Evaluation of atmospheric land exchange inverse model evaporative stress index utilizing soil climate analysis network stations in Alabama

Corey Walker

Follow this and additional works at: https://louis.uah.edu/uah-theses

Recommended Citation

This Thesis is brought to you for free and open access by the UAH Electronic Theses and Dissertations at LOUIS. It has been accepted for inclusion in Theses by an authorized administrator of LOUIS.
Evaluation of Atmospheric Land Exchange
Inverse Model Evaporative Stress Index
Utilizing Soil Climate Analysis Network
Stations in Alabama

Corey Walker

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Earth System Science
in
The Department of Atmospheric and Earth Science
to
The Graduate School
of
The University of Alabama in Huntsville

August 2023

Approved by:
Dr. Lee Ellenburg, Research Advisor
Dr. John Mecikalski, Committee Chair
Dr. Vikalp Mishra, Committee Member
Dr. John Christy, Committee Member
Dr. Lawrence Carey, Department Chair
Dr. Rainer Steinwandt, College Dean
Dr. Jon Hakkila, Graduate Dean
Abstract

Evaluation of Atmospheric Land Exchange Inverse Model Evaporative Stress Index Utilizing Soil Climate Analysis Network Stations in Alabama

Corey Walker

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Earth System Science

Atmospheric and Earth Science
The University of Alabama in Huntsville
August 2023

The Atmospheric Land Exchange Inverse (ALEXI) model uses Geostationary Operational Environmental Satellite (GOES) thermal bands to derive daily evapotranspiration estimates at the continental scale. These data are used to derive a weekly Evaporative Stress Index (ESI) representing standardized anomalies of the actual to potential evapotranspiration ratio (AET/PET). ESI is functionally related to the root zone soil moisture content. As the AET/PET ratio approaches unity, the greater the available water at the root zone is for agricultural fields, whereas the opposite is true during periods of drought. However, the relationship between ALEXI ESI and root zone soil moisture may not be linear or consistent across space or time. Therefore, there is a desire to better understand the relationship between ALEXI ESI and real soil moisture values, as well as an empirical knowledge of what ESI means in relation to point measurements of soil moisture on the ground. This thesis will provide such discussion on ESI and weekly-corrected, in-situ, volumetric soil moisture values from a collection of SCAN stations across Alabama. The implications of this
research will lead to a better understanding of how ALEXI ESI can be used as an indicator for agriculture and flash drought in the Southeastern U.S.
Acknowledgements

Funding for this research was provided by the National Oceanic and Atmospheric Administration (NOAA). I would first like to thank Dr. John Mecikalski for taking a chance on me and providing this opportunity to learn, work, and grow as a scientist. I would also like to thank my research advisor, Dr. Lee Ellenburg. He has provided constant mentoring, helping drive my research down the right path with great patience. I have developed a better understanding of Python, statistics and the field of hydrology because of this man, and for that I am grateful.

There are others who I would also like to thank for their support in this work. One would be Dr. Christopher Hain, who is a research scientist with the National Aeronautics and Space Administration (NASA). He provided me with prompt deliveries of ESI data on a web server when I needed it and sat down with me often to go over technical details I did not understand. I would also like to thank Dr. Vikalp Mishra for the many times he has offered advice and understanding of the ALEXI model. I would like to particularly thank him for letting me read his thesis work so that I could get a better understanding of what a great, final product looks like.

Another individual that I need to thank is Dr. John Christy. In a lot of ways, I cannot believe that my work will be read and criticized by the person leading the State Climatology Office for Alabama. In his class I gained a better understanding of the climate, drought and how data pertaining to these topics are created. I am appreciative of his willingness to teach me the importance of skepticism and making sure data are at the forefront of scientific argument and discovery.

I would also like to take a moment to appreciate my wife, Emily for being patient with me during the process of graduate school and for supporting me emotionally and financially. At the time of this thesis being published, we were working
on raising our first child together, whose name is Reid. He is now two years old at the time of this being published, and I hope he will one day find this document and see his name being mentioned. I love our son so much, and cannot believe that I get the opportunity to know and guide him as his father.

In particular, I just want to say how appreciative I am for all of my family, especially both of my grandmothers, MeMe and Nanny, my PawPaw, Mom, Dad, and step mother, Kareena, for the time spent raising me into the person I am today. All of the people mentioned above were major contributions to my success and allowed me to move forward towards my dream of becoming a scientist, and for that I am deeply thankful.

Lastly, I just want to take a second to reflect on the person that I was before and after the experience of graduate school. I am very proud of the person you’ve become, Corey. I am proud of the skills, knowledge and personal development that has come from you taking a leap of faith into this work. Here’s to you for making a small step forward and contributing to scientific knowledge. You did it! Yay!
Table of Contents

Abstract ................................................................. ii

Acknowledgements ...................................................... v

Table of Contents ...................................................... ix

List of Figures .......................................................... x

List of Tables ........................................................... xv

Chapter 1. Introduction and Purpose ................................. 1
   1.1 Background ....................................................... 1
   1.2 Research Questions ............................................. 6

Chapter 2. Background Review ........................................ 8
   2.1 Evapotranspiration .............................................. 8
   2.2 Evapotranspiration Equations .................................. 10
   2.3 ALEXI ............................................................ 13
   2.4 ESI .............................................................. 16
   2.5 SCAN .......................................................... 22
Chapter 3. Data and Methods ............................................ 26
  3.1 Summary ......................................................... 26
  3.2 Data .......................................................... 27
    3.2.1 ESI ......................................................... 27
    3.2.2 Volumetric Soil Moisture ................................. 29
    3.2.3 Drought Data .............................................. 31
    3.2.4 Soil Reports ............................................... 34
    3.2.5 Land Cover Data .......................................... 37
  3.3 Methods ....................................................... 39
    3.3.1 Statistical Assessment .................................... 39
    3.3.2 Accuracy Assessment ..................................... 42
    3.3.3 Flash Drought Assessment ................................. 43

Chapter 4. Results .................................................... 44
  4.1 Data Statistics ................................................ 44
  4.2 Statistical Results ............................................ 46
  4.3 Accuracy Results .............................................. 82
  4.4 Flash Drought Assessment Results ............................ 101

Chapter 5. Summary and Discussion ................................. 121
5.1 Summary of Results ........................................... 121

5.1.1 Question 1 .................................................... 122

5.1.2 Question 2 .................................................... 123

5.1.3 Question 3 .................................................... 125

5.2 Limitations and Recommendations .............................. 126

References ............................................................. 130
## List of Figures

1.1 Standardized Precipitation Index (SPI), Precipitation (PCP) and United States Drought Monitor (DM) data compared to ESI at 2 (CO2), 4 (CO4), and 8 (CO8) week composites (means) and change anomalies computed across 1, 2, 3 and 4 week (WK) difference schemes [50]. .......................................................... 5

2.1 The global annual mean energy flow calculated from CERES Re- analysis from March 2000 to May 2004. .................................................. 9

2.2 A) Climatology (thick blue line) for 38 years of root zone soil moisture versus time series for year 2000 (dotted blue line) and daily precipitation (red bars) [41] ......................................................... 11

2.3 Component H from canopy and surface as related to T_s and T_c, left and time integrated measurement of H at t_1 and t_2 and its influence on the slope of the boundary layer height relationship θ_s(z), right [27]. .............................................................. 15

2.4 28 day ESI composite for six months of the growing season, left and lower resolution Palmer Index, right in 2002 [28]. .................. 18

2.5 Top figure: Temperature Anomaly (black), Dew-point Depression (blue), Wind Speed (red) and Cloud Cover Anomaly (Green). Bottom Figure: US Drought Monitor Category (black line), Precipitation (red), ESI (blue, green and yellow separated by intervals of two, four and eight weeks respectively). ........................................ 21

2.6 A typical SCAN site [1]. ................................................................. 23

2.7 θ_v sensors at three depths in the soil column [1]. .................................. 24

3.1 All SCAN Sites Across Alabama [9]. .................................................. 28

3.2 2016 flash drought event identified at SCAN station 2174 (highlighted in red) by negative SMVI (blue) values over 21 days. .... 33

3.3 Soil Hydrologic Groups by Soil Type [66]. ....................................... 36
3.4 NLCD Land Cover Classes by Type at Station 2115 for the ESI Pixel Domain. 38

4.1 All Stations Pearson R Values by Month and Depth. 48

4.2 Pearson R Values for 2053:AL:SCAN by Month and Depth (bottom panel) and Calculated Dominance Values of 8 Land Cover Classes using HHI (top panel). 50

4.3 Pearson R Values for 2179:AL:SCAN by Month and Depth (bottom panel) and Calculated Dominance Values of 8 Land Cover Classes using HHI (top panel). 51

4.4 All Stations Average Pearson R Values by Month and Depth for Agricultural Land Cover (Agland) Type. 56

4.5 All Stations Average Pearson R Values by Month and Depth for Forest Land Cover Type. 57

4.6 All Stations Average 4inANOM Pearson R Values by Month and Dominant Land Cover Type. 59

4.7 Significant $p < 0.05$ 2inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type. 61

4.8 Significant $p < 0.05$ 4inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type. 62

4.9 Significant $p < 0.05$ 8inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type. 63

4.10 Significant $p < 0.05$ 20inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type. 64

4.11 Significant $p < 0.05$ 40inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type. 65

4.12 4inANOM vs 1 week ESI at 2053:AL:SCAN during Drought Year 2006. 68

4.13 Month 10 4inANOM Week Lag Pearson R Values at 2078:AL:SCAN Shifted to the Right and Left of 1 Week ESI by Increments of 1 Week. 70
4.14 All Stations 4inANOM Average Pearson R Values for Week Lag Interval (WLI) by Month. .................................................. 71
4.15 All Stations 4inANOM Seasonal Pearson R Values for Week Lag Interval (WLI) .................................................. 73
4.16 All Stations Average 2inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type. ................................. 73
4.17 All Stations Average 2inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month. ........................................ 75
4.18 All Stations Average 4inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type. ................................. 76
4.19 All Stations Average 4inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month. ........................................ 76
4.20 All Stations Average 8inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type. ................................. 77
4.21 All Stations Average 8inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month. ........................................ 77
4.22 All Stations Average 20inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type. ................................. 78
4.23 All Stations Average 20inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month. ........................................ 78
4.24 All Stations Average 40inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type. ................................. 79
4.25 All Stations Average 40inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month. ........................................ 79
4.26 Count of Total Soil Moisture Sensors by Hydrologic Soil Type. . 81
4.27 Confusion Matrix of ESI Prediction Accuracy for 2inANOM. . . 82
4.28 Confusion Matrix of ESI Prediction Accuracy for 4inANOM. . . 83
4.29 Confusion Matrix of ESI Prediction Accuracy for 8inANOM. . . 83
4.30 Confusion Matrix of ESI Prediction Accuracy for 20inANOM. . . 84
4.31 Confusion Matrix of ESI Prediction Accuracy for 40inANOM. . . 84

xii
4.32 Overall Accuracy for ESI Prediction by Depth ............... 85
4.33 All Stations Binned 1 Week ESI Prediction Accuracy by Depth . 86
4.34 ESI Prediction Accuracy by All Depths and Months ............. 87
4.35 ESI Prediction Accuracy by Month and Depth for Agricultural Land Cover Type .................................................. 90
4.36 ESI Prediction Accuracy by Month and Depth for Forest Land Cover Type .................................................. 91
4.37 All Stations Average 4inANOM ESI Prediction Accuracy by Month and Dominant Land Cover Type .......................... 92
4.38 Month 10 4inANOM Week Lag Accuracy Values at 2078:AL:SCAN Shifted to the Right and Left of 1 Week ESI by Increments of 1 Week .................................................. 93
4.39 All Stations 4inANOM Seasonal Accuracy Values for Week Lag Interval (WLI). DJF = December, January, February. MAM = March, April, May. JJA = June, July, August. SON = September, October, November .................................................. 95
4.40 All Stations Count of Binned 1 Week ESI Values ................. 96
4.41 All Stations 2inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type .................................................. 97
4.42 All Stations 4inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type .................................................. 97
4.43 All Stations 8inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type .................................................. 98
4.44 All Stations 20inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type .................................................. 98
4.45 All Stations 40inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type .................................................. 99
4.46 All Stations Binned 1 Week ESI Value versus 4inANOM Percentile Rank for Growing Season (Months 5-11) .............. 102
4.47 All Stations Binned 1 Week ESI Value versus 4inANOM Percentile Rank and Soil Type for Growing Season (Months 5-11) .... 103
4.48 Months 5-11 Total SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Hydrologic Soil Type. ............................... 106
4.49 Months 5-11 Total SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Dominant Land Cover. ............................... 107
4.50 Count of 224 SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Month. ......................................................... 108
4.51 Month 5-11 ESI True Negative Prediction of Flash Drought by Variable and Index. .......................................................... 110
4.52 ESI True Negative Prediction of Flash Drought by Variable and Index for Significant Station Months. ................................. 111
4.53 Months 5-11 95% Confidence Interval 1 Week ESI Change Anomaly (2 Week Difference) versus Flash Drought Duration Months. . 113
4.54 Months 5-11 95% Confidence Interval 4 Week ESI Change Anomaly (3 Week Difference) versus Flash Drought Duration. ........ 114
4.55 Months 5-11 95% Confidence Interval 8 Week ESI Change Anomaly (3 Week Difference) versus Flash Drought Duration Months. . 115
4.56 Months 5-11 95% Confidence Interval 8 Week ESI Change Anomaly (4 Week Difference) versus Flash Drought Duration. ........ 116
4.57 Months 5-11 8 Week ESI Change Anomaly (4 Week Difference) versus Flash Drought Duration. Number of Flash Droughts Reaching Each Week shown in Each Interquartile Range. ................. 117
4.58 Station 2180 Flash Drought Event versus 8 Week Change Anomaly (4 Week Difference), 4inANOM Soil Moisture Percentiles, and USDM Category). ......................................................... 119
4.59 Station 2174 Flash Drought Event versus 8 Week Change Anomaly (3 Week Difference), 4inANOM Soil Moisture Percentiles, and USDM Category). ......................................................... 120


## List of Tables

3.1 Soil Hydrologic Groups and Runoff Potential [57] ..................... 35  
3.2 Correlation Coefficient Strength Interpretation ..................... 40  

4.1 Description of Statistics for ESI Data alongside Soil Moisture Observations for Each Depth. ........................................ 45  
4.2 Description of Station Period Of Record and Number of Week Observations. ..................................................... 45  
4.3 All Stations Pearson R Correlation by Month and Depth ................. 47  
4.4 All Stations Pearson R Correlation by Month and Depth Descriptive Statistics .................................................. 47  
4.5 2053:AL:SCAN (Strongly Agricultural) Pearson R Correlation by Month and Depth ............................................. 52  
4.6 2179:AL:SCAN (Strongly Forest) Pearson R Correlation by Month and Depth ................................... 53  
4.7 2053:AL:SCAN (Strongly Agricultural) Pearson R Correlation by Month and Depth Descriptive Statistics ...................... 55  
4.8 2179:AL:SCAN (Strongly Forest) Pearson R Correlation by Month and Depth Descriptive Statistics .............. 55  
4.9 Dominant Land Cover by Station ........................................... 66  
4.10 All Stations ESI Prediction Accuracy by Depth and Month ............. 88  
4.11 All Stations ESI Prediction Accuracy by Depth and Month Descriptive Statistics ........................................ 88  
4.12 Statistics for 224 SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor including CPC Soil Moisture Drought Category Classifications (Column 1), Drought Duration (Column 2), and Total Change in 4inANOM Percentile During Event (Column 3) . 104
Chapter 1. Introduction and Purpose

1.1 Background

Currently, 38.5 thousand farms operate on 8.2 million acres in Alabama, with significant production in livestock. Alabama is number 24 in the United States for head of cattle, with 1.3 million being counted in 2021 across its industry [14]. Alabama’s crop outputs are also noteworthy, with the state producing 610 million pounds of peanuts, 690 thousand bales of cotton and 2.17 million tons of hay in 2021 [14]. Agricultural output at this scale creates large demand on water resources. In 2015, Alabama farms used 26 and 223 million gallons of water per day for livestock and irrigation purposes alone [3]. However, one challenge that exists is that farms that irrigate account for only 4 percent of the total. Currently, around 96 percent of Alabama farms rely on rain fed irrigation methods [13]. Because of this, monitoring surface water resources is hugely important to Alabama’s economic and food security.

