THESIS

EXPLORING THE CONTRIBUTION OF CROP WATER USE TO REMOTELY SENSED ESTIMATES OF SOIL SALINITY IN IRRIGATED AGRICULTURE

Submitted by

Brian D. Craig

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2023

Master's Committee:

Advisor: José L. Chávez
Co-Advisor: Timothy K. Gates

Gregory L. Butters

ABSTRACT

EXPLORING THE CONTRIBUTION OF CROP WATER USE TO REMOTELY SENSED ESTIMATES OF SOIL SALINITY IN IRRIGATED AGRICULTURE

Globally, 72% of the world's water withdrawals are used for agriculture. As the world's population continues to grow and increase its caloric intake, agricultural producers must provide more food and fiber with the same amount of water and soil, or less, due to expanding urbanization and climate change. In Colorado (CO, U.S.A.), agricultural producers in the South Platte and Arkansas River Basins, for instance, have been offered water transfer programs to temporarily or permanently transfer their water shares to municipalities and industry. Another challenge agricultural growers face is soil salinization, which needs to be monitored. In the Arkansas River basin, upflux from saline shallow groundwater tables consistently contributes to crop evapotranspiration (ET), leaving salts in the vadose zone. These salts accumulate over decades to the point where crop yields decline, threatening agricultural sustainability. Remote sensing is an economical tool to monitor salinity (e.g., soil electrical conductivity; EC, dS m^{-1}) at large spatial scales. Existing remote sensing models that predict EC mostly utilize vegetation indices (VIs), which are arithmetic combinations of vegetation reflectances captured by discrete spectral bands. In this study, two additional explanatory variables were investigated: 1) the actual crop ET (ET_a, mm d^{-1}), and 2) the crop water stress index (CWSI). Calculations of ET_a were performed using Landsat satellite multispectral imagery and a surface energy balance approach. This research was conducted over two growing seasons in commercial maize fields located within the Fairmont Drainage District near Swink, CO. Results indicate that models including

ET_a or CWSI with existing VIs improve the accuracy of soil EC mapping over models including VIs alone. The developed EC models are accurate within ±1 dS m^{-1} (Root Mean Squared Error), which is considered well within the precision required to make pragmatic field and ditch company-level management decisions.

ACKNOWLEDGMENTS

I want to thank the United States Agency for International Development (USAID) that funded the U.S.-Pakistan Center for Advanced Studies in Water, through which this study was funded at Colorado State University, and to my advisors Dr. José Chávez and Dr. Tim Gates for giving me the opportunity to be a part of the research team and guiding me through the process. My education at Colorado State University has prepared me for an excellent career in water resources. All my professors and advisors at CSU are worthy of thanks. I want to thank A.J. Brown for all the guidance and assistance he gave me through the six-year process. From field techniques to statistical interpretation, this study would not be what it is today without A.J.'s participation. Further, Dr. Aymn Elhaddad showed great patience in teaching me how to use ReSET over email. Luke Engler, Rustin Jensen, Luke Stark, Ian Stockdill and Jerod Berg provided excellent assistance in the field and in the lab. The Brown family, Mike Bartolo and Lane Simmons provided friendly hospitality in Rocky Ford and made me feel at home. I want to thank my parents for supporting me in numerous ways throughout the graduate school process. A final thanks goes to my sun hat for keeping me cool in the scorching heat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Worldwide State of Agriculture

Globally, 72 percent of the world's water withdrawals are used for agriculture (Food and Agriculture Organization of the United Nations, 2020). Even though the majority of water use is applied to the agricultural sector, 3.2 billion people live in agricultural areas with high to very high water shortages or scarcity (Food and Agriculture Organization of the United Nations, 2020). As the world's population continues to grow, and developing countries increase their caloric intake, agricultural producers must provide more food and fiber with the same amount of water and soil, or less, due to expanding urbanization and climate change. Indeed, the Food and Agriculture Organization of the United Nations (FAO) estimates that farmers will need to produce 50% more food, livestock feed, fiber and biofuel by 2050 over 2012 levels, while agricultural land available per capita has declined 22 percent from 2000 to 2019 (Food and Agriculture Organization of the United Nation, 2021). Of the over 13 billion hectares of land on Earth, about 50% is completely unusable for crop production, while only about 11% of land is considered arable (Acquaah, 2005).

There are three strategies to increase crop production and meet the goals of the FAO: 1) clear new land; 2) increase the number or frequency of crops planted in a given area; and 3) intensify the yield of existing cropped land. However, farmers are often focused on maximizing economic yield rather than crop yield, so the cost of increasing per-hectare crop production must be considered when attempting to increase yield. This thesis is related to item 3, intensifying the yield of existing cropped land, through the identification and management of salinized lands. Irrigated agriculture, the setting for this work, is particularly important for sustaining or

increasing yields. Irrigated lands produce approximately 40% of the available food supply while taking up only about 20% of available cropland (Food and Agriculture Organization of the United Nation, 2021).

State of Agriculture in Colorado

According to the FAO (2021), water stress is high in basins where intense irrigation and densely populated cities compete for water resources. This description matches the state of Colorado, where agriculture in the South Platte and Arkansas River Basins (and beyond) are competing for resources with the growing Front Range population. The option of 'buy and dry,' where municipalities buy farms for their water rights, fallow the fields, and use the water for municipal and industrial use, has become more common in the 21st century and is putting more pressure on farmers to maximize crop production. The Sustainable Development Goal indicator 6.4.2 (Figure 1), which includes an overall measure of physical water scarcity, indicates a High level of water stress in the Colorado River Basin (which supplies significant water to the Arkansas and South Platte River basins), indicating that 75-100% of total renewable freshwater resources are withdrawn per year. This same study indicates a Low level of water stress (25-50% of total renewable freshwater resources are withdrawn per year) in the Missouri River Basin, which includes the Arkansas and South Platte River basins (Food and Agriculture Organization of the United Nation, 2021) (Figure 1). However, this categorization is skewed by the states east of Colorado that receive significantly more precipitation than Eastern Colorado, and if the analysis were further quantized down to the basins in Eastern Colorado, Eastern Colorado may be classified as having higher water stress. The introduction of irrigation to the fertile soil found in the lower Arkansas and South Platte basins in the late 1800s has created a widespread agriculturally based economy in these areas. In 2021, maize was harvested in

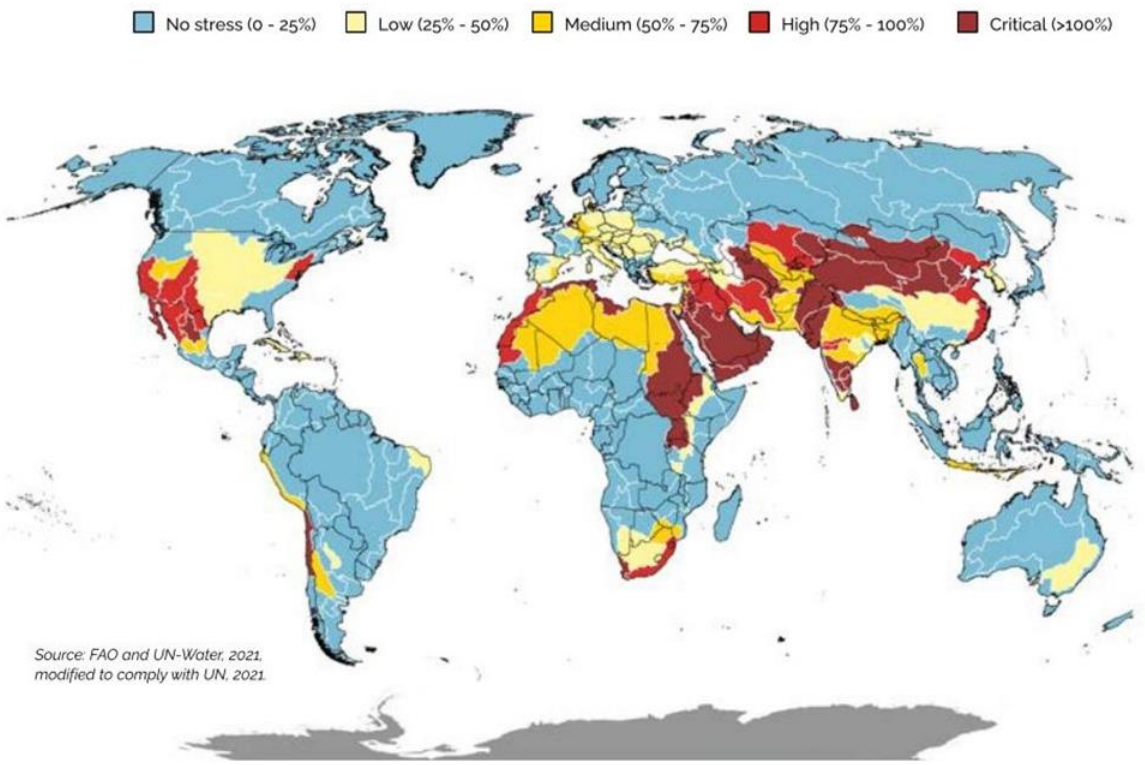

Figure 1: Sustainable Development Goal indicator 6.4.2, indicating High (75-100%) water stress in Western Colorado, where 75-100% of renewable water resources are used each year. Taken from FAO (2021).

558 thousand hectares (kha) throughout Colorado (NASS, 2021), and over 109 kha are irrigated in the Lower Arkansas River Basin alone (Gates et al. 2012). However, groundwater tables in the basin are high during the growing season, with depths of less than 2 m in 24% of a region of the Arkansas River basin ranging from approximately Fowler to Las Animas (Gates et al., 2016). High groundwater tables often lead to increased root zone soil salinity. In the Arkansas River basin, upflux consistently contributes to crop ET, bringing bedrock salts with it, and sprinklers that use groundwater pull from highly saline groundwater (Gates et al., 2012). Gates et al. (2012) found that average field salinity values (EC_e) ranged from 1.8 to 12 dS m^{-1} throughout the Arkansas River Basin in Colorado. The upper two-thirds of that range can significantly reduce crop yield.

Soil Salinity Development, Effects and Mitigation

Soil Salinity Development

Soil salinity refers to the total salt concentration in the soil solution (i.e., aqueous liquid phase of the soil and its solutes). Salts may originate through natural (primary salinization) or anthropogenic (secondary salinization) means. The most common source of soil salinity is the geochemical weathering and subsequent transport of salts from the rocks in the Earth's upper strata. Other natural sources include atmospheric deposition, seawater intrusion, and rising groundwater from saline aquifers, while anthropogenic sources include applied water, chemical fertilizers, and livestock waste (Acquaah, 2005).

Saline agricultural soils often result from the combination of a shallow groundwater table and an abundance of salts in the native soil and bedrock (Scudiero et al., 2014). Salts dissolve into the groundwater and rise through the water column through capillary action (Jiang et al., 2019). As the crops take up water, salts precipitate out of the water and into the soil. The intensity of salt accumulation within the crop root zone due to evapotranspiration (ET) depends upon the fraction of irrigation and precipitation that passes through the root zone, which is called the leaching fraction (LF). Sufficient leaching fractions indicate water is leaving the root zone and will decrease the total salts in the root zone due to their removal through the leached water.

Another source of salts in agricultural soils is the salts found in the applied water. Flood irrigation, a common practice in the Arkansas River Basin, involves significant return flows (water that was applied to the field and either directly runs back to the river, or leaches into the groundwater and enters the river through groundwater transport). Salts in the soil dissolve into this water before it leaves the field, increasing its electrical conductivity. When that water is later

diverted off the river for use on another downstream field, those salts are present in the field and able to be deposited when the crop roots take up water.

Although soil salinization occurs in all climates, soil salinization is common in arid and semi-arid soils where irrigation and rainfall are insufficient to leach salts (i.e., where precipitation is less than evaporation) (Aboelsoud et al., 2022). Other field conditions that result in salinization include areas where poor drainage and thus shallow water tables exist, where there is upslope recharge and downslope discharge, and where saline soils formed naturally from marine deposits. Thus, soil salinization is a concern in almost all arid and semi-arid agricultural areas.

Soil Salinity Effects

Salinity present within the crop root zone reduces the germination rate and decreases the osmotic potential in the root zone. Since osmotic potential is a negative value, this increases the osmotic potential gradient between the soil and crop roots (Ayers and Westcot, 1985). This larger gradient reduces crop water uptake, and thus reduces ET as well (NRCS, 1998). In practice, overwatering (either by watering more frequently than necessary or sustaining field capacity soil moisture for extended periods of time) is often seen in areas of high soil salinity, since applying extra water is thought to sustain yields and is often easier than other salinity solutions (in the short term). However, overwatering is a cost to the irrigators. This practice, when there is not efficient drainage, exacerbates the accumulation of salts, and can lead to root rot or oxygen deprivation. Root zone salinity can also cause ion toxicity and destroy the cells' internal structure (Zhu et al., 2021), which impacts crop physiological processes such as photosynthesis and transpiration, and lead to higher erosion rates (Metternicht and Zinck, 2003; Allbed and Kumar, 2013; Jiang et al., 2019;). The tolerance of a crop to salinity is traditionally characterized by plotting a crop's relative yield as a continuous function of soil salinity levels (Ayers and

Westcot, 1985). Even slightly saline soils can reduce plant growth, depending on the irrigation method (Ayers and Westcot, 1985; Davis et al., 2019). Soil salinity will also reduce a crop's yield and leaf area index (LAI) (Cui et al., 2022).

Early estimates stated that salt-affected soils made up approximately 23% of cultivated land, or about 350 million hectares (M ha) (Massoud, 1977). Currently, the FAO estimates that salinization and sodification (the accumulation of sodium, specifically) of soils removes 0.3 to 1.5 M ha of cropland per year, while reducing productivity on a further 20 to 46 million ha. (Food and Agriculture Organization of the United Nation, 2021) (Figure 2). Agricultural economic losses from salinization and sodification vary, but two global estimates state annual losses of approximately $12 billion USD (Ghassemi et al., 1995) and $27.3 billion USD (Qadir et al., 2014). Since salinization and sodification of soils reduce the soil structure, these processes also contribute to the erosion of topsoil, which is an issue for cereal

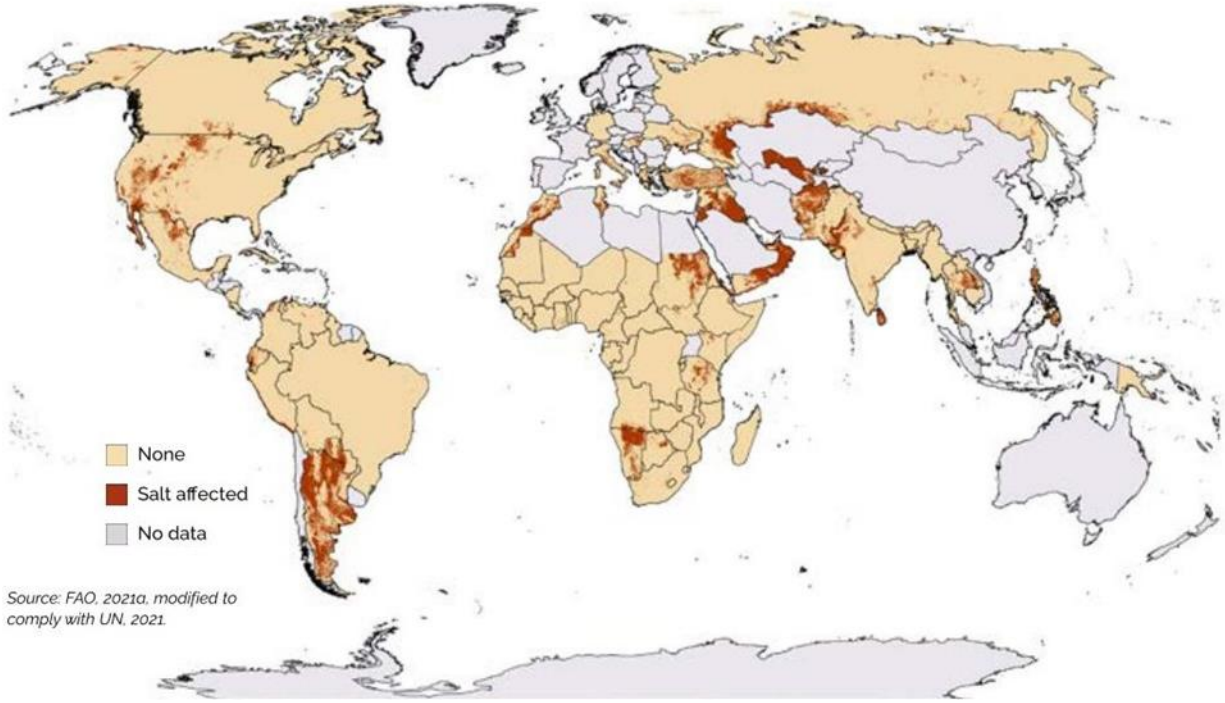

Figure 2: A global map of salt-affected land, taken from FAO 2021.

crops, where 7.6 million tons of topsoil is lost per year (Food and Agriculture Organization of the United Nation, 2021).

Soil Salinity Mitigation

Assessing soil salinity across multiple spatial scales is a key element in agronomic and water resources decision-making (Aboelsoud et al., 2022). Once salinity levels have been mapped, three approaches are common for managing soil salinity: properly-calculated crop water requirements and leaching fractions (Aboelsoud et al., 2022), timed irrigation, and/or crop selection based on salt tolerance. Irrigation and drainage methods for mitigating soil salinity include not over-watering crops, increasing field drainage, conducting off-season or cold-period irrigations (Cui et al., 2022), and finding water sources with lower salinity. Good drainage will allow saline water to leave the field rather than evaporate or transpire, where salts precipitate out of solution and into the root zone. These assessment and mitigation efforts are particularly crucial to producers and decision-makers in semi-arid or arid areas such as the Southwest and High Plains of the United States.

Many irrigation districts are not prepared to handle the problem of soil salinity. Local irrigation scheduling practices were often developed decades ago, before fields developed salinization issues. Thus, most areas have not tuned their irrigation application practices to saline conditions (Marino et al., 2019). The prevalence of flood irrigation throughout the world, if not paired with adequate drainage and managed effectively, will exacerbate salinity issues.

Electrical Conductivity

Soil salinity is commonly quantified by measuring the electrical conductivity (EC) of the soils, in the units of deci-Siemens per meter (dS m^{-1}). There are many sub-classifications of EC. Apparent

EC (EC_a) is the electrical conductivity of the bulk soil while preserving its structure and porosity. EC_a varies based upon a number of factors including soil salinity levels, texture, water content, soil bulk density, organic matter, clay minerology, cation exchange capacity, and temperature. Thus, EC_a is often used to direct soil sampling when mapping soil salinity, rather than being an absolute measurement of soil salinity itself.

A more standardized form of EC is measured by mixing a sample of soil with distilled water until it reaches saturation, filtering out the soil water, measuring the EC of the saturation extract at a given temperature (t, °C), and then converting that to the EC that would be observed at 25 degrees Celsius (°C). This is the electrical conductivity of the saturation paste extract at 25 °C, or EC_{e25}. The formula for EC_{e25} is in Eq. 1, where EC_{e25} is the electrical conductivity of the saturated paste extract corrected to 25 °C, and EC_{et} is the electrical conductivity of the saturated paste extract at a temperature t. EC_{e25} is commonly shortened to EC_e and will be referred to as such in this paper.

$$EC_{e25} = f_t \cdot EC_{et}, where\ f_t = 0.4470 + 1.4034 \cdot exp(-\frac{t}{26.815}) \quad (1)$$

Crop tolerance to salinity experiments generally used mixtures of NaCl and $CaCl_2$ salts (Maas et al., 1983; Maas, 1993). Gypsum ($CaSO_4 \cdot 2H_2O$), which is also found in some agricultural soils, is a relatively insoluble salt; in farm conditions, gypsum is not readily dissolved in the soil water. Because of this, if the soil salts contain a high proportion of gypsum, the laboratory analyst will detect EC_e at a higher level of salinity than the plant observes in field conditions. To counteract this error, a correction can be applied to measured EC_e values to give the value of soil electrical conductivity that the plant is experiencing in field conditions. EC_{eg} is the abbreviation for that EC reading of a paste extract with a gypsum correction such as that found in Brown (2021) (Eq. 2).

$$EC_{eg} = 0.9131 * EC_e - 0.7256 \quad for\ EC_e > 3.08 \text{ dS m}^{-1} \qquad (2)$$

$$= EC_e \qquad for\ EC_e < 3.08 \text{ dS m}^{-1}$$

There are three categories of salt-affected soils: saline, sodic, and saline-sodic. The U.S. Salinity Laboratory Staff (1954) classified saline soils as those with an EC_e > 4 dS m^{-1} and an exchangeable sodium percentage (ESP) of < 15%. Metternicht and Zinck (1997) categorized soil salinity into four ECe levels: non-saline (<4 dS m^{-1}), slightly saline (4-8 dS m^{-1}), moderately saline (8-16 dS m^{-1}) and strongly saline (>16 dS m^{-1}). Although 4 dS m^{-1} is the threshold for classification, many crops experience stress at EC_e < 4 dS m^{-1}. Even a slightly saline soil with an EC_e of 4 dS m^{-1} can result in a crop yield loss between 25 and 50 percent for maize, onion and potatoes (Ayers and Westcot, 1985).

Spatial and Temporal Variability of Soil Salinity

Soil salinity is a temporally and spatially dynamic property (Metternicht and Zinck, 1997; Corwin and Scudiero, 2019; Peng et al., 2019; Cui et al., 2022), since it is subject to a complex interaction of edaphic, pedogenic, anthropogenic, meteorological, topographic, and biological factors, and is thus difficult to monitor. Previous studies of soil EC have shown that the coefficient of variation of soil salinity is generally over 60% (Corwin et al., 2003). Furthermore, the scarcity of water is causing a shift from flood irrigation to pressurized irrigation (e.g., sprinkler, micro-sprinkler, and drip irrigation), which increases the spatial and temporal complexity of field-scale salinity distribution in the root zone.

General maps of soil type, giving common ranges of soil salinity, do exist, most notably those by the Natural Resources Conservation Service (NRCS), found in the Soil Survey Geographic Database (SSURGO). However, these maps are based on the soil parent material, often list a

range of salinities too broad to make farm management decisions, and do not consider anthropogenic forces. Corwin et al., (2017) found that root zone salinity in SSURGO was only accurate in 5 out of 22 surveyed fields within California's San Joaquin Valley.

Monitoring soil salinity at moderate values where vegetation is still present, but yields are significantly reduced has proven difficult, even within small areas in which management is essentially uniform (Delavar et al., 2020). As spatial scales increase, variations in management, pests, disease, climate, and soil properties can have a greater impact on vegetation growth than salinity, thus increasing the complexity of salt mapping at regional scales.

Characterizing spatial variability of soil salinity is a significant area of soil science and irrigation management because of its broad influence on field- to landscape-scale processes including crop yield and solute transport (Cui et al., 2022). These processes must be understood in order to optimize scarce water resources, to develop regulatory guidelines, and for assessing the trends of climate change. Technology necessary to solve this problem without expensive and impractical soil sampling has only emerged in the past few decades with the adoption of remote sensing (RS), global positioning systems (GPS), geographic information systems (GIS), and a more thorough understanding of EC_a. These technological advancements will help move the agricultural community toward the adoption of real-time digital information that will help "develop effective cross-sectoral policies that 'leave no one behind'" (Food and Agriculture Organization of the United Nation, 2021).

Methods of Detecting Soil Salinity

Corwin and Scudiero (2019) outlined six methods of detecting soil salinity: (1) qualitative visual assessment relating crop appearance to salinity level; (2) soil samples for laboratory analysis; (3)

four-electrode sensors or electromagnetic resistance (ER); (4) electromagnetic induction (EMI); (5) time-domain reflectometry (TDR); and (6) multi- and hyper-spectral remote sensing imagery. Qualitative visual assessment, though quick and inexpensive, does not provide accurate values of soil salinity levels for yield prediction or field-scale or district-wide water use planning, partially because it is a qualitative assessment, and partially because many factors affect crop health (Bai et al., 2018). Soil sampling, on the other hand, gives precise, accurate values of soil salinity by way of measuring soil EC. However, soil samples cannot be interpolated over more than a few meters and are very labor- intensive (Metternicht and Zinck, 2003; Li et al., 2015; Peng et al., 2019; Cui et al., 2022).

ER and EMI are both advantageous in that the user can collect a steady stream of data, allowing for constant movement around the field and quicker data collection over large areas. ER involves introducing an electrical current into the soil through current electrodes at the surface, then measuring the difference in the current flow potential at potential electrodes placed in the vicinity. This method is well suited for field-scale measurements because it measures a large volume of soil but is less reliable in dry or stony soils because good contact between the soil and the electrodes is required. ER works best in a damp, fallow field, which is often a practical limitation.

EMI involves a transmitter coil located at one end of the instrument that induces an eddy-current into the soil. The magnitude of these loops is directly proportional to the electrical conductivity in the vicinity of that loop. Each loop generates a secondary electromagnetic field that is proportional to the current flowing within the loop. A fraction of the secondary electromagnetic field from each loop is intercepted by the receiver coil of the instrument and the sum of these

signals forms an output voltage, which is related to the depth-weighted soil EC_a. EMI equipment is very expensive (>$10,000), which constrains it mostly to research purposes.

The TDR technique is based on the time for a voltage pulse to travel through a soil probe and back. This travel time is a function of the dielectric constant of the porous media in which the probe rests. TDR was later used to measure EC_a based upon the attenuation of the applied voltage as it travels through the soil. Due to its stationary nature, invasive installation, and the fact that it is measuring a very small portion of the soil, TDR is not realistic for field-scale salinity mapping.

Evapotranspiration and Crop Water Stress

Evapotranspiration (ET) is the rate at which water leaves the soil and plant through evaporation and transpiration. There are many different forms of ET. An observed ET rate is called Actual ET and is denoted ET_a. Reference ET (ET_r) is the maximum ET rate that could be expected from a well-watered grass or alfalfa reference field, based on weather conditions. Crop ET (ET_c) is the maximum ET rate that could be observed from a well-watered crop of a different type. ET_c is calculated by multiplying ET_r by a crop coefficient (K_c). K_c can be calculated either as a single or a dual coefficient.

ET_a can be a proxy for overall crop health, as a well-watered, unstressed crop will transpire more than a crop under stress. However, since ET_a is a proxy for overall crop health, rather than a singular stressor, it is difficult to separate different crop stressors by looking at ET_a rates alone.

Crop Water Stress Index

Another method of quantifying crop water use is through the Crop Water Stress Index (CWSI) (Jackson et al., 1981; Jackson et al., 1988). CWSI is a measure of the relative transpiration rate

occurring from a plant at the time of measurement in relation to its potential transpiration rate. It is expressed as a number between 0 and 1, with 0 indicating no water stress (fully evapotranspiring crop) and 1 indicating no evapotranspiration at all. CWSI can be calculated either using the evaporative fraction (ET_a /ET_c), shown in Equation 3:

$$CWSI = 1 - \frac{ET_a}{ET_c} \quad (3)$$

where ET_a is the actual ET and ET_c is the crop (potential) ET, or calculated using temperature measurements (Jackson et al., 1981), such as in Equation 4:

$$CWSI = \frac{(T_s - T_a) - (T_s - T_a)_{ll}}{(T_s - T_a)_{ul} - (T_s - T_a)_{ll}} \quad (4)$$

where T_s is the crop canopy temperature, T_a is the air temperature, $(T_s - T_a)_{ll}$ is the non-water-stressed baseline, and $(T_s - T_a)_{ul}$ is the non-transpiring baseline.

Electrical Conductivity's Effect on Evapotranspiration and Yield

Soil salinity, quantified by EC_e, reduces ET_a rates and yield, though the means described on page 5. Ayers and Westcot (1985) determined that the point at which soil salinity begins to reduce maize yield is at an EC_e of 1.7 dS m^{-1}, and 50% yield loss is obtained at 5.9 dS m^{-1}. In the Arkansas River Valley, Morway and Gates (2012) found that an upstream region of the basin had fields with mean EC_e of 3.4 to 4.6 dS m^{-1} and yield loss of 6%. Meanwhile, a downstream region of the basin had fields with mean EC_e of 4.5 to 8.0 dS m^{-1} and yield loss of 17%.

More recently, using weighing lysimeters, Liu et al. (2022) found that the combined effects of water stress and soil salinity reduced maize ET_a between 64 and 83 percent relative to ET_c. Marino et al. (2020) found that in a pistachio orchard with EC_e values ranging from 2 to 14 dS m^{-1}, pistachio ET_a was reduced by about 30% over the growing season. Thus, it is reasonable to expect that reductions in crop water use or ET will correlate with increasing salinity levels.

Remotely Sensed Imagery

There are three levels of remotely sensed multispectral imagery data: ground-level data from handheld or ground-based sensors (i.e., proximal RS), airborne data from low-flying aircraft, and satellite-based data. All three platforms collect multispectral data that can be correlated to biophysical processes, including leaf area index (LAI), vegetative fraction (VF), and biomass (kg). Building off these successes, many researchers have attempted to relate remotely sensed data to soil salinity, water stress, and water use (Corwin and Scudiero, 2019; Aboelsoud et al., 2022), as found in the rest of this chapter.

Early research into remote sensing of crop health discovered that a simple way to detect areas of crop stress is by using fractional vegetation cover (Jackson, 1986; Zhang et al., 2021) or an analog vegetation index such as the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973; Fernández-Buces et al., 2006; Delavar et al., 2020). Looking at these processes more closely, salt stress decreases photosynthetic activity, stunting growth and causing wilting, which leads to changes in the reflectance properties of crops (Zhang et al., 2021), particularly by increasing reflectance in the visible bandwidths, and decreasing reflectance in the near-infrared (NIR) bandwidths (Scudiero et al., 2015).

Multi- and hyper-spectral imagery have emerged as a preferred method of estimating soil salinity due to its repeatability, scalability, and relatively inexpensive datasets (Davis et al., 2019; Cui et al., 2022). Common multispectral RS platforms include satellites such as the Moderate Resolution Imaging Spectroradiometer (MODIS) (Lobell et al., 2010; Zhang et al., 2015; Huete et al., 2002), Landsat (Metternicht and Zinck, 1997; Fernández-Buces et al., 2006; Scudiero et al., 2014, 2015; Bai et al., 2018;), Sentinel-2 (Bannari et al., 2018; Davis et al., 2019), and the Indian Remote Sensing satellite (IRS) (Abbas et al., 2013). Once an accurate model for

predicting EC has been developed for a region and crop type, this same model can be applied annually or more frequently. Remotely sensed data are scalable from handheld radiometers with 3 meter (m) or smaller footprint, to Unmanned Aerial Vehicles (UAVs) or drones that can cover dozens of hectares in a single flight at sub-meter resolution, to satellites that can cover entire watersheds with 30- to 500-m spatial pixel resolution. Depending on the needs of the user, multispectral imagery are inexpensive relative to other methods of detecting soil salinity. Many satellite datasets, including the popular Landsat satellite series are free to download from the USGS website. UAV or unmanned aerial system (UAS) companies offer their services throughout the country, and handheld multispectral cameras are significantly cheaper than EMI devices.

Mapping Soil Salinity using Remotely Sensed Imagery

There are three methods of inferring soil properties from satellite imagery: physical models relating spectra directly to soil properties; empirical models based on satellite and ground databases; and semi-empirical models based on a combination of empirical data and physical models.

Basic physical methods of remotely identifying salts include detection of the distinct white crusts found in extremely saline areas (Metternicht and Zinck, 2003; Abbas et al., 2013; Matinfar et al., 2013), or by measuring the soil backscattering coefficient, which varies based upon the dielectric properties of the soil (Jiang et al., 2019). However, these methods of detecting bare soil salinity usually only work in areas with $EC_e > 20$ dS m^{-1}, which causes immense yield loss in the majority of plants, and is rarely cropped in practice (Scudiero et al., 2015). These methods are also generally qualitative rather than quantitative since surface conditions do not always directly correlate with crop root zone salinity levels.

Empirical models of detecting soil salinity mostly involve testing the correlation of many RS spectral band combinations against some form of EC in order to find one that correlates the best with salinity or sodicity. Some studies have done this with satellite imagery in field settings (Metternicht and Zinck, 1997; Bouaziz et al., 2011; Aldabaa et al., 2015; Davis et al., 2019; Aboelsoud et al., 2022), while others utilize soil samples and lab spectrometry (Liu et al., 2015). Davis et al. (2019) compared the application of surface reflectance data from bands of the Landsat 8 Operational Land Imager (OLI) to the Sentinel-2 Multispectral Instrument (MSI). They found that both satellites can predict salinity greater than 4 dS m^{-1} on a single day with an RMSE of 1.15 and 1.17 dS m^{-1} and R^2 of 0.54 and 0.67 for OLI and MSI, respectively, with n = 100 pixel locations.

Whereas detection of surface salts only captures information in the top few centimeters of soil, RS-based vegetation indices relate to the soil properties throughout the crop root zone. Though somewhat empirical, past research has narrowed in on crop canopy reflectance as the most accurate quantitative method for remotely sensing soil salinity (Scudiero et al., 2014). Perhaps the most commonly used crop canopy reflectance index is NDVI, which has been implemented in many soil salinity models (Fernández-Buces et al., 2006; Khan & Abbas, 2013; Scudiero et al., 2014; Ivushkin et al., 2019; Zhu et al., 2021; Cui et al., 2022) The most difficult part of the crop canopy method is that crop reflectance is a means of monitoring overall vegetation health, rather than discriminating between different stress types (Scudiero et al., 2014). However, Scudiero et al. (2016) were able to predict soil EC_e in fields containing a variety of crops and within the Western San Joaquin Valley in California with a Root Mean Square Prediction Error (RMSPE) of 3.63 dS m^{-1} and an R^2 of 0.73. A comprehensive list of vegetation indices used to detect soil salinity can be found in Corwin and Scudiero (2019).

Other studies mix different data types together to predict EC_e. Ivushkin et al. (2019) used thermal data, vegetation indices and Light Detection and Ranging (LiDAR) measurements of crop height to predict EC_e, with an R^2 of 0.46 throughout their study region.

