Optical Sensing of Vegetation Water Content: A Synthesis Study

Ying Gao¹, Jeffrey P. Walker¹, Mahdi Allahmoradi², Alessandra Monerri¹, Dongryeol

Ryu², Thomas Jackson³

¹Department of Civil Engineering, Monash University, Australia

²Department of Infrastructure Engineering, University of Melbourne, Australia

³Hydrology and Remote Sensing Lab, USDA ARS

Journal: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

2nd Revision submitted on 17 Oct 2014

Corresponding author:

Ying Gao

Room 156, Bld 60,

Dept. of Civil Engineering,

Monash University,

Clayton, VIC 3800

Email: ying.gao@monash.edu

Phone: +61 3 9905 4957

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

22 **Abstract**

23 Vegetation Water Content (VWC) plays an important role in parameterizing the vegetation
24 influence on microwave soil moisture retrieval. During the past decade relationships have
25 been developed between VWC and vegetation indices from satellite optical sensors, in order
26 to create large-scale VWC maps based on these relationships. Among existing vegetation
27 indices, the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference
28 Water Index (NDWI) have been most frequently used for estimating VWC. This work
29 compiles and inter-compares a number of equations developed for VWC derivation from
30 NDVI and NDWI using satellite data and ground samples collected from field campaigns
31 carried out in the United States, Australia and China. Four vegetation types are considered: a)
32 corn; b) cereal grains; c) legumes and d) grassland. While existing equations are reassessed
33 against the entire compiled data sets, new equations are also developed based on the entire
34 data sets. Comparing with existing equations, results show superiorities for the new equations
35 based on statistical analysis against the entire data set. $NDWI_{1640}$ and NDVI are found to be
36 the preferred indices for VWC estimation based on the availability and the error statistics of
37 the compiled data sets. It is recommended that the new equations can be applied in the future
38 global remote sensing application for VWC map retrieval.

39

40 **Keywords** - Vegetation Water Content (VWC), Normalized Difference Vegetation Index
41 (NDVI), Normalized Difference Water Index (NDWI)

42

1. Introduction

Over the past three decades it has been shown that the Vegetation Water Content (VWC) is an important variable in climatic, agricultural and forestry applications [1-4]. In passive microwave remote sensing, a vegetation canopy over the soil absorbs the emission of the soil and adds to the total radiative flux with its own emission. With an estimate of VWC, the vegetation optical depth and transmissivity can be modelled [5]. Thus VWC plays a particularly important role in soil moisture retrieval by parameterizing the effects of vegetation on the observed land surface emission.

Spatially distributed VWC information over large regions is not readily available. One approach is to use relationships with spectral reflectance measured by optical satellites with an appropriate function in order to map VWC (e.g. [4, 6-9]). These functions have been developed using relationships between the remotely sensed indices available from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors (with 16 day repeat at 30m resolution), or the MODerate resolution Imaging Spectroradiometer (MODIS) (with daily repeat at 250m resolution), together with ground-based spectral and VWC measurements.

The Normalized Difference Vegetation Index (NDVI) proposed by Rouse et al. [10] for estimating VWC is one of the most widely used indices:

$$NDVI = \frac{NIR_{860} - RED_{650}}{NIR_{860} + RED_{650}}, \quad \text{Eq. 1}$$

where NIR is the reflectance in the near infrared channel (centred at 860 nm) and RED is the reflectance in the red band visible (VIS) channel (centred at 650 nm). A drawback of using NDVI for this application is that it saturates when vegetation coverage become dense (when leaf area index reach around 5 [4, 11]) and is no longer sensitive to changes in vegetation.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

66 The saturation of NDVI was also observed by Chen et al. [6] for $VWC > 3\text{kg/m}^2$ for corn.
67 Moreover, RED and NIR are located respectively in the strong chlorophyll absorption region
68 and the high reflectance plateau of vegetation canopies, meaning that NDVI represents
69 chlorophyll rather than water content [6, 12]. Nevertheless, Jackson et al. [4] suggested that
70 for specific canopy types (such as grasslands) within specific regions and when supported by
71 ground sampling, it is still possible to establish useful VWC functions based on NDVI.