To help farmers mitigate and prepare for dry climate periods, the Alabama Office of the State Climatologist (AOSC) has an interest in providing timely maps of drought on a state-wide scale. For example, AOSC actively maintains a Lawn and Garden [7], Crop Moisture, [6], and Palmer Drought Severity Index [8]. The federal government also provides similar maps for Alabama that are used to assess
drought conditions and aid farmers [15]. For example, the United States Drought Monitor (USDM) is used to trigger programs such as those providing disaster recovery, water infrastructure, eligibility for low-interest loans and tax deferral on forced livestock sales due to economic losses from drought [20]. End users have an interest in these maps being accurate given the economic incentives that could be lost otherwise. Between 2010 and 2019, $236,050,457 was paid out by the USDA to farms in Alabama for crop losses due to drought [23]. This number may need to be adjusted depending on the accuracy of the assessments used in determining payouts.

One challenge that currently exists for drought assessments is the ability for maps to capture flash drought. These are droughts that occur with rapid onset due to lower-than-normal rates of precipitation, accompanied by abnormally high temperatures, winds, and radiation [19]. Flash droughts occur on the week to month timescale as opposed to conventional droughts which take place over years and decades and therefore limit time for mitigation and preparation [51]. Tools like the USDM offer a weekly snapshot of the evolution of a set of drought categories representing multiple hydrologic data sets interpolated over the Continental United States (CONUS). Because of the averaging that occurs to produce a spatially large data set, the USDM has been shown to respond 1-2 weeks after flash drought events in Oklahoma, Arkansas, Wisconsin and Ohio [33]. Ford and others also showed that in-situ soil moisture detected flash drought events in Oklahoma with 2-3 week lead times prior to the USDM [37]. This means that
new tools are needed to determine flash droughts before they are identified by products with increased lag time.

Flash droughts are problematic because they create exceptional impact to the agricultural economy in the Southeast. For example, the September 2019 event was historic, with six percent of the region in drought at the beginning of September rising to over forty four percent at the end [43]. This flash drought caused grasslands to die off, which meant that farmers had to feed livestock expensive hay. Water supplies in small streams and lakes dried up, and dry soils caused the ground to harden, making fall planting of grains harder. Additionally, blueberry and peach harvests were lost from intense heat [40]. This flash drought occurred due to a lack of rainfall from tropical storms and hurricanes, causing states like Mississippi, Georgia and Alabama to see their driest September on record in the last 125 years [22].

To mitigate the economic impacts of flash drought, prediction of key variables such as soil moisture have become relevant to stakeholders, but accurately measuring it over large spatial domains is challenging. For example, in the state of Alabama, there are only eighteen SCAN sites measuring volumetric soil moisture at the plant root-zone (5-100 cm), thus the majority of areas in the state do not have data descriptive of their soil moisture conditions. Heterogeneous soil conditions also cause soil moisture to behave differently from one location to the next. For example, coarser soils like sands have wider pore spacing than soils with a high percentage of silt or clay leading to increased infiltration rates and drying [30]. Because of this, there are large gaps in spatial coverage that would
be expensive or laborious to understand and rectify. Therefore, models are being tested across a variety of conditions to extend our understanding of soil moisture, such as ALEXI ESI (more extensively discussed in Chapter 2).

Otkin 2013 showed that ESI is related to precipitation anomalies through the use of a visualization method similar to a Hovmöller diagram. Figure 1.1 [50] shows an example of this approach, where a vertical time-series of composited ESI change anomalies are visually inspected alongside precipitation and drought measurement variables. This approach has shown that ESI can represent flash drought prior to onset, and that there are specific relationships to precipitation depletion. However, the downside of a qualitative approach like this is that it does not provide descriptive statistics on what is typically seen for precipitation across a wider climatology of flash droughts. These figures also do not provide information on other water resource variables like soil moisture. Therefore, one solution might be found in comparing various ESI change anomalies to real measurements of soil moisture across a set of identified flash drought events.

Yin in 2018 evaluated drought events using a blended ESI, satellite and soil moisture model [69]. With this research, ESI was defined in relationship to drought categories of soil moisture and precipitation. However, it was suggested that errors in detection of drought events using this methodology arose from data quality issues arising from seasonal accuracy errors in microwave soil moisture retrieval, as well as model bias and data noise. Otkin evaluated a flash drought event in 2015 using ESI, but suggested errors in drought detection due to bias in modeled soil moisture profiles [52]. The Yin and Otkin studies show the increased
Figure 1.1: Standardized Precipitation Index (SPI), Precipitation (PCP) and United States Drought Monitor (DM) data compared to ESI at 2 (CO2), 4 (CO4), and 8 (CO8) week composites (means) and change anomalies computed across 1, 2, 3 and 4 week (WK) difference schemes [50]. As an example, $\Delta$ESI.02WK represents the ESI change anomaly using a 2 week composite, and the corresponding column represents the difference of 1WK applied between anomaly time stamps. Z-Anomaly represents the number of standard deviations the value is from the mean.
importance of determining the if and when water resources like soil moisture can be detected with remote sensing products such as ESI.

1.2 Research Questions

This research aims to address the demand from farmers in Alabama seeking a better understanding, prediction, and monitoring capability of flash drought. To achieve this, the research will investigate three scientific questions pertaining to the relationship between ESI and soil moisture:

1.) What is the statistical relationship between ESI and in-situ soil moisture retrievals at 18 SCAN sites across Alabama? 2.) How accurate is ESI in generalizing in-situ soil moisture conditions across varying soil and land cover characteristics? 3.) What value can ESI provide in monitoring flash drought events identified by soil moisture retrievals? The hypothesis of this study posits that remotely sensed ESI and in situ soil moisture sensors can be utilized to improve the understanding of flash drought across the Southeastern United States.

The reason this study is important is because it will provide greater understanding of what ALEXI ESI values mean in relation to root-zone soil moisture conditions across time and space at increased scale. This in turn could help farmers, scientists and other stakeholders quantify soil moisture without the need for ancillary meteorological data, such as those measured by a ground-based, in-situ sensors. In order to answer these questions, knowledge governing the ALEXI ESI product, such as concepts from thermal remote sensing and hydrologic physics need to be discussed. There will also be a need to understand current research
being utilized to characterize drought as well as the equipment, tools and equations used to retrieve soil moisture data.

This study aims to answer three important science questions surrounding the relationship between ESI and in-situ soil moisture retrievals across Alabama. The hypothesis suggests that the use of ESI and in-situ soil moisture retrievals can lead to a better understanding of flash drought across the Southeastern United States. To validate this claim, this work begins by highlighting important background research in Chapter 2. Then, a detailed analysis of data and methods is described in Chapter 3. Chapter 4 will provide a thorough overview and description of results. Lastly, Chapter 5 will summarize and outline recommendations for future studies.
Chapter 2. Background Review

2.1 Evapotranspiration

ET is water moving to the atmosphere from the soil through a combination of evaporation from the surface and transpiration from plant parts such as stems, leaves, branches and canopies [42]. As ET occurs, water goes through a phase change from liquid to gas causing a release of heat from the surface, thereby causing cooling. This latent energy flux represented by $\lambda E$ can be calculated using Eqs. 2.1 and 2.2, where $\lambda$ is the specific latent heat of evaporation, $E$ is the evaporation rate, $R_n$ is net radiation, $H$ is the sensible heat flux, and $G$ is soil heat conduction, each in Wm$^2$ [42]:

$$R_n = \lambda E + H + G \quad \text{(2.1)}$$

$$\lambda E = R_n - H - G. \quad \text{(2.2)}$$

Researchers have developed multiple large-scale atmospheric reanalysis studies to approximate $\lambda E$ for the global energy budget using short-term, numerical weather models. For example, Figure 2.1 [42] shows the ensemble average of the Clouds and the Earth’s Radiant Energy System (CERES) reanalysis between 2000-2004, which provide an estimate of the net global $\lambda E$ of 80 Wm$^2$. It should
Figure 2.1: The global annual mean energy flow calculated from CERES Reanalysis from March 2000 to May 2004. Arrows represent directional fluxes in Wm$^{-2}$ [42]. Arrows describe the direction of radiation flow as input (down) or output (up) with respect to the Earth System.

It be noted that accurately describing the global energy budget is problematic due to parameterization errors that arise from models with differing assumptions. For example, $\lambda E$ approximated by CERES in Figure 2.1 differs from Japanese and National Center for Environmental Prediction / National Center for Atmospheric Research Reanalysis outputs of 83.1 and 90.2 Wm$^{-2}$ respectively [61]. Atmospheric reanalysis models have been useful in describing uncharacteristic, rapid drying of land due to precipitation deficits and heat waves. For ex-
ample, Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) data has been used to characterize the number of flash droughts across CONUS from 1980-2017 [41]. That study showed that ET has a strong relationship to soil moisture during flash droughts in the central United States. Figure 2.2 [41] shows the relationship with modeled soil moisture and ET for a flash drought event in 2000 for a grid cell in Oklahoma. The vertical lines representing a period flash drought starting on day 217 in 2000 shows a drop in 20 soil moisture percentiles within 20 days, corresponding directly to a drop in ET for the same time period. This paper showed that physically, ET is related to soil moisture. This means that when there is a decrease in rainfall, there is a subsequent drop in ET over plant canopies due to roots not having access to available water.

2.2 Evapotranspiration Equations

In the hydrology space, there has been much progress in estimating point measurements of $\lambda E$ from various surfaces. For example, Penman (1948) and others described $\lambda E$ over open water using a modified energy budget equation [53]. Then, Monteith and others built on Penman’s model with an additional stomatal resistance variable to describe $\lambda E$ over land [42]. Eqs. 2.3 and 2.6 show the Penman and the modified Penman-Montieth respectively, where $\rho$ is the mean air density, $C_p$ is the specific heat at constant pressure, $e_s$ is the saturation vapor pressure, $e$ is the vapor pressure, $T_a$ is the atmospheric temperature, $r_h$ is
Figure 2.2: A) Climatology (thick blue line) for 38 years of root zone soil moisture versus time series for year 2000 (dotted blue line) and daily precipitation (red bars) [41].

b) Soil moisture percentile (orange line) with dashed, horizontal lines representing 20-40 drought category percentiles.

c) ET for year 2000 (thin blue line) versus climatology (thick blue line).
the aerodynamic resistance to heat transfer from the surface to the air, $\Delta$ is the saturation vapor pressure deficit and $\gamma$ is the psychrometric constant:

$$\lambda E = \frac{(R_n - G) + \rho C_p \cdot [e_s(T_a) - e]}{\Delta + \gamma^*},$$

(2.3)

where

$$\gamma^* = \frac{r_v}{r_h} \cdot \gamma$$

(2.4)

and $r_v$ is the aerodynamic resistance to water vapor transfer from the surface to the atmosphere.

Under the assumption that you have a big leaf canopy [42], $r_v$ can be separated into its component parts of stomatal resistance $r_s$, and aerodynamic resistance $r_h$:

$$\gamma^* = \left(\frac{r_s + r_h}{r_h}\right) \cdot \gamma = \left(1 + \frac{r_s}{r_h}\right) \cdot \gamma.$$  

(2.5)

Combining Eqs. 2.3 and 2.5 yields the Penman-Monteith, or:

$$\lambda E = \frac{(R_n - G) + \rho C_p \cdot [e_s(T_a) - e]}{\Delta + (1 + r_s/r_h) \cdot \gamma}.$$  

(2.6)

As shown in Eqs. 2.3 and 2.6, the difference between Penman and Penman-Monteith is the inclusion of the stomatal resistance variable in the denominator, which more accurately describes the physical relationship between plant canopies and the atmosphere. In 1972 the Penman-Monteith equation was modified for
saturated soils where the energy terms dominate over the physical variables [54].

As shown below in Eq. 2.7, the Priestly-Taylor equation simplifies the terms found in Eqs. 2.3 and 2.6 by using a derived constant $\alpha$ to describe aerodynamic resistance values:

$$\lambda E = \alpha \Delta \Delta \gamma \cdot (R_n - G).$$  \hfill (2.7)

It has been found that this constant varies with location on the Earth with an average of 1.26 [54], and a wide range of possibilities from 1 to 1.5 [31]. The advantage of Eq. 2.7 is that it requires less measured parameters, allowing for more wide-scale applications in remote sensing, such as satellite retrieval, where aerodynamic and vapor pressure deficit (VPD) terms are unknown to the sensor. However, the disadvantage can be less accurate results that come from abstracting these terms away into a single parameter.