Many salinity prediction models use linear regression. Nonetheless, their forms change, based on whether the researcher is doing multiple linear regression (MLR), simple linear regression (SLR), least squares regression, or other regression types. Nonlinear regression methods include artificial neural networks (ANN), support vector machine regression (SVR) or random forest (RF). Cui et al. (2022) found that ANN and RF models outperformed MLR models. They were able to predict soil salt content under sunflower crops to within 8% by volume with R^2 greater than 0.7. Jiang et al. (2019) found that SVR outperformed ANN in the Yanqi Basin of China over numerous crops (including maize), with an RMSPE of 1.36 dS m^{-1} and an R^2 of 0.88. Their study is one of the few known studies to use ET as a model parameter.

Some studies recommend transforming EC with a natural log transformation. Lobell et al. (2010) found that vegetation indices such as EVI correlated more strongly with ln(EC) than EC without the transformation.

Advancing Modeling of Salinity Mapping

Although salinity modeling has come very far, EC models still often do not have enough accuracy to make field management decisions, even when calibrated to local conditions. It is widely accepted that soil salinity is difficult to predict using single-factor models due to the number of compounding variables (Jiang et al., 2019), so any new model creation should consider multivariate analysis.

Corwin and Scudiero (2019) summarized four challenges of using satellite imagery to detect soil salinity: (1) the spectral signature of a crop changes through phenological stages; (2) different crops have different spectral signatures; (3) other stressors such as nutrient deficiency and water stress create a similar spectral response; and (4) surface reflectance changes with different soil backgrounds. Because of these difficulties, many researchers focus on creating soil salinity prediction models that are calibrated to a specific crop in a specific region.

In order to better predict salinity at large spatial scales with enough accuracy to make field management decisions, it is necessary to find new biophysical factors that can increase the accuracy of salinity prediction models. As described in the section *Soil Salinity Effects*, soil salinity reduces ET_a. Although Jiang et al. (2019) created predictive models of soil salinity that included ET as one of the explanatory variables, there are few other instances of salinity modeling with ET as an independent variable.

There are numerous RS of ET_a models that are based on the land surface energy balance equation [Eq. (5)], including METRIC (Allen et al., 2007b, 2007a), SEBAL (Bastiaanssen et al., 1998a; Bastiaanssen et al., 1998b), and ReSET (Elhaddad et al., 2011). Since these models use satellite multispectral imagery that are used to calculate vegetation indices, RS of ET_a products can be used in tandem with vegetation indices for salinity mapping.

$$R_n = G_0 + H + LE \quad (5)$$

ReSET ET_a results have been verified in healthy crop conditions. The American Society of Civil Engineers (ASCE) Environmental and Water Resources Institute (EWRI) has a standard equation for calculating ET_r (Walter et al., 2005). ReSET-derived seasonal ET_a values were compared to the ASCE-EWRI ET_a with maize crop coefficients in the South Platte River basin in Colorado for the 2006 growing season. There, ReSET estimated seasonal ET_a within 5% of the

ASCE-EWRI ET_c estimates (Elhaddad et al., 2011) for fully irrigated (non-stressed) corn fields. Since the ASCE-EWRI method assumes ideal growing conditions, only healthy growth locations were chosen for this comparison. In the Palo Verde Irrigation District in California, the maximum annual ET for alfalfa fields estimated by ReSET was only 1.5% less than ET estimated from the Penman-Monteith reference ET obtained from a nearby California Irrigation Management Information System (CIMIS) weather station (Elhaddad et al., 2011).

Hypothesis and Objectives

It is necessary to find new explanatory variables that can improve the accuracy of remotely sensed salinity mapping models for irrigated crops in semi-arid regions, particularly those with saline bedrock and high water tables, where salinity accumulation is ongoing and the effects of soil moisture make it more difficult to predict salinity (Liu et al., 2015). Repeated EC measurements of these areas prone to salinization need to be available to farm managers for a cost-effective price. By improving the accuracy of EC mapping using RS models, water and farm managers can view salinity changes in real time and make better management decisions for salinity mitigation, water efficiency and yield maximization.

ET has been used in at least one salinity model study (Jiang et al., 2019) but results have not been as rigorously tested as have been many other vegetation indices or in combination with other potential explanatory variables. Therefore, this research includes RS-based ET_a and CWSI values of maize obtained from the ReSET model along with common vegetation indices such as NDVI and CRSI in the mapping of EC for maize. Since crop spectral parameters change throughout the growth season (Cui et al., 2022), this research evaluates each growth stage individually to attempt to identify which growth stage may provide the best information for

salinity prediction models. It is expected that the high water table may interfere with prediction capabilities, as the water table level creates a complex array of stressors and benefits to the crop. Indeed, Liu et al. (2015) found that the predictive capability of the visible and NIR spectra increases as the influences of moisture are removed. This research hypothesis claims that ET_a or CWSI in combination with vegetation indices will improve soil salinity mapping modeling on maize fields in contrast to models based solely on vegetation indices. Specific objectives are as follows:

- Determine the correlation between ET_a and three forms of EC: EC_a, EC_e, and EC_{eg};
- Create linear regression models to estimate EC in irrigated maize using multiple explanatory variable (e.g., ET_a rates, CWSI and standard vegetation indices);
- Assess the usefulness of adding ET_a rates to linear regression models made up of vegetation indices in predicting soil salinity in surface irrigated maize, for each growth stage as well as the total season.

CHAPTER 2: METHODS

Study Location

Lower Arkansas River Valley

The Lower Arkansas River Valley (LARV) (Figure 3) is an area of intense irrigated agriculture in Southeast Colorado. LARV geology is comprised of a series of sedimentary formations from the Cambrian to Tertiary age, and its alluvial aquifer has a strong hydraulic connection to the Arkansas River and its tributaries (Gates et al., 2016). Most of the water in the Arkansas River comes from mountain snowpack, and a trans-basin diversion run by the U.S. Bureau of Reclamation called the Frying Pan – Arkansas Project brings supplemental water into the basin. Releases from Pueblo Reservoir (Figure 3) have a large impact on the flow regime in the LARV. There are about 109,000 irrigated hectares in the LARV, spread out of 14,000 separate fields (Gates et al., 2012). Most fields are irrigated using flood irrigation methods, while in 2014, about

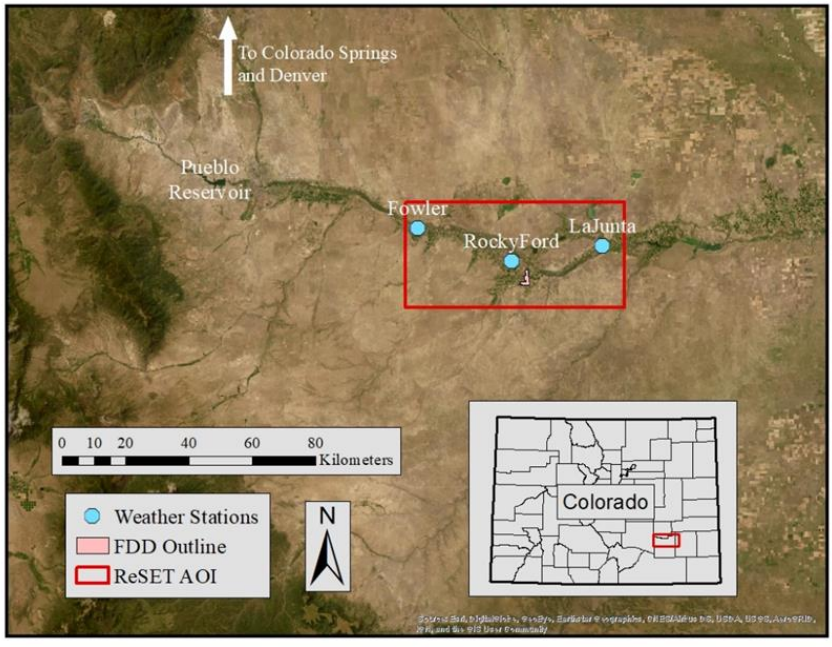

Figure 3: The Lower Arkansas River Valley (LARV) beginning at Pueblo Reservoir and extending through the study area. The Arkansas River flows from West to East.

21 percent of fields were irrigated with sprinklers or drip lines (Osborn et al., 2017). Inefficient irrigation and canal seepage have contributed to an increase in the height of the water table, pushing salts upward into the root zone of many crops (Gates et al., 2012, 2016; Osborn et al., 2017). Many fields within the LARV are negatively affected by salts, with yield losses estimated to range from 0 to 89%, averaging around 19% throughout the region (Burkhalter and Gates, 2005; Morway and Gates, 2012). Furthermore, many portions of the LARV contain naturally-occurring gypsum ($CaSO_4$), which is relatively insoluble in field conditions, but may be detected in the soil saturation paste extract when testing for EC_e. This makes EC_e values from fields with gypsum difficult to interpret without taking into account the presence of gypsum.

Fairmont Drainage District

The Fairmont Drainage District (FDD) is a collection of fields within the LARV and near the town of Swink (Figure 4). This area is delineated by the fields with an underground series of tile drains that were installed in the early twentieth century (Figure 5). The Rocky Ford Colorado Agricultural Meteorological Network (CoAgMet) (CSU, 2023) weather station approximately 8 km from the FDD has registered an average of 101.5 mm of rain per year over a recent 19 year period (2002 – 2020). Five maize fields within the FDD make up the study area. Similar to other areas in the FDD and throughout the LARV, these fields contain gypsum and are underlain by a high water table. The soils range from silty clay loam to clay loam. Tillage practices and cropping density are consistent throughout FDD maize fields.

The field of most interest in the FDD is Muth 2 (M2). This field has a zone of very high soil salinity (EC_e >10 dS m^{-1}) that is approximately 700 m^2 and has the largest range of salinity of any field. This is important for salinity prediction modeling purposes, as the large range will be useful for calibrating a model that can predict a wide range of EC values. M2 was planted with

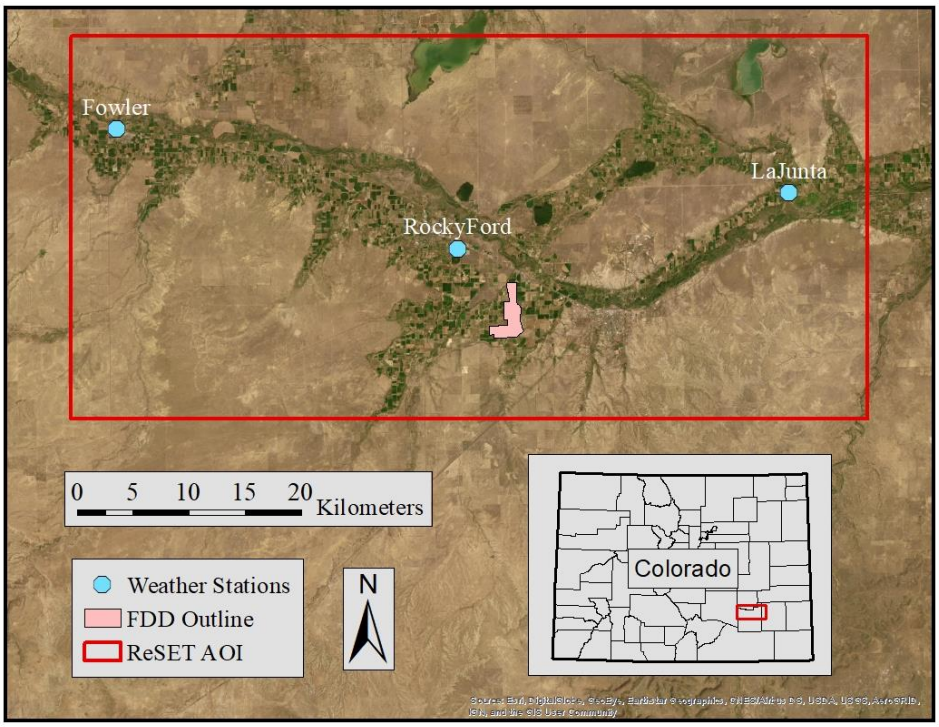

Figure 4: The Fairmont Drainage District within Colorado, and the ReSET Area of Interest (AOI) along with the three weather stations used for ReSET calculation.

maize in both 2018 and 2019. In 2018, M2 was side-dressed on June 18, and irrigated on May 5 (before planting), June 21-26, July 12, and July 20. In 2019, M2 was irrigated on July 23 and August 31. The other fields of focus were Muth 3 (M3), Muth 6 (M6), Muth 8 (M8), and Golden 2 (2019 only) (Figure 5). Although many other fields were surveyed, some did not have a large enough range of salinity to accurately calibrate the soil EC measurements.

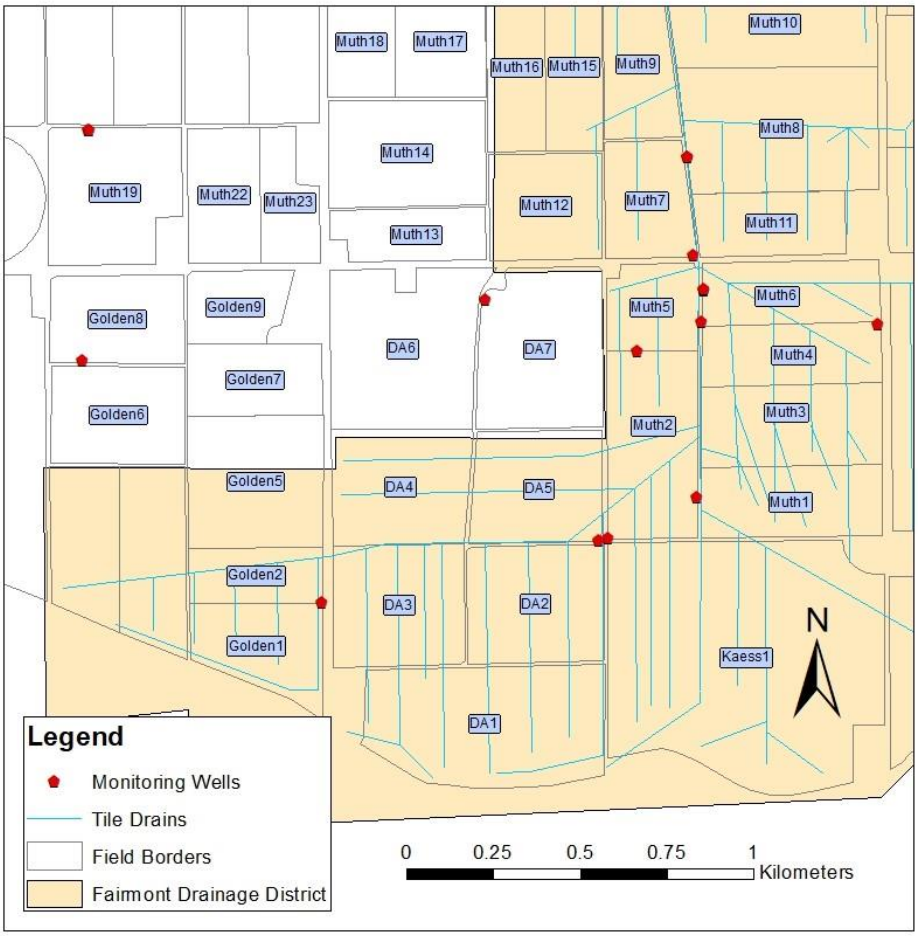

Figure 5: Field names in the Fairmont Drainage District and surrounding area.

Soil Properties

Soil Electrical Conductivity

Electromagnetic induction (EMI) techniques were used to measure geospatial apparent soil electrical conductivity (EC_a) for each field via an EM38-MK2 electrical conductivity meter (Geonics Ltd., Mississaugua, Ontario, Canada). The surveys of each field were conducted over the periods of May – October 2018 and May – July 2019, which are primarily during the maize growing seasons. A total of 18 fields were surveyed in 2018 (Figure 6). Of those 18 fields, 6 were again surveyed in 2019, plus 6 additional fields (Figure 7).

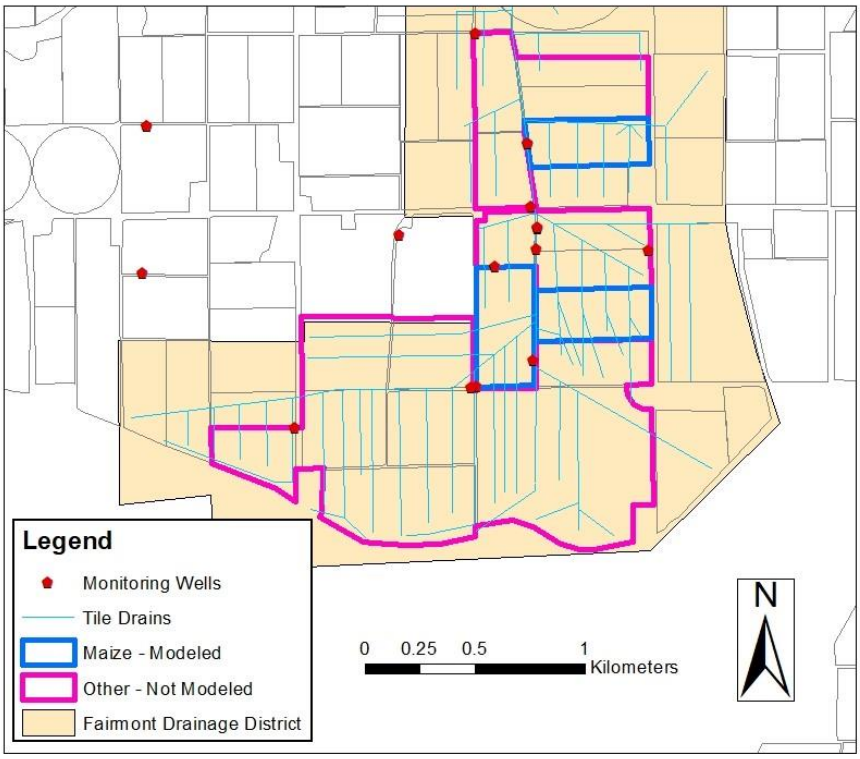

Figure 6: A map of all fields surveyed for EC in 2018, with the locations of monitoring wells and tile drains.

The protocols for collecting EC_a data in each field follow the guidelines from Corwin and Scudiero (2020). Briefly, metadata, including the field name, row orientation, and crop type, were collected to categorize each field, EMI survey strategies were created to avoid edge effects, soil water content was checked to estimate if field soil water levels were appropriate for surveyance (> 70% field capacity) by digging a hole and feeling the soil by hand, and then the survey was performed. Felds were surveyed by foot using an EM38-MK2 (Geonics Ltd., Mississaugua, Ontario, Canada), Trimble Global Navigation Satellite System (GNSS) receiver with a Differential Global Positioning System (DGPS) correction subscription (Sunnyvale, CA, USA), and Juniper Allegro CX handheld computer (Logan, UT, USA). This system was used to gather a continuous stream of EC_a measurements, recording one reading every four seconds,

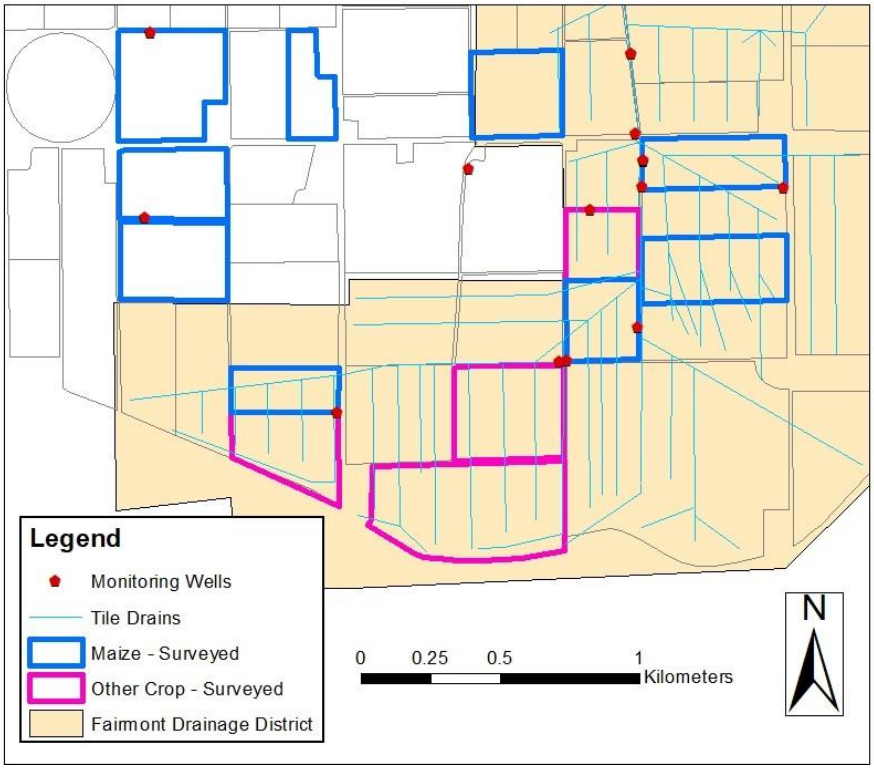

Figure 7: A map of all fields surveyed for EC in 2019, with the locations of monitoring wells and tile drains.

averaging to approximately 1,000 locations of EC_a measurement in each field.

EM38-derived EC_a measurements serve only as a surrogate to characterize the spatial variation of soil properties that influence EC within a field (i.e., saturation percentage (SP), water content (θ), bulk density (ρb), organic matter (OM), cation exchange capacity (CEC), clay content and mineralogy, temperature, and total dissolved solids (United States Salinity Laboratory Staff, 1954; Stiven and Khan, 1966; Rhoades et al., 1990; Corwin and Lesch, 2005; Gavlak et al., 2005). As a result, EMI-based soil survey products must be calibrated to a desired soil property using directed soil sampling at discrete locations. To choose these sampling locations, the EC_a data for each field in the FDD were entered into the Electromagnetic Sampling Analysis and Prediction model (ESAP, ver. 2.3.5, USDA ARS) to obtain the model-based sampling design locations. ESAP uses a response surface sampling design strategy which creates a 3-D surface of

the EC_a measurements and based on the range and variation, selects locations that characterize the EC_a variation while maximizing the distances between adjacent sampling locations (Lesch et al., 2002). The resulting sampling designs resulted in a minimum of 2 and a maximum of 12 soil sample locations per field, totaling 75 locations in the district (Figure 6) in 2018.

At each location, soil samples were collected using an 8 cm diameter soil auger at 0.3, 0.6, 0.9, and 1.2 m depths, with the intention of capturing the maize root zone. In the laboratory, a saturation extract of each soil sample was prepared and the EC_e was measured using the method presented in Rhoades (1996). The water content of the field at the time of extraction, measured from a separate portion of the soil sample, and the water content of the saturated paste were determined gravimetrically. EC_e is then calculated by calibrating the EC_a readings to the soil paste extract EC_e values. Analysis of Covariance (ANOCOVA) modeling was used to transform EC_a into EC_e.

Gypsum Correction

Sparingly soluble gypsum, which is rarely dissolved in the soil under field conditions, dissolves during laboratory preparation of saturated paste extract, inflating estimation of salinity when measuring EC_e (Brown, 2021). Because of this, a transformation developed by Brown (2021) was used to correct EC_e values, converting them to EC_{eg}:

$$EC_{eg} = 0.9131(EC_e) - 0.7256 \quad if\ EC_e > 3.08, \qquad (6)$$

$$= EC_e \qquad if\ EC_e < 3.08$$

Once this transformation was applied, three different versions of EC were available for use in linear regression analysis: EC_a, EC_e, and EC_{eg}. It was hypothesized that each version of EC would have a different level of correlation with ET, vegetation indices, and other explanatory variables used in a multi-variate regression approach.

Sodium Adsorption Ration

The Sodium Adsorption Ratio (SAR), a proxy for sodicity, when paired with EC, quantifies the level of flocculation or dispersion in the soil. The level of flocculation or dispersion varies based on a complex set of soil characteristics including the types of ions present, their ionic strength, adsorption of the ions, pH, and temperature (Sumner, 1993). In 2021, SAR was calculated at 24 locations throughout M2, M3, M6, and M8. SAR data provide a general idea of sodicity levels at those locations. Unfortunately, SAR could not be used in the modeling portion of this analysis because we did not have a large enough spatial coverage (only 24 sample sites) to interpolate between locations and obtain a value for each pixel in the RS imagery. The linear regression analysis was conducted using 350 locations.

Remote Sensing

When plants experience stress, their photosynthetic activity decreases, causing an increase in visible (e.g., red band) reflectance and reduced near-infrared band reflectance (Scudiero et al., 2015; Zhang et al., 2021). Crops under stress also wilt, decreasing visible leaf area and increasing the percent reflectance from the soil. Therefore, vegetation indices, arithmetic combinations of satellite-sensed spectral bands that measure these properties, change as a crop experiences stress, and can be used as an indicator of crop stress.

Here, we used Landsat 7 and Landsat 8 satellite data over the two growing seasons of 2018 and 2019 to calculate vegetation indices and actual evapotranspiration (ET_a) rates. Images were downloaded from EarthExplorer (USGS, 2023) and visually inspected for cloud cover over the FDD. Images with cloud cover exceeding about ten percent of the FDD were discarded. Level 2 images, which are corrected for atmospheric effects and thus represent surface reflectance values, were used for calculation of vegetation indices. Level 1 images were used for input to the

ReSET model for ET_a calculation, described on page 30. Landsat 7 and 8 have 16-day orbital return intervals offset by 8 days, resulting in one satellite pass every 8 days. Landsat 7 is equipped with the Enhanced Thematic Mapper Plus (ETM+) sensor which measures light in the visible and near infrared spectra at 30 m spatial resolution, plus a thermal band at 60 m spatial resolution. Landsat 8 includes both the Operational Land Imager (OLI) sensor measuring visible and near infrared light, and the Thermal InfraRed Sensor (TIRS), measuring thermal wavelengths. OLI has a resolution of 30 m, except for the panchromatic band at 15 m, while TIRS has a resolution of 100 m, resampled to 30 m during USGS postprocessing.

Satellite Image Dates

There were 16 usable satellite image dates in 2018: May 12, 20 and 28; June 5, 13, and 21; July 7, 15, 23 and 31; August 16 and 24; September 17; and October 3, 19 and 27. 5. The other 2018 growing season images were collected on June 29, August 8, September 1, 9 and 25, and October 11, and had too much cloud cover to be included in the analysis.

There were 15 usable image dates in 2019: May 15 and 31; June 24; July 2, 18 and 26; August 3, 11 and 19; September 4, 12 and 20; and October 6, 14 and 22. The other 2019 growing season images, collected on May 23, June 8 and 16, July 10, August 27, and September 28, had too much cloud cover to be included in the analysis.

Vegetation Indices

Vegetation indices were calculated using ESRI ArcMap 10.5.1 and Landsat 7 and 8 Level 2 imagery. The vegetation indices used in this study were the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973; Fernández-Buces et al., 2006; Delavar et al., 2020; Zhu et al., 2021), and the Canopy Response to Salinity Index (CRSI) (Scudiero et al., 2015). NDVI was included because it is arguably the most commonly understood and used vegetation index, while

CRSI was chosen because published literature has shown it to have potential to be the best vegetation index for predicting soil root zone salinity (Corwin and Scudiero, 2019). Many studies have shown that these two indices are highly sensitive to soil salt content (Qi et al., 2020; Wei et al., 2020). The Optimized Soil-Adjusted Vegetation Index (OSAVI) (Rondeaux et al., 1996) was originally included but later disregarded because it did not differ enough from NDVI to expand the prediction space of linear models.

ReSET-Raster

Remote Sensing of EvapoTranspiration-Raster (ReSET-Raster or simply ReSET) is an ESRI ArcGIS toolbox extension that is run on ArcMap 10.3.1. ReSET calculates daily ET_a in mm day^{-1} using a remotely sensed energy balance, and a network of weather station-based raster variables developed at Colorado State University (Elhaddad et al., 2011). ReSET can be run using data from Landsat 5, 7 or 8. ReSET is based on the basic surface energy balance equation (SEB):

$$R_n = G_0 + H + LE \quad (7)$$

where R_n = net radiation, G_0 = soil heat flux, H = sensible heat flux, and LE = latent heat flux, with all terms having the units of W m^{-2}. ReSET uses the same theoretical basis as its two predecessors, Mapping Evapotranspiration with Internalized Calibration (METRIC) (Allen et al., 2007a, 2007b) and SEBAL (Bastiaanssen et al., 1998a; Bastiaanssen et al., 1998b). These SEB models are based on the rationale that ET represents a change in the state of water, from liquid to vapor. This vaporization process requires available energy (net radiation minus the energy into the ground) in the environment (Elhaddad et al., 2011). The basic assumption for H is that the air temperature gradient between the near-ground surface and the adjacent air layer changes linearly with the ground surface temperature. Based on this, two extreme pixels are chosen (in an

agricultural field) to represent the maximum and minimum ET. At the "hottest" pixel in the satellite image, ET is assumed to be 0, so LE is 0 and $H = R_n - G$. At the "coldest" pixel, and in the calibrated mode (METRIC), ET is assumed to be $1.05 \cdot ET_r$, and LE is calculated as follows:

$$H = (R_n - G) - 1.05\lambda_v ET_r \quad (8)$$

where λ_v is the latent heat of vaporization (J kg^{-1}). The difference between SEBAL or METRIC and ReSET is that ReSET has the ability to create rasters of weather variables from multiple weather stations through interpolation, which creates multiple dT (temperature difference with elevation) functions spread throughout the area of interest (AOI). The AOI is input by the user as a polygon shapefile.

Elhaddad et al. (2011) compared ReSET-derived seasonal ET_a values over maize fields to ASCE-EWRI Penman-Monteith (PM)-derived ET_a seasonal estimates for healthy maize in the South Platte River basin in Colorado for the 2006 growing season, using Equation 9:

$$ET_a = K_c \frac{\Delta(R_n - G) + \rho_a c_p \frac{(e_s - e_a)}{r_a}}{\Delta + \gamma(1 + \frac{r_s}{r_a})} \quad (9)$$

where K_c is the crop coefficient, R_n is the net radiation, G is the soil health flux, ρ_a is the mean air density at constant pressure, c_p is the specific heat of the air, Δ represents the slope of the saturation vapor pressure-temperature relationship, γ is the psychrometric constant, and r_s and r_a are the bulk surface and aerodynamic resistances. ReSET estimated seasonal ET_a to within 5% of the ASCE-EWRI PM ET estimates. Since the ASCE-EWRI PM method assumes ideal growing conditions, only healthy growth locations were chosen for this comparison. In the Palo Verde Irrigation District in California, the maximum annual ET_a for alfalfa fields estimated by ReSET was only 1.5% less than ET_a estimated from the Penman-Monteith ET_r obtained from a nearby California Irrigation Management Information System (CIMIS) weather station. Because these

studies compared ReSET-derived ET_a in healthy crops, there may be inaccuracies in ReSET-derived ET_a when calculated in areas with known salinity stress, waterlogging, or other stressors. In this study, ReSET was applied for the 2018 and 2019 growing seasons, with a total of 14 overpass dates in 2018 and 15 overpass dates in 2019, as stated in section 1.3.2. ReSET was run in the Auto ReSET mode, where the coldest and hottest pixels were selected automatically by the model, with non-interpolated weather parameters, and in the calibrated mode, where ET_r from weather stations is used to set the maximum ET value for the AOI. This essentially makes the model function as METRIC. The inputs to ReSET include Landsat 7 or 8 Level 1 imagery and the accompanying metadata file, a digital elevation model (ASTER USGS), the daily wind run (with a minimum of 2.5 m s^{-1}), and both hourly and daily ET_r. Wind run and ET_r were taken from CoAgMet weather stations in Fowler (FWL01), Rocky Ford (RFD01), and La Junta (LJT01). Weather data were averaged across the three weather stations to get a mean value for this Arkansas River Basin. To achieve model convergence, a daily wind run of 2.5 m s^{-1} was used whenever the daily wind run was less than 2.5 m s^{-1}. Initial surface albedo was left at 0.025, per model recommendations.

Near-Surface Remote Sensing

Infrared Thermometers

Apogee SI-111 infrared thermometers (IRTs) were installed at the datalogger locations in 2018 (three in M2 and one in M3) (Figure 8) and 2019 (Figure 9). The IRT sensor measured crop canopy temperature every 1 hour, on the hour. The IRTs were placed facing northeast at 45 degrees below the horizontal and 1 m above the crop canopy to minimize the effects of shadows

and soil background. IRTs were raised at least once every 10 days to ensure they were 1 m above the crop canopy as the crops grew.

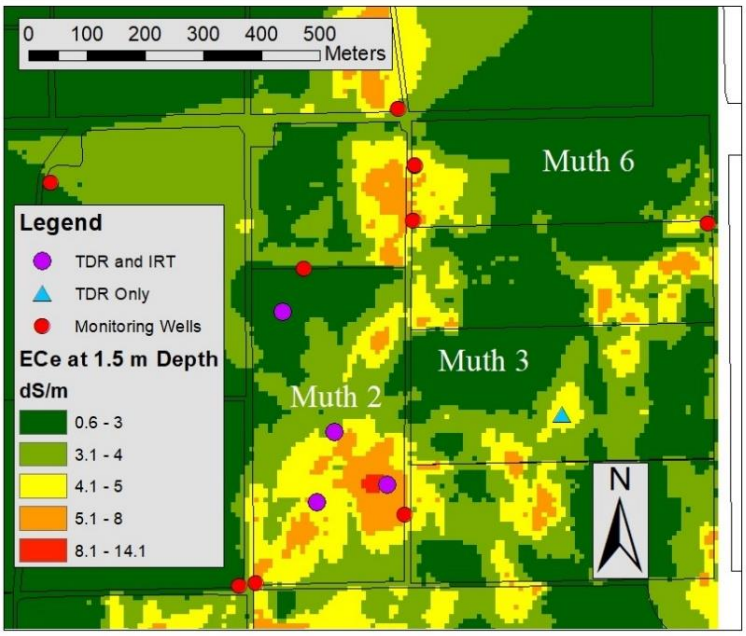

Figure 8: Field-level data collection locations in the Fairmont Drainage District (FDD) in 2018, laid over electrical conductivity of the paste extract, EC_e.

