2020 Final Report

Quantifying Diffuse Recharge at Camp Bullis, TX: Integrating Soil Water, Evapotranspiration, and Remote Sensing

Submitted to:

Edwards Aquifer Authority

Submitted by:

Alexander Y. Sun, Tara Bongiovanni, Todd G. Caldwell, Michael H. Young

Bureau of Economic Geology, Jackson School of Geosciences, The University of Texas at Austin

October 2020

Executive Summary

In the karst terrain of the Edwards Aquifer System in central Texas, groundwater recharge occurs through a combination of discrete (R_{ds}) and diffuse (R_{df}) mechanisms. The majority of recharge occurs as R_{ds} from perennial base flow in major rivers and episodically from losing stream reaches in the numerous creeks, while a smaller proportion occurring through infiltration and percolation through the vadose zone is R_{df}. Whereas R_{ds} can be estimated from discharge measurements above and below the recharge zone (Puente, 1978), R_{df} is highly dynamic both spatially and temporally and is difficult to quantify due to the complex nature of soils and varied degrees of karstic porosity in the unsaturated zone (e.g., the vadose zone) above the aquifers. Thus, R_{df} presents one of the largest uncertainties in the water budget used to model these aquifers. The ability to quantify R_{df} could significantly improve long-term groundwater management and land conservation practices. In this project, we conducted multiscale field observations, remote sensing, and numerical modeling studies to evaluate a water-balance-based estimate of R_{df}. The water balance components were independently studied using numerous hydrometeorological datasets, including precipitation (PPT), potential evapotranspiration (PET), actual ET, and soil water storage.

ET was measured directly using the eddy-covariance method at three sites, each operated on the Camp Bullis Training Site of Joint Base San Antonio (Camp Bullis) from 2016 through 2018. These sites were chosen to cover the range of vegetation and leaf-area indices (LAI) common across the major drainages crossing the Edwards/Trinity recharge zone, including (1) closed cedar woodland, (2) open oak-savanna, and (3) open grassland. Satellite-derived ET was developed for Camp Bullis using vegetation indices, which are good indicators of vegetation greenness, canopy structure, and temporal dynamics of water use. Of the four indices tested, the Enhanced Vegetation Index (EVI) was found most stable for our sites. Using an empirical model between measured ET and leaf area index, monthly maps of ET-estimated from EVI were derived from 2004 to 2018 for the entire Cibolo Basin (Basin #7). Finally, we compared these values to a spatially explicit vadose zone model (HYDRUS-1D) at a 500-m scale. The model was validated at all monitoring sites and the monthly EVI-ET estimates taken from Camp Bullis. Recharge calculated from HYDRUS-1D was assumed to be R_{df}.

Results (Figure ES-1) of the analyses highlight the nuanced differences between the traditional estimate of recharge using the USGS method (*Puente, 1978*) and the HYDRUS-1D model. Note the former includes both discrete and diffuse recharge and is estimated over the Edwards only, while the latter assumes only diffuse recharge and doesn't distinguish between geologic outcrop of the Edwards or Trinity Group rocks (aquifer outcrop). The results show significant year-over-year variability in recharge. Following the 2011 drought, recharge remained very low for four additional years, showing the impact of that drought event on long-term recharge. The analyses highlight the importance of interannual variability on hydrologic response, and how separating diffuse recharge from total recharge can provide a more holistic understanding of the hydrology of the region of the Edwards and Trinity aquifers.

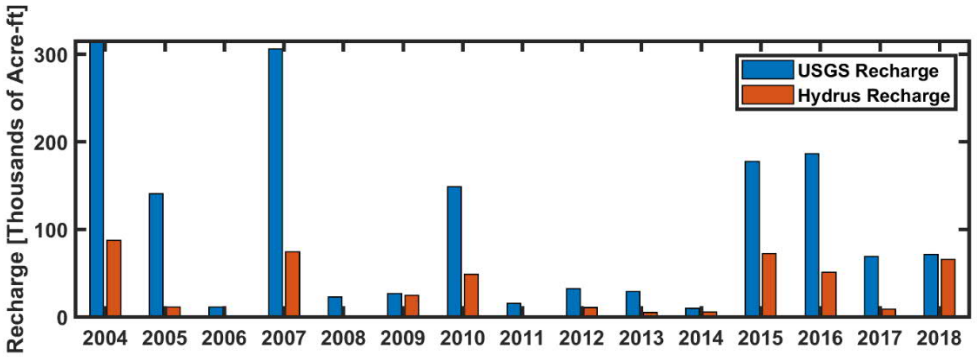


Figure ES-1: Annual total recharge in Cibolo Basin (discrete and diffuse) estimated from the USGS, and diffuse recharge estimated from HYDRUS.

The following list summarizes key findings of this study:

1. The ratio of annual diffuse recharge to precipitation (R_{df}/PPT), estimated from the HYDRUS model, ranged from 0 to 8%. This equates to a volume of annual diffuse recharge that ranges from essentially 0 acre-ft in 2008 to 87,000 acre-ft in 2004, which is equivalent to 8% of the total precipitation for that year. When compared to the established Puente method, the R_{df} ranged from 0 to 90% of total recharge (diffuse and discrete), with an average of 32%. However, when compared to other water balance methods, including the water balance method developed in this study, R_{df} accounts for between 0 and 30% of the total recharge. Our calculated percentage is similar to the 26% (R_{df}/Total Recharge) reported by Ockerman (Ockerman, 2007).
2. Of the four precipitation products evaluated, those that use rain gage data to refine and correct the precipitation estimates (e.g., RainVieux) perform better than those that do not. North American Land Data Assimilation System (NLDAS) was particularly poor with high biases, whereas gridMET and PRISM were both consistent. Results indicate that using rain gages as control points for precipitation estimates is critical when rainfall is locally intense and highly variable. Moreover, because diffuse recharge originates from precipitation, errors in precipitation will compound throughout the water balance.
3. Calculating PET from meteorological variables requires air temperature and humidity, wind speed, and solar radiation. Of these, wind speed and maximum relative humidity have the most uncertainty and lowest correlation to the gridded products available, whereas air temperature and minimum relative humidity are most highly correlated. In particular, gridMET-derived PET is more comparable to the meteorological stations than NLDAS, which has a monthly bias nearly 3 times higher than other methods.
4. The uncertainty and temporal dynamics in soil water storage are difficult to constrain in the water budget, but are potentially extremely important. For example, results showed that variations in monthly soil water storage can be more than 40% of ET. Of the NLDAS models tested here (Noah, Mosaic, and VIC), the change in soil water storage can range from ±16 to ±50 mm/month when assuming the soil depth is 40 cm.
5. The regression equation used to extend the eddy covariance measurements of ET to regional estimates is more accurate between EVIs of 0.23 and 0.44, which represents at least 64% of the region during any given month from 2000 to 2018, and more than 80% of the region ~90% of the time. Yearly ET/PPT can vary greatly between years and can be anywhere between 0.45 for a wet year to 1.35 for a dry year.

Recommendations and future research needed:

- Maintain/expand onsite weather and soil monitoring across EAA Region—The field measurements conducted for EAA focused on processes only at Camp Bullis. To better understand how recharge changes across the entire EAA region, especially in the drier, western regions, and to better parameterize other basins (e.g., Nueces-West Nueces River Basin), we recommend redeploying existing EC stations in different geographic areas of EAA's Region and maintaining the stations for at least a full year, preferably longer, to observe interannual variability.
- Automate blending of in-situ data and satellite remote sensing (or land surface models) - Research findings show that diffuse recharge can be a significant part of the groundwater recharge component of the water balance, but that variability in time and space complicates projections. To better understand how the intensive data collection at Camp Bullis can be extrapolated to other areas, we recommend developing and implementing an approach that will produce a soil moisture map of high spatial and temporal resolution over the EAA region.
- Apply multivariate regression or machine learning algorithms to predict weather/hydrology variables of interest, including diffuse recharge – Results from this project showed that soil water storage is critically important to diffuse recharge, but significant uncertainties are present in the current soil water storage estimates, due to uncertainties on soil hydraulic properties, soil horizon/types, and soil water content observations. Future effort should consider developing a data analytics workflow that optimally merges high-resolution satellite remote sensing data, outputs of land surface models, and in-situ data, to create high-resolution soil moisture maps across large EAA areas with lower uncertainty.
- Assess land use change on diffuse recharge – Results of this study indicate that approximately 10-12% of total precipitation is recharged through diffuse and discrete routes, and that approximately 30% of total recharge is through diffuse routes (or, ~4% precipitation). We recommend overlaying population projections, estimates of roads, physical structures, etc., and other diversion features onto the diffuse recharge tomograms obtained from numerical models (e.g., HYDRUS), and then determine whether associated land-use change could alter the diffuse recharge because of reduced pervious surfaces, or discrete recharge because of surface water routing toward or away from river reaches that are known to recharge aquifers.
- Investigate the role of epikarst, which comprises highly weathered carbonate bedrock immediately beneath the land surface or beneath the soil, in diffuse recharge. Hydraulic properties of epikarst are highly heterogeneous. The varied degree of karst development, existence of low permeability evaporite layers in the Glen Rose, and faulting likely play an important role in determining where recharge occurs more efficiently or less efficiently. Future studies should add epikarst representation in models to further constrain recharge estimates.

Table of Contents

List of Tables

List of Figures

1. Introduction

Climate, topography, geology, vegetation, and soil exert control on groundwater recharge. Mean annual precipitation explains 80% of the variation in recharge across Texas (Keese et al., 2005), but the diversity of these five factors makes local recharge estimates difficult to quantify, particularly in karst terrains where the landscape is underlain by soluble rock that contains dissolution features such as fractures, caves, and sinkholes. Karst aquifers, such as the Edwards Aquifer System in Texas, are some of the most productive groundwater resources in the world. The Edwards Aquifer System, consisting of a set of karst aquifers that are variably connected, including the Edwards-Trinity (Plateau) Aquifer, the Edwards (Balcones Fault Zone) Aquifer, and the Edwards (Washita Prairies) Aquifer (Hackett, 2019), provides most water resources across central Texas.

In the karst terrain of the Edwards/Trinity aquifers, recharge occurs through a combination of discrete (R_{ds}) and diffuse (R_{df}) mechanisms (Figure 1-1). All recharge originates as precipitation that either infiltrates in the contributing area above the recharge zone, runs off the contributing area, or falls directly on the recharge zone, becoming R_{ds} (Figure 1-2). Since the 1980s, annual groundwater recharge to the Edwards/Trinity aquifers has been estimated using a water-balance equation in which recharge within each watershed is the difference between measured streamflow above and below the recharge area (Puente, 1978). Focused recharge or R_{ds} occurs through faults, fractures and dissolution features in ephemeral stream channels as storm water flows across the recharge zone (Figure 1-2). R_{df} initiates as precipitation infiltrates into the soil in the inter-stream areas where it is stored, and then slowly released downward by gravity, or upward by evapotranspiration. In the latter case, water is used by vegetation as transpiration (T) or lost through the soil surface as evaporation (E). Combining the two becomes evapotranspiration (ET). As pore water pressures increase in the vadose zone, gravity pulls the water downward, deeper within and below the root zone and into the epikarst where it can become R_{df}.

Most groundwater recharge to the Edwards/Trinity aquifers originates episodically as R_{ds} from losing stream reaches while a smaller, less temporally dynamic proportion is R_{df} (Maclay, 1995; Marclay, 1995; Wong et al., 2012). Whereas R_{ds} can be estimated from discharge measurements above and below the recharge zone, R_{df} is difficult to quantify due to the complex nature of soils, varied degrees of karstic porosity in the unsaturated zone above the water table (e.g., the vadose zone) of the aquifers, and the different domains in the soil/upper epikarst that store or rapidly transmit water depending on degree of saturation and precipitation rate. The diffuse component is assumed minor in comparison to R_{ds} because potential ET (PET) is quite high in central Texas. Maclay (1995) suggested 85–90% of all precipitation is lost to ET and most recharge estimates assume R_{df} on the annual scale is a constant percentage of annual precipitation. However, such approaches neglect changes in soil water, epikarst, and groundwater storage (Wilcox, 2008), suggesting a water-balance approach may be more appropriate. For example, Dugas et al. (1998) quantified ET using the Bowen ratio method and estimated that stream flow was 20% of precipitation; by differencing alone, ET would be 80% of precipitation. However, their measured values of ET were significantly lower than others (approximately 65% of precipitation), suggesting R_{df} can be significant.

To date, research results are conflicted on the source of water for ET, with some findings suggesting water sourced from epikarst, (Wilcox and Huang, 2010; Wilcox et al., 2005; Wong and Banner, 2010), even though plant available water may be disconnected from deeper sources (Kukowski et al., 2013; Schwartz et al., 2013). These previous studies have sought to understand the extent to which soil moisture

contributes water for root uptake, given the common perception that soils in the Texas Hill Country are shallow with little water holding capacity (Wilcox et al., 2007; Woodruff and Wilding, 2008). It is known that the water storage capacity of soil and epikarst influence long-term ET, more so than differences in woody cover (Heilman et al., 2014; Schwartz et al., 2013), with little evidence of deeper soil water usage (i.e., tapping groundwater) or evidence of deep rooting (Heilman et al., 2009; Schwinning, 2008). These studies suggest a decoupling of epikarst water storage from plant-available water, even though most water loss is still through ET. As the studies noted above suggest, diffuse recharge may play a more significant role in karst aquifer recharge in areas like the Edwards-Trinity Plateau. Even within the general category of diffuse recharge, the soil/epikarst has at least two distinct domains where the flow differs: matrix dominated and fissure/macropore dominated. Several studies have suggested a piston-like process where water stored in the soil matrix is pushed downward by subsequent precipitation, during relatively dry conditions, or where "new" water from precipitation could mostly bypass the matrix and flow through fissures/macropores during wet conditions (Perrin et al., 2003; Aquilina et al., 2006; Schwinning, 2008); Thus, there is substantial memory in near-surface soil/epikarst (0–5 m) storage that can partition rainfall into plant-available water that moves upward through ET or downward through recharge, potentially at different rates.

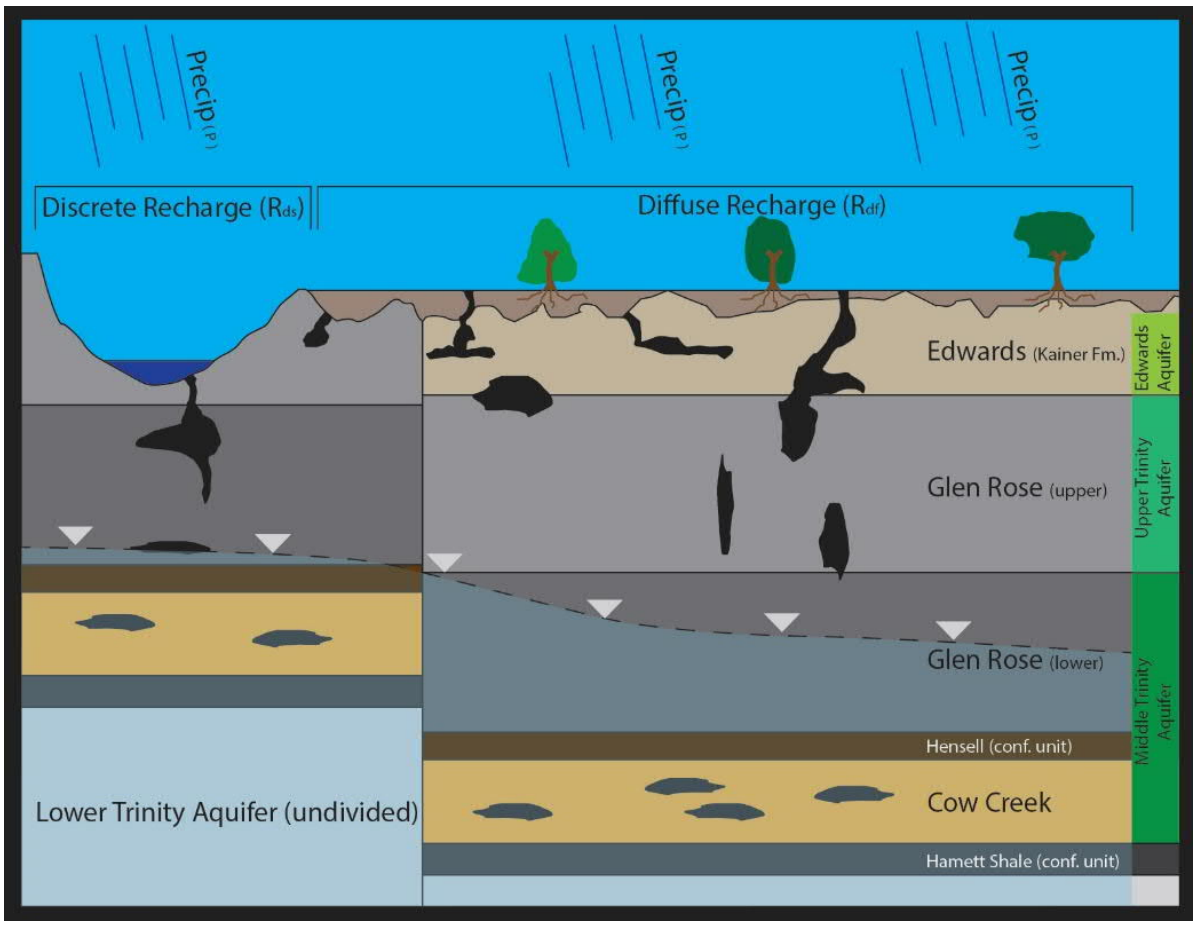


Figure 1-1: Conceptualization of discrete and diffuse recharge mechanisms, and the geologic and hydrologic units of the Edwards-Trinity aquifer system (from Gary and Smith, 2017).

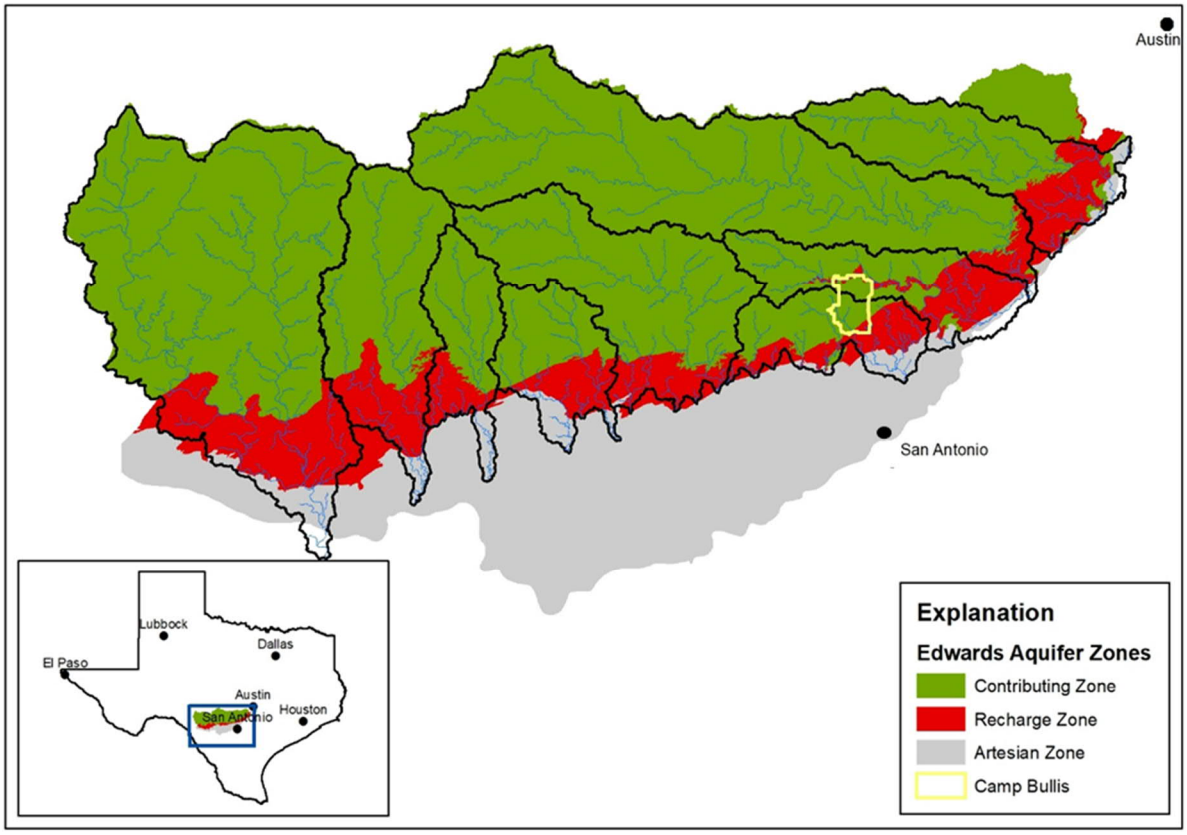


Figure 1-2: The location of the study area in Camp Bullis within the nine major drainage basins of the Edwards/Trinity aquifers. The primary aquifer zones include the drainage areas of the contributing zones to the north, the recharge zone of exposed Edwards Limestone over the Balcones Fault system, and the artesian or confined zone.

This project developed and evaluated a water balance–based estimate of inter-stream (or diffuse) recharge (R_{df}) through multiscale observation and modeling. It focuses on karstic landscapes in which the presence of woody vegetation points to substantial water loss pathways through ET. The goal of this study is, therefore, to investigate the R_{df} component of the water balance and improve recharge estimates by exploring actual evapotranspiration (ET) in more detail. This goal also required that we compare different methods of estimating the water balance components and then use the best method based on spatio-temporal stability and comparisons with field-collected data. Once the best methods were determined, they were incorporated into a workflow used to achieve the above-stated goal. The steps in the workflow include:

1) Develop a regression relationship between point-scale ET estimates using eddy covariance measurements with remote sensing data;

2) Implement the relationship to produce basin-scale maps;

3) Evaluate the components of the water balance including precipitation, ET, runoff, and storage; and

4) Quantify the recharge and compare to other methods of estimation.

2. Materials and Methods

2.1 Study Area

Camp Bullis (28,000 acres or ~12,730 hectares), a U.S. Department of Defense training facility in northern Bexar County, Texas, overlays both the Edwards and Trinity aquifers. The extent of Camp Bullis covers two drainage basins: Cibolo/Dry Comal Creek Basins (henceforth called Cibolo) and the area between the Medina River and Cibolo/Dry Comal Creek Basins (henceforth called Medina-Cibolo) (Figure 1-2).

Based on normal precipitation (the average over the past 30 years), in Cibolo Basin, the highest annual rainfall occurs in the northwest part and gradually decreases toward the southeast. Medina-Cibolo has higher rainfall in the north with less annual rainfall in the south (PRISM, 2018) (Figure 2-1). The mean-annual precipitation for the Cibolo and Medina-Cibolo Basins are 879 and 857 mm/year respectively. The landcover for Cibolo Basin is mostly savanna (42%) and evergreen forests (37%) with isolated areas of developed lands (13%) and grasslands (6%) (Yang et al., 2018). The Medina-Cibolo Basin has distinctly more development, especially in the eastern portion of the basin (37%). The remainder of the landcover is evergreen (cedar) forest (36%) and savanna (23%). Within the Camp Bullis area, above average rainfall was recorded in 2016 and 2018, whereas below average rainfall was recorded in 2017.

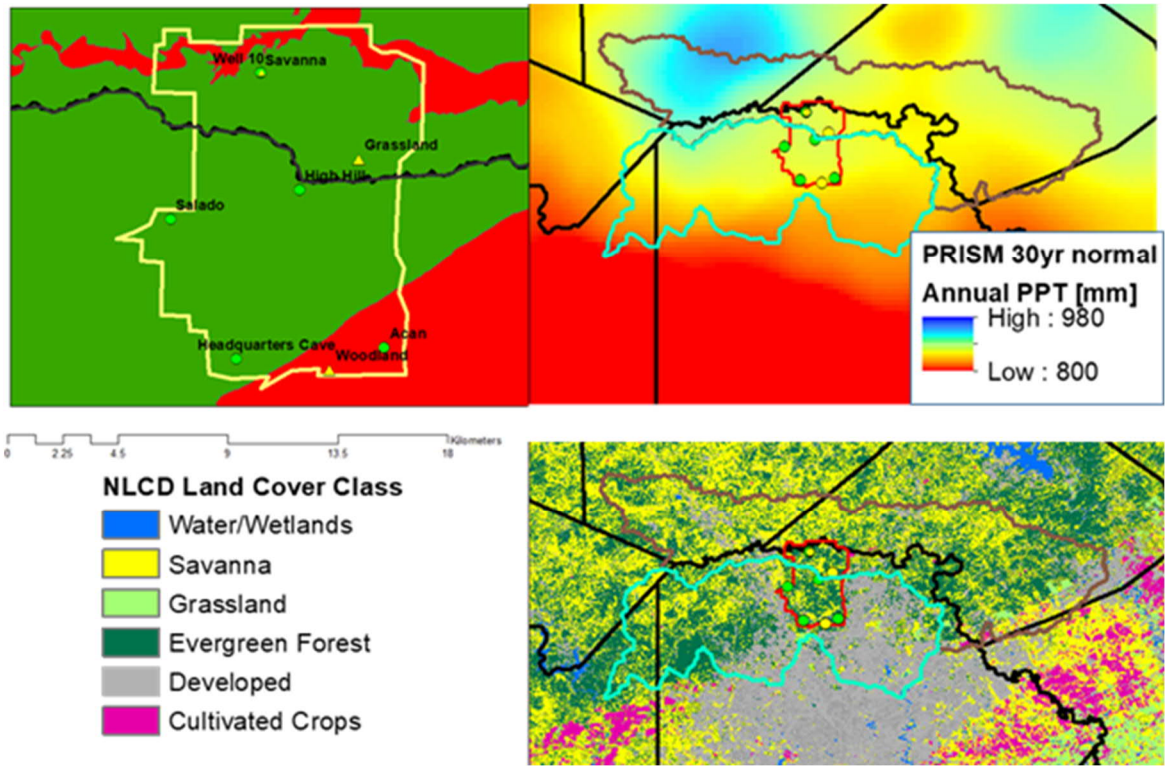


Figure 2-1: The Camp Bullis study region and the two major drainages, with Cibolo to the north and the area between Medina and Cibolo to the south. The green dots are Edwards Aquifer Authority PET (potential evapotranspiration) stations and the yellow dots are eddy covariance stations. The upper right figure is the 30-year normal annual precipitation (PRISM, 2018). The bottom figure is the landcover from the National Land Cover Database (NLCD) (Homer et al., 2020).

The Edwards Aquifer is contained within the Cretaceous limestones of the Edwards Group, which is capped by the Del Rio Clay and overlies the Trinity Aquifer. The Trinity Aquifer consists of three parts, including the upper member of the Glen Rose Formation at the top, the lower member of the Glen Rose Formation and the Cow Creek Limestone Member in the middle, and the Hosston Formation and overlying Sligo Formation at the bottom part (Mace et al., 2000). The Edwards Plateau is exposed Edwards Limestone in the drainage area and is bounded by the Balcones Fault Zone, forming the southern boundary of the Texas Hill Country. Displacement across the high-angle normal fault network can result in offset of the Edwards Group of over 100 m (Maclay and Small, 1983), creating flow paths with significantly variable permeability (Figure 1-1). The upper and lower confining units of the Edwards are the Del Rio Clay and Glen Rose respectively, which sit atop the Trinity Group. The Edwards Group (Kainer Formation) covers the southeastern third of Camp Bullis; the northern two-thirds is primarily Glen Rose limestone (Gary et al., 2013).

Soils of the Edwards aquifer Contributing Zone are generally calcareous, loamy skeletal soils derived from interbedded dolostone and limestone and are more generally called the Brackett-Eckrant-Real group. The more resistant layers of the Glen Rose Formation are interbedded with softer "marl" that erodes more easily, resulting in a tread-and-riser landform pattern (Wilcox et al., 2007; Woodruff and Wilding, 2008) with shallow soils of the series on the treads and deeper soils on hillslopes and valley bottoms.