2.3 ALEXI

ALEXI is a 5-10 km spatial resolution CONUS model of daily surface energy fluxes utilizing the 10-11 µm thermal infrared “window” channel data as input from GOES. ALEXI is based on a two-source energy balance partitioning the total surface radiometric temperature, $T_{RAD}$ into characteristic soil and canopy temperatures [46], $T_s$ and $T_c$, modified by the fractional cover of leaf area and vegetation clumping factor at viewing angle $f_\theta$:

$$T_{RAD} \approx f_\theta T_c + [1 - f_\theta] T_s.$$  \hfill (2.8)
Therefore, Eq. 2.8 is an aggregation of radiance values, which are related to component H fluxes at the surface and canopy. These H fluxes are directed into the atmosphere causing an aerodynamic temperature, $T_a$ which closes a system used to derive the height of the atmospheric boundary layer (ABL) [27]. As shown in Figure 2.3, total H flux depends on the temperature produced by the fraction of canopy cover, and the temperature of the surface viewed by the sensor viewing angle. The amount of H flux entering the atmosphere can be resisted over the surface, plant stomata, and canopy through various physical processes [27]. For example, lower wind speeds lead to higher atmospheric resistance to H flux and therefore a shallow boundary layer height. However, higher wind speeds decrease atmospheric resistance to H flux and deepen the height of the boundary layer.

ALEXI works by calculating an integral of H flux on two, separate model runs between time one, “t1” and time two, “t2” alongside a system of equations described by Anderson in 1997 [29] and Mecikalski in 1999 [46]. The ALEXI model considers the turbulent H flux driving the height (z) of ABL across time as shown on the right side of Figure 2.3. This “time-integrated” approach was developed to correct errors in sensor calibration, atmospheric corrections, and the specification of surface emissivity required for prior models using a single time step [46]. It has been approximated that the best time to run ALEXI is at 1.5 and 5.5 hours past local sunrise based on evidence of a strong correlation between changes in radiometric surface temperature measured by GOES and a total daytime H flux measured at a First International Satellite Land Surface Climatology Project Field Experiment (FIFE) site in Kansas in early July [29].
Figure 2.3: Component H from canopy and surface as related to $T_s$ and $T_c$, left and time integrated measurement of H at $t_1$ and $t_2$ and its influence on the slope of the boundary layer height relationship $\theta_s(z)$, right [27].
Alongside modeled H flux, a system of ALEXI equations provide outputs of λE for the canopy and the soil [27]. λE model outputs have been evaluated at FIFE sites with a Root Mean Square Difference of 53 Wm$^{-2}$ against measured data [46]. In 2004, ALEXI was evaluated at the disaggregated scale of 30 m against tower and surface based instruments within 10% error [26]. More recently, disaggregated ALEXI at 70 m resolution yielded correlations of 0.80 and Root Mean Square Error of 0.81 mm/day against 26 eddy covariance sites [32].

From a physical perspective, as the slope of H linearly increases between time-steps $t_1$ and $t_2$, the amount of λE entering the atmosphere must decrease, as energy is conserved [58] via the Bowen Ratio. This means that as H flux increases, there is a subsequent fall in λE. This causes plants to heat up via a loss of a cooling mechanism, especially if the surface moisture drops below the fraction of available water necessary for transpiration. Using ALEXI’s modeled λE over a modified Priestly Taylor equation yields ESI, which can be used to determine plant stress at a CONUS scale [27].

2.4 ESI

AET is a measure of λE flux into the atmosphere given the limited amount of available water on the surface. PET is the demand from the atmosphere, or maximum amount of energy that could move via λE flux provided non-limiting, available water. ESI is a measure of AET over PET describing moisture stress of the plant-soil system. Here, AET is the instantaneous λE flux modeled by the ALEXI model and PET is provided by calculating a modified Priestly Taylor
equation. PET and ESI are calculated for the soil and canopy components (subscript ‘c’ and ‘s’, respectively) as shown in Eqs. 2.9 - 2.13 below [27], where \( f_g \) is the fraction of green vegetation in the scene:

\[
\text{PET}_c = \frac{\alpha_c}{\lambda} f_g \frac{\Delta}{\Delta + \gamma} R_{nc} \quad (2.9)
\]

\[
\text{PET}_s = \frac{\alpha_s}{\lambda} \frac{\Delta}{\Delta + \gamma} R_{ns} \quad (2.10)
\]

\[
\text{ESI}_c = 1 - f_{PETc} = 1 - \frac{\text{AET}_c}{\text{PET}_c} \quad (2.11)
\]

\[
\text{ESI}_s = 1 - f_{PETs} = 1 - \frac{\text{AET}_s}{\text{PET}_s} \quad (2.12)
\]

\[
\text{ESI} = 1 - f_{PET} = 1 - \frac{E}{PET} = 1 - \frac{\text{AET}_c + \text{AET}_s}{\text{PET}_c + \text{PET}_s}. \quad (2.13)
\]

Physically, a number of 0 for ESI means that plants have ample moisture conditions and no stress. However, a number of 1 means instantaneous \( \lambda E \) has been cut off due to water limiting conditions [27]. Figure 2.4 [28] demonstrates how Anderson and others used this method to calculate a CONUS wide change in ESI with respect to a 28 day average, comparing it to a lower resolution model like the Palmer Index with good spatial correlation. This work showed that ESI could work on par with other drought indices without the need for antecedent precipitation measurements.
Figure 2.4: 28 day ESI composite for six months of the growing season, left and lower resolution Palmer Index, right in 2002 [28].
ESI is routinely calculated as a composite representing a weekly anomaly in comparison to a climatology. For example, Otkin and others in 2013 described how ESI, is calculated on 2, 4 and 8 week composite anomalies using clear day values subtracted by the climatology mean, and set over a standard deviation σ [50]. Eq. 2.14 defines ESI, where v(w, y) is equal to the clear sky composite for week “w” at year “y”, and the second term in the numerator defines the mean of all composites in the study climatology:

$$ESI_c = \frac{v(w, y) - \frac{1}{n_y} \sum_{y=1}^{n_y} v(w, y)}{\sigma_w}.$$  \hspace{1cm} (2.14)

Using Eq. 2.14, Otkin found that extreme flash drought events over eastern Indiana and western Ohio in 2007 were related to ESI as well as various meteorological conditions. As shown in Figure 2.5 [50], ESI tends to fall on 2, 4 and 8 week composites as a function of increasing temperature, decreasing precipitation and low cloud cover, with more severe cases showing higher winds and increasing dew-point depression [50]. Otkin also found that negative changes in ESI composite anomalies related to a decrease in range and pasture condition as reported by the National Agricultural Statistics Service, providing evidence that crop health can be monitored by ESI.

Figure 2.5 also shows how ESI was used as an “early-warning” of crop stress for this drought event, which is a basis for this study given the hypothesis in Chapter 1. As seen in Figure 2.5, at the beginning of May the rainfall increases, but plants are showing stress as depicted by the falling ESI values. This was not only due to sunny days and elevated wind speeds as shown in the top panel, but
also soils with higher runoff potentially causing a lack of groundwater recharge [21]. The figure also shows that USDM provides a slower response to flash drought detection, providing evidence of the potential for ESI to be used as an indicator of drought prior to its detection from other tools.

In regards to simulated soil moisture estimates, ESI has been evaluated against three North American Land Data Assimilation System (NLDAS) land surface models including Noah, Variable Infiltration Capacity, and Mosaic [49]. Results show that ESI is correlated to anomalies in modeled soil moisture and 2 m dew point depression. However, it should be noted that NLDAS models, while able to represent seasonal and inter-annual variability, often struggle with estimating soil moisture anomalies on smaller temporal scales. In particular, an analysis of 121 SCAN sites showed that NLDAS displayed large biases when compared to in situ observations due to model errors [68].

ESI has been used to derive fraction of available water ($f_{AW}$) measurements in the root zone and validated with real soil moisture measurements [39]. For example, the ALEXI model was used to predict values of $f_{AW}$ using four unique relationships (Noah Land Surface, linear, nonlinear and blend technique), and compared to values produced by Oklahoma Mesonet stations with “reasonable” results (less than 20 percent error). Mishra in 2013 [47] built on this research to force a Decision Support System for Agrotechnology Transfer model to output crop yield estimates using the ALEXI-modeled $f_{AW}$ for an area in North Alabama. Results showed that ALEXI-based crop yield estimates were similar to observed values.
Figure 2.5: Top figure: Temperature Anomaly (black), Dew-point Depression (blue), Wind Speed (red) and Cloud Cover Anomaly (Green). Bottom Figure: US Drought Monitor Category (black line), Precipitation (red), ESI (blue, green and yellow separated by intervals of two, four and eight weeks respectively).
2.5 SCAN

SCAN is a collection of 210 automated data collection sites across CONUS reporting various meteorological conditions to a central database via telemetry [11]. A typical SCAN site contains scientific measuring devices attached to tower, allowing variables like air temperature, relative humidity, solar radiation, wind speed and direction, liquid precipitation, barometric pressure and volumetric soil moisture, $\theta_v$, across several depths at specific locations. Figure 2.6 [1] demonstrates what a typical SCAN site looks like.

SCAN sites use a system of dielectric constant measuring devices for retrieval of $\theta_v$ across five depths as shown in Figure 2.7 [1]. This means they calculate volumetric soil moisture by using its relationship to the complex dielectric permittivity, $K^*$, or the ability of the soil to permit an electric field provided varying degrees of wetness. Changes in the $\theta_v$ are related to changes in $K^*$ using the equations provided by Topp (1980) [60], where the real dielectric permittivity, $\epsilon_r$, is measured and the imaginary dielectric permittivity, $\epsilon_i$, is calculated:

$$K^* = \epsilon_r - j\epsilon_i,$$

(2.15)

where

$$j = \sqrt{-1}$$

(2.16)

$$\epsilon_i = \epsilon_{rel} + \frac{\sigma_{dc}}{2 \cdot \pi f \epsilon_v}.$$  

(2.17)
Figure 2.6: A typical SCAN site [1].
and $\epsilon_{rel}$ is the Molecular relaxation, $\sigma_{dc}$ is the direct current electrical conductivity, and $\epsilon_v$ is the dielectric permittivity in a vacuum.

Topp (1980) used an empirical model to relate the actual dielectric permittivity, $K_a$ to $\theta_v$ as shown in Eq. 2.18 using least squares regression, assuming $K_a$ is roughly equal to the measured $\epsilon_r$:

$$\theta_v = -5.31 \times 10^{-2} + 2.92 \times 10^{-2}K_a - 5.5 \times 10^{-4}K_a^2 + 4.3 \times 10^{-6}. \quad (2.18)$$

This equation was found to be unaffected by soil bulk density, texture, and salinity and showed only a small standard error of 1.3% against real measurements. These errors could be calibrated for better estimation across soil types [62]. However, Seyfried in 2007 showed that using the dielectric equations can lead to upwards and downwards trends in estimations of $\theta_v$ due to temperature changes [56], al-
though typically within the specifications of the manufacturer’s accuracy of 1 to 5% [35].
Chapter 3. Data and Methods

3.1 Summary

Weekly ESI composites were obtained for each station pixel over all Alabama SCAN sites from 1 January 2001 to 31 December 2020. These composites were merged with available 1-week soil moisture anomalies across each station sensor depth. To ensure data quality, z-scores were calculated and anomaly values were removed that were 3.5 standard deviations from the monthly mean. Pedon and web soil survey reports were utilized to classify hydrologic soil groups at each station sensor depth. To classify land cover data, Multi-Resolution Land Consortium (MLRC) data for 8 eras within the 2001-2020 period were used. This data was averaged for the 20 year climatology and dominance was calculated using the Herfindahl-Hirschman Index (HHI). Lastly, ESI was compared to soil moisture conditions using Pearson R and accuracy assessments and analyzed across time, soil, and land cover to address science questions 1 and 2 in Chapter 1.

In order to classify soil moisture, percentile rank conversions at the 4-inch anomaly depth (4inANOM) were converted to Climate Prediction Center (CPC) drought categories. Then, the Soil Moisture Volatility Index (SMVI) was employed to identify a unique data set of flash droughts spanning 20 years. After, ESI composites and change anomalies were evaluated on their ability to predict
and monitor flash drought initiations across five specific time periods: the start of the event, 1 and 2 days prior, and 1 and 2 days after. To assess the behavior of ESI during flash drought events, ESI change anomalies were cumulatively summed and plotted across all flash drought events. To gain further insights, flash drought case studies were examined alongside cumulatively summed ESI change anomalies and USDM data to provide an answer to science question 3.

3.2 Data

3.2.1 ESI

ESI is generated for the global land mass at 4 and 12 week composites to reduce noise in daily retrievals [38] and can be downloaded as tif files from the NASA SERVIR global web portal [4]. However, for this study 1 week ESI composites were provided by Dr. Chris Hain at NASA SPoRT upon request in order to compare to soil moisture profiles at week timescales. ESI was obtained for latitudes and longitudes representing 18 in situ sites across Alabama as shown in Figure 3.1 [9] using a python script [63]. The ESI data used in this study represents a 20 year climatology from 1 January 2001 to 31 December 2020. ESI values representing no data were dropped from the data set. While ESI is affected by cloud cover, it is corrected by NASA SPoRT prior to data uploads using a gap filling technique described by Anderson and others in 2007 [27].
Figure 3.1: All SCAN Sites Across Alabama [9].
ESI change anomalies were calculated for each SCAN station climatology using a method created by Otkin et al. (2013) [50]. First, ESI composites (ESI\(_c\)) were computed over 1, 2, 3, and 4 week periods. For example, ESI\(_2\) would represent the 2 week ESI composite using Eq. 2.14. Then, the change anomaly (ΔESI\(_c\)(w)) for 2, 4, and 8 week (w) change intervals are calculated using Eq. 3.1, where v(w, y) represents the clear day ESI\(_c\) for week and year in the station climatology, and the second term in the numerator represents the mean difference for all years:

\[
\Delta\text{ESI}_c(w_1, w_2, y) = \frac{(v(w_2, y) - v(w_1, y)) - \frac{1}{ny} \sum_{y=1}^{ny} v(w_2, y) - v(w_1, y)}{\sigma(w_1, w_2)}. \tag{3.1}
\]

In the previous 2 week example, if the difference interval used between week 1 (w\(_1\)) and week 2 (w\(_2\)) were a 4 week period, this would yield ΔESI\(_2\)(4), representing a 2 week ESI composite with a 4 week difference between each data point. Change anomalies calculated by eq. 3.1 represent a smoother understanding of the relative to climatology change in vegetative stress over time.

