Correction for rapid-growth vegetation and testing of an upscaling method with a COSMOS probe

Samantha Lou Irvin
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, Remote Sensing Commons, and the Soil Science Commons

Recommended Citation
Irvin, Samantha Lou, "Correction for rapid-growth vegetation and testing of an upscaling method with a COSMOS probe" (2013). Graduate Theses and Dissertations. 13136.
http://lib.dr.iastate.edu/etd/13136

This Thesis is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Correction for rapid-growth vegetation and testing of an upscaling method with a COSMOS probe

by

Samantha Lou Irvin

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Environmental Science

Program of Study Committee:
Brian K. Hornbuckle, Major Professor
    Amy Kaleita
    Kristie Franz
    Trenton Franz

Iowa State University
    Ames, Iowa
    2013

Copyright © Samantha Lou Irvin, 2013. All rights reserved.
TABLE OF CONTENTS

LIST OF TABLES ......................................................... v
LIST OF FIGURES ....................................................... ix
ACKNOWLEDGEMENTS .................................................... xiv
ABSTRACT ................................................................. xv
CHAPTER 1. INTRODUCTION ............................................. 1
  1.1 Soil Moisture ....................................................... 1
  1.2 Soil Moisture Measurement ........................................ 3
    1.2.1 Remote sensors ................................................ 4
    1.2.2 COSMOS ....................................................... 5
  1.3 First Research Question ......................................... 9
    1.3.1 Parts of the first research question ........................ 10
  1.4 Upscaling ......................................................... 10
  1.5 Second Research Question ....................................... 11
    1.5.1 Parts of the second research question ....................... 11
CHAPTER 2. SOIL AND VEGETATION MEASUREMENTS ................. 13
  2.1 Introduction ..................................................... 13
  2.2 Soil Samples ..................................................... 15
    2.2.1 Calculations .................................................. 18
  2.3 Vegetation ....................................................... 19
    2.3.1 Allometric relationship ..................................... 19
    2.3.2 Vegetation measurements ................................... 21
    2.3.3 Total biomass ............................................... 24
2.3.4 Results ................................................................. 27
2.4 Vegetation Variability .................................................. 28
  2.4.1 Methods ............................................................. 33
  2.4.2 Results ............................................................. 34
2.5 Conclusions ............................................................ 42

CHAPTER 3. CALIBRATION OF COSMOS WITH CROPS .............. 44
  3.1 Introduction ............................................................ 44
  3.1.1 Question ............................................................ 46
  3.2 Methods ............................................................... 47
    3.2.1 Depth of measurement .......................................... 47
    3.2.2 $N_0$ determination ............................................. 49
  3.3 Results ............................................................... 53
    3.3.1 $\kappa$ values .................................................... 53
    3.3.2 $N_0$ changes ..................................................... 54
    3.3.3 Calibration ....................................................... 55
  3.4 Conclusions .......................................................... 64

CHAPTER 4. VALIDATION OF FEATURE SPACE INTERPOLATION METHOD .......................................................... 66
  4.1 Introduction ........................................................... 66
    4.1.1 Question ........................................................ 69
  4.2 Methods ............................................................... 69
    4.2.1 Field data ......................................................... 69
    4.2.2 Soil moisture data .............................................. 71
    4.2.3 Validation data .................................................. 73
    4.2.4 Weighting function ............................................. 74
  4.3 Results ............................................................... 76
  4.4 Conclusions .......................................................... 77
**LIST OF TABLES**

Table 1.1  A comparison of different methods that can be used to measure soil moisture. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

Table 1.2  Adapted from *Zreda et al.* (2012): Nuclear properties of ten elements contributing most to macroscopic scattering cross–section in terrestrial rocks. $A =$ atomic mass (g mol$^{-1}$); $\sigma_{sc} =$ elastic scattering cross–section (barns); NC = number of collisions to thermalize a 1-2 MeV neutron; $\zeta =$ average log decrement of energy per neutron collision; SP = elemental stopping power (computed as $\zeta \cdot \sigma_{sc}$, in cm$^{-1}$); $C =$ concentration of elements in dry “average rock” (ppm). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

Table 2.1  Allometric relationship data. Slope ($c_1$) (kg plant$^{-1}$ m$^{-3}$), y–intercept ($c_0$) (kg plant$^{-1}$), and $R^2$ data for 2011 maize measurements and 2012 soybean measurements when comparing the relationship of $c_1(S_d^2 Z_C) + c_0 = m$ for the 30 plants removed from the edge of the field on VC days. 26

Table 2.2  Variances and covariances found for each of the 30 plant mass fits made in 2011 and 2012. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

Table 2.3  The resulting maize means of the 105 plants at each location where $S_d$ (mm) is the mean stem diameter ($\sigma_{sd}$ standard deviation), $Z_C$ (mm) is the mean canopy height ($\sigma_{zc}$ standard deviation), $S_d^2$ (mm$^2$) is the mean stem diameter squared ($\sigma_{sd^2}$ standard deviation), and $S_d^2 Z_C$ (mm$^3$) is the mean stem diameter squared times canopy height ($\sigma_{sd^2 zc}$ standard deviation). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
Table 2.4 The resulting soybean means of the 105 plants at each location where $S_d$ (mm) is the mean stem diameter ($\sigma_{sd}$ standard deviation), $Z_c$ (mm) is the mean canopy height ($\sigma_{zc}$ standard deviation), $S_d^2$ (mm$^2$) is the mean stem diameter squared ($\sigma_{sd^2}$ standard deviation), and $S_d^2 Z_c$ (mm$^3$) is the mean stem diameter squared times canopy height ($\sigma_{sd^2 zc}$ standard deviation).

Table 3.1 Dates in–situ soil samples were taken on VC days with their respective vegetation water equivalents ($W_s$) (cm); calculated $z^*$ (using (3.4) and assuming $\kappa = 1$) (cm); the average, vertically weighted volumetric soil moisture (cm$^3$ cm$^{-3}$) down to that depth using our in–situ samples (standard error of measurements in parentheses); and the new calibrated $N_0$ (cph) for that date (error in parentheses).

Table 3.2 Effective depths ($z^*$) (cm) found with (3.4), using extreme $\kappa$ values.

Table 3.3 The $N_0$ values calculated after 2011 VC maize field days (cph); $N_0$ values using fits (Figure 3.7) from models (cph); in–situ soil moistures we found for field days in 2011 (cm$^3$ cm$^{-3}$); COSMOS soil moistures found with an uncorrected, constant $N_0$ of 2898 cph (cm$^3$ cm$^{-3}$); as well as the COSMOS soil moistures found with our modeled $N_0$ values (Figure 3.8) (cm$^3$ cm$^{-3}$). Differences between the calculated $N_0$ values and modeled $N_0$ values are shown in parentheses, as are the differences between in-situ soil moisture and uncorrected and modeled COSMOS soil moistures (+ means it was higher than the measured value, – means it was lower than the measured value).
Table 3.4 The $N_0$ values calculated after 2012 VC soybean field days (cph); $N_0$ values using fits (Figure 3.10) from models (cph); in–situ soil moistures we found for field days in 2012 (cm$^3$ cm$^{-3}$); uncorrected COSMOS soil moistures using a constant $N_0$ of 2898 cph (cm$^3$ cm$^{-3}$); and the COSMOS soil moistures found with our modeled $N_0$ values (Figure 3.11) (cm$^3$ cm$^{-3}$). Differences between the calculated $N_0$ values and modeled $N_0$ values are shown in parentheses, as are the differences between in–situ soil moisture and uncorrected and modeled COSMOS soil moistures (+ means it was higher than the measured value, − means it was lower than the measured value).

Table 4.1 Soil moistures for each centroid for the five UV sampling days. Standard errors are shown in parentheses.

Table 4.2 Fresh mass vegetation ($M_f$) amounts (kg m$^{-2}$) found via the fits in Chapter 3 (Figure 3.9) for the five UV sampling days. Using a fit for growing vegetation’s fresh and water masses, the water percentages of the $M_f$ values were found and used to create the vegetation water equivalent ($W_s$) (cm). The $N_0$ values (cph) found for UV sampling days using the vegetation correction function created with data from Chapter 3 (Figure 3.10). Also, the recalculated soil moistures (cm$^3$ cm$^{-3}$) for those days with these $N_0$ values using each respective day’s moderated neutron counts (Figure 3.11). Using our vegetation–dependent COSMOS modeling information, these are considered the “true” values for $N_0$ and $\theta_v$ on these five days. Each day’s $z^*$ is also shown, which assumes all the other data in this table is true. $z^*$ is found with (3.4) by assuming $\theta_v$ is uniform throughout the first 30 cm and $\rho_b$ does not vary by depth or time.
Table 4.3 COSMOS probe footprint areal soil moistures for cluster groups. $A =$ areal soil moisture found with no weighting function (cm$^3$ cm$^{-3}$). $W =$ areal soil moisture found with weighting function (cm$^3$ cm$^{-3}$).

Table 4.4 The difference between the true soil moisture (Table 4.2) and the upscaled soil moistures (Table 4.3) in cm$^3$ cm$^{-3}$ for all cluster groups on all UV days. Positive values represent differences when the upscaled soil moisture overestimated the true soil moisture while negative values represent when the upscaled soil moisture is lower than the true soil moisture.
### LIST OF FIGURES

| Figure 1.1 | Figure from *Crow et al.* (2012) that shows the dominant physical controls on soil moisture spatial variability as a function of scale. Grey shading of bars shows the relative importance of each control at various scales. |
| Figure 1.2 | Google Map image of the IVS. Grey center dot is the location of the COSMOS probe, while the 18 black dots are the suggested calibration locations. |
| Figure 2.1 | The soil sampling instrument used to take soil samples. |
| Figure 2.2 | PVC pipe soil cores were placed on with meter stick to measure the 5 cm increments. A putty knife was used to divide the cores and place the segments into the cans (both instruments shown in the image). |
| Figure 2.3 | The drying oven used to dry soil samples at 105°C for 24 hours. |
| Figure 2.4 | Digital calipers used to measure stem diameter and the 1.6 m aluminum conduit pole with tape measure used to measure canopy height. |
| Figure 2.5 | A few examples of the allometric relationship of $c_1(S_d^2 Z_c) + c_0 = m$ for fresh, dry, and water masses. There were 30 plants taken each VC day we sampled vegetation to be used in this relationship. The masses of the plants around the COSMOS probe were projected with these relationships and the plants’ respective $S_d$ and $Z_c$ values. Error bars shown are error due to instrument errors. |
| Figure 2.6 | Graph of the areal maize vegetation masses over the field in 2011. Data points were determined by (2.6) with error bars found with (2.7). |
Figure 2.7 Areal soybean vegetation masses during 2012. Data points were determined by (2.6) with error bars found with (2.7). Note that soybean plants lose their leaves as they senesce. This effect can be seen with the lack of a peak in the dry matter during the senescent sampling day and by a plateau of dry matter instead. As the plants were losing their leaves while the pods matured, the dry matter stayed at a more constant level than it did for senescent maize plants.

Figure 2.8 Soybean plants on September 25, 2012. All the leaves have fallen off the plants, leaving only pods on the stems.

Figure 2.9 The sampling locations for each of the fields used for PV days. In Figure 2.9(a), Location 1 is the center dot, Location 2 is the southern dot, Location 3 is the northern dot. In Figure 2.9(b), Location 1 is the northern dot, Location 2 is the southeast dot, Location 3 is the west dot.

Figure 2.10 Some examples of the original plant measurements. The mean for all the figures is shown with the dashed black line while each green bar represents a different plant. We measured 15 consecutive plants in 7 adjacent rows in both the maize and soybean fields. ($S_d$ being stem diameter in mm and $Z_c$ being canopy height in cm).

Figure 2.11 Side–by–side boxplot of the variance in $S_d^2Z_c$ of the 105 plants measured at 3 locations on 3 days in a maize field. Outliers are shown as open circles, the dark horizontal lines represent the median values, upper and lower box edges mark the first and third quartiles, and the whiskers represent the respective upper and lower 25% of the data. Horizontal axis points are read as “Location, Date” where 1,1 would represent maize location 1, on sampling date 1 (June 26, 2012).
Figure 2.12 Side–by–side boxplot of the variance in $S_d^2 Z_c$ of the 105 plants measured at 3 locations on 3 days in a soybean field. Outliers are shown as open circles, the dark horizontal lines represent the median values, upper and lower box edges mark the first and third quartiles, and the whiskers represent the respective upper and lower 25% of the data. Horizontal axis points are read as “Location, Date” where 1,1 would represent soybean location 1, on sampling date 1 (June 28, 2012).

Figure 2.13 A few examples of how $\sigma_m^2$ varied depending on $n$ - the number of plants sampled (all units kg plant$^{-1}$). After $n = 30$, $\sigma_m^2$ levels out and becomes almost constant.

Figure 3.1 Soil moisture profiles for each of the sampling days in 2011 as well as two of the 2012 days. Each grey line represents one location, the black line is the average of the 18 locations with standard errors shown as the error bars.

Figure 3.2 Soil moisture profiles for eight of the sampling days in 2012. Each grey line represents one location, the black line is the average of the 18 locations with standard errors shown as the error bars.

Figure 3.3 Fresh matter and water column densities of maize plants versus recalibrated $N_0$ values for 2011.

Figure 3.4 A plot of fresh matter and water column densities for soybean plants in 2012 against recalibrated $N_0$ values. Three bare soil condition samples were taken at the beginning of the year, which proved to have variable $N_0$ values.
Figure 3.5  Soybean plants on June 5, 2012. These plants were too small to be measured accurately so this day was categorized as a bare soil day. Remaining maize residue can also be seen here. This residue has been worked into the top soil now due to planting practices and has decomposed over time, but it can still be seen how challenging it would have been to measure the total mass of residue.

Figure 3.6  The original areal fresh matter vegetation over the season along with a one–term Gaussian fit for the 2011 maize vegetation.

Figure 3.7  A rough estimation of how $N_0$ would vary over the 2011 crop season. This was found by combining fits to Figures 2.6 and 3.3.

Figure 3.8  Soil moisture calculations by COSMOS during the 2011 crop season. This figure displays a comparison between the ‘original’, uncorrected soil moisture values found when using the initial calibration of a constant $N_0$ (black line) and the modeled soil moisture after using a vegetation–dependent variable $N_0$ (red line). In–situ soil moisture values are shown as blue dots on the days they were taken. The bottom figure is the difference between the uncorrected soil moisture and modeled soil moisture. The gap in the series in early July has to do with a short period of missing data.

Figure 3.9  2012 soybean fresh matter vegetation with a one–term Gaussian fit.

Figure 3.10  A rough estimation of how $N_0$ would vary over the 2012 crop season. This was found by combining fits to Figures 2.7 and 3.4.
Figure 3.11 Soil moisture calculations by COSMOS during the 2012 crop season. This figure displays a comparison between the ‘original’, uncorrected soil moisture values found when using the initial calibration of a constant $N_0$ (black line) and the modeled soil moisture after using a vegetation–dependent variable $N_0$ (red line). In–situ soil moistures are again shown as blue dots on their respective days. The bottom figure shows the difference between the uncorrected soil moistures and the modeled soil moistures. 62

Figure 3.12 A comparison of the two seasons of vegetation–dependent $N_0$ values. The 2011 $N_0$ values are displayed in blue while the 2012 $N_0$ values are displayed in purple. 64

Figure 4.1 All 10 m locations with soil and topographic data within the COSMOS footprint (350 m radius), COSMOS probe (center dot), centroids of the 4 cluster groups. 70

Figure 4.2 All 10 m locations with soil and topographic data within the COSMOS footprint divided into respective cluster groups. Each cluster’s centroid is the larger dot of the same color. Duplicate centroids are represented by the same color. 73

Figure 4.3 Figure 4.3(a) Cumulative distribution of the how neutron counts are weighted for a COSMOS probe. Figure 4.3(b) We used the same cut off as Zreda et al. (2008), that of 86% or $1 - e^{-2}$, at 350 m and used points between the dry (solid) and saturated (dashed) conditions shown to determine a single function. 75
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. Brian K. Hornbuckle for his guidance, patience and support throughout this research and the writing of this thesis. His insights and words of encouragement were invaluable to my graduate school experience.

I would also like to thank my committee members for their efforts and contributions to this work. Their guidance, knowledge, and valuable insights on the Feature Space Interpolation method and the COSMOS project were assets I could not have done without. I am also grateful to all of those who helped with my collection of data by either assisting me in the field or supplying the materials and knowledge I needed: Jason Patton, Ben Carr, Ciera Hancin, Andrew Kumler, Crystal Wang, Andrew Spencer, and Gavin Simmons. I also need to thank Dan Fortin for his guidance with the statistical work in my thesis.

Finally, I would like to thank my family and friends for their unwavering love and support throughout the last two years. None of this would have been possible without all of those mentioned above. I am extremely grateful to all of them and everyone else who assisted me in my pursuit of my Master’s.
ABSTRACT

COSMOS (COsmic–ray Soil Moisture Observing System) probes are a developing instrument in the field of remote sensing and are currently the only source of field scale soil moisture data. A COSMOS probe is essentially a hydrogen particle counter with a measurement depth ranging from 10-70 cm and a spatial resolution approaching 700 m in diameter. A majority of the hydrogen detected is in the form of water molecules. As vegetation contains both hydrogen in the vegetative dry matter and water, in addition to sitting above the ground, it may influence what is measured by the COSMOS probe as soil moisture. We hypothesized that COSMOS probes located in areas with rapid growth vegetation need to be checked for a vegetation effect on their soil moisture measurements.

In order to account for vegetation, however, a method of monitoring and modeling the amount of vegetation was needed. We assumed that an allometric relationship of stem diameter and canopy height would be able to model the moisture in a plant. In–situ soil samples were taken concurrently with the vegetation measurements and used to recalibrate the COSMOS probe’s maximum counting rate parameter, \( N_0 \). \( N_0 \) was thought to be a site–specific constant that took background hydrogen into account. However, we found that the vegetation hydrogen still needs to be accounted for because as the vegetation grew, \( N_0 \) decreased and as the vegetation senesced, \( N_0 \) increased. This discovery led us to create a vegetation–corrected calibration of \( N_0 \) that could be used to remove the vegetation water column effect from what the COSMOS probe is observing over the crop season.

Aside from remote sensors, estimation of soil moisture over large areas requires either numerous, exhaustive point samples, or some method of upscaling a smaller number of measurements. Upscaling point samples to a larger area can be difficult. A new method for upscaling point scale measurements is the Feature Space Interpolation (FSI) method. FSI uses K–means clustering to separate a larger area into cluster groups. The location thought to best represent
the cluster is identified for sampling purposes. Soil moisture samples can be taken at each of these points as a representation of the whole cluster. With our vegetation-corrected calibration of the COSMOS probe, we were able to validate the areal soil moisture within the COSMOS footprint found with the FSI method. We found that FSI can upscale soil moisture values to within 0.063 cm³ cm⁻³ of what the COSMOS probe reports for soil moisture.
CHAPTER 1. INTRODUCTION

1.1 Soil Moisture

Soil moisture is a measure of the water content in soil. While soil moisture accounts for only a small fraction of the world’s water, it is an important part of the water cycle. It provides water for growing crops, a prominent factor of life in the Midwest, and can be a contributing factor for droughts and floods. Soil moisture has also been found to play a key role in groundwater recharge (Dripps, 2012; Salvucci, 2001), surface and subsurface runoff (Dunne and Black, 1970; Bonta, 1998), dust emissions (Laurent et al., 2008), climate variability (Seneviratne et al., 2010), stream flow (Berg and Mulroy, 2006), short-term weather prediction (Drusch, 2007), and long-term terrestrial water cycle trends (Jung et al., 2010).

