Hyperspectral remote sensing for early detection of wild carrot in Carrot (Daucus carota) seed production – A feasibility study

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

in

Horticultural Science

at Massey University, Manawatū, New Zealand.

Sunmeet Singh Bhatia

2023

Abstract

Hyperspectral remote sensing for early detection of wild carrot in Carrot (Daucus carota) seed production – A feasibility study

Carrot (*Daucus carota*) seed production is an important part of the NZ vegetable seed industry with exports of $33.4 million NZD in 2020. Most carrot seed production is based in the Canterbury region, but there is a desire by key stakeholders to expand carrot seed production in the Hawke's Bay region of NZ. However, presence of weed wild carrot (*Daucus carota* subsp. *carota*) in the region acts as a significant constraint to carrot seed production. Wild carrot plants can crosspollinate with carrot crop plants, causing genetic contamination in a crop where genetic purity is of critical importance. The current weed management strategy of manual scouting and rouging is resource intensive and ineffective in achieving appropriate control of wild carrot in the region.

Airborne hyperspectral remote sensing is a technology that has proven its ability in plant species identification and can do so at a high spatial scale in a short period of time. This makes the technology a promising candidate for a superior alternative weed control method. This project aimed to test the feasibility of the technology to identify wild carrot plants in carrot seed crop fields and nearby areas within the crop's isolation distance (2000m). This involved creating spectral libraries of dominant plant species/materials, including wild carrot, in the area of interest. The methodology involved conducting a survey and collecting airborne hyperspectral data. Further, ground-based collection of GNSS enabled accurate GPS locations of wild carrot plants in the survey area, acted as training and validation data for subsequent classification analysis. The ground truth data was also used for a pixel composition analysis – which also helped understand the environmental context of wild carrot plants. The data was analysed in an image processing software (ENVI®, v5.6). The analysis involved two levels of classification algorithms. A first order classification – minimum distance classification (MDC) – was used to classify the data into broad land surface cover types. The classification was successful with an overall accuracy of 96%. The second order classification was a soft classification algorithm which employed spectral unmixing – mixture tuned matched-filtering (MTMF). This method allows sub-pixel classification when the target surface is smaller than pixel size, as in this case. MTMF helped create a model which predicted potential locations of wild carrot plants at a threshold level of 5% of pixel area (surface area - 0.05m2) and a producer's accuracy of 70% (Omission error rate – 30%) for patches above

the threshold surface area. These predicted locations were projected on appropriate RGB base layers to create wild carrot weed maps. The biggest limitation was likely the $1m^2$ spatial resolution of the hyperspectral camera employed in the study, which dictated the 5% pixel threshold level.

These detection threshold and accuracy levels are lower than in other similar studies, however they are likely acceptable in the current context and can help mitigate wild carrot damage in carrot seed production in the Hawke's Bay. The study has helped identify areas of future research to further improve the detection threshold and accuracy levels. These include identifying relationships between environmental context related parameters and wild carrot manifestation, acquiring higher spatial resolution data (lower altitude flights, unmanned aerial vehicle (UAV) mounted cameras, deploying of image fusion techniques using separate high spectral (hyperspectral) and spatial resolution (RGB/multispectral) imagery.

Acknowledgements

I'd like to thank Ivan Lawrie and Phil Rolston from Seed Industry Research Centre for support which allowed this research to happen.

I'd like to thank my main-supervisor and mentor Craig McGill for his guidance, support and his innate desire and ability to find solutions to problems.

Thanks to my other supervisors Gabor Kereszturi for his advice that this was a worthwhile idea to pursue. His initial support allowed me to learn Hyperspectral remote sensing related concepts and skills, which were essential for me to finish the study.

Thanks to Nitin Bhatia, for allowing me to work at Massey Agrifood Digital lab (especially with their good coffee) and for his technical input, especially in the thesis writing process.

Thanks to Stefan Carter and Eduardo Sandoval for help with the data pre-processing and chats over coffee.

I appreciate support of my friends especially Tom, Caroline, Flo, Laise, Sha, and Fena. Thanks, Tom, for his advice that feeling underprepared and out of depth after a technical project meeting is very common in graduate studies. Thanks, Caroline, for the laughs, support and continuing to ask about my thesis progress, keeping me on track.

Thanks to my management at School of Agriculture and Environment – Janet, Paul, Danny and Chris for your support and encouragement. I could not have completed this without your support and encouragement.

Finally, thanks to my family spread across three continents. Thank you for all the hard work you have done so I can have these opportunities. None of this is possible without you.

Contents

List of figures

List of tables

Chapter 1 Introduction

Carrots are an important root crop globally. The carrot storage root is rich in carotenoids, anthocyanins, various vitamins, minerals and dietary fibre, all of which have a positive relationship with human health. Carrot consumption has been linked to a reduction in cardiovascular disease risk, diabetes risk, anticarcinogen and overall immune system benefits (da Silva Dias, 2014; Que et al., 2019). A growing desire for human health optimisation through nutrition in the form of vegetables, with the aim of also securing nutritional security in low and medium income countries (Keatinge et al., 2011), and increasing awareness of nutritional attributes in carrots means cultivated carrots are an important and growing crop worldwide (Que et al., 2019). This increased production will require a consistent supply of high-quality seed to support it. This will drive increased carrot seed production.

Carrot (*Daucus carota* L.) seed production is an important part of the New Zealand vegetable seed industry. In 2020 New Zealand carrot seed exports reached $33.4 million, up from $7.6 million in 2010 (FreshFacts, 2021). New Zealand accounts for around 50% of the carrot seed produced globally (Dewi Preece, 2023). Most of this production is based in the Canterbury region of the South Island (Merfield et al., 2001). The Canterbury Plains have features that make it favourable for seed production (Hampton et al., 2012). This includes fertile alluvial soils, irrigation, and the presence of seed production infrastructure. The climate is favourable for seed production, especially for the crucial seed drying period at the end of seed development, and is characterised by relatively low rainfall (low moisture), warm days and cold nights (Hampton et al., 2012). However, there is pressure on the area available for seed production due to competing land uses such as dairying and urbanisation (Dynes et al., 2010).

There is therefore significant interest in expanding seed production in Hawke's Bay, with carrot seed being a key crop of interest (Phil Rolston, Seed Industry Research Centre, personal communication, 2021). The temperate plains of the east coast of both North and South islands share similar climate and geological features, making Hawke's Bay an area suitable for seed production (Hampton et al., 2012). However, a number of factors can act as constraints to carrot seed production in Hawke's Bay. Seed growers, rural professionals and seed industry bodies in the region have identified one of these factors as the presence of wild carrot (*Daucus carota* subsp. *carota* L.) plants in the carrot seed production area (Hugh Ritchie, Drumpeel Farms, personal communication, 2021).

Wild carrot is native to Central Asia and parts of Europe, specifically areas with a temperate and semi-arid/arid climate (Iorizzo et al., 2013). It has been introduced into various areas across the globe, including New Zealand (Rong et al., 2010). Carrot grown as a root vegetable is also *Daucus carota* but a different subspecies (subsp. *sativus*). The cultivated carrot has been domesticated from the wild carrot (Hauser & Bjørn, 2001). Figure 1-1 shows where the wild carrot is typically found growing along a fence and a road.

Figure 1-1: Wild carrot plants (Source: Robert Southward, 2022; AgResearch, 2022)

Wild carrot has been shown to freely cross-pollinate with commercial carrot varieties (Mandel et al., 2016). This will likely compromise the quality of the seed crop, especially in hybrid seeds where genetic purity is a crucial parameter of seed quality. Another issue is that transfer of pollen from the crop varieties to the wild carrot may allow the weed plants to inherit traits that further contribute to their environmental hardiness that further complicates their control (Kiran & Pandey, 2020).

Currently for seed production, the production and surrounding areas are manually inspected for the presence of wild carrot plants which are then rogued. Removal of the entire plant (roguing)

has been shown to be the only effective long-term control strategies for wild carrot weeds (Rome & Lucero, 2019).

Carrot pollen can potentially travel a significant distance (2000m isolation distance) (AsureQuality, 2021). This makes for a large area that is required to be manually inspected. Manual inspection of a large area for the weed plants is a time-consuming, labour intensive and likely an expensive process. Furthermore, there is a greater likelihood of weed plants being undetected in a manual inspection over a large area. Figure 1-2 illustrates the area within the isolation distance requirements in a hypothetical carrot field.

Figure 1-2: The area within the isolation distance requirements of a carrot seed production field

Cost and time mean that there is interest in and potential for alternative methods for the identification of wild carrot contaminants that are more effective and efficient than the current weed management strategy.

Hyperspectral remote sensing has been widely applied for the detection and mapping of weed plants in an agricultural context (Mundt et al., 2005; Thorp & Tian, 2004). The technology has the potential to be part of an alternative management strategy for wild carrot in Hawke's Bay and other areas.

The purpose of the following literature review is to confirm there is a knowledge gap on alternative methods for control of wild carrot near carrot seed crops and identify a research question relating to wild carrot weeds as a constraint in carrot seed production in Hawke's Bay. The review will first look at the biology of the carrot crop and the wild carrot, the current state of knowledge of wild carrot and its pollination risk and confirm that the presence of wild carrot weeds in Hawke's Bay acts as a constraint in carrot seed production, secondly look at the efficiency and effectiveness of the current weed management methods and finally to explore an alternative method that may be superior and a better fit-for-purpose compared with current weed management methods.

Chapter 2 Literature Review

2.1 Carrot

The carrot (*Daucus carota* subsp. *sativus*) crop plant belongs to the family Apiaceae. The Apiaceae are one of the largest families of seed producing plants with 466 genera and 3820 species (Spooner, 2019). The Apiaceae are considered to be a monophyletic family – but closely related to Araliaceae, Pittosporaceae and Myodocaepaceae. The Apiaceae are part of the order Apiales which contains 5400 species. Species from Apiaceae are found globally but the most diverse distribution is in the temperate regions of the northern hemisphere (Spooner, 2019).

The carrot plant has been selected from its wild counterpart wild carrot (Iorizzo et al., 2013) which is an annual plant, however, the carrot plant in use for food production is a biennial plant (Wohlfeiler et al., 2022). This means it can take up to two years to finish its lifecycle from seed to seed (Simon, 2019), however, in New Zealand the production time can be less (2.1.2 Carrot seed production in New Zealand). The carrot plant has tripinnate, complex, alternative leaves. It is characterised by having a taproot which is the edible part of the plant. The carrot crop is a biennial seed crop, undergoing one year of vegetative growth. This is when the edible taproot development occurs. Flower initiation occurs in the spring of the following year. Cultivars bred for a temperate climate usually require a period of cold exposure (vernalisation) to initiate flowering, once the plant has sufficient leaves (usually 8-12) and the tap root is greater than 4-8 mm (Simon, 2019; Wohlfeiler et al., 2022). This is followed by seed development and harvest in late summer/autumn. The carrot crop plant produces individual flowers, similar to other species in the Apiaceae family, on terminal branches in compound umbels. There is an order of hierarchy of the umbels based on the sequence of flowering. The first umbel to flower is the primary umbel and is usually terminal to the main stalk. The main stalk produces branches to form secondary umbels and the branches from these stalks form tertiary umbels (George, 2009). The flower in carrot crop plants are considered to be generalist and can be pollinated by hundreds of insect species (Broussard et al., 2017). Figure 2-1 shows the lifecycle of the carrot crop plant from seed germination to seed production.

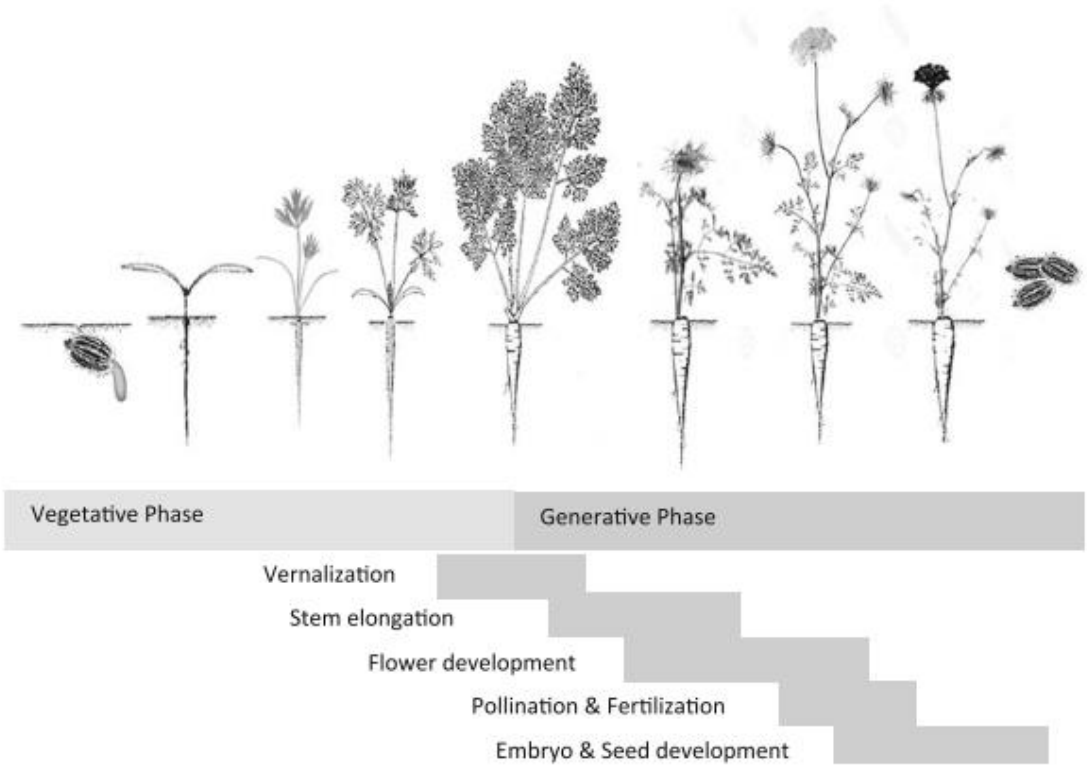

Figure 2-1: Carrot crop plant life cycle (Linke et al., 2019).

2.1.2 Carrot seed production in New Zealand

In New Zealand, carrot seed crop is planted directly from seed in January-February (mid-late summer). This is followed by a period of taproot development and overwintering. Flower initiation happens in November (spring) followed by seed set and seed development into the next year. Finally seed harvest takes place in March for early varieties through to April for late varieties. The production time is around 13-14 months, and the major crop development/management stages occur a few weeks earlier in the South Island relative to the North Island (Richard van Garderen, South Pacific Seeds (NZ) Ltd., personal communication, 2022).

2.2 Wild carrot

Wild carrot is said to be an annual or biennial plant (Wohlfeiler et al., 2022). Wild carrot has been reported to flower without a period of vernalisation (Simon, 2019).

Wild carrot can grow up to 1m tall. It has tripinnate leaves with white flowers clustered in the umbels, with sometimes red/purple flowers in the centre of the umbel. In contrast to cultivated carrots it has a thickened white tap root (AgResearch, 2023) . Wild carrot is found throughout the North Island, and on the West Coast, Southland and Canterbury regions of South Island (AgResearch, 2023). It appears alongside roadsides, other waste areas, and at times is found in

pastures. It is found in abundance on dry, hill slopes where there is a lack of other vegetation (AgResearch, 2023).

2.3 Vegetable seed and the New Zealand vegetable seed industry

Vegetables constitute a crucial part of human dietary requirements and overall sustenance. Most vegetable production systems depend on high quality vegetable seeds to propagate and continue vegetable production at scale. Seed and its production underpin the development of high quality propagation material, including trueness-to-type (genetic or varietal purity). Specific genetic traits, allows individual vegetable plants to exhibit desirable characteristics, including nutritional value and high performance (Pervez et al., 2009; Wright, 1980). The vegetable seed industry is crucial to food and nutritional security for humans globally. The industry is also seen as a means to improve the lives and livelihood of farmers in the developing world – as vegetable seed is generally a higher value crop relative to popular high-volume crops like rice and grain. This can allow farmers a greater return on their land and improve their economic situation (Mengistu & Yamoah, 2010; Vira et al., 2020).

The vegetable seed industry is economically important in New Zealand (Preece, 2023). Carrot, radish (*Raphanus sativus* L.), beetroot (*Beta vulgaris* L.) and other brassicas are the highest dollar value by export vegetable seed crops for the industry and contributed ~$115 million of export revenue in 2020 (Scott, 2021). The major export markets for the New Zealand seed industry include continental Europe, Australia, China, Japan, and the US (Scott, 2021).

2.4 Hybrid vegetable seed production

Hybrid seed production is where the genetics of the seed produced are strictly controlled by facilitating crosspollination between known male and female lines that have been selected to give the offspring specific desirable traits (Wright, 1980). Thus ensuring the genetic purity of the seed produced, which is a crucial quality parameter in hybrid seed systems.

Hybrid seed systems are essential in global vegetable production because they provide for traits like high seed vigour, resistance to pests/diseases, tolerance of abiotic stresses like salinity and drought, and high yield of high quality end products (Colombo & Galmarini, 2017). Another reason behind the success of hybrid seeds is from a commercial perspective. The stable/known genetics of a hybrid seed crop allows for the protection of biological intellectual property (Colombo &

Galmarini, 2017). This allows seed companies to control the sale and distribution of seed lines, making them a commercially viable product.

Hybrid seed systems tend to be resource intensive and require high inputs. Initial purchasing costs in hybrid seed is higher than in conventional seed. Farmers are also unable to save progeny due to intellectual property regulations (Colombo & Galmarini, 2017). These costs associated with hybrid seed systems act as constraints for farmers, especially for smallholder farmers in developing countries (Li et al., 2012)

2.5 New Zealand carrot seed sector

Hybrid carrot seed production is an important component of the New Zealand vegetable seed industry. New Zealand is the biggest producer of hybrid carrot seed globally and carrot seed is the largest hybrid seed industry in New Zealand (Midlands Seed, 2023). The industry had a value of over $33 million in 2020 (FreshFacts, 2022). There is a growing demand for carrot crop due to increasing awareness of its nutritional content globally (Que et al., 2019; Takagi et al., 2017). High quality hybrid seed will underpin this growing demand for carrots. The global carrot seed market is forecasted to have a compound annual growth rate of around 7% for the period of 2022-2030 (DataIntelo, 2022). This means that there is a potential opportunity for the New Zealand seed sector to expand its carrot production further and fill the predicted increasing demand for hybrid carrot seed. Most of the current hybrid carrot seed production is based in the Canterbury region of the South Island of New Zealand due to favourable soils and weather conditions. However, a changing climate has meant that in recent times the region has experienced variable weather patterns which is not ideal for carrot seed production (Richard van Garderen, South Pacific Seeds (NZ) Ltd, personal communication, 2023). This includes lower heat units during summer, which has resulted in inadequate pollination which in turn has resulted in a failure of the carrot seed crop to reach its full yield potential (Richard van Garderen, South Pacific Seeds (NZ) Ltd, personal communication, 2023).

This and other competing land use demands have driven the industry to investigate other regions in New Zealand into which carrot seed production can be expanded. One of the regions that fulfil the criteria needed for carrot seed production is Hawke's Bay, specifically, similar ideal climatic and soil conditions for seed production combined with a higher median annual average

temperature (NIWA, 2021). The industry has found this region will have adequate heat units which will maximise pollination and thus allow the crop to reach its yield potential (Richard van Garderen, South Pacific Seeds (NZ) Ltd, personal communication, 2023).

Currently around 100ha of hybrid carrot seed is planted in Hawke's Bay. This is worth around $1.8 million NZD to the industry in the region. The industry hopes to add a further 50 to 100ha to the area already under production (Note - this expansion is likely to be slowed down due to the effects of Cyclone Gabrielle in February 2023) (Richard van Garderen, South Pacific Seeds, 2023, Personal communication).

However, the presence of wild carrot in parts of Hawke's Bay is a constraint to this expansion. An outcrossing % of as low as 0.02% can cause the crop to be rejected resulting in significant economic losses (Richard van Garderen, South Pacific Seeds (NZ) Ltd, personal communication, 2023).

2.6 Wild carrot – a constraint to carrot seed production

The presence of wild carrot can be a constraint to carrot seed production in Hawke's Bay as it can cause genetic contamination within the hybrid carrot seed crop. For carrot crops grown in an area with wild carrot populations, there are reports of individuals in the carrot crop showing early flowering and intermediate root morphology (between wild carrot and carrot crop) (Mandel et al., 2016; Rong et al., 2010). Carrot and wild carrot individuals with intermediate morphological traits like leaf shape and root colour were found in carrot crop and carrot seed producing areas in northern Europe in the Netherlands (Wijnheijmer et al., 1989) and Denmark (Hauser et al., 2004). This is evidence of pollination contamination within the carrot crop from wild carrot resulting in crop-weed hybridisation. Such weed-crop hybridisation has been seen in other crop species. Crop-weed hybridisation has been reported in garden radish (*Raphanus sativus*) and its weedy relative wild mustard (*Sinapis arvensis* L.) (Eber et al., 1998), with cultivated lettuce (*Lactuca sativa* L.) and its wild relative (*Lactuca serriola* L.) (Uwimana et al., 2012), between cultivated rice (*Oryza sativa* L.) and its weedy/wild relatives weedy rice (*Orzya sativa* f. *spontanea* Rosh.) and common wild rice (*Orzya rufipogon* Griff.), and between sugar beet (*Beta vulgaris* subsp. *vulgaris* L.) and weedy forms of wild sea beet (*Beta vulgaris* subsp. *maritima* (L.) Arcangeli.) (Groot et al., 2003)

The likelihood of crop-weed hybridisation is dependent on two factors. The first factor is the degree of cross fertility between the crop/weed variety. The second factor is the overlap of flowering periods of the crop and the weed (Stewart, 2002). In the Hawke's Bay region, flowering periods for most carrot seed crops occur in late spring/early summer. Wild carrot plants in Hawke's Bay have also been observed flowering during this time (Mitchell Wolting South Pacific Seeds (NZ) Ltd, personal communication, December 2021). This indicates that there is an overlap in the flowering of the two subspecies. As suggested earlier, there is evidence of crop-weed hybridisation in wild carrot/carrot meaning that there is a considerable degree of cross fertility between the two subspecies (Mandel et al., 2016). These two factors suggest that there is a high likelihood of crop-weed hybridisation and subsequently of genetic contamination of the seed crops by wild carrot plants, considerably reducing the value of the carrot seed crop for producing a subsequent crop with the desired traits of the variety and hence economic value. The loss of desired traits may also render the carrots produced less valuable for human consumption.

2.6.1 Risk management - Isolation distance

Wild carrots will be an issue if the weed contaminants (sources of pollen) are growing within the isolation distance, which is usually defined in certification schemes for the maintenance of varietal purity such as the OECD seeds schemes (OECD, 2022). Isolation distance is the minimum separation distance between a seed crop and another crop of the same species that will deter crosspollination and reduce genetic contamination in the seed crop (George, 2009). The rationale for isolation distance is that cross pollination risk reduces with distance at close to an exponential rate (Stewart, 2002). The current isolation distance requirement for a carrot seed crop is 2000m. Figure 2-2 below shows the rate of decrease in pollination risk by distance in a radish crop.

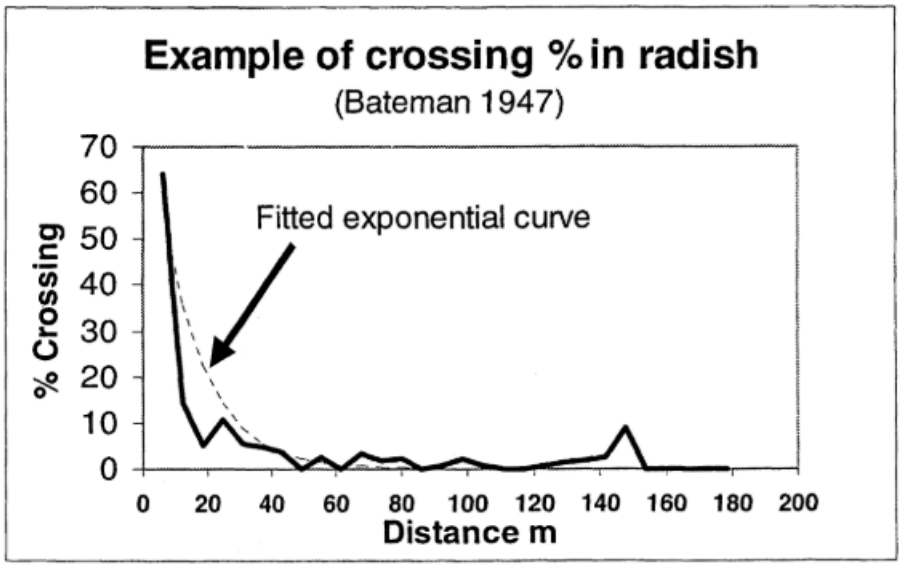

Figure 2-2: Cross pollination between radish crops at different isolation distances. Pollination risk declines exponentially with distance (Stewart, 2002)

2.7 Environmental impacts – more persistent wild carrot populations

Another carrot crop-wild carrot issue comes from a conservation/land management perspective. There is evidence suggesting there has been gene introgression (permanent incorporation of genes) into wild carrot populations from carrot crops (Mandel et al., 2016; Rong et al., 2010), indicating gene flow from cultivated carrot seed crops to wild carrot plants (Hauser et al., 2004; Wijnheijmer et al., 1989). Wild carrot plants growing close to the carrot crop fields have been found to possess traits intermediary between carrot crop plants and wild carrot (Hauser et al., 2004; Rong et al., 2010) with some more similar to the cultivated carrots than the original wild carrot populations (Mandel et al., 2016), suggesting gene introgression from cultivated carrot to the wild populations.

This gene flow may cause transfer of specific genes which enable the weed plants to inherit traits which promote their fitness and enhances their environmental hardiness. This may lead to evolution of aggressive weeds which will further complicate their control (Uwimana et al., 2012) .

Further, this gene transfer from crop plants to wild counterparts is also an issue from a breeding/production standpoint. Wild carrot is considered as an important gene pool for carrot breeding (Rong et al., 2010). Controlled introduction of genetic diversity from wild carrot is said to enhance genetic variation in cultivated carrot, thereby facilitating improvement of cultivar traits (Rong et al., 2010). A hypothesis explaining the development of orange carrots (tap root) is the

spontaneous hybridisation of the earlier yellow and white-root cultivars with a wild carrot subspecies (Hauser et al., 2004)

2.8 Genetic contamination mitigation methods

There are a range of mitigation methods that can be used to ensure varietal purity in hybrid seed lines. These have different effectiveness and given the stringent genetic purity requirements of 0.02% a number may not be feasible for use in hybrid carrot seed production.