Vegetation on the Edwards aquifer is a mix of grassland savannas and woodlands, with the latter encroaching into grasslands over the past 150 years (van Auken, 2000). Woody vegetation at Camp Bullis consists primarily of Texas live oak (*Quercus fusiformis*) and evergreen Ashe juniper (*Juniperus ashei*) with an understory of curly mesquite shortgrass (*Hilaria belangeri*); bunchgrasses such as purple three-awn (*Aristida purpurea*), sideoats grama (*Bouteloua curtipendula*), and red grama (*Bouteloua trifida*); and sedges (*Carex texensis*). Both live oak and juniper are evergreen, transpiring throughout the winter season. Grasses, however, are seasonally dormant.

2.2 Diffuse Recharge and Quantification of the Water Balance

The availability of water resources and its accounting begins with a water budget that is applied over some representative volume. The water balance equation is shown as:

$$P + Q_{in} = ET + dS + Q_{out} \qquad \textit{Eq. 2.2-1}$$

in which P is precipitation, Q_{in} is flux in as either surface or subsurface flow, dS is change in storage either in the soil or epikarst, and Q_{out} is flux out as either surface runoff (RO) or diffuse recharge (R_{df}). If Q_{in} is negligible, R_{df} is obtained as the residual of the water balance equation:

$$R_{df} = P - ET - dS - RO \qquad \textit{Eq. 2.2-2}$$

In this research, we quantify these components at varying temporal and spatial scales, obtaining representative estimates of R_{df}. We assess several products to get accurate values of P across the study areas, and we use eddy covariance (described in Section 2.3) to measure ET at specific areas and satellite data to upscale it. In this framework, dS is measured at the same locations as the eddy covariance (i.e., only a few specific locations) and RO is essentially unmeasured. However, using a numerical model, in this case HYDRUS, we can estimate each component, compare the upscaled fluxes, and resolve the potential error in R_{df}. Figure 2-2 shows the generalized workflow.

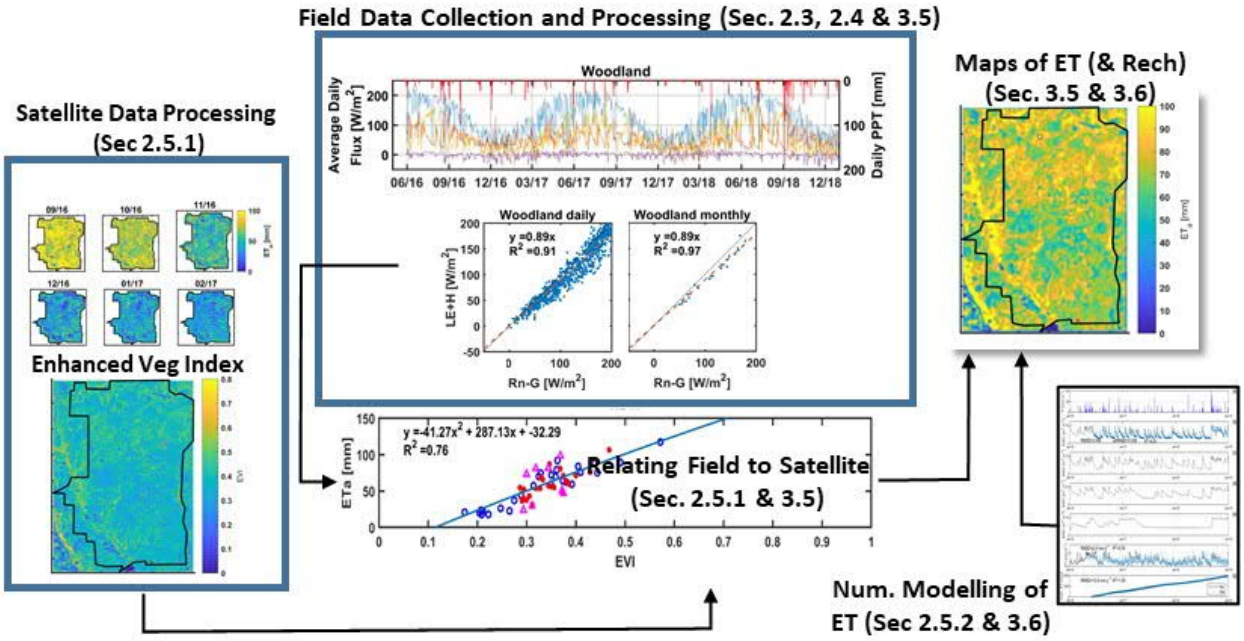


Figure 2-2 Generalized workflow of data collection and analysis

2.3 Field Data Collection

As inferred above, the approach used in this research uses field data as control points for satellite remote sensing and other products available from public and private third parties. Here, field data collection primarily focuses on eddy covariance (EC) as a direct measurement of ET upwind of the instrument (known as the fetch) over an area with a radius approximately 100 times the instrument height above the canopy (known as the flux footprint). This footprint depends on instrument height, surface roughness, and thermal stability. EC towers consist of a three-dimensional anemometer to measure temperature and wind speed and direction, and an open-path infrared gas analyzer (IRGA) to measure gas concentration (i.e., humidity and carbon dioxide). Instruments are installed approximately 1.2 m above the height of the vegetation. The anemometer rapidly samples rotating eddies of air to determine the total concentration of vertical flux, which is represented as the covariance of the vertical component of the velocity (measured with the anemometer) and the concentration of water (or carbon) molecules (measured with the IRGA). Measurements are usually collected every 0.1 seconds and averaged over 30 minutes to obtain 48 ET flux estimates per day.

Three EC sites were operated from 2016 to 2018 on the Camp Bullis property on three different landcovers (Figure 2-1), so that differences in ET between the sites could be attributed to differences in vegetative function. These three sites include (i) a closed cedar woodland (Woodland) in the southern portion, (ii) an open oak savanna (Savanna) located in the northern portion, and (iii) a centrally located open grassland (Grassland). At the Woodland and Savanna sites, an open-path analyzer and sonic anemometer (IRGASON, Campbell Scientific, Logan, UT) with an electronics module (EC100, Campbell Scientific) was used to measure high-frequency (10 Hz) water vapor and CO_2 fluxes, as well as the three-dimensional wind vector. At the Woodland site, the IRGASON was installed on scaffolding at a height of 8.6 m facing south (173°), with measurements starting in May 2016. At the Savanna site, the IRGASON was mounted on a tower at a height of 7.3 m facing south-southeast (130°), with measurements starting in April 2016. At the

Grassland site, an LI-7500a (LI-COR, Lincoln, NE) measured 10-Hz water vapor and CO_2 concentrations, and a 3D-sonic anemometer (CSAT3, Campbell Scientific) facing south-southeast (170°) measured the wind vector; both instruments were installed on a tripod at a height of 2.5 m, with measurements beginning in April 2017. Calibration of the EC systems occurred biannually, during November and May of each season, using a dew point generator (LI-610, LI-COR), a zero-gas reference, and a 500-ppm CO_2 standard. Data collection ceased in January 2019 at all sites.

Additional meteorological and soil moisture measurements were collected at each site every 30 minutes. All stations were equipped with a tipping-bucket rain gage (TE525, Texas Electronics, Dallas, TX) and a net radiometer (NR-LITE or CNR4, Kipp & Zonen, Delft, Netherlands). Soil moisture and temperature were recorded using water content reflectometers (CS655, Campbell Scientific) at a depth of 5 cm under the canopy and in the interspace at the Woodland site. At the Savanna site, CS655 sensors were installed at 5 (2), 10, and 20 cm. At the Grassland site, five water content sensors (model TDR315, Acclima, Meridian, ID) were used to measure soil moisture and soil temperature at 5 (2), 10, 20, and 50 cm depths. Soil heat flux was measured using soil heat flux plates (HFP01, Hukseflux Delft, Netherlands) installed at a depth of 8 cm and two averaging thermocouples (TCAV, Campbell Scientific) installed above and below the heat flux plate.

Much like the water balance (Eq. 2.3-1), EC fluxes are validated against the energy budget (units of W/m^2) given as:

$$Rn - (H + LE + G) = 0 \qquad \textit{Eq. 2.3-1}$$

in which net radiation (*Rn*) is balanced by sensible heat (*H*), the latent heat of vaporization (*LE*), and soil heat flux (*G*) according to conservation of energy. Additional details are presented in Section 2 of the Appendix. Here, we force energy balance closure, assuming any extra energy is evenly split between the measured EC fluxes of H and LE.

Lastly, five PET stations (HOBO, Onset Computer Corporation, Bourne, MA) operated by Edwards Aquifer Authority (EAA) provide 15-minute data of air temperature, atmospheric pressure, dew point temperature, relative humidity, wind speed and direction, soil moisture, and precipitation from 2016 to 2018. The EAA-operated stations are shown in Figure 2-1 and include Headquarters Cave and Acan in the southern portion of Camp Bullis, Salado and High Hill in the middle, and Well 10 which is very near the Savanna site in the northern extent of Camp Bullis.

2.4 Datasets and Processing

2.4.1 Precipitation

Multiple precipitation datasets were compared to in-situ data measurements taken at Camp Bullis. One such dataset, known as the Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate mapping system, combines topography and gage data to produce daily precipitation at 4 km resolution (Daly et al., 2008). Daily data for PRISM are available to 1981, but PRISM has a historical monthly dataset that is modeled by starting with the 30-year normal going back to 1895. Another, the North American Land Data Assimilation System Phase 2 (NLDAS-2), is made available at 0.125° spatial resolution and 1-hour temporal resolution from 1979 to the present. Precipitation from NLDAS is derived by disaggregating daily precipitation gage data from the Climate Prediction Center, located at the

National Centers for Environmental Prediction (NCEP) and using the PRISM algorithm with an orographic enhancement, whereas other data from NLDAS-2 like wind speed, air temperature, and specific humidity are derived from NCEP North American Regional Reanalysis (NARR) (Mesinger et al., 2006). RainVieux, a third precipitation dataset, produces hourly gage-adjusted Nexrad rainfall data at a 4-km spatial resolution for the San Antonio segment of the Balcones Fault Zone of the Edwards Aquifer from 2003 to present. The final precipitation dataset used in this study was gridMET (Abatzoglou, 2013). The gridMET dataset, which is a daily dataset at ~4-km (1/24°) spatial resolution and spans from 1979 to the present, blends the spatial attributes of PRISM with the temporal attributes of NLDAS-2. PRISM and Gridmet are available at daily timesteps while NLDAS and Vieux are available in hourly timesteps.

Monthly precipitation from PRISM, NLDAS-2, RainVieux, and gridMET were compared to observations from our five PET stations and three EC stations. The coefficient of determination (R^2) and root mean squared difference (RMSD) were used to assess performance. R^2 is defined as:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \right)^2 \qquad \textit{Eq. 2.4-1}$$

in which $\bar{X}$ and $\bar{Y}$ are the mean of the *X* and *Y*, and *n* is the total number of observations. An R of ~1 indicates a perfect linear correlation. RMSD provides an error measurement in the comparable units and a minimal value being ideal:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n}} \qquad \textit{Eq. 2.4-2}$$

2.4.2 Potential Evapotranspiration (PET)

The Penman-Monteith method (Allen et al., 1998) requires maximum and minimum temperature, daily average dewpoint temperature (equivalently, vapor pressure or vapor pressure deficit or relative humidity), wind speed and downward shortwave radiation. Temperature, relative humidity, wind speed, and solar radiation for the five meteorological stations were compared on a daily basis to NLDAS and gridMET using linear regression. Pressure, which is an output at the meteorological stations, is not included in the gridMET dataset but can be estimated using elevation:

$$P = 101.3\left(\frac{293 - 0.0065z}{z}\right)^{5.26} \qquad \textit{Eq. 2.4-3}$$

According to Allen et al. (Allen et al., 1998), a high R^2 and a slope between 0.7 and 1.3 indicates good relationship between data. NLDAS and gridMET wind speed, which are both expressed at 10 m height, were converted to 2 m assuming a vegetation height of 0.12 m and logarithmic wind profile. PET between the five meteorological stations was also compared to gridMET and NLDAS. NLDAS uses the modified Penman scheme of Mahrt and Ek (1984) while gridMET uses the Penman-Montieth equation (Allen et al., 1998). A more detailed explanation of Equation 2.4-4 can be found in Section 1 of the Appendix:

$$PET = \frac{\Delta(Rn - G) + \rho c_p\left(\frac{e_s - e_a}{r_a}\right)}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)} \qquad \textit{Eq. 2.4-4}$$

2.4.3 Soil Water Storage

The change in soil water storage (dS) from 0–10 cm was obtained from the 5-cm soil moisture sensor for each meteorological station and compared to the monthly 0–10 cm dS simulated by Noah (Chen et al., 1996; Ek et al., 2003), Mosaic (Koster and Suarez, 1992), and VIC (Liang et al., 1994; Wood et al., 1997), which are three different land-surface models incorporated into the NLDAS-2 suite from 2016 to 2018. The soil water (in mm) estimated at the beginning of the month was subtracted from the soil water estimated at the end of the month to get the monthly change in storage (dS). Because the average soil depth to a restrictive layer is less than 1 m (Soil Survey Staff, 2014) at the sites used in this study, soil water from 0–10 cm and 10–40 cm were summed to get soil water from 0-40 cm. The monthly dS of the land surface models was compared to the monthly dS using both shallow and deep soil moisture sensors. The NLDAS-Noah model was then used to determine dS in the water balance equation for the long-term recharge study.

2.4.4 Evapotranspiration

2.4.4.1 Eddy Covariance Data Processing

Data from the Savanna and Woodland sites, using the IRGASONs, were processed using Campbell Scientific's EasyFlux program. The system with the LI-COR equipment (at the Grasslands site) was processed similarly using LI-COR's EddyPro software. Section 2 of the Appendix describes the EC processing steps that include many corrections, including frequency corrections and air density, fluctuations in sonic temperature and humidity, and axis and tilt rotation. Further processing such as gap filling must be done, which is also described in Appendix Section 2. Due to factors still being studied in the scientific community, the surface energy fluxes are often underestimated by 10–30% relative to the available energy (Foken et al., 2012; Wilson et al., 2002a). LE and H can be adjusted to force closure while maintaining a constant Bowen ratio (β), which is the ratio between H and LE as shown in Equations 2.4-5 and 2.4-6 (Blanken et al., 1997; Lee, 1998; Twine et al., 2000):

$$LE_{corr} = \frac{(R_n - G)}{1 + \beta} \qquad \textit{Eq. 2.4-5}$$

$$H_{corr} = LE_{corr} \times \beta \qquad \textit{Eq. 2.4-6}$$

Energy balance (Eq 2.3-1) states the difference between R_n and G (the available energy) should equal the sum of the sensible heat flux (H) and latent heat flux (LE) (the surface energy flux). Thus, ET has a direct relationship with the latent heat flux described in Equation 2.4-7:

$$ET = \frac{LE}{\lambda \rho_w} \qquad \textit{Eq. 2.4-7}$$

in which LE is latent heat flux, λ is the latent heat of vaporization of water, and ρ_w is the density of the water.

2.5 Satellite-Derived and Numerical Model Estimates of ET

2.5.1 Satellite-Derived Evapotranspiration

Vegetation indices (VI) and pixel quality from Landsat 7 and 8 Collection 1 for Camp Bullis were acquired through USGS's Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) (https://espa.cr.usgs.gov/). The spectral indices include normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), modified soil adjusted vegetation index (MSAVI), and enhanced vegetation index (EVI). Each index is an indicator of vegetation greenness, plant health, and canopy cover. The high chlorophyll absorption in green vegetation reflects little radiation into the visible spectrum but more into the near-infrared due to leaf tissue and water content scattering the light. NDVI is the simplest vegetation index and uses the near-infrared and red bands. SAVI adds a constant soil-brightness correction factor, whereas MSAVI improves SAVI by using a recursive soil-brightness function. Finally, EVI not only accounts for soil background, but also includes an atmospheric resistance term and a blue band that accounts for background noise, atmospheric noise, and saturation. Further explanation of the vegetation indices is described in Section 3.1 of the Appendix. The VIs were masked for clouds and interference using the quality-assessment band.

A monthly maximum composite of each VI at each EC station was determined using monthly averages measured within a radius that approximates the upwind footprint containing 90% of the ET flux. A second-order polynomial was implemented to relate the monthly observed ET to the composite VI, using the VI corresponding to 0 (no lag), +1 (month), and −1 (month) lags. Lags were investigated to determine if the plants stopped transpiring before there was a reduction in the VIs. The final regression error for EVI was found using a 1000-run bootstrap that sampled 33% of the monthly ET data for training and the remaining 67% for validation.

The regression equation between EVI and ET was implemented from 2000 to 2018 on the two major drainage basins that overlie Camp Bullis. Google Earth Engine was used to find the maximum monthly EVI from Landsat 5, 7, and 8 Collection 1 over the major drainages. The maximum monthly VI composite for the basins was filled at time (t) by linearly interpolating at t by first using t ±1 then t ±2. The regression was then applied to the gap-filled composite EVI to develop monthly ET maps. The upper end of ET was limited to the potential evapotranspiration (PET) obtained from gridMET. If VI-derived ET was less than zero, it was replaced with zero.

Multiple gridded ET products were compared using the RMSD, R^2, and mean bias error (MBE) for the results of the monthly VI-derived ET for the two study basins for 2016 through 2018. The MBE is used to find the average bias and can be found using:

$$MBE = \frac{\sum_{i=1}^{n}(X_i - Y_i)}{n} \qquad \textit{Eq. 2.5-1}$$

in which *X* and *Y* are the simulated and observed values respectively. Products used for comparison include those from land-surface models like NLDAS-Noah, NLDAS-Mosaic, and NLDAS-VIC. The mean of the land surface model was also evaluated against the ET, as Xia et al. (Xia et al., 2014) found that using a mean ensemble of the land surface model performs better than the individual models. Finally, the evapotranspiration product from the Moderate Resolution Imaging Spectroradiometer (MODIS) (MOD16A2) for an 8-day composite with a 500-m resolution was also used for comparison. A description of these models can be found in Section 3.2 of the Appendix.

2.5.2 Numerical Model Estimates

For numerical simulations, we used HYDRUS-1D (Simunek et al., 2012; Simunek et al., 2016), which simulates water flow in variably saturated porous media using a finite element method to solve a one-dimensional Richards equation based on a mass-conservative iterative scheme (Celia et al., 1990). The one-dimensional simulations were solved over a 500 m grid parameterized with soil data at 30 m resolution, spatially explicit (4 km) daily rainfall totals from RainVieux, potential evapotranspiration from gridMET (Abatzoglou, 2013), and vegetation class (Friedl et al., 2010) and weekly leaf area index from MODIS (Myneni et al., 2015) at 500 m resolution. The 30 m soil data were scaled up to 500 m using nearest neighbor method, while the 4 km rainfall and potential evapotranspiration were scaled down to 500 m, also using the nearest neighbor method.

The model was parameterized with hydraulic soil properties obtained from the POLARIS data set (Chaney et al., 2019) that derived 30 m probabilistic soil properties from terrain and soils data in the National Cooperative Soil Survey Soil Geographic Database (SSURGO) and gridded SSURGO database (Soil Survey Staff, 2014). In addition, the SSURGO depth of soil to restrictive horizon was rasterized onto the POLARIS grid and used to define the maximum depth of the soil profile within a given POLARIS soil layer. POLARIS consists of six soil layers with boundaries at depths of 5, 15, 30, 60, 100, and 200 cm. Monthly fluxes (ET) were compared between the HYDRUS model and ET-EVI. Recharge estimates are classified as R_{df}. Basin recharge was found by summing values from each 500 m x 500 m cell over the whole drainage basins. These recharge estimates were then compared to the total recharge from the USGS-(Puente, 1978) method.

Model performance was assessed using RMSD (eq. 2.4-2); unbiased root mean squared difference (ubRMSD), in which the mean values of both the simulation and observations is removed; coefficient of determination (R^2); and Nash-Sutcliffe model efficiency (NSE) (Nash and Sutcliffe, 1970) as:

$$NSE = 1 - \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \overline{X}_i)^2}} \qquad \textit{eq. 2.5.2-1}$$

in which X and Y are daily observed and simulated values respectively, and $\bar{X}$ is the mean of the observations. For a model to perform better than using $\overline{X}$, NSE must be >0; perfect efficiency is achieved with an NSE of 1. A threshold value of NSE >0.5 is generally suggested (Moriasi et al., 2007; Ritter and Muñoz-Carpena, 2013), below which the relationship is not considered acceptable.

More detail on the vadose-zone numerical modeling is presented in Section 5 of the Appendix.

2.5.3 Additional Recharge Modeling

Recharge modeling was performed using multiple approaches, ranging from analytical solutions to fully numerical solutions, to cross-validate the EVI-ET results.

2.5.3.1 Water Balance

Recharge is estimated using a water balance as the difference between water inputs and outputs, namely, precipitation as input and ET, runoff, and soil water storage as output from 2000 to 2018. Precipitation, ET, and soil water storage were discussed above. Runoff was evaluated at yearly and monthly time scales from discharge measurements taken at gage 08185000 (Cibolo Creek in Selma, TX) for the Cibolo

watershed and gage 08178700 (Salado Creek at Loop 410, San Antonio, TX) for the Medina Cibolo watershed. We then compared our calculated recharge to estimates from the USGS and methods described in Section 2.5.3.3.

3. Results

3.1 Precipitation

In a direct comparison between gridded precipitation datasets of varying spatial resolution, PRISM, RainVieux, and gridMET precipitation products yielded far better performance statistics than NLDAS when compared to the EAA/UT operated stations (

Table 1). Whereas the RMSD for PRISM, RainVieux, and gridMET were on the order of 10–20 mm/month and R^2 was near 1, the RMSD for NLDAS monthly rainfall was generally 80–100 mm with low correlation ($0.4 < R^2 < 0.6$) to gages. The minimum RMSD for NLDAS was 68 mm with a maximum R^2 <0.60. The average RMSD for PRISM was 23 mm, whereas RMSD for both RainVieux and gridMET was 19.5 mm. We found that the PRISM, RainVieux, and gridMET precipitation products were highly correlated to the precipitation stations operated during this study (R^2 >0.90), with the smallest RMSD of 5.1 mm obtained for the RainVieux dataset. A major difference between these precipitation products is that NLDAS assumed an average of 185 mm of rainfall from September to October 2018 (Figure 3-1), whereas PRISM, RainVieux, and gridMET estimated 700, 599, and 703 mm, respectively; the monitoring stations measured an average of 671 mm with a standard deviation of 66 mm. We note that PRISM and gridMET products are first released as a provisional product; final products are then released at least 1 month later. Also, the spatial scale of NLDAS is much larger at ~12.5 km versus RainVieux and gridMET at ~4 km. RainVieux tends to underpredict precipitation by an average of 4.9 mm/month, whereas PRISM and gridMET tend to overpredict precipitation by an average of 6.5 and 8.5 mm respectively. The RMSD for the southern stations (Woodland, Headquarters Cave, and Acan) was smallest for the gridMET precipitation product. Results therefore indicate that NLDAS should not be used further, whereas the performances of the other 3 products were comparable.

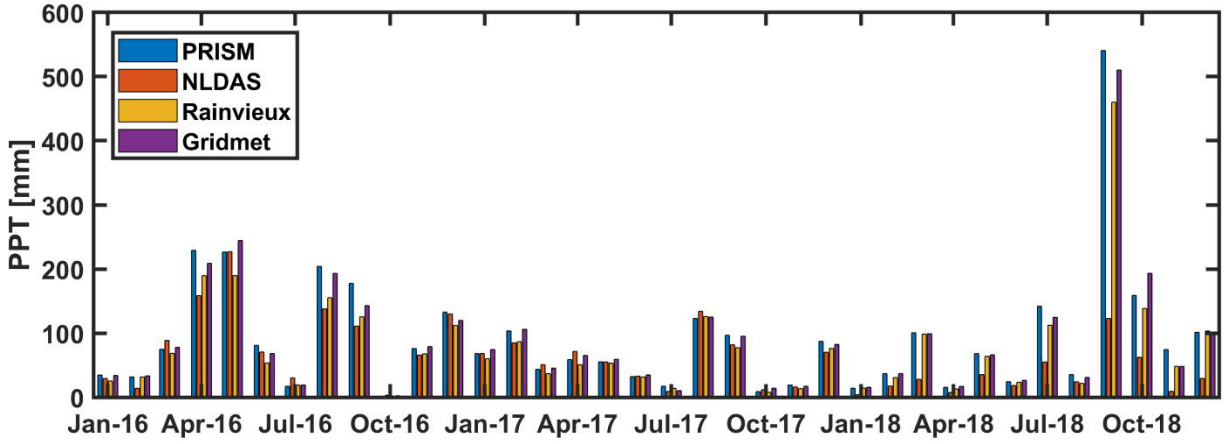


Figure 3-1: Precipitation (PPT) for the Camp Bullis area from 2016 to 2018.

We assume that the rain gages used in this study are accurate, level, and properly functioning; however, tipping-bucket rain gages tend to under-collect due to inefficient tipping under high-intensity events and raindrops splashing out of the small-diameter orifices of the gage itself (WMO, 2008). Here, we note that radar corrections for precipitation typically use data from ground gages, and the RainVieux product likely ingests EAA gages into its correction algorithm; thus, they are not independent. The gages used at the EC sites managed by The University of Texas are not used to correct RainVieux, which explains why the RMSDs are generally higher and correlation slightly lower for the other precipitation products. Also, neither PRISM nor gridMET were any less accurate than RainVieux; however, both are more readily available at no-cost at daily timesteps and with a spatial resolution (4 km), similar to RainVieux (an 800 m PRISM product is available for a cost). The latency for both gridMET and PRISM is 2–3 days, and both datasets include other ancillary variables such as daily maximum/minimum air temperature and humidity. Wind speed and solar radiation are available only from gridMET.

3.2 Potential Evapotranspiration (PET)

The cumulative distributions of maximum, minimum, and average air temperature and humidity, average wind speed, and solar radiation from the five EAA PET stations, NLDAS, and gridMET indicate good correspondence between air temperatures and solar radiation at each of the sites, but discrepancies in windspeed were higher (Figure 3-2). In particular, the cumulative error in PET (Figure 3-2f) shows that PET from NLDAS is consistently higher. Comparisons of temperature and minimum relative humidity between gridMET and the EAA stations were quite good ($R^2 > 0.7$) (Table 2). The correlation for windspeed, however, varies from $R^2 = 0.23$ for Headquarters Cave to $R^2 = 0.72$ for Well 10. Wind speed obtained from NLDAS and gridMET products are consistently overpredicted when compared to the EAA meteorological stations. The lowest correlation ($R^2 = 0.29$) was recorded for maximum relative humidity. Results show that PET for gridMET is more comparable to the meteorological stations than NLDAS (Table 3). The average RMSD for PET using a daily timestep is 1.54 mm/day for gridMET and 3.80 mm/day for NLDAS. On a monthly timestep, the average RMSD for gridMET is 37 versus 109 mm/month for NLDAS. Overall, NLDAS seems to yield a significantly high bias in PET. Time series of the monthly values of the sites closest to the EC towers are shown in Figure 3-3.