### 3.2.2 Volumetric Soil Moisture

Daily averages of Volumetric Soil Moisture θ\(_v\) were pulled from the National Water and Climate Center (NWCC) report generator [9] for each SCAN site in Alabama for all available depths and dates in the ESI climatology period. These values were then averaged to 1 week means using a simple running aver-
age. Values were returned only when there was at least a 3-day window period with data for each week. A Z-score representing the number of standard deviations above or below monthly station means were calculated for each soil moisture depth using Eq. 3.2:

\[ Z = \frac{\theta_{td} - \bar{\theta}_{vmd}}{\sigma_{vd}}. \]  

(3.2)

Here, \( \theta_{td} \) represents weekly \( \theta_v \) for timestamp \( t \) at depth \( d \), \( \bar{\theta}_{vmd} \) represents the mean \( \theta_v \) for a particular month and depth in all station years, and the denominator represents the standard deviation of all \( \theta_{vd} \) values in the station climatology. Values that were above or below 3.5 were excluded based on the assumption that they were an outlier and not representative of a station’s monthly climatology. A total of 14 values in the 4inANOM sensor across all station period of records met this criteria and were therefore discarded.

Weekly soil moisture anomaly values at sensor depth \( d \) were calculated for all available depths at each station using Eq. 3.3:

\[ d\text{ANOM} = \frac{\theta_{td} - \bar{\theta}_{vwd}}{\sigma_{vd}}. \]  

(3.3)

In this case, \( d\text{ANOM} \) represents a dry (less than zero), or wet (greater than zero) soil moisture anomaly value as compared to the same week across the station climatology. Here \( \theta_{td} \) represents \( \theta_v \) at timestamp \( t \) and sensor depth \( d \), \( \bar{\theta}_{vwd} \) represents the mean \( \theta_{vd} \) for a particular week in all station years, and the denominator represents the standard deviation of all \( \theta_{vd} \) values in the station climatology.
Note the difference between eqs. 3.2 and 3.3 is the second term in the numerator. This means that first, Z-scores were applied to test for values that were outside of monthly climatologies. After values above and below 3.5 standard deviations were thrown out, a more general Z-score representing weekly climatologies was selected for better comparison to the ESI data set. A more relaxed monthly Z-score was applied at first because of the limited number of weekly samples that exist and the concern that sharper flash drought events might be excluded on accident.

3.2.3 Drought Data

The USDM is created by a consortium of federal agencies (National Drought Mitigation Center, USDA, and NOAA) and public partners like the University of Nebraska Lincoln. Drought experts look at a variety of data sets considering hydrologic variables such as streamflow, precipitation, soil moisture, snow cover and meltwater runoff [17]. Experts work together to publish a CONUS scale map by dividing land area into five drought classifications: abnormally dry (D0), moderate (D1), severe (D2), extreme (D3), and exceptional (D4) drought [2]. These data have been released every Thursday for over 23 years beginning on 4 January 2000 and are provided as shape files for download by year [16]. USDM Data was extracted at point locations for each nearest week for the 20 year ESI climatology using a python tool [64]. For comparison, drought classification for each station climatology were calculated by ranking dANOM values into percentiles and applying CPC soil moisture definitions [2].
Flash droughts were identified for each station using the four inch soil moisture sensor with a modified SMVI from Osman and others in 2021 [48]. To get SMVI, the running 1 week dANOM is subtracted by the three week dANOM for each time step in the climatology. A flash drought is said to occur when 1) the running 1 week mean falls below the 3 week mean for a period of 3 weeks (21 days) and 2) descends to the 20th percentile (D1) drought category before a recovery of at least 2 weeks. Recovery is defined as positive increases in the 1 week mean (at or above the 3 week mean) for a period of 2 weeks. This is because the assumption is that a 1 week recovery period is not enough time for vegetation to respond from a lack of available soil moisture. If the one week mean increases positively above the three week mean for a period of 1 week and then relapses below it again, it is assumed that the flash drought has not recovered entirely. Therefore, 2 consecutive weeks of above zero SMVI are required for recovery. Figure 3.2 is an example of how a flash drought event can be identified using SMVI for station 2174 in West Alabama.
Figure 3.2: 2016 flash drought event identified at SCAN station 2174 (highlighted in red) by negative SMVI (blue) values over 21 days. 4inANOM (green) descends to a D3 USDM category based on CPC soil moisture definitions before recovery in November.
3.2.4 Soil Reports

Soil hydrologic groups can be used to describe the runoff potential of surfaces during precipitation events based on soil groupings. Table 3.1 [57] describes these soil groups based on their ability to runoff. Group A soils have a greater capacity to infiltrate, meaning less runoff while group D has a higher runoff potential with lower infiltration rates. Soils are placed in hydrologic groups by their percentage of silt, clay and sand as shown in Figure 3.3 [66]. Typically, a SCAN site on the USDA web portal [10] will contain a link to a soil pedon report by clicking on “station metadata.” Pedon reports are text files that contain site specific descriptive data about soils across layers, such as mineral or chemical values and soil taxonomy classifications [45]. These pedon reports can be used to identify lab textures of soils at each soil moisture sensor depth. This study was able to generate hydrologic soil groupings at 9 of the 18 SCAN sites using pedon reports. If a SCAN site did not have a pedon report, the typical profiles from USDA’s web soil survey [18] were extracted at the location of the SCAN site and were used to generate a hydrologic soil grouping for sensor depth instead.
Table 3.1: Soil Hydrologic Groups and Runoff Potential [57].

<table>
<thead>
<tr>
<th>Soil Group</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td><em>Lowest Runoff Potential</em>. Includes deep sands with very little silt and clay, also deep, rapidly permeable loess.</td>
</tr>
<tr>
<td>B</td>
<td><em>Moderately Low Runoff Potential</em>. Mostly sandy soils less deep than A, loess less deep or less aggregated than A, but the group as a whole has above-average infiltration after thorough wetting.</td>
</tr>
<tr>
<td>C</td>
<td><em>Moderately High Runoff Potential</em>. Comprises shallow soils and soils containing considerable clay and colloids, though less than those of group D. The group has below-average infiltration after pre-saturation.</td>
</tr>
<tr>
<td>D</td>
<td><em>Highest Runoff Potential</em>. Includes mostly clays of high swelling percent, but the group also includes some shallow soils with nearly impermeable sub-horizons near the surface.</td>
</tr>
</tbody>
</table>
Figure 3.3: Soil Hydrologic Groups by Soil Type [66].
3.2.5 Land Cover Data

MLRC is a group of 10 federal agencies that work closely together to maintain a CONUS scale land cover database for a variety of modeling, environment and management applications [67]. The MLRC National Land Cover Database (NLCD) Viewer tool [5] can be used to subset 30 m resolution CONUS images of land cover using an area of interest box. The MLRC NLCD Viewer sends a .zip file containing eras of land cover data to a valid email making it easy to analyze smaller images at convenient scales. Land Cover data generated by MLRC exists across 8 eras (2001, 2004, 2006, 2008, 2011, 2013, 2016, 2019). For this study, an AOI was drawn over the state of Alabama and 5km subsets of ESI shape files [34] representing each SCAN station were reclassified into 8 land cover classes - Wetland, Water, Agricultural, Forest, Barren, Herbaceous, Developed, and Shrub via a python tool [24]. Figure 3.4 shows an example of the land cover diversity that exists at station 2115 in East Alabama.
Figure 3.4: NLCD Land Cover Classes by Type at Station 2115 for the ESI Pixel Domain.
Land cover dominance for each subset was calculated using HHI, an economics statistic generally used by federal regulators to understand market concentration [55]. In this case, increased values of HHI indicate land cover dominance of a particular class, whereas decreased values indicate increased class diversity over a subset. HHI can be calculated for a subset using Eq. 3.4 where \( LCC_i \) represents the land cover class in subset \( i \) for \( n \) total classes:

\[
HHI = \sum_{i}^{n} LCC_i^2. \tag{3.4}
\]

3.3 Methods

3.3.1 Statistical Assessment

The Pearson R correlation coefficient is used to express the linear relationship between two variables \( x \) and \( y \) [59]. The Pearson R is calculated using Eq. 3.5 [12], where \( i \) is an iteration of \( x \) or \( y \) in a population, \( \bar{x} \) and \( \bar{y} \) are their respective means, and \( n \) represents the total samples:

\[
\text{Pearson R} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{3.5}
\]

Table 3.2 [25] is a chart suggesting commonly used Pearson R correlation coefficient interpretations across different research fields. For this study, the Dancey and Reidy interpretation will be used as a means to standardize the language of interpretation for chapter 4.
Table 3.2: Correlation Coefficient Strength Interpretation.

<table>
<thead>
<tr>
<th>Correlation Coefficient</th>
<th>Dancey &amp; Reidy (Psychology)</th>
<th>Quinnipiac University (Politics)</th>
<th>Chan YH (Medicine)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+1</td>
<td>Perfect</td>
<td>Perfect</td>
<td>Perfect</td>
</tr>
<tr>
<td>0.9</td>
<td>Strong</td>
<td>Very Strong</td>
<td>Very Strong</td>
</tr>
<tr>
<td>0.8</td>
<td>Strong</td>
<td>Very Strong</td>
<td>Very Strong</td>
</tr>
<tr>
<td>0.7</td>
<td>Strong</td>
<td>Very Strong</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.6</td>
<td>Moderate</td>
<td>Strong</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.5</td>
<td>Moderate</td>
<td>Strong</td>
<td>Fair</td>
</tr>
<tr>
<td>0.4</td>
<td>Moderate</td>
<td>Strong</td>
<td>Fair</td>
</tr>
<tr>
<td>0.3</td>
<td>Weak</td>
<td>Moderate</td>
<td>Fair</td>
</tr>
<tr>
<td>0.2</td>
<td>Weak</td>
<td>Weak</td>
<td>Poor</td>
</tr>
<tr>
<td>0.1</td>
<td>Weak</td>
<td>Negligible</td>
<td>Poor</td>
</tr>
<tr>
<td>0</td>
<td>Zero</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

P-values test if the correlation coefficient calculated between variables can be used to model the relationship across a population of data [44]. P-values can be looked up on a t-table using a calculated test statistic, t. The t calculation is shown in Eq. 3.6, where the numerator here assumes 2 degrees of freedom for n points being compared, and uses the calculated Pearson R (r) correlation coefficient:

\[ t = \frac{r \sqrt{n - 2}}{\sqrt{1 - r^2}} \]  

(3.6)

P-values are compared against a user defined alpha value representing an accepted significance level. If the p-value is less than the alpha, the null hypothesis which represents no significant relationship between variables across a population can be rejected. If this is the case, it could be said that two variables have a “significant” linear relationship.

This study used the Pearson R to compare one week ESI and dANOM across depths, stations, months, soil type and dominant land cover. Correlations were compared to by month and land cover dominance to identify linear relationships. P-values were calculated for monthly SCAN station correlations and were compared across spatial tests of significance by dominant land cover.

Time lagged Pearson R correlations between dANOM and ESI were compared using a week lag interval (WLI) methodology. A WLI describes the number of days the dANOM ending timestamp is prior to or ahead of the ESI timestamp. For example, if the ESI date is 14 January, this implies that ESI was computed for
the 7 days prior to the 14 January timestamp, or from 7 January until 14 January. If the WLI is -1, this implies that the rolling mean calculated for dANOM was computed between January 6th and January 13th and compared to the January 14th ESI timestamp.

To look for linear relationships between ESI and dANOM, Pearson R values were compared across +/- 30 WLI. These correlations were analyzed to identify synchronized time correlations across case stations. WLI correlations were also compared across hydrologic soil, depth and season months to look for relationships explained by physical characteristics.

### 3.3.2 Accuracy Assessment

ESI was compared to two classes of predicted dANOM conditions and an accuracy methodology described by Fawcett (2006) [36] was applied. In this case, ESI values above and below zero are mapped to predictions of dANOM classes “wet” and “dry.” Here, a “wet” dANOM value is above zero and a “dry” dANOM is below.

Accuracy is defined by Eq. 3.7 where TP is true positive, TN is true negative, FP is false positive, and FN is false negative:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.
\]  

(3.7)

Accuracy was compared by depths, stations, months, hydrologic soil grouping and dominant land cover. Accuracy was also binned into 12 ESI categories and
compared to analyze how ESI performs at predicting dANOM classes across a broad range of wet and dry conditions.

3.3.3 Flash Drought Assessment

A soil moisture based flash drought data set was created by calculating SMVI across each station dANOM climatology. Identified flash drought events were analyzed statistically by drought duration, USDM drought categories, and four inch percentile delta. Counts of flash drought events from 2002-2020 were plotted by land cover and hydrologic soil group. Events occurring during the growing season (months 5-11) were grouped and compared to look for physical relationships.

ESI change anomalies at 1, 2, 4, and 8 week composites along 1, 2, 3, and 4 week differences were calculated and compared to dANOM classes of dry and wet using accuracy at 5 time indices: the start, 1 and 2 days prior, and 1 and 2 days after the start of each flash drought event. These anomalies were also cumulatively summed across all flash drought events and compared to drought duration. Case studies using ESI change anomalies for measuring flash drought events at SCAN stations were tested visually with plots.
Chapter 4. Results

4.1 Data Statistics

Table 4.1 provides a set of descriptive statistics for the 1 Week ESI composites and soil moisture data used for analysis. There were a total of 4305 missing soil moisture values spread out across 5 depths. Total missing week values for each depth can be found by subtracting the ESI count by the depth count. For example, 2inANOM had the least missing values because 9674 - 9147 is 504, while the most were found at 40inANOM because 9674 - 8193 is 1481. Missing values occur due to sensor issues that go unresolved for each sensor depth over a period of time, as well as week means that violated the 3-day window period when calculated as described in section 3.1.2. Table 4.2 provides information on observation weeks for each station's observation record (minimum date - maximum date). The longest observation record is 20 years at station 2057, with the number of usable observation weeks being 809 (15 years). The shortest is located at 2175 spanning 10 years of record and contained 317 usable weeks (6 years).
Table 4.1: Description of Statistics for ESI Data alongside Soil Moisture Observations for Each Depth.