2.8.1 Isolation distance

As mentioned earlier, use of an appropriate isolation distance between 2 crop fields is the primary method to mitigate genetic contamination through cross-pollination (Thomison & Geyer, 2016; Warner & Lewis, 2019). The risk of cross pollination reduces at an exponential rate with distance between crop and pollen contaminant source (Stewart, 2002).

2.8.2 Border zones around crops

It has been reported that the likelihood pollen contamination from outside a crop decreases exponentially from the edges to the centre of the crop field (Staniland et al., 2000). Thus a practice to mitigate pollen contamination is to discard seed from the edges of the field as majority of the pollen contaminants is likely to have pollinated the crops on the edges (Stewart, 2002). However, given the value of the hybrid carrot seed, the low levels of contamination permitted and that the crop is insect rather than wind pollinated the feasibility of this approach in carrots is likely to be limited.

2.8.3 Physical barriers and hedges

A cultural method of mitigation of pollination contamination is use of physical barriers like hedges. It is believed to mitigate crop contamination by limiting the movement of wind and insect borne pollen (Stewart, 2002), but may not achieve the high purity requirements for hybrid carrot seed production.

2.8.4 Planting dates

Manipulating the flowering time of the crops by adjusting planting dates is another method that can be used to reduce genetic contamination. This usually involves planting the crop so the flowering period does not occur during peak wild carrot flowering of contaminant pollen source like other varieties of the crop or compatible weed species (Thomison & Geyer, 2016). However,

this can prove to be difficult in a carrot seed production context where the wild carrot populations flower through large time periods of the year.

2.8.5 Good housekeeping/hygiene

An effective housekeeping and hygiene system in place is crucial to minimise genetic contamination in seed crops. This includes adequate weed control in the fields, clearance of excess plant material from past crops, maintaining adequate hygiene in machinery used in crop management. This may also include practising a high level of record keeping including the planting/harvest dates, isolation distance considerations, weed/pest control records, all of which have implications on possible genetic contamination (Warner & Lewis, 2019).

Figure 2-3 shows the various sources of contamination at different crop production stages and the mitigation strategies used to mitigate the risk from the contamination sources (Warner & Lewis, 2019).

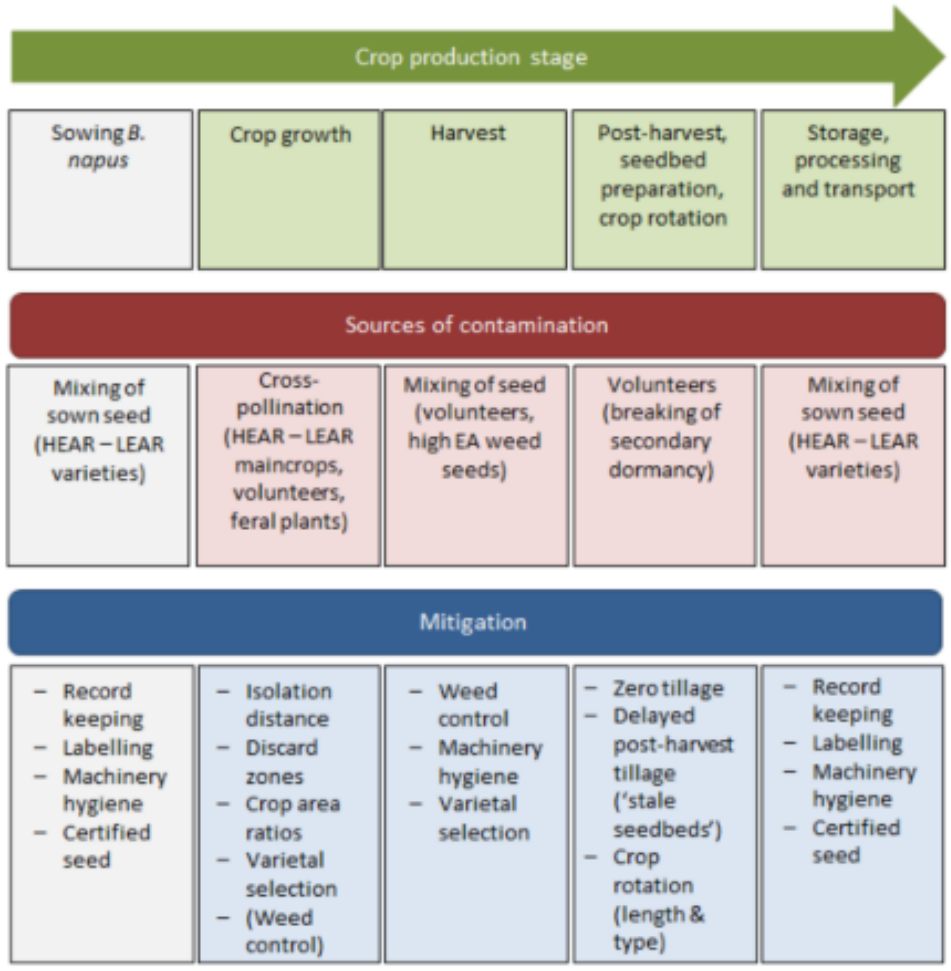

Figure 2-3: Mitigation strategies for genetic contamination from different sources at different crop production stages (Warner & Lewis, 2019)

2.9 Current weed/genetic contamination management strategy

Consultation with seed growers and rural professionals has revealed that the current method of wild carrot control in Hawke's Bay is to manually scout the nearby areas within the isolation distance and manually remove the plant. This process of pulling the plant out from the ground, including the roots is known as rouging. Rouging as a way to manage weeds is used in weed management of different weeds in many ecosystems, including weed management in cocoa production (Aboh & Effiong, 2019), removal of many aquatic weeds in a wetlands environment (Datta, 2009), and removing weeds in rice cultivation (Karim et al., 2004). It is also employed as a way to remove off type individuals in seed production systems (Chapke & Tonapi, 2016; Lavanya, 2002; Nair & Kumar, 2021). Rouging is employed in wild carrot control because its underground storage roots allow it to recover from other traditional control methods like cultural methods like mowing. However, this method is labour intensive (Hugh Ritchie, Drumpeel Farms, personal communication, November 2021) and error prone as there is a greater likelihood of weed plants being undetected in a manual inspection over a large area. Further, given the continued abundance of wild carrot in the region, the method is likely ineffective, as has been observed when rouging has been employed overseas (Rome & Lucero, 2019) and in carrot seed producing areas of Canterbury (Chris Smith, Foundation for Arable Research, personal communication, October 2022). Wild carrot remains ubiquitous in regions where it is actively controlled and scouting and rouging is the prime mode of control (Mandel et al., 2016; Rome & Lucero, 2019). In conclusion the current wild carrot control strategy of rouging is expensive and ineffective and there is a current gap in the knowledge on alternative wild carrot detection and control methods. Thus, there is a need to explore alternative weed management techniques.

2.9.1 Application of hyperspectral remote sensing in agriculture

Hyperspectral imaging and remote sensing have been employed for a wide range of applications in an agricultural context. These include determining nutrient content of various agricultural crops such as the nitrogen content of leaves in potatoes (Oliver et al., 2013), in cereals (Fu et al., 2020), monitoring phosphorus in rice cultivation (Mahajan et al., 2017) and management of nutrient loss such as monitoring N leaching in arable cropping systems (Zhao et al., 2020). Other applications are monitoring crop yield such as estimating seed weight in soy bean plants (Wójtowicz et al., 2016), developing yield prediction models in citrus fruit production (Ye et al., 2007), plant health

determination such as detecting greasy spot (*Mycosphaerella citri*) in citrus trees (Du et al., 2004), abiotic stress determination such as estimating salt stress in a sugarcane fields (Hamzeh et al., 2013) and early detection of drought stress in cereals (Römer et al., 2012).

Hyperspectral imagery is able to detect differences in biophysical properties of leaves of different weed species (Pignatti et al., 2019). The technology was able to detect/differentiate weed species *Amaranthus retroflexus* L. and *Cyperus rotundus* L. from a maize crop and from each other (Pignatti et al., 2019), detect spotted Knapweed (*Centaurea stoebe* L.) in forest areas of Idaho and Montana in the United States (Lass et al., 2002), Hoary cress (*Cardaria draba* L.) in an agricultural area of southwestern Idaho, United States (Mundt et al., 2005). These reports all providing evidence to support the hypothesis that the hyperspectral imaging could potentially be used to identify wild carrot plants in different ecosystems. Further, although limited, there is some evidence of the use case of the technology of identification at a variety/sub-species level (Tu et al., 2018).

Collection of hyperspectral data using a sensor mounted on an aircraft is called airborne hyperspectral remote sensing. The plane flies ~600-1500 m above the areas of interest and collects hyperspectral data (Veraverbeke et al., 2018). Airborne hyperspectral remote sensing allows data collection at a high spatial scale (Meerdink et al., 2019). This will allow data collection over a large surface area in relatively short period of time, allowing the large areas surrounding the carrot seed crop to meet the isolation distance requirement of 2000 m to be surveyed for the wild carrot. Once surveyed wild carrot rouging can focus on the areas where the wild carrot has been found, increasing roguing efficiency.

2.10 Hyperspectral remote sensing

For efficient and effective control, an alternative method is needed that can identify all the wild carrot plants over a large surface area (at least within the isolation distance for carrot crops) in a short period of time. Hyperspectral remote sensing technology has the potential to be an alternative method.

2.10.1 Principles of hyperspectral remote sensing

Hyperspectral remote sensing is a technology that collects high resolution spectral data remotely across a wide range of the electromagnetic spectrum, typically ranging from blue to short-wave Infrared region (~300-2500nm) (Manea & Calin, 2015; Xie et al., 2014).

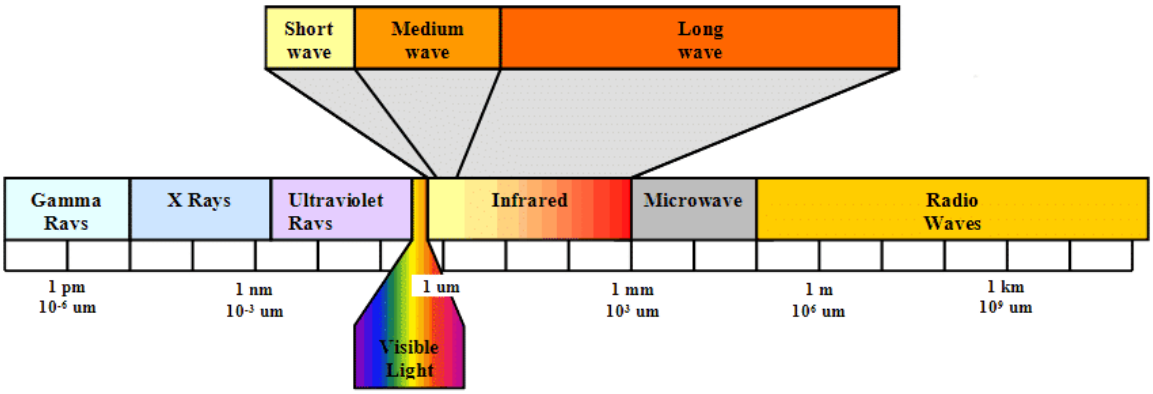

Figure 2-4: Electromagnetic spectrum (Source: (Isse & Ghouch, 2016))

This ability to collect high resolution spectral data means there is the potential for the technology to discriminate and identify the wild carrot plants among other weed species and in relevant environments (roadsides, ditches, cultivated paddocks). A hyperspectral camera/sensor collects spectral reflectance data, which is the proportion of radiation energy reflected at different wavelengths from the surface (Thorp & Tian, 2004; Wójtowicz et al., 2016). This reflectance data is high resolution as the data is collected in narrow contiguous bands across the relevant parts of the electromagnetic spectrum (Veraverbeke et al., 2018).

2.10.2 Multispectral versus hyperspectral

Multispectral and hyperspectral are two popular forms of sensors used in optical remote sensing. The major point of difference between the two sensor types is the number and size of bands. Multispectral sensors employ a lower number of bands (3-10 generally), and the width of the bands is bigger (10-20 nm). Hyperspectral sensors usually employ hundreds of narrow-contiguous bands (Srivastava et al., 2020). The figures below give an illustration of the number and width of bands across the electromagnetic spectrum in multispectral (Figure 2-5) and hyperspectral (Figure 2-6) imagery. Hyperspectral data has a greater number of bands that are narrow and contiguous, relative to multispectral data. Figure 2-5 shows multispectral bands (high band width, lower number) juxtaposed with hyperspectral bands narrow, contiguous and overlapping (not shown in figure 2-5, see Figure 9-4 in appendix for another illustration).

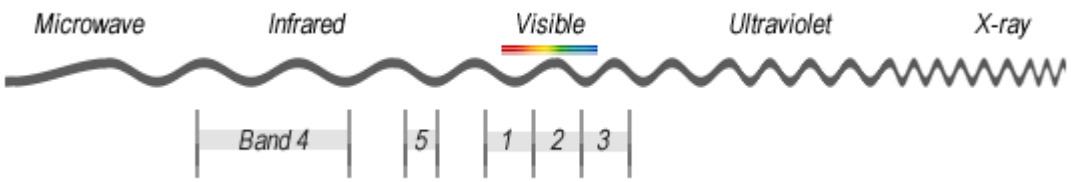

Figure 2-5: Multispectral data bands – low number of high bandwidth bands (Source: GIS geography, 2022)

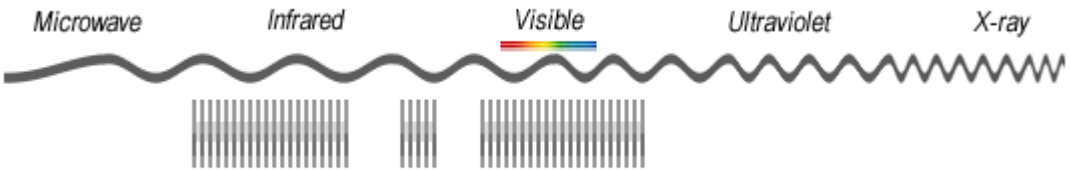

Figure 2-6: Hyperspectral data bands – high number of narrow (low bandwidth), contiguous bands (Source: GIS Geography, 2022)

Having a greater number of bands that are narrow and contiguous allows the collection of data with a high spectral resolution with more spectral detail relative to multispectral data (GIS Geography, 2022). This high spectral resolution data allows for a greater ability to discriminate between individual vegetation species relative to multispectral imaging (Adam et al., 2010).

Figure 2-7 shows a high-resolution spectral data curve collected using a hyperspectral sensor and a low-resolution spectral data curve collected using a multispectral sensor.

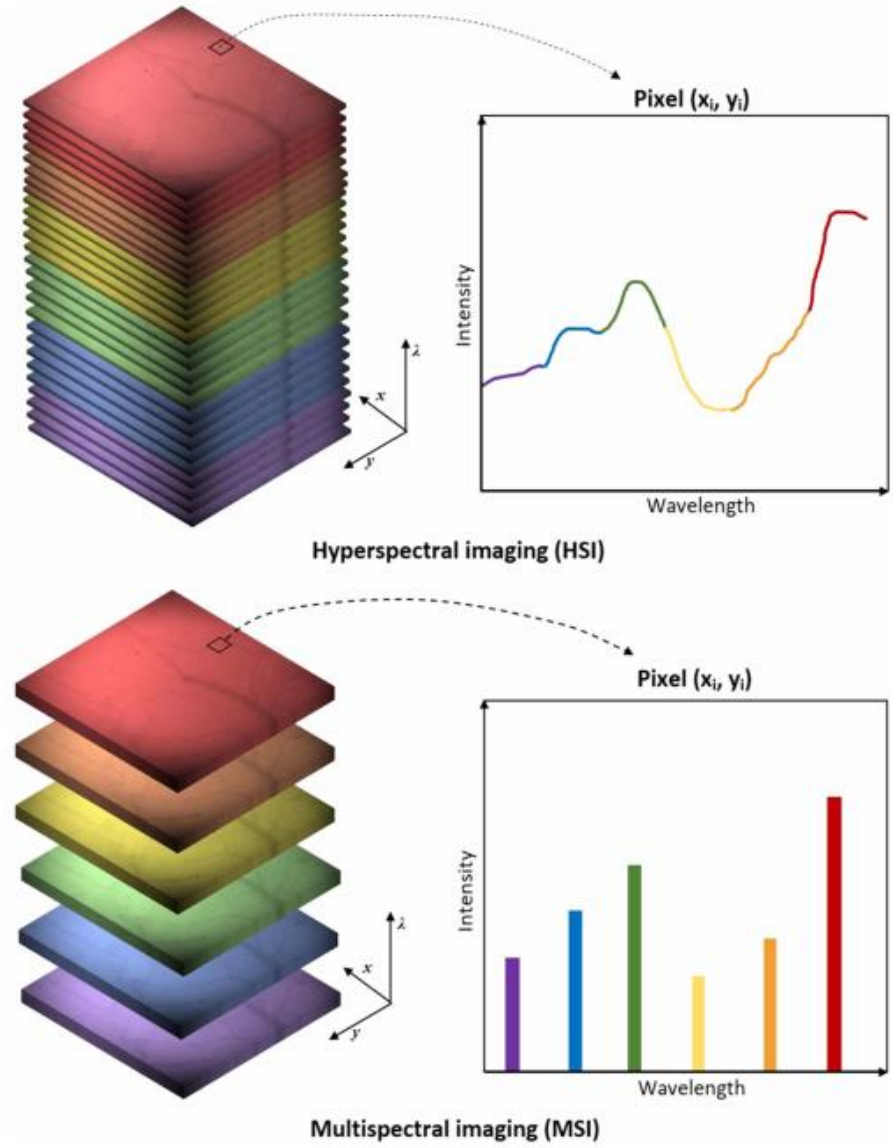

Figure 2-7: Output generated: Multispectral vs Hyperspectral sensors Illustrating the difference between a multispectral and hyperspectral spectral profile (Source: Giannoni et al., 2018)

2.10.2.1 Hyperspectral sensors

There are two main types of airborne sensors – active and passive sensors. The AsiaFENIX sensor (used in this study) is a passive sensor – which means that it does not have its own light source and records the natural light reflected from the underlying surface. As opposed to active sensors which have their own source of light and measure the reflected light they emit (Sarath et al., 2014). Figure 2-8 shows an illustration of difference in active and passive sensors. The figure shows passive sensors on left-hand side of the picture which use external source of light (the sun in this case). Active sensors on the right-hand side of the picture and have their own light source.

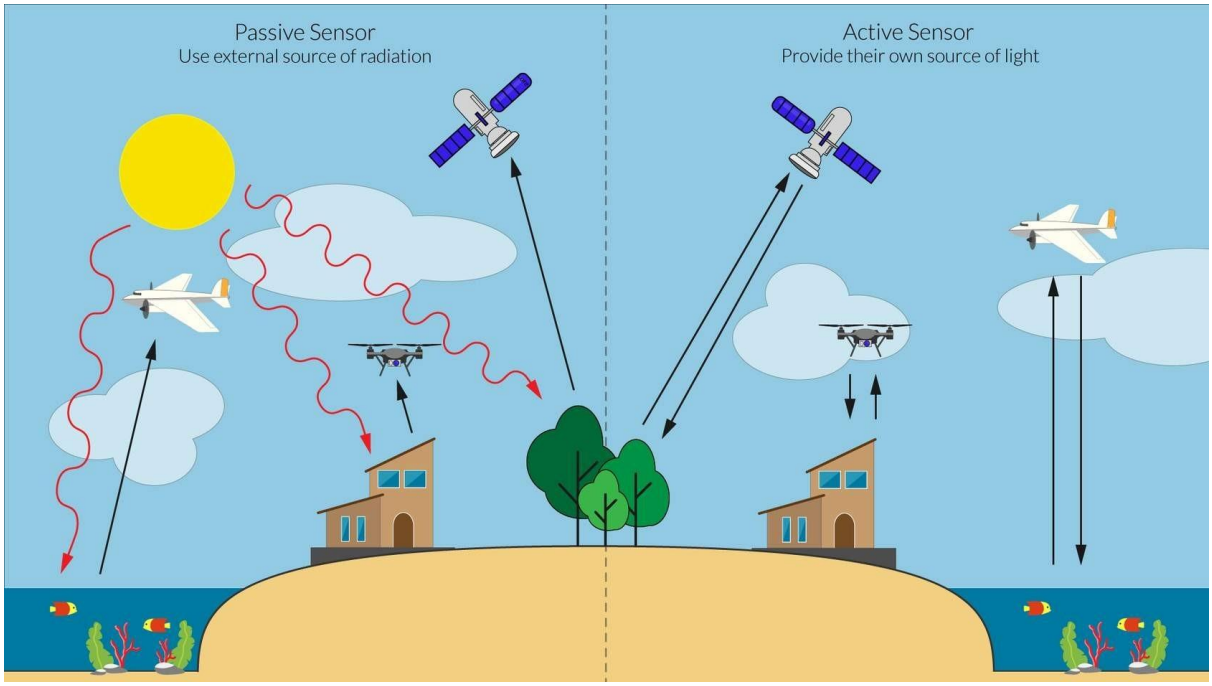

Figure 2-8: Passive sensors (Sensors on left hand side of the figure – use external light source) versus Active (Sensors on right hand side of the figure – have their own light source) in remote sensing (Source: Reef Resilience Network (2023))

Further, passive sensors are sub-divided into push broom (Figure 2-9) and whisk broom sensors (Figure 2-10). Push broom sensors hold a line of detectors on a perpendicular line to the flight direction of the aircraft. As the aircraft moves forward the data is collected one line at a time (Gibson et al., 2000a). All the pixels that fall onto the same line are measured simultaneously. Since these sensors have multiple small detectors, they must be calibrated appropriately and regularly to minimise variation between detectors (L3Harris Geospatial Solutions, 2017). This is different to another form of sensor known as Whisk broom sensors where there is only one detector and light is reflected onto it by a mirror moving back and forth. This sensor collects data from a single pixel at one time (Gibson et al., 2000a). Whisk broom sensors tend to be more expensive and can incur wear and tear at a faster rate due to their moving parts relative to push broom sensors. Figures 2-9 and 2-10 illustrate the workings of the push broom and whisk broom sensors (L3Harris Geospatial Solutions, 2023).

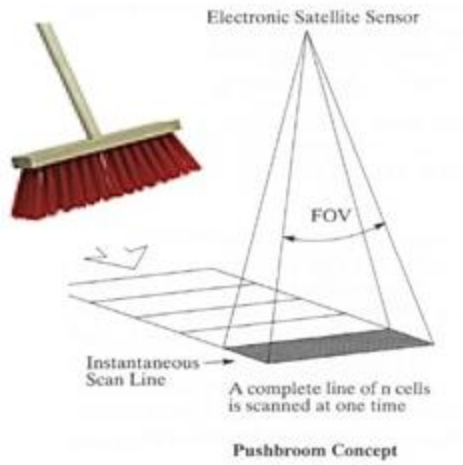

Figure 2-9: Push broom sensor concept (Source: (L3Harris Geospatial Solutions, 2023))

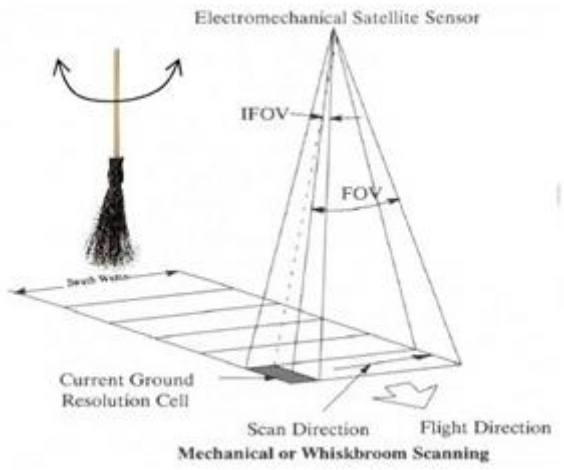

Figure 2-10: Whisk broom sensor concept (Source: (L3Harris Geospatial Solutions, 2023))

2.10.2.2 GNSS-GPS

The collection of aerial survey data is usually accompanied by ground truth data collection which aids in model development and validation of the method processes.

Ground truth data is usually collected using an accurate GPS device. Like a GNSS enabled GPS device. Global navigation satellite system (GNSS) enabled GPS device allows for a greater degree of accuracy and reliability on the location of interest (< 100 cm). The ground data will allow for the development of spectral libraries for dominant weed species, including wild carrot, and provide a foundation for the spectral unmixing and classification methods to be applied. This method of collecting aerial hyperspectral data and ground based validation with a view to discriminate between crops and weeds and between different weed species has been used before (Pignatti et al., 2019).

2.10.3 Data pre-processing

Hyperspectral remote sensing is an analytical technique which includes variables that influence the hyperspectral data. These variables need to be accounted for, to have reliable final output - spectral reflectance values. Camera, surface, and the light source are the main variables (Amigo & Santos, 2020)

Camera, surface and the light source are the main analytical variables/elements in hyperspectral remote sensing (Amigo & Santos, 2020). The photons received onto the camera/sensor from the

light source (in this case is the Sun) are influenced by the atmosphere, viewing geometry, and particles in the atmosphere. This makes their reading on the sensor subject to fluctuations. The photons that do reach the target surface are further influenced by the physical and chemical (material composition) and physical characteristics (smoothness/roughness, relief) of the target surface (Amigo & Santos, 2020). The camera/sensor detect the reflected radiance from the target surface (and from the background) but are also subject to instrument's spectral, spatial, and radiometric resolution. Thus, the recorded imagery is influenced by quality of the instrument too.

The influence of these individual aspects of the data collection process on the recorded data is a combination of effects accumulated across the 3 aspects. These effects may be further categorized into distortions in geometry and distortions in spectral signal.
Preprocessing refers to operations that need to be undertaken to minimize the effects of the spectral and geometrical distortions in the spectral signal due to the camera, surface and light source and prepare the data for the main image analysis. These are generally divided into three steps:

- Radiometric correction,
- Atmospheric correction and
- Geometric correction.

2.10.3.1 Radiometric correction

Radiometric correction refers to transformation of recorded raw data or Digital number (DN), which are sensor recorded radiance quantified into electronic intervals depending upon the radiometric resolution of the sensor and are not physically interpretable, into physical units that are comparable and transferable. Here digital numbers (DN's) are converted into radiance units. (EO College, 2021).
A radiometric correction is performed using a process called linear transformation - this requires use of correction coefficients, which are applied to every image pixel. Correction coefficients are determined for each detector, during their radiometric calibration. (EO College, 2021).