72 The Normalized Difference Water Index (NDWI), which utilizes the shortwave infrared
73 (SWIR) together with NIR, has been shown to have a better correlation with leaf water
74 content than the vegetation indices employing VIS and NIR [6, 12]. Compared to NDVI, it
75 has been found that the saturation of this SWIR-based spectral index occurs later [6, 13]. The
76 NDWI proposed by Gao [12] used a SWIR band centred at 1240 nm. This wavelength
77 became available with the launch of MODIS. Previous to this the SWIR bands at 1640 nm
78 and 2130 nm, which are available from Landsat, had been used to demonstrate that the water
79 absorption was dominant and thus sensitive to VWC variations [4, 6]. Therefore, the
80 following NDWI indices are also considered in this work:

81
$$NDWI_{1240} = \frac{NIR_{860} - SWIR_{1240}}{NIR_{860} + SWIR_{1240}} \quad \text{Eq. 2}$$

82
$$NDWI_{1640} = \frac{NIR_{860} - SWIR_{1640}}{NIR_{860} + SWIR_{1640}} \quad \text{Eq. 3}$$

83
$$NDWI_{2130} = \frac{NIR_{860} - SWIR_{2130}}{NIR_{860} + SWIR_{2130}}, \quad \text{Eq. 4}$$

84 where the subscript refers to the wavelength (nm).

85 Although many empirical relationships between VWC and the aforementioned vegetation
86 indices have been established for different vegetation categories and from different field
87 campaigns around the world, there has been no study to synthesize or inter-compare the data

and relationships derived from these different field campaigns, and to recommend a best relationship for global remote sensing applications, such as the Soil Moisture Active Passive (SMAP) satellite mission that needs a global VWC map as input for generating the soil moisture products. Currently, in order to obtain VWC information from optical sensing observations, many options are available as to which vegetation index and which model to apply based on the literature. Consequently, it is the intention of this investigation to synthesize the body of work available from literature and our own recently collected data sets into more robust models for VWC estimation. Statistical analysis is performed for both the new models and the existing models using the combined data sets, upon which a recommendation of vegetation index and model is made for both specific types of land cover and general categories.

2. Data Sources

Data from eight different studies [4, 6-9, 14-16] are analysed in this paper. These studies were chosen because 1) the vegetation indices they analysed were either NDVI or/and NDWI, which have been found to be the best for VWC estimation; and 2) the analysis was based on one or more specific land cover types and provided a vegetation type specific model to relate the index to VWC. The sources of VWC and vegetation index data provided in each study are summarized in Table 1. The data were from the following field campaigns: SMEX02 and SMEX05 in the U.S.A. [9, 15], NAFE'05 [17], NAFE'06 [18], AACES-1 and -2 [19], SMAPEX-1, -2 and -3 [20] in Australia, and the Weishan experiment [16] in China. The locations of these experiments are indicated in Fig.1.

The basic information of these field campaigns, including location, season, major crop types and ancillary data measured are summarized in Table 2. Most of the campaigns were conducted in spring or summer, except AACES-2 and SMAPEX-1, which were in winter. In

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

112 terms of crop types, the experiments in Australia had a more diverse range, including barley,
113 wheat, corn, lucerne and grasslands. For the two campaigns in the U.S.A., SMEX02 and 05,
114 the crop types included corn and soybean, being the only major crops in the experiment area.
115 While all the campaigns sampled VWC, Leaf Area Index (LAI) and dry biomass, however
116 ground-based surface reflectance was only measured in the NAFE, AACES and SMAPEX
117 campaigns. As a result, except for Maggioni et al. [8] and Allahmoradi et al. [14] which used
118 calculated vegetation indices from field spectrometer measurements, the rest of the studies
119 relied on either Landsat or MODIS to provide spectral data for calculation of the vegetation
120 indices.

121 Landsat 5 (TM sensor) and Landsat 7 (ETM+ sensor) have 8 frequency bands. Apart from
122 band 6 and band 8, which have a resolution of 60 m and 15 m respectively, all other bands
123 have a resolution of 30 m. In Eq. 1, RED and NIR correspond to band 3 (630-690 nm) and
124 band 4 (760-900 nm), respectively. For SWIR in Eq. 2-4, band 5 (1550-1750 nm) and band 7
125 (2080-2350 nm) are used to cover SWIR₁₆₄₀ and SWIR₂₁₃₀. SWIR₁₂₄₀ is not available from
126 Landsat. Moreover, because of the infrequent temporal coverage of TM and ETM+, it is
127 difficult to rely on them for estimating VWC for most applications [4]. However data from
128 MODIS on the Terra and Aqua satellites are available daily, and are free to access. The
129 resolution of MODIS is 250 m for bands 1 and 2 (centred at 648 and 858 nm), and 500 m for
130 bands 3-7 (centred at 470, 555, 1240, 1640 and 2130 nm). RED and NIR correspond to band
131 1 and band 2 respectively, while SWIR₁₂₄₀, SWIR₁₆₄₀ and SWIR₂₁₃₀ correspond to band 5, 6
132 and 7 respectively. A summary of the spectral wavelengths used by the hand spectrometers
133 for the field campaigns considered in this study and their associated satellite bands for
134 calculating vegetation indices is presented in Table 3. For more details on the satellite data
135 processing please refer to the original publications listed in Table 1.