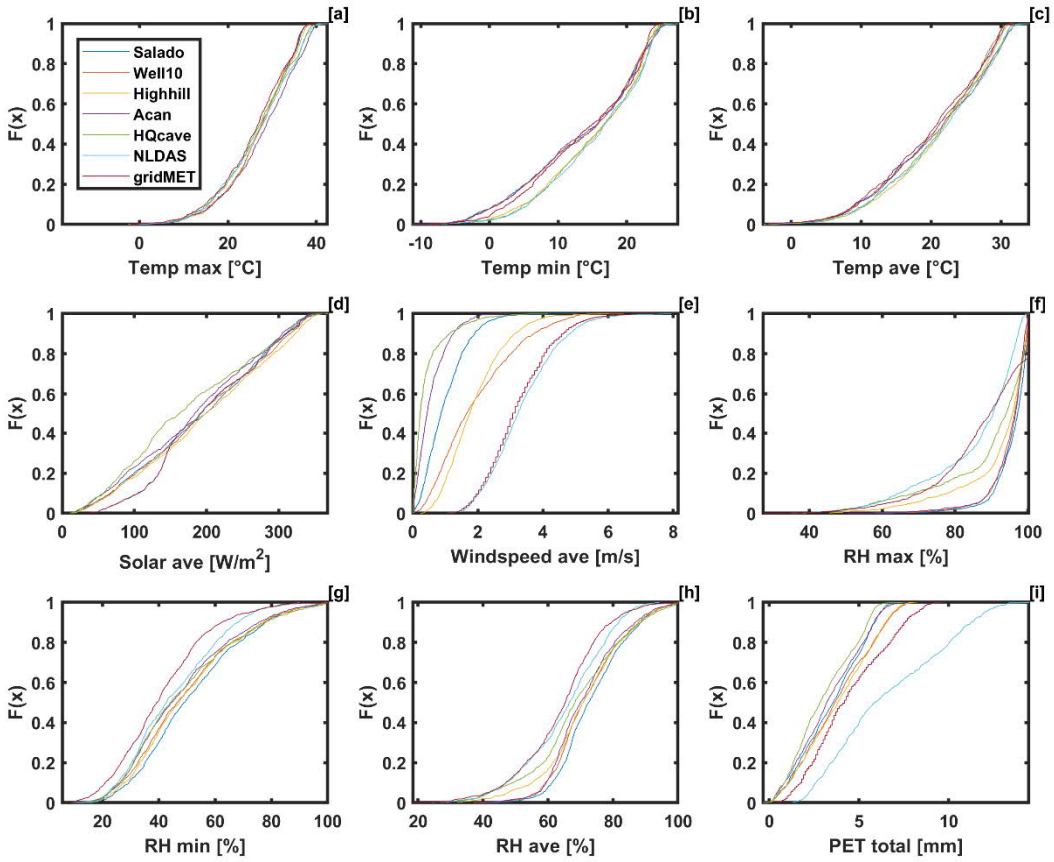


Figure 3-2: Comparison of daily meteorological variables from Edwards Aquifer Authority PET stations and land surface models (NLDAS and gridMET).

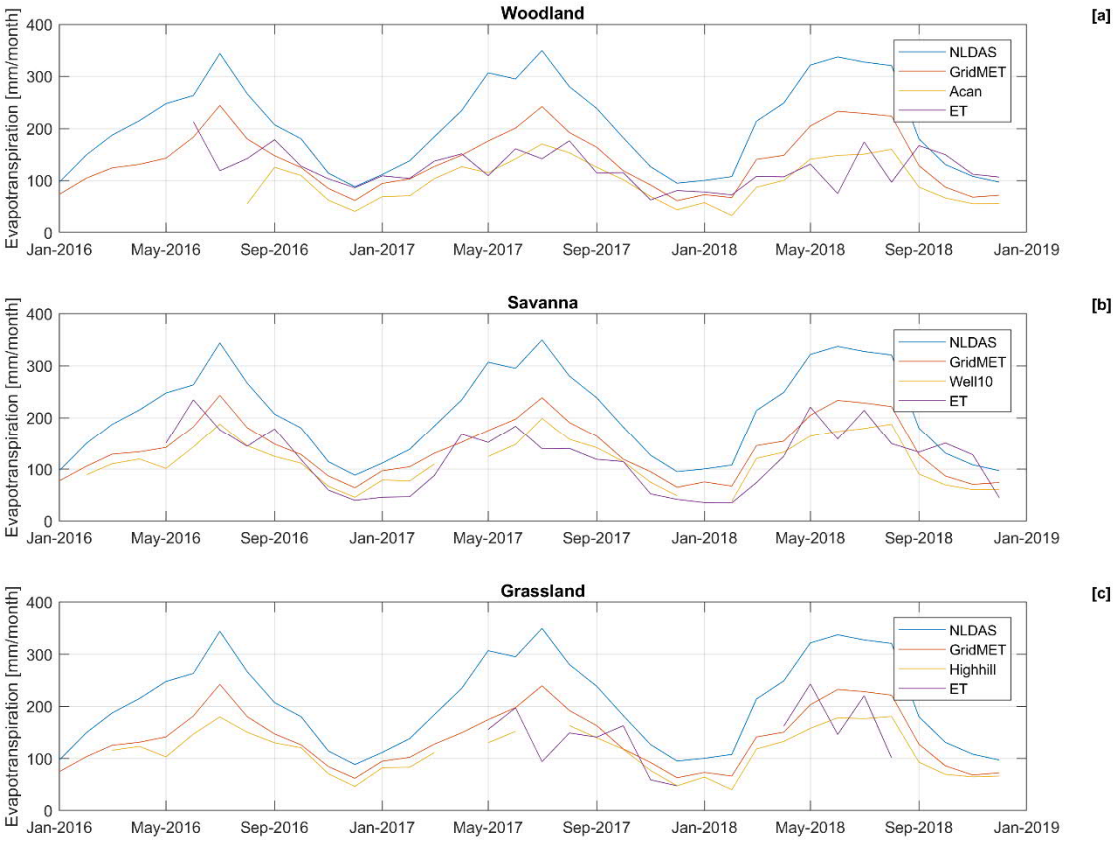


Figure 3-3: Time series comparison of ET between the land surface models in NLDAS, which use the modified Penman scheme of Mahrt and Ek (1984), gridMET, which uses Penman-Montieth evapotranspiration, Penman-Montieth evapotranspiration from the nearest Edwards Aquifer Authority site (Highhill), and ET from the eddy covariance site.

3.3 Soil Water Storage

For water balance estimates of R_{df}, the change in soil water storage may be extracted from the NLDAS model output suite which includes the Noah, Mosaic, and VIC land-surface models, all driven by the same atmospheric forcings. Table 4 shows the comparison to observations at PET and EC sites at Camp Bullis. When comparing the 0–10 cm monthly dS, on average VIC is the least comparable to the station data (R^2 of 0.50 and RMSD of 5.9 mm/month) whereas Noah is the most comparable (R^2 of 0.80 and RMSD of 4.6 mm/month). Correlation to Mosaic was mixed, with a slightly improved R^2 (0.82) but a poorer RMSD (5.13 mm/month). Results show that the bias is less than 1 mm for all sites. When soil depth is increased to 0–40 cm, VIC continues to perform the worst (R^2 of 0.68 and RMSD of 17.36 mm/month) whereas Noah and Mosaic perform similarly.

Figure SI1 shows the comparison of the average monthly change in soil water storage (dS) from 2016 to 2018 between gridded soil moisture products. The largest range of dS was found in Mosaic, whereas the lowest range was found in VIC. The change in soil water storage for the Cibolo Basin using Mosaic ranges

from −50 to 46 mm/month, whereas VIC ranges from −14 to 16 mm/month. The monthly storage depends on the timing of the rainfall during that month. For example, if a large rainfall event occurs on the last day of the month, soil water storage will have a substantial positive storage response.

The monthly average change in soil water storage (dS) for Noah from 2000 to 2018 for both basins is less than 0.1 mm with a standard deviation of 6.5 mm for Cibolo Basin and 7.8 mm for the Medina-Cibolo Basin (Figure SI2). The average difference between the two basins is 0.1 mm with a standard deviation of 5 mm. The largest increase in storage (36 mm) for the Medina-Cibolo Basin occurred in August 2001, with the largest decrease in storage (−27 mm) occurring in September 2001.

3.4 Evapotranspiration Rates Using Eddy Covariance

After precipitation, ET is the largest flux in the water balance in semiarid environments. Figure 3-4 shows the average daily fluxes of the four energy-balance components for the three sites: LE, H, R_n and G. Following precipitation, LE is initially high but then decreases as water becomes limited, resulting in H increasing. Net radiation is lowest in the winter when LE is generally lower as energy limitations prevail. For all sites, a sinusoidal response in H was observed for all seasons, with larger amplitudes observed in the summer months. The average daily G is less than 3 W/m^2, which is a small component of the energy balance. The seasonality of the LE/(R_n − G) and H/(R_n − G) are different for the Woodland and Savanna sites. For example, LE dominated the energy balance during the winter, and H dominated the energy balance during the spring and early summer (Figure SI3). The opposite behavior was observed for the Savanna and Grassland sites. The completeness of the data after spike removal varied from 67% for the Grassland site to 84% for the Woodland site.

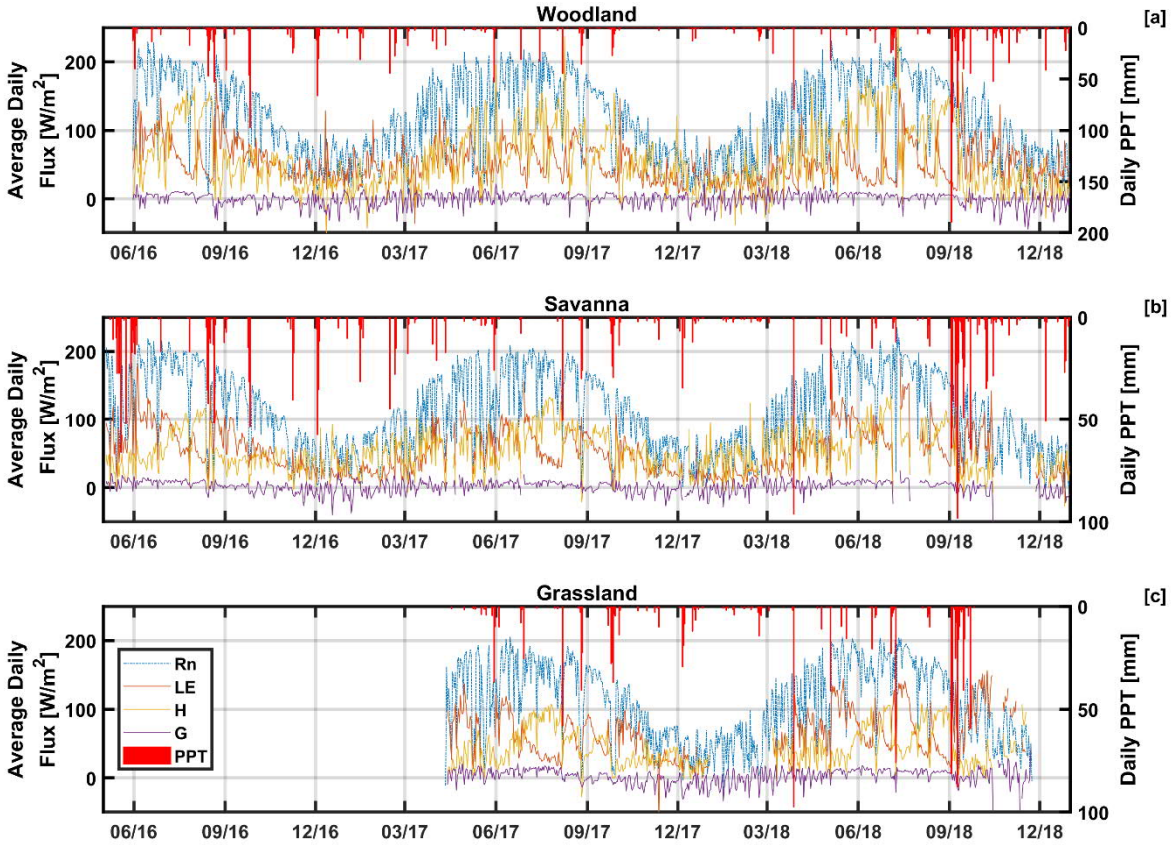


Figure 3-4: Average daily energy fluxes from the eddy covariance stations at the [a] Woodland, [b] Savanna, and [c] Grassland sites, in which R_n is the net radiation, LE is latent heat, H is sensible heat, and G is soil heat flux. PPT is precipitation.

The 30-minute (averaged) fluxes are accumulated into daily and monthly totals to illustrate the energy balance closure. Figure SI4 compares the average incoming and outgoing daily and monthly fluxes for the three sites and indicates energy balance closures of ~90%. These values are comparable to those found in Wilson et al. (Wilson et al., 2002b) which looked at energy balance closure at many eddy covariance sites around the world. For the daily values, the Grassland site had the highest R^2 but the lowest slope between the incoming (*Rn – G*) and outgoing flux (*LE + H*), indicating a systematic energy-balance issue. The slope is usually closer to 1 during the summer months. Though the average daily ET is comparable for the entire period (2.0 ±1.2 mm/day for the Savanna site, 2.0 ±1.0 for the Woodland site), ET is higher at the Savanna site during June through August and lower during December through February than at the Woodland site (Table 5 and Figure 3-5). On the other hand, the Grassland and Savanna sites tend to be better correlated throughout the year (Figure 3-6). The monthly ET rates for the three sites (Figure 3-7) indicate a bimodal distribution for the Savanna site, which is due to seasonal trends in vegetation. A tighter distribution at the Woodland site is due to ET being more persistent year-round.

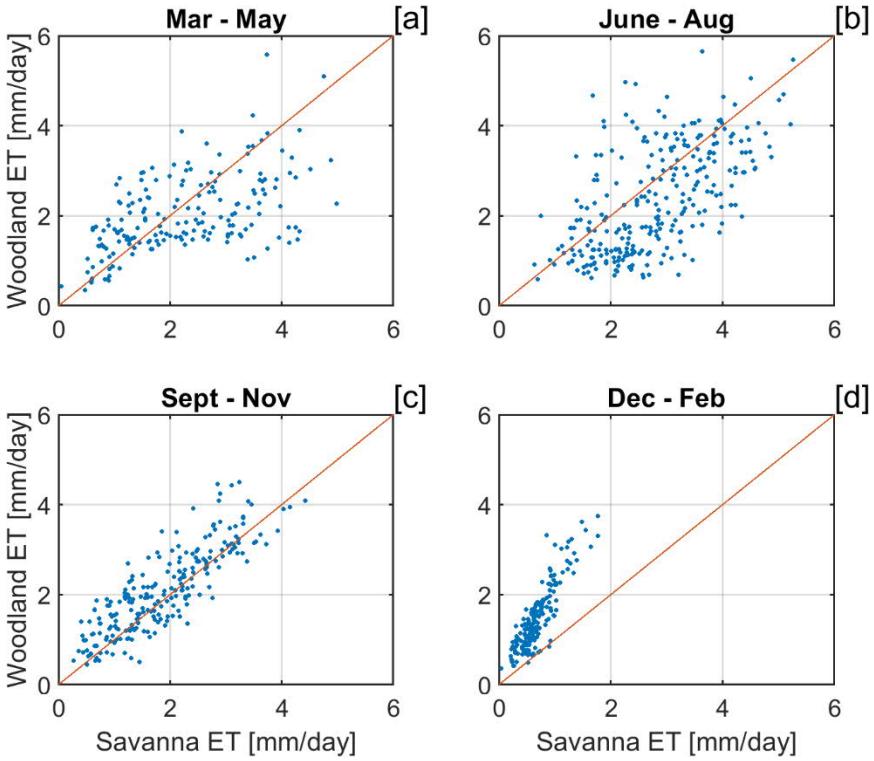


Figure 3-5: Comparison of the Woodland and Savanna daily Bowen-ratio-corrected evapotranspiration (ET_a) from the eddy covariance sites.

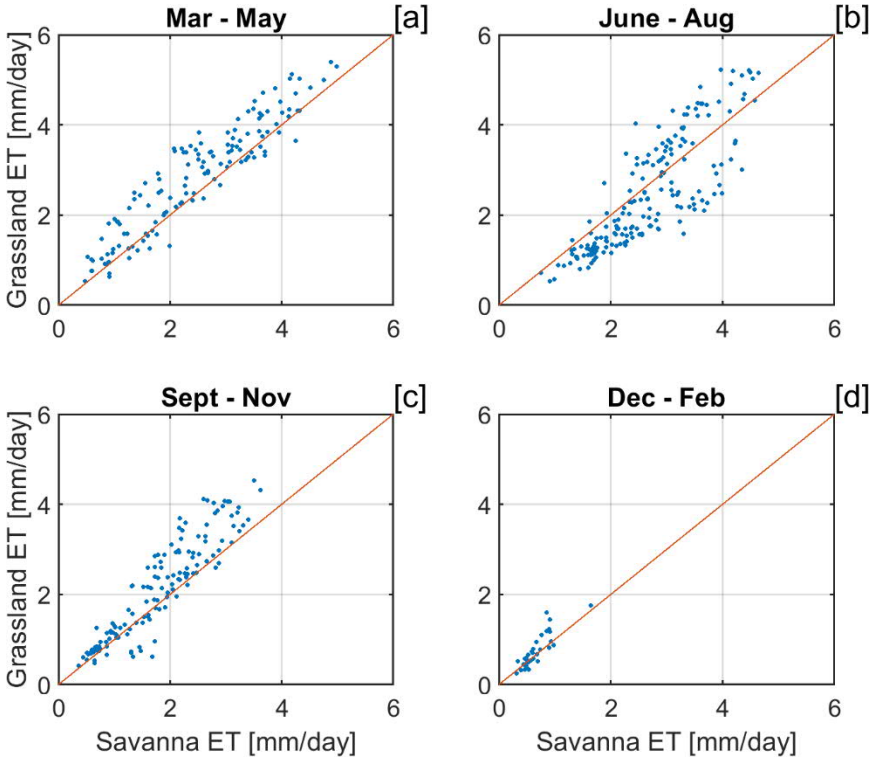


Figure 3-6: Comparison of the Grassland and Savanna daily Bowen-ratio-corrected evapotranspiration (ET) from the eddy covariance sites.

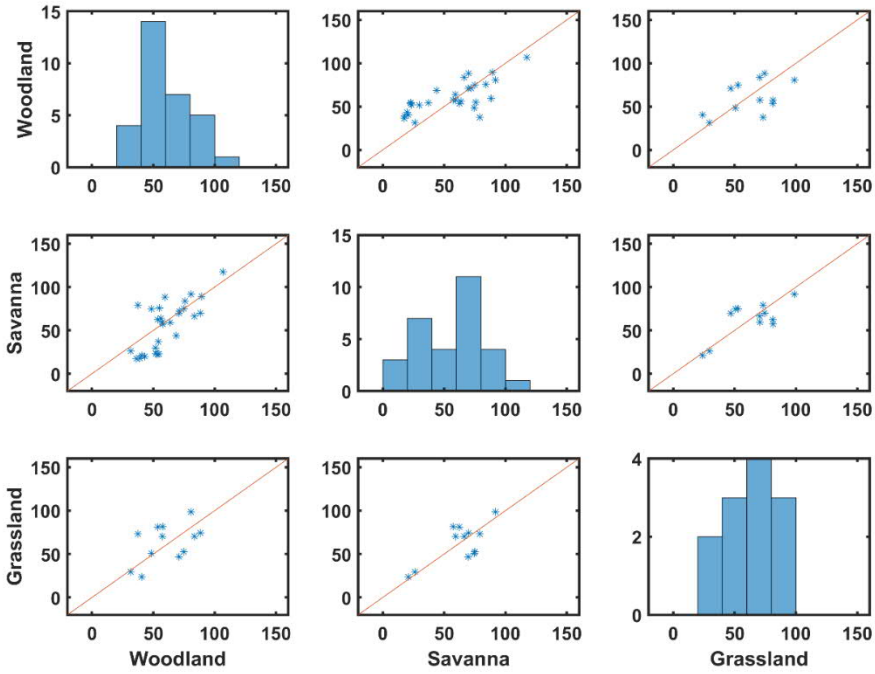


Figure 3-7: Intercomparison of Woodland, Savanna, and Grassland monthly Bowen-ratio-corrected evapotranspiration (ET) from eddy covariance.

3.5 Evapotranspiration Rates from Satellite Imagery Analysis

Figure SI5 shows the NDVI, SAVI, MSAVI, and EVI time series from Landsat for the three EC locations during the study period. The monthly maximum vegetation indices were used to create the regression versus monthly ET (Figure 3-8). In each case, ET is estimated (as the dependent variable) as a function of different vegetation indices, as determined by Landsat images obtained at the EC locations. The best regression performance occurred with zero lag and a Bowen-ratio-corrected ET (Table 6) (i.e., where energy balanced closure was forced). The vegetation indices SAVI, MSAVI, and EVI produced similar results (R^2 ~ 0.75) using the Bowen-ratio-corrected ET and zero lag. Conversely, correlation of ET with NDVI was lower (R^2 = 0.40), which may be due to the lack of a soil correction factors used in the other vegetation indices (see Table 7). Even though the R^2 is higher for the Grassland site than the Woodland site, the RMSD for the Woodland site is lower than the Grassland site by ~5 mm/month. The R^2 and RMSD for EVI, SAVI, and MSAVI are ~0.87 and ~10 mm/month, respectively, for the Savanna site. When using all data, we noted a slight decline in correlation for the Savanna site, but overall the statistics are comparable.

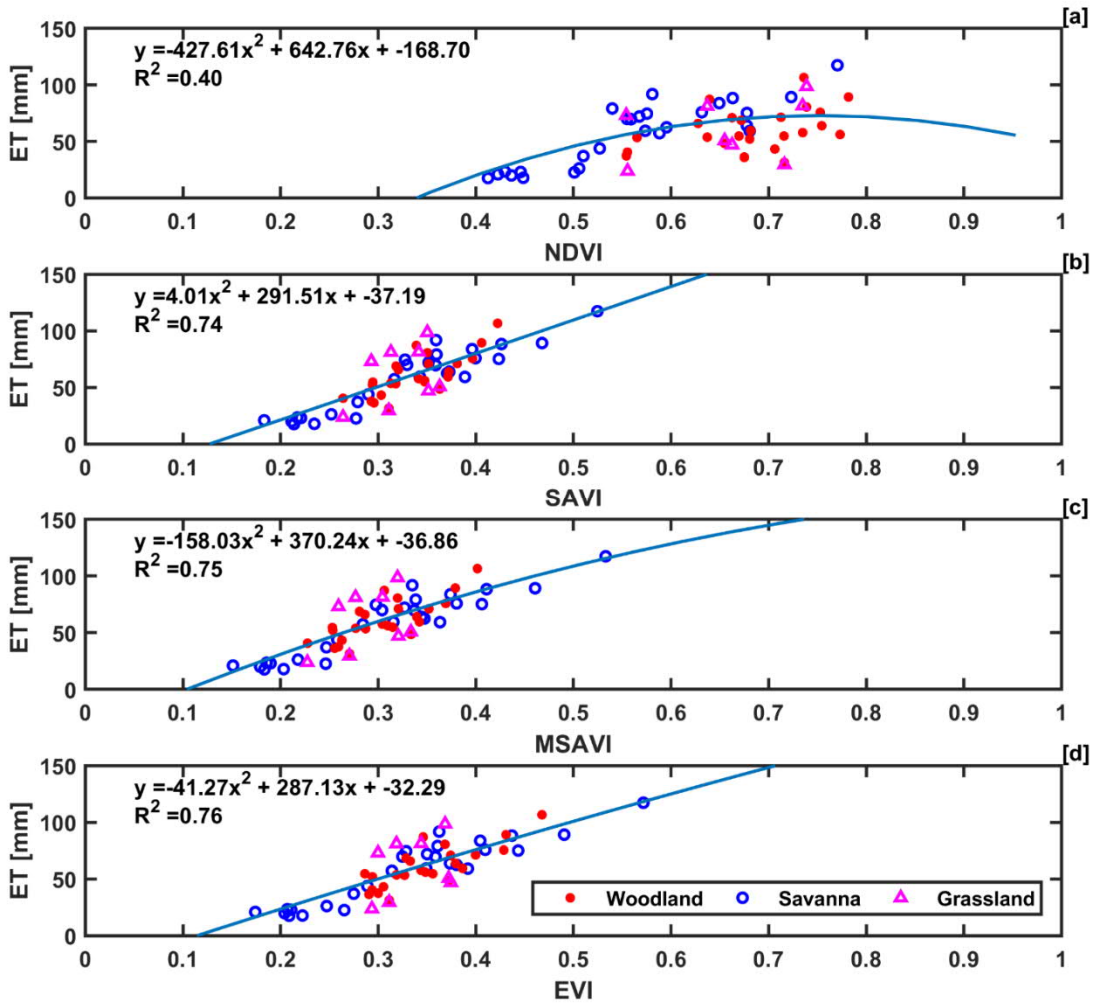


Figure 3-8: Regressions of Bowen-ratio-corrected monthly evapotranspiration (ET) versus the maximum monthly vegetation indices including [a] normalized difference vegetation index (NDVI), [b] soil adjusted vegetation index (SAVI), and [c] modified soil adjusted vegetation index (MSAVI), and [d] enhanced vegetation index (EVI) from Landsat using zero lag.

The regression error (Figure 3-9) was less than 5 mm/month for EVI-based ET estimates (range is between 0.23 and 0.44), but increased when EVI fell below 0.23 and rose above 0.44; we attribute this primarily to the lack of data at the lower and upper bounds, which is not surprising given that all ET measurements were taken in similar meteorological conditions. This EVI range was found to represent 64% of the Cibolo Basin region during any given month from 2000 to 2018, in which the data range represents more than 80% of the region ~90% of the time. The findings for the Medina-Cibolo Basin were worse, with only 55% of the region within the EVI range and only ~80% of the monthly data representing 80% of the region. The difference between the two basins is likely due to higher urbanization in the Medina-Cibolo Basin, leading to a large number of lower EVI values. The average RMSD over the 1000 runs for the bootstrapped (training) data was 10.8 mm/month versus a slight increase to 13.9 mm/month for the testing data. The average R^2 for the training data was 0.77 versus 0.72 for the testing data.

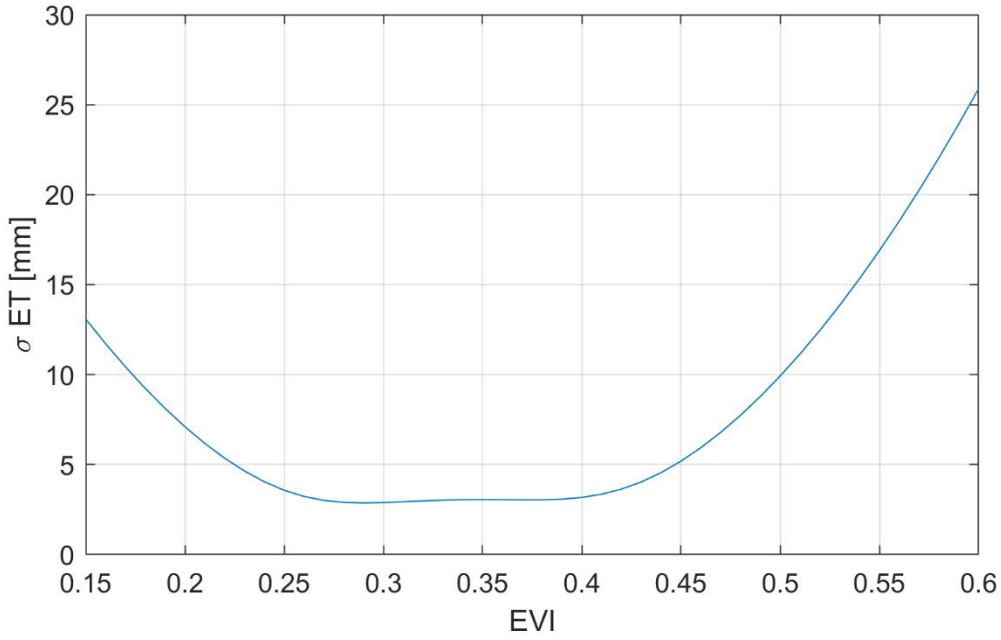


Figure 3-9: Regression error of the Bowen-ratio-corrected evapotranspiration (ET) versus the enhanced vegetation index (EVI) using zero lag.

Comparisons of ET using NLDAS-2 land surface models, MODIS ET, and EVI-ET (Figure 3-10 and Table 8) for Cibolo and Medina-Cibolo Basins are similar. For example, ET estimated with VIC is consistently lower than the EVI-ET with a bias of ~20 mm, but R^2 exceeds 0.70. Mosaic was not highly correlated with the EVI-ET (R^2 <0.50). The lower RMSD for the ensemble of NLDAS models (ENS) for Cibolo Basin was similar for Medina-Cibolo Basin when compared to the other NLDAS models. The highest R^2 and the lowest RMSD were found for MODIS when compared to the ET-EVI. The superior results for MODIS may be due to the similar methodology used to determine ET and the higher spatial resolution when compared to the other land surface products.