The spatial scale at which soil moisture information is desired can vary from point location, to field scale, to tens of square kilometers. A study examining the variation in soil moisture across the continental U.S.A. would be on a scale of tens of kilometers. However, a farmer who desires to know how wet their field is would have no use for such a large scale as it could not tell them about a small area accurately. Therefore, the spatial scale at which soil moisture is examined will depend on the individual situation.

Soil moisture can vary significantly over large spatial areas due to natural events such as droughts (Tang and Piechota, 2009). Large scale soil moisture and its variability is mostly moderated by weather and land cover (Hawley et al., 1983) whereas smaller scale moisture is more affected by topography (Mohanty et al., 2000a) and soil composition (Vinnikov et al., 1996) (see Figure 1.1). Soil moisture can vary significantly over small spatial areas due to the natural heterogeneity of soil caused by soil structure (Mohanty et al., 2000b), pore size (macro and micro) (Anderson and Stormont, 2006), soil texture (Hawley et al., 1983), organic matter
content (Fekete et al., 2012), and color (Jacobs et al., 2004). Topography is still important at larger scales though, in that it could affect a larger area’s variability. An area that consists of mostly flat ground would likely have less variability than an area that contains rolling hills and valleys (Charpentier and Groffman, 1992).

Figure 1.1 Figure from Crow et al. (2012) that shows the dominant physical controls on soil moisture spatial variability as a function of scale. Grey shading of bars shows the relative importance of each control at various scales.

Temporal resolution of soil moisture is also an important factor when looking at soil moisture. Soil moisture is a variable that can change over a short period of time, depending on the depth being examined. The top layer of soil is susceptible to drying and wetting from evaporation and rainfall events. Deeper layers of soil are not as easily changed by these factors and are more influenced by long-term trends. Therefore, it is usually the top segment of soil being examined for soil moisture determinations and higher temporal resolution could better capture or characterize rapidly changing variables. Generally speaking, the higher resolution of temporal data, the more effective it can be to examine various factors. Low temporal resolution is still useful, but may not be able to track minute changes.

Soil texture is an important factor in soil moisture (Hawley et al., 1983). The amount of clay, silt, and sand present in one area could mean the difference between soil having a high water holding capacity or a low water holding capacity. Soil with more clay is more likely to
retain more water as the small clay particles provide a larger surface area for adsorption of water molecules. Sandy soil, on the other hand, has larger particles and is less likely to retain water. Soil texture can change over an area as small as five meters.

Topography is another contributing factor to soil moisture determination and variability. Topography itself is variable and although changes may be small, on the order of a few meters difference, hills, valleys, and various degrees of slopes can be found in a field. This also affects soil moisture due to the natural effect of gravity with water moving down toward lower elevations. Erosion also becomes a factor in topography. Smaller soil particles, clay, are more likely to erode from wind and water and become settled in lower elevations rather than higher ones. This would change the soil texture and the soil’s capacity to hold water. Aspect plays a part in soil moisture as well. Due to the position of the Sun in the sky, its heat touches different areas at different intensities. A north–facing slope in the Northern Hemisphere would not receive as much sun as a south–facing slope; therefore, a north–facing slope may have a higher soil moisture than a south–facing slope. These factors, and more, are reasons why it can be difficult to determine soil moisture as they can all affect soil moisture and vary at small distances.

1.2 Soil Moisture Measurement

Soil moisture can be expressed as either a mass of water over the total mass of the dry soil (gravimetric, $\theta_g$) or volume of water compared to the whole volume of the section which includes air, soil, and water (volumetric, $\theta_v$). Soil moisture can range from completely dry to a level of saturation dependent on the type of soil. In the Midwest soils, soil moisture usually varies between 0.05-0.40 cm$^3$ cm$^{-3}$. A common way to measure in–situ soil moisture is to take a destructive sample of soil from an area and oven–dry it to find the gravimetric soil moisture (Gardner, 1986). Destructive sampling is considered the most accurate method of sampling, but it can be a time–consuming process. Destructive sampling is one example a of “point” scale soil moisture measurement. It should be noted that “point scale” can be defined differently depending on the source of sampling and therefore sample volume being examined.

Technological advancements have allowed for the creation of indirect methods of sampling
soil moisture by electrical devices which take less time and effort. Impedance probes are an example of a more indirect way to measure soil moisture. With these devices, one simply inserts the probe into the soil and, by sensing the dielectric constant of the soil, the probe estimates the volumetric water content. Time–domain reflectometry (TDR) is another common way to determine soil moisture at a location. With TDR, an instrument is buried at a certain depth of soil and a pulse of energy is sent down the prongs of the instrument. The travel time of the pulse back to the sensor is related to the electrical properties of the soil, which can be used to determine soil moisture.

1.2.1 Remote sensors

A new field of remote sensing is evolving to measure soil moisture at larger scales than in–situ measurements. One example is SMOS, Soil Moisture and Ocean Salinity, a passive L–band satellite launched by the European Space Agency (ESA) on November 2, 2009. SMOS uses a microwave radiometer to determine the level of salinity in the oceans as well as the soil moisture of the top ∼5 cm of the land surface. This satellite measures the dielectric constants of the water bodies and soil which are directly related to salt and water content. SMOS has a spatial footprint of 50 km and can provide global maps of soil moisture and ocean salinity every 2-3 days (Kerr et al., 2010). The National Aeronautics and Space Administration (NASA) has plans to launch a similar satellite, SMAP, Soil Moisture Active–Passive, in 2014 (Entekhabi et al., 2010). SMAP will use a combination of high–resolution L–band radar and radiometry to measure surface soil moisture and freeze/thaw states. These measurements will depend on emissivity of the natural brightness temperature of the top ∼5 cm of soil. Soil moisture measurements will have a spatial resolution of 10 km and freeze/thaw state will have a 3 km resolution. SMAP will also have a 2-3 day repeat cycle over the earth. Another satellite system in orbit is GRACE, Gravity Recovery And Climate Experiment, which consists of twin satellites launched in March 2002 (Tapley et al., 2004). The goal of GRACE is to map the change in the gravitational field of the earth which depends on the distribution of global mass. This distribution of global mass is related to changing distribution of water resources (soil moisture, snow, ice, etc.) on the planet. GRACE has a spatial resolution of 400 km and a 30 day repeat
Table 1.1  A comparison of different methods that can be used to measure soil moisture.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Spatial Resolution</th>
<th>Temporal Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravimetric samples</td>
<td>5 cm</td>
<td>hourly</td>
</tr>
<tr>
<td>Impedance probe</td>
<td>5 cm</td>
<td>hourly</td>
</tr>
<tr>
<td>TDR</td>
<td>30 cm</td>
<td>hourly</td>
</tr>
<tr>
<td>COSMOS</td>
<td>700 m</td>
<td>hourly</td>
</tr>
<tr>
<td>SMAP</td>
<td>10 km</td>
<td>2-3 days</td>
</tr>
<tr>
<td>SMOS</td>
<td>50 km</td>
<td>2-3 days</td>
</tr>
<tr>
<td>GRACE</td>
<td>400 km</td>
<td>30 days</td>
</tr>
</tbody>
</table>

cycle.

1.2.2 COSMOS

A developing land–based passive remote sensor is the cosmic–ray neutron probe for measuring soil moisture (Hydroinnova LLC, Albuquerque, NM), which is distributed across the continental U.S.A. in the COSMOS project (COsmic–ray Soil Moisture Observing System). The cosmic–ray neutron probe (here referred to as the COSMOS probe) sits in a stationary location on the earth, about two meters off the ground. One probe has a footprint close to 660 m diameter at sea level (Zreda et al., 2008). The depth at which a COSMOS probe senses down to varies depending on the current level of soil moisture. With completely dry soil, 0 cm$^3$ cm$^{-3}$ moisture by volume, a probe can sense down to nearly 70 cm while with saturated soil, 0.40 cm$^3$ cm$^{-3}$ moisture by volume, the probe only senses about 10 cm into the ground (Zreda et al., 2008). One COSMOS probe was installed at the Iowa Validation Site (IVS), south of Ames, IA, in September 2010. COSMOS probes are a fitting middle step between point samples and the large spatial resolutions of satellites. The footprint of a COSMOS probe is closer to field scale which makes it a useful tool for validating the upscaling of point measurements and could be upscaled to compare to a satellite’s measurement of a larger footprint. A comparison of these different methods used to measure soil moisture can be found in Table 1.1.

COSMOS probes use moderated fast neutrons (hereafter referred to as moderated neutrons) that originate naturally from cosmic–rays to determine soil moisture within their footprint. Fast neutrons are defined to be neutrons with energy on the order of 1-2 MeV (Hess et al., 1959) while
Table 1.2 Adapted from Zreda et al. (2012): Nuclear properties of ten elements contributing most to macroscopic scattering cross-section in terrestrial rocks. \( A \) = atomic mass (g mol\(^{-1}\)); \( \sigma_{sc} \) = elastic scattering cross-section (barns); \( NC \) = number of collisions to thermalize a 1-2 MeV neutron; \( \zeta \) = average log decrement of energy per neutron collision; \( SP \) = elemental stopping power (computed as \( \zeta \cdot \sigma_{sc} \), in cm\(^{-1}\)); \( C \) = concentration of elements in dry “average rock” (ppm).

<table>
<thead>
<tr>
<th>Element</th>
<th>( A )</th>
<th>( \sigma_{sc} )</th>
<th>( NC )</th>
<th>( \zeta )</th>
<th>( SP )</th>
<th>( C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1.0079</td>
<td>22.02</td>
<td>18</td>
<td>1.000</td>
<td>22.016</td>
<td>—</td>
</tr>
<tr>
<td>O</td>
<td>15.9994</td>
<td>4.232</td>
<td>149</td>
<td>0.120</td>
<td>0.508</td>
<td>487 875</td>
</tr>
<tr>
<td>C</td>
<td>12.011</td>
<td>5.551</td>
<td>113</td>
<td>0.158</td>
<td>0.875</td>
<td>87 638</td>
</tr>
<tr>
<td>Si</td>
<td>28.0855</td>
<td>2.167</td>
<td>257</td>
<td>0.070</td>
<td>0.151</td>
<td>281 367</td>
</tr>
<tr>
<td>Na</td>
<td>22.9898</td>
<td>3.28</td>
<td>211</td>
<td>0.085</td>
<td>0.277</td>
<td>23 206</td>
</tr>
<tr>
<td>Ca</td>
<td>40.078</td>
<td>2.83</td>
<td>364</td>
<td>0.049</td>
<td>0.139</td>
<td>70 963</td>
</tr>
<tr>
<td>Al</td>
<td>26.9815</td>
<td>2.83</td>
<td>247</td>
<td>0.072</td>
<td>0.109</td>
<td>58 015</td>
</tr>
<tr>
<td>Fe</td>
<td>55.847</td>
<td>11.62</td>
<td>505</td>
<td>0.035</td>
<td>0.411</td>
<td>28 980</td>
</tr>
<tr>
<td>Mg</td>
<td>24.305</td>
<td>3.71</td>
<td>223</td>
<td>0.080</td>
<td>0.297</td>
<td>13 436</td>
</tr>
<tr>
<td>K</td>
<td>39.0983</td>
<td>1.96</td>
<td>355</td>
<td>0.050</td>
<td>0.099</td>
<td>19 137</td>
</tr>
</tbody>
</table>

The moderated neutrons have energy on the order of 10 eV. These neutrons are best moderated and scattered by hydrogen particles. As can be seen in Table 1.2, adapted from Zreda et al. (2012), hydrogen has a stopping power more than 20 times larger than the next element. It takes 18 collisions with hydrogen particles before a neutron is thermalized. A COSMOS probe contains a tube of helium 3 gas and, when a neutron hits this tube or passes through it, the gas becomes ionized. An electron is then released and a high voltage source shapes the distribution of electrons which can be recorded by a datalogger. Count rates of moderated neutrons have been recorded to fall between 400 cph (counts per hour) and 6000 cph (Zreda et al., 2012). At IVS, we see counts on the order of 1000-2000 cph.

Desilets et al. (2010) used a program developed to model the interaction of moderated neutrons with the Earth’s surface to create an empirical equation relating soil moisture to the cumulative count of moderated neutrons over one hour.

\[
\theta(N) = \frac{a_0}{N/N_0 - a_1} - a_2
\]

(1.1)

is used to go from the moderated neutron count to volumetric soil moisture. \( N \) is the cumulative
count of moderated neutrons over one hour (cph); $N_0$ is the constant theoretical maximum counting rate of moderated neutrons that would occur if the soil were completely dry (cph) and is site–dependent; $a_0$, $a_1$, $a_2$ are universal constants - 0.0808, 0.372, and 0.115 respectively; and $\theta(N)$ is the volumetric soil moisture.

![Google Map image of the IVS. Grey center dot is the location of the COSMOS probe, while the 18 black dots are the suggested calibration locations.](image)

Figure 1.2  Google Map image of the IVS. Grey center dot is the location of the COSMOS probe, while the 18 black dots are the suggested calibration locations.

It had been thought that only a single calibration of each probe was needed for soil moisture determination after installation in order to find the site’s constant $N_0$. This calibration was done by taking 108 gravimetric soil samples at 18 locations around the probe. An image of the IVS probe and its 18 sampling locations can be seen in Figure 1.2. These 18 points are chosen to be of equal representation or weight in the probe’s footprint. This is needed because, as Zreda et al. (2008) shows, the influence of a single location depends on its distance from the probe. With the gravimetric soil moisture samples and bulk density data, volumetric soil moisture can be found. Using the current moderated neutron counts with the volumetric soil moisture, (1.1) can be solved for $N_0$.

However, COSMOS probes need to be checked for various effects before the moderated neutron counts can be used to find the current soil moisture (Zreda et al., 2012). These effects include: different probe size/composition, variation in cosmic–ray intensity, pressure changes,
temperature changes, relative humidity levels, water vapor, lattice water, organic matter, and other sources of hydrogen. The different probe size/composition is accounted for with a constant that is used in an intermediate step of correcting the raw moderated neutron counts measured by the probe. Variation in cosmic–ray intensity is accounted for by a cosmic–ray monitor found at a high altitude that is used to find the natural changes in incoming high–energy particle intensity (Zreda et al., 2012). Pressure is measured at the site and is used to correct the counts of moderated neutrons. Temperature and relative humidity measurements represent internal data box conditions; these are monitored because they could affect the electronics. Atmospheric water vapor also needs to be accounted for, which can be done by accounting for external air temperature, pressure, and relative humidity. The new generation of COSMOS probes will have temperature, pressure, and humidity sensors installed in them, while the old generation will be retro–fitted with temperature, pressure, and humidity sensors (Rosolem et al., 2013) to account for the atmospheric water vapor. Lattice water (water in the mineral grains and bound water) is also taken into account at each site. Lattice water is the water that is released at 1000°C preceded by drying at 105°C. Lattice water is taken to be a constant at each site and is used in determining the vertical sensitivity of the COSMOS probe. Organic matter becomes important when organic carbon exceeds ~1% by weight (Zreda et al., 2012). This is taken into account in Franz et al. (2013a) in a similar way that lattice water is taken into account. For the IVS, organic carbon is about half as important as lattice water and will mostly effect the probe depth. Hydrogen molecules that exist in vegetation is a correction factor that we examined.

We assumed that a major hydrogen pool would be vegetation that existed within the COSMOS probe’s footprint. The vegetation would contain water as well as hydrogen in the vegetative dry matter. To account for this vegetative hydrogen content, a method was needed to monitor and model the above–ground vegetation (water content and dry matter) within the footprint of the COSMOS probe. We decided to use an allometric relationship for this factor. We assumed that we could make a rough estimation of an individual plant’s mass by modeling it with the plant’s stem diameter and canopy height and that any leaves or reproductive parts would be directly related to the main core of the plant. As soil texture (and therefore soil water) is a large factor in the spatial variability of soil moisture at the field scale, we assumed
that spatial variation in the vegetation would behave in the same manner as the variation of soil texture. Therefore, we sampled vegetation in the same locations as soil moisture to monitor and record these variations.

We measured 5 maize plants in 2011 and 10 soybean plants in 2012 at each of the 18 sampling locations. The more plants that are measured, the better the characteristic and its variability are captured. However, limited resources restricted the number of measurements. With so few plants measured in one location compared to the number of plants within the whole field, we would need to find a way to estimate the error due to natural plant variability in our measurements.

1.3 First Research Question

As the count of moderated neutrons recorded by a COSMOS probe is inversely related to the presence of hydrogen, the presence of all forms of hydrogen is crucial. Vegetation contains some hydrogen within its dry matter, but more importantly vegetation contains water which consists of hydrogen and oxygen. When vegetation such as maize or soybean is present, this factor becomes crucial when accounting for the hydrogen present around the probe. Vegetation such as trees and grass does not change much over the course of the year. Crops, however, are rapid–growth plants and go from zero biomass to full growth (or 0 to \(\sim 7-10 \text{ mm water equivalents}\)) in a matter of months before being harvested and returning to zero biomass. Therefore, one would expect this rapid–growth vegetation to be a factor in the soil moisture calculations created by a COSMOS probe. After determining a method to account for the forms of above–ground vegetative hydrogen within the footprint, the question was:

*Would rapid–growth vegetation affect the COSMOS probe’s soil moisture calculation?*

Our hypothesis was that:

*The COSMOS probe would be sensitive to the vegetation due to the total mass of vegetative water as well as dry matter hydrogen in the vegetation found in the footprint and the vegetation’s natural seasonal variation.*
1.3.1 Parts of the first research question

We examined two different types of row crop vegetation in 2011 and 2012. We hypothesized that both vegetations would have an effect on the COSMOS probe, but that their effects would differ. This would be due to the naturally different characteristics of the two vegetation types, mainly the total mass of water, but perhaps also their physical structure.

1.4 Upscaling

Upscaling is the practice of using a few soil moisture measurements at a few small points to infer the soil moisture over a larger area. Upscaling can be done with a handful of soil moisture samples, dozens of samples, or more. The area soil moisture is being upscaled to could range from a few square meters to the size of a state county or larger. Taking advantage of the spatial scale of COSMOS probes, we realized we had the unique opportunity to test an upscaling method with the probe. After finding a way to correct (1.1) for vegetation effects, the COSMOS probe’s areal soil moisture could be used as a verification for a field–size resolution upscaling method.

The Feature Space Interpolation (FSI) method is a newly developed method of upscaling. FSI uses point sample measurements of soil moisture to obtain a larger scale view of the soil moisture in an area such as a field. A way to check the soil moisture FSI gives for an area would be to go out and sample the entire area. However, this is unpractical for any area larger than a few square meters. As a COSMOS probe reports the soil moisture within its footprint as one overall areal value, it is a superior method to check the FSI method’s upscaling of point samples.

The FSI method uses the K–means clustering algorithm (MacQueen, 1967) to organize a field–sized array of data collected at many locations into clusters. At the IVS, soil data were obtained at or adjusted to 10 m intervals throughout the field. These data were chosen because of their influences on soil moisture and include horizontal and vertical electromagnetic inductance (EMI), slope, aspect, elevation, and curvature. For this field’s data, the K–means cluster algorithm would divide all 10 m locations into $n$ clusters, as well as determine the
centroid of each cluster based on each location’s data of the above mentioned variables. We assumed that each centroid location was the best representation of the whole cluster.

1.5 Second Research Question

Now that various factors that can affect soil moisture have been determined, we need to test the FSI method. For our work, we will examine cluster groups of 1, 2, 3, and 4 for all locations within the COSMOS probe’s footprint. These clusters will be used to identify at which centroid locations gravimetric samples will be taken. The soil moisture at the centroids will be assigned to the whole cluster and will be used to predict the soil moisture within the COSMOS probe’s footprint which will be compared to the vegetation–corrected COSMOS soil moisture. Our research question was:

*How well does the FSI method’s upscaled soil moisture value match the COSMOS soil moisture for the same day?*

1.5.1 Parts of the second research question

While examining the FSI method’s ability to upscale soil moisture, we will be looking at 4 different cluster groups. We did this in order to test which cluster group can better upscale the soil moisture. We hypothesized that a 3 cluster group would be the optimal arrangement of cluster groups compared to a 1, 2, or 4 cluster group.

