegetation. Normal vegetation is characterised by a pronounced red edge absorption feature, identified by the depth of the continuum removed spectra, compared to the shallow depth of the stressed vegetation. Spectral measurements were made over grass canopies using field spectroscopy.

The use of spectral analysis techniques, such as stress ratios, red edge position estimation, derivative analysis and continuum removal analysis, have been shown to indicate stress in plants, as well as other vegetation parameters relating to plant health. The sources of stress discussed in this section include natural gas (Noomen et al., 2006; Smith et al., 2004b), methane (Noomen et al., 2006), ethane (Noomen et al., 2006), oil spills (Li et al., 2005), herbicide (Smith et al., 2005b), shade (Smith et al., 2005b), argon, nitrogen and waterlogging (Smith et al., 2004a). Only the latter study assesses the impact of soil moisture change, in this case an extreme increase, on the red edge feature. However, the impact of many of these sources of stress, including waterlogging, relate to oxygen displacement (Smith et al., 2004a). It could therefore be assumed that the techniques described in this section could possibly be used to identify areas of increased soil moisture over vegetated earthworks. However, relatively few of the aforementioned studies have attempted to apply REP detection techniques to heterogeneous vegetation covers (e.g. Li et al. (2005)) and therefore there is uncertainty as to how the developed techniques may work over vegetated earthworks. In addition, it is largely acknowledged that these techniques do not have the ability to differentiate between different sources of stress (Li et al., 2005; Smith et al., 2004a). In applying these techniques in an operational environment, it may therefore be necessary to make assumptions about the type of change being observed. Analysis of the derivative of the spectrum has the most promise for being used operationally due to its relative insensitivity to variations in illumination angle, background soil, and atmospheric attenuation (Dawson and Curran, 1998; Tsai and Philpot, 1998) and may therefore be better suited to multiple epoch airborne remotely sensed data, gathered from different sensor types.

2.4.2 Using vegetation as a bioindicator of soil moisture

Section 2.4.1 considers techniques that have the potential to identify vegetation stress, including discussion of how these methods might reflect short term responses of vegetation to changes in soil moisture conditions. However, this section is concerned with the long-term response of vegetation to soil moisture conditions in relation to plant species abundance.

Plant species are defined by their environment and can therefore be used to make inferences about local conditions (Hill et al., 2000; Schaffers and Sýkora, 2000). In this sense, plants can be seen as biological indicators and have been used as such by farmers for thousands of years (Diekmann, 2003). More recently, scientists have exploited the use of plants as bioindicators, citing the following advantages over conducting field measurements of environmental conditions: (1) plants express a summary of local environmental variables which can fluctuate strongly over time and space (Diekmann, 2003). This is particularly applicable to soil moisture where characterisations cannot be made using single measurements (Hawley et al., 1983). (2) Measurements of environmental variables tend to be made using expensive equipment and can often be laborious in comparison to floristic observations (Diekmann, 2003). (3) Plants tend to remain over time, therefore making it possible to make temporal observations of bioindicators (Diekmann, 2003).

The use of plants as bioindicators can be summarised by the concept of environmental gradients. This concept, outlined by Braak and Prentice (1988), states that all species occur in specific habitats based on their optimum environmental conditions. The theory explains that the abundance of a species is therefore a function of measureable environmental variables, such as soil moisture. A number of indicator scales have been proposed based on this theory. Examples include the C-S-R scales developed by Grime (2001). These scales are based on two categories that control the amount of dead or living plant matter in a habitat, notably stress (due to availability of light, water, nutrients, optimal temperature) and disturbance (due to erosion, fire, frost, being eaten etc). The scales group plant species as to whether they are a competitor (C), tolerant to stress (S), or a ruderal (R), this being plants that are first to colonise disturbed areas. Although the C-S-R scales have been successfully applied to study vegetation patterns across large habitats (608 ha) (Massant et al., 2009), they have no explicit link with soil moisture and therefore a link between these indictor scales and observed soil moisture may be tenuous.

Other examples of plant indicator scales tend to be purpose built for the study of a particular plant controlling variable (Hill et al., 2000). For example, Diekmann and Falkengren-Grerup (1998) developed an index for defining soil nitrogen content by forest tree species. Only two generalised indicator scales have been developed: Landolt

and Ellenberg indicator values (Hill et al., 2000). The former was developed as an extension of the Ellenberg values to be specifically tailored for Switzerland (Diekmann, 2003) and may therefore not reflect plant species outside this area.

The most widely used indicator scale is Ellenberg indicator values (Diekmann, 2003; Hill et al., 2000) which have been used as a standard for drawing comparisons with other scales (Massant et al., 2009; Diekmann and Falkengren-Grerup, 1998). Ellenberg indicator values were developed on the flora of central Europe (Ellenberg et al., 1991) and assign indicator values to plant species relating to specific environmental variables, namely, soil nitrogen content, pH, soil chloride concentration, light regime, temperature, continentally and soil moisture (Schaffers and Sýkora, 2000). The definitions of the soil moisture indicator values are summarised in Table 2.2.

Many studies have compared Ellenberg indicator values to field measurements (Schmidtlein and Sassin, 2004; Ewald, 2003; Schaffers and Sýkora, 2000; Ertsen et al., 1998). Schaffers and Sýkora (2000) tested the reliability of Ellenberg values to predict environmental conditions for vegetation in the Netherlands. They adopted two weighted averaging approaches to relate the indicator values to field measurements, species based and site based. For the plant based method, an optimal value for each environmental variable was calculated per species type. Correlations were then made between these optimal values and the species indicator values. For the site based method, average indicator values were calculated per site according to their abundance and compared against actual measurements. The results showed best results with the site based method, achieving strong correlations with soil moisture (R^2 between 0.8 and 0.9).

Despite the apparent success of Ellenberg value indicators there are a number of well documented criticisms associated with their development and subsequent application. One of the main criticisms is that they were developed for species in central Europe and should therefore warrant little meaning outside this area (Hill et al., 2000). This prompted Hill et al. (2000) to conduct a large scale survey of British plant species, in order to extend Ellenberg indicator values for the UK. Currently there have not been any published investigations into the reliability of these extended Ellenberg values but Hill et al. (2000) provided their own critical assessment.

Ellenberg moisture value	Definition
1	Indicator of extreme dryness, restricted to soils that often dry out for some time
2	Between 1 and 3
3	Dry-site indicator, more often found on dry ground than in moist places
4	Between 3 and 5
5	Moist-site indicator, mainly on fresh soils of average dampness
6	Between 5 and 7
7	Dampness indicator, mainly on constantly moist or damp, but not wet soils
8	Between 7 and 9
9	Wet-site indicator, often on water saturated, badly aerated soils
10	Indicator of shallow-water sites that may lack standing water for extensive periods
11	Plant rooting under water, but at least for a time exposed above, or plant floating on the surface
12	Submerged plant, permanently or almost constantly under water

Table 2.2 Moisture scale and definitions as defined by Ellenberg et al. (1991).

One of the principal acknowledgements made by Hill et al. (2000) was that no effort was made to ensure their vegetation sampling was unbiased. In fact, the contrary was adopted to deliberately represent particular vegetation types (Hill et al., 2000). Two surveys were included in their work, one biased towards common vegetation types, seldom recording rarer species, and the other skewed toward rare species and special habitats. They also discussed their modified indicator values in terms of being plant-centred, meaning they referred to the immediate environment in which they were found,

which can have implications when recording indicator values over relatively large plots. However, they acknowledged the fact that indicator values worked best when averaged out for a specific area and therefore the issue of plant centred definitions may not be appropriate.

A thorough review of bioindicators was conducted by Diekmann (2003), with particular emphasis on Ellenberg indicator values. The first criticism is that they rely on field observations and can therefore have a degree of subjectivity. This can be of particular relevance where the field study is conducted by inexperienced ecologists. It is therefore essential to have adequate knowledge of plant species in the study area. The review also highlighted the potential for circularity, in that indicator values are driven by floristic observations, which are then used to make inferences about habitat characteristics. The suggestion was that indicator values must be calibrated to environmental measurements. However, this could defeat the object of using bioindicators as a replacement for laborious field measurements of environmental components. To a degree this could be compensated for by drawing comparisons between indicator values and coincidental field measurements. Further criticisms include the fact that the response of plant species to environmental conditions may change over time and space (Diekmann, 2003). Again, the suggestion is to calibrate the indicator values with measured values, which may detract from the operational capabilities of using bioindicators.

2.4.2.1 Remote sensing of Ellenberg indicator values

The application of bioindicators is growing, and as a result, they are becoming increasingly accepted (Diekmann, 2003). Despite this, field based methods for mapping species abundance tend to be laborious and therefore assessments are made at a small scale, typically stand level (Schmidtlein, 2005). As a result, Schmidtlein (2005) explored the use of remotely sensed data to map bioindicators over wide areas by exploiting the strong relationship between plant assemblages and reflectance.

Authors have noted that plant species composition is difficult to detect using remote sensing due to the differences between ecological mapping methods and the methods employed in remote sensing (Lewis, 1998). Mapping bioindicators, however, skips the vegetation classification step which tends not to exploit all the spectral information provided by imaging spectroscopy and as a result leads to a loss of

ecological information that might otherwise be gained (Schmidtlein and Sassin, 2004). Despite this, there have only been a limited number of attempts to map bioindicators using remotely sensed data, perhaps due to the inadequacy of sensor resolution to identify particular species. This is supported by Schmidtlein (2005) who explained that attempts have only been made where vegetation cover is confined to one or two species, where more complex vegetation canopies are not adequately represented by coarse resolution sensors. Schmidtlein (2005) overcame this by mapping indicator values directly as a function of vegetation reflectance, using fine spectral resolution imaging spectroscopy (see Figure 2.18 for the resulting mapped soil moisture indicator values).

The method adopted by Schmidtlein (2005) was to make field observations of species type and relative abundance in 46 relevés, or quadrats, with a radius of 1 m. Average Ellenberg values were then calculated for each plot, which were weighted around their relative abundance. The position of these sites were recorded using differential GPS and were located within imagery obtained using an airborne AVIS-2 (Airborne Visible near Infrared Imaging Spectrometer) sensor, with a spatial resolution of approximately 2 m and a spectral coverage from 411-868 nm over 64 bands. Partial least squares (PLS) regression was then used to model the indicator values from the imagery. This method was favoured over multiple linear regression for hyperspectral data as the latter takes into account just a few spectral bands that are not collinear (Schmidtlein, 2005; Huang et al., 2004; Schmidtlein and Sassin, 2004). Reasonable correlations were found when compared to observed soil moisture values (R^2 of 0.58).

The correlations experienced by Schmidtlein (2005) were not as strong as the predictions made using the direct remote sensing techniques reported in Section 2.2. This is largely due to the fact that the direct remote sensing techniques were largely applied to areas with bare soils or sparse vegetation canopies. However, this could also be attributed to the fact that bioindicators are floristic gradients, and would therefore be a closer representation of floristic characteristics rather than actual soil attributes. This is highlighted by the results experienced by Schmidtlein (2005); for example, areas of high soil moisture were found to be characterised by a relatively high absorption in the green region of spectrum, suggesting low chlorophyll content in wetland plant species. However, other areas of high moisture content exhibited the opposite trend, with absorption in the red and a strong change in reflectance over the red edge, which

identified the differences between poor and rich wetland vegetation. In addition, the response of vegetation to different changes in environmental conditions can be the same. For example, the effects of low nutrient supply can have the same effects as low water supply (Schmidtlein and Sassin, 2004). However, ecologists tend to treat spatial variation in plant species as a result of long-term conditions and therefore bioindicators can be seen as expressing the overall characteristics of a soil rather than expressing short-term changes in environmental conditions (Schmidtlein and Sassin, 2004).

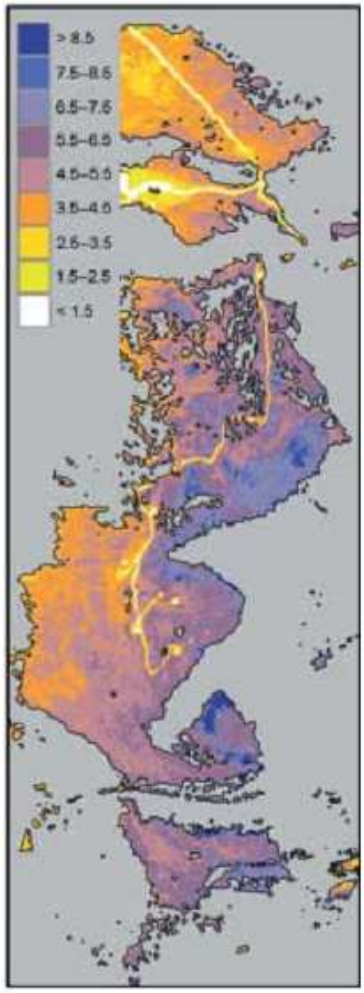


Figure 2.18 Map of soil moisture predicted by mapping Ellenberg moisture indicator values over alpine pasture, near Salzburg, Austria, using airborne hyperspectral data. From Schmidtlein (2005). Scale corresponds to the Ellenberg categorical values detailed in Table 2.2.

Despite these criticisms, the fact that reasonable soil moisture maps can be predicted over complex vegetated sites gives this work great potential for monitoring transport corridor earthworks. Furthermore, sensors, including CASI (Compact Airborne Spectrographic Imaging), AVIRIS (Airborne Visible InfraRed Imaging Spectrometer) and HyMap (Hyperspectral Mapper), are available that have a greater spectral range than the AVIS sensor used by Schmidtlein (2005), and could therefore hold further potential for this technique. Schmidtlein (2005) also identified the advantage of using field measurements of reflectance, rather than extracting spectra from the imagery, as this provided targets with more homogenous vegetation cover, and helped to enhance the overall model. However, this would require considerable field work and would be labour consuming for use over a transport corridor stretching tens of kilometres.

Although there is clearly potential with mapping bioindicators through imaging spectroscopy, there are a number of shortcomings associated with Schmidtlein's (2005) methods. Topographic normalisation was not applied to the imagery prior to analysis, which could have implications on illumination effects. Despite this, subsequent error analysis did not identify any significant relationship between model errors and topographic effects on illumination, despite some overestimation of soil moisture in areas of shadow. Furthermore, no consideration was given to the effects of canopy structure. At the leaf level, reflectance is largely governed by pigments, water content, carbon content, and has a significant control over reflectance at the canopy level. However, different ecosystems have very different reflectance properties at the canopy level. Leaf area index (LAI) has the strongest influence on canopy reflectance, which relates to the amount of green leaf area per unit ground area. Mean leaf angle also has a strong influence on canopy reflectance, for example, vegetation with vertically orientated foliage has very different reflectance properties to vegetation with horizontally orientated foliage, such as deciduous trees (Asner, 1998).

The importance of canopy structure is reinforced by Schmidtlein's study (2005) which characterised wetter areas by a small LAI, such as bog mosses, but the opposite could also be found with dense, higher canopies. Despite this, the study acknowledged the flexibility of the PLS approach to modelling, as it is able to take into account these differences by handling variations over a number of input spectral bands. In addition,

this highlights the fundamental advantage of multi/hyperspectral remote sensing over simple band ratio methods that do not embrace the same spectral detail.

Studies have taken into account the effects of canopy structure on reflectance through the use of canopy radiative transfer, or geometric-optical models (Bateson et al., 2000). There are a number of such models with varying degrees of complexity, from simple 1D models, to complex 3D models. Jacquemoud et al. (2000) discussed the various merits of such models and concluded that, for operational purposes, a 1D model is sufficient as it is relatively simple to run and requires a small number of parameters, but carries enough detail to improve a sensor's representation of a canopy. The 1D models PROSAIL, PROSIAPI, and PROKUUSK were all considered easy to compute, producing very few differences (Jacquemoud et al., 2000). Among the models tested, PROSAIL is considered to be the most widely used.

There are two approaches to deriving vegetation parameters for spectral reflectance, the empirical or statistical approach, and the physical approach (Darvishzadeh et al., 2008). The statistical approach is similar to that used by Schmidtlein (2005) by which correlations are made between the target variable (Ellenberg indicator values) and spectral reflectance. Such methods can be considered sensor specific and are dependent on the conditions in which sampling was undertaken, and as a result can change over time (Darvishzadeh et al., 2008; Meroni et al., 2004). The physical method involves the use of radiative transfer models that describe variation in spectral reflectance as a function canopy, leaf and soil background characteristics. As the physical method is based on the laws of physics it can offer an important link between canopy reflectance and the vegetation characteristics of interest (Darvishzadeh et al., 2008; Meroni et al., 2004). However, there is an argument to say that mapping soil moisture indicator values for slope stability in transport corridors is not concerned with changes over time, merely an overall characterisation of soil conditions and may not benefit from radiative transfer modelling.

Inputs to radiative transfer models require information about variables such as chlorophyll concentration, carotenoid concentration, water thickness, dry matter content, and a leaf structure parameter (Bateson et al., 2000). However, to quantify these parameters prior to model calculation would defeat the object of providing an operational technique. Furthermore, and possibly most crucially, radiative transfer

models define the canopy as a horizontally homogenous layer and can therefore only be suitable for largely homogenous vegetation canopies (Darvishzadeh et al., 2008; Meroni et al., 2004), whereas the nature of vegetated earthworks can often be complex, heterogeneous canopies.

2.5 Conclusion

This chapter has provided a critique of direct, indirect and surrogate techniques for characterising soil moisture using remotely sensed data, which is summarised at the end of this Section in Table 2.3. Of these techniques, the direct methods, using active microwave sensors are most well established and are considered to have the greatest potential for being used operationally. Specifically, soil moisture estimations made using these sensors have been shown to have good correlations with observed soil moisture. In addition, using microwave remotely sensed data at low frequencies (longer wavelengths) it is possible to penetrate vegetation canopies. However, due the inherent limitations on the spatial resolution that can be achieved by such sensors, both using satellite and airborne platforms, these techniques are unable to offer a viable solution to monitoring transport corridor earthworks. Active microwave sensors, such as TerraSAR-X, have the ability to provide finer spatial resolution data but these systems use high frequency (shorter wavelengths) microwaves that are highly susceptible to the effects of surface roughness and vegetation cover. In order to make reliable estimations using these systems, highly parameterised models are needed to correct for these effects, which require accurate information regarding vegetation characteristics and soil properties. As a result, these methods offer little operational potential over earthworks with complex vegetation covers.

Remotely sensed data using thermal sensors have also showed promise for characterising soil moisture distribution. Again, high correlations have been found with observed soil moisture measurements but studies tend to be restricted to bare soils as electromagnetic energy in this spectral domain cannot penetrate vegetation canopies. Successful attempts have been used to exploit information from red/near infrared vegetation indices to inform soil moisture retrieval from thermal sensors. Although good correlations were found in semi arid areas, where vegetation cover is sparse,

poorer results were obtained over denser canopies, conditions which are more comparable with temperate environments, found in UK transport corridor environments.

Direct soil moisture retrieval using reflectance measurements in the visible, near and shortwave infrared share characteristics with estimations made using thermal sensors. Whereas estimations made using these sensors have provided excellent correlations with observed soil moisture, electromagnetic energy over these wavelengths are unable to penetrate vegetation canopies and, therefore, estimations have been limited to bare soils. To an extent, the presence of vegetation has shown to be a limiting factor for all the techniques discussed in this Chapter relating to direct estimations of soil moisture. As a result, the characterisation of soil moisture distribution using indirect techniques was explored.

Indirect techniques using terrain analysis of DEMs, namely TWIs, are well established in hydrology and slope stability studies for delineating soil moisture distribution. Furthermore, when these calculations are applied to fine spatial resolution DEMs generated from ALS data, they have potential for use in monitoring transport corridor earthworks. However, these techniques have shown to have mixed results depending on a number of factors, including the spatial resolution of the DEM and the overall wetness conditions when validating observed soil moisture measurements are collected. Therefore, soil moisture characterisation using these techniques should be used in conjunction with other sources of data to improve predictions. Suggestions are to incorporate an estimation of solar radiation which has shown to improve correlations with observed soil moisture.

Despite the best efforts of vegetation removal techniques, the accuracy of a DEM generated from ALS datasets can be affected by the presence of vegetation. As a result, techniques are explored that use surrogate measures from remotely sensed data to characterise soil moisture conditions below, specifically, the techniques discussed refer to the use of vegetation canopy reflectance. The advantage of using these methods is that they can be used over vegetated areas which has been a major limiting factor for the other techniques discussed in this chapter. However, there have been few reported attempt to verify the estimation of soil moisture using these techniques.

The techniques using vegetation reflectance as surrogate measures of soil moisture conditions fit broadly into two categories: (1) methods for identifying

vegetation stress, and (2) techniques for mapping plant indicator values. There are a number of examples where spectral analysis techniques have been used to detect the onset of plant stress. In particular, derivative analysis has shown to be particularly sensitive to different sources of stress and have the advantage of being relatively unaffected by differences in illumination which may occur when comparing data from different sensors, or different epochs of data, for the same location. However, authors acknowledge that these techniques are unable to differentiate between different sources of stress, and therefore assumptions may have to be made when using them to detect changes in vegetation due to differences in soil moisture content. Despite the range of studies using spectral analysis for stress detection, very few have compared their results to actual observed soil moisture, with focus tending to be on the effects of extreme soil moisture conditions, i.e. waterlogging. As a result, the potential for using these technique for characterising soil moisture is unknown.

The other vegetation reflectance analysis technique considered in this chapter is mapping plant indicator values. Ellenberg indicator values have been shown to be the most widely used, which include specific categories detailing optimal soil moisture conditions for each plant species. Ellenberg values have been mapped using airborne imaging spectroscopy with relative success for central Europe, which is the area they were originally developed for. Ellenberg values have been extended for UK plant species but currently there have been no reported attempts to map these using remotely sensed data. The limitation of this technique is that it requires a detailed survey of species abundance. Despite this drawback, its ability to extrapolate environmental conditions, including soil moisture, over wide areas gives it great potential. In addition, the sensors that have been used to apply this method have a fine spatial resolution which would be adequate for monitoring transport corridor earthworks.

Technique	References	Correlations	Spatial resolution	Advantages	Limitations
Direct					
Passive microwave	(Sahoo et al., 2008; Jackson et al., 1999)	Very high Sahoo et al. (2008) R^2 of 0.81.	Footprints typically ~50 km with satellite sensors	Relatively unaffected by atmospheric conditions High correlations with observed soil moisture	Affected by surface roughness and presence of vegetation Limited to poor spatial resolutions
Active microwave	(Pierdicca et al., 2010; Le Hegarat-Mascle et al., 2002; Quesney et al., 2000)	Very high Quesney et al. (2000) R^2 of 0.89. Le Hegarat-Mascle et al. (2002) R^2 between 0.63-0.85.	Footprint typically above 7 m using airborne SAR. TerraSAR-X satellite has a potential footprint of 1 m but using X band which is highly affected by vegetation	Relatively unaffected by atmospheric conditions High correlations with observed soil moisture Higher spatial resolution than passive systems	Affected by surface roughness and presence of vegetation Require complex radiative transfer models with accurate information on vegetation parameters.
Thermal imagery	(Minacapilli et al., 2009; Crow et al., 2008; Lu et al., 2007; Sandholt et al., 2002)	High Minacapilli et al. (2009) R^2 of 0.75 Crow et al. (2008) R^2 of 0.84 Sandholt et al. (2002) R^2 of 0.81	Spatial resolution up to 2 m using airborne sensors	Can provide fine scale imagery High correlations with observed soil moisture Some success over vegetated areas	Require highly parameterised physically based models Highly perturbed by the presence of vegetation Success over vegetated areas reliant on accurate information on vegetation cover

Visible, near and shortwave infrared	(Lihua et al., 2005; Ben-Dor et al., 2002; Lobell and Asner, 2002; Weidong et al., 2002)	Very high Lihua et al. (2005) R^2 of 0.9 Weidong et al. (2002) R^2 between 0.7-0.85. Ben-Dor et al. (2002) between 0.65-0.85.	Studies conducted with 8 m spatial resolution but there are airborne sensors with the ability to provide resolutions up to 1 m	High correlations with observed soil moisture Sensors in this spectral domain can provide high spatial resolutions	Estimations affected by soil properties Models need to be tailored to specific soil types Only applicable to bare soil Largely restricted to laboratory studies
Indirect					
Topographic wetness index	(Tenenbaum et al., 2006; Schmidt and Persson, 2003; Western et al., 1999a)	Variable Schmidt and Persson (2003) R^2 of 0.64 Western et al. (1999a) R^2 between 0.42-0.61 (with potential solar radiation) Tenenbaum et al. (2006) R^2 between 0.56-0.8	Up to 0.5 m	Intrinsic relationship with areas of slope instability Capability of high spatial resolution using ALS systems Potential for good correlations when incorporated with potential solar radiation	Accuracy of DEM generation affected by presence of vegetation Variable results according terrain, DEM resolution, overall wetness conditions during observed soil moisture sampling

Potential solar radiation	(Du et al., 2007; Fu and Rich, 2002; Western et al., 1999a)	Unknown With TWI Western et al. (1999a) found R^2 between 0.42-0.61	Reported attempts up to 5 m but could be applied to ALS DEM with grid resolutions of less than 1 m	Improves estimations made using TWI Better characterisation of soil moisture distribution during overall drier conditions Some models are simple to compute	Very few comparisons with observed soil moisture Some models are highly parameterised
Surrogate					
Plant stress indicators	(Noomen et al., 2006; Smith et al., 2004a; Zarco-Tejada et al., 2003; Kokaly and Clark, 1999; Tsai and Philpot, 1998)	Unknown	Techniques can be applied to airborne hyperspectral sensors, some with a spatial resolution of less than 1 m	Relate to transport network inspectors identifying increased biomass due to increase in soil moisture Can be used over vegetated earthworks Potential to be applied to high spatial resolution airborne imagery Potential to be applied to different vegetation types	Potential for monitoring soil moisture unknown Few reported attempts to compare against observed soil moisture Success over complex vegetation covers unknown Unable to differentiate between different sources of stress

Mapping biological indicator values	(Feldmeyer-Christe et al., 2007; Schmidtlein, 2005)	Largely unknown Schmidtlein (2005) found R^2 of 0.58	Has been applied at a spatial resolution of 2 m but has potential to be applied to finer scale imagery of less than 1 m	Can be used over vegetated earthworks Potential to be applied to high spatial resolution airborne imagery Strong correlations found between indicator values and observed soil moisture	Mapping indictors has not been assessed for UK Require field observations of vegetation abundance Uncertainties over tress canopies

Table 2.3 Summary of the techniques discussed in Chapter 2 with reference to their advantages, disadvantages and potential spatial resolution.

3 Methodology for characterising soil moisture in transport corridor environments

Chapter 2 identified a number of techniques that have the potential to characterise soil moisture in a transport corridor environment. These techniques included terrain analysis calculations, such as TWI and potential solar radiation. Reflectance from vegetation canopies in the visible and near infrared region of the spectrum have been shown to be useful for detecting different sources of stress (Noomen et al., 2006; Smith et al., 2005b). However, there has been little work that explicitly focuses on the influence of changes in soil moisture on vegetation reflectance, and no reported attempts to apply such techniques to transport earthwork infrastructure. Mapping bioindicators have also been shown to have potential for charactering soil moisture (Schmidtlein, 2005) but there have only been a few reported attempts to do this using remotely sensed data and no attempts have been made over a transport corridor. Furthermore, there have been no reported attempts to map and validate the Ellenberg bioindicator values that have been extended for the UK in any environment.

This chapter details the methods that will be used to assess the potential of the aforementioned techniques applied to a transport corridor environment. In addition, methods designed to overcome the shortcomings associated with these individual techniques are explained. Some of the techniques considered here are conceptual and are therefore initially tested at an earthwork scale, using terrestrial remote sensing technologies. The techniques deemed successful, as well as other techniques that could not be assessed at a fine scale, are then applied at a broader scale, using airborne remotely sensed data for a UK transport corridor.

A summary of the methods described in this chapter can be found in Figure 3.1. This shows two distinct sections to the overall workflow, the earthwork scale experiment (described in Section 3.2) and the transport corridor scale work (Section 3.3). The flow diagram indicates that the techniques with potential for characterising soil moisture in a transport corridor environment, as identified in Chapter 2, are carried on to an experimental stage. This begins with a series of earthwork scale studies which are used to assess specific techniques, namely, terrain analysis techniques applied to a fine spatial resolution DEM (TWI, solar radiation and aspect) (see Section 3.2.1) and

spectral analysis of field spectroscopy to identify vegetation stress as a response to artificial wetting (Section 3.2.2). The diagram continues to show how the terrain analysis work is carried forward to inform how best to apply these techniques to a transport corridor scale. Conversely, the field spectroscopy study is not carried on and therefore does not link with the transport corridor scale work. This is due to the uncertainties associated with applying such techniques to relatively complex vegetation cover in transport corridor environments, which is later explained in Section 5.1.2. The lower part of the diagram, summarising the transport corridor work, shows how analysis is carried out using airborne remote sensing systems, namely lidar and CASI datasets. The terrain analysis techniques are applied to the lidar data in a similar fashion to the earthwork scale experiment (Section 3.3.1). The CASI data is used to map Ellenberg indicator values (Section 3.3.2), a technique that could not be assessed at a fine scale due to the relatively homogenous vegetation cover of the test embankment and therefore appears here for the first time. The diagram then shows how the terrain analysis techniques are brought together in an integrated model and compared against observed soil moisture measurements (Section 3.3.3). Although Figure 3.1 provides a coarse summary of the methods employed, a more detailed version, including a series of sub diagrams detailing the individual work packages, is provided in Appendix 1.

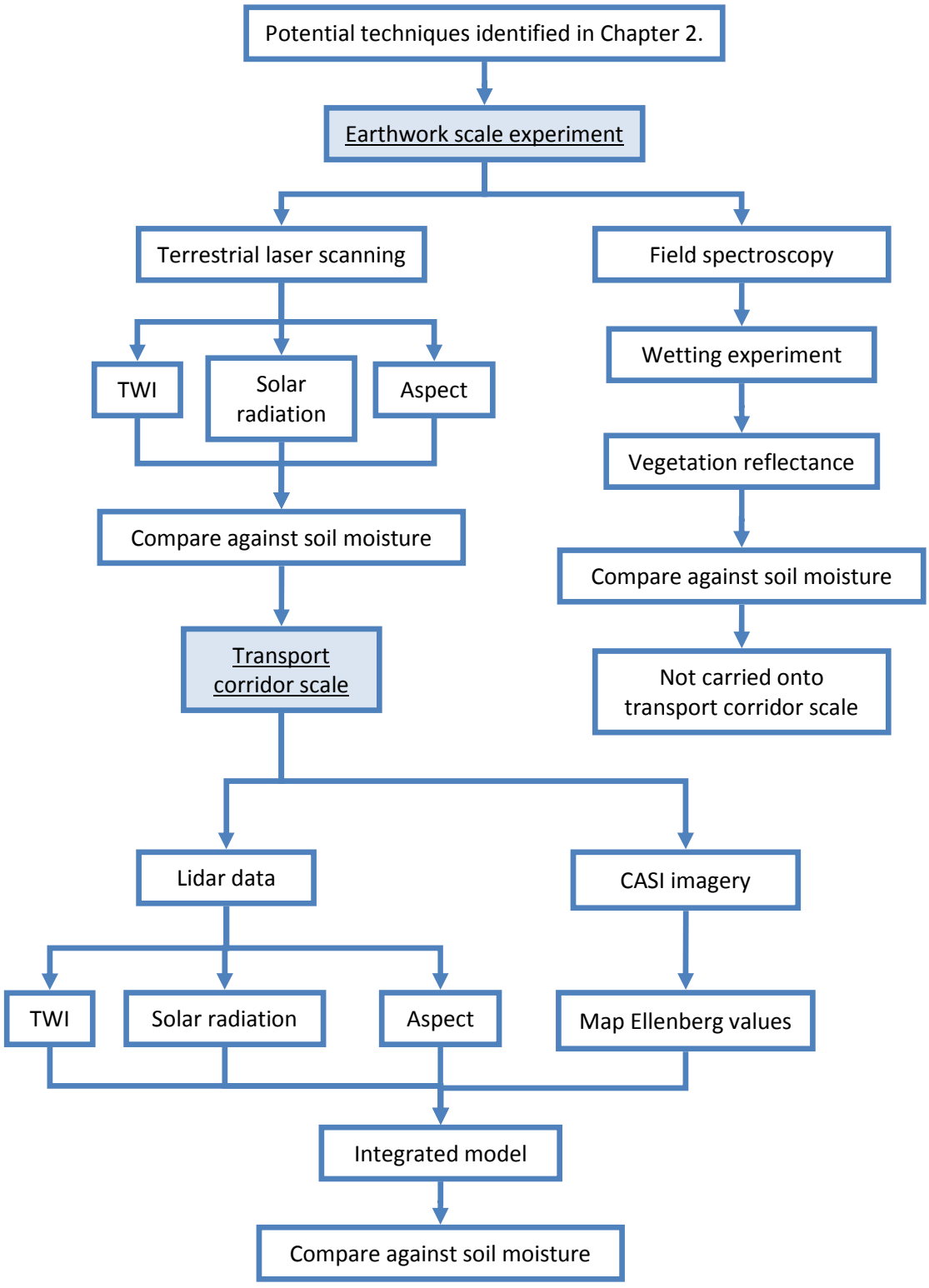


Figure 3.1 Flow diagram summarising the methods described in this Chapter. A more comprehensive set of flow diagrams detailing the methods can be found in Appendix 1.

3.1 Study sites

The analysis undertaken at an earthwork scale was carried out on a test embankment located at Nafferton Farm, Northumberland, which is managed by Newcastle University's School of Civil Engineering and Geosciences (Hughes et al., 2009). The embankment was constructed to establish a facility for engineering and biological research for the purposes of determining the effects of climate change on earthwork embankments (BIONICS, 2009). The earthwork measures 90 m in length, 29 m across and 6 m high, and has an approximate northwest, southeast facing aspect for either side of the earthwork (Figures 3.2 and 3.3). On the face of each slope are four separate panels, two constructed to the requirements of modern earthwork structures, and two poorly constructed to mimic the poorly compacted nature of historically constructed earthworks typical of transport infrastructure in the UK.

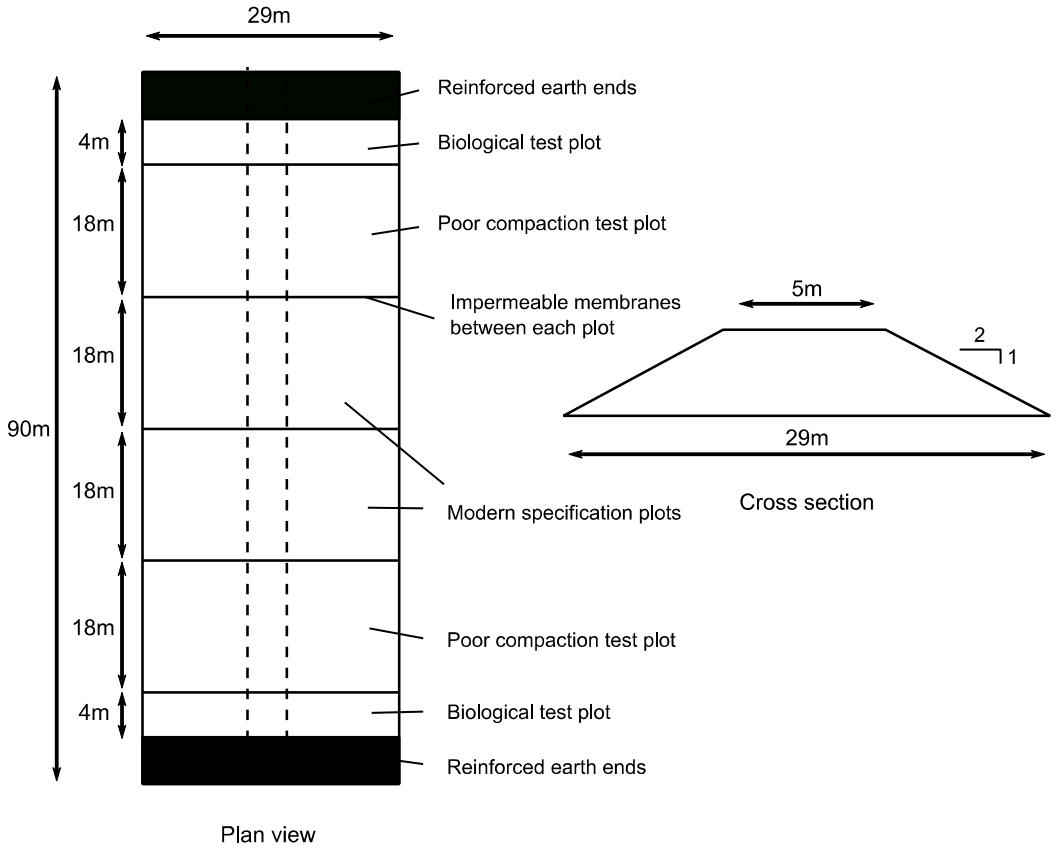


Figure 3.2 Diagram of the BIONICS test embankment located at Nafferton Farm, Northumberland. Modified from BIONICS (2009).

Figure 3.3 The BIONICS test embankment viewed from an approximately easterly direction. Various markers on the slope indicate the location of embedded instruments used in other studies. Blue piping, used for a sprinkler system, mark the approximate edges of the earthwork test plots.

The test site used for the transport corridor scale study was a four kilometre stretch of the railway line near Haltwhistle, located between Carlisle in the west and Newcastle upon Tyne to the east (Figure 3.4). This railway was constructed over 100 years ago and has been identified by Network Rail as having a history of instability (Hall, 2009) with numerous minor landslides occurring over recent years (for example see BBC (2007)). The transport corridor lies largely within the floodplain of the South Tyne River, with some stretches of embankment earthworks located within close proximity of the river (approximately fifteen metres). Vegetation conditions within the test site are representative of those found on the UK rail network, and across Northern England in particular (Miller et al., 2009). A mixture of grasses, agricultural weeds, small shrubs and deciduous trees dominate, but bare earth slopes, and stretches of dense

deciduous woodland are also relatively commonplace (Figures 3.5 and 3.6). This presents a relatively challenging environment in which to characterise soil moisture. As discussed in Chapter 2, many other studies that have attempted to estimate the distribution of soil moisture have been carried out over bare earth or sparse vegetation canopies, such as Burke et al. (1998) due to the complexities presented by multiple canopy layers (Schmugge and Jackson, 1994). Specifically, this can have an influence on optical remote sensing techniques where the signal can be confused by differences in leaf angle relative to the sun and the sensor position (Wolf et al., 2010). In addition, dense vegetation, as depicted in Figure 3.6 can obscure the ground reducing the effectiveness of laser scanning systems (Reutebuch et al., 2003). Resulting DEMs interpolated from such datasets are therefore likely to have a degree of uncertainty.

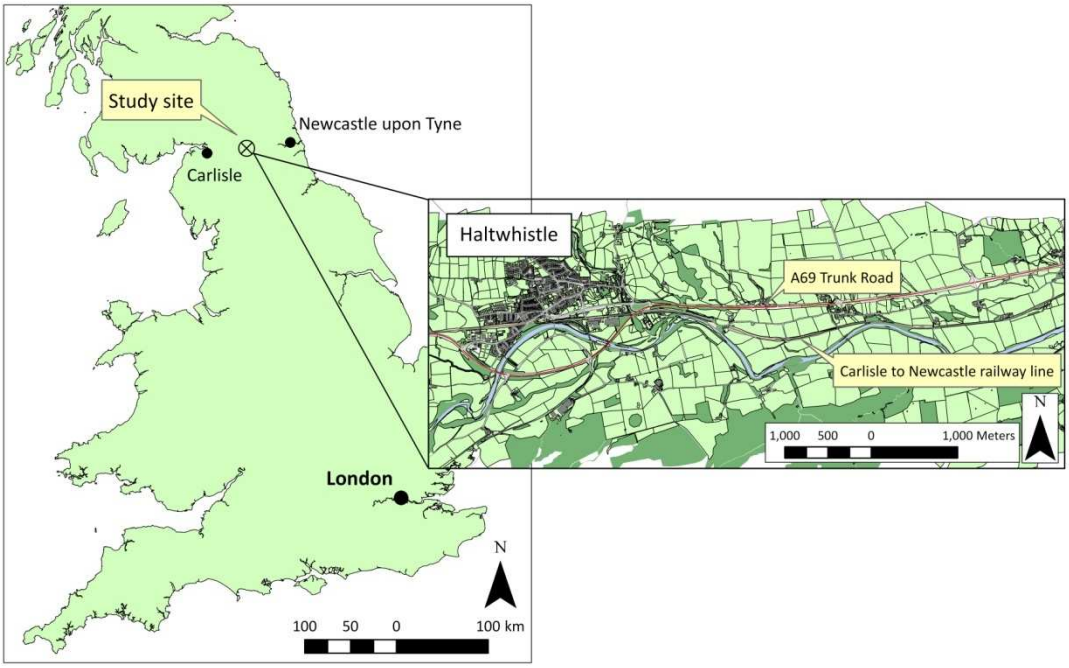


Figure 3.4 Map showing the location of the Haltwhistle transport corridor. Map data © Crown Copyright 2010. An Ordnance Survey EDINA supplied service.

Figure 3.5 Part of a cutting earthwork approximately 2.5 km east of Haltwhistle. Here, tall shrubs dominate as well as large areas of bare earth. Much of this earthwork is riddled with rabbit holes presenting an additional problem to the network operator. Note the leaning fence posts indicating slope movement.

Figure 3.6 A stretch of the Carlisle to Newcastle railway line approximately two kilometres east of Haltwhistle. This example indicates the heterogeneous nature of the vegetation cover, including tall trees, shrubs and wildflowers.

3.2 Earthwork scale

3.2.1 Terrestrial laser scanning data acquisition and processing

Terrestrial laser scanning data was acquired using a Leica HDS2500, which has a range of 50 m to 100 m, and a precision of ± 6 mm. This scanner has a fixed head and must be orientated using a resection approach using specially adapted control targets, the position of which were observed using a Leica Total Station. To ensure sufficient coverage of the earthwork, seven scan positions were set up towards the south of the embankment and five towards the north of the embankment (Figure 3.7). Data collection was carried out by a summer scholarship student but unfortunately the data from one of the scans was missing, which is identified in Figure 3.7. The raw point clouds were processed using Leica's *Cyclone* software which allows for data acquisition, registration and basic editing. A vegetation removal algorithm was then applied in TerraScan following the concept provided by Haugerud and Harding (2001).

The point data was then exported to ArcGIS where an ordinary kriging routine was used to interpolate between points to generate a raster grid surface. Kriging is a technique synonymous with geostatistics and is used to make predictions at unobserved locations (Trauth, 2007). In this case, elevation values are predicted between the point observations made by the laser scanner. These predictions are based on the notion that the elevation values are spatially correlated, i.e. closer observations are more similar than those far away (Trauth, 2007). Like other interpolation methods, including inversed distance weighted (IDW) (discussed later in Section 3.3.1), spatial correlation is summarised through locally weighted averaging. Kriging differs in this respect as the weights used are a measure of the spatial variance between neighbouring samples which provides an optimal interpolation with minimal variance (Desmet, 1997). Spatial variance is characterised in the kriging routine using a semi-variogram, this being half the squared difference between the observations plotted against the distance separating them (Trauth, 2007). Using this plot the spatial variance can be defined as the distance, or range, at which the semi-variance levels out, known as the sill (Figure 3.8). For detailed descriptions of the calculations made in the kriging routine refer to Trauth (2007).

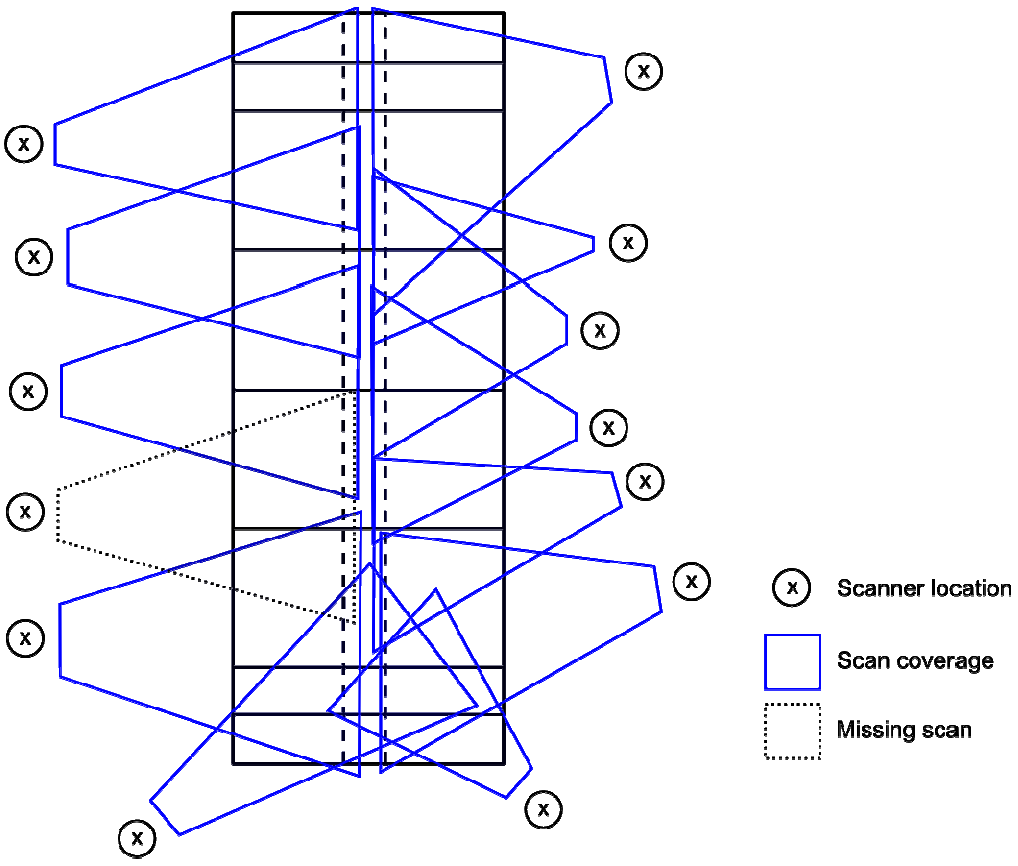


Figure 3.7 Diagram of the BIONICS test embankment with location of terrestrial laser scanner and subsequent coverage. Figure also includes the location of a missing scan.

Kriging interpolation was preferred to deterministic approaches, such as inverse distance weighted, as it is based on the statistical relationship between the measured points and as a result tends to produce a better representation of a terrain surface (Rees, 2000). Ordinary kriging using a spherical model was used. Ordinary kriging was preferred over universal kriging as it does not vary the prediction of spatial variance in the data (summarised by the semi-variogram), unlike universal kriging for which spatial variance is assumed to vary throughout the dataset which is more suited to the representation of heterogeneous terrain (Rees, 2000). A search radius of 12 points was used in this interpolation to ensure small scale variation is taken into account without becoming too computationally intensive (a smaller radius may have taken into account more detailed variation but would take a long time to compute). The radius was

determined by number of points rather than distance as the density of the elevation points provided by TLS varied over space. A spherical model was used to fit the semi-variogram. This was chosen by analysing the output from a number of different model types, including Gaussian and exponential. This type of model was chosen as it seemed to offer a better fit than other model types, such as exponential, although there was little visual difference between the Gaussian and spherical models. The ordinary kriging routine produced a DEM with a cell size of 35 cm which was the default set by ArcGIS, based on the length of the input extent divided by 250.

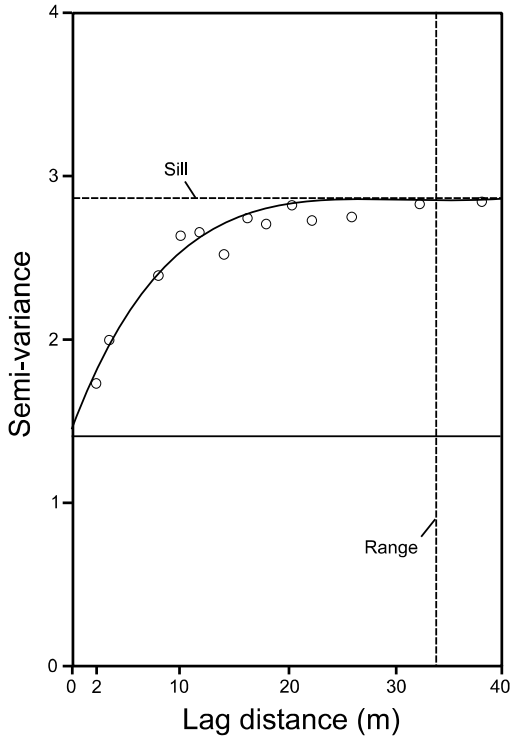


Figure 3.8 Example of a semi-variogram with labelled sill and range. The sill is the point at which the semi variance of the observations levels out, indicated by the solid interpolation line through the points. The range refers to the lag, or distance at which the sill occurs, which in this case is approximately 33 m. Adapted from Curran (1988).

As well as using the kriging routine, a thin plate spline interpolation method called AnuDEM interpolation (Hutchinson, 2009) was also evaluated. This technique is designed to construct a smooth surface from irregularly spaced elevation points. This is

done by applying a spatially variable smoothing function, or roughness penalty, which is dependent on a measure of roughness based on the input elevation values. See Hutchinson (2008) and (1989) for a full description of the calculations made in the AnuDEM routine. As well as removing spurious pits, this interpolation method is favoured in hydrological studies as it allows for the calculation of continuous flow through a given terrain (Hutchinson, 2008; Bishop and McBratney, 2002). Furthermore, the relatively smooth surface constructed using the AnuDEM routine reduces the tendency to produce spurious surface oscillations in complex areas (Hutchinson, 2008). The kriging and thin plate spline techniques were compared as the AnuDEM routine would not normally be employed for creating a DEM, with kriging being used to represent a standard interpolation technique.

One of the aims of the earthwork scale study is to explore the sensitivity of terrain analysis calculations to changes in the input DEM resolution. To do this, the original grid DEM was degraded to 1 m and 2 m using bilinear interpolation. This approach interpolates the new output grid cell value based upon an average, weighted by distance, of the four nearest input cells. Other techniques are available, including nearest neighbour, cubic convolution and majority resampling, but bilinear interpolation was used to be consistent with other similar studies (Chaubey et al., 2005). The resolution is degraded twice in order to characterise the effect of DEM grid resolution on the terrain analysis calculations described in Sections 3.2.1.1 and 3.2.1.2. Simple visual interpretation is used to assess whether changes in the DEM resolution produces artefacts in the calculations. The effect of resolution on reproducing surface roughness is assessed by comparing interpolated elevation values taken from a transect across the earthwork. The differences between these values are used to provide a quantitative description of the effects of DEM resolution. Following Schmidt and Persson (2003), the quality of the resulting DEMs are assessed by calculating root mean square error with the original point elevation values derived from the laser scanning data. This analysis is also carried out to quantify the difference between the two DEM interpolation techniques.

3.2.1.1 Topographic wetness index

The TauDEM plug-in software, developed by Tarboton (2002) for ArcGIS, was used to calculate TWI. TauDEM generates a wetness index through a ratio of the predefined raster grids, slope and specific catchment area. This software is favoured as it accommodates the Dinf flow direction routine which means that flow can be directed in an infinite number of single directions. This routine apportions flow between two downslope pixels according to how close this flow direction is to the direct angle to the downslope pixel. This has advantages over other flow direction procedures, such as the D8 routine, where flow is restricted to one of eight cardinal directions, potentially leading to grid bias and unrealistic flow dispersion (Tarboton, 1997). Other, more complicated algorithms, such as the DEMON (digital elevation model network extraction) method, were also available that demonstrate a more realistic model of flow dispersion (Wilson and Gallant, 2000). However, this technique suffers from being complicated to compute and can often create artefacts in hydrologically complex areas, for example, saddle shaped terrain (Tarboton, 1997). This is because such techniques specify flow direction over a locally fit plane per pixel according to the lowest corner estimated by averaging the elevation values for the adjoining pixels. Problems arise where certain combinations of elevation values can result in inconsistent or counter intuitive flow directions (Tarboton, 1997). Take for example the scenario in Figure 3.9 whereby flow from cells A and B are forced against each other, whereas the more likely flow direction would be the lowest grid elevation values located towards the north and south of cell A.

The first step in calculating a TWI involves filling all pits that are present in the DEM by raising their elevation to the level of the lowest point around the edge. The reason for this is that pit grid cells are completely surrounded by higher terrain and tend to be artefacts and if not removed can interrupt the flow of hydraulic conductivity throughout the scene (Tarboton, 2000). Flow direction and contributing area are then calculated from the DEM with filled pits using the Dinf procedure. The following describes these calculations in greater detail, based on Tarboton's paper (1997) which outlines the method.

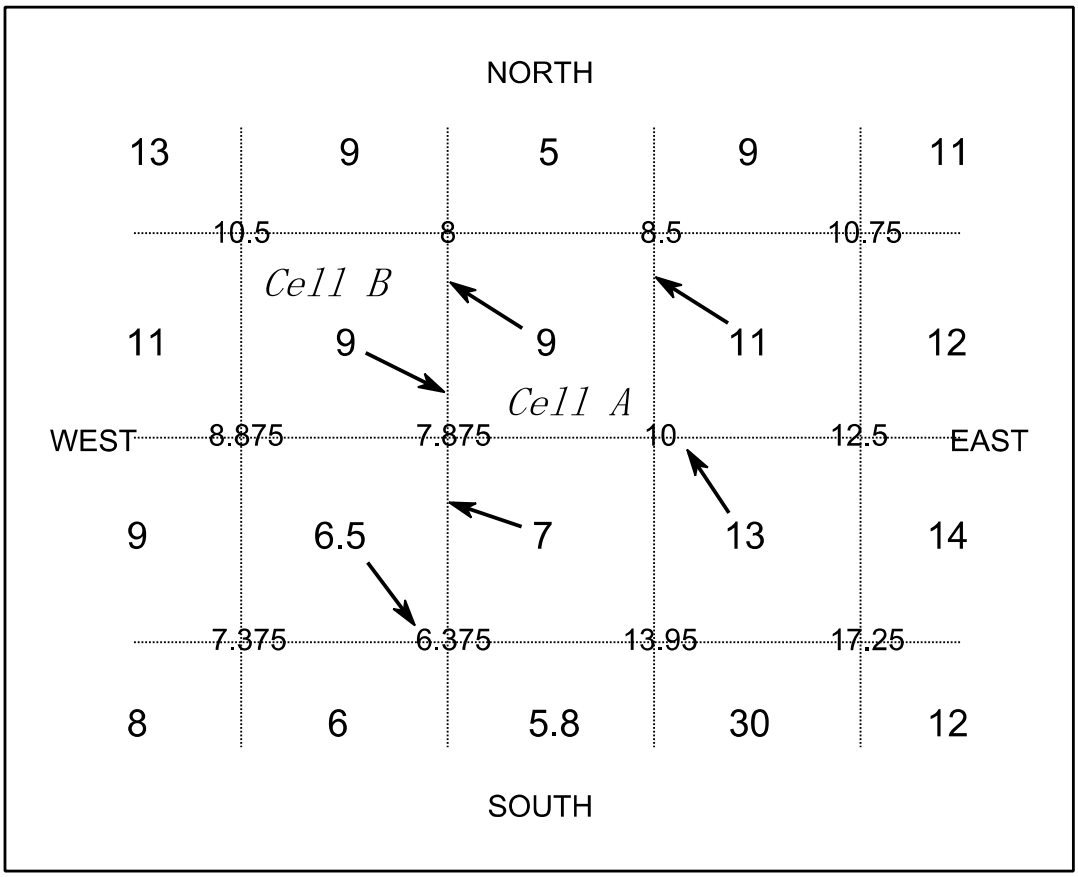


Figure 3.9 Hypothetical subset of a DEM demonstrating flow dispersal determined using locally fitted planes, as used in algorithms such as DEMON. Elevation values are given in the centre of each grid square. Corner elevations, denoting the locally fitted plane, are given as smaller numbers in between the grid squares. Arrows show flow direction as determined using DEMON, where the expected flow direction (from grid square A) would be towards the north and south where elevation values are lowest. Adapted from Tarboton (1997).

A single flow direction is assigned to each cell, which is represented as a continuous value between 0 and 2 π. This direction is defined as the steepest downward slope over eight triangular facets on a 3 x 3 pixel window centred on the cell of interest (Figure 3.10). A downslope vector is calculated for each triangular facet which can lie within or outside a 45° angle range of the centre point of the facet. The flow direction for the central pixel is taken as the direction of the steepest downslope vector from all eight facets.

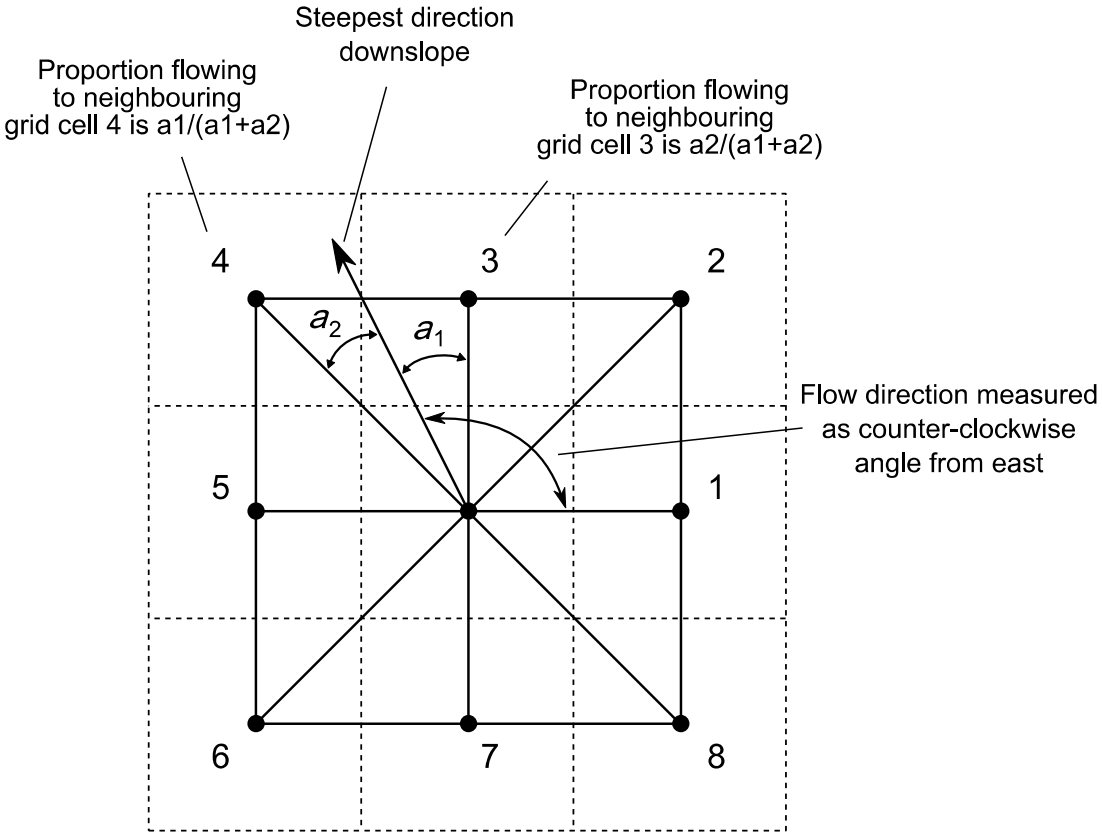


Figure 3.10 Diagram of a 3 x 3 pixel window on which flow direction is calculated using eight triangular facets. Outline of the actual grid cells are represented by the dotted lines.