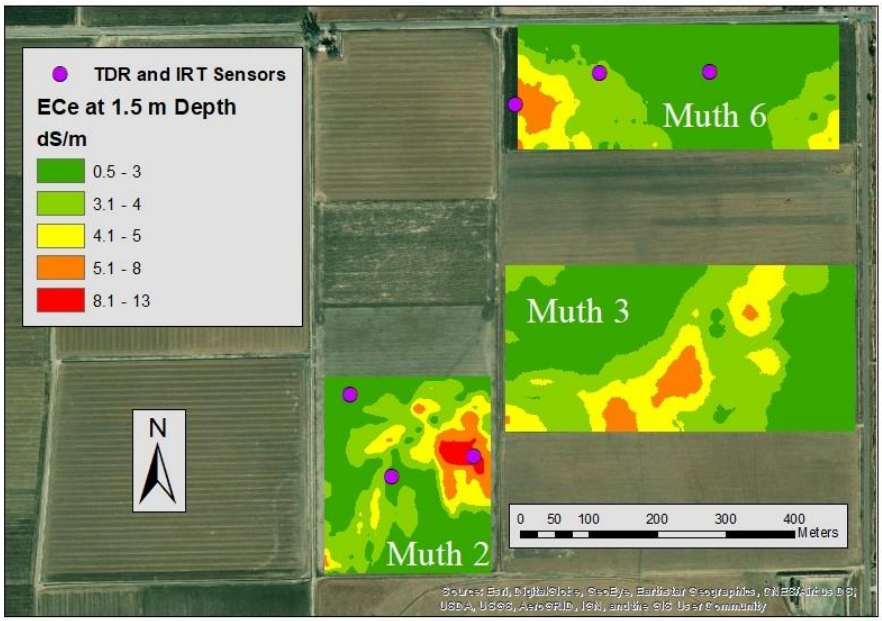

Figure 9: Field-level data collection locations in the Fairmont Drainage District in 2019, laid over electrical conductivity of the paste extract, EC_e.

Temperature data were used to compute the crop water stress index (CWSI) [Eq. (16)], which is described in the *Crop Water Stress Index* section. The IRT-derived CWSI was compared to RS-derived CWSI [Eq. (15)] to evaluate which CWSI product to use to evaluate water stress in each growth stage. Whenever comparing IRT-derived products to satellite RS data, the 11:00 MDT data point, which is the time closest to the satellite overpass, was used to reduce differences due to readings at different times of the day (Lakshmi and Zehrfuhs, 2002).

Applied Water and Water Stress

Total Applied Water

Irrigation water is delivered to fields in the FDD via the Catlin Canal, which diverts water off the Arkansas River. Concrete-lined head ditches distribute this water from the Catlin Canal to individual fields, and siphon tubes move the water from the ditch to alternating furrows in the field. Surface runoff is collected in an unlined tail ditch, where it either seeps into the ground or discharges to Timpas Creek, a tributary of the Arkansas River. M2 applied irrigation flows were measured using trapezoidal flumes and drainage flows with 3-cfs ramp flumes, each equipped with HOBO water level loggers. Total applied water was calculated for M2 in 2018 as the sum of precipitation at the Rocky Ford CoAgMet station (RFD01) and on-field irrigation.

Soil Water Balance

Periods of crop water stress were determined using a soil water balance within a Microsoft Excel spreadsheet. This spreadsheet model calculates soil moisture depletion (SMD) [Eq. (10)] on a daily timestep for the 2018 and 2019 growing seasons (Aboitiz et al., 1986). SMD on the jth day is the sum of the SMD from the previous day (SMD_{j-1}) and daily ET_{aj}, minus the total applied water (TAW_j) and upflux (U_j) (Jensen and Allen, 2016).

$$SMD_j = SMD_{j-1} + ET_{aj} - TAW_j - U_j \quad (10)$$

TAW is defined as follows:

$$TAW = P_g + I_g - DP - RO \quad (11)$$

where P_g is the gross precipitation, I_g is the gross irrigation, DP is deep percolation and RO is runoff. DP was assumed 0 in this study, and RO was assumed 0 on days without an irrigation. U is defined according to Jensen and Allen (2016) as:

$$\begin{aligned} U &= U_{max} && when\ W < W_s, \\ &= 0 && when\ W > W_c \\ &= U_{max}\frac{W_c - W}{W_c - W_s} && when\ Ws \leq W \leq W_c \end{aligned} \quad (12)$$

where W is the water depth in the soil root zone, W_c is the critical water depth (the water depth at which upflux > 0), and W_s is the steady water storage (the water storage at which U_j no longer increases).

If the total SMD at the end of the day is greater than the maximum allowed depletion (MaxAD) [Eq. (13)], then the crop is stressed. MaxAD is computed as:

$$MaxAD = \frac{ManAD * AW * Rz_j}{100} \quad (13)$$

where ManAD is the Management Allowed Depletion (percent), AW is the available moisture (a constant), and Rz_j is the root zone depth on the jth day (m). This soil water balance approach starts when SMD is 0, such as after a large rainfall or irrigation event. ReSET-derived ET_a was used in this analysis. On dates that didn't have a satellite overpass and thus no ET_a value, ET_a was calculated by the following method: ET_c, which is K_c multiplied by ET_r, on that date was multiplied by the Evaporative Fraction (EF) from the closest satellite overpass date to determine ET_a on that date:

$$EF = \frac{ET_a}{ET_c} = \frac{ET_a}{ET_r \cdot K_c} \quad (14)$$

Crop Water Stress Index

The Crop Water Stress Index (CWSI) (Jackson et al., 1981; Jackson et al., 1988) was calculated using two methods. The first method uses EF derived from ReSET ET_a [Eq. (8)] (Brutsaert, 1992) and is abbreviated $CWSI_{sat}$:

$$CWSI_{sat} = 1 - EF \quad (15)$$

The second method, taken from Taghvaeian et al. (2014), is a ratio of the difference between canopy temperature, T_c, and air temperature, T_a, minus the lower limit (ll) [Eq. (17)] of the temperature difference (i.e. a well-watered crop), over the difference between the upper limit (ul) [Eq. (18)] of the temperature difference (wilting crop) and lower limit, and is abbreviated $CWSI_{IRT}$:

$$CWSI_{IRT} = \frac{(T_c - T_a) - (T_c - T_a)_{ll}}{(T_c - T_a)_{ul} - (T_c - T_a)_{ll}} \quad (16)$$

$$(T_c - T_a)_{ll} = -1.99(VPD) + 3.04 \quad (17)$$

$$(T_c - T_a)_{ul} = -1.99(VPG) + 3.04 \quad (18)$$

VPD is the vapor pressure deficit and VPG is the vapor pressure gradient. T_c was taken from the IRTs, and T_a was obtained from the Rocky Ford CoAgMET station. When CWSI gets above 0.299, the maize crop was assumed to be water-stressed (Qin et al., 2021) and should be irrigated to avoid permanent yield loss.

Because of the 8 km distance to the weather station, and the calibration of the upper and lower temperature differences, computed CWSI was sometimes out of its range of 0 to 1. If CWSI was ever computed less than 0, the value was set to 0. If CWSI was ever computed greater than 1, the value was set to 1. In 2019, the upper limit temperature difference $(T_c - T_a)_{ul}$ was computed above 5 in about 10% of the data, which is unreasonable, and was corrected down to a value of 5.

This correction is standard practice in maize fields. Following the work of Qin et al. (2021), the maize was determined to be stressed when CWSI > 0.299.

Seasonal Crop Water Stress

To determine when the crop was experiencing stress, and thus properly irrigate crops to avoid over-watering the crop and exacerbating salinity accumulation, the percent of dates that a crop was stressed within a given growth stage, as well as through the whole season, was calculated. If either the spreadsheet SMD/MAD method or the CWSI method indicated crop stress, then the crop was assumed stressed. The number of dates that the crop was stressed in a given growth stage was summed in order to get an idea of how much water stress was present during each growth stage. This information was compared to the growth stages in which salinity modeling was most successful in order to get an idea of how water stress affects the ability to predict soil salinity.

Water Table Depth

Water table depth has many effects on crop health and yield. Although crops need water, a water table depth that is near or extends into the root zone may lead to waterlogging and associated root rot (Food and Agriculture Organization of the United Nations, 2020). Furthermore, a high water table will contribute to further salinization of soils, as it brings saline groundwater by upflux into the root zone, at which time the crop removes the water, leaving most of the salt in the soil to accumulate.

Water table depth measurements were collected in 11 (2018) or 12 (2019) observation wells at 15-minute intervals by HOBO pressure transducers (Figure 5). These instruments were installed in the FDD by previous graduate students in 2015, so there is a longer period of manual readings for calibration of the automated readings than the scope of this study. Manual readings of depth

to water table were taken at each site on ten days in May, June and September 2015, March, May, June and August 2016, and May 24, 2018. Since water table depth data are only available at 11 (2018) or 12 (2019) points around the edges of fields, water table depth had to be interpolated to determine water table depth within the fields. For satellite overpass dates, estimates of water table depth measurements were interpolated using the cubist kriging method in ArcGIS, and the interpolated water depth values were resampled to the 30-m pixel size. Water table depth data were used to help interpret model results by classifying whether the water table affected the root zone of the crops.

Crop Yield

Crop yield data were collected at points in two fields in 2018 and 8 fields in 2019. To determine the locations, EC_a data and a sampling density were input to ESAP, which gave 38 optimized sampling points. At each location, the maize ears of three consecutive plants were collected. Later, kernel weight to the nearest tenth of a gram was obtained using a scale. To convert these point measurements into a raster layer, crop yield measurements were interpolated in ArcGIS. In 2018, yield data were interpolated using kriging with a 1.9-m pixel size and an 80-m search radius. In 2019, yield data were interpolated using kriging with a 4-m pixel size and 80-m search radius, due to the minimal density of points. This yield raster was then resampled to the 30-m Landsat pixel using the cubist method (Quinlan, 1992).

Data Analysis and Statistics

Seasonal Data

Since VI, ET, and CWSI values all change throughout the growing season, while salinity remains relatively constant, VI, ET, CWSI-to-salinity relationships will vary with time in the growing season when the satellite passes over the fields. To account for the dynamic nature of the explanatory variables being considered, and building off the work by Lobell et al. (2010), Scudiero et al. (2015), and Zhang et al. (2015), summed indices over standard periods of time were developed to create a standardized vegetation index that does not vary based on when in the growth season the satellite flies over the fields.

These integrated metrics were created in two ways. In one method, called the interpolated method, VI and CWSI values were linearly interpolated between satellite overpass dates and then summed over the 170-day growing period to obtain a seasonal total metric, as well as summed over each growth stage (initial, development, middle, and late), for a total of five values. Values of ET_a between dates were calculated by multiplying the ET_r on that date by the EF_r ($EF_r = ET_a ET_r^{-1}$) on the closest satellite overpass date. The growth stage lengths outlined in Hoffman et al. (2007), Table 8.3 of 30, 40, 50 and 50 days were used for the initial, development, middle and late growth stages, respectively, for a total growing season of 170 days. In the second method, metrics were not interpolated between satellite dates, so only the Landsat-derived values on the satellite overpass dates were included in the integration. This is called the non-interpolated method.

In order to remove compounding crop stressors, the methods outlined in Scudiero et al. (2014, 2015) for using multiple years of data in a single analysis were followed. The idea behind this method is that salinity changes more gradually than other stressors such as drought, nutrient, pest

or weed stressors. Therefore, if we use the maximum or 'healthiest' value for a given variable at each location, we will be isolating soil salinity stress from more temporally-variant stressors. In this method, the maximum value of an index over a given time period, be it NDVI, CRSI, ET_a, or CWSI is chosen to represent that location. For example, when working in the development stage, the maximum development stage NDVI at a pixel in either 2018 or 2019 would be chosen to represent the development stage NDVI for that location. Once this is done for each index, this dataset would then be regressed against salinity metrics.

A paired t-test was performed on the 2018 and 2019 EC_e datasets, which showed that they were statistically different at a significance level of $p < 0.05$. Thus, combined multi-year datasets were not used because the changing salinity values meant that a single salinity value could not be used for multiple years of model input data.

Correlation Coefficients

The Pearson's correlation coefficient r is commonly used for model parameter selection (Aboelsoud et al., 2022; Cui et al., 2022), Eq. (19):

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (19)$$

where x and y are two random variables, and i is the index within the dataset, ranging from 1 to n, the number of points in the dataset. Here, r between NDVI, CRSI and ET_a and EC_e was calculated for each satellite date. The results, as well as averages by date and averages by VI are found in Table 5 and Table 6. Values of r were also calculated for the indices summed over each growth stage. These results are found in Table 7 and Appendix A. Values of r were used to determine which EC variable (EC_a, EC_e, and EC_{eg}) and which growth stage (initial, development, middle, late, or total season) should be used in the subsequent modeling analysis, by choosing the pairs that had higher r values and thus stronger linear relationships to each other (Table 1).

Test of Data Distribution Normality

In order to use ordinary lease squares (OLS) regression, the dependent variable distribution must be normal. Using the EC-type and correlating growth stage pairs from the correlation coefficient analysis, ten EC datasets were checked for normality by creating a quantile-quantile (q-q) plot: EC_a and EC_{eg} for 2018, and EC_a, EC_e, and EC_{eg} for 2019, and the natural log transformation of those five datasets. Log transformation was analyzed instead of removing outliers because the outliers were often the high EC values, and the purpose of this study was to locate areas of high salinity. A q-q plot (Marden, 2004) projects the quantiles of a normal distribution on one axis and the quantiles of the EC distribution on the other axis. If the dataset is normal, the points will be spread along the 1:1 line. If the points deviate significantly from the 1:1 line, the dataset is not normal. These normality analyses were confirmed by plotting a histogram of the EC dataset. Combinations of the year and EC type were disregarded if they were not normal.

All independent variables were also log-transformed to see if certain variables lead to stronger model predictive capability as log-transformed variables rather than with no transformation.

Model Setup

An array of linear regression models to predict EC were created in RStudio version 2021.09.0 and R version 4.1.2 for the growth stage and EC pairs found via the correlation coefficient analysis explained in the *Correlation Coefficients* section (Table 1). Multiple Linear Regression (MLR) is a linear regression method in which the optimal combination of multiple independent variables are used together to predict the dependent variable (Cui et al., 2022). MLR is commonly used in multi-factor regression analysis. For each EC and growth stage pair, three models were run using the lm() function in RStudio. The first model is the 'base scenario,' and had only NDVI and CRSI available for selection as explanatory variables. The second model

included all variables available for selection (i.e., NDVI, CRSI, ET_a and CWSI), and the third model had NDVI, CRSI and either CWSI or ET_a available for selection. CRSI or ET_a was chosen based on which variable appeared more frequently and/or in the higher-ranking model when all variables were available for selection. Within each of those models, the software iterated through all the combinations of single variables and their interactions, and the equations were ranked by Akaike Information Criterion (AIC) (Cavanaugh and Neath, 2019) using the dredge function. AIC is a mathematical method for evaluating how well a model fits the independent variable data and penalizes a model that uses more terms to get to the good fit. AIC results in a single value for each model that considers the number of independent variables used to build the model and the maximum likelihood estimate of the model. Thus, AIC chooses the model that explains the most variation using the fewest independent variables, limiting redundancy in explanatory variables. For each linear regression, the model with the lowest AIC was chosen.

All models were run in two different ways: with all variables log-transformed and with no variables log-transformed. The best-performing models were re-run using a combination of log-transformed and not-transformed variables. The Pearson r values were used to determine whether to use the not-transformed or log-transformed version of each variable, by choosing the variable with the higher r value. Finally, the top model in each year (2018 and 2019) was chosen for a piecewise regression analysis, with the data broken up into two pieces based on the type of EC chosen (EC_a, EC_e, or EC_{eg}, log-transformed or without transformation).

Table 1: EC and Growth Stage Combinations used in the modeling analysis.

Year	Growth Stage	EC Type
2018	Initial	$\ln(EC_a)$
	Development*	$\ln(EC_{eg})$*
	Mid	EC_{eg}
	Mid	$\ln(EC_{eg})$
	Late	$\ln(EC_a)$
	Total	EC_{eg}
2019	Mid	$\ln(EC_e)$
	Mid	$\ln(EC_{eg})$
	Development	EC_a
	Mid	EC_e
	Mid*	EC_{eg}*
	Total	$\ln(EC_e)$

**Growth Stage and EC Type Pair was used in piecewise analysis*

Model Evaluation

Root Mean Square Error (RMSE), RMSE normalized to the mean (NRMSE), coefficient of determination (R^2), and mean bias error (MBE) were recorded for each model, although R^2 is only appropriate for linear models, and was thus disregarded when models were not linear. RMSE, NRMSE, R^2, and MBE are common model accuracy evaluation metrics (Cui et al., 2022). RMSE characterizes the deviation between predicted and measured EC, while R^2 gives the proportion of the variation in the dependent variable that is predicted by the independent variables.

Leave One Field Out (LOFO) analysis (Scudiero et al., 2016) was conducted on each field in the dataset, in which the model is trained on all fields except one, and the final field is used as the testing dataset, and RMSE is recorded. In 2018, LOFO analysis was conducted on M2, M3 and M8, while in 2019, LOFO analysis was conducted on M2, M3, M6, M8, and Golden 2. RMSE

for the entire dataset and RMSPE for the M2 LOFO procedure were recorded, as well as those two statistics normalized to the mean.

The best-fitting EC model for each year (2018 and 2019) was determined as the model with the lowest NRMSE of the whole dataset and/or lowest M2 LOFO NRMSPE. MBE was also considered when choosing the best model, but all MBEs were small and similar in direction (positive or negative), so this did not add much value to the choosing of the best models. These models are reported in Chapter 3: Results.

CHAPTER 3: RESULTS

Soil Properties

Soil Electrical Conductivity

A map of 2018 EM38-derived EC_e in the study area can be found in Figure 10. For the MLR EC modeling, pixels were extracted from Landsat-derived multispectral reflectance imagery collected over M2, M3 and M8. A map of 2019 EM38-derived EC_e is found in Figure 11. In 2019, due to more knowledge of the FDD, refined field techniques, plus additional available labor, more fields were added to the MLR dataset. M2, M3, M6, M8, and Golden 2 were maize fields whose satellite pixels were used in the MLR model. The northern portion of M2 was not cropped in 2019 due to the private farmer's preference, and thus was not included in the survey.

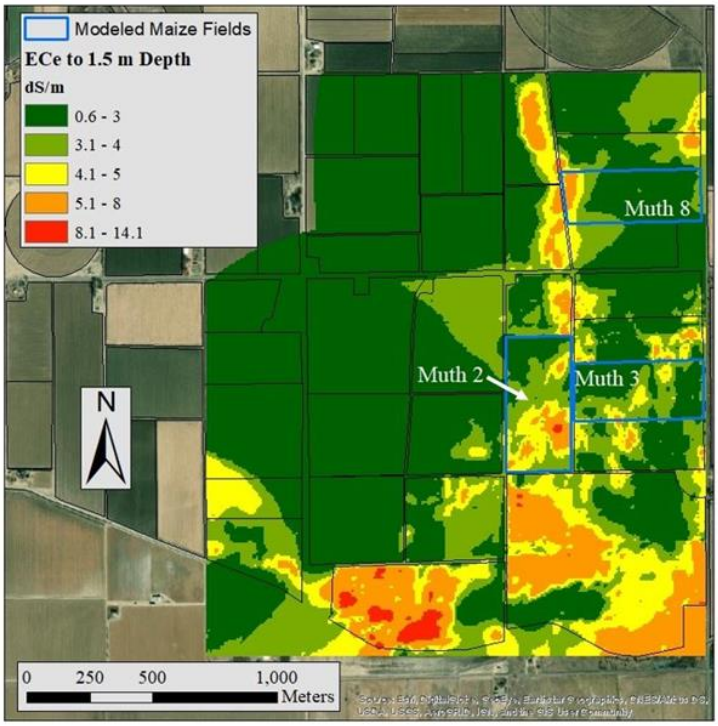

Figure 10: Map of the 2018 EC_e survey derived from the EM-38 in dS m^{-1}. The fields outlined in blue were used in the multiple linear regression model.

Figure 11: Map of the 2019 EC_e survey derived from the ME-38 in dS m^{-1}. Muth 2, Muth 3, Muth 6, Muth 8 and Golden 2 were used in the multiple linear regression model.

Frequency histograms of EC_a, EC_e, and EC_{eg} and their log transformations are found in Figure 12 and Figure 13, respectively. Upon visual inspection, each 2018 EC dataset was deemed to have a more normal distribution when subjected to a log transformation. For 2019, EC_e and EC_{eg} data depicted a normal distribution with a log transformation as well. However, the 2019 EC_a survey (data) showed a normal distribution without transformation.

Sodium Adsorption Ratio

Soil samples analyzed for Na, Ca, Mg and the calculation of the sodium adsorption ratio (SAR) were collected at 24 locations in 2021 and were subsequently analyzed (Figure 14). Given that sodicity occurs when EC_e > 4 dS m^{-1} and SAR > 13, only one of the 24 points could possibly be classified as saline-sodic or sodic, depending on its pH level, which was not analyzed. Hence, SAR data were not considered in the remainder of the analysis.

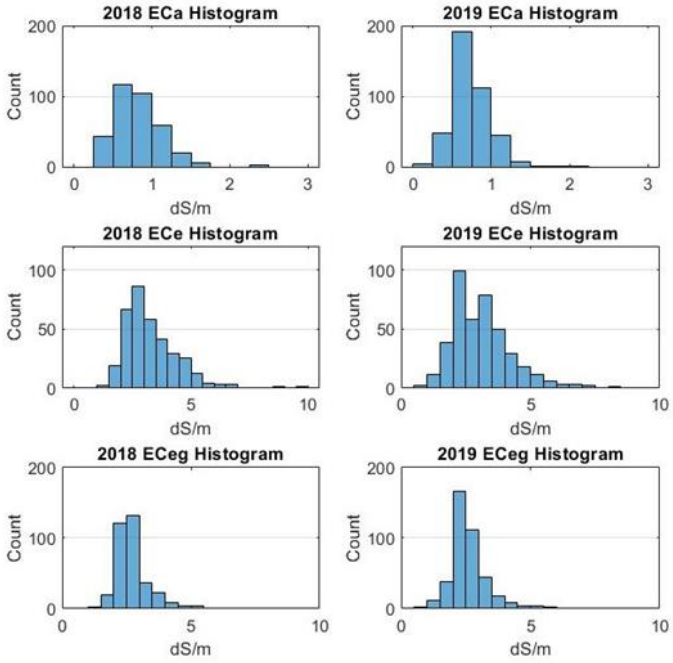

Figure 12: Frequency histograms of electrical conductivity (EC) for 2018 (left) and 2019 (right). EC_a is the apparent electrical conductivity (dS m^{-1}). Plots from 2018 included n = 311 data points, and plots from 2019 included n = 411 data points.

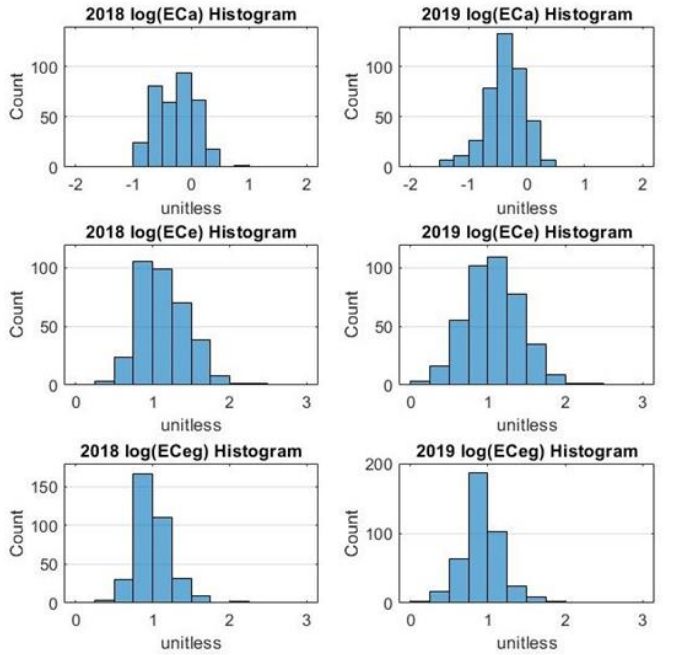

Figure 13: Frequency histograms of log-transformed electrical conductivity (EC) for 2018 (left) and 2019 (right). Plots from 2018 included n = 311 data points, and plots from 2019 included n = 411 data points.

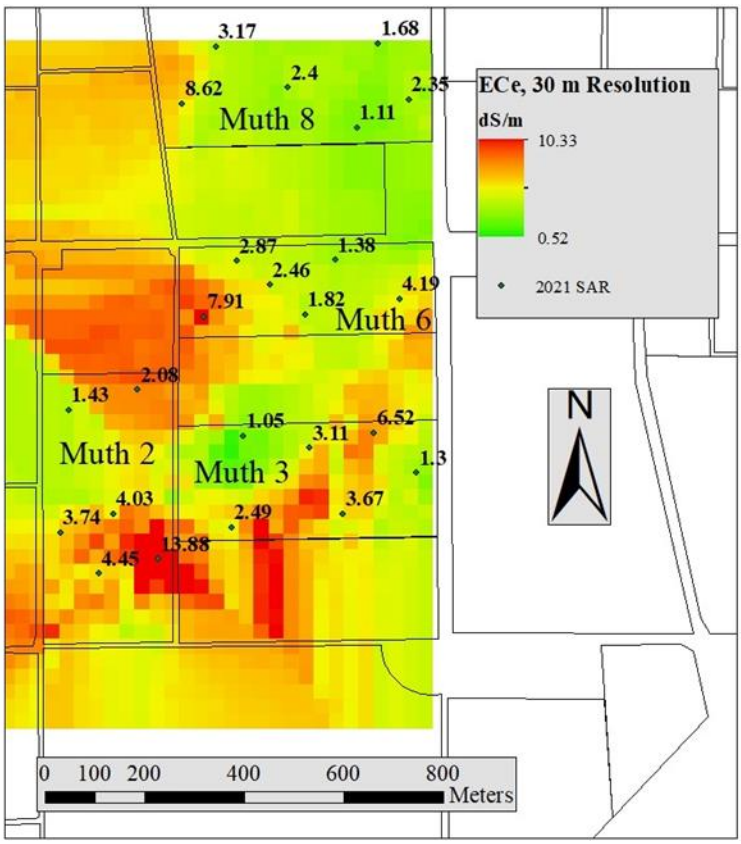

Figure 14: Location and Sodium Adsorption Ratio (SAR) values for 24 soil samples analyzed for Na, Ca and Mg in 2021.

Water Observations

2018 Total Applied Water in Muth 2

A cumulative plot of applied water for M2, calculated as the sum of gross precipitation and gross irrigation amounts, is shown in Figure 15. Also plotted are daily and cumulative gross irrigation and gross precipitation. Growing season precipitation totaled 172.7 mm, while the gross irrigation events totaled 1,152.8 mm. Total applied water, the sum of precipitation and irrigation, was 1,325.5 mm. Other fields are not plotted because only M2 had flumes on both the applied (head) and tail side of the fields.

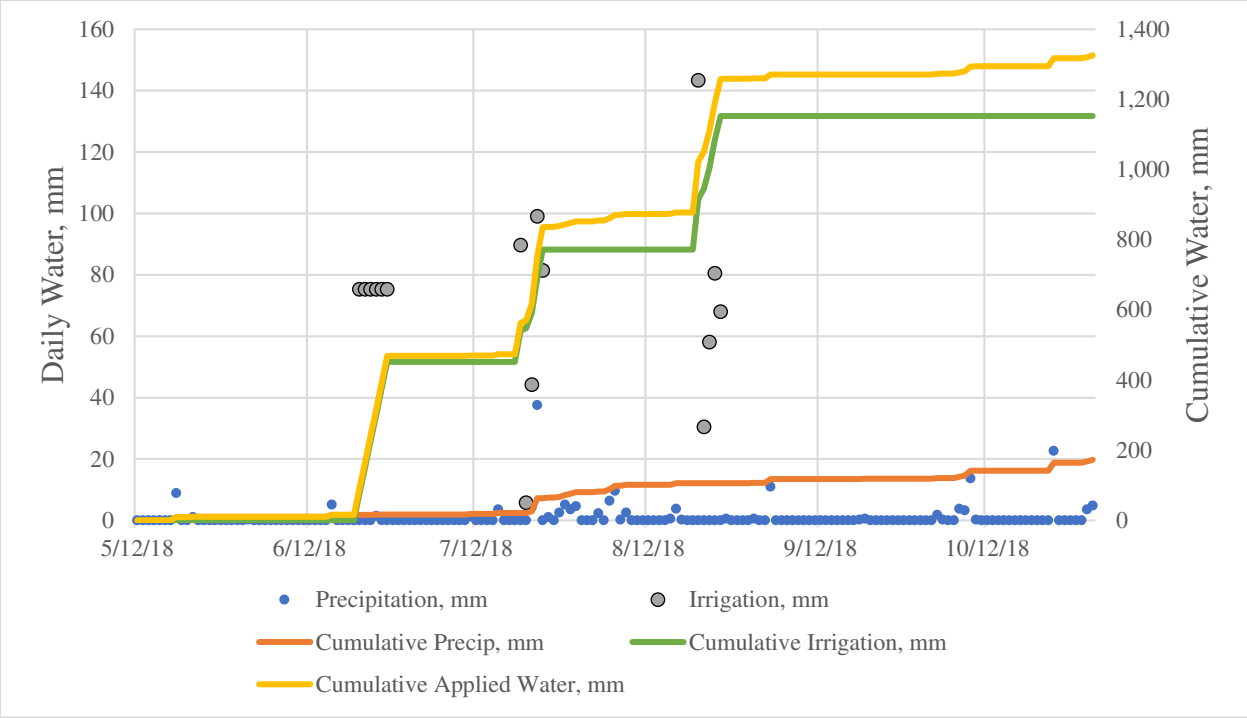

Figure 15: Daily and cumulative gross irrigation, precipitation, and total applied water for Muth 2 in 2018.

2018 Water Table Depth

The water table depth over the study area did not vary significantly over the 2018 growing season. When analyzing the average water table depth on satellite overpass dates from May 28 to August 24, 2018, the minimum average water table depth was 0.94 m, while the maximum was 1.26 m. The standard deviation of the average water table depth on satellite overpass dates was 0.12 m. Therefore, the water table depth throughout the FDD on July 7, 2018, shown in Figure 16, is representative of season-long conditions. Because the water table depth in the FDD was regularly within or near the maize root zone, capillary rise most likely had a large effect on the soil water balance and may be a source of crop stress via partial rootzone waterlogging. The water table depth was closer to the surface in the down-gradient areas of the tile drain system,

indicating that the drains may not be transporting enough groundwater to maintain water tables outside of the crop root zone.

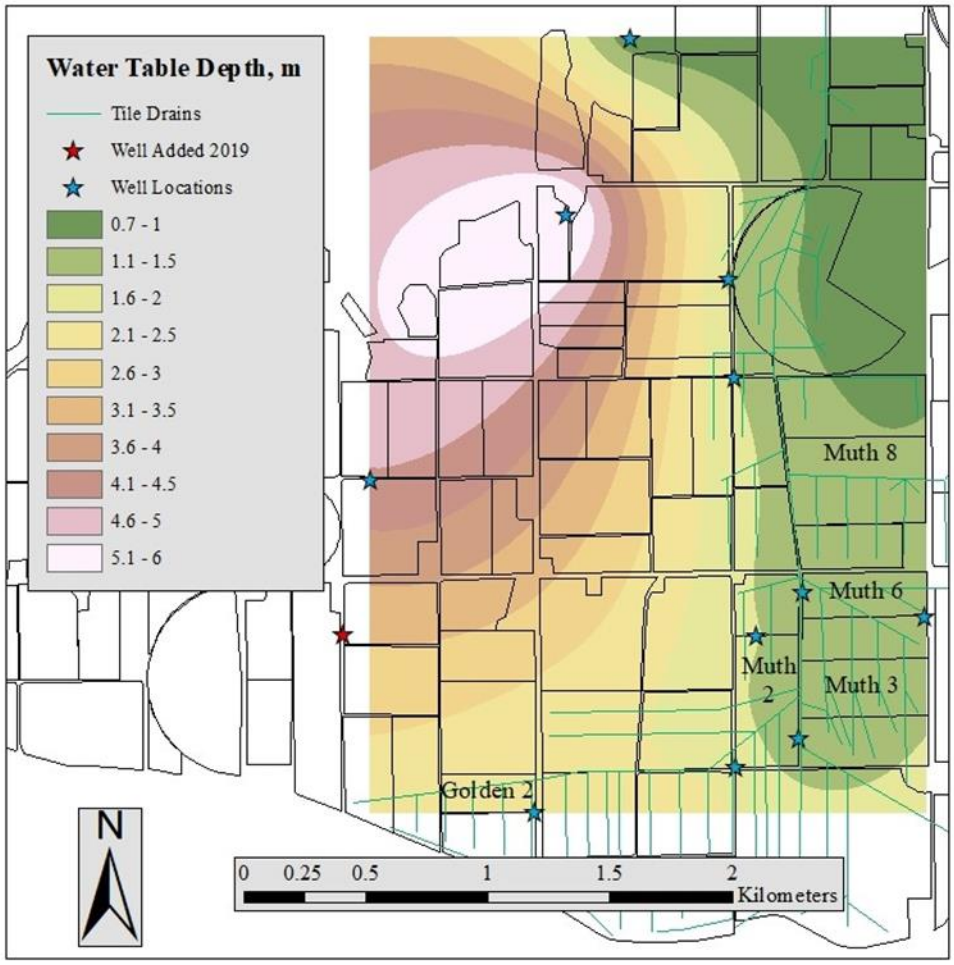

Figure 16: Interpolated color contours of water table depth in the Fairmont Drainage District on July 7, 2018, with the locations of observation wells indicated by green stars symbols.

Crop Water Stress Calculations

Satellite-Derived Crop Water Stress Index

Descriptive statistics of $CWSI_{sat}$ values for all satellite imagery pixels for the three fields used in the 2018 analysis and the five fields used in the 2019 analysis are reported in Table 2. Not only will lack of adequate water levels in the soil result in high CWSI values, but also soil salinity stress will contribute to elevated CWSI values (Gu et al., 2021). Given the arid environment and

water scarcity in the FDD, the crops likely needed water when $CWSI_{sat}$ > 0.299. In both years, the mean $CWSI_{sat}$ of each growth stage increased as the season progressed. In 2018, mean $CWSI_{sat}$ was above 0.299 in the mid and late crop growth stages, indicating that if water was available, the farmer should have irrigated more in the mid growth stage, and possibly in the late growth stage as well. In 2019, mean $CWSI_{sat}$ was above 0.299 in the development, mid and late growth stages, indicating that if water was available, the farmer should have applied more water in the development stage, and possibly in the mid and late growth stages as well.

