International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

Chief Editor
Dr. J.B. Helonde

Executive Editor
Mr. Somil Mayur Shah

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

PREDICTION AND DETECTION OF HONEY HARVESTS FROM REMOTE SENSING AND WEATHER DATA

Tristan Campbell[1*], Peter Fearns[2], Kingsley Dixon[2], Kenneth Dods[3]
[1*]School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, Western Australia, Australia
[2]School of Molecular and Life Sciences, Curtin University, Perth, Western Australia, Australia
[3]ChemCentre, Perth, Western Australia, Australia

DOI: 10.5281/zenodo.3596319

ABSTRACT

There have been many efforts to use remotely sensed data to map or predict the production of honey. Many studies recommend that weather data are incorporated into the model. Here we assess the ability of satellite and weather data to predict the volume of honey produced from *Corymbia calophylla* (Myrtaceae) in southwest Australia.

Utilising honey harvest data over 8 years, it was found that January NDVI could predict a 'good' harvest (more than 40 kg of honey harvested per hive) to 79% accuracy. Poor harvests (less than 20 kg of honey) and moderate harvests (between 20 and 40 kg of honey) were not distinguishable.

Assessing weather for January and February showed that the weather data from January was highly influential. Good harvests occurred after a cool, dry January, moderate harvests after a warmer, wetter January and poor harvests associated with warmer, drier January. Using a decision-tree approach, the combination of January weather and NDVI classified good harvests to 90% accuracy. Classification into the three quality levels achieved 69% accuracy from the overlapping data for poor and moderate years.

This study used monthly weather data. Addition of daily weather data and apiary health variables may improve the predictive accuracy.

KEYWORDS: Remote sensing; Corymbia calophylla; honey; prediction

1. INTRODUCTION

In nectar productive regions, such as the biodiverse region of the southwest of Australia, beekeeping has blossomed in recent years, with the number of registered beekeepers growing from 660 in 2010 to over 3,000 in 2019 [1] and significant international interest in local varietals. For example honey produced from Western Australian endemic species has some of the highest antimicrobial properties known for honey [2]. This has led to the 'farm gate' price for monofloral marri honey (produced by European honey bees (*Apis mellifera*) from the *Corymbia calophylla* trees) reaching prices of AUD 20/kg in 2018, compared to a decade earlier at of AUD 3/kg [3].

While annual honey yields from *Corymbia calophylla* forests can be exceptionally high, being over 70 kg per hive in some cases, it can also 'fail' in some seasons, with declines marri honey produced due to depleted flowering events [4]. In addition to the financial loss to the beekeeper, these failed years often result in bees starving and hive loss. Thus developing tools that can accurately predict flowering events in major nectar producing native species has major commercial benefits.

Developing predictive modelling of honey flow is a major area of interest with high commercial value to apiarists worldwide. Since the advent of remote sensing, there have been efforts to use remotely sensed data to map or predict the timing and volume of honey production for different regions for predictive hive placement.

For example, a European study of honey yield against temporal variations in the Normalized Difference Vegetation Index (NDVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor [5] showed a consistent time delay between the Start of Season (SOS) 'green-up' indicator and peak honey flow

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

time. The NDVI SOS measurement was shown to be a reasonably consistent predictor of peak honey flow timing ($R^2 = 0.76$). However, no correlation was found between the SOS and the start of honey flow as this is more dependent on the appearance of flowers on isolated plants, whereas the peak honey flow is a result of more widespread flowering and therefore more closely tied to the MODIS sensor scale.

In Australia, eucalypts across south-eastern Australia yielded a broad, qualitative correlation between changes in vegetation indices derived from MODIS data and honey yield data, with recommendations made for more detailed studies on local variations in floristic communities and the impact of climatic variables [6]. Similar work using the 'BeeBox' web-based portal [7] showed a stronger correlation between annual peak Enhanced Vegetation Index (EVI) from MODIS and *Eucalyptus tricarpa* (ironbark) winter honey flow over a 15-year period [8]. Again, this was a qualitative correlation and no quantitative model formed.

Spectral indices for predicting canola yield have shown that NDVI and EVI actually decrease over canola (*Brassica napus*) fields at the peak of flowering, despite the increase in floral cover and biomass [9]. As a result, a 'Normalized Difference Yellowness Index' (or NDYI) was developed, which provided a better yield predictor for canola than NDVI [10].

Higher resolution airborne hyperspectral data were shown to have a high degree of accuracy in detecting flowering in native vegetation in Africa [11]. Using 60 cm pixel hyperspectral data from 400 – 990 nm, this study was 77 – 97% accurate in detecting peak flowering depending on the plant species involved.

Previous spectral reflectance work by the authors has established the potential for direct detection of flowers in the peak nectar producing tree *Corymbia calophylla* (marri) using drone [12] and satellite platforms [13]. The success of the approach was influenced by the combination of the strong spectral separation of flowers of the target species from leaves and groundcover of these forests (see Figure 1), particularly in the visible and ultra-violet (UV) spectra, combined with a lack of other large canopy dominant species that flower at the same time and the large volume of flowers of marri during the flowering season (see Figure 2). Using the high degree of spectral separation of marri flowers, an algorithm (Marri Flowering Index (MFI)) was developed using MODIS data (Band 10 for the visible spectrum and Band 8 for the near-UV) that separated poor honey yield years from moderate to good years with an 80% accuracy for a trial location.

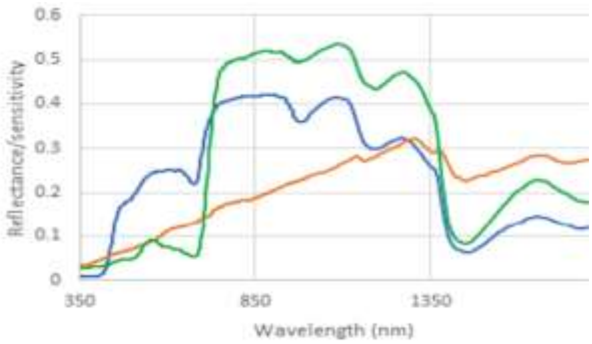

Figure 1: Median reflectance spectra collected from Corymbia calophylla forests using a handheld spectroradiometer. Blue line represents Corymbia calophylla flowers, green line represents leaves from Corymbia calophylla and other species present, orange line represents ground spectra. For a more complete discussion of the data, refer to Campbell and Fearns [13]. The full spectral data are available as a published dataset [14].

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Figure 2: Corymbia calophylla (marri) forest in bloom.