To illustrate this procedure Tarboton (1997) gives an example using a singular triangular facet (Figure 3.11). The downwards slope is represented by the vector S_1, S_2 as described in Equations 3.1 and 3.2.

$$S_1 = (e_0 - e_1)/d_1$$

Equation 3.1

$$S_2 = (e_1 - e_2)/d_2$$

Equation 3.2

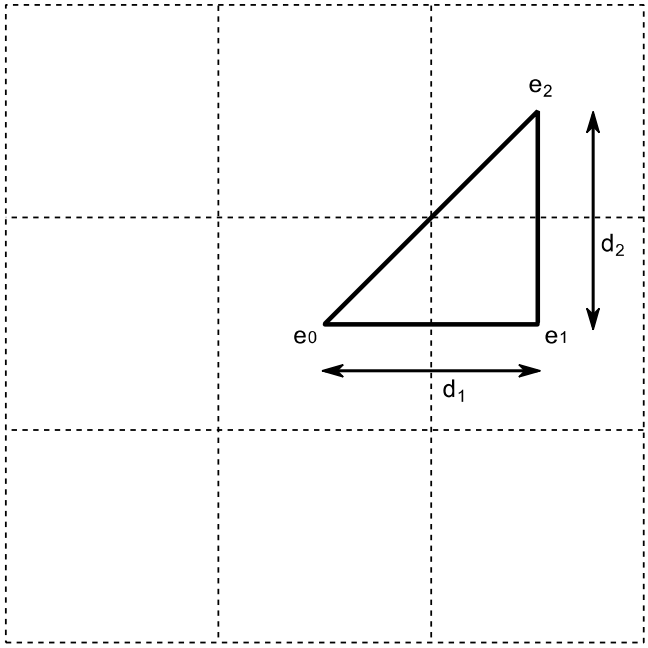


Figure 3.11 Definition of variables for the calculation of slope direction for a single triangular facet.

where, $e_{0,}$ e_1 and e_2 are elevations, and d_1 and d_1 are distances between the pixels as denoted in Figure 3.11. The values produced in Equations 3.1 and 3.2 can then be used to calculate slope direction (r) and magnitude (s) using Equations 3.3 and 3.4 respectively.

$$r = \tan^{-1}(s_2/s_1)$$

Equation 3.3

$$s = \sqrt{{s_1}^2 + {s_2}^2}$$

Equation 3.4

The downslope direction can be applied to each of the eight facets by using an appropriate transformation for each of the corner elevations shown in Figure 3.10. Relative transformations for each facet are shown in Table 3.1 and correspond to the elevation values used in Equations 3.1 to 3.4. To obtain the overall flow direction angle (r_g), the angle of the largest downward slope is adjusted to represent the angle counter-clockwise from east using Equation 3.5 (Tarboton, 1997).

$$r_g = a_f r' + a_c \pi/2$$

Equation 3.5

where, r' is the slope direction (from Equation 3.3) of the triangular facet with the largest slope magnitude (from Equation 3.4) and the multiplier a_f and constant a_c depend on the facet of interest, as listed in Table 3.1.

Facet	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$\mathbf{e_0}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$	$e_{i,j}$
$\mathbf{e_1}$	$e_{i,j+1}$	$e_{i-1,j}$	$e_{i-1,j}$	$e_{i,j-1}$	$e_{i,j-1}$	$e_{i+1,j}$	$e_{i+1,j}$	$e_{i,j+1}$
$\mathbf{e_2}$	$e_{i-1,j+1}$	$e_{i-1,j+1}$	$e_{i-1,j-1}$	$e_{i-1,j-1}$	$e_{i+1,j-1}$	$e_{i+1,j-1}$	$e_{i+1,j+1}$	$e_{i+1,j+1}$
$\mathbf{a_c}$	0	1	1	2	2	3	3	4
$\mathbf{a_f}$	1	-1	1	-1	1	-1	1	-1

Table 3.1 Table showing the transformation factors for elevation values for each facet used in the calculation of slope and angle.

An example of the calculation of flow direction using TauDEM is given in Figure 3.12. The same elevation values are used as in the demonstration of the DEMON algorithm over a saddle shaped terrain in Figure 3.9. Unlike the counter intuitive flow directions predicted by DEMON, TauDEM directs flow towards the north and south of cell A as would be expected in reality. This demonstrates the effectiveness of this method for representing flow over complex terrain features. Over simpler terrains, there has shown to be little difference between the methods, although the DEMON model is said to be difficult to program and often requires site specific parameterisation (Tarboton, 1997).

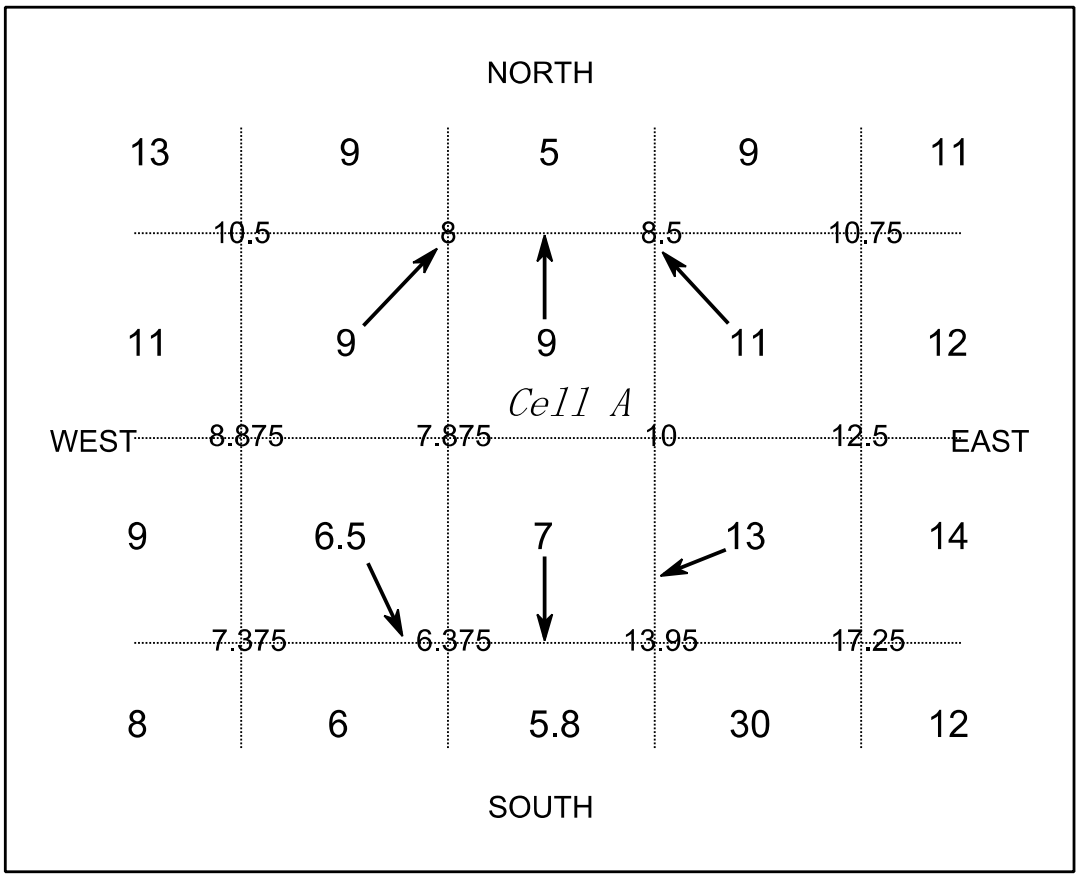


Figure 3.12 Hypothetical subset of a DEM as used in Figure 3.9. Here, it is used to demonstrate flow dispersal using the TauDEM Dinf method as proposed by Tarboton (1997). Elevation values are given in the centre of each grid square. Corner elevations, denoting the locally fitted plane, are given as smaller numbers in between the grid squares. Arrows show flow direction generally in a north and south direction where elevation values are lowest. Adapted from Tarboton (1997).

Specific catchment area is defined as its own area (one pixel), plus the area of the upslope neighbours that have some proportion flowing into the pixel of interest. If the downslope angle lies between two cardinal directions then a proportion of the flow is assigned according to how close the flow direction angle is to the direct angle of the neighbouring cells. This function is described in the two equations given in Figure 3.10 and is demonstrated in Figure 3.13. The direction of steepest slope has been calculated using the Dinf function as is displayed as an arrow between the centres of cells three and four. Let the angle from this arrow to the centre of cell three (angle a_1) be 30° and

the angle to cell four (angle a_2) be 15°, therefore equalling a 45° angle between the directions of cells three and four. Using the equations in Figure 3.13, the proportion of flow to cell three is 0.33 (15°/(30°+15°)) and the proportion flowing to cell four is 0.66 (30°/(30°+15°)). In this way, more flow is proportioned to cell four as this is closer to steepest downslope direction.

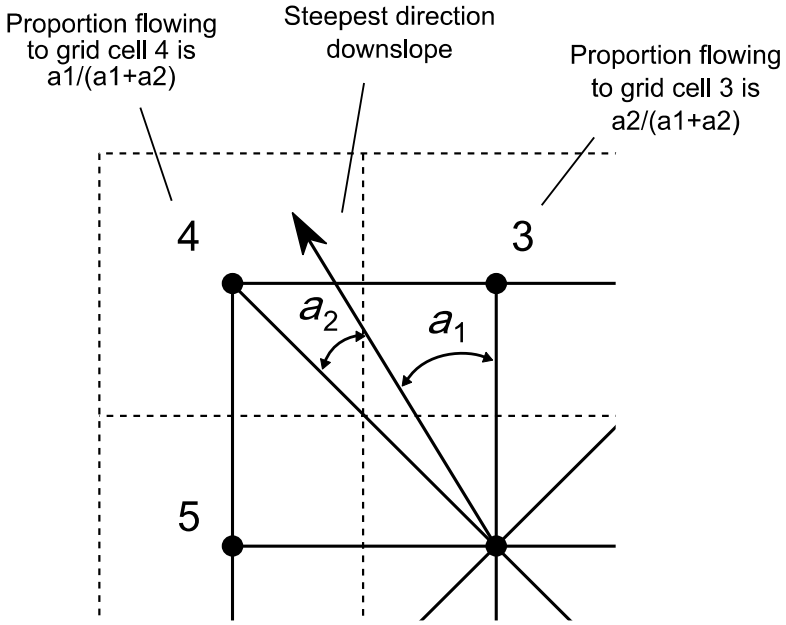


Figure 3.13 Diagram showing how flow is proportioned between two downslope grid cells based on their proximity to the angle of the steepest downslope direction. Adapted from Tarboton (1997).

The wetness index is finally generated through a ratio between the calculation of slope and specific catchment area. This is related to the original wetness index proposed by Beven and Kirkby (1979) (see Section 2.3.1), where catchment area is divided by slope, but in TauDEM slope is divided by catchment area (Equation 3.6). This is done to avoid errors where slope is zero, otherwise, the resulting wetness index value for flat areas would be zero (Tarboton, 2002).

$$\text{TauDEM TWI} = \frac{\text{Dinf Slope grid}}{\text{Dinf Specific Catchment Area grid}}$$

Equation 3.6

3.2.1.2 Potential solar radiation calculation

The Area Solar Radiation function in ArcGIS, developed by Fu and Rich (1999) was used to calculate potential solar radiation. As explained in Section 2.3.2, total radiation for a given DEM grid square is calculated by overlaying a representation of the viewshed with the sunmap and skymap. The concept behind these representations is detailed in Section 2.3.2 using Figures 2.8 to 2.10. The specific calculations that define the Area Solar Radiation function are given below.

Total (or global) radiation is calculated for each of the sunmap and skymap sectors using Equation 3.7 (ESRI, 2007b).

$$\text{Global}_{\text{tot}} = \text{Dir}_{\text{tot}} + \text{Dif}_{\text{tot}}$$

Equation 3.7

Where, $\text{Global}_{\text{tot}}$ is the total radiation, Dir_{tot} is direct radiation, and Dif_{tot} is diffuse radiation. Dir_{tot} is defined as the sum of the direct radiation for all the sunmap sectors. Direct radiation is defined by the solar constant, or solar flux outside the atmosphere at the average Earth to Sun distance. This is taken to be 1367 W m^{-2} which is derived from the World Radiation Centre (ESRI, 2007b). Diffuse radiation is estimated using a predefined constant relating to whether the dominant cloud conditions are overcast or clear. For this study, the conditions were set as overcast as field experience has shown this to be the dominant cloud conditions over the test embankment. As a sensitivity test, the calculation of solar radiation was carried out using different conditions but no visible difference was found.

The method for calculating potential solar radiation, as developed by Fu and Rich (2002; 1999), is used in this study as a black box. This method is used because it is more comprehensive than other techniques, namely the potential solar radiation index used by Western et al. (1999a), as it takes into account factors such as time of year, latitude of the site and subsequent sun angle. In an attempt to mitigate for the potential pit falls of using black box techniques, tests were carried out to assess the sensitivity of the model by changing the input parameters such as the period for which the calculation is made and cloud conditions.

The Area Solar Radiation function requires the input of the mean latitude for the input DEM. This was calculated by ArcGIS as 54.9˚. The function was run twice for

each interpolated DEM, with the time configuration set at two different periods, one calculating solar radiation for a whole year (2008), and the other for the specific day in which the ground truth soil moisture measurements were taken (24th November 2008). This was done to test the ability of the two configurations to characterise soil moisture gathered on a specific day.

The calculation of aspect was also carried out. This metric is a common input to previous terrain analysis studies, such as Western et al. (1999a), and is used as a more simplified quantification of potential solar radiation. Aspect was calculated for the kriging and AnuDEM interpolated DEMs using ArcGIS. This quantifies aspect for each grid cell in degrees from north by identifying the slope direction of the maximum difference in elevation value from a 3 x 3 cell neighbourhood (ESRI, 2007a).

3.2.1.3 Soil moisture measurements

In situ soil moisture was collected using a Theta portable impedance probe (www.delta-t.co.uk). This probe consists of four stainless steel rods which are inserted into the soil. The rods then use a simplified voltage standing wave method to measure relative impedance, which is used to infer the dielectric constant of the soil relating directly to the volumetric water content of the soil (Famiglietti et al., 1999). Following Cosh et al. (2005) the probe was calibrated using gravimetric soil moisture measurements taken from soil samples located across the embankment (Section 4.1.7). Calibrations made in this way have shown to yield better results than generalised calibrations made with the same soil type from a different area (Cosh et al., 2005).

Measurements were made across half the embankment, covering a well constructed and poorly constructed panel on both sides of the earthwork. Only half the earthwork could be studied due to the construction of covers, used for an ongoing climate experiment towards one end of the embankment (see Hughes et al. (2009)). Measurements were restricted to the earthwork slopes as the top of the embankment is flat and is not liable to the same forces of instability. For the terrain analysis experiment, measurements were taken approximately every two metres based on a grid sampling strategy. This scheme was adopted as authors have identified tenfold improvements to precision when measuring soil properties on a sampling grid rather than randomly selected points (McBratney and Webster, 1983). This is because

systematic sampling ensures any spatial variability in the data will be taken into account (McBratney and Webster, 1983), a characteristic synonymous with the distribution of soil properties (Famiglietti et al., 1999; Famiglietti et al., 1998; Western et al., 1998; McBratney and Webster, 1983). The position of each point measurement was recorded using real time kinematic (RTK) GPS, with the ability to provide centimetre level positional accuracy (Montillet et al., 2007).

Ideally, the ground truth data would have been gathered periodically throughout the year to test the sensitivity of the terrain indices to characterise soil moisture during different wetness conditions. However, this could not be achieved in the time frame for the study at hand, as focus was diverted to the transport corridor scale experiments detailed later in Section 3.3. This means that comparisons made with the subsequent soil moisture measurements must not be assumed to be a representative of conditions throughout the year. The measurements were made on 24th November 2008 during cold, wet and windy conditions. This time of year was chosen as other studies have reported increased topographic organisation of soil moisture to occur during overall wetter conditions and may therefore improve correlations with the terrain analysis calculations (Tenenbaum et al., 2006; Schmidt and Persson, 2003; Western et al., 1999a).

3.2.1.4 Analysis techniques

Elevation values and the terrain analysis metrics were compared for the kriging and AnuDEM interpolated DEMs, and the subsequent spatially degraded resolution DEMs. Initially, qualitative comparisons were made through visual analysis. Quantitative analysis was then made by extracting values from a series of representative transects across the earthwork. The elevation values taken from these transects were also compared to the original TLS point elevation data, providing a measure of residual error. This was done to determine which interpolation technique best reproduces the original TLS elevation values. This was not, however, a measure of the overall accuracy of the interpolation techniques. Such analysis could be undertaken by comparing against additional observations of elevation (Chang et al., 2004). This was not undertaken as it was not the overall focus of the study at hand.

To facilitate a comparison between the soil moisture measurements and the terrain analyses, the raster outputs from the TWI and potential solar radiation

calculations were converted to point shapefile datasets. These point datasets were then joined to the soil moisture point dataset based on spatial location. As the points do not overlap exactly, a new value was assigned as an average of the closest points to that particular location using bilinear interpolation. Point measurements were compared rather than raster grids following the example of Schmidt and Persson (2003) and Tenenbaum et al. (2006).

Authors have drawn comparisons between wetness indices and soil moisture measurements using ordinary least squares correlation coefficients (R^2 values) (Tenenbaum et al., 2006; Schmidt and Persson, 2003; Western et al., 1999a). In addition, some authors have considered multiple regression models to combine terrain indices to predict soil moisture values (Western et al., 1999a). As a result, this investigation primarily used ordinary least squares regression correlation coefficients and multiple regression modelling to compare soil moisture measurements with the estimations made using the terrain analysis calculations. Multiple regression models were made using stepwise forward regression as this only enters variables into the prediction based on their significance, in this case where the p value is less that 0.05, representing the 95% confidence level.

The ordinary least squares (OLS) regression procedure, adopted in the studies cited above, is a global regression technique. This type of statistical analysis is vulnerable to a phenomena known as the Simpson's Paradox. This paradox occurs where the relationship between model parameters varies spatially. These local variations are ignored in global regression which might otherwise lead to a change, or even reverse, in trend between the observed and predicting variables. As a result, geographically weighted regression (GWR) was also considered in the analysis, which provides a spatial component to the ordinary regression procedure (Fotheringham et al., 2002). This is a local regression technique that exploits the widely regarded theory that points closer to one another are more alike than those far away (Fotheringham et al., 2002; Brunsdon et al., 1998) (see Appendix 2 for a description of GWR).

To facilitate a comparison between the soil moisture measurements and the terrain analysis calculations, a GWR model was defined with a fixed Gaussian kernel, as the spatial density of the observations does not change significantly. Bandwidth was

selected using AIC_c cross validation, producing an optimal bandwidth of 5.8 m. A Monte Carlo significance test was used to test for spatial variability of the metrics.

The analysis techniques discussed in this section have all assumed the relationships between the terrain calculations and observed soil moisture are linear. However, some authors have noted that the relationship between TWI and soil moisture is non linear (Western et al., 1999a). Despite this, there are few examples of studies attempting to deal with this relationship to improve overall soil moisture estimations. Western et al. (1999a) explained that a lowess (Locally Weighted Scatterplot Smoother) regression procedure significantly improves overall correlations. Although this may be the case, the use of such locally fitting curves introduces a large number of model parameters, which can reduce the stability of the model for making predictions outside the model training area. Other non linear solutions include cubic or quadratic fits, but again this would introduce additional parameters into the model. Expressing a metric as the natural logarithm may also provide a solution for linearising the relationship without introducing any additional parameters as with the cubic and quadratic solutions. Despite the simplistic nature of this method there have been no reported attempts to improve correlations by expressing the TWI by the natural logarithm. This could have the potential to improve estimations, particularly over uniform transport earthwork slopes, as moisture would tend to concentrate towards the toe of the slope, rather than gradually increasing downslope.

The reason why previous studies have not explored the natural logarithm of the TWI for use in soil moisture estimation models is that the original equation, put forward by Beven and Kirkby (1979), uses this function to express their index (as described in Equation 2.1). However, the wetness index used in the TauDEM tool does not refer at any time to the use of the natural logarithm in expressing any of the input calculations, specific catchment area and slope grid (see Section 3.2.1.2 or Tarboton (2002)). Therefore, this study will explore the use of the natural logarithm to improve the relationship between the wetness index in TauDEM and observed soil moisture.

3.2.2 Hyperspectral reflectance data acquisition and processing

3.2.2.1 Experiment design and sampling strategy

This part of the earthwork scale experiment aimed to investigate the response of embankment grasses to artificial wetting through hyperspectral analysis techniques. Three plots were set up on the north facing uncompacted panel of the test embankment (see Figure 3.2 for a diagram of the test embankment). The uncompacted panel was chosen as it was designed to mimic the characteristics an earthwork typically found in the UK (see Section 3.1). The plots were located on the north facing slope as construction of a climate simulator was taking place on the opposite south facing slope. One plot was wetted each day for a month with 20 litres of water, one was covered to stop water infiltration from rainfall, and the other was left as a control plot. The amount of water used to wet the wetted plot was chosen as this was the amount initially needed to fully saturate the soil. This was done to enforce soil oxygen displacement in an attempt to induce plant stress, similar to the methods of Smith et al. (2004a). The plots were placed at least 5 metres apart in an attempt to limit movement of water from one plot to the next.

Each plot measured 1 x 1 m, which was chosen as it is similar to the pixel size of the aerial imagery collected for the transport corridor scale analysis (see Section 3.3.2). This meant that any sub pixel variations in soil moisture, or vegetation reflectance, could be assessed. Each plot was subdivided into a grid of twenty five 20 x 20 cm subplots. Measurements of soil moisture and spectral reflectance were taken within each of these subplots to ensure the whole square was sampled. Grid sampling was chosen over random sampling to aid repeatability, but more importantly, as McBratney and Webster (1983) explained, systematic grid sampling can almost always improve the precision of sampling soil properties, as they have an inherent spatial component which random sampling does not take into account. *In situ* soil moisture contents were collected using a Theta portable impedance probe, as described in Section 3.2.1.3 An average of three measurements were taken at each point to take into account any variation in the sensor response.

3.2.2.2 Spectral reflectance measurements

Spectral reflectance was measured using an ASD Field Spec Pro® spectroradiometer, which has a spectral range from 350 to 2500 nm covering the visible, near infrared and shortwave infrared. Measurements were taken from a height of approximately 1.5 m using an 8° foreoptic producing a field of view of approximately 20 cm^2. Analysis of the spectra was restricted to the 350 nm to 1000 nm domain, as this is the spectral range that a number of potentially operational airborne sensors cover, including the sensor described later in Section 3.2.2.4. Multiples of five spectra were gathered and averaged for each measurement point to suppress any atmospheric changes and to take into account small scale variations at the sampling point. When recording these multiple readings, care was taken to rotate the probe to take into account any blind spots in the field of view caused by broken fibre optic cables (FSF, 2008). Reflectance measurements were made in raw digital number mode, which tends to improve the signal to noise ratio compared to measurements made in White Reference Mode (FSF, 2008). Corrections for absolute reflectance were made using coincidental reference measurement taken from a white Spectralon panel, which had previously been calibrated by the Natural Environment Research Council's Field Spectroscopy Facility at Edinburgh University.

3.2.2.3 Spectral processing techniques

A five point weighted mean moving average filter was applied to the data to suppress any noise from the instrument without losing any spectral detail (Section 2.4.1.1). This weighting was selected as it has previously been shown to provide sufficient smoothing without losing any spectral detail (Smith et al., 2005b). Continuum removal analysis was then performed to highlight the chlorophyll absorption feature. The line of continuum is defined by the minimum peak reflectance in the red (550 nm) and the peak reflectance in the near-infrared (750 nm) parts of the spectrum. The continuum line is then removed by dividing the filtered spectra by the value of the line at that position (Section 2.4.1.1). The derivative of the smoothed spectra was then calculated by dividing successive values by the wavelength interval separating them (Equation 2.2). This accentuates the maximum point of inflection which is indicative of the red-edge feature (Section 2.4.1.1).

The position of the red-edge was located on the derivative of the spectra using the Lagrangian technique which has been shown to be strongly correlated to chlorophyll content (Section 2.4.1.1) (Dawson and Curran, 1998). REP was also estimated using the linear interpolation technique (Clevers et al., 2004) which was applied to the original spectra, without any derivative calculation. Stress ratios were also applied to the data which exploit the twin peak feature at the maximum point of inflection of the first derivative of the spectra, which have been shown to represent vegetation stress, notably at 702 nm and 725 nm (Section 2.4.1.1) (Smith et al., 2004b). These spectral analysis techniques were compared to the soil moisture measurements using regression analysis following similar studies such as Smith et al. (2004b).

The methods discussed in this Section have mainly referred to techniques borne out of hyperspectral analysis, which utilise the fine spectral detail offered by specific sensors. However, attempts were also made to correlate soil moisture measurements using simple band ratios developed for broad band sensors, namely the Normalised Difference Vegetation Index (NDVI). Although there are a whole host of vegetation indices that can be applied in spectroscopy (see Bannari (1995) for a comprehensive review), the focus in this study is to exploit the fine spectral detail offered by hyperspectral sensors, whereas simple band ratios tend to take into account reflectance from just a small number of regions across the electromagnetic spectrum. Despite this, the NDVI was tested because, if successful, the characterisation of soil moisture using simple band ratios could be applied to imagery from a much less complex sensor. Such imagery tends to be available at a relatively low cost, compared to hyperspectral sensors and therefore increases the potential for becoming operational.

3.2.2.4 Sensor simulation

The CASI (Compact Airborne Spectrographic Imaging) airborne sensor provides hyperspectral imagery, over 32 bands covering the visible and near infrared regions of the spectrum. The sensor is flown by operators such as the Geomatics Group and has been used in numerous vegetation and marine monitoring surveys on behalf of the Environment Agency, demonstrating its potential for use in operational monitoring programmes. The field spectroscopy data gathered in this experiment was used to

simulate the CASI sensor in order to investigate the potential for the spectral analysis techniques to be scaled up to airborne imaging spectroscopy.

The CASI sensor was simulated using a normalised Gaussian point spread function relating to the spectral range of CASI sensor's bands, as denoted by the operator, Geomatics Group. This approach is adopted as the CASI sensor has shown to have a Gaussian spectral response because the grating within the pushbroom sensor head is curved (Milton and Choi, 2004). Milton and Choi (2004) explained that the response of hyperspectral sensors are often rounded due to the effects of diffraction, scattering in the optical system and the non-rectangular nature of the detector.

The CASI sensor is a programmable system making it possible to centre the bands over different wavelengths. For this study, the default bandset defined by the Geomatics Group for the analysis of vegetation was used (Table 3.2). This was chosen because the main operation for the sensor in this study was the analysis of vegetation reflectance. The ASD field spectroscopy measurements described in Section 3.2.2.2 were used to simulate CASI reflectance.

The simulation begins by calculating the standard deviation (σ) for each band centre (defined by central wavelength in nm using Equation 3.8 (Milton and Choi, 2004).

$$\sigma = \frac{FWHM}{2\sqrt{2\ln 2}}$$

Equation 3.8

where, the full width half maximum (*FWHM*) corresponds to the spectral resolution of the sensor at a particular bandwidth. This function is defined as the spectral interval at which the sensor's response equals one half its maximum value (Figure 3.14) which was provided by the Geomatics Group (Table 3.2). Although sensors receive reflectance beyond the FWHM, the signal is weak and unreliable, and therefore insignificant (Campbell, 2006). Using the calculation of σ, the wavelength values (at 1nm intervals to match the ASD spectrometer) surrounding the central wavelength of the CASI band are expressed as a Gaussian point spread function using Equation 3.9 (van der Meer et al., 2006; Castleman, 1996).

$$\text{Gaussian PSF} = EXP^{-\frac{(x_i - x_o)^2}{2\sigma^2}}$$

Equation 3.9

where, x_i is the wavelength of interest and x_0 is the position of the central wavelength which is assumed to be the peak of the Gaussian function (Mather, 2004). These values are then normalised by dividing by the total of the Gaussian point spread function values covering the range of CASI band. This provides a numerator per wavelength which can be multiplied by the corresponding reflectance value provided by the ASD spectroscopy data, simulated spectra per CASI band can then be obtained by summing these values. An example of how this method works using hypothetical data is provided in Appendix 3.

Band	Wavelength centre (nm)	FWHM (+/- nm)
001:	397.6	4
002:	405.3	3.7
003:	412.7	3.7
004:	420.1	3.7
005:	427.6	3.8
006:	435.4	4
007:	443.6	4.2
008:	452.3	4.5
009:	461.5	4.7
010:	471.2	5
011:	481.6	5.3
012:	492.5	5.7
013:	504.2	6
014:	516.5	6.4
015:	529.7	6.8
016:	543.6	7.2
017:	558.6	7.7
018:	574.5	8.3
019:	591.7	8.9
020:	610.1	9.6
021:	630.1	10.4
022:	651.6	11.2
023:	675	12.2
024:	700.3	13.2
025:	727.8	14.3
026:	757.5	15.5
027:	789.6	16.7
028:	824.3	18
029:	861.5	19.2
030:	901.2	20.5
031:	943.4	21.7
032:	988	22.9

Table 3.2 Default bandset defined by the Geomatics Group for vegetation analysis.

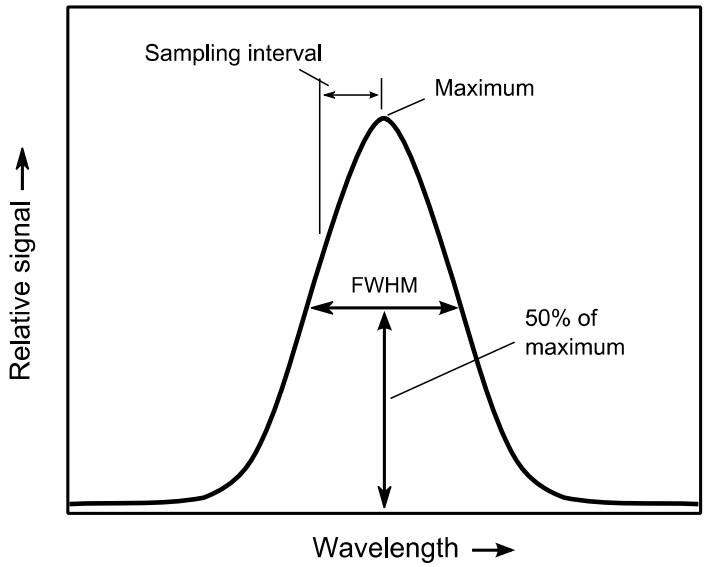


Figure 3.14 Definition of field width half maximum used in the CASI sensor simulation. From Campbell (2006).

3.2.2.5 Analysis of spectral techniques

The spectral analysis techniques were applied to the field spectroscopy data and the CASI simulated spectra. These were then compared to the soil moisture measurements using OLS regression analysis. The regression models were then applied to the reflectance data and compared to additional soil moisture measurements for validation.

The spectral analyses described in Section 3.2.2.3 refer to published techniques. In addition to these already established techniques, PLS regression was used to explore the potential for building a new model, specifically designed to predict soil moisture using vegetation reflectance. This type of regression was employed as it can cope with a large number of noisy predictors that are not truly independent (Carrascal et al., 2009), as is often the case with imaging spectroscopy data (Næs et al., 2002). PLS is similar to principal components analysis (PCA) whereby new variables, or components, are created which are linear combinations of the predicting variables that explain the total variance of the data. These new variables are uncorrelated which, therefore, solve the colinearity issue associated with OLS, which would likely produce a well fitted, but unstable model using collinear predictors (Næs et al., 2002). The drawback of PCA is that it requires the user to choose the number of components to be used in the final model. This means that the user must accept the risk that some of the components may

have little relevance for predicting the observed values (Næs et al., 2002). Although the use of loading plots can make this an intuitive process, it remains a subjective decision and does not carry much statistical rigour.

Whereas PCA creates components that best describe the predicting variables, PLS is based on explaining the variance in both the predicting and observed variables. It does this by maximising the covariance between the observed variable and all possible combinations of the predicting variables. The resulting components are therefore more directly related to the original observed values (Næs et al., 2002). Generally speaking, the regression coefficient will increase with the number of components used in the model, as a greater variation in the predictor and response variables are explained.

PLS differs from PCA regression by selecting the optimum number of components based on a leave one out cross validation, producing a predicted R^2 for each sub model with varying numbers of components. This correlation coefficient tends to be higher for models with just a few components as they tend to be more stable. However, there can also be cases where a model with more components has a much better fit at a sacrifice of just a marginally smaller cross validation R^2 score. For example, Figure 3.15 shows the model selection plot for a set of chemical data (this example was chosen as it provides a clearer demonstration than the model selection plots calculated in this study). Here, the optimum model, with two components, is selected based on the highest cross validation R^2 score. However, the user is more likely to choose a model with four components as this has a comparable cross validation correlation but with a much improved fit for the original data.

One of the criticisms of PLS regression is the lack of statistical rigour for testing the significance of the model parameters (Davies, 2001). Realising this, Martens (2000) developed an uncertainty test to complement the PLS routine. Their test is used to estimate the standard errors of the regression coefficient estimates in the PLS model. Similar to cross validation, the test uses the jack-knife principle where one (or more) sample is removed at a time to create a number of sub-models (see Martens and Martens (2000) for a full description of the technique). The regression coefficients for each sub-model are then computed giving information about the variability of the parameter values and an estimate of the standard error. The estimated standard error can then be used to calculate T-test values to test the significance of the variables in the model (Næs

et al., 2002). The uncertainty test, embedded in the Unscrambler® multivariate analysis software (CAMO, 2010), denotes insignificant predicting variables where the uncertainty limits go below zero.

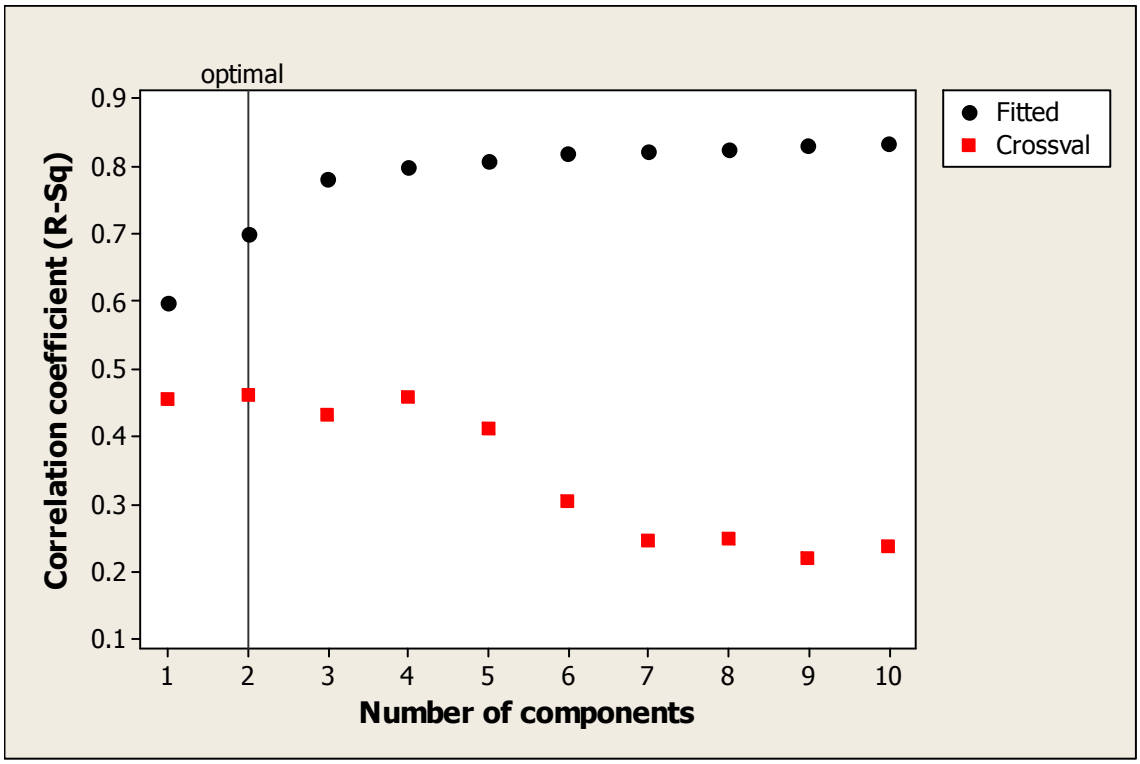


Figure 3.15 Example plot demonstrating the selection of optimum number of components for a PLS model using the correlation coefficient (R^2) scores for the original model and the model validated using leave one out cross validation. Data taken from MiniTab example data with elemental concentrations acting as the predicting variable and the response being wine aroma. Data from Frank and Kowalski (1984).

PLS regression was applied to the CASI simulated spectra which had been transformed using the expression in Equation 3.10.

$$\log_{10}(1/R)$$

Equation 3.10

where, R is relative reflectance. This transformation routine was born out of chemical spectroscopy as it is generally assumed that analysis procedures, such as PLS

regression, do not work well with nonlinear data. Equation 3.10 is held to be the best effort for linearising near infrared reflectance data (Reeves, 2009). In terms of spectral reflectance of vegetation, it has been said that there is a near linear relationship between transformed spectra (using Equation 3.10) and spectral absorption by chlorophyll pigments within a leaf (Kumar et al., 2006; Schmidtlein, 2005). As a result, this transformation is commonly used to prepare remotely sensed spectroscopy data for PLS analysis (Smith et al., 2003; Smith et al., 2002). However, it has also been suggested that such transformations not have a great influence on any subsequent spectral analysis (Reeves, 2009). Despite this, the transformation in Equation 3.10 was carried out to enable consistency with the remote sensing literature.

The transformation routine discussed above is carried out in spectroscopy to linearise the relationship between spectral reflectance and absorbing component (Kumar et al., 2006). In addition, such techniques reduce the effect of differences in sample thickness and light scattering (Næs et al., 2002). Another simple technique that achieves the same effect is the use of the derivative of the spectra (Næs et al., 2002), which is described in Section 2.4.1.1. The motivation behind this is that the derivative of the spectrum can minimise the effects of bi-directional reflectance caused by differences in the orientation of the sun, sensor, or target. Myneni et al. (1995) also demonstrated that derivatives of the spectra can be direct indicators of absorbing components in the leaves of dense canopies, making the technique an ideal candidate for use in the analysis of dense vegetation in transport corridors.

The PLS regression models built using the original CASI simulated spectra (without transformation), $Log_{10}(1/R)$ transformed spectra, and derivative of the spectra, were compared. This was done to judge which technique, if any, is best at preparing vegetation spectral data for PLS regression analysis. The motivation for this is that examples in the literature, including Schmidtlein (2005), Smith et al. (2003), Townsend et al. (2003), and Wolter et al. (2008), have not settled on one recommended transformation technique. The analysis was undertaken within the Unscrambler® software to take advantage of the Martens' uncertainty test. Using each set of spectra (original and transformed) as the predicting variables and the soil moisture measurements as the observations, PLS regression was initially run to determine which spectral bands were significant predictors. These bands were then selected for a second

PLS regression analysis, using the same observed soil moisture measurements, to determine the optimum number of components required for describing the variation in the data without becoming over fitted. For validation, this model was then used to predict additional soil moisture measurements.

3.3 Transport corridor scale

3.3.1 Lidar data acquisition and processing

Lidar data was collected from a helicopter-mounted Optech ALTM 2033 sensor in July 2007, producing a minimum point density of 20 points per m^2. The data were processed in TerraScan, where non-ground and vegetation points were classified and removed, following the concept detailed by Haugerud and Harding (2001). Two separate DEMs were generated from this data, the first using inverse distance weighted (IDW) interpolation and the second using the AnuDEM interpolation routine as described in Section 3.2.1.

These two interpolation techniques were compared in an attempt to repeat the methods carried out an earthwork scale, as detailed in Section 3.2.1, with IDW being used to represent a standard interpolation technique. This technique was used instead of the computationally intensive kriging interpolation used in Section 3.2.1 as it is relatively quick and simple to compute over the much larger area of the transport corridor. In contrast, kriging interpolation often stalled when applied to the transport corridor lidar data and could only be reliably used to produce coarse resolution DEMs (5 m). IDW interpolation estimates grid elevation from the lidar point elevation data by averaging values within a neighbourhood surrounding each grid cell. A weighting is applied that gives more influence to points closer to the centre of the cell of interest. IDW was applied with a power of two using a variable radius distance set to take an average 20 points. A power of two was selected as a compromise between higher powers, which give a higher weighting to points closer to the centre of a cell typically resulting in an increase in topographic detail, and lower powers that give more influence to points further from the centre, leading to a loss of topographic detail resulting in a smoother surface. This power is the default used in ArcGIS (ESRI, 2007a) and has been used in other studies, including Yue et al. (2007). The points were averaged using a variable radius distance as the density of lidar point data was not distributed evenly

across the scene. An average of the nearest 20 points was specified to match the minimum density of the lidar data. This was deemed reasonable as a 1 m grid resolution DEM was desired (based on the results of the test embankment work detailed later in Section 4.1.2) and, therefore, this ensured that there were at least 20 points per grid square to interpolate from.

Calculations of TWI and potential solar radiation were applied to the two DEMs in an attempt to characterise soil moisture distribution using the methods outlined in Sections 3.2.1.1 and 3.2.1.2. In addition to visual analysis, comparisons were drawn by comparing various profiles of elevation, TWI and potential solar radiation values across the earthworks. These values could then be used to quantify the differences between the two interpolated DEMs.

3.3.1.1 Soil moisture sampling strategy

Soil moisture measurements were gathered over the transport corridor using the Theta portable impedance probe (introduced in Section 3.2.1.3). The locations of these measurements were decided using cluster analysis of terrain analysis calculations. K-means clustering was used, despite the disadvantage of having to subjectively predefine the number of clusters used in the analysis. Although there are techniques, such as hierarchical clustering, that do not require a predefined number of clusters, they are computationally intensive (Wu et al., 2009) and proved to be problematic when applied to the extensive raster datasets used in this study. To overcome this, K-means clustering was repeated with an increasing number of clusters defined each time. The optimum number of clusters was determined through analysis of silhouette plots generated after each run. These plots provide a measure of how close each point in one cluster is to the points in neighbouring clusters which is defined by the squared Euclidean distance, +1 being very distinct from neighbouring clusters and - 1 indicating misassignment (Lletí et al., 2004). The overall success of the run can be determined by the average of these values (Lletí et al., 2004).

Analysis of the silhouette plots indicated two main clusters. These clusters were mapped which helped to identify one particular location, a cutting at Whitchester, that accommodated both clusters and was readily accessible for study. Three subsequent soil moisture campaigns were carried out in this area. The first occurred on 18th June 2009

following a period of little rainfall and involved a transect of readings across the profile of the cutting with measurements made every 2 m. This sampling strategy was chosen as it takes into account both downslope variation in soil moisture and the influence of the north/south slope aspects and can give an indication of small scale variations in soil moisture distribution without being too laborious. The second soil moisture campaign was carried on 9th July 2009 again following a period of dry weather. On this occasion measurements were taken approximately every 5 m to create a grid of measurements across the earthwork where access was possible. Some areas were inaccessible due to dense vegetation. this is a potential limitation as soil moisture contents are likely to be underrepresented under dense vegetation which is likely to influence soil moisture distribution through processes such as rainfall interception. The overall sampling strategy was chosen in order to maximise the spatial coverage of the soil moisture measurements following the guidance of McBratney and Webster (1983). The third campaign was carried out on 30th July 2009 following a period of heavy rainfall. The sampling strategy here was similar to that used on 18th June 2009 with a series of transects across the profile with 2 m spacing between measurement points. On each occasion the position of the measurements were recorded using real-time kinematic GPS.

3.3.1.2 Analysis techniques

The soil moisture measurements were compared to the terrain analysis calculations, for both the IDW and AnuDEM interpolated DEMs using linear regression. Stepwise multiple regression was then used to create a model predicting soil moisture from the terrain analysis metrics. Non linear transformation of the TWI metric was also considered (see Section 3.2.1.4 for further details). GWR was also used to consider the spatial variability in the relationship between soil moisture and the metrics. A GWR model was defined with a fixed Gaussian kernel, as the spatial density of the observations does not change significantly. Bandwidth was selected using AIC_c cross validation (see Appendix 2 for details on GWR). This produced an optimal band width of 18 m. A Monte Carlo significance test was used to test the spatial variability of the parameters used in the GWR model.

3.3.2 Airborne imagery acquisition and processing

High spectral resolution CASI imagery was collected over the Haltwhistle transport corridor in September 2007 onboard a fixed wing aircraft at a height of 1100 m giving a pixel resolution of 60 cm. Radiometric calibration of the imagery was undertaken by applying the empirical line method (Smith and Milton, 1999) to coincidental field spectroscopy measurements taken over three 6 x 6 m tarpaulin targets (black, grey and white) (Figure 3.16). The three targets were positioned at least twelve metres from each other to ensure any signal taken from the imagery was purely from one target at a time and had no influence from nearby targets.

Smith and Milton (1999) demonstrated that the empirical line method provides adequate results (errors in the region of a few percent) and is relatively easy to compute, compared to complicated radiative transfer models. However, they stress the importance of using at least two targets. Figure 3.17 shows an example of the radiometric calibration for band 1 using this method, illustrating a strong relationship (R^2 of 1) between the reflectance recorded using field spectroscopy and the radiance recorded by the airborne CASI sensor (see Appendix 4 for the quality of the radiometric correction for all 32 bands).

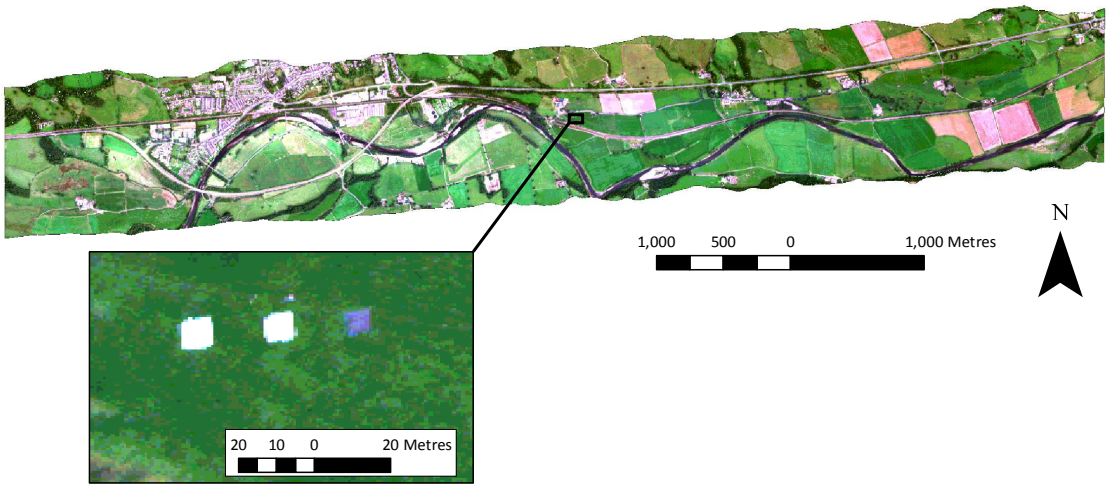


Figure 3.16 Location of the black, grey and white targets used in the empirical line radiometric calibration.

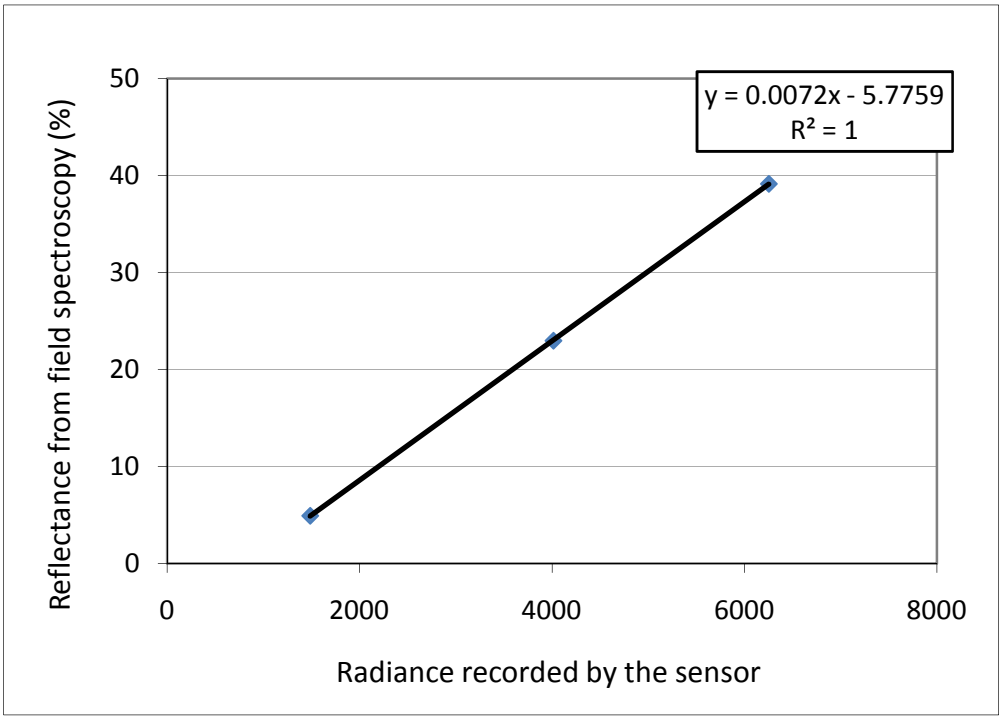


Figure 3.17 Example of the empirical line method for radiometric calibration of CASI band 1, with regression equation and correlation coefficient.

Geometric calibration was carried out by the operator, Geomatics Group (Environment Agency). Following the procedure outlined in Brown et al. (2003) the CASI imagery is synchronised with the onboard navigation system using a Precise Positioning Service (PPS) from the GPS. The CASI system is then geocorrected using navigational data using bespoke software from Itres (www.itres.com), which allows for geocorrection using a fixed elevation surface. The navigational system is integral to this process which consists of a dual frequency Novatel Millenium GPS receiver providing positional data which is post processed using ground GPS data referenced to the Ordnance Survey passive and active networks in the UK. A rudimentary quality assessment is made by comparing the imagery to map overlays from Ordnance Survey data. The system is calibrated each time the instrument is replaced in the aircraft, and periodical test flights are flown to ensure this calibration is still valid.

The quality assessment, referred to by Brown et al. (2003), was deemed insufficient for a precision engineering study due to the unknown accuracy of the Ordnance Survey data. Therefore geometric validation was carried out against independent ground control points collected in the field using RTK GPS. The maximum

RMSE quoted by the operator (1.5 pixels) was found to be exceeded in two of the CASI strips (RMSE of 1.7 and 2.4 pixels), which was reflected by poor mean residual errors and standard deviations quoted in Tables 3.3 and 3.4. This information was reported and as a result the operator made a rudimentary adjustment by applying a global shift to the data to bring it in line with the ground control points, reducing the aforementioned RMSE to 1.1 and 0.98 respectively. This is reflected in Figure 3.18 which shows the change in residuals for one CASI strip before and after correction, and Tables 3.4 and 3.5 which show changes in mean and standard deviation of the residuals for two CASI strips. The justification for the use of this method was that the residuals for each point tended to follow the same pattern. This was not a favourable solution and suggests perhaps this form of data collection and processing is not accurate enough for application in precision engineering. A more robust solution would be to collect coincidental GPS measurements of clearly visible targets and use them as control points to carry out geometric correction of the raw imagery. This was not done because the data ordered from the operator was expected to be within the quoted accuracy.

	Old		**New**	
	x	**y**	**x**	**y**
Flight line 1	-0.51	-2.09	-0.35	0.03
Flight line 2	-0.4	-1.41	-0.27	0.7

Table 3.3 Mean residuals between GCPs located using RTK GPS and identified in two CASI imagery flight lines before and after correction by the operator.

	Old		**New**	
	x	**y**	**x**	**y**
Flight line 1	0.54	0.82	0.7	0.66
Flight line 2	0.52	0.62	0.57	0.59

Table 3.4 Standard deviation of residuals between ground RTK GPS location and location identified in two CASI imagery flight lines before and after correction by the operator.

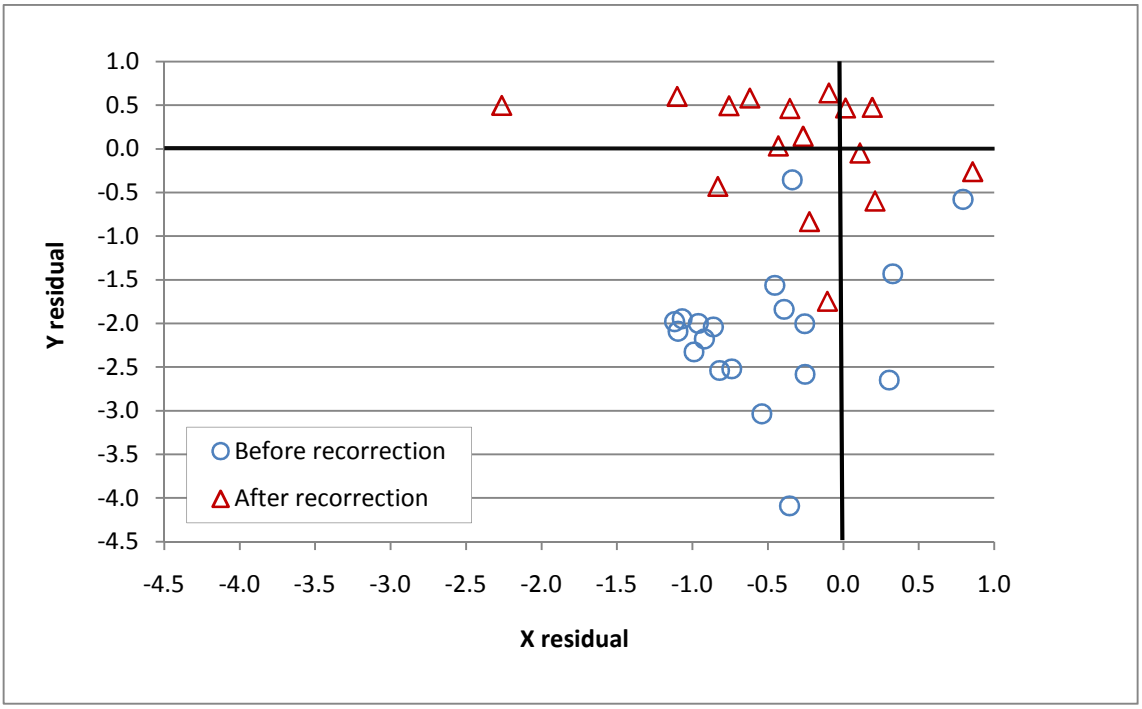


Figure 3.18 Plot showing change in residual error following correction by the operator for ground control points located using RTK GPS and identified in the CASI imagery.

Two strips covering the transport corridor earthworks and surrounding fields and structures were merged using ERDAS IMAGINE software. To prepare the imagery for subsequent analysis a number of steps were performed to reduce the amount of noise within the data and improve the overall signal from the vegetation. Firstly, a Minimum Noise Fraction (MNF) transformation was applied in the ITT ENVI software which can be used to identify spectral anomalies. The MNF routine begins with a forward rotation of the spectral data which orders the image data into different components of varying quality (van der Meer et al., 2006). It does this using a cascaded principal components transform whereby new components are selected as the signal to noise ratio decreases. In ENVI, the degree of noise per component can be visualised as a grey-scale image where the pixel value is considered to have a linear relationship with the level of noise. Typically, the images in first few MNF components are much more coherent than later component images that are often speckled, with little or no spatial coherence. By examining the images and associated eigenvectors the number of components with

coherency can be chosen. These components can then be selected as a spectral subset in an inverse MNF transform which places the MNF bands back in their original data space, which in this case corresponds to each of the 32 CASI bands. The spectral subset chosen for the Haltwhistle data were bands one to seven, which corresponds to an eigenvalue of more than five.

3.3.2.1 Mapping Ellenberg values

3.3.2.1.1 Vegetation sampling strategy

To date there have been no reported attempts to map Ellenberg values derived for UK plant species, such as the typology developed by Hill et al. (2000). Indeed, apart from Schmidtlein (2005), there have been few attempts to map Ellenberg values using imaging spectroscopy. One possible reason for this is the uncertainty surrounding the adequate number of vegetation species abundance observations that are required to characterise the range of Ellenberg values that exist at a given site. For example, Schmidtlein (2005) explained that variation in plant species composition and cover was assessed in 46 relevés (a uniform area used to make vegetation observations, such as a quadrat) positioned randomly throughout the study site. However, no reference is made as to how this sampling strategy was chosen.

One approach that may be suited to defining adequate sample size for measuring species abundance is the use of the nested plot technique (Mueller-Dombois and Ellenberg, 1974) (Figure 3.19). This technique involves counting the number of species that occur in the smallest sub plot, in the case of Figure 3.19, plot number one measuring 0.5 x 0.5 m. The number of species are then recorded in the subsequent sub plots, up to number nine, with an increasing sample area covered. A species-area curve can then be constructed where the number of species recorded is plotted against sample size. Mueller-Dombois and Ellenberg (1974) explained that adequate sample size can be defined as the point at which number of species recorded does not increase significantly with an increase in plot size (being defined as the point at which a 10% increase in plot size yields only 10% more species). Despite this, the species-area curve tends to be used in modern ecology for monitoring biodiversity, with little mention to the method being used for defining sample size.

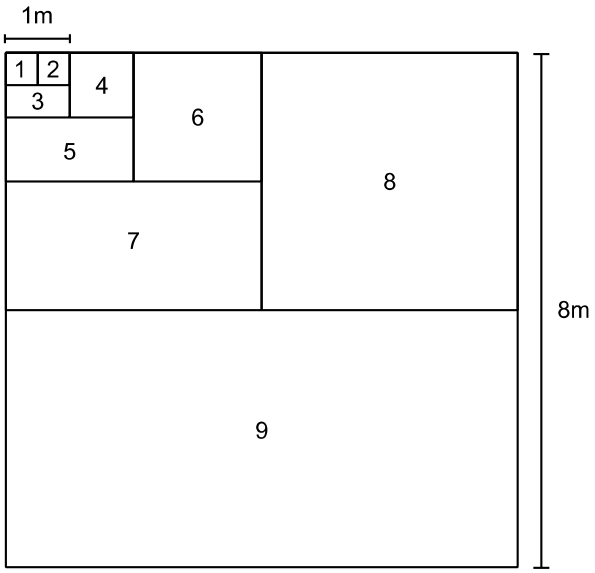


Figure 3.19 System of nested plots. Each subplot from 1 includes the area of the previous subplot, therefore all plots are square. From Mueller-Dombois and Ellenberg (1974).

Stohlgren et al. (1995) explained a number of limitations associated with the nested-plot technique. Firstly, the number of species recorded is added from each previous sub plot, starting from the initial 0.5 x 0.5 m square. As a result, each measure is not truly independent from the last, which can lead to autocorrelation and a strong bias around the first sub plot. Secondly, if a habitat is not strictly homogenous then the number of species measured per plot may be influenced by the shape of the area being examined. For example, circular or square plots will record fewer species as they have a reduced perimeter to surface area ratio, compared to a long, thin rectangle plots. In an attempt to overcome these issues, Stohlgren et al. (1995) suggested the use of a series of long thin plots that do not overlap. This type of sampling plot is known as a modified-Whittaker plot as it is based on the widely used Whittaker plot which first made use of long, elongated sub plots.