The average EVI-ET values for both basins are similar, the largest difference being only 7.6 mm/month (Figure SI6). The average EVI-ET for both basins is ~51 mm/month, with a standard deviation of 14.7 and 13.3 mm/month for Cibolo and Medina-Cibolo Basins, respectively. Results showed a slightly higher average ET during the summer and a slightly lower ET during the winter in Cibolo Basin than in the Medina-Cibolo Basin. This may be due to the larger area of savanna vegetation in the Cibolo Basin than in the Medina-Cibolo Basin (42% versus 23%).

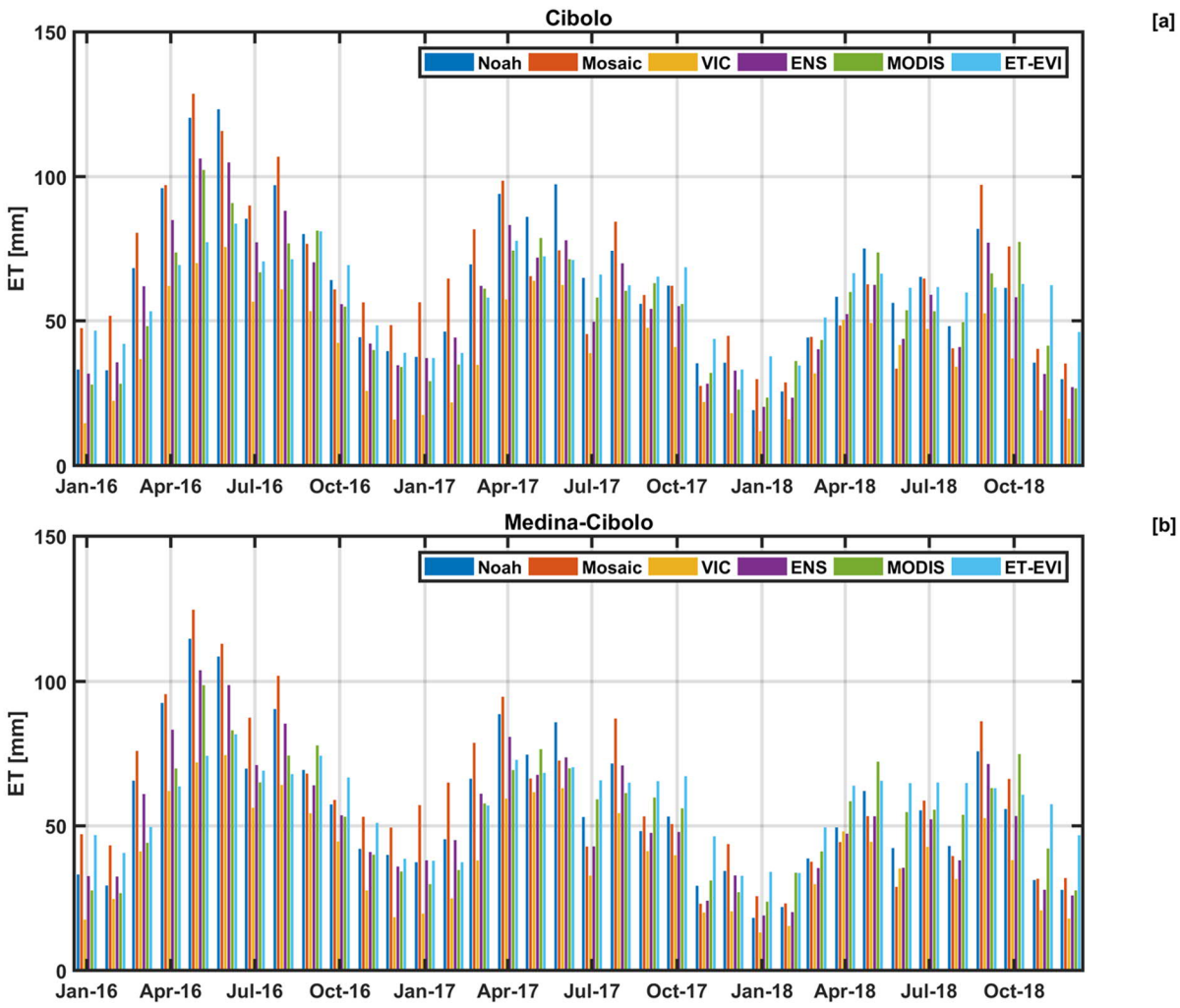


Figure 3-10: Monthly evapotranspiration (ET) comparison over the Cibolo and Medina-Cibolo Basins from January 2016 through December 2018, in which Noah, Mosaic, and VIC use NLDAS-2 forcings and ENS in the ensemble mean of the land surface models, MODIS is from MODIS MOD16A2 8-day composite, and EVI-ET is from the regression using the enhanced vegetation index.

3.6 Numerical Modeling of the Vadose Zone

HYDRUS was parameterized with available datasets for soil hydraulic properties (POLARIS), total soil depth (SSURGO), and root distribution (Noah model). Atmospheric boundary conditions for rainfall came from RainVieux, PET from gridMET, and LAI from MODIS. To preserve transferability throughout the major drainage areas, we limited model calibration and used the three monitoring sites for validation. Thus, the model could be scaled to larger areas and used to validate the satellite-derived monthly ET maps. We present these results first as field validation at these sites, then as watershed validation.

3.6.1 Numerical Simulations at EC Sites

At the site level, HYDRUS was first parameterized using default values and soil properties from the 30-m POLARIS dataset (Table 9). Although this dataset shows considerable variability over the continental U.S.,

the values at each EC site are essentially identical at a restrictive layer depth of 60 cm. POLARIS classified each EC site as loam to a 15-cm depth, silty clay at 15–30 cm, and clay at 30–60 cm. However, field observations found the bedrock at 40 cm for the Savanna site and 6 cm at the Woodland site, so these results are not encouraging. Furthermore, the saturated conductivity (K_s) increases with depth, despite substantial increases in clay content. The POLARIS algorithm begins with the soil map unit, which is a polygon, and adds heterogeneity based on terrain characteristics (e.g., slope, aspect, drainage accumulation). Given the tread-riser geomorphology of the Edwards Aquifer (Woodruff and Wilding, 2008), the POLARIS algorithm likely could not define enough topographic variation to change the within-map unit properties at each EC site. The only variations between sites are the atmospheric forcings (precipitation and PET), root distribution, and LAI.

Despite similar soil properties, the numerical simulations represented soil water content and daily ET fluxes reasonably well (Table 10). The default HYDRUS parameterization level (P1) using a monthly leaf area index (LAI) for each vegetation class over the study period (2016–2018) resulted in an average RMSD for observed soil water content of 0.080 m^3/m^3, a ubRMSD of 0.056 m^3/m^3, an R^2 of 0.60, and an NSE of 0.46. Simulated daily ET fluxes averaged for all sites had an RMSD of 0.123 mm, an R^2 of 0.34, and an NSE of 0.28 (Table 10). For a model to perform better than the mean value, the NSE must be >0. This condition occurred at all sites, although few performed above the acceptable limits used in watershed modeling (NSE >0.5) (Moriasi et al., 2007; Ritter and Muñoz-Carpena, 2013), even though soil moisture was kept within a fixed range and with lower variation than watershed discharge. Thus, we are using a combination of metrics to evaluate these models.

The next parameterization level (P2) used weekly updated MODIS-based LAI that lowered the soil water content RMSD and ubRMSD to 0.069 and 0.048 m^3/m^3, respectively, and the daily ET-flux RMSD to 0.110 mm while improving correlations. Next, we manually adjusted the vegetation water uptake from −800 to −1200 cm to minimize the ET flux RMSD (P3) and added a more realistic karst bottom boundary with a seepage face of −10 cm (P4); however, none of these changes greatly altered model performance, though they did improve the conceptual model. This final parameterization (P4) produced acceptable results for soil water content when evaluated using an RMSD and ubRMSD of 0.072 and 0.052 m^3/m^3 respectively, an R^2 of 0.64, and an NSE of 0.52 (Table 10).

At the individual site level, the P4 model performed satisfactorily for soil water content at all three sites, whereas daily ET fluxes were only satisfactory at the Woodland site (Table 11). These results are shown for the Woodland (Figure 3-11), Savanna (Figure 3-12), and Grassland (Figure 3-13) sites over the 3-year study period. At the Woodland site, with only 6 cm of actual soil, results showed surprisingly good model performance for soil water content at a 5-cm depth, with a daily ET-flux RMSD of 0.9 mm/d (Figure 3-11d) and a ubRMSD of 0.038 m^3/m^3 (Figure 3-11b). Observations of soil water content at the Savanna site indicated more water storage during winter, sustaining soil moisture into summer. The model could not reproduce this trend, leading to a greater magnitude of ET fluxes even during the winter, when plants used all available soil water faster than observations (Figure 3-12f). Like the Savanna site, Grassland site soil water content data also suggest some prolonged storage over winter, though the model simulated a more dynamic range (Figure 3-13) and a daily ET-flux RMSD of 1.2 mm/d. In all cases, the model seems to sufficiently reproduce soil water content (unbiased RMSD of 0.05 m^3/m^3) and ET flux (RMSD of 0.12 mm/d). The cumulative ET fluxes compared well, despite no site-specific calibration.

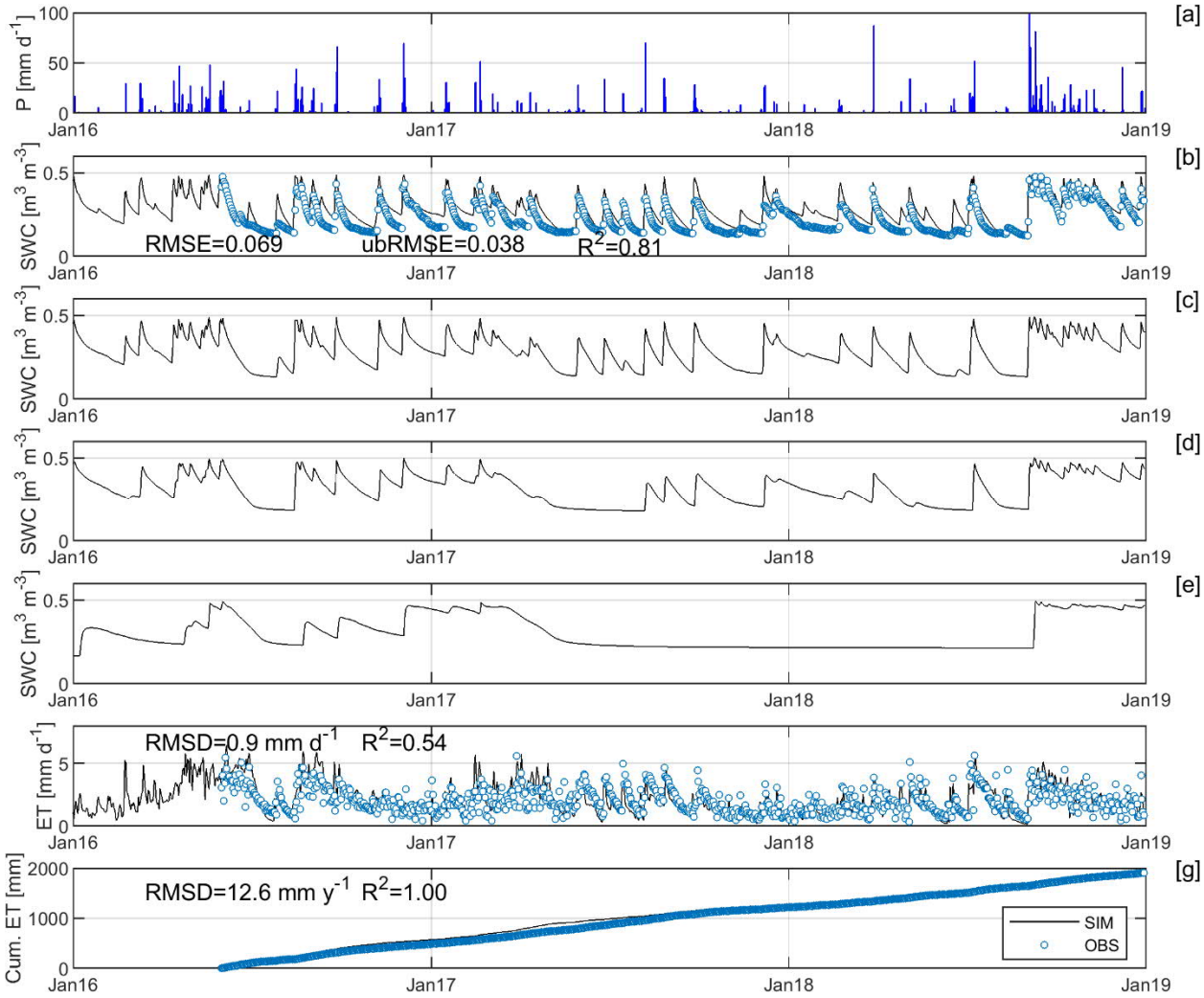


Figure 3-11: Model results from HYDRUS for the Woodland site including soil water content at [b] 5 cm, [c] 10 cm (no sensor), [d] 20 cm (no sensor), [e] 50 cm (no sensor), and [f] daily ET flux and [g] cumulated ET.

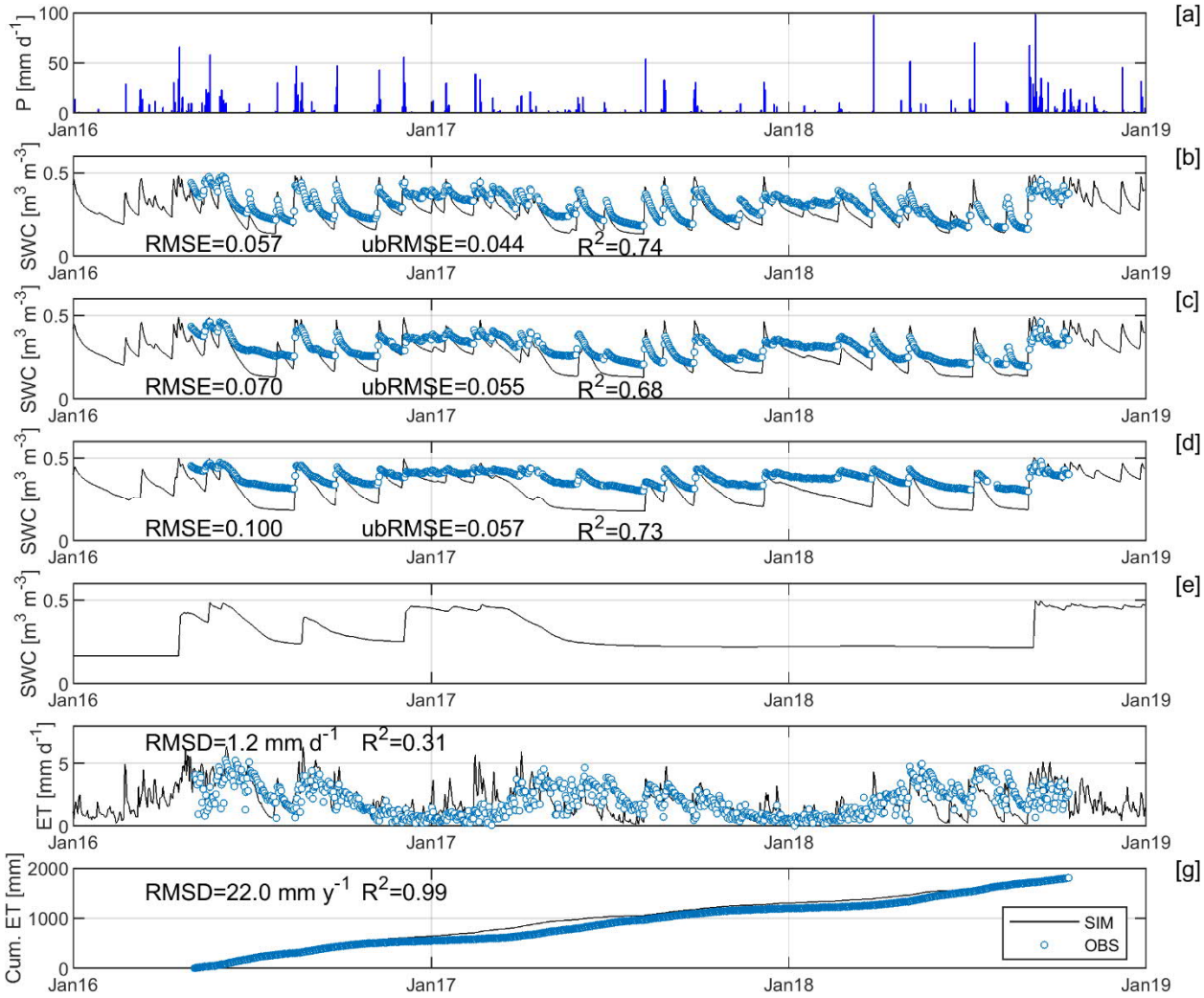


Figure 3-12: Model results from HYDRUS for the Savanna site including soil water content at [b] 5 cm, [c] 10 cm, [d] 20 cm, [e] 50 cm (no sensor), and [f] daily ET flux and [g] cumulated ET.

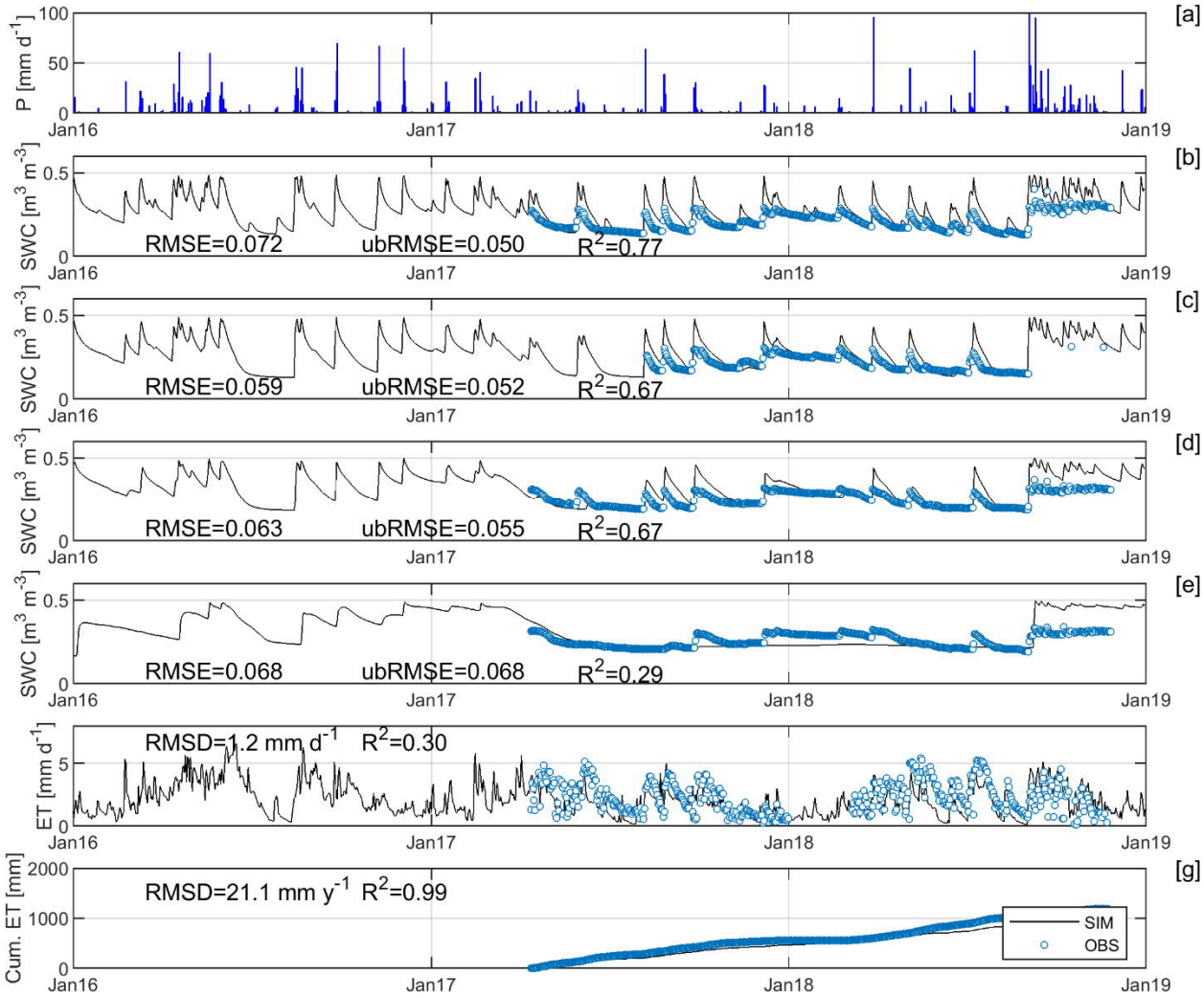


Figure 3-13: Model results from HYDRUS for the Grassland site including soil water content at [b] 5 cm, [c] 10 cm, [d] 20 cm, [e] 50 cm, and [f] daily ET flux and [g] cumulated ET.

3.6.2 Upscaled Numerical Simulations of ET Flux Versus ET-EVI

As indicated above, EVI-ET is the relationship between measured ET from the EC towers, essentially at point scale, and EVI determined from Landsat imagery for the same footprint. This relationship was determined across the relatively short time period when the towers were operational, and is shown in Figure 11d, where ET is the dependent variable and EVI is the independent variable. The HYDRUS model also represents conditions at the EC tower sites. Thus, Table 12 compares the monthly and annual ET fluxes from 2004 to 2018 between the two methods, P1/P4 HYDRUS simulations and remotely sensed EVI-ET at the three EC sites. Results showed that HYDRUS simulated ET responded more dynamically to rainfall events than did EVI-ET (Figure 3-14). For example, during wet periods like fall 2007, the simulated ET peaked at nearly 150 mm/month at all three sites, whereas EVI-ET never surpassed 100 mm/month. Conversely, the 2011 drought stressed vegetation throughout Texas, and ET clearly decreased at both the Savanna (Figure 3-14b) and Grassland (Figure 3-14c) sites. The EVI-ET at the Woodland site, however,

ranged consistently 50–80 mm/month regardless of drought conditions, which suggests that vegetation response to drought at the Woodland site is less sensitive than at the other sites, also leading to a lower correlation (R^2=0.31) between Woodland site simulations and ET-EVI. Annually, the RMSD between EVI-ET and the model ranged from 83.2 mm/year at the Grassland site and 153 mm/year at the Woodland site (Figure 3-15). Annual correlation was also very high for both the Savanna and Grassland sites, but not the Woodland site (Table 12), which is unusual considering the best HYDRUS model performance was observed at the Woodland site.

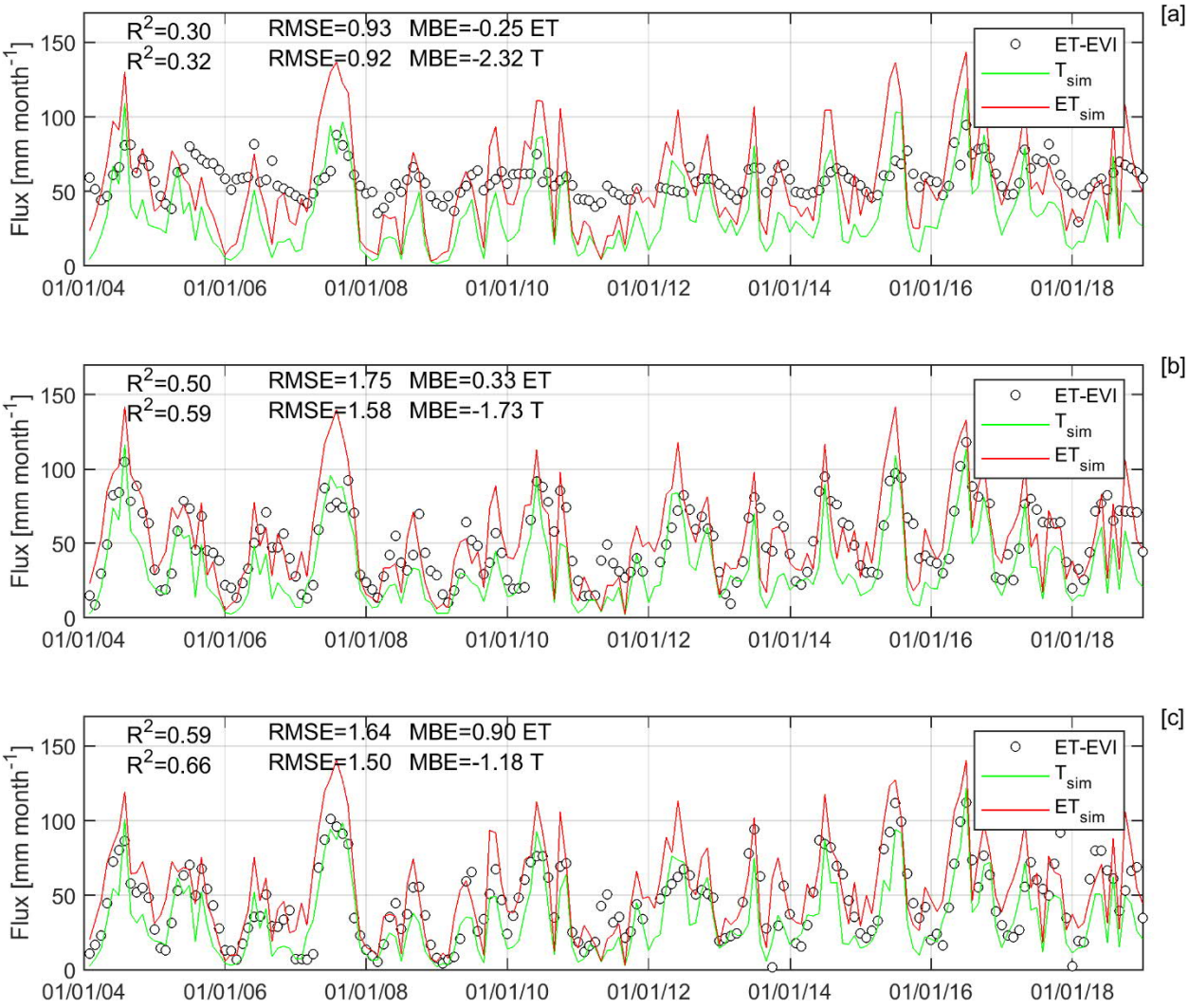


Figure 3-14: Monthly ET and transpiration (T_{sim}) flux comparison between HYDRUS and remotely sensed ET using EVI (ET-EVI) at [a] Woodland, [b] Savanna, and [c] Grassland field sites.