Table 2: Satellite-derived Crop Water Stress Index ($CWSI_{sat}$) statistics for the 2018 and 2019 growing seasons, for the total growing season and for each stage of the growing season. The 2018 dataset included n = 351 data points, while the 2019 dataset included n = 411 data points.

		Growth Stage				
CWSI Statistics		Total	Initial	Development	Mid	Late
2018	Minimum	0.09	0.00	0.00	0.17	0.26
	Mean	**0.43**	**0.11**	**0.14**	**0.44**	**0.58**
	Maximum	0.77	0.72	0.49	0.80	0.99
	Standard Deviation	0.14	0.16	0.12	0.10	0.16
2019	Minimum	0.12	0.00	0.05	0.26	0.00
	Mean	**0.38**	**0.27**	**0.34**	**0.41**	**0.41**
	Maximum	0.68	0.87	0.72	0.65	0.77
	Standard Deviation	0.10	0.25	0.11	0.07	0.14

Infrared Thermometer-Derived Crop Water Stress Index

Maize canopy temperatures measured with IRTs were used to calculate $CWSI_{IRT}$ using the approach indicated in Eq. (16). This analysis was done to verify the accuracy of $CWSI_{sat}$ calculations that used ReSET ET_a [Eq. (15)]. The 2018 IRT-based crop canopy temperature readings were taken at a medium-salinity level location (EC_e of 3.4 dS m^{-1}) that coincided with the middle of a Landsat satellite pixel in M2. Field M2 was chosen because it has the largest area of high salinity values. The minimum, mean, maximum and standard deviation of these $CWSI_{IRT}$ calculations, taken from July 5 to October 5, 2018, were 0.00, 0.37, 1.00, and 0.28, respectively.

These values indicate water stress at this location, since the threshold for water stress in maize is 0.299 (Qin et al., 2021).

2018 Muth 2 Soil Water Balance

Daily values of SMD, dMAD, $CWSI_{sat}$, $CWSI_{IRT}$, and days with wetting events greater than 3 mm were plotted for a medium-level salinity location (EC_e = 3.4 dS m^{-1}) from field M2 (Figure 17). Maize is considered stressed if SMD > dMAD or CWSI > 0.299 (Qin et al., 2021). As SMD increases, CWSI values increase as well. CWSI values should drop to 0 after irrigations, when the SMD is 0 mm. Interpolated $CWSI_{sat}$ was created by multiplying ET_r from the Rocky Ford Research Station (Figure 4) on a given day by the EF [Eq. (14)] from the nearest satellite overpass date, and then subtracting that value from 1 [Eq. (15)]. The $CWSI_{sat}$ values did not follow the SMD pattern well during the crop growth season at this location. On the other hand, $CWSI_{IRT}$ matched the SMD pattern quite well. This could be explained by the footprint of each measurement. $CWSI_{sat}$ uses the 100 m Landsat thermal pixel to calculate CWSI, while $CWSI_{IRT}$ has a smaller footprint according to the IRT field-of-view, which may give more accurate measurements of crop conditions. Since the fields in the FDD are heterogeneous, this difference in measurement footprint will affect the readings. Furthermore, the IRT was placed at a 45-degree angle (oblique view) relative to the surface, meaning that it only measures crop canopy temperature. Landsat satellites, on the other hand, view the ground from nadir, meaning the satellite captures the spectral response of the soil as well as the canopy. For these reasons, when evaluating the percent of days where the crop was stressed, $CWSI_{IRT}$ was given preference over $CWSI_{sat}$.

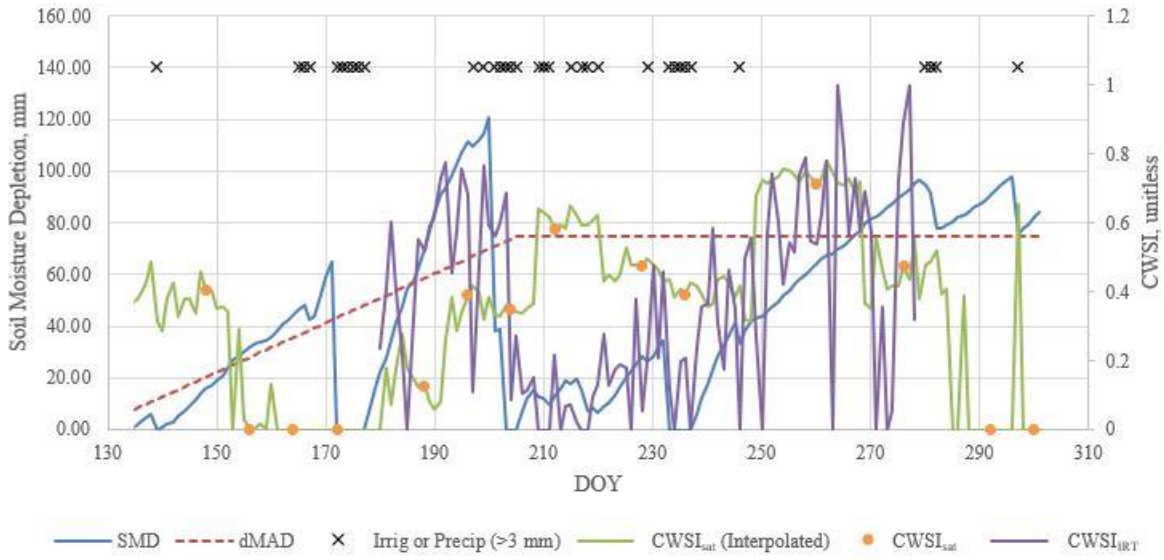

Figure 17: Crop Water Stress indices for 2018 in a medium-salinity (EC_e = 3.4 dS m^{-1}) location in Muth 2, including Soil Moisture Depletion (SMD), Management Allowed Depletion (dMAD), Crop Water Stress derived from Landsat ET_a ($CWSI_{sat}$), and CWSI derived from the IRT ($CWSI_{IRT}$). A black X symbol indicates the date in which at least 3 mm of water was received by the field.

Table 3 gives the number of days in each growth stage where the maize in M2 was considered water-stressed using both the SMD and the $CWSI_{IRT}$ metrics. Since the IRTs could not be placed in the field until after side-dressing, which was in the middle of the crop rapid development stage, $CWSI_{IRT}$ data are not available in the initial stage, hence, there is a different number of days available for analysis in the development stage. This analysis shows that maize at this location was significantly water-stressed during the development growth stage with 75% of days crossing at least one of the two stress metric thresholds ($CWSI_{IRT}$ or SMD). Similarly, the maize at this location was significantly water-stressed during the late growth stage with 100% of days crossing at least one of the two stress metric thresholds ($CWSI_{IRT}$ or SMD). Over the entire growing season (May 15 to October 28, 2018), the maize at this location was considered water-stressed by at least one of the two selected metrics on 61% of days.

Table 3: Number of days and percent of the growth stage in 2018 in Muth 2 where the maize was considered water-stressed using the SMD metric, $CWSI_{IRT}$ metric, or either.

Metric		Total	Initial	Development	Mid	Late
SMD	Days Stressed	69	9	25	0	35
	Pct of Days Stressed	41%	33%	63%	0%	70%
$CWSI_{IRT}$	Days Stressed	56	-	19	13	24
	Pct of Days Stressed	57%	-	86%	26%	89%
Either	Days Stressed	102	9	30	13	50
	Pct of Days Stressed	61%	33%	75%	26%	100%

Maize Yield

In 2018, maize yield in M2 and M3 ranged from 4,809 to 19,255 kg ha^{-1}. Per the FAO, "good commercial [maize] grain yield" is 6,000 to 9,000 kg ha^{-1} (FAO, 2023), indicating that yields in the FDD range from below to well above good commercial maize grain yield. Yield in kg ha^{-1} and EC_e in dS m^{-1} are mapped in Figure 18, indicating that salinity is reducing yields. This is confirmed by the plot in Figure 19, which indicates a logarithmic downward trend in yield with EC_e, with an RMSE of 1,965 kg ha^{-1}, or 15% of average yield.

In 2019, maize yield in the FDD ranged from 1,495 to 22,003 kg ha^{-1}. Yield in kg ha^{-1} and EC_e in dS m^{-1} are mapped in Figure 20. Note that yield and EC_e follow a spatial pattern that indicates that salinity is reducing yields. This fact is confirmed in Figure 21, which, like the 2018 data, shows a logarithmic downward trend in yield with EC_e with an RMSE of 3,032 kg ha^{-1} or 25% of average yield.

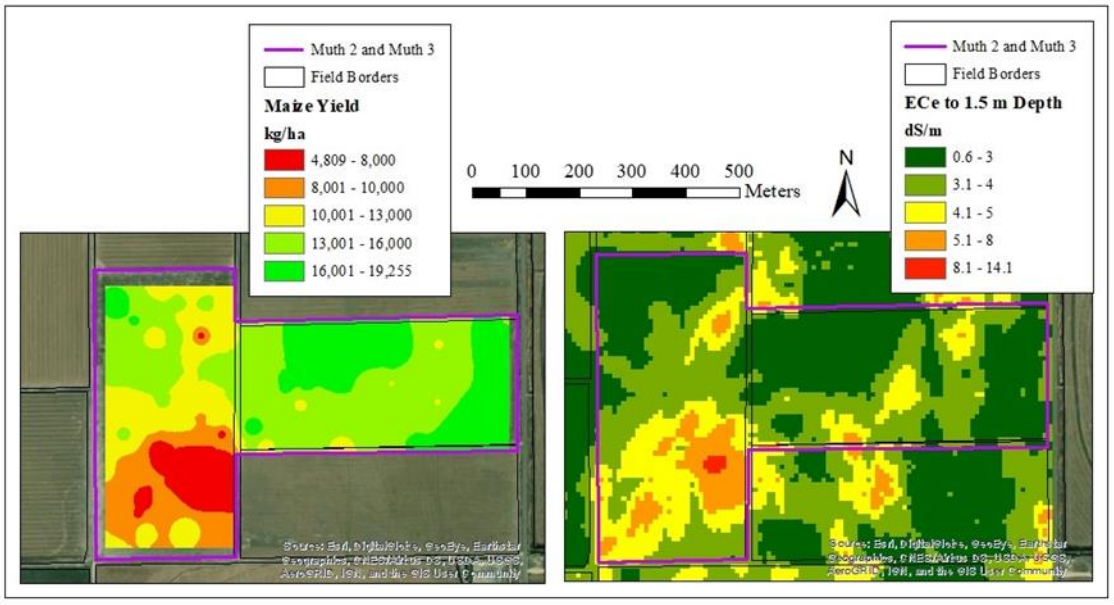

Figure 18: Maize yield in kg ha^{-1} for Muth 2 and Muth 3 in 2018 is shown on the left, with EC$_e$ in dS m^{-1} for those same fields shown on the right. Notice that yield follows a similar spatial pattern to EC$_e$, indicating that increased salinity in soils was reducing yield.

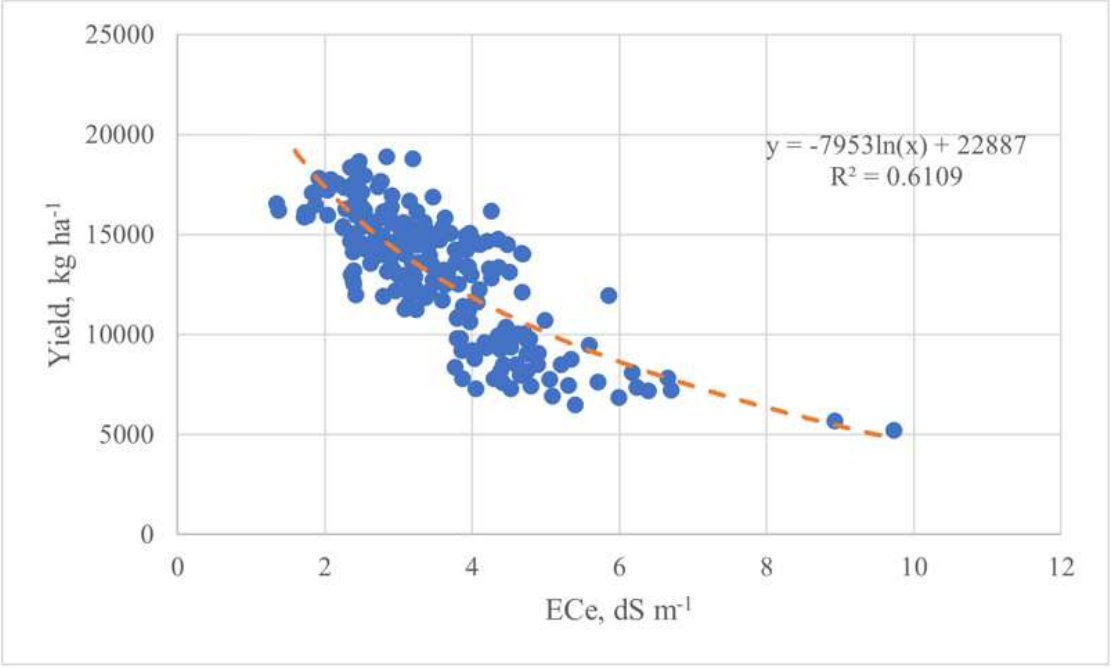

Figure 19: Maize yield vs. EC$_e$ for the 2018 fields Muth 2 and Muth 3 surveyed in the FDD, with a sample size of 223 data points.

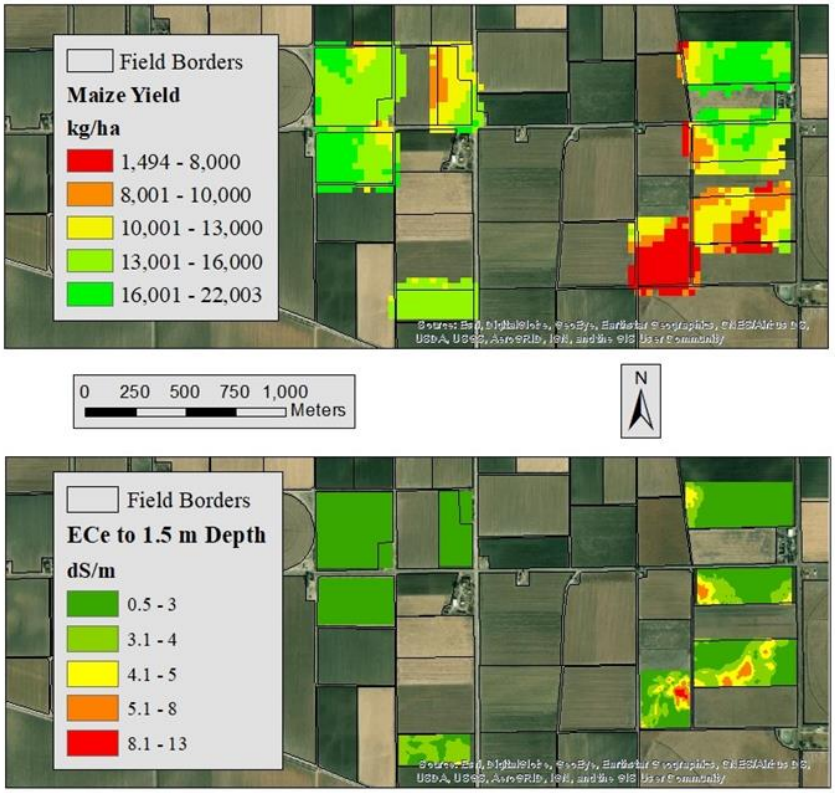

Figure 20: Maize yield in kg ha^{-1} in fields Muth 2, 3, 6, 8, 19 and 23, and Golden 2 and 8 is shown in the top plot, with EC$_e$ dS m^{-1} for those same fields shown on the bottom. Notice that yield follows a similar spatial pattern to EC$_e$, indicating that increased salinity in soils was reducing yield.

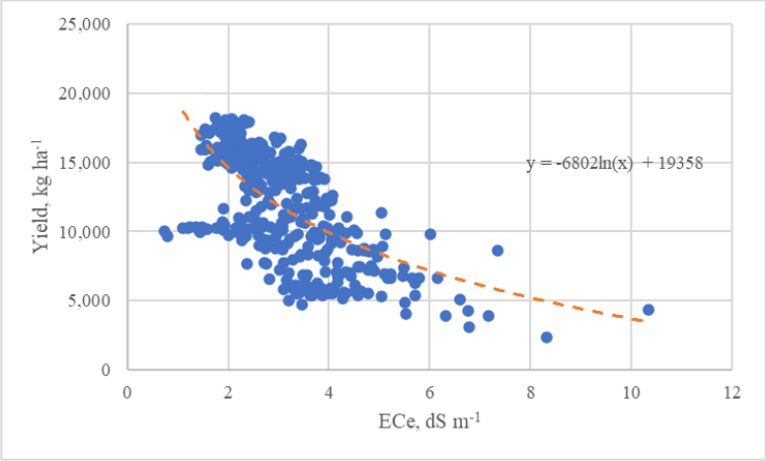

Figure 21: Maize yield vs. EC$_e$ for the 2019 fields Muth 2, 3, 6, 8, 19 and 23, and Golden 2 and 8, with a sample size of 401 data points.

Soil Electrical Conductivity Modeling

The independent variables in MLR modeling of EC were NDV I, CRSI, ET_a, and $CWSI_{sat}$. The dependent variables were EC_a, EC_e, and EC_{eg}. Descriptive statistics for daily NDVI, CRSI, ET_a, $CWSI_{sat}$, EC_a, EC_e, and EC_{eg} in 2018 and 2019 can be found in Table 4. The statistics of the two years' independent variables, NDVI, CRSI, ET_a and $CWSI_{sat}$, were not directly compared for two reasons. First, two of the fields (Golden 2 and M6) were not in the 2018 dataset but were added in 2019. Second, 16 cloud-free images were collected in 2018, with an average day of year (DOY) of 206.5, while 15 cloud-free images were collected in 2019, with a later average DOY of 223. The difference in average DOY between the two years is mostly because some images are unusable due to cloud cover.

Table 4: Descriptive statistics of daily NDVI, CRSI, ETa, ECa, ECe, and ECeg from the 2018 and 2019 datasets. The 2018 dataset included n = 4,789 data points over three fields and 16 satellite image dates. The 2019 Muth 2, 3 and 8 dataset included n = 4,253 data points over 15 satellite image dates, while the full 2019 dataset included n = 6,022 data points over five fields and 15 satellite image dates.

		NDVI *unitless*	**CRSI** *unitless*	**ETa** *mm/day*	**CWSI** *unitless*	**ECa** *dS/m*	**ECe** *dS/m*	**ECeg** *dS/m*
2018	**Minimum**	0.10	0.60	0.00	0.09	0.38	1.35	1.35
	Mean	0.52	0.82	4.08	0.43	0.83	3.32	2.75
	Maximum	0.92	0.95	10.39	0.77	2.37	9.73	8.15
	Standard Deviation	0.26	0.07	2.06	0.14	0.30	1.12	0.74
2019, Muth 2, 3 and 8 Only	**Minimum**	-	-	-	-	0.19	0.74	0.74
	Mean	-	-	-	-	0.71	3.13	2.62
	Maximum	-	-	-	-	2.19	10.33	8.70
	Standard Deviation	-	-	-	-	0.27	1.33	0.91
2019, All Fields	**Minimum**	0.11	0.69	0.00	0.12	0.19	0.74	0.74
	Mean	0.56	0.86	3.55	0.38	0.73	3.10	2.62
	Maximum	0.90	0.99	8.53	0.68	2.19	10.33	8.70
	Standard Deviation	0.26	0.07	1.80	0.10	0.25	1.18	0.80

However, the dependent variables, EC_a, EC_e, and EC_{eg}, are more comparable, since EC data comes from a single survey taken before crop emergence, i.e., around the same date in both 2018 and 2019, and because salinity does not change with NDVI, CRSI, CWSI and ET_a which fluctuate by a large amount throughout the growing season. When analyzing the three fields that were surveyed for EC in 2018 and 2019 (M2, M3 and M8), all three mean EC metrics slightly decreased from 2018 to 2019 (Table 4). The decrease in mean EC_e and EC_{eg} indicates a possible improvement in salinity conditions.

However, the ANOCOVA transformation of EM38 EC_a to EC_e had an RMSE of 1.3 dS m^{-1} (Brown, 2021), so the change of mean EC_e and EC_{eg} of -0.19 and -0.13 dS m^{-1} (respectively) means that there was not a discernable change in mean EC_e or EC_{eg}. Maximum EC_e and EC_{eg} increased from 2018 to 2019, increasing by 0.60 and 0.55 dS m^{-1}, respectively. Although this change is less than the RMSE of the EC_e calculation (1.3 dS m^{-1}), it is much more substantial than the change in mean EC, inviting the possibility that the areas with the worst salinity problems are continuing to accumulate salts.

Correlation Coefficients

Daily Correlation Coefficients

A single-date correlation coefficient analysis was conducted to indicate how correlation between variables changes within a crop growth stage. It is not intended as the final basis for model selection. All r values are expected to be negative (inverse correlation) because salinity negatively affects crop health, and these are metrics of crop health. Table 5 shows the Pearson correlation coefficient (r) for 2018 values of NDVI, CRSI and ET_a with EC_e for each satellite overpass date.

Table 5: r values for NDVI, CRSI and ET_a versus EC_e for all 2018 satellite overpass dates categorized by maize growth stage. Statistically significant r values at $p < 0.05$ are shown in bold boxed values. Means highlighted in gray boxes showed the strongest correlations. Red and blue shading increases with the magnitude of the r value, with red indicating a positive correlation and blue indicating a negative correlation. Orange shading is included to easily view growth stages. Each date included n = 351 data points.

		Pearson's r against EC_e for 2018 Data			
Growth Stage	**Date**	**NDVI**	**CRSI**	**ET_a**	**Mean**
Initial	**28-May**	-0.09	-0.05	**-0.18**	**-0.11**
	5-Jun	0.06	0.10	**-0.33**	-0.06
Development	**13-Jun**	**0.35**	**0.36**	**0.12**	**0.27**
	21-Jun	**-0.20**	**-0.19**	**-0.22**	**-0.21**
Mid	**7-Jul**	**-0.46**	**-0.35**	**-0.22**	**-0.35**
	15-Jul	**-0.54**	**-0.23**	**-0.32**	**-0.36**
	23-Jul	**-0.41**	**-0.39**	**0.11**	**-0.23**
	31-Jul	**-0.20**	**-0.11**	-0.03	**-0.11**
	16-Aug	**-0.15**	-0.06	-0.01	-0.07
Late	**24-Aug**	**-0.32**	**-0.28**	-0.05	**-0.22**
	17-Sep	**-0.26**	-0.04	0.02	-0.09
	3-Oct	-0.03	**0.15**	0.06	0.06
	19-Oct	0.03	0.10	0.05	0.06
	27-Oct	0.07	**0.14**	0.02	0.08
	Mean	**-0.15**	-0.06	-0.07	

Means by date and by metric are shown on the edges. Means highlighted in gray boxes showed the strongest correlations. Red and blue shading increases with the magnitude of the r value, with red indicating a positive correlation and blue indicating a negative correlation. Orange shading is included to easily view growth stages. Statistically significant r values at $p < 0.05$ are shown as bold boxed values. All values in the development stage, and most of the values in the mid stage are statistically significant.

The mid maize growth stage revealed the greatest inverse correlation with EC_e for all three metrics. This corresponds with the period of least crop water stress, as shown in Table 3. NDVI shows the greatest inverse correlation with EC_e when looking at the entire growing season, while

ET_a shows the least inverse correlation and CRSI in the middle. However, ET_a shows the strongest inverse correlation with EC_e in the initial and development stages.

Table 6 shows the r values for the 2019 data for each satellite overpass date. Similar to the 2018 analysis, means by date and by metric are shown on the edges. Means highlighted in gray boxes

Table 6: r values for NDVI, CRSI and ET_a versus EC_e for all satellite overpass dates categorized by maize growth stage. Statistically significant R values at $p < 0.05$ are shown in bold boxed values. Means highlighted in gray boxes showed the strongest correlations. Red and blue shading increases with the magnitude of the R value, with red indicating a positive correlation and blue indicating a negative correlation. Orange shading is included to easily view growth stages. Each date included n = 411 data points.

		Pearson's r against EC_e for 2019 Data			
Growth Stage	**Date**	**NDVI**	**CRSI**	**ET_a**	**Mean**
Initial	**15-May**	0.01	-0.03	-0.03	-0.02
	31-May	**0.18**	**0.32**	**0.22**	**0.24**
Development	**24-Jun**	**0.22**	**0.29**	**-0.11**	**0.13**
	2-Jul	-0.06	**0.15**	**0.25**	**0.11**
	18-Jul	**-0.20**	**-0.19**	**-0.26**	**-0.21**
Mid	**26-Jul**	-0.09	**-0.11**	**-0.13**	**-0.11**
	3-Aug	-0.09	**-0.13**	**-0.15**	**-0.12**
	11-Aug	-0.06	**-0.13**	**-0.21**	**-0.13**
	19-Aug	**-0.37**	**-0.30**	**-0.25**	**-0.31**
	4-Sep	**-0.51**	**-0.30**	**-0.24**	**-0.35**
Late	**12-Sep**	**-0.56**	**-0.12**	**-0.23**	**-0.31**
	20-Sep	**-0.50**	**-0.30**	**-0.35**	**-0.38**
	6-Oct	**-0.41**	**-0.33**	**-0.18**	**-0.31**
	14-Oct	**-0.32**	**0.38**	**-0.14**	-0.02
	22-Oct	**-0.11**	**0.35**	-0.05	0.06
	Mean	**-0.19**	-0.03	**-0.12**	

showed the strongest inverse correlations. Red and blue shading increases with the magnitude of the r value, with red indicating a positive correlation and blue indicating a negative correlation. Orange shading is included to easily view growth stages. Statistically significant r values at p <

0.05 are shown as bold boxed values. The majority of r values throughout the growing season are statistically significant, including all but one value in the late growth stage.

NDVI, CRSI and ET_a within the latter half of the mid growth stage and the first half of the late growth stage show the greatest inverse correlation to EC_e. Overall, NDVI shows the greatest inverse correlation with EC_e, particularly in the mid and late growth stages. CRSI and ET_a both reveal low inverse correlation in the early and late parts of the season but higher inverse correlation in the mid growth stage and the first part of the late growth stage.

Seasonal Correlation Coefficients

Seasonal correlations were used to determine which model parameters and explanatory variables should be used for MLR modeling. The natural log of all four independent variables (NDVI, CRSI, ET_a and CWSI) are included in the seasonal correlation analysis. The EC type and growth stage combinations that showed the strongest correlation against each independent variable in the modeling are found in Table 7. All R values are statistically significant at $p < 0.05$. Full seasonal correlation tables can be found in Appendix A.

The strongest r values (inverse for NDVI, CRSI and ET_a, but positive for CWSI) were found in 2019 using data that were not interpolated between satellite dates. These correlations occurred with EC_a in the development stage for each of the 8 independent variables used. NDVI, CRSI and ET_a had equal correlation with EC_a whether transformed or not transformed, while CWSI had higher correlation without a transformation than with a transformation.

Correlations in 2018 were lower than correlations in 2019 for all combinations of interpolation, transformation, and explanatory variables. Correlations between NDVI and EC were stronger with NDVI interpolated between satellite overpass dates. Correlations between CRSI or ET_a and

Table 7: The strongest correlation for each independent variable, classified by year and whether the independent variables were interpolated or not between satellite dates. All r values are statistically significant at $p < 0.05$. Correlations with NDVI, CRSI and ETa are inverse while correlations with CWSI are positive. Color shading indicates magnitude, with red being negative and blue being positive. The full seasonal correlation tables can be found in Appendix A. Data in 2018 had a sample size of n = 351, while data in 2019 had a sample size of n = 411 data points.

Year	Interpolation	Independent Variable List	Strongest Correlation		
			Type of EC	Growth Stage	r
2018	Interpolated	NDVI	EC_{eg}	Development	-0.54
		CRSI	$\ln(EC_{eg})$	Development	-0.34
		ET_a	EC_{eg}	Development	-0.37
		CWSI	$\ln(EC_{eg})$	Total	0.28
		ln(NDVI)	EC_{eg}	Development	-0.57
		ln(CRSI)	$\ln(EC_{eg})$	Development	-0.34
		$\ln(ET_a)$	EC_{eg}	Development	-0.38
		ln(CWSI)	$\ln(EC_{eg})$	Development	0.34
	Not Interpolated	NDVI	EC_{eg}	Development	-0.50
		CRSI	EC_a	Mid	-0.35
		ET_a	EC_{eg}	Development	-0.39
		CWSI	EC_{eg}	Development	0.39
		ln(NDVI)	EC_{eg}	Development	-0.54
		ln(CRSI)	EC_a	Mid	-0.35
		$\ln(ET_a)$	EC_{eg}	Development	-0.40
		ln(CWSI)	$\ln(EC_{eg})$	Development	0.30
2019	Interpolated	NDVI	EC_e	Total	-0.68
		CRSI	EC_a	Development	-0.65
		ET_a	EC_a	Development	-0.40
		CWSI	EC_a	Development	0.40
		ln(NDVI)	EC_e	Total	-0.69
		ln(CRSI)	EC_a	Development	-0.65
		$\ln(ET_a)$	EC_a	Development	-0.42
		ln(CWSI)	EC_a	Development	0.35
	Not Interpolated	NDVI	EC_a	Development	**-0.72**
		CRSI	EC_a	Development	**-0.71**
		ET_a	EC_a	Development	**-0.59**
		CWSI	EC_a	Development	**0.59**
		ln(NDVI)	EC_a	Development	**-0.72**
		ln(CRSI)	EC_a	Development	**-0.71**
		$\ln(ET_a)$	EC_a	Development	**-0.59**
		ln(CWSI)	EC_a	Development	**0.55**

EC were stronger without interpolation, and correlations between CWSI and EC did not show a clear trend whether to interpolate between satellite overpass dates or not.

In 2019, explanatory variables had the strongest correlation with EC in the form of EC_a in 14 of 16 of the correlations found in Table 7, with the two other correlations being with EC_e. Out of the 16 explanatory variable correlations in 2018, two were with EC_a, 9 were with EC_{eg}, and 5 were with $\ln(EC_{eg})$.

Two of the four explanatory variables used in the MLR modeling were NDVI and ET_a. These two variables are expected to be somewhat redundant, since the NDVI (greenness) of the crop depends significantly on the transpiration rate of the crop, which is captured in ET_a. Therefore, Table 8 is presented to show the correlation between NDVI at ET_a.

Table 8: r values for NDVI versus ET_a for 2018 and 2019 data, both interpolated between satellite dates and only satellite overpass data. The correlation for the total season and each individual growth stage is presented. The mean of each growth stage and each dataset is presented on the edges of the table. The red color is a color scale shown to emphasize strong correlations. The starred value is the only value that is not statistically significant at $p < 0.05$.

	Total	Initial	Development	Mid	Late	Mean
2018 Interpolated	0.60	0.06*	0.79	0.13	0.67	0.45
2018 Not Interpolated	0.60	0.25	0.78	-0.12	0.79	0.46
2019 Interpolated	0.12	0.24	0.28	0.38	0.31	0.26
2019 Not Interpolated	0.19	0.25	0.65	0.20	0.32	0.32
Mean	0.38	0.20	0.62	0.15	0.52	

Values of r between NDVI and ET_a in the development and late growth stages, as well as during the total growing season show the strongest correlation. Values of r using 2018 data were stronger than r values using 2019 data. Values of r from data only on satellite overpass dates were higher than r values from data with interpolation between satellite overpass dates.

Although this analysis shows statistically-significant correlation between NDVI and ET_a, these datasets were still used as explanatory variables in MLR modeling of EC. Although NDVI and ET_a are both related to crop transpiration or growth, one is a measure of the greenness of the crop, and one is the measure of water moving from liquid to vapor form. Past studies have used independent variables that are heavily related. Bouaziz et al., (2011) created an equation that uses a vegetation index that includes the NIR bandwith, as well as the NIR band as the two independent variables. Wei et al. (2020) used machine learning to create models that have independent variables of hyperspectral bands very close together on the electromagnetic spectrum. The study also uses those same spectral bands in vegetation indices that have the same structure and similar bandwidths. Blue, Green, Red and NIR bands appear in many different arithmetic combinations within each model. Therefore, there is precedent in the literature to use NDVI and ET_a in the same MLR models.

Data Distribution Normality

Data distribution normality is a requirement for MLR modeling. Here, a quantile-quantile plot was used to assess normality. See the section *Test of Data Distribution Normality* for a more in-depth description of quantile-quantile plots. Out of the five combinations of year and EC type chosen for analysis (2018 EC_a and EC_{eg} and 2019 EC_a, EC_e, and EC_{eg}, Table 1), all combinations except for 2019 EC_a displayed a more normal distribution with a log transformation, as shown by the point cloud being closer to the 1:1 line in Figure 22 and Figure 23. This shift to normality with a log transformation was noticeable particularly at higher values of the EC variables, which are more important in EC prediction modeling (Figure 22, Figure 23). Therefore, only the models that used transformed values of EC variables were used, except for those using 2019 EC_a.

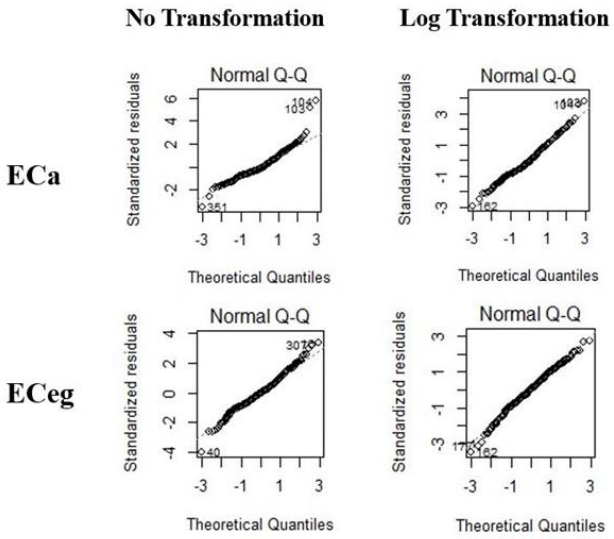

Figure 22: Normal Quantile-Quantile (Q-Q) plots for 2018 EC_a and EC_{eg} data. The left plots are data with no transformation, and the right plots are data with a log transformation. The data have a sample size of n = 351.