<table>
<thead>
<tr>
<th>ESI</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>18inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>9674.00</td>
<td>9147.00</td>
<td>8762.00</td>
<td>9031.00</td>
<td>8888.00</td>
<td>8193.00</td>
</tr>
<tr>
<td>mean</td>
<td>0.02</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>std</td>
<td>1.01</td>
<td>0.95</td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>min</td>
<td>-3.50</td>
<td>-3.02</td>
<td>-3.14</td>
<td>-3.59</td>
<td>-3.07</td>
<td>-3.59</td>
</tr>
<tr>
<td>25%</td>
<td>-0.63</td>
<td>-0.71</td>
<td>-0.70</td>
<td>-0.67</td>
<td>-0.65</td>
<td>-0.72</td>
</tr>
<tr>
<td>50%</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.11</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>75%</td>
<td>0.70</td>
<td>0.73</td>
<td>0.73</td>
<td>0.69</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>max</td>
<td>3.35</td>
<td>3.07</td>
<td>2.03</td>
<td>3.40</td>
<td>2.93</td>
<td>2.74</td>
</tr>
</tbody>
</table>

Table 4.2: Description of Station Period Of Record and Number of Week Observations.

<table>
<thead>
<tr>
<th>Station</th>
<th>Minimum Date</th>
<th>Maximum Date</th>
<th>Observation Count (Weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2175:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-12-30</td>
<td>317</td>
</tr>
<tr>
<td>2180:AL:SCAN</td>
<td>2010-05-14</td>
<td>2020-12-30</td>
<td>351</td>
</tr>
<tr>
<td>2181:AL:SCAN</td>
<td>2011-04-30</td>
<td>2020-12-30</td>
<td>361</td>
</tr>
<tr>
<td>2177:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-12-30</td>
<td>392</td>
</tr>
<tr>
<td>2176:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-12-30</td>
<td>403</td>
</tr>
<tr>
<td>2182:AL:SCAN</td>
<td>2010-08-14</td>
<td>2020-12-30</td>
<td>408</td>
</tr>
<tr>
<td>2178:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-12-30</td>
<td>448</td>
</tr>
<tr>
<td>2179:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-12-30</td>
<td>454</td>
</tr>
<tr>
<td>2173:AL:SCAN</td>
<td>2010-04-09</td>
<td>2020-12-30</td>
<td>450</td>
</tr>
<tr>
<td>2174:AL:SCAN</td>
<td>2010-04-16</td>
<td>2020-09-30</td>
<td>467</td>
</tr>
<tr>
<td>2055:AL:SCAN</td>
<td>2002-04-23</td>
<td>2020-12-30</td>
<td>535</td>
</tr>
<tr>
<td>2114:AL:SCAN</td>
<td>2006-05-21</td>
<td>2019-10-08</td>
<td>644</td>
</tr>
<tr>
<td>2115:AL:SCAN</td>
<td>2006-05-21</td>
<td>2020-12-30</td>
<td>660</td>
</tr>
<tr>
<td>2113:AL:SCAN</td>
<td>2006-05-28</td>
<td>2020-12-30</td>
<td>686</td>
</tr>
<tr>
<td>2053:AL:SCAN</td>
<td>2002-02-12</td>
<td>2020-12-30</td>
<td>766</td>
</tr>
<tr>
<td>2056:AL:SCAN</td>
<td>2002-04-23</td>
<td>2020-12-30</td>
<td>809</td>
</tr>
<tr>
<td>2057:AL:SCAN</td>
<td>2002-04-30</td>
<td>2020-12-30</td>
<td>809</td>
</tr>
</tbody>
</table>
4.2 Statistical Results

Table 4.3 provides Pearson R correlation by month and depth for all stations combined. There are correlation maximums in the early and later months (4-5 and 10-11) across each depth. Correlation minimums tend to occur in the winter months (12-2) and there exists a trough between the monthly maximums (5-8) where correlation tends to fall. The overall descriptive statistics are found in Table 4.4. The best correlation is found in the 4inANOM depth at $\sim 0.53$ in month 11 while the worst is found at 40inANOM at month 1 at $\sim 0$. The 2inANOM shows the greatest mean value while the 8inANOM shows the tightest standard deviation. Therefore, depending on the statistic and the objective of analysis, different depths might be used to better understand the relationship between soil moisture and ESI. It should be noted as well that this analysis becomes more complex over a wide range of soil and land cover characteristics as will be shown in this chapter.
Table 4.3: All Stations Pearson R Correlation by Month and Depth.

<table>
<thead>
<tr>
<th>Month</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.147845</td>
<td>0.080224</td>
<td>0.036096</td>
<td>0.005102</td>
<td>-0.001079</td>
</tr>
<tr>
<td>2</td>
<td>0.155177</td>
<td>0.180764</td>
<td>0.188772</td>
<td>0.087306</td>
<td>0.061009</td>
</tr>
<tr>
<td>3</td>
<td>0.209725</td>
<td>0.207795</td>
<td>0.226498</td>
<td>0.117040</td>
<td>0.036514</td>
</tr>
<tr>
<td>4</td>
<td>0.261975</td>
<td>0.255022</td>
<td>0.246198</td>
<td>0.222607</td>
<td>0.130911</td>
</tr>
<tr>
<td>5</td>
<td>0.279318</td>
<td>0.258146</td>
<td>0.240307</td>
<td>0.200247</td>
<td>0.150698</td>
</tr>
<tr>
<td>6</td>
<td>0.221373</td>
<td>0.200369</td>
<td>0.167314</td>
<td>0.212439</td>
<td>0.105352</td>
</tr>
<tr>
<td>7</td>
<td>0.227329</td>
<td>0.181525</td>
<td>0.164201</td>
<td>0.113307</td>
<td>0.195919</td>
</tr>
<tr>
<td>8</td>
<td>0.273758</td>
<td>0.200589</td>
<td>0.213268</td>
<td>0.210317</td>
<td>0.177186</td>
</tr>
<tr>
<td>9</td>
<td>0.413648</td>
<td>0.393461</td>
<td>0.381183</td>
<td>0.358760</td>
<td>0.330872</td>
</tr>
<tr>
<td>10</td>
<td>0.459643</td>
<td>0.458771</td>
<td>0.437033</td>
<td>0.413700</td>
<td>0.381078</td>
</tr>
<tr>
<td>11</td>
<td>0.488489</td>
<td>0.531818</td>
<td>0.427634</td>
<td>0.414055</td>
<td>0.347435</td>
</tr>
<tr>
<td>12</td>
<td>0.072597</td>
<td>0.126772</td>
<td>0.107815</td>
<td>0.123432</td>
<td>0.132502</td>
</tr>
</tbody>
</table>

Table 4.4: All Stations Pearson R Correlation by Month and Depth Descriptive Statistics.

<table>
<thead>
<tr>
<th></th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.267573</td>
<td>0.256271</td>
<td>0.236360</td>
<td>0.206526</td>
<td>0.170700</td>
</tr>
<tr>
<td>std</td>
<td>0.127680</td>
<td>0.135976</td>
<td>0.123377</td>
<td>0.130697</td>
<td>0.123982</td>
</tr>
<tr>
<td>min</td>
<td>0.072597</td>
<td>0.080224</td>
<td>0.036096</td>
<td>0.005102</td>
<td>-0.001079</td>
</tr>
<tr>
<td>25%</td>
<td>0.196088</td>
<td>0.181334</td>
<td>0.166535</td>
<td>0.116107</td>
<td>0.094266</td>
</tr>
<tr>
<td>50%</td>
<td>0.244652</td>
<td>0.204192</td>
<td>0.219883</td>
<td>0.205282</td>
<td>0.141600</td>
</tr>
<tr>
<td>75%</td>
<td>0.312900</td>
<td>0.291975</td>
<td>0.279944</td>
<td>0.256645</td>
<td>0.229657</td>
</tr>
<tr>
<td>max</td>
<td>0.488489</td>
<td>0.531818</td>
<td>0.437033</td>
<td>0.414055</td>
<td>0.381078</td>
</tr>
</tbody>
</table>
Figure 4.1: All Stations Pearson R Values by Month and Depth.
Figure 4.1 is a plot of Table 4.3 showing the variation in correlation by depth over month. Visually, there is a weak (refer to the Dancy & Reidy classification scheme in Table 3.2), maximum peak at month 5 for the 2inANOM depth, whereas month 11 shows a moderate 4inANOM value. This suggests that some soil moisture depths may be more linearly related with ESI depending on the month that values are compared. Across months, sensors higher in the soil profile show a stronger correlation while deeper sensors correlate less. However, there are some examples of this relationship flipping. For example, in month 7, the 40inANOM climbs from 0.1 to 0.2 in correlation and overtakes the three sensor depths above it. Evidence described in the following paragraphs suggest this could be explained by land cover dominance. Although, one explanation is that ESI is best sensing the drought signature in months 9-11 where soils are typically the driest. Therefore, cloud cover affects the ESI soil moisture relationship as winter rains begin causing the signal to decrease significantly.
Figure 4.2: Pearson R Values for 2053:AL:SCAN by Month and Depth (bottom panel) and Calculated Dominance Values of 8 Land Cover Classes using HHI (top panel).
Figure 4.3: Pearson R Values for 2179:AL:SCAN by Month and Depth (bottom panel) and Calculated Dominance Values of 8 Land Cover Classes using HHI (top panel).
Table 4.5: 2053:AL:SCAN (Strongly Agricultural) Pearson R Correlation by Month and Depth.

<table>
<thead>
<tr>
<th>Month</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.220730</td>
<td>0.153376</td>
<td>0.096305</td>
<td>0.083786</td>
<td>0.136976</td>
</tr>
<tr>
<td>2</td>
<td>0.205409</td>
<td>0.306906</td>
<td>0.285971</td>
<td>0.206550</td>
<td>0.236205</td>
</tr>
<tr>
<td>3</td>
<td>0.145241</td>
<td>0.141899</td>
<td>0.006111</td>
<td>0.036445</td>
<td>-0.056430</td>
</tr>
<tr>
<td>4</td>
<td>0.245073</td>
<td>0.294667</td>
<td>0.255980</td>
<td>0.095737</td>
<td>0.291462</td>
</tr>
<tr>
<td>5</td>
<td>0.344874</td>
<td>0.581866</td>
<td>0.410980</td>
<td>0.285632</td>
<td>0.333610</td>
</tr>
<tr>
<td>6</td>
<td>0.227440</td>
<td>0.388689</td>
<td>0.305494</td>
<td>0.368293</td>
<td>0.373114</td>
</tr>
<tr>
<td>7</td>
<td>-0.149352</td>
<td>-0.203115</td>
<td>-0.070477</td>
<td>-0.316073</td>
<td>-0.276762</td>
</tr>
<tr>
<td>8</td>
<td>0.315568</td>
<td>0.339502</td>
<td>0.430092</td>
<td>0.417665</td>
<td>0.243112</td>
</tr>
<tr>
<td>9</td>
<td>0.383856</td>
<td>0.486753</td>
<td>0.455235</td>
<td>0.156236</td>
<td>0.538892</td>
</tr>
<tr>
<td>10</td>
<td>0.337375</td>
<td>0.596977</td>
<td>0.516758</td>
<td>0.040720</td>
<td>0.457415</td>
</tr>
<tr>
<td>11</td>
<td>0.479186</td>
<td>0.541153</td>
<td>0.303542</td>
<td>0.089044</td>
<td>0.491051</td>
</tr>
<tr>
<td>12</td>
<td>0.077721</td>
<td>0.254425</td>
<td>0.273906</td>
<td>0.205092</td>
<td>0.189511</td>
</tr>
</tbody>
</table>

52
Table 4.6: 2179:AL:SCAN (Strongly Forest) Pearson R Correlation by Month and Depth.

<table>
<thead>
<tr>
<th>Month</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.092914</td>
<td>-0.022852</td>
<td>-0.094730</td>
<td>-0.032361</td>
<td>0.048641</td>
</tr>
<tr>
<td>2</td>
<td>0.032365</td>
<td>-0.042931</td>
<td>0.045711</td>
<td>0.158157</td>
<td>0.138681</td>
</tr>
<tr>
<td>3</td>
<td>0.183272</td>
<td>0.151261</td>
<td>0.081451</td>
<td>0.30265</td>
<td>0.277754</td>
</tr>
<tr>
<td>4</td>
<td>0.173507</td>
<td>0.049455</td>
<td>-0.030735</td>
<td>0.262245</td>
<td>0.383376</td>
</tr>
<tr>
<td>5</td>
<td>0.006028</td>
<td>0.051424</td>
<td>-0.122580</td>
<td>-0.120609</td>
<td>-0.129740</td>
</tr>
<tr>
<td>6</td>
<td>0.128667</td>
<td>0.323582</td>
<td>0.180714</td>
<td>0.148425</td>
<td>0.137683</td>
</tr>
<tr>
<td>7</td>
<td>0.107030</td>
<td>0.157903</td>
<td>0.075303</td>
<td>0.101946</td>
<td>0.198589</td>
</tr>
<tr>
<td>8</td>
<td>0.543660</td>
<td>0.507735</td>
<td>0.455159</td>
<td>0.375613</td>
<td>0.153655</td>
</tr>
<tr>
<td>9</td>
<td>0.467351</td>
<td>0.553946</td>
<td>0.399381</td>
<td>0.608459</td>
<td>0.570543</td>
</tr>
<tr>
<td>10</td>
<td>0.489012</td>
<td>0.531992</td>
<td>0.500005</td>
<td>0.090006</td>
<td>0.060060</td>
</tr>
<tr>
<td>11</td>
<td>0.583575</td>
<td>0.667777</td>
<td>0.593613</td>
<td>0.64367</td>
<td>0.606465</td>
</tr>
<tr>
<td>12</td>
<td>0.121131</td>
<td>0.263426</td>
<td>0.254636</td>
<td>0.182740</td>
<td>-0.089405</td>
</tr>
</tbody>
</table>
Figures 4.2 and 4.3 provide plots of 8 land cover class percentages (top panels) as well as the monthly station correlations (bottom panels) at station 2053 and 2179 by depth. Note that 2053 shows a strongly agricultural dominated pixel, with a mean 20 year coverage of 79%, whereas 2179 provides a similar plot of station 2179, containing a mean 20 year forest coverage of 68%. The monthly correlation values for each station is found in tables 4.5 and 4.6. Note the strongly negative correlation across depths in month 7 for station 2053 and a more positive case at station 2179.

The correlation time series for 2053 and 2179 show distinct maximums, minimums and deficits. For example, the 2053 pixel has maximum peaks in month 5 and 10, showing a moderate 4inANOM in both places at \( \sim 0.58 \) and \( \sim 0.59 \) respectively. All sensors fall below zero correlation in month 7, with the 20inANOM showing a weak correlation maximum in the negative direction at \( \sim -0.31 \), suggesting an inverse relationship during this time. For station 2179, the deeper sensors (20 and 40inANOM) seem to peak at month 4 and month 10 at \( \sim 0.26 \) and \( \sim 0.38 \). Notice that for the forest example, the deeper sensors play a stronger role by overtaking the three sensor depths above them during the maximum correlation event at month 10.