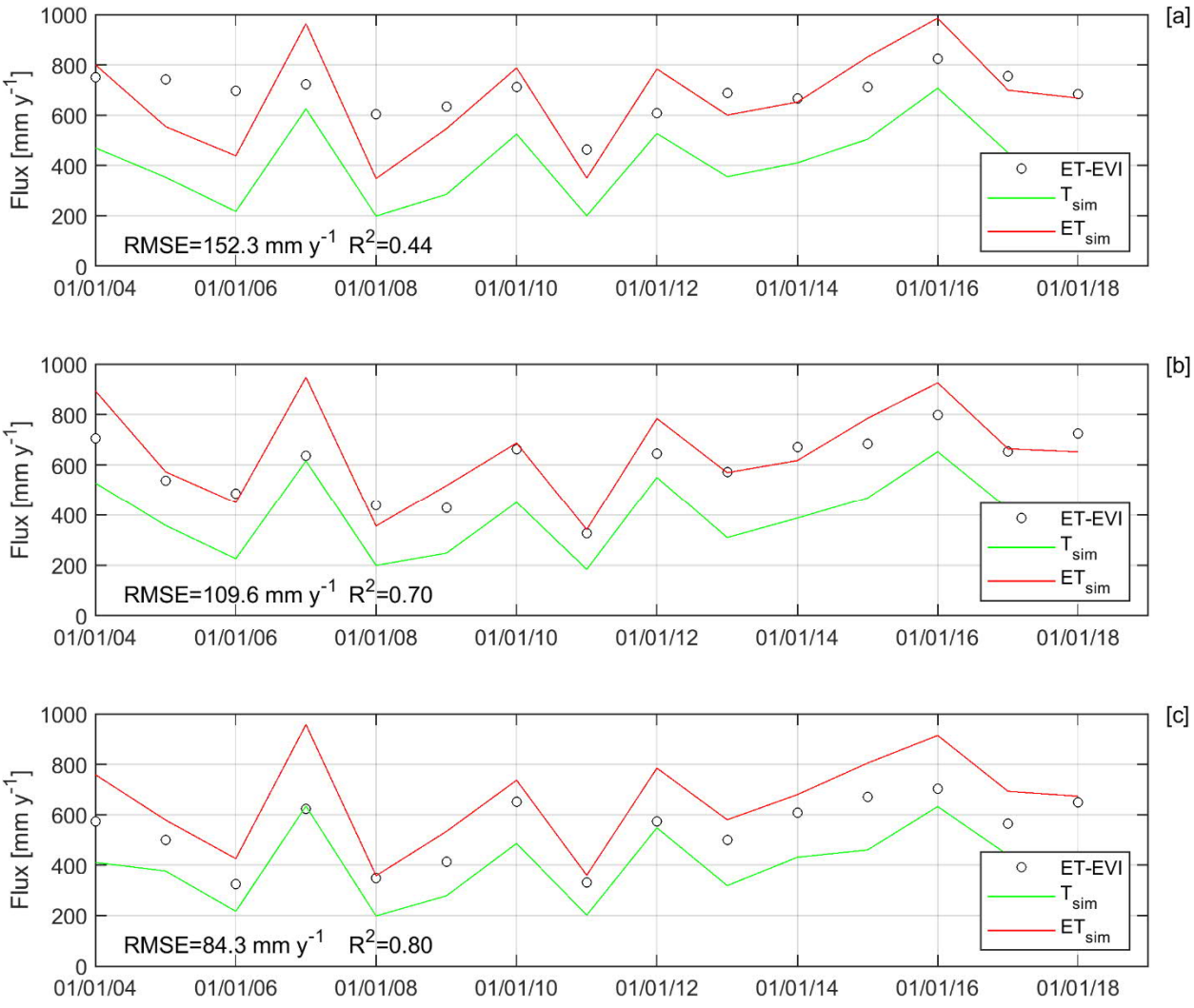


Figure 3-15: Annual ET and Transpiration (T_{sim}) flux comparison between HYDRUS and remotely sensed ET using EVI (ET-EVI) at [a] Woodland, [b] Savanna, and [c] Grassland field sites.

Finally, the P4 parameterization was applied across the entire Cibolo major drainage basin, with each HYDRUS model representing a cell at 500 m resolution (i.e., resulting in 4093 total simulations). The monthly (Figure 3-16) and annual (Figure 3-17) ET fluxes were then totaled and compared to the EVI-ET fluxes determined at the same resolution. Results of these comparisons indicate a mean R^2 of 0.36 (Figure 3-16b), a mean RMSE of 14 mm/month (Figure 3-16d), and an MBE of −2.5 mm/month (Figure 3-16f). The RMSE was lower in the central area of the watershed (where it was calibrated) and increased to the west with increasing aridity. Annually, the mean R^2 was 0.43 (Figure 3-17b) with a mean RMSE of 70 mm/year (Figure 3-17d) and an MBE of −27 mm/year (Figure 3-17f) over Cibolo Basin.

To conclude, both the HYDRUS model and remotely sensed EVI-ET have benefits, drawbacks and uncertainties. For example, we found that the soil hydraulics in the model were poorly represented by POLARIS. In the future, other products such as SoilGrids250m (Hengl et al., 2017), which produces 250 m soil depth, texture, and bulk density maps, could be transformed into more representative hydraulic properties (Schaap et al., 2001; Zhang and Schaap, 2017). Regardless, the correspondence between the model and EVI-ET suggests either or both could be used to constrain future estimates of R_{df}.

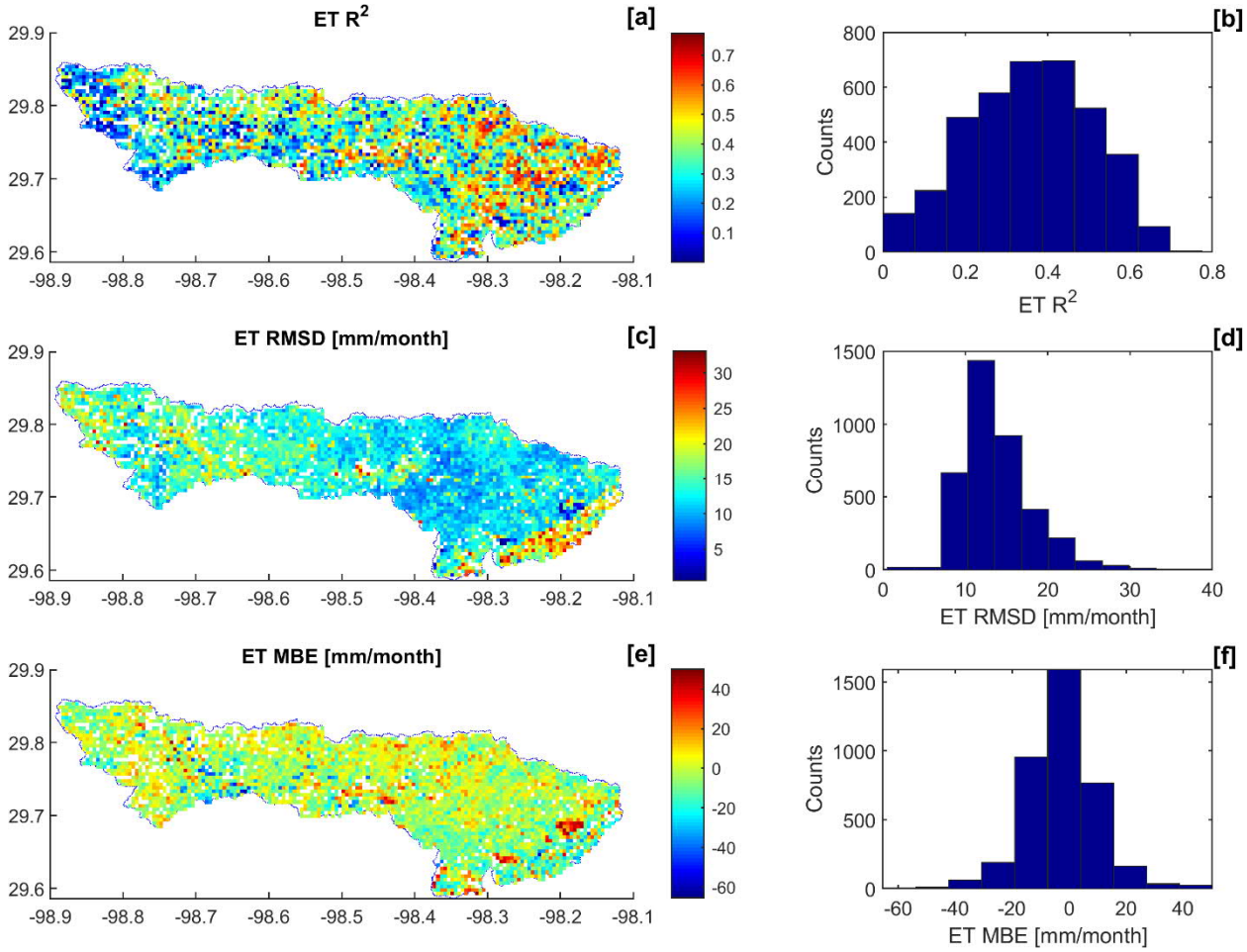


Figure 3-16: Monthly [a and b] correlation (R^2), [c and d] root-mean squared difference (RMSD), and [e and f] mean bias error (MBE) between HYDRUS-ET and EVI-ET from 2004 to 2018 for Cibolo Basin.

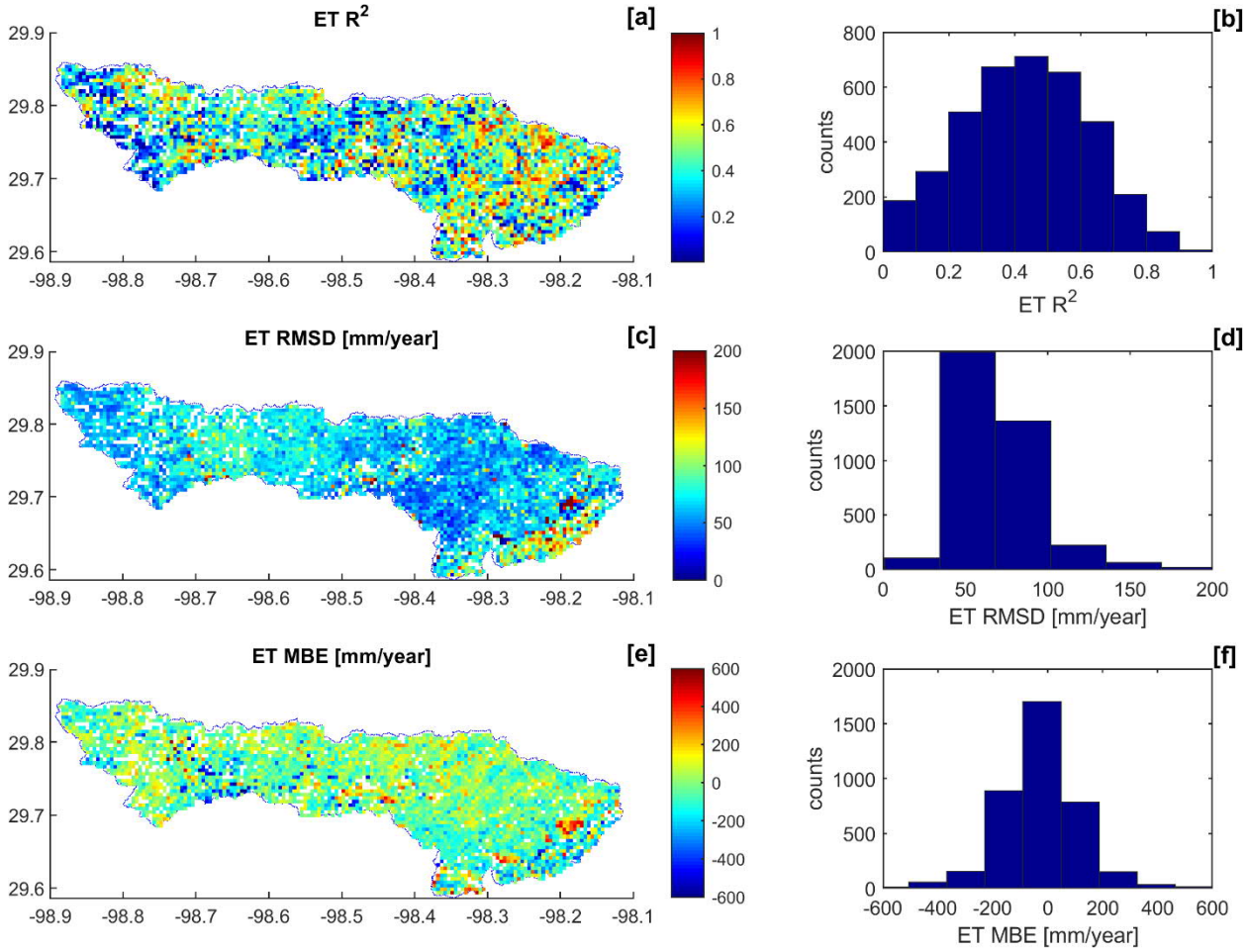


Figure 3-17: Annual [a and b] correlation (R^2), [c and d] root-mean squared difference (RMSD), and [e and f] mean bias error (MBE) between HYDRUS-ET and EVI-ET from 2004 to 2018 for Cibolo Basin.

3.6.3 Diffuse Recharge Dynamics from Numerical Simulations

Revisiting the three EC sites, HYDRUS simulations matched both the observed soil water content and EC-ET fluxes over the study period, and highlight highlights temporal dynamics of R_{df} over the past 10 years (Figure 3-18). For example, the impact of the 2011 drought is evident through 2015: R_{df} is unchanged with time until soil water storage increases to approximately 200 mm during most of 2015, before fall precipitation events lead to nearly 100 m of diffuse recharge in just a few days. Similar precipitation events occurring in fall 2018 also produced nearly 200 mm of R_{df}. Thus, there appears to be a critical threshold in soil water storage that needs to be met before significant R_{df} is possible. Thus, R_{df} cannot be estimated as a constant percentage of precipitation depth, both of which are expressed in units of length. As an illustration, the volume of R_{df} from Cibolo Basin ranges from 87,000 acre-ft (8% of precipitation) in 2004 to essentially 0 acre-ft in 2008, when nearly 300,00 acre-ft of precipitation was estimated (Figure 3-19c). The estimate of diffuse recharge from HYDRUS between 2004 and 2018 is approximately 32% of the total (discrete and diffuse) recharge estimated from the USGS method (Puente, 1978) for the Cibolo Basin (Figure 3-19b). Accuracy of the diffuse recharge estimate from HYDRUS can be refined by measuring baseflow in Cibolo Creek, which is assumed to be equal to diffuse recharge in the contributing zone, but the result showing that about 32% of total recharge is through diffuse processes does appear reasonable.

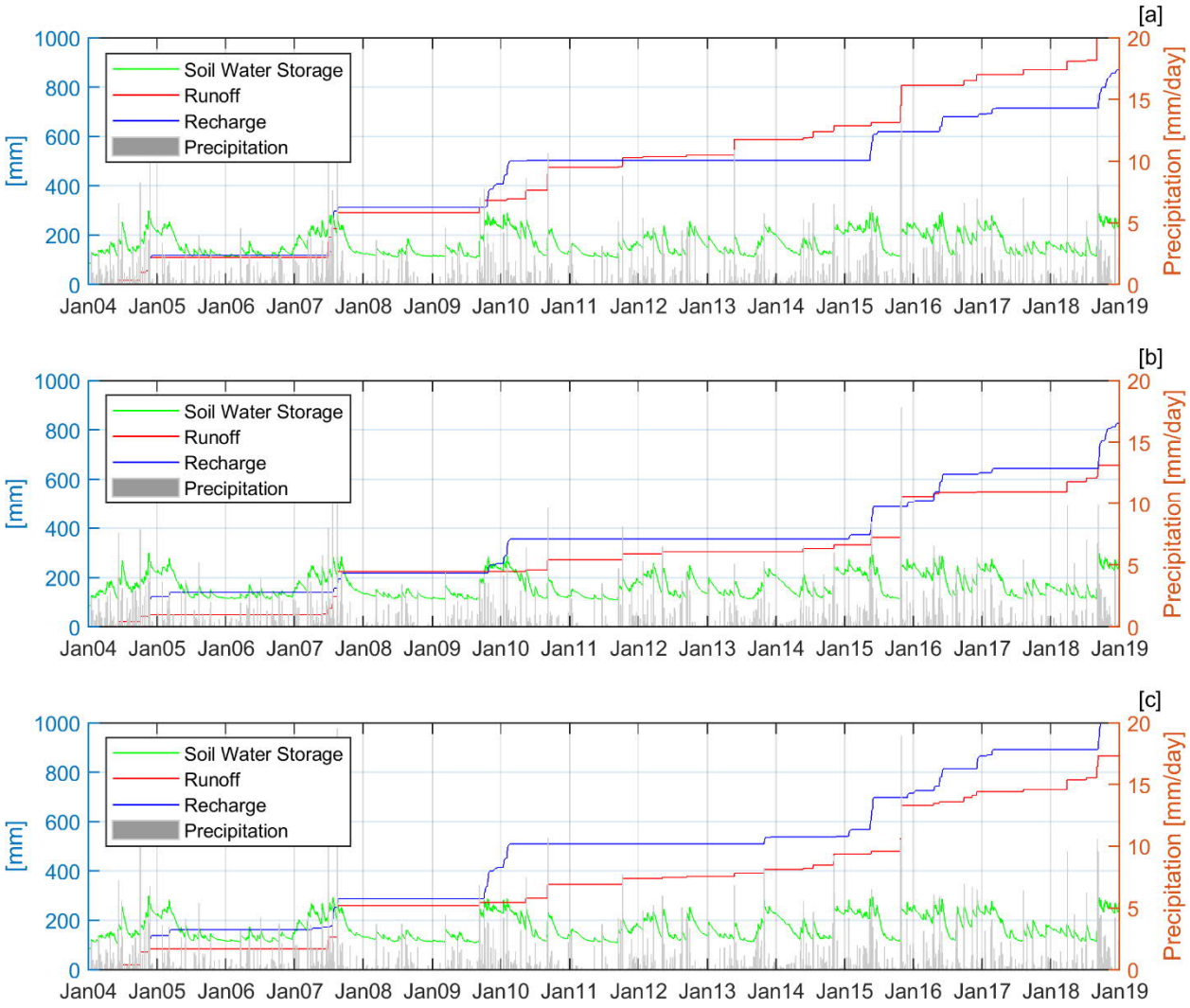


Figure 3-18: HYDRUS simulations of cumulative fluxes of runoff and diffuse recharge (Recharge), and precipitation rate and soil water storage at the [a] Woodland, [b] Savanna, and [c] Grassland eddy covariance sites.

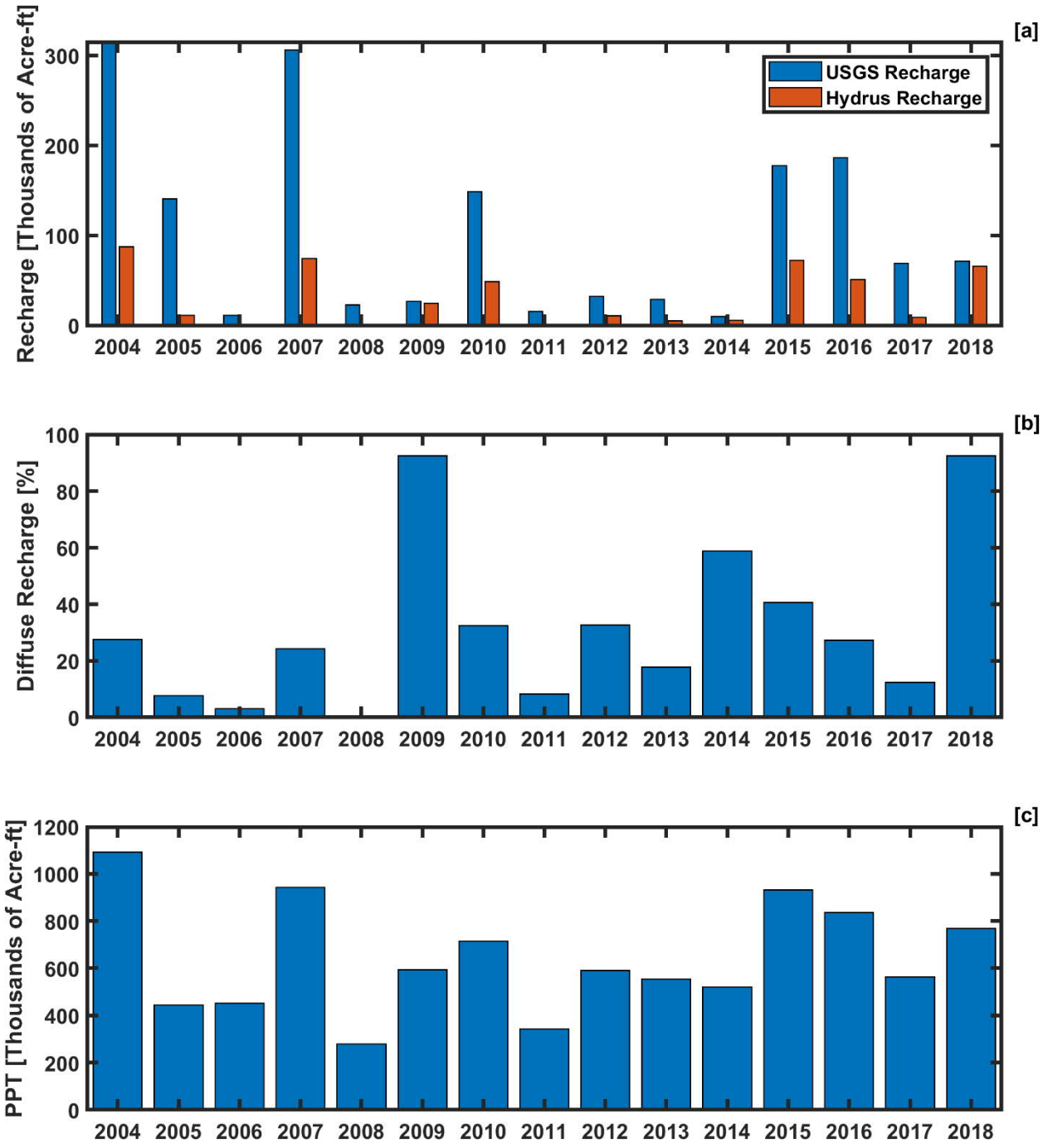


Figure 3-19: Cibolo Basin [a] annual total recharge from the USGS and R_{df} from HYDRUS, [b] annual percentage of R_{df} of total USGS recharge, and [c] precipitation volume from RainVieux.

3.6.4 Comparison of Field Measured ET (EC-ET), EVI-ET and HYDRUS-ET

We compared measured ET from the eddy covariance method as a validation point for both the EVI route and numerical modeling route. In both comparisons—EC-ET versus EVI-ET, and EC-ET versus HYDRUS-ET—we regressed (predicted) EVI-ET and HYDRUS-ET onto the (observed) EC-ET, and determined correlation,

standard error, and the p-level indicating significance. All tests were conducted at a significance level of 0.05.

We have conducted statistical analyses of the data (SigmaStat, v13) starting with (1) analysis of monthly observed ET fluxes versus those estimated using EVI, then (2) we compared monthly EC-ET and ET-EVI, (3) compared monthly EC-ET and HYDRUS-ET (for each of three sites), and finally (4) we compared daily EC-ET and HYDRUS-ET for the Savanna site, to understand whether high resolutions measurements and estimates are comparable. All results are found in the Appendix, Section 7.

(1) Monthly EC-ET fluxes as a function of EVI measurements – Results show that the relationship is significant to $P<0.001$ ($R^2 = 0.758$, StdErr = 11.92 mm/month)
(2) Monthly EC-ET fluxes versus ET-EVI – Results show the relationship is significant to $P<0.001$ ($R^2 = 0.739$, StdErr = 10.64 mm/month). In this case, the regression equation (slope = 0.739) indicates that the EVI-estimated ET is underestimated by approx. 26%
(3) Monthly EC-ET fluxes versus monthly HYDRUS-ET – Results of linear regression showed the following: Woodland ($R^2=0.805$, StdErr=13.4 mm/month; $P<0.001$); Savanna ($R^2 = 0.276$; StdErr = 24.4 mm/month; $P = 0.005$); Grassland ($R^2 = 0.631$; StdErr = 14.4 mm/month; $P = 0.019$). In the case of Woodland growth forms, HYDRUS-ET overestimated EC-ET by 46%; in the cases of Savanna and Grasslands, HYDRUS-ET underestimated EC-ET by 42% and 36%, respectively. Figure 3-20 shows the results of these regression calculations.
(4) Daily EC-ET fluxes versus Daily HYDRUS-ET (Savanna site) – relationship is significant to $P<0.001$ ($R^2 = 0.307$). We also noted that the regression equation (slope = 0.605) indicates that the HYDRUS-ET is underestimating EC-ET by around 40%, pointing to additional sources of available water below the modeling domain.

In general, the workflow, from measuring ET using ground-based measurement techniques, and then upscaling those techniques using a vegetative index (e.g., EVI in this case) or comparing the results to a numerical model shows significant promise. Model results for each of the three growth forms, measured during the field project, were well simulated and indicated significance when compared to field measurements. Clearly, the underestimation with HYDRUS is a concern but extending the model domain downward, into the epikarst or karst material could improve the results further.

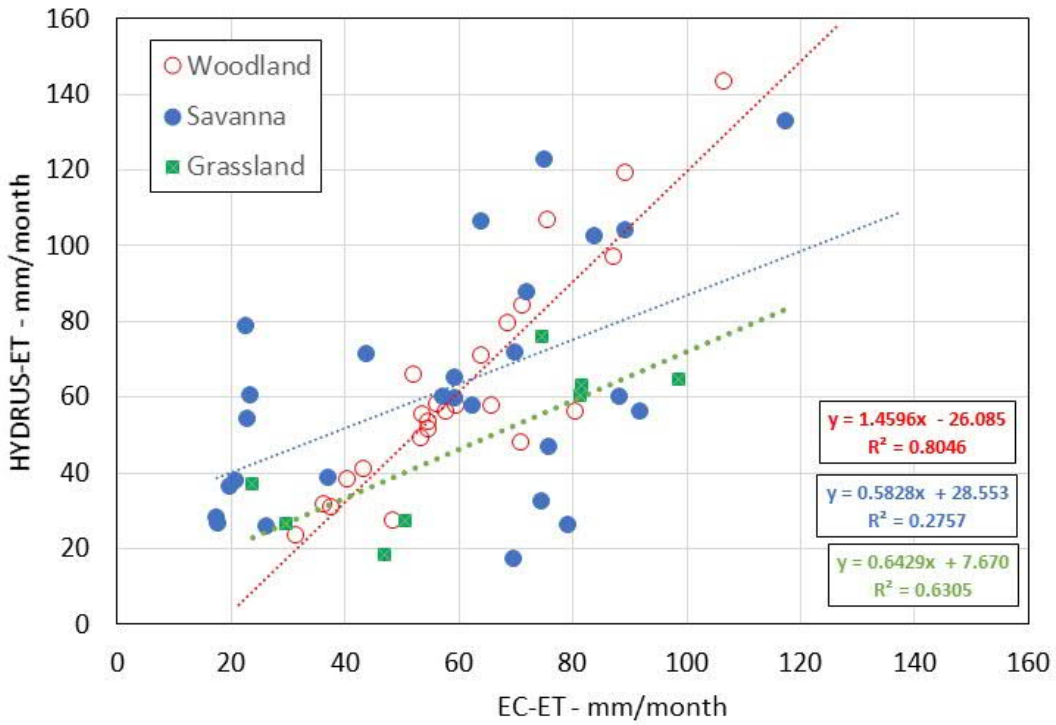


Figure 3-20: Results of linear regression between monthly (cumulative) observed EC-ET and predicted HYDRUS-ET results.

3.7 Recharge Modeling

3.7.1 Soil Moisture

Figure SI7 shows the results of the water flux using the soil moisture method compared to the results from the EC method. The water flux using the soil moisture method at the Woodland site was more similar to the results from the EC method than the Savanna site with an R^2 of 0.66 and an RMSD of 55.61 mm/month (Table 13). The bias was also lower at the Woodland site (7.28 mm/month) than the Savanna site (13.54 mm/month). The comparison using the meteorological station data indicates a more negative bias at Salado, Headquarters Cave, and High Hill when using soil moisture to estimate water flux (Table 14).

3.7.2 Water Balance

From 2016 to 2018, high outflows were recorded at Cibolo Creek in spring 2016 and September 2018 (Figure SI8). Discharge was also recorded at Salado Creek in spring 2016 and September 2018, but flows were higher in winter 2017. From 2000 to 2018, discharge at Cibolo Creek in Selma, TX was above 202,000 acre-ft for 6 out of 19 years, with the highest discharge volume recorded at 154,000 acre-ft in 2003 (mostly occurring during July). The second highest discharge took place in 2007, but over a period extending from March through August.