In an attempt to bring a degree of robustness to the vegetation survey, preliminary sampling was undertaken to construct a species-area curve for earthworks in the Haltwhistle transport corridor. This was done using the modified-Whittaker plot but adapted to fit in the confines of a cutting earthwork near Whitchester (Figure 3.20). Within each sub plot, the number of vegetation species was recorded (see Appendix 5

for details). This equated to ten small sub plots, two medium sub plots, one large sub plot, and the plot encompassing all the sub plots and the space in between. The average species number for each plot size was then calculated. This was plotted against relative sample area to produce a species-area curve (Figure 3.21). Scheiner (2003) explained that a power function best describes the species-area curve, which is supported in the present example with an R^2 of 0.95 compared to R^2 values of 0.66 and 0.92 for exponential and logarithmic fitted curves respectively. However, it could also be argued that the logarithmic fit is better suited as it reaches a plateau rather than the assumption made by a power curve that the total number of species would increase indefinitely with sample area.

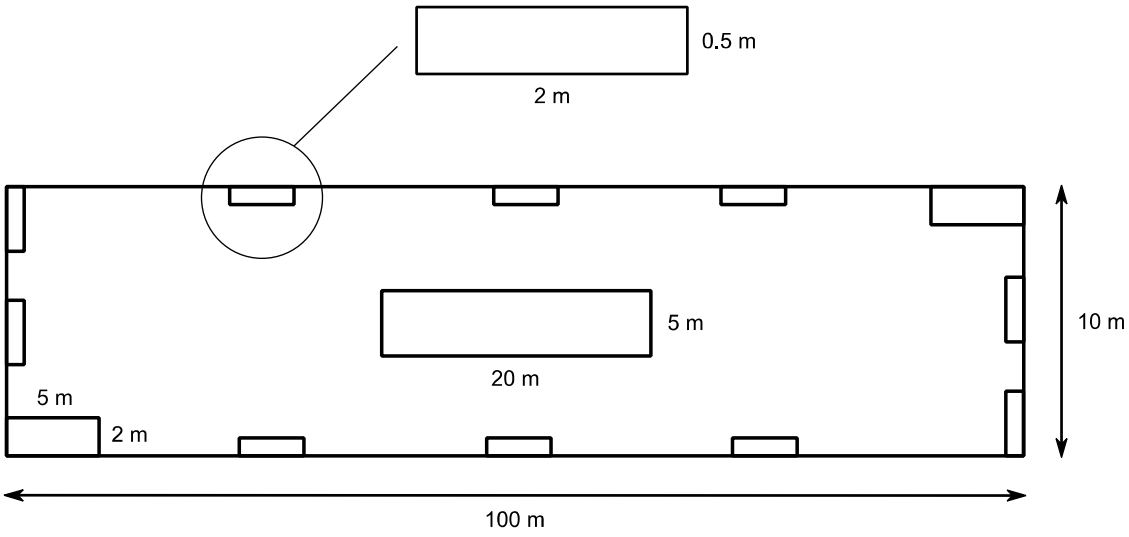


Figure 3.20 Diagram of the modified-Whittaker plot used to record species abundance over a cutting earthwork near Whitchester (not to scale) in order to create a species-area curve used to define adequate sample size. The rectangles refer to different sized sub plots for which species type is recorded. The rectangle encompassing all the sub plots is the total area. Further explanation is provided in Appendix 5.

The species area-curve in Figure 3.21 shows that the number of species does not increase significantly after a sample area of 100 m^2. Therefore, this value is used as the adequate sample area for assessing species abundance in the transport corridor. Species type and abundance was assessed following the example of Schmidtlein (2005). The first step was to define the study area by species composition. Two main vegetation

covers were identified: (1) a mixture of agricultural weeds, wild flowers and shrubs, and (2) deciduous woodland. A decision was made to ignore deciduous woodland in the Ellenberg mapping analysis as tall tree canopies can not only complicate the spectral signal from bidirectional reflectance and shadowing (Wolf et al., 2010; Feldmeyer-Christe et al., 2007), but also mask smaller species on the forest floor that may be a better indicator of soil moisture conditions. This is an important limitation on the use of Ellenberg values which is discussed in greater detail in Chapters 6 and 7.

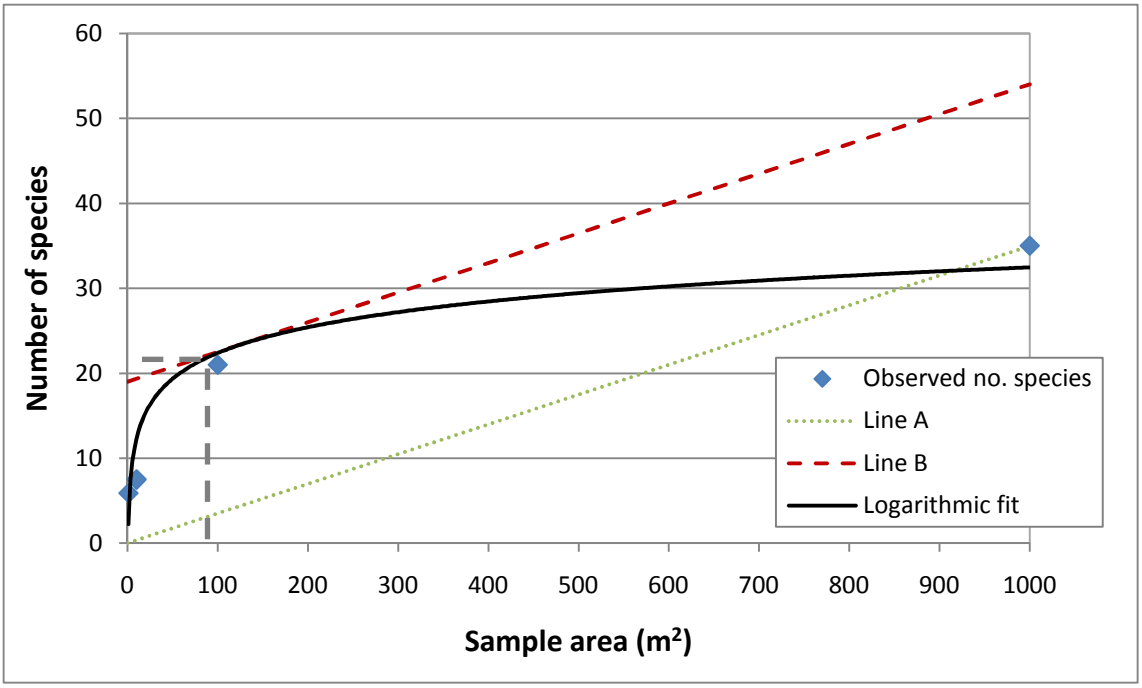


Figure 3.21 Species-area curve created using observations of plant species within a modified-Whittaker plot located on a cutting earthwork near Whitchester. The observed number of species are taken from the modified-Whittaker plot sampling. The solid line refers to a logarithmic fit on the observed number of species points. Line A refers to an increase in species and sample area in 10% steps. Line B is a tangent parallel to the 10% line (Line A) where it touches the species-area curve. The point at which Line B intercepts the logarithmic fit is taken to be the adequate sample area, indicated by the grey dashed line.

The two vegetation cover types, along with additional covers of bare earth and manmade surfaces, were classified using the CASI imagery. This was done so that random sampling points could be positioned over areas of agricultural weeds, wild flowers and shrubs, which were used to sample species abundance for the calculation of weighted Ellenberg values. Classification was carried out using a maximum likelihood classification routine. Many other classification approaches exist but these were not explored as this was not a listed objective for the study at hand. Furthermore, there is suggestion that adequate selection of training sites is more influential than the choice of classification procedure (Campbell, 2006; Mather, 2004).

The training areas used in the maximum likelihood classification routine were sampled using a mixture of ground-based knowledge of the study area and 5 cm aerial photography taken at the same time as the lidar data discussed in Section 3.3.1. The minimum area covered by each training area per class was selected using simple heuristics, which suggests 10 to 30 pixels per class, multiplied by the number of bands used in the classification (Lillesand and Kiefer, 2000), which in the case of the CASI sensor is 32 bands. A fuzzy convolution filter was applied to the resulting classification to reduce the speckled effect and improve overall interpretation. The classification accuracy was then assessed at 300 points located randomly throughout the study site, again using a mixture of field-based knowledge and the 5 cm aerial photography to determine the ‘true’ class of a pixel.

The overall classification accuracy was 90% with a kappa statistic 0.87 (Table 3.5). This suggests a high level of classification accuracy, however, discrepancies did exist. For instance, areas of bare ground were sometimes misclassified as vegetation, especially where these were adjacent to clusters of vegetation, or where vegetation was partially overhanging. In addition, ballast, which forms the basis of the railway track bed, often spilt over onto bare earthwork slopes. This led to misclassification of the ballast into the bare earth category. The classification of the weeds, wild flowers and shrubs land cover was shown to largely successful with some misclassification due to overhanging deciduous trees. The user’s accuracy for this class was 94.3% showing that this classification could be used to delineate the Ellenberg mapping study area with a certain level of accuracy.

	Unclassified	Bare ground	Deciduous wood	Manmade	Grass and shrub	Total
Unclassified	0	0	0	0	0	0
Bare ground	0	27	0	2	2	31
Deciduous wood	0	1	70	0	9	80
Manmade	0	9	0	75	0	84
Grass and shrub	0	0	6	0	99	105
Total	0	37	76	77	110	

	Producer's accuracy	User's accuracy
Bare ground	73.0%	87.1%
Deciduous wood	92.1%	87.5%
Manmade	97.4%	89.3%
Grass and shrub	90.0%	94.3%

Table 3.5 Error matrix showing the classification accuracy of the maximum likelihood classification routine applied to the airbone CASI imagery. The clasification was carried out in order to delineate agricultural weeds, wild flowers and shrubs for the random position of vegetation abundance plots.

The classification discussed above was used to randomly position plots over grasses, weeds and shrubs throughout the study area. These plots were then used to measure species abundance in order to calculate weighted average Ellenberg values. Schmidtlein (2005) recorded species abundance at random plots which consisted of three 1 m^2 circular relevés, positioned approximately 1 m from the centre of the plot, the position of which was measured using differential GPS. This strategy was adopted in the present study, but 1 x 1 m quadrats were employed, with 10 cm grid squares (representing 1% of total coverage). This was done to help improve an otherwise subjective measure of percentage species abundance as traditional methods for

recording species abundance involve an estimate of relative percentage cover per species. More accurate measures of species abundance can be made by harvesting the vegetation at each plot, then weighing the relative masses for each species (Kent and Coker, 1995). However, this was deemed impractical due to the need for additional personnel to collect field samples, as well as extended laboratory work. In addition, measuring species abundance through harvesting tends to be reserved for studies that are primarily concerned with biomass, or the yield and productivity of particular species (Kent and Coker, 1995).

Thirty-three plots were positioned over grasses, weeds and shrubs landcover throughout the study area using randomly generated coordinates, which were located in the field using RTK GPS. This ensured that an adequate area was sampled (approximately 100 m^2), as defined by the modified-Whittaker plot analysis. However, it later emerged that four of the plots could not be used due to misclassification of vegetation type, or inaccurate soil moisture measurements. Once species composition and abundance were recorded, a weighted average of the Ellenberg indicator value for soil moisture could be calculated for each plot using Equation 3.11.

$$E = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} x_i}$$

Equation 3.11

Where, *E* is the weighted average Ellenberg value, *x* refers to the relative Ellenberg value for the specific vegetation type (*i*), as defined by Hill et al. (1999), and *w* is the weight per species based on their relative abundance in the plot. For example, take a plot with a relative abundance of Bramble (50%), Stinging Nettle (20%) and Bracken (30%), with corresponding Ellenberg moisture values of 6, 6 and 5 respectively (Hill et al., 1999). Using Equation 3.11 the resulting weighted average Ellenberg value for this plot would be:

$$\text{Weighted average} = \frac{(6 * 50\%) + (6 * 20\%) + (5 * 30\%)}{100} = 5.7$$

3.3.2.1.2 Linking Ellenberg values with spectral reflectance using PLS

The processed reflectance data, described in Section 3.3.2, was linked to the average Ellenberg scores using PLS regression, following Schmidtlein (2005). As described in Section 3.2.2.5, this type of regression was employed as it can cope with a large number of noisy predictors that are not truly independent. Prior to analysis, the spectral reflectance data was transformed using $Log_{10}(1/R)$, and the derivative of the spectra. Details of these transformations, and the motivations behind their use, can be found in Section 3.2.2.5.

Using the selected CASI bands considered to be significant, and optimum number of components necessary to represent the variation in the data, Ellenberg soil moisture values could be mapped and compared to observed soil moisture data. As the West to East transport corridor typically includes earthworks with north or south facing slopes, separate models were also considered for the two dominating aspects. Ideally, additional plots would have been recorded to validate the Ellenberg indicator value models. However, due to the lengthy time needed to collect plot abundance data, as well as the need to collect validating soil moisture data at each point within a day, this was not possible. As a substitute, leave one out cross validation was carried out. The ability of the model to predict Ellenberg indicator values outside the observation point was assessed by comparing the RMSE for model calibration with the RMSE for the cross validation.

Part of the original plan in this part of the study was to collect field based spectroscopy at each of the Ellenberg plots. This would have provided an ideal reflectance dataset, with reduced atmospheric or bi-directional reflectance effects associated with airborne data, to fully understand the potential for mapping Ellenberg values using spectroscopy. However, damage by a third party to the field spectrometer, and lengthy estimated times of repair, meant that this was not possible, but remains an interesting scope for future study.

3.3.3 Model integration

The experimental methods described above are designed to identify metrics that are able to help characterise soil moisture distribution. Some of these metrics, namely the terrain calculations, have been discussed in terms of using multiple regression to improve

estimations by integrating the calculations. The final step of this methodology is to create a multiple regression model that incorporates all the metrics, derived from analysis of both the airborne laser scanning and spectroscopy datasets, to characterise soil moisture across the transport corridor. Multiple regression was used as predictions can be made outside the confines of the model calibration area. Multiple regression also provides a rudimentary measure of success with the R^2 calibration coefficient, which is comparable with the analysis techniques used in assessing the accuracy of the individual metrics throughout this methodology.

One of the limitations of using multiple regression is that is does not provide a thorough measure of model error. This could have particular bearing in this study as error propagation might occur where the remotely sensed metrics with their own levels of error are integrated, therefore compounding subsequent errors. In this respect it may have been interesting to explore a Bayesian modelling strategy whereby predictions are based on probabilities (Dungan, 1999). In the present case, areas of increased soil moisture can be identified by taking into account the probability of its occurrence predicted by the remotely sensed metrics. This strategy also has the ability to take into account known uncertainty in the metrics by assigning a probability that the metrics are not true. This method was not thoroughly explored as this type of conditional modelling has the tendency to be less accurate locally (Dungan, 1999) which could be significant when concerned with a component that is highly variable over space, such as soil moisture content.

Other modelling procedures include neural networks. This technique mimics the way in which the brain works, by using a number of simple processing units called neurons (Mather, 2004). Using neural networks the input predicting variables, in this case remotely sensed metrics, are inputted into the neurons with a specific weighting. The neuron then sums these weighted input variables and applies a simple calculation, such as a threshold. In the example of a threshold, if the input value is greater than a preset value then the output is sent onto the next layer of neurons, or to the final output or prediction. This is known as a feed-forward network (Næs et al., 2002). The issue with this type of model is that it does not learn itself and requires input information from the user. This means that the model must be trained by the user and modified accordingly to produce a more desirable result, in this case, a better prediction of soil

moisture distribution. Techniques to overcome this dependence on user input include back-propagation. Here, errors associated with the model output are used to redefine the weights applied to the model inputs. One known issue with this kind of model is that despite speedy computation times for the final model, training times can be very lengthy (Mather, 2004) which may be problematic when dealing with high spatial resolution data over large areas. In addition, models can become 'over trained' losing effectiveness over areas outside the training area (Mather, 2004).

In order to create the integrated model, stepwise linear regression was carried out using observed soil moisture measurements, gathered on 30th July 2009, as the dependent variable, and the remotely sensed metrics as the predictor variables. This method was favoured over standard multiple regression, as variables are entered into the model providing they fit a predefined criterion. This criterion is based on the significance of the predicting variable, which for the present study was defined as the probability of the F value (0.05 for entry and 0.1 for removal). The resulting model was used to map predicted soil moisture distribution across the corridor, and was validated using soil moisture measurements gathered on 18th June 2009.

As well as the global linear multiple regression model a spatially variable mode using GWR was also considered. However, it must be noted that this was not carried out in an attempt to improve upon the original global model, but to explore the spatial variability of the relationships between soil moisture and the predicting variables. This is because the model parameters estimated in the GWR routine are specifically tailored to the extent of the input calibration, or training data. Therefore any predictions made outside this area would be unstable and unreliable.

3.4 Summary

This Chapter has detailed the methods that were used to assess the potential of the remote sensing techniques identified in Chapter 2 that have the potential to characterise soil moisture over transport corridor earthworks. As explained at the start of this chapter, some of the techniques are conceptual and were therefore tested at an earthwork scale, using terrestrial remote sensing technologies. The details of these methods were outlined in Section 3.2. The techniques deemed successful, as well as other techniques that could not be assessed at a small scale, were then applied to

airborne remotely sensed data at a transport corridor scale using the methods detailed in Section 3.3. Although many of the techniques used here are not new, the novelty lies in their integration and use within the context of a transport corridor environment. Furthermore, attempts are made to cover some of the shortcomings associated with some of the techniques. In particular, there is a distinct lack of robustness in designing the sample strategy and subsequent statistical analysis of vegetation abundance data. It is hoped that the methods detailed here will provide a basis for applying such techniques in precision engineering.

4 Earthwork scale study results

This chapter details the results of the earthwork scale study following the application of the methods outlined in Section 3.2. This is broken down according to the type of data being investigated. Section 4.1, covers the terrain analysis techniques applied to terrestrial laser scanning data. Section 4. 2 relates to the spectral analysis techniques applied to field spectroscopy data. At the start of each of these sections a very brief overview of the relevant methods are given in order to bring the results into context.

4.1 Terrain analysis experiment

Section 3.2.1 detailed the methods used in the terrain analysis experiment. In brief, terrain analysis calculations were applied to DEMs generated from TLS point elevation data using kriging and AnuDEM interpolation techniques. These terrain analysis metrics were then compared to observed soil moisture measurements. The following details the results of this work.

4.1.1 Impact of DEM interpolation on representing terrain

Due to the relatively uniform relief of the test embankment, there was little difference between the elevation values interpolated using the kriging and AnuDEM techniques, with an average difference of 1 cm. This is reflected in Figure 4.1 which shows a plot of elevation values taken from the two DEMs for a transect running across the earthwork (the location of this transect is shown in Figure 4.3). The maximum difference between the two elevation profiles was 7 cm, which occurred towards the summit of the earthwork, at a distance of 13 m from point 1 (see Figure 4.3 for the location of point 1). Comparisons with the original TLS point elevation data showed very small residuals, with an average of -0.46 cm and 0.59 cm for the AnuDEM and kriging DEMs respectively. This corresponds to RMSE values of 0.68 cm and 0.77 cm respectively.

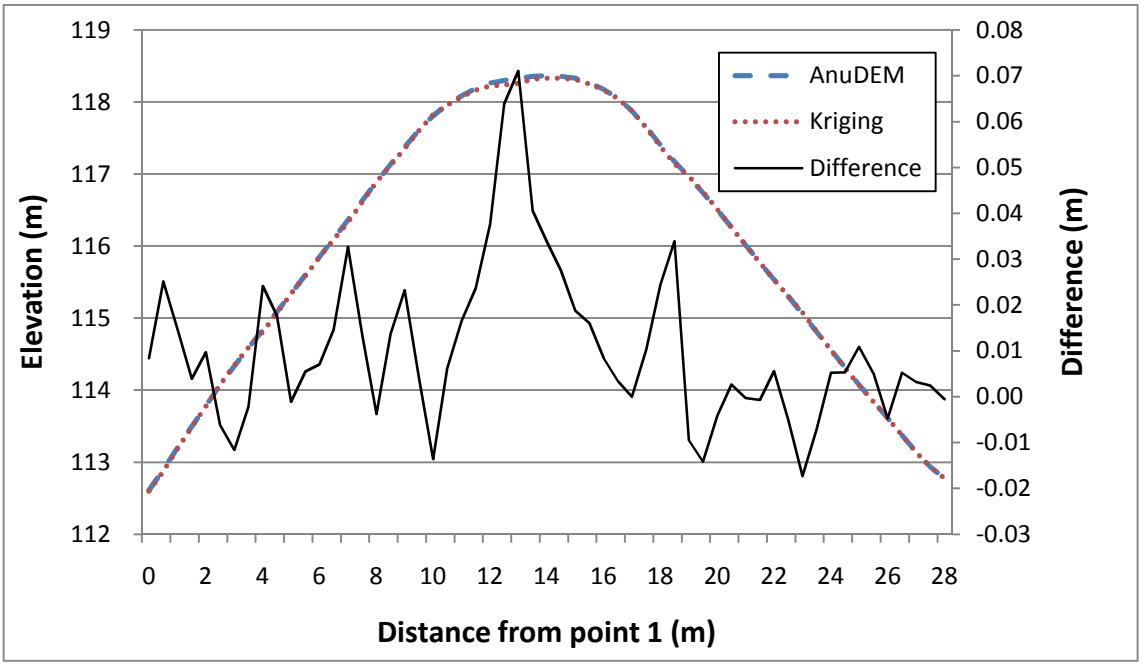


Figure 4.1 Profile of elevation values and differences across transect 1 for the kriging and AnuDEM interpolated DEMs. See Figure 4.3 for location of the transect.

Elevation values derived from the two interpolation techniques were also taken from transect 2 drawn along the contour of the embankment slope (location shown in Figure 4.3). Plots of these values, shown in Figure 4.2, demonstrated the degree of smoothing enforced by AnuDEM. This resulted in a maximum difference of approximately 5 cm at approximately 2.5 m from position y (location marked on Figure 4.3), where the AnuDEM method has suppressed the hollow feature that is identified in the kriging DEM. Despite these visual differences, a T and F-test showed there was no significant difference between the two samples (p value > 0.05). The average residuals, when comparing the DEMs to elevation values from the original TLS data, are 0.46 cm and 0.39 cm, corresponding to an RMSE of 0.68 cm and 0.62 cm for AnuDEM and kriging respectively. Due to the lack of difference between the two datasets, these results cannot be used to suggest which interpolation technique is more accurate for reproducing the original TLS elevation values.

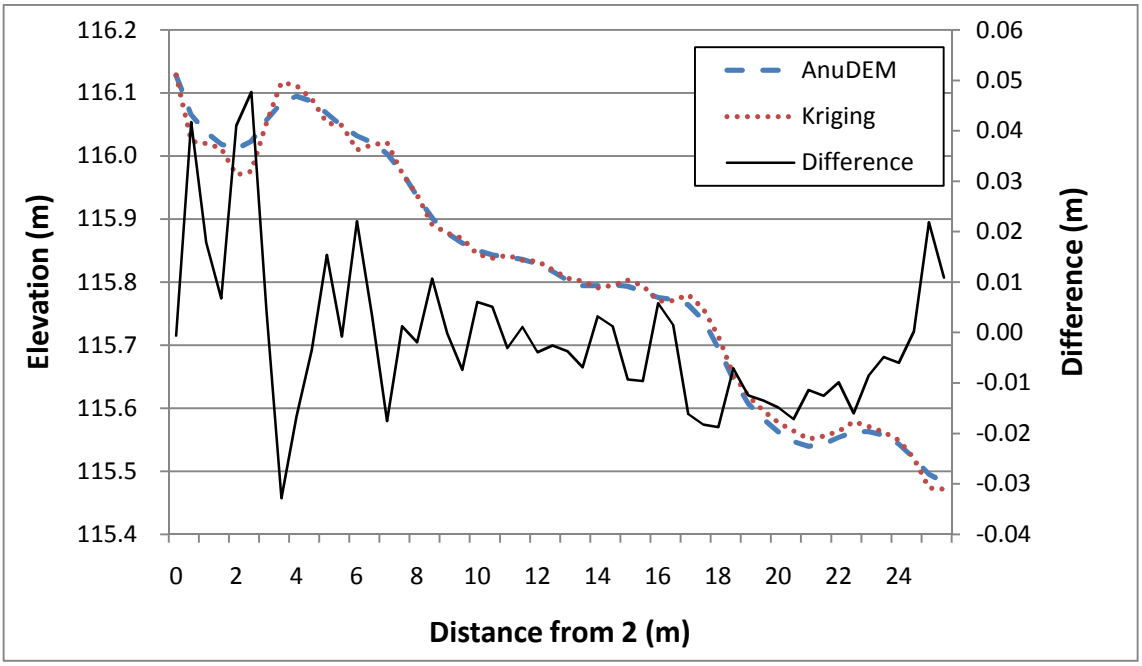


Figure 4.2 Profile of elevation values across transect 2 for the kriging and AnuDEM interpolated DEMs. See Figure 4.3 for location of the transect.

Hillshade analysis was used to make a visual assessment of how the two interpolation techniques represented the terrain of the test embankment. This was based on the assumption that increased terrain variability would be represented by an increase in the variation in hillshade values. Figures 4.3 and 4.4 show hillshades of the AnuDEM and kriging DEMs respectively. Although AnuDEM identified fine scale topographic features, such as the panel membranes and embedded instruments, these features were more prominent in the kriging interpolated DEM. This again reflected the smoothing effect enforced by the AnuDEM interpolation technique. As a result, kriging interpolation is more useful for depicting fine scale terrain, whereas the AnuDEM method creates a smoother surface. Figure 4.3 also identifies the location of missing data in the TLS dataset, which is noted in Section 3.2.1.

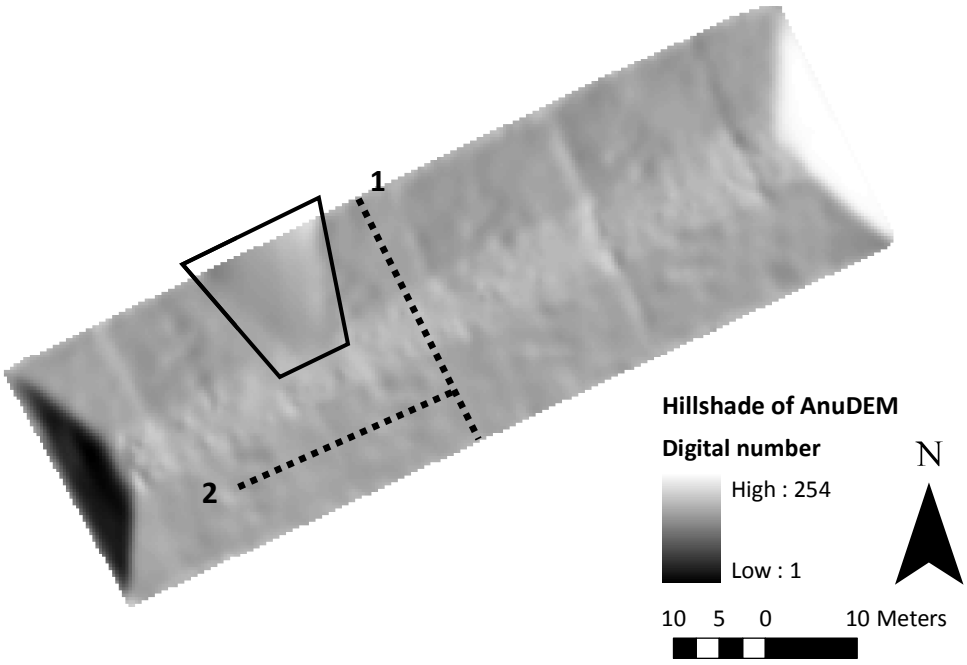


Figure 4.3 Hillshade of the AnuDEM interpolated DEM for the test embankment. Transects 1 and 2 refer to the location of the profile graphs in Figures 4.1 and 4.2 respectively. Quadrilateral shape shows the location of missing TLS data.

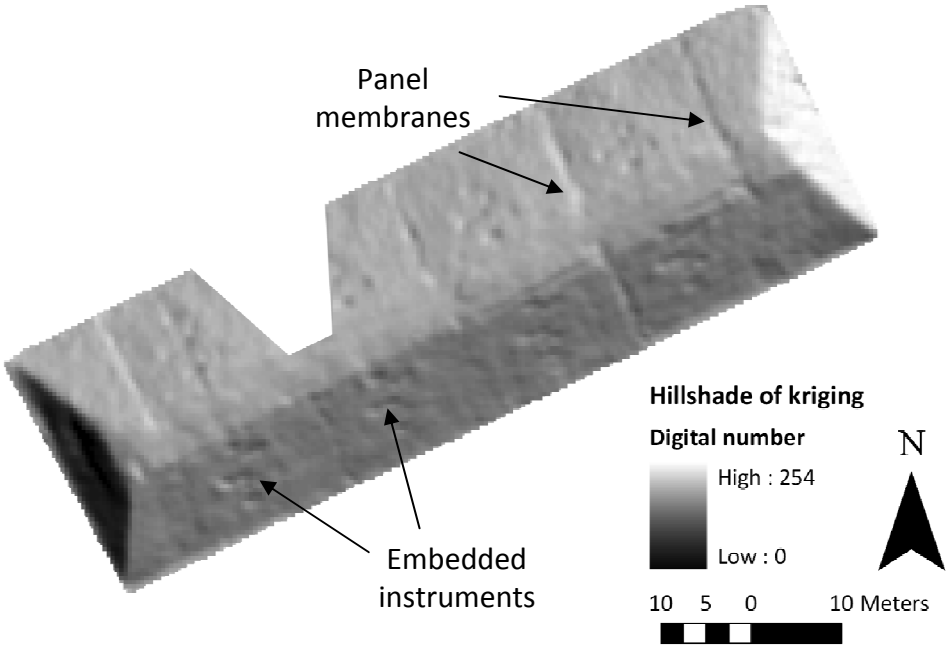


Figure 4.4 Hillshade of the kriging interpolated DEM for the test embankment. Markers identify topographic details that are more prominent in the kriging DEM compared to the AnuDEM DEM in Figure 4.3.

4.1.2 Impact of DEM interpolation on TWI calculation

Despite there being no significant difference between the elevation values for the two interpolated DEMs, there were marked differences in the TWI calculations. When analysing the TWI results it should be noted that high TWI values correspond to potentially drier areas. This is a result of the index being inverted to avoid errors where the slope is zero (see Section 3.2.1.1 for more details). Visually, TWI calculated using the AnuDEM interpolated DEM was noticeably smoother than the TWI for the kriging interpolated DEM (Figures 4.5 and 4.6 respectively). The kriging TWI included a number of seemingly random peaks across the earthwork, which are likely to be a response to the fine scale topographic detail identified using the kriging interpolation routine. There were also differences in the range of TWI values calculated, with 0 to 4.04 for the kriging DEM and 0 to 1.26 for the AnuDEM DEM. The high values found with the kriging DEM were attributed to a handful of grid cells located on the steep slopes at the southwest and northeast facing ends of the embankment. These extreme values occur due to the calculation of very small contributing areas which are the result of a flatter upslope neighbouring grid square taking all the upslope contribution. This did not occur with the AnuDEM DEM as the smoothing algorithm ensured continuous flow downslope.

Profiles of the TWI calculations for the different interpolated DEMs once again demonstrated the smoothness of the AnuDEM DEM compared to the kriging DEM. Moving downslope from the shoulder of the south facing slope (point 3 as indicated in Figure 4.6), the TWI values gradually decreased for the AnuDEM DEM (Figure 4.7). This trend should correlate with an expected downslope increase in soil moisture. In contrast, the profile of TWI values for the kriging DEM was more erratic. Moving downslope, there were three large peaks between 0 m and 4 m from point 3 with a TWI value of more than 0.25, indicating drier soils towards the shoulder of the slope, with the bottom half of the slope having relatively low values of approximately 0.05, indicating uniformly high soil moisture (Figure 4.7). The difference between the AnuDEM and kriging TWIs is also shown in Figure 4.7. This shows that there is little difference between the two calculations from 4 m to 11 m from point 3. In contrast, there is a great deal of difference between values from 0 m to 4 m. Here, the kriging

DEM shows three peaks TWI values at 0 m, 2 m and 4 m from position z, with a range of over 0.25 covering nearly the entire range of kriging TWI values.

The difference in TWI values was the result of a difference in the calculation of specific catchment area. Specific catchment area refers to the number of cells that flow into a cell of interest. In this example, values are represented by the cell size in m^2, therefore, for the AnuDEM DEM with a grid resolution of 35 cm (see Section 3.2.1 for details), a cell with a contribution of just one upstream grid cell will have a specific catchment area of 0.7 m^2 (i.e. the value in metres of itself and one upstream cell). The profiles in Figure 4.8 demonstrate that specific catchment area increased linearly downslope. In contrast, specific catchment area for the kriging DEM increases rapidly to 10 m^2 (representing approximately 22 grid cells) at approximately 5 m from the shoulder of the slope (Figure 4.8). This difference can be attributed to fine scale differences between the two input DEMs. The kriging DEM identified fine scale topographic features, which act as small local minima leading to the TWI being reset at these locations, resulting in a pattern of sporadic increases and decreases around the location of these features. Conversely, the AnuDEM suppressed these features and allowed flow to continue downslope resulting in a gradual increase in contributing area and a gradual decrease in TWI values (indicating an increase in potential surface wetness).

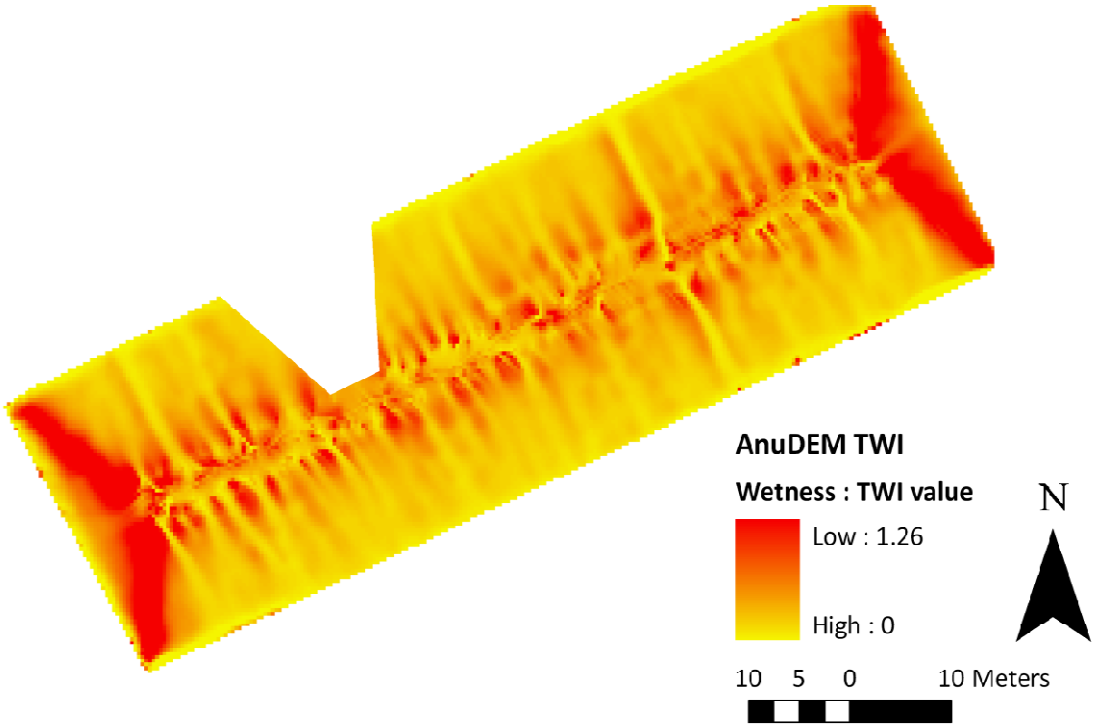


Figure 4.5 TWI calculated for the test embankment using the AnuDEM interpolated DEM. High TWI values correspond to low wetness.

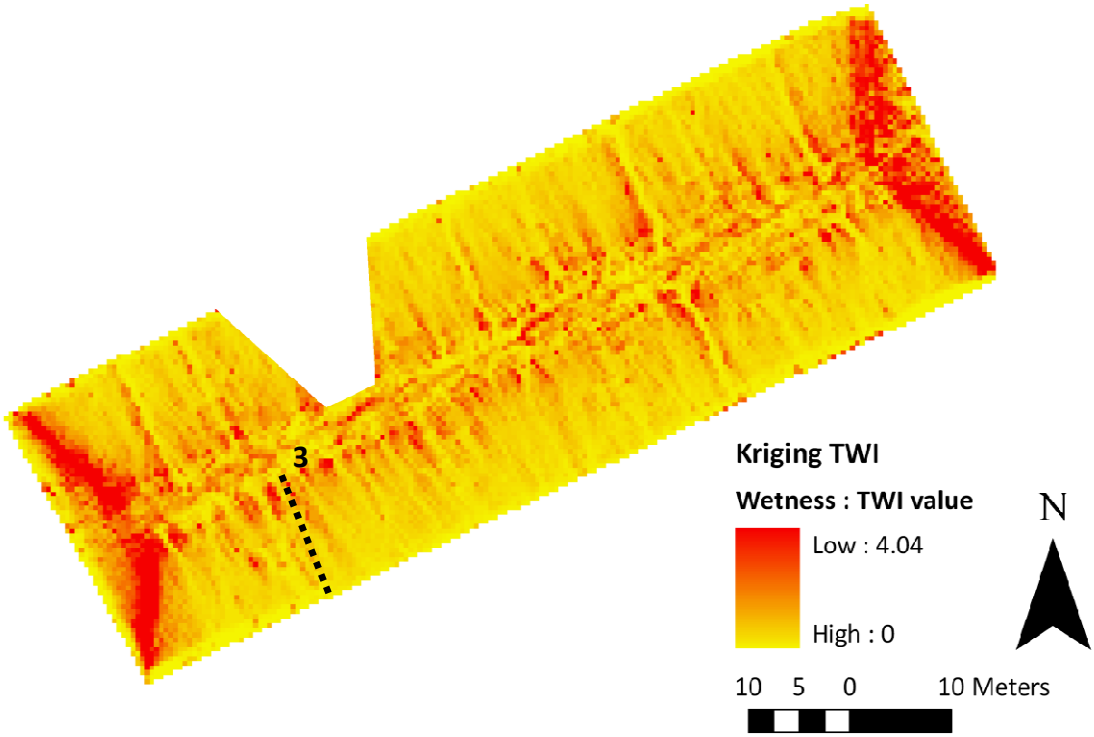


Figure 4.6 TWI calculated for the test embankment using the kriging interpolated DEM. Transect 3 shows location of profiles for Figures 4.7 and 4.8.

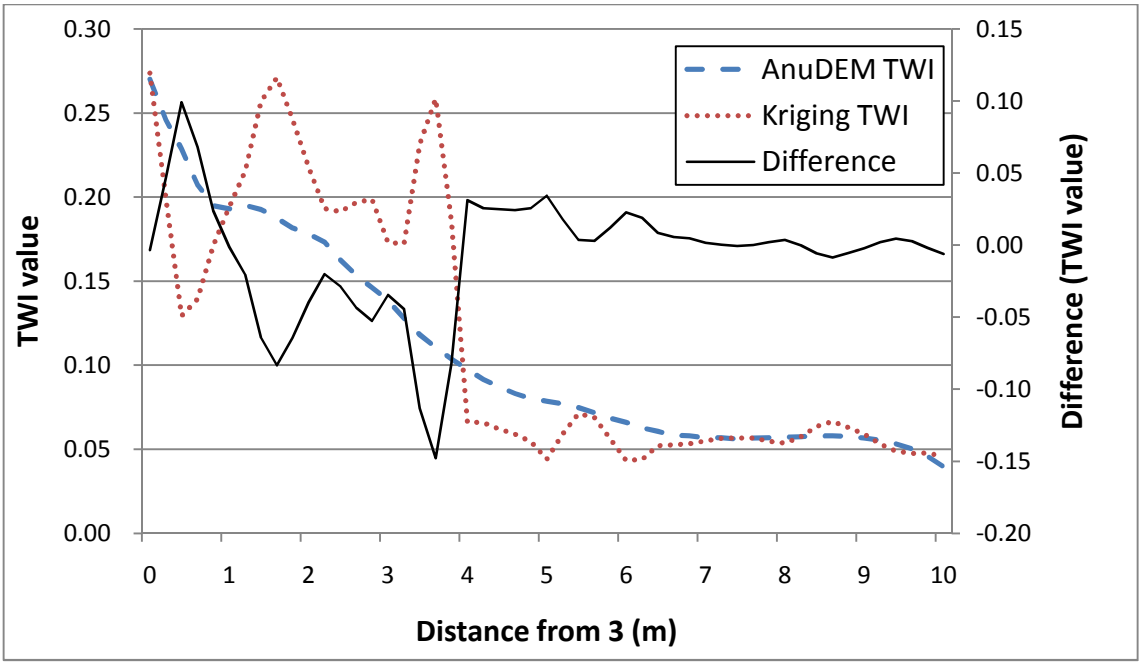


Figure 4.7 Plots of TWI values for kriging and AnuDEM interpolated DEMs for transect 3 shown in Figure 4.6.

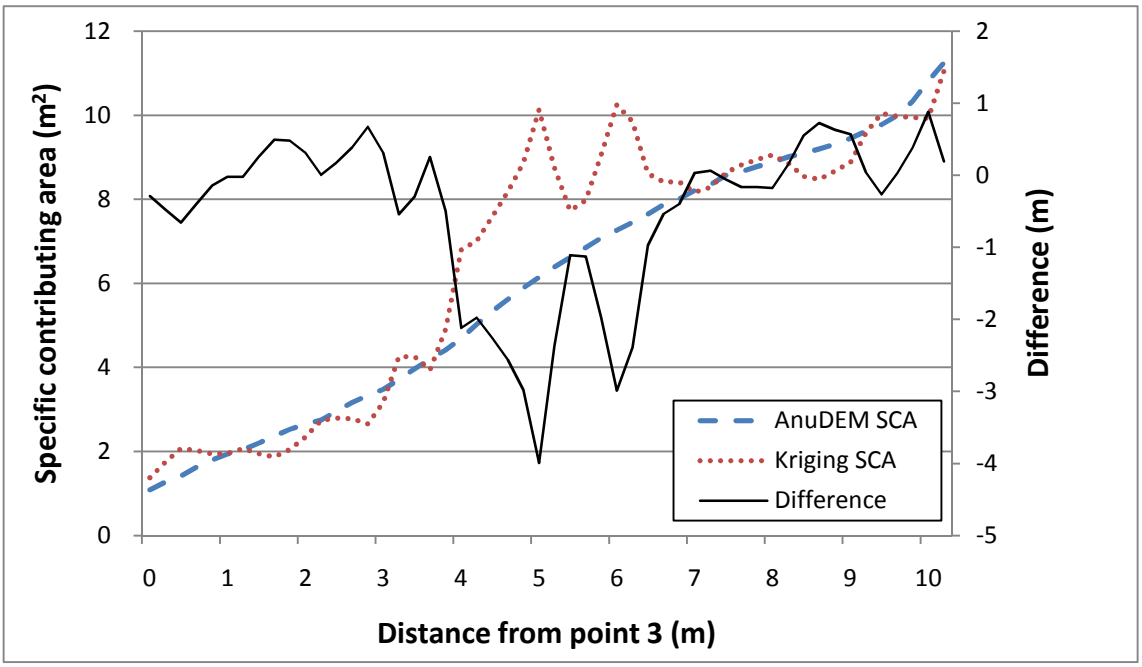


Figure 4.8 Profile plots of specific catchment area (SCA) for kriging an AnuDEM interpolated DEMs for transect 3 shown in Figure 4.6.

4.1.3 Impact of DEM interpolation on potential solar radiation calculation

The calculation of potential solar radiation showed that the north facing slope of the earthwork received less radiation than the south facing slope. Specifically, the calculations made for an entire year (2008) show a rise from approximately 600,000 WH/m^2 on the north facing slope to 1,000,000 WH/m^2 for the south facing slope (Figure 4.9). The calculations made for 24th November 2008 indicated a rise from 50 WH/m^2 on the north facing slope to approximately 190 WH/m^2 for the south facing slope (Figure 4.10). This increase is relatively large increase, with a percentage increase of 280%, compared to the calculation made for a year, with a percentage increase of 67%. This is because the calculation was made for 24th November, a time of the year in which the sun is relatively low meaning the north facing slope receives little direct radiation. As a result, there is a greater difference between the radiation received on either side of the embankment.

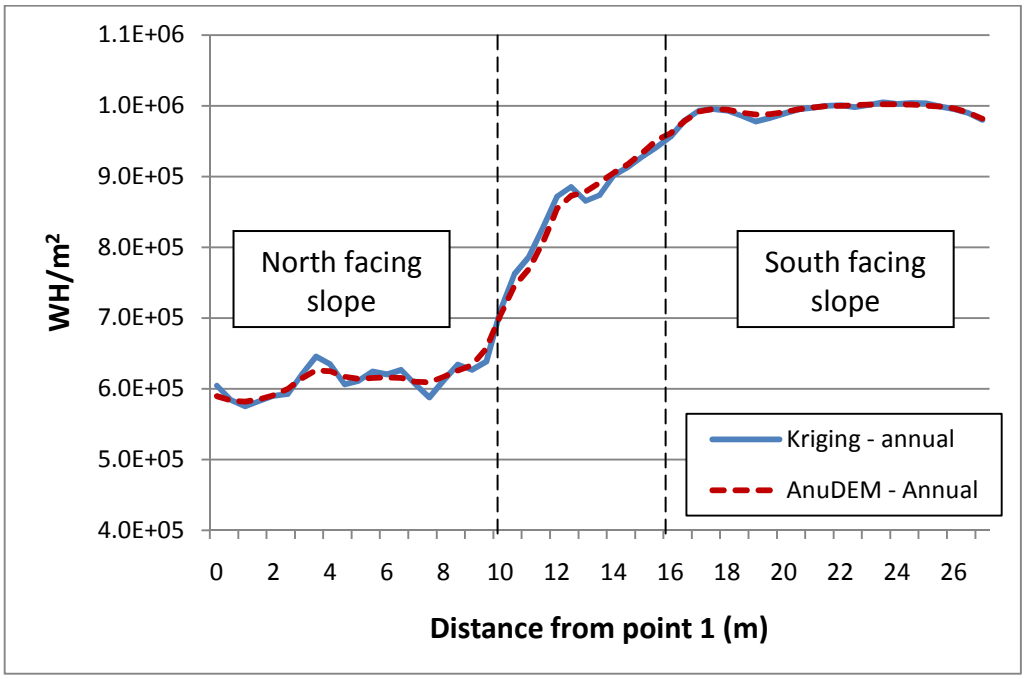


Figure 4.9 Plot of potential solar radiation values calculated for an entire year (2008) across transect 1 (see Figure 4.3 for location). Labels show corresponding north and south facing slope aspects.

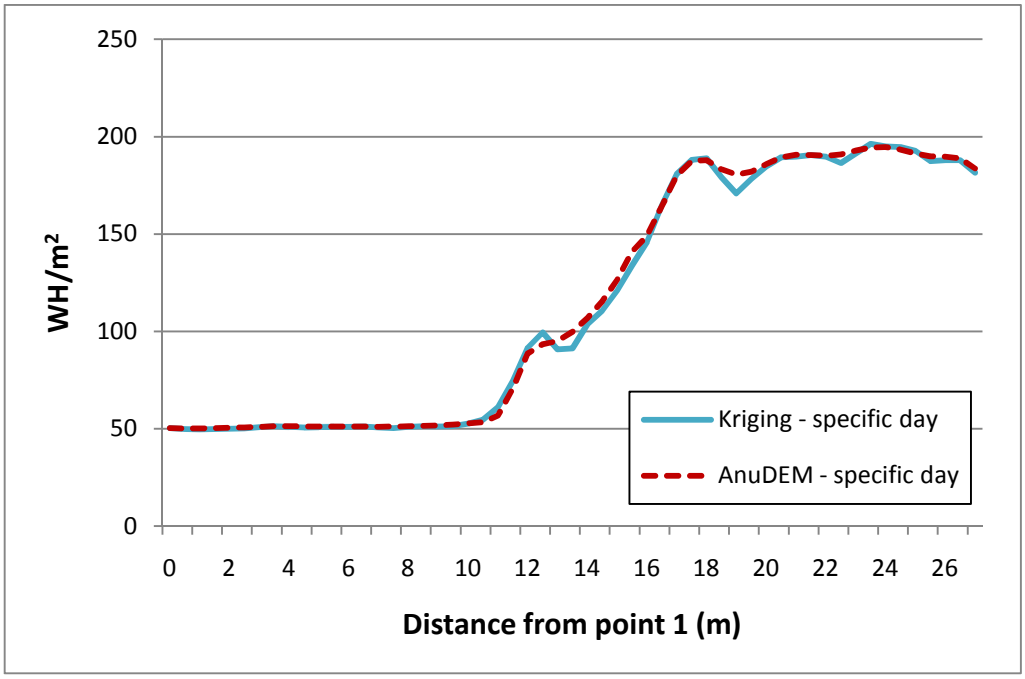


Figure 4.10 Plot of potential solar radiation values calculated for 24th November 2008 across transect 1 (see Figure 4.3 for location). Units are expressed as watt hours per metre square.

The north facing slope lies between 0 m and 10 m from point 1. The calculation of solar radiation for this portion of the earthwork showed more variation when made for an entire year compared to the calculation made for 28th November 2008 (Figures 4.9 and 4.10). This is because the calculation made for an entire year includes the summer period where the sun is high enough for the both slopes to receive direct radiation. This increase in variation is apparent in Figures 4.11 and 4.12 that show maps of solar radiation values calculated for 28th November 2008 and an entire yearsrespectively, using the kriging DEM.

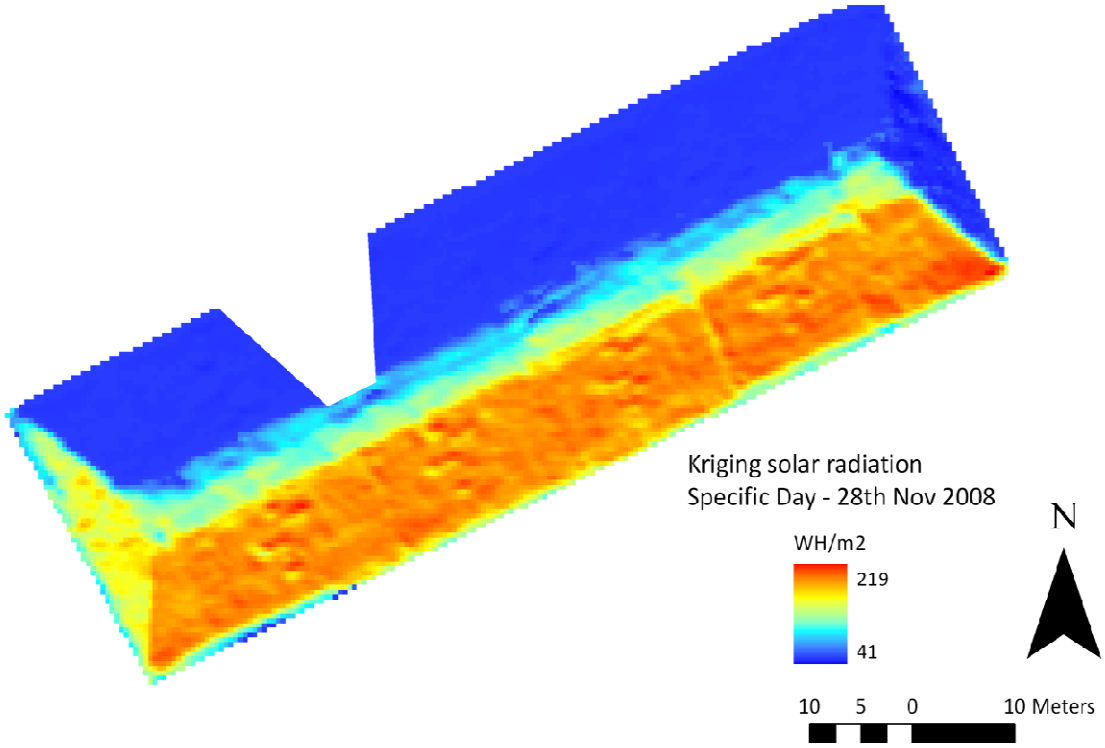


Figure 4.11 Potential solar radiation calculated for 28th November 2008 using the kriging interpolated DEM.

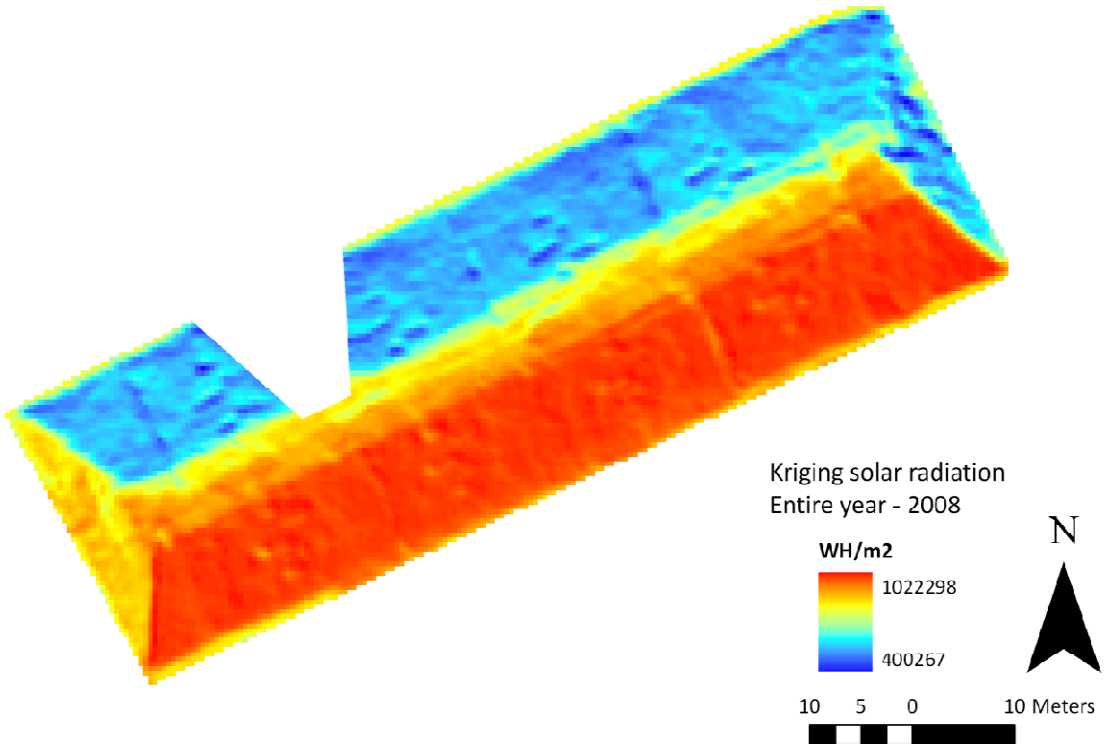


Figure 4.12 Potential solar radiation calculated for the year 2008 using the kriging interpolated DEM.

The examples shown in Figures 4.11 and 4.12 show that the calculation of potential solar radiation made over an entire year demonstrated more variation for the north facing slope. This was also apparent when the calculation was applied to the AnuDEM DEM, as demonstrated in the profiles shown in Figures 4.9 and 4.10. However, this variation was not as prominent as the calculations made using the kriging DEM, with the latter showing a range of 135,765 WH/m^2 over a distance of 0 m to 10 m from x, compared to a smaller range of 122,135 WH/m^2 for the AnuDEM DEM. This was also apparent in the maps of solar radiation values calculated using the AnuDEM DEM (Figures 4.13 and 4.14) which were generally smoother and revealed less response over the embedded instruments and panel membranes (the locations of which are shown in Figure 4.4). Similar to the calculations of TWI discussed in Section 4.1.2, this is due to the smoothing enforced by the AnuDEM interpolation technique. The fine scale topographic features were suppressed by this smoothing, leading to a more uniform calculation of solar radiation.

Whereas this smoothing is preferable in the calculation of TWI, to ensure continuous flow downslope, this may be less accurate for the calculation of solar radiation. This is because solar radiation at a given DEM grid square is largely governed by its aspect and the presence of features that may obscure direct radiation, i.e. shadowing. Therefore, it would be preferable to have a DEM that represented the terrain as accurately as possible. Although a T-test showed no significant difference between the two (Section 4.1.1), the profiles of elevation values (Figures 4.1 and 4.2) and hillshade analysis (Figures 4.3 and 4.4) showed that the kriging DEM better represented fine scale topographic features. Therefore, a DEM created using this interpolation method is better suited to the calculation of the potential solar radiation.

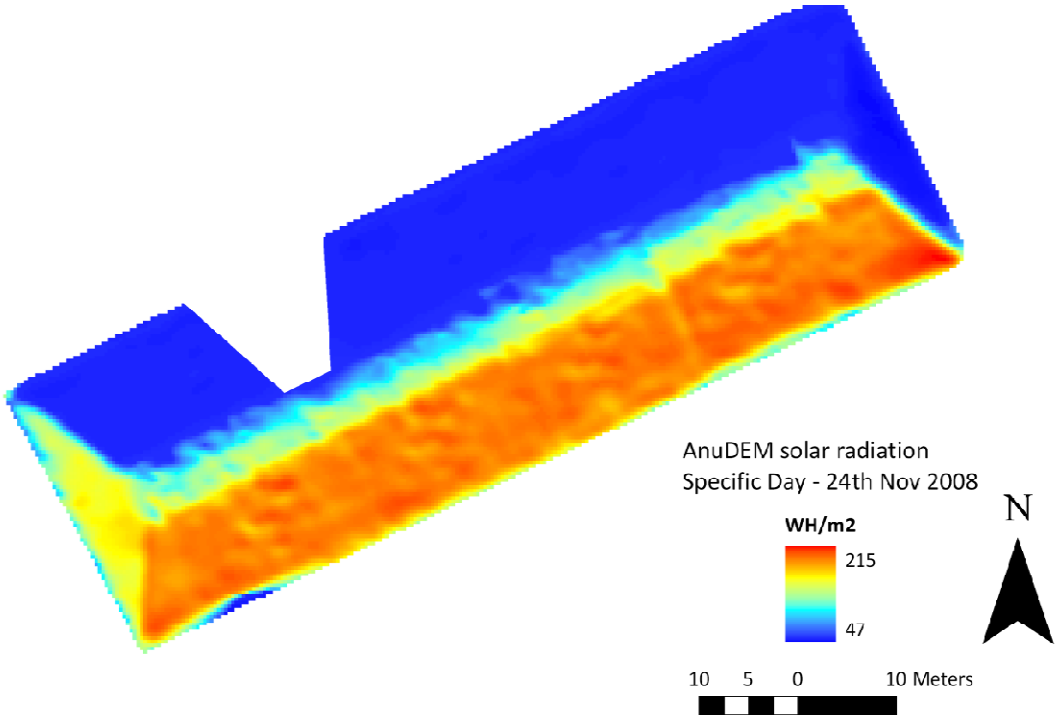


Figure 4.13 Potential solar radiation calculated for 28th November 2008 using the AnuDEM interpolated DEM.