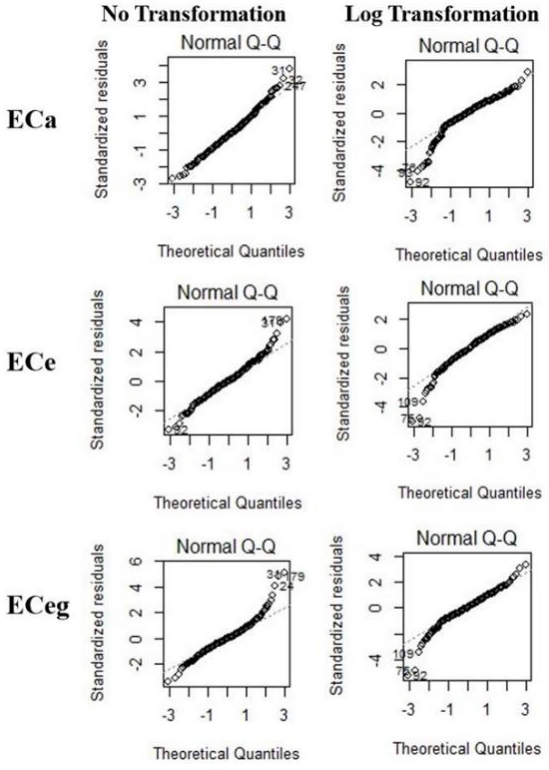

Figure 23: Normal Quantile-Quantile (Q-Q plots for 2019 EC_a, EC_e, and EC_{eg} data. The left plots are data with no transformation and the right plots are data with a log transformation. The data have a sample size of n = 411.

2018 EC Modeling Results

The 2018 EC modeling dataset included 16 satellite images, each with 35 satellite pixels with a 30 m pixel size. In total, 56 MLR models to predict EC_a or EC_{eg} were created in the R language (Appendix C) using the lm() command. The modeler entered the dependent variable array and any independent variable arrays that the regression could select from, while indicating that interactions were also available for selection. The regression analysis selects the EC model with the lowest AIC (Cavanaugh and Neath, 2019). An expansion on AIC is found in the *Model Setup* section of Chapter 2.

After accounting for data distribution normality in the datasets and eliminating models that were very similar but did not improve the statistical results, 33 models were identified for further analysis. Out of the 33 identified models, 30 can be simplified to ten sets of parameters (associated with growth stage, EC type, integration or not, and data transformation type), each with three different combinations of input variables available for selection by the model (CWSI and ET_a both available, one of the two available, or neither available). The model in which neither CWSI nor ET_a is available for selection by the model is considered the base scenario. The remaining three of the 33 models were created after the original 30 models were evaluated, by conducting piecewise MLR modeling of the models with the best RMSE statistics.

Mean Bias Error of 2018 EC Models

All MBEs in the modeling of EC using the 2018 datasets were positive and less than 0.06 dS m^{-1}, indicating that the models were consistently and slightly overestimating EC.

Addition of Crop Water Use or Stress Metrics to Explanatory Variables

When considering all ten of the models and their parameter sets (EC type predicted, transformation used, growth stage used, interpolated data or not), the average NRMSE over all

fields and M2 NRMSPE decreased slightly when adding either CWSI or ET_a to the variables available for selection (Table 9). A slight decrease in average NRMSE and NRMSPE in M2 alone also occurred when both CWSI and ET_a were available for selection. These results indicate that adding a measure of crop water use or stress, whether CWSI or ET_a, only marginally increased EC model prediction accuracy.

Table 9: Average Normalized Root Mean Square Error (NRMSE) over all fields and Average Muth 2 Normalized Root Mean Square Prediction Error (NRMSPE) for all 2018 EC models (EC_a and EC_{eg}), separated by which variables were available for selection by the model. Ten models were used in this calculation, each with a sample size of n = 351 data points. Darker red cells indicate lower error values.

	Explanatory Variables Used		
	NDVI and CRSI Only	NDVI, CRSI and Either CWSI or ET_a	NDVI, CRSI, CWSI and ET_a
Average NRMSE	0.24	0.23	0.22
Average M2 NRMSPE	0.27	0.24	0.24

Data Transformation Type

Six of the ten EC models formed two pairs of three models in which the same parameters (EC type predicted, growth stage used, transformation used, interpolation or not) were included in the EC model, except for a variation by transformation type. Three EC models included all log-transformed variables, while three included a combination of log-transformed and non-transformed variables. Table 10 demonstrates that in 2018, a combination of log-transformed and non-transformed input variables had slightly better EC prediction accuracy than an EC model in which all variables were log-transformed, demonstrated by a decrease in NRMSE or NRMSPE of about 1 to 2% of the mean EC. This transformation analysis used a small sample size, though, with only three equations making up each row.

Interpolation Between Satellite Dates

Six of the ten EC models formed two pairs of three models in which the same parameters (EC type predicted, interpolation or not, growth stage used, transformation used) were inserted into the EC model, except a variation by whether the data were interpolated between satellite dates before integrating or not. Removing interpolation between satellite dates, and only integrating

Table 10: Average Normalized Root Mean Square Error (NRMSE) and Average Muth 2 Normalized Root Mean Square Prediction Error (NRMSPE) for 2018 EC modeling, separated by transformations used. Each row in the table includes three models, each with a sample size of n = 351 data points. Darker red cells indicate lower error values.

Variables Available for Selection	NDVI and CRSI Only		NDVI, CRSI and Either CWSI or ET_a		NDVI, CRSI, CWSI and ET_a	
Statistical Metric	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE
All Terms Log-Transformed	0.25	0.28	0.23	0.25	0.23	0.25
Combination of Log and No Transformation	0.23	0.26	0.22	0.24	0.21	0.23

the data from satellite overpasses slightly lowered the NRMSE and M2 NRMSPE statistics, as shown in Table 11. The result indicates that interpolating data between satellite dates might not improve the EC model prediction accuracy and should not be used in future maize EC prediction modeling.

Table 11: Average EC Normalized Root Mean Square Error (NRMSE) and Average Muth 2 EC Normalized Root Mean Square Prediction Error (NRMSPE) for 2018 EC modeling, separated by whether the data were interpolated between satellite dates or not. Each row of the table is the average of four EC models' results, each with a sample size of n = 351 data points. Darker red cells indicate lower error values.

Variables Available for Selection	NDVI and CRSI Only		NDVI, CRSI and Either CWSI or ET_a		NDVI, CRSI, CWSI and ET_a	
Statistical Metric	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE
Interpolation	0.23	0.27	0.22	0.24	0.21	0.23
No Interpolation	0.23	0.25	0.22	0.23	0.21	0.23

Best EC Prediction Equation in 2018

The best-performing single EC modeling equation [Eq. (20)] for the 2018 dataset was as follows:

$$\ln(EC_{eg}) = -0.94 - 0.07(CWSI) + 1.19(CRSI) - 0.84[\ln(NDVI)] - 4.91(CWSI)[\ln(NDVI)] \quad (20)$$

Eq. (20) predicts $\ln(EC_{eg})$ using data from the crop development stage that did not include interpolation between satellite dates. CWSI, CRSI and ln(NDVI) plus their interactions were available for model selection of explanatory variables. Eq. (20) resulted in an RMSE of 0.49 dS m^{-1} for the entire dataset (NRMSE of 18% of average EC_{eg} over all fields), and an RMPSE of 0.58 dS m^{-1} (NRMSE of 19% of average M2 EC_{eg}) for M2 LOFO prediction. The MBE was 0.04 dS m^{-1}. For reference, studies reviewed in the Introduction chapter had EC_e RMSE values between 1.15 and 3.63 dS m^{-1}. A 1:1 scatter plot of predicted vs observed EC values for Eq. (20) is shown in Figure 24, and a 1:1 density plot is shown in Figure 25.

When modeling EC_{eg} using the same parameters as in Eq. (20) (development stage, natural log transformation of EC_{eg}), but with data interpolated between satellite dates, and considering the

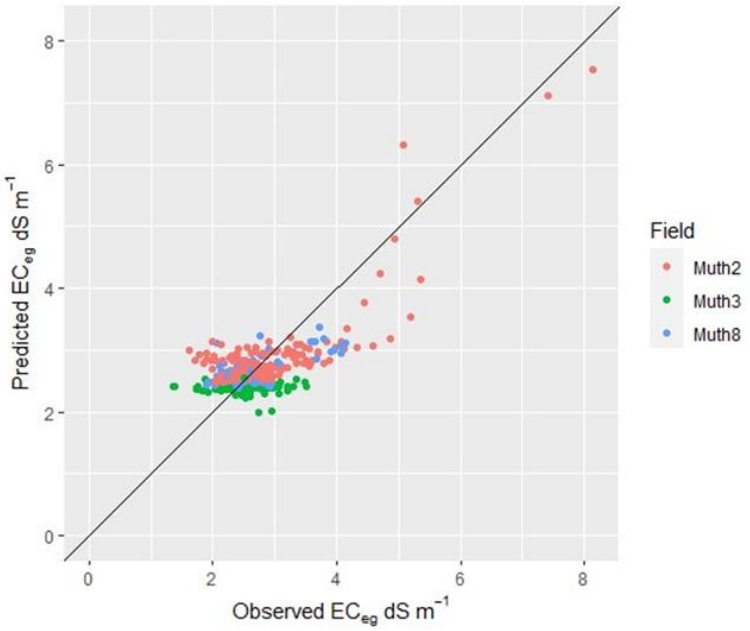

Figure 24: A 1:1 scatter plot of 2018 Eq. (20) model predicted EC_{eg} vs observed EC_{eg} values. Eq. (20) was developed using data from the crop growth development stage for satellite overpass dates only. The sample size is n = 351.

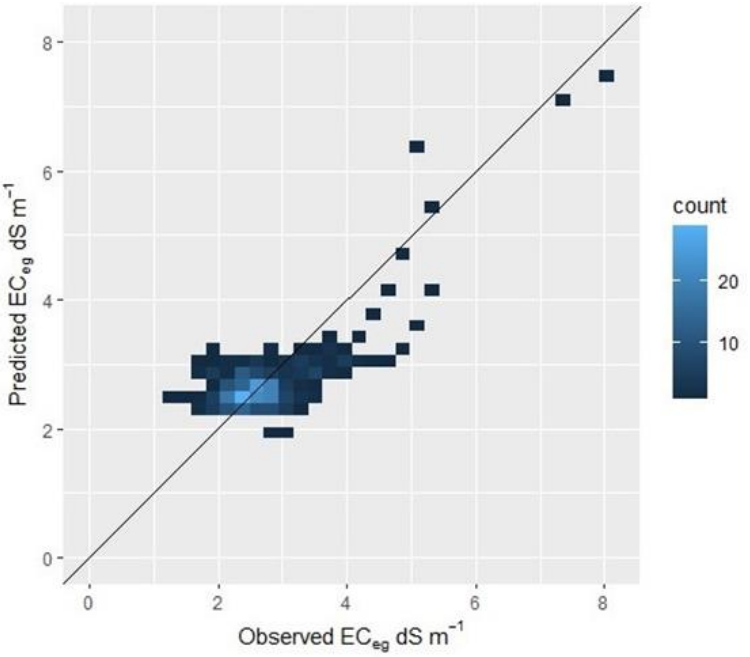

Figure 25: A 1:1 density plot of the 2018 Eq. (20) model predicted EC_{eg} vs observed EC_{eg} values. The color of the pixel indicates the number of data points in that area. Eq. (20) was developed using data from the crop growth development stage for satellite overpass dates only. The sample size is n = 351.

same explanatory variables available for selection, the RMSE of the whole dataset increased from 0.49 to 0.51 dS m^{-1} and the RMSPE of M2 LOFO prediction jumped from 0.58 to 0.65 dS m^{-1}. This decrease in predictive capability provides additional evidence that interpolation between satellite dates should not be used when creating EC prediction models.

Piecewise Regression of 2018 EC

The data and variables [EC_{eg}, CWSI, CRSI, ln(NDVI) and (CWSI)(ln(NDVI))] used in the regression to compute Eq. (20) were input into a piecewise regression. Visual inspection of Figure 24 and Figure 25 indicates a split in EC model fitting capability at about EC_{eg} = 4 dS m^{-1}. This EC_{eg} value could be the point at which salinity becomes the dominant stressor, or the split could be at 4 dS m^{-1} because that is where there begins to be fewer points available to fit. Therefore, all 2018 data were split at 4 dS m^{-1} for the piecewise regression. There were 333 data points with EC_{eg} < 4 dS m^{-1} and 18 data points with EC_{eg} > 4 dS m^{-1}. The resulting piecewise EC_{eg} model is:

$$\ln(EC_{eg}) = -0.34 + 0.81(CRSI) - 0.67[\ln(NDVI)] \quad for\ EC_{eg} < 4 \qquad (21a)$$

$$= -8.36 + 5.87(CRSI) - 2.37[\ln(NDVI)] \quad otherwise \qquad (21b)$$

When comparing Eq. (21) to Eq. (20), the RMSE decreased from 0.49 dS m^{-1} (NRMSE of 18% of average EC_{eg} over all fields) to 0.44 dS m^{-1} (NRMSE of 16% of average EC_{eg} over all fields). The M2 RMSPE value decreased from 0.58 dS m^{-1} (NRMSE of 19% of average M2 EC_{eg}) to 0.51 dS m^{-1} (NRMSPE of 17% of average M2 EC_{eg}). The MBE of the piecewise model decreased from 0.04 dS m^{-1} (non-piecewise model) to 0.03 dS m^{-1}. When comparing the individual parts of the piecewise equation to those from the single equation, the RMSE of data with EC_{eg} values below 4 dS m^{-1} stayed the same at 0.44 dS m^{-1} (NRMSE of 17% of average EC_{eg}). The MBE of the regression using data below 4 dS m^{-1} increased from 0.003 dS m^{-1} using

one equation [Eq. (20)] to 0.04 dS m^{-1} using a two-part equation [Eq. (21)] equations. The RMSE of data with EC_{eg} above 4 dS m^{-1} decreased from 1.05 dS m^{-1} (NRMSE of 21% of average EC_{eg}) with a single equation [Eq. (20)] to 0.52 dS m^{-1} (NRMSE of 11% of average EC_{eg}) using a two-part equation [Eq. (21)]. The MBE of the regression using data above 4 dS m^{-1} decreased from 0.79 dS m^{-1} using one equation [Eq. (20)] to 0.02 dS m^{-1} using a two-part equation [Eq. (21)]. For reference, studies reviewed in the Introduction chapter had EC_e RMSE values between 1.15 and 3.63 dS m^{-1}, though some of these studies included areas with EC_e of up to 20 dS m^{-1}. The 1:1 plots of model-predicted vs observed EC_{eg} values from Eq. (21) are shown in Figure 26 and Figure 27.

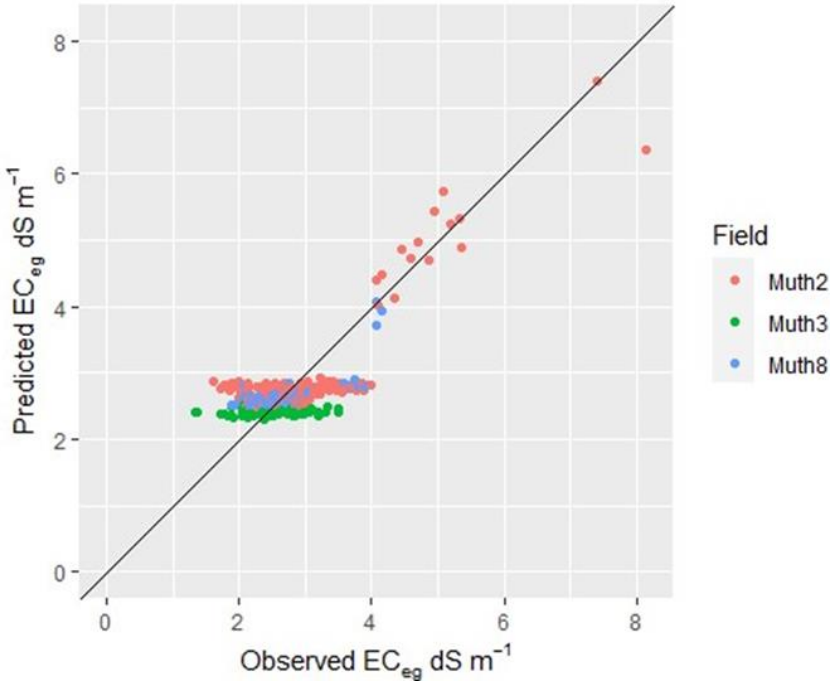

Figure 26: A 1:1 scatter plot of the 2018 model-predicted [Eq. (21)] vs observed EC_{eg} values for the two-part piecewise regression. Eq. (21) was developed using data from the development stage for satellite overpass dates only. The sample size was n = 351.

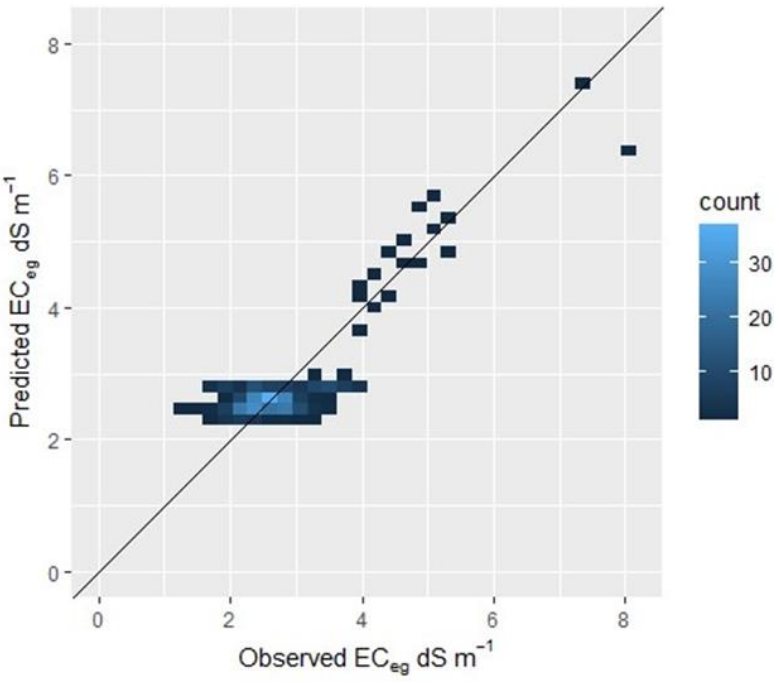

Figure 27: A 1:1 density plot of the 2018 model-predicted [Eq. (21)] vs observed EC_{eg} values for the two-part piecewise regression. The color of the pixel indicates the number of data points in that area. Eq. (21) was developed using data from the development stage for satellite overpass dates only. The sample size was n = 351.

2019 Modeling Results

The 2019 EC modeling dataset included 15 satellite images each with 411 satellite pixels of 30 m pixel size. In total, 67 MLR models to predict EC_a, EC_e, or EC_{eg} were created in R. The lm() command in the R language was used to create these linear EC models (Appendix C). The modeler entered the dependent variable array and any independent variable arrays that the regression could select from, while indicating that interactions were also available for selection. The regression creates the EC model with the lowest AIC (Cavanaugh and Neath, 2019). An explanation of AIC is found in the *Model Setup* portion of Chapter 2.

After accounting for distribution normality in the datasets and eliminating EC models that were very similar but did not produce good statistics, 33 resulting EC models were analyzed. Similar to 2018, within these 33 EC models, 30 could be simplified to ten sets of parameters (growth

stage, EC type, integration or not, data transformation type), each with three different combinations of input variables available for selection by the model (both CWSI and ET_a available, one of the two available, or neither available). The model in which neither CWSI nor ET_a is available for selection by the model is considered the base scenario. The remaining three of the 33 models were created after the original 30 models were evaluated, by conducting piecewise MLR modeling of the EC models with the best RMSE statistics.

One of the 33 EC models stood out as having particularly poor results. When using 2019 data from the total growing season rather than from an individual growth stage, NDVI showed a strong correlation with EC_e. Therefore, EC_e models were developed from explanatory variables obtained from the entire growing season and interpolated between satellite dates. This EC_e model was included in the final iteration of results but performed poorly compared to the other ten models. The NRMSE for this 2019 EC_e model, when using NDVI, CRSI and either CWSI or ET_a was 0.33 dS m^{-1}, while the average for the other ten models was 0.25 dS m^{-1}, with a maximum of 0.28 dS m^{-1}. When considering M2 data only, the NRMSPE for this EC_e model was 0.78 dS m^{-1}, while the average for the other ten models was 0.24 dS m^{-1}, with a maximum of 0.41 dS m^{-1}.

Mean Bias Error of 2019 EC Prediction Models

All MBE values in the 2019 EC modeling were positive and below 0.13 dS m^{-1}, except for the case of the model created using the total crop growth season, which had MBE values between 0.15 and 0.18 dS m^{-1}. This result indicates that the EC models were consistently and slightly overestimating the three types of EC.

Addition of Crop Water Use or Stress Metrics to Explanatory Variables

When looking at all 10 of the parameter sets analyzed (related to EC type, transformation used, growth stage used, and interpolation or not), the average NRMSE and average M2 NRMSPE of

the EC prediction models decreased slightly when adding either CWSI or ET_a to the variables available for selection (Table 12). This decrease occurred for both the full dataset NRMSE and the M2 NRMSPE. This slight NRMSE decrease occurred again in the average 2 NRMSPE, when both CWSI and ET_a were available for selection, but not in the full dataset NRMSE. These results indicate that adding a measure of crop water use or stress, whether CWSI or ET_a, the EC model prediction accuracy increased modestly. This increase in accuracy is similar to those values found for the 2018 analysis (Table 9), where adding a crop water use or stress metric also slightly improved the EC prediction accuracy.

Table 12: Average Normalized Root Mean Square Error (NRMSE) over all field and Average Muth 2 Normalized Root Mean Square Prediction Error (RMSPE) of EC for 2019 EC models, separated by transformations used. Eleven models were evaluated, each with n = 411 data points. Darker red cells indicate lower error values.

	Explanatory Variables Available for Selection		
	NDVI and CRSI Only	NDVI, CRSI and Either CWSI or ETa	NDVI, CRSI, CWSI and ETa
Average NRMSE	0.27	0.25	0.25
Average M2 NRMSPE	0.31	0.29	0.27

Data Transformation Type

Eight of the ten model parameter sets (EC type being predicted, growth stage used, data transformation type, interpolation or not) formed two pairs of four parameter sets in which the same parameters formed the EC model except a variation by data transformation type. Four EC models included two equations with all explanatory variables log-transformed and two equations with no data transformations, while the other four EC models included a combination of log-transformed and untransformed variables. Results in Table 13 reveal that in 2019 there was no discernable improvement in model RMSE or RMSPE of EC when some dependent variables

were transformed (but not all), in contrast to a model in which all variables were subject to the same transformation (all log-transformed or none transformed). This analysis considered only four equations (each with a sample size of n = 411 pixels), hence the sample size of n = 4 is small. This result was different than the results found for the 2018 data analysis (Table 10), in which using a combination of data transformations improved the EC model prediction, although with only six models evaluated.

Table 13: Average Normalized Root Mean Square Error (NRMSE) over all fields and Average Muth 2 Normalized Root Mean Square Prediction Error (RMSPE) for 2019 EC models, separated by transformations used. Four models are evaluated in this table, each with a sample size of n = 411 data points. Darker red cells indicate lower error values.

Variables Available for Selection	NDVI and CRSI Only		NDVI, CRSI and Either CWSI or ETa		NDVI, CRSI, CWSI and ETa	
Statistical Metric	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE
All log-transformed or not transformed	0.25	0.26	0.24	0.22	0.24	0.22
Combination of Log and No Transformation	0.25	0.26	0.24	0.22	0.24	0.22

Interpolation of Explanatory Variables Between Satellite Dates

Eight of the ten EC models formed two pairs of four EC models in which the same EC model parameters (growth stage, type of EC predicted, transformation used, interpolation or not) were used except for a variation by whether the data were interpolated between satellite dates before integrating. Four models used data that were interpolated between satellite dates, and four models were identical except they used data that were not interpolated between satellite dates. Removing interpolation of explanatory variables between satellite dates, and only integrating the data from satellite overpasses generally slightly lowered the NRMSE and average M2 NRMSPE

statistics, as shown in Table 14. This result indicates that interpolating explanatory variables between satellite overpass dates did not improve EC model prediction accuracy and therefore it should not be used in future maize EC prediction modeling. This result mirrors the results found in the 2018 data analysis (Table 11).

Table 14: Average Normalized Root Mean Square Error (NRMSE) over all fields and Average Muth 2 Normalized Root Mean Square Prediction Error (RMSPE) of EC for 2019 EC models, separated by whether the data were interpolated between satellite dates. Eight models are evaluated in this table, each with a sample size of n = 411 data points. Darker red cells indicate lower error values.

Variables Available for Selection	NDVI and CRSI Only		NDVI, CRSI and Either CWSI or ET_a		NDVI, CRSI, CWSI and ET_a	
Statistical Metric	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE	Average NRMSE	Average M2 NRMSPE
Interpolation	0.27	0.27	0.25	0.28	0.24	0.22
No Interpolation	0.25	0.26	0.24	0.23	0.24	0.23

Best EC Prediction Equations in 2019

The best-performing EC equation [Eq. (22)] for the 2019 dataset was:

$$ln(EC_{eg}) = 178.10 + 259.87(ln(CWSI)) - 179.29(ln(CRSI)) - 183.85(ln(NDVI)) - 265.12(ln(CWSI))(ln(CRSI)) - 270.11(ln(CWSI))(ln(NDVI)) + 186.09(ln(CRSI))(ln(NDVI)) + 275.80(ln\,(CWSI)(ln(CRSI))(ln(NDVI)) \quad (22)$$

Eq. (22) predicts $\ln(EC_{eg})$ using data obtained from the crop mid growth stage, and only including data acquired from satellite overpass dates (i.e., no interpolation between satellite overpass dates), with a sample size of n = 411. Variables available for model selection included CWSI, CRSI and NDVI plus their interactions. Although Eq. (22) is long, the AIC method does take the number of terms into account when ranking models, and weighs models with many

terms negatively. Eq. (22) resulted in an EC_{eg} RMSE of 0.59 dS m^{-1} (NRMSE of 23% of average EC_{eg}) for the entire dataset, and an RMPSE of 0.71 dS m^{-1} (NRMSE of 21% of M2 average EC_{eg}) for M2 LOFO prediction. For reference, the studies reviewed in the Introduction chapter had EC_e RMSE values between 1.15 and 3.63 dS m^{-1}. The MBE value was 0.06 dS m^{-1}. A 1:1 scatter plot of EC_{eg} predicted vs observed values from Eq. 22 is provided in Figure 28, and a 1:1 density plot is provided in Figure 29.

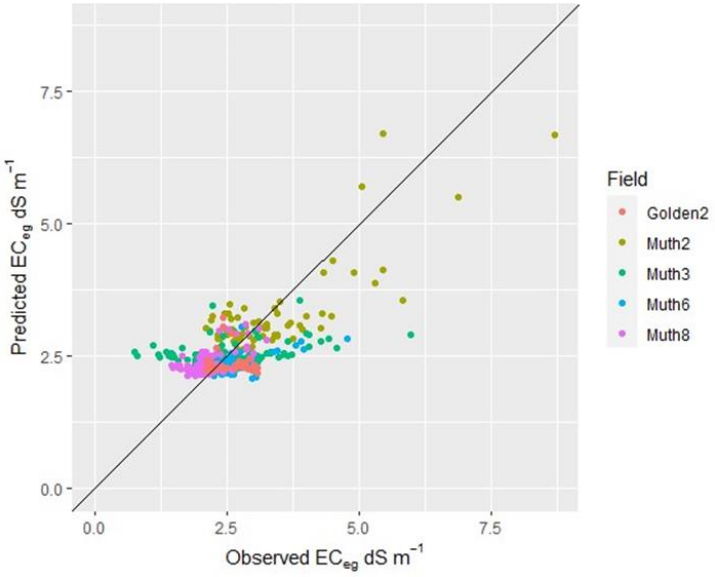

Figure 28: A 1:1 scatter plot of the 2019 EC_{eg} model-predicted [Eq. (22)] vs. observed EC_{eg} values. Eq. (22) was the best statistically-performing equation in 2019 and was developed using data from the mid growth stage from satellite overpass dates only. The sample size is n = 411.

In both the 2018 and 2019 dataset analyses, using explanatory variables obtained from the crop development growth stage showed higher correlation with EC than from other growth stages. For this reason, data from the development stage was used in an EC prediction model. Using the 2019 dataset alone, these high correlations were obtained with EC_a. EC_a is affected by a suite of parameters including water content and other transient properties and is not a direct metric of

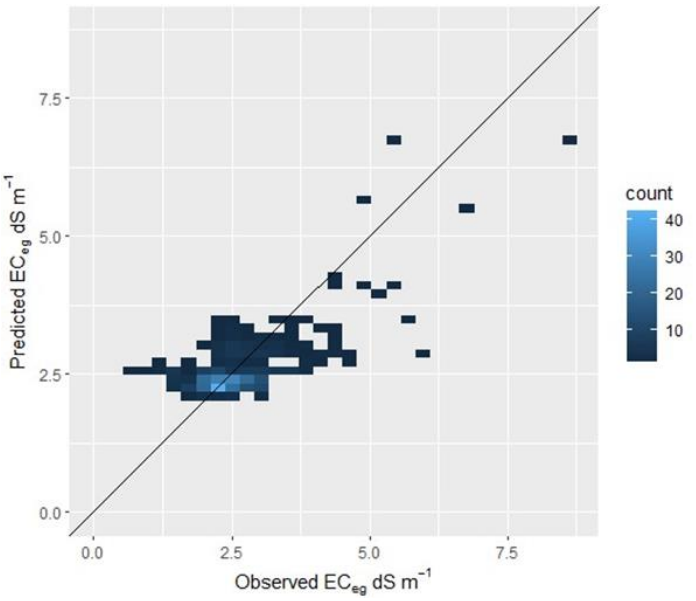

Figure 29: A 1:1 density plot of the 2019 model-predicted [Eq. (22)] vs. observed EC_{eg} values. Eq. (22) was the best statistically-performing equation in 2019 and was developed using data from the mid growth stage from satellite overpass dates only. The sample size is n = 411.

salinity level (Corwin and Lesch, 2005; Gavlak et al., 2005; Rhoades et al., 1990; Stiven and Khan, 1966; United States Salinity Laboratory Staff, 1954). Thus, an EC_a model may not provide sufficient robustness to effectively predict soil salinity levels. However, EC_a may be useful for directing field-level soil salinity sampling, similar to how the EM38 measures EC_a and then results from soil sample are used to account for other transient properties in calibrating EC_a to EC_e (see the *Soil Electrical Conductivity* section of Chapter 2). An equation predicting EC_a using data from the crop growth development stage, that only included data from satellite overpasses, produced the second-best suite of statistics:

$$EC_a = 1.70 + 0.43(CWSI) - 0.99(NDVI) \quad (23)$$

Variables available for selection in this EC_a model included CWSI, CRSI and NDVI plus their interactions. Eq. (23) resulted in an RMSE of 0.17 dS m^{-1} (NRMSE of 23% of average EC_a) for the entire dataset, and an RMPSE of 0.26 dS m^{-1} (NRMSE of 26% of M2 average EC_a) for M2

LOFO prediction. The MBE was 0.00 dS m^{-1}. A 1:1 scatter plot of Eq. (23) predicted vs observed EC_a values is found in Figure 30, and a 1:1 density plot is found in Figure 31.

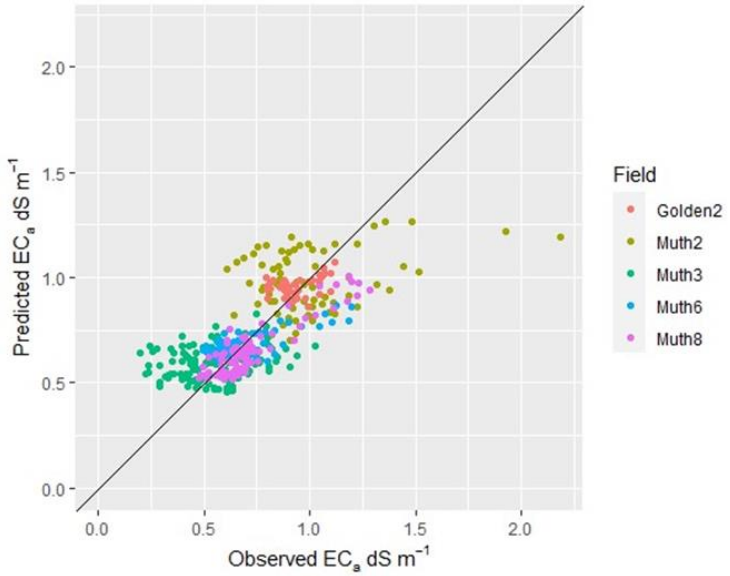

Figure 30: A 1:1 scatter plot of the 2019 model-predicted [Eq. (23)] vs. observed EC_a values. Eq. (23) was developed using data from the development stage from satellite overpass dates only. The sample size was n = 411.

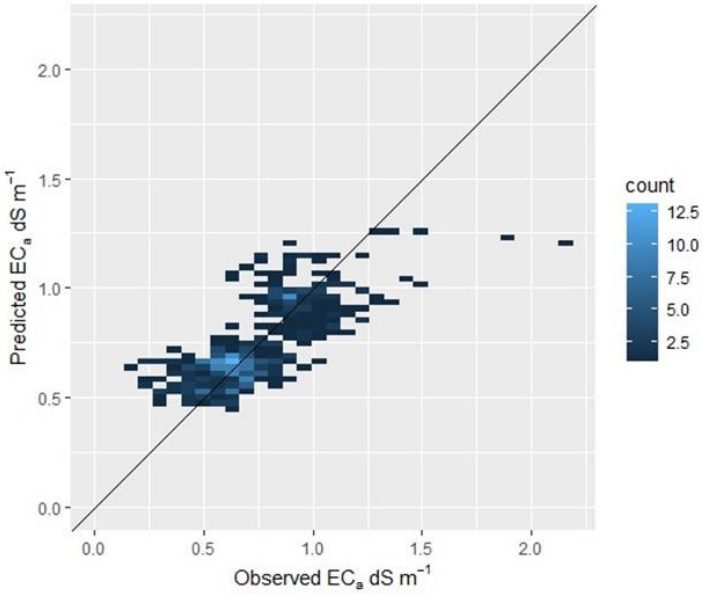

Figure 31: A 1:1 density plot of 2019 model-predicted [Eq. (23)] vs. observed EC_a values. The color of the pixel indicates the number of data points in that area. Eq. (23) was developed using data from the development stage on satellite overpass dates only. The sample size was n = 411.