Precipitation dominates the monthly water-balance components (Figure 3-21 and Figure 3-22). For 60% of the time, precipitation is more than 40,500 acre-ft for the Cibolo Basin and 24,300 acre-ft for the Medina-Cibolo Basin. For Cibolo Basin, ET is consistent over time and ranges from an average of <32,400

acre-ft per month during December through February to >52,700 acre-ft during May through July. For the Medina-Cibolo Basin, ET is slightly lower, with an average monthly ET of <24,300 acre-ft during December through February to >38,500 acre-ft per month during May through July. Yearly ET/PPT was estimated to range between 0.45 for a very wet year and 1.35 for a dry year. Values greater than 1.0 indicate that groundwater or surface water could be contributing water to ET. A caveat of this method is that, during the study period (2016–2018), the ETo 0.9 for the region and did not include years when the yearly ET exceeded precipitation. Also, runoff could have contributed to flow in Cibolo Creek, in addition to base flow. For example, discharge at the 0818500 gage (Cibolo Creek at Selma) was at least 20% of precipitation for a total 4 months during the study period, which means that either runoff occurred, or that precipitation events led to pulses of diffuse recharge, as seen in the HYDRUS simulations and described in Section 3.6.3. Though changes in soil water storage (dS) never exceed ET in a given month, there are times when dS/ET exceeds 40%. The highest average recharge volumes (>24,300 acre-ft) were estimated to occur during May, September, and October, with the lowest recharge (<0 acre-ft) occurring during August.

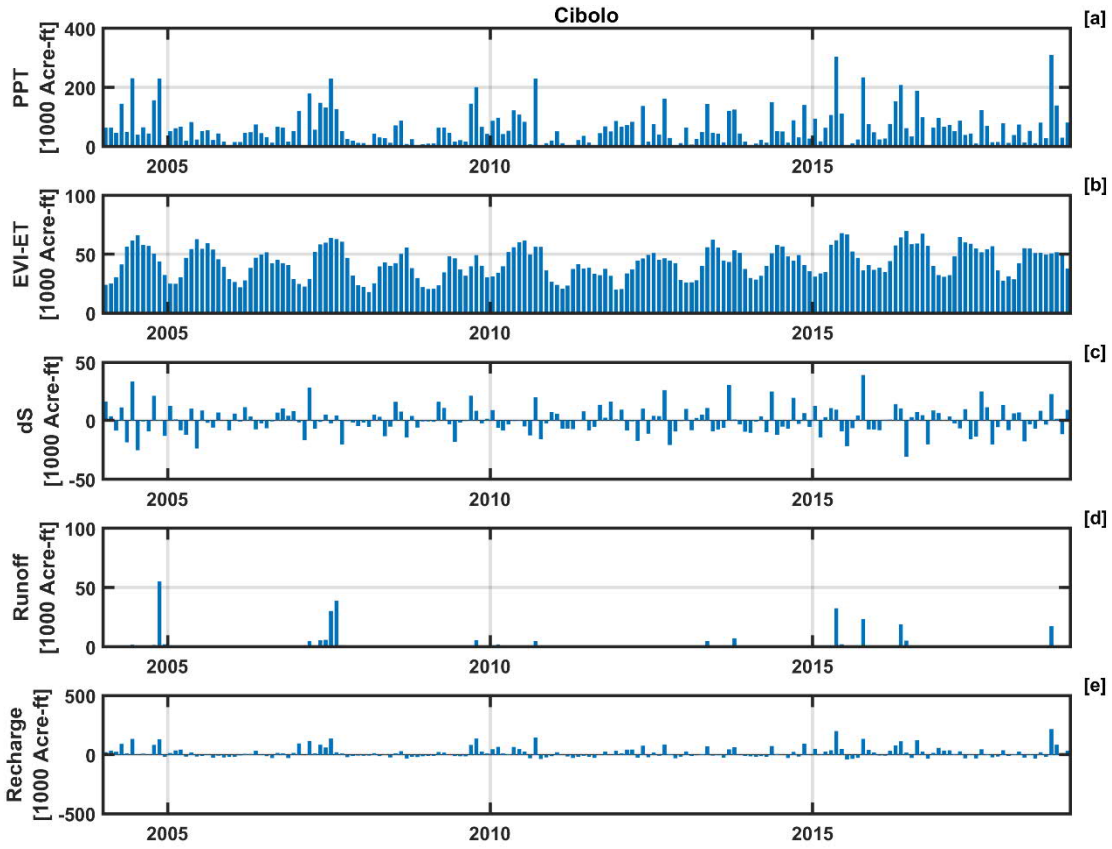


Figure 3-21: Monthly components of the water balance for the Cibolo Basin, which include [a] precipitation from gridMET (PPT), [b] evapotranspiration (ET) from the EVI, [c] soil water storage (SWS) from NLDAS-2 Noah, [d] runoff from USGS gage 08185000 (Cibolo Creek at Selma, TX), and [e] recharge.

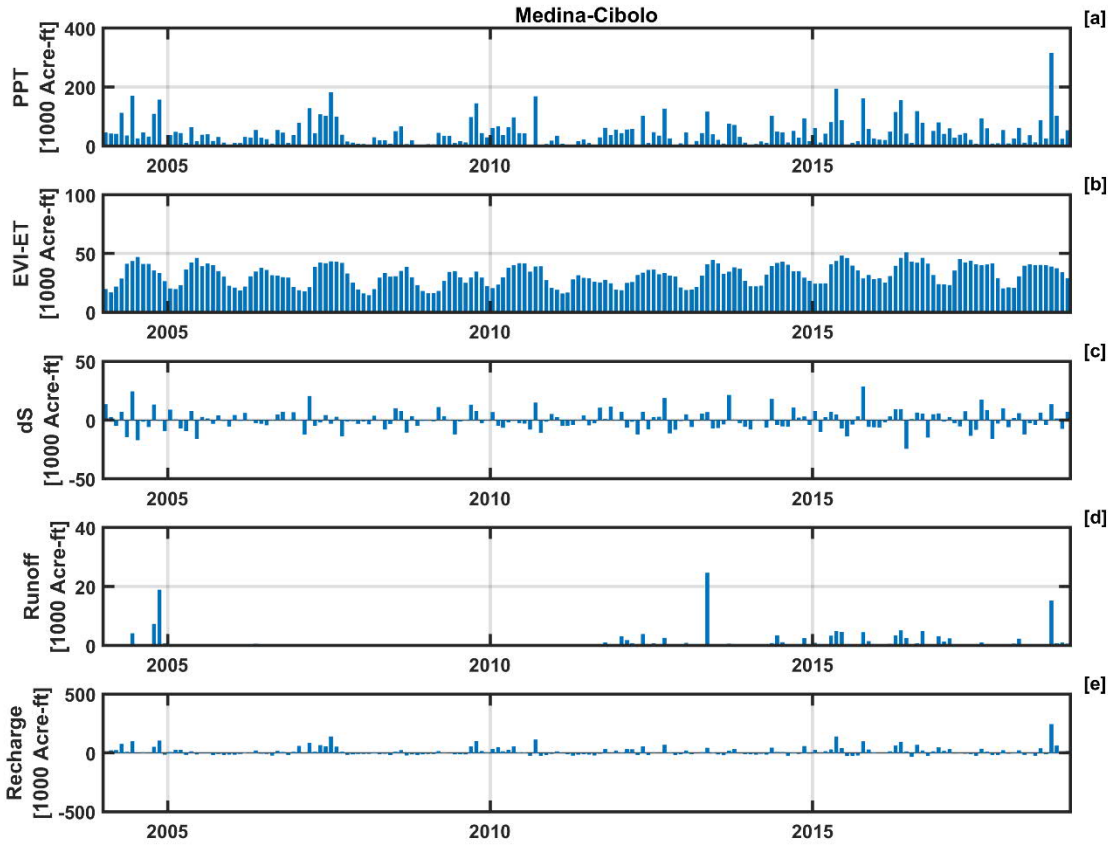


Figure 3-22: Monthly components of the water balance for the Medina-Cibolo Basin, which include [a] precipitation from gridMET (PPT), [b] evapotranspiration (ET) from the EVI, [c] soil water storage (SWS) from NLDAS-2 Noah, [d] runoff from USGS gage 08178700 (Salado Creek at Loop 410, San Antonio, TX), and [e] recharge.

4. Discussion and Conclusions

4.1 Precipitation

Recharge is limited by the amount of precipitation in the area, which dominates the monthly water balance components. It is crucial that precipitation estimates are accurate. Products corrected using rain gage data have an advantage when rain-gage networks are dense, especially when rainfall is scattered or locally heavy as is common for Texas. Partnerships with schools to host precipitation gages would help bolster the number of stations in an area while also engaging the next generation of decision-makers.

4.2 Potential Evapotranspiration

Results show that, of the weather station variables used to calculate PET, wind speed and maximum relative humidity correlate the lowest to the gridMET products, whereas temperature and minimum

relative humidity are very highly correlated. We found that gridMET PET was more comparable to the meteorological stations than NLDAS, which estimated an average of 3 times more PET on a monthly basis.

4.3 Soil Moisture

Soil water storage is a source of uncertainty in the water budget but can be extremely relevant. Monthly storage increases or decreases were found to be more than 40% of ET during some months. Of the NLDAS products used (Noah, Mosaic, and VIC), the change in soil water storage can range from ±50 mm/month to ±16 mm/month, assuming the soil depth is 40 cm. We recognize that the soil depth in our analyses is a limiting factor, if the roots extend below the soil layer and into the epikarst, utilizing additional source of water. Land surface models are still unable to handle the types of heterogeneity observed in epikarst. When comparing these products to conditions at soil water monitoring stations, even though the correlation is acceptable and the bias is low, the RMSD for monthly soil water storage can be more than 20 mm/month. Properly characterizing the soil profile at weather stations and installing soil moisture sensors in shallow and deep soils would provide insight into this component of the water balance.

4.4 Evapotranspiration

Evapotranspiration accounts for the most water loss from the hydrologic system, and thus accurate measurements are extremely important. We found an EVI with zero lag was the most suitable option tested to relate satellite measurements to monthly measured ET. The regression equation to extend the EC station measurements of ET to the regional scale was more accurate between EVIs of 0.23 and 0.44. Siting future EC stations in regions that can extend this range could improve area-wide ET estimates. We found that overall, ET rates were very similar across the three different landcovers studied, with the Savanna site transpiring more during the summer months and less in the winter compared to the Woodland site. To extend this methodology to other basins, installing EC stations across a broader climatological zone would be ideal. Measurements taken during extreme events (either droughts or very wet conditions) are also beneficial. Continuous monitoring would help capture these extreme periods. Another improvement to the ET estimates could be to consider increasing the time resolution of satellite remote-sensing imagery. For example, Planet Labs is a private vendor that provides commercial multi-band imagery at very high spatial (less than 4 m) and temporal (~ daily) resolution.

4.5 Recharge

The most recent Groundwater Availability Model (GAM) of the area (Jones et al., 2011) provides a list of 10 previous recharge studies (their Table 5-1). These studies lump direct and diffuse recharge into a single number, so we cannot directly compare them with our diffuse recharge (R_{df}) values; yet; relative comparisons are useful. For example, the range of total recharge as a percent of PPT, across all studies is between 1.5% - 11%. Our estimates of 11.2% and 13.5% (ET-EVI) for the Cibolo and Medina-Cibolo Basins are higher, but closer to the more recent study by Ockerman (2007). We note that the methods, field areas and measurement periods all differ, so direct comparisons are difficult to make. Negative monthly recharge values using the water balance method may indicate that plants are extracting water from the epikarst layers below the soil, but the changes in recharge from month to month or year to year are likely influenced by weather events or conditions that may have occurred months or years earlier. As we indicate above, the amount of yearly diffuse recharge ranged from 87,000 acre-ft in 2004, which accounted for 8% of the total precipitation, to essentially 0 acre-ft in 2008. Memory from the previous

month or year likely explains this large range, but our measured dataset was too short, forcing us to rely on estimates of precipitation products. Future studies on the impact of how month-over-month or year-over-year conditions might impact recharge would be valuable.

The HYDRUS model, as any other model, does have some limitations. First, the model only simulates vertical (1-D) flow and therefore active groundwater flow or lateral flow between 1-D domains are not considered. Second, the datasets used to create the model array were originally collected at different resolutions, which could perpetuate error when up- and down-scaling. For example, the 30 m soil data, which was scaled up to 500 m resolution using the nearest neighbor method would lose information during this transfer. Likewise, the 4 km rainfall and potential evapotranspiration data were scaled down to 500 m resolution, also using the nearest neighbor method. These original data products already include some averaging of spatial information, and using the nearest neighbor approach does not bring back the spatial information. In an ideal scenario, an accurate high-resolution data set would be used for all input, but this creates additional computational complexity.

Partitioning the diffuse recharge into specific aquifers would require a significantly more sophisticated below-ground study than what was done here. The emphasis in this study, and as described above, was on the amount of water percolating below root zone. Figure 4-1 shows the percentages of total diffuse recharge projected to occur above the Contributing (Figure 4-1a) and Recharge Zones (Figure 4-1b). This was done by summing the HYDRUS-modeled recharge for each 500 m cell, identifying the zone represented by the cell and expressing that sum as a percent of total R_{df} for each basin. The remaining amount of R_{df} occurring in the Artesian Zone is 9.3% and 29.3% of the total R_{df}. The results show temporal differences between basins and with time, and illustrate that the percentages of area overlying the three zones do not translate directly into the percent of diffuse recharge, just given different precipitation and vegetation across the EAA region. Additional research would be needed to understand to where the diffuse recharge flows after percolating below the root zone.

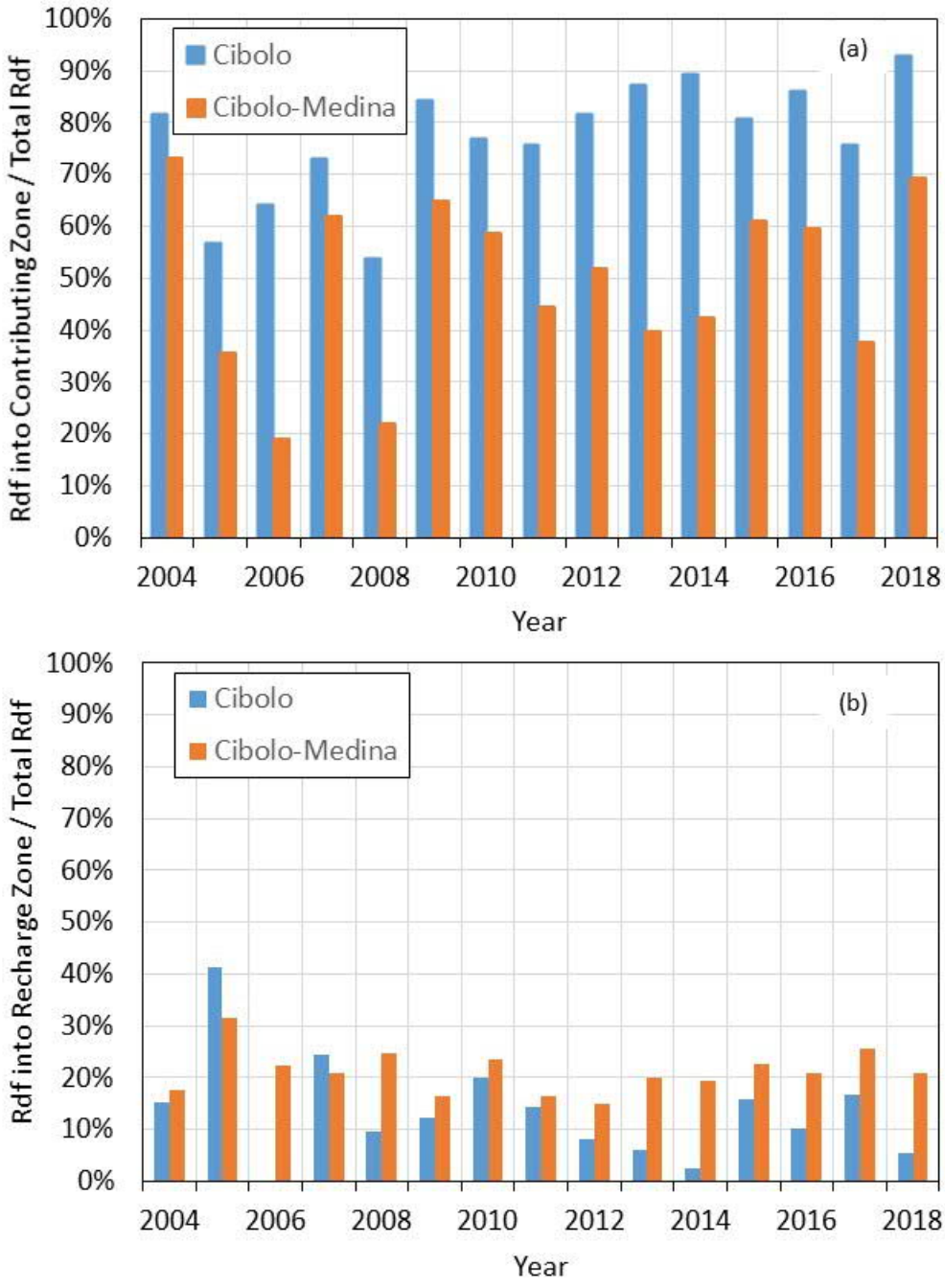


Figure 4-1: Percentage of Total Rdf occurring above the Contributing Zone (a) and the Recharge Zone (b), using estimates of diffuse recharge determined using HYDRUS.

5. Recommendations

5.1 Maintain/expand onsite weather and soil monitoring across EAA Region

The field measurements conducted for EAA focused on processes only at Camp Bullis, though the results were upscaled more regionally using satellite remote sensing and other meteorological products that populated the entire Cibolo Basin (also known as Basin #7). The research highlighted several insights that could be applied to other basins in the EAA Region:

- The ET estimates between the three Eddy Covariance towers installed and operated at Camp Bullis were comparable (~2.1 mm/day when averaged across the entire monitoring period), leading to the general conclusions that diffuse recharge was also similar. One likely explanation is that the weather conditions (e.g., incoming solar radiation, temperature and precipitation) across Camp Bullis were similar. To better understand how recharge changes across the entire EAA region, especially in the drier, western regions, and to better parameterize other basins (e.g., Nueces-West Nueces River Basin), we recommend redeploying existing EC stations in different geographic areas of EAA's Region and maintaining the stations for at least a full year, preferably longer, to observe interannual variability.

- The five weather stations installed by EAA at Camp Bullis, as well as other precipitation gages operated within the EAA Region, are used as correction points for RainVieux, improving precipitation estimates. Densifying the monitoring network by installing additional stations would further reduce uncertainty in precipitation estimates. New station locations could be chosen either through a grid pattern or other statistical techniques, like mean relative differencing (Vachaud et al., 1985).

5.2 Automate blending of in-situ data and satellite remote sensing (or land surface models)

Research findings show that diffuse recharge can be a significant percentage of the groundwater recharge component of the water balance, but that variability in time and space complicates projections. It is unclear how well the intensive data collection at one area, Camp Bullis, can be extrapolated to other areas, because of different land use, soil properties, etc. Thus, we recommend developing and implementing an approach that will produce a soil moisture map of high spatial and temporal resolution. To accomplish this recommendation, the approach would blend both satellite remote sensing or land surface models, and in-situ data, to produce high-resolution soil moisture maps across large EAA areas. The approach would use results from the Camp Bullis project, plus research conducted elsewhere in Texas, in which techniques are being developed to project soil moisture status up to 72 hours into the future. The ability to automate this would allow for better irrigation scheduling, improved management of dam operations and water releases, and more effective flood risk mitigation activities that would protect lives and property. The process automates the delivery of soil moisture data from the National Land Data Assimilation System (NLDAS), and uses the relationships between vegetation indices and ET developed in this work, but expanded across the EAA service area. With the combined monitoring program and machine learning approaches (described below), we can produce better soil moisture maps that would be available elsewhere, and make them usable to EAA scientists.

5.3 Apply multivariate regression or machine learning algorithms

Future effort should consider developing a data analytics workflow that blends high-resolution satellite remote sensing data, outputs of land surface models, and in-situ data, to enable generation of high-resolution soil moisture maps across large EAA areas. Results from the Camp Bullis project showed that knowledge of soil water storage is critically important to diffuse recharge, but significant uncertainties are present in the current storage estimates, due to uncertainties in soil hydraulic properties, soil horizon/types and depths, and soil water content observations. Furthermore, we recognize that our inability to measure water distribution in, and physical properties of, the epikarst introduces uncertainties as well. Results from the Camp Bullis project also suggest that many existing data products (e.g., precipitation) may manifest different degrees of uncertainty. In addition, process-based approaches alone (e.g., HYDRUS) may encounter equifinality issues, in which different parameter values yield the same output, because of the parameter uncertainty and model conceptualization errors. Recent data analytics approaches, especially newer methods using machine learning, enable us to exploit linkages in multisource data to develop more-constrained regression models. For example, SoilGrids250, a new dataset generated through use of data and information from 150,000 observed soil profiles worldwide and Quantitative Random Forest and Gradient Boosting methods (Hengl et al., 2017) reports significant improvement over other methods. When combined with physical insights and process-based models, the regression models, like the EVI-ET relationship developed here, become more robust. We recommend adapting the gridded soil moisture data from the National Land Data Assimilation System (NLDAS) for the EAA service area, and automating the delivery of that gridded product to EAA scientists involved with estimating diffuse recharge to aquifer systems. NLDAS includes a quality-controlled soil moisture dataset compiled by NASA from modeling and re-analyses. It is often used to support various modeling activities, but also can provide useful information on soil moisture status so that local stakeholders (from hydrologists to irrigators) can estimate hydrologic response and/or make decisions on land use, such as irrigation scheduling. The NLDAS land surface models used to create the soil moisture maps are provided on grids of ~8.5 mile (14 km, 0.125 deg) resolution. Thus, the NLDAS gridded soil moisture maps need to be downscaled and bias-corrected before they can be used for diffuse recharge estimates.

5.4 Assessing land use change on diffuse recharge

Results of this study indicate that approximately 10-12% of total precipitation is recharged through diffuse and discrete routes, and that approximately 30% of total recharge is through diffuse routes (~4% precipitation). The percentages here are averaged across both time and space. The heterogeneous nature of the geology and weather patterns across the EAA Boundary, and the memory that the near-surface soil and geological systems carries through year-over-year add significant uncertainty to any estimate. One change that is almost certain to occur is that population will continue to grow across the Contributing and Recharge Zones of the Edwards Aquifer. What is less certain is whether and to what extent the associated land-use change will alter the diffuse recharge because of impervious surfaces or discrete recharge because of surface water routing toward or away from river reaches that are known to recharge aquifers. We thus recommend extending the approach used in this study, in which we overlay population projections, estimates of roads, physical structures, etc., and other diversion features onto the diffuse recharge tomograms obtained from numerical models (e.g., HYDRUS). The two outcomes, combined, would show where development might interrupt diffuse recharge or enhance or mitigate discrete recharge. The results could then be used for long-term future projections of recharge, given different

climate scenarios (baseline, wet periods, dry periods, etc.). Depending on the needs for EAA, the projections could be updated every 5 years and synced with the TWDB state water plans, so that water demand and potential water supply are co-projected.

5.5 Integrating dynamic recharge maps into groundwater models

Results of this study reveal significant spatial and temporal variability of diffuse recharge in EAA regions, resulting from variabilities in precipitation, ET, land surface, soil properties, and epikarst properties. It is recommended that a coupled surface water/groundwater modeling approach be taken in future studies to integrate recharge estimates with groundwater availability modeling for EAA regions. Coupled surface water/groundwater modeling has been increasingly used at the global and regional scales to assess groundwater sustainability under current and future climates (Wada and Bierkens, 2014). The rich set of historical and real-time monitoring data collected by EAA can be combined with high-resolution hydrometeorological gridded maps mentioned in the above sections to enable these types of studies.

6. Data Availability

The final dataset was posted to the Texas Data Repository Dataverse (Bongiovanni, 2020). This includes the raw 30-minute flux data and metadata for the Woodland, Savanna, and Grassland sites, using the Biological, Ancillary, Disturbance and Metadata (BADM) protocol, which is standardized across AmeriFlux, Fluxnet and other networks (Law et al., 2008). The BADM site metadata includes general site information, sensor type, and sensor calibration information. The 30-minute data includes heat flux, gas, footprint, and meteorological data.

7. Tables

Table 1: Root mean squared difference (RMSD) and correlation (R^2) for monthly rainfall between the gridded precipitation products and the weather stations.

	RMSD (mm/month)				R^2			
Site Name	PRISM	NLDAS	RainVieux	gridMET	PRISM	NLDAS	RainVieux	gridMET
Salado	28.6	79.0	11.6	19.9	0.97	0.40	0.99	0.98
Well 10	17.4	84.3	23.0	18.5	0.99	0.39	0.97	0.98
High Hill	27.9	72.5	18.7	19.6	0.97	0.48	0.97	0.97
Acan	18.7	103	25.7	11.2	0.98	0.46	0.98	0.99
Headquarters Cave	15.1	105	17.3	13.4	0.98	0.40	0.99	0.99
Woodland	20.4	85.5	22.4	17.5	0.97	0.42	0.97	0.97
Savanna	29.1	68.7	24.1	23.8	0.93	0.50	0.96	0.95
Grassland	26.6	73.7	13.6	31.3	0.96	0.60	0.98	0.97
Mean	23.0	84.0	19.6	19.4	0.97	0.46	0.98	0.98

Table 2: Correlation coefficient (R^2) and the slope of regression equation for the daily meteorological variables between the PET stations and the gridded datasets of air temperature (Temp), incoming solar radiation (Solar), wind speed, and relative humidity (RH).

Product	Stat	Station	Temp max (C)	Temp min (C)	Temp mean (C)	Solar (W/m^2)	Windspeed (m/s)	RH max (%)	RH min (%)	RH mean (%)
gridMET	R^2	Salado	0.93	0.88	0.93	0.69	0.57	0.25	0.75	0.65
		Well10	0.93	0.88	0.93	0.70	0.72	0.27	0.76	0.65
		Highhill	0.93	0.90	0.94	0.68	0.71	0.34	0.76	0.67
		Acan	0.94	0.88	0.94	0.64	0.48	0.24	0.74	0.62
		Hqcave	0.94	0.90	0.94	0.72	0.23	0.33	0.74	0.67
	slope	Salado	0.94	1.02	0.98	0.98	0.48	0.20	1.00	0.66
		Well10	0.94	1.00	0.97	1.00	0.94	0.22	1.04	0.69
		Highhill	0.96	0.89	0.93	1.00	0.72	0.47	1.02	0.82
		Acan	1.00	1.02	1.00	0.93	0.33	0.22	1.02	0.68
		Hqcave	1.00	0.87	0.95	1.00	0.27	0.60	1.04	0.94
NLDAS	R^2	Salado	0.91	0.85	0.94	0.70	0.57	0.15	0.70	0.69
		Well10	0.91	0.86	0.95	0.70	0.70	0.20	0.71	0.70
		Highhill	0.91	0.92	0.95	0.68	0.71	0.58	0.70	0.74
		Acan	0.91	0.84	0.94	0.64	0.48	0.12	0.73	0.69
		Hqcave	0.92	0.93	0.96	0.72	0.25	0.66	0.73	0.76
	slope	Salado	0.90	1.13	1.01	0.98	0.48	0.16	1.00	0.70
		Well10	0.90	1.11	1.00	0.99	0.96	0.20	1.04	0.74
		Highhill	0.92	1.00	0.96	1.00	0.72	0.62	1.01	0.89
		Acan	0.96	1.10	1.02	0.94	0.31	0.15	1.01	0.70
		Hqcave	0.96	0.98	0.97	1.01	0.26	0.82	1.04	0.98

Table 3: Root mean squared difference (RMSD), correlation coefficient (R^2), mean bias error (MBE) for monthly and daily potential evapotranspiration (PET) between the gridded products and the EAA PET stations.

	gridMET PET (mm/month)			NLDAS (mm/month)		
Station	*RMSD*	R^2	*MBE*	*RMSD*	R^2	*MBE*
Salado	37.79	0.92	−32.97	112.10	0.89	−102.29
Well10	31.69	0.96	−28.46	105.08	0.92	−95.97
Highhill	29.36	0.95	−24.57	101.80	0.91	−91.63
Acan	41.05	0.93	−35.16	109.73	0.90	−97.73
Hqcave	46.84	0.93	−41.48	117.94	0.91	−105.93
	gridMET PET (mm/day)			NLDAS (mm/day)		
Station	*RMSD*	R^2	*MBE*	*RMSD*	R^2	*MBE*
Salado	1.53	0.73	−1.09	3.87	0.66	−3.36
Well10	1.35	0.78	−0.91	3.57	0.69	−3.10
Highhill	1.35	0.74	−0.82	3.58	0.67	−3.07
Acan	1.66	0.69	−1.15	3.85	0.62	−3.26
Hqcave	1.85	0.70	−1.41	4.14	0.65	−3.57

Table 4: Root mean squared difference (RMSD), correlation coefficient (R^2), and mean bias error (MBE) for monthly soil water storage between the gridded soil moisture products and the PET and Eddy covariance sites.