3. Methodology

Existing equations for NDVI and NDWI are summarized in Table 4. Lucerne in Allahmoradi et al. [14] is grouped with soybean in a category referred to as legumes, due to their similar spectral behaviour. In addition to the equations, the data series of sampled VWC and calculated vegetation indices have also been digitized from their original graphs and replotted in Figs. 2-5, according to the category of vegetation type and vegetation index. The red-dotted lines indicate the newly established equation based on all the available data sets. It should be noted that the equations and data sets from Huang et al. [15] are not included in the NDVI and NDWI₁₆₄₀ plots for corn and soybean, since the same SMEX02 data sets as Chen et al. [6] were used.

A recommended function is provided for the categories where multiple data sets are present (Table 4). These functions were developed based on all the available data sets for a certain category. For NDVI, exponential equations were chosen due to the notable upward trend which matches with the saturating behaviour of NDVI over the higher range of VWC. For the rest of vegetation indices, either linear or quadratic equations were provided. It should be noted that no recommended equation is given for NDVI₂₁₃₀ for corn, because the two available studies applied the same data set but with different source of spectral data. Also, for those categories with only one data set available (NDVI₁₂₄₀ for cereal grains and grassland, and NDVI₂₁₃₀ for legumes), the recommended equation would be the same as the one developed from its original study.

Statistical analysis is carried out to assess the correlation and VWC retrieval performance of all equations. Since R^2 was provided with most existing equations, they are directly quoted here in Table 4. However, not all studies gave RMSE as the VWC retrieval error. Therefore RMSE is calculated here for all existing equations, based on their digitized data sets, both

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

160 against their own data sets and against the entire synthesized data sets for each vegetation
161 category (Table 4).

162 **4. Data Comparisons**

163 4.1 NDVI

164 It can be seen in Fig. 2a that both the data and the equations from Jackson et al. [4] and Chen
165 et al. [6] agree well for corn, especially in the higher VWC range (3-5 kg/m²). In comparison,
166 the data from Allahmoradi et al. [14] are focused on a lower range of VWC (1-2 kg/m²) and a
167 limited number of samples were used in its equation derivation. However, these data still fall
168 approximately into the range of the data from [4] and [6]. It is also clear that NDVI becomes
169 saturated for VWC above about 3 kg/m², which is consistent with most previous studies (eg.
170 [4, 6, 12].

171 For cereal grains (Fig. 2b), Allahmoradi et al. [14] had a greater number of samples,
172 including barley, wheat and oats. While the winter wheat data sets from Yi et al. [16] agree
173 with the data from [14] in the lower range of VWC (<1.5 kg/m²), the VWC of winter wheat
174 reached to 3-4 kg/m² with an NDVI of 0.6-0.8, making it significantly higher compared with
175 [14] (0.5-2.5 kg/m²) for the same NDVI range. To explain this, [16] pointed out that there
176 were significant solar and zenith angular effects on the surface reflectance data from MODIS
177 after the wheat heading stage, meaning that NDVI would be unable to detect crop growth
178 during this phase. As a result, the data with high VWC values (circled by a red-dotted line in
179 Fig. 2b) from [16] are considered to be outliers, and not used in the subsequent analysis.

180 For legumes (Fig. 2c), the equations from Jackson et al. [4] and Allahmoradi et al. [14] are
181 similar to each other, as are the underlying data sets. For grassland (Fig. 2d), the equations
182 from Maggioni et al. [8] and [14] are the only ones available for estimating VWC. Although

the number of data points of [8] are very limited, they still fall into the same range as the data of [14].

4.2 NDWI₁₂₄₀

For the land cover categories of corn and legumes (Fig. 3a and Fig. 3c), only two studies are available for comparison: Huang et al. [15] and Allahmoradi et al. [14]. Although their NDWI was calculated from different sources, MODIS and field spectrometer MSR-16, the equations and underlying data sets match well with each other. This is because the MSR-16 was set to match with the MODIS bands during the NAFE and SMAPEX experiments. As noted previously, [14] is the only study to have used NDWI₁₂₄₀ to estimate VWC for both cereal grains and grassland (Fig. 3b and Fig. 3d). Thus until now the MODIS SWIR bands, especially at the 1240 nm recommended by Gao [12], have not been fully assessed and evaluated for estimating VWC.