When using data interpolated between satellite dates, but with the same independent variables available for selection, $\ln(ET_a)$ was chosen over CWSI using the AIC ranking. For this EC_a model, the RMSE increased from 0.17 to 0.18 dS m^{-1} but the RMSPE for M2 LOFO decreased from 0.26 to 0.22 dS m^{-1}. Although these two statistics are on par with the EC_a model shown in Eq. (23), this model was deemed not as good of an EC_a predictor as Eq. (23) because the average LOFO RMSPE of all fields increased from 0.20 ds m^{-1} to 0.24 ds m^{-1}.

Piecewise Linear Regression of EC

Upon visual inspection of the 1:1 plots created from Eq. (22), there is an apparent trend change at about 4 dS m^{-1}. Therefore, a piecewise linear regression to predict EC_{eg} was created using a data split at 4 dS m^{-1}. Similar to Eq. (22), this EC_{eg} model used data from the mid growth stage to predict $\ln(EC_{eg})$, with only data from satellite overpass dates (i.e., no interpolation). The same explanatory variables were available for selection (ln(CWSI), ln(CRSI), ln(NDVI) and their interactions). There were 391 data points with EC_{eg} < 4 dS m^{-1} and 20 data points with EC_{eg} > 4 dS m^{-1}. The resulting piecewise EC equation is:

$$\ln(EC_{eg}) = -15.28 + 17.16[\ln(CRSI)] + 19.64[\ln(NDVI)] -$$

$$20.17[\ln(CRSI)][(\ln(NDVI)] \qquad for\ EC_{eg} < 4\ dS\ m^{-1} \qquad (24a)$$

$$= 2.36 - 1.04[\ln(NDVI)] \qquad otherwise \qquad (24b)$$

When comparing the errors of modeling EC_{eg} using Eq. (22) to the errors resulting from the piecewise Eq. (24), the RMSE of the predictions decreased from 0.59 dS m^{-1} (NRMSE of 19% of average EC_{eg}) to 0.52 dS m^{-1} (NRMSE of 20% of average EC_{eg}). The M2 EC_{eg} estimation RMSPE decreased from 0.71 dS m^{-1} (NRMSE of 21% of average M2 EC_{eg}), to 0.61 dS m^{-1} (NRMSPE of 18% of average M2 EC_{eg}). The MBE value slightly decreased from 0.06 dS m^{-1} to 0.05 dS m^{-1}. When comparing the individual parts of the piecewise equation to those from the

single equation, the RMSE of data with EC_{eg} values below 4 dS m^{-1} stayed the same at 0.49 dS m^{-1} (NRMSE of 20% of average EC_{eg}). The MBE of data with EC_{eg} values below 4 dS m^{-1} increased from 0.01 dS m^{-1} when using one equation [Eq. (22)] to 0.05 dS m^{-1} when using the piecewise equation [Eq. (24)]. The RMSE of data with EC_{eg} above 4 dS m^{-1} decreased from 1.52 dS m^{-1} (NRMSE of 30% of average EC_{eg}) using a single equation [Eq. (22)] to 0.87 dS m^{-1} (NRMSE of 17% of average EC_{eg}) using a two-part equation [Eq. (24)]. The MBE of data with EC_{eg} values above 4 dS m^{-1} decreased from 1.18 dS m^{-1} when using one equation [Eq. (22)] to 0.06 dS m^{-1} when using the piecewise equation [Eq. (24)]. For reference, the RMSE of studies reviewed in the Introduction had EC_e RMSE values between 1.15 and 3.63 dS m^{-1}. The 1:1 plots of model predicted vs observed EC_{eg} values for Eq. (24) are found in Figure 32 and Figure 33.

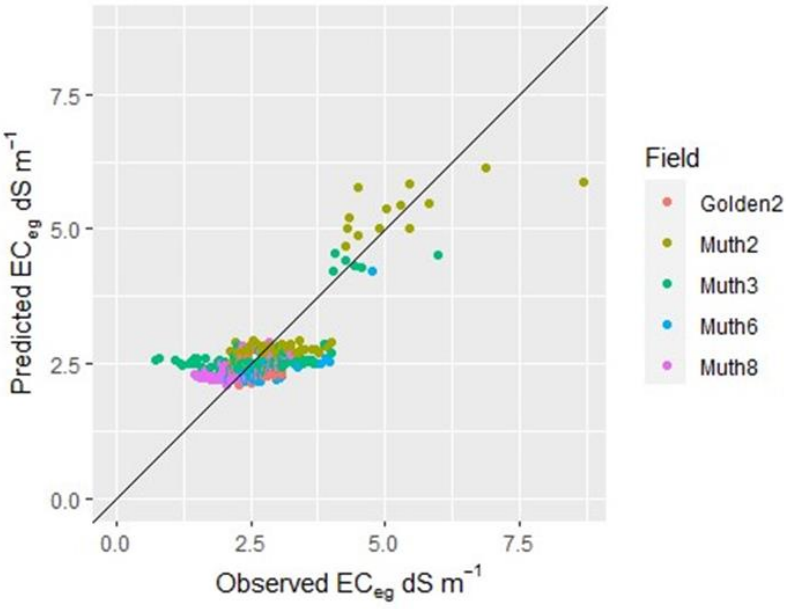

Figure 32: A 1:1 scatter plot of the 2019 EC_{eg} model-predicted [Eq. (24)] vs. observed EC_{eg} values. Eq. (24) was developed using data from the mid crop growth stage on satellite overpass dates only. The sample size is n = 411.

Upon visual inspection of the 1:1 plots for Eq. (23), there is an apparent trend change at about 1.15 dS m^{-1}. Therefore, a piecewise linear regression to predict EC_a was created with the data

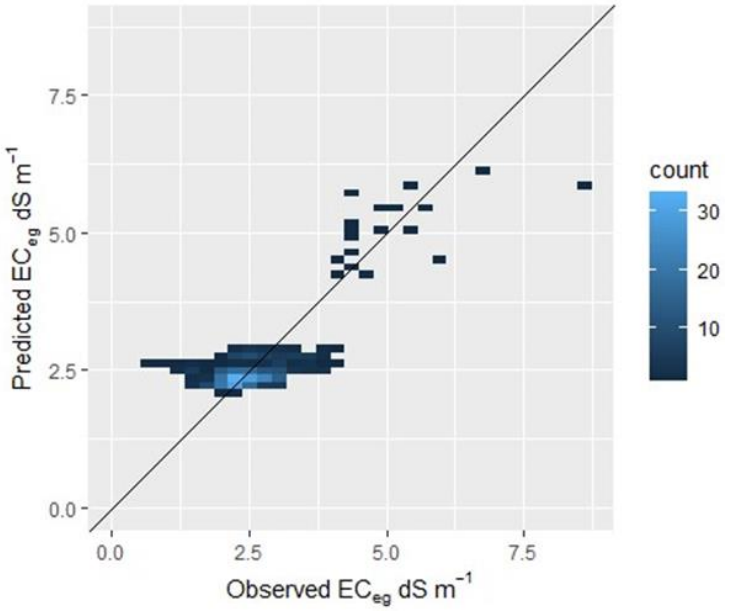

Figure 33: A 1:1 density plot of 2019 model-predicted [Eq. (24)] vs observed EC_{eg} values. The color of the pixel indicates the number of data points in that area. Eq. (24) was developed using data from the mid growth stage on satellite overpass dates only. The sample size is n = 411.

split at 1.15 dS m^{-1}. Similar to Eq. (23), this EC_a model used data from the development growth stage to predict EC_a, and from satellite overpass dates only (i.e., no interpolation). The same explanatory variables were available for selection (CWSI, NDVI, and their interactions). There were 390 data points with EC_a < 1.15 dS m^{-1} and 21 data points with EC_a > 1.15 dS m^{-1}. The resulting piecewise EC_a equation is:

$$EC_a = 1.45 + 0.40(CWSI) - 0.77(NDVI) \quad for\ EC_a < 1.15\ dS\ m^{-1} \qquad (25a)$$

$$= 2.58 - 1.35(NDVI) \quad otherwise \qquad (25b)$$

When comparing these new results to the results from Eq. (23), the RMSE of the model predicted EC_a values decreased from 0.17 dS m^{-1} (NRMSE of 23% of average EC_a) to 0.15 dS m^{-1} (NRMSE of 20% of average EC_a). The M2 EC_a RMPSE value decreased from 0.26 dS m^{-1} (NRMSE of 26% of M2 average EC_a) to 0.10 dS m^{-1} (NRMSE of 19% of M2 average EC_a). The new MBE remained 0.00 dS m^{-1}. When comparing the individual parts of the piecewise

equation to those from the single equation, the RMSE of data with EC_a values below 1.15 dS m^{-1} decreased from 0.15 dS m^{-1} (NRMSE of 21% of average EC_a) using a single equation [Eq. (23)] to 0.14 dS m^{-1} (NRMSE of 20% of average EC_a) using a two-part equation [Eq. (25)]. The MBE of data with EC_a values below 1.15 dS m^{-1} decreased from 0.33 dS m^{-1} when using one equation [Eq. (23)] to 0.00 dS m^{-1} when using the piecewise equation [Eq. (25)]. The RMSE of data with EC_a above 1.15 dS m^{-1} decreased from 0.39 dS m^{-1} (NRMSE of 29% of average EC_a) using a single equation [Eq. (23)] to 0.21 dS m^{-1} (NRMSE of 16% of average EC_a) using a two-part equation [Eq. (25)]. The MBE of data with EC_a values above 1.15 dS m^{-1} increased from -0.02 dS m^{-1} when using one equation [Eq. (23)] to 0.00 dS m^{-1} when using the piecewise equation [Eq. (25)]. For reference, the RMSE of studies reviewed in the Introduction had EC_e RMSE values between 1.15 and 3.63 dS m^{-1}. A 1:1 scatter plot of predicted vs observed EC_a values for Eq. (25) is found in Figure 34, and a 1:1 density plot is found in Figure 35.

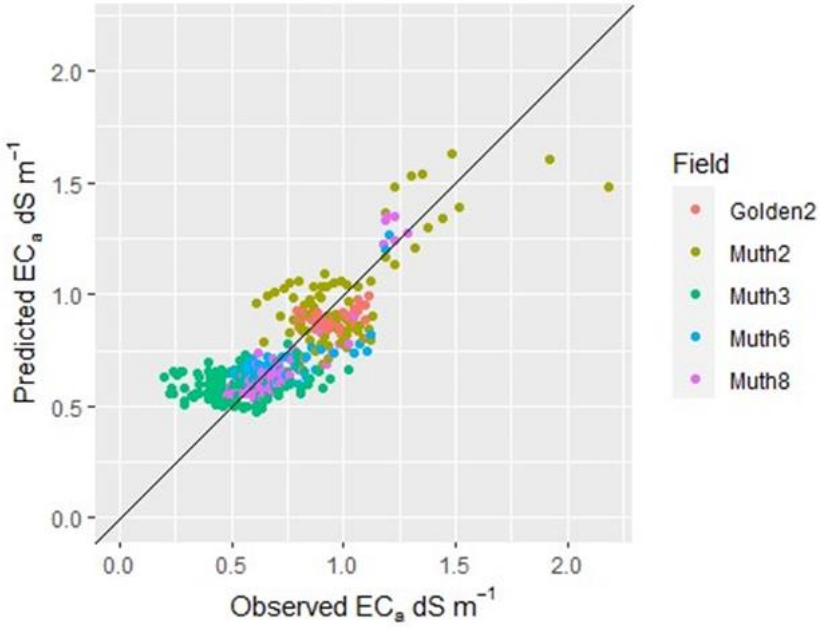

Figure 34: A 1:1 scatter plot of the 2019 model-predicted [Eq. (25)] vs. observed EC_a values. Eq. (25) was developed using data from the development stage on satellite overpass dates only. The sample size was n = 411.

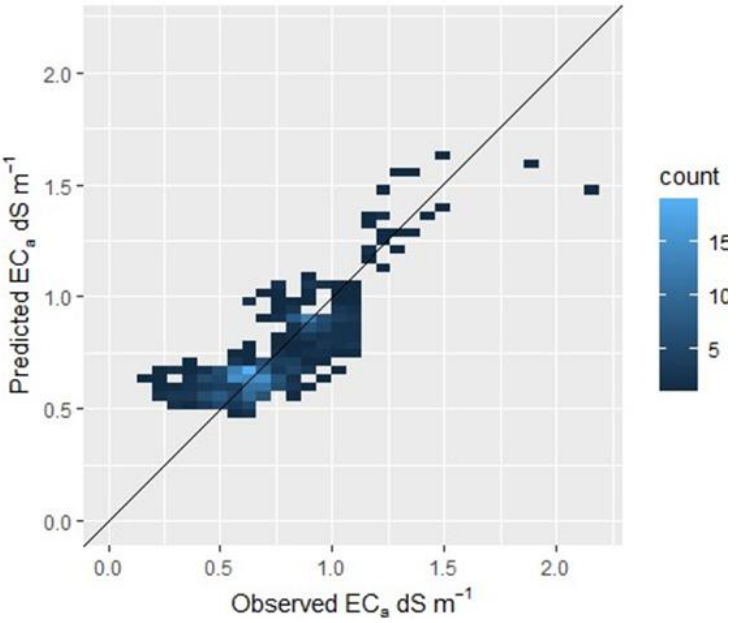

Figure 35: A 1:1 density plot of 2019 model-predicted [Eq. (25)] vs. observed EC_a values. The color of the pixel indicates the number of data points in that area. Eq. (25) was developed using data from the development stage on satellite overpass dates only. The sample size was n = 411.

CHAPTER 4: DISCUSSION

Crop Root Zone Salinity

Based on the soil salinity classifications given by Metternicht and Zinck (1997) and the 2018 and 2019 EM38 surveys, there are areas within the FDD that classify as moderately saline, although most of the FDD is non-saline or slightly saline. The few points evaluated for sodicity do not indicate a high level in the FDD.

Change in mean EC_e and EC_{eg} derived from the EM38 from the summer of 2018 to the summer of 2019 in the three fields that were surveyed in both years (M2, M3 and M8) was very small (Table 4), considering that EM38 measurements are subject to an estimated uncertainty of up to +/- 1.3 dS m^{-1} (Brown, 2021). The maximum EC_e and EC_{eg} in the fields studied increased from 2018 to 2019 by a more substantial amount, 0.60 and 0.55 dS m^{-1}, respectively, yet still less than the estimated +/- 1.3 dS m^{-1} uncertainty level for measurements, according to Brown (2021). This indicates that areas that are prone to significant salinity accumulation might be continuing to see this accumulation. The three studied fields (M2, M3 and M8), particularly the southern end of M2, need drainage rehabilitation such as replacing old tiles with perforated corrugated PE flexible pipes in order to lower water table levels, reduce waterlogging conditions and salinity, and therefore promote better conditions to produce better crop yields.

Effect of Crop Growth Stages

When selecting the crop growth stage and type of EC that correlated most strongly with each combination of year, interpolation, and explanatory variable (e.g., “2018, not interpolated, ln(NDVI)”), the development stage resulted in the best correlation in 27 of the 32 combinations.

This result strongly indicates that the crop development stage was the best stage to collect data for predicting soil salinity (EC_e or EC_{eg}) or EC_a in the FDD maize fields. The importance of using the development growth stage for EC mapping was further confirmed in the EC modeling, when two of the three top-performing equations included data from the crop development stage. The development (rapid growth, vegetative) stage has been shown to be an ideal time in the growing season to deficit irrigate, i.e., limit the amount of water available to the crop to conserve limited water supplies with minimal reduction in yield (Comas et al., 2019). When deficit irrigation occurs, the stress on the crop increases. Stress can also increase by introducing salinity to the available water. Therefore, it is counterintuitive to think that the development stage is an ideal growth stage to predict salinity in the root zone, since the effects of increased stress during the development stage do not affect yield as significantly as during other growth stages. One explanation of increased EC modeling capability during the development stage is that it may be early enough in the season to where the maize roots are not dense and deep enough to sufficiently access upflux from the shallow water table observed in the FDD, yet late enough in the season that the crop has enough vegetative cover to significantly influence the nadir satellite reflectance measurements.

2018 vs. 2019 Data Quality

Graduate students were the primary researchers and data collectors with significant assistance from undergraduate students. Students learned and were better prepared for data collection during the second year of the study. Therefore, the data from the second year of the study is a more robust dataset. Field data collection methods require familiarity with field conditions and equipment. In the second year of data collection, all tasks were conducted more smoothly and

efficiently by the graduate and undergraduate students. For example, in 2019, all researchers knew better how to hold the EM38 over the fields in a consistent manner, which is necessary in order to get consistent electromagnetic readings of the soil profile. Through practice, researchers were able to dig holes for soil samples and soil moisture sensors without disturbing the surrounding soil column and causing topsoil to fall into the hole. In the lab, through practice, researchers and technicians possibly became more consistent with the water content used to create soil paste extract samples as more samples were analyzed. Particularly, the consistency at which 2019 correlation coefficients were highest using data from the crop growth development stage points to the likelihood of it being the best stage for modeling maize soil EC. Any future modeling of maize soil EC for salinity mapping and monitoring should include, similar to this study, a rigorous exploration of the importance of the crop rapid growth or development stage for assessment of soil salinity using remote sensing techniques.

Specifically, researchers and field technicians should become familiar with all field techniques before collecting the final datasets for EC model calibration. For EM38 use, users should be able to identify when soil moisture content in fields is at or just below field capacity and need to be able to always walk the fields with the EM38 at a consistent angle and height above the fields. Anyone collecting EM38 data needs to be consistent in whether they are holding the EM38 over the furrow or over the bed. When installing soil moisture sensors, it is important to dig the hole in a manner that leaves the soil profile intact with no air gaps in the side of the hole when installing the TDR soil water content sensors. Practice digging holes for soil sampling is also recommended, so that soil from the top of the soil profile does not collapse into the hole when collecting soil at depth. Flumes to measure applied irrigation flows and volumes must be installed in areas that will not be affected by backwater from downstream culverts, which can

become clogged. Flumes also must be installed to ensure that surface water does not flow under or around the flume.

Need for Data Transformation and Interpolation

A normal distribution in the dependent variable dataset is a requirement of MLR modeling, a condition which often is not met with field data. In this study, transformations were required to achieve the data distribution normality requirement for MLR. The natural log transformation was performed to achieve normality in the EC data distribution, as has been successfully used in many other salinity prediction or evaluation studies (Corwin and Lesch, 2014; Aldabaa et al., 2015; Corwin and Lesch, 2017; Brown, 2021).

The natural log transformation was applied either to all the independent and dependent variables, or to only certain variables based on their correlation to EC. In 2018, the RMSE of modeled EC decreased when each independent variable was selectively transformed, rather than transforming all variables. However, in 2019, there was no difference between the two methods. The differing results across the two years indicate that each independent variable should be analyzed for transformation type, rather than assuming that all variables should be under the same transformation.

Interpolating temporal explanatory variables between satellite overpasses may be desirable in order to better characterize the crop growth and water use throughout most of the crop growing season. However, in both 2018 and 2019, interpolation of independent variable values between satellite overpass dates did not improve the RMSE of the EC model predictions.

Addition of Crop Water Use or Crop Water Stress Metrics to Salinity Prediction Models

In 2018 and 2019, average modeled EC NRMSE decreased only by 1% and 2%, respectively, when adding either ET_a or CWSI to the explanatory (independent) variables available for selection, and was lowered by only 1% and 0%, respectively, again when both metrics were available for selection. In M2 specifically, where the highest EC values were measured, this decrease in EC NRMSE was slightly more substantial, i.e., 3% in 2018 and 2% in 2019 when adding either ET_a or CWSI and an additional 0% and 2%, respectively, when adding both metrics. This is a key finding of this study. Both NDVI and CRSI are proven good predictors of soil salinity within the maize root zone. In the maize fields found in the FDD, adding actual crop water use (ET_a) or a water stress metric (CWSI) only marginally improved the EC model accuracy. This marginal improvement may stem from the demonstration in Table 8 that NDVI and ET_a are significantly correlated. Adding ET_a to a model that already contains NDVI may not provide the model with significantly more predictive power in the explanatory variables. Future soil salinity (EC) prediction models, applied to irrigated agricultural settings, should incorporate a crop water stress or water use metric to improve the EC model prediction accuracy, but should explore other water stress or water use metrics to see if the increase in predictive accuracy can be increased from that found in this study.

Ability of Spectral Data to Predict Soil Salinity

Previous studies have shown that it is feasible to predict crop root zone soil salinity using multi-spectral radiation responses from vegetated agricultural surfaces (Metternicht & Zinck, 1997; Bouaziz et al., 2011; Scudiero et al., 2014; Aldabaa et al., 2015; Davis et al., 2019; Zhu et al., 2021; Aboelsoud et al., 2022) This work confirms the theory that spectral surface radiation

responses can assist in predicting soil salinity. Then, adding a crop water use or stress metric (in this study, ET_a or CWSI) to vegetation indices derived from remote sensing may improve the predictive capability of soil salinity models. One significant caveat remains, in that these EC models were developed using data from an area with a shallow groundwater table (< 1.27 m) (Figure 16). A shallow groundwater table will significantly affect crop transpiration and water stress. Therefore, the predictive ability of ET_a and CWSI in MLR models predicting EC may change if a future study uses the explanatory variables discussed in an area with a deeper groundwater table.

In models created using data from 2018, the best-performing piecewise equation was able to predict EC_e to within 0.44 dS m^{-1}, or 16% of average EC_e. In models created using data from 2019, one best-performing piecewise equation was able to predict EC_{eg} to within 0.52 dS m^{-1}, or 20% of average EC_{eg}, while the other was able to predict EC_a to within 0.15 dS m^{-1}, or 20% of average EC_a. All of these EC NRMSE values are well within the accuracy required to make field management decisions, since significant maize yield changes occur over a range of several dS m^{-1} (Ayers and Westcot, 1985; Davis et al., 2019; Cui et al., 2022). Very few studies have used RS data to create EC_{eg} prediction models, which are important because an estimated 20% of salt-affected land has gypsum present in the soil (Brown, 2021), and RS-derived models can be applied over a large spatial extent at multiple times per month, unlike expensive and labor-intensive field-level data collection.

Uses for This Work

Multi-field soil salinity maps derived from multispectral satellite imagery can aid in many aspects of farm management (Anderson et al., 2012), including crop selection (Scudiero et al.,

2016), economic analyses (Qadir et al., 2014; Osborn et al., 2017), water budgets (Anderson et al., 2012), the placement of tile drains (Aboelsoud et al., 2022), modeling of non-point source pollution to rivers (Corwin et al., 1998), delineation of site-specific management units for salinity management (Scudiero et al., 2016), monitoring of degraded water reuse (Khan & Abbas, 2013), and monitoring of the impacts of climate change (i.e. greater occurrence of extended drought or flash flooding) on soil salinity (Corwin, 2021). Although not as relevant to fields in Colorado, identifying soil salinity can also help in expanding the use of brackish coastal groundwater to supplement freshwater supplies (Zhu et al., 2021). These planning benefits are available both to the individual farmer managing a few fields, or for a district manager, who needs to plan and size inter-farm drainage solutions for long-term irrigation management.

Maps of EC_e and EC_{eg} are excellent tools for monitoring salinity buildup because they can be repeated many times for relatively little cost. Landsat satellites fly over North America every sixteen days (combining data from two satellites results in an image every eight days), and the data are free to download off the USGS website. Once an EC_e or EC_{eg} model has been calibrated to a specific region, the model can be run using the most recent satellite imagery and the local agricultural weather station. Through this repeated estimation of salinity, farm managers can identify trends in salinity buildup or reduction and plan their reclamation projects accordingly. However, the ability to accurately detect changes in soil salinity over short time periods (< 1 year, when changes in EC are less than 1 dS m^{-1}) remains unproven.

Specific to the FDD, it appears that the subsurface tile drains are not as efficient as intended, since the groundwater table is closer to the surface in the lower parts of the tile drain system. This study and others like it can help identify areas where tile drain rehabilitation should focus. Since tile drain rehabilitation can be an expensive endeavor, FDD field managers may want to

begin their work rehabilitating tile drains near M2 and in the center line of the tile drain network, before expanding out to the areas where salinity is not as great of a concern.

The EC models shown in this study were developed using field data and remote sensing data collected over an area with a shallow groundwater table (< 1.27 m) (Figure 16). A shallow groundwater table will significantly affect crop transpiration and water stress. Therefore, the predictive ability of ET_a and CWSI in MLR models predicting EC may change if a future study uses these explanatory variables in an area with a deeper groundwater table.

CHAPTER 5: RECOMMENDATIONS AND CONCLUSION

Recommendations for Future Work

Crop Water Use Calculations

When developing models to predict soil salinity, in agricultural areas, it is important to estimate and understand the temporal changes in crop water stress in the fields being monitored so that when a crop displays stress the modeler can identify whether the stress is salinity- or water deficit-driven stress. When conducting a soil water balance using soil water content sensors' data, from TDR sensors for instance, to determine periods of crop water stress, future studies should include deep percolation calculations such as those that can be done with a water flux model such as Root Zone Water Quality Model 2 (USDA ARS, 2023) Deep percolation is a significant term in water balances of fields that are flood-irrigated, and if ignored, will add substantial inaccuracy to the water balance calculation and water stress periods (deficits beyond those allowed for the crop of interest) may not be determined well. Water flux models often outperform water balance models but require more stringent soil hydraulic properties (Liu et al., 2006), which were not available for this study. If the researcher instead chooses to use a water balance using flumes to measure applied water and runoff, combined with ET_a measurements, it is important to confirm that the flumes are sized properly for the anticipated flows, and that the flumes are installed securely to the ground to withstand a season's worth of stressors including high flows, rodents chewing on cords or zip ties, sun exposure on plastics, and other stressors. Also, researchers should make sure that the flumes are not in an area that will receive backwater from nearby culverts. In this study, flumes on the tail-side of M2 reached critical flows in 2019

(but not in 2018), so a water balance calculation was not able to be performed in that year, as the researchers did not know the volume of water directly running off of the fields.

Remotely sensed ET_a is useful in determining crop water use over large spatial areas throughout the entire crop growth season. ReSET-based ET_a estimates' accuracy has been evaluated on healthy, relatively unstressed crops. These ET_a error estimations, found in Elhaddad et al. (2011), include calculating ET_a with ReSET to within 5% of Penman-Monteith and K_c-derived ET_a in an alfalfa field, and to within 14% of a sorghum lysimeter (Elhaddad et al., 2011). However, ReSET products have not been evaluated in areas of high salinity stress, such as the saline areas of the FDD. Therefore, it is unclear whether ReSET is the ideal source of ET_a for predictive modeling of root zone salinity. Furthermore, it is unclear whether the satellite images contained true hot and cold (i.e., extreme low and high ET_a in nearby agricultural areas) conditions. It is recommended to run ReSET on areas of at least 500 km^2 (as done in this study) to increase the chances of including true extreme high and low ET_a conditions. However, it is advantageous to use ReSET in this study and similar studies, since it uses readily-available data (public satellite imagery and public weather data). When calculating ET_a using ReSET or any other energy balance approach that relies on extreme pixel selection (such as METRIC or SEBAL), the heat flux assumption of $H = R_n - G_0$ only applies if no rain events have occurred within a week prior to the image acquisition date. If a rain event has occurred, LE needs to be estimated at that pixel through a soil water balance for ET_a, and placed in the full equation, $H = R_n - G_0 - LE$.

Another method to calculate ET_a is a near-surface energy balance using the aerodynamic temperature derived from canopy temperature using near-surface IRTs, rather than satellite imagery (Costa-Filho et al., 2021). This surface energy balance would use the aerodynamic temperature derived from canopy temperature from field-level IRTs, and vegetation indices such

as NDVI and OSAVI, from Landsat 7 and 8, and near-field weather data. Unfortunately, the near-surface energy balance model was not used in this study because only a few IRTs were installed which did not cover a significant portion of the fields of interest. If properly implemented, ET_a could be calculated using this near-surface energy balance at a higher spatial resolution than the 100 m satellite pixel size. These ET_a data could aid in developing a soil EC predictive model with finer spatial resolution, which could better capture the spatial dynamics of root zone salinity in fields with highly spatially-varying conditions (water table depth, applied water, gypsum concentrations, fertilizer applications, etc.).

Recommendations for Future Different Studies

Future studies that aim to diverge from this study may want to combine multiple seasons of field data into a single model to predict salinity. It would be interesting to compare the soil EC models developed in this study to EC models created from multiple years of data but only a single EC survey, so that the data from the least-stressed year at any given pixel is input to the model. This may allow the model to parse out the effects of more transient stressors including weeds, water stress, planting errors, and others that vary more significantly than salinity from year to year and would be similar to the work conducted by Lobell et al. (2010), Scudiero et al. (2015), and Zhang et al. (2015). This method of creating an EC model with multiple years of satellite imagery but only one EC survey was not used, because a paired t-test showed that the two EC_e datasets were statistically different at $p < 0.05$, meaning that one "measured" EC_e dataset could not be used to model EC spatially for both years.

Boundary edge modeling (Corwin and Scudiero, 2019) could also help parse out stress from sources other than salinity. In this technique, the yield vs. salinity relationship is plotted and only the points at the top of that point cloud (indicating high yield relative to salinity) are used. In this

way, only the healthiest crops in a given salinity class are input into the model. This study did not use boundary-edge modeling because the purpose of boundary-edge modeling is to evaluate locations where salinity is by far the most dominant (or only) stressor. The location of this study, the FDD, has multiple crop stressors present, so boundary-edge modeling was not appropriate. Other studies have improved modeling accuracy by breaking the dataset by vegetation cover and modeling each vegetation class separately (Zhang et al., 2021). However, that was not investigated in this study because of the limited number of data points with low vegetation cover or with EC_e greater than 5 dS m^{-1}. A study that includes more areas of high salinity, limiting vegetation cover, could lend itself well to vegetation cover class modeling.

Although untested in the literature analyzed for this study, dividing the dataset by water table depth and modeling on each subset of the data may improve EC model accuracy. This study's dataset was limited in this regard, as all areas of study had a water table depth between 1.3 and 0.7 m depth. Relative to the root zone of maize (usually up to 1.2 m), there was no significant variation in the water table depth throughout our study region.

Conclusion

A simple, easily-repeatable model for frequently predicting crop root zone soil salinity is necessary in order to manage the ever-growing problem of salt accumulation in agricultural fields, particularly in arid or semi-arid environments. This mission to predict soil salinity will become even more pressing as salinity continues to accumulate in crop root zones, and the world's population and per-capita caloric intake continue to rise. Remote sensing lends itself well to this task, as remotely sensed imagery can have a large spatial scope and data collection can be repeated multiple times per month or even more frequently.

This study aimed to accomplish three objectives: 1) determine the correlation between ET_a and three forms of EC: EC_a, EC_e, and EC_{eg}; 2) create linear regression models to estimate EC in irrigated maize using multiple explanatory variables (e.g., ET_a rates, CWSI and standard vegetation indices); and 3) assess the usefulness of adding ET_a rates and CWSI to linear regression models made up of vegetation indices in predicting soil salinity in surface irrigated maize, for each growth stage as well as the total season.

Regarding objective 1, this study revealed that ET_a has a strong negative correlation with salinity. Using 2018 data, the correlation between ET_a and EC metrics was on par with the correlation between CRSI and EC metrics, but not as strong as the correlation between NDVI and EC metrics (Table 5, Table 7). Using 2019 data, this correlation was not as strong as that with either NDVI or CRSI with EC metrics (Table 6, Table 7). Similarly, CWSI has a strong positive correlation with salinity. Using 2018 or 2019 data, the correlation between CWSI and EC was not as strong as NDVI or CRSI (Table 5, Table 6, Table 7). ET_a generally had stronger correlation with EC than CWSI (Table 5, Table 6, Table 7). It is important to note that the data analyzed in this report were collected in an area with a shallow groundwater table (< 1.27 m) (Figure 16). A shallow groundwater table will significantly affect crop transpiration and water stress. Therefore, the correlations between ET_a, CWSI and EC may change if a future study uses these explanatory variables in an area with a deeper groundwater table.

Regarding objective 2, this study created linear regression models to predict EC_a, EC_e and EC_{eg}. The best statistically-performing single equation using 2018 data predicted EC_{eg} with an RMSE of 0.49 dS m^{-1} (18% of mean EC_{eg}). When breaking the data into two pieces and creating a piecewise linear regression, this model's RMSE decreased to 0.44 dS m^{-1}, or 16% of average EC_{eg}. The RMSE of data with EC_{eg} above 4 dS m^{-1} decreased from 1.05 dS m^{-1} (NRMSE of 21%

of average EC_{eg}) with a single equation [Eq. (20)] to 0.52 dS m^{-1} (NRMSE of 11% of average EC_{eg}) using a two-part equation [Eq. (21)]. Two equations were deemed the best statistically-performing equations using 2019 data. The first predicted EC_a with an RMSE of 0.17 dS m^{-1} or 23% of average EC_a. When this same model was created using piecewise linear regression, the RMSE decreased to 0.15 dS m^{-1} or 20% of average EC_a. The RMSE of data with EC_a above 1.15 dS m^{-1} decreased from 0.39 dS m^{-1} (NRMSE of 29% of average EC_a) using a single equation [Eq. (23)] to 0.21 dS m^{-1} (NRMSE of 16% of average EC_a) using a two-part equation [Eq. (25)]. Since EC_a is a measure of total soil conductivity and not salinity, a second equation was listed in the best statistically-performing equations. This equation predicted EC_{eg} with an RMSE of 0.59 dS m^{-1} or 23% of average EC_{eg}. When this same model was created using piecewise linear regression, the RMSE decreased to 0.52 dS m^{-1} or 20% of average EC_{eg}. The RMSE of data with EC_{eg} above 4 dS m^{-1} decreased from 1.52 dS m^{-1} (NRMSE of 30% of average EC_{eg}) using a single equation [Eq. (22)] to 0.87 dS m^{-1} (NRMSE of 17% of average EC_{eg}) using a two-part equation [Eq. (24)]. All of these RMSEs are well within the range necessary to make field-level management decisions. These RMSEs are also lower than many other EC prediction RMSEs derived from crop canopy reflectance (Ivushkin et al., 2019; Jiang et al., 2019; Scudiero et al., 2016).