2.10.3.2 Atmospheric correction

The recorded radiance of a given pixel tends to be a combination of radiation from the target surface, radiance reflected from the atmosphere, and from the adjacent surfaces. In the presence of the Earth's atmosphere, however, the recorded radiance is distorted by the atmospheric

scattering and absorption due to the gases and aerosols. Scattering reduces the radiant energy of radiation passing through the atmosphere in the incident direction by redistributing it towards other directions resulting in distorted remote measurement of the surface reflected radiation (Bhatia et al., 2015; Liang et al., 2020). Another process that distorts the at-sensor radiance is absorption. Like scattering, absorption is a function of wavelength. However, unlike scattering, absorption represents a transformation of the radiation into another form of energy. The spectral ranges at which radiation is absorbed by atmospheric constituents are known as absorption features. Light in a strong absorption feature cannot penetrate the atmosphere, and thus cannot reach a sensor (Gibson et al., 2000b).

Atmospheric correction is the process of adjusting for the phenomena of absorption effects and scattering. It is used to retrieve surface-reflected radiation (actual reflected radiation from target surface) from at-sensor radiance (radiation perceived at the sensor). It can help quantify these energy components to obtain an estimation of surface reflectance values (Hadjimitsis & Themistocleous, 2008). For example - when the energy contribution from the target surface is 25 units, the contribution from the atmosphere and its corresponding effects is 60 units, then final surface reflectance value recorded at the sensor abord the aircraft would be 85 units (target + atmospheric effects contribution) (Hadjimitsis & Themistocleous, 2008). Thus, it is crucial to perform an atmospheric correction and adjust for atmospheric effects.

2.10.3.2.1 Absorption effects

Absorption effects refers to the phenomena of selective absorption of radiation of different wavelengths by gaseous components of the atmosphere. Different molecules absorb different wavelengths (Gibson et al., 2000b). Radiation of all wavelengths in the electromagnetic spectrum reach the top of atmosphere, but only a few reach the earth's surface due to these absorption effects. Undesired absorption effects tend to add noise to the data (Amigo & Santos, 2020).

2.10.3.2.2 Scattering

Scattering of radiation by physical particles is a common natural phenomenon. Scattering is the reason an object can be seen under shade (Gibson et al., 2000b). There are 2 types of scattering – selective and non-selective scattering. Selective scattering or molecular scattering is where the particle size is one tenth or smaller than the wavelength of the radiation (Gibson et al., 2000b).

Non-selective scattering happens via scattering from particles like dust aerosols, smoke, pollen, whose size is 10x or greater than the wavelengths of the radiation (Gibson et al., 2000b). Figure 2-11 illustrates the scattering phenomenon.

Figure 2-11: Scattering phenomenon: a. causing illumination of areas under shadow, b. reason for different colours/appearance in the sky (Gibson et al., 2000b)

As illustrated in Figure 2-11, non-selective scattering causes illumination of regions in shadow whereas selective scattering of radiation of different wavelengths causes degrees of scattering in light of different colors. This gives the sky a blue appearance during the day (blue light most scattered) and the red/orange during sunrise/sunset (red/orange most scattered).

Also, due to scattering, the diffused solar radiation enters the instantaneous field of view (IFOV) of a sensor which is called path radiance. Furthermore, the scattering causes radiation reflected from the background of the target surface to enter the IFOV of a sensor. In the presence of the atmosphere, at-sensor radiance is, therefore, composed of three radiance components: 1)

distorted surface-reflected radiance; 2) path radiance; and 3) background radiance (Bhatia et al., 2018; Dennison et al., 2006)

In an ideal world, the radiation recorded from a single pixel in hyperspectral data will reflect the true surface reflectance value of the pixel. However the aforementioned effects of the phenomenon of scattering can cause reflectance of nearby pixels (background of target surface) to be reported as reflectance of pixel in question, compromising the data quality (Bhatia et al., 2018; Dennison et al., 2006)

For this study, this could mean that data from a pixel with wild carrot plants can be contaminated by data from nearby pixels, which could interfere with identification of spectral signatures associated with wild carrot plants. This could also mean pixels that do not contain wild carrot could have wild carrot spectral signatures emanating from nearby pixels containing wild carrot. This would lead to greater number of false positives and reduce the overall accuracy of the classification model. This suggests that scattering effects have important implications for the classification analysis and could eventually undermine the study's aim of detection of wild carrot plants.

Thus, the phenomenon of scattering is an issue in hyperspectral spectral data analysis and must be adjusted for by using atmospheric correction processes to maintain data integrity/quality and prevent undermining the wider aim of the study to detect wild carrot plants.

A 'radiative transfer model' approach will be taken to perform atmospheric correction on the data in this study. This approach to atmospheric correct takes into account both scattering and absorption phenomena and removes these effects from the at sensor radiance and gives a more accurate surface reflectance values (Bhatia et al., 2018).

2.10.3.3 Geometric correction

Geometric correction is used to translocate a geometrically distorted image to a standard reference - like coordinates on a map. These geometric distortions can be caused by factors like the surface topography, curvature of the earth, movement of the sensor (roll, yaw, and pitch angles) the tilt of the sensor and the altitude of the platform (EO College, 2022)

The correction process takes place in 2 parts -

1. Finding a correct set of coordinates in the image to be geometrically corrected and in the reference map.
2. Interpolating the distorted image to the correct reference points.

Geometric correction establishes a mathematical relationship between the coordinate system of the image and a standard geographical coordinate system (Japan Association of Remote Sensing, 1996). A popular way to identify this relationship is using a geo correction method which is based on using geometric control points (GCPs) that are common among both the reference map and the image to be processed. GCPs are usually permanent elements like roadways, airport runways, permanent road markings and buildings/structures (Amigo & Santos, 2020).

Another method is to georeference from an input geometry data file that has map locations for every pixel for the hyperspectral data (L3Harris Geospatial Solutions**,** 2023).

2.10.4 Classification algorithms

Weed management using hyperspectral remote sensing will require identification of weed plants and creation of weed maps locating the plants. Species identification and mapping work using hyperspectral data typically involves analysis of the data by employing statistical tools known as classification algorithms (Roth et al., 2015).

A pixel is the smallest spectral data unit in a hyperspectral image. Hyperspectral data has reflectance values at different wavelengths (bands) for each individual pixel. This allows calculation of basic statistics like mean, data minimum/maximum, range, and standard deviation. Classification algorithms in hyperspectral data are statistical tools. Using these basic statistics clustering of pixels, into a small number of classes is carried out. The pixels belonging to one class have similar properties (Mohd Hasmadi et al., 2009) – in this case this means that the areas represented by pixels in the same class have similar land surface cover. Most classification is based on the spectral response from different plants/materials. Classification algorithms will classify/allocate pixels into pre-defined classes based on their spectral properties. Usually the end product of application of a classification algorithm is a land surface cover map used for various applications (Zhao et al., 2022), and in this case: weed maps.

There are two broad categories of classification methods - unsupervised and supervised. Unsupervised classifications use algorithms that do not have user's input through training data. These algorithms can be used to create land surface cover maps without collecting any field ground-truth data and without any prior knowledge of the ground cover of the area of concern (Mohd Hasmadi et al., 2009). Supervised image classification uses algorithms that require input from the user in form of training data identifying individual classes/materials. These algorithms require prior knowledge of the land surface cover types. Training data can be ground-truth data or could be spectral data from a spectral library (Mohd Hasmadi et al., 2009)

2.10.5 Mixture-tuned matched filtering (MTMF)

Mixture-tuned matched filtering (MTMF) is a supervised classification method. MTMF as a classification method to detect invasive plants and weed species has been used widely in literature across various ecosystems and for various plant species. MTMF was used to detect cheatgrass (*Bromus tectorum* L.) in the intermountain rangelands of Washington in the United States (Noujdina & Ustin, 2008), hoary cress (*Cardaria draba* L.) (Mundt et al., 2005) and leafy splurge (Glenn et al., 2005) in the agricultural regions of Idaho in the United States and perennial pepperwood (*Lepidium latifolium* L.) in the wetland and riparian regions of California in the United States (Andrew & Ustin, 2008).

Spectral data at a specific pixel is a combination of spectral responses of its constituent plant species and materials. Spectral unmixing is a process used for decoupling the spectral response of an individual endmember. Endmember spectra is spectra from material at its purest form (Shi & Wang, 2014).

MTMF is a form of classification algorithm that employs and utilises spectral unmixing. Spectral unmixing is the process of decomposing the spectral signature of a mixed pixel into a set of endmembers (Shi & Wang, 2014) – endmembers could be individual species like wild carrot plants or material like soil/rocks. In simple terms, it is the process of decomposing spectral response from a pixel that is mixed (has different materials, many plant species) to spectra of its constituent materials/plant species.

MTMF uses spectra of the endmember of interest, which in this case are wild carrot plants, as a target spectra (Noujdina & Ustin, 2008). It then decouples the target spectra from other spectra – giving an indication of likelihood of presence and abundance of the target endmember (wild carrot

in this case) in the pixel. The input data required to apply MTMF is minimum noise fraction (MNF) transformed data. MNF is a two-step principal component analysis. In the first step, MNF decorrelates and rescales the noise in the data. This allows differentiation between bands that are noise whitened (first few bands) and bands which are almost most fully noise (last few bands). In the second step it consolidates the spectral information of the full dataset into fewer transformed bands (Underwood et al., 2007). This captures most of the variability and valuable spectral information - like unique spectral reflectance features which may help detect target - of the entire data set in fewer and less noisy bands (Underwood et al., 2007; Mundt et al., 2005). This technique greatly reduces the amount of data to be processed, eliminates most of the noise from the data, and makes the bands more distinguishable. All this has positive ramifications for and streamlines the classification process (Hamada et al., 2007).

MTMF projects the MNF input data onto the target data vector and tries to identify the spectral component of the target within each pixel (Mundt et al., 2005). The output of MTMF includes two sets of scores for each pixel – the MF score and the infeasibility score. The MF score gives an estimate of the relative match of the pixel to the target spectrum and approximate sub-pixel abundance of the target endmember. A score of 1.0 is a perfect match – suggesting there is a high likelihood of wild carrot presence on this pixel (Noujdina & Ustin, 2008). The infeasibility score helps reduce the number of false positives. The infeasibility score is calculated using the noise variance of the input data and indicates the feasibility of the MF scores (Mundt et al., 2005; L3Harris Geospatial Solutions, 2022). A high infeasibility score is associated with a low likelihood of presence of the target endmember at a pixel (Noujdina & Ustin, 2008; L3Harris Geospatial Solutions, 2022).

2.11 Aims of the study

The purpose of this review was to identify knowledge gaps and develop the research question(s) relating to wild carrot weeds as a constraint in carrot seed production in the Hawke's Bay. This review has identified that presence of wild carrot is a constraint to carrot seed production in the Hawke's Bay. Further, the current wild carrot control strategy of manual scouting and rouging is expensive, time-consuming, and given the continued abundance of wild carrot, ineffective. Literature evidence has suggested that hyperspectral remote sensing, using associated classification/spectral unmixing algorithms, has the potential to provide a superior alternative

weed control approach and would be a worthy technology to explore further with the following **research question**:

Is Hyperspectral remote sensing a feasible method for early detection of wild carrot in carrot seed production?

This study therefore aims to:

1. Identify areas with wild carrot populations in the Hawke's Bay and collect hyperspectral data.
2. Develop a database of the environmental context of the wild carrot populations.
3. Assess the utility of MTMF as a classification algorithm to detect and classify wild carrot plants.
4. Assess possibility of developing weed maps with accurate potential wild carrot GPS locations.

Chapter 3 Materials and methods

3.1 Introduction to methodology/data analysis schematic

To determine the efficacy of hyperspectral remote sensing in detection of wild carrot - initially an appropriate data collection site is located – this is an area where wild carrot is growing, preferable in a region where carrot seed is produced. This is followed by data collection.

There are two broad steps in the data collection process – the first step in the data collection process is to acquire hyperspectral data of the area including wild carrot plants, the surrounding vegetation and other features such as roads and / or water ways. Where a large area needs to be surveyed this is achieved through an aerial survey (3.2 Airborne hyperspectral data collection). The second part of the data collection process is collection of ground truth data. Collection of aerial survey data is usually accompanied by ground truthing – especially in novel studies. The ground data aids both model development and validation of the method. Figure 3-1 shows a schematic of the data collection/analysis methodology used in this study.

Specifically, Figure 3.1 shows primary initial data collection steps of ground truth data collection and aerial hyperspectral data collection. Data acquired in these two steps was analysed following the series of steps given in the figure culminating in the production of weed maps with potential locations of wild carrot occurrences.

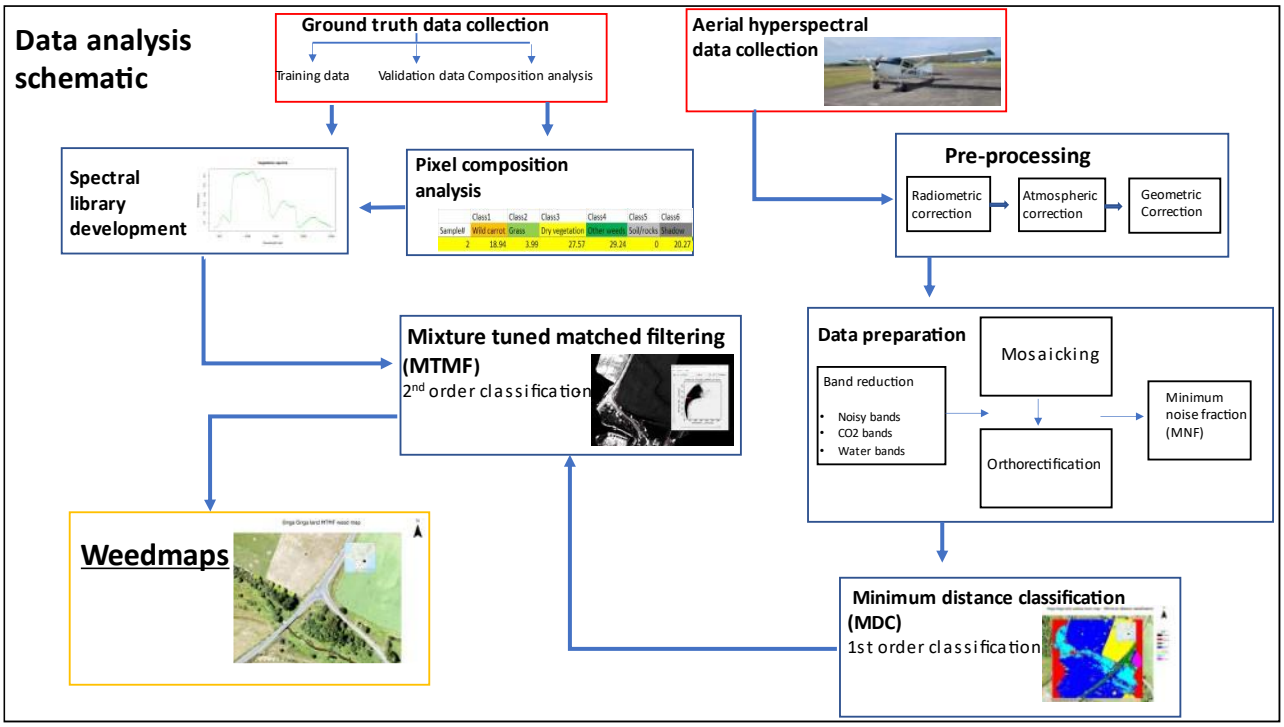

Figure 3-1: Weed detection with hyperspectral data - Methodology and data analysis schematic.

3.1.1 Research location

This research was undertaken in Hawke's Bay, specifically near OngaOnga (wild carrot 'Hotspot1' co-ordinates: -39.909, 176.413). The research location was selected in discussion with South Pacific Seeds (Mitch Wolting, South Pacific Seeds (NZ) Ltd, personal communication, November 2021) as an area where wild carrot has been found. The areas in and near OngaOnga are also carrot seed producing areas hence the spectral data captured will have wild carrot in a carrot seed production environment.

3.2 Airborne hyperspectral data collection

Collection of airborne hyperspectral data is the initial step of the data collection methodology adopted in this study. The data was collected to enable identification and detection of wild carrot populations in the wild carrot hotspots present near carrot seed producing areas within Hawke's Bay.

The methodology is based on flying a hyperspectral camera mounted on a fixed-wing aircraft over the study area. The camera on the plane is AsiaFENIX airborne hyperspectral sensor (Table 3-1). In addition to hyperspectral images, RGB images were collected using Nikon D810 RGB camera. This is a higher resolution camera used for accurate georeferencing to locate ground points (Table 3-1).

Airborne hyperspectral data collection was followed by field-based identification and accurate GPS location tracking of wild carrot plants using a Leica Zeno 20, and RGB image collection of wild carrot plants using a Sony SLT-A57 Camera with a Sony SAL1855 Lens (Table 3-1)

Table 3-1: Equipment used in hyperspectral, RGB and ground-truth data collection and equipment specifications.

Equipment model name	**Purpose**	**Specifications**
AsiaFENIX	Hyperspectral sensor – airborne	Range – 370-2500 nm , spectral resolution – 3-5 nm (NIR region) and 8-11 nm (SWIR region), Spatial resolution – 1m
Nikon D810	Aircraft mounted-RGB camera – airborne - accurate georeferencing	Resolution: The D810 has a 36.3-megapixel CMOS image sensor. ISO range: of 64-12,800, which can be expanded to 32-51,200, allows shooting in low-light conditions. Autofocus and continuous shooting systems capable.
Leica Zeno 20	RTK-GPS	Accuracy <100cm
Sony SLT-A57 Camera with a Sony SAL1855 Lens	RGB camera – ground truthing	Sensor: 23.5 x 15.6mm (APS-C type), Exmor APS HD CMOS sensor with RGB primary colour filter Resolution: Total: 16.7 megapixels; Effective: 16.1 megapixels

The AsiaFENIX sensor is a passive sensor. This means it does not have its own light source and records the natural light reflected from the underlying surface. The sensor is in the pushbroom sensor category as described in the literature review chapter (see 2.10.2.1 Hyperspectral sensors)

The GPS device used is a global navigation satellite system (GNSS) enabled device, which allows for a high degree of accuracy and reliability on the location of the plants (< 100 cm).

Figure 3-2 shows the fixed wing aircraft used to capture the hyperspectral data. Figure 3-3 and Figure 3-4 show the aerial view from the aircraft collecting hyperspectral data - also visible is the orthoimage RGB camera attached to the aircraft wing strut.

Figure 3-2: Fixed wing aircraft used for Hyperspectral data collection (Photo: Stefan Carter, Massey University)

Figure 3-3: Hyperspectral data collection – view from the aircraft (Photo: Stefan Carter, Massey University)

Figure 3-4: Hyperspectral data collection – view from the aircraft (Photo: Stefan Carter, Massey University)

The aerial hyperspectral data collection (aerial survey) and part of data pre-processing (see 2.10.3.1 Radiometric correction, 2.10.3.2 Atmospheric correction) was undertaken by Massey Agrifood Digital Lab (MAF), Massey University, Palmerston North, New Zealand.

3.3 Field work

3.3.1 Field sites

The data collection sites included two sites which are hotspots for wild carrot presence near the carrot seed production areas in OngaOnga, Hawke's Bay. The wild carrot hotspots were located at: -39.909, 176.413 ('Hotspot 1' - intersection of State Highway 50 and OngaOnga Road); -39.934, 176.517 ('Hotspot 2' on OngaOnga Road). The hotspots were chosen as these areas were confirmed locations with wild carrot populations growing relatively close to carrot seed production paddocks (Mitch Wolting, South Pacific Seeds (NZ) Ltd, personal communication, 2021) and thus allowed collection of hyperspectral and ground truth data of wild carrot growing in a carrot seed production environment, which fits the sampling criteria learnt from literature review (see 2.9 Current weed/genetic contamination management strategy, 2.10.1 Principles of hyperspectral remote sensing and 2.10.2.2 GNSS-GPS). Figure 3-5 below shows the two wild carrot hotspots in OngaOnga, Hawke's Bay.

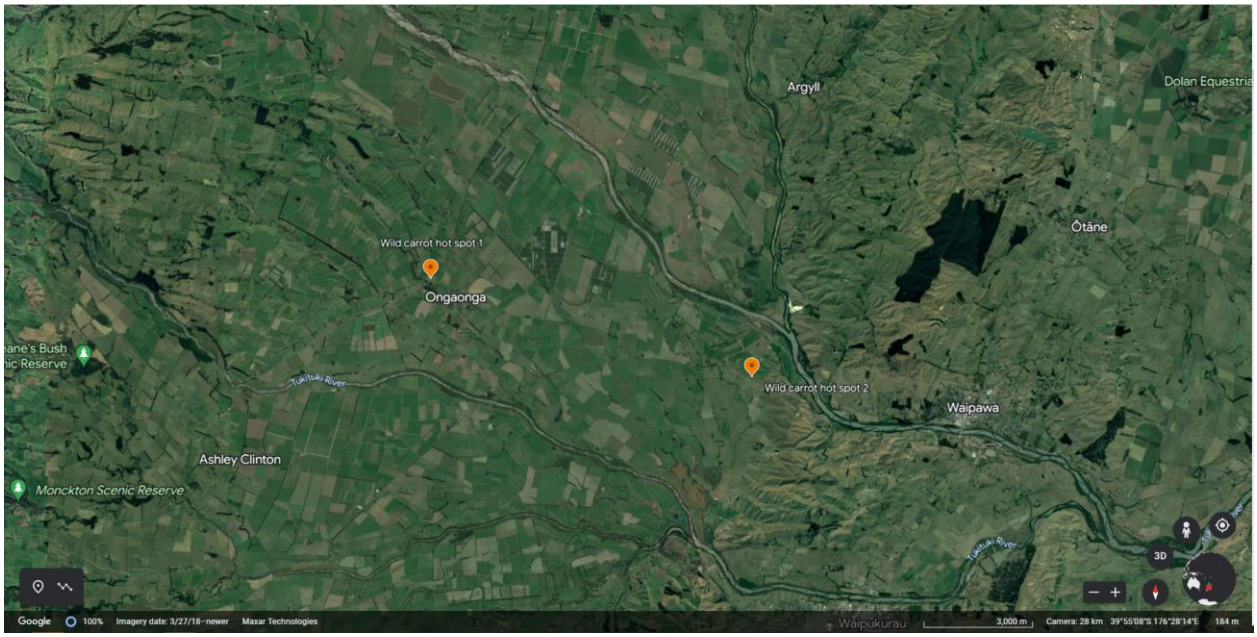

Figure 3-5: Map giving the locations of the two wild carrot hotspots (data collection sites) in OngaOnga, Hawke's Bay - Hotspot 1 (-39.909, 176.413), Hotspot 2 (-39.934, 176.517)

Figure 3-6 below shows the flight lines over the two wild carrot hotspots in OngaOnga, Hawke's Bay.

Figure 3-6: Flight lines over the two wild carrot hotspots in OngaOnga, Hawke's Bay.

3.3.2 Ground-truth data

Ground truth data is collected to overlay and match the field data (wild carrot GPS locations) to the hyperspectral data collected in the aerial survey. The ground truth data acts as training and validation data for the classification analysis. This is a crucial step in the data analysis and

classification analysis performed with an aim to detect and identify wild carrot plants in the hyperspectral data.

Ground truth data was collected at the two known wild carrot hotspots (Figure 3-5). The wild carrot patches were identified, and an accurate RTK-GPS location was captured using GPS Leica Zeno 20. The model is GNSS-environment enabled system and can collect data to a high level of accuracy (up to 1 cm) (Leica Geosystems, 2022) (Figure 3-7).

Part of the ground truth points acted as training data for the subsequent classification (MDC) and spectral unmixing algorithms (MTMF). Training data in this case was used to create the MTMF model which in turn predicts the potential wild carrot locations. RGB images were captured of each wild carrot patch from a nadir angle for a pixel composition analysis. The RGB pictures were collected using Sony SLT-A57 Camera using a Sony SAL1855 Lens. A measuring tape was added to the scene before acquiring the picture to provide scale and to ascertain an approximate surface area coverage of the images so the composition analysis data can be seen as pixel compositional data. The RGB image covered a surface area of around 1.5×1.5 m, which included the wild carrot plants and the surrounding vegetation or growing environment. A total of 65 wild carrot patches were sampled (Figure 3-8) with 24 samples being in hotspot 1 and 41 samples being in hotspot 2. A further 8 images were taken which represented the most common background for the wild carrot plants.

Figure 3-7: Ground truth points - data collection. Obtaining the RTK-GPS location using the GPS Leica Zeno 20

Figure 3-8: Ground truth data point – Wild carrot patch and surrounding vegetation RGB image with the tape measure for scale in the foreground.

3.4 Data preprocessing

In hyperspectral remote sensing there are variables that influence the final output data. These variables need to be accounted for, to have reliable final output (spectral reflectance values). Camera, surface and the light source are the main variables (Amigo & Santos, 2020)

Preprocessing refers to operations that need to be undertaken to minimize the effects of the spectral and geometrical distortions in the spectral signal due to the camera, surface and light source and prepare the data for the main image analysis. These are generally divided into three steps:

- Radiometric correction,
- Atmospheric correction and
- Geometric correction.

3.4.1 Radiometric correction

For this study a radiometric correction was performed using a process called Linear transformation - this requires use of correction coefficients, which are applied to every image pixel. Correction coefficients are determined for each detector, during their radiometric calibration. (EO College,

2021). CalGeo Pro (Specim, Spectral Imaging Ltd) was the software used for radiometric correction in this study.

3.4.2 Atmospheric correction

An atmospheric correction is needed to remove the effects of atmospheric phenomena like absorption and scattering from the final surface reflectance data (see 2.10.3.2 Atmospheric correction).

A 'radiative transfer model' approach was taken to perform atmospheric correction on the data in this study. This approach aims to model how the radiation interacts with the gases and particles in the atmosphere and tries to adjust the at sensor radiance accordingly. This approach takes into account both scattering and absorption phenomena, removes these effects from the at sensor radiance and gives more accurate surface reflectance values (Bhatia et al., 2018; Dasf et al., 2018).

3.4.3 Geometric correction

There are factors in hyperspectral data collection process - like the surface topography, curvature of the earth, movement of the sensor (roll, yaw and pitch angles) the tilt of the sensor and the altitude of the platform - that cause geometric distortions in hyperspectral images.

Geometric correction is the process is used to translocate a geometrically distorted image to a standard reference - like coordinates on a map.

The correction process takes place in 2 parts -

1. Finding a correct set of coordinates in the image to be geometrically corrected and in the reference map.
2. Interpolating the distorted image to the correct reference points.

Geometric correction in this study was performed by georeferencing an input geometry data file that has map locations for every pixel for the hyperspectral data. A GLT file (geographic lookup table) was used as the input geometry data file. The GLT file was developed using the aeroplane's navigation data and a georeferenced digital elevation model (DEM). The tool used was 'Georeference from GLT' in ENVI® v 5.6 software (Manufactured by: L3Harris Geospatial. Based in: Boulder, Colorado, United States). Here the tool uses the georeferencing information from the GLT file to remove geometric distortions and applies appropriate georeferencing to the data.