4.3 NDWI₁₆₄₀

The most frequently used index for VWC estimation is NDWI₁₆₄₀. It is also the preferred index for estimating VWC, mainly because SWIR bands are sensitive to changes in water content of plant canopies, and SWIR₁₆₄₀ has been available on Landsat for many years. For corn (Fig. 4a), all studies obtained NDWI₁₆₄₀ from Landsat except for Allahmoradi et al. [14]. However, Chen et al. [6] applied both Landsat and MODIS data to calculate NDWI₁₆₄₀ and compared the two sets of data. Although only the Landsat data sets are included here (Fig. 4a), the analysis in [6] showed that the data sets derived from MODIS were similar to those derived from Landsat, but with a small shift. This shift could be due to that the centre wavelength of Landsat Band 5 being slightly higher than MODIS Band 6, which were used to calculate SWIR₁₆₄₀. It can be seen in Fig. 4a, that all equations and data sets match well.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The data sets for legumes (Fig. 4c) and grassland (Fig. 4d) also have a good agreement. For cereal grains (Fig. 4b), similar winter wheat outliers as those of the NDVI analysis can be observed. This is consistent with the previous discussion that the outliers could be due to the angular effects at late growth stage during the experiment period.

4.4 NDWI₂₁₃₀

The NDWI₂₁₃₀ index has not received as much attention in the literature as NDWI₁₆₄₀. However, it is also a valuable index in estimating VWC since it is available from both Landsat and MODIS. In Fig. 5a, both the VWC field data of Chen et al. [6] and Huang et al. [15] are from SMEX02 while the NDWI₂₁₃₀ were derived from MODIS and Landsat respectively. This graph confirms the phenomenon noted in [6]: that the data sets derived from MODIS are consistent with those derived from Landsat, but with a small shift (approximately 0.1-0.4 for NDWI) towards the left. This means that MODIS-derived NDWI is generally larger than the Landsat-derived value for the same type of vegetation in the same area. This is due to the larger centre wavelength of Landsat (Landsat Band 7 compared with MODIS Band 7 for calculating SWIR₂₁₃₀).

For the remaining categories, a separate calculation of NDWI₂₁₃₀ was performed using field data from the AACES campaigns because NDWI₂₁₃₀ was not considered in Allahmoradi et al. [14]. This is the only experiment that has NDWI₂₁₃₀ data available. For cereal grains (Fig. 5b) there were not enough data from AACES to establish an equation for barley and wheat. Similarly, for the studies conducted by Maggioni et al. [8] and Yi et al. [16], a limited number of samples were presented, although they still provided equations. However, it is suggested that the newly established equation based on the combined data sets from [8] and [16] should still be used with caution. Conversely, there are enough samples from AACES to

establish a relationship for grassland (Fig. 5d), with several samples from [8] also falling in the similar range.

5. Results and Discussion

The performance statistics of all equations, including R^2 and RMSE are listed in Table 4. Comparing the two RMSE values of the existing equations, the RMSE for the original data sets and RMSE for the combined data sets, it can be seen that the latter is generally much larger. This means that each of these equations may be representative for a specific data set at a specific location, but fail to capture well the conditions of other areas. Therefore the proposed new equations, with smaller error against the combined data sets, are expected to be more robust when used for VWC estimation globally, as required by satellite soil moisture missions.

Comparing the R^2 and RMSE of different indices for each type of land cover, the most suitable index for estimating VWC was identified for that specific land cover. As can be seen in Table 4, the recommended equation for $NDWI_{1640}$ performs the best in estimating VWC for corn, providing the highest R^2 (0.87) and the lowest RMSE (0.51 kg/m²). NDVI also works well for corn based on the large range of available data sets and the relatively high correlation ($R^2=0.8$). In the case of cereal grains, the recommended equation for $NDWI_{2130}$ performs the best in terms of R^2 (0.84), although the retrieval error is slightly higher than other indices (RMSE=0.55 kg/m² compared with 0.4~0.5 kg/m² for other indices). For legumes, $NDWI_{1240}$ and $NDWI_{1640}$ performed much better than the other two indices, both with an R^2 of 0.76 and a RMSE of around 0.2 kg/m². While for grassland NDVI worked the best according to its highest R^2 (0.52 compared with 0.2~0.4 for other indices), although all indices had a similar retrieval accuracy (RMSE $\approx$ 0.3 kg/m²).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Disregarding the vegetation types, the new equations for NDVI and NDWI₁₆₄₀ are considered to be best for VWC estimation in general at the current stage. This is because: 1) the amount of historical data for these two indices are larger and therefore allow a more reliable equation to be established; and 2) performance statistics show a better correlation for NDVI and NDWI₁₆₄₀ in general. There are at least three studies for NDVI for each land cover type, and as many as six studies for NDWI₁₆₄₀, while for NDWI₁₂₄₀ and NDWI₂₁₃₀ there are only one or two studies available. Amongst these, there is a preference for using NDVI, as the R^2 for all the NDVI equations are above 0.5, even for the highly scattered grassland data, while for NDWI₁₆₄₀ the R^2 ranges from 0.57 to as high as 0.87, but is only 0.2 for grassland. Moreover, since NDVI is readily available from MODIS satellite, it is more convenient for VWC retrieval than NDWI₁₆₄₀. Nevertheless, it should be noted that the model performance might vary over time or throughout the growing season of the crops. However, there are insufficient data sets to demonstrate this. Therefore long-term experiments are needed to address this issue.