Descriptive statistics for stations 2053 and 2179 can be found in Tables 4.7 and 4.8, revealing greater complexity in the relationship between soil moisture and ESI. The strongly forested case exhibits higher maximum values overall at every depth compared to the agricultural case. Despite this, the mean values are generally higher for the agricultural area, with the exception of the 20inANOM
**Table 4.7:** 2053:AL:SCAN (Strongly Agricultural) Pearson R Correlation by Month and Depth Descriptive Statistics.

<table>
<thead>
<tr>
<th></th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>count</strong></td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td>0.236101</td>
<td>0.323591</td>
<td>0.272491</td>
<td>0.139103</td>
<td>0.250680</td>
</tr>
<tr>
<td><strong>std</strong></td>
<td>0.163086</td>
<td>0.226674</td>
<td>0.180710</td>
<td>0.189826</td>
<td>0.233771</td>
</tr>
<tr>
<td><strong>min</strong></td>
<td>-0.149352</td>
<td>-0.203115</td>
<td>-0.070477</td>
<td>-0.316973</td>
<td>-0.275762</td>
</tr>
<tr>
<td><strong>25%</strong></td>
<td>0.190435</td>
<td>0.229163</td>
<td>0.216062</td>
<td>0.073020</td>
<td>0.176377</td>
</tr>
<tr>
<td><strong>50%</strong></td>
<td>0.236257</td>
<td>0.323204</td>
<td>0.294756</td>
<td>0.126036</td>
<td>0.283834</td>
</tr>
<tr>
<td><strong>75%</strong></td>
<td>0.339250</td>
<td>0.500353</td>
<td>0.415758</td>
<td>0.226327</td>
<td>0.396689</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>0.479186</td>
<td>0.596977</td>
<td>0.516758</td>
<td>0.417665</td>
<td>0.538892</td>
</tr>
</tbody>
</table>

**Table 4.8:** 2179:AL:SCAN (Strongly Forest) Pearson R Correlation by Month and Depth Descriptive Statistics.

<table>
<thead>
<tr>
<th></th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>count</strong></td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
<td>12.000000</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td>0.228557</td>
<td>0.256053</td>
<td>0.194911</td>
<td>0.277218</td>
<td>0.241564</td>
</tr>
<tr>
<td><strong>std</strong></td>
<td>0.230199</td>
<td>0.247445</td>
<td>0.243105</td>
<td>0.261825</td>
<td>0.253824</td>
</tr>
<tr>
<td><strong>min</strong></td>
<td>-0.092914</td>
<td>-0.042981</td>
<td>-0.122580</td>
<td>-0.120609</td>
<td>-0.129740</td>
</tr>
<tr>
<td><strong>25%</strong></td>
<td>0.088364</td>
<td>0.050931</td>
<td>0.027350</td>
<td>0.136805</td>
<td>0.115422</td>
</tr>
<tr>
<td><strong>50%</strong></td>
<td>0.151087</td>
<td>0.210664</td>
<td>0.131082</td>
<td>0.222493</td>
<td>0.176122</td>
</tr>
<tr>
<td><strong>75%</strong></td>
<td>0.472766</td>
<td>0.513792</td>
<td>0.413325</td>
<td>0.433825</td>
<td>0.430168</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>0.583575</td>
<td>0.657777</td>
<td>0.593613</td>
<td>0.696066</td>
<td>0.608485</td>
</tr>
</tbody>
</table>
Figure 4.4: All Stations Average Pearson R Values by Month and Depth for Agricultural Land Cover (Agland) Type.

sensor. This observation may suggest that soil moisture in strongly forested areas has a better overall correlation with ESI. However, correlation patterns between the two cases differ significantly and show wide variation across depths. For example, the standard deviation in the forest case is higher across every depth than the agricultural station. This might suggest that correlation is more variable to the overall mean in forest cases when compared to agricultural examples.

Figures 4.4 and 4.5 display the average Pearson R values for all stations by month and land cover dominance. These plots highlight visual patterns along
Figure 4.5: All Stations Average Pearson R Values by Month and Depth for Forest Land Cover Type.
troughing months (5-8) where correlations typically decrease across depth. For example, the 2 and 4inANOM sensors exhibit a lower average correlation in agricultural pixels during month 7, while the 40inANOM sensor increases and overtakes the 4 and 20inANOM sensor values above it. During these months, the 2 and 4inANOM sensors show shallower declines in forest-covered pixels and the 40inANOM sensor value increases significantly. Figure 4.6 provides a specific plot of the average 4inANOM correlation for all stations by month and land cover dominance. Notice the visual decline from a moderately positive correlation in month 7 for agricultural fields while the correlation is maintained for the forest type. These results could be explained by the idea that rooting depths typical to agricultural fields may be less representative of soil moisture profiles in month 7 due to lower ET. However, forest root systems have access to deeper available water and are able to maintain ET during the summer months causing a greater correlation between soil moisture and ESI.
Figure 4.6: All Stations Average 4inANOM Pearson R Values by Month and Dominant Land Cover Type.
Figures 4.7-4.11 show statistically significant Pearson R correlations by station, month and dominant land cover. Months 9-11 tend to show better agreement between ESI and soil moisture by depth based on the number of total dots counted across plots. This would make sense with previous plots where maximum correlations were found in month 11. During month 7 forest cover tends to dominate in statistical significance compared to agricultural land based on the fact that there are 5 green dots compared to 1 red dot in the figure for this time. This might provide further evidence that deeper root systems sampled in this study are better able to ET during and maintain ESI’s statistically significant correlation with soil moisture. However, it is important to note that forest cover is present in pixels more often than agricultural land, as shown in the Table 4.9 and that this study did not perform any analysis with ET itself. Therefore, this relationship may only exist because of bias caused by an unequal sampling of dominant land cover classes.
Figure 4.7: Significant $p < 0.05$ 2inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type.
Figure 4.8: Significant $p < 0.05$ 4inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type.
Figure 4.9: Significant \( p < 0.05 \) 8inANOM versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type.
Significant Stations (p<0.05) 20inANOM versus 1 Week ESI Pearson R by Month

Figure 4.10: Significant $p < 0.05$ 20inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type.
Figure 4.11: Significant $p < 0.05$ 40inANOM Versus 1 week ESI Pearson R Station Correlations by Month and Dominant Land Cover Type.
Table 4.9: Dominant Land Cover by Station.

<table>
<thead>
<tr>
<th>Dominant Land Cover</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>agland 2177:AL:SCAN</td>
</tr>
<tr>
<td>1</td>
<td>agland 2053:AL:SCAN</td>
</tr>
<tr>
<td>2</td>
<td>agland 2174:AL:SCAN</td>
</tr>
<tr>
<td>3</td>
<td>agland 2078:AL:SCAN</td>
</tr>
<tr>
<td>4</td>
<td>agland 2180:AL:SCAN</td>
</tr>
<tr>
<td>5</td>
<td>agland 2178:AL:SCAN</td>
</tr>
<tr>
<td>6</td>
<td>agland 2181:AL:SCAN</td>
</tr>
<tr>
<td>7</td>
<td>forest 2179:AL:SCAN</td>
</tr>
<tr>
<td>8</td>
<td>forest 2173:AL:SCAN</td>
</tr>
<tr>
<td>9</td>
<td>forest 2175:AL:SCAN</td>
</tr>
<tr>
<td>10</td>
<td>forest 2114:AL:SCAN</td>
</tr>
<tr>
<td>11</td>
<td>forest 2055:AL:SCAN</td>
</tr>
<tr>
<td>12</td>
<td>forest 2113:AL:SCAN</td>
</tr>
<tr>
<td>13</td>
<td>forest 2057:AL:SCAN</td>
</tr>
<tr>
<td>14</td>
<td>forest 2056:AL:SCAN</td>
</tr>
<tr>
<td>15</td>
<td>forest 2115:AL:SCAN</td>
</tr>
<tr>
<td>16</td>
<td>forest 2182:AL:SCAN</td>
</tr>
<tr>
<td>17</td>
<td>forest 2176:AL:SCAN</td>
</tr>
</tbody>
</table>
The relationship between ESI and soil moisture is hard to explain since different parts of the year show varying results. For example, Figure 4.12 shows an example of 1 week ESI versus 4inANOM for a case year (2006) at station 2053. ESI tends to increase and decrease with soil moisture. This means that for most of the time series, ESI mimicks 4inANOM in its up and down motions. However, in December the ESI signal departs from soil moisture by moving up while 4inANOM goes down. In some cases, there also seems to be a time component where ESI lags behind or proceeds after the soil moisture signal. For example, there is an ESI response to 4inANOM between months 1-3 because it is shifted to the right of the soil moisture signal visually, while other months (6, 7) show an ESI signal to the left of the soil moisture time series.
\textbf{Figure 4.12:} 4inANOM vs 1 week ESI at 2053:AL:SCAN during Drought Year 2006.
Figure 4.13 provides evidence from station 2078 showing how the Pearson R correlation can be improved by shifting the 4inANOM week mean to the left of the ESI time series using the WLI methodology described in section 3.2.1. For example, when comparing 1 week ESI and 4inANOM at the same WLI (0), the Pearson R is lower than a similar comparison at -5 WLI. This research might suggest that ESI could have a strong linear relationship with earlier soil moisture signals depending on various unknown factors. An explain for station 2078's maximum correlation at -5 WLI could be physical characteristics caused by the B type soil surrounding the sensor. This soil type produces similar results which are shown later in this study across depth. The better relationships found at extreme WLI values (-30 and 15) could represent ESI capturing cyclical patterns common to the soil moisture climatology.
Figure 4.13: Month 10 4inANOM Week Lag Pearson R Values at 2078:AL:SCAN Shifted to the Right and Left of 1 Week ESI by Increments of 1 Week.
Figure 4.14: All Stations 4inANOM Average Pearson R Values for Week Lag Interval (WLI) by Month.
Grouping the data by month across all station data shows that the location of maximum correlation across WLI values are typically constant. For example, Figure 4.14 shows WLI versus average Pearson R peaking at 0 WLI for 9 out of 12 months. This could suggest that ESI tends to correlate well with soil moisture at 1 to 1 WLI overall, while some stations show better relationship with ESI at later or earlier soil moisture signals. However, some months diverge from this pattern with maximum correlations at extreme WLI. For example, during month 12 the maximum average Pearson R peaks at $\sim -25$ WLI, whereas month 1 provides an even earlier maximum at -30.

Figure 4.15 shows that comparing WLI and Pearson R by seasonal groupings creates wide dispersion in the 95% confidence interval (blue shading). Of interest is the MAM months where the WLI’s close to 0 show large variability from the mean, or a lack of precision. This is also true for many WLI’s in the DJF months, where extremely negative and positive values tend to be imprecise, or provide a larger 95% confidence interval. The highest Pearson R values tend to exist at the later seasons with JJA and SON showing better precision and weak to moderate Pearson R values close to 0 WLI.
Figure 4.15: All Stations 4inANOM Seasonal Pearson R Values for Week Lag Interval (WLI). DJF = December, January, February. MAM = March, April, May. JJA = June, July, August. SON = September, October, November. The blue line is the mean Pearson R. The blue shading is the 95% confidence interval.
The next group of figures compare WLI, month and Pearson R for all stations broken down by depth and soil type. For example, Figures 4.16 and 4.17 are a set of plots corresponding to the 2inANOM depth, whereby the top panel contains WLI on the x axis and the bottom panel contains a similar comparison except by month instead. There is a similar setup for Figures 4.18 and 4.19 at the 4inANOM depth. This figure setup repeats all the way to 4.26 ending at the 40inANOM depth to make it easy to look at Pearson R by soil type and WLI as well as across month in a co-located way.

One feature that exists for the top panel of each depth set is that the maximum average Pearson R is located at similar WLI values for each soil type. For example, in Figures 4.22 (20inANOM) and 4.24 (40inANOM), the peak of the C soil type tends to exist at the right of each plot, or 3-5 WLI. This could suggest that the ESI signal is more predictive of, or has a better linear relationship with weekly soil moisture means calculated in C soils and future time stamps. Similarly, A type soils correspond to \( \sim 1 \text{ WLI} \) in Figure 4.16 (2inANOM), 4.18 (4inANOM), 4.20 (8inANOM), and 4.22 (20inANOM) but diverge from this pattern in Figure 4.25 (40inANOM). This could suggest that ESI and A type soils have a similar relationship to ESI and C type soils, but the A type soils also depend on depth.
Figure 4.16: All Stations Average 2inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type.

Figure 4.17: All Stations Average 2inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month.
Figure 4.18: All Stations Average 4inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type.

Figure 4.19: All Stations Average 4inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month.
Figure 4.20: All Stations Average 8inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type.

Figure 4.21: All Stations Average 8inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month.
Figure 4.22: All Stations Average 20inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type.

Figure 4.23: All Stations Average 20inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month.
Figure 4.24: All Stations Average 40inANOM Week Lag Interval versus 1 Week ESI Pearson R Values by Soil Type.

Figure 4.25: All Stations Average 40inANOM versus 1 Week ESI Pearson R Values by Soil Type and Month.
There are soil groups that perform similarly across depths. For example, B type soils tend to peak at or around 0 WLI for all depths, with D soils following a very close pattern. B and D soils are grouped together by their higher percentage of silt as shown in Figure 3.4. A and C soils share similar patterns but diverge at the 40inANOM depth. A hypothesis to explain these groupings could be related to soil texture. For example, because silty soils (types B and D) have a larger surface area than sandy soils (types A and C), they hold moisture longer causing soil moisture signals in the past to match better with a drier ESI signal in the future. However, soils with more sand have higher infiltration rates, causing an increased drying speed and a better linear relationship with ESI at a 1 to 1 WLI. One problem, however is the unequal amount of soil types by sensor count as shown in Figure 4.26.