3.5 Data analysis/Image processing

3.5.1 Composition analysis

Composition analysis involves quantifying the field-based data (RGB images) into the proportions of its constituent plant species/materials or components. In this instance the images were divided six classes that represented different components. These were – 1. wild carrot plants, 2. broad leaf weeds, 3. grass, 4. dry vegetation (dry vegetation here was defined as vegetation where majority of the tissue has undergone chlorosis and/or necrosis.), 5. soil/rocks and 6. shadow. Each data point (wild carrot patch RGB image) covers an area of ~1.5*1.5 m and the components of the image are quantified based on the surface area they cover within the total area of the image (eg: 10% rocks, 50% Grass, 30% dry vegetation, 10% Wild carrot - Figure 3-9). Here the software adds points randomly on the image and the user (human) manually assigns a single class to the point based on the plant species/materials where the point occurs. Table 3-2 below shows the different software used for different data analysis steps during the study.

Table 3-2: Software used for Image/data analysis, composition analysis, GIS and map creation across the study.

Software	**Purpose**
ENVI® v 5.6	Image/data analysis
Jmicrovision v1.27	Composition analysis
ArcGIS Pro 2.5.0	GIS, map creation
RStudio v.4.2.1	Statistical analysis

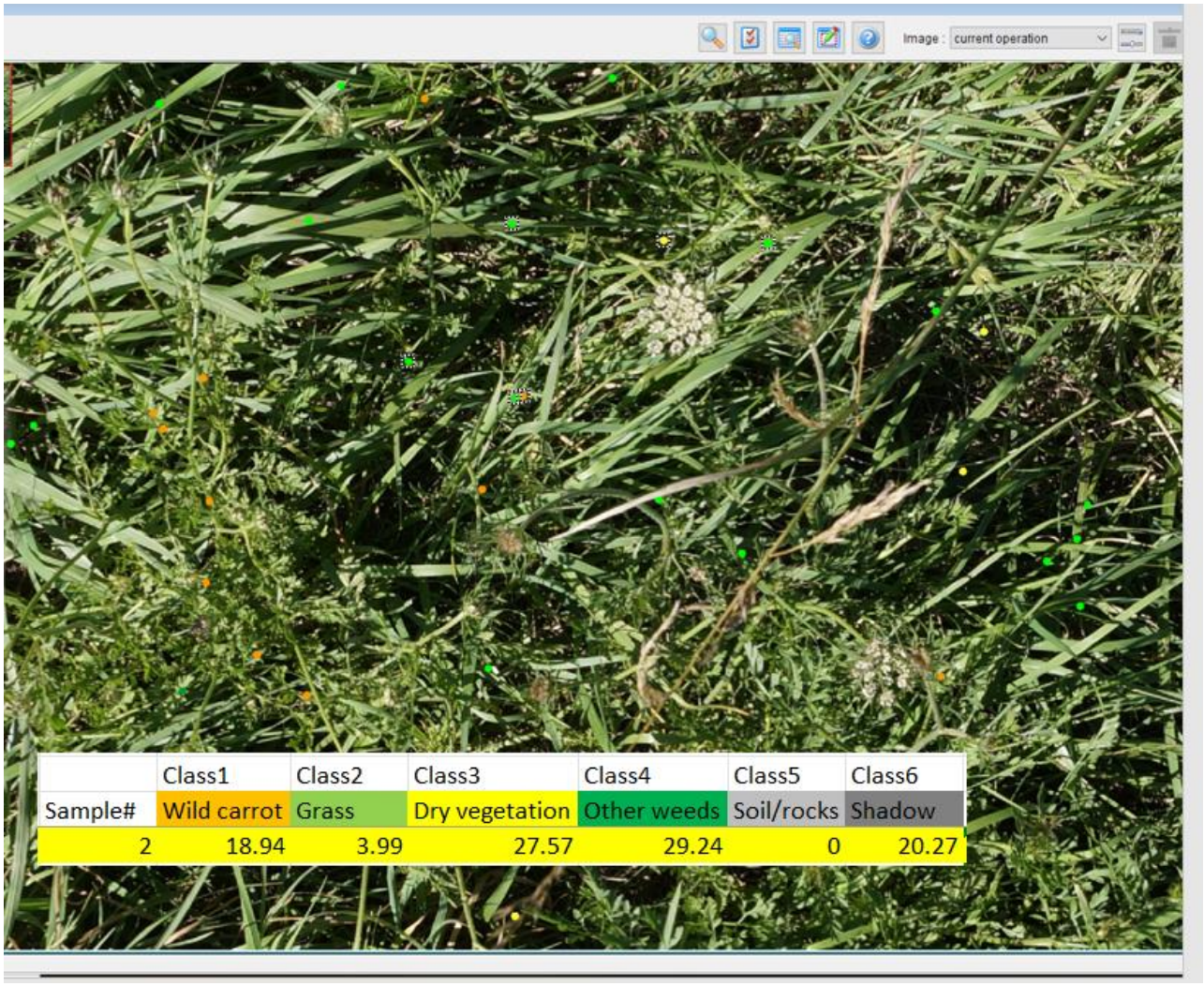

Figure 3-9: Composition analysis of wild carrot patch RGB image taken in OngaOnga, Hawke's Bay (The white flower head of the wild carrot is visible in the patch)

3.5.1.1 Statistical analysis

A statistical analysis was carried out for the compositional analysis data. All the analysis was carried out using RStudio v.4.2.1.

A boxplot was created for the wild carrot patch pixel composition data using function 'boxplot()'.

Analysis of variance (on ranks) was carried out to test the difference between the grass and dry vegetation composition of 'small', 'medium' and 'large' wild carrot patches. A normality test was conducted on the sample means to ensure the sample sets belonged to a population with normal distribution. The Shapiro-Wilkes normality test was conducted using the 'shapiro.test()' function.

The test showed non-normal distribution; thus a non-parametric test (no assumptions about sample distribution) was used to test if sample sets belonged to different populations (statistically significant difference). The Kruskal-Wallis Rank Sum Test was used to test if the difference

between sample sets was statistically significant and they belonged to different populations using the ‘kruskal.test()’ fuction. ‘bartlett.test()’ function was used to test if the sample sets had equal variance, which is an assumption that must be satisfied for the Kruskal-Wallis test to be valid. See appendix for all outputs.

3.5.2 Geo-orthorectification

Acquisition of geo-data from different sources like airborne imagery and on-ground GPS points can have different positioning accuracies and have georeferencing issues when the data are overlayed (Habib et al., 2017; EO College, 2022). Image registration is the process which allows geometric alignment of two images into one coordinate system which means that the corresponding pixels represent the same objects (Jin, 2018)

The hyperspectral images acquired in this study had inaccurate georeferencing and they were registered to orthorectified images with accurate georeferencing. The orthoimages were obtained from a RGB camera mounted on the aircraft. The data was collected at the same time as the hyperspectral data. The high resolution RGB images taken with the aircraft-mounted cameras are stored with precise coordinates using the plane’s GPS. Further a photogrammetry software, Pix4D v4.8.2, was used with these images to create orthoimages with accurate georeferencing. The hyperspectral data was georeferenced by registering it to the RGB orthoimages. This allows alignment of the hyperspectral image into the coordinate system of the orthoimage which has accurate georeferencing.

The image registration was done using the 'SPEAR Image-to-Map' Registration tool in ENVI® v5.6. The process involved plotting corresponding accurate points of physical locations/features such as road markings and physical structures such as building edges within the image to be registered and the reference image. Within the tool, the image to be registered (hyperspectral image) is the Warp image while the reference orthoimage (RGB-frame image) is the base image. These matching feature points are referred to as tie points or geographical control points (GCPs) (Schläpfer & Richter, 2002). The tie points are used in a computation process which results in a geometric transformation between the two – with the final result being an accurately georeferenced warp image (hyperspectral image).

3.5.3 Data preparation

3.5.3.1 Band reduction

Hyperspectral data includes collection of data in narrow contiguous bands that represent spectral characteristics of the underlying surface in specific wavelength range across a wide range of electromagnetic spectrum. For FENIX sensor, the spectral range is ~370-2500 nm and the number of bands is 449.

However, not all the information collected and condensed into the hyperspectral data is relevant and for some bands there is a low signal to noise ratio. Further, because of the contiguous nature of bands, there are bands that contain information overlaps with previous bands and can at times contain redundant information.

The process of eliminating these bands with redundant information is called 'Band reduction'. There are also wavelengths/bands which heavily interact with atmospheric water or carbon dioxide, and do not contain any meaningful information. Band reduction allows saving computational resources by ensuring only data that is relevant is analyzed. Band reduction also improves the process of finding bands with important information by reducing the amount of original data to be analyzed.

The tool used in this for band reduction was 'Resize data' in ENVI® v 5.6, which allows making spatial and/or spectral (removing bands) subsets of original data set. Within the tool the hyperspectral image was chosen as the input file. The option of 'Spectral subset' was chosen which revealed all the bands 449 – representing different wavelengths - available in the hyperspectral image. Then the bands to be removed were deselected from the tool and a hyperspectral file with fewer bands was created.

The tool 'Band animation' was used to visually identify bands with low signal to noise ratio (bands to be removed). The bands that were removed from the data set were – Bands 1-13 (377-417 nm), 175-176 (972-978 nm), 235-265 (1310-1478 nm), 320-351 (1786-1953 nm), 440-449 (2449-2497 nm).

Figure 3-10 and Figure 3-11 below shows a scene from the hyperspectral data set visualised in the 'Band animation' tool at two different bands – 174 (969nm) (kept as part of data analysis), 175 (972nm) (removed from data analysis due to low signal to noise ratio).

Figure 3-10: Part of Study area visualised in band 174 (969nm) (Used in analysis - High signal to noise ratio)

Figure 3-11: part of Study area visualised in band 175 (972nm). Not used in analysis - Low signal to noise ratio)

3.5.3.2 Mosaicking

Mosaicking is a common process in remote sensing where two or more images are combined together to form a single image, where the borders or the seamlines of the two images are indistinguishable and the images maintain their georeferencing integrity (Li et al., 2019). Within the context of this study the limitations imposed by flight paths and the geographical distance between the scenes meant that the scenes of interest occurred across three different individual hyperspectral images. In order to analyze these scenes all together, they were mosaicked into one file.

The tool used was called 'Mosaicking' in ENVI® 'Classic' v5.6. Three hyperspectral images were mosaicked into one – their images represented the wild carrot hotspots under study.

The process of mosaicking can cause misalignments in the image which can cause compromised georeferencing (Li et al., 2022; Santellani et al., 2018). In order to offset mosaicking induced misalignment and georeferencing issues, a manual georectification process was performed where the ground truth data points were manually moved on the hyperspectral image based on their occurrence on the RGB orthoimages.

3.5.3.3 Minimum noise fraction transformation (MNF)

The acceptable form of input data for classification algorithms used in this study (MDC and MTMF) is MNF transformed data. The 'Forward MNF Estimate Noise Statistics' tool in ENVI® v 5.6 was used for MNF transformation of data being analysed (Figure 3-13). Within the tool the hyperspectral image was chosen as the input file. The option of 'Spectral subset' was chosen which revealed all the bands 449 – representing different wavelengths - available in the hyperspectral image. The number of bands for the output MNF file, which would capture most of the relevant information (Signal, minimal noise) was chosen as 10 based on experiences with similar data and similar studies reported in the literature (Hamada et al., 2007; Arslan, 2018). The tool will aggregate variation and relevant information from all the bands into these 10 bands.

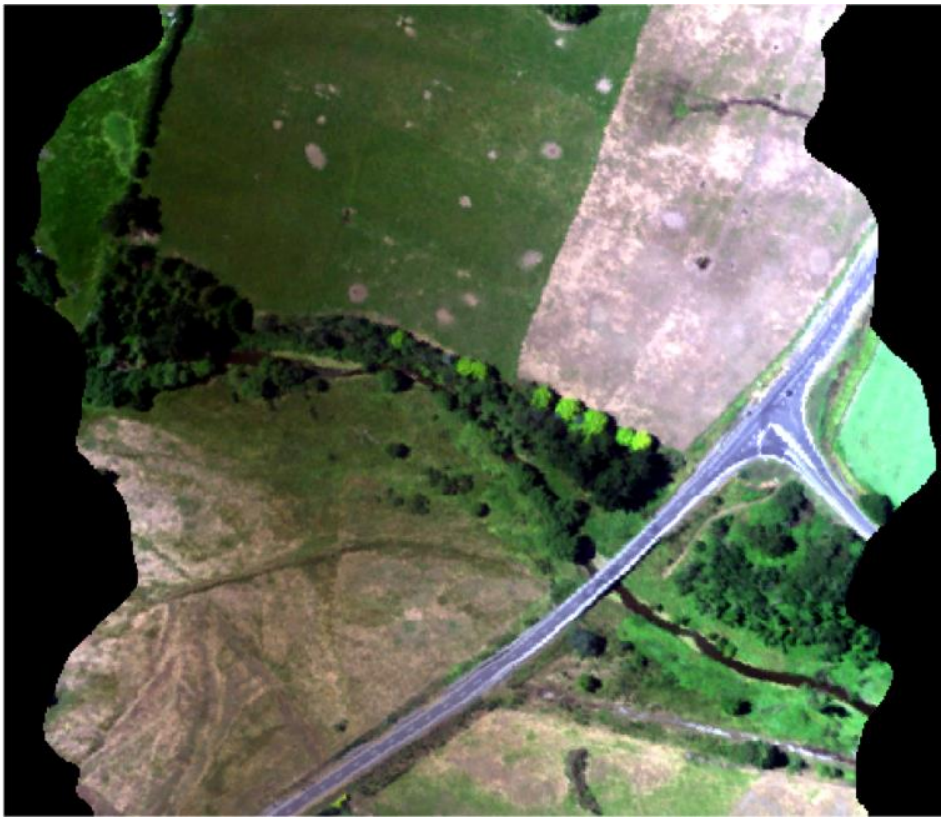

Figure 3-12: Hyperspectral data – RGB image from original data

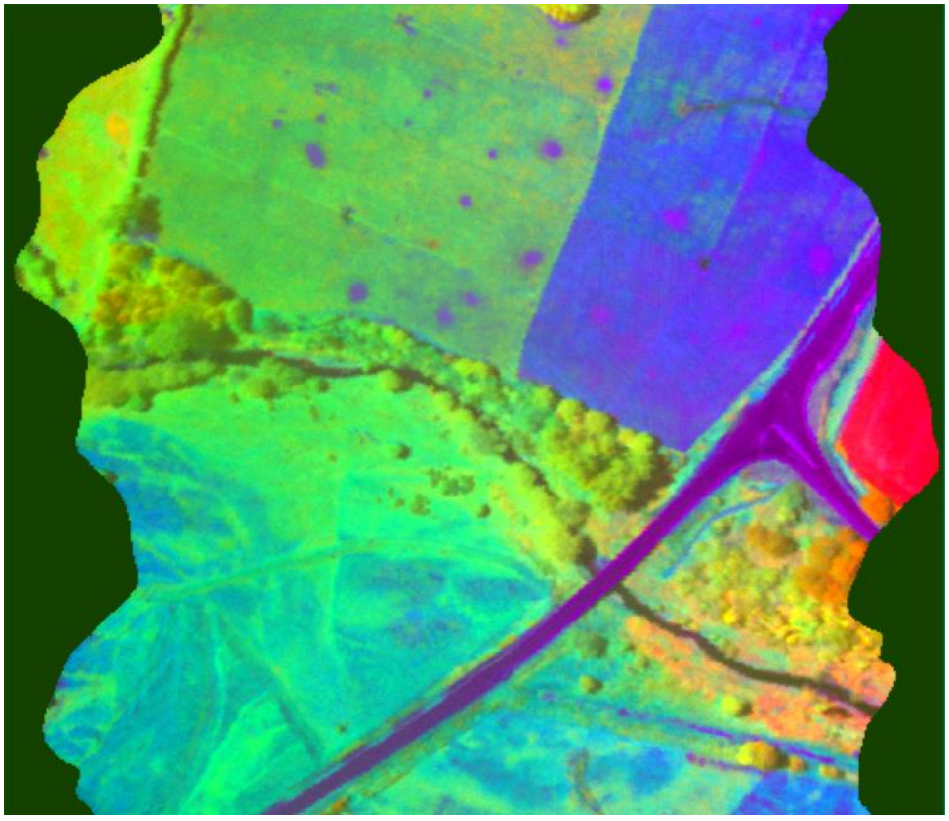

Figure 3-13: Hyperspectral data – post MNF transformation

3.5.4 Development of spectral libraries

A spectral library is a database of spectral reflectance data from different materials. For this study this is spectra from endmembers of wild carrot plants, dry vegetation, other dominant components of wild carrot patches.

The tool is called 'Spectral library builder' in ENVI® v 5.6. Here the spectral library of individual classes was formed using the pixels/samples with their highest % composition, which for wild carrot was around 25%. Regions of interest (ROI) containing the individual sample pixels were used as training data.

3.5.5 1st order classification – Minimum distance classification (MDC) - ENVI

Minimum distance classification (MDC) is a widely used classification algorithm for classifying land use surface cover. The MDC algorithm assigns individual pixels to the classes where the distance between the mean of the classes and the spectral response of the pixel is lowest (Gibson et al., 2000a) (Figure 3-14)

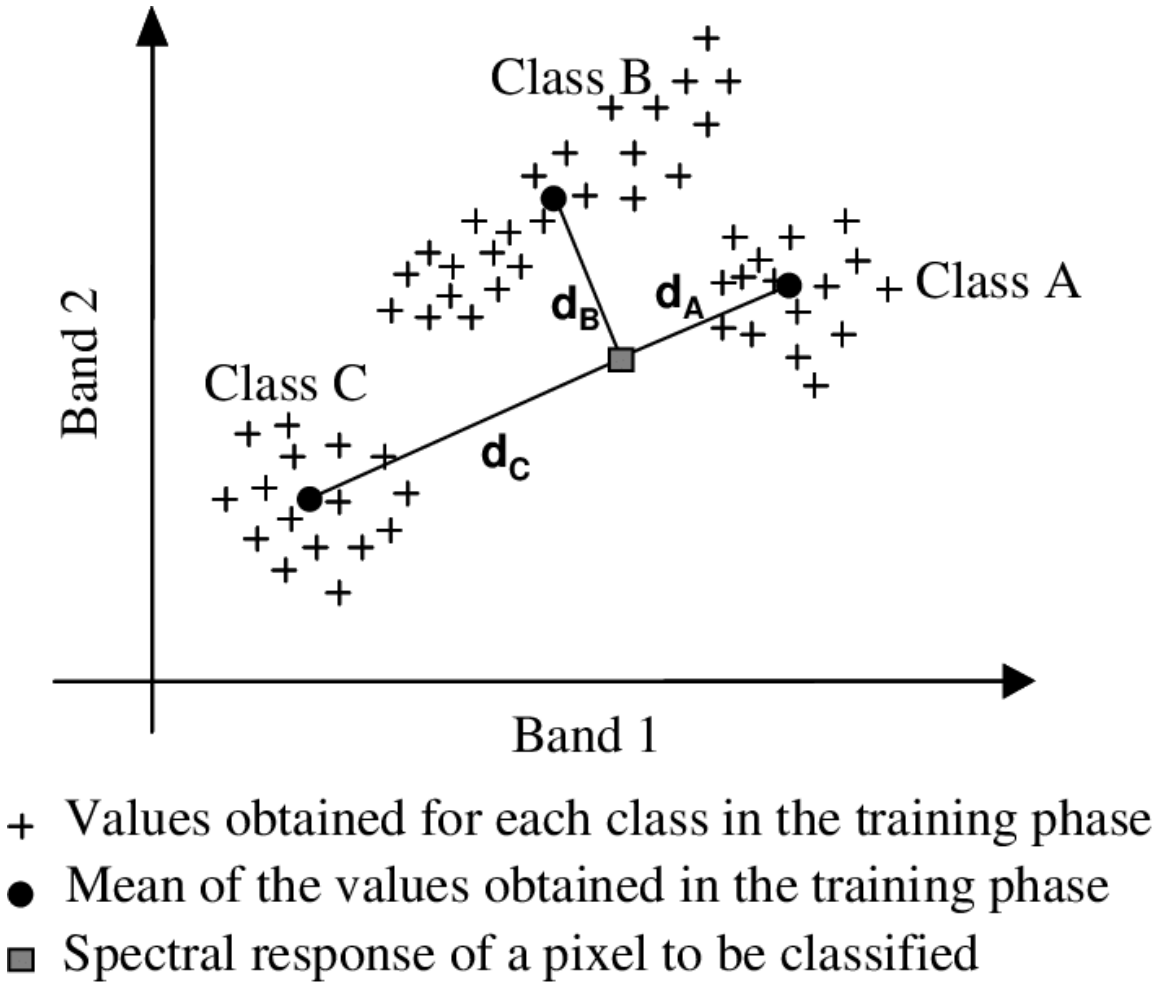

Figure 3-14: Minimum distance classification illustration (Gonçalves et al., 2009)

In the Figure 3-14 the square in the middle is the spectral response to be classified, the data used to train classes are represented by the cross (+) symbol and the black dots represent the mean of the training data. The pixel is assigned to the class, the mean of which is at the minimum distance to the pixel.

The classification was carried out in image analysis software ENVI®, version 5.6. For the analysis the following land use classes were defined – 1. pasture/vegetation, 2. bare soil, 3. road, 4. trees/forest, 5. shadow/background. The training data used were ROI (region of interest) polygons manually created that were best representative of the land use class. This was done by collecting polygons spread across the entire image. This ensures that the polygons capture both expected inter-class and intra-class variability (Figure 3-15). The polygons varied in shape/size, but each polygon included at least 5-10 pixels. At least 25 polygons per class were used as training data. The tool used in the software is labelled 'Minimum distance classification'. Within the tool the MNF transformed image was the input file. The training ROI polygons were used as training data for

each class. The resulting output is a classified image with different colours representing different classes and land surface cover types.

The analysis of the accuracy of the classification was done using a tool in the software labelled 'Confusion matrix using ground truth ROIs'. Within the tool the classified image is the input file. The validation ROI polygons were used as validation data. The resulting output is a confusion matrix giving the accuracy and error levels of the classification process. The validation data used was produced using the same method as the training data, ensuring that the validation and training polygons did not overlap by having the training data polygons as an active layer when creating validation data polygons.

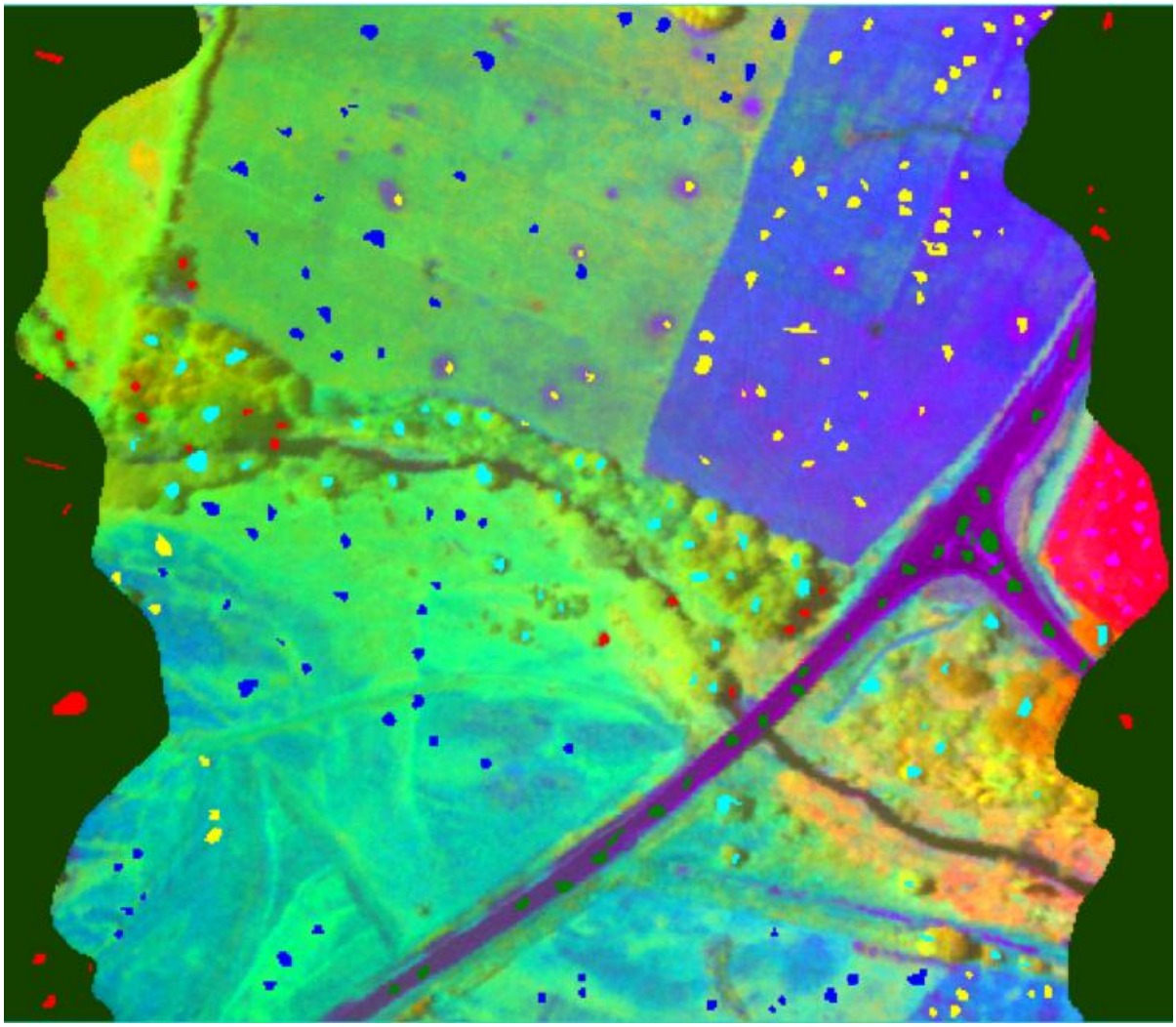

Figure 3-15: Hyperspectral data (MNF) image with ROI polygons training the algorithm in identifying individual classes/land use types.

3.5.6 2nd order classification – Spectral unmixing – Mixture tuned matched filtering (MTMF) - ENVI

The second order classification involved using a spectral unmixing analysis using a tool within the software called 'Mixture tuned matched filtering'. The input data used is the MNF processed data. The training data for the algorithm was the spectral library file containing endmember spectra of wild carrot plants. The endmember spectra in spectral library files were the spectra of samples comprising of the biggest pixel area coverage (% composition) of wild carrot. The spectral library for wild carrot included samples with wild carrot pixel coverage of between 12-25%. This included eight samples combined across the two wild carrot hotspots.