The objective of this study was to test the hypothesis that satellite-borne remotely sensed data supplemented by the inclusion of local weather conditions has the ability to accurately predict marri flowering events across a range of vegetation communities and climatic regions for the southwest of Australia, a source of high quality mono-floral varietal honeys [2]. Marri trees typically flower over late summer (February is generally the prime honey producing month [15]) and an interrogation of the Western Australian state governments botanical database [15] shows that there are rarely other canopy species flowering at this time [13]. Developing such a predictive model of honey production provides an invaluable tool for apiarists to plan hive placement to maximise yield and apiary profitability.

2. MATERIALS AND METHODS

2.1 Honey harvest and other site data

Honey harvest data were collected from across the south-west of Australia (see Figure 3), with over 350 km between the northern and southern sites. Most sites are clustered centrally, within 65 km of the Western Australian state capital city of Perth. Data were from two apiarists; one ‘commercial’ apiarist with approximately 700 hives and one ‘hobby’ apiarist with approximately 50 hives. Both apiarists use the Italian subspecies of *Apis mellifera*, with queen bees bred from the ‘Better Bee’ queen breeding program (based on Western Australia’s Rottnest Island). The ‘northern apiary sites’ (i.e. those more than 50 km north of Perth) are typically used for end of season harvests by the commercial apiarist as the marri trees typically finish flowering later in the season further north [16].

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Figure 3: Apiary site locations indicated by yellow markers. Site labels correspond to those listed in Table 1. Coordinates are in WGS84.

Data included the number of hives at the apiary site for the season, dates the hives were moved onto and off the apiary site and the total weight of marri honey harvested from the apiary site for the season. This resulted in data from 58 harvests between 2010 – 2018 across 16 apiary sites, with not all apiary sites being utilised each year. Other apiary variables, such as queen bee quality and hive condition or strength (for example measures of adult bee population and brood, as per Delaplane, Van Der Steen [17]) and when the marri trees started to flower, were not recorded. All apiary sites were at least three kilometres from other known apiary sites, and therefore outside the forage range for honey bees when nectar is moderately to readily available from other apiary sites [18].

Harvest data were reduced to the average weight of honey harvested per hive per year per apiary site. The summary of the honey harvest data presented in Table 1 highlights the difficulty in obtaining continuous, or at least annual, data for all sites. Not all apiary sites are utilised every year, however the absence of data should not be interpreted as implying likely low yield. For example, sometimes a site is not utilised because better yield is expected elsewhere. The practice of apiarists locating hives at the sites expecting higher yields each year limits the number of data points representing failed honey flow due to no flowers, thus limiting the potential for verification of the "null result" in the satellite-derived data. There are four "null results" in the honey harvest data used, all of which occur from 2010 - 2012. It is important to note that in many years, the harvest data ranged from poor to good depending on the location, indicating a high spatial variability in nectar production for a given year across different locations.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Table 1: Summary of average marri honey harvest per hive by apiary site and year, with the number of hives at the apiary site in brackets after the weight. Red = poor harvest (< 20 kg per hive), yellow = moderate harvest (20 - 40 kg per hive) and green = good harvest (> 40 kg per hive)

Site	2010	2011	2012	2013	2014	2015	2016	2017	2018
101	*Not used*	*Not used*	*Not used*	*Not used*	*Not used*	39.3 (2)	20.3 (4)	12.5 (4)	35.0 (4)
102	*Not used*	*Not used*	*Not used*	*Not used*	*Not used*	36.3 (2)	0.0 (4)	4.0 (4)	30.0 (4)
103	*Not used*	*Not used*	*Not used*	*Not used*	*Not used*	*Not used*	6.0 (3)	10.0 (4)	28.1 (5)
201	*Not used*	*Not used*	*Not used*	49.6 (112)	*Not used*	52.2 (112)	*Not used*	*Not used*	*Not used*
202	*Not used*	*Not used*	*Not used*	65.6 (112)	*Not used*	45.5 (112)	*Not used*	*Not used*	*Not used*
204	0.0 (112)	16.1 (112)	*Not used*	71.0 (112)	*Not used*	*Not used*	*Not used*	26.8 (112)	40.2 (112)
205	16.1 (112)	0.0 (112)	*Not used*	*Not used*	*Not used*	48.2 (112)	*Not used*	13.4 (112)	38.8 (112)
206	34.8 (112)	*Not used*	0.0 (112)	18.8 (112)	*Not used*	29.5 (112)	*Not used*	*Not used*	16.1 (112)
207	37.5 (112)	*Not used*	0.0 (112)	18.8 (112)	*Not used*	29.5 (112)	*Not used*	*Not used*	16.1 (112)
208	18.0 (224)	*Not used*	8.0 (224)	16.1 (112)	16.1 (112)	12.1 (112)	*Not used*	*Not used*	16.1 (112)
209	25.4 (112)	*Not used*	*Not used*	18.8 (112)	*Not used*	20.1 (112)	*Not used*	*Not used*	16.1 (112)
210	34.8 (112)	*Not used*	*Not used*	56.3 (112)	*Not used*	67.0 (112)	*Not used*	*Not used*	*Not used*
211	16.1 (112)	*Not used*	*Not used*	49.6 (112)	*Not used*	52.2 (112)	*Not used*	*Not used*	*Not used*
212	*Not used*	8.9 (112)	*Not used*	42.9 (112)	*Not used*	32.1 (112)	*Not used*	11.6 (112)	26.8 (112)
213	*Not used*	*Not used*	17.1 (112)	*Not used*	*Not used*	*Not used*	*Not used*	44.2 (112)	*Not used*
Annual average	*22.8*	*8.3*	6.3	*40.8*	*16.1*	*38.6*	*8.8*	17.5	*26.3*

Previous work on correlating MODIS data with marri honey harvests by Campbell and Fearns [12] utilised an apiary site surrounded in all directions by several kilometres of relatively uniform, mature marri forest. It was expected that this would show the greatest change due to flowering intensity from flowers resulting in a larger

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

proportion of the pixel's area with the floral signature compared with mixed land-use areas. However not all apiary sites have such consistent forest coverage.

To assess the impact of canopy cover on the vegetation indices and the MFI metric across a broader range on canopy cover densities than the previous study (i.e. including younger regrowth forest and partially cleared land rather than only mature forest), vegetation structure products produced by AusCover [19] were used to represent the mature canopy cover for each site. The AusCover vegetation structure data were derived from remotely sensed Landsat data, resulting in a ground resolution of approximately 30 m. While the uncertainty for the vegetation cover has been proven to be as high as 20% [20], the AusCover data were considered sufficiently accurate to distinguish complete forest cover from partially cleared land or land with some remnant trees.