An important consideration is the impact of VWC error on soil moisture retrieval accuracy. According to the analysis in Parinussa et al. [21], the higher the vegetation optical depth is, the greater the influence on soil moisture retrieval error. As vegetation optical depth can be linearly related to VWC through a vegetation parameter ‘b’ (the slope of the regression line for VWC versus vegetation optical depth) [5], thus a higher VWC can also result in a higher soil moisture retrieval error. Combining the results of Jackson et al. [5] and Parinussa et al. [21], it can be inferred that for vegetation such as corn, which can reach a VWC of as high as 4-5 kg/m² during its mature stage, a VWC error of 0.5 kg/m² will lead to a change of approximately 0.2 m³/m³ for soil moisture retrieval accuracy for C-band, X-band, or Ku-band microwave instruments. However, for vegetation water content less than 1.5 kg/m² such as legumes and grassland, a 0.5 kg/m² VWC error has almost no influence on the error of soil

moisture retrieval. Therefore, for soil moisture related remote sensing applications special attention needs to be paid for vegetation types such as corn and cereal grains, especially as they approach maturity. An example VWC map from MODIS-derived NDVI and the recommended equations from this paper is given in Fig. 6 for the SMAPEX-3 campaign. The VWC equations are applied on the basis of a Landsat derived landcover map, which is strongly reflected in the VWC distribution across the study site.

6. Conclusion

This study combined and inter-compared all available data sets and developed equations for estimating VWC from NDVI, NDWI₁₂₄₀, NDWI₁₆₄₀ and NDWI₂₁₃₀, based upon land cover type. Analyses led to several conclusions:

- 1) There were marked similarities among the data sets and equations developed from most field campaigns for each type of vegetation, but some significant differences exist, especially for cereal grains.
- 2) According to the performance statistics and the number of data sets available, NDWI₁₆₄₀ and NDVI are the two preferred vegetation indices for VWC estimation. Despite that NDVI is theoretically less suitable for estimating VWC when compared with the NDWI, it still provided a reliable estimate for VWC. Moreover, NDVI maps are readily available from the MODIS satellite, making operational implementation a relatively simple task.
- 3) The MODIS SWIR bands, especially at 1240 nm wavelength, have not been fully utilized for estimating VWC. More studies with larger number of VWC samplings are still needed, especially for cereal grains and grassland, to further evaluate the relationship between NDWI₁₂₄₀ and VWC.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

300 Additionally, this synthesis study recommended a new set of equations for VWC estimation
301 of four different vegetation types (corn, cereal grains, legumes and grassland), which will be
302 more reliable than the equations developed from single data sets. These equations can be
303 directly applied to satellite data in order to obtain VWC information for soil moisture
304 retrieval or other climatic and agricultural applications.

305
306 **Acknowledgement**

307 The authors would like to acknowledge all the researchers who have previously worked on
308 VWC and VI relationships, including those involved in the field campaigns listed in this
309 paper. The authors would also like to acknowledge research funding from the Australian
310 Research Council (DP0556941, DP0557543, DP0879212 and DP0984586) which supported
311 NAFE’05 and 06, AACES1-2 and SMAPEX1-3 field experiments. Special thanks are also
312 given to the Faculty of Engineering, Monash University for providing a postgraduate
313 scholarship for Ying Gao.

314 **References**

315 [1] C. J. Tucker, "Remote sensing of leaf water content in the near infrared," *Remote Sensing of*
316 *Environment*, vol. 10, pp. 23-32, 1980.
317 [2] J. Peñuelas, I. Filella, C. Biel, L. Serrano, and R. Savé, "The reflectance at the 950–970 nm
318 region as an indicator of plant water status," *International Journal of Remote Sensing*, vol. 14,
319 pp. 1887-1905, 1993.
320 [3] S. J. Pyne, P. L. Andrews, and R. D. Laven, *Introduction to wildland fire*: Wiley, 1996.
321 [4] T. J. Jackson, D. Y. Chen, M. Cosh, F. Q. Li, M. Anderson, C. Walthall, P. Doriaswamy, and
322 E. R. Hunt, "Vegetation water content mapping using Landsat data derived normalized
323 difference water index for corn and soybeans," *Remote Sensing of Environment*, vol. 92, pp.
324 475-482, Sep 2004.
325 [5] T. J. Jackson and T. J. Schmugge, "Vegetation effects in the microwave emission of soils,"
326 *Remote Sensing of Environment*, vol. 36, pp. 203-212, Jun 1991.
327 [6] D. Y. Chen, J. F. Huang, and T. J. Jackson, "Vegetation water content estimation for corn and
328 soybeans using spectral indices derived from MODIS near- and short-wave infrared bands,"
329 *Remote Sensing of Environment*, vol. 98, pp. 225-236, Oct 2005.
330 [7] M. H. Cosh, J. Tao, T. J. Jackson, L. McKee, and P. O'Neill, "Vegetation water content
331 mapping in a diverse agricultural landscape: National Airborne Field Experiment 2006,"
332 *Journal of Applied Remote Sensing*, vol. 4, May 2010.