Across monthly average Pearson R values in the 2, 4, and 8inANOM (Figures 4.17, 4.19 and 4.21), there are dips in the D soil type in month 7. However, types B and A seem to maintain their strength during this time. In Figure 4.23 (20inANOM), Pearson R values by soil type climb up and down erratically and lack similarity across previous depths. Figure 4.25 (40inANOM) is also unique when compared to other plots. This is because the drop is experienced by type B and C soils, while D and A maintain their strength. Because of this, there does not seem to be a clear pattern of what soil types perform better across monthly climatologies. However, there are interesting features across depths, such as the B type soils having the most strength in month 4, as well as the A type soils having a distinct maximum at month 11 for 3 out of 5 depth cases.
Figure 4.26: Count of Total Soil Moisture Sensors by Hydrologic Soil Type.
4.3 Accuracy Results

Confusion matrix plots can be used to visually characterize ESI’s ability to predict wet (above 0) and dry (below 0) classes of dANOM soil moisture. This is shown in Figures 4.27-4.31 for each depth. A pattern that exists for all 5 depths is a higher true negative and positive detection of soil moisture classes (wet=0, dry=1). True negative detection tends to dominate across plots followed by true positive values. This might suggest that ESI is able to predict wet classes more accurately than dry. However, as shown in Table 4.1, the mean of ESI is 0.02 suggesting a lean towards wet ESI values in the data. This could skew the results towards a higher number of true negative detections.
Figure 4.28: Confusion Matrix of ESI Prediction Accuracy for 4inANOM.

Figure 4.29: Confusion Matrix of ESI Prediction Accuracy for 8inANOM.
Figure 4.30: Confusion Matrix of ESI Prediction Accuracy for 20inANOM.

Figure 4.31: Confusion Matrix of ESI Prediction Accuracy for 40inANOM.
Figure 4.32: Overall Accuracy for ESI Prediction by Depth.

The overall accuracy of ESI prediction on soil moisture class is weakly impacted by depth. For example, accuracy tends to be slightly higher in the upper soil column as shown in Figure 4.32. The minimum and maximum accuracy is at 40inANOM and 2inANOM with a value of 57% and 60% respectively. The relationship described in Figure 4.32 suggests that accuracy falls as depth increases, or that ESI less accurately predicts the lower soil depths overall. One explanation for this result is that the thermal channel used by ALEXI may be impacted with depth because of the increased mass that the signal must travel through in order to be captured by the satellite sensor. This physical limitation to emission may describe why a thermal signal originating at deeper depths is less accurate at characterizing soil moisture.
Figure 4.33: All Stations Binned 1 Week ESI Prediction Accuracy by Depth.

ESI bins in the middle of the data distribution perform worse, whereas bins on the extremes show distinct maximums. For example, as shown in Figure 4.33, there is a clear peak for all depths in the dry ESI bins between -2.0 and -1.5, where accuracy reaches 70%. However, accuracy is more dispersed on the right side of the Figure or lacks a clear maximum. Bins that perform well on the wet side of ESI are between 1.0 and 2.0, but only the 2inANOM goes above 65% at values near 2.0. Bins between -0.5 and 0.5 show distinct lows, with 40inANOM reaching a distinct minimum of 45%.
Figure 4.34: ESI Prediction Accuracy by All Depths and Months.

Figure 4.34 is a plot of Table 4.10 and shows similar patterns across monthly values of accuracy by depth when compared to Pearson R (Figure 4.1). For example, maximum values exist in months 4-5 and 9-11. In addition, there are top-to-bottom relationships where accuracy tends to be higher in the upper soil depths and smaller in the deeper layers. There is also a flip in this relationship at month 7, showing a stronger 40inANOM that overtakes four depths above it. Lastly, there is a month 7 dip that may be explained by land cover dominance (Figures 4.35 and Figure 4.36). Table 4.11 shows that the maximum accuracy was found in the 2inANOM sensor at $\sim 0.58$ during month 11. Minimums were found in the 20 and 40inANOM sensor, which tied at $\sim 0.48$ during month 1.
Table 4.10: All Stations ESI Prediction Accuracy by Depth and Month.

<table>
<thead>
<tr>
<th>Month</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.519914</td>
<td>0.514532</td>
<td>0.519914</td>
<td>0.482239</td>
<td>0.482239</td>
</tr>
<tr>
<td>2</td>
<td>0.548431</td>
<td>0.553888</td>
<td>0.540246</td>
<td>0.515689</td>
<td>0.527967</td>
</tr>
<tr>
<td>3</td>
<td>0.576316</td>
<td>0.578947</td>
<td>0.581579</td>
<td>0.530263</td>
<td>0.494737</td>
</tr>
<tr>
<td>4</td>
<td>0.616480</td>
<td>0.579858</td>
<td>0.594100</td>
<td>0.582909</td>
<td>0.547304</td>
</tr>
<tr>
<td>5</td>
<td>0.567531</td>
<td>0.581173</td>
<td>0.590723</td>
<td>0.568895</td>
<td>0.556617</td>
</tr>
<tr>
<td>6</td>
<td>0.589905</td>
<td>0.552050</td>
<td>0.558360</td>
<td>0.548896</td>
<td>0.534700</td>
</tr>
<tr>
<td>7</td>
<td>0.549096</td>
<td>0.545220</td>
<td>0.552972</td>
<td>0.538760</td>
<td>0.559432</td>
</tr>
<tr>
<td>8</td>
<td>0.589896</td>
<td>0.545319</td>
<td>0.566122</td>
<td>0.595840</td>
<td>0.561664</td>
</tr>
<tr>
<td>9</td>
<td>0.640407</td>
<td>0.630241</td>
<td>0.644219</td>
<td>0.649301</td>
<td>0.630241</td>
</tr>
<tr>
<td>10</td>
<td>0.659164</td>
<td>0.610932</td>
<td>0.647374</td>
<td>0.637728</td>
<td>0.621651</td>
</tr>
<tr>
<td>11</td>
<td>0.660777</td>
<td>0.656066</td>
<td>0.630153</td>
<td>0.654888</td>
<td>0.599529</td>
</tr>
<tr>
<td>12</td>
<td>0.536117</td>
<td>0.527088</td>
<td>0.546275</td>
<td>0.522573</td>
<td>0.532731</td>
</tr>
</tbody>
</table>

Table 4.11: All Stations ESI Prediction Accuracy by Depth and Month Descriptive Statistics.

<table>
<thead>
<tr>
<th>Month</th>
<th>2inANOM</th>
<th>4inANOM</th>
<th>8inANOM</th>
<th>20inANOM</th>
<th>40inANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>6.500000</td>
<td>0.587836</td>
<td>0.572943</td>
<td>0.581003</td>
<td>0.568998</td>
</tr>
<tr>
<td>std</td>
<td>3.605551</td>
<td>0.047566</td>
<td>0.042247</td>
<td>0.041781</td>
<td>0.056250</td>
</tr>
<tr>
<td>min</td>
<td>1.000000</td>
<td>0.519914</td>
<td>0.514532</td>
<td>0.519914</td>
<td>0.482239</td>
</tr>
<tr>
<td>25%</td>
<td>3.750000</td>
<td>0.548929</td>
<td>0.545295</td>
<td>0.551298</td>
<td>0.528341</td>
</tr>
<tr>
<td>50%</td>
<td>6.500000</td>
<td>0.583106</td>
<td>0.566418</td>
<td>0.573850</td>
<td>0.558895</td>
</tr>
<tr>
<td>75%</td>
<td>9.250000</td>
<td>0.622462</td>
<td>0.588613</td>
<td>0.603113</td>
<td>0.606312</td>
</tr>
<tr>
<td>max</td>
<td>12.000000</td>
<td>0.660777</td>
<td>0.656066</td>
<td>0.647374</td>
<td>0.654888</td>
</tr>
</tbody>
</table>
Agricultural and forest covered pixels show distinct depth patterns. For example, in Figure 4.35 the maximum accuracy values are during months 4 and 11 and occur in the 2inANOM depth. In Figure 4.36, the deeper 20inANOM layer overtakes the maximum accuracy values between months 9-11, suggesting that ESI is able to predict soil moisture classes better in deep layers with dominant forest cover. This is similar to Pearson R (Figures 4.10, 4.11) where upper layers in the soil column play a strong role in months 9-11, while deeper layers provide a stronger signal during the same time period in forest dominated pixels. Figure 4.37 summarizes the 4inANOM accuracy by land cover dominance and suggests that the dip in month 7 is due to agricultural. This dip is also dip shown in a previous plot (Figure 4.12) describing the difference between Pearson R values over agricultural and forest covered pixels.
Figure 4.35: ESI Prediction Accuracy by Month and Depth for Agricultural Land Cover Type.
Figure 4.36: ESI Prediction Accuracy by Month and Depth for Forest Land Cover Type.
Figure 4.37: All Stations Average 4inANOM ESI Prediction Accuracy by Month and Dominant Land Cover Type.
Figure 4.38: Month 10 4inANOM Week Lag Accuracy Values at 2078:AL:SCAN Shifted to the Right and Left of 1 Week ESI by Increments of 1 Week.
One difference with the Pearson R analysis is that while WLI may improve station correlations, it does not have significant impacts on accuracy. Figure 4.38 shows WLI and accuracy at station 2078 for month 10. Improvements follow the same up and down pattern as Pearson R in Figure 4.13, but accuracy only increases by $\sim 2$ percent between WLI 0 and -5, while the Pearson R example showed increases of $\sim 11$ percent. Figure 4.40 shows that the majority of the ESI data is in the middle of the histogram. Therefore, accuracy assessments using WLI may be impacted by the bin effects shown in Figure 4.33, where ESI fails to make accurate predictions above 55% until values are below -0.5 and above 0.5.

Figure 4.39 shows that WLI fails to produce the u-shape curve in accuracy for each of the seasonal groupings shown in a similar Pearson R plot in Figure 4.15. The variability in the 95% confidence interval is also unified and regular across WLI and season suggesting little change in improvement. This is different than the Pearson R example which showed examples of wide and precise patterns of the 95% confidence interval for specific WLI. After these findings, analysis was discontinued for WLI’s affect on accuracy because of time constraints and a desire to answer other scientific questions. Therefore, the accuracy analysis by depth and soil type is unlike previous examples with Pearson R (Figures 4.16-4.25) where figures containing WLI versus accuracy and soil type are missing.
Figure 4.39: All Stations 4inANOM Seasonal Accuracy Values for Week Lag Interval (WLI). DJF = December, January, February. MAM = March, April, May. JJA = June, July, August. SON = September, October, November.
Figure 4.40: All Stations Count of Binned 1 Week ESI Values.
Figure 4.41: All Stations 2inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type.

Figure 4.42: All Stations 4inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type.
Figure 4.43: All Stations 8inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type.

Figure 4.44: All Stations 20inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type.
Figure 4.45: All Stations 40inANOM 1 Week ESI Prediction Accuracy by Month and Soil Type.
One difference between analyzing accuracy versus Pearson R by soil type is that the values lie on different scales. For example, for each of the Pearson R plots (Figures 4.17, 4.19, 4.21, 4.23, and 4.25), the scale is between -0.2 and 0.6 and values fall between this range. This is different from each of the accuracy plots (figures 4.41-4.45) where the scale is between 0.2 and 0.8 and most values fall between 0.4 and 0.6. For accuracy and Pearson R, soil types show similarities across month. For example, the shape of the accuracy and Pearson R lines for type A, B, and D soils are similar regardless of depth. This would suggest that on average, accuracy and Pearson R are typically stable across months, depths and soil type.

The lower depths (20inANOM and 40inANOM) show distinct minimums and maximums at specific months and soil types. Like Pearson R (Figure 4.23 and 4.25), there are characteristic dips in month 6 associated with accuracy (Figure 4.44) for soil type C. However, the dip is associated with the 40inANOM in Pearson R and the 20inANOM in accuracy for the C soil type. There is also a strong increase in accuracy for the A soil type in month 6 for the 40inANOM. However, there are characteristic differences between the 20 and 40inANOM plots for the C soil type that exist for Pearson R and accuracy assessments. This is probably caused by the limited samples that exist for soil moisture sensors as shown in Figure 4.26. Note that there are no examples of 2, 4 or 8inANOM C type soils shown.
4.4 Flash Drought Assessment Results

Figure 4.46 provides a quantitative relationship between 4inANOM soil moisture percentiles binned by ESI values. This graph provides evidence that as ESI increases, so does the mean 4inANOM soil moisture mean. This graph also provides evidence that ranges of soil moisture values could be identified from their associated ESI signal. Therefore, the results suggest that ESI may be of value to end users who wish to define soil moisture characteristics over a 5km pixel during the growing season. However, one problem with this suggestion is that there is large uncertainty shown by error bars. An explanation for this error could be that a point measurement of soil moisture is not representative of a 5km study domain. This means that data collected by a soil moisture sensor is specific to a soil type, root system and land cover that may not be found on average over the ESI pixel.

Hydrologic soil processes might explain the variability found when comparing ESI and soil moisture percentiles. For example, Figure 4.47 shows that the mean percentile rank of the D soil type is higher than types A and B until the 0.45-0.68 ESI bin, but it falls below the A and B mean after this bin. An explanation for this result is that when ESI returns a wetter bin, there is also high infiltration rates in the A and B soils. This would cause an overall higher mean percentile rank value because water is entering the soil column. However, B and D soils have a tendency to allow more rainfall runoff in wetter environments causing a smaller percentile rank mean in the same soil column. Inversely,
when ESI is dry, D type soils tend to stay wet because of their increased surface area and ability to hold onto moisture causing an elevated mean in comparison to drier B and A soils. Therefore, the use of ESI to observe hydrologic processes associated with soil physics causing runoff or infiltration should be explored.
Figure 4.47: All Stations Binned 1 Week ESI Value versus 4inANOM Percentile Rank and Soil Type for Growing Season (Months 5-11).
Table 4.12: Statistics for 224 SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor including CPC Soil Moisture Drought Category Classifications (Column 1), Drought Duration (Column 2), and Total Change in 4inANOM Percentile During Event (Column 3).