With flowering marri trees being from 10 – 40 m high depending on age [21], the proportions of plant cover for the 10 – 30 m and > 30 m intervals were summed (the AusCover dataset provides fraction of vegetation cover from between 0 – 5 m height, 5 – 10 m height, 10 – 30 m height and greater than 30 m height). As the typical forage range for honey bees with moderate nectar availability is 1 – 2 km [18], the AusCover mean mature tree canopy cover data from this process were averaged to the nearest 1 km MODIS pixel extents to assist with assessing the relative impact of changes in vegetation phenology per pixel. Examples of this process are shown in Figure 4.

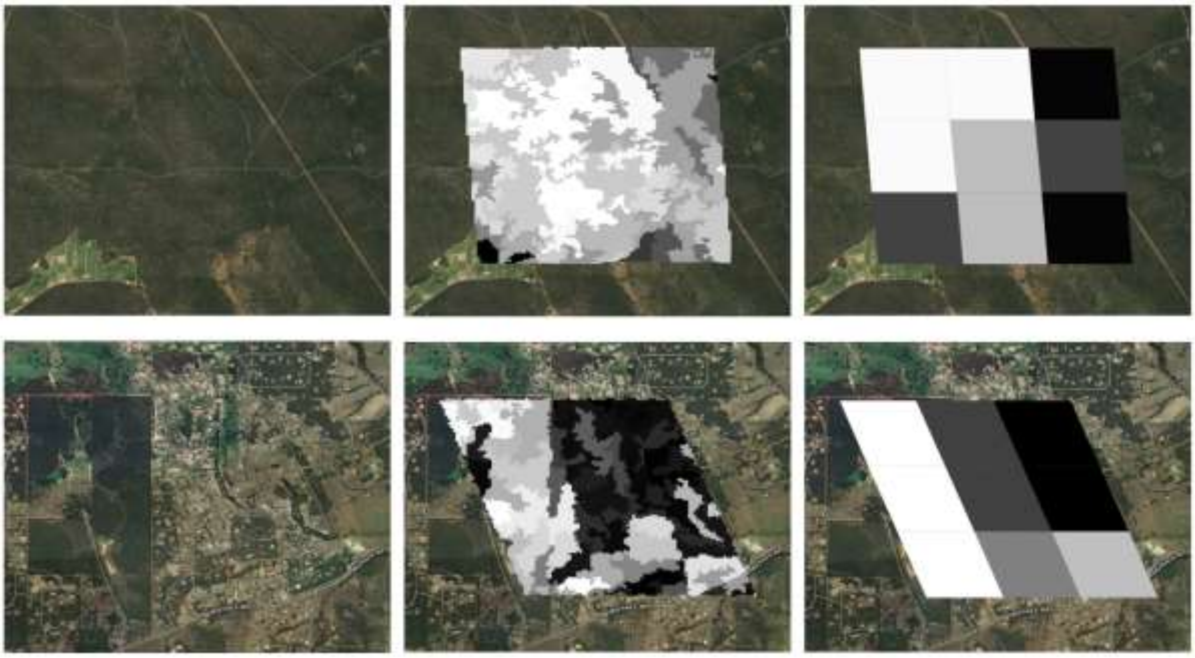

Figure 4: Comparison of aerial imagery (left panels) with AusCover vegetation height data at ~ 30 m spatial resolution (middle panels) and AusCover mean per (1km) MODIS pixel (right panels) for forest (top row) and mixed land use (bottom row) apiary locations.

2.2 MODIS derived indices and honey yields

Satellite-derived data from the MODIS sensor were used in this study in preference to other sensors with higher spatial resolution (e.g. Landsat or Sentinel-2) for the following reasons:

- The narrower spectral bands for MODIS compared with Landsat and their specific wavelengths result in a significantly higher spectral separation for marri flowers [13].
- The daily temporal resolution of MODIS, compared with the 16-day temporal resolution of Landsat, is more suited to accurately measuring maximum values over a 1-month period.
- While Sentinel 2 has a better temporal resolution than Landsat (5 days versus 16 days), this data is only available from after March 2017. This gives two years of Marri harvest data to analyse.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

Previous studies correlating MODIS NDVI and EVI to honey harvests in the southeast Australian tree *Eucalyptus tricarpa* (ironbark) showed a degree of qualitative correlation between vegetation indices peaks in winter and the peak quarterly honey production for the same site [8]. However, no quantitative assessment was performed. In addition, marri trees are a late summer to early autumn flowering species (the annual low period for NDVI in Australia). It was expected that seasonal variations in vegetation indices would be larger than spectral variations due to flower coverage, as demonstrated by Campbell and Fearns [13], who estimated that NDVI and EVI require minimum canopy cover of 50% and 60% respectively for changes in flower coverage to be detectable. Even at 100% canopy cover, it was estimated in this previous study that these two indices still require 15% and 44% flower coverage of the pixel respectively for detection of flowers to be viable.

Over the apiary sites used, the average canopy cover was only 22%. Accordingly, three sites with the highest canopy coverage and years of honey harvest data were initially selected to test the ability of NDVI and EVI to detect marri flowers. These three sites had canopy coverages ranging from 38% to 53%. The remaining 13 sites with lower canopy coverage were then assessed.

Maximum monthly NDVI and EVI data for these sites were calculated from the MODIS 1 km pixel size database (as an indication of the highest level of canopy chlorophyll for the month [22]).These maximum values were extracted from the same MODIS pixels used to calculate the Marri Flowering Index (MFI), with the MFI calculated by dividing the value of MODIS Band 10 by Band 8, and as described in Campbell and Fearns [13].

2.3 Weather data

To achieve a greater understanding of the causes of the variations in honey harvest weight, average maximum temperature and monthly rainfall data for January and February (the month immediately preceding the main flowering period and the month of the main flowering period [21]) were obtained from the Australian Government's Bureau of Meteorology (BOM) website (http://www.bom.gov.au/climate/data/) for the weather stations closest to the apiary sites. Note that the distance between the source of weather data and harvest data was up to 40 km in some cases as apiary sites can be situated in remote, unpopulated locations.

These weather data were plotted both against the honey harvest weight data for each site and year, as well as plotted against each other, with points coloured by harvest quality to locate multiple weather factors that work together to affect the honey harvest quality.

3. RESULTS

3.1 MODIS derived indices and honey yields

With the increase in flower coverage over the flowering period resulting in less photosynthetically active vegetation at the top of the canopy, it would be expected that if the vegetation indices were directly detecting the presence of flowers then the vegetation index would decrease due to vegetation being occluded by flowers (as per spectral separation analysis by Sulik and Long [9] and Campbell and Fearns [13]). However, as the February NDVI data in Figure 5 show, there is a weak positive correlation between February maximum NDVI and marri honey harvest weight (R = 0.73).