Although NDVI and ET_a are strongly correlated (Table 8) they were both still included in MLR model. The two variables have slightly different physical analogues, with NDVI being a measure of greenness, and ET_a being a measure of crop transpiration (plus evaporation). Furthermore, many past studies have used explanatory variables that are strongly correlated (Bouaziz et al., 2011; Wei et al., 2020), giving precedent in past literature.

Regarding objective 3, it was determined that a metric of water use or stress such as ET_a or CWSI slightly increases the predictive ability of remotely sensed models for predicting crop root zone soil salinity in maize in shallow groundwater table areas of southeastern Colorado (Table 9 and Table 12). By adding either CWSI or ET_a to a model with explanatory variables of NDVI and CRSI, the RMSE of the entire model decreased by 1 to 2%. When focusing on the high salinity field M2, this decrease in RMSE was greater, at 2 to 3% of average EC. By adding both CWSI and ET_a to a model with NDVI and CRSI, the RMSE decreases were greater, at 2% of average EC in the whole model dataset, and 3 to 4% when focusing on the high salinity field M2. This marginal improvement may stem from the demonstration in Table 8 that NDVI and ET_a are significantly correlated. Adding ET_a to a model that already contains NDVI may not provide the model with significantly more predictive power in the explanatory variables. Future soil salinity (EC) prediction models applied to irrigated agricultural settings should incorporate a crop water stress or water use metric to improve the EC model prediction accuracy but should explore other water stress or water use metrics to see if the increase in predictive accuracy can be increased from that found in this study.

In this study, interpolating data between satellite images before creating the linear regression increased RMSE of predictions in the high salinity field M2 by up to 5% of mean EC (Table 11 and Table 14), and increased RMSE in the whole modeling area by up to 2% (Table 11 and Table 14). Therefore, it is recommended that future studies do not interpolate data between satellite overpass dates. By considering data transformations (in this study, the natural logarithm) on a variable-by-variable basis, rather than transforming all variables in the same manner, the RMSE of linear regression models of EC decreased by up to 2% of average EC (Table 10 and

Table 13). It is recommended that future studies evaluate data transformations on a variable-by-variable basis.

The majority of the strongest correlations between the explanatory variables of vegetation indices, ET_a, and CWSI and the independent variable EC occurred during the development stage (Table 7). One explanation of increased EC modeling capability during the development stage is that it may be early enough in the season to where the maize roots are not dense and deep enough to use water from the (shallow) water table, yet late enough in the season that the crop has enough vegetative cover to significantly influence the nadir satellite reflectance measurements. Therefore, when creating linear regression models to predict maize root zone soil salinity, researchers should focus on predicting soil salinity during the crop rapid growth or vegetative development stage, with an accurate assessment of the soil water balance and groundwater depth for use when interpreting results. Similar to this study, future research into this field must include consistent, rigorous measurements of the data needed for a soil water balance (precipitation, irrigation, soil moisture, ET_a, deep percolation, surface runoff), and further, researchers should be well-versed in their chosen tools used to measure EC. Satellite data of crop canopy reflectance should not be interpolated between satellite dates, and data transformations should be performed on a variable-by-variable basis. With these EC modeling suggestions in mind, this study showed that linear regression models created from crop canopy reflectance and using vegetation indices and water metrics such as ET_a and CWSI can successfully predict maize root zone soil salinity to well within the accuracy need for agricultural management decisions.

REFERENCES

Abbas, A., Khan, S., Hussain, N., Hanjra, M. A., & Akbar, S. (2013). Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Physics and Chemistry of the Earth*, *55–57*, 43–52. https://doi.org/10.1016/j.pce.2010.12.004

Aboelsoud, H. M., AbdelRahman, M. A. E., Kheir, A. M. S., Eid, M. S. M., Ammar, K. A., Khalifa, T. H., & Scopa, A. (2022). Quantitative Estimation of Saline-Soil Amelioration Using Remote-Sensing Indices in Arid Land for Better Management. *Land*, *11*(7), 1–19. https://doi.org/10.3390/land11071041

Aboitiz, M., Labadie, J. W., & Heermann, D. F. (1986). *Stochastic Soil Moisture Estimation and Forecasting for Irrigated Fields*. *22*(2), 180–190.

Acquaah, G. (2005). *Principles of Crop Production: Theory, Techniques, and Technology*. New York, New York: Pearson Prentice Hall.

Aldabaa, A. A. A., Weindorf, D. C., Chakraborty, S., Sharma, A., & Li, B. (2015). Combination of proximal and remote sensing methods for rapid soil salinity quantification. *Geoderma*, *239*, 34–46. https://doi.org/10.1016/j.geoderma.2014.09.011

Allbed, A., & Kumar, L. (2013). Soil Salinity Mapping and Monitoring in Arid and Semi-Arid Regions Using Remote Sensing Technology: A Review. *Advance in Remote Sensing*, *2*(December), 373–385. https://doi.org/10.4236/ars.2013.24040

Allen, R. G., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., … Robison, C. W. (2007a). Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Applications. *Journal of Irrigation and Drainage Engineering*, *133*(4), 395–406. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(395)

Allen, R. G., Tasumi, M., Morse, A., Trezza, R., Wright, J. L., Bastiaanssen, W., … Robison, C. W. (2007b). Satellite-Based Energy Balance for Mapping Evapotranspiration with Internalized Calibration (METRIC)—Model. *Journal of Irrigation and Drainage Engineering*, *133*(4), 395–406. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(395)

Anderson, M. C., Allen, R. G., Morse, A., & Kustas, W. P. (2012). Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sensing of Environment*, *122*, 50–65. https://doi.org/10.1016/j.rse.2011.08.025

Ayers, R. S. and Westcot, D. W. (1985). *Water Quality for Agricultura FAO 29*.

Bai, Y., Kissoudis, C., Yan, Z., Visser, R. G. F., & van der Linden, G. (2018). Plant behaviour under combined stress: tomato responses to combined salinity and pathogen stress. *Plant Journal*, *93*(4), 781–793. https://doi.org/10.1111/tpj.13800

Bannari, A., El-Battay, A., Bannari, R., & Rhinane, H. (2018). Sentinel-MSI VNIR and SWIR bands sensitivity analysis for soil salinity discrimination in an arid landscape. *Remote Sensing*, *10*(6), 1–20. https://doi.org/10.3390/rs10060855

Bastiaanssen, W. G. M., Menenti, M., Feddes, R., & Holstag, A. (1998). A remote sensing Surface Energy Balance Algorithm for Land (SEBAL). *Journal of Hydrology*, *212–213*, 198–212.

Bastiaanssen, W. G. M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J. F., Roerink, G. J., & Van Der Wal, T. (1998). A remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation. *Journal of Hydrology*, *212–213*(1–4), 213–229. https://doi.org/10.1016/S0022-1694(98)00254-6

Bouaziz, M., Matschullat, J., & Gloaguen, R. (2011). Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *Comptes Rendus - Geoscience*, *343*(11–12), 795–803. https://doi.org/10.1016/j.crte.2011.09.003

Brown, A. J. (2021). *Improving the Assessment of Soil-Water Salinity Hazard in Irrigated Fields Affected by Environmental Gypsum*.

Brutsaert, W. (1992). *Energy Budget to Determine Daily Evaporation (hi • od) -Bw (hi • d. 97*, 377–382.

Burkhalter, J. P., & Gates, T. K. (2005). Agroecological Impacts from Salinization and Waterlogging in an Irrigated River Valley. *Journal of Irrigation and Drainage Engineering*, *131*(2), 197–209. https://doi.org/10.1061/(asce)0733-9437(2005)131:2(197)

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdisciplinary Reviews: Computational Statistics*, *11*(3), 1–11. https://doi.org/10.1002/wics.1460

Comas, L. H., Trout, T. J., DeJonge, K. C., Zhang, H., & Gleason, S. M. (2019). Water productivity under strategic growth stage-based deficit irrigation in maize. *Agricultural Water Management*, *212*(November 2017), 433–440. https://doi.org/10.1016/j.agwat.2018.07.015

Corwin, D. L. (2021). Climate change impacts on soil salinity in agricultural areas. *European Journal of Soil Science*, *72*(2), 842–862. https://doi.org/10.1111/ejss.13010

Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and Electronics in Agriculture*, *46*(1–3 SPEC. ISS.), 11–43. https://doi.org/10.1016/j.compag.2004.10.005

Corwin, D. L., & Lesch, S. M. (2014). A simplified regional-scale electromagnetic induction - Salinity calibration model using ANOCOVA modeling techniques. *Geoderma*, *230–231*, 288–295. https://doi.org/10.1016/j.geoderma.2014.03.019

Corwin, D. L., & Lesch, S. M. (2017). Validation of the ANOCOVA model for regional-scale ECato ECecalibration. *Soil Use and Management*, *33*(2), 178–190. https://doi.org/10.1111/sum.12262

Corwin, D. L., Lesch, S. M., Shouse, P. J., Soppe, R., & Ayars, J. E. (2003). Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agronomy Journal*, *95*(2), 352–364. https://doi.org/10.2134/agronj2003.0352

Corwin, D. L., Loague, K., & Ellsworth, T. R. (1998). *GIS-Based Modeling of Non-Point Source Pollutants in the Vadose Zone*. *53*(1), 34–38.

Corwin, D. L., & Scudiero, E. (2019). Review of soil salinity assessment for agriculture across multiple scales using proximal and/or remote sensors. In *Advances in Agronomy* (Vol. 158, pp. 1–130). Cambridge, MA: Academic Press.

Corwin, D. L., & Scudiero, E. (2020). Field-scale apparent soil electrical conductivity. *Soil Science Society of America Journal*, *84*(5), 1405–1441. https://doi.org/10.1002/saj2.20153

Corwin, D. L., Yemoto, K., Kevin, W. C., Banuelos, G., Skaggs, T. H., Lesch, S. M., & Scudiero, E. (2017). Evaluating oilseed biofuel production feasibility in California's San Joaquin Valley using geophysical and remote sensing techniques. *Sensors (Switzerland)*, *17*(10), 1–25. https://doi.org/10.3390/s17102343

Costa-Filho, E., Chávez, J. L., Zhang, H., & Andales, A. A. (2021). An optimized surface aerodynamic temperature approach to estimate maize sensible heat flux and evapotranspiration. *Agricultural and Forest Meteorology*, *311*(May). https://doi.org/10.1016/j.agrformet.2021.108683

CSU. (2023). CoAgMET: Colorado's Mesonet. Retrieved from https://coagmet.colostate.edu/
Cui, X., Han, W., Zhang, H., Cui, J., Li, G., Conservation, W., & Collins, F. (2022). Estimating soil salinity under sunflower cover in the Hetao Irrigation District based on UAV remote sensing. *Land Degradation and Development*, 0–2. https://doi.org/10.1002/ldr.4445
Davis, E., Wang, C., & Dow, K. (2019). Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: a case study of agricultural lands in coastal North Carolina. *International Journal of Remote Sensing*, *40*(16), 6134–6153. https://doi.org/10.1080/01431161.2019.1587205
Delavar, M. A., Naderi, A., Ghorbani, Y., Mehrpouyan, A., & Bakhshi, A. (2020). Soil salinity mapping by remote sensing south of Urmia Lake, Iran. *Geoderma Regional*, *22*, e00317. https://doi.org/10.1016/j.geodrs.2020.e00317
Elhaddad, A., Garcia, L. A., & Chávez, J. L. (2011). Using a Surface Energy Balance Model to Calculate Spatially Distributed Actual Evapotranspiration. *Journal of Irrigation and Drainage Engineering*, *137*(1), 17–26. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000276
FAO. (2023). Land & Water Databases & Software. Retrieved June 22, 2023, from https://www.fao.org/land-water/databases-and-software/crop-information/maize/en/
Fernández-Buces, N., Siebe, C., Cram, S., & Palacio, J. L. (2006). Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. *Journal of Arid Environments*, *65*(4), 644–667. https://doi.org/10.1016/j.jaridenv.2005.08.005
Food and Agriculture Organization of the United Nation. (2021). *The State of the World's Land and Water Resources for Food and Agriculture - Systems at a Breaking Point* (M. Kay, S. Bunning, & J. Burke, Eds.). Rome, Italy.
Food and Agriculture Organization of the United Nations. (2020). The State of Food and Agriculture 2020. Overcoming water challenges in agriculture. In *The State of Food and Agriculture 2020*. https://doi.org/10.4060/cb1447en
Gates, T. K., Garcia, L. A., Hemphill, R. A., Morway, E. D., & Elhaddad, A. (2012). *Irrigation Practices, Water Consumption, & Return Flows in Colorado's Lower Arkansas River Valley*. (221), 116.
Gates, T. K., Steed, G. H., Niemann, J. D., & Labadie, J. W. (2016). Data for improved water management in Colorado's Arkansas River Basin: hydrological and water quality studies. *Colorado Water Institute Special Report No. 24*, (24), 122.
Gavlak, R., Horneck, R., & Miller, R. O. (2005). *Soil, Plant and Water Reference Methods for the Western Region* (3rd ed., Vol. 1).
Ghassemi, F., Jakeman, A., & Nix, H. (1995). *Salinisation of land and water resources: Human causes, extent, management and case studies*. Wallingford, U.K.: CAB International.
Gu, S., Liao, Q., Gao, S., Kang, S., Du, T., & Ding, R. (2021). Crop water stress index as a proxy of phenotyping maize performance under combined water and salt stress. *Remote Sensing*, *13*(22), 1–18. https://doi.org/10.3390/rs13224710
Hoffman, G. J., Evans, R. G., Jensen, M. E., Martin, D., & Elliott, R. (2007). *Design and Operation of Farm Irrigation Systems* (2nd ed.). St. Joseph, MI: American Society of Agricultural and Biological Engineers.
Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213. https://doi.org/10.1080/0965156x.2013.836857

Ivushkin, K., Bartholomeus, H., Bregt, A. K., Pulatov, A., Franceschini, M. H. D., Kramer, H., … Finkers, R. (2019). UAV based soil salinity assessment of cropland. *Geoderma*, *338*(September 2018), 502–512. https://doi.org/10.1016/j.geoderma.2018.09.046

Jackson, R. D. (1986). Remote Sensing of Biotic and Abiotic Plant Stress. *Annual Review of Phytopathology*, *24*(1), 265–287. https://doi.org/10.1146/annurev.py.24.090186.001405

Jackson, R. D., Idso, S. B., Reginato, R. J., & Pinter, P. J. (1981). *Canopy temperature as a crop water stress indicator. 17*(4), 1133–1138.

Jackson, R. D., Kustas, W. P., & Choudhury, B. J. (1988). A reexamination of the crop water stress index. *Irrigation Science*, *9*(4), 309–317. https://doi.org/10.1007/BF00296705

Jensen, M. E., & Allen, R. G. (2016). *Evaporation, Evapotranspiration, and Irrigation Water Requirements* (Second). Retrieved from https://ascelibrary.org/doi/book/10.1061/9780784414057

Jiang, H., Rusuli, Y., Amuti, T., & He, Q. (2019). Quantitative assessment of soil salinity using multi-source remote sensing data based on the support vector machine and artificial neural network. *International Journal of Remote Sensing*, *40*(1), 284–306. https://doi.org/10.1080/01431161.2018.1513180

Khan, S., & Abbas, A. (2013). Using Remote Sensing Techniques for Appraisal of Irrigated Soil Salinity. *Unpublished Project. CSIRO Land and Water, Locked Bag 588, Wagga, Wagga, NSW 2678, Austrailia*, 2632–2638.

Lakshmi, V., & Zehrfuhs, D. (2002). Normalization and comparison of surface temperatures across a range of scales. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(12), 2636–2647. https://doi.org/10.1109/TGRS.2002.805069

Lesch, S. M., Rhoades, J. D., Corwin, D. L., Robinson, D., & Suarez, D. L. (2002). *ESAP-SaltMapper version 2.30 R. User manual and tutorial guide* (149th ed.). Riverside, CA: USDA-ARS.

Li, J., Pu, L., Zhu, M., Dai, X., Xu, Y., Chen, X., … Zhang, R. (2015). Monitoring soil salt content using HJ-1A hyperspectral data: A case study of coastal areas in Rudong County, Eastern China. *Chinese Geographical Science*, *25*(2), 213–223. https://doi.org/10.1007/s11769-014-0693-2

Liu, M., Shi, H., Paredes, P., Ramos, T. B., Dai, L., Feng, Z., & Pereira, L. S. (2022). Estimating and partitioning maize evapotranspiration as affected by salinity using weighing lysimeters and the SIMDualKc model. *Agricultural Water Management*, *261*(December 2021), 107362. https://doi.org/10.1016/j.agwat.2021.107362

Liu, Y., Pan, X., Wang, C., Li, Y., & Shi, R. (2015). Predicting soil salinity with Vis-NIR spectra after removing the effects of soil moisture using external parameter orthogonalization. *PLoS ONE*, *10*(10), 1–13. https://doi.org/10.1371/journal.pone.0140688

Liu, Y., Pereira, L. S., & Fernando, R. M. (2006). Fluxes through the bottom boundary of the root zone in silty soils: Parametric approaches to estimate groundwater contribution and percolation. *Agricultural Water Management*, *84*(1–2), 27–40. https://doi.org/10.1016/j.agwat.2006.01.018

Lobell, D. B., Lesch, S. M., Corwin, D. L., Ulmer, M. G., Anderson, K. A., Potts, D. J., … Baltes, M. J. (2010). Regional-scale assessment of soil salinity in the Red River Valley using multi-year MODIS EVI and NDVI. *Journal of Environmental Quality*, *39*(1), 35–41. https://doi.org/10.2134/jeq2009.0140

Maas, E. V., Hoffman, G. J., Chaba, G. D., Poss, J. A., & Shannon, M. C. (1983). Salt sensitivity of corn at various growth stages. *Irrigation Science*, *4*(1), 45–57.

https://doi.org/10.1007/BF00285556
Maas, E. V. (1993). *Testing Crops for Salinity Tolerance*. (94), 234–247.
Marden, J. I. (2004). Positions and QQ plots. *Statistical Science*, *19*(4), 606–614. https://doi.org/10.1214/088342304000000512
Marino, G., Zaccaria, D., Lagos, L. O., Souto, C., Kent, E. R., Grattan, S. R., … Snyder, R. L. (2020). Effects of salinity and sodicity on the seasonal dynamics of actual evapotranspiration and surface energy balance components in mature micro-irrigated pistachio orchards. *Irrigation Science*, *39*(1), 23–43. https://doi.org/10.1007/s00271-020-00694-z
Marino, G., Zaccaria, D., Snyder, R. L., Lagos, O., Lampinen, B. D., Ferguson, L., … Sanden, B. L. (2019). Actual evapotranspiration and tree performance of mature micro-irrigated pistachio orchards grown on saline-sodic soils in the san joaquin valley of California. *Agriculture (Switzerland)*, *9*(4). https://doi.org/10.3390/agriculture9040076
Massoud, F. (1977). Basic principles for prognosis and monitoring of soil salinity and sodicity. *Proceedings of the International Conference on Managing Saline Waters for Irrigation*, 432–454. Lubbock, TX: Texas Tech University.
Matinfar, H. R., Alavi Panah, S. K., Zand, F., & Khodaei, K. (2013). Detection of soil salinity changes and mapping land cover types based upon remotely sensed data. *Arabian Journal of Geosciences*, *6*(3), 913–919. https://doi.org/10.1007/s12517-011-0384-6
Metternicht, G. I., & Zinck, J. A. (2003). Remote sensing of soil salinity: Potentials and constraints. *Remote Sensing of Environment*, *85*(1), 1–20. https://doi.org/10.1016/S0034-4257(02)00188-8
Metternicht, G., & Zinck, J. A. (1997). Spatial discrimination of salt- and sodium-affected soil surfaces. *International Journal of Remote Sensing*, *18*(12), 2571–2586. https://doi.org/10.1080/014311697217486
Morway, E. D., & Gates, T. K. (2012). Regional Assessment of Soil Water Salinity across an Intensively Irrigated River Valley. *Journal of Irrigation and Drainage Engineering*, *138*(5), 393–405. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000411
NASS, U. (2021). 2021 State Agriculture Overview: Colorado. Retrieved from https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=COLORADO
NRCS. (1998). *Soil Quality Resource Concerns: Salinity*. Retrieved from https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5293778.pdf
Osborn, B., Orlando, A. S., Hoag, D. L., Gates, T. K., & Valliant, J. C. (2017). *The economics of irrigation in Colorado's Lower Arkansas River Valley.* (32).
Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., & Shi, Z. (2019). Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma*, *337*(August 2018), 1309–1319. https://doi.org/10.1016/j.geoderma.2018.08.006
Qadir, M., Quillérou, E., Nangia, V., Murtaza, G., Singh, M., Thomas, R. J., … Noble, A. D. (2014). Economics of salt-induced land degradation and restoration. *Natural Resources Forum*, *38*(4), 282–295. https://doi.org/10.1111/1477-8947.12054
Qi, G., Zhao, G., & Xi, X. (2020). Soil salinity inversion of winter wheat areas based on satellite-unmanned aerial vehicle-ground collaborative system in coastal of the yellow river delta. *Sensors (Switzerland)*, *20*(22), 1–18. https://doi.org/10.3390/s20226521
Qin, A., Ning, D., Liu, Z., Li, S., Zhao, B., & Duan, A. (2021). Determining threshold values for a crop water stress index-based center pivot irrigation with optimum grain yield.

Agriculture (Switzerland), *11*(10). https://doi.org/10.3390/agriculture11100958
Quinlan, J. R. (1992). Learning with continuous classes. *Australian Joint Conference on Artificial Intelligence*, *92*, 343–348.
Rhoades, J. D. (1996). Salinity: Electrical Conductivity and Total Dissolved Solids. In *Methods of Soil Analysis, Part 3* (pp. 417–435). Madison, WI: Soil Science Society of America and American Society of Agronomy.
Rhoades, J. D., Shouse, P. J., Alves, W. J., Manteghi, N. A., & Lesch, S. M. (1990). Determining Soil Salinity from Soil Electrical Conductivity using Different Models and Estimates. *Soil Science Society of America Journal*, *54*(1), 46–54. https://doi.org/10.2136/sssaj1990.03615995005400010007x
Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2), 95–107. https://doi.org/10.1016/0034-4257(95)00186-7
Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring Vegetation Systems in the Great Plains with ERTS. *Journal of Agricultural and Food Chemistry*, 309–317. https://doi.org/10.1021/jf60203a024
Scudiero, E., Skaggs, T. H., & Corwin, D. L. (2014). Regional scale soil salinity evaluation using Landsat 7, Western San Joaquin Valley, California, USA. *Geoderma Regional*, *2–3*(C), 82–90. https://doi.org/10.1016/j.geodrs.2014.10.004
Scudiero, E., Skaggs, T. H., & Corwin, D. L. (2015). Regional-scale soil salinity assessment using Landsat ETM+ canopy reflectance. *Remote Sensing of Environment*, *169*, 335–343. https://doi.org/10.1016/j.rse.2015.08.026
Scudiero, E., Skaggs, T. H., & Corwin, D. L. (2016). Comparative regional-scale soil salinity assessment with near-ground apparent electrical conductivity and remote sensing canopy reflectance. *Ecological Indicators*, *70*, 276–284. https://doi.org/10.1016/j.ecolind.2016.06.015
Stiven, G. A., & Khan, M. A. (1966). Saturation Percentage as a Measure of Soil Texture in the Lower Indus Basin. *Journal of Soil Science*, *17*(2), 255–263.
Sumner, M. E. (1993). Sodic soils: New perspectives. *Australian Journal of Soil Research*, *31*(6), 683–750. https://doi.org/10.1071/SR9930683
Taghvaeian, S., Chávez, J. L., Bausch, W. C., DeJonge, K. C., & Trout, T. J. (2014). Minimizing instrumentation requirement for estimating crop water stress index and transpiration of maize. *Irrigation Science*, *32*(1), 53–65. https://doi.org/10.1007/s00271-013-0415-z
United States Salinity Laboratory Staff. (1954). Diagnosis and Improvement of Saline and Alkaline Soils. In L. A. Richards (Ed.), *United States Department of Agriculture* (Vol. 60). https://doi.org/10.2136/sssaj1954.03615995001800030032x
USDA ARS. (2023). Root Zone Water Quality Model. Retrieved August 23, 2023, from https://www.ars.usda.gov/plains-area/fort-collins-co/center-for-agricultural-resources-research/rangeland-resources-systems-research/docs/system/rzwqm/
USGS. (2023). EarthExplorer. Retrieved from https://earthexplorer.usgs.gov/
Walter, I., Allen, R. G., Elliott, R., Itenfisu, D., Brown, P., Jensen, M. E., … Wright, J. L. (2005). The ASCE Standardized Reference Evapotranspiration Equation. In *ASCE-EWRI*. https://doi.org/10.3406/arasi.1965.919
Wei, G., Li, Y., Zhang, Z., Chen, Y., Chen, J., Yao, Z., … Chen, H. (2020). Estimation of soil salt content by combining UAV-borne multispectral sensor and machine learning algorithms. *PeerJ*, *2020*(4), 1–24. https://doi.org/10.7717/peerj.9087

Zhang, J., Zhang, Z., Chen, J., Chen, H., Jin, J., Han, J., … Wei, G. (2021). Estimating soil salinity with different fractional vegetation cover using remote sensing. *Land Degradation and Development*, *32*(2), 597–612. https://doi.org/10.1002/ldr.3737

Zhang, T. T., Qi, J. G., Gao, Y., Ouyang, Z. T., Zeng, S. L., & Zhao, B. (2015). Detecting soil salinity with MODIS time series VI data. *Ecological Indicators*, *52*, 480–489. https://doi.org/10.1016/j.ecolind.2015.01.004

Zhu, K., Sun, Z., Zhao, F., Yang, T., Tian, Z., Lai, J., … Long, B. (2021). Relating hyperspectral vegetation indices with soil salinity at different depths for the diagnosis ofwinter wheat salt stress. *Remote Sensing*, *13*(2), 1–21. https://doi.org/10.3390/rs13020250

Appendix A: Full Pearson's r Correlation Tables

Table A1: Full Pearson's r table for the 2018 dataset that included interpolation between satellite dates. Magnitudes greater than 0.10 are statistically significant at $p < 0.05$. The color scale indicates the magnitude of the value. All correlations have a sample size of n = 351 data points.

2018, Interpolated		**ECa**	**ECe**	**ECeg**	**ln(ECa)**	**ln(ECe)**	**ln(ECeg)**
NDVI	Total	-0.06	-0.29	-0.37	0.02	-0.28	-0.38
	Initial	0.24	0.07	-0.03	0.29	0.05	-0.05
	Development	-0.26	-0.43	**-0.54**	-0.17	-0.38	**-0.52**
	Mid	-0.43	-0.40	-0.35	-0.42	-0.37	-0.33
	Late	0.08	-0.17	-0.22	0.15	-0.18	-0.26
CRSI	Total	0.27	0.00	-0.13	0.35	0.00	-0.15
	Initial	0.25	0.07	-0.01	0.29	0.05	-0.04
	Development	0.04	-0.19	**-0.33**	0.15	-0.16	**-0.34**
	Mid	-0.22	-0.27	-0.27	-0.18	-0.24	-0.26
	Late	0.36	0.08	-0.04	0.44	0.09	-0.07
ETa	Total	0.05	-0.13	-0.25	0.13	-0.10	-0.25
	Initial	-0.25	-0.16	-0.13	-0.25	-0.16	-0.12
	Development	-0.03	-0.23	**-0.37**	0.06	-0.20	**-0.36**
	Mid	0.11	-0.05	-0.17	0.18	-0.03	-0.17
	Late	0.25	0.04	-0.09	0.33	0.06	-0.10
CWSI	Total	0.01	0.18	**0.27**	-0.07	0.16	**0.28**
	Initial	0.25	0.17	0.16	0.24	0.14	0.13
	Development	0.01	0.22	0.36	-0.08	0.18	0.35
	Mid	-0.11	0.05	0.17	-0.18	0.03	0.17
	Late	-0.25	-0.04	0.09	-0.33	-0.06	0.10
ln(NDVI)	Total	-0.06	-0.29	-0.37	0.03	-0.27	-0.38
	Initial	0.31	0.12	0.01	0.36	0.12	0.00
	Development	-0.29	-0.46	**-0.57**	-0.19	-0.40	**-0.54**
	Mid	-0.42	-0.39	-0.34	-0.41	-0.36	-0.33
	Late	0.13	-0.12	-0.19	0.20	-0.13	-0.22

2018, Interpolated, Continued		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
ln(CRSI)	Total	0.27	0.00	-0.13	0.36	0.01	-0.15
	Initial	0.26	0.08	0.00	0.30	0.06	-0.03
	Development	0.04	-0.19	**-0.33**	0.15	-0.16	**-0.34**
	Mid	-0.22	-0.27	-0.26	-0.18	-0.24	-0.26
	Late	0.37	0.09	-0.03	0.45	0.10	-0.06
ln(ETa)	Total	0.05	-0.12	-0.25	0.14	-0.09	-0.24
	Initial	-0.27	-0.19	-0.17	-0.26	-0.17	-0.14
	Development	-0.04	-0.24	**-0.38**	0.06	-0.20	**-0.37**
	Mid	0.13	-0.03	-0.15	0.21	0.00	-0.15
	Late	0.28	0.09	-0.04	0.36	0.11	-0.05
ln(CWSI)	Total	0.03	0.18	0.26	-0.04	0.17	0.28
	Initial	0.22	0.11	0.07	0.23	0.09	0.05
	Development	0.06	0.21	**0.31**	-0.01	0.20	**0.34**
	Mid	-0.08	0.08	0.18	-0.15	0.06	0.19
	Late	-0.21	-0.01	0.11	-0.29	-0.03	0.12

Table A2: Full Pearson's r table for the 2018 dataset that included only data from satellite overpass dates. Magnitudes greater than 0.10 are statistically significant at $p < 0.05$. The color scale indicates the magnitude of the value. All correlations have a sample size of n = 351 data points.

2018, Not Interpolated		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
NDVI	Total	-0.08	-0.30	-0.38	0.00	-0.28	-0.39
	Initial	0.25	0.04	-0.10	0.32	0.03	-0.13
	Development	-0.27	-0.41	**-0.50**	-0.17	-0.36	**-0.48**
	Mid	-0.44	-0.33	-0.26	-0.45	-0.28	-0.21
	Late	0.14	-0.11	-0.18	0.21	-0.12	-0.22
CRSI	Total	0.28	0.01	-0.13	0.37	0.03	-0.15
	Initial	0.26	0.05	-0.07	0.31	0.03	-0.11
	Development	0.12	-0.11	-0.27	0.23	-0.08	-0.27
	Mid	**-0.35**	-0.31	-0.27	**-0.32**	-0.27	-0.24
	Late	0.38	0.10	-0.04	0.46	0.11	-0.07

2018, Not Interpolated, Continued		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
ETa	Total	0.08	-0.10	-0.24	0.16	-0.07	-0.25
	Initial	-0.06	-0.10	-0.13	-0.04	-0.10	-0.14
	Development	-0.09	-0.26	**-0.39**	0.00	-0.22	**-0.38**
	Mid	0.16	0.00	-0.13	0.23	0.03	-0.14
	Late	0.30	0.06	-0.08	0.39	0.07	-0.10
CWSI	Total	-0.08	0.10	0.24	-0.16	0.07	0.25
	Initial	0.06	0.08	0.14	0.02	0.07	0.13
	Development	0.08	0.24	**0.39**	-0.02	0.18	**0.35**
	Mid	-0.16	0.00	0.13	-0.23	-0.03	0.14
	Late	-0.30	-0.06	0.08	-0.39	-0.07	0.10
ln(NDVI)	Total	-0.08	-0.31	-0.39	0.00	-0.28	-0.40
	Initial	0.31	0.08	-0.07	0.38	0.08	-0.10
	Development	-0.31	-0.45	**-0.54**	-0.20	-0.38	**-0.51**
	Mid	-0.43	-0.31	-0.25	-0.44	-0.27	-0.20
	Late	0.19	-0.07	-0.15	0.26	-0.07	-0.18
ln(CRSI)	Total	0.28	0.02	-0.12	0.37	0.03	-0.15
	Initial	0.27	0.06	-0.07	0.32	0.04	-0.10
	Development	0.12	-0.11	-0.27	0.23	-0.07	-0.27
	Mid	**-0.35**	-0.31	-0.27	**-0.32**	-0.27	-0.24
	Late	0.39	0.11	-0.03	0.46	0.11	-0.06
ln(ETa)	Total	0.09	-0.09	-0.24	0.17	-0.06	-0.24
	Initial	-0.11	-0.15	-0.20	-0.07	-0.13	-0.18
	Development	-0.10	-0.27	**-0.40**	-0.01	-0.22	**-0.39**
	Mid	0.17	0.02	-0.12	0.24	0.05	-0.12
	Late	0.36	0.12	-0.03	0.44	0.14	-0.05
ln(CWSI)	Total	-0.04	0.13	0.24	-0.12	0.11	0.26
	Initial	-0.04	-0.01	0.01	-0.05	-0.01	0.02
	Development	0.01	0.17	**0.30**	-0.07	0.14	**0.30**
	Mid	-0.14	0.02	0.14	-0.21	-0.01	0.15
	Late	-0.22	0.01	0.13	-0.30	0.00	0.15

Table A3: Full Pearson's r table for the 2019 dataset that included interpolation between satellite dates. Magnitudes greater than 0.09 are statistically significant at $p < 0.05$. The color scale indicates the magnitude of the value. All correlations have a sample size of n = 411 data points.