A 1 cluster group would rely on one location to determine soil moisture and, while the purpose of the FSI method is to limit the number of sampling points and find the best location to sample at in order to accurately upscale soil moisture, we reasoned that there is too much natural variability in a field to rely on one location to inform on the whole area. A cluster group of 2 could sort locations into high and low elevations or relative clay–heavy or clay–deficient areas, which would be better than relying on one location. However, a 3 cluster group would be able to better separate locations into high, low, and middle (or side slope) elevations or high, medium, and low clay content areas. This would be a better representation of natural field variability. By that logic, a 4 cluster group would be even better than a 3 cluster group as it
could separate locations into high points or low points and north-facing slopes or south-facing slopes. However, we felt that a 4 cluster group would not be significantly better than a 3 cluster group. The difference in soil moisture between a north-facing and south-facing slope would not be as diverse as the difference between a high and low point, or a high and middle point.
CHAPTER 2. SOIL AND VEGETATION MEASUREMENTS

2.1 Introduction

Scientists often are making measurements and collecting data as a natural part of their work. A key factor for data collection is reproducibility. To ensure results are relatable and usable, an agreed upon method of data collection or sampling is critical. Having a consistent way of doing things is therefore important in scientific work. This reduces error and variability due to inconsistencies. For scientists working with naturally changing variables of study, perhaps temporally or spatially varying, this is especially important because they need to know if observed changes are natural effects or are due to some sort of error in the sampling method. In our study, we were working with soil moisture and vegetation changes. Both of these parameters can be spatially and temporally variable so knowing the error associated with the samples is critical. Vegetation and soil moisture can change quickly so that one day of difference in sampling could give different results. It is therefore important to know if the difference in that one day is natural or a sampling error. For the COSMOS probe, we want to know if there is a difference between the in–situ soil moisture and what the probe is reporting for soil moisture. If there is a difference, we want to test if the difference is related to vegetation.

A COSMOS probe reports the volumetric soil moisture within its footprint (Desilets et al., 2010). This soil moisture is measured down to a certain depth of soil, which depends on the current level of soil moisture (Zreda et al., 2008). In order for us to know how accurate the COSMOS probe’s report of soil moisture is, we will have to know what the actual volumetric soil moisture is within the probe’s footprint. The depth of soil seen by the probe (86% cumulative sensitivity) usually varies between 15 and 20 cm (Franz et al., 2012a), so we will need to take soil samples down to 30 cm in order to capture the same depth and a little more in case of drier
soil conditions (Zreda et al., 2008). This will give us a good sample of the in–situ soil moisture within the footprint, which we can compare to what the COSMOS probe calculates the soil moisture to be for that period. If there is a difference in the soil moisture, we want to test if the difference is due to vegetation, meaning we will need a way to quantify the vegetation as well.

Agricultural land crops have a large impact on an area. Crops can affect an individual’s livelihood, the U.S. economy, the weather, and erosion and runoff. Knowing how to model crops is useful because it can provide data for forecasting yields, weather, and more. In this case, we want to be able to model crops so we can quantify the amount of biomass within the COSMOS probe’s footprint. We will need to find a way to determine the equivalent hydrogen content in the vegetation present, as neutron scattering is largely dominated by the presence of hydrogen. Without knowledge on the quantity of vegetation hydrogen present, we won’t be able to determine if any discrepancies in the probe’s soil moisture values are tied to the presence of vegetation.

Modeling crops can be a complex endeavor. Throughout one field, plants can be at different stages of growth at the same time. While one field location may be at growth stage $V_5$ (vegetative stage 5) (Abendroth et al., 2011; Pedersen, 2009), one 15 m away might only be at $V_3$ or maybe $V_6$ (Martin et al., 2005). Field to field variability can be even larger. Different fields are planted at different times and with different hybrids that grow at different rates or under different conditions. Even if a model can be made that accounts for those decision factors over a region, there are more natural factors such as rain, heat stress, soil moisture availability, and wind that can contribute to a plant’s growth and size at any time during the season and can vary across field distances. With soil moisture, locations with more available soil moisture will likely produce larger plants than locations with little plant available soil moisture as the plants have better access to water while growing. Soil moisture itself can be affected by different things, such as soil texture and elevation which can change over a scale of meters. The COSMOS probe has a set sampling schematic that was designed so each sampling location has equal weighted area representation in the overall areal soil moisture. This design was set without studying the natural variability of the soil around the probe. In sampling
vegetation, we will assume that any spatial variability of the soil is also tied to the variability of the vegetation at the same points and that the sampling design accounts for this (Zreda et al., 2012). This is because soil moisture is tied to soil texture and elevation (among other things) and plant growth is tied to soil moisture (among other things).

Thus, in order to collect data to check for a vegetation effect on a COSMOS probe, we need to account for three things.

1. The actual volumetric soil moisture in the top 30 cm of soil to be able to compare it to what the COSMOS probe is reporting.

2. Total plant and/or water mass per area so we can test if it has an effect on the probe.

3. Assume that the given sampling location design will account for spatial variability within the probe’s footprint and that it accurately accounts for the horizontal sensitivity of the COSMOS probe.

2.2 Soil Samples

The accepted way to determine soil moisture is through destructive sampling (Gardner, 1986). This method gives a direct soil moisture which is usually needed or used to calibrate other soil moisture probes. We took destructive soil samples around 18 locations within the COSMOS probe’s footprint on days classified as VC sample days (Vegetation Calibration). These locations were predetermined to be of equal representation of the probe’s footprint. An image of the spread of these locations can be found in Figure 1.2. These 18 locations are divided into 3 rings with radii of 25, 75, and 225 m from the COSMOS probe and into 6 radials of 60 degrees. It should be noted that due to the nature of destructive sampling, we could not sample at the exact same location each time. Therefore, a sample “at” one of the 18 locations is defined to be a sample taken within 2 m of that exact coordinate. Due to the presence of row crops, it should also be stated that in 2011 (maize) we took our soil samples in the middle of the row. With maize plants getting to the size they do and with how unwaivering they are, it was difficult to take soil samples in the rows between them as they grew. The middle of the row became the easiest location for us. In 2012 (soybean), we took our soil samples closer
to the plants than the middle of the row. This was because the soil surface in the middle of
the rows was usually quite hard, as a result of the drought we experienced and the continual
exposure to sunlight. We had a difficult time getting the soil sampler into the ground if we did
not sample closer to the plants. These soil samples therefore may have had a larger presence
of root matter, though it was not quantified.

![Image of soil sampling instrument](image)

**Figure 2.1** The soil sampling instrument used to take soil samples.

Soil samples were taken at the 18 points, down to 30 cm with the instrument shown in
Figure 2.1. As the COSMOS probe’s depth of measurement is usually 10 to 20 cm, this ensures
the soil samples extend to that depth with a cushion of error in case the soil is drier than normal
and the probe reads deeper down than 20 cm. The samples were split at 5 cm increments and
put into separate soil cans. Some compaction of the soil was seen, especially on wetter soil days.
We tried to minimize this error by partitioning the sample out at increments that were not
simply 5 cm long but rather a proper ratio representative of the sample taken. An example of
the instrumentation used for this can be found in Figure 2.2. Five centimeters is a reasonable
divide of the soil profile we took because it allows for fewer cuts of the soil sample which means
less soil that could be lost during the cutting. Also, 5 cm gives some resolution of the soil
profile, which is useful as *Franz et al. (2012b)* found that vertical moisture heterogeneity can
Figure 2.2  PVC pipe soil cores were placed on with meter stick to measure the 5 cm increments. A putty knife was used to divide the cores and place the segments into the cans (both instruments shown in the image).

...affect the depth of measurement. The volumetric soil moisture of a 30 cm soil sample could change considerably from top to bottom if the top portion is saturated or dried out compared to the bottom part. A wetter top layer means a shallower depth seen by the probe while a drier top would result in a deeper depth of measurement (Franz et al., 2012b).

The cans were sealed to minimize moisture and sediment loss and the number of the can was recorded for the site and depth from which it was taken. After all samples were taken, the mass of each of the cans was recorded. This is the wet mass of the samples. The lids were removed and the samples were then placed in a drying oven such as the one shown in Figure 2.3 set at 105°C for 24 hours to dry. After 24 hours, the mass of the entire can was taken again with the now dry soil. This is the dry mass. These masses were recorded, along with the mass of each empty can.
2.2.1 Calculations

We were able to determine gravimetric and volumetric soil moisture as well as bulk density for each 5 cm increment of the 30 cm soil core. The gravimetric soil moisture (g g\(^{-1}\)) was found by dividing the mass of the water by the dry weight of the soil:

$$\theta_g = \frac{m_w}{m_d}$$  \hspace{1cm} (2.1)

where \(\theta_g\) is the gravimetric soil moisture, \(m_w\) is the mass of the water (wet soil mass minus dry soil mass) (g), and \(m_d\) is the dry soil mass (g).

The volumetric soil moisture (cm\(^3\) cm\(^{-3}\)) is related to the gravimetric soil moisture with:

$$\theta_v = \frac{\rho_b}{\rho_w} \theta_g$$  \hspace{1cm} (2.2)

where \(\theta_v\) is the volumetric soil moisture, \(\rho_w\) is the density of water (1 g cm\(^{-3}\)), and \(\rho_b\) is the
bulk density \((g \ cm^{-3})\). The bulk density was found by:

\[
\rho_b = \frac{m_d}{v_s}
\]  \hspace{1cm} (2.3)

or the mass of dry soil divided by the total volume of the sample (air, water, and soil), \(v_s\) \((cm^3)\). The volume of the sample was found using the length of the sample and the size of the sampling probe: 5 cm and radius of 0.953 cm or \(\sim 15 \ cm^3\). The volumetric soil moisture needs to be found in order to relate the in–situ measurements to the measurements of soil moisture created by the COSMOS probe through (1.1). We now have a method for collecting data to determine the real soil moisture in the probe’s footprint. This can be used to compare to the probe’s calculated soil moisture.

2.3 Vegetation

In order to determine if vegetation affects the probe’s soil moisture reading, we will need a way to account for and model the vegetation and its hydrogen content. We decided to use an allometric relationship to find the mass of plants. Allometry is a method of relating different characteristics of an object in order to learn about one facet while making easier measurements. We decided that stem diameter, \(S_d\), and canopy height, \(Z_c\), would be two characteristics we would measure. Stem diameter was defined to be the diameter of the stem of the plant. The location on the plant at which this was measured varied and is explained later. Canopy height was defined to be the height from the ground to the tallest part of the plant. We wanted to use these easily measurable characteristics to help us infer what the mass of the plant would be without having to cut down numerous plants every time we sampled. We chose these characteristics to match previous work done by our research group.

2.3.1 Allometric relationship

In the summer of 2011, our research group had two undergraduate students, Andrew Spencer (University of Idaho) and Crystal Wang (Emory University), working with us. Andrew and Crystal had small research projects they worked on that summer. They looked at finding an
allometric relationship of $S_d$ and/or $Z_c$ to project plant mass. They also researched how many plants were needed to make this relationship.

Using Reddy et al. (1998) as a guide, they adapted the relationship of canopy height to plant mass for their research. They changed the allometric relationship to $S_d$ squared times $Z_c$ instead of just $Z_c$, and looked at the correlation of fresh mass and water mass to this relationship. Their relationship became:

$$m = c_1(S_d^2 Z_c) + c_0$$

(2.4)

where $m$ can be either fresh mass ($m_f$) or water mass ($m_w$), and $c_1$ and $c_0$ are variables tied to the slope and y–intercept of the relationship respectively. They hypothesized that these relationships would be linear. They examined multiple combinations of $S_d$ and $Z_c$ such as $S_d$, $Z_c$, $S_d^2$, $Z_c^2$, $S_d Z_c$, $S_d^2 Z_c$, and $S_d Z_c^2$. The relationship of $S_d^2 Z_c$ was chosen because it was a rough volumetric model of the plants and had the best correlation to mass in the initial runs they made that season. $S_d$ was squared because they noted that while leaves will contain water, $S_d$ will have a greater percentage of water and thus a greater impact on the vegetative water content. Niklas and Enquist (2002) showed that leaf mass can be scaled as stem diameter squared, which also supports the use of $S_d^2 Z_c$ as it would account for the stem and leaves as related to plant height. Andrew and Crystal kept the $S_d^2 Z_c$ relationship when looking at fresh mass as well so that they could test the same principle for the two parameters.

Andrew and Crystal also determined the number of plants needed to make a stable relationship. They measured 10-40 plants’ $S_d$ and $Z_c$ and then measured those plants’ fresh masses before drying them out to determine their respective dry masses and water contents. They input this data to a software program, MATLAB® (2010a, MathWorks, Natick, Massachusetts), to select random observations, ran it 30 times per $n$ value (10-40), and checked the coefficient of correlation of the relationship (2.4). If the coefficient changed by more than the second significant figure, the relationship was determined to be unstable. Stability was obtained at $n = 30$ plants and maintained for $n > 30$. Therefore, they chose 30 plants as the necessary number of plants needed for a stable relationship.
2.3.2 Vegetation measurements

To ensure our data could be comparable, we needed to have a consistent method of making measurements. We would be making measurements of the soil moisture as well as maize and soybean vegetation over the course of two summer seasons. Maize and soybean plants are inherently different from one another, so the same method of making measurements could not be used in 2011 and 2012. At each of the 18 sampling locations (Figure 1.2), we measured vegetation as well as soil moisture on the VC days. We measured five consecutive plants in 2011 (maize) that were chosen to be the best representation of that location. In 2012 (soybean), we measured ten plants. These plants corresponded to row distances of \( \sim 0.88 \) m for maize and \( \sim 0.41 \) m for soybean. The \( S_d \) and \( Z_c \) values of these plants would be averaged together for a single representation of vegetation at that location. Martin et al. (2005) showed in their work with maize that plant variation decreases when averaging plants together in groups of four. As their work was with maize and averaged at most four plants together, we measured five plants, assuming to decrease the variation even more. For soybean, we knew they would be planted more densely than maize so we decided to double the number of plants measured.

For maize plants, the stem is an ellipse, not a perfect circle. For the stem diameter of maize plants, we measured the diameter along the thick diameter of the stem using a digital caliper (shown in Figure 2.4). At the beginning of the crop season, from vegetative stage \( V_e \) (emergence) to \( V_t \) (tassel), we measured the stem diameter immediately below the highest collared leaf. We assumed that at the early stages of growth, this area would be easy to distinguish and would change rapidly as the plants changed vegetative stages. However, as the maize plants reached maturity, the highest collared leaf was no longer easy to distinguish nor easy to reach. We assumed that area of the plant would not be changing as much, as the plant was focusing on developing the areas where the reproductive ears are now. Therefore, after the plants reached reproductive stages \( R_1 \) to \( R_6 \), we measured maize stem diameter directly below the first collared leaf above the primary ear.

For soybean plants in 2012, stem diameter was measured at roughly the same location on the plants throughout the vegetative stages with a digital caliper. Soybean plants can have
multiple branches as they grow and no two plants grow the exact same, unlike maize plants that consistently have the one stem and all leaves and ears come off of that stem. The multiple branches of soybean plants meant that it would be difficult to determine which branch was the main stem as the plants grew. To be able to have a consistent location for stem diameter measurement on every soybean plant, we kept our stem diameter measurements below the first junction of branches. Therefore, our measurement location was above the dicote of each plant or, if that was no longer on the plant, below the lowest trifoliate. When the plants had reached the early reproductive stages, this location was still easily distinguishable so it remained the location to measure stem diameter. For the September 25 field day, however, the plants had all reached the last reproductive stage for soybeans and no longer had any leaves on them. The easily discernible location to measure stem diameter at for each plant became the lowest joint in the stem, which was roughly the same spot the dicote would have existed had the leaves not all fallen off.

Canopy height for maize plants was considered to be the distance from the ground to the highest point of the plant. In the early vegetative stages of growth, we used a tape measure to measure this distance. In the later vegetative stages and reproductive stages, this became more difficult to do as the plants had reached heights above two meters. We then created a pole with centimeter tick marks on it that could be held up to each plant and read off by looking up. This introduced some error, as we could not get an eye–level reading of the height. However, it would have been difficult to carry a stool or ladder throughout the field to assist us and get us level with the top of each plant so this was a systematic error we did not concern ourselves with trying to fix.

Soybean canopy height measurements were made with the same instrument throughout the season. We used a 1.6 m aluminum conduit pole with a metric tape measure adhered along its length, pictured in Figure 2.4. Canopy height was again defined as the distance between the ground and the highest point of the plant. Unlike maize plants though, which grow almost completely straight up, soybean plants grow more outward. Therefore, a simple measurement that was based on the height of the natural plant was determined to not be as easy or accurate as it was for maize plants. We gently extended the entire soybean plant straight up and marked
the canopy height as the tallest point any part of the plant stretched straight up to.

Now that methods had been determined for the stem diameter and canopy height parameters, we needed a method for determining the mass of a plant so that we could relate plant mass to stem diameter and canopy height. To do this, we destructively removed 30 plants from the edge of the field, again on these VC days. We took plants similar in size and growth stage to those we saw in the center of the field. We took plants from the edge instead of the center of the field because it made our work easier. We did not have to carry 30 plants, along with our sampling equipment, out from the center of the field. Most importantly, this meant that there was less of a possibility of losing parts of the plant, and lessened the time between removing the plant and finding its mass so that the least moisture possible was lost from the plant.

For the 30 plants, we measured stem diameter and canopy height in the same methods as described previously before cutting them down. We then took the plants to a farm dryer. The plants were placed in paper bags to dry for a minimum of three days at 60°C. Depending on current vegetative stage, we divided the plants into their respective leaves, stems, and reproductive parts so that each plant took up two to three paper bags. For each batch of paper
bags that we used, we found the original mass of 10-20 empty bags, and put those bags in the
cart with the plant bags to be dried. The dried mass was found for these 10-20 bags of the
same batch and that data was used to subtract pre–dried bag and dried bag mass from the
combined plant and bag masses. This left us with the fresh mass and dried mass of each of
the 30 plants’ stems, leaves, and reproductive parts which, when combined, gave us the total
masses for the individual plants. The plants’ water masses could be found from this data where
\( m_{total} = m_f = m_d + m_w \) (fresh mass is equal to dry mass and water mass).

### 2.3.3 Total biomass

As it had already been determined that the relationship of (2.4) was the most accurate for
prediction of a plant’s mass, we stayed consistent and used that relationship throughout our
work. We used this relationship for all three categories of plant mass: fresh mass, dry mass,
water mass for each group of 30 plants removed from the edge of the field. Some example
figures of this are shown in Figure 2.5. The slopes \( c_1 \) from (2.4), y–intercepts \( c_0 \) from (2.4),
and \( R^2 \) values for each type of plant mass taken on each VC day are shown in Table 2.1 as a
display of the variability and patterns seen. This allometric relationship allowed us to project
the mass of a plant based on measurements of stem diameter and canopy height. We used
this method to project a mass for the composite 5 maize/10 soybean plants at each of the 18
locations using their mean \( S_d \) and \( Z_c \) values.