The output of the tool is a set of scores – MF and infeasibility scores. These scores give an indication of likelihood of presence of wild carrot on an individual pixel. Based on the scores of the biggest wild carrot samples (in terms of % composition), a threshold level of these scores were set and based on these threshold levels the likely locations of wild carrot were plotted. These locations were then overlayed on a more appropriate RGB base layer (RGB orthoimage) to produce weed maps predicting the likely locations of wild carrot plants.

3.5.7 ArcGIS - Mapping

The weed maps with the MTMF outputs overlayed, and which indicated potential wild carrot locations were developed using the ESRI GIS software ArcGIS pro v 2.5.0. The base layer of the map was the ortho RGB images captured during hyperspectral data collection.

Chapter 4 Results

4.1 Composition analysis

The composition analysis of the wild carrot ground truth samples was carried out by the point counting method. Composition analysis is a systematic way to elucidate the plant species/materials that cover surface area represented by a pixel. From the composition analysis results the patches were divided into 3 patch types based on the wild carrot %composition – small, medium and large patches (Table 4-1). Small patches constitute patches where the wild carrot coverage is below 5% of the total pixel area - this represents a surface area coverage that is below 0.05 m^2, medium patches that constitute 5-10% of the pixel (surface area coverage of 0.05-0.1 m^2), and large patches that constitute >10% of the pixel area (over 0.1 m^2 surface area coverage). Analysis showed that most of the wild carrot samples/patches belonged to the small patch group. With 47 of the 65 patches belonging to the small group, 13 patches belonging to the medium group and 5 belonging to the large group. Figure 4-1 shows a boxplot of distribution of wild carrot patch sizes in % pixel composition. The boxplot is skewed to the right with the median (3.1%) occurring towards the left in the inter-quartile range. This means that majority of the wild carrot patches belonged to the 'small' group and most patches were smaller than the mean patch size.

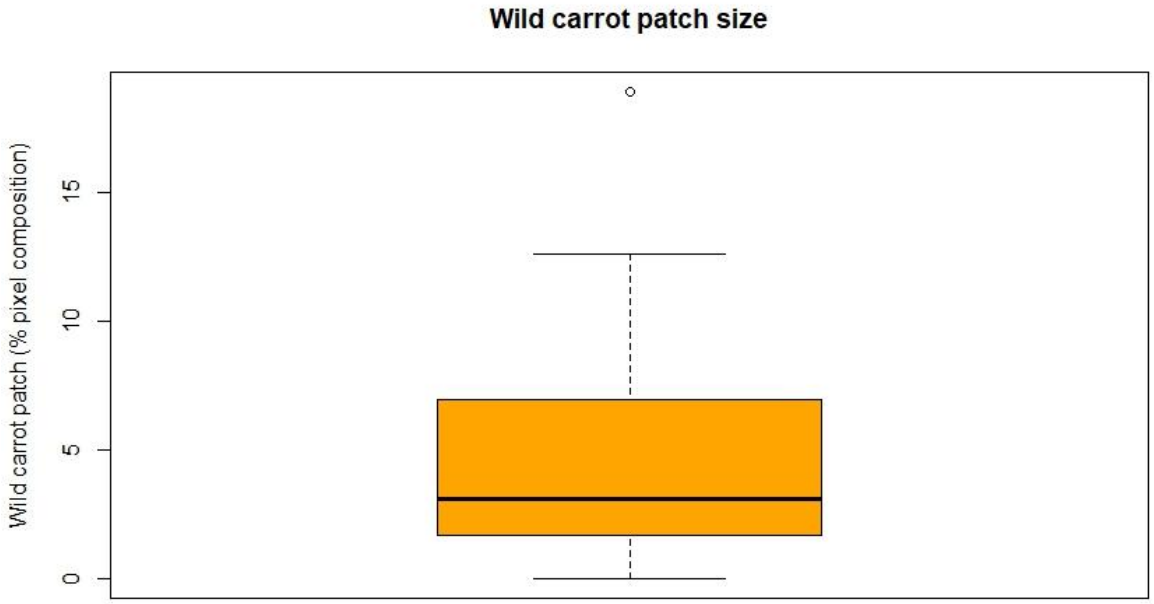

Figure 4-1: Boxplot of wild carrot patch sizes

The contribution from other vegetation was determined to provide the environmental context around the wild carrot patches – on average there was 41% grass coverage in small wild carrot patch pixels, 23% in the medium patch pixels, and 13% in the large patches. The Kruskal-Wallis test (one-way ANOVA on ranks) showed that this difference in grass coverage based on patch sizes is almost statistically significant (P-value: 0.050) (Figure 9-2).

For the small patches – the coverage of dry vegetation was around 32%, and around 40% for medium and large patches. However, the Kruskal-Wallis test (one-way ANOVA on ranks) (Figure 9-2) showed that this difference in dry vegetation coverage based on patch sizes was not statistically significant.

Soil coverage was similar for all sizes of wild carrot patches.

Table 4-1: Composition analysis results of wild carrot patches showing the wild carrot percentage and percentages of other composition classes within each pixel/patch.

		% Composition				
	# of Samples	Wild carrot	Grass	Dry vegetation	Other weeds	Soil/rocks
Small (<5)	47	<5	41.2	32.5	14.3	7.6
Medium (5-10)	13	5-10	23.5	39.6	18.5	6.8
Large (>10)	5	>10	13.1	39.4	20.3	8.8

4.2 Minimum distance classification (MDC)

A first order classification was carried out as a higher-level classification which was able to classify land surface cover types broadly. The analysis shown below was carried out on scenes depicting the two wild carrot hotspots in OngaOnga, Hawke's Bay.

Figure 4-3 and Figure 4-5 show the results from a minimum distance classification (MDC) analysis on a land surface cover map of the wild carrot hotspots. The different colours portray different classes that represent different land surface cover types. The classes include pasture, trees/forest, cropping, bare soil/tilled fields, roads and shadow/background.

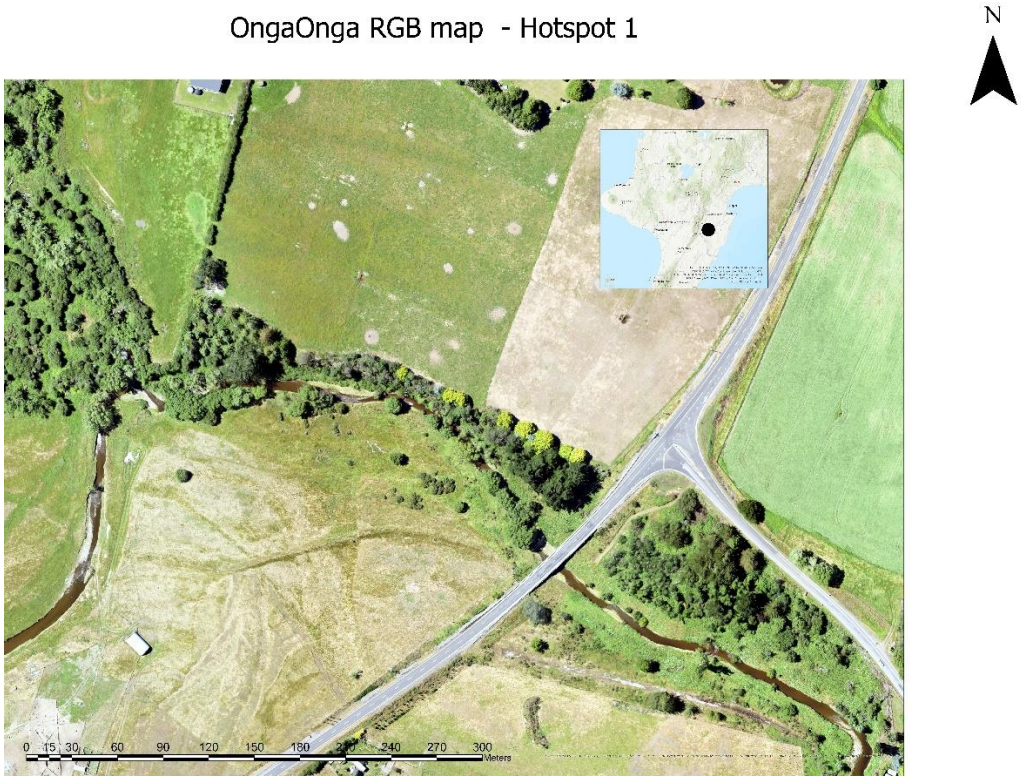

Figure 4-2: RGB map of wild carrot Hotspot 1 (Coordinates: -39.909, 176.413, intersection of State Highway 50 and OngaOnga Road)

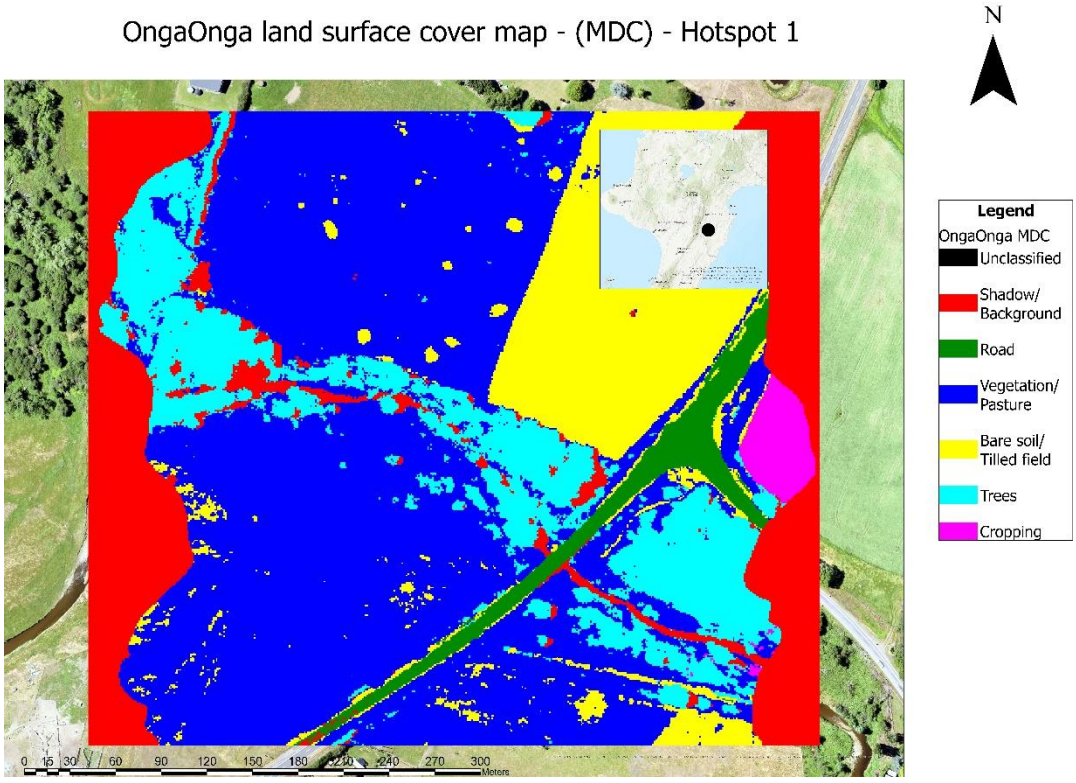

Figure 4-3: Minimum distance classification (MDC) land surface cover map of wild carrot Hotspot 1 (Coordinates: -39.909, 176.413, intersection of State Highway 50 and OngaOnga Road)

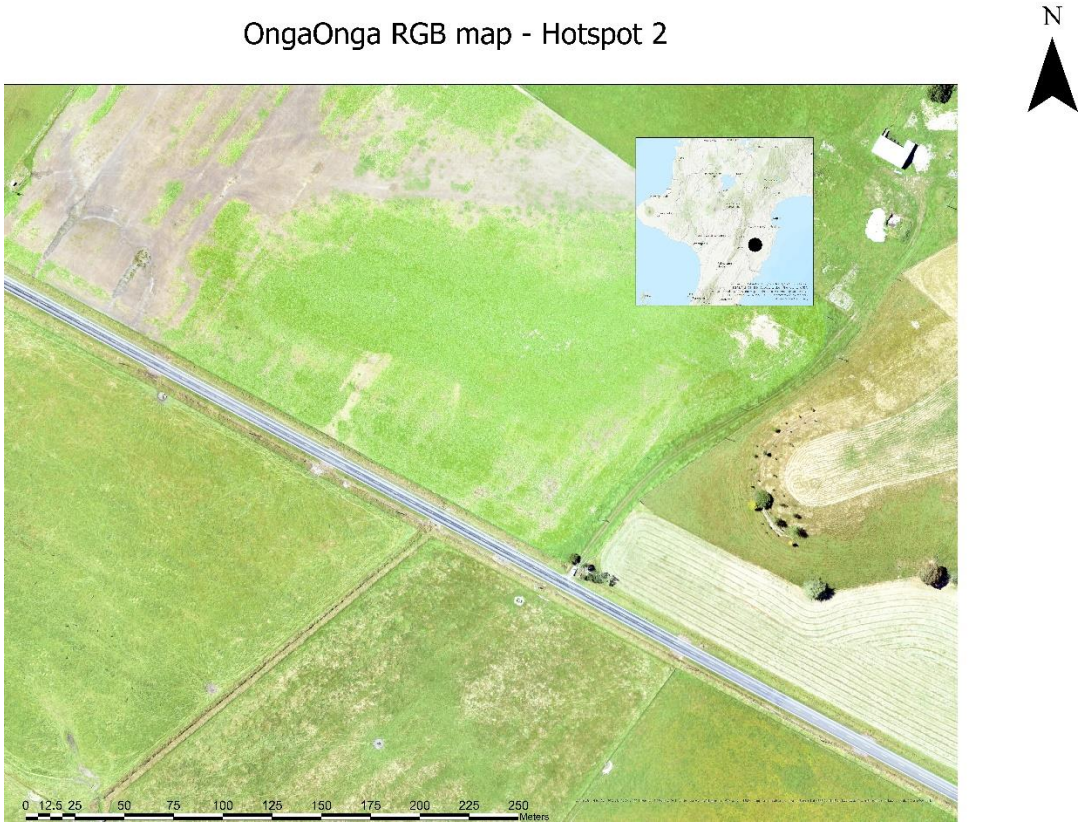

Figure 4-4: RGB map of wild carrot Hotspot 2 (Coordinates: -39.934, 176.517, on OngaOnga Road).

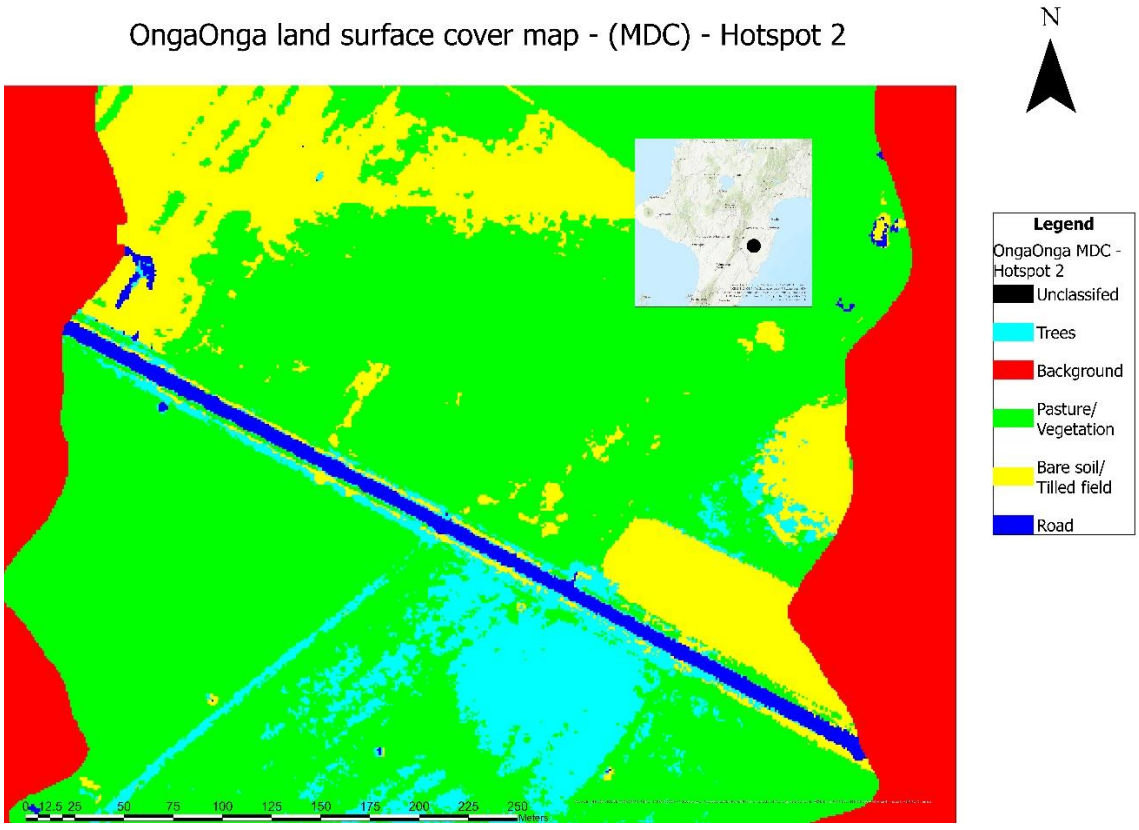

Figure 4-5: Minimum distance classification (MDC) land surface cover map of wild carrot Hotspot 2 (Coordinates: -39.934, 176.517, on OngaOnga Road).

A validation protocol was performed on the results from the minimum distance classification analysis (Figure 9-1). The results from the validation protocol showed that the algorithm was able to classify the land surface types to an overall accuracy of 96%. With a kappa coefficient of 0.95.

Note – The classifier has incorrectly classified part of a pasture paddock as 'Trees' in Figure 4-5. Showing the potential limitations of the classification method which must be taken into consideration in further analysis and interpretation of the classification results.

4.3 Mixture tuned matched filtering (MTMF)

MTMF was used as a 2nd order classification algorithm for the classification and identification of wild carrot plants in the areas of interest.

The output for MTMF classification is a pair of scores for each individual pixel – the MF and infeasibility scores. These both give an indication of the likelihood of presence of wild carrot in an individual pixel. An MF score of 1 indicates a perfect match and a low infeasibility score corresponds to a high likelihood of wild carrot occurrence (L3Harris Geospatial Solutions, 2023). Figure 4-6 shows a graph for MF versus infeasibility scores of the wild carrot patches/samples. The samples are assigned to their wild carrot patch groups.

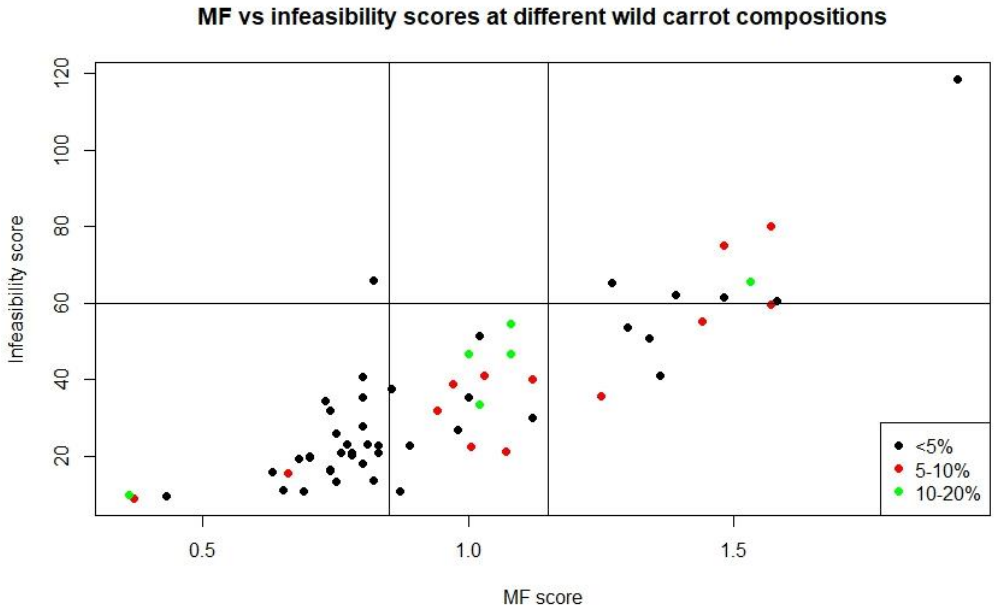

Figure 4-6: MF versus infeasibility scores of different wild carrot patches – based on their size - small (<5% - black circles), medium (5-10% - red circles) and large (10-20% - green circles).

A threshold level of MF and infeasibility was selected for the model – these were selected based on occurrence (based on their MF versus Infeasibility scores) of the biggest wild carrot patches (red and green dots in the box) in Figure 4-6. These threshold levels were used to generate likely locations of wild carrot plants within these thresholds. Threshold levels selected were: MF score – 0.9 - 1.1. Infeasibility score 60.

Figure 4-7 shows the model prediction of wild carrot based on the input threshold MF of 0.9 to 1.1 and infeasibility score of 60. The inset graph is a scatter plot of MF versus infeasibility scores of all the pixels on the MTMF analysed scenes. The red area highlighted in the inset graph represents all the pixels that fall into the above specified MF (0.9-1.1) and infeasibility score (60) threshold levels. The pixels highlighted in red in the map represent the pixels the model, within the defined threshold limits, predicts as having potential wild carrot plants.

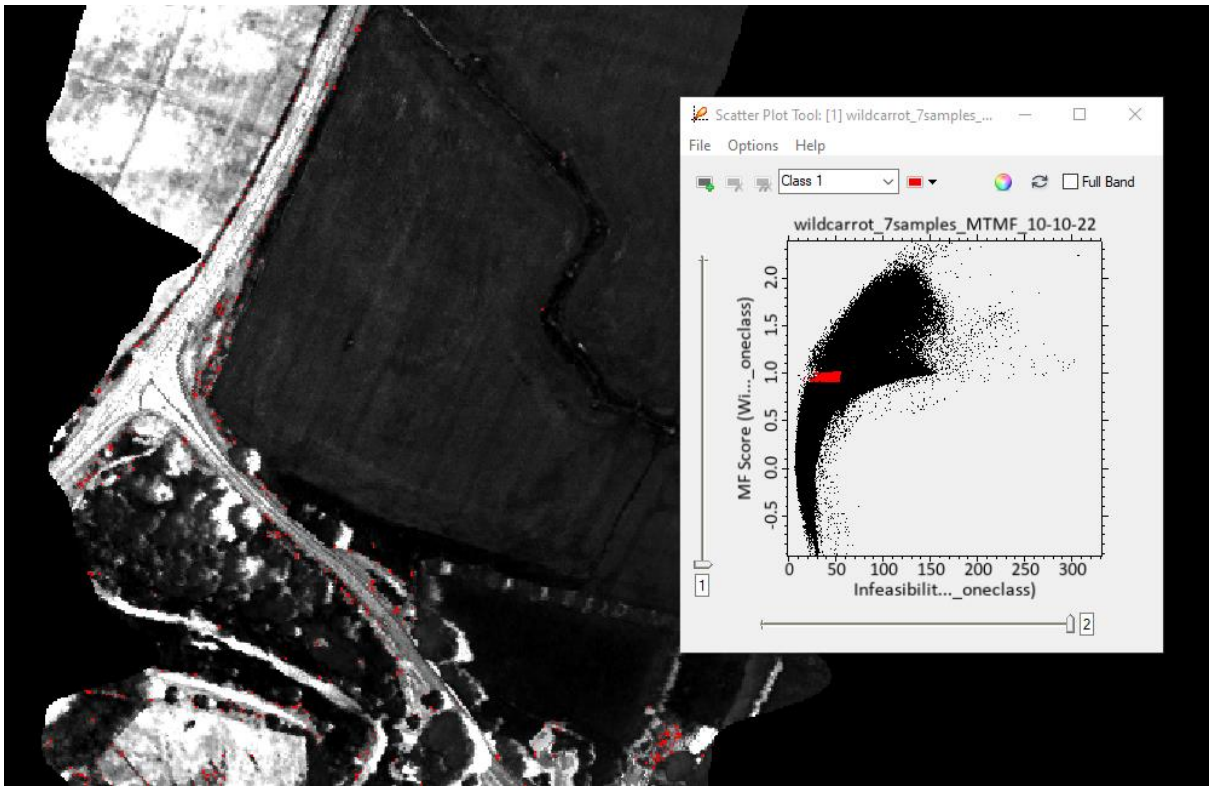

Figure 4-7: MTMF model - wild carrot plant predictions

Omission rates were calculated by using some of the medium and large patches as validation data. The patches that occurred – within a distance of one pixel to the model prediction– were considered as predicted by the model. The model was able to predict 7 out of the 10 samples (validation data) ('medium' and 'large' patches that were not used as training data) with the MF

range 0.9 – 1.1 and an infeasiblity score of 60 or below. This is an omission error rate of 30% (Producer's accuracy: 70%, where Producer's accuracy = 1 – omission error). The MTMF analysis was carried out on the two hotspots combined by Mosaicking the two scenes together as described in 3.5.3.2 Mosaicking.

4.4 Weed maps

The above MTMF model wild carrot plant predictions were overlayed onto an RGB base layer of the two OngaOnga wild carrot hotspots and potential weed maps with possible locations of the wild carrot plants were detected. Figure 4-8 and Figure 4-9 are the weed maps with the predicted wild carrot locations.