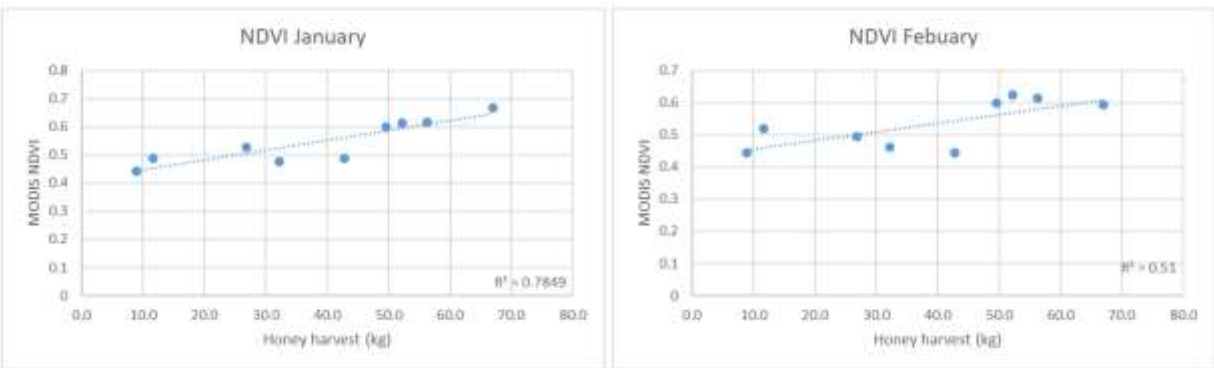

Figure 5: MODIS NDVI versus annual marri honey yield

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

The correlation between NDVI and honey harvest is stronger in January (R = 0.88) with increased chlorophyll content, as measured by the NDVI data, linked to more active green floral bud formation in the pre-flowering stage with resultant higher flowering, nectar production and concomitant increase in honey harvest. Marri buds are held in terminal corymbs and can be visible as a lighter green 'canopy' over the tree compared with marri leaves.

The January data shows that data points can be divided into two parts based on marri honey harvests above or below 45 kg per hive for the season. Below this harvest level there is a very poor correlation between honey harvest and NDVI (R = 0.40). Above this harvest (the four largest honey harvests), there is a stronger linear correlation between honey harvest and the January maximum NDVI (R = 0.99) (Figure 6). It is possible that for lower nectar production years, the variation in chlorophyll content is not sufficient to be measurably different by MODIS NDVI. For years with exceptionally high nectar production, the chlorophyll rises above the detection threshold and this linear relationship is measurable.

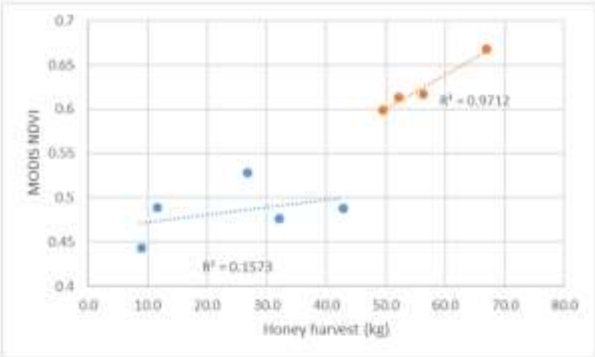

Figure 6: Two zones of different relationships between January NDVI and annual marri honey yield

The data analysis was then extended to include all honey harvest data points, regardless of degree of canopy cover. The result, shown in Figure 7, indicates a weak positive correlation between January NDVI and honey harvest (R = 0.46). The box plot of the data in Figure 7 does show that 'good' harvest years (honey harvest weight > 40 kg) have generally higher January NDVI. This was confirmed by an ANOVA analysis, which showed that there is no significant difference between the January MODIS NDVI data for poor (< 20 kg) and moderate harvest amounts (20 – 40 kg) whereas good harvests were separable from poor and moderate harvests (ρ<0.05 for good vs moderate and good vs poor years but not moderate vs poor years). Note that these ranges for poor, moderate and good harvest years are based on the average weight of honey harvested per 'box', or full-depth super (the upper boxes in a hive where honey is deposited), of honey (average weight of 20 kg per 'box' [16]).

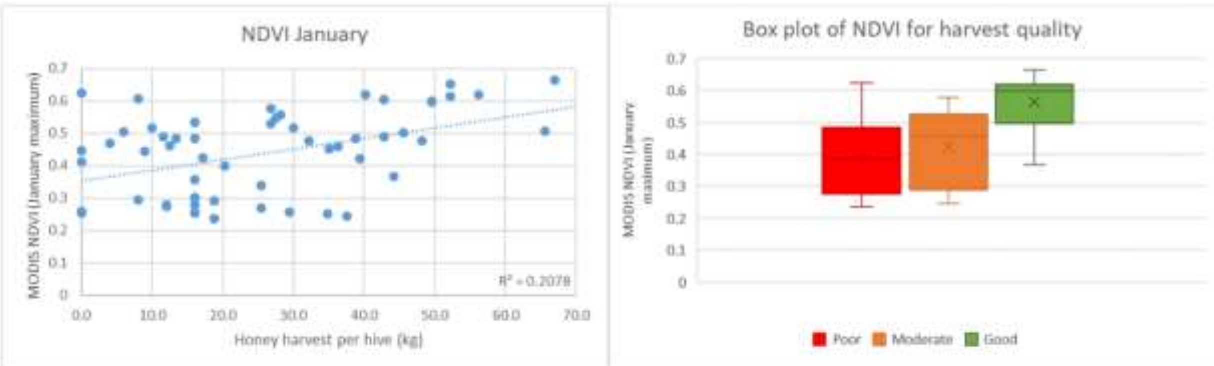

Figure 7: Maximum January MODIS NDVI for all honey harvest data points from *Error! Reference source not found.*

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

To assess the effect of site or year specific effects, the data in Figure 7 were grouped into apiary sites and years (Figure 8 and Table 2). These groups show that there is significantly better relationship between harvest volume and NDVI by year than by site (median R^2 0.75 by year vs 0.13 by site), and that the positive relationships dominate by year as the harvest volume increases (there are no negative correlations above 45 kg per hive harvest weight). However, the majority of the relationship between NDVI and honey harvest by year is likely to be a result of increasing NDVI with increasing canopy cover rather than from the honey harvest itself (see Figure 9).