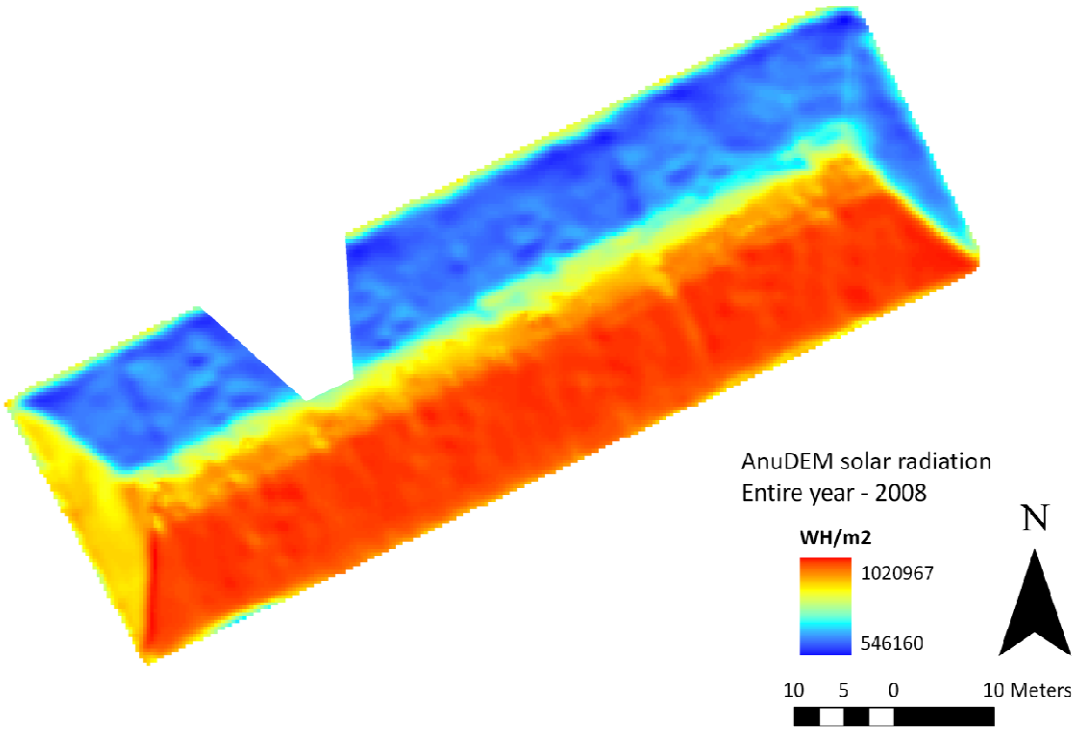


Figure 4.14 Potential solar radiation calculated for an entire year (2008) using the AnuDEM interpolated DEM.

4.1.4 Impact of DEM interpolation on the calculation of aspect

The difference in the calculation of aspect using the kriging and AnuDEM interpolated DEMs followed similar trends to the differences found in the calculation of potential solar radiation (Section 4.1.3). Aspect calculated using the kriging DEM showed a great response to the panel membranes and embedded instruments. This resulted in a greater variation in aspect which is illustrated in the transect of values shown in Figure 4.15. This transect shows that the transition of values for the AnuDEM DEM is smoother than the kriging DEM, which is a reflection of the smoothing enforced by the AnuDEM interpolation routine. This suggests that the kriging DEM is better suited to the calculation of aspect over the test embankment, as this interpolation routine better represents fine scale terrain features. These features could be important where localised variations in solar radiation result in variations in soil moisture, such as a hollow feature which retains moisture as it does not receive direct radiation.

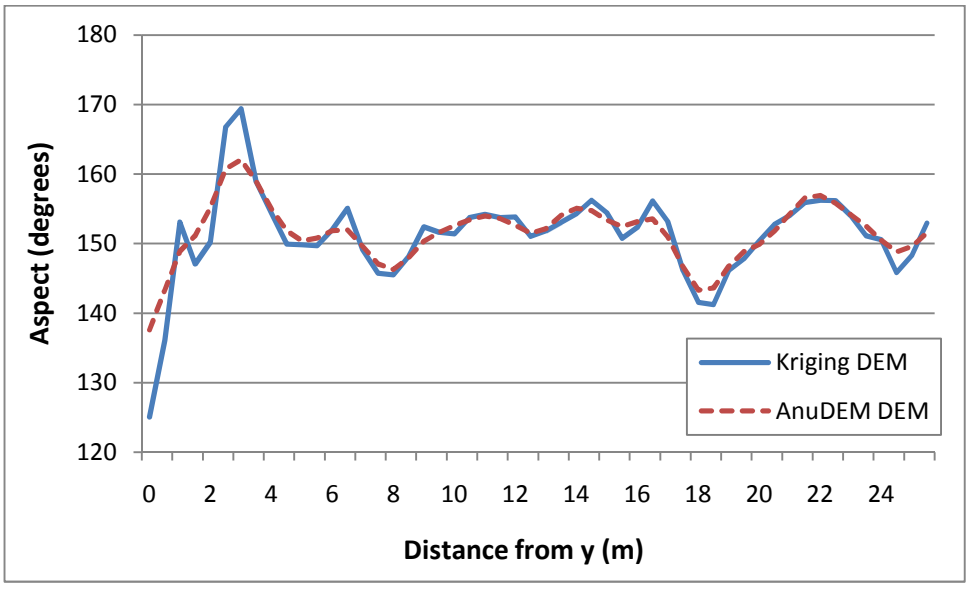


Figure 4.15 Profile of aspect values (degrees) taken from transect 2 running along the contour of the test embankment. Values range between 125 and 170 as the transect was taken from the south facing slope, location of which is shown in Figure 4.3.

4.1.5 Impact of earthwork orientation on potential solar radiation calculation

Due to the west to east orientation of the test embankment, the main variation in potential solar radiation was polarised by the north and south facing slope aspects. This dominant feature masks small scale variations that may exist on the slopes themselves. Generally, it was found that the influence of small scale features was augmented by calculating potential solar radiation for an entire year, therefore taking into account direct radiation of the north facing slope during summer periods, using a kriging interpolated DEM, which represented the fine scale topographic features. A test was carried out to see how the calculation might perform on a north to south orientated earthwork, by rotating the DEM in a GIS environment.

The calculations of potential solar radiation applied to the rotated DEMs demonstrated a response to smaller scale terrain features, rather than the polarisation exhibited in the original DEM orientation. A map of potential solar radiation calculated using the kriging DEM demonstrates this well (Figure 4.16) with variation being reported over the embedded instruments and panel membranes. This was also reflected in a transect of values drawn across the contour of the earthwork (Figure 4.17) which showed a fluctuation in solar radiation calculated for 28th November 2008 at 2.5 m, 20 m and 38.5 m in response to the panel membranes identified by the kriging DEM. The difference in solar radiation over these features reaches a maximum of 67 WH/m^2 (20 m from point 4) which represented over 85% of the total range in values calculated across the transect. This showed that variation in solar radiation calculated for a north to south orientated earthwork was controlled by the presence of topographic features on the slopes rather than a difference in slope aspect as seen with the original west to east orientated earthworks.

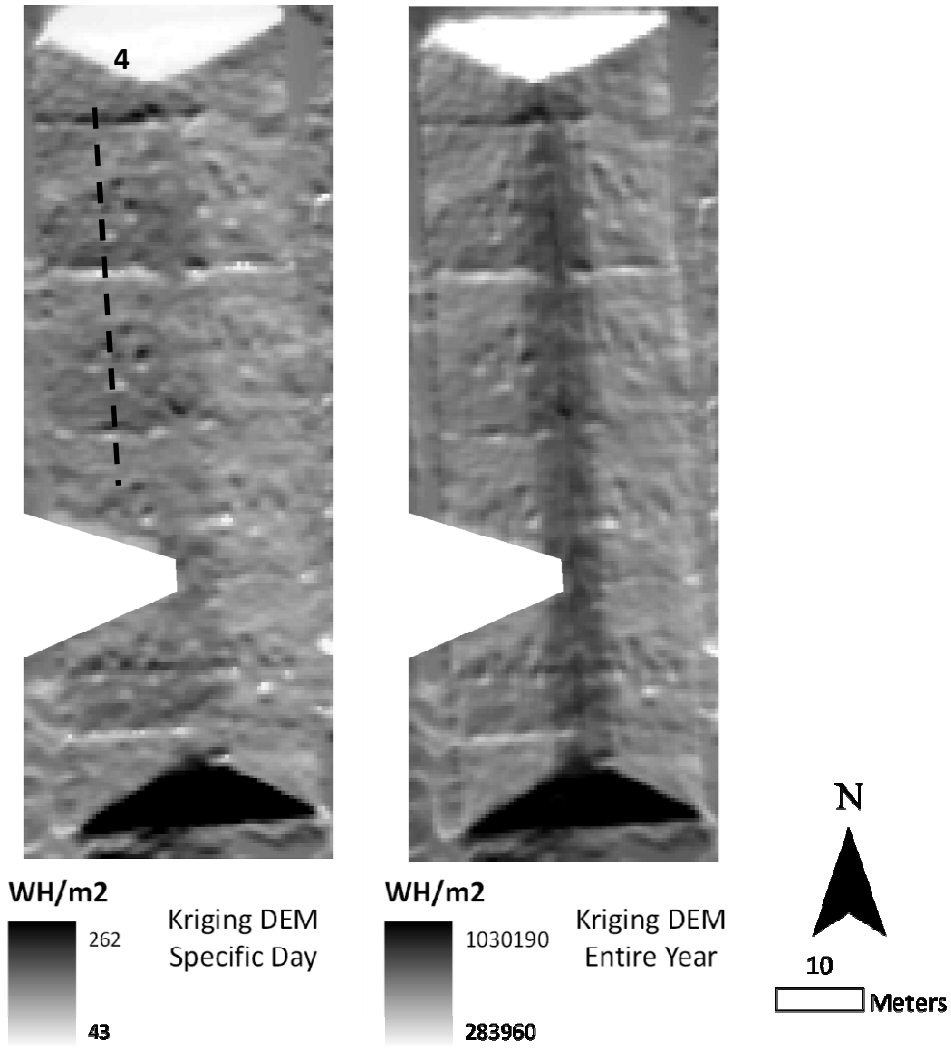


Figure 4.16 Potential solar radiation calculated for 28th November 2008 and an entire year (2008) using the kriging DEM rotated to be orientated approximately north to south. The location of transect 4 is also shown.

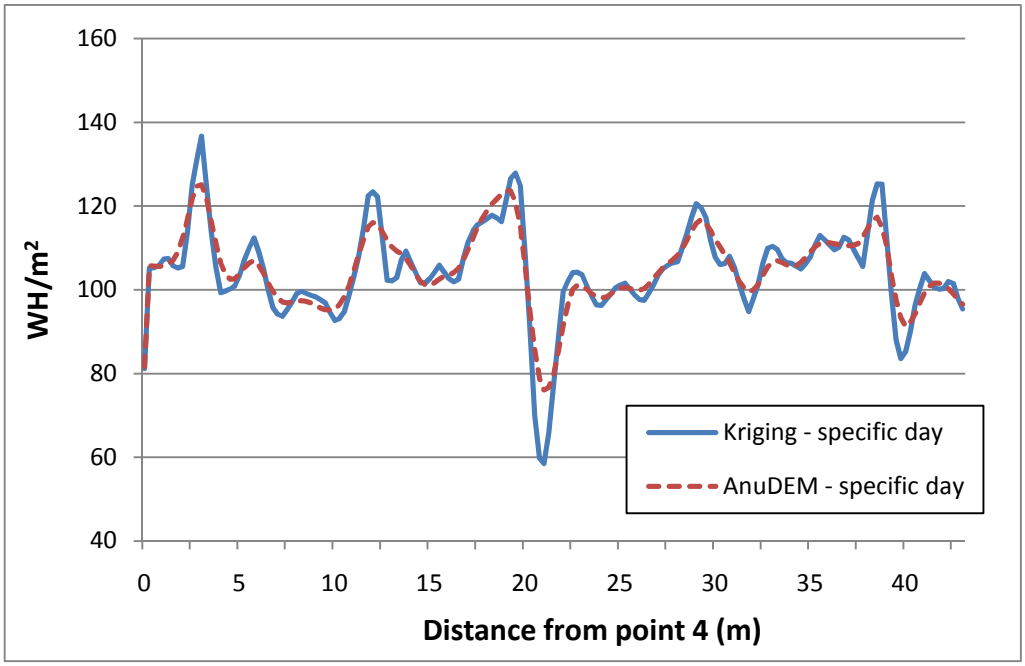


Figure 4.17 Transect of potential solar radiation values across the contour of the simulated north to south orientated earthwork. The location of transect 4 is shown in Figure 4.16.

The map of solar radiation calculated using the AnuDEM DEM also depicted a response to the panel membranes and embedded instruments (Figure 4.18). However, this response was less pronounced as seen with the kriging DEM due to the suppression of topographic features in the AnuDEM interpolation. This is illustrated by the transect values in Figure 4.17. In comparison to the values calculated using the kriging DEM, the AnuDEM showed a smoother variation in values over the features at 2.5 m, 20 m and 38.5 m. For example, at a distance of 20 m from point 4 the difference in solar radiation was 42 WH/m^2 compared to 67 WH/m^2 for the kriging DEM. This again suggests that a DEM created using kriging interpolation is better suited to the calculation of potential solar radiation. There was no significant difference between the distribution of values calculated for 28th November 2008 and an entire year. This was indicated by a Pearson correlation of over 0.98 for both the AnuDEM and kriging calculations and an F-test indicating that sample variances were not significantly different ($p < 0.005$). This is because direct radiation is distributed uniformly across the earthwork in a similar fashion to the calculation made over an entire year. However, the variation in solar radiation over a year has previously been shown to influence soil

moisture distribution patterns (Western et al., 1999a). If the calculations made here do not vary significantly for 28th November 2008 or an entire year then it is unlikely that predicted solar radiation would be a strong predictor of soil moisture. Therefore, the use of this terrain analysis for characterising soil moisture may be limited to earthworks with north and south facing aspects, or other topographic features that result in significantly different solar radiation distributions.

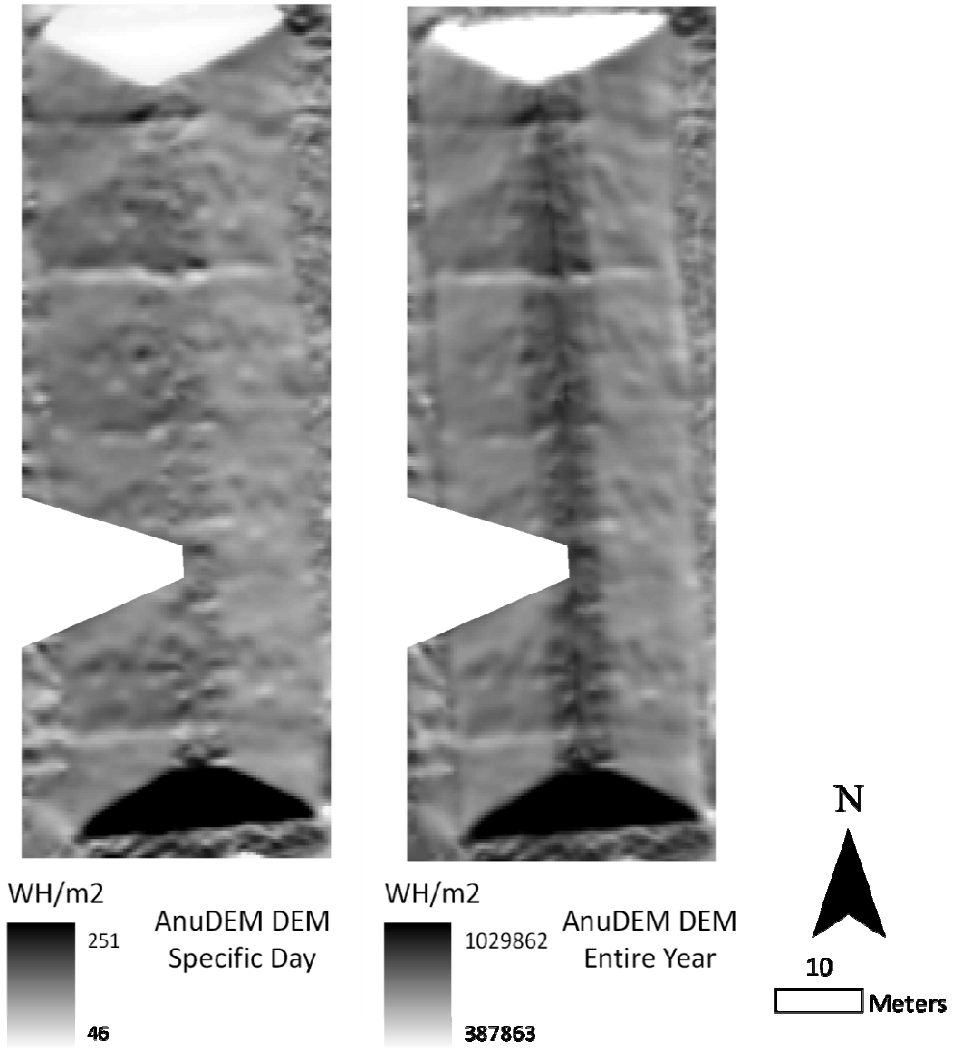


Figure 4.18 Potential solar radiation calculated for 28th November 2008 and an entire year (2008) using the AnuDEM DEM rotated to be orientated approximately north to south.

4.1.6 Impact of DEM resolution on the terrain analysis calculations

As discussed in Section 4.1.1, the original grid DEMs created using the AnuDEM and kriging interpolation methods depicted fine scale topographic features such as embedded instruments and panel membranes. When the DEM resolution was degraded to 2 m this information was lost, but the overall shape of the earthwork was retained. As a result, the calculation of TWI did not show a response to topographic features such as panel membranes or embedded instruments. Interestingly, at this resolution there was little difference between the calculations of TWI using the two DEMs. Figure 4.19 shows a transect of TWI values across the profile of the earthwork demonstrating the similarity between the two profiles. A T-test showed that there was no significant difference between the two sample means ($p < 0.005$), with a Pearson correlation of 0.98. Furthermore, a F-test suggested there was no evidence to suggest that the variance of the two samples were significantly different. This was due to small differences in the TWI calculation, particularly at a distance of 10 m from point 1 (location shown in Figure 4.3) where the AnuDEM TWI was approximately 11% higher than the kriging TWI.

In Section 4.1.2, the differences in the calculation of TWI between the two interpolated DEMs was found to be due to a difference in the calculation of specific catchment area. However, at 10 m from point 1 there was no difference between specific catchment area as the only contribution this grid cell was receiving was from itself, i.e. one grid square (equating to 2 m). Rather, the differences are found in the calculation of slope grid with the kriging DEM showing a slope grid value of 0.51 compared to 0.48 for the AnuDEM DEM. Although the difference is small, it shows the kriging DEM at this grid square to be steeper than the AnuDEM DEM, resulting in a higher TWI value, corresponding to a lower potential wetness. This is again the result of the smoothing mechanism of the AnuDEM interpolation routine ensuring that fine scale topographic features are suppressed.

A T-test showed there to be no significant difference between the means of potential solar radiation values calculated using the two interpolated DEMs with a grid resolution of 2 m ($p < 0.005$). An F-test also indicated that there was no difference between the variance of the two datasets. This is because the resolution of the DEM is too coarse to identify topographic detail, such as panel membranes and embedded

instruments that was evident in the finer resolution DEMs (as discussed in Section 4.1.1.3). This means that interpolation method has little or no influence on the calculation of potential solar radiation for the test embankment at grid resolutions coarser than 2 m.

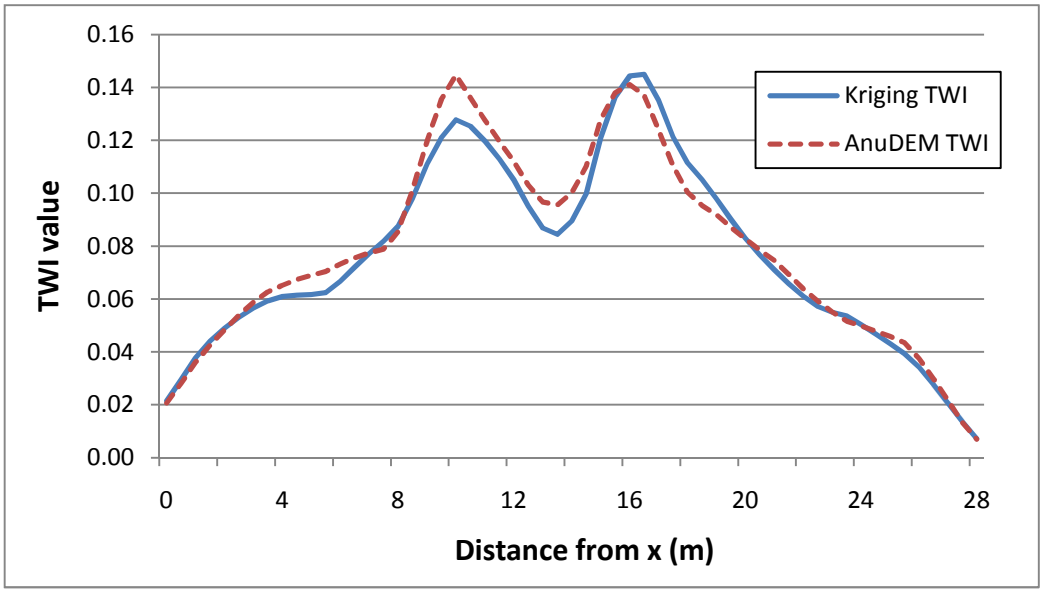


Figure 4.19 Transect of TWI values across the profile of the test embankment calculated for the kriging and AnuDEM DEMs. Location of transect 1 is shown in Figure 4.3.

Some of the fine scale topographic features that were identified in the original resolution DEMs (both kriging and AnuDEM interpolated) were also apparent in the 1 m degraded DEMs. Although less pronounced than in the original DEM, the hillshade of the 1 m kriging DEM demonstrated this with variations found around the panel membranes (Figure 4.20). A transect of elevation points across the contour of the kriging and AnuDEM interpolated DEMs revealed subtle differences (Figure 4.21). Unlike the 2 m resolution DEMs, the smoothing of topographic features again becomes apparent at a grid resolution of 1 m. For example, the panel membrane features at 10 m and 25 m from point 5 are clearly shown, with a reduction in elevation of over 10 cm in both DEM profiles. However, the surface roughness depicted in the kriging DEM from approximately 17 m to 23 m from point 5 has been smoothed out in the AnuDEM profile.

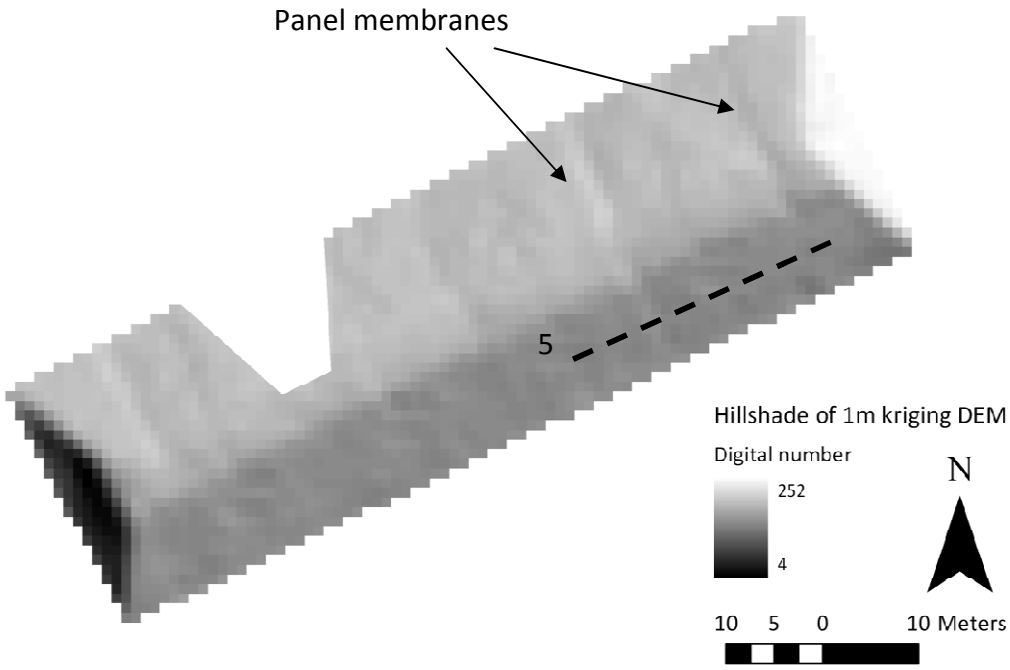


Figure 4.20 Hillshade analysis of the 1 m degraded kriging DEM with annotations identifying the panel membrane features and the location of transect 5.

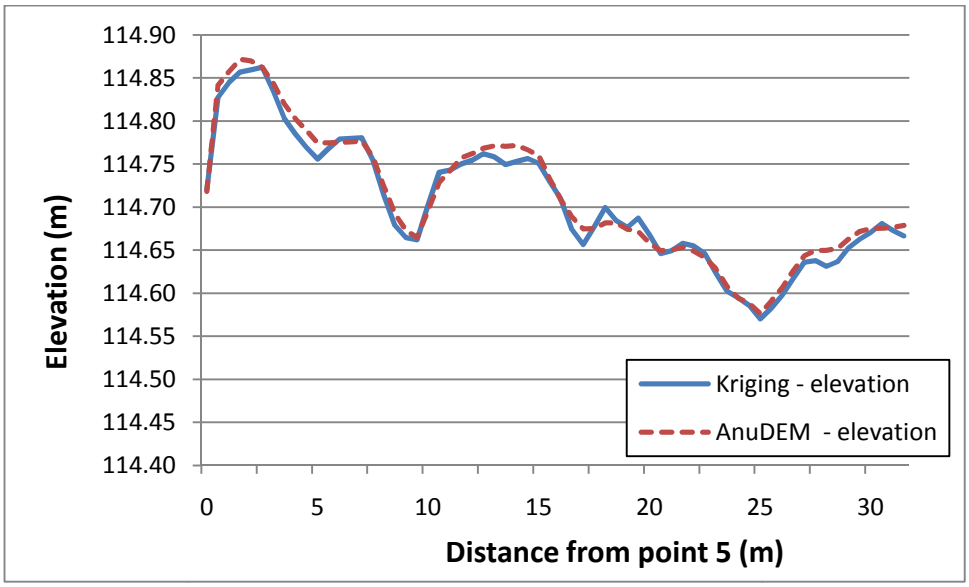


Figure 4.21 Profile of elevation values from transect b along the contour of the test embankment for the 1 m kriging and AnuDEM DEMs. The location of transect 5 is shown in Figure 4.20.

Respective T and F-tests showed that there was no significant difference ($p < 0.005$) between the calculation of potential solar radiation using the 1 m kriging or AnuDEM interpolated DEMs. However, there was a difference with the calculation of TWI. Values taken from transect 3 (the location of which is identified in Figure 4.6) showed the AnuDEM wetness index to gradually decrease downslope, whereas the kriging TWI was less consistent, with a fluctuation at approximately 6 m from point 3 where the value was 15% higher (Figure 4.22). This is because the kriging DEM represented finer scale features resulting in fluctuations in specific catchment area, slope grid and subsequent TWI. This is similar to the findings made with the original resolution DEM in Section 4.1.2 demonstrating that at a grid resolution of 1 m, TWI should be calculated using an AnuDEM interpolated DEM.

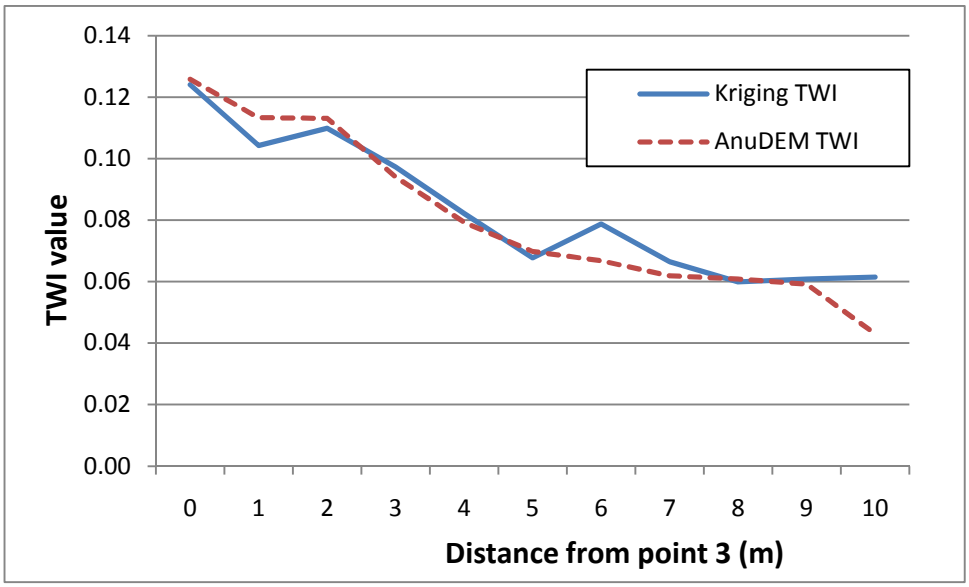


Figure 4.22 Downslope profile of TWI values for transect z calculated using kriging and AnuDEM DEMs. High TWI values represent low wetness. The location of transect 3 is found in Figure 4.6.

4.1.7 Description of the test embankment soil moisture measurements used to compare against the terrain analysis calculations

The soil moisture measurements used throughout this study, at both the earthwork and transport corridor scale, were made using a Theta portable impedance probe. This probe was calibrated using gravimetric soil moisture measurements taken from soil samples located across the test embankment. A scatter plot comparing the measurements can be found in Figure 4.23. The results of the calibration indicated an R^2 value of 0.85 which proved to be better than the results reported by Cosh et al. (2005) who showed correlation coefficients ranging from R^2 of 0.61 to 0.84. The average residual was 2.9% resulting in a total RMSE of 1.7%. Although these values are very small, the residuals in some cases reached almost 10%. Deviations of this magnitude could have a significant effect on the comparisons made with the remotely sensed metrics used in this study. In addition, sampled soil moisture was poorly represented between 0 - 20% and therefore there is uncertainty over the accuracy of the probe during dry conditions.

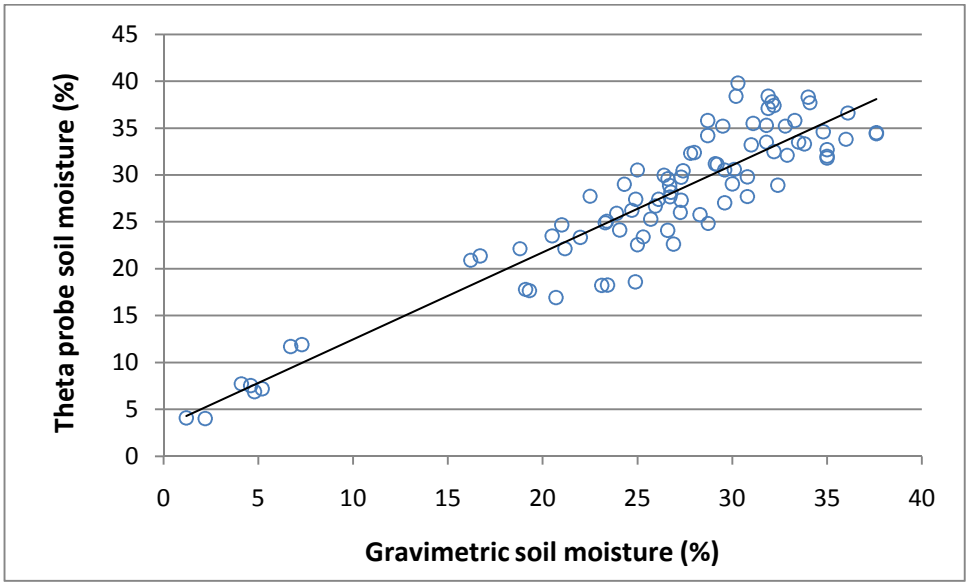


Figure 4.23 Scatter plot between Theta probe and gravimetric soil moisture measurements for the calibration of the Theta probe.

Observed soil moisture measurements were obtained on 24th November 2008 for comparison with the terrain analysis calculations. Table 4.1 shows the descriptive statistics for this dataset. Local weather data indicated a number of rainfall events over

the few weeks before the measurements were taken, resulting in overall wet conditions on the embankment with a mean volumetric soil moisture content of 41.1%. The mean soil moisture value masks a wide range with a minimum and maximum of 18.3% and 50.3% respectively, and a standard deviation of 5.2%. The measure of skewness shows that the distribution of this data is negatively skewed. This is reflected in the histogram shown in Figure 4.24 which shows a bulk of the measurements to be distributed around 43% and a tail tending towards lower soil moisture values, below 30%.

Descriptive statistics	
Mean	41.1
Standard Deviation	5.16
Skewness	-0.90
Minimum	18.3
Maximum	50.3
Count	311

Table 4.1 Table showing descriptive statistics for the test embankment soil moisture measurements gathered on 24th November 2008. Values are in volumetric soil moisture (%).

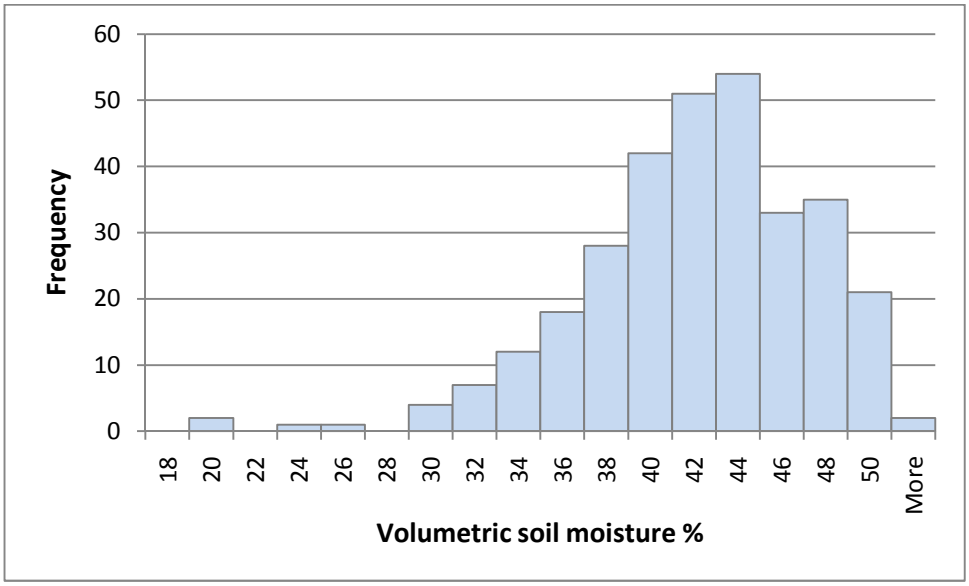


Figure 4.24 Histogram showing the distribution of soil moisture measurements for the test embankment gathered on 24th November 2008.

Figure 4.25 shows the location and value of the soil moisture measurements. The first pattern that emerged was a slight increase in overall values for the north facing slope, illustrating the influence of aspect. This was reflected in an average soil moisture content of 43% and 38.9% for the north and south facing slopes respectively. Values were shown to vary more on the south facing slope with a range of 30.2% compared to 25% for the north facing slope. The variation in soil moisture contents were higher than expected and shows that there was a lack of spatial organisation, which may be the result of a rainfall event just prior to data collection. This may have implications on the comparison between these measurements and the terrain analysis metrics because the metrics base their prediction on patterns of soil moisture distribution. Therefore, poor correlations might be expected if the observed soil moisture values exhibit little or no pattern.

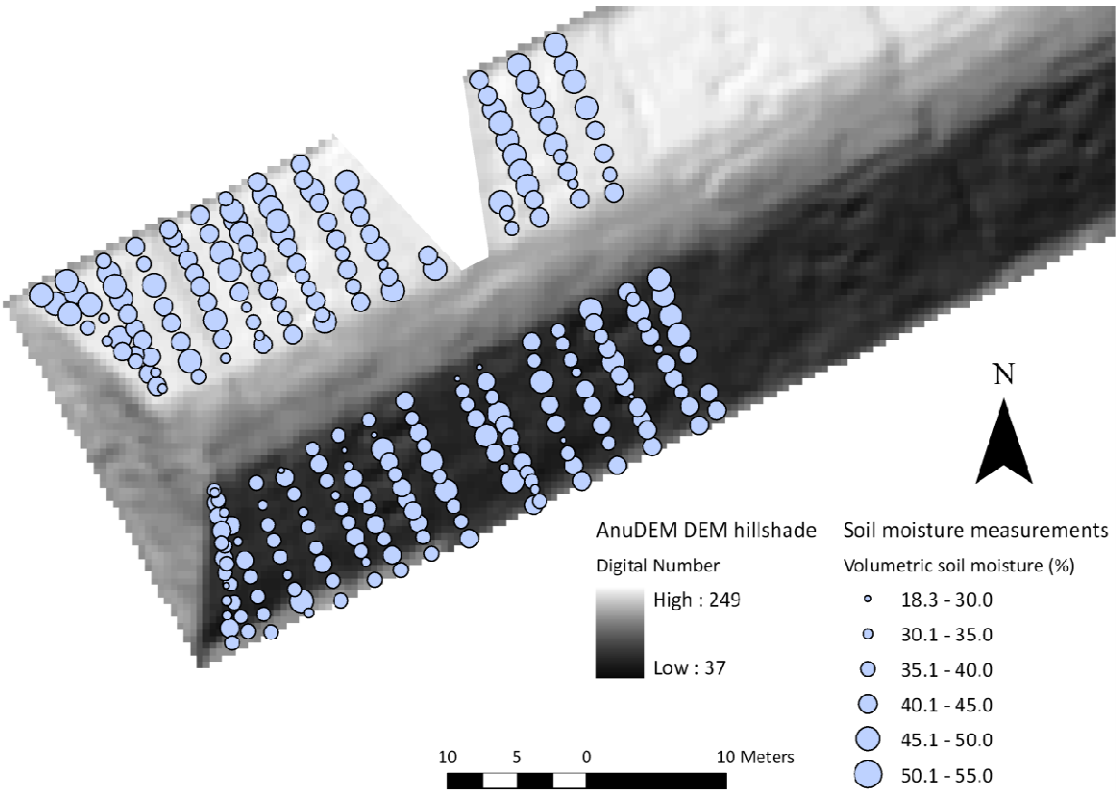


Figure 4.25 Location and values of soil moisture (% volumetric) for the test embankment. Values displayed as graduated symbols.

Figure 4.26 shows the relationship between the soil moisture measurements and elevation. This shows that overall there was a trend of increasing soil moisture

downslope, which was more prominent on the north facing slope. Although this is an expected trend, as it is assumed that moisture will migrate downslope over time, the magnitude of this trend is relatively small. The south facing slope, for example, shows an increase in soil moisture content from the top to the bottom of the slope of just 2.5%. This difference cannot be considered significant as the standard deviation for the total dataset was 5% (Table 4.1). The difference between values at the top and bottom of the north facing slope was above the standard deviation of the dataset, showing an increase of 7.5%. Despite this, overall the downslope trends were poor. The reason for the lack of trend may again be attributed to the rainfall that occurred just prior to the measurements, leaving little time for downslope movement and resulting topographic organisation of soil moisture values.

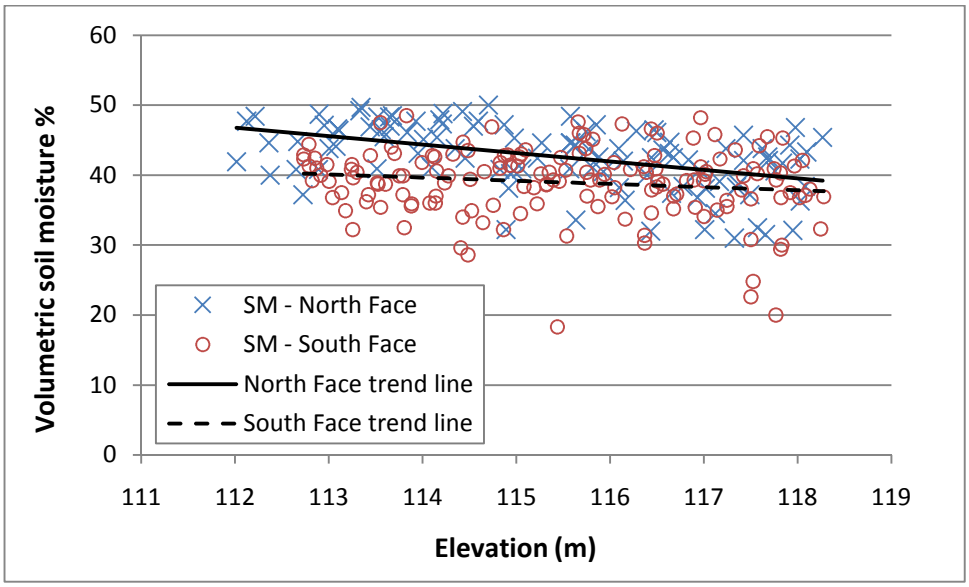


Figure 4.26 Scatter plot of elevation against soil moisture measurements for the north and south facing slopes with trend line calculated using linear regression.

4.1.8 Regression analysis between terrain analysis calculations and observed soil moisture measurements

Ordinary least squares regression between the various terrain analysis calculations and soil moisture measurements generally showed poor correlations (see Table 4.2). The TWI calculations showed almost no correlation with observed soil moisture which was not improved when the calculations were expressed as the natural logarithm. The

calculation of potential solar radiation and aspect performed better than the TWIs but they still showed poor regression coefficients with observed soil moisture, all with an R^2 below 0.16 suggesting that these metrics explain less than 16% of the variation in observed soil moisture. These findings show that individual terrain analysis calculations cannot be used to predict soil moisture over the test embankment.

Topographic metric	**R^2 correlation coefficient with soil moisture**	**Significance (p-value)**
TWI (AnuDEM DEM)	0.06	0.000
TWI (kriging DEM)	0.043	0.000
Potential solar radiation - entire year (AnuDEM DEM)	0.156	0.000
Potential solar radiation - entire year(kriging DEM)	0.155	0.000
Potential solar radiation - specific day (AnuDEM DEM)	0.149	0.000
Potential solar radiation - specific day (kriging DEM)	0.144	0.000
Aspect (AnuDEM DEM)	0.141	0.000
Aspect (kriging DEM)	0.148	0.000
Natural logarithm of TWI (AnuDEM DEM)	0.058	0.000
Natural logarithm of TWI (kriging DEM)	0.047	0.000

Table 4.2 Regression correlation coefficients for terrain calculations against soil moisture measurements.

Stepwise linear regression showed best estimations could be made using a mixture of potential solar radiation calculated using the kriging DEM and TWI calculated using the AnuDEM DEM. However, the correlation coefficient remained poor with an R^2 of 0.26 (p value < 0.000). A GWR model with a bandwidth of 5.8 m was built using these two metrics. This significantly improved the correlation with an R^2 of 0.56. The Akaike information criterion showed that a locally fitted GWR model was better suited than a global regression model. This was supported by the Monte Carlo test for significance that showed the relationship between soil moisture and the terrain calculations to vary significantly over space (significant to 0.1% level). Mapping the GWR model residuals revealed no spatial patterns which suggested that perhaps the reason for the poor correlations found were due to the lack of spatial organisation of the

observed soil moisture values. In addition, mapping the predicted values from the GWR model revealed that the independent soil moisture measurements used to build the model was dominated by the influence of potential solar radiation (Figure 4.27). This is reinforced by an average predicted soil moisture content of 43.5% for the north facing slope compared to an average of 31.7% for the south facing slope, representing a difference of 11.8%. By comparison, the average difference in the observed soil moisture measurements was just 4.1% meaning the GWR model overestimated the influence of solar radiation. This may also be due to the lack of downslope trend found in the observed soil moisture measurements meaning the TWI has reduced influence in the model. This is supported by the scatter plot of predicted soil moisture using GWR against elevation in Figure 4.28, showing that there was only a slight downslope trend in values.

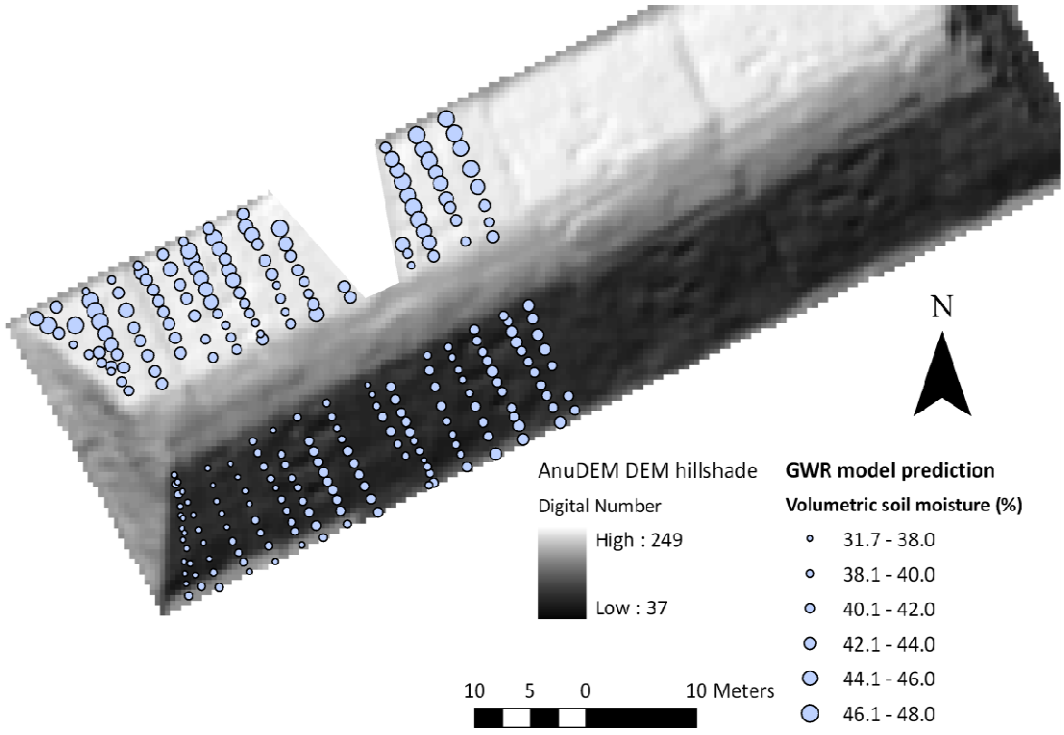


Figure 4.27 Predicted soil moisture using the TWI (calculated using the AnuDEM interpolated DEM) and potential solar radiation (using the kriging interpolated DEM) metrics in a GWR model.

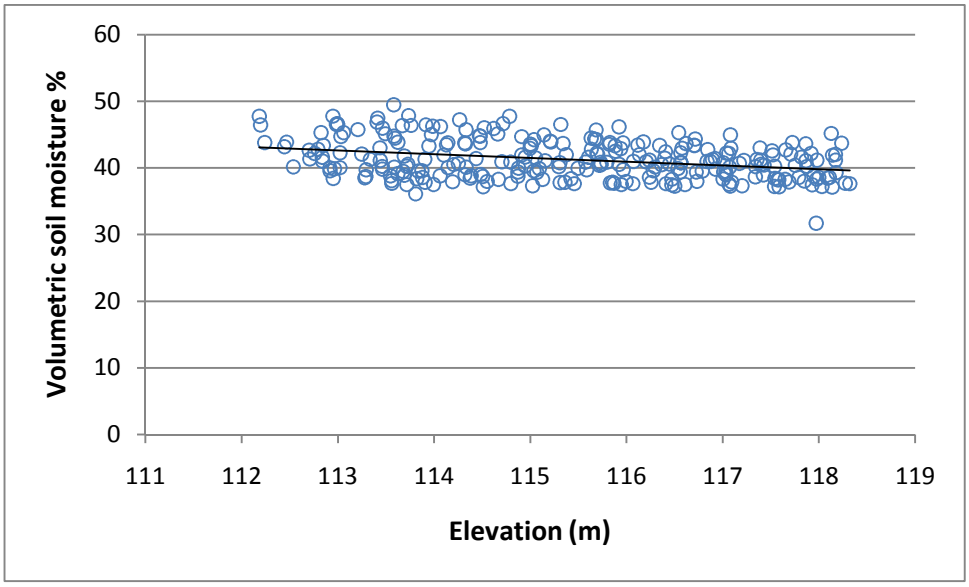


Figure 4.28 Scatter plot of predicted soil moisture values using the GWR model against elevation. Trend line was fitted using linear regression indicating a slight downslope trend in the data.

4.2 Hyperspectral response of vegetation to artificial wetting

The methods used in this experiment were detailed in Section 3.2.2. In summary, three plots were set out on the test embankment, one wetted, one covered and one left as a control, in an attempt to exhibit a range of soil moisture contents. Systematic measurements of soil moisture content and hyperspectral reflectance were made over a period of a month. Spectral analysis techniques were then used to assess the potential for using vegetation spectral reflectance data to characterise soil moisture content.

4.2.1 Soil moisture measurements

The vegetation reflectance experiment was carried out over the period 4th March to 3rd April 2009. Weather records for this period showed that rainfall gradually decreased (Figure 4.29). This meant that the soil moisture content of the embankment was naturally falling, which was reflected in the change in values for the control plot (Figure 4.29). Soil moisture also gradually decreased in the covered plot, although no more than the control plot. Rather than increasing soil moisture in the wetted plot, artificial wetting ensured soil moisture values remained above 40%.

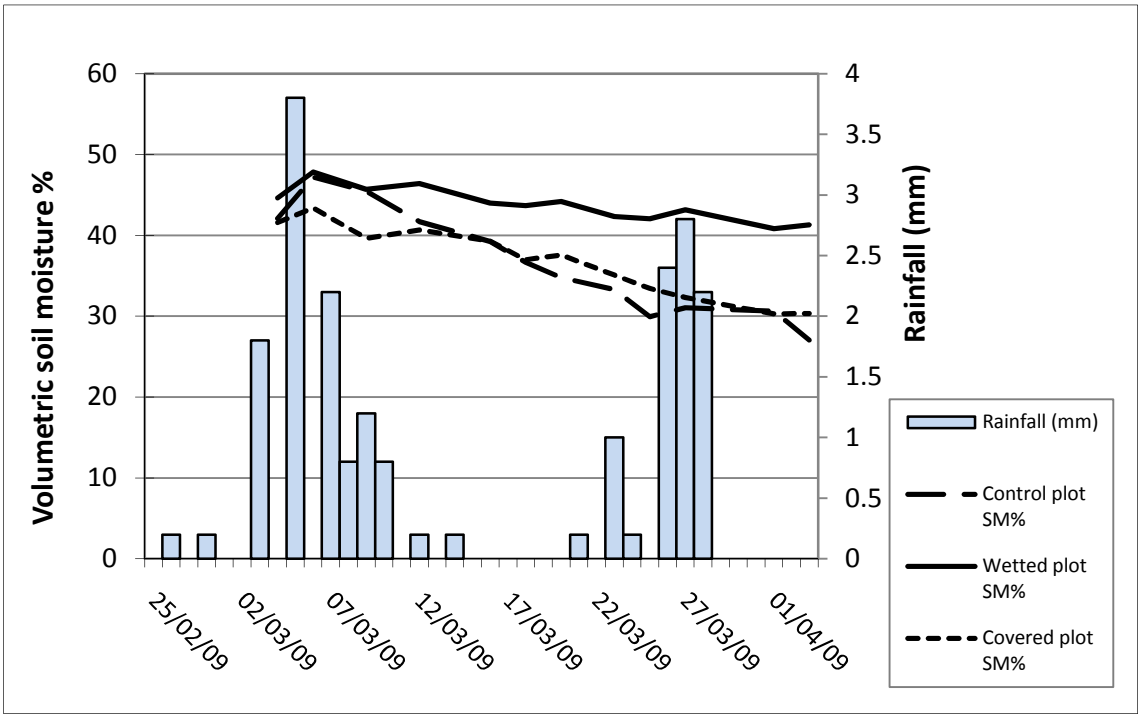


Figure 4.29 Graph showing rainfall and soil moisture content per plot over the duration of the experiment. Trend line is fitted to the rainfall data using linear regression.

T-tests demonstrated that the soil moisture values for the wetted plot were significantly different to the control and covered plots ($p < 0.005$). The control plot showed some response to rainfall events, as illustrated in Figure 4.29, with a slight increase in soil moisture content of 5% following the rainfall event on 5th March 2009. The wetted plot, in comparison, demonstrated a smaller increase of 3% in response to the same rainfall event. The covered plot registered an increase of less than 2% to the same rainfall event but showed no increase in response to any other rainfall event. This suggested that the cover was successful in limiting the influence of natural rainfall. However, a T-test showed that there was no significant difference between the soil moisture measurements for the control and covered plots ($p > 0.05$). Soil moisture content in the covered plot was expected to reduce more than the control plot due to the limitation of rainfall enforced by the cover. However, it is likely that the cover limited evapotranspiration, similar to the effects of a greenhouse, resulting in the plot retaining its moisture. This was due to inadequate ventilation, as the cover had to be tied down during periods of high wind speeds (Figure 4.30).

Figure 4.30 Photograph of the covered plot with the plastic sheeting tied down during periods of high wind.

4.2.2 Spectral analysis techniques

Analysis of the original spectra (before derivative calculation) did not reveal any obvious response to changes in soil moisture for the three experimental plots. This is illustrated in the spectral profiles in Figure 4.31 which show no particular pattern in the change of reflectance for the wetted plot over time. The derivative of the spectra indicated some signs of change, particularly over the wetted plot where the red-edge feature had weakened over the duration of the experiment, from 0.64 to 0.47 at 720 nm (Figure 4.32). The control plot also exhibited a change in the derivative of the spectra at 720 nm with a decrease from 0.78 to 0.57. The covered plot, however, did not indicate any significant difference. Each of the spectral analysis techniques showed a response to changes in soil moisture. This is illustrated by the trend lines in Figure 4.33 which show a gradual decrease in both soil moisture and the derivative stress ratio for the wetted plot. The covered plot also demonstrated a strong response to a change in soil moisture content with an increase in the stress ratio from 1.37 on 10th March to 1.65 on 3rd April.

The control plot also indicated a negative, if less pronounced, relationship with soil moisture, with an increase from 1.49 on 10th March to 1.65 on 3rd April. These negative relationships suggest that the vegetation in the control and covered plots developed over the experimental period, indicated by an increase in stress ratio. In contrast, the positive relationship found in the wetted plot suggests that vegetation is becoming stressed in response to waterlogging.

A fault in the spectral measurements made on 7th March resulted in a large decrease in reflectance values for all plots, resulting in erroneous values calculated by each of the analysis techniques (demonstrated in Figure 4.33 by a peak in the derivative stress ratio). As a result, data from day three is excluded from all further analysis.

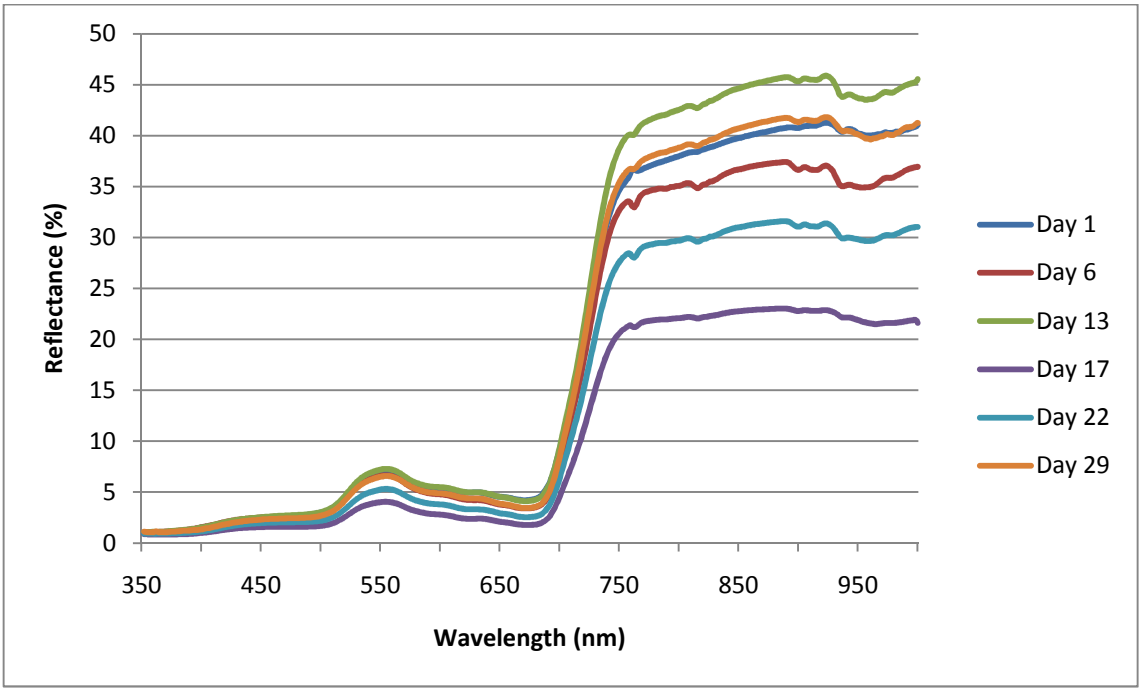


Figure 4.31 Spectral profiles for the wetted plot for six days throughout the experimental period. Profiles are shaded chronologically in order to identify any temporal trend in reflectance values.

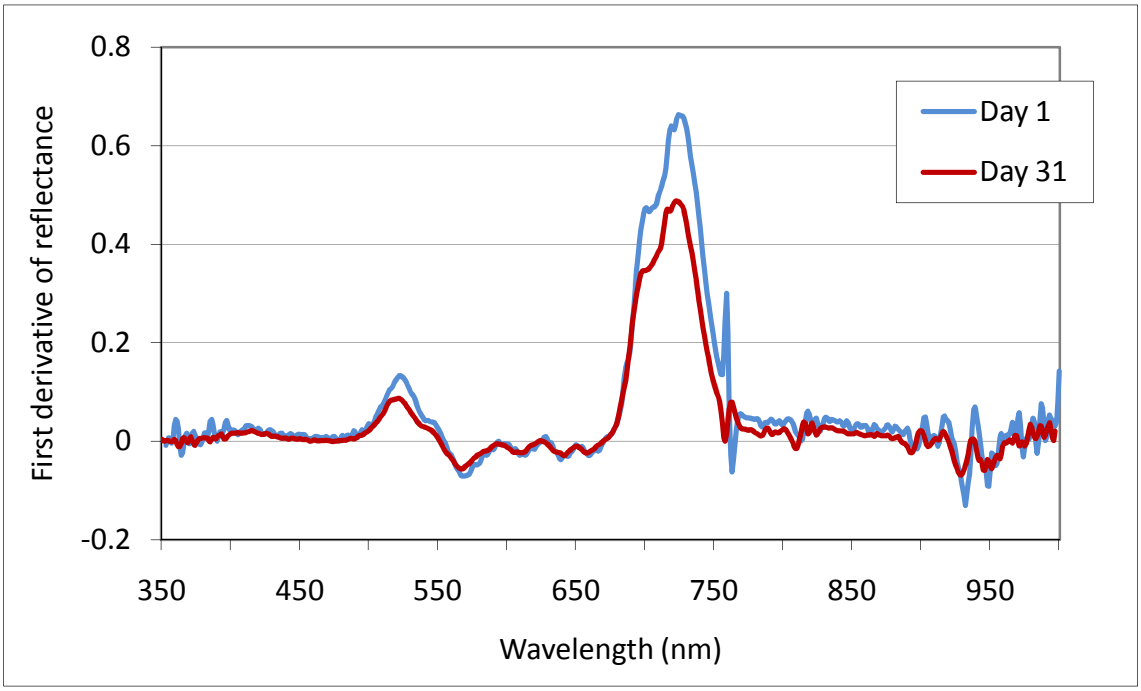


Figure 4.32 First derivative of spectral reflectance for the wetted plot at day 1 and day 31 (start and end of the wetting period).