However, we needed to be able to look at the total mass of plants in the footprint. To find
this, we needed to scale our measurements up by plant density to give us the areal vegetation
mass of the field. Plant density, \( \rho_P \), was determined by

\[
\rho_P = \frac{n_p \cdot n_r}{L \cdot W}
\]

where \( \rho_P \) is plant density (plants m\(^{-2}\)), \( n_p \) is the number of plants counted in one row, \( n_r \) is
the number of rows measured across, \( L \) is the length of the row in which we counted our \( n_p \)
(m), and \( W \) is the distance across \( n_r \) rows (m). This gave us a value in plants per area, which
we averaged among the 18 sites for a field density. When the field average plant density is
multiplied by our projected plant masses \( m \), found with (2.4) relationships) (kg plant\(^{-1}\) we
Figure 2.5  A few examples of the allometric relationship of \( c_1 (S_d^2 Z_c) + c_0 = m \) for fresh, dry, and water masses. There were 30 plants taken each VC day we sampled vegetation to be used in this relationship. The masses of the plants around the COSMOS probe were projected with these relationships and the plants’ respective \( S_d \) and \( Z_c \) values. Error bars shown are error due to instrument errors.
Table 2.1 Allometric relationship data. Slope \((c_1)\) (kg plant\(^{-1}\) m\(^{-3}\)), y–intercept \((c_0)\) (kg plant\(^{-1}\)), and R\(^2\) data for 2011 maize measurements and 2012 soybean measurements when comparing the relationship of
\[c_1(S_d^2 Z_c) + c_0 = m\]
for the 30 plants removed from the edge of the field on VC days.

<table>
<thead>
<tr>
<th>Date</th>
<th>Fresh mass ((m_f))</th>
<th>Dry mass ((m_d))</th>
<th>Water mass ((m_w))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>Intercept</td>
<td>(R^2)</td>
</tr>
<tr>
<td>2011 - Maize</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June 24</td>
<td>266</td>
<td>0.0076</td>
<td>0.90</td>
</tr>
<tr>
<td>July 8</td>
<td>366</td>
<td>-0.0078</td>
<td>0.80</td>
</tr>
<tr>
<td>August 18</td>
<td>506</td>
<td>0.47</td>
<td>0.69</td>
</tr>
<tr>
<td>September 13</td>
<td>466</td>
<td>0.46</td>
<td>0.77</td>
</tr>
<tr>
<td>2012 - Soybean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June 19</td>
<td>1260</td>
<td>0.00085</td>
<td>0.90</td>
</tr>
<tr>
<td>July 6</td>
<td>934</td>
<td>0.0073</td>
<td>0.84</td>
</tr>
<tr>
<td>July 20</td>
<td>1290</td>
<td>0.0024</td>
<td>0.95</td>
</tr>
<tr>
<td>August 1</td>
<td>1480</td>
<td>-0.0031</td>
<td>0.92</td>
</tr>
<tr>
<td>August 15</td>
<td>1650</td>
<td>0.0061</td>
<td>0.93</td>
</tr>
<tr>
<td>September 25</td>
<td>360</td>
<td>0.013</td>
<td>0.62</td>
</tr>
</tbody>
</table>
can scale up to a mass per area value, $M$ (kg m$^{-2}$) (2.6) for each location.

$$M = \rho_p \cdot m$$  (2.6)

To have a sole field mass per area value, we then averaged the 18 $M$ values together. This relationship was used for all 3 different masses, giving us an $M_f$, $M_d$, and $M_w$ for each sample day (areal fresh matter, areal dry matter, and areal water mass respectively).

2.3.4 Results

It can be seen in Table 2.1 that our $y$–intercepts ($c_0$ values) were not zero as they should have been for a linear relationship. This could have affected our mass projections, which would have propagated to our areal masses. Our $R^2$ values were mostly above 0.70 for both years and all mass types. For modeling naturally grown variables, we thought this was good. Our (2.4) relationship did the best with the soybean plants on all but the senescent vegetation day. The $R^2$ values for the growing soybean sampling days were all above 0.80 which shows that this relationship would be good at predicting the mass of a plant. The $R^2$ values for maize were not as good as for soybean; however, to be consistent in our work, we kept the (2.4) relationship when looking at projecting a maize plant’s mass despite this.

Our $M$ values for 2011 and 2012 are shown in Figures 2.6 and 2.7, respectively. Error bars shown and how we determined them will be discussed in the following section. Focusing on the maize relationship (Figure 2.6), we can see that dry matter peaked during the senescent stage of growth (September 13, 2011) while fresh matter and water content both peaked during the date identified to be the peak vegetation sampling date (August 18, 2011). It can also be seen that water composed a majority of the plants’ mass toward the beginning of the season as the plants were starting to grow and develop. For soybean plants (Figure 2.7), this growth relationship was found to behave similarly. The plants initially were composed of mostly water as the plants were developing and, as they grew, took up less water and accumulated more dry matter. It should be noted that it is normal for soybean plants to lose their leaves during the last reproductive stages. This was observed in our collection of data. On our senescent vegetation day, September 25, 2012, all the plants we sampled had lost all of their leaves.
Figure 2.6 Graph of the areal maize vegetation masses over the field in 2011. Data points were determined by (2.6) with error bars found with (2.7).

Figure 2.8 shows what the soybean plants looked like on this day. In Figure 2.7 this loss of leaves can be seen with a plateau of dry matter from the observed peak vegetation and senescent vegetation sampling dates, whereas the maize had a peak of dry matter during the senescent stage. This was plausible because the soybean plants should have been accumulating more dry matter during this time as their pods finished developing; however, with the leaves falling off, the plants were also losing dry matter which results in a plateau of dry matter.

2.4 Vegetation Variability

Little work has been done on the variability of plants throughout a field. Martin et al. (2005) noted many variables that can contribute to extensive plant variability over distances as small as 30 m and even 1 m (Lengnick, 1997; Raun et al., 1998; Solie et al., 1999). Some of these are: depth of planting, surface crusting, soil texture, random soil clods, distance between seeds, insect damage, moisture availability, and surface residue. With all these factors and more present in fields, within row plant variability should be expected. However, despite all these causes of spatial variability in plants, Porter et al. (1998) showed that the spatial variability of
Figure 2.7  Areal soybean vegetation masses during 2012. Data points were determined by (2.6) with error bars found with (2.7). Note that soybean plants lose their leaves as they senesce. This effect can be seen with the lack of a peak in the dry matter during the senescent sampling day and by a plateau of dry matter instead. As the plants were losing their leaves while the pods matured, the dry matter stayed at a more constant level than it did for senescent maize plants.

yield in one year is three times less than the temporal variability found in yield for soybeans and four times less than the temporal yield variability seen in maize.

Martin et al. (2005) examined by–plant spatial yield variability in maize. They measured the distance between consecutive plants and then harvested each plant and determined the by–plant yield in 15 m transects of rows. In their 2004 work in Ames, IA, they examined by–plant yields and variability as well as yields and variability averaged over two, three, and four plants. They found that as they averaged the yield over more plants, the difference in the yield prediction decreased. The ratio of maximum yield prediction to minimum yield prediction went from 10.6 to 2.6 when looking at one plant versus four plants averaged. Looking at the corresponding row distances of these transects of plants, they found that the error of the yield prediction approached constant at distances > 1 m and that the true yield mean of the row could be estimated at a distance of 0.5-0.6 m. Our 5 maize plants per location fit into this area
with a mean row distance of $\sim 0.88$ m, allowing us to claim we had enough data for an estimate of a mean value. *Niklas and Enquist (2002)* also found that when taking a representative sample of plants for use with allometric relationships, five plants is enough for a data set.

We also did our own examination of the error associated with our 5 maize/10 soybean plant measurements due to the natural variability of the vegetation. As it would be impractical as well as impossible to measure every single plant at the IVS, we had to find a method that still allowed us to gather sufficient data so that we felt we had captured an area and its natural variability. We decided to choose three locations in a maize and a soybean field where we would take measurements (Figure 2.9) on days classified as PV days (Plant Variability). These three areas were chosen so that they were distinctly different field conditions. For example, one area was a depression, one was on a hill, and the third on a hill slope. This would allow us to characterize any difference in plant growth depending on field location. In the maize field, there were two hybrids planted and one of the three locations in this field was located in a different hybrid than the other two locations. At each location, we measured 15 consecutive plants in seven consecutive rows. This amounted to an area of $\sim 14$ m$^2$ for maize and $\sim 3.2$ m$^2$
for soybean, which we felt was small enough that there was no significant soil or topography variations yet still had enough plants that we felt we measured a large sample. These 105 plants were defined to give the true field average of our variables at that one location. Some examples of the natural variation in these 105 plants and their parameters are shown in Figure 2.10.

Our hypothesis was that vegetation variability would only be due to unaccountable error which would be a combination of such things as genetics and random planting differences. We assumed that plants experience the same rainfall, same solar radiation, and same soil properties. We therefore expected the vegetation variability to depend on three things:

1. The error in projections of individual plant mass due to $S_d^2 Z_c$ fits on VC days.

2. The in–field spatial variability of plants.

3. The intrinsic variability of the plants at each location in the field.

Item 1 could be tracked with data on our 30 plant relationship. Item 2 would be accounted for by our assumption that the 18 sampling locations would capture this spatial variability already. Item 3 would still need to be determined. We will therefore be focusing on items 1 and 3 in the following section.

(a) Sampling locations for the maize field.  (b) Sampling locations for the soybean field.

Figure 2.9  The sampling locations for each of the fields used for PV days. In Figure 2.9(a), Location 1 is the center dot, Location 2 is the southern dot, Location 3 is the northern dot. In Figure 2.9(b), Location 1 is the northern dot, Location 2 is the southeast dot, Location 3 is the west dot.
Figure 2.10  Some examples of the original plant measurements. The mean for all the figures is shown with the dashed black line while each green bar represents a different plant. We measured 15 consecutive plants in 7 adjacent rows in both the maize and soybean fields. ($S_d$ being stem diameter in mm and $Z_c$ being canopy height in cm).
2.4.1 Methods

To find the error in our vegetation amounts ((2.6), Figures 2.6 and 2.7), we had to find the error in its two parts: plant density and the (2.4) by–plant mass values. Using Beers (1962) as a guide and assuming independence between $\rho_P$ and $m$, the error of (2.6) is

$$\sigma_M = \sqrt{\sigma_{\rho_P}^2 \cdot m^2 + \sigma_m^2 \cdot \rho_P^2}$$  \hspace{1cm} (2.7)

where $\sigma_{\rho_P}$ is the error associated with plant density, $m$ is the projected plant mass (can be $m_f$, $m_d$, or $m_w$), $\sigma_m$ is the error due to the projected plant mass (again $m_f$, $m_d$, or $m_w$), and $\rho_P$ is plant density. As we wanted the error estimate of our mean values for $M$ and not their standard deviations, we divided $\sigma_M$ by the square root of the number of locations at which samples were taken and averaged over (18) to obtain the standard error of our estimated $M$ values.

2.4.1.1 Plant Density

Plant density was found to be $7.43 \pm 0.52$ plants m$^{-2}$ for maize and $32.4 \pm 1.6$ plants m$^{-2}$ for soybean at the 18 locations displayed in Figure 1.2. These were found with (2.5). $\sigma_{\rho_P}$ is the standard deviation of the 18 measurements (shown as the $\pm$ values). However, we wanted the standard error of our measurements, not just the standard deviation. Therefore, for $\sigma_{\rho_P}^2$ we used the standard deviation of the 18 measurements divided by the square root of the number of samples taken.

2.4.1.2 Individual Plant Mass

The second part of determining the error in our $M$ values has to do with finding the error due to the individual plant mass used as $m$ in (2.6). This will be composed of two parts: how accurate a projected mass might be due to the 30 plant relationship (2.4), and the natural variability of plants. The error due to $m$ ($\sigma_m^2$) will be found by

$$\sigma_m^2 = \text{var}(c_0) + \text{var}(c_1)\text{var}(\pi) + \text{var}(c_1)\pi^2 + \text{var}(\pi)c_1^2 + 2\text{cov}(c_0,c_1)$$  \hspace{1cm} (2.8)

where $m$ is the mass of interest ($m_f$, $m_d$, or $m_w$), $\text{var}(c_0)$ is the variance of a specific 30 plant fit’s intercept, $\text{var}(c_1)$ is the variance of a specific 30 plant fit’s slope, $\pi$ is an abbreviation for
the mean $S^2_d Z_c$, $\var(\bar{x})$ is tied to the natural plant-to-plant variation, and $\cov(c_0, c_1)$ is the covariance of the slope and intercept for a specific 30 plant fit.

The $\var(c_0)$, $\var(c_1)$, and $\cov(c_0, c_1)$ terms are all related to the goodness of the 30 plant mass fits, or item 1 in our list of vegetation variability factors. These variances and covariances were found with a simple analysis of the (2.4) relationships in RStudio® (RStudio, Inc., Boston, Massachusetts). For each VC sampling day that vegetation measurements were made (four dates in 2011, six dates in 2012), and each mass type ($m_f, m_d, m_w$), there were different $\var(c_0)$, $\var(c_1)$, and $\cov(c_0, c_1)$ values. These are shown in Table 2.2.

The remaining error in $m$ is tied to plant-to-plant variability (item 3 in our list). For this, we will use our PV data on the 3 locations of 105 plants for both maize and soybean, which we defined as our “true” field averages. Summaries of the natural development and change in $S_d, S^2_d, Z_c,$ and $S^2_d Z_c$ for the three maize locations on three PV sample days are shown in Table 2.3. Similarly, the results of the natural variation of $S_d, S^2_d, Z_c,$ and $S^2_d Z_c$ for the three soybean locations are displayed in Table 2.4.

To determine the variance due to plant variability ($\var(\bar{x})$) and justify our choice of using only 5 maize/10 soybean plants to sample at each location, we will define $\var(\bar{x})$ to be

$$\var(\bar{x}) = \frac{\eta^2}{n} \quad (2.9)$$

where $\eta$ is our best estimate of the intrinsic plant variability (due to genetics and random planting differences), and $n$ is the number of plants we chose to measure at each of the 18 sampling locations (5 maize, 10 soybean).

### 2.4.2 Results

It can be seen that all the $c_1$ and $c_0$ characteristics had a negative covariance (Table 2.2). For soybean, $\var(c_1)$ for the senescent sampling day was noticeably different from the rest. For fresh mass and water mass, $\var(c_1)$ was the smallest while for dry mass it was the largest. This was not as noticeable in maize.

As these PV sampling dates differed from the previous VC sampling dates, we needed to determine a way to use these vegetation measurements together. For both maize and soybean,
Table 2.2 Variances and covariances found for each of the 30 plant mass fits made in 2011 and 2012.

<table>
<thead>
<tr>
<th>Date</th>
<th>Fresh mass ($m_f$)</th>
<th>Dry mass ($m_d$)</th>
<th>Water mass ($m_w$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>var($c_1$)</td>
<td>var($c_0$)</td>
<td>covar($c_0, c_1$)</td>
</tr>
<tr>
<td>2011 - Maize</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June 24</td>
<td>275</td>
<td>6.38 e-5</td>
<td>-0.127</td>
</tr>
<tr>
<td>July 8</td>
<td>1,174</td>
<td>9.58 e-4</td>
<td>-1.03</td>
</tr>
<tr>
<td>August 18</td>
<td>4,112</td>
<td>3.0 e-3</td>
<td>-3.34</td>
</tr>
<tr>
<td>September 13</td>
<td>2,333</td>
<td>1.2 e-3</td>
<td>-1.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012 - Soybean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June 19</td>
<td>6,568</td>
<td>2.67 e-7</td>
<td>-0.037</td>
</tr>
<tr>
<td>July 6</td>
<td>5,963</td>
<td>4.94 e-6</td>
<td>-0.144</td>
</tr>
<tr>
<td>July 20</td>
<td>3,299</td>
<td>5.82 e-6</td>
<td>-0.114</td>
</tr>
<tr>
<td>August 1</td>
<td>6,362</td>
<td>2.8 e-5</td>
<td>-0.355</td>
</tr>
<tr>
<td>August 15</td>
<td>6,968</td>
<td>2.4 e-5</td>
<td>-0.345</td>
</tr>
<tr>
<td>September 25</td>
<td>2,826</td>
<td>9.22 e-6</td>
<td>-0.128</td>
</tr>
</tbody>
</table>
Table 2.3  The resulting maize means of the 105 plants at each location where \( \bar{S}_d \) (mm) is the mean stem diameter (\( \sigma_{sd} \) standard deviation), \( \bar{Z}_c \) (mm) is the mean canopy height (\( \sigma_{zc} \) standard deviation), \( S^2_d \) (mm\(^2\)) is the mean stem diameter squared (\( \sigma_{sd^2} \) standard deviation), and \( S^2_d \bar{Z}_c \) (mm\(^3\)) is the mean stem diameter squared times canopy height (\( \sigma_{sd^2 \cdot zc} \) standard deviation).

<table>
<thead>
<tr>
<th></th>
<th>( \bar{S}<em>d ) (( \sigma</em>{sd} ))</th>
<th>( \bar{Z}<em>c ) (( \sigma</em>{zc} ))</th>
<th>( S^2_d ) (( \sigma_{sd^2} ))</th>
<th>( S^2_d \bar{Z}<em>c ) (( \sigma</em>{sd^2 \cdot zc} ))</th>
<th>( \bar{S}<em>d ) (( \sigma</em>{sd} ))</th>
<th>( \bar{Z}<em>c ) (( \sigma</em>{zc} ))</th>
<th>( S^2_d ) (( \sigma_{sd^2} ))</th>
<th>( S^2_d \bar{Z}<em>c ) (( \sigma</em>{sd^2 \cdot zc} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>July 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location 1</td>
<td>29.8 (5.7)</td>
<td>2 000 (250)</td>
<td>23.8 (3.6)</td>
<td>2 480 (150)</td>
<td>23.1 (3.7)</td>
<td>2 470 (150)</td>
<td>22.9 (3.0)</td>
<td>2 550 (140)</td>
</tr>
<tr>
<td>Location 2</td>
<td>27.5 (4.2)</td>
<td>2 050 (160)</td>
<td>23.4 (3.7)</td>
<td>2 510 (130)</td>
<td>22.9 (3.0)</td>
<td>2 550 (140)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location 3</td>
<td>28.1 (6.7)</td>
<td>1 970 (420)</td>
<td>25.4 (5.1)</td>
<td>2 560 (300)</td>
<td>25.1 (4.7)</td>
<td>2 600 (220)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>( S^2_d ) (( \sigma_{sd^2} ))</th>
<th>( S^2_d \bar{Z}<em>c ) (( \sigma</em>{sd^2 \cdot zc} ))</th>
<th>( S^2_d ) (( \sigma_{sd^2} ))</th>
<th>( S^2_d \bar{Z}<em>c ) (( \sigma</em>{sd^2 \cdot zc} ))</th>
<th>( S^2_d ) (( \sigma_{sd^2} ))</th>
<th>( S^2_d \bar{Z}<em>c ) (( \sigma</em>{sd^2 \cdot zc} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location 1</td>
<td>919 (326)</td>
<td>1 880 000 (760 000)</td>
<td>579 (170)</td>
<td>1 450 000 (460 000)</td>
<td>549 (168)</td>
<td>1 370 000 (450 000)</td>
</tr>
<tr>
<td>Location 2</td>
<td>773 (233)</td>
<td>1 610 000 (560 000)</td>
<td>559 (167)</td>
<td>1 410 000 (440 000)</td>
<td>533 (134)</td>
<td>1 370 000 (370 000)</td>
</tr>
<tr>
<td>Location 3</td>
<td>833 (348)</td>
<td>1 750 000 (860 000)</td>
<td>671 (247)</td>
<td>1 760 000 (710 000)</td>
<td>654 (224)</td>
<td>1 730 000 (640 000)</td>
</tr>
</tbody>
</table>
Table 2.4  The resulting soybean means of the 105 plants at each location where $S_d$ (mm) is the mean stem diameter ($\sigma_{sd}$ standard deviation), $Z_c$ (mm) is the mean canopy height ($\sigma_{zc}$ standard deviation), $S_d^2$ (mm$^2$) is the mean stem diameter squared ($\sigma_{sd^2}$ standard deviation), and $S_d^2 Z_c$ (mm$^3$) is the mean stem diameter squared times canopy height ($\sigma_{sd^2 zc}$ standard deviation).