2019, Interpolated		**ECa**	**ECe**	**ECeg**	**ln(ECa)**	**ln(ECe)**	**ln(ECeg)**
NDVI	Total	-0.42	**-0.68**	-0.58	-0.29	**-0.60**	-0.51
	Initial	0.15	0.22	0.13	0.17	0.26	0.17
	Development	-0.52	-0.31	-0.32	-0.47	-0.31	-0.31
	Mid	-0.56	-0.63	-0.56	-0.46	-0.56	-0.50
	Late	-0.08	-0.48	-0.37	0.03	-0.43	-0.32
CRSI	Total	-0.57	-0.57	-0.47	-0.47	-0.54	-0.44
	Initial	0.02	0.33	0.27	0.02	0.34	0.30
	Development	**-0.65**	-0.33	-0.32	**-0.62**	-0.34	-0.32
	Mid	-0.45	-0.60	-0.49	-0.35	-0.56	-0.46
	Late	-0.21	-0.52	-0.39	-0.12	-0.49	-0.37
ETa	Total	-0.27	-0.11	-0.13	-0.24	-0.09	-0.11
	Initial	-0.19	0.16	0.11	-0.18	0.19	0.14
	Development	**-0.40**	-0.17	-0.18	**-0.39**	-0.16	-0.18
	Mid	-0.28	-0.23	-0.23	-0.24	-0.22	-0.22
	Late	-0.05	-0.11	-0.11	-0.01	-0.09	-0.09
CWSI	Total	0.27	0.11	0.13	0.24	0.09	0.11
	Initial	0.23	-0.06	-0.02	0.24	-0.09	-0.05
	Development	**0.40**	0.17	0.18	**0.39**	0.16	0.18
	Mid	0.28	0.23	0.23	0.24	0.22	0.22
	Late	0.05	0.11	0.11	0.01	0.09	0.09
ln(NDVI)	Total	-0.44	**-0.69**	-0.60	-0.30	**-0.61**	-0.52
	Initial	0.16	0.24	0.14	0.18	0.27	0.18
	Development	-0.53	-0.32	-0.33	-0.48	-0.31	-0.32
	Mid	-0.57	-0.64	-0.57	-0.46	-0.56	-0.50
	Late	-0.08	-0.49	-0.38	0.02	-0.43	-0.33
ln(CRSI)	Total	-0.57	-0.57	-0.47	-0.47	-0.54	-0.44
	Initial	0.02	0.33	0.28	0.03	0.34	0.30
	Development	**-0.65**	-0.33	-0.32	**-0.62**	-0.34	-0.32
	Mid	-0.45	-0.59	-0.49	-0.35	-0.56	-0.46
	Late	-0.21	-0.52	-0.39	-0.12	-0.49	-0.37

2019, Interpolated, Continued		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
ln(ETa)	Total	-0.29	-0.14	-0.15	-0.25	-0.11	-0.12
	Initial	-0.23	0.08	0.04	-0.23	0.11	0.07
	Development	**-0.42**	-0.20	-0.21	**-0.40**	-0.18	-0.19
	Mid	-0.29	-0.25	-0.25	-0.24	-0.22	-0.22
	Late	-0.06	-0.12	-0.11	-0.01	-0.10	-0.09
ln(CWSI)	Total	0.24	0.08	0.10	0.21	0.05	0.08
	Initial						
	Development	**0.35**	0.11	0.14	**0.35**	0.11	0.14
	Mid	0.26	0.22	0.22	0.22	0.20	0.21
	Late						

Table A4: Full Pearson's r table for the 2019 dataset that included only data from satellite overpass dates. Magnitudes greater than 0.09 are statistically significant at $p < 0.05$. The color scale indicates the magnitude of the value. All correlations have a sample size of n = 411 data points.

2019, Not Interpolated		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
NDVI	Total	-0.44	-0.68	-0.56	-0.32	-0.61	-0.50
	Initial	0.11	0.18	0.11	0.13	0.22	0.15
	Development	**-0.72**	-0.55	-0.46	**-0.67**	-0.56	-0.47
	Mid	-0.55	-0.66	-0.60	-0.44	-0.58	-0.52
	Late	-0.03	-0.44	-0.33	0.07	-0.39	-0.28
CRSI	Total	-0.55	-0.56	-0.45	-0.47	-0.54	-0.43
	Initial	-0.02	0.28	0.25	-0.02	0.29	0.27
	Development	**-0.71**	-0.49	-0.41	**-0.67**	-0.50	-0.41
	Mid	-0.38	-0.57	-0.47	-0.28	-0.53	-0.43
	Late	-0.17	-0.51	-0.38	-0.08	-0.48	-0.36
ETa	Total	-0.29	-0.15	-0.16	-0.26	-0.13	-0.14
	Initial	-0.17	0.14	0.08	-0.16	0.17	0.12
	Development	**-0.59**	-0.34	-0.30	**-0.57**	-0.34	-0.31
	Mid	-0.16	-0.14	-0.16	-0.13	-0.12	-0.14
	Late	-0.08	-0.14	-0.12	-0.03	-0.12	-0.10
CWSI	Total	0.29	0.15	0.16	0.26	0.13	0.14
	Initial	0.17	-0.14	-0.08	0.16	-0.17	-0.12
	Development	**0.59**	0.34	0.30	**0.57**	0.34	0.31
	Mid	0.16	0.14	0.16	0.13	0.12	0.14
	Late	0.08	0.14	0.12	0.03	0.12	0.10

2019, Not Interpolated, Continued		ECa	ECe	ECeg	ln(ECa)	ln(ECe)	ln(ECeg)
ln(NDVI)	Total	-0.46	-0.69	-0.58	-0.33	-0.62	-0.51
	Initial	0.13	0.20	0.12	0.15	0.24	0.17
	Development	**-0.72**	-0.55	-0.47	**-0.67**	-0.55	-0.47
	Mid	-0.56	-0.66	-0.61	-0.44	-0.57	-0.52
	Late	-0.03	-0.44	-0.33	0.07	-0.39	-0.28
ln(CRSI)	Total	-0.55	-0.56	-0.45	-0.47	-0.54	-0.43
	Initial	-0.02	0.28	0.26	-0.01	0.30	0.28
	Development	**-0.71**	-0.49	-0.41	**-0.67**	-0.50	-0.41
	Mid	-0.38	-0.57	-0.47	-0.28	-0.53	-0.43
	Late	-0.17	-0.51	-0.38	-0.08	-0.48	-0.36
ln(ETa)	Total	-0.31	-0.18	-0.18	-0.27	-0.15	-0.15
	Initial	-0.23	0.03	-0.02	-0.20	0.08	0.03
	Development	**-0.59**	-0.37	-0.33	**-0.56**	-0.37	-0.33
	Mid	-0.17	-0.15	-0.17	-0.14	-0.12	-0.15
	Late	-0.08	-0.14	-0.13	-0.04	-0.12	-0.11
ln(CWSI)	Total	0.26	0.12	0.13	0.23	0.10	0.12
	Initial						
	Development	**0.55**	0.27	0.26	**0.55**	0.29	0.27
	Mid	0.15	0.13	0.15	0.12	0.11	0.14
	Late	0.06	0.13	0.12	0.02	0.12	0.10

Appendix B: Best Single Equations for each EC Type

The best single equations for each type of EC using 2018 data are as follows:

$$\ln(EC_{eg}) = -0.94 - 0.07(CWSI) + 1.19(CRSI) - 0.84[\ln(NDVI)] - 4.91(CWSI)[\ln(NDVI)] \quad (B1)$$

Eq. (B1) predicts $\ln(EC_{eg})$ using data from the crop development stage that did not include interpolation between satellite dates. CWSI, CRSI and ln(NDVI) plus their interactions were available for model selection of explanatory variables. Eq. (B1) resulted in an RMSE of 0.49 dS m^{-1} for the entire dataset (NRMSE of 18% of average EC_{eg} over all fields), and an RMPSE of 0.58 dS m^{-1} (NRMSE of 19% of average M2 EC_{eg}) for M2 LOFO prediction. The MBE was 0.04 dS m^{-1}.

$$\ln(EC_a) = 9.49 - 37.20(CWSI) - 3.27(CRSI) + 0.06(NDVI) + 13.05(CWSI)(CRSI) + 6.90(CWSI)(NDVI) + 0.20(CWSI)(NDVI) - 3.21(CWSI)(CRSI)(NDVI) \quad (B2)$$

Eq. (B2) predicts ln(EC_a) using data obtained from the crop late growth stage, and only including data acquired from satellite overpass dates (i.e., no interpolation between satellite overpass dates), with a sample size of n = 351. Variables available for model selection included CWSI, CRSI and NDVI plus their interactions. Eq. (B2) resulted in an EC_a RMSE of 0.22 dS m^{-1} (NRMSE of 27% of average EC_a) for the entire dataset, and an RMSPE of 0.25 dS m^{-1} (NRMSE of 25%) for M2 LOFO prediction. The MBE value was 0.03 dS m^{-1}.

The best single equations for each type of EC using 2019 data are as follows:

$$ln(EC_{eg}) = 178.10 + 259.87(ln(CWSI)) - 179.29(ln(CRSI)) - 183.85(ln(NDVI)) - 265.12(ln(CWSI))(ln(CRSI)) - 270.11(ln(CWSI))(ln(NDVI)) + 186.09(ln(CRSI))(ln(NDVI)) + 275.80(ln\,(CWSI)(ln(CRSI))(ln(NDVI)) \quad (B3)$$

Eq. (B3) predicts ln(EC_{eg}) using data obtained from the crop mid growth stage, and only including data acquired from satellite overpass dates (i.e., no interpolation between satellite overpass dates), with a sample size of n = 411. Variables available for model selection included CWSI, CRSI and NDVI plus their interactions. Although Eq. (B3) is long, the AIC method does take the number of terms into account when ranking models, and weighs models with many terms negatively. Eq. (B3) resulted in an EC_{eg} RMSE of 0.59 dS m^{-1} (NRMSE of 23% of average EC_{eg}) for the entire dataset, and an RMPSE of 0.71 dS m^{-1} (NRMSE of 21% of M2 average EC_{eg}) for M2 LOFO prediction. The MBE value was 0.06 dS m^{-1}.

$$EC_a = 1.70 + 0.43(CWSI) - 0.99(NDVI) \quad (B4)$$

Eq. (B4) predicts EC_a using data obtained from the crop development growth stage that only included data from satellite overpasses. Variables available for selection in this EC_a model included CWSI, CRSI and NDVI plus their interactions. Eq. (B4) resulted in an RMSE of 0.17 dS m^{-1} (NRMSE of 23% of average EC_a) for the entire dataset, and an RMPSE of 0.26 dS m^{-1} (NRMSE of 26% of M2 average EC_a) for M2 LOFO prediction. The MBE was 0.00 dS m^{-1}.

$$ln(EC_e) = -306.84 + 586.78(CWSI) + 316.71(\ln(CRSI)) + 327.60(\ln(NDVI)) - 600.32(CWSI)(\ln(CRSI)) - 616.07(CWSI)(\ln(NDVI)) - 337.12(\ln(CRSI))(\ln(NDVI)) + 630.64(CWSI)(\ln(CRSI))(\ln(NDVI)) \quad (B5)$$

Eq. (B5) predicts $\ln(EC_e)$ using data from the crop mid growth stage that only included data from satellite overpasses. Variables available for selection in this EC_e model included CWSI, CRSI and NDVI plus their interactions. Eq. (B5) resulted in an RMSE of 0.84 dS m^{-1} (NRMSE of 27% of average EC_e) for the entire dataset, and an RMSPE of 0.84 dS m^{-1} (NRMSE of 19% of M2 average EC_e) for M2 LOFO prediction. The MBE was 0.12 dS m^{-1}.

Appendix C: Full EC Modeling Results

Table C1: Full EC model results for 2018 multiple linear regression (MLR) modeling with CRSI and NDVI, plus their interaction available for model selection. The bold row was deemed the best statistically-performing equation. Grey boxes indicate that the model and results were not used in the final analyses. Green boxes indicate the best statistic in the column, and yellow boxes indicate the second-best statistic in the column. The sample size for each model was n = 351 data points.

Model Parameters				CRSI and NDVI Available for Model Selection							
Transformation	**Stage**	**Interpolation**	**Type of EC**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Avg LOFO RMSE**	**NRMSE**	**M2 NRMSE**	**Avg LOFO NRMSE**	**MBE**
None	Total	Yes	EC_{eg}	0.32	0.61	0.75		**0.22**	**0.25**		0.00
	Initial		EC_a	0.17	0.27	0.33		**0.33**	**0.33**		0.00
	Development		EC_{eg}	0.42	0.56	0.72		**0.20**	**0.24**		
	Total	No	EC_{eg}	0.30	0.61	0.76		**0.22**	**0.25**		0.00
	Initial		EC_a	0.17	0.27	0.33		**0.33**	**0.33**		0.00
	Development		EC_{eg}	0.47	0.53	0.62		**0.19**	**0.20**		0.00
Log-log	Total	Yes	EC_{eg}	0.32	0.6	0.76		**0.22**	**0.25**		0.05
	Initial		EC_a	0.22	0.27	0.35		**0.33**	**0.35**		
	Development		EC_{eg}	0.35	0.54	0.79		**0.20**	**0.26**		
	Total	No	EC_{eg}	0.30	0.60	0.77		**0.22**	**0.25**		0.05
	Initial		EC_a	0.25	0.27	0.35		**0.33**	**0.35**		0.04
	Development		EC_{eg}	0.38	0.50	0.61		**0.18**	**0.20**		
Combination	**Development**	**No**	**$ln(EC_{eg})$**	**0.38**	**0.50**	**0.61**	**0.48**	**0.18**	**0.20**	**0.17**	**0.05**
	Late		$ln(EC_a)$	0.37	0.25	0.30	0.23	**0.31**	**0.30**	**0.28**	0.03
	Mid		EC_{eg}	0.14	0.68	0.91	0.60	**0.25**	**0.30**	**0.22**	
	Development	Yes	$ln(EC_{eg})$	0.35	0.54	0.70	0.49	**0.20**	**0.23**	**0.18**	0.05

Table C2: Full EC modelresults for 2018 multiple linear regression (MLR) modeling with CRSI, NDVI and either CWSI or ET_a plus their interactions available for model selection. The bold row was deemed the best statistically-performing equation. Grey boxes indicate that the model and results were not used in the final analyses. Pink boxes indicate it was part of the piecewise regression of the best equation. Green boxes indicate the best statistic in the column, and yellow boxes indicate the second-best statistic in the column. The sample size for each model was n = 351 data points.

Model Parameters				CRSI, NDVI, and either CWSI or ET_a Available for Model Selection								
Transformation	**Stage**	**Interpolation**	**Type of EC**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Avg LOFO RMSE**	**CWSI or ET_a**	**NRMSE**	**M2 NRMSE**	**Avg LOFO NRMSE**	**MBE**
None	Total	Yes	EC_{eg}	0.43	0.56	0.69		CWSI	**0.20**	**0.23**		0.00
	Initial		EC_a	0.24	0.26	0.31		ET_a	**0.32**	**0.31**		0.00
	Development		EC_{eg}	0.53	0.51	0.64		ET_a	**0.18**	**0.21**		
	Total	No	EC_{eg}	0.44	0.54	0.66		CWSI	**0.20**	**0.22**		0.00
	Initial		EC_a	0.20	0.27	0.32		CWSI	**0.33**	**0.32**		0.00
	Development		EC_{eg}	0.57	0.48	0.58		CWSI	**0.17**	**0.19**		0.00
Log-log	Total	Yes	EC_{eg}	0.41	0.55	0.69		CWSI	**0.20**	**0.23**		0.05
	Initial		EC_a	0.29	0.26	0.32		ET_a	**0.32**	**0.32**		0.04
	Development		EC_{eg}	0.42	0.51	0.65		ET_a	**0.18**	**0.21**		
	Total	No	EC_{eg}	0.41	0.53	0.67		CWSI	**0.19**	**0.22**		0.05
	Initial		EC_a	0.28	0.27	0.33		ET_a	**0.33**	**0.33**		0.04
	Development		EC_{eg}	0.42	0.49	0.60		ET_a	**0.18**	**0.20**		
Log-log	Total	No	EC_{eg}	0.41	0.53	0.66		CWSI	**0.19**	**0.22**		0.05
	Development		EC_{eg}	0.43	0.49	0.61		CWSI	**0.18**	**0.20**		
Combination	**Development**	**No**	**$ln(EC_{eg})$**	**0.39**	**0.49**	**0.58**	**0.47**	**CWSI**	**0.18**	**0.19**	**0.17**	**0.04**
Same as above				0.42	0.49	0.60	0.46	ET_a	**0.18**	**0.20**	**0.17**	0.04
Same as above. Piecewise, ln(ECeg) below 4				0.11	0.44	0.50	0.43	CWSI	**0.16**	**0.17**	**0.16**	0.04
Same as above. Piecewise, ln(ECeg) above 4				0.80	0.52	0.56	0.40	CWSI	**0.19**	**0.19**	**0.14**	0.02
Same as above. Piecewise combined					0.44	0.51	0.43	CWSI	**0.16**	**0.17**	**0.16**	0.03
Combination	Late	No	$ln(EC_a)$	0.49	0.22	0.25	0.21	CWSI	**0.27**	**0.25**	**0.26**	0.03
	Mid	No	EC_{eg}	0.19	0.66	0.87	0.59	ET_a	**0.24**	**0.29**	**0.21**	
	Development	Yes	$ln(EC_{eg})$	0.42	0.51	0.65	0.47	CWSI	**0.18**	**0.21**	**0.17**	0.04

Table C3: Full EC model results for 2018 multiple linear regression (MLR) modeling with CRSI, NDVI and both CWSI and ET_a plus their interactions available for model selection. The bold row was deemed the best statistically-performing equation. Grey boxes indicate that the model and results were not used in the final analyses. Green boxes indicate the best statistic in the column, and yellow boxes indicate the second-best statistic in the column. The sample size for each model was n = 351 data points.

Model Parameters				CRSI, NDVI, CWSI and ET_a Available for Model Selection								
Transformation	**Stage**	**Interpolation**	**Type of EC**	**Variables Available for Selection**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Avg LOFO RMSE**	**NRMSE**	**M2 NRMSE**	**Avg LOFO NRMSE**	**MBE**
None	Total		EC_{eg}		0.45	0.55	0.68		**0.17**	**0.18**		0.00
	Initial	Yes	EC_a		0.35	0.24	0.27		**0.43**	**0.24**		0.00
	Development		EC_{eg}		0.53	0.50	0.64		**0.18**	**0.21**		
	Total		EC_{eg}		0.45	0.54	0.66		**0.16**	**0.17**		0.00
	Initial	No	EC_a		0.23	0.26	0.31		**0.32**	**0.31**		0.00
	Development		EC_{eg}		0.58	0.47	0.57		**0.17**	**0.19**		0.00
Log-log	Total	Yes	EC_{eg}		0.44	0.54	0.68		**0.20**	**0.22**		0.04
	Initial	Yes	EC_a		0.38	0.24	0.29		**0.29**	**0.29**		0.03
	Development	Yes	EC_{eg}		0.44	0.50	0.65		**0.18**	**0.21**		
	Total	No	EC_{eg}		0.43	0.52	0.66		**0.19**	**0.22**		0.04
	Initial	No	EC_a		0.30	0.26	0.34		**0.32**	**0.34**		0.03
	Development	No	EC_{eg}		0.42	0.49	0.6		**0.18**	**0.20**		
Log-log	Total	No	EC_{eg}		0.42	0.53	0.66		**0.19**	**0.22**		0.04
	Development	No	EC_{eg}		0.48	0.47	0.58		**0.17**	**0.19**		
Combination	**Development**	**No**	**$ln(EC_{eg})$**	**ln(NDVI), CRSI, CWSI, $ln(ET_a)$**	**0.48**	**0.45**	**0.55**	**0.43**	**0.16**	**0.18**	**0.16**	**0.04**
	Late	No	$ln(EC_a)$	NDVI, CRSI, CWSI, ET_a	0.51	0.21	0.25		**0.26**	**0.25**		
	Mid	No	EC_{eg}	NDVI, CRSI, ln(CWSI), ET_a	0.19	0.66	0.87	0.59	**0.24**	**0.29**	**0.21**	
	Development	Yes	$ln(EC_{eg})$	ln(NDVI), CRSI, CWSI, $ln(ET_a)$	0.45	0.5	0.63	0.46	**0.18**	**0.21**	**0.17**	0.04

Table C4: Full EC model results for 2019 multiple linear regression (MLR) modeling with CRSI and NDVI plus their interactions available for model selection. The bold rows were deemed the two best statistically-performing equations. Grey boxes indicate that the model and results were not used in the final analyses. Green boxes indicate the best statistic in the column, and yellow boxes indicate the second-best statistic in the column. The sample size for each model was n = 411 data points.

Model Parameters				CRSI and NDVI Available for Model Selection							
Transformation	**Stage**	**Interpolated**	**Type of EC**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Average LOFO RMSE**	**NRMSE**	**M2 LOFO NRMSE**	**Average LOFO NRMSE**	**MBE**
None	Mid	Yes	EC_e	0.42	0.90	1.14		0.29	0.26		0.00
None	Mid	Yes	EC_{eg}	0.32	0.66	0.99		0.25	0.30		0.00
None	Development	Yes	EC_a	0.41	0.19	0.26		**0.26**	**0.26**		0.00
None	Mid	No	EC_e	0.43	0.88	1.02		0.28	0.23		0.00
None	Mid	No	EC_{eg}	0.36	0.64	0.91		0.24	0.27		0.00
None	Development	No	EC_a	0.52	0.17	0.27		**0.23**	**0.27**		0.00
Log-log	Mid	Yes	$\ln(EC_e)$	0.36	0.91	1.17		**0.29**	**0.27**		0.12
Log-log	Mid	Yes	$\ln(EC_{eg})$	0.26	0.66	0.95		**0.25**	**0.28**		0.07
Log-log	Development	Yes	$\ln(EC_a)$	0.37	0.19	0.28		0.26	0.28		0.02
Log-log	Mid	No	$\ln(EC_e)$	0.35	0.88	1.01		**0.28**	**0.23**		0.12
Log-log	**Mid**	**No**	**$\ln(EC_{eg})$**	**0.28**	**0.63**	**0.89**		**0.24**	**0.27**		**0.07**
Log-log	Development	No	$\ln(EC_a)$	0.47	0.17	0.27		**0.23**	**0.27**		0.02
Combination	Development	Yes	EC_a	0.42	0.19	0.26	0.19	**0.26**	**0.26**	**0.26**	0.00
Combination	Development	No	EC_a	0.53	0.17	0.27	0.16	0.23	0.27	0.22	0.00
Combination	**Development**	**No**	**EC_a**	**0.52**	**0.17**	**0.27**	**0.16**	**0.23**	**0.27**	**0.22**	**0.00**
Combination	Development	No	EC_a								
Combination	Total	Yes	$\ln(EC_e)$	0.11	1.13	3.49	2.26	**0.36**	**0.80**	**0.73**	0.18
Combination	Late	No	$\ln(EC_{eg})$	0.13	0.74	2.58	1.86	**0.28**	**0.77**	**0.71**	0.09
Combination	Initial	No	$\ln(EC_e)$	0.19	1.08	3.46	2.30	**0.35**	**0.79**	**0.74**	0.20
Combination	Mid	No	$\ln(EC_e)$	0.35	0.88	1.01	0.84	**0.28**	**0.23**	**0.27**	0.12
Combination	Mid	No	$\ln(EC_{eg})$	0.28	0.63	0.89	0.60	0.24	**0.27**	**0.23**	0.07

Table C5: Full EC model results for 2019 multiple linear regression modeling with CRSI, NDVI and either CWSI or ET_a plus their interactions available for model selection. The bold rows were deemed the two best equations. Grey boxes indicate that the model was not used in the final analyses. Pink boxes indicate it was part of the piecewise regression of the best equations. Green boxes indicate the best statistic in the column, and yellow boxes indicate the second-best statistic. The sample size for each model was n = 411.

Model Parameters				CRSI, NDVI, and Either CWSI or ET_a Available for Model Selection								
Transformation	**Stage**	**Interpolated**	**Type of EC**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Average LOFO RMSE**	**CWSI or ETa**	**NRMSE**	**M2 LOFO NRMSE**	**Average LOFO NRMSE**	**MBE**
None	Mid	Yes	EC_e	0.48	0.85	0.94		CWSI	0.27	0.21		0.00
None	Mid	Yes	EC_{eg}	0.41	0.62	0.81		CWSI	0.20	0.19		0.00
None	Development	Yes	EC_a	0.48	0.18	0.22		CWSI	**0.24**	**0.22**		0.00
None	Mid	No	EC_e	0.49	0.83	0.83		CWSI	0.27	0.19		0.00
None	Mid	No	EC_{eg}	0.44	0.59	0.73		CWSI	0.23	0.22		0.00
None	Development	No	EC_a	0.54	0.17	0.26		CWSI	**0.23**	**0.26**		0.00
Log-log	Mid	Yes	$ln(EC_e)$	0.38	0.86	1.80	1.10	CWSI	**0.28**	**0.41**	**0.35**	0.12
Log-log	Mid	Yes	$ln(EC_{eg})$	0.30	0.62	0.83	0.58	ET_a	**0.24**	**0.25**	**0.22**	0.07
Log-log	Development	Yes	$ln(EC_a)$	0.42	0.18	0.22		ET_a	0.24	0.22		0.02
Log-log	Mid	No	$ln(EC_e)$	0.37	0.84	0.83	0.78	CWSI	**0.27**	**0.19**	**0.25**	0.12
Log-log	**Mid**	**No**	**$ln(EC_{eg})$**	**0.31**	**0.59**	**0.71**	**0.56**	**CWSI**	**0.23**	0.21	**0.21**	**0.06**
Same as above. Piecewise, ECeg below 4				**0.11**	**0.49**	**0.5**	**0.47**	**CWSI**	**0.19**	**0.15**	**0.18**	**0.05**
Same as above. Piecewise, ECeg above 4				**0.38**	**0.35**	**0.97**	**0.73**	**CWSI**	**0.13**	**0.29**	**0.28**	**0.06**
Same as above. Piecewise combined					**0.52**	**0.61**	**0.5**	**CWSI**	**0.20**	**0.18**	**0.19**	**0.05**
Log-log	Development	No	$ln(EC_a)$	0.5	0.17	0.26		ET_a	**0.23**	**0.26**		0.02
Combination	Development	Yes	EC_a	0.49	0.18	0.22	0.18	$ln(ET_a)$	**0.24**	**0.22**	**0.24**	0.00
Combination	Development	No	EC_a	0.55	0.17	0.26	0.15	$ln(ET_a)$	0.23	0.26	0.20	0.00
Combination	**Development**	**No**	**EC_a**	**0.54**	**0.17**	**0.26**	**0.15**	**CWSI**	**0.23**	**0.26**	**0.20**	**0.00**
Combination	Development	No	EC_a	0.54	0.19	0.26	0.15	$ln(ET_a)$	0.26	0.26	0.20	0.00
Same as above. Piecewise, ECa below 1.15				0.50	0.14	0.17	0.13	CWSI	**0.19**	**0.17**	**0.18**	0.00
Same as above. Piecewise, ECa above 1.15				0.31	0.21	0.26	0.14	CWSI	**0.29**	**0.26**	**0.19**	0.00
Same as above. Piecewise combined					0.15	0.19	0.13	CWSI	**0.20**	**0.19**	**0.18**	0.00
Combination	Total	Yes	$ln(EC_e)$	0.23	1.01	3.43	2.25	CWSI	**0.33**	**0.78**	**0.72**	0.15
Combination	Late	No	$ln(EC_{eg})$	0.13	0.74	2.58	1.86	ET_a	**0.28**	**0.77**	**0.71**	0.09
Combination	Initial	No	$ln(EC_e)$	0.24	1.04	3.43	2.29	CWSI	**0.34**	**0.78**	**0.74**	0.15
Combination	Mid	No	$ln(EC_e)$	0.37	0.84	0.84	0.78	CWSI	**0.27**	**0.19**	**0.25**	0.12
Combination	Mid	No	$ln(EC_{eg})$	0.31	0.59	0.72	0.56	CWSI	**0.23**	**0.22**	**0.21**	0.06

Table C6: Full EC model results for 2019 MLR modeling with CRSI, NDVI, CWSI and ET_a plus their interactions available for model selection. The bold rows were deemed the two best equations. Grey boxes indicate that the model was not used in the final analyses. Green boxes indicate the best statistic in the column, and yellow indicates the second-best. Sample size for each model was n = 411.

Model Parameters				CRSI, NDVI, CWSI and ET_a Available for Model Selection								
Transformation	**Stage**	**Interpolated**	**Type of EC**	**Variables Available for Selection**	**R^2**	**RMSE**	**M2 LOFO RMSE**	**Average LOFO RMSE**	**NRMSE**	**M2 LOFO NRMSE**	**Average LOFO NRMSE**	**MBE**
None	Mid	Yes	EC_e		0.48	0.85	0.94		0.27	0.21		0.00
None	Mid	Yes	EC_{eg}		0.41	0.62	0.81		0.20	0.19		0.00
None	Development	Yes	EC_a		0.52	0.17	0.21		**0.23**	**0.21**		0.00
None	Mid	No	EC_e		0.49	0.83	0.83		0.27	0.19		0.00
None	Mid	No	EC_{eg}		0.44	0.59	0.73		0.23	0.22		0.00
None	Development	No	EC_a		0.54	0.17	0.26		**0.23**	**0.26**		0.00
Log-log	Mid	Yes	$ln(EC_e)$		0.38	0.86	0.98		**0.28**	**0.22**		0.12
Log-log	Mid	Yes	$ln(EC_{eg})$		0.30	0.62	0.83	0.58	**0.24**	**0.25**	**0.22**	0.07
Log-log	Development	Yes	$ln(EC_a)$		0.42	0.18	0.22		0.24	0.22		0.02
Log-log	Mid	No	$ln(EC_e)$		0.37	0.84	0.83		**0.27**	**0.19**		0.12
Log-log	**Mid**	**No**	**$ln(EC_{eg})$**		**0.31**	**0.59**	**0.71**	**0.56**	**0.23**	**0.21**	**0.21**	**0.06**
Log-log	Development	No	$ln(EC_a)$		0.52	0.17	0.26		0.23	0.26		0.02
Combination	Development	Yes	EC_a	ln(NDVI), ln(CRSI), $ln(ET_a)$, ln(CWSI)	0.51	0.17	0.21	0.17	**0.23**	**0.21**	**0.23**	0.00
Combination	Development	No	EC_a	ln(NDVI), ln(CRSI, $ln(ET_a)$, CWSI	0.55	0.17	0.26	0.15	0.23	0.26	0.20	0.00
Combination	**Development**	**No**	**EC_a**	**NDVI, CRSI, CWSI, $ln(ET_a)$**	**0.54**	**0.17**	**0.26**	**0.15**	**0.23**	**0.26**	**0.20**	**0.00**
Combination	Total	Yes	$ln(EC_e)$	ln(NDVI), ln(CRSI), $ln(ET_a)$, CWSI	0.24	1.01	3.43	2.25	**0.33**	**0.78**	**0.72**	0.15
Combination	Late	No	$ln(EC_{eg})$	ln(NDVI), ln(CRSI), ET_a, ln(CWSI)	0.13	0.74	2.58	1.86	**0.28**	**0.77**	**0.71**	0.09
Combination	Initial	No	$ln(EC_e)$	NDVI, CRSI, ET_a, CWSI	0.29	1.01	3.39	2.29	**0.33**	**0.77**	**0.74**	0.14
Combination	Mid	No	$ln(EC_e)$	ln(NDVI), ln(CRSI), $ln(ET_a)$, CWSI	0.37	0.84	0.84	0.78	**0.27**	**0.19**	**0.25**	0.12
Combination	Mid	No	$ln(EC_{eg})$	ln(NDVI), CWSI, $ln(ET_a)$, CWSI)	0.31	0.58	0.72	0.56	**0.22**	**0.22**	**0.21**	0.06

LIST OF ABBREVIATIONS

ANOCOVA: Analysis of Covariance
AOI: Area of Interest
ASCE: American Society of Civil Engineers
C: Degrees Celsius
CIMIS: California Irrigation Management Information System
CoAgMet: Colorado Agricultural Meteorological Network
CRSI: Canopy Response Salinity Index
CSU: Colorado State University
CWSI: Crop Water Stress Index
$CWSI_{IRT}$: Crop Water Stress Index derived from infrared thermometers
$CWSI_{sat}$: Crop Water Stress Index derived from satellite imagery
DGPS: Differential Global Positioning System
DOY: Day of Year
DP: Deep Percolation
dS/m: Decisiemens per meter
EC: Electrical Conductivity
EC_a: Apparent Electrical Conductivity
EC_e: Electrical Conductivity of the Soil Paste Extract
EC_{eg}: Gypsum-Corrected Electrical Conductivity of the Soil Paste Extract
EF: Evaporative Fraction
EMI: Electromagnetic Induction
ER: Electromagnetic Resistance
ESP: Exchangeable Sodium Percentage
ET: Evapotranspiration
ET_a: Actual Evapotranspiration
ET_c: Crop Evapotranspiration
ET_r: Reference Evapotranspiration
EVI: Enhanced Vegetation Index
EWRI: Environmental and Water Resources Institute
FAO: Food and Agriculture Organization of the United Nations
G: Soil Heat Flux
GIS: Geographic Information System
GNSS: Global Navigation Satellite System
GPS: Global Positioning System
H: Sensible Heat Flux
Ha: Hectares
IRS: Indian Remote Sensing
IRT: Infrared Thermometer
LAI: Leaf Area Index
LE: Latent Heat Flux
LF: Leaching Fraction
LiDAR: Light Detection and Ranging
LOFO: Leave One Field Out

M ha: Mega Hectares or Million Hectares
M2: Muth 2
M3: Muth 3
M6: Muth 6
M8: Muth 8
MAD: Management-Allowed Depletion
MBE: Mean Bias Error
MDT: Mountain Daylight Time
METRIC: Measuring Evapotranspiration with Internalized Calibration
MODIS: Moderate Resolution Imaging Spectroradiometer
MSI: Multispectral Instrument
NDVI: Normalized Difference Vegetation Index
NRCS: National Resources Conservation Service
NRMSE: Normalized Root Mean Square Error
NRMSPE: Normalized Root Mean Square Prediction Error
OLI: Operational Land Imager
OLS: Ordinary Least Squares
PM: Penman-Monteith
q-q: Quantile-Quantile
R_n: Net Radiation
RF: Random Forest
RMSE: Root Mean Square Error
RMSPE: Root Mean Square Prediction Error
RS: Remote Sensing or Remotely Sensed
SAR: Sodium Adsorption Ratio
SEB: Surface Energy Balance
SMD: Soil Moisture Depletion
SSMU: Site-Specific Management Unit
SSURGO: Soil Survey Geographic Database
SVR: Support Vector Machine Regression
TAW: Total Applied Water
TDR: Time-Domain Reflectometry
U: Upflux
UAS: Unmanned Aerial System
UAV: Unmanned Aerial Vehicle
VF: Vegetative Fraction
VI: Vegetation Index
VPD: Vapor Pressure Deficit
VPG: Vapor Pressure Gradient