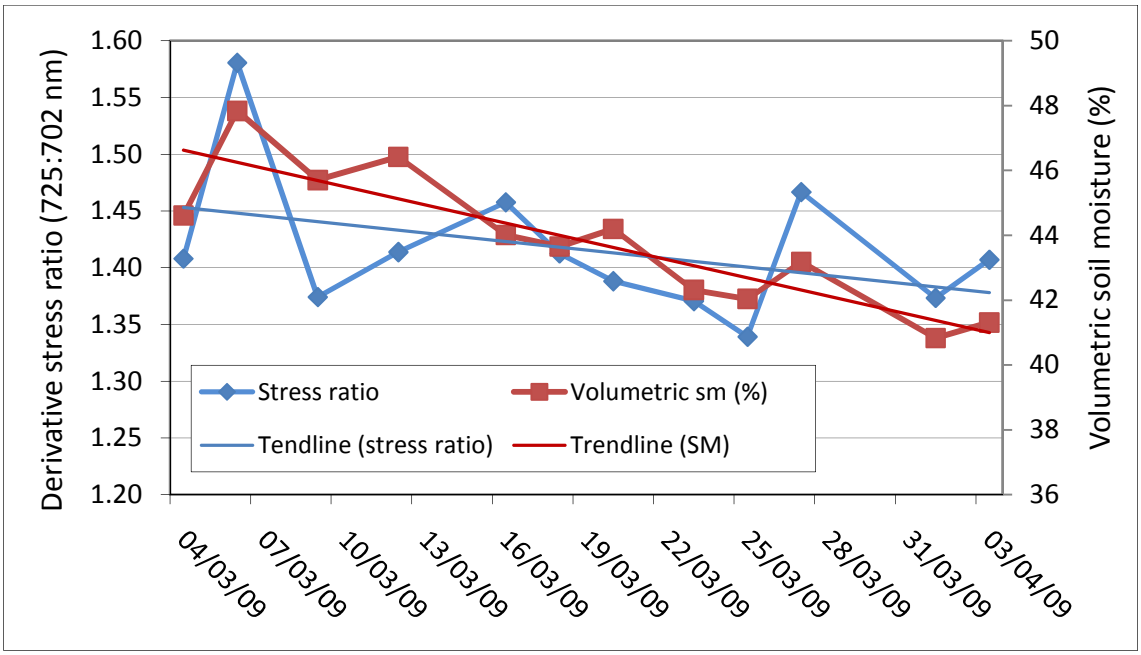


Figure 4.33 Volumetric soil moisture and 725:702nm derivative stress ratio over the duration of the experiment for the wetted plot.

4.2.3 Regression analysis

Table 4.3 shows the regression correlation coefficients for the hyperspectral analysis techniques compared against soil moisture measurements for all three plots together. The resulting regression models were also validated using an independent set of reflectance and soil moisture measurements. All correlations were significant with 99% confidence ($p < 0.000$). The results demonstrated R^2 correlation coefficients above 0.5 for the stress ratio, Lagrangian REP, linear interpolated REP and NDVI analysis techniques, showing that the metrics explained over 50% of the variation in observed soil moisture. Upon validation, the linear interpolated REP and NDVI were shown to be the most successful in predicting soil moisture, which were also amongst the simplest to compute. However, the correlations coefficients reported were all very similar. This is supported by Pearson correlation matrix shown in Table 4.4 which shows a number of the metrics to be significantly correlated (at 95% confidence), including the linear interpolated REP and NDVI.

	Stress ratio	Continuum removal (675 nm)	Lagrangian REP	Linear interpolated REP	NDVI
R^2	0.61	0.49	0.57	0.66	0.57
R^2 validation	0.52	0.51	0.55	0.60	0.58

Table 4.3 Regression correlation coefficients (R^2 values) for spectral analysis techniques using field spectroscopy data compared against soil moisture measurements for all the plots.

	Stress ratio	Continuum removal	Lagrangian REP	Linear REP
Continuum removal	-0.44*			
Lagrangian REP	0.83**	-0.29		
Linear REP	0.24	-0.69**	0.15	
NDVI	0.43*	-0.99**	0.28	0.77**

Table 4.4 Pearson correlation matrix between the spectral analysis metrics. Significance at 95% confidence indicate by *, significance at 99.5% confidence indicated by **.

4.2.4 Airborne sensor simulation

The regression analysis was repeated for the spectral analysis techniques applied to CASI simulated spectra for all three plots together. Interestingly, improvements were made to the regression correlation coefficients for nearly all the spectral analysis techniques, bar the linear interpolated REP (Table 4.5). This indicates that the CASI sensor holds sufficient spectral detail to successfully apply these hyperspectral analysis techniques. This is significant as it would mean the techniques could be scaled up to an airborne platform allowing for their application over large areas.

Overall the Lagrangian REP and derivative stress ratio analysis were shown to have the highest correlation coefficient with soil moisture using CASI simulated spectra. However, Pearson correlation analysis showed these two metrics to be significantly related (Table 4.6). Incidentally, Table 4.6 shows that all the metrics were significantly correlated with each other at a 99% confidence interval. As a result, the analysis carried out here cannot suggest with confidence which is the best technique to use for characterising soil moisture.

PLS regression was also considered for predicting soil moisture using the CASI simulated spectra. Prior to this analysis, the simulated spectra underwent a $\log_{10}(1/R)$ transformation (see Section 3.2.1). Martens' uncertainty test (Martens and Martens, 2000) showed that just two simulated CASI bands, at 901 nm and 943 nm, had any significant ability to predict observed soil moisture. These bands were placed into a PLS regression model in order to select the optimum number of components for the final model. However, results for the model calibration were poor with an R^2 of 0.48 and subsequent (leave one out) cross validation showed that the model was unstable with an R^2 of 0.37.

	Stress ratio	Continuum removal	Lagrangian REP	Linear interpolated REP	NDVI
R^2	0.66	0.6	0.68	0.64	0.57
R^2 validation	0.60	0.52	0.61	0.59	0.58

Table 4.5 Regression correlation coefficients (R^2 values) for spectral analysis techniques compared against soil moisture measurements for the CASI simulated spectra.

	Stress ratio	Continuum removal	Lagrangian REP	Linear REP
Continuum removal	-0.72			
Lagrangian REP	0.98	-0.80		
Linear REP	1.00	-0.69	0.98	
NDVI	0.78	-0.99	0.84	0.75

Table 4.6 Pearson correlation matrix between the spectral analysis metrics applied to the CASI simulated spectra. All correlations are significance at a 99% confidence interval.

PLS regression was also considered using the first derivative of the CASI simulated spectra. Martens' uncertainty test (Martens and Martens, 2000) indicated six simulated CASI bands as having significant predicting power. These bands correspond to the wavelengths 675, 700, 728, 758, 901 and 943 nm, and are highlighted in the example spectral plot shown in Figure 4.34. PLS regression analysis using these bands suggested that an optimum six components are necessary to describe the variation in both the predicting simulated CASI bands and the observed soil moisture values. The plot in Figure 4.35 shows the regression correlation coefficients behind this model selection. It indicates that the calibrated model, with six components, produced a good fit with an R^2 value of 0.82. The plot also shows that this model had the highest validated R^2 score of 0.72, indicating that it was the most stable model in the cross validation analysis. To further test the predictive ability of this model, the PLS regression model was compared against an independent set of soil moisture observations producing an R^2 of 0.72 which was significant to the 99% confidence level.

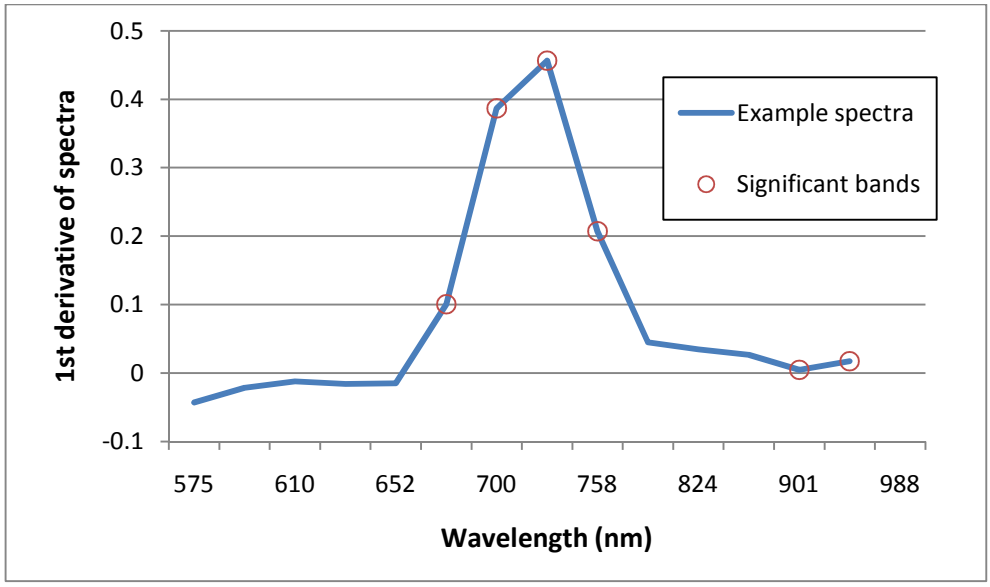


Figure 4.34 Example first derivative spectral plot for simulated CASI spectra with the location of significant predicting bands selected using Martens' uncertainty test (Martens and Martens, 2000).

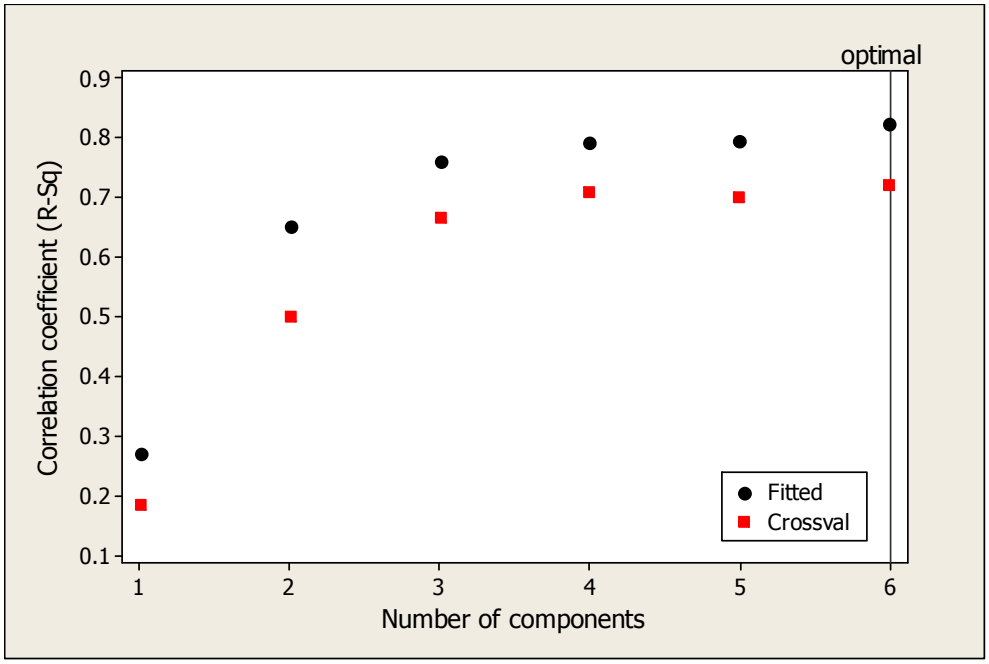


Figure 4.35 Plot showing the PLS model selection plot which includes the correlation coefficients (R^2) for the fitted and the leave one out cross validated models, with increasing numbers of components.

The spectral analysis techniques described in this section were also applied to the airborne CASI imagery (Section 3.3.2) for the entire transport corridor (Figure 3.4) and compared to the observed soil moisture values detailed in Section 3.3.1.1 and later in Section 5.1.2. In all cases, the techniques indicated no correlation with observed soil moisture. This result was expected as the heterogeneous vegetation canopy found in the Haltwhistle transport corridor is much more complex compared to the relatively homogenous vegetation of the test embankment.

4.3 Earthwork scale study conclusions

The soil moisture measurements used to test the ability of the terrain analysis calculations to predict soil moisture conditions exhibited considerable spatial variability and showed only a small degree of topographic organisation. Little downslope trend in soil moisture values was found to be present. This may have been the result of collecting soil moisture measurements shortly after a rainfall event, leaving little or no time to allow soil moisture values to become spatially organised. This only became apparent when the rainfall data was obtained. It was not possible to repeat the experiment as construction relating to another study on the embankment had begun, which changed the topography of the earthwork and, therefore, new TLS data would be required in addition to new soil moisture measurements. Future studies should consider recent weather conditions when considering the collection of validating soil moisture data. In addition, consideration could be given to the temporal variation in patterns of soil moisture distribution which may include the effects of different weather regimes and changes in topography.

Multiple linear regression using a combination of TWI and potential solar radiation produced best correlations with observed soil moisture, although coefficients remained poor with an R^2 of 0.26. GWR vastly improved estimations and revealed spatial variability in the relationship between the terrain attributes and soil moisture. The lack of trend in the residuals of this model supports the suggestion that the observed soil moisture measurements were lacking in spatial organisation.

The response of vegetation spectral reflectance to changes in soil moisture were much more encouraging. Reasonable predictions of soil moisture could be made using a range of spectral analysis techniques, including relatively simple calculations that could

be made using reflectance values from broad band sensors, namely NDVI and linear interpolated REP. Some of these techniques also showed potential for being scaled up to imagery acquired by the CASI airborne sensor. In particular the Lagrangian REP and the derivative stress ratio calculated using CASI simulated spectra performed well when compared to observed soil moisture measurements, with R^2 values of 0.68 and 0.66 respectively.

PLS regression was used to build a model incorporating six of the CASI simulated spectral bands. An R^2 correlation coefficient of 0.72 was found when comparing predicted soil moisture using the PLS generated model with observed soil moisture measurements not used to train the model.

4.4 Summary

The soil moisture measurements gathered in all the experiments reported in this chapter indicated a great deal of spatial and temporal variation, sometimes exhibiting little spatial organisation. This supports claims by other studies that soil moisture is a difficult component to quantify (Famiglietti et al., 1999; Famiglietti et al., 1998). As a result, many of the techniques used in this study have produced relatively poor correlations with observed soil moisture. Despite this, the results detailed in this chapter give an indication of which techniques work best in the context of transport corridor environments.

Differences in the DEM interpolation and subsequent TWI calculation showed that the AnuDEM interpolation routine is necessary to produce a reasonable depiction of soil moisture distribution at both the earthwork and transport corridor scales. Furthermore, the relationship between soil moisture and the TWI was shown to be non linear. Therefore, expressing the TWI as the natural logarithm improves the correlations. At an earthwork scale, a number of spectral analysis techniques were shown to have the ability to represent changes in soil wetness. Tests using reflectance from a simulated airborne sensor also indicated the potential for these analysis techniques to be scaled up. However, these techniques were only tested over homogenous vegetation cover, so there is uncertainty over their performance in heterogeneous environments, such as a transport corridor.

5 Transport corridor scale study results

This chapter details the transport corridor scale study results following the application of the methods described in Section 3.3. Similar to the results chapter for the earthwork scale study (Chapter 4) this is broken down according to the type of data being investigated. Section 5.1 covers the terrain analysis work applied to airborne lidar data and Section 5.2 details the analysis of airborne spectral reflectance data for mapping Ellenberg indicator values. At the start of each of these sections a very brief overview of the relative methods are given in order to bring the results into context. The results of integrating the terrain analysis and vegetation spectral reflectance metrics are reported in Section 5.3.

5.1 Terrain analysis

A detailed description of the methods used in terrain analysis work at a transport corridor scale can be found in Section 3.3.1. In brief, terrain analysis calculations were applied to DEMs generated from airborne lidar point elevation data using IDW and AnuDEM interpolation techniques. These terrain analysis metrics were then compared to a series of observed soil moisture datasets, which were collected under different overall soil moisture conditions or using different spatial sampling strategies. The following details the results of this work.

5.1.1 DEM interpolation

Hillshade analysis of the two interpolated DEMs showed that the IDW DEM held more fine scale topographic detail (Figure 5.1) compared to the smoother AnuDEM DEM (Figure 5.2). Examination of the elevation profiles taken from each DEM over an example cutting earthwork supported this claim (Figure 5.3). Specifically, the IDW profile showed a response to a feature about 18 m from point 1 (profile line marked in Figure 5.2) where the surface flattened out slightly, compared to the smoother surface of the AnuDEM DEM. This translated itself into the TWI as the index responded to the feature in the IDW DEM, by resetting the calculation of upslope contributing area. The result was a peak in TWI values half way down the slope, at approximately 15 m from point 1 (Figure 5.4), representing an increase in TWI value of 0.33 over a distance of

2.5 m. This would correspond to a low soil moisture value half way down the slope, where in reality it is more likely to increase further downslope. The AnuDEM TWI also showed a response to this topographic feature but in this case it was much less pronounced with an increase of just 0.02 over the same distance. This is because the AnuDEM interpolation has smoothed out the topographic feature enabling upslope contributing area to continue to increase downslope.

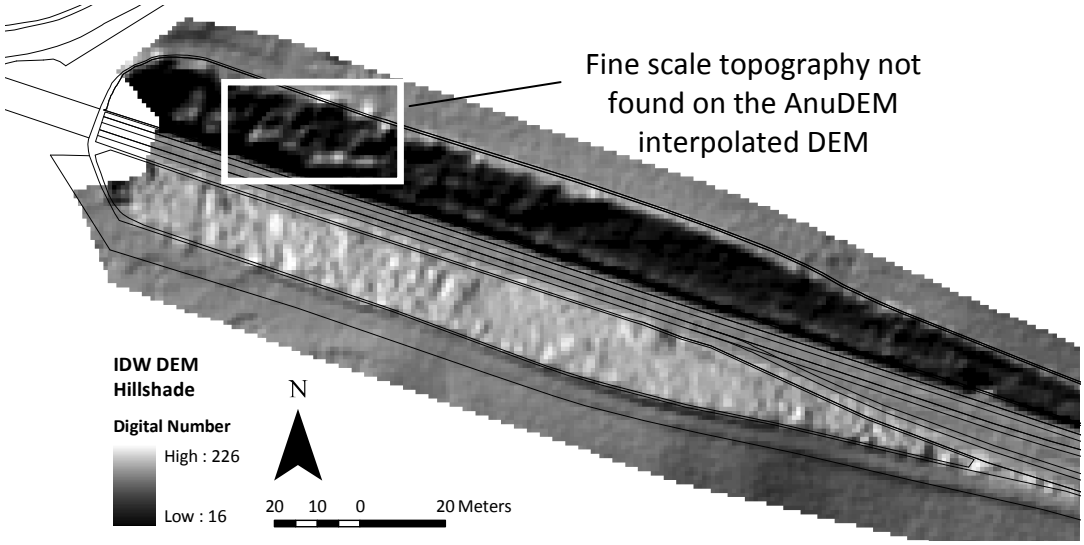


Figure 5.1 Hillshade of the IDW interpolated DEM for a cutting at Whitchester.

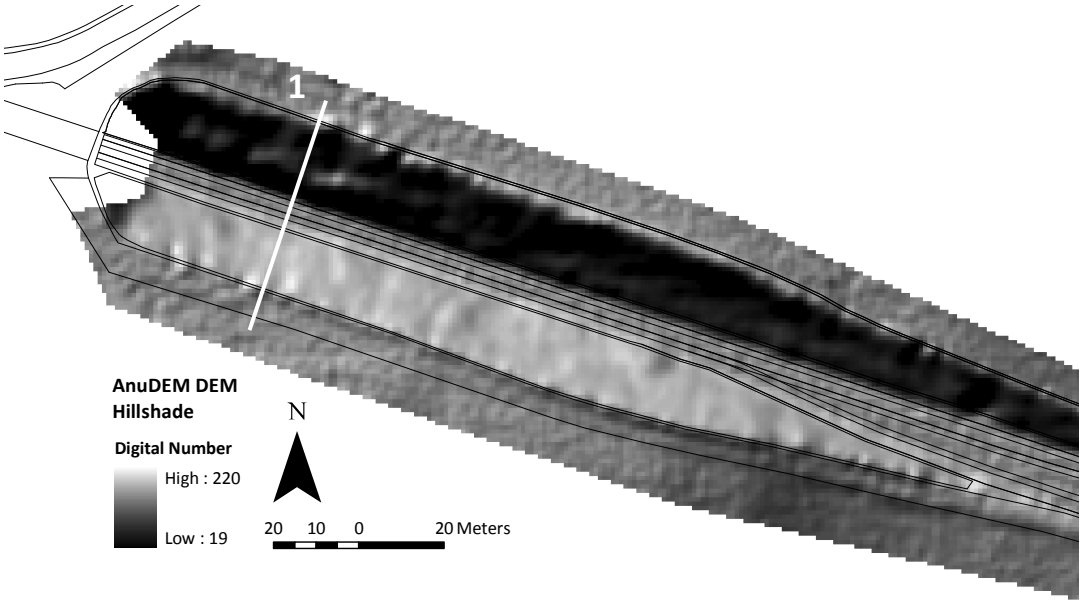


Figure 5.2 Hillshade of the AnuDEM interpolated DEM for a cutting at Whitchester with line identifying the location of the elevation profiles in Figures 5.3 and 5.4.

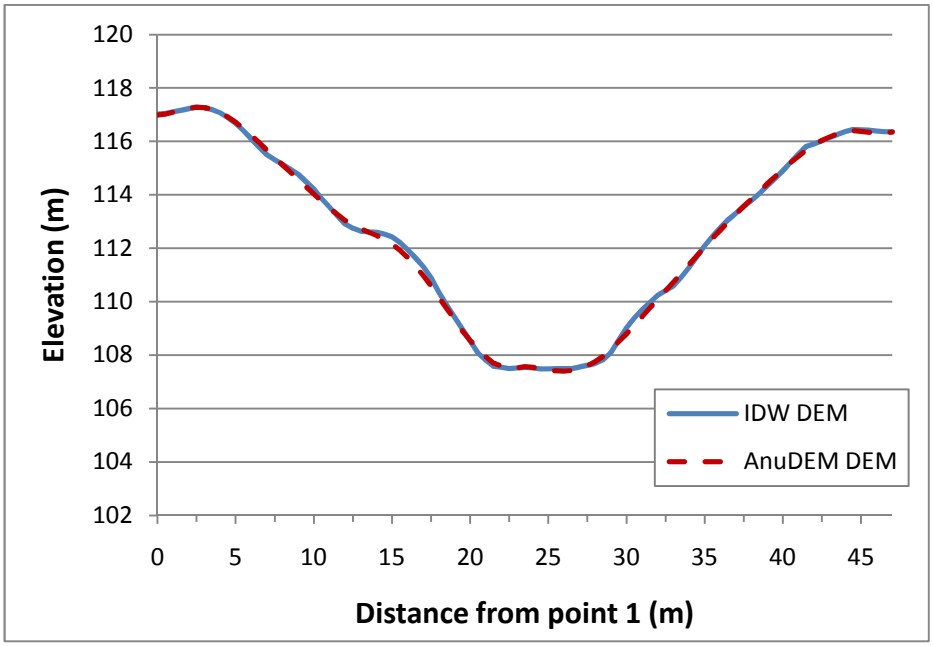


Figure 5.3 Transect of elevation values across the Whitchester cutting for the IDW and AnuDEM interpolated DEMs. Location of transect 1 is shown in Figure 5.2.

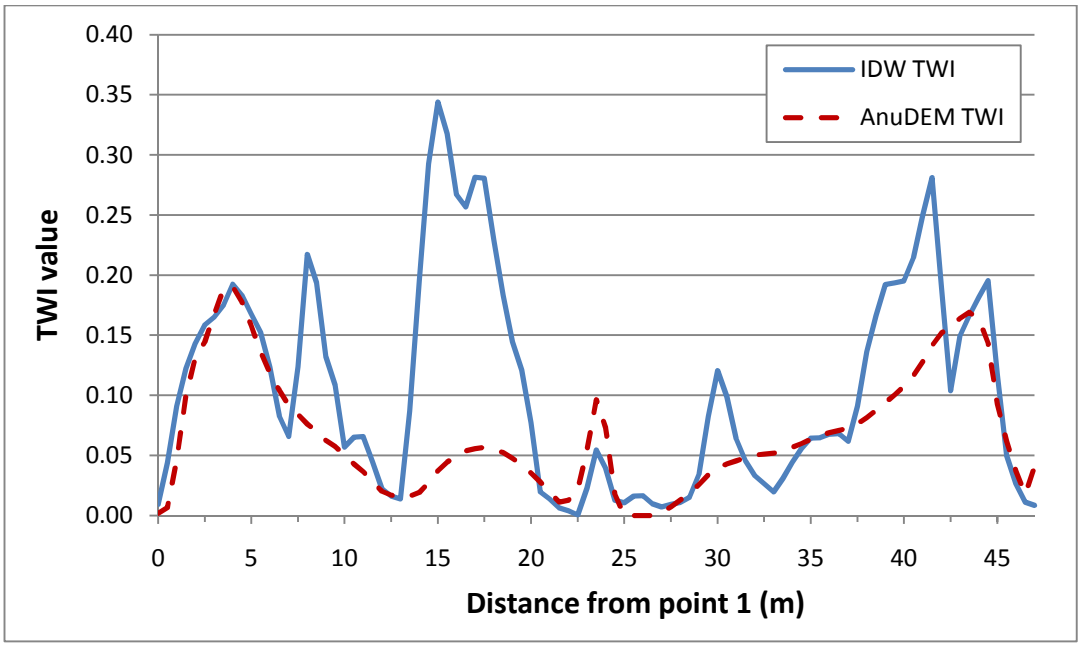


Figure 5.4 Transect of TWI values across the Whitchester cutting calculated for the IDW and AnuDEM interpolated DEMs. Location of transect 1 is shown in Figure 5.2.

5.1.2 Soil moisture measurements

Figures 5.5 - 5.7 show maps of the three soil moisture measurement campaigns. Visual analysis showed a small degree of topographic organisation on 7th July 2009 (Figure 5.5), with little discernable pattern between the north and south facing slopes. One pattern that did emerge was an increase in the soil moisture content of surrounding pasture fields and also towards the toe of the earthwork slopes. However, this was not consistent across the whole sample area. Conversely, the measurements taken on 18th June and 30th July 2009 (Figures 5.6 and 5.7 respectively) demonstrated high soil moisture content in the surrounding fields, and a general increase downslope with moisture concentrating towards the toe of the earthwork slopes. Additionally, the soil moisture readings for the north facing slopes were generally higher than the south facing slopes.

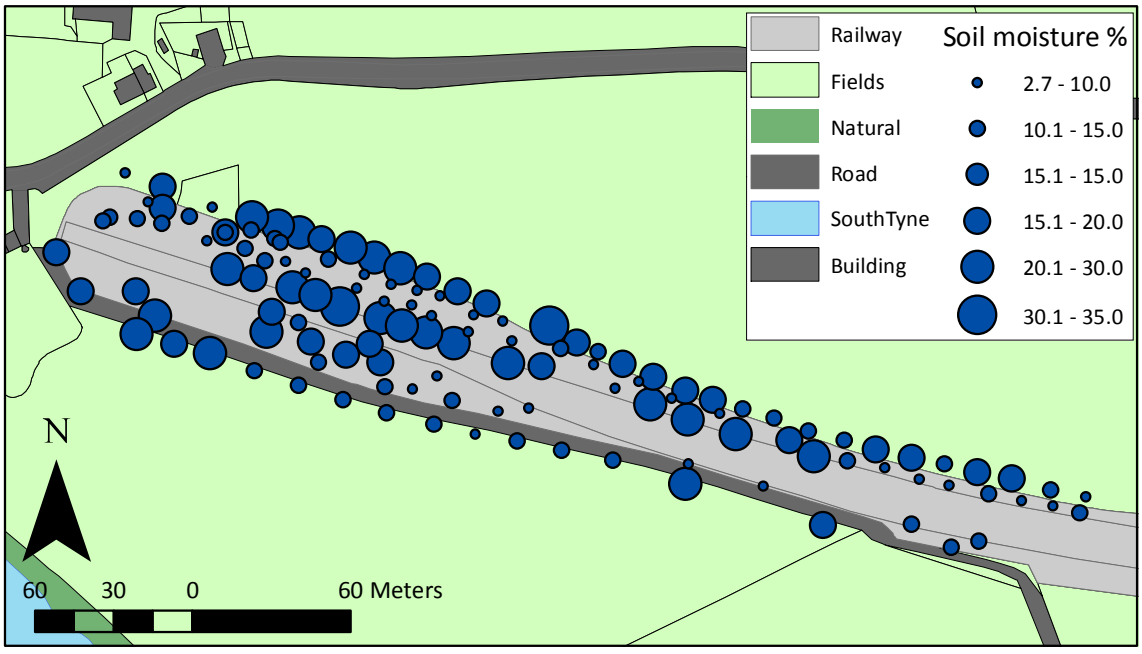


Figure 5.5 Distribution of soil moisture measurement points displayed as graduated symbols for 7th July 2009. Most of the measurements were taken within a cutting earthwork meaning the downslope direction is towards the middle of the railway feature.

Figure 5.6 Distribution of soil moisture measurement points for 18th June 2009. Smaller maps show soil moisture values displayed as graduated symbols for two selected areas. The soil moisture values for the smaller map towards the left were taken in a cutting, therefore, downslope direction is towards the middle of the railway feature.

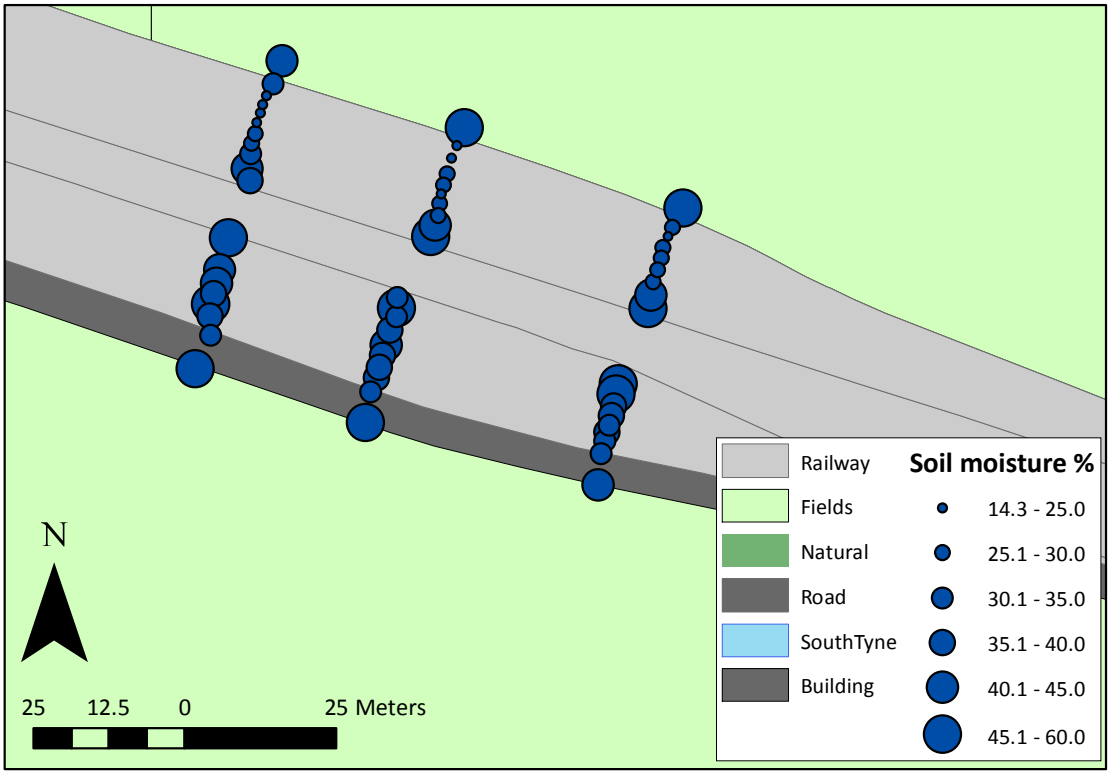


Figure 5.7 Distribution of soil moisture measurement points displayed as graduated symbols for 30th July 2009. All measurements were taken within a cutting earthwork meaning the downslope direction is towards the middle of the railway feature.

Table 5.1 shows the descriptive statistics for the three soil moisture datasets. The average soil moisture content for the measurements made on 7th July 2009 was 14.2%. This was relatively low compared to the measurements made on 18th June 2009 and 30th July 2009 with averages of 22.8% and 35% respectively. The standard error, standard deviation, and range of the datasets increased with an increase in average soil moisture. This shows that there was greater variation in the measurements made during wetter conditions.

	7th July 2009	18th June 2009	30th July 2009
Mean	14.2	22.8	34.9
Standard Error	0.6	0.9	1.3
Standard Deviation	6.3	9.1	9.6
Range	31.0	41.0	45.3
Minimum	2.7	8.2	14.3
Maximum	33.7	49.3	59.6
Count	128	96	57

Table 5.1 Descriptive statistics for the three soil moisture datasets gathered over the Haltwhistle transport corridor. Values are expressed as % volumetric soil moisture content.

	7th July 2009	18th June 2009	30th July 2009
Average	14.2	22.8	34.9
North facing slopes	15.2	25.9	39.8
South facing slopes	12.9	19.8	30.5

Table 5.2 Average soil moisture values (% volumetric) on the north and south facing slopes for the three soil moisture datasets.

The increased variation in soil moisture content experienced in wetter conditions is spatially organised. This is illustrated by the scatter plot shown in Figure 5.8 which shows the relationship between soil moisture and elevation. Specifically, the trend lines, fitted using linear regression, show increasingly steep gradients with an increase in average soil moisture conditions. This shows that during wetter conditions, the relationship between soil moisture and elevation was stronger with a gradient of -1.02 compared to -0.08 during dry conditions. In addition, all three datasets showed a variation in soil moisture content between north and south facing slopes (Table 5.2). This variation increased during higher average soil moisture conditions. For example, on 30th July when the average soil moisture content was 34.9% the difference between the average values for the north and south facing slopes was 10%. This is in contrast to the measurements made on 7th July when the average soil moisture content was 14.2% the difference was just 2.3%.

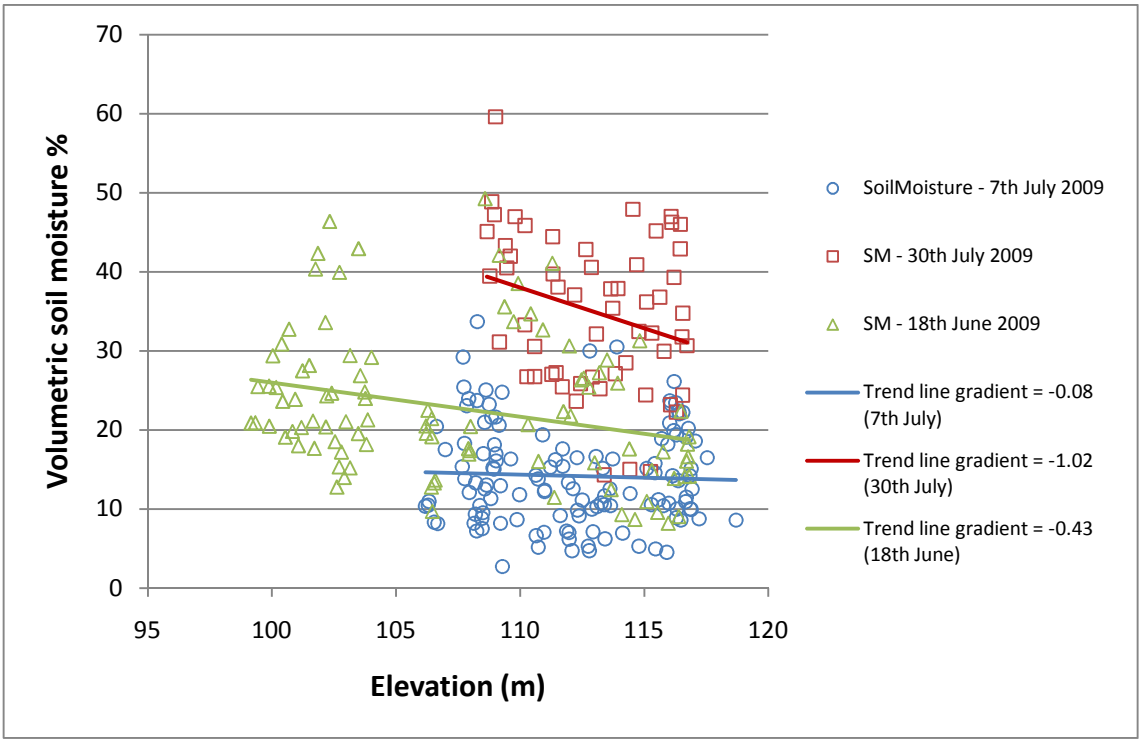


Figure 5.8 Scatter plot of soil moisture measurements against elevation for 7th July 2009 with average soil moisture content, trend line and the gradient of the trend line.

5.1.3 Terrain analysis regression results

Figure 5.8 shows that the soil moisture measurements collected on 30th July 2009 increased downslope. Therefore, it was expected that the TWI calculations would provide a reasonable predictor of this pattern. However, poor correlations were found between soil moisture and the TWI calculated using the IDW interpolated DEM, with an R^2 of 0.19. The AnuDEM TWI gave a poor, but improved, R^2 correlation coefficient of 0.23. The best correlations were found with potential solar radiation calculated using the IDW DEM, with an R^2 value of 0.27. Stepwise regression using both the AnuDEM TWI and potential solar radiation achieved an R^2 value of 0.5. All correlations were significant at a 99% confidence level.

Previous studies comparing the wetness index to soil moisture measurements have assumed linearity (Tenenbaum et al., 2006). Analysis of the soil moisture measurements revealed that in some cases, moisture tended to concentrate towards the toe of the slope rather than increase steadily downslope. This is illustrated by the

profiles in Figure 5.9 which show soil moisture measurements from the 30th July 2009 dataset. Here, soil moisture does not increase significantly until an elevation of 110 m, which occurs approximately two thirds of the way downslope.

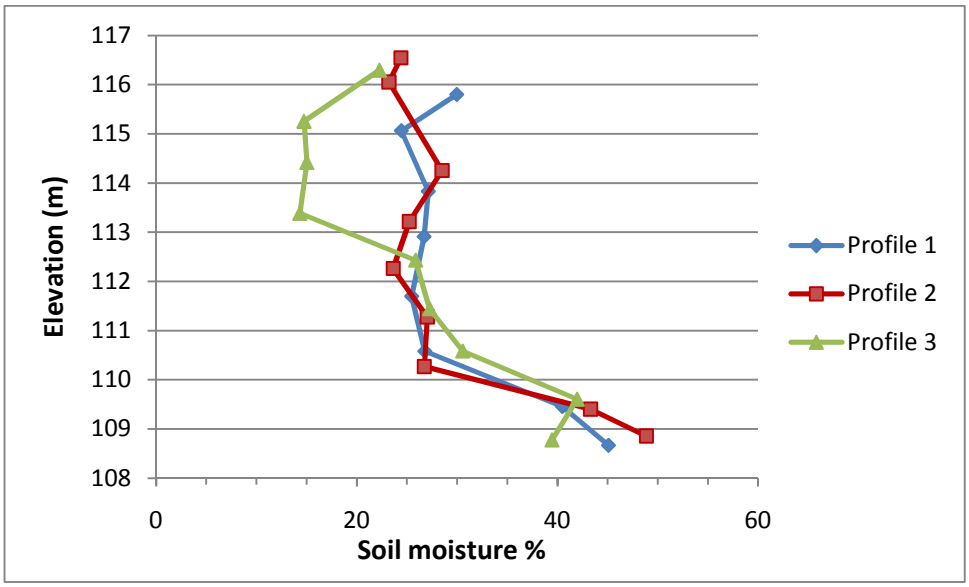


Figure 5.9 Plot showing soil moisture values (from the 30th July 2009 dataset) for three profiles running down an earthwork slope. In each case, soil moisture does not increase significantly until an elevation of approximately 110 m.

On the basis of the trends exhibited by the profiles shown in Figure 5.9 a number of non linear fitting methods were considered to help improve correlations between TWI and soil moisture. Cubic and quadratic fits to the TWI values worked best, achieving R^2 values of 0.38 and 0.31 respectively. However, these solutions introduced additional model parameters which could reduce their effectiveness in an eventual model to predict soil moisture. This is because whereas the parameters may suit the data used to form the model, they are less likely to be suited to a set of independent observations. As a result, the TWI was expressed as a natural logarithm which introduces just one additional parameter which therefore has a greater chance at being suited to an independent dataset. This improved the R^2 correlation coefficient to 0.33. Subsequent stepwise multiple regression of the natural logarithm of TWI and potential solar radiation further improved the relationship with observed soil moisture,

with an R^2 of 0.65 being achieved which was which was significant at the 99% confidence level.

GWR further improved correlations with observed soil moisture using the natural logarithm of TWI and potential solar radiation with an R^2 of 0.76 being achieved. This suggests that a GWR model better describes the variation in observed soil moisture than the global linear regression model. However, the Monte Carlo test for significance showed that potential solar radiation was the only parameter that had a significant spatially variable relationship (with 99% confidence) with soil moisture.

Analysis of the GWR model residuals showed that the model tended to over predict where observed soil moisture was high, and under predict where soil moisture was low. This is illustrated in Figure 5.10 which shows the relationship between observed soil moisture and the GWR model residuals. The model residuals revealed some under predictions towards the shoulders of the earthwork slopes. The residuals for the OLS regression between soil moisture and the natural logarithm of the TWI were polarised by the north and south facing slopes. For example, in one area the average residual for the south facing slope was 4.6 compared to -4.2 on the north facing slope, with an average residual of -0.8. This highlighted the importance of including the calculation of potential solar radiation in modelling soil moisture using terrain analysis.

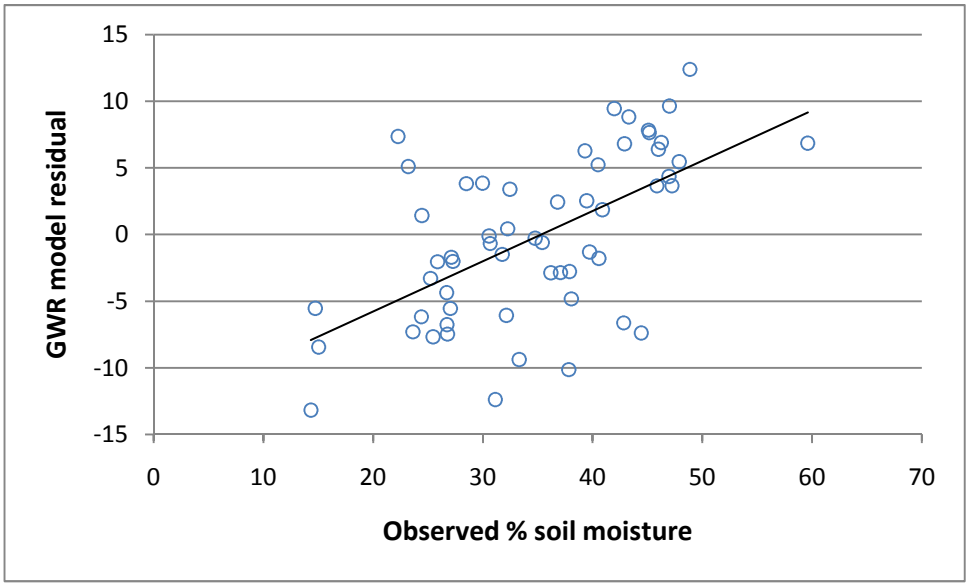


Figure 5.10 Scatter plot showing the relationship between observed soil moisture and the predictions made using GWR.

5.2 Vegetation reflectance analysis

The methods relating to this section are described in Section 3.3.2. In summary, vegetation species abundance was sampled in order to derive weighted Ellenberg indicator values across the transport corridor. PLS regression was then used to link the observed Ellenberg values with the airborne CASI multispectral data in order to map the indicator values across the study area. These were then compared to observed soil moisture measurements.

5.2.1 Observed Ellenberg indicator values

In total, 45 different species types were found at the sample point locations (see Appendix 6 for a list of these species). In comparison, a total of 36 species were found in the modified-Whittaker plot sampling used to define adequate sample size (see Section 3.3.2.1.1 for more details and Appendix 5 for a list of these species). This meant that the plots used to observe Ellenberg values were more species rich than the modified-Whittaker plot. Ideally, the modified-Whittaker plot sampling should be a perfect representation of the diversity of the total study site. However, this was difficult to predict as the number of species is not always obvious, particularly when considering indistinct grasses and small wildflowers. If the modified-Whittaker plot did accurately represent the diversity of the total plot, it is likely that the subsequent species-area curve (Figure 3.21) would suggest that a larger area should be sampled.

At each sampling point, abundance was recorded in three plots to give a measure of variation. The average standard deviation for each sampling point was 0.47 (Ellenberg moisture value). Table 5.3 shows summary statistics for the weighted average Ellenberg values. The average value was 5.63 with a positive skew towards higher values, which is illustrated in the histogram in Figure 5.11 showing the distribution of the values. This was due to an abundance of Himalayan Balsam (*Impatiens glandulifera*) at two sample locations. This species has a high Ellenberg value of eight, as it prefers wetter soils, but also tends to dominate local vegetation (Wadsworth et al., 2000). The range of values recorded across the embankment was 2.45, which is much smaller than the values reported by Schmidtlein (2005) who showed a range of 4.3 to 9.5. The latter range was found as the study site included areas of raised bog and wet fens which skewed the distribution of values away from a mean

of 6.0. Overall, higher values were expected to be found in the case of Schmidtlein (2005), as the study site was located over upland pastures with high average annual rainfall rates (2000 mm per year).

Weighted average Ellenberg indicator values	
Mean	5.63
Range	2.45
Minimum	4.82
Maximum	7.27
Count	32

Table 5.3 Summary statistics for the weighted average Ellenberg indicator values.

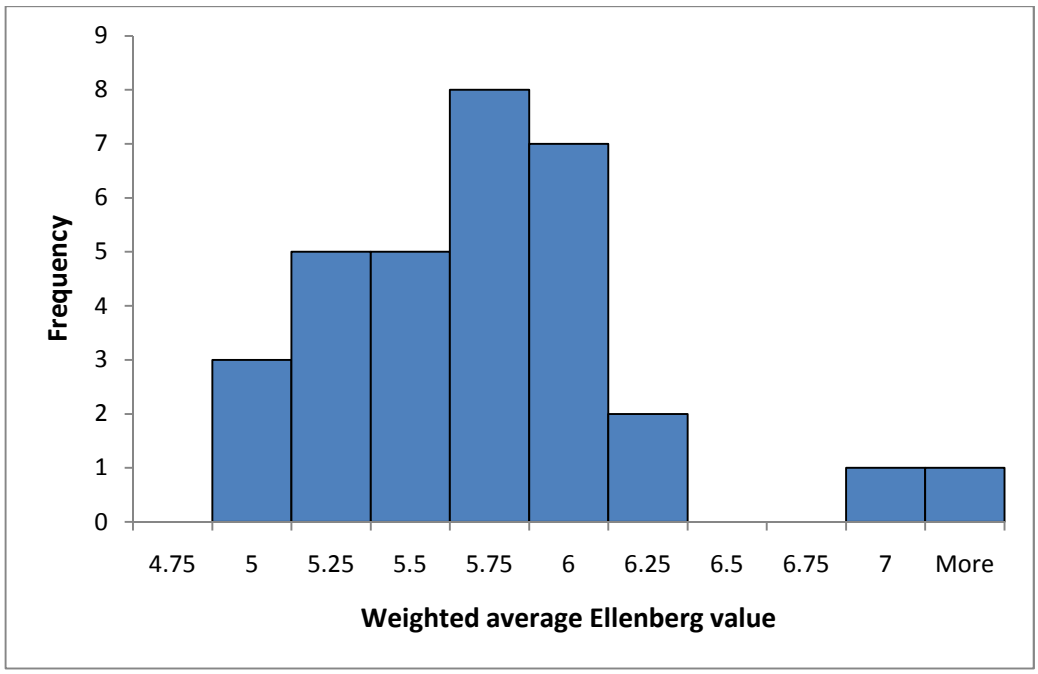


Figure 5.11 Histogram of weighted average Ellenberg indicator values.

5.2.2 Mapping Ellenberg values

The $\log_{10}(1/R)$ transformed CASI spectral bands (see Section 3.3.2 for details on this transformation) were linked to the observed Ellenberg moisture values using PLS. The Martens' uncertainty test (Martens and Martens, 2000) in the Unscrambler® software (Section 3.2.2.5) indicated that there were no significant bands, and therefore any subsequent PLS model built using these predictors would be unstable. For example,

small changes in the data, i.e. one sample being left out at a time, leads to different bands becoming important, therefore increasing the variance in the model coefficients and instability of the model. This was likely to be due to the heterogeneity of the predicting variables, leading to colinearity.

The results of the Martens' uncertainty test suggested that the CASI bands could not be used to build a reliable PLS model to predict Ellenberg values. However, Næs et al. (2002) explained that the selection of significant predicting variables could also be made using a stepwise regression routine. Forward stepwise regression, based on entering significant variables at a 95% confidence level ($p < 0.05$), was run in MatLab. This routine has the advantage of allowing the user to define a number of different variable combinations. In this way, the user can interactively view changes to the stepwise regression results as different predictors, in this case CASI bands, are removed and included, showing the change in significance of the input predicting bands. This analysis identified eight significant bands that cover the range of the CASI sensor, with six of the bands clustering around the red and near infrared reflectance and absorbance features (Figure 5.12). Using these significant predicting bands, a strong correlation coefficient can be found using ordinary least squares regression with observed Ellenberg soil moisture values (R^2 of 0.86). However, a large degree of colinearity existed between the eight predicting variables, with only bands 29 and 30 shown to have a significant difference with other predicting bands (95% confidence interval). As a result, this model would be prone to instability, and therefore, a PLS regression routine was employed.

To add statistical rigour to the stepwise regression band selection process, the predicting bands were all tested for significance using Martens' uncertainty test. The results indicated that each of the eight selected variables were significant predictors of observed Ellenberg moisture indicator values. This was because, unlike the model using all 32 bands, the model using just eight predictors was more reliable, a result of there being less autocorrelation. PLS regression using the eight significant predicting variables identified seven components as being the optimum number to describe the maximum variation in the data without becoming over fitted. This was illustrated by a correlation coefficient for the cross validation of 0.64, and the highest correlation coefficient for the original fitted model (R^2 of 0.85).

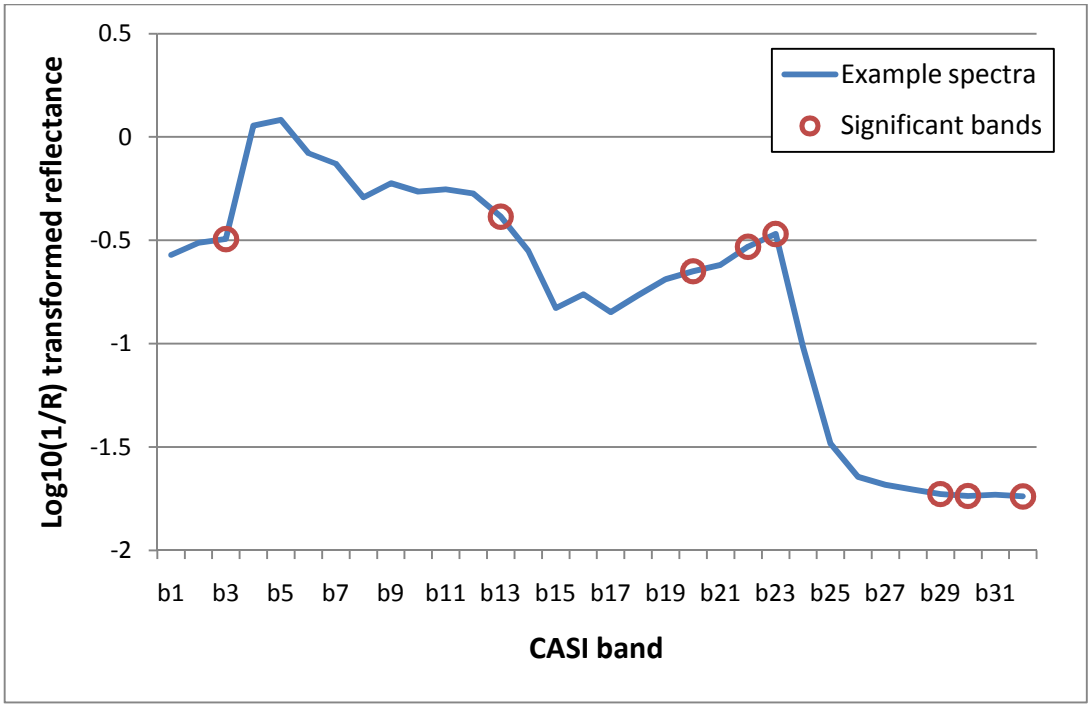


Figure 5.12 Graph showing example spectra (transformed by $Log_{10}(1/R)$) with location of significant bands chosen using forward stepwise regression.

It would have been preferable to compare the results of this model to additional Ellenberg values, rather than validating the model using cross validation. However, this would have required additional vegetation abundance data. This was not possible as validating soil moisture data has to be sampled within a day, ideally within the space of just a few hours, due to the temporally variable nature of this parameter (Famiglietti et al., 1998). This therefore restricted the number of sites where vegetation species abundance could be sampled.

As discussed in Section 3.3.2.1, Ellenberg moisture values were not mapped for areas with tree canopies due to the complexities of canopy bidirectional reflectance (Wolf et al., 2010), as well as issues with gathering the observed Ellenberg values. In addition, it was decided that mapped values should be restricted to the range of observed values used to build the PLS model, as predictions outside this range are likely to be unreliable. As a result, the map of modelled values in Figure 5.13 indicated large areas of no data values, mainly representing woodland or manmade features, and speckles of no data values depicting values predicted outside the observed range.

The resulting map of predicted Ellenberg moisture values (Figure 5.13) potentially reveals a great deal more information, with regard to soil moisture, than the terrain analysis calculations discussed in Section 5.1. In comparison to the uniform distribution of the terrain analysis results, the mapped Ellenberg values indicated variability across the earthworks. Figure 5.13 illustrates this with higher values predicted towards the west of the cutting earthwork. These scores represented an abundance of Stinging Nettles (*Urtica dioica*) and Brambles (*Rubus fruticosus*) that have a relatively high Ellenberg moisture values of six. In contrast, relatively low values were predicted for the east of the earthwork. This represented a mixture of grasses and wildflowers, including Heath Bedstraw (*Galium sterneri*) and Common Ragwort (*Senecio jacobaea*), that have lower Ellenberg moisture values of four. In addition, some areas showed an increase in Ellenberg values down the slope of the earthwork cutting. For example, transect 3 shows values increasing from 5.2 at the shoulder of the slope to 6.1 at the toe (Figure 5.14).

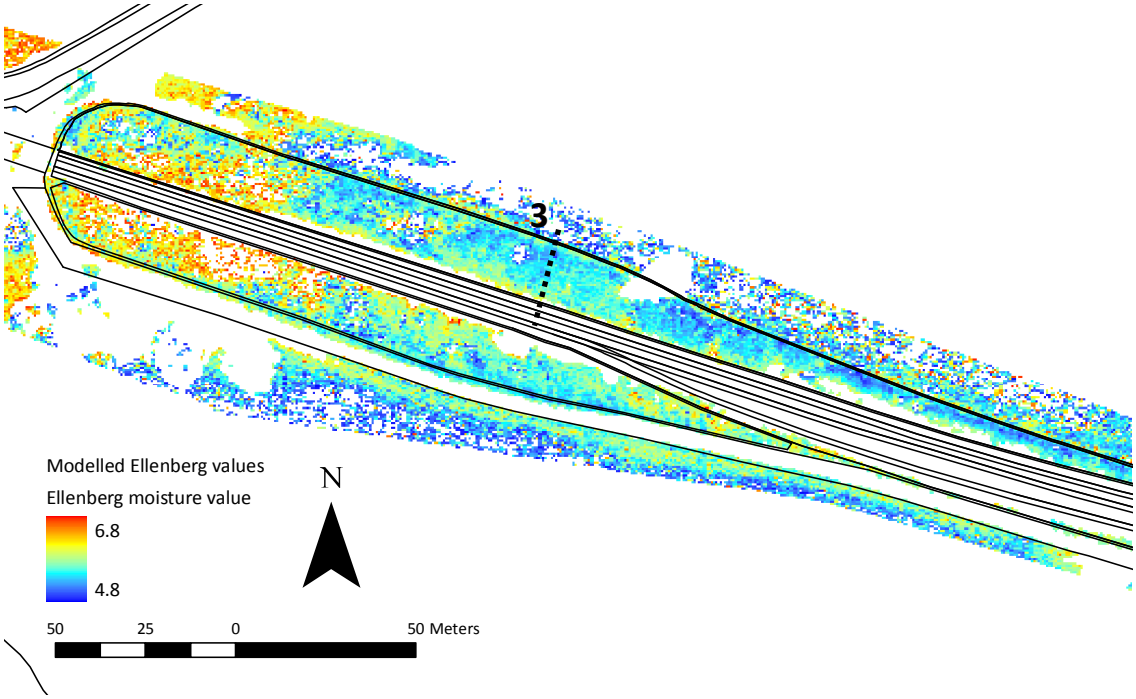


Figure 5.13 Mapped Ellenberg moisture values for a cutting earthwork. Mapped values are restricted to mask out tree canopies and manmade surfaces, which are represented by the clear (white) areas. Transect 3 refers to the location of the profile drawn in Figure 5.14.

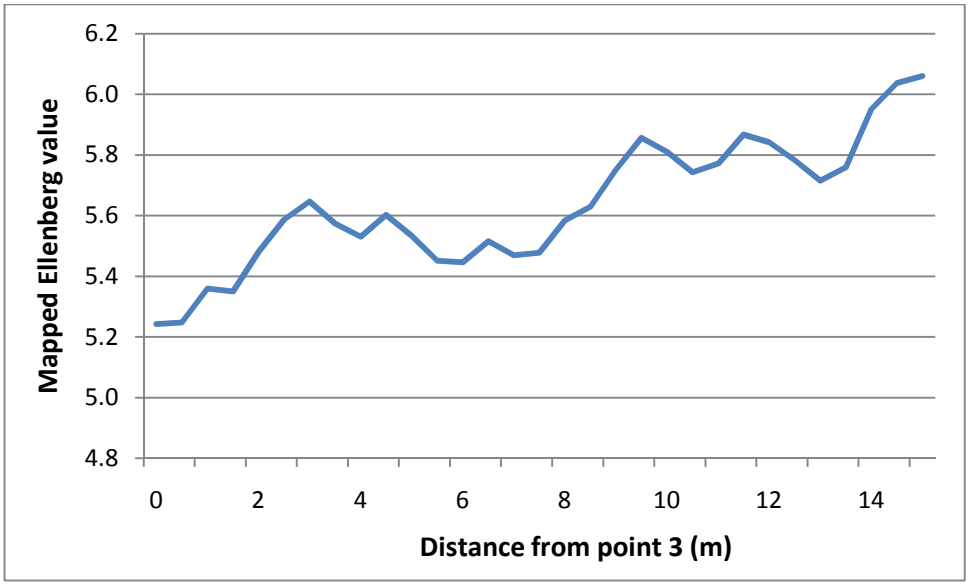


Figure 5.14 Profile of mapped Ellenberg values for transect 3 running down the slope of the cutting earthwork. The location of the transect is found in Figure 5.13.