<table>
<thead>
<tr>
<th></th>
<th>June 28</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_d$ ($\sigma_{sd}$)</td>
<td>$Z_c$ ($\sigma_{zc}$)</td>
<td>$S_d$ ($\sigma_{sd}$)</td>
<td>$Z_c$ ($\sigma_{zc}$)</td>
</tr>
<tr>
<td>Location 1</td>
<td>4.1 (1.1)</td>
<td>340 (55)</td>
<td>4.9 (1.3)</td>
<td>580 (64)</td>
</tr>
<tr>
<td>Location 2</td>
<td>4.7 (1.5)</td>
<td>380 (91)</td>
<td>5.1 (2.1)</td>
<td>620 (76)</td>
</tr>
<tr>
<td>Location 3</td>
<td>4.0 (1.1)</td>
<td>350 (66)</td>
<td>4.6 (1.8)</td>
<td>580 (95)</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_d^2$ ($\sigma_{sd^2}$)</td>
<td>$S_d^2 Z_c$ ($\sigma_{sd^2 zc}$)</td>
<td>$S_d^2$ ($\sigma_{sd^2}$)</td>
<td>$S_d^2 Z_c$ ($\sigma_{sd^2 zc}$)</td>
</tr>
<tr>
<td>Location 1</td>
<td>18 (9.1)</td>
<td>6 600 (3 800)</td>
<td>25 (12)</td>
<td>15 000 (7 500)</td>
</tr>
<tr>
<td>Location 2</td>
<td>24 (15)</td>
<td>9 700 (6 500)</td>
<td>30. (20.)</td>
<td>19 700 (13 700)</td>
</tr>
<tr>
<td>Location 3</td>
<td>17 (9.5)</td>
<td>6 300 (4 200)</td>
<td>24 (16)</td>
<td>15 000 (10 500)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>June 28</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$S_d$ ($\sigma_{sd}$)</td>
<td>$Z_c$ ($\sigma_{zc}$)</td>
<td>$S_d$ ($\sigma_{sd}$)</td>
<td>$Z_c$ ($\sigma_{zc}$)</td>
</tr>
<tr>
<td>Location 1</td>
<td>5.2 (1.5)</td>
<td>810 (100)</td>
<td>5.2 (1.5)</td>
<td>810 (100)</td>
</tr>
<tr>
<td>Location 2</td>
<td>6.0 (2.1)</td>
<td>830 (95)</td>
<td>6.0 (2.1)</td>
<td>830 (95)</td>
</tr>
<tr>
<td>Location 3</td>
<td>5.2 (1.4)</td>
<td>790 (74)</td>
<td>5.2 (1.4)</td>
<td>790 (74)</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_d^2$ ($\sigma_{sd^2}$)</td>
<td>$S_d^2 Z_c$ ($\sigma_{sd^2 zc}$)</td>
<td>$S_d^2$ ($\sigma_{sd^2}$)</td>
<td>$S_d^2 Z_c$ ($\sigma_{sd^2 zc}$)</td>
</tr>
<tr>
<td>Location 1</td>
<td>29 (13)</td>
<td>24 700 (12 200)</td>
<td>29 (13)</td>
<td>24 700 (12 200)</td>
</tr>
<tr>
<td>Location 2</td>
<td>40. (24)</td>
<td>34 200 (21 900)</td>
<td>40. (24)</td>
<td>34 200 (21 900)</td>
</tr>
<tr>
<td>Location 3</td>
<td>29 (13)</td>
<td>23 400 (11 600)</td>
<td>29 (13)</td>
<td>23 400 (11 600)</td>
</tr>
</tbody>
</table>
we examined the variance of these PV days’ 105 plants dependent on location and sampling date. The hope was that we would find that the variances did not change significantly over time or location, so that we could assume a constant variance of vegetation due to growth ($\eta$). To examine these variances, we made side-by-side boxplots of the variances of each group of 105 plants’ $S_d^2 Z_c$ relationships. These are shown in Figure 2.11 for maize and Figure 2.12 for soybean. For maize, there is no easily distinguishable pattern to the variances. This implies that any variance seen in maize is due to random variables of plant growth and not the location of planting or the day of sampling. Therefore, we were able to use a constant value for $\eta$, identified to be the residual standard error (RSE) of the composite 945 maize plants’ $S_d^2 Z_c$ data (936 degrees of freedom).

Maize

![Maize Boxplot](image)

**Figure 2.11** Side–by–side boxplot of the variance in $S_d^2 Z_c$ of the 105 plants measured at 3 locations on 3 days in a maize field. Outliers are shown as open circles, the dark horizontal lines represent the median values, upper and lower box edges mark the first and third quartiles, and the whiskers represent the respective upper and lower 25% of the data. Horizontal axis points are read as “Location, Date” where 1,1 would represent maize location 1, on sampling date 1 (June 26, 2012).

For soybean it can be seen in the boxplot that there are some patterns in the variance.
Figure 2.12 Side–by–side boxplot of the variance in $S^2_d Z_c$ of the 105 plants measured at 3 locations on 3 days in a soybean field. Outliers are shown as open circles, the dark horizontal lines represent the median values, upper and lower box edges mark the first and third quartiles, and the whiskers represent the respective upper and lower 25% of the data. Horizontal axis points are read as “Location, Date” where 1,1 would represent soybean location 1, on sampling date 1 (June 28, 2012).

For instance, the variances can be seen to be related to the sampling date. We reasoned that the difference here between the maize plants and soybean plants has to do with maize being a determinate plant while soybean is an indeterminate plant. If a maize plant loses a leaf or ear during growth, that leaf or ear does not grow back. Maize plants stop growing at a predetermined stage. They grow one main ear in a season. Soybean plants, however, can continuing producing pods and there is no determined maximum number of pods that can be developed, so one soybean plant could have 20 pods while another could only have 15. If a soybean plant loses a leaf or pod at some point during growth, it may regrow that part. Therefore, we were not surprised that the soybean variances changed more noticeably throughout the season. As time goes on, the chance of soybean plants having a different number of pods or leaves gets larger and the difference itself could get larger.
A dependence on sampling location can also be seen in Figure 2.12. Discovering that the soybean plant variation did depend on the date and location of sampling made matters more complicated as we would have to find different variances for different days in the season and sampling locations. As the PV sampling dates did not match with the VC sampling dates, it would have been difficult to relate the variances. However, for simplicity and consistency, we decided to ignore this variance variability and again use the RSE with 936 degrees of freedom as our soybean $\eta$ value. We therefore found or assumed that the variability in item 3 (random variation) was equal to $\eta$.

To see how our choice of $n$ influenced our error, we plotted $\sigma^2_m$ as a function of $n$. Some examples are shown in Figure 2.13. It can be seen that at a certain $n$ value ($\sim 30$), $\sigma^2_m$ levels out and becomes essentially constant. This means that any number of plants sampled greater than $\sim 30$ will not give any better approximation of the vegetation. It could be said that we therefore introduced a larger error in $M$ by only choosing 5 maize/10 soybean plants to measure per location. However, a look at the difference between $m$ values and corresponding $\sigma^2_m$ values shows that $\sigma^2_m$ is consistently at least $10^{-2}$ less than $m$. For instance, maize $m$ values ranged from $10^{-2}$-$10^{-1}$ kg plant$^{-1}$ and maize $\sigma^2_m$ values ranged from $10^{-4}$-$10^{-3}$ kg plant$^{-1}$. In the case of soybean, $m$ values ranged from $10^{-3}$-$10^{-2}$ kg plant$^{-1}$ while $\sigma^2_m$ ranged from $10^{-7}$-$10^{-4}$ kg plant$^{-1}$. Thus, we felt that we had managed to capture $\sigma^2_m$ well.

We can now quantify the error of $M$ with (2.7). Using an $n$ of 5 for our 2011 maize days and an $n$ of 10 for our 2012 soybean days, we were able to solve (2.7) and (2.8) to find the error bars in $M$ (Figures 2.6 and 2.7). For maize, our $M$ values were $10^{-1}$ to $10^0$ kg m$^{-2}$, and $\sigma_M$ values were $10^{-2}$ to $10^{-1}$ kg m$^{-2}$. For soybean, our $M$ values were $10^{-2}$ to $10^0$ kg m$^{-2}$, and $\sigma_M$ values were $10^{-3}$ to $10^{-2}$ kg m$^{-2}$. The highest errors seen was 0.24 kg m$^{-2}$ for maize and 0.050 kg m$^{-2}$ for soybean which were respectively equal to 3.8% and 2.2% compared to the $M$ values for those days. Both of these were for fresh mass and they were for August 18, 2011 and August 15, 2012, respectively. If these were water equivalents, that is an error of at most 0.24 mm of water. With our soil samples being taken at 5 cm increments, an error of 0.24 mm does not make any larger of a difference as we can only be accurate to a 5 cm $z^*$. 
Figure 2.13 A few examples of how $\sigma_m^2$ varied depending on $n$ - the number of plants sampled (all units kg plant$^{-1}$). After $n = 30$, $\sigma_m^2$ levels out and becomes almost constant.
2.5 Conclusions

We now have a method for sampling soil moisture. This will allow us to compare the COSMOS probe’s reported soil moisture to the in–situ soil moisture. By being able to find the in–situ soil moisture, we can recalibrate the probe and find the corresponding \( N_0 \). If the recalibrated \( N_0 \) values are constant, then the probe’s neutron counts are not affected by the presence of vegetation and \( N_0 \) can be calibrated just one time. If, however, \( N_0 \) is found to vary, then work will need to be done to take the presence of vegetation into account.

With this work, we also have a method for measuring vegetation. By using an allometric relationship, we can find the mass of an individual plant without removing it from the field. This involves removing some plants, but we chose these plants so that we can optimize our data. We removed 30 plants from the edge of the field so that we did not lose much moisture in the plants before weighing and our allometric relationship would be as accurate as possible. By taking measurements of a few plants at key locations around the COSMOS probe, we also have a way to quantify the vegetation mass in different areas with this allometric relationship. With data on plant density within the probe’s footprint, we can scale up these individual plant measurements and find an areal amount of vegetation mass. This will give us a way to quantify the amount of vegetation present in the COSMOS probe’s footprint so that if \( N_0 \) is found to vary, we can test if \( N_0 \) is tied to the presence of vegetation.

We also found a way to determine the accuracy of our vegetation measurements. Our scaled up vegetation mass per area \( (M) \) values depend on two things: plant density and estimated plant mass using the allometric relationship. The estimated plant mass itself depends on two things: how good the projection of mass due to the allometric relationship is and natural plant variation. By taking samples of large amounts of vegetation in different areas of a field (105 plants at 3 locations), we were able to claim we had found “average” plant values in an area. An examination of the variance in these plants depending on location and time led us to claim that the natural variance of plants was constant. We also examined the variance due to our 30 plant allometric relationships and the variance in plant density. We found that the error in our vegetation mass values was small, normally less than 5%. We feel this is acceptable as there
are many different things that could have affected our plant measurements and, considering the values of $M$, this error is small.
CHAPTER 3. CALIBRATION OF COSMOS WITH CROPS

3.1 Introduction

Soil moisture is an important reservoir of the water cycle. Though small on the global scale, soil moisture and its fluxes play a role in crop production, weather, floods, and droughts. Knowing soil moisture levels, therefore, is important. Normal methods of sampling, usually on a point scale level and by hand, can be exhausting and time consuming. Depending on the resolution desired, sampling could take hours. Remote sensing has started to make it easier to measure soil moisture at large spatial scales. Remote sensing involves using an instrument that determines soil moisture by means other than direct sampling. Remote sensors can be active or passive. Active sensors send out signals and use these signals to gather the data needed for an observation. Passive sensors simply sense natural emissions of the medium. A developing field–scale remote sensor is COSMOS, COsmic–ray Soil Moisture Observing System. COSMOS is a network of probes distributed across the U.S.A.

COSMOS probes are stationary passive sensors that use the natural occurrence of moderated neutrons from cosmic–rays to determine volumetric soil moisture in a stationary footprint. Cosmic–rays are a primary form of galactic rays of solar origin. Neutrons are considered a secondary cosmic–ray flux particle (Zreda et al., 2012). There are three main classifications of neutrons: high energy neutrons, fast neutrons, and epithermal/thermal neutrons. Fast neutrons are defined to be neutrons on the order of 1-2 MeV (Hess et al., 1959) whereas moderated fast neutrons are on the order of 10 eV. COSMOS probes use these moderated fast neutrons (here referred to a moderated neutrons) to determine soil moisture. Neutrons have a lifespan of ∼15 minutes (Zreda et al., 2012; Anton et al., 1989). The equilibrium concentration of neutrons depends on the production rate and the moderating efficiency. The production of neutrons de-
pends on the mineral composition of the medium. The production rate increases as a function of $A^{2/3}$, where $A$ is the atomic mass of the element present (Geiger, 1956). Neutrons can travel a few tens of g cm$^{-2}$ in matter such as soil, but a few hundred meters in air before they are thermalized or slowed. The spatial variation of neutrons depends on both the strength of the geomagnetic field and variable atmospheric pressure. The temporal variation of neutrons is due mainly to solar activity and barometric pressure changes. Fast neutrons undergo elastic collisions with nuclei which causes them to lose energy and become moderated neutrons and are eventually absorbed in inelastic nuclear collisions which is called moderation. The slow time of a fast neutron is on the scale of $10^{-4}$ s, which is basically instantaneous (Glasstone and Edlund, 1952). The moderation of fast neutrons depends on three factors. One factor is the elemental scattering cross-section. The second factor is the logarithmic decrement of energy per collision. The third factor is the number of atoms of an element per unit mass of material.

For a COSMOS probe, these moderated neutrons are counted with a polyethylene–shielded detector. Within this gas–filled detector is enriched $^3$He for the IVS probe (or $^{10}$B for others). When a neutron goes through the $^3$He, it creates ionization which results in an electrical pulse (Zreda et al., 2012). These pulses are recorded on a cumulative hourly basis giving a counts per hour (cph) value. These counts are then used in (1.1) to determine the areal soil moisture. Neutrons are able to be counted in a variety of land conditions, some of these being: bare soil, soybean canopy, maize canopy, soybean residue, and maize residue. COSMOS probes are able to give soil moisture under all of these conditions. These conditions are all observed at the Iowa Validation Site (IVS) in Ames, IA.

A COSMOS probe has a footprint on the scale of 330 m radius at sea level (Zreda et al., 2008). For the IVS at an elevation of 316 m, it has been established to be a radius of $\sim$350 m. This footprint has been determined to not change significantly based on soil moisture levels (Zreda et al., 2008). However, as Zreda et al. (2008) also show, the measurement depth of the probes varies with soil moisture. Under a completely dry soil, the probe would theoretically sense down to almost 70 cm while with a saturated soil the probe would only sense about 12 cm down. Knowing how the measurement depth varies with soil moisture is something we will have to take into account.
COSMOS probes are essentially hydrogen counters, as hydrogen particles are what best scatter and absorb neutrons (Table 1.2) (Zreda et al., 2012). COSMOS probes need a few corrections made before their data can be used. One of these corrections deals with the level of humidity at the location. The more humid it is, the more water molecules are in the air and therefore the more hydrogen occurs in the atmosphere. A correction has been determined for this. Rosolem et al. (2013) shows how to correct for water vapor with current probes that have a flux tower nearby with data on barometric pressure levels, air temperature, and humidity. Subsequent probes will have temperature, pressure, and humidity sensors within them to account for water vapor and current probes will be fit with the same sensors. Another influence on the probes is that cosmic-rays do not come down to earth at the same intensity all the time. A correction is in place to account for the natural variability of cosmic-rays (Zreda et al., 2012). Our research group had the unique opportunity to look an additional correction factor. The IVS is an agricultural field that is under an every other year rotation of maize and soybean. With a COSMOS probe planted in the middle of the IVS, we had the chance to examine the effect this vegetation might have on a probe.

3.1.1 Question

Our research question was:

*Would rapid-growth vegetation affect the COSMOS probe’s soil moisture calculation?*

Our hypothesis was that:

*The COSMOS probe would be sensitive to the vegetation due to the total mass of vegetative water as well as dry matter hydrogen in the vegetation found in the footprint and the vegetation’s natural seasonal variation.*

3.1.1.1 Parts of the Question

We would be examining two different types of row crop vegetation in 2011 and 2012. We hypothesized that both types of vegetation would have an effect on the COSMOS probe, but
that their effects would differ. This would be due to the naturally different characteristics of
the two vegetation types, mainly their peak size but also perhaps their physical structure.

3.2 Methods

In order to look at what effect vegetation might have on a COSMOS probe, we needed
ways to determine how much vegetation is present and what the actual soil moisture is for
comparison. Methods behind how we sampled for soil moisture and measured vegetation are
detailed in Sections 2.2 and 2.3.2.

3.2.1 Depth of measurement

As a COSMOS probe measures down to different depths depending on the current level of
soil moisture (Zreda et al., 2008), a method was needed to take that factor into account. We
could not use our full 30 cm soil core, unless it was found that the probe would be measuring
down to 30 cm. We had soil sample data at 5 cm increments down to 30 cm for each of the
18 locations on each day we took soil samples. Using equation (4) from Franz et al. (2012b),
and assuming a constant lattice water, we modified an equation to find the effective measuring
depth. This equation was

\[
\phi(z^*) = W_s + \int_0^{z^*} (\rho_b(z) \tau + \theta(z)) dz
\]  

(3.1)

where \(z^*\) is the effective measuring depth (cm), \(W_s\) is the ponded surface water (cm), \(\rho_b\) is
the bulk density of the soil (g cm\(^{-3}\)), \(\tau\) is the gravimetric weight fraction of lattice water
(determined to be 0.045 g g\(^{-1}\) at the IVS), and \(\theta\) is the volumetric water content of the soil
(cm\(^3\) cm\(^{-3}\)). \(\phi(z^*)\) is also defined in Franz et al. (2012b) from particle modeling simulations to
be

\[
\phi(z^*) = 5.8 - 0.0829 z^* \quad 0 \leq z^* \leq 70
\]  

(3.2)

which is the 86% cumulative vertical sensitivity contour (cm), or the depth seen into soil given
a certain water content.

This equation, (3.1), assumes that the vegetative water column can be turned into a solid
pond of water. For vegetation, we used our plant mass per area values found in Section 2.3.4
(Figures 2.6 and 2.7). As $W_s$ is a water equivalent (cm of H$_2$O), we needed to convert our measurements to that. For this, we used

$$W_s = \frac{18}{2} \frac{6}{100} M_d + M_w. \quad (3.3)$$

Since 1 kg water m$^{-2}$ is equal to 1 mm of equivalent depth, converting $M_w$ is straightforward. To convert $M_d$, we accounted for the fact that $\sim$6% of above ground plant dry matter mass is hydrogen (Epstein and Bloom, 2005), and that water mass is a 2 g hydrogen of 18 g water conversion. This gives us the total fresh mass in water equivalents (cm).

As the vegetation exists in rows and is not necessarily uniform, we added a factor that was able to take the distribution and structure of vegetation into account. Therefore, our final equation to determine the probe’s depth of measurement is

$$5.8 - 0.0829 z^* = \kappa W_s + \int_0^{z^*} (\rho_b(z) \tau + \theta(z)) dz \quad (3.4)$$

where $\kappa$ is a factor from 0 to 1 that accounts for the distribution and structure of vegetation. A $\kappa$ value of 1 would be completely uniform vegetation.