- [8] V. Maggioni, R. Panciera, J. P. Walker, M. Rinaldi, and V. Paruscia, "A multi-sensor approach for high resolution airborne soil moisture mapping," presented at the 30th Hydrology and Water Resource Symposium [CD-ROM], The Institute of Engineers Australia, Launceston, Australia, 4-8 December, 2006, 2006.
- [9] M. T. Yilmaz, E. R. Hunt, and T. J. Jackson, "Remote sensing of vegetation water content from equivalent water thickness using satellite imagery," *Remote Sensing of Environment*, vol. 112, pp. 2514-2522, May 2008.
- [10] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," in *Proceedings of the Third ERTS Symposium*, Washington DC, 1973, pp. 309-317.
- [11] J. A. Gamon, C. B. Field, M. L. Goulden, K. L. Griffin, A. E. Hartley, G. Joel, J. Penuelas, and R. Valentini, "Relationships between NDVI, canopy structure, and photosynthesis in 3 Californian vegetation types," *Ecological Applications*, vol. 5, pp. 28-41, Feb 1995.
- [12] B.-C. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sensing of Environment*, vol. 58, pp. 257-266, 1996.
- [13] D. A. Roberts, R. O. Green, and J. B. Adams, "Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS," *Remote Sensing of Environment*, vol. 62, pp. 223-240, 1997.
- [14] M. Allahmoradi, J. P. Walker, A. W. Western, D. Ryu, T. Jackson, and E. Kim, "Estimating vegetation water content using optical remote sensing: application for soil moisture retrieval," *Remote Sensing of Environment*, 2013 (submitted).
- [15] J. Huang, D. Y. Chen, and M. H. Cosh, "Sub-pixel reflectance unmixing in estimating vegetation water content and dry biomass of corn and soybeans cropland using normalized difference water index (NDWI) from satellites," *International Journal of Remote Sensing*, vol. 30, pp. 2075-2104, 2009.
- [16] Y. H. Yi, D. W. Yang, D. Y. Chen, and J. F. Huang, "Retrieving crop physiological parameters and assessing water deficiency using MODIS data during the winter wheat growing period," *Canadian Journal of Remote Sensing*, vol. 33, pp. 189-202, Jun 2007.
- [17] O. Merlin, J. P. Walker, J. D. Kalma, E. J. Kim, J. Hacker, R. Panciera, R. Young, G. Summerell, J. Hornbuckle, M. Hafeez, and T. Jackson, "The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution," *Advances in Water Resources*, vol. 31, pp. 1444-1455, Nov 2008.
- [18] R. Panciera, J. P. Walker, J. D. Kalma, E. J. Kim, J. M. Hacker, O. Merlin, M. Berger, and N. Skou, "The NAFE'05/CoSMOS data set: Toward SMOS soil moisture retrieval, downscaling, and assimilation," *Ieee Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 736-745, Mar 2008.
- [19] S. Peischl, J. P. Walker, C. Rüdiger, N. Ye, Y. H. Kerr, E. Kim, R. Bandara, and M. Allahmoradi, "The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment," *Hydrology and Earth System Sciences*, vol. 16, pp. 1697-1708, 2012/06/22 2012.
- [20] R. Panciera, J. P. Walker, T. J. Jackson, D. A. Gray, M. A. Tanase, D. Ryu, A. Moneris, H. Yardley, C. Rudiger, X. L. Wu, Y. Gao, and J. M. Hacker, "The Soil Moisture Active Passive Experiments (SMAPEX): Toward soil moisture retrieval from the SMAP mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 490-507, Jan 2014.
- [21] R. M. Parinussa, A. Meesters, Y. Y. Liu, W. Dorigo, W. Wagner, and R. A. M. de Jeu, "Error Estimates for Near-Real-Time Satellite Soil Moisture as Derived From the Land Parameter Retrieval Model," *Ieee Geoscience and Remote Sensing Letters*, vol. 8, pp. 779-783, Jul 2011.

Fig. 1. Locations of the field campaigns compiled in this study.

Fig. 2. Data sets and models for VWC estimation using NDVI.

Fig. 3. Data sets and models for VWC estimation using NDWI₁₂₄₀.

Fig. 4. Data sets and models for VWC estimation using NDWI₁₆₄₀.