The mapped Ellenberg moisture values were compared against soil moisture measurements gathered coincidentally with the species abundance data. This data was collected on 13th August 2009 during relatively dry conditions resulting in an average soil moisture content of 14.5%. Previous work in this Chapter (Section 5.1.2) showed greater spatial organisation of soil moisture during higher overall wetness conditions. Therefore, the measurements made on 13th August were not expected to exhibit a high degree of spatial organisation. This was demonstrated by comparing the measurements against relative elevation values (Figure 5.15). Overall there is a trend of increased soil moisture with a reduction in elevation, however, the data depicts a wide spread from the trend line, in some cases exceeding 20% (volumetric soil moisture). In addition, there was little difference between north and south facing slopes with average soil moisture contents of 14.7% and 14.2% respectively which resulted in a T-test result of no significant difference between the two sample means (p value > 0.05).

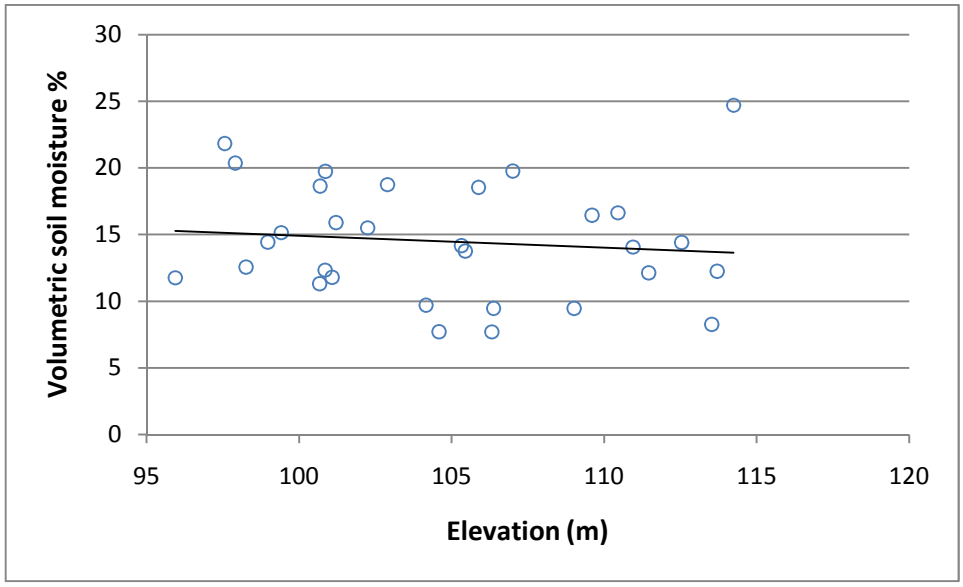


Figure 5.15 Scatter plot showing the relationship between soil moisture and elevation for the measurements made on 13th August 2009. Trend line fitted using linear regression.

The resulting R^2 correlation coefficient between the Ellenberg values and observed soil moisture was 0.4 (significant at the 99% confidence level). Despite not being a strong correlation it does represent a marked improvement on the results found in the terrain analysis work (Section 5.1.3), that found individual correlations up to an R^2 of just 0.33. In addition, this correlation was higher than expected in light of the poor spatial organisation of this soil moisture dataset as explained above. Despite this, the results obtained here cannot be used to reliably suggest that soil moisture can be characterised by mapping Ellenberg indicator values alone.

5.3 Integrated model

The results of comparing both the terrain analysis metrics (Section 5.1.3) and the mapped Ellenberg indicator values (Section 5.2.2) showed that no single metric could provide a reliable estimation of soil moisture. As a result, an integrated model was developed which could exploit information regarding topographic control on soil moisture and the information provided by species abundance. The terrain analysis metrics used in this model were determined by those that gave the best correspondence with observed soil moisture, namely, potential solar radiation calculated using the

kriging DEM and the natural logarithm of the TWI calculated using the AnuDEM DEM.

The terrain analysis metrics were integrated with the mapped Ellenberg values using OLS multiple regression, using observed soil moisture measurements from 30th July 2009. This resulted in an R^2 correlation coefficient of 0.67. This model was used to map soil moisture across the Haltwhistle transport corridor, with an example shown in Figure 5.16. Predicted soil moisture from the resulting map was validated against additional soil moisture measurements, gathered on 18th June 2009, giving an R^2 of 0.48. The model was further validated using cross validation. This procedure demonstrated a RMSE (root mean square error) in volumetric soil moisture of 5.5% for the original model calibration and 5.9% for the cross validation, suggesting that the original model error does not change significantly when applied to new observations.

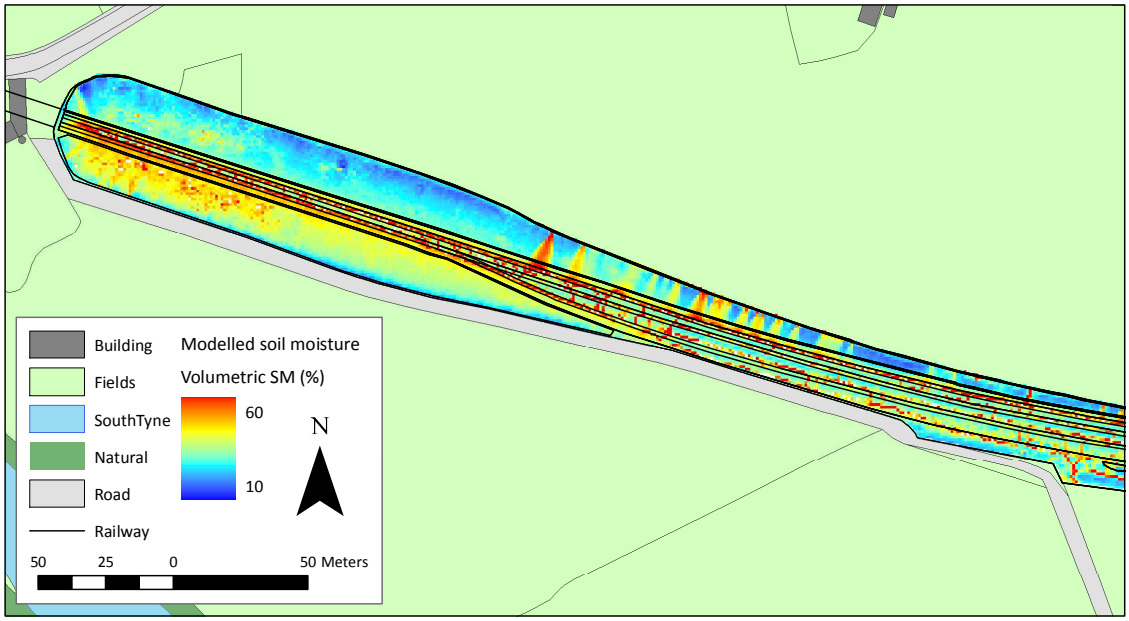


Figure 5.16 Modelled soil moisture using OLS multiple regression model integrating the natural logarithm of the TWI, potential solar radiation, and mapped Ellenberg moisture indicator values, for a cutting earthwork.

5.4 Transport corridor scale study conclusions

The observed soil moisture measurements collected over the transport corridor demonstrated two main patterns. The first was an increase in wetness downslope. The second was an increase in wetness on north facing slopes. These trends were shown to be stronger with higher average soil moisture contents.

Comparisons made between the remotely sensed metrics and observed soil moisture showed that no single metric could be used to characterise soil moisture. This was illustrated by generally poor correlation coefficients (less than R^2 of 0.5) . Despite being poor, they were improvements on the correlations found in the earthwork scale experiments (reported in Section 4.1.8). This was mainly due to the increase in spatial organisation of the observed soil moisture data used to draw comparisons. In addition, the transport corridor provided more variance in terrain compared to the relatively uniform relief of the test embankment used in the earthwork scale study. Observed soil moisture was shown to increase in a non-linear fashion downslope, with moisture tending to gather at the toe of earthwork slopes. As a result, the correlation between TWI and observed soil moisture was found to improve when the index was expressed as the natural logarithm.

The vegetation reflectance analysis work showed that Ellenberg indicator values can be successfully mapped over a UK transport corridor, with a cross validation R^2 correlation coefficient of 0.64. Comparisons with observed soil moisture measurements showed this metric to be a better individual predictor of soil moisture than the terrain analysis metrics, with an R^2 of 0.40. An integrated model, using mapped Ellenberg values, TWI calculated using the AnuDEM and potential solar radiation calculated using IDW DEM, was shown to provide the best estimate of soil moisture, with an R^2 of 0.67. The resulting map of predicted soil moisture distribution showed influence from both terrain analysis calculations, including a predicted increase in wetness on the north facing slope, increase in wetness downslope and in areas of contribution from neighbouring fields, as well as more subtle influences of the Ellenberg mapping, with an increase in wetness over water loving species. Validation of this model using cross validation showed a reasonable RMSE of 5.89 volumetric soil moisture (%). Validation using additional soil moisture measurements produced an R^2 of 0.48.

5.5 Summary

The soil moisture measurements gathered over the transport corridor indicated a great deal of spatial and temporal variation, sometimes exhibiting little spatial organisation. This supports claims by other studies that soil moisture is a difficult component to quantify (Famiglietti et al., 1999; Famiglietti et al., 1998). As a result, many of the techniques used in this study have produced relatively poor correlations with observed soil moisture. Despite this, the results detailed in this chapter give an indication of which techniques work best in the context of transport corridor environments.

Differences in the DEM interpolation and subsequent TWI calculation showed that the AnuDEM interpolation routine is necessary to produce a reasonable depiction of soil moisture distribution at the transport corridor scale. Furthermore, the relationship between soil moisture and the TWI was shown to be non linear. Therefore, expressing the TWI as the natural logarithm improves the correlations.

The results for both the terrain analysis and mapping of Ellenberg values showed no single metric to be a strong predictor of observed soil moisture, although mapped Ellenberg moisture values were shown to have the strongest correlation. In order to achieve a reasonable characterisation of soil moisture distribution, the metrics had to be integrated. Best results were found using a multiple regression model with the natural logarithm of the TWI (using the AnuDEM DEM), potential solar radiation (using the IDW DEM), and mapped Ellenberg values as the predictors, with an R^2 of 0.67. However, a reasonable prediction can be made using just the terrain analysis, with an R^2 of 0.65,which is limited to just one source of remotely sensed data.

The results reported in this chapter have a number of implications both in terms of scientific development of the techniques themselves, as well as what the results might mean for transport corridor earthwork stability. These issues are discussed in Chapter 6.

6 Analysis and discussion

This chapter provides an analysis and discussion of the results presented in Chapters 4 and 5. This is divided into two main sections relating to the work carried out at the earthwork and transport corridor scales respectively, with a summary section discussing the implications of the results found in this study as a whole.

6.1 Earthwork scale study

The results from this body of work, carried out over the test embankment, is analysed and discussed in the subsections below. These subsections (Sections 6.1.1 and 6.1.2 respectively) relate to the two distinct experiments carried out at this scale, terrain analysis and hyperspectral analysis of vegetation.

6.1.1 Terrain analysis experiment

The terrain analysis calculations applied to DEMs interpolated from TLS data revealed a number of important findings that could be used to develop an optimal strategy for implementing the terrain analysis calculations in transport corridor environments. These findings are discussed below.

6.1.1.1 Impact of DEM interpolation

The main topographic features present on the test embankment were membranes dividing each panel section and embedded instruments used to take geotechnical measurements. A DEM interpolated from the TLS point data using kriging was better at representing these features than the AnuDEM interpolated DEM. The kriging interpolation technique also revealed a greater degree of variation as a response to small scale changes in terrain (Section 4.1.1). This was due to the smoothing effect that is enforced by the AnuDEM interpolation technique. As a result, kriging interpolation is more useful for depicting fine scale terrain.

Although the fine scale detail identified in the kriging DEM makes it an ideal candidate for general monitoring of earthwork assets, monitoring deformation for example (Miller et al., 2008), it makes it less useful for characterising soil moisture using terrain analysis. Specifically, TWI values showed little downslope trend, with

seemingly erratic responses to fine scale topographic features using the kriging DEM. In contrast, TWI values calculated using the AnuDEM interpolated DEM was smoother and showed values decreasing gradually downslope, which represents a gradual increase in soil moisture. This is much closer to what is expected in reality implying that the AnuDEM interpolation is better suited for the calculation of TWI. This was an important finding as some studies, such as Schmidt and Persson (2003), used a kriging interpolated DEM for the calculation of TWI. However, these findings were made using terrestrial laser scanning data at a fine scale. Therefore, TWI calculation may not be as sensitive to DEM interpolation using broader scale DEMs that do not represent fine scale topographic features.

The calculation of potential solar radiation was also shown to be sensitive to DEM interpolation (Section 4.1.3). Although both the kriging and AnuDEM interpolated DEMs demonstrated a response in potential solar radiation values around the panel membrane features and embedded instruments, this response was more prominent using the kriging DEM. Again, this was due to the AnuDEM method suppressing fine scale topographic features, restricting their impact on the calculation of potential solar radiation. Therefore, a DEM created using kriging interpolation is better suited to representing the effects of terrain on potential solar radiation and subsequent control on soil moisture distribution. There have not been any reported attempts to compare the impact of DEM interpolation on the calculation of potential solar radiation proposed by Fu and Rich (1999). Although the differences found here have shown to be relatively small compared to other terrain analysis calculations (TWI), this study demonstrates the importance of considering which DEM interpolation method to use, which is rarely discussed in the literature.

The calculation of potential solar radiation made over an entire year showed greater variation in response to topographic features compared to the calculation made for a specific day, which showed variation only in response to the north and south facing slope aspects. It is expected that fine scale topographic features would have an influence on soil moisture distribution, following studies such as Rajkai and Rydén (1992) that showed features such as track marks, with similar spatial dimensions to the panel membranes found at the test embankment, exhibited a strong influence on the distribution of soil moisture. Based on this assumption, potential solar radiation

calculated for an entire year may provide a better characterisation of the variation in soil moisture. Resulting correlation coefficients were found to be higher for the annual calculation (0.16) than the specific day calculation (0.15). However, this difference is not significant and any conclusions based on these results cannot be stated with much confidence.

This work has shown that the individual terrain analysis calculations are better suited to DEMs created using specific interpolation techniques. However, it is not common practice to apply multiple interpolation routines to the same data due to issues concerning processing time, data storage, or simply that operators have trust in the interpolation method being adopted. This study shows that users need to consider which interpolation routine best suits their needs. It has been shown that whereas kriging provides a better depiction of terrain and subsequent potential solar radiation calculation, AnuDEM interpolation should be used when calculating TWI. In terms of expected soil moisture distribution patterns, the recommendation is that two different interpolation methods may be necessary to make an accurate characterisation using the two different terrain analysis calculations. The interpolation technique was shown to affect the calculation of TWI greater than potential solar radiation. Therefore, if one interpolation technique is to be chosen to reduce processing and storage demands then the AnuDEM routine should be used.

Kriging was applied in this study using a search radius of 12 points, to ensure small scale variation are taken into account without becoming too computationally intensive (Section 3.2.1). Increasing this search radius would suppress the impact of finer scale topographic features, similar to the AnuDEM method. However, the latter routine also has the advantage of ensuring calculations of flow accumulation increase downslope and is therefore a more accurate representation of reality (Hutchinson, 2008).

6.1.1.2 Impact of earthwork orientation on potential solar radiation calculation

Once the earthwork was rotated to produce a north to south orientation, the calculation of potential solar radiation demonstrated a greater response to small scale terrain features than the original earthwork orientation (Section 4.1.5). This showed that variation in solar radiation calculated for a north to south orientated earthwork was

controlled by the presence of topographic features on the slopes rather than a difference in slope aspect, as seen with the original west to east orientated earthworks. This was illustrated by a variation in estimated potential solar radiation over the panel membrane features representing 85% of the total range in values for the north to south orientated earthwork. In comparison, the maximum variation over the same feature as a proportion to the total range represented just 44% for the original west to east orientated earthwork.

This test was only carried out on a hypothetical basis in a GIS environment and should be tested using a real north to south orientated earthwork using soil moisture observations. This was not considered when choosing the study site because the test embankment used in this study is the only one of its kind in the UK, leaving no alternative. A real world earthwork could have been used but the test embankment had the advantage of unlimited access, improved safety, as well as availability of complementary measurements, such as weather data.

6.1.1.3 Impact of DEM resolution on the terrain analysis calculations

The terrain calculations discussed in this section have all referred to fine scale topographic features, such as the panel membranes and embedded instruments. However, when the DEMs were degraded to a grid resolution of 2 m, this information was lost. Despite this, the overall terrain analysis patterns that were exhibited in the finer resolution work was retained. For instance, TWI values were shown to decrease downslope, potential solar radiation and aspect were polarised by the north and south facing slopes. Therefore, if a DEM was generated for an entire transport corridor at this resolution, only broad scale patterns of the influence of terrain on soil moisture distribution could be represented. Features smaller than this scale would have little or no influence on resulting terrain analysis calculations. This could be significant, as earthwork assets, such as drainage ditches, have shown to have a strong influence on soil moisture content (Perry et al., 2003b). Despite the loss of detail, this coarser resolution has the advantage of requiring less computer storage and processing time to calculate the terrain analysis techniques and therefore has the potential to be applied to a much greater area.

A compromise between the fine scale detail depicted in the original resolution DEM and the 2 m DEM could be found when the grid resolution was degraded to 1 m. Here, some influence from the small scale topographic features was retained which is

illustrated by the hillshade analysis in Figure 4.20 (Section 4.1.6) but due to the coarser resolution, was relatively quick to compute compared to the original scale DEM. This could be significant when considering the practicalities of processing data for an entire transport corridor.

6.1.1.4 Comparison between observed soil moisture measurements and terrain analysis calculations

The soil moisture measurements, gathered to test the predictability of the terrain analysis calculations, were taken on 28th November 2008. This time of year was chosen as it tends to be wetter than average, and previous studies have shown a higher degree of topographic organisation of soil moisture during wetter conditions (Tenenbaum et al., 2006; Western et al., 1999a). Weather data for the test embankment confirmed that the month of November was wetter than average, reflected in the relatively high mean volumetric soil moisture content of 41.1% (relative to previous soil moisture contents of 27.0% in February and April). It should be noted, however, that laser scanning data cannot be easily collected during times of precipitation. Therefore, operational difficulties may be encountered when attempting to obtain airborne lidar data during wetter periods when there is a greater chance of a precipitation event. Despite this, it can be assumed that terrain would not change significantly, except in the case of an extreme event such as a landslide, and therefore the lidar data can be collected on the nearest rain free date.

Generally, the distribution of soil moisture values followed expected patterns with higher values on the north facing slope and an increase in moisture downslope. However, these patterns were not as strong as expected. For example, soil moisture measurements made down the profile of the earthwork indicated a degree of variation which did not always follow the general downslope trend. Similarly, there was only a small amount of difference between average soil moisture content of the north and south facing slopes with average soil moisture contents of 43% and 38.9% respectively. As a result, poor correlations were found between observed soil moisture and the TWI and potential solar radiation metrics, with correlation coefficients less than 0.2. The reason for the lack of topographic organisation, and resulting correlation with observed soil moisture, was due to heavy rainfall event occurring just prior to the measurements. This

left little time for the downslope movement of water, which has been shown to be a lengthy process (Wilson and Gallant, 2000).

The lack of observed topographic organisation of soil moisture was confirmed by poor regression correlation coefficients for all the terrain analysis metrics (correlation coefficients less than 0.2). The TWIs calculated using both the interpolated DEMs illustrated little or no explanation of the distribution of observed soil moisture (R^2 less than 0.1). This was also the case for the TWIs expressed by the natural logarithm, carried out following the findings made in the transport corridor scale study, which revealed a non linear relationship between TWI and soil moisture.

Although results were poor, the calculation of potential solar radiation for an entire year was shown to have better correspondence with soil moisture (R^2 of 0.16) compared to the calculation made for a specific day (R^2 of 0.15 respectively). This was not expected as it was assumed that the potential solar radiation calculated for a specific day would provide a better characterisation of solar radiation at the time of year in which the validating soil moisture measurements were taken. The fact that a non-specific calculation of solar radiation for an entire year provides a better representation of the distribution of soil moisture means that the calculation does not need to be repeated for different times of the year. This increases the practicality of using this metric in an operational context. However, a full temporal study would need to be carried out to confirm this, comparing the calculation to soil moisture values observed throughout the year, under different wetness conditions. This was beyond the scope of the present study as the primary focus is on scaling up the characterisation of soil moisture to a transport corridor rather than focussing on variations at a small earthwork scale.

Integrating the TWI and potential solar radiation using stepwise multiple linear regression improved the correlation with soil moisture to an R^2 of 0.26. Despite making a marked improvement on the correlation coefficients for the individual metrics, this correlation is poor. GWR helped to further improve the correlation, producing an R^2 of 0.56. This shows that locally fitted regression was better suited to the data compared to global regression. A map of soil moisture values predicted using this model showed that the only discernable pattern was an increase in soil moisture on the north facing slope, with little or no trend running downslope. This reinforced the earlier notion that

the poor correlations found in this study were due to poor organisation of the soil moisture. Again, this would need to be supported by more soil moisture measurements gathered at different times of the year to test whether the poor correlations found were due to poor spatial organisation or due to the metrics not fully representing the distribution of soil moisture over the test embankment. In addition, although the GWR model may provide an improved correlation with soil moisture, questions are raised over the model's transferability outside the study area. This is because the regression parameters are tailored to localised clusters of values and may be unstable when applied to other areas.

6.1.1.5 Summary of the test embankment terrain analysis experiments

Although the terrain analysis metrics showed poor correlations with observed soil moisture, important lessons were learnt. Firstly, if observed soil moisture measurements are to be used to build a statistical model using the terrain calculations, they must be made not only during wetter overall conditions, as stated by other authors (Tenenbaum et al., 2006; Western et al., 1999a), but also with sufficient time after a rainfall event to allow for downslope organisation to take place. This study showed that measurements made within hours of a rainfall event did not allow sufficient time for moisture to organise. Future work, with a greater emphasis on temporal patterns of soil moisture distribution, could help to demonstrate the adequate amount of time required for soil moisture to spatially organise before making measurements.

This study has highlighted the complex nature of soil moisture distribution and that a temporal dimension may have to be taken into account to enable an accurate characterisation. This is a limitation that should be noted when using terrain analysis for the future characterisation of soil moisture, suggesting the use of quasi-dynamic wetness indices may be more applicable, such as Borga et al. (2002). However, users of these calculations must also consider the application for which the soil moisture characterisation is being used. For instance, in flood modelling it may be necessary to monitor the changes in soil moisture distribution over time to coincide with extreme rainfall events, but for slope stability studies a single snap shot characterising moisture distribution may be adequate to focus maintenance work in areas susceptible to soil

moisture accumulation, such as areas of water ingress from areas surrounding earthworks (Perry et al., 2003b).

6.1.2 Hyperspectral response of vegetation to artificial wetting

This part of the earthwork scale study focussed on the analysis of hyperspectral reflectance from earthwork grasses in response to artificial changes in soil moisture. This work was carried out on the same test embankment as the terrain analysis (Section 6.1.1) but acts as a standalone study.

6.1.2.1 Comparisons between observed soil moisture and spectral reflectance

Analysis of the spectra did not reveal any obvious response to changes in soil moisture. The derivative of the spectra, however, indicated some signs of change, particularly over the wetted plot where the red-edge feature had weakened over the duration of the experiment (Figure 4.32). The red-edge feature is indicative of chlorophyll production (Zarco-Tejada et al., 2003) therefore demonstrating that changes in soil moisture were having an effect on the vegetation. The fact that the derivative of the spectra indicated a response, and not the original spectra, was an important finding. The derivative of a spectral band is based on the previous and next bands. This restricts the application of derivative techniques to sensors with sufficient spectral resolution, i.e. hyperspectral sensors (Tsai and Philpot, 1998). Imagery from such sensors, both airborne and spaceborne, tends to be expensive to acquire relative to imagery from a broad band sensor. This may therefore limit the number of data collection campaigns a network operator can afford to carry out, reducing its effectiveness as an operational monitoring technique.

All spectral analysis techniques were shown to be reasonable predictors of soil moisture. Overall, the linear interpolated REP showed to be the best predictor with a validated R^2 correlation coefficient of 0.60. This was an interesting finding as it was the simpler of the two methods for calculating REP and would be less computationally demanding if applied to large airborne hyperspectral datasets. Continuum removal analysis was shown to have the poorest correlation with soil moisture (validated R^2 of 0.51), despite showing the largest visual response to changes in soil moisture over the wetting period. This implies that the continuum removed spectra were not being fully exploited because comparisons with observed soil moisture were made using the value

at the lowest point on the continuum removed trough, at 675 nm. Future work may focus on enhancing the information gained from continuum removed spectra. This study attempted to develop metrics based on ratios and spectral angle of the continuum removed spectra but found no improvement to the correlation with soil moisture.

A correlation matrix showed little significant different between the spectral analysis techniques. As a result, it was difficult to determine which technique is best using the findings made in this study. The success of the techniques may therefore be assessed by their ease of calculation. For example, linear interpolated REP and derivative stress ratios were much simpler to compute compared to the Lagrangian REP and may therefore be more applicable to larger datasets. However, techniques such as Lagrangian REP have the advantage of requiring no *a priori* knowledge of the spectrum (Dawson and Curran, 1998). The derivative stress ratio requires knowledge of the spectra in order to sample reflectance from the double peak feature (Zarco-Tejada et al., 2003). However, this study has shown that the position of this feature defined by Smith et al. (2004b) using derivative stress ratios can be applied to embankment vegetation for characterising soil moisture. Further work may help to support these findings by testing its robustness over a range vegetation types.

6.1.2.2 Potential for scaling up the techniques to an airborne sensor

The spectra gathered in this experiment were used to simulate the signal from the CASI sensor in an attempt to test the potential for scaling up the techniques to an airborne sensor (Section 4.2.4). Interestingly, improvements were made to the correlation coefficients for all the spectral analysis techniques except the linear interpolated REP. This implies that the spectral resolution of the CASI sensor is adequate for detecting the response of vegetation to changes in soil moisture. Therefore, the derivative stress ratio, Lagrangian REP, linear interpolated REP and NDVI could potentially be applied to a sensor that is routinely flown by operators such as the Environment Agency's Geomatics Group. Specifically, the techniques derivative stress ratio, NDVI, Lagrangian and linear REP estimation showed a validating R^2 correlation coefficient between 0.58 and 0.61. Although there are no similar spectral analysis studies to draw comparisons with, these correlations are comparable with reported correlations with

topographic indices, such as Tenenbaum et al. (2006) who found R^2 values typically in the region of 0.56.

It was of particular interest to note that the NDVI calculated using the simulated spectra showed good correspondence with soil moisture. This index is easy to compute and is widely applied to imagery with red and near infrared capabilities. In addition, the derivative stress ratio showed good correspondence with observed soil moisture despite using central wavelengths that do not correspond directly to the double peak feature found in the spectra. This means the default bands of the CASI sensor provide sufficient detail to represent the double peak feature for the vegetation of the test embankment. This is despite the earthwork including a mixture of grass and wildflower species (Glendinning et al., 2009) where previous work has shown that the wavelength position of the double peak feature varies from species to species (Cho and Skidmore, 2006; Clevers et al., 2004; Smith et al., 2004a; Zarco-Tejada et al., 2003). This demonstrated the robustness of the original derivative stress ratio defined by Smith et al. (2004b) for use over a mixture of vegetation species which could prove crucial in the development of an operational methodology.

Despite encouraging results from the analysis of the CASI simulated spectra, the simulation used in this study is no replacement for testing the techniques using a sensor with reduced spectral resolution. For example, the smoothing process carried out in the sensor simulation may have retained spectral detail which might otherwise be lost using a reduced spectral resolution sensor. In addition, only spectral resolution was examined in this simulation. No consideration was given the spatial resolution, or field of view of the sensor. The field of view for the field spectroscopy data was just 20 cm^2 (Section 3.2.2.2) whereas the spatial resolution of airborne CASI imagery tends to range between 1 m and 2 m and is therefore likely to capture reflectance from a wider range of vegetation canopy types. An attempt was made to simulate this by taking an average of multiple field spectroscopy measurements taken over the metre squared plots. However, fine scale detail may have been retained.

6.1.2.3 Partial least squares regression analysis

Section 6.1.2 has focussed on the assessment of existing spectral analysis techniques for the characterisation of soil moisture. In addition to these techniques, the development of

a new technique applied the CASI simulated spectra was explored using PLS regression, that is specifically designed to characterise soil moisture. Prior to the PLS analysis, the spectra was transformed using $Log_{10}(1/R)$ following Reeves (2009) and Schmidtlein (2005). A subsequent Martens' uncertainty test (Martens and Martens, 2000) applied to the transformed spectra showed only two bands to have significant ability to predict soil moisture. However, a subsequent model built using these two bands produced a cross validated correlation coefficient of just 0.37 (Section 4.2.4).

Although disappointing, the poor results were likely to be a result of the $Log_{10}(1/R)$ transformation. This method was used as it has been employed in other vegetation spectral analysis studies such as Smith et al. (2003; 2002) and Schmidtlein (2005). Normally, reflectance in the red and near infrared would be expected to exhibit the greatest sensitivity over a vegetation target but once the spectra were transformed, the blue and green regions of the spectrum became more prominent. This meant that the transformation was actually detrimental to the PLS regression routine as it suppressed the signal from the red and near infrared and accentuated the noisy blue region of the spectrum. This is an issue that has been little explored in vegetation spectroscopy and shows that such transformations, which were designed for spectro-chemical studies (Reeves, 2009; Næs et al., 2002), should not necessarily be applied in vegetation analysis. This raises the need for a study to assess the impact of different transformation techniques on the hyperspectral reflectance analysis of vegetation, similar to the review conducted by Reeves (2009) for chemical spectroscopy.

Næs et al. (2002) explained that spectra can also be transformed using the derivative of the spectrum. The derivative of the spectrum has already been shown in this study to have a strong response to changes in vegetation following artificial changes in soil moisture. This finding was reinforced by the Martens' uncertainty test (Martens and Martens, 2000) which indicated six simulated CASI bands as having significant predicting power. These bands corresponded to both the red edge feature and the specific bands used in the derivative stress ratio. As a result, these bands are more likely to demonstrate a response to changes in the vegetation. A subsequent PLS regression model, built using these significant bands, produced a strong correlation with observed soil moisture which was consistent upon validation. This suggested that the model was stable and had the potential to make predictions outside the model training area.

The success of the PLS regression model could be a significant finding for the characterisation of soil moisture over transport corridor earthworks. However, a number of simplifications were made in order to facilitate this earthwork scale test. Primarily, the vegetation of the plots were relatively homogenous, with low canopy grasses dominating. Although this is a characteristic of modern constructed earthworks, the predominantly aging infrastructure of the UK has a much more complex mix of grasses, shrubs and trees. As well as providing a number of different leaf structures, this also produces a much more complex canopy in terms of bi-directional reflectance (Wolf et al., 2010), which would not have been taken into account in this model. In addition, field observations of many real-world earthworks show that areas of bare ground are commonplace. Background reflectance from soil within the sensor's instantaneous field of view is likely to reduce the signal from the vegetation canopy and destabilise the model which in this case has been constructed exclusively using vegetation.

6.1.2.4 Hyperspectral vegetation analysis summary

This work has demonstrated that spectral analysis techniques have the potential for measuring changes in vegetation caused by changes in soil moisture content. In addition, sufficient evidence has been given to consider the application of these techniques to airborne sensors, with a coarser spectral resolution. This work also showed that PLS regression can be used to build a stable model for predicting changes in soil moisture. However, there is uncertainty as to the ability of these techniques to perform over more complex environments, such as tree canopies, which are likely to be present within real world transport corridors.

6.1.3 Earthwork scale study summary

The results of the earthwork scale study have shown poor correlations between the terrain analysis metrics and observed soil moisture, whereas the spectral analysis study has generally shown good correspondence. However, these findings are not necessarily a direct reflection of how these techniques might perform at the transport corridor scale using airborne remotely sensed data. For instance, the relatively homogenous terrain of the test embankment meant there was little variability to be captured by the terrain analysis calculations. In contrast, a real world transport corridor would be expected to

have earthworks with a range of different slope angles, aspects, sizes and other influences such as flow contribution from surrounding areas. The increase in topographic variation would also be expected to have an influence on the distribution of soil moisture. This is supported by the range in soil moisture contents measured in the transport corridor scale study. For example, the range in soil moisture content found in the transport corridor was 45% (volumetric soil moisture content) (Section 5.1.2) compared to a range of 32% for the test embankment during similar weather conditions (Section 4.1.7). The suggestion here is that soil moisture is easier to model in an environment with increased variability.

In contrast to the terrain analysis experiment, the spectral analysis study was carried out in a more controlled environment. Here, the variation in soil moisture was artificially induced using the control, covered and wetted plots. A range of soil moisture values could therefore be enforced, allowing for a more informed model relating the spectral analysis techniques to observed soil moisture content. In addition, the spectral reflectance data used in this experiment was relatively free from the complexities normally found in real world transport corridors, including large shrub and tree canopies.

6.2 Transport corridor scale

The results from transport corridor scale work is analysed and discussed in the subsections below. This is divided into two distinct sections addressing terrain analysis and vegetation reflectance analysis (Sections 6.2.1 and 6.2.2 respectively). The findings made using these two distinct experiments were used to inform an integrated model which is analysed and discussed in Section 6.2.3.

6.2.1 Terrain analysis

The findings made by applying the terrain analysis calculations at the transport corridor scale broadly followed the observations made at the earthwork scale. This helped to support some of the claims made at an earthwork scale for the optimal implementation of the calculations within transport corridor environments. These findings are discussed below in Section 6.2.1.1. In addition, it was found that improvements could be made when applying the calculations at the transport corridor scale in comparison with the

earthwork scale. This is discussed in Section 6.2.1.2 together with an analysis and discussion of the observed distribution of soil moisture and what this the implications of these findings mean for slope stability analysis as a whole.

6.2.1.1 Impact of DEM interpolation

As with the analysis undertaken over the test embankment, two separate DEMs were generated for the transport corridor using two different interpolation techniques. The AnuDEM routine was used as it allows for the calculation of continuous flow downslope and is therefore, in terms of hydrological modelling, a more accurate representation of reality (Hutchinson, 2008). IDW interpolation was also explored at this scale as opposed to the kriging technique used in the test embankment study. This was because the kriging interpolation routine was too computationally demanding for generating a DEM with a grid resolution of 1 m over a 7 km transport corridor, often leading to lengthy processing times (over two days) and was unstable and prone to crashing using both ArcGIS and TerraScan software. This finding has implications in terms of using the methods developed here in an operational context. Various interpolation techniques may have relative advantages and disadvantages in terms of representing terrain, or making characterisations of soil moisture, but their operational use is limited if the demand on computer processing exceeds the capacity of current desktop computers.

Following similar trends to the earthwork scale results (Section 6.1.1.1), the AnuDEM interpolation technique produced a much smoother representation terrain compared to the IDW method, which held more fine scale topographic detail. This difference translated itself into a difference in the calculation of TWI. Specifically, a TWI calculated using the AnuDEM DEM showed a smooth decrease downslope, whereas TWI for the IDW DEM appeared spatially erratic, with wetness predicted to increase and decrease several times before reaching the bottom of the slope. This was similar to the results found in the test embankment scale study which were made at a much finer grid resolution. This showed that although the 1 m resolution DEM for the transport corridor would suppress a degree of topographic detail, the IDW interpolation retained enough detail to affect the flow of moisture downslope predicted by the TWI. As a result, it is suggested that the AnuDEM interpolation technique should always be

used when generating a DEM for TWI calculation. Reiterating the statement made in Section 6.1.1.1, this finding is important as some studies have relied upon other interpolation techniques, such as kriging, to generate DEMs for the calculation of TWI (Schmidt and Persson, 2003). Despite this, Schmidt and Persson (2003) reported no similar erratic patterns in the calculation of TWI. This could be due to the fact that they made their calculations using a DEM with a grid resolution of 2 m. This means that less small scale topographic variation would be captured by the kriging interpolation than if a 1 m DEM was used, leading to less variation in the calculation of TWI. In addition, Schmidt and Persson (2003) drew comparisons with point soil moisture measurement by taking an average TWI value using a filter with a radius of three cells which is likely to suppress any small scale variation.

6.2.1.2 Comparisons between observed soil moisture measurements and the terrain analysis calculations

The terrain calculations, TWI and potential solar radiation, made using the two interpolated DEMs were compared against observed soil moisture measurements. Measurements made during relatively dry periods demonstrated little topographic organisation, which was consistent with the findings made by other studies, including Tenenbaum et al. (2006) and Western et al. (1999a). The low degree of spatial organisation specifically related to the middle of the earthwork slopes or in locations where slopes are relatively flat. However, patterns did emerge including a general increase in soil moisture found on north facing slopes and at the toe of earthwork slopes. The latter trend showed that during periods of little or no rainfall, water retained within an earthwork migrated downslope and concentrated at the bottom of the slopes.

This trend was more apparent in the soil moisture measurements taken following a relatively wet period, with generally little change over the centre of the slope but a marked increase at the toe. This finding could be significant in terms of slope geotechnics as areas of soil moisture convergence are likely to increase the strain at specific points in the earthwork and increase the risk of instability (Pack et al., 1998). In addition, current slope stability models, including the coupled SHETRAN -FLAC model (Kilsby et al., 2009; Rouainia et al., 2009; Ewen et al., 2000), tend to characterise soil moisture using a single value for a whole slope. High soil moisture values were also found in flat fields bordering the transport corridor earthworks. This could also have

geotechnical implications with cutting earthworks as contributions of soil moisture are likely to be made from surrounding land (Perry et al., 2003b). This is particularly significant where rainfall onto saturated soils can lead to direct runoff onto the earthwork slopes which has been a major cause of recent failures (e.g. BBC (2007)). The TWI applied in this study indicated potential for the this phenomena in a cutting east of Haltwhistle which is demonstrated in the example map for the integrated model shown in Figure 5.16 (Section 5.3) which is discussed later in Section 6.2.3.

Despite the apparent organisation of the observed soil moisture measurements, regression correlation coefficients with the terrain analysis calculations were poor. Overall, potential solar radiation was the best predictor (R^2 of 0.27) which showed that earthwork orientation, and the subsequent influence of solar radiation, was the dominating factor on soil moisture distribution. This was an unusual finding as previous work has shown potential solar radiation to be less dominant in wetter conditions (Tenenbaum et al., 2006; Western et al., 1999a; Isard, 1986). The influence of this metric was a result of the west to east orientation of the transport corridor, with earthworks having predominately north and south slope facing aspects. Despite its influence here, potential solar radiation is unlikely to have the same control in transport corridors with other orientations, which was demonstrated in the earthwork scale study when the test embankment was rotated. Therefore, characterisation of soil moisture distribution using potential solar radiation may have to be tailored to suit specific orientations. This could be tested by applying the methods used in this study to transport corridors with a variety of earthwork orientations and comparing against observed soil moisture measurements. This was not considered in the initially choice of study site as preference was given to a stretch of railway with a history of instability.

A potential flaw in the calculation of potential solar radiation is that it does not take into account the vegetation canopy. Dense vegetation canopies, particularly broad-leafed deciduous trees, mask the soil surface below, limiting the amount of direct solar radiation that can reach the soil surface as well as intercepting rainfall. Future refinement of the DEM used to make this calculation but this would involve the use of optical ray tracing models (Disney et al., 2000) which require a high degree of parameterisation and are therefore beyond the scope of the present study. The use of full waveform lidar could provide an interesting direction for future study. This technology

can provide more detailed information about the vegetation canopy (Hyde et al., 2005) and have been shown to provide accurate measures of canopy geometry using the aforementioned ray tracing models (Hancock et al., 2008).

Although poor, a TWI calculated using the AnuDEM interpolated DEM provided a better predictor of soil moisture (R^2 of 0.23) compared to the IDW DEM (R^2 of 0.19). This supported the suggestion that the AnuDEM interpolation technique is better suited to making hydrological calculations (Hutchinson, 2008) but the small correlation coefficients give this statement little gravity. The TWI calculation showed a gradual decrease downslope implying a linear increase in soil moisture content. However, the soil moisture measurements showed values to remain relatively constant with a sharp rise at the toe of the earthwork slopes (Figure 5.9). In order to capture this trend the TWI was expressed by the natural logarithm, which improved correlations with soil moisture for the AnuDEM DEM to an R^2 of 0.33. This was a potentially significant finding as other studies using the TauDEM version of the TWI calculation (Tenenbaum et al., 2006) have not considered solving the non linear relationship with soil moisture. Furthermore, expressing the metric by the natural logarithm is simple to compute and does not introduce multiple additional parameters such as cubic and quadratic expressions.

Combining the natural logarithm of the TWI and potential solar radiation using stepwise multiple regression produces an R^2 of 0.65. This was comparable to other studies, such as Western et al. (1999a), who found a combination of TWI and potential solar radiation explained up to 61% of the variation of soil moisture during wetter conditions. The difference with the present study is that the characterisation of soil moisture has been made at a fine spatial resolution of 1 m compared to Western et al. (1999a) with a grid resolution of 5 m. This implies that the terrain analysis metrics used in this study and by Western et al. (1999a) are robust at both DEM grid resolutions. In addition, the grid resolution used on this study is sufficient for monitoring transport corridor earthworks with widths typically in the order of tens of metres (Section 2.5).

The correlation coefficient for the multiple regression model (0.65) was also a marked improvement on the correlations between observed soil moisture and the individual terrain analysis calculations (between 0.27 and 0.33). Other studies, such as Tenenbaum et al. (2006), have used TWI as the sole predictor of soil moisture

distribution, achieving correlation coefficients of approximately 0.56. The findings made in this study show that this correlation could have been improved if potential solar radiation were taken into account. Moreover, this would have been possible using the same DEM used to calculate the TWI. Tenenbaum et al. (2006) found best results using a relatively coarse DEM, with a grid resolution of 11.25 m. They also compared observed soil moisture to a TWI calculated using a much finer, 0.5 m DEM. This produced poorer correlations ranging from an R^2 of 0.56 during wetter conditions and no correlation during dry conditions. This study has shown that a reasonable characterisation of soil moisture can be made at a comparable spatial resolution of 1 m.

One potential explanation for the poor correlations found by Tenenbaum et al. (2006) at this spatial scale is that they were applying the terrain analysis calculations within an urbanised catchment. Urban areas provide more complex hydrological environments to model due to the potential range of permeable and impermeable surfaces which is not taken into account by the terrain analysis metrics. Hydrological features within an urbanised catchment also occur at different scales which may not be captured by even high spatial resolution (< 1 m) DEMs. For instance, kerbs act as a barrier forcing water down the sides of roads (Hollis, 1988) but are often less than 20 cm in height which tends to be beyond the vertical accuracy achieved by airborne lidar campaigns (Hodgson and Bresnahan, 2004).

Tenenbaum et al. (2006) found best correlations when the TWI was averaged around a 11 x 11 kernel. Therefore, these results may be misrepresentative of the accuracy of a resulting map of predicted soil moisture with a grid resolution matching the input DEM. This study did not average TWI values around a kernel, meaning observed correlations with soil moisture are representative for a map of predicted values with a grid resolution of 1 m.

A GWR model was built to predict observed soil moisture using the terrain calculations, natural logarithm of TWI and potential solar radiation. This helped to improve the correlation coefficient to an R^2 of 0.76. Therefore, a locally fitted regression was better than global regression for describing the relationship between soil moisture distribution and the terrain analysis attributes. The improvement in correlation coefficient reflects the complex spatial relationship between soil moisture and the terrain analysis metrics. For example, whilst there is a general negative trend between

potential solar radiation and observed soil moisture, the relationship changes downslope as observed soil moisture tends to gather at the toe of earthwork slopes. This was an important finding as previous studies have not considered a spatially local model to describe the relationship between terrain analysis calculations and soil moisture. Future work should therefore consider spatially variable relationships, in order to refine existing terrain analysis calculations. This is particularly poignant when considering a parameter (soil moisture) which has consistently been stated as having innate spatial variability (Famiglietti et al., 1999; McBratney and Webster, 1986; Hawley et al., 1983), where studies attempting to model its distribution have largely relied on global models (Tenenbaum et al., 2006; Western et al., 1999a).

Monte Carlo significance testing showed that potential solar radiation was the only parameter that had a spatially significant relationship (with 99% confidence) with soil moisture. The spatial dependence between soil moisture and potential solar radiation reflects the influence of the predominately north and south facing slopes. This is demonstrated by the GWR model residuals which were shown to be an average of 4.6 (% volumetric soil moisture content) on the south facing slope compared to -4.2 on the north facing slope, with an average residual of -0.8 (Section 5.1.3). As a result, future work should consider the localised relationships between potential solar radiation and soil moisture, in order to make better informed predictions of soil moisture distribution.

6.2.2 Vegetation reflectance analysis

The analysis undertaken here could not be assessed at the earthwork scale due to the relatively homogenous nature of vegetation on the test embankment. Therefore, the results discussed here are irrespective of the hyperspectral analysis results found at the earthwork scale. The work carried out here was the first reported attempt to map Ellenberg biological indicator values in UK transport corridor environments. As a results, separate sections (6.2.2.1 and 6.2.2.2 respectively) discuss the observed Ellenberg values and the mapping of these values using airborne imaging spectroscopy. An additional section (Section 6.2.2.3) addresses the use of this technique as a predictor of soil moisture distribution.

6.2.2.1 Observed Ellenberg indicator values

Thirty-three plots were used to measure species abundance in order to derive weighted average Ellenberg indicator values. The variation in species found in the species abundance plots demonstrated the heterogeneous nature of vegetation cover within the transport corridor. However, slope stability models, including SLIP4EX (Greenwood, 2006) and SHETRAN-FLAC (Kilsby et al., 2009; Rouainia et al., 2009; Ewen et al., 2000), characterise vegetation cover as a single category, normally bare soil, grass, or trees. Although this study is not concerned with the relationship between slope stability and vegetation, the observed variation in vegetation type found in the study area suggests that these models insufficiently describe the actual distribution of vegetation cover over a transport corridor earthwork slope. This is of particular significance where vegetation type changes down a slope profile from shallow root types (grasses and wildflowers) to deep root types (tall shrubs and trees) which have shown to have different influences on slope stability (Glendinning et al., 2009).

The number of plots used to measure species abundance was based on a species-area curve constructed from observations made within a modified-Whittaker plot following Stohlgren et al. (1995) and Mueller-Dombois and Ellenberg (1974). It later emerged that four of the plots could not be used due to misclassification of vegetation type or inaccurate soil moisture measurements (Section 3.3.2.1). On reflection, it would have been better practice to over sample in case any plots needed to be discarded. As a result, it is possible that the full range of vegetation species, and subsequent Ellenberg values, was not sampled. This could have implications on the success of the mapping of Ellenberg values where it is necessary to use observations that cover the full range of the parameter that is being sampled, and may account for some of the uncertainties found in the modelling process. In addition, it was noted in Section 6.2.2 that an improved validation of the Ellenberg values could have been made if an additional set of data could be collected. However, this would have required additional vegetation abundance data which was not possible as validating soil moisture data has to be sampled within a day, preferably within the space of just a few hours, due to the temporally variable nature of this parameter (Famiglietti et al., 1998). This could have been possible with additional labour on the ground and remains a potential for future studies with more resources.

6.2.2.2 Mapping Ellenberg values

PLS regression was used to link the CASI spectra to the Ellenberg moisture indicator values. Prior to analysis, the spectra was transformed using the $Log_{10}(1/R)$ transformation. Similar to the findings made in the earthwork scale study (Section 6.1.2.3), the initial Martens' uncertainty test (Martens and Martens, 2000) did not show any bands to be significant. The uncertainty test also failed to show any significant predicting bands when the spectra was transformed using the first derivative, a transformation which produced meaningful predictive bands in the earthwork scale study (Section 6.1.2.3). This further reinforces the uncertainty over which transformation technique, if any, should be used in hyperspectral remote sensing of vegetation. Therefore, an alternative approach was adopted whereby predicting bands were selected using stepwise regression (Næs et al., 2002). This identified eight significant bands from the $Log_{10}(1/R)$ transformed spectra which were also found to be significant using the Martens' uncertainty test. Again, this implies uncertainty over the correct procedure for applying PLS regression in vegetation spectroscopy. Future work should be directed at addressing these shortcomings as the results from both the earthwork and transport corridor scale studies have indicated a strong potential for the use of PLS regression.

PLS regression using the significant bands from stepwise regression produced a strong correlation with the Ellenberg indicator values, with an R^2 of 0.85. The latter model used eight components to describe the relationship between Ellenberg moisture values and the significant CASI bands. Normally, it is held that a large number of model components would make it unstable for making predictions outside the training area (Næs et al., 2002). However, this model was chosen specifically due to its high cross validation correlation coefficient (R^2 of 0.64). The large number of components required to create a meaningful model represents the multi-dimensionality of the relationship between the Ellenberg moisture values and the predicting CASI bands. Næs et al. (2002) explained that using near infrared reflectance data, good predictions can normally be made using a few components. However, this is related to chemical spectroscopy studies where relationships may be simpler than studies examining vegetation reflectance. Therefore, future studies using PLS regression to estimate

vegetation parameters from spectral reflectance should consider a greater number of model parameters than expected.

Analysis of the mapped Ellenberg moisture values revealed trends that were consistent with field observations of vegetation type. This was encouraging as it further supported the stability of the PLS model. In addition, field observations showed a trend between high predicted soil moisture and an overall denser canopy. This canopy has a larger leaf area, associated with vegetation species such as Brambles and Stinging Nettles. The larger leaf area of these vegetation types restricts evaporation from the soil surface below the canopy, leading to overall wetter soils than the relatively less dense canopies of grasses and wildflowers towards the east of the earthwork. Although this was not the intention of the Ellenberg mapping procedure this could prove to be a useful finding. Specifically, slope stability models, such as SHETRAN-FLAC (Kilsby et al., 2009; Rouainia et al., 2009; Ewen et al., 2000), characterise soil moisture based on the dominating vegetation cover (bare ground, grass or trees) and their relative impact on the interception of rainfall due to leaf area. The observations of leaf area made here are purely subjective but it would be interesting to pursue this matter further by attempting to predict leaf area index using the CASI imagery (Boegh et al., 2002). However, this would have involved extensive field observations and as a result, was beyond the timescale of the present study.

6.2.2.3 Comparisons between observed soil moisture measurements and mapped Ellenberg values

The mapped Ellenberg moisture values were compared against observed soil moisture measurements, producing an R^2 value of 0.4 (significant at the 99% confidence level). This was poorer than the results found in similar studies, such as Schmidtlein (2005), who found an R^2 value of 0.58. This difference was due to the range of observed Ellenberg moisture values found over the study site. The present study found a relatively limited range of 4.8 to 6.8, compared to Schmidtlein (2005) with a range of 4.3 to 9.5. The greater range found in the latter study was due to the presence of raised bogs and wet fens which are unlikely to be found on transport corridor earthworks. Despite the smaller range in observed values, the correlation coefficient demonstrated that the mapped Ellenberg moisture values can provide some explanation of the distribution of soil moisture over earthworks. However, as the correlation was not

strong the suggestion was that additional metrics, the terrain analysis metrics, are required to make a reasonable characterisation of soil moisture. This was similar to the findings discussed in Section 6.2.1 which showed that the terrain analysis calculations were poor predictors on their own and needed to be incorporated to produce any meaningful predictions. Other attempts to map soil moisture distribution using terrain analysis such as Western et al. (1999a) also concluded that best correlations are found when a combination of metrics are used. However, studies using mapped biological indicator values, such as Schmidtlein (2005), tended not to consider the integration of other remotely sensed metrics. This study shows that the estimation of soil moisture using mapped Ellenberg values can be improved by integrating terrain analysis calculations, as discussed in Section 6.2.3.

One major limitation of the Ellenberg mapping procedure is that it was not applied to woodland areas. Such areas were not considered following the example of Feldmeyer- Christe et al. (2007) and Schmidtlein (2005). As mentioned in Section 3.3.2.1, this was because tall tree canopies not only complicate the spectral signal with bidirectional reflectance and shadowing (Wolf et al., 2010) but also mask smaller species on the forest floor that may be a better indicator of soil moisture conditions. Consideration could be given to the incorporation of radiative transfer models into the Ellenberg mapping procedure. Such models are used to model spectral reflectance over vegetation canopies in an attempt to remove the complications caused by issues such as bidirectional reflectance (Cescatti, 1997). This improves the signal from the vegetation canopy and has been shown to improve the characterisation of vegetation using remotely sensed data (Jacquemoud et al., 2009). This was not explored as such models require a number of input parameters (Section 2.2.2) which may detract from the potential operational use of these techniques.

6.2.3 Integrated model

The terrain analysis work showed that metrics derived from remotely sensed data have little predictive power on their own. This was confirmed by R^2 poor regression correlation coefficients of 0.33 for the natural logarithm of the TWI and 0.27 for the calculation of potential solar radiation. However, when the two metrics were integrated using stepwise multiple regression the R^2 value improved to 0.65 with 99% confidence.

The inclusion of mapped Ellenberg moisture values further improved this correlation, although marginally, with an R^2 value of 0.67. As discussed earlier in Section 6.2.1.2, this is comparable to the best correlations found in other studies (Tenenbaum et al., 2006; Schmidt and Persson, 2003; Western et al., 1999a) but in this case predictions were made at a spatial resolution sufficient for monitoring transport corridor earthworks. In addition, and perhaps more significantly, these predictions were made over densely vegetated areas, a characteristic that has restricted the use of other remote sensing techniques, including thermal inertia mapping, active and passive microwave systems, and spectral reflection from the visible, near and shortwave infrared (Chapter 2).

A map of predicted soil moisture, made using the integrated multiple regression model (natural logarithm of the TWI, potential solar radiation, mapped Ellenberg values) clearly showed the influence of the TWI with a gradual increase in predicted moisture running down the profile of earthwork slopes. Furthermore, the influence of the TWI resulted in a number of points within a cutting earthwork where contribution from neighbouring fields has directed flow into the earthwork. It is important that these features have been retained as slope stability studies, such as Pack et al. (1998), depict areas of potential instability by moisture concentration in areas of topographic convergence. Similarly, Perry et al. (2003b) explained that the infiltration of water into cutting earthworks from surrounding areas can result in the cess heaving process. This is a significant slope process whereby infiltration of water into an earthwork softens the clay foundations, weakening the soil above. The integrated model also showed the influence of the potential solar radiation calculation, suggesting that effects of cess heaving will be accentuated on north facing slopes where predicted soil moisture contents are higher that the south facing slope.

The correlation coefficients for the multiple regression models show that the inclusion of mapped Ellenberg values only slightly improved the overall prediction of soil moisture (improvement in correlation coefficient from 0.65 to 0.67). Potential users of this work could therefore conclude that a reasonable characterisation of soil moisture can be made using just the terrain analysis calculations. This could have greater operational potential as the terrain analyses require just one airborne dataset (lidar) and relatively little field work, compared to the extensive vegetation sampling associated

with the Ellenberg values. However, the integrated model showed subtle variations in modelled soil moisture which originate from the mapped Ellenberg value metric. When this particular metric is taken out of the integrated model, leaving the natural logarithm of the TWI and potential solar radiation, this pattern is lost. Potential users of this methodology must carefully consider whether these subtle variations are important to justify the use of the Ellenberg mapping procedure. In terms of slope stability analysis this could be crucial as small scale areas of increased moisture may represent areas of water inundation from water sources such as leaking pipes, ponds or lakes, blocked drains, and increases in groundwater level (Preuth et al., 2010; Perry et al., 2003b; Fukuoka, 1980), which might otherwise not be detected by the terrain analysis metrics.

Predicted soil moisture using the integrated model was validated against additional soil moisture measurements giving an R^2 of 0.48 which was significant at the 99% confidence level. The reason for this relatively poor correlation may be due to the fact that the validating measurements were gathered following a period of relatively dry weather when the degree of topographic organisation was not at its greatest. In addition, some of the measurements were taken on earthworks containing a mixture of fly ash and ballast which were shown to produce erroneous measurements. This was because the presence of ballast created pockets of air within the soil matrix being sampled which would change the electromagnetic impedance recorded by the probe to estimate soil moisture content.

The measurements were taken using a portable Theta impedance probe (Section 3.2.1.2) which have been used in a number of studies into the spatial distribution of soil moisture, including Cosh et al. (2005), Famiglietti et al. (1999), Western et al. (1999a), and Tenenbaum et al. (2006). However, these studies do not refer to any difficulties in obtaining meaningful measurements. This implies that transport corridor earthworks containing fly ash and ballast present an additional challenge to the study of soil moisture. Calibration of the Theta probe prior to data collection did not take into account this type of material (Section 4.1.7) as permission was not given to take soil samples from the transport corridor earthworks. Future work should endeavour to capture the full variability in soil type, particularly in transport corridor environments. However, this would require extensive field sampling which would have detracted from

the original aim of this study, to provide a more efficient characterisation of soil moisture distribution.

Due to the lack of confidence in the validating soil moisture dataset, the integrated model was also validated using cross validation. This procedure demonstrated an RMSE of 5.48% (volumetric soil moisture) for the original model calibration and an RMSE of 5.89% for the cross validation, suggesting model error does not change significantly when applied to new observations. Although this supports the overall robustness of the model it is not as reliable as testing the model against actual observations. In addition, no consideration has been given to error propagation throughout the modelling process. Future work should consider systematically removing samples from the metrics used in the integrated model to test its sensitivity. For example, the PLS regression model used to map the Ellenberg values was not validated against additional field observations. Therefore, tests could be done to remove particular observations and see how this affects the overall PLS model. In turn, the resulting mapped Ellenberg values could integrated with the other metrics to test for differences with the original integrated model. If the original model was stable, differences with the new model would be expected to be small. This analysis was not carried out due to the extensive processing that was undertaken for each remotely sensed metric. In order to carry out a rigorous analysis of error propagation this would take a great deal of processing time and was not within the time frame of the present study.

6.2.4 Implications for mapping slope stability

The overall aim of this study was to characterise soil moisture as an indicator of slope instability in transport corridor environments. Although the characterisation of soil moisture has been achieved, there has been little mention of its application in mapping slope stability. This work has been carried out as part of the wider EPSRC (Engineering and Physical Sciences Research Council) funded project ‘Remote asset inspection for transport corridor environments’ (reference number EP/D023726/1). The aim of this project was to provide an assessment of slope instability by extracting a number of key parameters using a range of remote sensing devices (Miller et al., 2009). These parameters include slope gradient, slope aspect, vegetation type, and soil moisture.

Following other slope stability studies (Borga et al., 2002; Pack et al., 1998) soil moisture was characterised using a TWI.

Specifically, the remotely measured slope parameters, slope gradient and vegetation type, are used as an input into a finite element numerical modelling procedure. This procedure involves the coupling of the SHETRAN hydrological and the FLAC TP geotechnical models, to simulate the response of different slopes to a range of different climate scenarios (Kilsby et al., 2009; Miller et al., 2009; Rouainia et al., 2009). The influence of soil moisture, here modelled using a TWI, is later incorporated using a weighting factor based on wetter soils having a greater risk of slope instability. As this study has shown, the calculation of TWI provides a limited characterisation of soil moisture distribution, particularly as the calculation was made using a DEM interpolation technique (IDW) that is not 'hydrologically correct', unlike AnuDEM (Hutchinson, 2008). In addition, non linear relationships between the TWI and soil moisture were not considered. Therefore, the findings made in this study, using an incorporated TWI, potential solar radiation, and mapped Ellenberg values model, can help to improve the characterisation of the soil moisture parameter for input into the wider 'Remote asset inspection for transport corridor environments' project.