To find $z^*$ for our data, we used each day’s total vegetation value (converted to equivalent water depth in cm) for a total ponded surface water value. This gave us a daily $W_s$ (Table 3.1) to use with the constant $\tau$ and a $\kappa$ value of 1. The remaining variables depend on $z^*$, the depth seen by the probe. We averaged each day’s 18 locations’ $\rho_b$ and $\theta_v$ values at 5 cm increments. Using an iterative approach, we started with the sum of all 6 averaged increments’ values for $\rho_b$ and $\theta_v$ and with 30 cm for $z^*$, then the sum of the 0-25 cm values with $z^* = 25$ cm, and so on until we found a depth for each day that best satisfied (3.4). No 5 cm depth fully satisfied the equation; however, as Franz et al. (2012b) notes, guesses within a few centimeters of the true $z^*$ will not cause a significant error. Therefore, we felt confident that our method was sufficient. The $z^*$ values found for each sampling day are shown in Table 3.1.

The soil moisture values were weighted vertically using the method described by Franz et al. (2012b). These weighted soil moistures are the values shown in Table 3.1. Only the samples that fell above the $z^*$ found for that day were used for the weighting method and they were averaged across the 18 locations for each depth. The errors in these soil moistures values were
Table 3.1  Dates in–situ soil samples were taken on VC days with their respective vegetation water equivalents ($W_s$) (cm); calculated $z^*$ (using (3.4) and assuming $\kappa = 1$) (cm); the average, vertically weighted volumetric soil moisture ($\theta_v$) ($\text{cm}^3 \text{ cm}^{-3}$) down to that depth using our in–situ samples (standard error of measurements in parentheses); and the new calibrated $N_0$ (cph) for that date (error in parentheses).

<table>
<thead>
<tr>
<th>Date</th>
<th>$W_s$</th>
<th>$z^*$</th>
<th>mean $\theta_v$</th>
<th>$N_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011 - Maize</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 19</td>
<td>0 cm</td>
<td>15 cm</td>
<td>0.253 (0.011) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3345 (46) cph</td>
</tr>
<tr>
<td>June 24</td>
<td>0.087 cm</td>
<td>15 cm</td>
<td>0.255 (0.013) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3232 (49) cph</td>
</tr>
<tr>
<td>July 8</td>
<td>0.25 cm</td>
<td>20 cm</td>
<td>0.153 (0.011) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>2982 (62) cph</td>
</tr>
<tr>
<td>August 18</td>
<td>0.55 cm</td>
<td>15 cm</td>
<td>0.158 (0.012) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>2973 (61) cph</td>
</tr>
<tr>
<td>September 13</td>
<td>0.44 cm</td>
<td>15 cm</td>
<td>0.165 (0.011) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3023 (57) cph</td>
</tr>
<tr>
<td>November 14</td>
<td>0 cm</td>
<td>15 cm</td>
<td>0.231 (0.008) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3169 (38) cph</td>
</tr>
<tr>
<td>2012 - Soybean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>March 14</td>
<td>0 cm</td>
<td>15 cm</td>
<td>0.218 (0.013) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3080 (52) cph</td>
</tr>
<tr>
<td>May 9</td>
<td>0 cm</td>
<td>15 cm</td>
<td>0.250 (0.013) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3283 (51) cph</td>
</tr>
<tr>
<td>June 5</td>
<td>0 cm</td>
<td>15 cm</td>
<td>0.249 (0.013) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3313 (51) cph</td>
</tr>
<tr>
<td>June 19</td>
<td>0.022 cm</td>
<td>15 cm</td>
<td>0.241 (0.012) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3264 (49) cph</td>
</tr>
<tr>
<td>July 6</td>
<td>0.072 cm</td>
<td>15 cm</td>
<td>0.159 (0.008) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3305 (51) cph</td>
</tr>
<tr>
<td>July 20</td>
<td>0.11 cm</td>
<td>20 cm</td>
<td>0.127 (0.008) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3092 (53) cph</td>
</tr>
<tr>
<td>August 1</td>
<td>0.13 cm</td>
<td>20 cm</td>
<td>0.143 (0.008) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3055 (47) cph</td>
</tr>
<tr>
<td>August 15</td>
<td>0.20 cm</td>
<td>20 cm</td>
<td>0.161 (0.009) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3061 (50) cph</td>
</tr>
<tr>
<td>September 25</td>
<td>0.046 cm</td>
<td>15 cm</td>
<td>0.181 (0.010) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3218 (55) cph</td>
</tr>
<tr>
<td>October 11</td>
<td>0 cm</td>
<td>20 cm</td>
<td>0.171 (0.011) $\text{cm}^3 \text{ cm}^{-3}$</td>
<td>3277 (58) cph</td>
</tr>
</tbody>
</table>

found by determining the standard error of the samples. The standard error is the standard deviation divided by the square root of the number of samples. Figure 3.1 and Figure 3.2 show the soil profiles for every sampling day in 2011 and 2012. It can be seen that soil moisture changed considerably by depth and location. Each grey line represents one of the 18 sampling locations while the black line is the mean of the 18 locations at each depth and the error bars are the standard errors.

3.2.2 $N_0$ determination

In order to test if $N_0$ was indeed a constant and therefore independent of vegetation present, we recalibrated the IVS probe’s $N_0$ value with the soil moisture data collected on all our
Figure 3.1  Soil moisture profiles for each of the sampling days in 2011 as well as two of the 2012 days. Each grey line represents one location, the black line is the average of the 18 locations with standard errors shown as the error bars.
Figure 3.2  Soil moisture profiles for eight of the sampling days in 2012. Each grey line represents one location, the black line is the average of the 18 locations with standard errors shown as the error bars.
sampling days in 2011 and 2012. We rearranged (1.1) to solve for \( N_0 \), using our \( z^* \) soil moisture as \( \theta \). For \( N \) we found the average moderated neutron count for each respective sampling day for a 24 hour period, running midnight to midnight local time. This time period was used because it ensured data on moderated neutron counts for every day of sampling. As the COSMOS probe calculates the cumulative counts of moderated neutrons over a 60 minute period, any data collected and recorded at any more or less than that interval, such as 59 or 61 minutes, is removed immediately and a gap is left in that collection period. This became problematic when we looked at using the average counts over the \( \sim 4 \) hour period it took to sample the soil moisture each day. Although that 4 hour period would have moderated neutrons counts that more accurately matched with what was captured with our in–situ soil moisture measurements, some days had no available moderated neutron counts data in that period. For this reason, we used the midnight to midnight \( N \). By averaging the moderated neutron counts over a 24 hour period, we removed the diurnal pattern seen in the counts. Our results of the calibrated \( N_0 \) values are shown in Table 3.1.

The error in our \( N_0 \) values was found by the same method as in Section 2.4.1. From Beers (1962) we used the general rule for combining independent errors

\[
s_V = \sqrt{\left( \frac{\partial V}{\partial x} \right)^2 s_x^2 + \left( \frac{\partial V}{\partial y} \right)^2 s_y^2}
\]

(3.5)

where \( V \) is some function of two variables \( x \) and \( y \), \( s_V \) is the standard deviation of \( V \), \( \frac{\partial V}{\partial x} \) is the partial derivation of \( V \) with respect to \( x \), \( s_x \) is the standard deviation of \( x \), \( \frac{\partial V}{\partial y} \) is the partial derivation of \( V \) with respect to \( y \), and \( s_y \) is the standard deviation of \( y \). In this case, we used the rearranged (1.1) that solved for \( N_0 \) as our function \( V \) where \( x \) and \( y \) are \( N \) and \( \theta \). For the standard deviation of \( \theta \) we used the standard error of our soil moisture measurements that are shown in Table 3.1. The standard error was defined to be the standard deviation of the samples divided by the square root of the number of samples (18 per 5 cm depth). For the standard error of \( N \), we found the standard deviation of the daily moderated neutrons counts and divided it by the square root of the number of measurements made for that day (24 for a full day of counts).
3.3 Results

3.3.1 κ values

While we used a κ = 1 value in our determinations of \( z^* \), we did a quick study of how \( z^* \) would change depending on the vegetation distribution. We used three extreme κ values: 1 for completely uniform vegetation, 0.5 as a middle value, and 0.1 for very non-uniform vegetation.

The \( z^* \) values for each sampling day in 2011 and 2012 are shown in Table 3.2 for the three κ values. We found that with these three κ values, only one \( z^* \) changed, that of August 18, 2011. This was the date that corresponded to peak maize vegetation. It can be reasoned that this should be the first or only date to show a change in \( z^* \) depending on κ because of the total mass of vegetation present on that date (peak maize vegetation). Thus, we concluded that vegetation distribution may not be a large factor in determining probe sensing depth. For more information on the spatial distribution of hydrogen, see Franz et al. (2013b).

Table 3.2 Effective depths (\( z^* \)) (cm) found with (3.4), using extreme κ values.

<table>
<thead>
<tr>
<th>Date</th>
<th>κ = 0.1</th>
<th>κ = 0.5</th>
<th>κ = 1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011 - Maize</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 19</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>June 24</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>July 8</td>
<td>20 cm</td>
<td>20 cm</td>
<td>20 cm</td>
</tr>
<tr>
<td>August 18</td>
<td>20 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>September 13</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>November 14</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>2012 - Soybean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>March 14</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>May 9</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>June 5</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>June 19</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>July 6</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>July 20</td>
<td>20 cm</td>
<td>20 cm</td>
<td>20 cm</td>
</tr>
<tr>
<td>August 1</td>
<td>20 cm</td>
<td>20 cm</td>
<td>20 cm</td>
</tr>
<tr>
<td>August 15</td>
<td>20 cm</td>
<td>20 cm</td>
<td>20 cm</td>
</tr>
<tr>
<td>September 25</td>
<td>15 cm</td>
<td>15 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>October 11</td>
<td>20 cm</td>
<td>20 cm</td>
<td>20 cm</td>
</tr>
</tbody>
</table>
3.3.2 $N_0$ changes

As has been shown in Table 3.1, $N_0$ was found to be non–constant. To see how $N_0$ varied, we plotted each year’s collective fresh matter and water column densities against the recalibrated $N_0$ values. The results of this can be seen in Figures 3.3 and 3.4 for 2011 and 2012, respectively. An interesting fact to note: three collection dates at the beginning of the season in 2012 were all considered to be bare soil conditions, compared to only one capture of a beginning season bare soil for maize in 2011. These were: March 14, May 9, and June 5, 2012. On the March 14 date, there was quite a bit of remaining maize residue from the 2011 harvest. On May 9 there was less residue present due to the natural decomposition of the residue. The residue was determined to be too difficult to measure, hence why these two days were considered bare soil days as well as the November 14, 2011 date with the “fresh” maize residue. While the IVS was planted on May 16-17,2012 with plant emergence on May 22, on June 5 these plants were too small to be accurately measured and modeled with our method so we assumed bare soil conditions. A view of the plants on June 5 is included in Figure 3.5.

While we labeled several days as bare soil conditions (May 19 and November 14, 2011; March 14, May 9, June 5, October 11, 2012), we must acknowledge that these days were not
Figure 3.4 A plot of fresh matter and water column densities for soybean plants in 2012 against recalibrated $N_0$ values. Three bare soil condition samples were taken at the beginning of the year, which proved to have variable $N_0$ values.

void of vegetation. Rather, we did not measure above–ground standing vegetation on these days. It should be noted that despite the lack of above–ground upright vegetation, there was still above–ground residue and below–ground root masses that would have been additional sources of hydrogen. For modeling purposes, we assumed one zero–vegetation day at the beginning of each season and one at the end of each season. For 2011, we only had one day for each. For 2012, we only had one end of season day but three beginning of season days. Considering the presence of the root and residue biomass, June 5 was chosen to be the zero–vegetation day as it was assumed that the leftover roots and residue from the 2011 maize had decomposed the most by this period and the small growth of new soybean plants was deemed insignificant vegetation mass at this time.

3.3.3 Calibration

Now that it has been shown that $N_0$ is not a constant but rather a vegetation–dependent variable, a correction factor for this must be taken into account. We were able to use our data and previous mentioned results to create such a correction factor.
Figure 3.5  Soybean plants on June 5, 2012. These plants were too small to be measured accurately so this day was categorized as a bare soil day. Remaining maize residue can also be seen here. This residue has been worked into the top soil now due to planting practices and has decomposed over time, but it can still be seen how challenging it would have been to measure the total mass of residue.

3.3.3.1 Maize, 2011

By fitting a one–day time step curve to Figure 2.6, the amount of fresh vegetation present throughout the growing period can be roughly modeled with a one term Gaussian fit

\[ a_1 \exp \left( - \left( \frac{x - b_1}{c_1} \right)^2 \right) \]  

(Figure 3.6) on a daily interval. It was decided to use the fresh mass of the three masses because the fresh mass consists of both the dry matter and water masses and is the purest form of the vegetation. This fresh vegetation mass model can be used to determine the subsequent \( N_0 \) values, also with a model estimation. For this, a curve with steps of 0.1 kg m\(^{-2}\) was fit up to the peak vegetation modeled in our relationship of fresh vegetation mass and \( N_0 \) (Figure 3.3),
and a second linear relationship was found from the peak vegetation back to bare soil conditions by route of the senescent vegetation. These fits were, respectively:

\[ N_0 = 20.1 M_f^2 - 189.9 M_f + 3360.8 \]  \hspace{1cm} (3.7)

\[ N_0 = -31.3 M_f + 3169.2 \]  \hspace{1cm} (3.8)

The fresh mass relationship was again used here because it was the most basic, and easy to relate to parameter. All the other factors used to compared \( N_0 \) to (hydrogen equivalents, water equivalents, dry matter, water mass) are found through the fresh vegetation. By combining the relationships of fresh mass over time and \( N_0 \) over fresh mass, a rough model can be made for \( N_0 \) over time, as shown in Figure 3.7.

\( N_0 \) now is a vegetation–dependent variable that can be used with Level 2 data (http://cosmos.hwr.arizona.edu/Probes/StationDat/016/corcounts.txt) on corrected moderated neutron counts to find volumetric soil moisture. By averaging the corrected moderated neutron counts to a daily value to match the fits, the true, vegetation–removed daily soil moisture within the COSMOS probe’s footprint can be calculated. A look at this new adjusted soil moisture and how it compares to the original soil moisture levels found with no vegetation correction (constant \( N_0 \)) is shown in Figure 3.8. As a check of how accurate our fits were, we found
Figure 3.7  A rough estimation of how $N_0$ would vary over the 2011 crop season. This was found by combining fits to Figures 2.6 and 3.3.

Figure 3.8  Soil moisture calculations by COSMOS during the 2011 crop season. This figure displays a comparison between the ‘original’, uncorrected soil moisture values found when using the initial calibration of a constant $N_0$ (black line) and the modeled soil moisture after using a vegetation–dependent variable $N_0$ (red line). In–situ soil moisture values are shown as blue dots on the days they were taken. The bottom figure is the difference between the uncorrected soil moisture and modeled soil moisture. The gap in the series in early July has to do with a short period of missing data.
the fitted $N_0$ values over time and COSMOS adjusted soil moisture for each sampling date to compare to our original recalibrated $N_0$ and in–situ soil moisture values. Table 3.3 shows these comparisons and it can be noted that the largest variation in $N_0$ was 95 cph on June 24 when comparing calculated $N_0$ values to modeled values. Also, with an uncorrected probe (constant $N_0$), soil moisture values differed from 0.015 to 0.108 cm$^3$ cm$^{-3}$ whereas the modeled soil moisture values’ differences only ranged from 0.001 to 0.028 cm$^3$ cm$^{-3}$. The modeled soil moisture has a better range of soil moisture differences than the uncorrected soil moisture. A maximum difference of 0.028 cm$^3$ cm$^{-3}$ is good considering all the possible errors that could have occurred in the models we made. 0.028 cm$^3$ cm$^{-3}$ is also a much better error than 0.108 cm$^3$ cm$^{-3}$ when soil moisture contents usually only vary $\sim 0.300$ cm$^3$ cm$^{-3}$.

3.3.3.2 Soybean, 2012

Figure 3.9 2012 soybean fresh matter vegetation with a one–term Gaussian fit.

The same method was implemented for the 2012 data to find $N_0$ over the soybean crop season. A one term Gaussian curve was fit to the fresh matter vegetation (Figure 2.7) that was stepped by daily increments from June 5 to October 11 and is shown in Figure 3.9. After creating a fit to the vegetation mass over time, we created a fit for the relationship of $N_0$ to vegetation (Figure 3.4). As with maize, we used a second–order polynomial to model the
Table 3.3 The $N_0$ values calculated after 2011 VC maize field days (cph); $N_0$ values using fits (Figure 3.7) from models (cph); in–situ soil moistures we found for field days in 2011 (cm$^3$ cm$^{-3}$); COSMOS soil moistures found with an uncorrected, constant $N_0$ of 2898 cph (cm$^3$ cm$^{-3}$); as well as the COSMOS soil moistures found with our modeled $N_0$ values (Figure 3.8) (cm$^3$ cm$^{-3}$). Differences between the calculated $N_0$ values and modeled $N_0$ values are shown in parentheses, as are the differences between in-situ soil moisture and uncorrected and modeled COSMOS soil moistures (+ means it was higher than the measured value, – means it was lower than the measured value).

<table>
<thead>
<tr>
<th>Date</th>
<th>calculated $N_0$</th>
<th>modeled $N_0$</th>
<th>in–situ $\theta$</th>
<th>uncorrected COSMOS $\theta$</th>
<th>modeled COSMOS $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 19</td>
<td>3345</td>
<td>3337 (−8)</td>
<td>0.253</td>
<td>0.145 (−0.108)</td>
<td>0.251 (−0.002)</td>
</tr>
<tr>
<td>June 24</td>
<td>3232</td>
<td>3137 (−95)</td>
<td>0.255</td>
<td>0.167 (−0.088)</td>
<td>0.227 (−0.028)</td>
</tr>
<tr>
<td>July 8</td>
<td>2982</td>
<td>3003 (+21)</td>
<td>0.153</td>
<td>0.136 (−0.017)</td>
<td>0.157 (+0.004)</td>
</tr>
<tr>
<td>August 18</td>
<td>2973</td>
<td>2966 (−7)</td>
<td>0.158</td>
<td>0.143 (−0.015)</td>
<td>0.157 (−0.001)</td>
</tr>
<tr>
<td>September 13</td>
<td>3023</td>
<td>3003 (−20)</td>
<td>0.165</td>
<td>0.140 (−0.025)</td>
<td>0.161 (−0.004)</td>
</tr>
<tr>
<td>November 14</td>
<td>3169</td>
<td>3159 (−10)</td>
<td>0.231</td>
<td>0.164 (−0.067)</td>
<td>0.228 (−0.003)</td>
</tr>
</tbody>
</table>
growing vegetation and a first–order polynomial for the senescent vegetation. These equations were:

\[ N_0 = 77.5 M_f^2 - 289.3 M_f + 3320.1 \]  \hspace{1cm} (3.9)

\[ N_0 = -94.6 M_f + 3282.7 \]  \hspace{1cm} (3.10)

By combining our two relationships, we were able convert \( N_0 \) to a vegetation–dependent variable again, shown in Figure 3.10.

With this vegetation–dependent \( N_0 \) parameter for the crop season, we were able to adjust the soil moisture calculated by COSMOS over the 2012 season as well. Using our new \( N_0 \) values and Level 2 data (http://cosmos.lwr.arizona.edu/Probes/StationDat/016/corcounts.txt) on a daily scale, we looked at the new adjusted soil moistures against the soil moisture originally reported from COSMOS with the original \( N_0 \) parameter found in September 2010. Figure 3.11 shows our adjusted–\( N_0 \) soil moisture levels with those found with a constant \( N_0 \). Unlike maize where a large deviation can be seen in the two soil moistures, these soil moistures have a similar pattern throughout the season. This is not surprising as the original calibration was done with soybean plants covering the field. Therefore, the original \( N_0 \) is actually calibrated for the presence of some soybean vegetation.