0–10 cm monthly storage (mm/month)										
		Salado	Well10	Highhill	Acan	Hqcave	Woodland	Savanna	Grassland	Average
RMSD	Noah	4.75	4.50	3.08	8.02	4.35	5.16	4.24	2.74	4.60
	Mosaic	4.99	5.65	4.34	8.45	4.53	4.59	4.55	3.94	5.13
	VIC	4.61	3.98	3.71	7.55	6.51	9.13	7.39	4.57	5.93
MBE	Noah	−0.33	−0.19	−0.10	−0.37	−0.47	0.25	−0.32	−0.26	−0.22
	Mosaic	−0.44	−0.18	0.04	−0.36	−0.67	0.03	−0.41	−0.32	−0.29
	VIC	−0.18	−0.16	−0.38	−0.34	−0.62	0.56	−0.15	0.24	−0.13
R^2	Noah	0.75	0.78	0.90	0.45	0.81	0.91	0.87	0.92	0.80
	Mosaic	0.88	0.77	0.90	0.48	0.83	0.91	0.86	0.91	0.82
	VIC	0.17	0.50	0.63	0.36	0.48	0.64	0.52	0.66	0.50
0–40 cm monthly storage (mm/month)										
		Salado	Well10	Highhill	Acan	Hqcave	Woodland	Savanna	Grassland	Average
RMSD	Noah	6.30	12.10	12.14	29.46	24.32	NaN	14.12	9.80	15.46
	Mosaic	11.28	14.92	10.20	29.08	17.99	NaN	8.57	9.56	14.51
	VIC	12.43	12.98	12.36	30.59	24.95	NaN	16.66	11.59	17.36
MBE	Noah	−0.24	−0.44	−0.65	−1.18	−3.19	NaN	−0.42	−0.66	−0.97
	Mosaic	−1.12	−0.47	−0.02	−1.20	−2.82	NaN	−0.48	−2.33	−1.21
	VIC	−0.54	−0.44	−1.25	−1.18	−2.65	NaN	0.01	−0.51	−0.94
R^2	Noah	0.92	0.77	0.82	0.43	0.66	NaN	0.94	0.86	0.77
	Mosaic	0.90	0.77	0.89	0.50	0.80	NaN	0.94	0.91	0.82
	VIC	0.47	0.74	0.87	0.33	0.64	NaN	0.86	0.85	0.68

Table 5: Daily evapotranspiration (ET) comparison using the correlation (R^2), mean, and standard deviation (Stdev) between the Woodland and Savanna site.

		Mean ET (mm/day)	Stdev ET (mm/day)	Mean ET (mm/day)	Stdev ET (mm/day)
	R^2	Woodland		Savanna	
Mar–May	0.29	2.03	0.85	2.25	1.14
June–Aug	0.36	2.36	1.17	2.79	0.98
Sept–Nov	0.72	2.10	0.93	1.87	0.89
Dec–Feb	0.77	1.48	0.72	0.68	0.31
Total	0.44	2.04	1.01	2.00	1.18

Table 6: Correlation (R^2) of the vegetation indices (normalized difference vegetation index [NDVI], soil adjusted vegetation index [SAVI], modified soil adjusted vegetation index [MSAVI], and enhanced vegetation index [EVI]) versus the Bowen-corrected and uncorrected evapotranspiration using lags of –1 month, no lag, and 1 month.

	Uncorrected Bowen ratio			Bowen correction		
Lag	−1	0	1	−1	0	1
NDVI	0.17	0.35	0.30	0.18	0.40	0.32
SAVI	0.30	0.71	0.47	0.32	0.74	1.46
MSAVI	0.31	0.73	0.49	0.32	0.75	0.47
EVI	0.29	0.72	0.46	0.32	0.76	0.44

Table 7: Correlation (R^2) and root mean squared difference (RMSD) of the vegetation indices (normalized difference vegetation index [NDVI], soil adjusted vegetation index [SAVI], modified soil adjusted vegetation index [MSAVI], and enhanced vegetation index [EVI]) versus the Bowen-corrected evapotranspiration using the equation found using all the data for the individual sites.

	R^2				RMSD [mm/month]			
	Woodland	Savanna	Grassland	All	Woodland	Savanna	Grassland	All
NDVI	0.14	0.76	0.50	0.40	18.67	16.41	23.20	18.38
SAVI	0.58	0.87	0.68	0.74	11.82	10.20	17.97	12.17
MSAVI	0.59	0.88	0.69	0.75	11.62	10.10	17.39	11.94
EVI	0.63	0.87	0.71	0.76	11.23	10.12	16.93	11.71

Table 8: Root mean squared difference (RMSD), correlation (R^2) and mean bias error (MBE) of monthly evapotranspiration comparison over the Cibolo and Medina-Cibolo Basins from January 2016 through December 2018, in which Noah, Mosaic and VIC use NLDAS-2 forcings, ENS in the ensemble mean of the land surface models, and MOD16 from MODIS MOD16A2 8-day composites.

	Cibolo					Medina-Cibolo				
	Noah	Mosaic	VIC	ENS	MOD16	Noah	Mosaic	VIC	ENS	MOD16
R^2 (–)	0.74	0.45	0.83	0.70	0.83	0.61	0.37	0.76	0.58	0.81
RMSD (mm)	16.14	19.98	20.56	13.31	10.10	15.80	20.63	19.99	15.36	9.87
MBE (mm)	3.83	6.12	−19.06	−3.04	−3.70	−1.29	3.11	−17.96	−5.38	−3.91

Table 9: Soil texture and hydraulic properties extracted from the POLARIS (Chaney et al., 2019) dataset for each EC sites.

	Layer	Sand	Silt	Clay	Bulk Density	θ_r	θ_s	α	n	K_s
Woodland	(cm)	(%)	(%)	(%)	(g/cm^3)	(m^3/m^3)	(m^3/m^3)	(1/cm)	(–)	(cm/h)
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.636
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.582
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.639
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.827
Savanna										
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.724
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.669
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.735
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.915
Grassland										
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.709
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.653
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.706
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.899

θ_s = saturated water content; θ_r = residual water content; α = inverse air-entry pressure; K_s = saturated conductivity

Table 10: Vadose-zone model statistics for root mean square difference (RMSD), unbiased RMSD (ubRMSD), correlation (R^2) and Nash-Sutcliffe model efficiency (NSE) for soil water content (SWC) and ET fluxes using four model parameterizations: (P1) default parameterization with monthly LAI climatology, (P2) default parameterization with weekly LAI updated from Modis, (P3) P2 with some manipulation to vegetation parameters, and (P4) the final model using P3 with a seepage-face bottom boundary.

		Site 1: Woodland		Site 2: Savanna		Site 3: Grassland		Average	
Model Parameterization	Units	SWC (m^3/m^3)	ET (mm/d)	SWC (m^3/m^3)	ET (mm/d)	SWC (m^3/m^3)	ET (mm/d)	SWC (m^3/m^3)	ET (mm/d)
P1: default parameterization, monthly LAI	RMSD	0.084	1.11	0.089	1.37	0.066	1.22	0.080	1.23
	ubRMSD	0.044	1.10	0.062	1.36	0.063	1.13	0.056	1.20
	R^2	0.78	0.41	0.54	0.23	0.47	0.38	0.60	0.34
	NSE	0.49	0.40	0.36	0.12	0.52	0.32	0.46	0.28
P2: default parameterization, MODIS LAI	RMSD	0.069	0.83	0.075	1.21	0.064	1.25	0.069	1.10
	ubRMSD	0.037	0.83	0.054	1.21	0.054	1.18	0.048	1.08
	R^2	0.81	0.55	0.61	0.27	0.56	0.28	0.66	0.37
	NSE	0.54	0.55	0.45	0.12	0.54	0.14	0.51	0.27
P3: P2, optimized veg parameters	RMSD	0.067	0.87	0.083	1.23	0.065	1.23	0.071	1.11
	ubRMSD	0.038	0.86	0.058	1.23	0.060	1.16	0.052	1.09
	R^2	0.81	0.54	0.59	0.28	0.50	0.30	0.64	0.38
	NSE	0.59	0.54	0.42	0.16	0.54	0.19	0.51	0.30
P4: P3, seepage face	RMSD	0.067	0.88	0.082	1.21	0.066	1.20	0.072	1.10
	ubRMSD	0.038	0.87	0.058	1.21	0.061	1.15	0.052	1.08
	R^2	0.81	0.54	0.60	0.30	0.51	0.32	0.64	0.39
	NSE	0.59	0.54	0.43	0.20	0.54	0.22	0.52	0.32

* model did not converge, h_a set to −0.1 cm

Table 11: Vadose-zone model results (P4) at each EC site compared against observed soil water content and daily ET fluxes for root mean square difference (RMSD), unbiased RMSD (ubRMSD), correlation (R^2) and Nash-Sutcliffe model efficiency (NSE).

Site	Vegetation	Obs.	Units	RMSD	ubRMSD	R^2	NSE
1	Woodland	5/all	(m^3/m^3)	0.069	0.038	0.81	0.56
		ET	(mm/d)	0.860	0.850	0.54	0.53
2	Savanna	5	(m^3/m^3)	0.057	0.044	0.74	0.63
		10	(m^3/m^3)	0.070	0.054	0.68	0.53
		20	(m^3/m^3)	0.100	0.057	0.73	0.31
		All	(m^3/m^3)	0.078	0.056	0.61	0.45
		ET	(mm/d)	1.170	1.170	0.31	0.18
3	Grassland	5	(m^3/m^3)	0.072	0.050	0.77	0.51
		10	(m^3/m^3)	0.059	0.052	0.67	0.78
		20	(m^3/m^3)	0.063	0.055	0.67	0.51
		50	(m^3/m^3)	0.068	0.068	0.29	0.29
		All	(m^3/m^3)	0.066	0.059	0.53	0.53
		ET	(mm/d)	1.210	1.160	0.30	0.17

Table 12: Monthly fluxes of ET and transpiration (T) comparison between satellite-based EVI-ET and vadose-zone model at each site using (P1) default parameters in HYDRUS and a climatology of monthly LAI and (P4) the vegetation optimized with a seepage face bottom boundary.

		P1: Default parameters, LAI climatology, free drainage								
		ET flux monthly (mm)			T flux monthly (mm)			ET flux annual (mm)		
Site	Vegetation	R^2	RMSD	MBE	R^2	RMSD	MBE	R^2	RMSD	MBE
1	Woodland	0.25	9.6	−0.4	0.24	9.6	−2.8	0.41	175	−2
2	Savanna	0.36	19.7	5.4	0.42	18.9	−7.9	0.67	125	−67
3	Grassland	0.46	18.8	11.5	0.51	17.9	−1.9	0.78	100	−135
	mean	0.36	16.1	5.5	0.39	15.5	−4.2	0.62	134	−68
		P4: P3/Seepage face								
		ET flux monthly (mm)			T flux monthly (mm)			ET flux annual (mm)		
Site	Vegetation	R^2	RMSD	MBE	R^2	RMSD	MBE	R^2	RMSD	MBE
1	Woodland	0.31	0.92	−0.18	0.33	0.91	−2.25	0.46	15.3	1.3
2	Savanna	0.52	1.72	0.41	0.60	1.56	−1.65	0.70	11.1	−5.2
3	Grassland	0.60	1.61	0.98	0.66	1.49	−1.10	0.81	8.3	−11.6
	mean	0.48	1.41	0.40	0.53	1.32	−1.67	0.66	11.6	−5.2

MBE = mean bias error between satellite and model at monthly and annual totals

Table 13: Correlation (R^2), root mean squared difference (RMSD), and mean bias error (MBE) between the monthly water flux using an analytical solution with soil moisture (SM) as the input versus field precipitation (PPT)–eddy covariance evapotranspiration (ET) for the Savanna and Woodland site.

	Water flux (SM) vs PPT-ET		
	R^2	RMSD (mm/month)	MBE (mm/month)
Woodland	0.66	55.61	7.28
Savanna	0.59	62.63	13.54

Table 14: Correlation (R^2), root mean squared difference (RMSD), and mean bias error (MBE) between the monthly water flux using an analytical solution with soil moisture (SM) as the input versus precipitation (PPT)–evapotranspiration from EVI (ET-EVI) for the sites operated by the Edwards Aquifer Authority.

	Water flux (SM) vs PPT-ET-EVI		
Site	R^2	RMSD (mm/month)	MBE (mm/month)
Salado	0.69	52.42	−3.95
Well 10	0.45	92.39	0.63
Highhill	0.67	60.40	−18.01
Acan	0.59	70.24	9.99
HQcave	0.77	55.81	−1.95

8. References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol., 33(1): 121-131. doi: 10.1002/joc.3413

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidlines for computing crop water requirements, Irrigation and Drainage Paper No. 56. FAO, Rome, Italy, 300 pp.

Blanken, P. et al., 1997. Energy balance and canopy conductance of a boreal aspen forest: partitioning overstory and understory components. Journal of Geophysical Research: Atmospheres, 102(D24): 28915-28927.

Bongiovanni, T., 2020. Edwards Aquifer Authority Camp Bullis Eddy Covariance Data. Texas Data Repository Dataverse. DOI:10.18738/T8/OXB5NY

Celia, M.A., Bouloutas, E.T., Zarba, R.L., 1990. A general mass-conservative numerical solution for the unsaturated flow equation. Water Resour. Res., 26(7): 1483-1496.

Chaney, N.W. et al., 2019. POLARIS Soil Properties: 30-m Probabilistic Maps of Soil Properties Over the Contiguous United States. Water Resour. Res., 55(4): 2916-2938. DOI:doi: 10.1029/2018WR022797

Chen, F. et al., 1996. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. Journal of Geophysical Research: Atmospheres, 101(D3): 7251-7268.

Daly, C. et al., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int. J. Climatol., 28(15): 2031-2064. doi: 10.1002/Joc.1688

Dugas, W.A., Hicks, R.A., Wright, P., 1998. Effect of removal of Juniperus ashei on evapotranspiration and runoff in the Seco Creek watershed. Water Resour. Res., 34(6): 1499-1506. doi:10.1029/98wr00556

Ek, M. et al., 2003. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale ETa model. Journal of Geophysical Research: Atmospheres, 108(D22).

Foken, T., Leuning, R., Oncley, S.R., Mauder, M., Aubinet, M., 2012. Corrections and data quality control, Eddy covariance. Springer, pp. 85-131.

Friedl, M.A. et al., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ., 114(1): 168-182. doi:10.1016/j.rse.2009.08.016

Gary, M.O., Rucker, D.F., Smith, B.D., Smith, D.V., Befus, K., 2013. Geophysical investigations of the Edwards-Trinity aquifer system at multiple scales: Intrepreting airborne and direct-current resistivity in karst. In: Land, L., Doctor, D.H., Stephenson, J.B. (Eds.), Sinkholes and the Engineering and Enviromental Impacts of Karst. National Cave and Karst Research Institute, Carlsbad, NM, pp. 195-206.

Hackett, C.C., 2019. Storage dynamics of the upper Nueces River alluvial aquifer: Implications for recharge to the Edwards Aquifer, Texas.

Heilman, J.L. et al., 2014. Water-storage capacity controls energy partitioning and water use in karst ecosystems on the Edwards Plateau, Texas. Ecohydrology, 7(1): 127-138. doi: 10.1002/Eco.1327

Heilman, J.L., McInnes, K.J., Kjelgaard, J.F., Owens, M.K., Schwinning, S., 2009. Energy balance and water use in a subtropical karst woodland on the Edwards Plateau, Texas. J. Hydrol., 373(3-4): 426-435. DOI:doi: 10.1016/j.jhydrol.2009.05.007

Hengl, T. et al., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One, 12(2): e0169748. DOI:10.1371/journal.pone.0169748

Homer, C. et al., 2020. Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. Int. J. Photogramm. Remote Sens., 162: 184-199. doi: 10.1016/j.isprsjprs.2020.02.019

Keese, K.E., Scanlon, B.R., Reedy, R.C., 2005. Assessing controls on diffuse groundwater recharge using unsaturated flow modeling. Water Resour. Res., 41(6). doi: 10.1029/2004WR003841

Koster, R.D., Suarez, M.J., 1992. Modeling the land surface boundary in climate models as a composite of independent vegetation stands. Journal of Geophysical Research: Atmospheres, 97(D3): 2697-2715.

Kukowski, K.R., Schwinning, S., Schwartz, B.F., 2013. Hydraulic responses to extreme drought conditions in three co-dominant tree species in shallow soil over bedrock. Oecologia, 171(4): 819-830. doi: 10.1007/s00442-012-2466-x

Law, B.E. et al., 2008. Terrestrial carbon observations: Protocols for vegetation sampling and data submission. FAO, Rome.

Lee, X., 1998. On micrometeorological observations of surface-air exchange over tall vegetation. Agricultural and Forest Meteorology, 91(1): 39-49.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7): 14415-14428.

Mace, R.E., Chowdhury, A.H., Anaya, R., Way, S.-C., 2000. Groundwater availability of the Trinity aquifer, Hill Country area, Texas: numerical simulations through 2050. Texas Water Development Board Report, 353: 117.

Maclay, R.W., 1995. Geology and hydrology of the Edwards Aquifer in the San Antonio area, Texas, US Geological Survey.

Maclay, R.W., Small, T.A., 1983. Hydrostratigraphic subdivisions and fault barriers of the Edwards Aquifer, South-Central Texas, USA. J. Hydrol., 61(1-3): 127-146.

Mahrt, L., Ek, M., 1984. The influence of atmospheric stability on potential evaporation. Journal of Climate and Applied Meteorology, 23(2): 222-234.

Marclay, R.W., 1995. Geology and hydrology of the Edwards aquifer in the San Antonio area, Texas: U.S. U.S. Geological Survey, pp. 64.

Mesinger, F. et al., 2006. North American regional reanalysis. Bull. Am. Meteorol. Soc., 87(3): 343-360. doi:10.1175/BAMS-87-3-343

Moriasi, D.N. et al., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE, 50(3): 885-900. DOI:10.13031/2013.23153

Myneni, R.B., Knyazikhin, Y., Park, T., 2015. MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD15A2H.006

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. J. Hydrol., 10(3): 282-290. DOI:10.1016/0022-1694(70)90255-6

Ockerman, D.J., 2007. Simulation of streamflow and estimation of ground-water recharge in the Upper Cibolo Creek Watershed, south-central Texas, 1992-2004. 2328-0328, US Geological Survey.

PRISM, 2018. Oregon State Univeristy. Oregon State Univeristy.

Puente, C., 1978. Method of estimating natural recharge to the Edwards Aquifer in the San Antonio area, Texas. Water-Resources Investigations Report 78-10, Austin, TX. DOI:10.3133/wri7810

Ritter, A., Muñoz-Carpena, R., 2013. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. J. Hydrol., 480: 33-45. DOI:10.1016/j.jhydrol.2012.12.004

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol., 251(3-4): 163-176.

Schwartz, B.F. et al., 2013. Using Hydrogeochemical and Ecohydrologic Responses to Understand Epikarst Process in Semi-Arid Systems, Edwards Plateau, Texas, USA. Acta Carsologica, 42(2-3): 315-325.
Schwinning, S., 2008. The water relations of two evergreen tree species in a karst savanna. Oecologia, 158(3): 373-83. DOI:doi: 10.1007/s00442-008-1147-2
Simunek, J., van Genuchten, M.T., Sejna, M., 2012. Hydrus: Model Use, Calibration, and Validation. Transactions of the Asabe, 55(4): 1261-1274.
Simunek, J., van Genuchten, M.T., Sejna, M., 2016. Recent Developments and Applications of the HYDRUS Computer Software Packages. Vadose Zone J., 15(7). DOI:doi:10.2136/vzj2016.04.0033
Soil Survey Staff, 2014. Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United State. United States Department of Agriculture, Natural Resources Conservation Service.
Twine, T.E. et al., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agricultural and Forest Meteorology, 103(3): 279-300.
Vachaud, G., Passerat de Silans, A., Balabanis, P., Vauclin, M., 1985. Temporal stability of spatially measured soil water probability density function. Soil Science Society of America Journal, 49(4): 822-828.
van Auken, O.W., 2000. Shrub invasions of North American semiarid grasslands. Annual Review Of Ecology And Systematics, 31: 197-215.
Wada, Y., Bierkens, M.F., 2014. Sustainability of global water use: past reconstruction and future projections. Environmental Research Letters, 9(10): 104003.
Wilcox, B.P., 2008. Juniperus woodlands and the water cycle on karst rangelands. In: Van Auken, O.W. (Ed.), Western North American Juniperus Communities. Ecological Studies. Springer New York, pp. 202-215. DOI:10.1007/978-0-387-34003-6_11
Wilcox, B.P., Huang, Y., 2010. Woody plant encroachment paradox: Rivers rebound as degraded grasslands convert to woodlands. Geophys. Res. Lett., 37. DOI:L07402, doi: 10.1029/2009gl041929
Wilcox, B.P., Owens, M.K., Knight, R.W., Lyons, R.K., 2005. Do woody plants affect streamflow on semiarid karst rangelands? Ecological Applications, 15(1): 127-136. DOI:doi: 10.1890/04-0664
Wilcox, B.P., Wilding, L.P., Woodruff, C.M., 2007. Soil and topographic controls on runoff generation from stepped landforms in the Edwards Plateau of Central Texas. Geophys. Res. Lett., 34(24). DOI:L24S24, doi:10.1029/2007GL030860
Wilson, K. et al., 2002a. Energy balance closure at FLUXNET sites. Agricultural and Forest Meteorology, 113(1): 223-243.
Wilson, K. et al., 2002b. Energy balance closure at FLUXNET sites. Agr. Forest Meteorol., 113(1-4): 223-243.
WMO, 2008. Guide to Meteorological Instruments and Methods of Observation. In: Organization, W.M. (Ed.). World Meteorological Organization, Geneva 2, Switzerland, pp. 680.
Wong, C., Banner, J.L., 2010. Response of cave air CO2 and drip water to brush clearing in central Texas: Implications for recharge and soil CO2 dynamics. Journal of Geophysical Research-Biogeosciences, 115. DOI:Artn G04018, doi 10.1029/2010jg001301
Wong, C.I., Mahler, B.J., Musgrove, M., Banner, J.L., 2012. Changes in sources and storage in a karst aquifer during a transition from drought to wet conditions. J. Hydrol., 468-469: 159-172. doi: 10.1016/j.jhydrol.2012.08.030
Wood, E.F., Lettenmaier, D., Liang, X., Nijssen, B., Wetzel, S.W., 1997. Hydrological modeling of continental-scale basins. Annual Review of Earth and Planetary Sciences, 25(1): 279-300.
Woodruff, C.M., Wilding, L.P., 2008. Bedrock, soils, and hillslope hydrology in the Central Texas Hill Country, USA: implications on environmental management in a carbonate-rock terrain. Environ. Geol., 55(3): 605-618.

Xia, Y. et al., 2014. Evaluation of multi-model simulated soil moisture in NLDAS-2. Journal of Hydrology, 512: 107-125.

Yang, L. et al., 2018. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. ISPRS journal of photogrammetry and remote sensing, 146: 108-123.