Since this study was conducted, the UK railways operator Network Rail has expressed an interest in using some of the developed methods. Specifically, there has been growing concern over the impact of blocked or defective drainage, leading to concentrations of soil moisture and risk of instability for transport corridor earthworks (Hall, 2009). Particular interest has been given to the potential for using mapped Ellenberg values to identify areas abundant in hydrophilic vegetation species which may be indicative of defective drainage. This would help to direct ground based network inspectors to assess areas with potential risk of soil moisture concentration. Further development may also include the use of a TWI to identify areas of moisture convergence that are not currently logged as having drainage installed in earthwork structures or in surrounding land.

6.3 Summary

This summary is divided into three distinct sections. The first addresses the overall success of characterising soil moisture in transport corridors using the developed

methodology. The second refers to the implications of the findings made in this study with specific reference to slope stability analysis in UK transport corridors. The third looks at the wider implications of this study in terms of benefitting other applications.

6.3.1 The success of characterising soil moisture

This chapter has discussed the implications of the results found in this study. This has included an assessment of individual techniques for characterising soil moisture at both an earthwork and transport corridor scale. At an earthwork scale, the terrain analysis calculations were shown to be poor predictors of soil moisture distribution. These poor results were the result of a combination of poorly organised observed soil moisture measurements and the lack of topographical variability that existed on the test embankment. This implies the terrain analysis technique is only applicable to areas with variable terrain. In addition, the soil moisture measurements have shown that future studies need to consider the time since the last rainfall event in order to obtain meaningful soil moisture observations.

The vegetation spectral analysis, however, was more encouraging. This work showed that a reasonable characterisation of changes in soil moisture content could be determined from vegetation reflectance. Furthermore, this experiment showed the potential for applying these techniques to an airborne sensor. However, further work needs to be carried out to assess the potential for these techniques to be applied to more complex vegetation covers, which exist over real world transport corridors.

At a transport corridor scale the individual terrain analysis calculations also demonstrated generally poor correlations with observed soil moisture measurements. Despite this, important findings were made. The calculation of TWI was shown to be better suited to the AnuDEM interpolated DEM. This is significant as this interpolation technique is not always adopted in studies that use calculations of TWI. In contrast, the calculation of potential solar radiation was shown to work best when applied to an IDW interpolated DEM which better represents fine scale topographic features. This is again an important finding as studies tend to use one interpolation technique to generate a DEM, regardless of the type of analysis it is being used for.

The analysis of vegetation reflectance over the transport corridor has revealed that Ellenberg indicator values can be robustly mapped for a UK transport corridor. This

is a feat that has yet to be reported in the literature and represents a progression in deriving land surface parameters at a high spatial resolution using remotely sensed data. In terms of characterising soil moisture, the mapped values showed fair, but by no means strong, correspondence with observed soil moisture. However, when they were incorporated with the terrain analysis metrics, the correlation with soil moisture was comparable to other studies at a coarser spatial resolution. In addition, these predictions were made over densely vegetated areas that have previously restricted other remote sensing techniques.

6.3.2 Implications in slope stability studies

As well as providing a reasonable characterisation of soil moisture, this study has demonstrated a number of implications in association with slope stability studies. Observed soil moisture was shown to gather towards the toe of earthwork slopes, resulting in a non linear relationship with the wetness index. The TWI also predicted areas in which moisture is contributed to cutting earthworks from surrounding fields. These areas of convergence, and subsequent concentration of soil moisture, raise their susceptibility to instability due to the potential for higher pore water pressures. In addition, this susceptibility is shown to increase on north facing slopes where soil moisture was shown to be higher.

This study has also raised questions over the integrity of current slope stability models. Firstly, soil moisture was shown to vary considerably downslope whereas models such as SHETRAN-FLAC characterise soil moisture by a single value. This is a generalisation which could have implications on modelling of the internal structure of earthworks as variability in surface soil moisture is likely to reflect a variation in pore pressure down the soil profile. Secondly, in the example of SHETRAN-FLAC, soil moisture is characterised by the dominating vegetation cover. This is again assumed to be constant, whereas field observations of vegetation species distribution in this study have demonstrated great variability. Although all models must make assumptions, the generalisations made by these models are unrealistic representations of conditions in transport corridors and are likely to affect the subsequent modelling undertaken for the ‘Remote asset inspection for transport corridor environments’ project that this work contributes to.

This study has shown potential for characterising soil moisture in vegetated transport corridors at a high spatial resolution (1 m). As reported in the literature review (Chapter 2) this has so far eluded the remote sensing community where efforts have largely been focussed on using microwave systems. Although techniques using these systems have provided higher correlations with soil moisture compared to this study (Baghdadi et al., 2006) they are made at the expense of a relatively poor spatial resolution which is too coarse for the monitoring of transport corridor earthworks. In addition, best results have been restricted to bare or sparsely vegetated areas which is contrary to the nature of earthwork infrastructure (Glendinning et al., 2009).

6.3.3 Implications for other applications

This study has focussed on the characterisation of soil moisture as an indicator of slope stability in transport corridors. However, mapping soil moisture has a wide range of applications. Walker et al. (2004) summed up the importance of measuring soil moisture by explaining that "the measurement of soil moisture content can be extremely useful, if not essential, in hydrologic, environmental and agricultural applications" (pages 85-86). Specifically, this study has helped to develop a methodology for characterising soil moisture at fine spatial scales (1 m) over vegetated areas. This has the potential to be applied to a host of different applications.

One example for the application of the methods developed in this study is for mapping malarial mosquito habitats. Research has shown that surface water habitats for certain species of mosquito (Anonpheles) in Eastern Africa tend to be smaller than 40 m^2 (Mutuku et al., 2009; Mushinzimana et al., 2006; Mutuku et al., 2006). This fine scale has meant that previous studies using broader spatial resolution sensors, such as Landsat TM (Bogh et al., 2007), have not been able to represent a large proportion of these habitats (Mushinzimana et al., 2006). The methods presented in this thesis could potentially be used to provide an improved classification of such habitats by incorporating information about moisture convergence, using a TWI, and vegetation characteristics, using mapped Ellenberg values.

Another application that could benefit from the methods developed in this study is in precision farming. Steps are continually being made to improve the spatial resolution of measuring soil parameters in the agricultural sector, particularly over large

areas (Bishop and McBratney, 2002). TWIs have been successfully used to estimate soil moisture distribution for farmland by Schmidt and Persson (2003). However, the findings made in the present study at both an earthwork and transport corridor scale (Sections 6.1.1.1 and 6.2.1 respectively) have shown that the calculation of TWI is better suited to a DEM interpolated using a thin plate spline routine (AnuDEM) rather than kriging as used by Schmidt and Persson (2003). In addition, the results have indicated that this estimation could be improved by incorporating a calculation of potential solar radiation.

The estimation of species richness for landscape conservation could also benefit from the developed methods. Luoto et al. (2002) demonstrated that species richness could be estimated over agricultural areas in Finland using a mixture of terrain analysis using a TWI and vegetation classification using broad band satellite imagery (Landsat TM). As with the previous example in precision agriculture, improvements could be made to their methodology by calculating TWI using a AnuDEM interpolated DEM, although it should be noted that Luoto et al. (2002) do not state the interpolation method used. More significantly, the methods developed for mapping Ellenberg values in this thesis may provide an improvement to the overall measure of species richness. This notion is based on previous work in the ecological literature which advocates the use of biological indicator values, such as those developed by Ellenberg et al. (1991), as an indicator of biodiversity and species richness (Diekmann, 2003; Ellenberg et al., 1991). Mapping these values over large areas using remotely sensed imagery could therefore provide major developments in this area.

7 Conclusions

The overall aim of this study was to explore the potential for using remotely sensed datasets to characterise soil moisture for a transport corridor. In order to achieve this overall aim, a number of objectives were set:

1. Identify and review existing techniques for characterising soil moisture from remotely sensed data and critically assess their operational capability in the context of transport corridor environments.
2. Carry out a proof of concept study for the techniques deemed applicable in objective one using ground based instruments at an earthwork scale and assess their potential for being scaled up to a transport corridor scale.
3. Based on the findings made in objectives one and two, apply the techniques to a transport corridor using airborne remotely sensed datasets.
4. Design a method for integrating the most successful techniques into an overall model for characterising soil moisture in transport corridor environments.

This concluding chapter explains how each of the objectives have been achieved. Specific reference is given to the key findings made and their implications for characterising soil moisture in transport corridor environments. Suggestions are also given as to how the findings made in this study can help to direct future work, as well as how this work might contribute to other applications. The final section of this chapter provides a short concluding statement to this thesis.

7.1 Objective 1

Objective 1 was addressed by the literature review in Chapter 2. This revealed a number of existing techniques that have been shown to have great potential for characterising soil moisture. In particular, the use of passive and active microwave system has proved to have the greatest potential for mapping soil moisture over wide areas (Famiglietti et al., 1999). However, these techniques share inherent limitations that render them inapplicable for use in monitoring transport corridor earthworks. Their primary limitation relates to the poor spatial resolution that these sensors often provide, which is too coarse to make characterisations over earthworks with widths in the region of tens of metres. An additional limitation is that the microwave signal from the soil is strongly

perturbed by the presence of vegetation (Notarnicola et al., 2006). As a result, studies are often restricted to areas with low homogeneous vegetation canopies or bare earth (Schmugge et al., 2002), whereas transport corridor earthworks tend to be covered in dense heterogeneous vegetation.

Chapter 2 identified techniques using two sources of remotely sensed data that have the potential to overcome the limitations associated with microwave sensors, airborne laser scanning and airborne imaging spectroscopy. Techniques applied to the former involve terrain analysis calculations that exploit the relationship between soil moisture distribution and topography, such as TWI and potential solar radiation. These calculations have been successfully applied to high spatial resolution (1 m) DEMs derived from airborne lidar (Tenenbaum et al., 2006) but have not been assessed within UK transport corridor environments. The techniques applied to imaging spectroscopy, including derivative stress ratios, continuum removal analysis, red edge position estimation, and mapping of biological indicator values, also have the potential for being applied to fine spatial resolution data (~1 m). Here, the techniques use measures of canopy spectral reflectance to make inferences about the vegetation which act as a surrogate for characterising the soil moisture below. This overcomes the limitations associated with techniques using other sensors that are restricted by the presence of vegetation, such as microwave systems and thermal sensors. Few reported attempts have explicitly attempted to characterise soil moisture using these spectral analysis techniques but some studies have shown a link with plant stress induced by waterlogging (Smith et al., 2004a; Pickerill and Malthus, 1998). The potential for characterising a number of environmental parameters, including soil moisture, by mapping biological indicator values has also been demonstrated (Schmidtlein, 2005) but not in the context of a UK transport corridor.

7.2 Objective 2

This objective was chiefly addressed by the test embankment study. Here, a proof of concept study was carried out to test the ability of the techniques identified in objective 1 for characterising soil moisture in a controlled environment using ground based sensors. This study could be divided into two distinct parts: terrain analysis using terrestrial laser scanning and the analysis of vegetation reflectance using a field

spectrometer. Varying success was found with the terrain analysis work. Although important findings were made, which helped to inform the implementation of these methods to the transport corridor scale, the poor correlations found with observed soil moisture meant that claims made did not have much statistical backing. This meant that many of the methods had to be evaluated once more at a transport corridor scale. The vegetation reflectance work, however, revealed more statistically significant findings meaning conclusions could be drawn with more confidence. The key findings from both experiments are summarised below.

7.2.1 Terrain analysis

Generally, poor correlations were found between the terrain analysis metrics and observed soil moisture. Despite this, important lessons had been made concerning their suitability for being scaled up to a transport corridor:

1. Validating soil moisture measurements must be taken during wet conditions to maximise the degree of topographic organisation. In addition, sufficient time needs to be given after a rainfall event before collecting measurements to give moisture enough time to become spatially organised.
2. The terrain analysis calculations, particularly TWI, were sensitive to DEM interpolation method. This study found that the thin plate spline interpolation method called AnuDEM (Hutchinson, 2009) provided a better characterisation of soil moisture distribution than kriging.
3. The terrain analysis calculations were also found to be sensitive to DEM grid resolution. A spatial resolution of 1 m was found to be sufficient for taking into account fine scale topographic features such as the test embankment panel membranes.
4. Improvements could be made to the prediction of soil moisture using the terrain analysis metrics by incorporating the calculation of TWI with potential solar radiation using stepwise regression (R^2 of 0.26).
5. A locally fitted model using GWR is better suited to describing the relationship between the terrain analysis metrics and soil moisture than global regression, resulting in an improved in correlation coefficient from 0.26 (global regression) to 0.56 (GWR). This improvement was mainly a result of the spatially variable

relationship between potential solar radiation and observed soil moisture due to the predominantly north and south facing slope aspects.

This study suggests that if the above criteria are followed, an improved characterisation of soil moisture can be made using digital elevation data. Although this study has focussed on the use of high spatial resolution DEMs, the same methodology is transferable at broader scales, outside the context of transport corridor environments. Examples were given in Chapter 2 of this thesis where broad scale DEMs have been used as an input to a range of hydrological and slope stability studies (Baum et al., 2005; Gritzner et al., 2001; Pack et al., 1998). The findings made in this study are likely to have implications for such studies, particularly where the availability of free elevation data is increasing (for example http://edc2.usgs.gov/geodata/index.php) and spaceborne lidar campaigns are reaching near global coverage (Rosette et al., 2008).

7.2.2 Vegetation reflectance

The aim of this experiment was to test the ability of the spectral analysis techniques identified in objective 1 to predict changes in soil moisture. Artificial soil moisture change was successfully induced through a periodic wetting of a small plot on the test embankment. Despite this, there was no obvious change in the vegetation spectra in response to these changes. In contrast, many of the spectral analysis techniques (derivative stress ratio, Lagrangian and linear interpolated red edge position analysis, NDVI) demonstrated good correlations with observed soil moisture (R^2 between 0.57 and 0.66). Additionally, the techniques applied to CASI simulated spectra showed similar correlations with observed soil moisture (R^2 of 0.57 to 0.68). This was an important finding as it demonstrated the potential for scaling these techniques up to an airborne sensor and, therefore, for application to real world transport corridors.

The derivative ratio uses vegetation reflectance values taken at specific wavelengths in order to represent the double-peak feature indicative of vegetation stress (Smith et al., 2004b; Zarco-Tejada et al., 2003). The wavelengths used in this study, 702 nm and 725 nm were taken from Smith et al. (2004b) but were developed for identifying a different source of plant stress using different vegetation types to those used in this study. This study has shown that this specific stress ratio can also be used to

characterise changes in soil moisture over the vegetation on the test embankment (R^2 of 0.61). In addition, this ratio was also tested using wavelengths centred on the default band position of the CASI sensor, 700 nm and 728 nm. Again, good results were found (R^2 of 0.66) indicating this technique can potentially be scaled up to an airborne sensor. These findings are promising for the development of a universal stress ratio although future work may help to further test the robustness of this ratio over a range vegetation types and sources of plant stress.

The novelty in using the spectral analysis techniques is that many have not been studied with direct reference to changes in soil moisture. Despite this, the techniques themselves are not new. In an attempt to further fulfil objective 2 and fully test the potential for using the spectral reflectance of vegetation to characterise soil moisture, a new technique was developed based on PLS regression modelling. It is generally accepted that spectral reflectance measurements should be linearised using the $Log_{10}(1/R)$ transformation prior to PLS regression (Reeves, 2009). However, this study found that this suppressed the signal from the red and near infrared regions of the spectrum, which are indicative of vegetation.

An alternative transformation was explored using the derivative of the spectra following Næs et al. (2002). This maximised the signal from the red and near infrared regions of spectrum, as well as suppressing the normally noisy bands in the blue region. The resulting PLS regression model demonstrated a strong fit with observed soil moisture and was shown to be stable under cross validation, with an R^2 of 0.72. Although this model can only be applicable to the vegetation found on the test embankment, the results show great potential for using such a technique. Furthermore, this model could have a significant bearing for modern highways earthworks, as they tend to be seeded with the same mix of grasses and wildflowers that were used on the test embankment (Glendinning et al., 2009).

7.3 Objective 3

The earthwork scale study sought to achieve objective 2 by providing a proof of concept for the techniques identified for objective 1. In doing so, this helped to set out optimal methods for applying the techniques to airborne remotely sensed dataset for a transport corridor. In addition, there were some techniques that could not be applied to the test

embankment, notably, mapping Ellenberg biological indicator values. Therefore, the reliability of their application could only be based on previous studies and have been tested for the first time in the context of transport corridor environments here. The key findings made when achieving this objective are summarised below.

Similar to the test embankment study, the terrain analysis metrics were found to be sensitive to DEM interpolation. The AnuDEM routine again proved to be the best interpolation method for representing moisture distribution using TWI. This is an important finding, particularly where previous studies, such as Schmidt and Persson (2003), have used other methods (kriging) to interpolate a DEM for the calculation of TWI.

The correlation coefficient between TWI and observed soil moisture could be improved by expressing the TWI as the natural logarithm (from 0.23 to 0.33). This accounted for the non linear increase in soil moisture down earthwork slopes. Although correlations between observed soil moisture and the terrain metrics, TWI (expressed as the natural logarithm) and potential solar radiation, were poor (R^2 of 0.33 and 0.27 respectively), a reasonable correlation could be found when the two were integrated using stepwise regression (R^2 of 0.65). This finding implies that other studies, such as Tenenbaum et al. (2006), could have improved their estimations of soil moisture distribution by incorporating a calculation of potential solar radiation.

The relationship between soil moisture and the terrain calculations could be further improved by using a locally fitted regression model. GWR using the natural logarithm of TWI and potential solar radiation produced a correlation coefficient of 0.76. Similar to the analysis undertaken at the test embankment (Section 7.2.1), this was because the relationship between potential solar radiation and soil moisture distribution varied significantly over space, a result of the north and south facing earthwork slopes.

The observed soil moisture measurements used for drawing comparisons with the remotely sensed metrics were found to be highly variable over space. Specifically, soil moisture was shown to increase non linearly down earthwork slopes, similar to the findings made in the test embankment study (Section 7.2.1). However, some slope stability models characterise soil moisture as a single value for an entire slope. This generalisation is likely to have implications on accurate modelling of the internal structure of earthworks. In addition, observed species abundance within the transport

corridor indicated a heterogeneous vegetation cover with a mixture of grasses, agricultural weeds, wildflowers, trees and shrubs. Existing slope stability models tend to characterise slope vegetation as a single cover and are therefore inadequate for application in transport corridor environments.

Ellenberg indicator values for soil moisture can be mapped in transport corridor environments using airborne CASI data. This is the first reported attempt to do so for values extended for the UK by Hill et al. (2000). By mapping Ellenberg values using imaging spectroscopy this study found that the $Log_{10}(1/R)$ transformation advocated by spectro-chemical studies (Reeves, 2009) is inadequate for application in multivariate vegetation spectral reflectance analysis. This study advises the use of the derivative of the spectra which provides a transformation routine that accentuates the red and near infrared and suppresses the typically noisy blue regions of the spectrum. Subsequent mapping of the Ellenberg values showed that high values coincided with the distribution of hydrophilic (water loving) vegetation types. This could be an important finding for network operators as earthwork inspectors are required to identify such species as indicators of water ingress into earthworks from sources such as leaking ponds or water pipes (Perry et al., 2003b) which may not be detectable using the terrain analysis metrics.

As discussed in Section 6.3.3 the mapping of Ellenberg indicator values has the potential to benefit applications outside the context of slope stability studies. This is largely due to the fact that Ellenberg values provide an indication of soil nitrogen content, pH, soil chloride concentration, light regime, temperature, continentally, as well as soil moisture (Schaffers and Sýkora, 2000). Knowledge of the distribution of these variables could be significant for a range of applications, particularly in the agricultural sector (Knops and Tilman, 2000) and water resources (Kang and Lin, 2007; Alexander et al., 2000).

7.4 Objective 4

The work carried out in order to achieve objective 3 demonstrated that the individual remotely sensed metrics, potential solar radiation, TWI and mapped Ellenberg values, were poor predictors of soil moisture with correlation coefficients below 0.35. This was because the information provided by the individual metrics were insufficient for

describing the complex nature of soil moisture distribution. Therefore, objective 4 was set out to provide a more informed characterisation of soil moisture by incorporating these metrics. The method for incorporating these metrics was stepwise regression, providing a simple model that can be applied outside the model calibration area, as well as ensuring the predicting variables used in the model are statistically significant.

Best results were found using a stepwise regression model incorporating TWI, potential solar radiation and mapped Ellenberg values, producing a correlation coefficient of 0.67. The results of this model demonstrated the influence of potential solar radiation with an increase in soil moisture content predicted for the north facing slopes. The influence of TWI indicated soil moisture to concentrate towards the toe of earthwork slopes. Perhaps more significantly, the TWI also showed areas of topographic convergence where moisture is predicted to contribute to cutting earthworks at specific points from neighbouring fields. In terms of slope stability analysis, this is an important finding as studies have noted the coincidence between areas prone to slope instability and areas of topographic convergence (Pack et al., 1998).

The influence from mapped Ellenberg values was more subtle. For example, there was a coincidence between an increase in observed soil moisture content and the abundance of hydrophilic vegetation (Section 7.3). When this metric was taken out of the integrated model, leaving just the terrain analysis metrics, this information was lost. Despite this, the improvement in correlation coefficient from 0.65 to 0.67 showed that the inclusion of mapped Ellenberg values only slightly improved the overall prediction of soil moisture. Therefore, a reasonable characterisation of soil moisture can be made using just the terrain analysis calculations. These calculations are applied to just one source of remotely sensed data, lidar, which therefore has greater potential for becoming operational due to reduced data acquisition costs and relatively little field work. However, the subtle variations in soil moisture predicted by the Ellenberg values may represent areas of water inundation from sources not detectable by the terrain analysis metrics, such as leaking pipes (Section 6.2.3).

The overall aim of objective 4 by providing an integrated model to characterise soil moisture was achieved using stepwise regression. As discussed, the relative success of this procedure is supported by a reasonable correspondence between the model and observed soil moisture. Stepwise regression was used because it only takes into account

metrics that make a significant improvement to the correlation with observed values without introducing many additional parameters. It is also quick to compute and can be applied as an equation to remotely sensed metrics measured outside the model training area. However, in order to further fulfil this objective different modelling procedures could have been explored. For example, stepwise regression does not provide a thorough measure of model error. As discussed in Section 3.3.3, this could have particular bearing in the present study as error propagation might occur where the remotely sensed metrics with their own levels of error are integrated, therefore compounding subsequent errors. A Bayesian approach to modelling has been discussed as a method to potentially overcome these issues. Although this approach may be interesting to explore, this method was not thoroughly explored here as this type of conditional modelling has the tendency to be less accurate locally (Dungan, 1999) which could be significant when concerned with a component that is highly variable over space such as soil moisture content.

Other modelling procedures cited in Section 3.3.3 include neural networks. This type of modelling may provide a more intelligent and informed model of soil moisture distribution and may well prove valuable to explore in the future. However, such an approach was not considered as although fast computation times are expected for the final model, training times can be very lengthy (Mather, 2004) which may be problematic when dealing with high spatial resolution data over large areas. In addition, models can become 'over trained' losing their effectiveness over areas outside the training area (Mather, 2004).

7.5 Future research

The results presented in this thesis have identified a number of directions for future research. The following provides a summary of these key points.

- Observed soil moisture contents were found to be highly variable over space. Specifically, soil moisture was shown to increase non linearly down earthwork slopes, similar to the findings made in the test embankment study (Section 7.2.1). However, some slope stability models characterise soil moisture as a single value for an entire slope. This generalisation is likely to have implications

on accurate modelling of the internal structure of earthworks and should be addressed in future work.

- Measurements of soil moisture content using the Theta soil moisture probe over fly ash and ballast that has proven to be problematic. Future work should consider the reliability of this piece of equipment, particularly in a transport corridor environment.
- A reasonable characterisation of soil moisture can be made (R^2 of 0.66) over controlled plots on the test embankment using derivative stress ratios developed by Smith et al. (2004b). Future work should be directed at further testing the robustness of this technique over different vegetation and plant stress types.
- A number of spectral analysis studies have employed the $\log_{10}(1/R)$ transformation (Schmidtlein, 2005; Smith et al., 2003; Smith et al., 2002). This study has found mixed results when applying this transformation technique. This uncertainty has been little explored in vegetation spectroscopy and shows that such transformations, which were designed in spectro-chemical studies (Reeves, 2009; Næs et al., 2002), should not necessarily be applied in vegetation analysis. This raises the need for a study to assess the impact of different transformation techniques on the hyperspectral reflectance analysis of vegetation, similar to the review conducted by Reeves (2009) for chemical spectroscopy.
- In addition to soil moisture, Ellenberg values also provide an indication of light, acidity, nitrogen content, and salinity. If these parameters were mapped with similar success to the moisture values in this study, then it would provide a significant tool for a range of land management applications. One such application is in water resources, where the development of the EU Water Framework Directive (Chave, 2001) require river basins to be characterised. Mapping Ellenberg values using imaging spectroscopy could provide an invaluable tool for providing such characterisations over wide areas.
- This study mapped Ellenberg values using airborne CASI imagery. The acquisition of such data can be costly but satellite borne sensors, such as WorldView 2 and Rapid Eye, could provide a cheaper alternative, at the expense of a coarser spatial resolution (1.8 m and 6.5 m respectively).

- The integrated model developed in this study showed reasonable correspondence with observed soil moisture (R^2 of 0.67). This model was created using stepwise regression creating a simple model that could be applied to areas outside the model training area. However, future work should consider different modelling strategies to characterise soil moisture using the remotely sensed metrics to give, for example, a better measure of the model errors. Proposed modelling directions include conditional modelling using Bayesian theory and neural networks.
- A reasonable characterisation of soil moisture has shown to be possible using the terrain analysis calculations (R^2 of 0.65) which are applied to just one source of remotely sensed data. This is an attractive proposition considering the costs associated with airborne data collection. However, the information provided by mapping Ellenberg values may provide crucial information regarding sources of increased soil moisture not detectable by terrain analysis. Potential users of the methods presented in this thesis must carefully consider this point before discounting the collection of CASI imagery and species abundance data for the mapping of Ellenberg values. Future work could help to reinforce these claims by examining imagery over areas of known leaks, similar to Pickerill and Malthus (1998).
- Mapping Ellenberg indicator values was limited to non-forested areas. This is restricting an otherwise potentially important technique for a multitude of applications and should therefore be given careful consideration in a more detailed study using canopy radiative transfer modelling to help improve the signal over complex vegetation canopy structures.
- Network Rail have expressed an interest in using mapped Ellenberg indicator values for identifying blocked or defective drainage. The methods developed in this study can be used to implement this procedure over entire transport networks. This could be achieved by increasing the number of species abundance observations over the whole range of existing vegetation stand types and linking them to airborne hyperspectral imagery using PLS regression. To ensure the success of this method, however, the future work outlined in the previous point should be addressed.

- This study contributes to the wider ‘Remote asset inspection for transport corridor environments’ project. Presently, soil moisture is quantified using a TWI calculation applied to a DEM interpolated using IDW. The results of this study have shown that the calculation of TWI can be improved by applying it to a DEM interpolated using the AunDEM routine and by expressing the metric as the natural logarithm. More significantly, this study has shown that vast improvements can be made to the accuracy of characterising this parameter if an integrated model, using a TWI, potential solar radiation, and mapped Ellenberg values, is used. Future work in this project should therefore be directed at utilising the integrated model designed in this study to refine the weighted assessment of risk.

7.6 Concluding statement

Many of the techniques used in this study have been previously developed. The novelty here is that often these techniques have not been assessed in the context of characterising soil moisture and very few have discussed the applicability to transport corridor environments. In addition, this study has examined multiple sources of remotely sensed data and thought about novel ways of exploiting the data provided by these technologies. In carrying out this work this study has provided a methodology for characterising soil moisture with reasonable accuracy at a high spatial resolution over vegetated areas, two assets that have so far eluded the remote sensing community.

As discussed throughout this thesis, soil moisture is spatially and temporally very variable and is therefore difficult to model. As a result, it is important to develop models that are informed by a number of sources to help represent the numerous processes that govern the distribution of soil moisture. In light of this, a simple statistical model has been developed in this study to integrate remotely sensed metrics to provide an improved characterisation of soil moisture. The emphasis now is to employ this model in the wider ‘Remote asset inspection for transport corridor environments’ project that this study contributes to so that improved assessments of instability risk can be made.

In Section 6.3 reference was made to a quote from Walker et al. (2004) stating the importance of measuring. As a result, the methods developed in this study have the potential to benefit a number of different applications including mapping malarial

mosquito habitats, precision farming, and mapping species richness. Events during the last winter (2009/10) have reiterated the importance of measuring soil moisture in the context of slope stability analysis in UK transport corridors. Specifically, in November 2009 following a period of heavy rainfall there were a number of reports of landslides causing disruption, delays and even a derailment on various UK road and railway networks (BBC, 2009a; 2009b; 2009c). Fortunately nobody was hurt during these events and the only damage caused was economic. However, this reinforces the importance of this study to help target areas of potential instability and prevent these events occurring.

References

Adams, M. L., Philpot, W. D. and Norvell, W. A. (1999) 'Yellowness index: an application of spectral second derivatives to estimate chlorosis of leaves in stressed vegetation', *International Journal of Remote Sensing,* 20, (18), pp. 3663-3675.

Alexander, R. B., Smith, R. A. and Schwarz, G. E. (2000) 'Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico', *Nature,* 403, (6771), pp. 758-761.

Asner, G. P. (1998) 'Biophysical and Biochemical Sources of Variability in Canopy Reflectance', *Remote Sensing of Environment,* 64, pp. 234-253.

Baghdadi, N., Aubert, M., Cerdan, O., Franchistéguy, L., Viel, C., Martin, E., Zribi, M. and Desprats, J. F. (2007) 'Operational mapping of soil moisture using synthetic aperture radar data: application to the Touch Basin (France)', *Sensors,* 7, pp. 2458-2483.

Baghdadi, N., Holah, N. and Zribi, M. (2006) 'Soil moisture estimation using multi-incidence and multi-polarization ASAR data', *International Journal of Remote Sensing,* 27, (10), pp. 1907-1920.

Bannari, A., Morin, D. and Bonn, F. (1995) 'A review of vegetation indices', *Remote Sensing Reviews,* 13, pp. 95-120.

Baranoski, G. V. G. and Rokne, J. G. (2005) 'A practical approach for estimating the red edge position of plant leaf reflectance', *International Journal of Remote Sensing,* 26, (3), pp. 503-521.

Bárdossy, A. and Lehmann, W. (1998) 'Spatial distribution of soil moisture in a small catchment. Part 1: geostatistical analysis', *Journal of Hydrology,* 206, (1-2), pp. 1-15.

Barling, R. D., Moore, I. D. and Grayson, R. B. (1994) 'A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content', *Water Resources Research,* 30, (4), pp. 1029-1044.

Bateson, C. A., Asner, G. P. and Wessman, C. A. (2000) 'Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis', *IEEE Transactions on Geoscience and Remote Sensing,* 38, (2), pp. 1083-1094.

Baum, R. L., Coe, J. A., Godt, J. W., Harp, E. L., Reid, M. E., Savage, W. Z., Schulz, W. H., Brien, D. L., Chleborad, A. F., McKenna, J. P. and Michael, J. A. (2005) 'Regional landslide-hazard assessment for Seattle, Washington, USA', *Landslides,* 2, (4), pp. 266-279.

BBC. (2007) 'Residents trapped by flash flood', [Online]. Available at: http://news.bbc.co.uk/1/hi/england/cumbria/6907783.stm (Accessed: 08/10/09).

BBC. (2009a) 'Landslide derails passenger train', [Online]. Available at: http://news.bbc.co.uk/1/hi/england/dorset/8386354.stm (Accessed: 04/06/10).

BBC. (2009b) 'Landslip prevention work starts ', [Online]. Available at: http://news.bbc.co.uk/1/hi/england/gloucestershire/8385646.stm (Accessed: 04/06/10).

BBC. (2009c) 'Trains suspended due to landslip', [Online]. Available at: http://news.bbc.co.uk/1/hi/england/essex/8322191.stm (Accessed: 04/06/10).

Beguería, S. (2006) 'Changes in land cover and shallow landslide activity: A case study in the Spanish Pyrenees', *Geomorphology*, 74, (1-4), pp. 169-206.

Ben-Dor, E., Patkin, K., Banin, A. and Karniele, A. (2002) 'Mapping of several soil properties using DAIS-7915 hyperspectral scanner data—a case study over clayey soils in Israel', *International Journal of Remote Sensing*, 23, (6), pp. 1043-1062.

Beven, K. J. (1997) 'TOPMODEL: a critique', *Hydrological Processes*, 11, (3), pp. 1069-1085.

Beven, K. J. and Kirkby, M. J. (1979) 'A Physically Based, Variable Contributing Area Model of Basin Hydrology ', *Hydrological Sciences Bulletin* 24, (1), pp. 43-69.

BIONICS (2009) *Biological and engineering impacts of climate change on slopes*. Available at: www.ncl.ac.uk/bionics (Accessed: 27/01/09).

Bishop, T. F. A. and McBratney, A. B. (2002) 'Creating Field Extent Digital Elevation Models for Precision Agriculture', *Precision Agriculture*, 3, (1), pp. 37-46.

Boegh, E., Soegaard, H., Broge, N., Hasager, C. B., Jensen, N. O., Schelde, K. and Thomsen, A. (2002) 'Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture', *Remote Sensing of Environment*, 81, (2-3), pp. 179-193.

Bogh, C., Lindsay, S. W., Clarke, S. E., Dean, A., Jawara, M., Pinder, M. and Thomas, C. J. (2007) 'High spatial resolution mapping of malaria transmission risk in the Gambia, West Africa, using Landsat TM satellite imagery', *The American Journal of Tropical Medicine and Hygiene*, 76, (5), pp. 875-881.

Borga, M., Fontana, G. D. and Cazorzi, F. (2002) 'Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index', *Journal of Hydrology*, 268, (1-4), pp. 56-71.

Boughton, W. (2005) 'Catchment water balance modelling in Australia 1960-2004', *Agricultural Water Management,* 71, (2), pp. 91-116.

Brown, K. M., Duncan, A., O'Dwyer, C., Davison, B., Hogarth, P., Butler, D. and Sampson, E. (2003) *Remote Sensing and Photogrammetry Society Annual General Meeting.* Nottingham, UK, 10-12 September 2003.

Brunsdon, C., Fotheringham, S. and Charlton, M. (1998) 'Geographically weighted regression - modelling spatial non-stationarity', *The Statistician,* 47, (3), pp. 431-443.

Burke, E. J., Gurney, R. J., Simmonds, L. P. and O'Neill, P. E. (1998) 'Using modelling approach to predict soil hydraulic properties from passive microwave measurements', *IEEE Transactions on Geoscience and Remote Sensing,* 36, (2), pp. 454-462.

CAMO (2010) *Multivariate Analysis*. Available at: www.camo.com (Accessed: 18/01/10).

Campbell, J. B. (2006) *Introduction to Remote Sensing.* London: Taylor and Francis.

Carrascal, L. M., Galván, I. and Gordo, O. (2009) 'Partial least squares regression as an alternative to current regression methods used in ecology', *Oikos,* 118, (5), pp. 681-690.

Carter, G. A. (1993) 'Responses of leaf spectral reflectance to plant stress', *American Journal of Botany,* 80, (3), pp. 239-243.

Castleman, K. R. (1996) *Digital Image Processing.* London: Prentice Hall.

Cescatti, A. (1997) 'Modelling the radiative transfer in discontinuous canopies of asymmetric crowns. I. Model structure and algorithms', *Ecological Modelling,* 101, (2-3), pp. 263-274.

Chang, H.-C., Ge, L., Rizos, C. and Milne, T. (2004) 'Validation of DEMs derived from radar interferometry, airborne laser scanning and photogrammetry by using GPS-RTK', *IEEE International Geoscience and Remote Sensing Symposium.* Anchorage, Alaska, 20-24 September. pp. 2815-2818

Chaubey, I., Cotter, A. S., Costello, T. A. and Soerens, T. S. (2005) 'Effect of DEM data resolution on SWAT output uncertainty', *Hydrological Processes,* 19, pp. 621-628.

Chave, P. (2001) *The EU Water Framework Directive: An Introduction.* London: IWA Publishing.

Cho, M. A. and Skidmore, A. K. (2006) 'A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method', *Remote Sensing of Environment,* 101, (2), pp. 181-193.

Clevers, J. G. P. W., Kooistra, L. and Salas, E. A. L. (2004) 'Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data', *International Journal of Remote Sensing,* 25, (19), pp. 3883 - 3895.

Collins, B. D. and Znidarcic, D. (2004) 'Stability Analyses of Rainfall Induced Landslides', *Journal of Geotechnical and Geoenvironmental Engineering,* 130, (4), pp. 362-372.

Cosh, M. H., Jackson, T. J., Bindlish, R., Famiglietti, J. S. and Ryu, D. (2005) 'Calibration of an impedance probe for estimation of surface soil water content over large regions', *Journal of Hydrology,* 311, (1-4), pp. 49-58.

Cracknell, A. P. and Hayes, L. (2007) *Introduction to Remote Sensing.* London: Taylor & Francis.

Crosson, W. L., Limaye, A. S. and Laymon, C. A. (2005) 'Parameter sensitivity of soil moisture retrievals from airborne C- and X-band radiometer measurements in SMEX02', *Geoscience and Remote Sensing, IEEE Transactions on,* 43, (12), pp. 2842-2853.

Crow, W. T., Kustas, W. P. and Prueger, J. H. (2008) 'Monitoring root-zone soil moisture through the assimilation of a thermal remote sensing-based soil moisture proxy into a water balance model', *Remote Sensing of Environment,* 112, (4), pp. 1268-1281.

Curran, P. J. (1988) 'The semivariogram in remote sensing: An introduction', *Remote Sensing of Environment,* 24, (3), pp. 493-507.

Curran, P. J., Dungan, J. L. and Peterson, D. L. (2001) 'Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies', *Remote Sensing of Environment,* 76, (3), pp. 349-359.

Cutler, M. E. J. and Curran, P. J. (1996) 'An observation of shifts in the position of the red edge at different spatial resolutions', *RSS96: Remote Sensing Science and Industry,* pp. 290-297.

Dai, F. C., Lee, C. F. and Ngai, Y. Y. (2002) 'Landslide risk assessment and management: an overview', *Engineering Geology,* 64, (1), pp. 65-87.

Darvishzadeh, R., Skidmore, A., Schlerf, M. and Atzberger, C. (2008) 'Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland', *Remote Sensing of Environment,* 112, (5), pp. 2592-2604.

Davies, A. M. C. (2001) 'Uncertainty testing in PLS regression', *Spectroscopy Europe,* 13, (2), pp. 16-19.

Dawson, T. P. and Curran, P. J. (1998) 'A new technique for interpolating the reflectance red edge position, Technical note', *International Journal of Remote Sensing,* 19, (11), pp. 2133-2139.

Desmet, P. J. J. (1997) 'Effects of Interpolation Errors on the Analysis of DEMs', *Earth Surface Processes and Landforms,* 22, (6), pp. 563-580.

Diekmann, M. (2003) 'Species indicator values as an important tool in applied plant ecology – a review ', *Basic and Applied Ecology,* 4, (6), pp. 493-506.

Diekmann, M. and Falkengren-Grerup, U. (1998) 'A New Species Index for Forest Vascular Plants: Development of Functional Indices Based on Mineralization Rates of Various Forms of Soil Nitrogen', *Journal of Ecology,* 86, (2), pp. 269-283.

Disney, M. I., Lewis, P. and North, P. R. J. (2000) 'Monte Carlo ray tracing in optical canopy reflectance modelling', *Remote Sensing Reviews,* 18, (2), pp. 163 - 196.

Drinkwater, M., McMullan, K., Marti, J., Brown, M., Martín-Neira, M., Rits, W., Ekholm, S., Lemanczyk, J., Kerr, Y., Font, J. and Berger, M. (2009) 'The SMOS payload: MIRAS', (Bulletin 137), [Online]. Available at: (Accessed: 12/02/10).

Du, J., Xie, S., Xu, Y., Xu, C.-y. and Singh, V. P. (2007) 'Development and testing of a simple physically-based distributed rainfall-runoff model for storm runoff simulation in humid forested basins', *Journal of Hydrology,* 336, (3-4), pp. 334-346.

Dungan, J. L. (1999) 'Conditional simulation: An alternative to estimation for achieving mapping objectives', in Stein, A., van der Meer, F. and Gorte, B.(eds) *Spatial Statistics in Remote Sensing.* London: Kluwer Academic Publishers, pp. 135-152.

Ekström, M., Fowler, H. J., Kilsby, C. G. and Jones, P. D. (2005) 'New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 2. Future estimates and use in impact studies ', *Journal of Hydrology,* 300, (1-4), pp. 234-251.

Ellenberg, H., Weber, H. E., Duell, R., Wirth, V., Werner, W. and Paulissen, D. (1991) 'Indicator values of plants in Central Europe', *Scripta Geobotanica* 18, pp. 1-248.

Ertsen, A. C. D., Alkemade, J. R. M. and Wassen, M. J. (1998) 'Calibrating Ellenberg indicator values for moisture, acidity, nutrient availability and salinity in the Netherlands ', *Plant Ecology,* 135, (1), pp. 1573-5052.

ESRI (2007a) *ArcGIS 9.2 Desktop Help*. Available at: http://webhelp.esri.com (Accessed: 30/03/10).

ESRI (2007b) *ArcGIS 9.2 Desktop Help: solar radiation analysis equations.* Available at: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Solar_radiation_analysis_equations (Accessed: 08/04/2010).

Ewald, J. (2003) 'The sensitivity of Ellenberg indicator values to the completeness of vegetation relevés ', *Basic and Applied Ecology,* 4, (6), pp. 507-513.

Ewen, J., Parkin, G. and O'Connell, P. E. (2000) 'SHETRAN: distribution river basin flow and transport modelling system.', *Journal of Hydrological Engineering,* 5, (3), pp. 250-258.

Famiglietti, J. S., Devereaux, J. A., Laymon, C. A., Tsegaye, T., Houser, P. R., Jackson, T. J., Graham, S. T., Rodell, M. and van Oevelen, P. J. (1999) 'Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) Hydrology Experiment', *Water Resources Research,* 35, (6), pp. 1839-1851.

Famiglietti, J. S., Rudnicki, J. W. and Rodell, M. (1998) 'Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas ', *Journal of Hydrology* 210, (1-4), pp. 259-281.

Feldmeyer-Christe, E., Ecker, K., Küchler, M., Graf, U. and Waser, L. (2007) 'Improving predictive mapping in Swiss mire ecosystems through re-calibration of indicator values', *Applied Vegetation Science,* 10, (2), pp. 183-192.

Filella, I. and Penuelas, J. (1994) 'The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status', *International Journal of Remote Sensing,* 15, (7), pp. 1459-1470.

Foody, G. M. (2003) 'Geographical weighting as a further refinement to regression modelling: An example focused on the NDVI-rainfall relationship', *Remote Sensing of Environment,* 88, (3), pp. 283-293.

Fotheringham, S., Brunsdon, C. and Charlton, M. (2002) *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships.* London: John Wiley & Sons.

Fourie, A. B. (1996) 'Predicting rainfall-induced slope instability', *Proceedings from the Institution of Civil Geotechnical Engineering,* 119, (4), pp. 211-218.

Frank, I. E. and Kowalski, B. R. (1984) 'Prediction of Wine Quality and Geographic Origin from Chemical Measurements by Partial Least-Squares Regression Modeling', *Analytica Chimica Acta,* 162, pp. 241-251.

Fredlund, D. G. and Xing, A. (1994) 'Equations for the soil-water characteristic curve', *Canadian Geotechnical Journal,* 31, (3), pp. 521-532.

FSF. (2008) Personal Communication, NERC Field Spectroscopy Facility, Edinburgh, 15th September 2008.

Fu, P. and Rich, P. M. (1999) 'Design and implementation of the Solar Analyst: an ArcView extension for modeling solar radiation at landscape scales', *Nineteenth Annual ESRI User Conference.* San Diego, USA, pp. 1-33.

Fu, P. and Rich, P. M. (2002) 'A geometric solar radiation model with applications in agriculture and forestry', *Computers and Electronics in Agriculture,* 37, (1-3), pp. 25-35.

Fukuoka, M. (1980) 'Landslides associated with rainfall', *Geotechnical Engineering,* 11, pp. 1-29.

Glendinning, S., Loveridge, F., Starr-Keddle, R. E., Bransby, M. F. and Hughes, P. N. (2009) 'Role of vegetation in sustainability of infrastructure slopes', *Proceedings of the ICE - Engineering Sustainability,* 162, (2), pp. 101-110.

Goldberg, D. E. and Miller, T. E. (1990) 'Effects of Different Resource Additions of Species Diversity in an Annual Plant Community', *Ecology,* 71, (1), pp. 213-225.

Grayson, R. B., Western, A. W., Chiew, F. H. S. and Bloeschl, G. (1997) 'Preferred states in spatial soil moisture patterns: Local and non-local controls ', *Water Resources Research,* 33, (12), pp. 2897.

Greenwood, J. R. (2006) 'SLIP4EX - A program or routine slope stability analysis to include the effects of vegetation, reinforcement and hydrological changes', *Geotechnical and Geological Engineering,* 24, (3), pp. 449-465.

Grime, J. P. (2001) *Plant strategies, vegetation processes, and ecosystem properties. Second edition.* Chichester: John Wiley & Sons.

Gritzner, M. L., Marcus, A. W., Aspinall, R. and Custer, S. G. (2001) 'Assessing landslide potential using GIS, soil wetness modelling and topographic attributes, Payette River, Idaho', *Geomorphology,* 37, (1-2), pp. 149-165.

Guardian (2003) *Train hits landslide after torrential rain*. Available at: http://www.guardian.co.uk/environment/2003/jan/02/weather.transportintheuk (Accessed: 18/08/2010).

Gyasi-Agyei, Y., Sibley, J. and Ashwath, N. (2001) 'Quantitative evaluation of strategies for erosion control on a railway embankment batter', *Hydrological Processes,* 15, pp. 3249-3268.

Hall, S. (2009) Personal communication, Network Rail, York, 11/11/09.

Hancock, S., Lewis, P., Disney, M. I., Foster, M. and Muller, J.-P. (2008) 'Assessing the accuracy of forest height estimation with long pulse waveform lidar through Monte-Carlo ray tracing', *SilviLaser*. Edinburgh, UK, September 17-19th, 2008. pp. 199-206.

Haugerud, R. A. and Harding, D. J. (2001) 'Some algorithms for virtual deforestation (VDF) of LiDAR topographic survey data', *International Archives of Photogrammetry and Remote Sensing*, 34, (3/W4), pp. 714-719.

Hawley, M. E., Jackson, T. J. and McCuen, R. H. (1983) 'Surface soil moisture variation on small agricultural watersheds', *Journal of Hydrology*, 62, (1-4), pp. 179-200.

Hill, M. O., Mountford, J. O., Roy, D. B. and Bunce, R. G. H. (1999) *ECOFACT 2a Technical Annex - Ellenberg's indicator values for British Plants*. Available at: http://www.ceh.ac.uk/products/publications/untitled.html (Accessed: 19/07/2010).

Hill, M. O., Roy, D. B., Mountford, J. O. and Bunce, R. G. H. (2000) 'Extending Ellenberg's indicator values to a new area: an algorithmic approach', *Journal of Applied Ecology*, 37, (1), pp. 3-15.

Hodgson, M. E. and Bresnahan, P. (2004) 'Accuracy of airborne lidar-derived elevation: empirical assessment and error budget', *Photogrammetric Engineering and Remote Sensing*, 70, (3), pp. 331-339.

Hollaus, M., Wagner, W. and Kraus, K. (2005) 'Airborne laser scanning and usefulness for hydrological models', *Advances in Geosciences*, 5, pp. 57-63.

Hollis, G. E. (1988) 'Rain, roads, roofs and runoff: Hydrology in cities', *Geography*, 73, pp. 9-18.

Huang, Z.-H., Zhou, G.-Y., Zhou, G.-Y., Morris, J., Silberstein, R. and Wang, X. (2005) 'Terrain analysis and steady-state hydrological modelling of a small catchment in southern China', *Ecology and Environment*, 14, (5), pp. 700-705.

Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R. and Foley, W. J. (2004) 'Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis', *Remote Sensing of Environment*, 93, (1-2), pp. 18-29.

Hughes, P. N., Glendinning, S., Toll, D. G., Gallipoli, D., Parkin, G. and Mendes, J. (2009) 'Full-scale testing to assess climate effects on embankments', *Proceedings from the Institution of Civil Engineers, Engineering Sustainability*, 162, (2), pp. 67-79.

Hutchinson, M. F. (1989) 'A new procedure for gridding elevation and stream line data with automatic removal of spurious pits', *Journal of Hydrology*, 106, (3-4), pp. 211-232.

Hutchinson, M. F. (2008) 'Adding the Z Dimension', in Wilson, J. P. and Fotheringham, A. S.(eds) *The Handbook of Geographic Information Science.* Oxford: Blackwell, pp. 144-168.

Hutchinson, M. F. (2009) *ANUDEM Version 5.2*. Available at: http://fennerschool.anu.edu.au/publications/software/anudem.php (Accessed: 27/07/2009).

Hyde, P., Dubayah, R., Peterson, B., Blair, J. B., Hofton, M., Hunsaker, C., Knox, R. and Walker, W. (2005) 'Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems', *Remote Sensing of Environment,* 96, (3-4), pp. 427-437.

Isard, S. A. (1986) 'Factors influencing soil moisture and plant community distribution on Niwot Ridge, Front Range, Colorado, U.S.A.', *Arctic and Alpine Research,* 18, (1), pp. 83-96.

Jackson, T. J., Le Vine, D. M., Hsu, A. Y., Oldak, A., Starks, P. J., Swift, C. T., Isham, J. D. and Haken, M. (1999) 'Soil moisture mapping at regional scale using passive microwave radiometry: the Southern Great Plains hydrology experiment', *IEEE Transactions on Geoscience and Remote Sensing*, 37, (5), pp. 2136-2151.

Jackson, T. J., Le Vine, D. M., Swift, C. T., Schmugge, T. J. and Schiebe, F. R. (1995) 'Large area mapping of soil moisture using the ESTAR passive microwave radiometer in Washita'92', *Remote Sensing of Environment*, 54, (1), pp. 27-37.

Jacquemoud, S., Bacour, C., Poilvé, H. and Frangi, J.-P. (2000) 'Comparison of four radiative transfer models to simulate plant canopies reflectance: direct and inverse mode ', *Remote Sensing of Environment*, 74, (3), pp. 471-481.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C. and Ustin, S. L. (2009) 'PROSPECT + SAIL models: A review of use for vegetation characterization', *Remote Sensing of Environment*, 113, (Supplement 1), pp. S56-S66.

Johnson, K. A. and Sitar, N. (1990) 'Hydrologic conditions leading to debris-flow initiation', *Canadian Geotechnical Journal*, 27, (6), pp. 789-801.

Kang, S. and Lin, H. (2007) 'Wavelet analysis of hydrological and water quality signals in an agricultural watershed', *Journal of Hydrology*, 338, (1-2), pp. 1-14.

Karam, M. A., Fung, A. K., Lang, R. H. and Chauhan, N. S. (1992) 'A microwave scattering model for layered vegetation', *IEEE Transactions on Geoscience and Remote Sensing*, 30, (4), pp. 767-784.

Kempeneers, P., De Backer, S., Debruyn, W., Coppin, P. and Scheunders, P. (2005) 'Generic wavelet-based hyperspectral classification applied to vegetation stress detection', *IEEE Transactions on Geoscience and Remote Sensing* 43, (3), pp. 610-614.

Kent, M. and Coker, P. (1995) *Vegetation Description and Analysis.* Chichester: John Wiley & Sons.

Kerr, Y. (2007) 'Soil moisture from space: Where are we?', *Hydrogeology Journal*, 15, (1), pp. 117-120.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J., Font, J. and Berger, M. (2001) 'Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission', *IEEE Transactions on Geoscience and Remote Sensing*, 39, (8), pp. 1729-1735.

Kilsby, C., Glendinning, S., Hughes, P. N., Parkin, G. and Bransby, M. F. (2009) 'Climate-change impacts on long-term performance of slopes', *Engineering Sustainability*, 162, (2), pp. 59-66.

Kim, S. and Jung, S. (2003) 'Digital terrain analysis of the dynamic wetness pattern on the Sulmachun watershed', *Diffuse Pollution Conference.* Dublin, Ireland, pp. 1-6.

Knops, J. M. H. and Tilman, D. (2000) 'Dynamics of soil nitrogen and carbon accumulation for 61 years after agricultural abandonment', *Ecology*, 81, (1), pp. 88-98.

Kokaly, R. F. and Clark, R. N. (1999) 'Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression', *Remote Sensing of Environment*, 67, (3), pp. 267-287.

Kumar, L., Schmidt, K., Dury, S. J. and Skidmore, A. K. (2006) 'Imaging spectroscopy and vegetation science', in van der Meer, F. and de Jong, S.(eds) *Imaging Spectroscopy: Basic Principals and Prospective Applications.* Dordrecht, Netherlands: Springer, pp. 111-154.

Kuria, D. N., Koike, T., Hui, L., Tsutsui, H. and Graf, T. (2007) 'Field-Supported Verification and Improvement of a Passive Microwave Surface Emission Model for Rough, Bare, and Wet Soil Surfaces by Incorporating Shadowing Effects', *IEEE Transactions on Geoscience and Remote Sensing*, 45, (5), pp. 1207-1216.

Kwak, Y., Jang, Y. and Kang, I. (2006) 'Web GIS management and risk evaluation of a road slope using a terrestrial LiDAR', *Web and wireless geographical information systems*, 3833, pp. 256-266.

Le Hegarat-Mascle, S., Zribi, M., Alem, F., Weisse, A. and Loumagne, C. (2002) 'Soil moisture estimation from ERS/SAR data: toward an operational methodology', *IEEE Transactions on Geoscience and Remote Sensing,* 40, (12), pp. 2647-2658.

Lewis, M. M. (1998) 'Numeric classification as an aid to spectral mapping of vegetation communities', *Plant Ecology,* 136, (2), pp. 133-149.

Li, L., Ustin, S. L. and Lay, M. (2005) 'Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, New Mexico', *Remote Sensing of Environment,* 94, (1), pp. 1-16.

Lihua, Z., Minzan, L., Jianying, S., Ning, T. and Xijie, Z. (2005) 'Estimation of soil moisture with aerial images and hyperspectral data', *IEEE International Geoscience & Remote Sensing Symposium* Seoul, Korea, 25-29 July. pp. 4516-4519.

Lillesand, T. M. and Kiefer, R. W. (2000) *Remote sensing and image interpretation.* Chichester: John Wiley & Sons.

Lim, M., Mills, J. P., Barr, S. L., Barber, D., Glendinning, S., Parkin, G., Hall, J. and Clarke, B. (2007) 'High resolution earth imaging for transport corridor slope stability risk analysis', *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* 36 (1/W51), pp. 1-6.

Lletí, R., Ortiz, M. C., Sarabia, L. A. and Sánchez, M. S. (2004) 'Selecting variables for k-means cluster analysis by using a genetic algorithm that optimises the silhouettes', *Analytica Chimica Acta,* 515, (1), pp. 87-100.

Lloyd, D. M., Anderson, M. G., Hussein, A. N., Jamaludin, A. and Wilkinson, P. L. (2001) 'Preventing landslides on roads and railways: a new risk-based approach', *Proceedings of ICE Civil Engineering,* 144, (3), pp. 129-134.

Lobell, D. B. and Asner, G. P. (2002) 'Moisture effects of soil reflectance', *Soil Science Society of America Journal,* 66, (3), pp. 722-727.

Lu, S., Ren, T., Gong, Y. and Horton, R. (2007) 'An improved model for predicting soil thermal conductivity from water content at room temperature', *Soil Science Society of America Journal,* 71, (1), pp. 8-14.

Luoto, M., Toivonen, T. and Heikkinen, R. (2002) 'Prediction of total and rare plant species richness in agricultural landscapes from satellite images and topographic data', *Landscape Ecology,* 17, (3), pp. 195-217.

Malthus, T. J. and Madeira, A. C. (1993) 'High resolution spectroradiometry: spectral reflectance of field bean leaves infected by Botrytis fabae', *Remote Sensing of Environment,* 45, (1), pp. 107-116.

Manning, L. J., Hall, J. W., Kilsby, C. G., Glendinning, S. and Andrerson, M. G. (2008) 'Spatial analysis of the reliability of transport networks subject to rainfall-induced landslides', *Hydrological Processes*, 22, pp. 3349-3360.

Martens, H. and Martens, M. (2000) 'Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR)', *Food Quality and Preference,* 11, (1-2), pp. 5-16.

Massant, W., Godefroid, S. and Koedam, N. (2009) 'Clustering of plant life strategies on meso-scale', *Plant Ecology,* 205, (1), pp. 47-56.

Mather, P. M. (2004) *Computer processing of remotely-sensed images – an introduction. Third edition.*Third ed Chichester: John Wiley and Sons.

McBratney, A. B. and Webster, R. (1983) 'How many observations are needed for regional estimation of soil properties?', *Soil Science* 135, (3), pp. 177-183.

McBratney, A. B. and Webster, R. (1986) 'Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates', *European Journal of Soil Science,* 37, (4), pp. 617-639.

McGinnity, B. T., Fitch, R. and Rankin, W. J. (1998) *Institution of Civil Engineers. Proceedings of the seminar "Value of Geotechnics in Construction".* London:Construction Research Communications Limited.

MEN (2008) *Landslide blocks rail line*. Available at: http://www.manchestereveningnews.co.uk/news/s/1076088_landslide_blocks_rail_line?rss=yes (Accessed: 03/07/2010).

Meroni, M., Colombo, R. and Panigada, C. (2004) 'Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations', *Remote Sensing of Environment*, 92, (2), pp. 195-206.

Merton, R. (1998) *Seventh annual JPL airborne Earth Science Workshop.* Pasadena, California, 12-16 January, 1998.NASA.

Mickoviski, S. B. and Van Beek, L. P. H. (2006) 'A decision support system for the evaluation of eco-engineering strategies for slope protection', *Geotechnical and Geological Engineering,* 24, pp. 483-498.

Miller, P. E., Hardy, A. J., Mills, J. P., Barr, S. L., Birkinshaw, S. J., Parkin, G., Glendinning, S. and Hall, J. W. (2009) 'Intelligent integration of multi-sensor data for risk assessment in transport corridor environments', *ISPRS Workshop: High-Resolution Earth Imaging for Geospatial Information.* Hannover, Germany, pp. 1-6.

Miller, P. E., Mills, J. P., Barr, S. L., Lim, M., Barber, D., Parkin, G., Clarke, B., Glendinning, S. and Hall, J. (2008) 'Terrestrial laser scanning for assessing the risk of slope instability along transport corridors', *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* 37 (B5), pp. 495-500.

Milton, E. J. and Choi, K. Y. (2004) 'Estimating the spectral response function of the CASI-2', *Remote Sensing and Photogrammetry Society Annual General Meeting.* Aberdeen, UK, 7-10 September 2004. pp. 1-11.

Minacapilli, M., Iovino, M. and Blanda, F. (2009) 'High resolution remote estimation of soil surface water content by a thermal inertia approach', *Journal of Hydrology,* 379, (3-4), pp. 229-238.

Montillet, J. P., Meng, X., Roberts, G. W., Taha, A., Hancock, C. and Ogundipe, O. (2007) 'Achieving centimeter-level positional accuracy in urban canyons with Locata technology', *Journal of Global Positioning Systems,* 6, (2), pp. 158-165.

Moran, M. S., Peters-Lidard, D., Watts, J. M. and McElroy, S. (2004) 'Estimating soil moisture at the watershed scale with satellite-based radar and land surface models', *Canadian Journal of Remote Sensing,* 30, (5), pp. 805-826.

Mueller-Dombois, D. and Ellenberg, H. (1974) *Aims and Methods of Vegetation Ecology.* London: John Wiley & Sons.