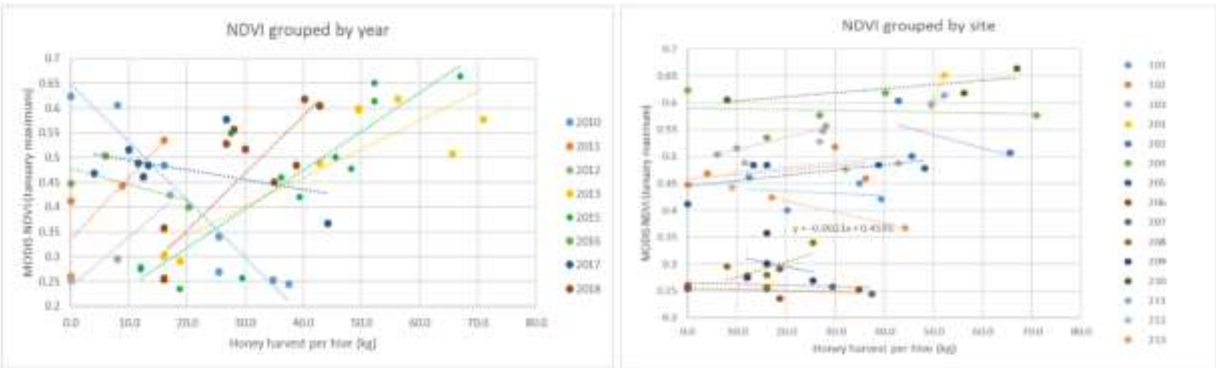

Figure 8: Maximum January NDVI vs honey harvest per hive, grouped into years (left frame) and apiary sites (right frame) with linear regression trendlines shown

Table 2: NDVI by year and site linear regression summary

YEAR	R^2	GRADIENT
2010	0.93	-0.012
2011	0.71	0.012
2012	0.94	0.010
2013	0.78	0.006
2015	0.18	-0.002
2016	0.42	-0.003
2017	0.75	0.008
2018	0.75	0.012
MEDIAN	***0.75***	***0.007***

SITE	R^2	GRADIENT
101	0.04	-0.0004
102	0.23	0.0008
103	0.98	0.0022
201	N/A	0.0208
202	0.03	-0.0025
204	0.03	-0.0002
205	0.37	0.0010
206	0.10	-0.0002
207	0.03	-0.0002
208	0.32	0.0028
209	0.07	0.0019
210	0.61	0.0008
211	N/A	0.0056
212	0.16	0.0008
213	N/A	-0.0021
MEDIAN	***0.13***	***0.0008***

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

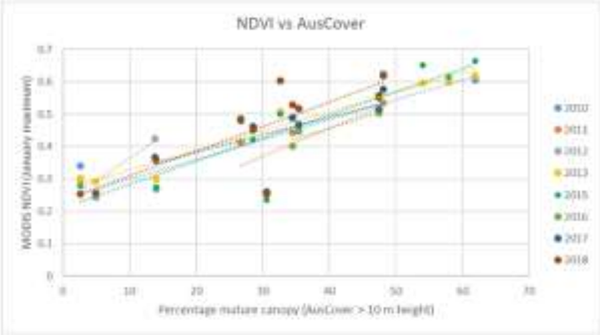

Figure 9: January maximum NDVI versus mature canopy cover (AusCover > 10 m height)

As MFI and honey harvest weight were reported by Campbell and Fearns [13] as being linked, the MFI was compared to the honey harvest data used for NDVI and EVI evaluation. A plot of MFI vs honey harvest data (Figure 10a) shows no correlation across the full range of harvest weights.

Given the range of canopy coverage as derived from AusCover data (2.6% – 61.9%), the MFI was multiplied by the canopy coverage in an effort to enhance the effect that canopy cover has on the index (Figure 10b). This scaled MFI (sMFI) showed a similar result to the maximum NDVI in January, with low correlation for poor to moderate years (<40 kg per hive) and generally higher sMFI for good years (>40 kg per hive). This is more apparent in the box plots in Figure 11, with the sMFI for good years being noticeably higher. However, an ANOVA analysis of the sMFI for the three different groups of harvest weight only yielded a value of $\rho < 0.05$ for moderate vs good years (the NDVI ANOVA analysis gave this result for both poor vs good and moderate vs good years).

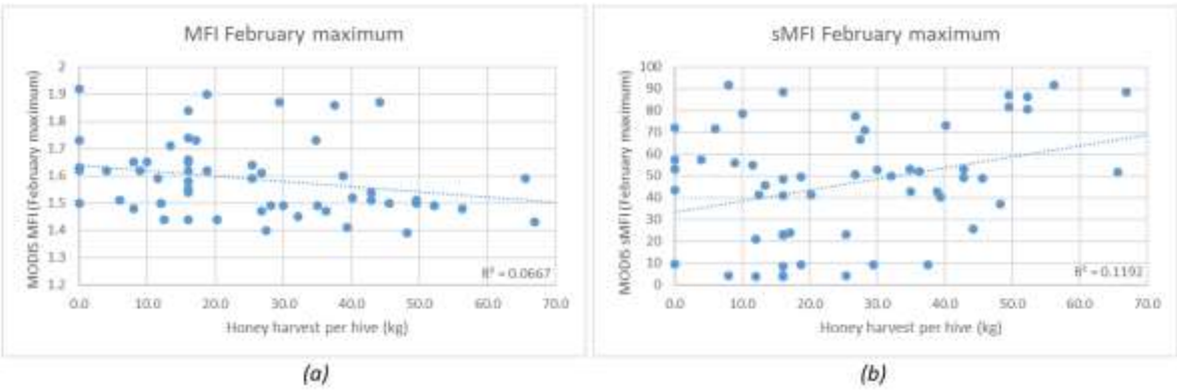

Figure 10: Marri Flowering Index (MFI) and scaled Marri Flowering Index (sMFI) vs honey harvest weight

Figure 11: Box plots of MFI and sMFI for different honey harvest weights

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

The assessment of linear regression for sMFI when grouped by year and site (Figure 12 and Table 3) again showed poor correlation between sMFI and honey harvest weight. Indeed, the median gradient for sMFI grouped by site in Figure 12 and Table 3 is negative, indicating that there is a negative correlation between sMFI and honey harvest weight. It may be that MFI can be effective in some cases where there is a high proportion of canopy cover (as found by Campbell and Fearns [13]). However the signal may too close to the noise floor for the MODIS data to be used reliably across areas of differing canopy covers, especially those areas with a lower proportion of canopy cover.