As we did for maize, we looked at our fits to the data and found the adjusted \( N_0 \) values and soil moisture values for the days we sampled. As we only used the June 5 bare soil location for the fits, we do not have data on the March and May dates. We can compare these new, adjusted values to the original ones that were found after days of data collection in the field (Table 3.4). It can be seen that the largest difference between calculated and modeled \( N_0 \) values (137 cph) was on July 6. Other items to note from Table 3.4 are that the uncorrected soil moistures differed from 0.027-0.101 cm\(^3\) cm\(^{-3}\), which is a similar range as the 2011 maize uncorrected differences (0.015-0.108 cm\(^3\) cm\(^{-3}\)). Modeled soil moisture differences from in–situ soil moistures only ranged from 0.001-0.024 cm\(^3\) cm\(^{-3}\) which is also similar to the range found in 2011 (0.001-0.028 cm\(^3\) cm\(^{-3}\)).

A comparison of maize versus soybean and the effect each had on \( N_0 \) shows that both vegetation types showed a hysteresis–like relationship of \( M \) versus \( N_0 \) (Figures 3.3 and 3.4).
Figure 3.10  A rough estimation of how \( N_0 \) would vary over the 2012 crop season. This was found by combining fits to Figures 2.7 and 3.4.

Figure 3.11  Soil moisture calculations by COSMOS during the 2012 crop season. This figure displays a comparison between the ‘original’, uncorrected soil moisture values found when using the initial calibration of a constant \( N_0 \) (black line) and the modeled soil moisture after using a vegetation-dependent variable \( N_0 \) (red line). In–situ soil moistures are again shown as blue dots on their respective days. The bottom figure shows the difference between the uncorrected soil moistures and the modeled soil moistures.
Table 3.4  The $N_0$ values calculated after 2012 VC soybean field days (cph); $N_0$ values using fits (Figure 3.10) from models (cph); in–situ soil moistures we found for field days in 2012 (cm$^3$ cm$^{-3}$); uncorrected COSMOS soil moistures using a constant $N_0$ of 2898 cph (cm$^3$ cm$^{-3}$); and the COSMOS soil moistures found with our modeled $N_0$ values (Figure 3.11) (cm$^3$ cm$^{-3}$). Differences between the calculated $N_0$ values and modeled $N_0$ values are shown in parentheses, as are the differences between in–situ soil moisture and uncorrected and modeled COSMOS soil moistures (+ means it was higher than the measured value, − means it was lower than the measured value).

<table>
<thead>
<tr>
<th>Date</th>
<th>calculated $N_0$</th>
<th>modeled $N_0$</th>
<th>in–situ $\theta$</th>
<th>uncorrected COSMOS $\theta$</th>
<th>modeled COSMOS $\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 5</td>
<td>3313</td>
<td>3309 (−4)</td>
<td>0.249</td>
<td>0.148 (−0.101)</td>
<td>0.248 (−0.001)</td>
</tr>
<tr>
<td>June 19</td>
<td>3264</td>
<td>3274 (+10)</td>
<td>0.241</td>
<td>0.152 (−0.089)</td>
<td>0.244 (+0.003)</td>
</tr>
<tr>
<td>July 6</td>
<td>3305</td>
<td>3168 (−137)</td>
<td>0.159</td>
<td>0.093 (−0.066)</td>
<td>0.135 (−0.024)</td>
</tr>
<tr>
<td>July 20</td>
<td>3092</td>
<td>3072 (−20)</td>
<td>0.127</td>
<td>0.097 (−0.030)</td>
<td>0.124 (−0.003)</td>
</tr>
<tr>
<td>August 1</td>
<td>3055</td>
<td>3051 (−4)</td>
<td>0.143</td>
<td>0.116 (−0.027)</td>
<td>0.142 (−0.001)</td>
</tr>
<tr>
<td>August 15</td>
<td>3061</td>
<td>3065 (−4)</td>
<td>0.161</td>
<td>0.130 (−0.031)</td>
<td>0.162 (0.001)</td>
</tr>
<tr>
<td>September 25</td>
<td>3218</td>
<td>3227 (−9)</td>
<td>0.181</td>
<td>0.120 (−0.061)</td>
<td>0.183 (0.002)</td>
</tr>
<tr>
<td>October 11</td>
<td>3277</td>
<td>3267 (−10)</td>
<td>0.171</td>
<td>0.105 (−0.066)</td>
<td>0.169 (−0.002)</td>
</tr>
</tbody>
</table>
64

Figure 3.12 A comparison of the two seasons of vegetation-dependent $N_0$ values. The 2011 $N_0$ values are displayed in blue while the 2012 $N_0$ values are displayed in purple.

This confirms our hypothesis that they would both have influences on the probe. $N_0$ values ranged from $\sim 2900$ cph to 3350 cph for maize and 3050 cph to 3300 cph for soybean. A comparison of the corrected $N_0$ values for both seasons is shown in Figure 3.12. It can be seen that both seasons’ $N_0$ values had a similar pattern to them, where $N_0$ values decrease somewhat sharply until close to the end of July and then increase somewhat sharply starting at the end of August with a small period of waviness between. The 2012 (soybean) values are consistently higher than the 2011 (maize) values with no overlap of the two seasons’ $N_0$ values. This figure shows that maize and soybean both had an effect on the $N_0$ values over their seasons and that this effect was of a similar pattern. However, the two crops effects were of different magnitudes. With soybeans present, the probe does not need to be adjusted as much as with maize present. This proves our hypothesis that the two crops would have different effects on the probe.

3.4 Conclusions

We have found a way to determine the depth seen into the soil ($z^*$) by a COSMOS probe under the presence of vegetation. By using data collected with in-situ soil samples and vegetation measurements around the probe, we can find $z^*$ in an iterative way. Vegetation mass
\( M_f \) has to be converted to water equivalents \( W_s \) which will depend on the amount of water in the vegetation and the amount of hydrogen in the dry matter. Converting \( M_f \) to \( W_s \) allows us to use the vegetation as ponded water. Then, with information on the bulk density and volumetric soil moisture at different depths, \( z^* \) can be determined.

With a \( z^* \) value, we can crop our 30 cm soil moisture samples to be what is actually seen by the COSMOS probe and weight it vertically. Then we are able to recalibrate the probe with these in-situ soil moistures. \( N_0 \) was found to vary and when plotted against vegetation, a relationship was determined. We used fits to the vegetation over time and \( N_0 \) dependent on vegetation to determine how \( N_0 \) varied over time with maize and soybean present. Our model of \( N_0 \) over time fit fairly well with the original recalibrated \( N_0 \) values found. The highest difference was 137 cph or \( \sim 4\% \). The recalibrated soil moisture values were also close to the in-situ values we found with the largest difference being 0.028 cm\(^3\) cm\(^{-3}\).

By creating \( N_0 \) as a function of time, the COSMOS probe can be corrected for the presence of vegetation. After adjusting the probe for the presence of vegetation, a vegetation effect can be seen on the seasonal soil moisture in 2011 with maize (Figure 3.8). This familiar variation of soil moisture is inversely related to the presence of vegetation. With no vegetation present, soil moisture is high. With the presence of vegetation though, soil moisture decreases, likely as the vegetation absorbs the water for its own purposes. Upon harvesting the vegetation, soil moisture values increase again. This pattern was not as noticeable in the probe’s soil moisture 2012 soybean time series (Figure 3.11). This may be due to the drought experienced that year or the fact that soybean plants are not as large as maize plants, thus they have a smaller effect on seasonal soil moisture patterns. Without the vegetation adjustment to \( N_0 \), the COSMOS probe underestimates soil moisture throughout the season. Maize and soybean crops both had effects on \( N_0 \). These effects were of a similar pattern but different magnitude, resulting in \( N_0 \) seasonal variations from 2900 cph-3350 cph for maize but only 3050 cph-3300 cph for soybean.
CHAPTER 4. VALIDATION OF FEATURE SPACE INTERPOLATION METHOD

4.1 Introduction

Point scale samples of soil moisture are a common method of determining soil moisture. Point scale measurements can be defined differently depending on the instrument used. Point scale measurements can be time and labor intensive and must be made at multiple locations in a certain area to quantify the variation of the soil moisture. Depending on the area being examined, a large number of soil samples may need to be taken in order to adequately cover the whole area. Upscaling of these point scale measurements can be useful as it would give the soil moisture over a larger area with fewer samples. Upscaling soil moisture measurements is the process of determining an areal soil moisture using an algorithm that requires a few measurements of soil moisture in order to infer the areal soil moisture. The Feature Space Interpolation (FSI) method is a newly developing method of upscaling that uses topographic and soils data on elevation, slope, aspect, curvature, and horizontal and vertical electromagnetic inductance (EMI) to upscale a few point measurements into an areal soil moisture. The following characteristics were chosen because they all have been found to influence soil moisture across areas (Famiglietti et al., 1998).

**Elevation** is a common and understandable concept. Elevation is the respective difference in height from a set point to other locations. Usually, elevation is given in reference to sea level. It is usually found that soil moisture is inversely related to elevation. Water tends to move to lower elevation areas due to percolation, infiltration, runoff, and erosion factors (Hawley et al., 1983; Henninger et al., 1976; Robinson and Dean, 1993). Gravimetric water potential also influences soil moisture based on elevation.
Slope, or slope angle, pertains to how steep a location is or how sharply an area changes from one location to another. As elevation can change over areas, slopes must occur between those elevations. The larger the difference in elevation is between two points and the smaller the distance is between them, the steeper the slope. The smaller the difference in elevation and the farther away from each other they are, the shallower the slope. Slope can affect the movement of water. Steeper slopes have water move across them faster than shallow slopes. Water may be able to infiltrate better on the shallower slopes and have less runoff than steep slopes. Slope has been shown to be a factor in determining soil moisture (Moore et al., 1988; Hills and Reynolds, 1969).

Aspect, or slope orientation, has to do with the position of a location in reference to the Sun. Solar radiation on a location is tied to its aspect and, therefore, as is the evapotranspiration that occurs. In the Northern Hemisphere, the Sun is positioned to the south year round for locations north of the Tropic of Cancer. Thus, south–facing locations receive more sunlight than north–facing locations. East–facing locations receive more sunlight and radiation in the morning and west–facing receive more after solar noon. Weeks and Wilson (2006) showed that north–facing areas in the Northern Hemisphere received significantly less radiation than south–facing or horizontal hills. The amount of sunlight, or radiation, can affect how an area dries. Reid (1973) found a correlation between aspect and soil moisture and Western et al. (1999) found radiation to be the best predictor of soil moisture in dry periods.

Curvature is the measure of the concavity or convexity of a landscape (Famiglietti et al., 1998) which would affect the flow and possible ponding of water. A concave area would likely retain more water than a convex area as it is shaped more like a bowl to collect water. A convex area would lose more water because it would be positioned for more runoff and have a smaller upslope contributing area. The curvature of an area has been found to influence soil moisture (Moore et al., 1988). Tomer et al. (2006) showed that surface curvature was the attribute most commonly correlated with soil moisture.

Soil texture plays a key role in determining soil moisture (Brady and Weil, 1999). Clay
particles are smaller in size than sand particles and therefore have a higher surface area. Water molecules are better able to attach to clay particles than sand particles because of this. A well–aerated, well–granulated soil would have more pore space and therefore more water holding capacity. The particle sizes of soil affect the ability of water to infiltrate, percolate, and evaporate or transpire. Electromagnetic inductance has been found to be tied to soil texture or particle size and therefore water holding capacities. Sand particles have a low conductivity, around the range of 1 mS m\(^{-1}\) (milliSiemens per meter). Clay particles have a larger conductivity between 10 and 1000 mS m\(^{-1}\) and silt fall between near the range of 10 mS m\(^{-1}\) (Grisso et al., 2009). Thus, EMI data is an easy way to determine soil texture.

While there are other possible factors that can directly affect soil moisture, these five were chosen because Van Arkel (2012) had the goal of employing the most dominate physical parameters that have an effect on the spatial variability of soil moisture. Soil moisture is complex and therefore can not be captured well by one predictor. Soil moisture is affected by so many things that to predict it multiple variables are needed.

FSI uses the K–means clustering algorithm to divide the desired area into \(n\) cluster groups. The K–means clustering algorithm finds similarly behaving clusters within input data. It then identifies critical samplings points, or centroids. MATLAB has a K–means clustering function. The algorithm finds an initial ‘mean’ of \(n\) clusters from randomly selected input data. Clusters are then created by associating each input vector to the closest mean. The algorithm then finds the geometric center of each arbitrary cluster and makes that the new mean until convergence is met. (If desired, additional information on K–means clustering can be found in MacQueen (1967).) Upon collection of the topographic and soil data of elevation, slope, curvature, aspect, and EMI, FSI divides each location data was obtained at into \(n\) groups of clusters depending, therefore, not on the physical distance between the points per se, but on the similar characteristics of the data. FSI also identifies the centroid of each cluster as well as how all the locations are assigned to the clusters.
4.1.1 Question

We had a unique opportunity to test FSI at the IVS. With the COSMOS probe now able to be corrected and adjusted for vegetation, we can have a field scale knowledge of soil moisture to which upscaled soil moisture values can be compared. Using the vegetation calibrated COSMOS probe as the ground truth, our research question for this was:

*How well does the FSI method’s upcaled soil moisture value match the COSMOS soil moisture for the same day?*

4.1.1.1 Parts of the Question

While examining the FSI method’s ability to upscale soil moisture, we will be looking at 4 different groups of clusters: 1, 2, 3, 4. We did this in order to test which group of clusters can better upscale the soil moisture. We hypothesized that the 3 cluster group would be the best at upscaling.

4.2 Methods

In order to test this research question, knowledge of multiple parameters is needed. We will need to know: the in–situ soil moisture; where the cluster centroids are; how the field is divided into clusters; EMI, elevation, curvature, slope, and aspect; and the true COSMOS vegetation–corrected soil moisture.

4.2.1 Field data

Topographic data were collected at a ∼20 m resolution with a GPS receiver on an all–terrain vehicle (Van Arkel, 2012). The IVS was divided into 10 m sections and the slope, planar curvature, and aspect for each point was derived from the elevation data using Surfer® (Golden Software, Inc., Golden, Colorado). Electromagnetic inductance (EMI) data were gathered at a ∼20 m resolution as well using an EMI sled pulled behind an all–terrain vehicle. Horizontal and perpendicular conductances (mS m⁻¹) were found and interpolated to 10 m points with the same method as the topographic data. Elevation, planar curvature, aspect, slope, horizontal
EMI, and perpendicular EMI data were now available at 10 m grid points for the IVS. For comparison with the COSMOS footprint (∼350 m), these data were cropped to an area similar to the COSMOS footprint. Only locations within the 350 m radius of the probe were used to ensure we were using data for the same area in both methods.

Figure 4.1 All 10 m locations with soil and topographic data within the COSMOS footprint (350 m radius), COSMOS probe (center dot), centroids of the 4 cluster groups.

A best matching unit (BMU) finder was used to determine the optimal sampling locations, or the K–means centroids, within the 350 m radius of the COSMOS probe. Four different clusters were examined: 1, 2, 3, 4 for a total of 10 possible sampling locations. In reality though, there were a few duplicate locations identified. The 2 cluster group had one match with a centroid identified in the 3 cluster group while the 4 cluster group had two centroids that were the same as the other two centroids in the 3 cluster group. This meant that we only sampled at seven centroids, shown in Figure 4.1, and these three sets of duplicate centroids used the same data for the point identified in two different cluster groups.

Three soil samples were taken at each of the seven centroids every day soil samples were collected for this project on UV days (Upscaling Validation). The three samples were taken
within ~0.5 m radius of one another and within ~2 m of the exact coordinate. This was to ensure a good capture of the current soil moisture as one sample could be abnormally high or low moisture, two might be completely different, but three allows for a respectable average. More than three would make a better average, but that would have added a significant amount of time and resources to collecting the samples. With only three samples at each centroid, all seven locations could be sampled within a three hour time frame. This small time frame meant that there was less of a chance of the soil naturally drying out between the first sampling location and the last.

Each 10 m location at which data were available within the COSMOS footprint was separated into a cluster based on K–means clustering determined by elevation, slope, aspect, curvature, and electromagnetic inductance. Figure 4.2 shows the break down for cluster assignments. Small differences can be seen with how the 10 m locations were divided into the clusters for these duplicate centroids. For the most part, the corresponding locations in the clusters for the duplicate centroids are the same.

4.2.2 Soil moisture data

We assumed that, as the centroid of each cluster was identified to be the best representation of that cluster, its in–situ soil moisture would be the best representation of soil moisture for that cluster. Therefore, every 10 m location for a centroid’s cluster was assigned that centroid’s soil moisture. Each centroid location’s three in–situ samples were averaged together at 5 cm increments, from 0-30 cm, to find the average $\rho_b$, $\theta_g$, and $\theta_v$. The $\theta_v$ used for each day will depend on $z^*$, which will be found with validation data described in the next section.

After determining the $z^*$ for each day, the centroid’s in–situ soil samples were abbreviated to that depth. The corresponding soil moistures for each centroid on all five sampling days is displayed in Table 4.1. The standard error of each centroids’ soil moisture value is shown in parentheses.
Table 4.1 Soil moistures for each centroid for the five UV sampling days. Standard errors are shown in parentheses.

<table>
<thead>
<tr>
<th>Cluster group</th>
<th>Cluster</th>
<th>June 12</th>
<th>June 28</th>
<th>July 11</th>
<th>July 27</th>
<th>August 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.252 (0.022)</td>
<td>0.219 (0.018)</td>
<td>0.172 (0.012)</td>
<td>0.119 (0.010)</td>
<td>0.244 (0.012)</td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td>0.240 (0.019)</td>
<td>0.211 (0.008)</td>
<td>0.141 (0.015)</td>
<td>0.125 (0.006)</td>
<td>0.244 (0.012)</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.288 (0.010)</td>
<td>0.209 (0.015)</td>
<td>0.173 (0.014)</td>
<td>0.135 (0.008)</td>
<td>0.226 (0.012)</td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td>0.288 (0.010)</td>
<td>0.209 (0.015)</td>
<td>0.173 (0.014)</td>
<td>0.135 (0.008)</td>
<td>0.226 (0.012)</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.255 (0.019)</td>
<td>0.225 (0.012)</td>
<td>0.204 (0.011)</td>
<td>0.175 (0.009)</td>
<td>0.234 (0.006)</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>0.254 (0.025)</td>
<td>0.238 (0.024)</td>
<td>0.196 (0.011)</td>
<td>0.123 (0.015)</td>
<td>0.243 (0.010)</td>
</tr>
<tr>
<td>4</td>
<td>a</td>
<td>0.274 (0.015)</td>
<td>0.203 (0.015)</td>
<td>0.178 (0.011)</td>
<td>0.132 (0.010)</td>
<td>0.208 (0.011)</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.254 (0.025)</td>
<td>0.238 (0.024)</td>
<td>0.196 (0.011)</td>
<td>0.123 (0.015)</td>
<td>0.243 (0.010)</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>0.255 (0.019)</td>
<td>0.225 (0.012)</td>
<td>0.204 (0.011)</td>
<td>0.175 (0.009)</td>
<td>0.234 (0.006)</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>0.256 (0.016)</td>
<td>0.266 (0.014)</td>
<td>0.145 (0.008)</td>
<td>0.120 (0.008)</td>
<td>0.255 (0.011)</td>
</tr>
</tbody>
</table>
Figure 4.2 All 10 m locations with soil and topographic data within the COSMOS footprint divided into respective cluster groups. Each cluster’s centroid is the larger dot of the same color. Duplicate centroids are represented by the same color.