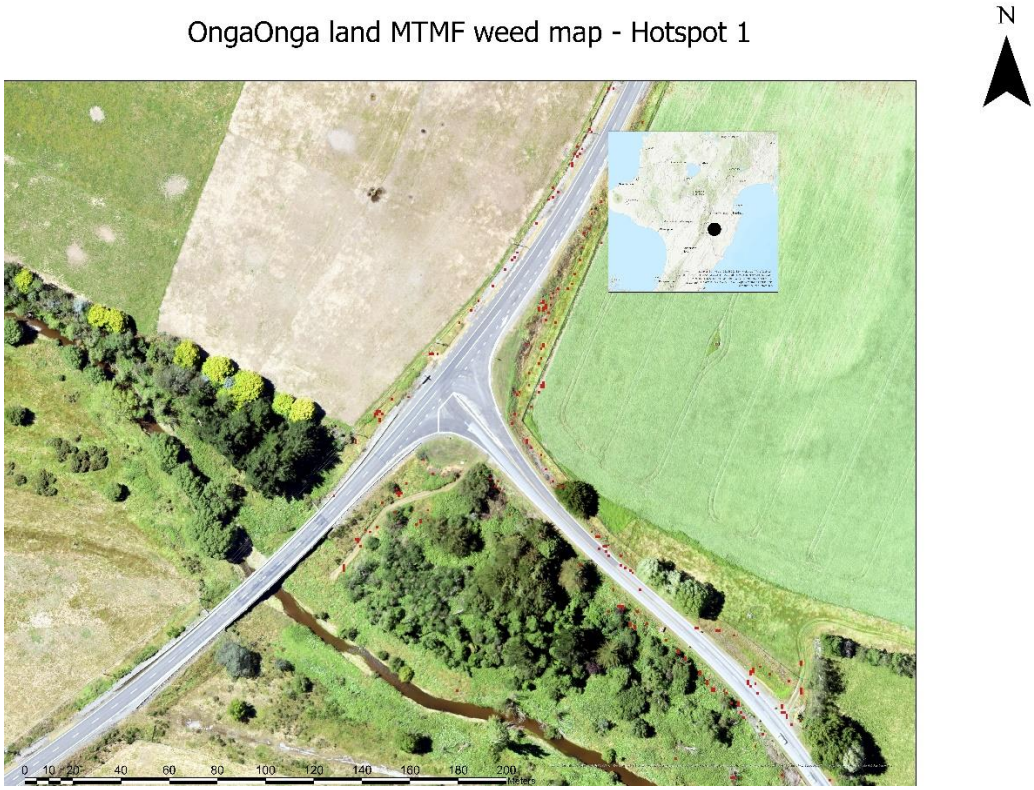

Figure 4-8: OngaOnga land MTMF weed map - MTMF model predicted potential wild carrot locations (red points), Hotspot 1 (Coordinates: -39.909, 176.413, intersection of State Highway 50 and OngaOnga Road)

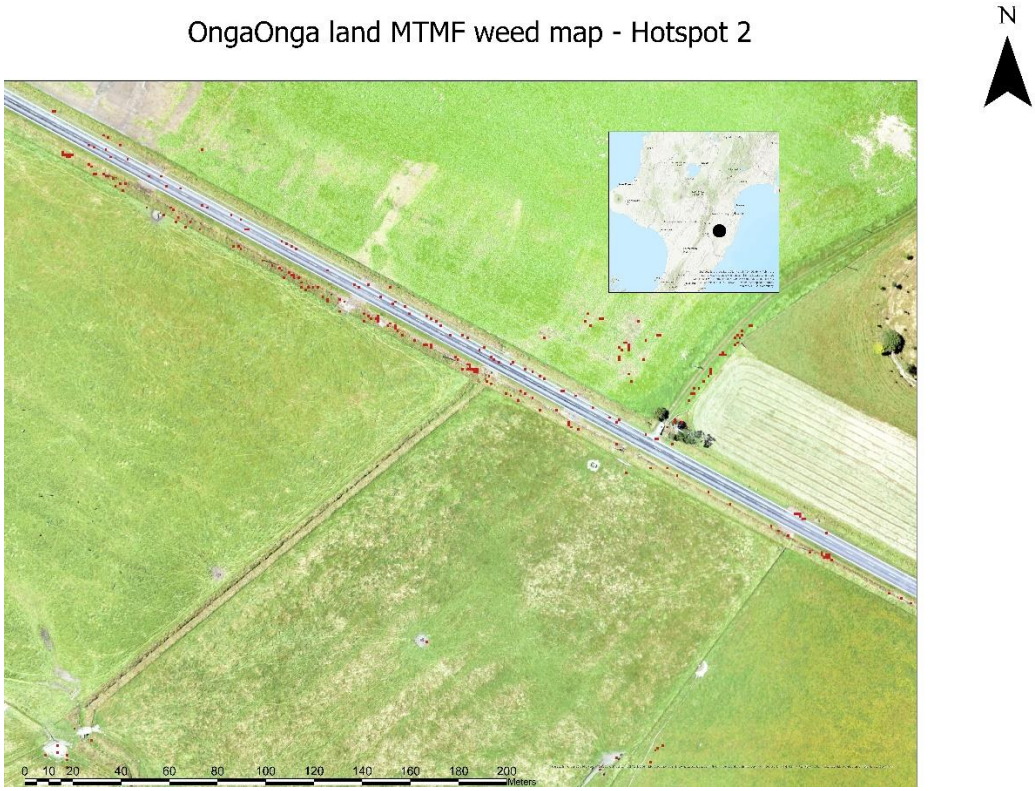

Figure 4-9: OngaOnga land MTMF weed map - MTMF model predicted potential wild carrot locations (red points), Hotspot 2(Coordinates: -39.934, 176.517, on OngaOnga Road).

Figure 4-8 and Figure 4-9 show potential wild carrot locations as predicted by the model generated. Predicted location points that appear on roads are obvious false positives. These are likely either predictions that have escaped being filtered through the classification process or a potential slight misalignment between the classification image (processed hyperspectral image) and the base layer – which in this case is the RGB image. This is likely the case especially in the hotspot 2 image where predictions which should appear on the side of the road appear on the road.

4.5 Environmental context

Composition analysis data was used to obtain a perspective on the environmental context of the wild carrot patches/samples in terms of the plant species and materials that formed the dominant features of the environment wild carrots are found in for the two wild carrot hotspots in the OngaOnga study area. The patches from both wild carrot hotspots were combined and analysed together with a broad view to understand the environmental context of wild carrot plants in the OngaOnga study area, which also allows use of the maximum number of samples possible for the analysis. The Figure 4-10; Figure 4-11, and Figure 4-12 show the environmental context of the wild carrot patches in terms of the soil/rocks, grass, and dry vegetation composition of the wild carrot patches.

Figure 4-10 shows the MF versus infeasibility scores of the wild carrot patches/samples and the different % soil/rock compositions of these patches.

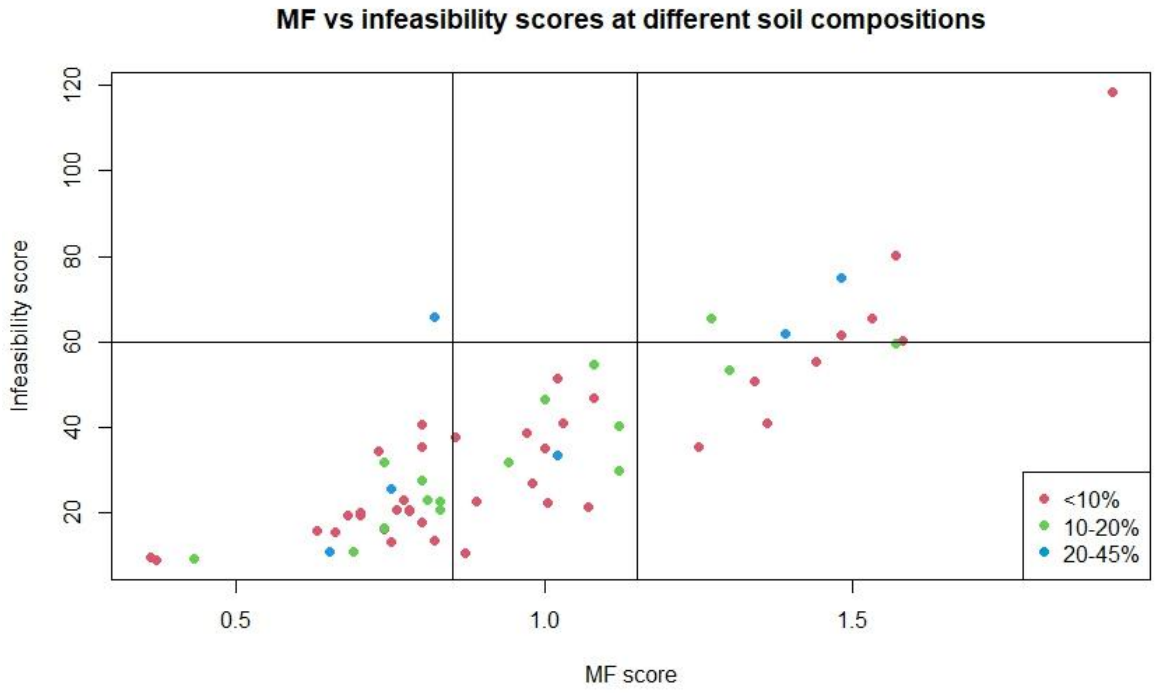

Figure 4-10: Environmental context – soil/rocks coverage - of wild carrot patches - red – small (<10%), green - medium (10-20%), blue- large (20-45%).

From Figure 4-10, there is no obvious pattern of soil coverage groups in the region indicating high likelihood of wild carrot patch occurrence (MF score – around 1, infeasibility score below 60). This indicates a lack of relationship between soil coverage and wild carrot patch occurrence.

The Figure 4-11 shows the MF versus Infeasibility scores of the wild carrot patches/samples and the different % grass compositions of these patches.

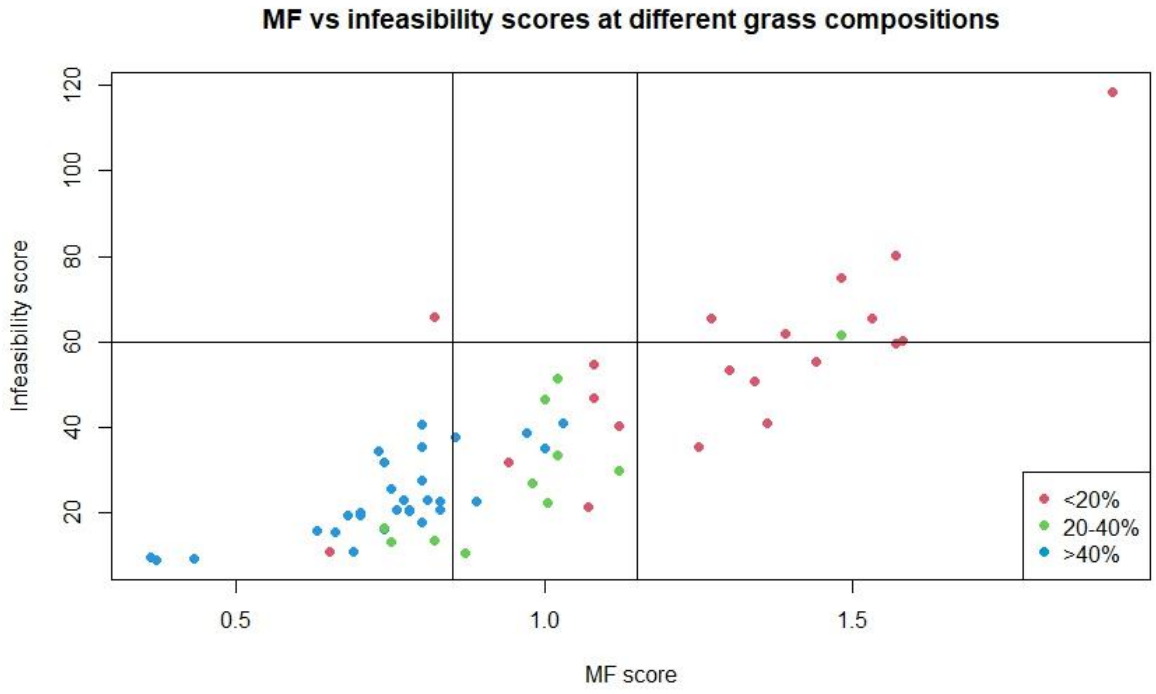

Figure 4-11: Environmental context – grass coverage - of wild carrot patches - red – small (<20%), green - medium (20-40%), blue - large (>40%).

Figure 4-11 shows that extremes of grass coverage groups (large and small) occur away from the region indicating high likelihood of wild carrot patch occurrence (in the area where MF score – ~0.9-1.1, infeasibility score below 60). However, most of the patches in the region still have small and medium coverage of grass, likely supporting the earlier observation of lower grass coverage in and around bigger wild carrot patches (see 4.1 Composition analysis).

The Figure 4-12 shows the MF versus infeasibility scores of the wild carrot patches/samples and the different % dry vegetation compositions of these patches.

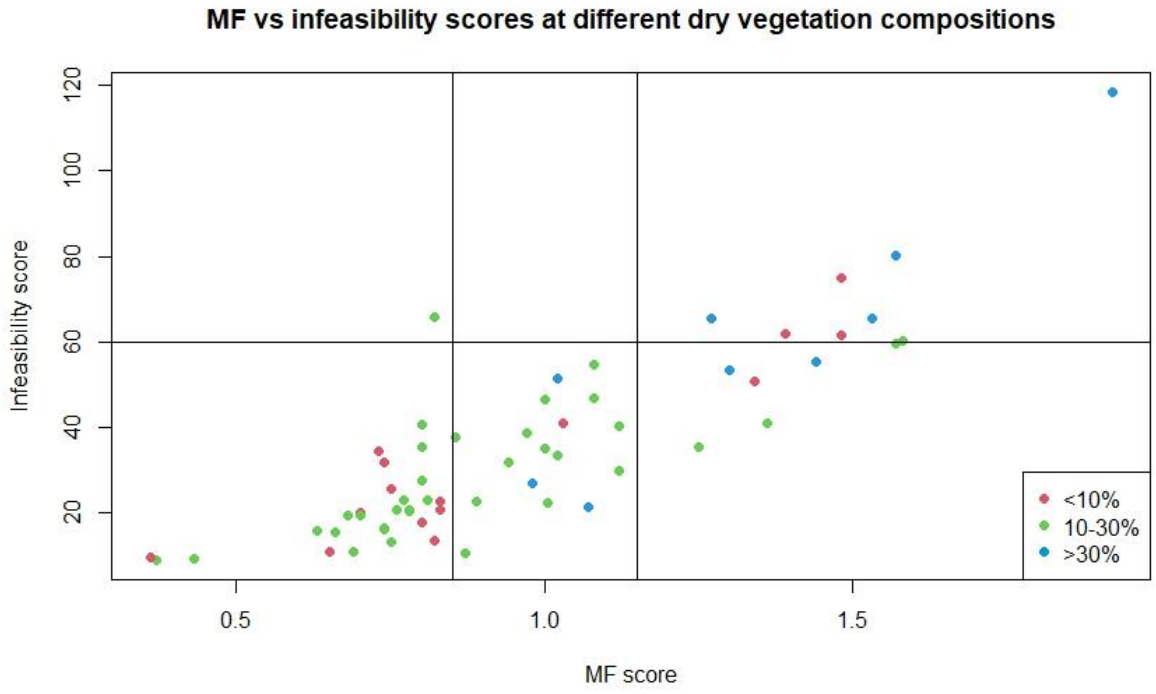

Figure 4-12: Environmental context - dry vegetation coverage - of wild carrot patches - Red – small (<10%), green - medium (10-30%), blue - large (>30%).

The occurrence of mainly medium and large group coverage of dry vegetation in the region indicating high likelihood of wild carrot patch (in the area where MF score – ~0.9-1.1, infeasibility score below 60) occurrence supports the earlier suggestion of wild carrot occurrence around dry vegetation (see 4.1 Composition analysis).

Chapter 5 Discussion

5.1 Composition analysis

Composition analysis and data from (Table 4-1) revealed that most (72%) of the sampled wild carrot patches were small - in terms of surface area coverage that is less than 0.05m^2 (5%) of a total 1 m^2 surface area on a pixel. 13 patches were in the medium group – with surface area coverage of 0.05-0.1 m^2. Only 5 were large patches with a surface area coverage of over 0.1 m^2. Considering the detection levels of vegetation in previous similar studies, it is possible that most of the samples found would not be detectable within the current configuration (spatial resolution – 1m^2) of hyperspectral remote sensing (small and most likely medium patches). However, presence of many small patches around few large patches suggests that a collection of patches tend to occur together in a clustered pattern across the same broad areas. Thus, in the current context of wild carrot control in carrot seed production, application of hyperspectral remote sensing to detect large wild carrot patches (with small patches around them in the same area) would likely still be useful and help mitigate the risk they pose.

Seed in the soil seed bank are the main source of new annual weeds (Dessaint et al., 1991). Carrot seed have been shown to be long lived in artificial storage (Solberg et al., 2020). Moreover a characteristic of weed species is persistence of their seed in the soil's seed bank (Dessaint et al., 1991; Thompson et al., 1993), especially in annual weeds (Ghersa & Martinez-Ghersa, 2000). This might indicate that wild carrot seed will also persist for a long period in the soil seed bank and give rise to new individuals over time, however this is speculative as there are no reports that confirm this for wild carrots specifically.

Further, localised redistribution of weed seeds in the soil means that seed tend to occur in clusters in soil (Stafford & Miller, 1996). This clustered nature of seed in the soil seed bank likely explains clustered occurrence of wild carrot patches in the wild carrot hotspots. In this context, this clustered occurrence of wild carrot can be leveraged for wild carrot control by detecting the large patches, which would likely have smaller patches in the same area. A limitation of this may be where wild carrot is newly invasive to an area and therefore small patches may predominate.

For wild carrot, hotspots were found to be areas that are not under active cultivation, but are still disturbed areas. In disturbed areas ecological succession has not caused native vegetation/ perennial weeds to become dominant (Zanin et al., 1997). This may explain the disproportionate

occurrence of wild carrot along roadsides (which are regularly disturbed for maintenance – spraying and mowing) and likely have a seedbank which replenishes the annual wild carrot plants in the absence of competition from other plants.

5.2 Ecological context

The ecological context analysis of the patches showed that the large patches had on an average 13.1% grass coverage. The medium patches show 23.5% grass coverage and the small patches had on average 41.2% grass coverage. This suggests that the micro areas dominated by grass had lower abundance of wild carrot and occurrence of large wild carrot patches corresponded to less grass coverage. It is possible that areas that promote grass growth are areas with lower abundance of wild carrot populations as the grass outcompetes the annual wild carrot populations. This may be because established perennial grasses tend to have a competitive advantage over annual weeds that are looking to get established. Perennials have established their root systems and therefore have a greater access to the resources available relative to annual weeds that need to establish annually (Borman et al., 1991; Torner et al., 2000).

Further the compositional analysis also showed that that the small wild carrot patches on an average have less dry vegetation in their environmental context with around 32% coverage relative to medium and large patches which have around on average 39.4% dry vegetation coverage. This suggests a possible positive relationship between presence of dry vegetation and large wild carrot patch occurrence.

Ryegrass species were among the dominant grass species around the wild carrot plants. This is not surprising given that ryegrass in the predominant pastoral grass species in Hawke's Bay (Charlton & Stewart, 1999). Ryegrass species are not known for their high drought tolerance and need well-watered moist soils to grow optimally and persist (He et al., 2017). This suggests that an area that does not provide ideal conditions for grass growth, likely due to dry conditions, is more likely to have wild carrot presence as opposed to areas that are more suitable for growth of vegetation like grass (AgResearch, 2023). Presence and abundance of specific plant species in microclimates based on their resource capture capability, and the spatial and temporal heterogeneity in availability of resources is a common ecological process (Avolio et al., 2019; Storkey & Neve, 2018). In the context of wild carrot in the Hawke's Bay, there would be greater likelihood of wild carrot occurrence during the drier periods of the year, especially in microclimates that do not

support vegetation like grass – e.g. roadsides that are disturbed regularly due to maintenance via mowing or spraying. This all suggests a possible negative relationship between soil moisture and wild carrot occurrence, that could be explored in the future using hyperspectral imaging.

The cultivated carrot was domesticated from the wild carrot which originates from the central Asia region (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan) (Iorizzo et al., 2013) where the climate is arid to semi-arid – thus carrots originated from an region low in moisture (Lioubimtseva & Henebry, 2009). This supports the argument that wild carrot plants can thrive in moisture deprived environments. This likely also explains the abundance of wild carrot in the carrot seed producing areas of Canterbury and Hawke's Bay. Both these east coast regions are characterised by having similar climatic conditions (Hampton et al., 2012) and being some of the driest regions in New Zealand (NIWA, 2022).

The propensity of wild carrot to emerge in dry conditions and not within already established vegetation suggests that there will be higher open soil coverage around wild carrot plants. There has been a lower reported emergence of annual weeds around vegetation residue due to their release of allelopathic compounds (Torner et al., 2000). This further suggests areas with less vegetation and vegetation residue coverage and more open soil would have a greater likelihood of wild carrot occurrences. However, the composition analysis results showed no correlation between wild carrot patch size occurrence and open soil coverage around the patches.

There was a relatively high presence of wild carrot plants in areas along roadsides as opposed to areas further away from roads. Wild carrot absence away from roads and close to cultivation areas could also be a result of standard weed control practices in and around a farming environment – especially in a seed production environment where for carrot seed production wild carrot would be removed (including rouging) as part of the management of the seed crop. Another possibility could be because areas just off the roads undergo more disturbance in terms of weed management activities like mowing and spraying, as opposed to areas away from roads. Annual weeds like wild carrot would establish here because of reduced perennial grass competition and the inherent nature of annual weeds where they allocate a disproportionate amount of their resources on reproduction allowing them to colonise disturbed areas quickly (Koziol et al., 2012; Van Kleunen et al., 2010).

In the Hawke's Bay another possible reason of presence of wild carrot on roadsides is due to shingle (small rocks) transported from along river valleys to build/repair roads (Richard van Garderen, South Pacific Seeds (NZ) Ltd, personal communication, 2023). It is possible this transport of shingle causes transfer of seed from seed banks of wild carrot in these river valleys, leading to greater manifestation of wild carrot along roadsides.

5.3 Minimum distance classification

A minimum distance classification was carried out as a first order classification. Here the objective was to classify the data into distinct land surface cover types. The classification was successful with an overall accuracy of 96% and a kappa coefficient of 0.95.

Overall accuracy refers to the ratio of total number of correctly classified pixels in all classes to the total number of reference pixels from validation data. The kappa coefficient is an accuracy indicator that measures how the results from the classification compare with values assigned by chance. The coefficient has a range of 0 to 1, a score of 0 suggests the ground truth and classified image have no agreement. While a score of 1 suggests there is total agreement between the ground reference data and the classified image (Ukrainski, 2019).

Use of supervised classification algorithms to create land surface cover models has been widely reported (Mirik et al., 2013; Mohd Hasmadi et al., 2009; Petropoulos et al., 2013; Roth et al., 2015). This can allow elimination of areas that are unlikely to have wild carrot plants. This includes areas such as roads, forests, buildings, and water bodies from the final classification analysis.

A successful land surface cover classification allowed the application of a technique called 'Image masking'. Image masking can be used to apply classification methods selectively to parts of data/image.

Use of image masking to selectively apply classification processes has been used to assess specific contaminated areas in a wetland ecosystem (Salem et al., 2005), for masking water bodies and clouds for land surface cover mapping in a conservation/land management context (Clark & Kilham, 2016), for masking non-vegetative pixels in context of mapping plant species in a forest ecosystem (Dadon et al., 2019). For the study the MDC classification allowed the ability to mask land surface cover types that would likely not have wild carrot plants. This includes roads, buildings, forest canopies/trees and water bodies. However, this is not appropriate in case of inaccurate classification - as in 'Hotspot 2' where part of paddock was incorrectly classified as

having trees - meaning the layer of 'trees' could not be used in the masking process (see Figure 4-4 and Figure 4-5).

Masking in the context of detection of wild carrot would help reduce potential false positives and would save resources analysing redundant data, increasing the commercial viability of hyperspectral remote sensing for identification of wild carrot contaminants, and potentially other weed contaminants, located in or near seed production sites. Reduced number of false positives in this context would also allow for mitigating resource expenditure used in scouting for plants in areas where they cannot exist.

5.4 Mixture tuned matched filtering

Mixture tuned matched filtering (MTMF) has been shown to be an effective classification algorithm in mapping of vegetation (Andrew & Ustin, 2008; Hamada et al., 2007) and weeds in general (Glenn et al., 2005; Mundt et al., 2005; Noujdina & Ustin, 2008), especially when in cases of subpixel classification where the weeds to be detected are smaller than the spatial resolution of the detection system as has been identified in this study (Williams & Hunt, 2002).

The MTMF analysis results show that the model was capable of detecting wild carrot plant patches that covered over 5% of the pixel surface area ($0.05m^2$) with an omission error rate of 30% (counting wild carrot samples in the 'Medium' and 'Large' groups only ie >5% surface area coverage). A detection threshold of 5% is within the range obtained in other studies for similar weedy plants, the omission error rate at over 30% is higher than other similar studies where the error rate was below 20% (Mundt et al., 2005; Noujdina & Ustin, 2008; Williams & Hunt, 2002; Yang & Everitt, 2010). This higher error rate means that a proportion of wild carrot would not be detected using this specific detection system where the spatial resolution is $1m^2$. It is uncertain how much pollination risk these undetected plants would pose but considering very low levels of contamination (0.02%) are acceptable in hybrid carrot seed production, the risk may be unacceptable. This is therefore a risk that requires further research. This study has demonstrated that this set up of hyperspectral imaging (type of sensor, aircraft carried, altitude of flight) can identify wild carrot patches – however, with limitations around patch size - justifying further research to reduce the spatial resolution for accurate detection.

5.5 Improvement of detection levels

As discussed, it is desirable to improve on detection thresholds for wild carrot plants using hyperspectral remote sensing as this would allow detection of more plants which are smaller in size. This will lead to creation of more accurate weed maps which enable greater mitigation of pollination risk posed by wild carrot plants and ultimately would reduce costs and allow production of carrot seed crop with required genetic purity levels.

5.5.1 Training data

The training data for the model included endmembers of wild carrot. However, the endmembers were not completely pure wild carrot pixels. The biggest wild carrot sample used for training data had a pixel coverage of around 25%.

In the detection of Hoary cress Mundt et al. (2005) reached user's accuracy of over 90% using MTMF algorithm, however their training data endmembers had high % pixel coverage (pure endmember pixels). MTMF was able to detect leafy splurge at a detection threshold of around 5%, however the endmember training data was high quality with 100% coverage 'pure' leafy spurge pixels (Williams & Hunt, 2002), as opposed to endmember training data in this study where the biggest training data sample had a 25% wild carrot coverage. This suggests that it is possible that the efficacy of the model in terms of its detection threshold limits and its errors of omission within those limits could be improved with training data of wild carrot plants that have a greater pixel surface area coverage.

5.5.2 Spatial resolution vs spectral resolution

A way of improving detection capacity would be to acquire data with higher spatial resolution. Acquiring data using an unmanned aerial vehicle (UAV), which can fly much closer to the ground, can help acquire data of greater spatial resolution and thus can help improve detection limits and accuracy by reducing the levels of spectral mixing (Näsi et al., 2018). However, there are trade-offs. Use of aircraft hyperspectral data can help acquire data over a much larger surface area per unit time relative to UAV data. This will likely be an issue in the current study's context of carrot seed production where due to the distance of pollen travel, data from a large area needs to be acquired in a relatively short period of time.

Further, sensors mounted on UAVs often have poor spectral resolution relative to aircraft mounted sensors (Mozgeris et al., 2018). This can cause a loss of spectral features important in detection of specific vegetation.