Figure 12: Maximum February sMFI grouped into years and apiary sites with linear regression trendlines

Table 3: sMFI by year and site linear regression summary

YEAR	R^2	GRADIENT
2010	0.51	1.69
2011	0.01	-0.27
2012	0.06	0.18
2013	0.62	1.26
2015	0.12	0.65
2016	0.06	0.18
2017	0.35	-0.79
2018	0.05	0.66
MEDIAN	*0.09*	*0.42*

SITE	R^2	GRADIENT
101	0.03	0.056
102	0.23	-0.124
103	0.98	-0.368
201	N/A	-2.907
202	0.28	0.267
204	0.01	0.016
205	0.38	-0.113
206	0.09	0.113
207	0.04	0.018
208	0.25	-0.074
209	0.09	0.301
210	0.63	-0.119
211	N/A	-0.878
212	0.18	-0.134
213	N/A	-0.185
MEDIAN	*0.21*	*-0.113*

3.2 Relationships between honey yields and weather data

Honey harvest weight versus weather data are shown in Figure 13, with broad trends indicated but there is still overlap between each weather metric and the harvest quality categories (poor < 20 kg, moderate 20 – 40 kg and good > 40 kg). This is borne out by the ANOVA analysis of the data, which is summarised in Table 4. The only weather data that has a statistically significant difference is the January median maximum temperature (i.e. the median of the daily maximum temperatures in January). For this dataset, good harvest years do not exceed 31°C degrees whereas poor to moderate harvest years range from 30.5°C to 37.0°C. While there are other clusters in the weather data (circled in red in Figure 13: higher January rainfall for moderate harvests and higher February rainfall for poor harvests), these only apply to some data points for a given harvest category and therefore are not the only factor influencing harvest quality.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Table 4: Monthly weather vs honey harvest ANOVA ρ-value results. Note that significant results (where ρ < 0.05) are highlighted in green.

COMPARISON	JANUARY TEMPERATURE	JANUARY RAINFALL	FEBRUARY TEMPERATURE	FEBRUARY RAINFALL
POOR VS MOD	$\rho > 0.05$	$\rho > 0.05$	$\rho > 0.05$	$\rho > 0.05$
POOR VS GOOD	$\rho < 0.05$	$\rho > 0.05$	$\rho > 0.05$	$\rho > 0.05$
MOD VS GOOD	$\rho < 0.05$	$\rho > 0.05$	$\rho > 0.05$	$\rho > 0.05$

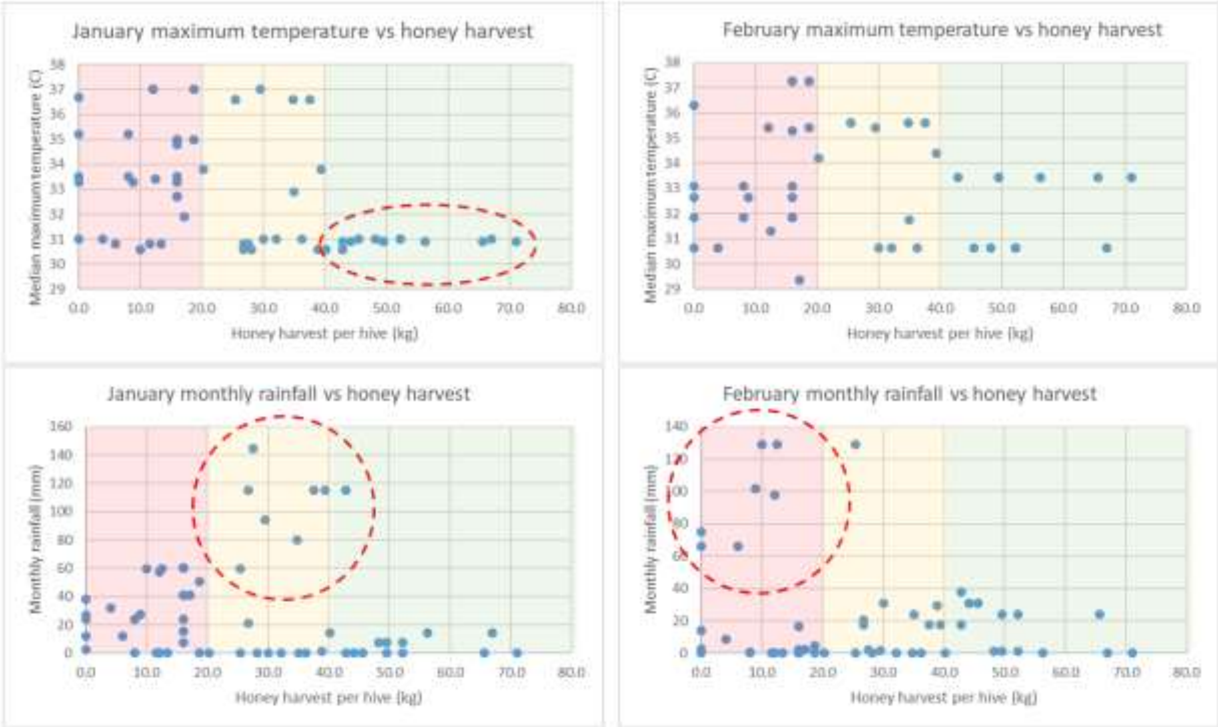

Figure 13: Monthly weather data vs honey harvest weight. Each data point is for a harvest site. Background is coloured by poor, moderate and good harvest classifications (red, yellow and green respectively)

To assess the effect of multiple weather and vegetation index variables on harvest quality, a series of cross-plots were produced, with the harvest quality grouped by colour. The results, shown in Figure 14, show the potential for combinations of measurements to improve on Boolean classification results compared to single variables alone (see combinations indicated in Figure 14). Interestingly, there are clusters and trends for the good and moderate harvest quality categories, but the poor harvest quality category is spread across the full range of results for all metrics except January rainfall. This indicates that there are a number of conditions that can cause harvest of poor quality, but only a few conditions that may support moderate or good quality harvests.

4. DISCUSSION

4.1 MODIS derived indices and honey yields

The correlations between NDVI and honey yield show that the increased vegetation activity associated with increased nectar production has a larger effect than flower coverage alone (i.e. direct detection of increased flower coverage with the NDVI data is not feasible within this dataset). Comparison of the ANOVA analysis of NDVI and scaled Marri Flowering Index (sMFI) compared to honey harvests showed that January NDVI is a more reliable indicator of honey harvest weight than the sMFI.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

To assess the accuracy of an NDVI cutoff point in classifying a year as good vs moderate or poor harvest quality, the percentages of years classified correctly for NDVI ranging between 0.4 and 0.6 were calculated. A cutoff at an NDVI of 0.5 was found to be the most accurate, with 79% of the 'good' years being above this point and 77% of the other years being below. Thus it is possible to predict honey harvest from satellite remotely sensed data, although the prediction is not reliable across all canopy covers nor harvest qualities.