4.2.3 Validation data

Using the polynomial fits for the 2012 data in Section 3.3.3.2 (Figure 3.10), the vegetation–corrected $N_0$ values for each validation day were found. These modeled $N_0$ values can be seen in Table 4.2. Using these modeled $N_0$ values and the $N$ values for each respective sampling day, we found what the soil moisture was on these days, assuming our vegetation–dependent $N_0$ fit is accurate. These $\theta_v$ values will be our “true” soil moisture values for each day, as we are assuming the COSMOS probe is our ground truth.
Table 4.2  Fresh mass vegetation ($M_f$) amounts (kg m$^{-2}$) found via the fits in Chapter 3 (Figure 3.9) for the five UV sampling days. Using a fit for growing vegetation’s fresh and water masses, the water percentages of the $M_f$ values were found and used to create the vegetation water equivalent ($W_s$) (cm). The $N_0$ values (cph) found for UV sampling days using the vegetation correction function created with data from Chapter 3 (Figure 3.10). Also, the recalculated soil moistures (cm$^3$ cm$^{-3}$) for those days with these $N_0$ values using each respective day’s moderated neutron counts (Figure 3.11). Using our vegetation–dependent COSMOS modeling information, these are considered the “true” values for $N_0$ and $\theta_v$ on these five days. Each day’s $z^*$ is also shown, which assumes all the other data in this table is true. $z^*$ is found with (3.4) by assuming $\theta_v$ is uniform throughout the first 30 cm and $\rho_b$ does not vary by depth or time.

<table>
<thead>
<tr>
<th></th>
<th>$M_f$</th>
<th>% water</th>
<th>$W_s$</th>
<th>$N_0$</th>
<th>$\theta_v$</th>
<th>$z^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>June 12</td>
<td>0.085</td>
<td>83</td>
<td>0.0078</td>
<td>3296</td>
<td>0.211</td>
<td>20</td>
</tr>
<tr>
<td>June 28</td>
<td>0.36</td>
<td>80</td>
<td>0.033</td>
<td>3226</td>
<td>0.190</td>
<td>25</td>
</tr>
<tr>
<td>July 11</td>
<td>0.86</td>
<td>78</td>
<td>0.077</td>
<td>3129</td>
<td>0.130</td>
<td>30</td>
</tr>
<tr>
<td>July 27</td>
<td>1.7</td>
<td>75</td>
<td>0.15</td>
<td>3052</td>
<td>0.124</td>
<td>30</td>
</tr>
<tr>
<td>August 8</td>
<td>2.2</td>
<td>73</td>
<td>0.19</td>
<td>3059</td>
<td>0.211</td>
<td>20</td>
</tr>
</tbody>
</table>

We also found each day’s respective areal fresh mass ($M_f$) value. These were found by our daily stepped model of vegetation over time (Figure 3.9). To determine the subsequent $W_s$ values, we modeled the fraction of fresh vegetation that was water mass using the data in Figure 2.7 during the growing season. The percent of fresh mass that was water is shown in Table 4.2. This allowed us to determine $W_s$ for these five days.

Assuming that $\theta_v$ was uniform in depth, and bulk density was a constant 1.3 g cm$^{-3}$, we were able to determine each day’s subsequent $z^*$. We used (3.4) with these assumptions and values for $\theta_v$, $\rho_b$, and $W_s$ and kept our assumption of $\kappa = 1$. We then determined the $z^*$ that best satisfied the data and used that as the day’s $z^*$. This will be needed for use of the in–situ soil samples so we know how much of the 30 cm soil cores to take for each day.

4.2.4  Weighting function

To determine the areal soil moisture over the COSMOS footprint, the composite of all the locations’ soil moisture needed to be taken. However, as shown in Zreda et al. (2008)
(Figure 4.3(a)), the area within a COSMOS probe’s footprint is not evenly weighted when determining soil moisture. The closer a location is to the probe, the more weight it carries in determining the overall soil moisture. Therefore, a weighting function needs to be used on all the 10 m locations before the areal soil moisture can be determined. To do this, we made an approximation of the function found in Zreda et al. (2008). Where the solid (dry soil) and dashed (wet soil) lines diverge in that figure, a point approximately halfway between the two was used for that distance so that a single function could be used no matter what the soil moisture was at the time. A comparison of our approximation to that of the cumulative distribution function found in Zreda et al. (2008) is shown in Figure 4.3.

![Graph](image)

(a) Figure 3 from Zreda et al. (2008).  
(b) Our approximation of the Zreda et al. (2008) figure.

Figure 4.3  Figure 4.3(a) Cumulative distribution of the how neutron counts are weighted for a COSMOS probe. Figure 4.3(b) We used the same cut off as Zreda et al. (2008), that of 86% or $1 - e^{-2}$, at 350 m and used points between the dry (solid) and saturated (dashed) conditions shown to determine a single function.

A weighting function was then created with our approximation of the footprint’s measurement area. This weighting function was used with each 10 m location’s soil moisture (determined by what cluster it belonged to) and distance from the COSMOS probe. The composite of these weighted soil moistures creates an areal soil moisture that depends on cluster assignments and centroid soil moistures. A table of these areal soil moistures using cluster groups can be found in Table 4.3 along with what the areal, non-weighted average soil moisture would
Table 4.3  COSMOS probe footprint areal soil moistures for cluster groups. A = areal soil moisture found with no weighting function (cm$^3$ cm$^{-3}$). W = areal soil moisture found with weighting function (cm$^3$ cm$^{-3}$).

<table>
<thead>
<tr>
<th>Cluster group</th>
<th>June 12</th>
<th>June 28</th>
<th>July 11</th>
<th>July 27</th>
<th>Aug. 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>W</td>
<td>A</td>
<td>W</td>
<td>A</td>
</tr>
<tr>
<td>1</td>
<td>0.252</td>
<td>0.252</td>
<td>0.219</td>
<td>0.219</td>
<td>0.172</td>
</tr>
<tr>
<td>2</td>
<td>0.266</td>
<td>0.269</td>
<td>0.210</td>
<td>0.210</td>
<td>0.159</td>
</tr>
<tr>
<td>3</td>
<td>0.271</td>
<td>0.274</td>
<td>0.220</td>
<td>0.219</td>
<td>0.186</td>
</tr>
<tr>
<td>4</td>
<td>0.261</td>
<td>0.262</td>
<td>0.232</td>
<td>0.229</td>
<td>0.180</td>
</tr>
</tbody>
</table>

have been for the cluster group.

4.3 Results

An examination of Table 4.3 suggests that weighting the soil moistures by distance from the probe may not have been necessary. Clearly, with a 1 cluster group, a weighting function does not change the mean soil moisture as every location is assumed to be the same soil moisture. Of cluster groups 2, 3, and 4, the largest difference between non–weighted soil moisture and weighted soil moisture is 0.003 cm$^3$ cm$^{-3}$.

Using the vegetation adjusted COSMOS soil moisture as the true soil moisture, a comparison of how the cluster groups compared to the true soil moisture on each of the five days is shown in Table 4.4. For all but one comparison, the FSI cluster groups overestimated the soil moisture. The largest overestimation was 0.063 cm$^3$ cm$^{-3}$ on June 12 for the 3 cluster group. The closest comparison of FSI soil moisture to the true COSMOS soil moisture was the 1 cluster group on July 27, which showed a difference of 0.006 cm$^3$ cm$^{-3}$. The cluster groups did the best upscaling on August 8 where the largest discrepancy was 0.016 cm$^3$ cm$^{-3}$. The worse upscaling comparison was on June 12 where the smallest discrepancy was 0.041 cm$^3$ cm$^{-3}$.

Comparing the four different cluster groups to each other, there is no obvious cluster group that upscaled soil moisture better. Cluster groups 2 and 3 had three upscaled soil moistures within 0.03 cm$^3$ cm$^{-3}$ of the true soil moisture while the 1 and 4 cluster groups each had two within that frame. However, the 2 and 3 cluster groups also had the largest differences, that of
Table 4.4  The difference between the true soil moisture (Table 4.2) and the upscaled soil moistures (Table 4.3) in cm$^3$ cm$^{-3}$ for all cluster groups on all UV days. Positive values represent differences when the upscaled soil moisture overestimated the true soil moisture while negative values represent when the upscaled soil moisture is lower than the true soil moisture.

<table>
<thead>
<tr>
<th>Cluster group</th>
<th>June 12</th>
<th>June 28</th>
<th>July 11</th>
<th>July 27</th>
<th>August 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+0.041</td>
<td>+0.029</td>
<td>+0.042</td>
<td>−0.006</td>
<td>+0.033</td>
</tr>
<tr>
<td>2</td>
<td>+0.058</td>
<td>+0.020</td>
<td>+0.031</td>
<td>+0.007</td>
<td>+0.022</td>
</tr>
<tr>
<td>3</td>
<td>+0.063</td>
<td>+0.029</td>
<td>+0.055</td>
<td>+0.016</td>
<td>+0.020</td>
</tr>
<tr>
<td>4</td>
<td>+0.051</td>
<td>+0.039</td>
<td>+0.049</td>
<td>+0.012</td>
<td>+0.020</td>
</tr>
</tbody>
</table>

0.058 and 0.063 cm$^3$ cm$^{-3}$. Cluster group 1 was the only to underestimated the soil moisture. The average differences of upscaled soil moisture to true soil moisture were 0.030, 0.028, 0.037 and 0.034 cm$^3$ cm$^{-3}$ for the four cluster groups, respectively. That comparison suggests that 3 cluster group was actually the worst at upscaling the soil moisture and that 2 cluster group was best.

4.4 Conclusions

After correcting for vegetation, the COSMOS probe’s soil moisture values are considered to be the ground truth. These soil moistures were found with our fit of $N_0$ over time (Figure 3.10) for five sampling days in 2012. We found the days’ $N_0$ values and, with moderated neutron data from http://cosmos.hwr.arizona.edu/Probes/StationDat/016/corcounts.txt, found the day’s soil moisture via (1.1). We also found the vegetation amounts for these five days with our vegetation model from Chapter 3 (Figure 3.9). The effective measuring depth ($z^*$) was also found by (3.4) for these days by converting $M_f$ to $W_s$, assuming a constant $\rho_b$, and assuming a uniform soil moisture profile.

With the FSI method and using topographic and soil data obtained at the IVS, locations within the COSMOS probe’s footprint were separated into clusters for four different groups. The centroid of each cluster was obtained and in–situ soil moisture found at the centroid was assigned to the whole cluster. Depending on the $z^*$ found, the in–situ samples’ volumetric
soil moisture data was used to that depth. We used a weighting function to combine the soil moisture at each location within the probe’s footprint for comparison to the true soil moisture.

Considering all the assumptions we made, we felt the FSI method did a respectable job of upscaling soil moisture. We assumed our vegetation and $N_0$ fits from Chapter 3 were accurate and true, which they may not have been. We also assumed that bulk density and soil moisture did not vary by depth, which looking at previous data we had, we knew was not accurate. We also assumed bulk density did not vary by time which it does (Logsdon, 2012). We thought the $z^*$ values found were a little deep, perhaps as a result of our assumptions of bulk density and soil moisture. Using $z^*$ of 25 and 30 cm seemed quite deep for us as the deepest we had seen in our previous work was 20 cm. We assumed that the soil moisture at each centroid best represented the soil moisture for the whole cluster, which it may not have. Soil moisture and how it varies is still not known perfectly though, so we thought that a maximum difference of 0.063 cm$^3$ cm$^{-3}$ was acceptable, although high. Upscaling soil moisture within a range such as 0.03 cm$^3$ cm$^{-3}$ would have been better as both SMOS and SMAP are accurate to ±0.04 cm$^3$ cm$^{-3}$ (Kerr et al., 2010; Entekhabi et al., 2010). A difference of 0.06 cm$^3$ cm$^{-3}$, when soil moisture values typically fall between 0.05 and 0.40 cm$^3$ cm$^{-3}$, leaves a lot of room for error.
CHAPTER 5. CONCLUSIONS

5.1 Soil and Vegetation Conclusions

By taking soil moisture samples around a COSMOS probe, we could recalibrate the probe in order to test if $N_0$ varied over time or was a constant. We took soil samples down to 30 cm as the probe usually only senses 15-20 cm into the soil. We also found a way to measure the vegetation present in the COSMOS probe’s footprint. We used an allometric relationship that tied the product of stem diameter squared and canopy height to plant mass. We used 30 plants from the edge of the field to make multiple allometric relationships throughout the 2011 and 2012 crop seasons. We also measured plants in areas around the COSMOS probe and, by using these allometric relationships, were able to project a mass of plants around the probe. Paired with plant density, we could scale our measurements up to a plant mass per area value. In trying to determine the error in our vegetation measurements, we took measurements of large areas of plants to determine the natural variability of plants in a field. Our choice of measuring 5 maize/10 soybean plants at each location around the probe instead of more may have resulted in a larger error than if we had measured 30 plants or more. However, we still found error values for our areal vegetation masses that were fairly small. Our largest error for maize was 0.24 kg m$^{-2}$ and for soybean was 0.050 kg m$^{-2}$.

5.2 COSMOS Calibration Conclusions

Using our soil moisture and vegetation data determined in the previous section, we were able to recalibrate the COSMOS probe on various sampling days. We found a way to approximate the depth of measurement of the probe ($z^*$) by converting fresh mass measurements to water equivalents and using in-situ bulk density and volumetric soil moisture data. We were then
able to determine what \( N_0 \) should be for each day we collected samples. \( N_0 \) was found to vary and, when compared to the vegetation present, a relationship was found between the two.

We created fits to our measurements of vegetation over time and \( N_0 \) over vegetation and used them to create a function of \( N_0 \) over time. Using this function and moderated neutron counts measured by the COSMOS probe, we were able to adjust the probe’s time series of soil moisture. Without the adjustment to \( N_0 \), the COSMOS probe was underestimating soil moisture levels throughout the season. The probe was more accurate in 2012 when soybean crop was again present as that is the vegetative cover that was present on the day of the initial \( N_0 \) calibration. We proved that \( N_0 \) is not a constant, but is in fact dependent on the presence of vegetation. The two types of vegetation observed had similarly patterned effects but different magnitudes of an effect on the probe.

### 5.3 FSI Validation Conclusions

Using our modeled \( N_0 \) values to adjust the COSMOS probe’s soil moisture readings, we now had true field–scale soil moisture at the IVS. We used this as our ground truth soil moisture for comparison of an upscaling method. FSI is an upscaling method that uses soil and topographic data, separated into cluster groups by the K–means clustering algorithm, to upscale point measurements to a field–scale soil moisture. We looked at four different cluster groups (1, 2, 3, 4) and took in–situ soil moisture measurements at specific locations on five days in 2012 at the IVS.

With our data on the true soil moisture and vegetation present, we were able to determine a \( z^* \) for the days we took in–situ soil samples at the cluster centroids. We were then able to use our 30 cm in–situ measurements to an appropriate depth. After assigning centroid soil moistures to the rest of the locations within the cluster, we used a weighting function dependent on distance from the probe to determine the corresponding upscaled COSMOS footprint soil moisture. A comparison of the upscaled soil moisture and true soil moisture resulted in soil moisture differences of 0.006 cm\(^3\) cm\(^{-3}\) to 0.063 cm\(^3\) cm\(^{-3}\). A look at how the four different cluster groups’ upscaled soil moisture varied showed that no one group systematically upscaled the soil moisture better than the rest.
5.4 Future Work

There are many possible areas of future work with this project. Relating to the vegetation measurements and variability, determination of above-ground biomass (residue) would be beneficial. We did not measure the residue on days we sampled where there was significant residue; we assumed bare soil conditions. This could affect our $N_0$ calculations because we did not account for all the vegetative hydrogen present. Therefore, the COSMOS probe would still be underestimating soil moisture during periods of residue cover. As maize plants leave a large amount of residue that takes a long time to decompose, this could be significant.

Related to residue determination would be root biomass. Root biomass is another form of vegetative hydrogen which was neglected in this study but could prove to be significant. As much as 90% of maize root biomass has been found to be located in the top 30 cm of soil profiles with almost 70% of all root biomass existing in the top 22.5 cm (Amos and Walters, 2006). Mitchell and Russell (1971) found that for soybeans 90% of root dry weight was concentrated in the upper 7.5 cm in the early stages of growth and in the top 15 cm during the rest of the season. As root mass has also been found to scale to leaf mass to the $3/4$ power, and approximately the same as stem mass (Enquist and Niklas, 2002), this could amount to significant root biomass and therefore significant hydrogen in the soil zone that the COSMOS probe is sensing. We had not thought to try to quantify root biomass and after we had realized it may be significant, we could not determine a good method for sampling it. Allometric relationships may be needed to determine the amount of root biomass produced by plants.

Our work suggested that soybean plants’ variability was time and/or location dependent. However, we assumed that it was both time and location independent in our approximation of the vegetation variability ($\eta$). To get a better result of the vegetation variability, measurements of the true field average (PV days) would need to be made at the same time as allometric plant measurements (VC days). While this could be a lot of work, it would allow for a better error estimate for soybeans. We did not see any time dependent component on maize plant variability, but there could be, or perhaps there is a location dependency.

An examination of the variability of the mean stem diameter squared times canopy height
for plants (var(\(x\))) showed that samples of more than 5 maize/10 soybean plants might also help reduce the error in the approximation of vegetation variability. Sampling 30 or more plants at each location could better account for the natural variability of plants in a field. This may increase the confidence in projected vegetation amounts.

Another area of research from this project would be a closer examination of the vegetation distribution factor (\(\kappa\)) and how it might affect \(z^*\). While a \(\kappa\) of 1 seemed to work for our row crops of maize and soybean, perhaps there is a better way to determine the plant uniformity. It seems counterintuitive that row crops could be modeled as completely uniform. We would think that non–row crops (such as grasses) would better fit the uniform distribution as they would not have the distinct row pattern to them. \(\kappa\) may be found to differ for different crops as well. Perhaps soybean plants fit a \(\kappa\) of 1 but maize plants better fit a \(\kappa < 1\).

It has been thought that the amount of vegetation within the probe’s footprint could be accounted for by directly modeling neutron scattering (Zreda et al., 2008; Desilets et al., 2010). We did a quick examination of the measurement of thermal neutrons by the IVS probe and found that thermal neutrons behaved more proportionally to the vegetation amounts, whereas moderated neutrons behaved inversely to the increased presence of hydrogen. If a relationship can be found between thermal neutron counts and vegetation amounts, this could prove to be a better method for estimation of the presence of vegetation compared to taking vegetation samples and measurements like we did.
BIBLIOGRAPHY


Mitchell, R. L., and W. J. Russell (1971), Root development and rooting patterns of soybean 
(*Glycine max* (L.) Merrill) evaluated under field conditions, *Agron. J.*, 63, 313–316, doi: 
10.2134/agronj1971.00021962006300020034x.

Mohanty, B. P., T. H. Skaggs, and J. S. Famiglietti (2000a), Analysis and mapping of field– 
scale soil moisture variability using high–resolution, ground–based data during the Southern 

Mohanty, B. P., J. S. Famiglietti, and T. H. Skaggs (2000b), Evolution of soil moisture spatial 
structure in a mixed vegetation pixel during the Southern Great Plains 1997 (SGP97) Hydrology 

Moore, I. D., G. J. Burch, and D. H. Mackenzie (1988), Topographic effects on the distribution 


Pedersen, P. (2009), *Soybean growth and development*, Iowa State University Extension, Ames, 
IA.


Raun, W. R., J. B. Solie, G. V. Johnson, M. L. Stone, R. W. Whitney, H. L. Lees, H. Sembiring, 

Reddy, V. R., Y. A. Pachepsky, and F. D. Whisler (1998), Allometric relationships in field– 

Reid, I. (1973), The influence of slope orientation upon the soil moisture regime, and its hy-