Flying the aircraft at a lower altitude can be another way of acquiring data with higher spatial resolution however there can be trade-offs in terms of sensor limitations. Using lower flying altitudes to acquire data with greater spatial resolution– outside the range of the sensor – can cause saturation of the sensor as it not capable of dealing with the increased energy interception and results in poor quality data (Dao et al., 2019; Markelin et al., 2008; Mozgeris et al., 2018).

Figure 5-1 shows a study comparing the UAVs and aircraft (ULA, Cessna) collecting hyperspectral data over a 4000ha test area. The figure shows that greater spatial resolution data can be acquired by flying at a lower altitude across all platforms. Secondly even though greater spatial resolution can be acquired in UAV platforms relative to aircraft mounted, the excess resources (time and cost) associated with UAV data collection might mean that the exercise ceases to be economically and agronomically justifiable - which might be the case in carrot seed production.

For each platform the four points represent 4 different altitudes of data collection, with the point furthest on the left being lowest altitude flight.

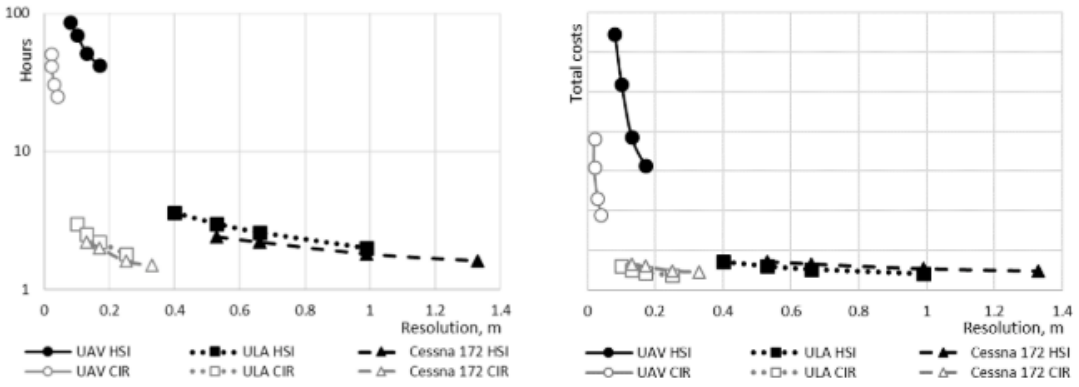

Figure 5-1: Costs and time resources associated with hyperspectral data collection with UAVs and aircrafts (Mozgeris et al., 2018)

5.5.3 Image fusion techniques

Another way to improve detection limits would be to use image fusion techniques. This technique usually involves use of two or more different sensors (Simone, 2002). This method of data

acquisition results in obtaining data that has both high spatial and spectral resolution. A specific image fusion technique called 'fusion based HS image super-resolution' is a technique which can allow acquisition of data that has both high spatial and spectral resolution (Vella et al., 2021). This usually involves fusion of hyperspectral data with high spectral resolution and RGB data with high spatial resolution (Vella et al., 2021). This technique can help improve spatial resolution of the hyperspectral data and can help improve detection limits in vegetation identification/detection work. Improved detection levels in this instance would once again allow creation of more accurate weed maps with potentially fewer or no "missed" plants and help mitigate pollination risk posed by wild carrot plants, leading to production of hybrid carrot seed crop of acceptable genetic purity levels.

5.5.4 Detection using environmental context

This study has given clues regarding possible relationships between wild carrot presence and environmental context related parameters like moisture content, dry vegetation presence and species composition in terms of type of plant species and their respective coverage in the wild carrot growing environment. Use of hyperspectral imagery to study environmental gradients has been reported in literature, including in an agricultural context for species detection (Andrew & Ustin, 2009; Wang et al., 2014). Future research may focus at leveraging hyperspectral data to further study environmental context/gradient related parameters like moisture content and species composition/coverage of wild carrot growing environment with a view to deduce a relationship between them and wild carrot occurrence.

This can help develop prediction models and allow detection of wild carrot plants below the threshold surface area coverage level required in the MTMF model from this study. Mapping of individual species by leveraging parameters around their environmental context/gradients using optical remote sensing, including hyperspectral remote sensing has been widely reported in the literature (Andrew & Ustin, 2009; Artigas & Yang, 2007; Landmann et al., 2018)

5.6 Model efficacy in practise

Most of the wild carrot patches sampled (47 out of 65) belonged to the 'Small' group (<5% pixel-surface area coverage). Considering the 5% pixel detection threshold level of the model, the model would be unlikely to be able to detect and identify these patches. However as mentioned earlier,

the small patches were observed to be around other patches, including bigger patches within the detection limits of the model. This suggests that in the current context of carrot seed production, a detection threshold of 5% pixel coverage may still provide value in terms of weed identification and thus its management by enabling effective scouting and hand rouging.

Chapter 6 Conclusions

The study set out to answer the following research question: *Is Hyperspectral remote sensing a feasible method for early detection of Wild carrot in Carrot seed production?*

To achieve this, four broad aims were defined within the context and limitations of the study. These included:

1. Identify areas with wild carrot populations in Hawke's Bay.
2. Develop a database of the environmental context of the wild carrot populations.
3. Assess the utility of MTMF as a classification algorithm to detect and classify wild carrot plants.
4. Assess possibility of developing weed maps with accurate potential wild carrot GPS locations.

Ground-truth data was used to undertake a compositional analysis of wild carrot patches and gave an insight into the growing environment of the wild carrot. It revealed that most of the wild carrot occurred in disturbed areas, and predominantly along roadsides. The data collection suggested that most wild carrot patches occur in a cluster – likely due to clustered seed in soil seed banks. Transport of seed from riverbeds in shingle used for roads may be a source of wild carrot seed, which is also consistent with wild carrot being found along the roadside. Large wild carrot patches occurred in areas where grasses were not the dominant species. This may be due to annual weeds like wild carrot having an inherent disadvantage when trying to establish amongst already established vegetation like grass. Further, large wild carrot patch occurrence was also in areas with a high proportion of dry vegetation this may be explained by the phylogenetic origins of wild carrot in moisture deprived arid and semi-arid regions of central Asia and reduced competition posed by absence of other vegetation.

Classification algorithms were applied to hyperspectral data collected in the region with the aim of detecting wild carrot plants. The second order classification applied called MTMF helped create a model which predicted possible locations of the wild carrot plants. The classifier was able to detect wild carrot patches amongst other vegetation which had a surface area of 0.05m^2 or greater (detection threshold of 5% of total pixel surface area). However, this detection threshold meant that it was unable to detect most of the wild carrot patches sampled (47/65, 72%) as the majority fell below this threshold. However, in the context of carrot seed production this may be

an acceptable threshold, especially considering the clustered nature of occurrence of wild carrot plants, but nonetheless the risk of not identifying isolated small patches remains. Regardless this work has demonstrated that hyperspectral remote sensing is a feasible method for detecting wild carrot weed populations amongst other weed populations, but that further research is needed to for hyperspectral remote sensing to be used commercially. The further research needed is discussed in the next section.

6.1 Future Research

The study identified potential future research which may help improve on the current model in terms of its detection threshold and accuracy levels. This includes collection of better endmember training data – collection of larger wild carrot patches for 'pure pixels' of wild carrot. Improving spatial resolution by changes in data collection methods – including aerial data collection at a lower altitude, other data analysis techniques like 'Image fusion techniques' which can allow collection of data with both high spatial and spectral resolution. It also includes using the data collected to further study environmental context related parameters like soil moisture content, dry vegetation composition, and species composition gradients in the wild carrot growing environment in order to look for relationships between these environmental context related parameters and wild carrot presence which may help detect and predict wild carrot manifestations. All these methods are likely to help improve on the detection thresholds of wild carrot plants achieved in this study – however there are also trade-offs and limitations associated with them – these include equipment compatibility, increased costs, sensor saturation and availability of appropriate wild carrot patches for 'pure pixels' - which must be taken into consideration.

Chapter 7 References

Aboh, C. L., & Effiong, J. B. (2019). Adoption of different weed management techniques among cocoa farmers in Akamkpa Local Government Area, Cross River State, Nigeria. *Global Journal of Pure and Applied Sciences*, *25*(1), 7. https://doi.org/10.4314/gjpas.v25i1.2

Adam, E., Mutanga, O., & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: A review. *Wetlands Ecology and Management*, *18*(3), 281–296. https://doi.org/10.1007/s11273-009-9169-z

AgResearch. (2023). *AgPest " Wild carrot and parsley dropwort*. AgPest. Retrieved January 30, 2023, from https://agpest.co.nz/?pesttypes=wild-carrot-and-parsley-dropwort

Amigo, J. M., & Santos, C. (2020). Preprocessing of hyperspectral and multispectral images. *Data Handling in Science and Technology*, *32*, 37–53. https://doi.org/10.1016/B978-0-444-63977-6.00003-1

Andrew, M. E., & Ustin, S. L. (2008). The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sensing of Environment*, *112*(12), 4301–4317. https://doi.org/10.1016/j.rse.2008.07.016

Andrew, M. E., & Ustin, S. L. (2009). Habitat suitability modelling of an invasive plant with advanced remote sensing data. *Diversity and Distributions*, *15*(4), 627–640. https://doi.org/10.1111/j.1472-4642.2009.00568.x

Arslan, N. (2018). Assessment of oil spills using Sentinel 1 C-band SAR and Landsat 8 multispectral sensors. *Environmental Monitoring and Assessment*, *190*(11). https://doi.org/10.1007/s10661-018-7017-4

Artigas, F. J., & Yang, J. S. (2007). Hyperspectral remote sensing of marsh species and plant vigour gradient in the New Jersey Meadowlands. *International Journal of Remote Sensing*, *1161*. https://doi.org/10.1080/01431160500218952

AsureQuality. (2021). Seed Crop Isolation Distance Project. https://scid.asurequality.com/

Avolio, M. L., Forrestel, E. J., Chang, C. C., Pierre, K. J. La, Burghardt, K. T., & Smith, M. D. (2019). Demystifying dominant species. *New Phytologist*, *223*, 1106–1126. https://doi.org/10.1111/nph.15789

Bhatia, N., Iordache, M., Stein, A., Reusen, I., & Tolpekin, V. A. (2018). Remote Sensing of Environment Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing. *Remote Sensing of Environment*, *204*, 472–484. https://doi.org/10.1016/j.rse.2017.10.008

Bhatia, N., Tolpekin, V. A., Reusen, I., Sterckx, S., Biesemans, J., & Stein, A. (2015). Sensitivity of Reflectance to Water Vapor and Aerosol Optical Thickness. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(6), 3199–3208. https://doi.org/10.1109/JSTARS.2015.2425954

Borman, M. M., Krueger, W. C., & Johnson, D. E. (1991). Effects of established perennial grasses on yields of associated annual weeds. *Journal of Range Management*, *44*(4), 318–322. https://doi.org/10.2307/4002390

Broussard, M. A., Mas, F., Howlett, B., Pattemore, D., & Tylianakis, M. (2017). Possible mechanisms of pollination failure in hybrid carrot seed and implications for industry in a changing climate. *PLoS ONE*, 1–23.

Chapke, R. R., & Tonapi, V. (2016). Principles of quality seed production and maintenance in sorghum. In *Best practices for sorghum cultivation and importance of value-addition* (Issue October, pp. 163–168). www.millets.res.in

Charlton, J. F. L., & Stewart, A. V. (1999). Pasture species and cultivars used in New Zealand – a list. *Proceedings of the New Zealand Grassland Association*, *61*, 147–166.

Clark, M. L., & Kilham, N. E. (2016). Mapping of land cover in northern California with simulated hyperspectral satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *119*, 228–245. https://doi.org/10.1016/j.isprsjprs.2016.06.007

Colombo, N., & Galmarini, C. R. (2017). The use of genetic, manual and chemical methods to control pollination in vegetable hybrid seed production: a review. *Plant Breeding*, *136*(3), 287–299. https://doi.org/10.1111/pbr.12473

da Silva Dias, J. C. (2014). Nutritional and Health Benefits of Carrots and Their Seed Extracts. *Food and Nutrition Sciences*, *05*(22), 2147–2156. https://doi.org/10.4236/fns.2014.522227

Dadon, A., Mandelmilch, M., Ben-Dor, E., & Sheffer, E. (2019). Sequential PCA-based classification

of mediterranean forest plants using airborne hyperspectral remote sensing. *Remote Sensing*, *11*(23). https://doi.org/10.3390/rs11232800

Dao, P. D., He, Y., & Lu, B. (2019). Maximizing the quantitative utility of airborne hyperspectral imagery for studying plant physiology : An optimal sensor exposure setting procedure and empirical line method for atmospheric correction. *International Journal of Applied Earth Observation and Geoinformation*, *77*, 140–150. https://doi.org/10.1016/j.jag.2018.11.010

Dasf, F., Adams, J., & Lewis, P. (2018). Decoupling Canopy Structure and Leaf Biochemistry : Testing the Utility of Directional Area Scattering. *Remote Sensing*, *10*(1911). https://doi.org/10.3390/rs10121911

DataIntelo. (2022). *Global carrot seed market by type (large carrot seeds, Cherry Carrot Seeds), by application (farmland, greenhouse, others) and by region (North America, Latin America, Europe, Asia Pacific and Middle East & Africa), forecast from 2022 to 2030*. Dataintelo. Retrieved April 11, 2023, from https://dataintelo.com/report/carrot-seed-market/

Datta, S. (2009). Aquatic Weeds and Their Management for Fisheries. *Researchgate.Net*, *March*, 22. https://www.researchgate.net/profile/Subhendu-Datta/publication/258931408_Aquatic_Weeds_and_Their_Management_for_Fisheries/links/00b7d5321adae291c6000000/Aquatic-Weeds-and-Their-Management-for-Fisheries.pdf

Dennison, P. E., Charoensiri, K., Roberts, D. A., Peterson, S. H., & Green, R. O. (2006). Wildfire temperature and land cover modeling using hyperspectral data. *Remote Sensing of Environment*, *100*, 212–222. https://doi.org/10.1016/j.rse.2005.10.007

Dessaint, Y. F., Chadoeuf, R., & Barralis, G. (1991). Spatial Pattern Analysis of Weed Seeds in the Cultivated Soil Seed Bank. *Journal of Applied Ecology*, *28*(2), 721–730.

Du, Q., French, J. V., Skaria, M., Yang, C., & Everitt, J. H. (2004). Citrus pest stress monitoring using airborne hyperspectral imagery. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *6*(1), 3981–3984. https://doi.org/10.1109/igarss.2004.1370000

Dynes, R. A., Burggraaf, V. T., Goulter, C. G., & Dalley, D. E. (2010). Canterbury farming: production, processing and farming systems. *Proceedings of the New Zealand Grassland Association*.

Eber, F., Boucherie, R., Broucqsault, L.-M., Bouchet, Y., & Chèvre, A.-M. (1998). Spontaneous hybridisation between vegetable crops and weeds. 1 : Garden radish (*Raphanus sativus* L.) and wild mustard (*Sinapis arvensis* L.). *Agronomie*, *18*, 489–497.

EO College. (2021). *Beyond the visible – introduction to hyperspectral remote sensing*. EO College. Retrieved December 7, 2021, from https://eo-college.org/courses/beyond-the-visible/

Freshfacts. (2022). https://unitedfresh.co.nz/assets/site/Fresh-Facts-2021.pdf

Freshfacts. (2021). https://unitedfresh.co.nz/assets/site/Fresh-Facts-2020.pdf

Fu, Y., Yang, G., Li, Z., Li, H., Li, Z., Xu, X., Song, X., Zhang, Y., Duan, D., Zhao, C., & Chen, L. (2020). Progress of hyperspectral data processing and modelling for cereal crop nitrogen monitoring. *Computers and Electronics in Agriculture*, *172*, 105321. https://doi.org/10.1016/j.compag.2020.105321

George, R. A. T. (2009). *Vegetable Seed Production* (3rd ed.). CAB International. https://books-google-co-nz.ezproxy.massey.ac.nz/books?hl=en&lr=&id=nWilRHvRiFIC&oi=fnd&pg=PR5&dq=isolation+distance+for+hybrid+vegetable+seed+production&ots=0IVroX0amy&sig=5FkCG79HRplPtV50RnKX0HeTgtY&redir_esc=y#v=onepage&q=isolation distance for hybrid v&f=false

Ghersa, C. M., & Martinez-Ghersa, M. A. (2000). Ecological correlates of weed seed size and persistence in the soil under different tilling systems: implications for weed management. *Field Crops Research*, *67*, 141–148.

Giannoni, L., Lange, F., & Tachtsidis, I. (2018). Hyperspectral imaging solutions for brain tissue metabolic and hemodynamic monitoring: Past, current and future developments. *Journal of Optics (United Kingdom)*, *20*(4). https://doi.org/10.1088/2040-8986/aab3a6

Gibson, P. J., Keating, J., & Power, C. H. (2000a). *Introductory remote sensing*. Routledge.

Gibson, P. J., Keating, J., & Power, C. H. (2000b). Principles of remote sensing. In *Introductory remote sensing*.

GISGeography. (2022, May 30). *Multispectral vs Hyperspectral Imagery explained*. GIS Geography. Retrieved September 4, 2022, from https://gisgeography.com/multispectral-vs-hyperspectral-imagery-explained/

Glenn, N. F., Mundt, J. T., Weber, K. T., Prather, T. S., Lass, L. W., & Pettingill, J. (2005). Hyperspectral data processing for repeat detection of small infestations of leafy spurge. *Remote Sensing of Environment*, *95*(3), 399–412. https://doi.org/10.1016/j.rse.2005.01.003

Gonçalves, L., Fonte, C., Júlio, E., & Caetano, M. (2009). Evaluation of Remote Sensing Image Classifiers with Uncertainty Measures. *Spatial Data Quality*, *June*, 163–177. https://doi.org/10.1201/b10305-19

Groot, M. H. M., van de Wiel, C. C. M., van Tienderen, P. H., & den Nijs, H. C. M. (2003). Hybridisation and introgression between crops and wild relatives. *COGEM Research*, *August 2016*, 55.

Habib, A., Xiong, W., He, F., Yang, H. L., & Crawford, M. (2017). Improving Orthorectification of UAV-Based Push-Broom Scanner Imagery Using Derived Orthophotos from Frame Cameras. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(1), 262–276. https://doi.org/10.1109/JSTARS.2016.2520929

Hadjimitsis, D. G., & Themistocleous, K. (2008). The importance of considering atmospheric correction in the pre-processing of satellite remote sensing data intended for the management and detection of cultural sites: A case study of the Cyprus area. *Digital Heritage - Procceedings of the 14th International Conference on Virtual Systems and Multimedia*, *October*, 9–12. https://www.researchgate.net/publication/257067491

Hamada, Y., Stow, D. A., Coulter, L. L., Jafolla, J. C., & Hendricks, L. W. (2007). Detecting Tamarisk species (Tamarix spp.) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery. *Remote Sensing of Environment*, *109*(2), 237–248. https://doi.org/10.1016/j.rse.2007.01.003

Hampton, J. G., Rolston, M. P., Pyke, N. B., & Green, W. (2012). Ensuring the long term viability of the New Zealand seed industry. *Agronomy Society of New Zealand Special Publication*, 129–140.

Hamzeh, S., Naseri, A. A., AlaviPanah, S. K., Mojaradi, B., Bartholomeus, H. M., Clevers, J. G. P. W., & Behzad, M. (2013). Estimating salinity stress in sugarcane fields with spaceborne hyperspectral: Vegetation indices. *International Journal of Applied Earth Observation and Geoinformation*, *21*(1), 282–290. https://doi.org/10.1016/j.jag.2012.07.002

Hauser, T. P., & Bjørn, G. K. (2001). Hybrids between wild and cultivated carrots in Danish carrot fields. *Genetic Resources and Crop Evolution*, *48*(5), 499–506. https://doi.org/10.1023/A:1012051731933

Hauser, T. P., Bjørn, G. K., Magnussen, L., & Shim SangIn, S. S. (2004). Hybrids between cultivated and wild carrots: a life history. In *Introgression from genetically modified plants into wild relatives* (pp. 41–51). https://doi.org/10.1079/9780851998169.0041

He, L., Hatier, J. H. B., & Matthew, C. (2017). Drought tolerance of two perennial ryegrass cultivars with and without AR37 endophyte. *New Zealand Journal of Agricultural Research*, *60*(2), 173–188. https://doi.org/10.1080/00288233.2017.1294083

Iorizzo, M., Senalik, D. O. A., Ellison, S. L., Grzebelus, D., Cavagnaro, P. F., Allender, C., Brunet, J., Spooner, D. M., van Deynze, A., & Simon, P. W. (2013). Genetic Structure And Domestication Of Carrot (*Daucus Carota* Subsp. *Sativus*) (Apiaceae). *American Journal of Botany*, *100*(5), 930–938. https://doi.org/10.3732/ajb.1300055

Isse, J. Y., & Ghouch, C. El. (2016). *Information Theoretic Similarity Measures for Robust Image Matching* (Issue May). https://doi.org/10.13140/RG.2.1.3209.1769

Japan Association of Remote Sensing. (1996.). *Geometric Correction*. Retrieved January 26, 2023, from http://sar.kangwon.ac.kr/etc/rs_note/rsnote/cp9/cp9-4.htm

Jin, X. (2018). *ENVI Automated Image Registration Solutions*. L3Harris Geospatial. https://www.nv5geospatialsoftware.com/portals/0/pdfs/ENVI_Image_Registration_Whitepaper.pdf

Karim, R. S. M., Man, A. B., & Sahid, I. B. (2004). Weed problems and their management in rice fields of Malaysia: An overview. *Weed Biology and Management*, *4*(4), 177–186. https://doi.org/10.1111/j.1445-6664.2004.00136.x

Keatinge, J. D. H., Yang, R. Y., Hughes, J., Easdown, W. J., & Holmer, R. (2011). The importance of vegetables in ensuring both food and nutritional security in attainment of the Millennium Development Goals. *Food Security*, *3*(4), 491–501. https://doi.org/10.1007/s12571-011-0150-3

Kiran, U., & Pandey, N. K. (2020). Transgenic food crops: public acceptance and IPR. In *Transgenic*

technology based value addition in plant biotechnology (pp. 273–307). INC. https://doi.org/10.1016/b978-0-12-818632-9.00012-5

Koziol, L., Rieseberg, L. H., Kane, N., & Bever, J. D. (2012). Reduced drought tolerance during domestication and the evolution of weediness results from tolerance-growth trade-offs. *Evolution*, *66*(12), 3803–3814. https://doi.org/10.1111/j.1558-5646.2012.01718.x

L3Harris Geospatial Solutions. (2023). *Georeference from Input Geometry*. Retrieved March 1, 2023, from https://www.l3harrisgeospatial.com/docs/georeferencefrominputgeometry.html

L3Harris Geospatial Solutions. (2022.). *Mixture tuned matched filtering*. Retrieved December 27, 2022, from https://www.l3harrisgeospatial.com/docs/mtmf.html

L3Harris Geospatial Solutions. (2017, March 6). *Push broom and whisk broom sensors*. L3Harris Geospatial. Retrieved January 10, 2023, from https://www.l3harrisgeospatial.com/Learn/Blogs/Blog-Details/ArtMID/10198/ArticleID/16262/Push-Broom-and-Whisk-Broom-Sensors

Landmann, T., Feilhauer, H., Shen, M., Chen, J., & Raina, S. (2018). Mapping the Distribution and Abundance of Flowering Plants Using Hyperspectral Sensing. In P. S. Thenkabail, J. G. Lyon, & A. Huete (Eds.), *Advanced applications in remote sensing of agricultural crops and natural vegetation* (pp. 69–78). CRC Press. https://doi.org/10.1201/9780429431166-4

Lass, L. W., Thill, D. C., Shafii, B., & PRATHER, Ti. S. (2002). Detecting Spotted Knapweed (*Centaurea maculosa*) with Hyperspectral Remote Sensing Technology 1. *Weed Technology*, *16*(2), 426–432. https://doi.org/10.1614/0890-037x(2002)016[0426:dskcmw]2.0.co;2

Lavanya, C. (2002). Sensitivity of sex expression and sex variation in castor (*Ricinus communis* L.) to environmental changes. *Indian Journal of Genetics and Plant Breeding*, *62*(3), 232–237. http://cat.inist.fr/?aModele=afficheN&cpsidt=14005905

Li, J., Lammerts, E. T., Jiggins, J., & Leeuwis, C. (2012). Farmers' adoption of maize (*Zea mays* L.) hybrids and the persistence of landraces in Southwest China: implications for policy and breeding. *Genetic Resources and Crop Evolution*, *59*, 1147–1160. https://doi.org/10.1007/s10722-011-9750-1

Li, X., Feng, R., Guan, X., Shen, H., & Zhang, L. (2019). Remote Sensing Image Mosaicking:

Achievements and Challenges. *IEEE Geoscience and Remote Sensing Magazine*, *7*(4), 8–22. https://doi.org/10.1109/MGRS.2019.2921780

Li, Y., Yin, H., Yao, J., Wang, H., & Li, L. (2022). A unified probabilistic framework of robust and efficient color consistency correction for multiple images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *190*(June), 1–24. https://doi.org/10.1016/j.isprsjprs.2022.05.009

Liang, S., Wang, J., & Wen, L. xiao. (2020). Atmospheric correction of optical imagery. In *Advanced Remote Sensing* (pp. 131–156). https://doi.org/10.1016/b978-0-12-815826-5.00004-0

Linke, B., Alessandro, M. S., Galmarini, C. R., & Nothnagel, T. (2019). Carrot Floral Development and Reproductive Biology. In *The Carrot Genome* (pp. 27–57). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-03389-7_3

Lioubimtseva, E., & Henebry, G. M. (2009). Climate and environmental change in arid Central Asia: Impacts, vulnerability, and adaptations. *Journal of Arid Environments*, *73*(11), 963–977. https://doi.org/10.1016/j.jaridenv.2009.04.022

Mahajan, G. R., Pandey, R. N., Sahoo, R. N., Gupta, V. K., Datta, S. C., & Kumar, D. (2017). Monitoring nitrogen, phosphorus and sulphur in hybrid rice (*Oryza sativa* L.) using hyperspectral remote sensing. *Precision Agriculture*, *18*(5), 736–761. https://doi.org/10.1007/s11119-016-9485-2

Mandel, J. R., Ramsey, A. J., Iorizzo, M., & Simon, P. W. (2016). Patterns of Gene Flow between Crop and Wild Carrot, Daucus carota (Apiaceae) in the United States. *PLoS ONE*, 1–19. https://doi.org/10.1371/journal.pone.0161971

Manea, D., & Calin, M. A. (2015). Hyperspectral imaging in different light conditions. *Imaging Science Journal*, *63*(4), 214–219. https://doi.org/10.1179/1743131X15Y.0000000001

Markelin, L., Honkavaara, E., Peltoniemi, J., Ahokas, E., & Kuittinen, R. (2008). Radiometric Calibration and Characterization of Large-format Digital Photogrammetric Sensors in a Test Field. *Photogrammetric Engineering & Remote Sensing*, *74*.