Fig. 5. Data sets and models for VWC estimation using NDWI₂₁₃₀.

Fig. 6. Example VWC map (Kg/m^2) from SMAPEX-3 for September 23, 2011, retrieved from a combination of MODIS-derived NDVI and a Landsat-derived land classification map.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Summary of literature used for this study, including the source of VWC, spectral data and derived vegetation indices.

Publication by author names	Year	Data Source				
		VWC	NDVI	NDWI _{I240}	NDWI _{I640}	NDWI _{I2130}
T. Jackson et al.	2004	SMEX02 (interpolated)	Landsat TM/ETM+	-	Landsat TM/ETM+	-
D. Y. Chen et al.	2005	SMEX02	Landsat TM/ETM+, Terra-MODIS	-	Landsat TM/ETM+, Terra-MODIS	Terra- MODIS
V. Maggioni et al.	2006	NAFE'05	100BX Radiometer	-	Aqua-MODIS	MODIS
Y. H. Yi et al.	2007	Weishan Experiment	Terra, Aqua-MODIS	-	Terra, Aqua- MODIS	Terra, Aqua- MODIS
M. T. Yilmaz et al.	2008	SMEX05	-	-	Landsat TM, AWiFS, ASTER	-
J. Huang et al.	2009	SMEX02	Landsat TM/ETM+	MODIS	Landsat TM/ETM+	Landsat TM/ETM+
M. H. Cosh et al.	2010	NAFE'06	-	-	Landsat TM	-
M. Allahmoradi et al.	2013	NAFE'06	CROPSCAN Multi-Spectral Radiometer (MSR-16)			
		AACES-1, -2	ASD Field Spectrometer (FieldSpec 3)			
		SMAPEX-1, -2, -3	CROPSCAN Multi-Spectral Radiometer (MSR-16)			

Table 2. Summary of campaign information.

Experiment [Source]	SMEX 02, 05 [9, 14]	NAFE'05 [17]	NAFE'06 [16]	Weishan [15]	AACES 1, 2 [18]	SMAPEX-1, -2, -3 [19]
Location	Walnut Creek watershed, Iowa, USA	Goulburn River catchment, NSW, Australia	Kyeamba/Yanco/ Yenda, NSW, Australia	Weishan Irrigation Zone, China	Murrumbidgee catchment, NSW, Australia	Yanco, NSW, Australia
Season	Summer	Spring		Spring	Summer and Winter respectively	Winter, Summer, and Spring respectively
Major crop types	Corn, soybean	Barley, wheat, corn, canola		Winter wheat	Barley, wheat, corn, canola	Barley, wheat, corn, canola, lucerne
Available ancillary data	VWC, LAI, dry biomass	VWC, LAI, vegetation height, surface reflectance		VWC, LAI, dry biomass	VWC, LAI, dry biomass, surface reflectance	VWC, LAI, dry biomass, surface reflectance

Table 3. Summary of spectral bands from field spectrometers used in the field campaigns of this paper, and current satellites that can be used for calculating the vegetation indices.

Field Spectrometer or Satellite	Wavelength (nm)				
	RED ₆₅₀	NIR ₈₆₀	SWIR ₁₂₄₀	SWIR ₁₆₄₀	SWIR ₂₁₃₀
100BX	630-690	760-900	NA	NA	NA
FieldSpec 3	630-670	820-880	1234-1246	1632-1648	2122-2138
MSR-16	630-670	820-880	1234-1246	1632-1648	NA
Landsat TM/ETM+	Band 3 630-690	Band 4 760-900	NA	Band 5 1550-1750	Band 7 2080-2350
MODIS	Band 1 620-670	Band 2 841-876	Band 5 1230-1250	Band 6 1628-1652	Band 7 2105-2155

Table 3. Equations for estimating VWC ('y') using the respective vegetation index ('x') according to individual studies in literature. Also shown is the recommended equation for each vegetation category where more than a single data set exists. Performance statistics are also provided.