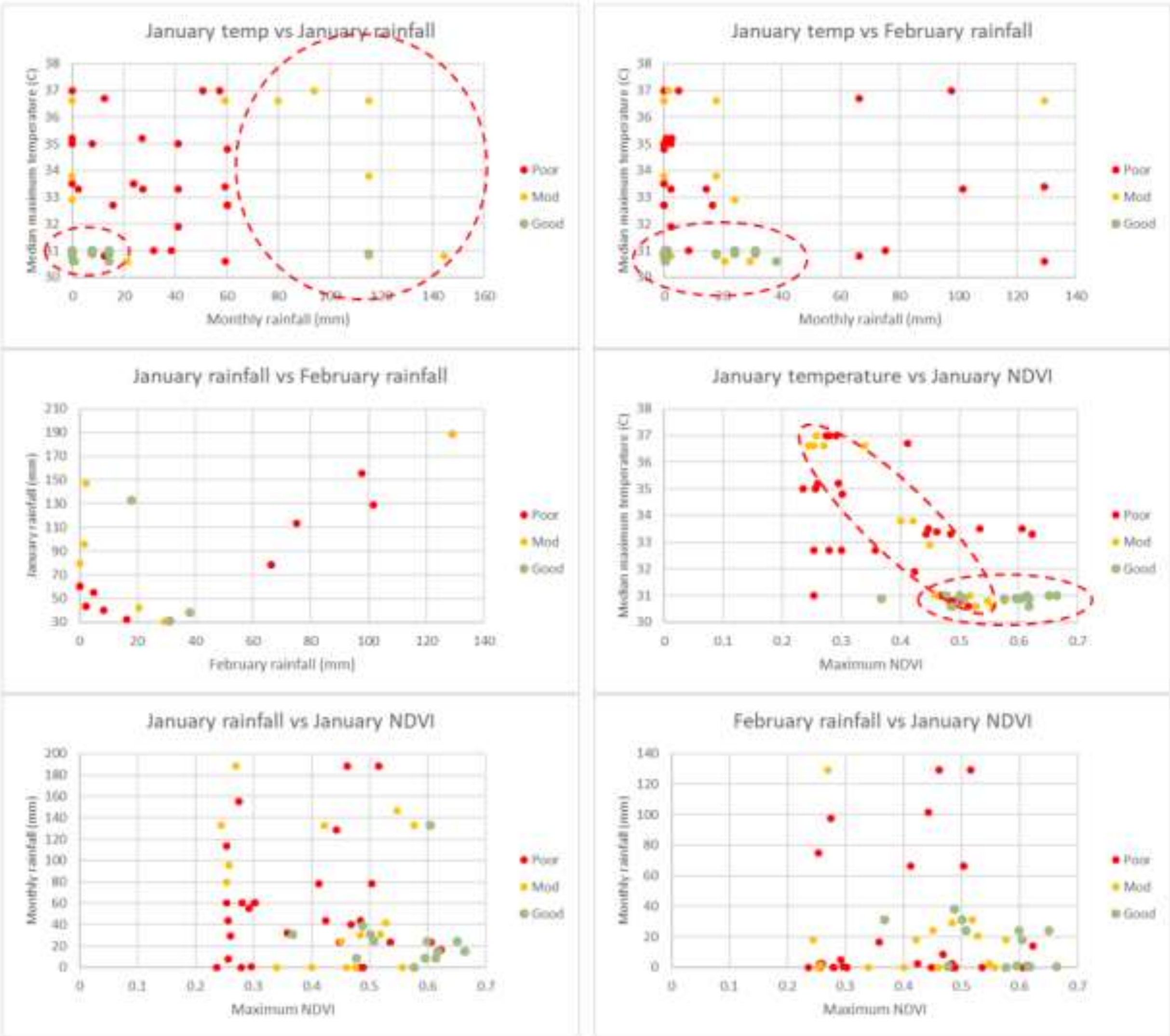

Figure 14: Cross-plots of weather and vegetation index data vs harvest quality. Red dashed lines highlight areas where good or moderate harvest datapoints cluster together well.

4.2 Relationships between honey yields and weather

The effectiveness of using multiple weather metrics to assign a honey yield quality was quantitatively tested by calculating the classification accuracy of different combinations of metrics, with the highest preceding this being a 78% accuracy in classifying good quality harvests using January maximum NDVI (note that moderate and poor quality harvest years were unable to be separated using NDVI). As summarised in Table 5, greater predictive accuracy is obtained using a combination of NDVI, median maximum temperature and monthly rainfall for January to classify a harvest as either good quality or of lower quality (accuracy of 90%). Of the 58 datapoints, two moderate quality years are classified as good years and four good years are classed as moderate

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

years. Adjusting any of the three criteria to improve the number of good years that are correctly classified also increases the number of poor and moderate harvest years within the criteria, reducing the overall accuracy.

After classification of good quality harvests, the next most accurate prediction is classification of harvests better than poor quality (i.e. moderate or good quality years), with at an accuracy of 72%. This uses a combination of the good quality harvest classification criteria as well as a set range of January temperature to January NDVI ratios for moderate years (see the moderate harvest quality range indicated by the red oval in Figure 14). As seen in Figure 13 and Figure 14, there is significant overlap between poor and moderate quality harvest years in the NDVI and weather metrics that is a limiting factor in the classification accuracy.

Moderate harvest quality years and poor quality harvest years were classified to 66% and 69% accuracies respectively, using a combination of the good harvest quality criteria and higher quality indices than poor harvest. As discussed above, there is significant overlap with the metrics used for poor harvest quality years and moderate harvest quality years, that limits the effectiveness of a Boolean classification approach. It may be that more detailed analysis of daily temperature, rainfall and/or NDVI data over the pre-flowering period may achieve better results than the monthly figures.

Attempts to create a combined classification approach for all three harvest quality levels yielded at best a classification accuracy of 69%, mostly due to the difficulty of separating moderate and poor harvest years.

The outcome of the inclusion of the weather data into the predictive model shows that the honey harvest weight prediction model based on satellite remotely sensed data improved the model's predictive capability. As only monthly weather data were used, it may be that a more detailed analysis of daily weather data may achieve greater accuracies for these poor and moderate harvest quality years. Greater predictability may also be achieved if apiary variables were incorporated in future models (such as queen bee quality and an index of hive strength as proposed by Delaplane, Van Der Steen [17]).