Muller, E. and Decamps, H. (2001) 'Modeling soil moisture-reflectance', *Remote Sensing of Environment,* 76, (2), pp. 173-180.

Mushinzimana, E., Munga, S., Minakawa, N., Li, L., Feng, C.-c., Bian, L., Kitron, U., Schmidt, C., Beck, L., Zhou, G., Githeko, A. and Yan, G. (2006) 'Landscape determinants and remote sensing of anopheline mosquito larval habitats in the western Kenya highlands', *Malaria Journal,* 5, (1), pp. 13.

Mutanga, O., Skidmore, A. K., Kumar, L. and Ferwerda, J. (2005) 'Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain', *International Journal of Remote Sensing,* 26, (6), pp. 1093 - 1108.

Mutuku, F. M., Alaii, J. A., Bayoh, M. N., Gimnig, J. E., Vulule, J. M., Walker, E. D., Kabiru, E. and Hawley, W. A. (2006) 'Distribution, description, and local knowledge of larval habitats of Anopheles gambiae S.L. in a village in western Kenya', *American Journal of Tropical Medicine and Hygiene,* 74, (1), pp. 44-53.

Mutuku, F. M., Bayoh, M. N., Hightower, A. W., Vulule, J. M., Gimnig, J. E., Mueke, J. M., Amimo, F. A. and Walker, E. D. (2009) 'A supervised land cover classification of a western Kenya lowland endemic for human malaria: associations of land cover with larval Anopheles habitats', *International Journal of Health Geographics,* 8, (1), pp. 19-32.

Myneni, R. B., Hall, F. G., Sellers, P. J. and Marshak, A. L. (1995) 'The interpretation of spectral vegetation indexes', *IEEE Transactions on Geoscience and Remote Sensing,* 33, (2), pp. 481-486.

Næs, T., Isaksson, T., Fearn, T. and Davies, T. (2002) *A User-Friendly Guide to Multivariate Calibration and Classification.* Chichester: NIR Publications.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K. and Nghiem, S. V. (2003) 'Soil moisture retrieval from AMSR-E', *IEEE Transactions on Geoscience and Remote Sensing,* 41, (2), pp. 215-229.

Noomen, M. F., Skidmore, A. K., van der Meer, F. D. and Prins, H. H. T. (2006) 'Continuum removed band depth analysis for detecting the effects of natural gas, methane and ethane on maize reflectance ', *Remote Sensing of Environment,* 105, (3), pp. 262-270.

Notarnicola, C., Angiulli, M. and Posa, F. (2006) 'Use of radar and optical remotely sensed data for soil moisture retrieval over vegetated areas', *IEEE Transactions on Geoscience and Remote Sensing,* 44, (4), pp. 925-935.

Pack, R. T., Tarboton, D. G. and Goodwin, C. N. (1998) 'The SINMAP approach to terrain stability mapping', *Proceedings Eighth International Congress - International Association of Engineering Geology.* Vancouver, British Columbia, 21-25 September, 1998. pp. 1157-1165.

Perry, J., Pedley, M., Brady, K. and Reid, M. (2003a) 'Embankment cuttings: condition appraisal and remedial treatment', *Proceedings of the Institute of Civil Engineers,* 156, (4), pp. 171-175.

Perry, J., Pedley, M. and Reid, M. (2003b) *Infrastructure embankments - condition appraisal and remedial treatment.* London: CIRIA.

Phipps, P. J. and McGinnity, B. T. (2001) 'Classification and stability assessment for chalk cuttings: the Metropolitan Line case study', *Quarterly Journal of Engineering Geology and Hydrogeology,* 34, (4), pp. 353-370.

Pickerill, J. M. and Malthus, T. J. (1998) 'Leak detection from rural aqueducts using airborne remote sensing techniques', *International Journal of Remote Sensing,* 19, (12), pp. 2427-2433.

Pierdicca, N., Pulvirenti, L. and Bignami, C. (2010) 'Soil moisture estimation over vegetated terrains using multitemporal remote sensing data', *Remote Sensing of Environment,* 114, (2), pp. 440-448.

Porporato, A., Laio, F., Ridolfi, L. and Rodriguez-Iturbe, I. (2001) 'Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress III. Vegetation water stress', *Advances in Water Resources,* 24, (7), pp. 725-744.

Preuth, T., Glade, T. and Demoulin, A. (2010) 'Stability analysis of a human-influenced landslide in eastern Belgium', *Geomorphology,* 120, (1-2), pp. 38-47.

Quesney, A., Le Hégarat-Mascle, S., Taconet, O., Vidal-Madjar, D., Wigneron, J. P., Loumagne, C. and Normand, M. (2000) 'Estimation of watershed soil moisture index from ERS/SAR data', *Remote Sensing of Environment,* 72, (3), pp. 290-303.

Rajkai, K. and Rydén, B. E. (1992) 'Measuring areal soil moisture distribution with the TDR method', *Geoderma,* 52, (1-2), pp. 73-85.

Rao, Y. S. (1992) *Microwave remote sensing of soil moisture*. Available at: http://www.csre.iitb.ac.in/ysrao/m_tech_civil1.pdf (Accessed: 16/03/10).

Ray, R. and Jacobs, J. (2007) 'Relationships among remotely sensed soil moisture, precipitation and landslide events', *Natural Hazards,* 43, (2), pp. 211-222.

Rees, W. G. (2000) 'The accuracy of Digital Elevation Models interpolated to higher resolutions', *International Journal of Remote Sensing,* 21, (1), pp. 7-20.

Reeves, J. B. (2009) 'Does the Spectral Format Matter in Diffuse Reflection Spectroscopy?', *Applied Spectroscopy,* 63, (6), pp. 669-677.

Reutebuch, S. E., McGaughey, R. J., Andersen, H.-E. and Carson, W. W. (2003) 'Accuracy of high-resoltuion lidar terrain model under a conifer forest canopy', *Canadian Journal of Remote Sensing,* 29, (5), pp. 527-535.

Ridley, A., McGunnity, B. and Vaughan, P. (2004) 'Role of pore water pressures in embankment stability', *Geotechnical Engineering,* 157, (4), pp. 193-198.

Rodriguez-Iturbe, I., Porporato, A., Laio, F. and Ridolfi, L. (2001) 'Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress I. Scope and general outline', *Advances in Water Resources,* 24, (7), pp. 295-705.

Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E. and Goldstein, R. M. (2000) 'Synthetic aperture radar interferometry', *Proceedings of the IEEE,* 88, (3), pp. 333-382.

Rosette, J. A. B., North, P. R. J. and Suárez, J. C. (2008) 'Vegetation height estimates for a mixed temperate forest using satellite laser altimetry', *International Journal of Remote Sensing,* 29, (5), pp. 1475 - 1493.

Rouainia, M., Davis, O., O'Brien, T. and Glendinning, S. (2009) 'Numerical modelling of climate effects on slope stability', *Engineering Sustainability,* 162, (2), pp. 81-89.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W. and Harlan, J. C. (1974) *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation.* Greenbelt, Maryland: NASA/GSFC.

Rouse, J. W., Smith, M. O. and Adams, J. B. (1973) 'Monitoring vegetation systems in the great plains with ERTS', *Third ERTS Symposium, NASA SP,* 351, pp. 39-317.

Sahoo, A. K., Houser, P. R., Ferguson, C., Wood, E. F., Dirmeyer, P. A. and Kafatos, M. (2008) 'Evaluation of AMSR-E soil moisture results using the in-situ data over the Little River Experimental Watershed, Georgia', *Remote Sensing of Environment,* 112, (6), pp. 3142-3152.

Sandholt, I., Rasmussen, K. and Andersen, J. (2002) 'A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status', *Remote Sensing of Environment,* 79, (2-3), pp. 213-224.

Schaffers, A. P. and Sýkora, K. V. (2000) 'Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: a comparison with field measurements', *Journal of Vegetation Science,* 11, (3), pp. 225-244.

Scheiner, S. M. (2003) 'Six types of species-area curve', *Global Ecology & Biogeography,* 12, (6), pp. 441-447.

Schmidt, F. and Persson, A. (2003) 'Comparison of DEM Data Capture and Topographic Wetness Indices ', *Precision Agriculture,* 4, (2), pp. 179-192.

Schmidtlein, S. (2005) 'Imaging spectroscopy as a tool for mapping Ellenberg indicator values', *Journal of Applied Ecology,* 42, (5), pp. 966-974.

Schmidtlein, S. and Sassin, J. (2004) 'Mapping of continuous floristic gradients in grasslands using hyperspectral imagery ', *Remote Sensing of Environment,* 92, (1), pp. 126-138.

Schmugge, T. and Jackson, T. J. (1994) 'Mapping surface soil moisture with microwave radiometers', *Meteorology and Atmospheric Physics,* 54, (1), pp. 213-223.

Schmugge, T. J., Kustas, W. P., Ritchie, J. C., Jackson, T. J. and Rango, A. (2002) 'Remote sensing in hydrology', *Advances in Water Resources,* 25, (8-12), pp. 1367-1385.

Schowengerdt, R. A. (2007) *Remote Sensing: models and methods for image processing.*Third ed London: Elsevier.

Schulz, W. H. (2005) 'Landslide susceptibility estimated from mapping using light detection and ranging (LIDAR) imagery and historical landslide records, Seattle, Washington', *Engineering Geology,* 89, (1-2), pp. 67-87.

Sharkov, E. A. (2003) *Passive Microwave Remote Sensing of the Earth: Physical Foundations.* Chichester: Praxis Publishing Ltd.

Shi, J., Jiang, L., Zhang, L., Chen, K.-S., Wigneron, J. P. and Chanzy, A. (2005) 'A parameterized multifrequency-polarization surface emission model', *IEEE Transactions on Geoscience and Remote Sensing,* 43, (12), pp. 2831-2841.

Sithole, G. and Vosselman, G. (2004) 'Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds', *ISPRS Journal of Photogrammetry and Remote Sensing,* 59, (1-2), pp. 85-101.

Smith, G. M. and Milton, E. J. (1999) 'The use of the empirical line method to calibrate remotely sensed data to reflectance', *International Journal of Remote Sensing,* 20, (13), pp. 2653-2662.

Smith, K. L., Colls, J. J. and Steven, M. D. (2005a) 'A facility to investigate effects of elevated soil gas concentration on vegetation', *Water, Air, and Soil Pollution,* 161, (1), pp. 75-96.

Smith, K. L., Steven, M. D. and Colls, J. J. (2004a) 'Spectral responses of pot-grown plants to displacement of soil oxygen', *International Journal of Remote Sensing,* 25, (20), pp. 4395-4410.

Smith, K. L., Steven, M. D. and Colls, J. J. (2004b) 'Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks', *Remote Sensing of Environment,* 92, (2), pp. 207-217.

Smith, K. L., Steven, M. D. and Colls, J. J. (2005b) 'Plant spectral responses to gas leaks and other stresses', *International Journal of Remote Sensing,* 26, (18), pp. 4067-4081.

Smith, M.-L., Ollinger, S. V., Martin, M. E., Aber, J. D., Hallett, R. A. and Goodale, C. L. (2002) 'Direct Estimation of Aboveground Forest Productivity through Hyperspectral Remote Sensing of Canopy Nitrogen', *Ecological Applications,* 12, (5), pp. 1286-1302.

Smith, M. L., Martin, M. E., Plourde, L. C. and Ollinger, S. V. (2003) 'Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor', *IEEE Transactions on Geoscience and Remote Sensing,* 41, (6), pp. 1332-1337.

Sørensen, R., Zinko, U. and Seibert, J. (2006) 'On the calculation of the topographic wetness index: evaluation of different methods based on field observations', *Hydrology and Earth System Sciences*, 10, (1), pp. 101-112.

Stohlgren, T. J., Falkner, M. B. and Schell, L. D. (1995) 'A Modified-Whittaker nested vegetation sampling method', *Plant Ecology*, 117, (2), pp. 113-121.

Tarboton, D. G. (1997) 'A New Method for the Determination of Flow Directions and Upslope Areas in Grid Digital Elevation Models', *Water Resources Research*, 33, (2), pp. 309-319.

Tarboton, D. G. (2000) *TARDEM, a suite of programs for the analysis of digital elevation data*. Available at: http://www.engineering.usu.edu/cee/faculty/dtarb/tardem.html (Accessed: 01/09/2009).

Tarboton, D. G. (2002) *Terrain analysis using digital elevation models (TauDEM)*. Available at: http://hydrology.neng.usu.edu/taudem/ (Accessed: 22/01/09).

Tenenbaum, D. E., Band, L. E., Kenworthy, S. T. and Tague, C. L. (2006) 'Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets', *Hydrological Processes*, 20, (2), pp. 219-240.

ter Braak, C. J. F. and Prentice, I. C. (1988) 'A theory of gradient analysis', *Advances in Ecological Research*, 18, pp. 271-317.

Thenkabail, P. S., Smith, R. B. and De Pauw, E. (2000) 'Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics', *Remote Sensing of Environment*, 71, (2), pp. 158-182.

Townsend, P. A., Foster, J. R., Chastain, R. A. and Currie, W. S. (2003) 'Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS', *IEEE Transactions on Geoscience and Remote Sensing*, 41, (6), pp. 1346-1354.

Trauth, M. H. (2007) *MATLAB Recipes for Earth Sciences*. Berlin: Springer.

Trought, M. C. T. and Drew, M. C. (1980) 'The development of waterlogging damage in wheat seedlings (Triticum aestivum L.) I. Shoot and roots growth in relation to changes in the concentrations of dissolved gases and solutes in the soil solution', *Plant and soil*, 54, (1), pp. 77-94.

Tsai, F. and Philpot, W. (1998) 'Derivative analysis of hyperspectral data', *Remote Sensing of Environment*, 66, (1), pp. 41-51.

Ulaby, F. T., Dubois, P. C. and van Zyl, J. (1996) 'Radar mapping of surface soil moisture', *Journal of Hydrology*, 184, (1-2), pp. 57-84.

Ustin, S. L., Smith, M. O., Jacquemoud, S., Vertraete, M. and Govaerts, Y. (1999) 'Geobotany: vegetation mapping for earth sciences', in Rencz, A. N.(ed), *Remote sensing for the earth sciences, Manual of Remote Sensing.* New York: John Wiley & Sons, pp. 189-248.

van der Meer, F., de Jong, S. and Bakker, W. (2006) 'Imaging spectroscopy: basic analytical techniques', in van der Meer, F. and de Jong, S.(eds) *Imaging Spectroscopy: Basic Principals and Prospective Applications.* Dordrecht, Netherlands: Springer, pp. 17-61.

Verstraeten, W. W., Veroustraete, F., Sande, C., Grootaers, I. and Feyen, J. (2006) 'Soil moisture retrieval using thermal inertia, determined with visible and thermal spaceborne data, validated for European forests', *Remote Sensing of Environment,* 101, (3), pp. 299-314.

Vicente-Serrano, S. M., Pons-Fernández, X. and Cuadrat-Prats, J. M. (2004) 'Mapping soil moisture in the central Ebro river valley (northeast Spain) with Landsat and NOAA satellite imagery: a comparison with meteorological data', *International Journal of Remote Sensing,* 25, (20), pp. 4325 - 4350.

Wadsworth, R. A., Collingham, Y. C., Willis, S. G., Huntley, B. and Hulme, P. E. (2000) 'Simulating the Spread and Management of Alien Riparian Weeds: Are They Out of Control?', *Journal of Applied Ecology,* 37, (1), pp. 28-38.

Walker, J. P., Willgoose, G. R. and Kalma, J. D. (2004) 'In situ measurement of soil moisture: a comparison of techniques', *Journal of Hydrology,* 293, (1-4), pp. 85-99.

Wang, Q., Ni, J. and Tenhunen, J. (2005) 'Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems', *Global Ecology and Biogeography,* 14, (4), pp. 379-393.

Weidong, L., Baret, F., Xingfa, G., Qingxi, T., Lanfen, Z. and Bing, Z. (2002) 'Relating soil surface moisture to reflectance', *Remote Sensing of Environment,* 81, (2-3), pp. 238-246.

Werninghaus, R., Balzer, W., Buckreuss, S., Mittermayer, J. and Mühlbauer, P. (2002) 'The TerraSAR-X-mission*', 5th European Conference on Synthetic Aperture Radar, Eusar2004.* Cologne, Germany, pp. 49-52.

Western, A. W., Blöschl, G. and Grayson, R. B. (1998) 'How well do indicator variograms capture the spatial connectivity of soil moisture?', *Hydrological Processes,* 12, (12), pp. 1851-1868.

Western, A. W., Grayson, R. B., Blöschl, G., Willgoose, G. R. and McMahon, T. A. (1999a) 'Observed spatial organization of soil moisture and its relation to terrain indices', *Water Resources Research,* 35, (3), pp. 797-810.

Western, A. W., Grayson, R. B. and Green, T. R. (1999b) 'The Tarrawarra project: high resolution spatial measurement, modelling and analysis of soil moisture and hydrological response', *Hydrological Processes,* 13, (5), pp. 633-652.

Whiting, M. L., Li, L. and Ustin, S. L. (2004) 'Predicting water content using Gaussian model on soil spectra', *Remote Sensing of Environment,* 89, (4), pp. 535-552.

Wigneron, J.-P., Calvet, J.-C., Pellarin, T., Van de Greind, A. A., Berger, M. and Ferrazzoli, P. (2003) 'Retrieving near-surface soil moisture from microwave radiometric observations: current status and future plans', *Remote Sensing of Environment,* 85, (4), pp. 489-506.

Wigneron, J.-P., Schmugge, T., Chanzy, A., Calvet, J.-C. and Kerr, Y. (1998) 'Use of passive microwave remote sensing to monitor soil moisture', *Agronomie,* 18, (1), pp. 27-43.

Wilson, J. P. and Gallant, J. C. (2000) *Terrain Analysis: Principals and Applications.* Chichester: John Wiley & Sons.

Wolf, A., Berry, J. A. and Asner, G. P. (2010) 'Allometric constraints on sources of variability in multi-angle reflectance measurements', *Remote Sensing of Environment,* 114, (6), pp. 1205-1219.

Wolter, P. T., Townsend, P. A., Sturtevant, B. R. and Kingdon, C. C. (2008) 'Remote sensing of the distribution and abundance of host species for spruce budworm in Northern Minnesota and Ontario', *Remote Sensing of Environment,* 112, (10), pp. 3971-3982.

Woodhouse, I. H. (2006) *Introduction to Microwave Remote Sensing.* Boca Raton, USA: Taylor and Francis.

Wu, J., Xiong, H. and Chen, J. (2009) 'Towards understanding hierarchical clustering: A data distribution perspective', *Neurocomputing,* 72, (10-12), pp. 2319-2330.

Xie, M., Esaki, T., Zhou, G. and Mitani, Y. (2003) 'Geographic information systems-based three-dimensional critical slope stability analysis and landslide hazard assessment', *ASCE Journal of Geotechnical and Geoenvironmental Engineers.,* 129, (12), pp. 1109-1118.

Yesilnacar, E. and Suzen, M. L. (2006) 'A land-cover classification for landslide susceptibility mapping by using feature components', *International Journal of Remote Sensing,* 27, (2), pp. 253-275.

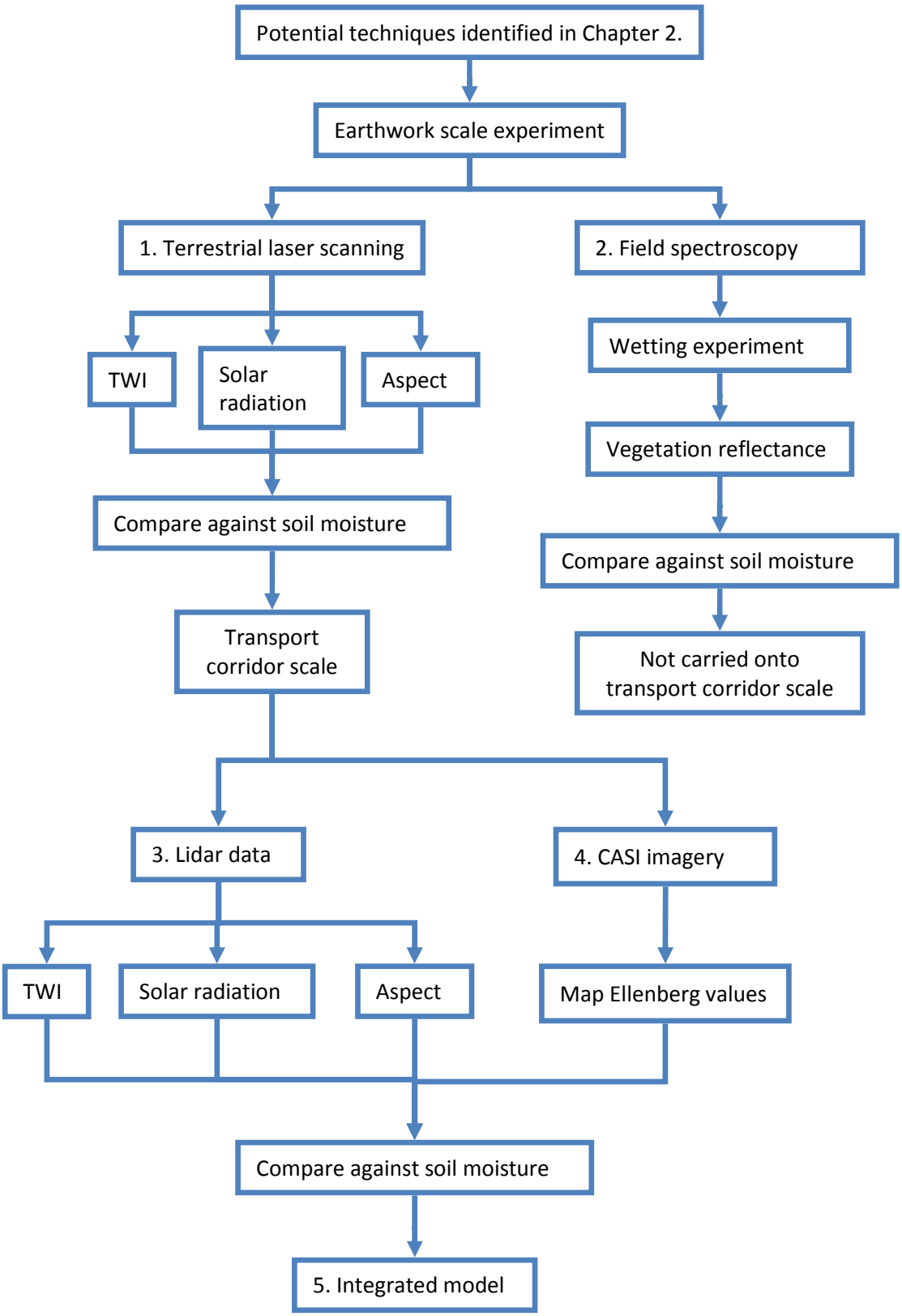
Yue, T.-X., Du, Z.-P., Song, D.-J. and Gong, Y. (2007) 'A new method of surface modeling and its application to DEM construction', *Geomorphology,* 91, (1-2), pp. 161-172.

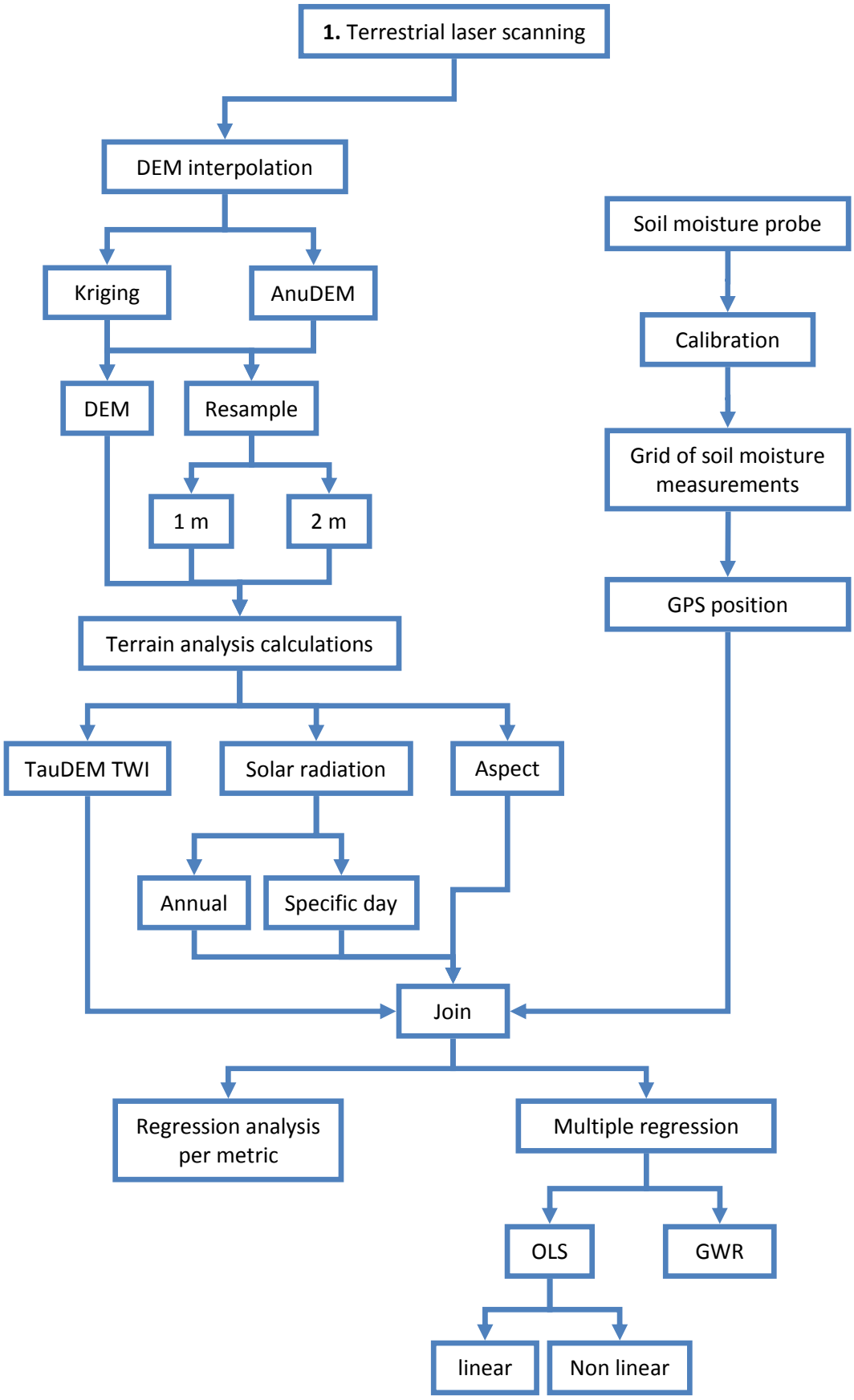
Zarco-Tejada, P. J., Pushnik, J. C., Dobrowski, S. and Ustin, S. L. (2003) 'Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects', *Remote Sensing of Environment*, 84, pp. 283-294.

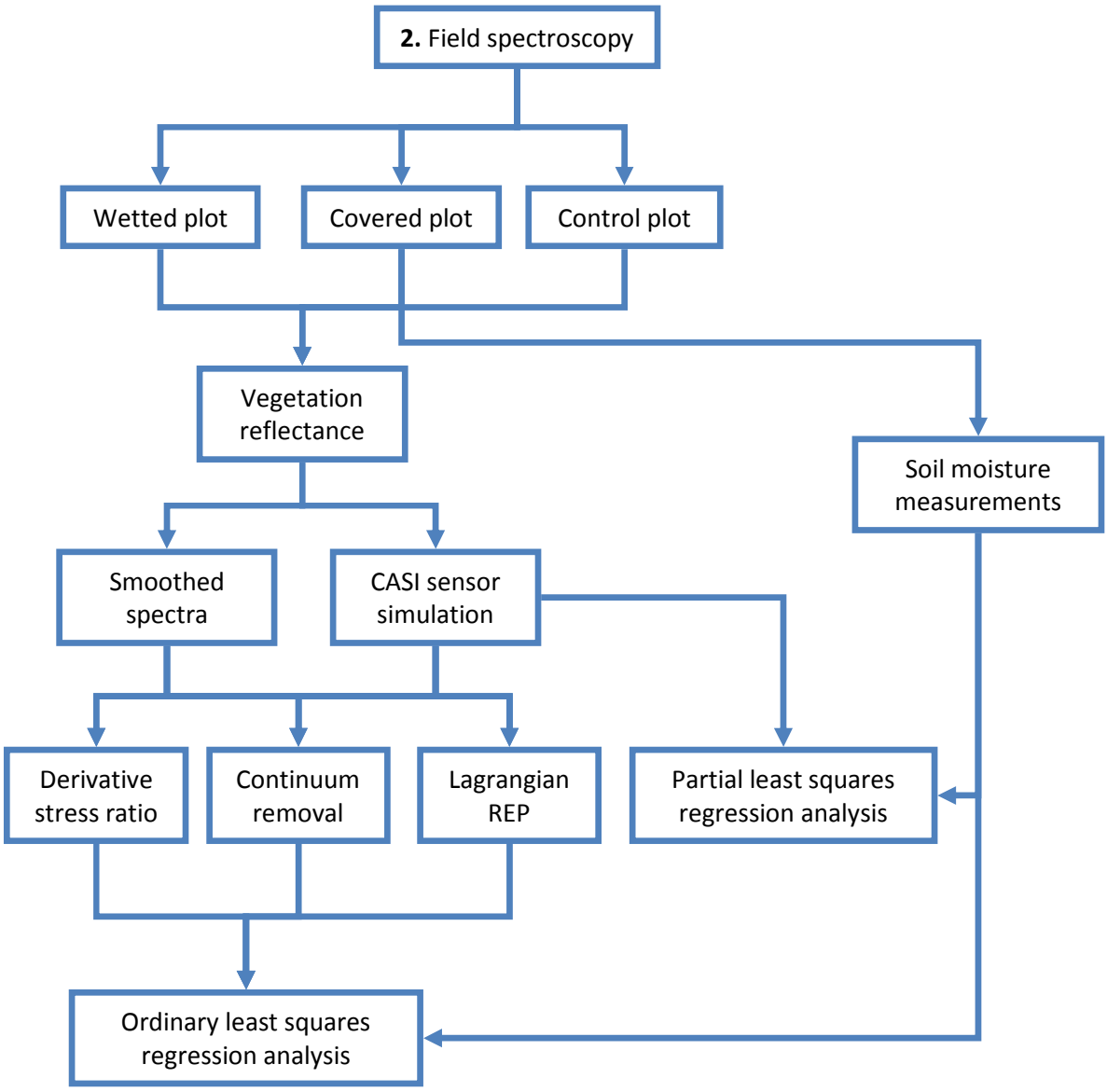
Zhang, W., Wu, L., Zhou, J., Qin, Z. and Zhang, D. (2005) 'Quantitative hill-slope stability assessment with a remote sensing and GIS based distribution modelling scheme', *IEEE International Geoscience and Remote Sensing Symposium.* pp. 2891-2895.

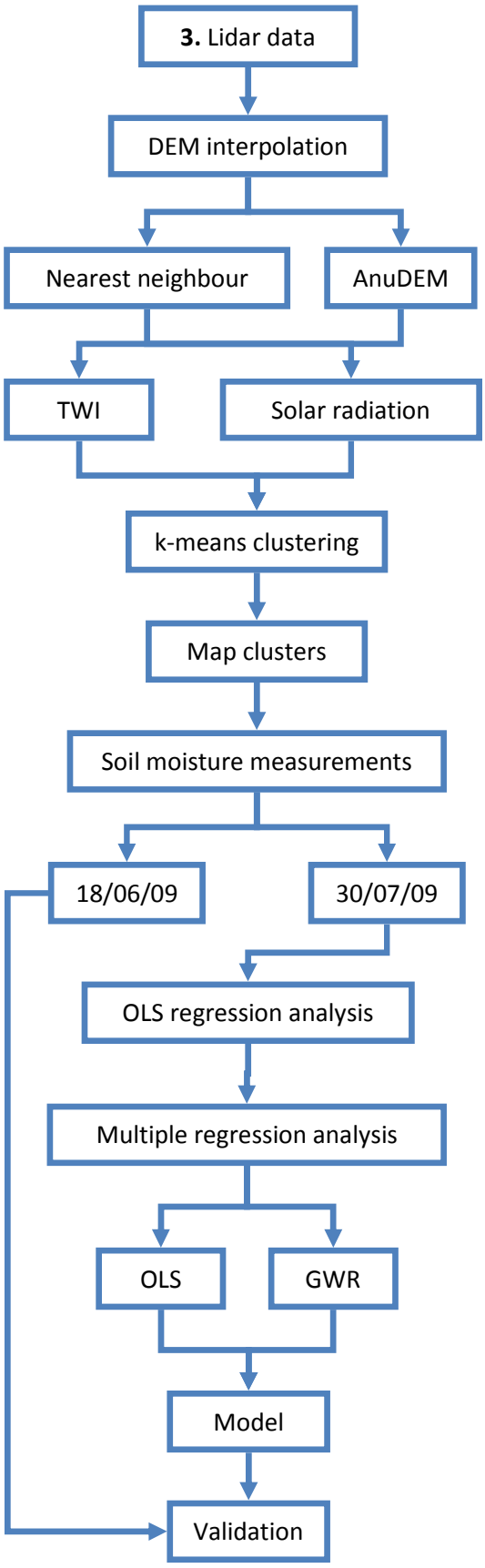
Appendix 1: Flow diagrams of methods

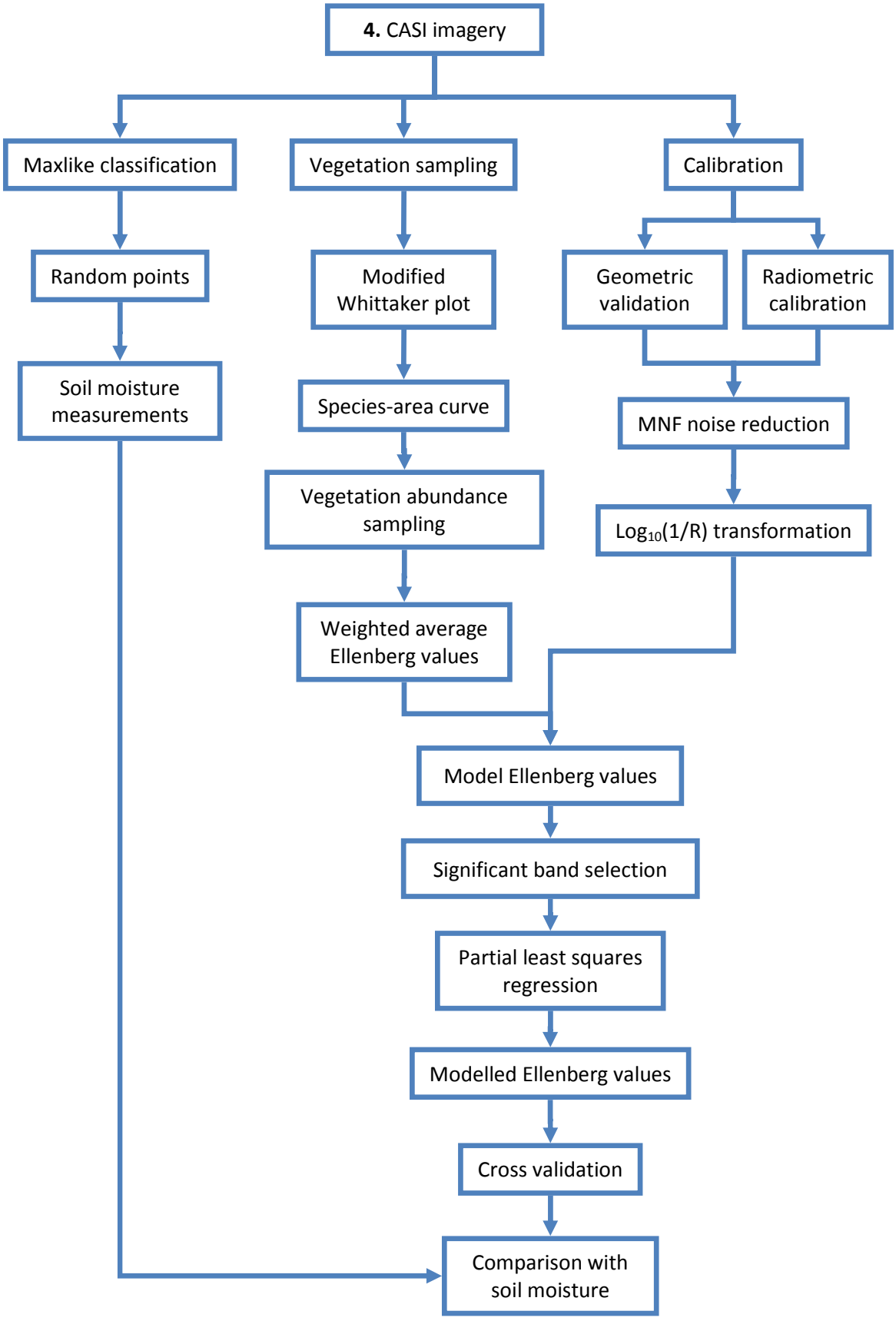
This appendix provides an extension to the flow diagram of methods shown at the start of Chapter 3 (Figure 3.1). The diagram given in Figure 3.1 (and repeated below) is designed to give the reader an overview of the methods used in this study. This includes five distinct sections, or work packages, which are labelled: (1) Terrestrial laser scanning, (2) field spectroscopy, (3) lidar data, (4) CASI data, and (5) integrated model. Following this initial overview, five more flow diagrams are given which provide a breakdown of the methods used in each of these distinct sections. All these methods are described in Chapter 3 but these diagrams will hopefully provide the reader with a clear structure of the methods used in this study to aid repeatability and improve clarity.

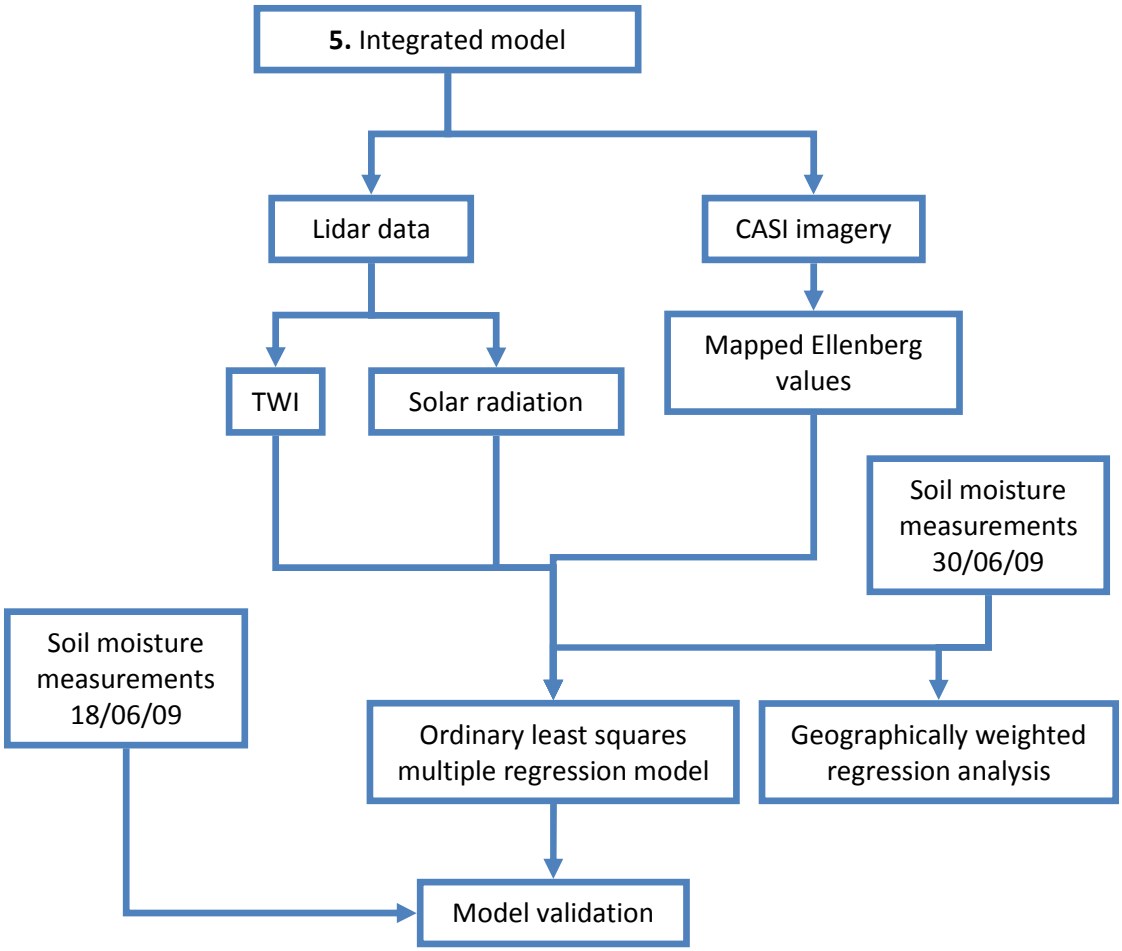

Potential techniques identified in Chapter 2.
Earthwork scale experiment
1. Terrestrial laser scanning
2. Field spectroscopy
TWI
Solar radiation
Aspect
Wetting experiment
Vegetation reflectance
Compare against soil moisture
Compare against soil moisture
Transport corridor scale
Not carried onto transport corridor scale
3. Lidar data
4. CASI imagery
TWI
Solar radiation
Aspect
Map Ellenberg values
Compare against soil moisture
5. Integrated model


1. Terrestrial laser scanning
DEM interpolation
Kriging
AnuDEM
DEM
Resample
1 m
2 m
Terrain analysis calculations
TauDEM TWI
Solar radiation
Aspect
Annual
Specific day
Soil moisture probe
Calibration
Grid of soil moisture measurements
GPS position
Join
Regression analysis per metric
Multiple regression
OLS
GWR
linear
Non linear


2. Field spectroscopy
Wetted plot
Covered plot
Control plot
Vegetation reflectance
Soil moisture measurements
Smoothed spectra
CASI sensor simulation
Derivative stress ratio
Continuum removal
Lagrangian REP
Partial least squares regression analysis
Ordinary least squares regression analysis


3. Lidar data
DEM interpolation
Nearest neighbour
AnuDEM
TWI
Solar radiation
k-means clustering
Map clusters
Soil moisture measurements
18/06/09
30/07/09
OLS regression analysis
Multiple regression analysis
OLS
GWR
Model
Validation


4. CASI imagery
Maxlike classification
Vegetation sampling
Calibration
Random points
Modified Whittaker plot
Geometric validation
Radiometric calibration
Soil moisture measurements
Species-area curve
MNF noise reduction
Vegetation abundance sampling
Log10(1/R) transformation
Weighted average Ellenberg values
Model Ellenberg values
Significant band selection
Partial least squares regression
Modelled Ellenberg values
Cross validation
Comparison with soil moisture


5. Integrated model
Lidar data
CASI imagery
TWI
Solar radiation
Mapped Ellenberg values
Soil moisture measurements 30/06/09
Soil moisture measurements 18/06/09
Ordinary least squares multiple regression model
Geographically weighted regression analysis
Model validation

Appendix 2: Geographically Weighted Regression

Section 3.2.1.4 refers to a type of locally variable regression known as Geographically Weighted Regression (GWR) which may not be familiar to all readers. The following provides a description of GWR.

The theoretical background to GWR can be explained in terms of the ordinary least squares, or global, regression model. This can be defined as:

$$y = \beta_1 x_1 + \ldots + \beta_p x_p + \varepsilon$$

where y is the dependent variable, x_1 to x_p are independent variables, β_1 to β_p are estimated coefficients, and ε is the error term (Wang et al., 2005). GWR allows for the estimate of local parameters rather than just the global parameters and can therefore be expressed as:

$$y = \beta_0(\mu, v) x_1 + \ldots + \beta_n(\mu, v) x_n + \varepsilon$$

where β_0 is the intercept and (μ, v) represents the coordinates of each observation.

In GWR, the parameters are estimated by weighting a contributing point based on its proximity to the point of interest. In this way, the weighting of an observation is no longer constant, but varies over space, with those observations closer to the point of interest having a higher weighting. The parameter estimation can be expressed as:

$$\beta^*(\mu, v) = (X^T W(\mu, v) X)^{-1} X^T(\mu, v) y$$

where, $\beta^*(\mu,v)$ is an estimate of β, X is the matrix of independent variables, and $W(\mu,v)$ is the weigh matrix which ensures that observations closer to the point of interest have more influence than those further away (Wang et al., 2005).

The model weighting can take the form of a number of different models, depending on the type of variable being analysed (Fotheringham et al., 2002). Often a Gaussian function with a fixed kernel bandwidth is used (Figure 1). However, this approach can be problematic as, by assuming a constant bandwidth across a study area, the degree of spatial variation can be exaggerated (Foody, 2003). To overcome this, a kernel with an adaptive bandwidth can be employed (Foody, 2003; Fotheringham et al.,

2002). Here, a bi-square function is used to specify the weight of the *j*th observation at a specific location of interest, regression point *i* is,

$$w_{ij} = \left[1 - \left(d_{ij}/b\right)^2\right]^2 \text{ when } d_{ij} \leq b$$

$$w_{ij} = 0 \text{ when } d_{ij} > b$$

where *dij* is the Euclidian distance between the locations of the sites and *b* is the kernel bandwidth.

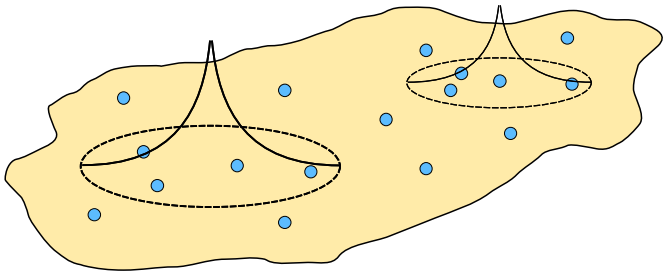


Figure 1. Graphical representation of the local neighbourhood, defined by kernel bandwidth, over which regression analysis is perform. In this example, the two neighbourhoods are of different sizes, meaning the calculation has a kernel with an adaptive bandwidth (note that the large circle is located over less dense points compared to the smaller circle). The curved lines running from the centre of the two neighbourhoods represent the relative weighting used in the regression analysis. In this case the weighting is defined by a Gaussian function.

The kernel bandwidth size can have an important influence on the model. For example, where kernel sizes are too big the GWR model can be no different from its global equivalent, and where kernels are too small, they may not estimate the parameters sufficiently where data is scarce (Wang et al., 2005). To account for this, a measure of how well the model fits the data can be made, such as the Akaike Information Criterion (AIC_c), which is used to compare the fit of the GWR model against the measure of fit for the global regression model (Foody, 2003). Fotheringham et al. (2002) define the AIC_c as:

$$AIC_c = 2n\log_e(\sigma^*) + n\log_e(2\pi) + n\left\{\frac{n + tr(S)}{n - 2 - tr(S)}\right\}$$

Where, n is the sample size, σ^* is the estimated standard deviation of the error term (ε), and $tr(S)$ is the trace of the hat matrix S (Wang et al., 2005). The $_c$ part of *AIC* refers to the fact that the criterion is corrected (Fotheringham et al., 2002). Wang et al. (2005) explains as a general rule that the lower the AIC_c number, the more the estimate represents reality. By using this rule, one would expect the AIC_c value for the GWR model to be lower than the value for the global regression model. As part of the GWR validation, an F test is performed using the results of the global regression, with the null hypothesis being that GWR makes no improvement on the estimation of the independent variable (Wang et al., 2005).

Appendix 3: Simulation of the CASI sensor

Spectral reflectance data collected using the ASD field spectroradiometer was used to simulate the response from the CASI sensor. This was done using the equations provided in Section 3.2.2.4. The following diagram shows how these calculations work for hypothetical data. The data used corresponds to CASI band 1. This is centred over the wavelength 398 nm and has a spectral range of +/- 4 nm.

1. Calculate the standard deviation (σ) based on the FWHM (Field Width Half Maximum) using:

$$\sigma = \frac{FWHM}{2\sqrt{2\ln 2}}$$

Therefore,

$$\sigma = \frac{4}{2\sqrt{2\ln 2}} = 1.7$$

2. Calculate the Gaussian point spread function for each wavelength covering the spectral range (FWHM) around the band centre

$$\text{Gaussian PSF} = EXP^{-\frac{(x_i - x_o)^2}{2\sigma^2}}$$

Where, x_i is the wavelength of interest and x_0 is the position of the central wavelength. For CASI band 1, with a wavelength of 398 nm this corresponds to the following values:

x_i (nm)	x_o (nm)	Gaussian PSF
394	398	0.06
395	398	0.21
396	398	0.50
397	398	0.84
398	398	1.00
399	398	0.84
400	398	0.50
401	398	0.21
402	398	0.06

3. The Gaussian point spread function is normalised by dividing by the total of all the Gaussian point spread values. The sum of the values in the above table equates to 4.23. Therefore the normalised point spread values are:

x_i (nm)	Normalised PSF
394	0.01
395	0.05
396	0.12
397	0.20
398	0.24
399	0.20
400	0.12
401	0.05
402	0.01

4. The normalised point spread function values are then used as a multiplier to simulate the response of the CASI sensor from field spectroscopy measurements. An example of this is given below using hypothetical field spectroscopy data:

x_i (nm)	Normalised PSF	Field spectroscopy (% reflectance)	Simulated CASI response
394	0.01	3.70	0.05
395	0.05	3.76	0.19
396	0.12	3.73	0.44
397	0.20	3.85	0.77
398	0.24	3.86	0.91
399	0.20	3.67	0.73
400	0.12	3.65	0.43
401	0.05	3.73	0.19
402	0.01	3.87	0.06

5. Finally, CASI simulated spectra for the central wavelength 398 nm can be calculated by summing the simulated CASI response. For the example above, this equates to 3.77%.

This process is repeated for each CASI band to simulate the entire response from the CASI sensor.

Appendix 4: CASI radiometric correction

Radiometric calibration of the CASI imagery was carried out by applying the empirical line method (Smith and Milton, 1999) to coincidental field spectroscopy measurements taken over three 6 x 6 m tarpaulin targets (black, grey and white) (Section 3.3.2). The following table shows the regression equations for each band with relative R^2 correlation coefficient:

CASI Band	Central wavelength (nm)	Regression equation	R^2 correlation coefficient
1	398	y = 0.0072x - 5.7759	1
2	405	y = 0.0075x - 2.7677	0.9994
3	413	y = 0.0094x - 4.0857	0.9993
4	420	y = 0.0109x - 6.2091	0.9999
5	428	y = 0.0126x - 5.5489	0.9993
6	435	y = 0.0124x - 5.0131	1
7	444	y = 0.0113x - 4.6365	0.9994
8	452	y = 0.01x - 3.0913	0.9999
9	462	y = 0.0103x - 3.8989	0.9999
10	471	y = 0.0105x - 3.0749	1
11	482	y = 0.0105x - 3.0558	0.9998
12	493	y = 0.0108x - 3.128	0.9999
13	504	y = 0.011x - 2.7955	0.9997
14	517	y = 0.0112x - 2.4983	0.9997
15	530	y = 0.0109x - 2.4335	0.9996
16	544	y = 0.011x - 2.6015	0.9997
17	559	y = 0.0112x - 2.326	0.9999
18	575	y = 0.0118x - 2.2385	0.9998
19	592	y = 0.0122x - 2.1158	0.9999
20	610	y = 0.0122x - 2.0664	0.9998
21	630	y = 0.0126x - 1.6787	0.9999
22	652	y = 0.0132x - 1.6441	0.9996
23	675	y = 0.0131x - 1.4088	0.9997
24	700	y = 0.0149x - 1.6909	0.9998
25	728	y = 0.0163x - 2.7653	0.9997
26	758	y = 0.0157x - 3.6425	0.9998
27	790	y = 0.0143x - 3.7689	0.9997
28	824	y = 0.018x - 4.1063	0.9997
29	862	y = 0.0164x - 4.0208	0.9998
30	901	y = 0.0224x - 3.9166	0.9997
31	943	y = 0.051x - 4.7153	0.9996
32	988	y = 0.0254x - 4.5207	0.9997

Appendix 5: Modified-Whittaker species sampling

Species type was recorded over a number of sub plots in order to calculate a species-abundance curve. The names and occurrence of each species are detailed in Table 1. The location and size of the sub plots are identified in Figure 1, which is also referred to in Section 3.3.2.1.1.

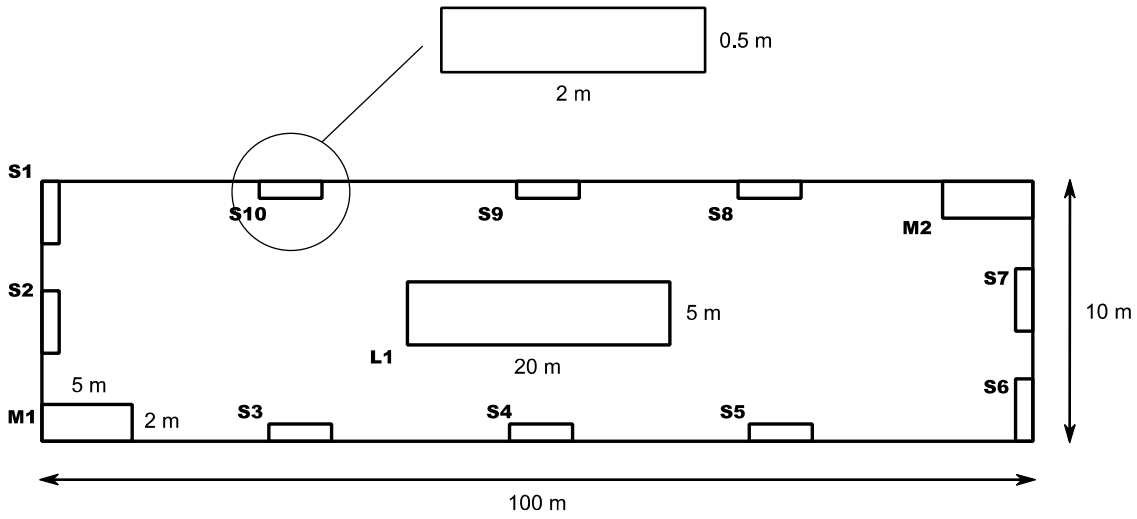


Figure 1. Diagram of the modified-Whittaker plot used to sample vegetation species type for construction of a species-area curve. Plot labels refer to the list of species name and occurrence reported in Table 1.

English name	*Latin name*	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	M1	M2	L1	Total area
Stinging Nettle	*Urtica dioica*	1	1	1					1		1	1	1		1
Creeping Thistle	*Cirsium arvense*	1	1			1				1	1	1		1	1
Cleavers	*Galium aparine*	1							1	1	1	1		1	1
Bramble	*Rubus fruticosus agg.*	1	1	1								1		1	1
Heath Bedstraw	*Galium sterneri*		1	1		1	1							1	1
Red Campion	*Silene dioica*		1	1							1	1			1
Common Dog Violet	*Viola riviniana*					1		1							1
Ribwort Plantain	*Plantago lanceolata*					1	1	1				1		1	1
Birdsfoot Trefoil	*Lotus corniculatus*					1									1
Tufted Vetch	*Vicia cracca*						1		1		1			1	1
Common Knapweed	*Centaurea nigra agg.*						1	1						1	1
Common Sorrel	*Rumex acetosa*						1		1					1	1
Common Ragwort	*Senecio jacobaea*								1			1		1	1
Rosebay Willowherb	*Chamerion angustifolium*											1	1	1	1
Germander Speedwell	*Veronica chamaedrys*												1	1	1
Elder	*Sambucus nigra*													1	1
Bracken	*Pteridium aquilinum*													1	1
Dog Rose	*Rosa canina*													1	1
Spear Thistle	*Cirsium vulgare*													1	1
Hawthorn	*Crataegus monogyna*													1	1
Dandelion	*Taraxacum officinale*													1	1
Pale Lady's Mantle	*Alchemilla xanthochlora*													1	1
Meadow Cranesbill	*Geranium pratense*														1
Ox-eye Daisy	*Leucanthemum vulgare*														1
Smooth Hawksbeard	*Crepis capillaris*														1
Field Forgetmenot	*Myosotis arvensis*														1
Hogweed	*Heracleum sphondylium*														1
Field Scabious	*Knautia arvensis*														1
Himalayan Balsam	*Impatiens glandulifera*														1

Continued overleaf…

Great Burnet	*Sanguisorba officinalis*															1
American Willowherb	*Epilobium ciliatum*															1
Meadow Foxtail	*Alopecurus pratensis*										1					1
Common Bent-Grass	*Agrostis capillaris*				1	1	1	1	1	1			1	1		1
Yorkshire Fog	*Holcus lanatus*	1	1	1	1	1	1	1	1	1	1	1	1	1		1
Cocksfoot	*Dactylis glomerata*						1	1	1	1			1	1		1
	Total	5	6	5	2	7	8	6	8	5	7	9	6	21		35

Table 1. List of plant species found in the modified-Whittaker plot. Location and size of the plots are reported in Figure 1.

Appendix 6: Species abundance sampling for Ellenberg values

Species abundance was sampled in 96 quadrats measuring 1 x 1 m. This data was used to calculate weighted average Ellenberg values for each site. Table 1 shows a list of all the plant species recorded and their corresponding Ellenberg value for the indication of soil moisture. Higher values corresponding to higher wetness.

English name	*Latin name*	Ellenberg value
American Willowherb	Epilobium ciliatum	6
Betony	Stachys arvensis	5
Birdsfoot Trefoil	Lotus corniculatus	4
Bracken	Pteridium aquilinum	5
Bramble	Rubus fruticosus agg.	6
Brown Bent-Grass	Agrostis vinealis	6
Cleavers	Galium aparine	6
Cocksfoot	Dactylis glomerata	5
Common Bent-Grass	Agrostis capillaris	4
Common Knapweed	Centaurea nigra agg.	5
Common Ragwort	Senecio jacobaea	4
Common Sorrel	Rumex acetosa	5
Common Toadflax	Linaria vulgaris	4
Creeping Buttercup	Ranunculus repens	7
Creeping Thistle	Cirsium arvense	6
Crested Dog's-Tail	Cynosurus cristatus	5
Curled Dock	Rumex Crispus	6
Daisy	Bellis perennis	5
Dog Rose	Rosa canina	5
Elder	Sambucus nigra	5
False Oat-Grass	Arrhenatherum elatius	7
Field Horsetail	Equisetum arvense	6
Germander Speedwell	Veronica chamaedrys	5
Greater Plantain	Plantago major	5
Hawthorn	Crataegus monogyna	5
Heath Bedstraw	Galium sterneri	4
Himalayan Balsam	Impatiens glandulifera	8
Hogweed	Heracleum sphondylium	5
Lesser Clubmoss	Selaginella selaginoides	7
Lilac	Syringa vulgaris	5
Ox-eye Daisy	Leucanthemum vulgare	4
Red Campion	Silene dioica	6
Red Dead-Nettle	Lamimum purpureum	5

Continued overleaf…

Ribwort Plantain	Plantago lanceolata	5
Rosebay Willowherb	Chamerion angustifolium	5
Rough Chervil	Chaerophyllum temulum	5
Spear Thistle	Cirsium vulgare	5
Square-stalked St. John's Wort	Hypericum tetrapterum	8
Stinging Nettle	Urtica dioica	6
Tall Fescue	Festuca arundinacea	6
Tufted Vetch	Vicia cracca	6
White Campion	Silene latifolia	4
White Clover	Trifolium repens	5
White Dead-Thistle	Lamimum album	5
Yorkshire Fog	Holcus lanatus	6

Table 1. List of plant species identified in the quadrats used for determining species abundance to calculate weighted average Ellenberg value.