	Publication by first author names	Equations and Statistics															
		NDVI	RMSE - org. data	R ² - org. data	RMSE - all data	NDWI ₁₂₄₀	RMSE - org. data	R ² - org. data	RMSE - all data	NDWI ₁₆₄₀	RMSE - org. data	R ² - org. data	RMSE - all data	NDWI ₂₁₃₀	RMSE - org. data	R ² - org. data	RMSE - all data
Corn	T. Jackson	$y=192.64x^3-417.46x^4+347.96x^3-138.93x^2+30.7x-2.82$	0.05	0.99	0.89	-	-	-	-	$y=9.82x+0.05$	0.05	0.98	0.61	-	-	-	-
	D. Y. Chen	$y=-17.75x^3+75.71x^4-73.46x^3+25.42x^2-0.83x-0.37$	0.62	0.72	0.75	-	-	-	-	$y=7.88x+0.58$	0.46	0.84	0.56	$y=6.67x+0.1$	0.61	0.72	-
	M. T. Yilmaz	-	-	-	-	-	-	-	-	$y=7.69x+0.75$	0.47	0.89	0.52	-	-	-	-
	J. Huang	-	-	-	-	$y=25.29x+1.31$	0.79	0.31	0.90	$y=7.71x+0.26$	0.25	0.62	2.31	$y=10.51x-4.11$	0.85	0.48	-
	M. H. Cosh	-	-	-	-	-	-	-	-	$y=9.39x+1.26$	0.34	0.71	1.21	-	-	-	-
	M. Allahmoradi	$y=68.22x^2-97.19x+35.37$	0.72	0.85	3.40	$y=19.66x+0.38$	0.64	0.89	1.51	$y=11.47x-1.23$	0.68	0.67	1.19	-	-	-	-
	Recommended	$y=0.098e^{4.225x}$	-	0.80	0.69	$y=25.37+1.1$	-	0.50	0.88	$y=7.84x+0.6$	-	0.87	0.51	-	-	-	-
Cereal grains	Y. H. Yi	$y=-51.73x^2+70.48x-20.24$	0.83	0.69	3.26	-	-	-	-	$y=12.5x-0.44$	0.74	0.76	2.23	$y=10.29x-1.98$	0.44	0.84	0.67
	V. Maggioni	$y=4.81x-0.55$	0.18	0.83	1.65	-	-	-	-	$y=13.2x^2+1.62x$	0.38	0.79	0.92	$y=10.99x-3.07$	0.45	0.75	0.72
	M. H. Cosh	-	-	-	-	-	-	-	-	$y=2.1x-0.51$	1.10	0.69	1.22	-	-	-	-
	M. Allahmoradi	$y=1.9x^2+0.02x+0.09$	0.42	0.54	0.48	$y=3.25x^2+5.31x+0.89$	0.40	0.62	-	$y=-0.82x^2+2.49x+0.62$	0.41	0.57	0.43	-	-	-	-
	Recommended	$y=0.078e^{3.510x}$	-	0.59	0.50	-	-	-	-	$y=2.45x+0.57$	-	0.57	0.43	$**y=12.38x-3.26$	-	0.84	0.55
Legumes	T. Jackson	$y=7.63x^4-11.41x^3+6.87x^2-1.24x+0.13$	0.03	0.99	0.45	-	-	-	-	$y=1.44x^2+1.36x+0.34$	0.02	0.97	0.17	-	-	-	-
	D. Y. Chen	$***y=2.06x-0.86$	0.22	0.46	0.43	-	-	-	-	$y=1.78x+0.28$	0.21	0.52	0.19	-	-	-	-
	M. T. Yilmaz	-	-	-	-	-	-	-	-	$y=2.22x+0.38$	0.12	0.87	0.20	-	-	-	-
	J. Huang	$y=0.89x-0.30$	0.29	0.08	0.35	$y=3.30x+0.63$	0.22	0.48	0.22	$y=0.85x+0.33$	0.17	0.31	0.25	$y=0.97x-0.01$	0.25	0.31	-
	M. Allahmoradi	$y=0.88x^2+0.18x-0.06$	0.36	0.96	0.31	$y=11.42x^2+5.29x+0.59$	0.24	0.96	0.23	$y=0.29x^2+1.36x+0.26$	0.22	0.96	0.20	-	-	-	-
	Recommended	$y=0.059e^{2.573x}$	-	0.51	0.31	$y=4.03x+0.68$	-	0.76	0.21	$y=1.74x+0.34$	-	0.76	0.18	-	-	-	-
Grassland	V. Maggioni	$y=0.21x+0.24$	0.03	0.92	0.34	-	-	-	-	$y=0.19x+0.32$	0.07	0.04	0.33	$y=0.78x+0.01$	0.02	0.9	0.38
	M. H. Cosh	-	-	-	-	-	-	-	-	$y=0.98x+0.28$	0.07	0.35	0.33	-	-	-	-
	M. Allahmoradi	$y=1.93x^2-0.21x+0.01$	0.29	0.56	0.30	$y=3.63x+0.80$	0.22	0.43	-	$y=0.17x^2+1.79x+0.5$	0.30	0.51	0.31	$y=0.94x+0.25$	0.33	0.37	0.32
	Recommended	$y=0.017e^{5.866x}$	-	0.52	0.33	-	-	-	-	$y=1.16x+0.45$	-	0.20	0.30	$y=0.74x+0.23$	-	0.31	0.31

* Literature that used interpolated data.
** Equation should be used with caution due to lack of data.
*** No data sets were presented in the original paper.