Table 5: Multivariate classification accuracy summary

TARGET	**METRIC(S)**	**ACCURACY**	**TARGET**	**METRIC(S)**	**ACCURACY**
GOOD HARVESTS	JAN NDVI > 0.5	79%	MODERATE HARVESTS	JAN NDVI > (-0.042*JAN TEMP+1.75)	45%
	JAN NDVI > 0.5 *AND IF* JAN TEMP < 31.5	83%		JAN NDVI < (-0.042*JAN TEMP+1.75) *AND IF* JAN NDVI > (-0.042*JAN TEMP+1.90) *AND IF* JAN NDVI < 0.5 *AND IF* JAN TEMP > 31.5 *AND IF* JAN RAIN > 40	66%
	JAN NDVI > 0.5 *AND IF* JAN TEMP < 31.5 *AND IF* JAN RAIN < 40	90%			
TARGET	**METRIC(S)**	**ACCURACY**	**TARGET**	**METRIC(S)**	**ACCURACY**
> POOR HARVEST	JAN NDVI > 0.5 *AND IF* JAN TEMP < 31.5 *AND IF* JAN RAIN < 40 *AND IF* JAN NDVI < (-0.042*JAN TEMP+1.75)	72%	POOR HARVEST	JAN NDVI > 0.5 *OR IF* JAN TEMP < 31.5 *OR IF* JAN RAIN < 40	69%

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

	AND IF JAN NDVI > (-0.042*JAN TEMP+1.90)				

5. CONCLUSIONS

Building on previous work in Europe and south-east Australia correlating satellite-derived vegetation indices to honey yields (summarised in the introduction sector of this paper), an assessment of NDVI versus honey harvest weight found a strong correlation between higher January MODIS NDVI and good quality marri honey harvests (> 40 kg of honey per hive). Using a Boolean classification approach with a criterion of maximum January NDVI > 0.5, and honey harvest data from 2010 – 2018 in southwest Australia across 16 wide-ranging apiary sites were classified correctly with 78% accuracy.

Incorporating climatic variables, particularly January monthly rainfall and median maximum daily temperature, with NDVI improved the classification of good quality harvest years to 90%. However, moderate and poor harvest quality years had a larger overlap of the weather and NDVI data for compared with good quality harvest years.

The results of this study will allow beekeepers to more accurately predict honey harvest quality prior to the start of honey production and adjust their preparation and site selection accordingly.

The approach used here in developing the predictive model can be readily applied to honey harvests from other floral sources, allowing beekeepers from other regions or countries to develop models suited to their own particular floral assemblages and climates.

6. ACKNOWLEDGEMENTS

This research is supported by the Beekeeping Industry Council of Western Australia (BICWA), the Western Australian government's Department of Primary Industries and Regional Development (DPIRD) and ChemCentre as part of the Grower Groups Research and Development Grant - Round 2: GGRD2 2016-1700179 (Industry Standards Optimising Storage And Supply Volume Of WA Mono-Floral Honey). The authors also wish to thank all Apiaries and Apiarists for use of their honey yield data and apiary sites for monitoring purposes.

7. AUTHOR CONTRIBUTIONS

Tristan Campbell and Peter Fearns conceived and designed the investigation. Tristan Campbell conducted the investigation and wrote the paper. Peter Fearns supervised the study and reviewed the paper. Kenneth Dods and Kingsley Dixon provided key input into the paper. Kenneth Dods acquired financial support for the investigation.

8. REFERENCES

1. Thomson, J., *Western Australia a sweet spot for beekeeping*. 2019, Department of Primary Industries and Regional Development.
2. Irish, J., S. Blair, and D. Carter, *The Antibacterial Activity of Honey Derived from Australian Flora.* PLoS ONE, 2011. **6**(3).
3. Hastie, H., *Sweet as: The rise and rise of West Australian honey*, in *The West Australian*. 2018: Perth, Western Australia. p. 1.
4. Painter, S., *Jarrah honey crisis as yield wiped out*, in *The West Australian*. 2010: Perth.
5. Blomstedt, W., *Mapping The Phenology of European Honey Bee Nectar Flows*, in *School of Geosciences*. 2014, University of Edinburgh: Scotland.
6. Webber, E., *Eucalypt leaf-flush detection from remotely sensed (MODIS) data*, in *Department of Infrastructure Engineering-Geomatics*. 2011, University of Melbourne.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

7. Winter, S., et al., *BeeBox Application User Manual*, R.I.R.a.D. Corporation, Editor. 2013.
8. Arundel, J., et al., *A web-based application for beekeepers to visualise patterns of growth in floral resources using MODIS data.* Environmental Modelling & Software, 2016. **83**: p. 116 - 125.
9. Sulik, J. and D. Long, *Spectral indices for yellow canola flowers.* International Journal of Remote Sensing, 2015. **36**(10): p. 2751-2765.
10. Sulik, J. and D. Long, *Spectral considerations for modeling yield of canola.* Remote Sensing of Environment, 2016. **184**: p. 161 - 174.
11. Landmann, T., et al., *Application of hyperspectral remote sensing for flower mapping in African savannas.* Remote Sensing of Environment, 2015. **166**: p. 50 - 60.
12. Campbell, T. and P. Fearns, *Simple remote sensing detection of Corymbia calophylla Flowers using common 3 –band imaging sensors.* Remote Sensing Applications: Society and Environment, 2018. **11**: p. 51-63.
13. Campbell, T. and P. Fearns, *Honey crop estimation from space: Detection of large flowering events in Western Australian forests*, in *ISPRS TC I Mid-term Symposium "Innovative Sensing – From Sensors to Methods and Applications"*. 2018, International Society for Photogrammetry and Remote Sensing: Karlsruhe, Germany. p. 79 - 86.
14. Campbell, T., *Ground-based spectroradiometer measurements of vegetation and groundcover of Corymbia calopylla forests in Western Australia*. 2019, Curtin University.
15. Herbarium, W.A., *Florabase - the Western Australian Flora*, D.o.P.a. Wildlife, Editor. 1998.
16. Leyland, D., *Review of historic marri harvest records*. 2015, Personnal communication: Chidlow, Western Australia.
17. Delaplane, K.S., J. Van Der Steen, and E. Guzman-Novoa, *Standard methods for estimating strength parameters of Apis mellifera colonies*. 2013, Taylor & Francis. p. 1-12.
18. Hagler, J., et al., *Foraging range of honey bees, Apis mellifera, in alfalfa seed production fields.* Journal of Insect Science, 2011. **11**(144).
19. Scarth, P., *Vegetation height and structure - derived from ALOS-1 PALSAR, Landsat and ICESat/GLAS, Australia coverage*, J.R.S.R. Project, Editor. 2009: Australia.
20. Guerschman, J.P., et al., *Evaluation of the MODIS-based vegetation fractional cover product*, CSIRO, Editor. 2012: New South Wales, Australia.
21. Brooker, M.I.H. and D.A. Kleinig, *Field guide to eucalypts.* Vol. Volme 2, South-western and southern Australia. 2001, Melbourne, Australia: Bloomings Books. 428.
22. Gitelson, A.A. and M.N. Merzlyak, *Remote estimation of chlorophyll content in higher plant leaves.* International Journal of Remote Sensing, 1997. **18**(12): p. 2691-2697.

IJESRT is licensed under a Creative Commons Attribution 4.0 International License.