Meerdink, S. K., Roberts, D. A., Roth, K. L., King, J. Y., Gader, P. D., & Koltunov, A. (2019). Classifying California plant species temporally using airborne hyperspectral imagery. *Remote Sensing of Environment*, *232*, 111308. https://doi.org/10.1016/j.rse.2019.111308

Mengistu, T., & Yamoah, C. (2010). Effect of sowing date and planting density on seed production of carrot (*Daucus carota* var . *sativa*) in Ethiopia. *African Journal of Plant Science*, *4*, 270–279.

Merfield, C. N., Hampton, J. G., Wratten, S. D., Prapanoppasin, P., & Yeeransiri, P. (2001). Seed production studies in carrot (*Daucus carota* L.) I. Effect of plant density on seed quality and yield. *Agronomy Society of New Zealand*, *13*(i), 12.

Midlands Seed. (2023). Midlands Seed - About Us. https://www.midlandsnz.com/our-products/seed/about-us

Mirik, M., Ansley, R. J., Steddom, K., Jones, D. C., Rush, C. M., Michels, G. J., & Elliott, N. C. (2013). Remote distinction of a noxious weed (Musk Thistle: *Carduus Nutans*) using airborne hyperspectral imagery and the support vector machine classifier. *Remote Sensing*, *5*(2), 612–630. https://doi.org/10.3390/rs5020612

Mohd Hasmadi, I., Pakhriazad, H. Z., & Shahrin, M. F. (2009). Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. *Malaysian Journal of Society and Space*, *5*(1), 1–10.

Mozgeris, G., Juodkiene, V., Jonikavičius, D., Straigyte, L., Gadal, S., & Ouerghemmi, W. (2018). Ultra-Light aircraft-based hyperspectral and colour-infrared imaging to identify deciduous tree species in an urban environment. *Remote Sensing*, *10*(10). https://doi.org/10.3390/rs10101668

Mundt, J. T., Glenn, N. F., Weber, K. T., Prather, T. S., Lass, L. W., & Pettingill, J. (2005). Discrimination of hoary cress and determination of its detection limits via hyperspectral image processing and accuracy assessment techniques. *Remote Sensing of Environment*, *96*(3–4), 509–517. https://doi.org/10.1016/j.rse.2005.04.004

Nair, R., & Kumar, S. (2021). Innovation in hybrid seed production of vegetable crops : A review. *The Pharma Innovation Journal*, *10*(1), 1270–1275.

Näsi, R., Honkavaara, E., Blomqvist, M., Lyytikäinen-Saarenmaa, P., Hakala, T., Viljanen, N., Kantola, T., & Holopainen, M. (2018). Remote sensing of bark beetle damage in urban forests at individual tree level using a novel hyperspectral camera from UAV and aircraft. *Urban Forestry and Urban Greening*, *30*, 72–83. https://doi.org/10.1016/j.ufug.2018.01.010

NIWA. (2021, May 12). *National and regional climate maps*. Retrieved February 1, 2023, from https://niwa.co.nz/climate/national-and-regional-climate-maps/national

Noujdina, N. V., & Ustin, S. L. (2008). Mapping Downy Brome (*Bromus tectorum*) Using Multidate AVIRIS Data. *Weed Science*, *56*(1), 173–179. https://doi.org/10.1614/ws-07-009.1

OECD. (2022). OECD schemes for the varietal certification of seed. https://www.oecd.org/agriculture/seeds/

Oliver, M. A., Bishop, T. F. A., & Marchant, B. P. (2013). Precision agriculture for sustainability and environmental protection. In *Precision Agriculture for Sustainability and Environmental Protection*. https://doi.org/10.4324/9780203128329

Pervez, M. A., Ayub, C. M., Khan, H. A., Shahid, M. A., & Ashraf, I. (2009). Effect of drought stress on growth , yield and seed quality of tomato (*Lycopersicon esculentum* L.). *Pakistan Journal of Agricultural Science*, *46*(3), 174–178.

Petropoulos, G. P., Vadrevu, K. P., & Kalaitzidis, C. (2013). Spectral angle mapper and object-based classification combined with hyperspectral remote sensing imagery for obtaining land use/cover mapping in a Mediterranean region. *Geocarto International*, *28*(2), 114–129. https://doi.org/10.1080/10106049.2012.668950

Pignatti, S., Casa, R., Harfouche, A., Huang, W., Palombo, A., Pascucci, S., Imaa, C. N. R., & Loja, C. S. (2019). Maize crop and weeds species detection by using UAV VNIR hyperpectral data. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 7235–7238.

Preece, D. (2023, January). *NZ's world-leading seed industry just keeps growing*. 1 News. Retrieved January 30, 2023, from https://www.1news.co.nz/2023/01/15/nzs-world-leading-seed-industry-just-keeps-growing/

Que, F., Hou, X. L., Wang, G. L., Xu, Z. S., Tan, G. F., Li, T., Wang, Y. H., Khadr, A., & Xiong, A. S. (2019). Advances in research on the carrot, an important root vegetable in the Apiaceae family. *Horticulture Research*, *6*(1). https://doi.org/10.1038/s41438-019-0150-6

Ravensdown. (n.d.). *Aerowork aerial spreading*. Ravensdown. Retrieved February 10, 2023, from https://www.ravensdown.co.nz/services/spreading/aerowork-aerial-spreading

Reef Resilience Network. (2023). *What is remote sensing?* Retrieved February 10, 2023, from https://reefresilience.org/management-strategies/remote-sensing-and-mapping/introduction-to-remote-sensing/what-is-remote-sensing/

Rome, C., & Lucero, C. (2019). Wild Carrot (*Daucus carota*) Management in the Dungeness Valley, Washington, United States: The Power of Citizen Scientists to Leverage Policy Change. *Citizen Science: Theory and Practice*, *4*(1), 1–9. https://doi.org/10.5334/cstp.201

Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C., Kersting, K., Rascher, U., & Plümer, L. (2012). Early drought stress detection in cereals: Simplex volume maximisation for hyperspectral image analysis. *Functional Plant Biology*, *39*(11), 878–890. https://doi.org/10.1071/FP12060

Rong, J., Janson, S., Umehara, M., Ono, M., & Vrieling, K. (2010). Historical and contemporary gene dispersal in wild carrot (*Daucus carota* ssp . *carota*) populations. *Annals of Botany*, *106*, 285–296. https://doi.org/10.1093/aob/mcq108

Roth, K. L., Roberts, D. A., Dennison, P. E., Alonzo, M., Peterson, S. H., & Beland, M. (2015). Remote Sensing of Environment Differentiating plant species within and across diverse ecosystems with imaging spectroscopy. *Remote Sensing of Environment*, *167*, 135–151. https://doi.org/10.1016/j.rse.2015.05.007

Salem, F., Kafatos, M., El-Ghazawi, T., Gomez, R., & Yang, R. (2005). Hyperspectral image assessment of oil-contaminated wetland. *International Journal of Remote Sensing*, *26*(4), 811–821. https://doi.org/10.1080/01431160512331316883

Santellani, E., Maset, E., & Fusiello, A. (2018). Seamless Image Mosaicking Via Synchronization. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*(2), 247–254. https://doi.org/10.5194/isprs-annals-IV-2-247-2018

Sarath, T., Nagalakshmi, G., & Jyothi, S. (2014). A Study on Hyperspectral Remote Sensing Classifications. *International Journal of Computer Applications. International Conference on Information and Communication Technologies*, 0975–8887.

Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: Parametric orthorectification. *International Journal of Remote Sensing*, *23*(13), 2609–2630. https://doi.org/10.1080/01431160110115825

Scott, A. (2021, February 3). *Seed exports looking good*. Farmers Weekly. Retrieved April 4, 2022, from https://www.farmersweekly.co.nz/news/seed-exports-looking-good/

Shi, C., & Wang, L. (2014). Incorporating spatial information in spectral unmixing: A review. *Remote Sensing of Environment*, *149*, 70–87. https://doi.org/10.1016/j.rse.2014.03.034

Simon, P. W. (2019). Classical and Molecular Carrot Breeding. In *The Carrot Genome* (pp. 137–147). https://doi.org/10.1007/978-3-030-03389-7_9

Simone, G. (2002). Image fusion techniques for remote sensing applications. *Information Fusion*, *3*, 3–15.

Solberg, S. Ø., Yndgaard, F., Andreasen, C., Bothmer, R. Von, Loskutov, I. G., & Asdal, Å. (2020). Long-Term Storage and Longevity of Orthodox Seeds : A Systematic Review. *Frontiers in Plant Science*, *11*(July), 1–14. https://doi.org/10.3389/fpls.2020.01007

Spooner, D. M. (2019). Daucus: Taxonomy, Phylogeny, Distribution. In *The Carrot Genome* (pp. 9–26). https://doi.org/10.1007/978-3-030-03389-7_2

Srivastava, P. K., Malhi, R. K. M., Pandey, P. C., Anand, A., Singh, P., Pandey, M. K., & Gupta, A. (2020). Revisiting hyperspectral remote sensing: Origin, processing, applications and way forward. In *Hyperspectral Remote Sensing: Theory and Applications* (pp. 3–21). https://doi.org/10.1016/B978-0-08-102894-0.00001-2

Stafford, J. V., & Miller, P. C. H. (1996). Spatially Variable Treatment of Weed Patches. *Precision Agriculture*, 465–474. https://doi.org/10.2134/1996.precisionagproc3.c50

Staniland, B. K., McVetty, P. B. E., Friesen, L. F., Yarrow, S., Freyssinet, G., & Freyssinet, M. (2000). Effectiveness of border areas in confining the spread of transgenic *Brassica napus* pollen. *Canadian Journal of Plant Science*, *80*(3), 521–526. https://doi.org/10.4141/P99-117

Stewart, A. V. (2002). A review of Brassica species , cross-pollination and implications for pure seed production in New Zealand. *Agronomy Society of New Zealand*, *32*, 63–82.

Storkey, J., & Neve, P. (2018). What good is weed diversity? *Weed Research*, 239–243. https://doi.org/10.1111/wre.12310

Takagi, T., Nagashima, H., & Noguchi, A. (2017). *Trends in carrot production and breeding in Japan*. 15–20. https://doi.org/10.17660/ActaHortic.2017.1153.3

Thomison, P., & Geyer, A. (2016). Managing “Pollen Drift” to Minimize Contamination of Non-GMO Corn. *Agriculture and Natural Resources, Ohio State University Extension*.

Thompson, K., Band, S. R., & Hodgson, J. G. (1993). Seed size and shape predict persistence in soil. *Functional Ecology*, *7*(2), 236–241.

Thorp, K. R., & Tian, L. F. (2004). A review on remote sensing of weeds in agriculture. *Precision Agriculture*, *5*(5), 477–508. https://doi.org/10.1007/s11119-004-5321-1

Torner, C., Sanchez, M. J., Satorre, E., Torner, C., Sanchez, M. J., Satorre, E., & A, C. F. (2000). A comparison of the growth patterns and the competitive ability of four annual weeds. *Agronomie*, *20*, 147–156.

Tu, Y., Bian, M., Wan, Y., & Fei, T. (2018). Tea cultivar classification and biochemical parameter estimation from hyperspectral imagery obtained by UAV. *PeerJ*, *2018*(5), 1–17. https://doi.org/10.7717/peerj.4858

Ukrainski, P. (2019, February 26). *Classification accuracy assessment. Confusion matrix method*. 50 North | GIS blog from Ukraine. Retrieved February 13, 2023, from http://www.50northspatial.org/classification-accuracy-assessment-confusion-matrix-method/#:~:text=Another%20accuracy%cla20indicator%20is%20the,image%20and%20the%20reference%20image

Underwood, E. C., Ustin, S. L., & Ramirez, C. M. (2007). A comparison of spatial and spectral image resolution for mapping invasive plants in coastal California. *Environmental Management*, *39*(1), 63–83. https://doi.org/10.1007/s00267-005-0228-9

Uwimana, B., Smulders, M. J. M., Hooftman, D. A. P., Hartman, Y., van Tienderen, P. H., Jansen, J., McHale, L. K., Michelmore, R. W., van de Wiel, C. C. M., & Visser, R. G. F. (2012). Hybridization between crops and wild relatives: The contribution of cultivated lettuce to the vigour of crop-wild hybrids under drought, salinity and nutrient deficiency conditions. *Theoretical and Applied Genetics*, *125*(6), 1097–1111. https://doi.org/10.1007/s00122-012-1897-4

Van Kleunen, M., Weber, E., & Fischer, M. (2010). A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters*, *13*(2), 235–245. https://doi.org/10.1111/j.1461-0248.2009.01418.x

Vella, M., Zhang, B., Chen, W., & Mota, J. F. C. (2021). Enhanced hyperspectral image super-resolution via RGB fusion and TV-TV minimization. *IEEE International Conference on Image Processing (ICIP)*, 0–4.

Veraverbeke, S., Dennison, P., Gitas, I., Hulley, G., Kalashnikova, O., Katagis, T., Kuai, L., Meng, R., Roberts, D., & Stavros, N. (2018). Hyperspectral remote sensing of fire: State-of-the-art and future perspectives. *Remote Sensing of Environment*, *216*, 105–121. https://doi.org/10.1016/j.rse.2018.06.020

Vira, A., Maddison, B., Wallace, C., Alfred, J., Bhatia, S., & Williams, Z. (2020). *Improving Systems , Improving Livelihoods : Telangana Seed Value-chain Development Project with India & New Zealand Seed Industry Collaboration*.

Wang, J., Brown, D. G., & Bai, Y. (2014). Investigating the spectral and ecological characteristics of grassland communities across an ecological gradient of the Inner Mongolian grasslands with in situ hyperspectral data. *International Journal of Remote Sensing*, *35*(20), 7179–7198. https://doi.org/10.1080/01431161.2014.967885

Warner, D. J., & Lewis, K. A. (2019). Evaluation of the risks of contaminating low erucic acid rapeseed with high erucic rapeseed and identification of mitigation strategies. *Agriculture (Switzerland)*, *9*(9), 1–20. https://doi.org/10.3390/agriculture9090190

Wijnheijmer, E. H. M., Brandenburg, W. A., & Ter Borg, S. J. (1989). Possible relationships between wild and cultivated carrots (*Daucus carota* L.) in the Netherlands. *Euphytica*, *40*(1), 147–154. https://doi.org/10.1007/BF02014765

Williams, A. P., & Hunt, E. R. (2002). Estimation of leafy spurge cover from hyperspectral imagery using mixture tuned matched filtering. *Remote Sensing of Environment*, *82*(2–3), 446–456. https://doi.org/10.1016/S0034-4257(02)00061-5

Wohlfeiler, J., Alessandro, M. S., Morales, A., Cavagnaro, P. F., & Galmarini, C. R. (2022). Vernalization requirement, but not post-vernalization day length, conditions flowering in Carrot (*Daucus carota* L.). *Plants*, *11*(8), 1–10. https://doi.org/10.3390/plants11081075

Wójtowicz, M., Wójtowicz, A., & Piekarczyk, J. (2016). Application of remote sensing methods in agriculture. *Communications in Biometry and Crop Science*, *11*(1), 31–50.

Wright, H. (1980). Commercial Hybrid Seed Production. In *Hybridization of Crop Plants* (pp. 161–176).

Xie, Q., Huang, W., Liang, D., Chen, P., Wu, C., Yang, G., & Zhang, J. (2014). Indices Derived From Airborne Hyperspectral Images in Winter Wheat. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(8), 3586–3594.

Yang, C., & Everitt, J. H. (2010). Mapping three invasive weeds using airborne hyperspectral imagery. *Ecological Informatics*, *5*(5), 429–439. https://doi.org/10.1016/j.ecoinf.2010.03.002

Ye, X., Sakai, K., Manago, M., Asada, S. I., & Sasao, A. (2007). Prediction of citrus yield from airborne hyperspectral imagery. *Precision Agriculture*, *8*(3), 111–125. https://doi.org/10.1007/s11119-007-9032-2

Zanin, G., Otto, S., Riello, L., & Borin, M. (1997). Ecological interpretation of weed flora dynamics under different tillage systems. *Agriculture, Ecosystems and Environment*, *66*(3), 177–188. https://doi.org/10.1016/S0167-8809(97)00081-9

Zhao, B., Ragnarsson, H. I., Ulfarsson, M. O., Cavallaro, G., & Benediktsson, J. A. (2022). Predicting Classification Performance for Benchmark Hyperspectral Datasets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 4180–4193. https://doi.org/10.1109/JSTARS.2022.3173893

Zhao, J., De Notaris, C., & Olesen, J. E. (2020). Autumn-based vegetation indices for estimating nitrate leaching during autumn and winter in arable cropping systems. *Agriculture, Ecosystems and Environment*, *290*, 106786. https://doi.org/10.1016/j.agee.2019.106786

Chapter 8 Glossary of terms

1. Annual plant - a plant species that completes its life cycle in a single year.
2. Biennial plant – a plant species that completes its life cycle in two years.
3. Hybrid seed – seeds where the genetics are strictly controlled by facilitating crosspollination between known male and female lines bred to give the offspring specific desirable traits.
4. Isolation distance - the minimum separation distance between a seed crop and another crop of the same species that will deter crosspollination and reduce genetic contamination in the seed crop. Usually enforced by a governing body.
5. Rouging - removing a plant from the ground in or near the seed production site, including the roots (usually as a form of weed/crop management).
6. Gene introgression – movement of genes from one plant species to gene pool of another.
7. Hyperspectral imaging – a technique that records/analyses radiation across a wide range in the electromagnetic spectrum at short intervals – usually from visible to short-wave infrared.
8. Remote sensing – process of collection of earth's surface data – remotely. Usually using a distant carrier (UAVs, aircrafts, satellites).
9. UAV – unmanned flying vehicle.
10. Training data – data used to train an algorithm to perform analysis and predict outcomes it is designed to do - in this study classification analysis
11. Validation data – data used to validate and test the efficacy of the developed model
12. Endmember - spectra from surface material in its purest form.
13. Endmember pixel – pixel with 'pure' spectra (spectra from only one material)
14. Pixel – smallest unit of spectral data in a hyperspectral image – representing a known surface area – in this study 1 m^2.
15. Classification – clustering of land surface areas into groups with similar land surface cover types.
16. Classification algorithm – statistical tools that allow clustering of pixels (land surface area) with similar properties based on their surface reflectance/spectra.
17. Supervised classification algorithm - classification algorithms that require user's input through training data.
18. Unsupervised classification algorithm - classification algorithms that do not require user's input through training data.
19. MF score - One of two output scores from MTMF algorithm giving a relative match of the pixel to the target spectra. A score of 0 is background and a score of 1 is a perfect match, indicating a high likelihood of target presence at the pixel (wild carrot in the study).
20. Infeasibility score - One of two output scores from MTMF algorithm using noise variance in the data to indicate the feasibility of the MF scores. A low infeasibility score (combined with high MF score) would be associated with high likelihood of target presence at the pixel (wild carrot in the study).
21. Error of omission – False negatives. (1- Correctly classified wild carrot pixels/total reference wild carrot pixels.)

22. Error of commission – False positives. (1- Correctly classified wild carrot pixels/total classified wild carrot pixels.)
23. Producer's accuracy: 1 - omission error
24. User’s accuracy: 1- commission error

Chapter 9 Appendix

Class Confusion Matrix

File

Confusion Matrix: C:\Users\sunme\OneDrive - Massey University\Sunmeet Study\Special topic\Assignment\Supervised\MDC\MDC_2

Overall Accuracy = (4973/5150) 96.5631%
Kappa Coefficient = 0.9575

Ground Truth (Pixels)

Class	Shadow/Backgr	Road	Grass/Pasture	Bare soil/Til	Trees
Unclassified	0	0	0	0	0
Shadow/Backgr	839	12	0	0	0
Road1	0	659	0	0	0
Grass/Pasture	1	0	1052	121	19
Bare soil/Til	0	6	0	1278	0
Trees1	16	0	0	0	844
Cropping1	0	0	0	0	0
Total	856	677	1052	1399	863

Ground Truth (Pixels)

Class	Cropping	Total
Unclassified	0	0
Shadow/Backgr	0	851
Road1	0	659
Grass/Pasture	0	1193
Bare soil/Til	2	1286
Trees1	0	860
Cropping1	301	301
Total	303	5150

Ground Truth (Percent)

Class	Shadow/Backgr	Road	Grass/Pasture	Bare soil/Til	Trees
Unclassified	0.00	0.00	0.00	0.00	0.00
Shadow/Backgr	98.01	1.77	0.00	0.00	0.00
Road1	0.00	97.34	0.00	0.00	0.00
Grass/Pasture	0.12	0.00	100.00	8.65	2.20
Bare soil/Til	0.00	0.89	0.00	91.35	0.00
Trees1	1.87	0.00	0.00	0.00	97.80
Cropping1	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

Ground Truth (Percent)

Class	Cropping	Total
Unclassified	0.00	0.00
Shadow/Backgr	0.00	16.52
Road1	0.00	12.80
Grass/Pasture	0.00	23.17
Bare soil/Til	0.66	24.97
Trees1	0.00	16.70
Cropping1	99.34	5.84
Total	100.00	100.00

Class	Commission (Percent)	Omission (Percent)	Commission (Pixels)	Omission (Pixels)
Shadow/Backgr	1.41	1.99	12/851	17/856
Road1	0.00	2.66	0/659	18/677
Grass/Pasture	11.82	0.00	141/1193	0/1052
Bare soil/Til	0.62	8.65	8/1286	121/1399
Trees1	1.86	2.20	16/860	19/863
Cropping1	0.00	0.66	0/301	2/303

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Shadow/Backgr	98.01	98.59	839/856	839/851
Road1	97.34	100.00	659/677	659/659
Grass/Pasture	100.00	88.18	1052/1052	1052/1193
Bare soil/Til	91.35	99.38	1278/1399	1278/1286
Trees1	97.80	98.14	844/863	844/860
Cropping1	99.34	100.00	301/303	301/301

Figure 9-1: Validation protocol results – confusion matrix - Minimum distance classification

```
        Kruskal-Wallis rank sum test

data:  STATSwildcarrot[, c("GrassSMALL", "GrassMEDIUM", "GrassLARGE")]
Kruskal-Wallis chi-squared = 5.9608, df = 2, p-value = 0.05077

> kruskal.test(STATSwildcarrot[, c("DVegSMALL", "DVegMEDIUM", "DVegLARGE")])

        Kruskal-Wallis rank sum test

data:  STATSwildcarrot[, c("DVegSMALL", "DVegMEDIUM", "DVegLARGE")]
Kruskal-Wallis chi-squared = 1.6852, df = 2, p-value = 0.4306
```

Figure 9-2: Kruskal-Wallis test (One-way analysis of variance by ranks) output - % grass (top) and dry vegetation (bottom) composition in wild carrot patches

```
        Bartlett test of homogeneity of variances

data:  STATSwildcarrot[, c("GrassSMALL", "GrassMEDIUM", "GrassLARGE")]
Bartlett's K-squared = 3.6648, df = 2, p-value = 0.16

> bartlett.test(STATSwildcarrot[, c("DVegSMALL", "DVegMEDIUM", "DVegLARGE")])

        Bartlett test of homogeneity of variances

data:  STATSwildcarrot[, c("DVegSMALL", "DVegMEDIUM", "DVegLARGE")]
Bartlett's K-squared = 0.4216, df = 2, p-value = 0.8099
```

Figure 9-3: Bartlett test output – test for equal variations in sample sets (assumption for Kruskal-Wallis test)

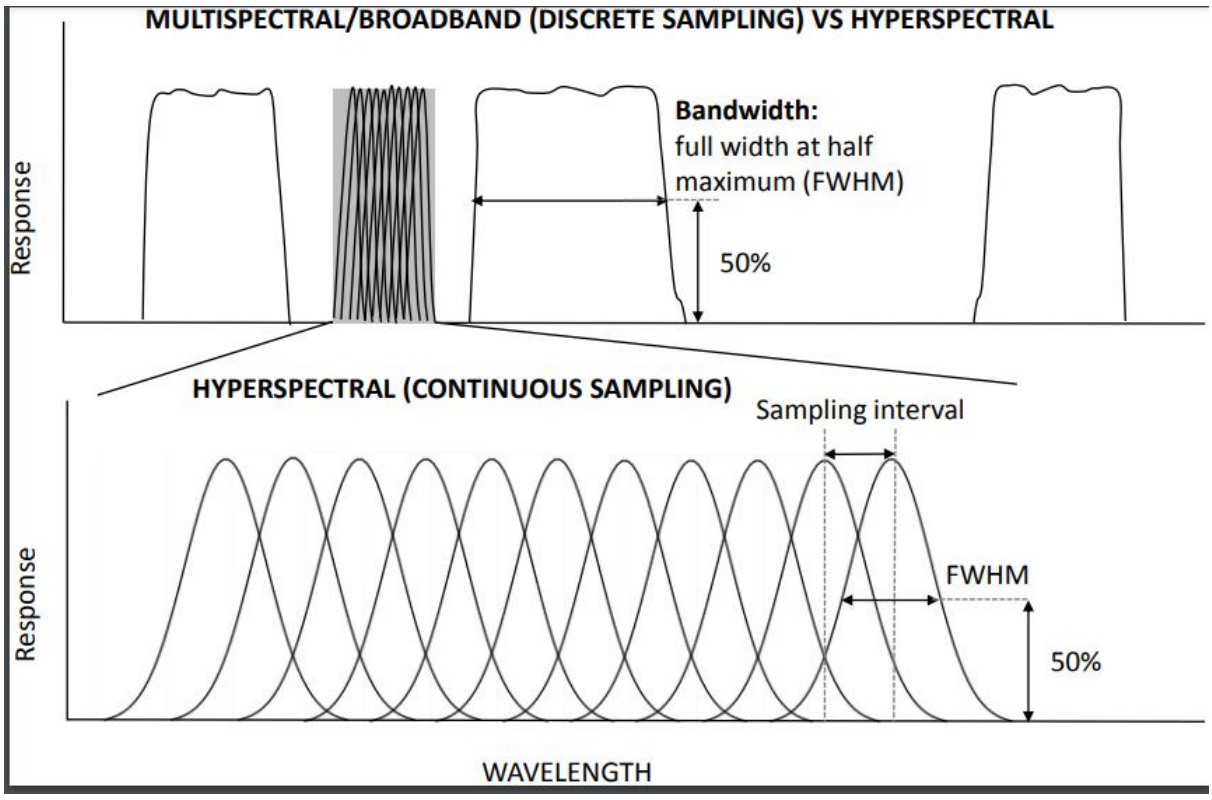

Figure 9-4: Illustration of difference in bandwidth and band number in Multi vs hyperspectral imaging