<table>
<thead>
<tr>
<th>USDM Category</th>
<th>Drought Duration (Weeks)</th>
<th>4inANOM Percentile Rank Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>224.000000</td>
<td>224.000000</td>
</tr>
<tr>
<td>mean</td>
<td>2.218750</td>
<td>5.669643</td>
</tr>
<tr>
<td>std</td>
<td>1.228001</td>
<td>3.310233</td>
</tr>
<tr>
<td>min</td>
<td>1.000000</td>
<td>3.000000</td>
</tr>
<tr>
<td>25%</td>
<td>1.000000</td>
<td>3.000000</td>
</tr>
<tr>
<td>50%</td>
<td>2.000000</td>
<td>5.000000</td>
</tr>
<tr>
<td>75%</td>
<td>3.250000</td>
<td>7.000000</td>
</tr>
<tr>
<td>max</td>
<td>4.000000</td>
<td>31.000000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-49.890558</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25.734391</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-97.386441</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-70.352089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-51.289660</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-28.075782</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.358137</td>
</tr>
</tbody>
</table>

This research is unique in that it shows how SMVI can be applied to soil moisture percentiles to detect flash drought events. For example, Table 4.12 provides a summary of descriptive statistics for 224 events identified using 4inANOM values. The Table is useful because it provides a snapshot of what flash droughts mean in the context of soil moisture records. For example, the mean 20 year decrease in recorded 4inANOM percentiles was 49.89 and the mean CPC defined USDM category was 2.21 for 20 years of data. This Table also provides an idea of what extreme cases might mean for states like Alabama, where the longest flash drought event occurred for 31 recorded weeks. Similarly the most extreme example showed a 97.39 decline in 4inANOM percentile ranks during its time series. Therefore, this work may be useful to end users who wish to understand flash drought events in the context of a soil moisture climatology (2002-2020).
SMVI identified flash droughts show expected physical relationships by soil type. This means that a higher number of events occurred in A type soils as shown in Figure 4.48. This relationship may be explained by the fact that A type soils tend to be higher in sand content followed by B and D. High sand content causes increased infiltration and therefore drying, whereas lower sand content would mean a greater ability to withstand the loss of available water due to greater surface area. One unexpected result is shown in Figure 4.49, where flash droughts occurred in forest covered pixels more often than Agricultural. However, this could be explained by the unequal sampling of forest cover compared to as shown in Table 4.3.
Figure 4.48: Months 5-11 Total SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Hydrologic Soil Type.
Figure 4.49: Months 5-11 Total SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Dominant Land Cover.
Figure 4.50: Count of 224 SMVI Identified Flash Droughts at 4 Inch Soil Moisture Sensor by Month.
Figure 4.50 provides a monthly count of all flash drought events identified with SMVI. In the context of previous analysis, flash droughts have occurred at high rates during periods where ESI shows high linear relationship with soil moisture and accuracy at predicting wet and dry classes. For example, ESI may be a good tool for monitoring flash droughts during months 4 and 9-11 because of its performance with soil moisture during that time. However, while the highest Pearson R and accuracy values tend to be in month 11 (figures 4.7, 4.34), flash droughts seem to have less occurrence during this month. Therefore, it is important to understand the implications of the data and its performance during various time periods before selecting ESI as a tool to monitor a flash drought.

Figure 4.51 provides an evaluation of different ESI variables ability to predict flash drought events at 5 time indices: the start, 1 and 2 days prior, and 1 and 2 days after. Six of the top 10 most accurate variables use the change anomaly methodology, and are at the prior2 or prior1 index. Figure 4.52 is a similar graph where flash droughts are indexed by significant stations months shown in Figure 4.3 across the 4inANOM sensor. Accuracy improves at the top variable between both graphs from $\sim 60$ percent to $\sim 76$ percent. All 10 of the most accurate variables are also at the prior2 and prior1 index and also use the change anomaly methodology. This may indicate that using ESI change anomalies over regular composites reduces noise in the data that increases prediction error. It also suggests that ESI prediction of flash drought can be improved by using significant station correlation months.
Figure 4.51: Month 5-11 ESI True Negative Prediction of Flash Drought by Variable and Index. Explanation: Change anomalies on the x-axis have an [index][composite][week difference] naming convention. For example, prior2esi2_4wk means the 2 week ESI composite with a 4 week difference applied, 2 weeks prior to flash drought occurrence. In equation 3.1, it could be described as prior2 ΔESI2(4).

There are also regular composites of ESI in the Figure with a [index][composite] naming convention. For example, startesi8 is the 8 week composite at the start of the flash drought event, or start ESI8 as defined by eq. 2.14.
Figure 4.52: ESI True Negative Prediction of Flash Drought by Variable and Index for Significant Station Months.
Figures 4.53-4.56 provide evidence that flash droughts can be monitored by various ESI change anomalies with expected results. This means that as the drought duration increases across events, the cumulative sum of each change anomaly example decreases on average. One pattern that emerges is that as the composite and difference increase, so does the scale of the y axis. Furthermore, larger composites with increased difference intervals typically produce wider top to bottom limits, or have greater curvature over similar flash drought periods. Another pattern that is apparent, is that change anomalies produced on smaller composites with shorter week differences (Figure 4.53) produce less precision in the 95% confidence interval at week 14 when compared to larger composites with a greater week difference (Figure 4.56). However, the number of events that reach week 14 are small (Figure 4.57) and precision may be impacted by rebounding or degenerating soil moisture values at week 7 where the variability is high.
Figure 4.53: Months 5-11 95% Confidence Interval 1 Week ESI Change Anomaly (2 Week Difference) versus Flash Drought Duration Months.
Figure 4.54: Months 5-11 95% Confidence Interval 4 Week ESI Change Anomaly (3 Week Difference) versus Flash Drought Duration.
Figure 4.55: Months 5-11 95% Confidence Interval 8 Week ESI Change Anomaly (3 Week Difference) versus Flash Drought Duration Months.
Figure 4.56: Months 5-11 95% Confidence Interval 8 Week ESI Change Anomaly (4 Week Difference) versus Flash Drought Duration.
Figure 4.57: Months 5-11 8 Week ESI Change Anomaly (4 Week Difference) versus Flash Drought Duration. Number of Flash Droughts Reaching Each Week shown in Each Interquartile Range.
ESI change anomalies show good visual results with soil moisture information for two flash drought cases. Figure 4.58 shows a SMVI identified flash drought event (highlighted in transparent red) that occurred at station 2180 between September and early November of 2016. That flash drought resulted in the 4inANOM soil moisture (blue line) fall by $\sim 90$ percentiles. The associated 8 Week ESI change anomaly (yellow line) fell from +5 to 0 during the flash drought and continued to -5 for a few weeks after it ended. During this highlighted event, the drought monitor recorded a D1 category, while the CPC soil moisture percentiles recorded the 4inANOM sensor as a D4. Figure 4.59 shows another flash drought event where the 4inANOM soil moisture sensor recorded a D2 at station 2174 between October and December. Like the prior example, the ESI change anomaly shows agreement with the 4inANOM soil moisture percentiles with a difference in drought monitor categories. Therefore, more work needs to be done to understand what ESI means in relationship to in-situ drought categories, and why the USDM can sometimes misrepresent it.
Figure 4.58: Station 2180 Flash Drought Event versus 8 Week Change Anomaly (4 Week Difference), 4inANOM Soil Moisture Percentiles, and USDM Category).
Figure 4.59: Station 2174 Flash Drought Event versus 8 Week Change Anomaly (3 Week Difference), 4inANOM Soil Moisture Percentiles, and USDM Category.
Chapter 5. Summary and Discussion

5.1 Summary of Results

This chapter describes the overall findings related to the three science questions proposed in chapter 1. The first subsection notes that the relationship between soil moisture and ESI varies depending on the month, with correlation maximums in early and later months and minimums in winter months. This section also shows that agricultural and forest dominance, sensor depth, and the WLI methodology can affect the correlation between soil moisture and ESI. The second subsection highlights that ESI’s accuracy in predicting soil moisture classes varies depending on the month of measurement with maximum and minimums following similar patterns with the best correlation months, but WLI does not improve accuracy. It also shows that deeper layers tend to have higher accuracy in forest-dominated pixels during July and October. The third subsection shows that soil moisture values can be estimated from ESI signals, but with high uncertainty due to large error bars. Flash drought events can be identified using SMVI and their occurrence can be tracked by ESI change anomalies with good performance in line with the correlation and accuracy assessments.
5.1.1 Question 1

1.) What is the statistical relationship between ESI and \textit{in-situ} soil moisture retrievals at 18 SCAN sites across Alabama?

- Soil moisture may be more or less linearly related with ESI depending on the month that values are compared.

- Correlation maximums tend to occur in the early and later months (April-May and September-November) across depth.

- Correlation minimums tend to occur in the winter months (December-February) across depth.

- There exists a trough where correlation decreases and then increases between monthly maximums (May-August).

- Agland dominance tends to cause shallower depths to correlate more during earlier and later months (April-May and September-November).

- Forest dominance tends to cause deeper depths to correlate more compared to shallower depths during July and October.

- Sensors higher in the soil profile typically show a stronger correlation while deeper sensors correlate less.

- There is a strong decline in correlation in July for agriculture examples while correlation is maintained for the same month in forest cases.
• Forest cover tends to dominate during statistically significant July relationships across depth.

• Using the WLI methodology can improve correlation.

• Comparing Pearson R across WLI and seasonal groupings creates wide dispersion in the 95% confidence interval during MAM and DJF months.

• Comparing Pearson R across WLI and soil groupings show similar behavior in type A and C soils as well as type B and D soils.

5.1.2 Question 2

2.) How accurate is ESI in generalizing in-situ soil moisture conditions across varying soil and land cover characteristics?

• ESI can predict soil moisture classes with higher or lower accuracy depending on the month it is measured.

• Accuracy maximums follow similar patterns with correlation in terms of maximum and minimum months, and shows a trough between maximums.

• ESI predictions show a high true negative and true positive detection of soil moisture classes across depth.

• ESI bins in the middle of the data distribution are less accurate, whereas bins on the extremes show distinct maximums.

• There are top-to-bottom relationships where accuracy tends to be higher in the upper soil column and lower in the deeper layers.
• Accuracy is similar to Pearson R, where upper layers in the soil column play a strong role in September-November for agricultural pixels.

• Forest dominance tends to cause deeper depths to have higher accuracy than shallower depths during July and October.

• Deeper layers provide a stronger accuracy signal during September-November in forest dominated pixels.

• While WLI may improve station correlation, it does not seem to have a significant impact on accuracy.

• WLI fails to produce the u-shaped improvement curve in accuracy for each of the seasonal groupings similar to the Pearson R plots.

• The shape of the accuracy time-series across month for type A, B, and D soils are similar regardless of depth.

• There are characteristic decreases in June associated with accuracy for soil type C in 20inANOM and increases in soil type A in 40inANOM.
5.1.3 Question 3

3.) What value can ESI provide in monitoring flash drought events identified by soil moisture retrievals?

- Ranges of soil moisture values could be identified from their associated ESI signal, but uncertainty is high due to large error bars.

- Mean percentile rank of the D soil type is higher than types A and B until the 0.45-0.68 ESI bin, but it falls below the A and B mean after this bin.

- SMVI can be applied to soil moisture percentiles to detect flash drought events.

- SMVI identified flash droughts show expected physical relationships by soil type where A type soils show the most events, followed by B and C.

- Flash droughts have occurred at high rates during periods where ESI shows high linear relationship with soil moisture and accuracy at predicting wet and dry classes.

- While the highest Pearson R and accuracy values tend to be in month 11, flash droughts seem to have less occurrence during this month.

- ESI prediction of flash drought can be improved by using significant station correlation months.
• When predicting flash drought initiation, 6 of the top 10 most accurate variables use the change anomaly methodology, and are at the prior2 or prior1 index.

• Change anomalies produced on smaller composites with shorter week differences produce less precision in the 95% confidence interval by week 14 compared to larger composites with a greater week difference.

• Evidence shows that as drought duration increases, the cumulative sum of tested change anomalies fall on average.

• ESI change anomalies applied to a cumulative sum show good visual results with soil moisture information for two flash drought cases.

• USDM misrepresents in-situ drought categories calculated by CPC soil moisture percentiles.

5.2 Limitations and Recommendations

In this study, a 5 km ESI product was used to compare soil moisture measurements from 18 SCAN stations across 5 depths, but the findings are subject to significant uncertainty. For example, Wanders in 2012 suggested that the validation of remotely sensed soil moisture products from satellites is hampered by the footprint of the spatial domain in comparison to in-situ sites [65]. This study did not examine how the location of the SCAN station could impact the Pearson R and accuracy assessments, as some soil moisture sensors were positioned at the corner or side of each 5 km ESI domain. Future research should investigate if
the distance between the SCAN station and the center of the ESI pixel create error, or at least describe the uncertainty caused by the size of the ESI domain in relation to a point measurement.

In addition, this study did not characterize soil classes across the entire 5 km study domain. This causes uncertainty surrounding the soil class data presented by each sensor location as being representative of the ESI pixel. All stations were also located in small areas mostly covered by Bermuda grass, and while larger land cover dominance seemed to affect station correlations and accuracy, a higher resolution ESI model is needed to limit the uncertainty caused by physical characteristics in future studies. This might also be rectified by extending the soil moisture coverage into specific areas, like putting a couple stations in dense forest to clarify the results of this study.

Another uncertainty regards the soil data provided by USDA. For example, 9 SCAN stations had available Pedon reports and information had to be filled in with USDA Web Soil Survey. Therefore, USDA could rectify this problem by completing Pedon assessments for all Alabama stations by sensor depth to limit having to fill in this information with other products. For all future sensors that are installed by the Alabama State Climatology office, associated soil reports could also be unified and processed like those provided by SCAN stations that do have Pedon assessments.

One physical factor that may contribute to the difficulty of ESI representing soil moisture is the surface energy budget. In a water limited environment, the fraction of PET that is evaporated will be linked to the available moisture
on the surface. However, in an energy limited state, the available moisture and the surface are no longer tightly coupled to the latent energy. This can occur by a decrease in incoming solar radiation or by saturating the surface or lower boundary layer. In the southeastern US, the regime of energy limited vs water limited is fluid. In the summertime, conditions favor water limited as the dense vegetation ignites a rapid hydrologic cycle. However, convective downpours can often create near energy limited conditions for short periods of time. This can perhaps explain some of the degradation in midsummer months.

In the analysis of flash drought, there is uncertainty about the depth of soil moisture measurements used. This study utilized 4inANOM data because all soil moisture sensors were located in Bermuda grass, and this sensor showed maximum performance overall. However, due to the discovered relationships between land dominance and the mismatch between USDM and CPC-derived drought categories calculated by soil moisture percentiles, it would be beneficial to apply SMVI to all depths to create a more comprehensive flash drought dataset. This might provide a better picture of what identified flash droughts mean for ESI across depth.

A final recommendation is to use ESI and SMVI defined flash drought data in concert. While case studies in this thesis show how cumulatively summed ESI change anomalies can mimick soil moisture anomalies across time at particular pixels, a regional picture is needed. For example, a study should be done where flash droughts are identified using SMVI and then ESI change anomalies are calculated and cumulatively summed over larger regions. This kind of coupled
methodology may offer a better understanding of flash droughts, especially if they can be identified at high resolution and then monitored over larger regions with ESI.
References