Zhang, Y., Schaap, M.G., 2017. Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). J. Hydrol., 547: 39-53. DOI:10.1016/j.jhydrol.2017.01.004

1. Potential Evapotranspiration Calculations

Variable	Description	Units	Eq/Constant/Input	Tall grass reference*
R_{solar}	Incoming shortwave/solar radiation	MJ m^{-2} h^{-1}	Input	
RH	Relative humidity		Input	
T	Temperature	°C	Input	
T_{dew}	Dew point temperature	°C	Input	
u_{hw}	Wind speed at h_w	m s^{-1}	Input	
P	Pressure	kPa	Input	
h_w	Height of wind measurement	m	Input	2
h_h	Height of humidity measurement	m	Input	2
h_v	Height of vegetation	m	Input	0.12
ET_o	Reference evapotranspiration	mm h^{-1}	Eq. 1-1	
e_s	Saturated vapor pressure	kPa	Eq. 1-2	
e_a	Actual vapor pressure	kPa	Eq. 1-3 or Eq. 1-4	
Δ	Slope of the saturation vapor pressure temperature relationship	kPa C^{-1}	Eq. 1-5	
λ	Latent heat of vaporization	MJ kg^{-1}	Eq. 1-6	
γ	Psychrometric constant	kPa C^{-1}	Eq. 1-7	
Ɛ	Ratio molecular weight of water vapor/dry air		0.622	
c_p	Specific heat of the air	MJ kg^{-1} °C^{-1}	1.01E–03	
r_a	Aerodynamic resistance	s m^{-1}	Eq. 1-8	Eq 1.9
z_{om}	Roughness length governing momentum transfer	m	Eq. 1-10	0.01476
z_{oh}	Roughness length governing transfer of heat and vapor	m	Eq. 1-11	0.001476
d	Displacement height		Eq. 1-12	0.08
K	Von Karman's constant	–	0.41	
r_s	Surface resistance	s m^{-1}	Eq. 1-13	70
Rl	Bulk stomatal resistance	s m^{-1}	100	100
LAI_a	Active Leaf area index	m^2 m^{-2}	Eq. 1-14	1.44
LAI	Leaf area index	m^2 m^{-2}	Eq. 1-15	2.88
ρ	Atmospheric density	kg m^{-3}	Eq 1-16	
R	Specific gas constant	kJ kg^{-1} K^{-1}	0.287	
T_{kv}	Mean virtual temperature	K	Eq. 1-17	

u_2	Wind speed at 2 meters	m s^{-1}	Eq. 1-18	Eq. 1.19
R_n	Net radiation	MJ m^{-2} h^{-1}	Eq. 1-20	
R_{ns}	Net shortwave radiation	MJ m^{-2} h^{-1}	Eq. 1-21	
A	Albedo		0.23	
R_{nl}	Net longwave radiation	MJ m^{-2} h^{-1}	Eq. 1-22	
Σ	Stefan-Boltzmann constant	MJ m^{-2} h^{-1}	2.042E–10	
f_{cd}	Fraction of cloud		Eq. 1-23	
R_{so}	Calculated clear-sky radiation	MJ m^{-2} h^{-1}	Eq. 1-24	
R_a	Extraterrestrial radiation	MJ m^{-2} h^{-1}	Eq. 1-25	
G_{sc}	Solar constant	MJ m^{-2} h^{-1}	4.92	
d_r	Inverse relative distance Earth–Sun		Eq. 1-26	
J	Julian day			
Φ	Latitude	Radians		
δ	Solar declination	Radians	Eq. 1-27	
ω_1	Solar time angle at beginning of period	Radians	Eq. 1-28	
ω_2	Solar time angle at end of period	Radians	Eq. 1-29	
Ω	Solar time angle at midpoint	Radians	Eq. 1-30	
T	Standard clock time at the midpoint of the period	Hours		
t_1	Length of calculation period	Hours		
L_z	Longitude of the center of the local time zone	Degrees		
L_m	Longitude of the measurement site	Degrees		
S_c	Seasonal correction for solar time	Hours	Eq 1-31	
G	Soil heat flux density	MJ m^{-2} h^{-1}	Eq 1-33	Eq 1-34

* Tall grass reference (ET_o) assumes $h_v = 0.12$ and $h_h = h_w = 2$ m

$$ET_o = \frac{\left(\frac{\Delta(R_n - G) + K_{time}\rho c_p \frac{e_s - e_a}{r_a}}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)}\right)}{\lambda} \qquad \textit{Eq. 1-1}$$

$$e_s = 0.6108 exp\left(\frac{17.27T}{T + 237.3}\right) \qquad \textit{Eq. 1-2}$$

This is the main equation to calculate e_a,

$$e_a = \frac{RH}{100} e_s \qquad \textit{Eq. 1-3}$$

If *RH* does not exist, then the following equation is used,

$$e_a = 0.6108 exp\left(\frac{17.27T_{dew}}{T_{dew} + 237.3}\right) \qquad \textit{Eq. 1-4}$$

$$\Delta = 4098 \frac{0.6108 exp\left(\frac{17.27\,T}{T+237.3}\right)}{(T+237.3)^2} \quad \text{Eq. 1-5}$$

$$\lambda = 2.501 - 2.361 \times 10^{-3} T \quad \text{Eq. 1-6}$$

$$\gamma = \frac{c_p P}{\varepsilon \lambda} \quad \text{Eq. 1-7}$$

$$r_a = \frac{ln\left(\frac{h_w - d}{z_{om}}\right) ln\left(\frac{h_h - d}{z_{oh}}\right)}{k^2 u_h} \quad \text{Eq. 1-8}$$

$$ra = \frac{208}{u_2} \quad \text{Eq. 1-9}$$

$$z_{om} = 0.123 h_v \quad \text{Eq. 1-10}$$

$$z_{oh} = 0.0123 h_v \quad \text{Eq. 1-11}$$

$$d = 0.67 h_v \quad \text{Eq. 1-12}$$

$$r_s = \frac{r_l}{LAI_a} \quad \text{Eq. 1-13}$$

$$LAI_a = 0.5 LAI \quad \text{Eq. 1-14}$$

$$LAI = 24 h_v \quad \text{Eq. 1-15}$$

$$\rho = \frac{P}{T_{kv} R} \quad \text{Eq. 1-16}$$

$$T_{kv} = \frac{T_k}{1 - 0.378 \frac{e_a}{P}} \quad \text{Eq. 1-17}$$

$$u_2 = u_{hw} \frac{ln(\frac{2-d}{z_{om}})}{ln(\frac{h_w - d}{z_{om}})} \quad \text{Eq. 1-18}$$

$$u_2 = u_h \frac{4.87}{ln(67.8 h_w - 5.42)} \quad \text{Eq. 1-19}$$

$$R_n = R_{ns} - R_{nl} \quad \text{Eq. 1-20}$$

$$R_{ns} = (1 - \alpha) R_{solar} \quad \text{Eq. 1-21}$$

$$R_{nl} = \sigma T_k^4 \left(0.34 - 0.14 \sqrt{e_a}\right) f_{cd} \quad \text{Eq. 1-22}$$

$$f_{cd} = a \frac{R_{solar}}{R_{so}} + b \quad \text{Eq. 1-23}$$

In which:

- $a = 1.35$
- $b = -0.35$
- $f_{cd} = f_{cd} (\beta > 0.3)$

$$R_{so} = (0.75 + 2 \times 10^{-5} z) R_a \quad \text{Eq. 1-24}$$

$$R_a = \frac{12}{\pi} G_{sc} d_r [(\omega_2 - \omega_1) \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) (\sin(\omega_2) - \sin(\omega_1))] \quad \text{Eq. 1-25}$$

$$d_r = 1 + 0.033 \cos\left(\frac{2\pi}{365} J\right) \quad \text{Eq. 1-26}$$

$$\delta = 0.409 \sin\left(\frac{2\pi}{365} J - 1.39\right) \quad \text{Eq. 1-27}$$

$$\omega_1 = \omega - \frac{\pi t_1}{24} \quad \text{Eq. 1-28}$$

$$\omega_2 = \omega - \frac{\pi t_1}{24} \quad \text{Eq. 1-29}$$

$$\omega = \frac{\pi}{12} [(t + 0.06667(L_z - L_m) + S_c) - 12] \quad \text{Eq. 1-30}$$

$$S_c = 0.1645 \sin(2b) - 0.1255 \times \cos(b) - 0.025 \times \sin(b) \quad \text{Eq. 1-31}$$

$$b = \frac{2\pi(J - 81)}{364} \quad \text{Eq. 1-32}$$

$$G = K_G \exp(-0.5\, LAI)\, R_n \quad \text{Eq. 1-33}$$

K_G = 0.4 during daytime R_n >0

K_G = 2.0 during nighttime R_n <0

$G = 0.1 R_n$ *for Rn >0, else* $G = 0.5 \times R_n$ *Eq. 1-34*

2. Eddy Covariance Data Processing

The EasyFlux software (Campbell Scientific, Logan, UT) removes spikes from 10-Hz data using signal strength, measurement output range, and flags which include diagnostic data. The 10-Hz data was then maximized for covariance to compensate for any lags and averaged to create 30-minute data. Axis rotation for tilt correction was implemented using the double-rotation method (Tanner and Thurtell, 1969). Averaging the 10-Hz data (block averaging) leads to attenuation of low frequencies. Eddies smaller than the measurement path or volume (line/volume averaging) lead to signal attenuation of high-frequency data. Frequency corrections account for block averaging (Kaimal et al., 1989), line averaging (Moore, 1986; Moncrieff et al., 1997b; van Dijk, 2002; Foken et al., 2012a), and sensor separation (Horst and Lenschow, 2009; Foken et al., 2012a). The frequency correction for sensor separation is 1 for the IRGASON due to the design. The SND correction named after Schotanus, Nieuwstadt, and DeBruin (1983) was conducted to correct the fluctuation of sonic temperature for humidity to obtain the fluctuations of the actual temperature (Schotanus et al., 1983; Van Dijk et al., 2004). The WPL correction named after Webb, Pearman, and Leuning (1980) was applied to the CO_2 and H_2O flux to correct for air-density changes (Webb et al., 1980). Footprint analysis was then implemented using the findings of Kljun et al. (2004) if the atmospheric stability (z/L), friction velocity (u_*), and aerodynamic height (z) were within certain ranges. Those ranges are $-200 \le z/L \le 1$, $u_* \ge 0.2$, and $z \ge 1$. For data that did not meet the atmospheric conditions, the algorithm described by Kormann and Meixner (2001) was applied.

For the LI-COR's EddyPro software, the 10-Hz data was flagged using statistical screenings which include spike detection, amplitude resolution, drop outs, absolute limits, and skewness and kurtosis (Vickers and Mahrt, 1997). The 10-Hz data was then maximized for covariance to compensate for any lags and averaged to create 30-minute data. Block averaging was used to extract the turbulent fluxes. Axis rotation for tilt correction was implemented using the double-rotation method (Wilczak et al., 2001). WPL and SND corrections were also applied (Webb et al., 1980; Van Dijk et al., 2004). Low-pass frequency correction (Moncrieff et al., 2004) and high-pass frequency correction (Moncrieff et al., 1997a) was employed.

For all systems, data was discarded if 10% of the 10-Hz data was missing in the 30-minute interval. Spike detection was implemented for a 13-day window based on the double-differenced time series, using the median of absolute deviation about the median (Papale et al., 2006). Gap filling was done for the missing data. If the gap was less than 2 hours, then the missing data was filled through linear interpolation. The missing R_n data was then filled using the nearby station. For gaps greater than 2 hours, R_n, air temperature, and vapor pressure deficits were used to find similar data within a ±7 day period (Reichstein et al., 2005).

Energy balance states the difference between R_n and G (the available energy) should equal the sum of the sensible heat flux (H) and latent heat flux (LE) (the surface energy flux). The sensible heat flux is a result of the heat energy that is due to the temperature gradient. In contrast, LE is related to phase change (evaporation). H and LE are found using Equations 2-1 and 2-2,

$$H = \rho_a c_p \overline{w'T'} \qquad \textit{Eq. 2-1}$$

$$LE = \lambda \overline{w'q'} \qquad \textit{Eq. 2-2}$$

in which c_p is the specific heat of air, ρ_a is the density of the air, λ is the latent heat of vaporization, $\overline{w'T'}$ is the covariance between the sonic temperature and the vertical wind, and $\overline{w'q'}$ is the covariance between the water density and the vertical wind. The eddy covariance system can provide the terms $\overline{w'T'}$ and $\overline{w'q'}$.

G is found by adding the soil heat flux measurements to a storage term ($\Delta_{storage}$) (Eq. 2-3),

$$\Delta_{storage} = \frac{[c_s \rho_s (T_{soil,f} - T_{soil,i}) + c_w \rho_w (T_{soil,f} q_{v,f} - T_{soil,i} q_{v,i})] D}{\Delta t} \qquad \textit{Eq. 2-3}$$

in which c_s is the specific heat of the soil, ρ_s is the soil bulk density, T_{soil} is the soil temperature, c_w is the specific heat of water, ρ_w is the density of water, q_v is the volumetric water content, D is the depth of the soil heat flux plate, and Δt is the averaging interval.

Due to factors still being studied in the scientific community, the surface energy fluxes are often times underestimated by 10–30% relative to the available energy (Wilson et al., 2002; Foken et al., 2012b). LE and H can be adjusted to force closure while maintaining a constant Bowen ratio (β), which is the ratio between H and LE as shown in Equations 2-4 and 2-5 (Blanken et al., 1997; Lee, 1998; Twine et al., 2000).

$$LE_{corr} = \frac{(R_n - G)}{1 + \beta} \qquad \textit{Eq. 2-4}$$

$$H_{corr} = LE_{corr} \times \beta \qquad \textit{Eq. 2-5}$$

3. Evapotranspiration Modeling

3.1 Vegetation Indices

Landsat surface reflectance spectral indices were derived from Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI)/Thermal Infrared Sensor (TIRS) scenes. Vegetation indices were derived using a combination of these spectral bands after removing contaminated (quality-flagged) pixels.

The normalized difference VI (NDVI) is used to quantify vegetation greenness and is useful for understanding vegetation density and assessing changes in plant health. The NDVI uses the red (R) and the near infrared (NIR) bands and can be found with Equation 3.1 (Rouse Jr et al., 1974).

$$NDVI = \frac{NIR - R}{NIR + R} \qquad \textit{Eq. 3-1}$$

The soil adjusted VI (SAVI) corrects NDVI for soil brightness in areas with less vegetation using *L* as a correction factor (Huete, 1988). Even though Huete (1988) found that the optimal *L* varies with vegetation density, a constant value was used to reduce the soil noise through a wide range of vegetation. Because *L* is usually larger than R, it would buffer the reflectance variations. The SAVI is calculated as:

$$SAVI = (1 + L)\frac{NIR - R}{NIR + R + L} \qquad \textit{Eq. 3-2}$$

The Modified Soil Adjusted VI (MSAVI) attempts to minimize the effect of bare soil on SAVI by implementing an inductive *L* function to maximize the reduction of soil effects on the vegetation signal (Qi et al., 1994). The inductive *L* is able to continue to reduce the soil background effect and increase the vegetation sensitivity by increasing the range compared to the SAVI. The MSAVI is

$$MSAVI = \frac{2 \times NIR + 1 - \sqrt{(2 \times NIR + 1)^2 - 8(NIR - R)}}{2} \qquad \textit{Eq. 3-3}$$

Like NDVI, the enhanced VI (EVI) quantifies vegetation greenness but is more sensitive in areas with dense vegetation. EVI incorporates *L* to account for canopy background, *C* for atmospheric resistance, and the blue band (B) for reducing background noise, atmospheric noise, and saturation (Liu and Huete, 1995). The EVI is calculated as:

$$EVI = G\frac{NIR - R}{NIR + C1 \times R - C2 \times B + L} \qquad \textit{Eq. 3-4}$$

The Landsat 7 and 8 Collection 1 spectral reflectance vegetation indices were acquired through USGS's Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) (https://espa.cr.usgs.gov/).

3.2 Evapotranspiration Comparison

North American Land Data Assimilation Systems Phase 2 (NLDAS-2) was used to force land-surface models that use different methodologies (Noah, Mosaic, and VIC). The VIC model came from the hydrologic community and is a macroscale semidistributed model (Liang et al., 1994; Wood et al., 1997). Whereas both the Noah (Chen et al., 1996; Ek et al., 2003a) and the Mosaic (Koster and Suarez, 1992) models were developed to be coupled with climate models, the Mosaic model accounts for the subgrid heterogeneity

of soil moisture and vegetation. The mean of the land-surface model was also evaluated against the ET, as some studies have found that using a mean ensemble of the land-surface model performs better than the individual models (Xia et al., 2014). The evapotranspiration product from Modis (MOD16A2) is an 8-day composite with a 500-m resolution.

4. ESPERE Approach and Results

ESPERE (estimation of effective rainfall and groundwater recharge) (Lanini et al., 2016) is a toolbox that provides several (semi)analytical methods for estimating recharge (Lanini et al., 2016). The inputs into the multiple models from gridMET include daily precipitation, minimum and maximum temperature (yielding a mean daily temperature), and daily PET. Three models predict recharge on an annual basis (Guttman and Zuckerman, 1995; Kessler, 1967; Turc, 1954), whereas other models predict recharge on a daily timestep (Dingman, 2002; Thornthwaite, 1948).

Model	Inputs	Output	Assumptions
Guttman	PPT	Yearly	Empirical formula developed in Israel
Turc	PPT, Temp	Yearly	
Kessler	PPT	Yearly	Emphasizes recharge during Sept-Dec and Jan-Apr
Thornthwaite	PPT, PET, Temp	Daily	Uses crop coefficient. When P_{eff} is less than zero, P_{eff} is zero
Dingman-Hamon	PPT, Temp	Daily	Calculates PET from T. When P_{eff} is less than zero P_{eff} is zero
ET-EVI	EVI, PET, PPT, SM, RO	Monthly	Extrapolation past the range of ET-EVI
HYDRUS-1D	PPT, PET, LAI, vegetation class, Soil	Daily	

Turc (1954), Kessler (1967), and Guttman and Zuckerman (1995) developed an empirical formula to estimate annual groundwater recharge. The formula was developed in a semiarid Mediterranean climate for a carbonate aquifer in Israel. The following equations are used depending upon the amount of precipitation.

$$R = 0.45(P - 180) \quad \textit{when } P < 600 \textit{ mm} \qquad \textit{Eq. 4-1a}$$

$$R = 0.88(P - 410) \quad \textit{when } 600 \textit{ mm} < P < 1000 \textit{ mm} \qquad \textit{Eq. 4-1b}$$

$$R = 0.97(P - 463) \quad \textit{when } P > 1000 \textit{ mm} \qquad \textit{Eq. 4-1c}$$

in which *R* is annual recharge and *P* is annual precipitation.

Turc (1954), Kessler (1967), and Guttman and Zuckerman (1995) also developed a relationship between annual precipitation and mean annual temperature (°C) by using discharge at the outlet for 254 catchments throughout the world to find annual effective rainfall (P_{eff}).

$$P_{eff} = P - \frac{P}{0.9 + \frac{P^2}{L^2}} \qquad \textit{Eq. 4-2}$$

$$L = 300 + 25\,T + 0.05T^2 \qquad \text{Eq. 4-3}$$

Turc (1954), Kessler (1967), and Guttman and Zuckerman (1995), using 22 years of observations of a karstic spring in Hungary, found that the annual discharge was correlated with the precipitation for the previous autumn (P_{sep_dec}) and the beginning of the year (P_{jan_apr}). However, Andreo et al. (2008) found that this equation is justified only for aquifers in Mediterranean climate conditions with two annual recharge periods. This equation uses the corrective precipitation rate (CPR), interannual mean of precipitation from September to December (*m*), correcting constant (*k*), determinative precipitation rate (*DPR*), and infiltration rate (*IR*).

$$CPR = \frac{\left(P_{sep_{dec}} - m\right)}{m} \qquad \text{Eq. 4-4a}$$

$$DPR = \frac{P_{jan\_apr}}{P} + k \qquad \text{Eq. 4-4b}$$

$$IR = 0.0003 \times DPR^3 - 0.0339 \times DPR^2 + 2.2778 \times DPR - 1.5758 \qquad \text{Eq. 4-4c}$$

$$R = IR \times P \qquad \text{Eq. 4-4d}$$

Thornthwaite (1948) used daily precipitation, potential evapotranspiration (*PET*) and average daily temperature to find daily effective rainfall. First, actual evapotranspiration (*ET*) is calculated by Equation 4-5a, in which *Kc* is the crop coefficient and *RU* is the soil-water storage. P_{eff} is calculated using Equation 4-5b. If P_{eff} is <0 then P_{eff} becomes 0. *RU* for the next timestep is calculated using Equation 4-5c. RU_{max} is assumed to be 48 mm.

$$ET_A(i) = min\left(Kc(i) \times ET_P(i); P(i) + RU(i)\right) \qquad \text{Eq. 4-5a}$$

$$P_{eff}(i) = P(i) - ET_A(i) + RU(i) - RU_{max} \qquad \text{Eq. 4-5b}$$

$$RU(i+1) = min\left(RU_{max};\ P(i) + RU(i) - ET_A(i)\right) \qquad \text{Eq. 4-5c}$$

Thornthwaite (1948), Turc (1954), and Dingman (2002) used daily precipitation and daily mean temperature to find daily effective rainfall. First, ET_P was calculated using Hamon, in which Δ is the sun declination and *I* is the latitude in radians (Eq. 4-6a–c). This method accounts for snow, which was ignored.

$$ET_p = \frac{29.8 \times D \times esat}{T + 273.2} \qquad \text{Eq. 4-6a}$$

$$esat = 0.611 \times exp\left(\frac{17.3 \times T}{T + 237.3}\right) \qquad \text{Eq. 4-6b}$$

$$D = 2 \times acos(-tan(\Delta) \times tan(l)) \qquad \text{Eq. 4-6c}$$

The next step is to calculate ET and RU (Eq. 4-6d–h). If $P \geq Kc \times ET_p$,

$$ET_a = Kc \times ET_p \qquad \text{Eq. 4-6d}$$

$$RU(i) = min\left(P(i) - Kc(i) \times ET_p(i) + RU(i-1); RU_{max}\right) \qquad \text{Eq. 4-6e}$$

If $P < Kc \times ET_p$,

$$ET_a = P + RU(i-1) - RU(i) \qquad \text{Eq. 4-6f}$$

$$RU(i) = RU(i-1) \times exp\left(\frac{Kc(i) \times ET_p(i) - P(i)}{RU_{max}}\right) \quad \textit{Eq. 4-6g}$$

Then P_{eff} is calculated from Equation 4.-6h. If P_{eff} is <0, then P_{eff} = 0.

$$P_{eff} = P(i) - ET_a(i) - (RU(i) - RU(i-1)) \quad \textit{Eq. 4-6h}$$

Comparison between the models in ESPERE and USGS recharge, with the recharge from the water balance, is shown in Figure A4-1. Guttman, Turc, Thornthwaite, and Dingman-Hamon have an R^2 >0.9 for both basins (Table A4-1). The Kessler method and USGS led to the biggest difference with an R^2 of ~0.55 and an RMSD >0.14 km^3. The Kessler method emphasizes recharge during the first 4 and last 4 months of the year. For this region, large precipitation events that produce recharge can occur outside of these months, especially in May, one of the highest months for precipitation. The Medina-Cibolo Basin had slightly less bias and lower RMSDs for all models. The Guttman model produced the highest bias but the lowest ubRMSD. Figure SI9 shows the model developed by Guttman and Zuckerman (1995) compared to the results of Cibolo and Medina-Cibolo Basins. Results show that recharge is higher for the model given the precipitation. This could be because the model was developed in Israel, which has different vegetation and climate. The recharge is positively biased for the Turc, Thornthwaite, and USGS methods and negatively biased for Guttman, Kessler, and Hamon. The Turc (1954) model was developed using areas throughout the world, which may explain the lower bias than the model developed by Guttman and Zuckerman (1995). The Turc (1954) model also tends to overpredict recharge when the average annual precipitation is less than 700 mm (Figure SI10). When comparing the models that run on a daily timestep (Thornthwaite and Dingman-Hamon) to the monthly recharge (Figure SI11), the correlation decreases slightly (Table A4-2), with the Dingman-Hamon model having a higher R^2. The RMSD for the monthly timestep is lower for the Thornthwaite model.

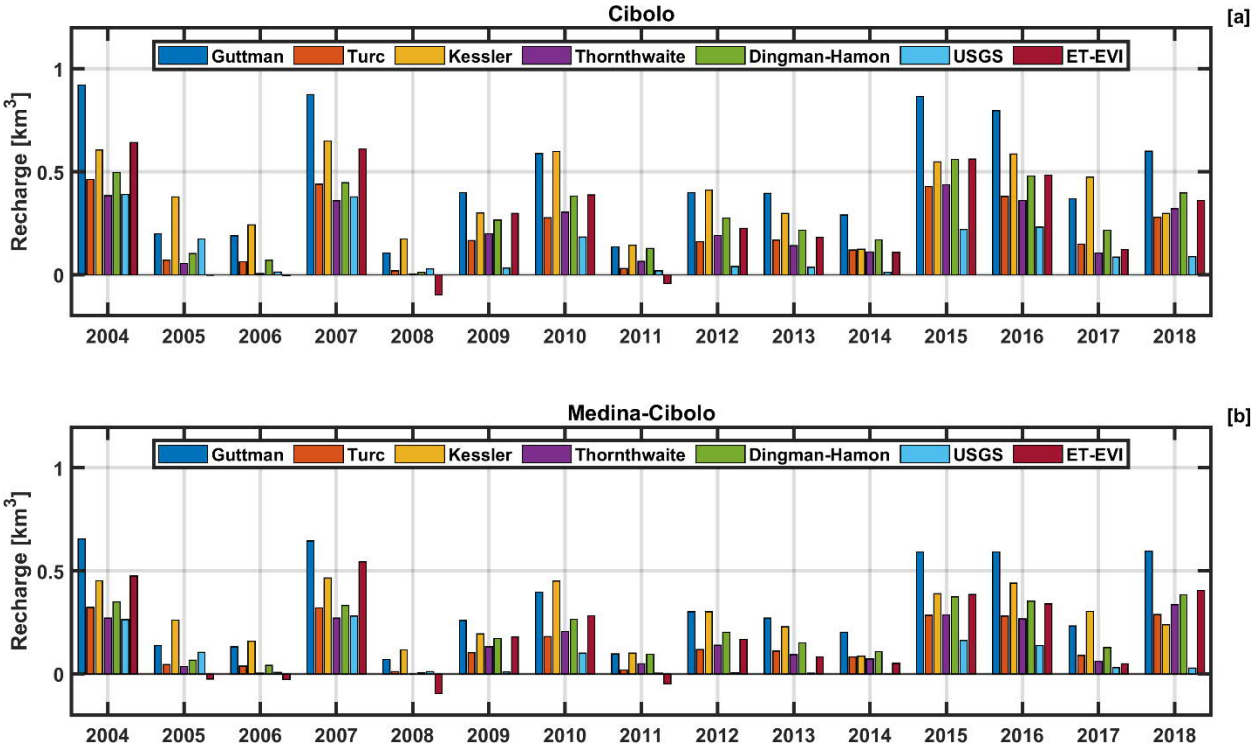


Figure A4-1: Total recharge comparison between existing recharge models and recharge found from the water balance method (ET-EVI).

Table A4-1: Root mean squared difference (RMSD), correlation (R^2), and mean bias error (MBE) of yearly recharge between models in ESPERE and USGS and the recharge found in this study for the two basins, Cibolo (Cib) and Medina-Cibolo (Med-Cib).

	MBE (10^{-6} ac-ft)		R^2		RMSD (10^{-6} ac-ft)		ubRMSD (10^{-6} ac-ft)	
	Cib	Med-Cib	Cib	Med-Cib	Cib	Med-Cib	Cib	Med-Cib
Guttman	-0.17	-0.12	0.79	0.78	0.18	0.13	0.05	0.03
Turc	0.04	0.03	0.78	0.78	0.09	0.08	0.07	0.07
Kessler	-0.10	-0.07	0.50	0.45	0.15	0.13	0.12	0.11
Thornthwaite	0.05	0.04	0.74	0.73	0.10	0.09	0.08	0.08
Dingman-Hamon	-0.02	-0.01	0.75	0.74	0.07	0.07	0.06	0.07
USGS	0.10	0.04	0.45	0.46	0.16	0.11	0.13	0.11

Table A4-2: Root mean squared difference (RMSD), correlation (R^2), and mean bias error (MBE) of monthly recharge between models in ESPERE and the recharge found in this study for the two basins, Cibolo (Cib) and Medina-Cibolo (Med-Cib).

	Bias (10^{-6} ac-ft)		R^2		RMSD (10^{-6} ac-ft)		ubRMSD (10^{-6} ac-ft)	
	Cib	Med-Cib	Cib	Med-Cib	Cib	Med-Cib	Cib	Med-Cib
Thornthwaite	-0.01	0.00	0.65	0.69	0.03	0.02	0.03	0.02
Dingman-Hamon	-0.02	-0.01	0.71	0.73	0.04	0.02	0.03	0.01

5. HYDRUS Modeling

For numerical simulations, we used the HYDRUS-1D model (Simunek et al., 2012; 2016) that simulates water flow in variably saturated porous media utilizing a Galerkin finite-element method based on the mass-conservative iterative scheme (Celia et al., 1990). The model solves a one-dimensional Richards equation,

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} K \left(\frac{\partial h}{\partial z} + 1 \right) - S \qquad \textit{Eq. 5.1}$$

in which t represents time (h), z is the vertical coordinate (positive upward) (cm), h denotes the pressure head (cm), K is the unsaturated hydraulic conductivity function (cm h^{-1}), and S is a sink term representing plant water uptake [$m^3\ m^{-3}\ h^{-1}$] defined by Feddes et al. (1978),

$$S(h) = \beta(h, z) S_p \qquad \textit{Eq. 5.2}$$

in which the root-water-uptake water-stress-response function β(h) is a prescribed dimensionless function of the soil water-pressure head and depth ($0 \leq \beta \leq 1$), and S_p is the potential water uptake rate. The water-stress-response function accounts for the reduction in water uptake rate as the soil water potential is reduced below an optimum value. Integrating Equation 3.2.2.4.2 over the depth domain, z, the actual transpiration rate, T_a, is,

$$T_a = PT \int_0^z \beta(h, z) S_p dz \qquad \textit{Eq. 5.3}$$

in which PT is the potential transpiration rate defined below. The root zone distribution with depth was regarded as non-uniform, with a higher root mass near the soil surface and decreasing root mass with depth. The non-uniform distribution, used for all simulations, can be approximated using a root zone distribution function (Vrugt et al., 2001) as,

$$\beta(z) = \left(1 - \frac{z}{Z_d}\right) exp\left[- \left(\frac{P_z}{Z_d}\right) |Z_m - z| \right] \qquad \textit{Eq. 5.4}$$

in which β(z) is the dimensionless spatial root distribution with depth, z ($z \geq 0$), Z_d is the maximum rooting depth (cm), P_z (–) is an empirical parameter, and Z_m is the depth at which root distribution is maximized (cm). Here, we define Z_m using the International Geosphere–Biosphere Programme (IGBP) vegetation classification (Loveland and Belward, 1997; Friedl et al., 2010) and the vegetation look-up table from Ek et al. (2003b) which specifies Z_d. We specify Z_m as $1/3 Z_d$, and P_s as 1.2.

We extract the atmospheric-limited potential evapotranspiration (PET) from the 4-km Gridmet data set (Abatzoglou, 2013) and partition it between potential canopy transpiration (PT) and soil evaporation (PE) based on exponential relationship to leaf area index (LAI) as proposed by Kemp et al. (1997),

$$PT = PET\left(1 - e^{-kLAI}\right) \quad \textit{Eq. 5.5}$$

$$PE = PET\left(e^{-kLAI}\right) \quad \textit{Eq. 5.6}$$

in which the extinction coefficient, *k*, is 0.39 (Ritchie, 1972); LAI was more recently updated weekly from a 500-m gridded MODIS-derived data set (Myneni et al., 2015).

For all simulations, we applied the daily total of precipitation accumulated from EAA RainVieux and applied it over the first 12-hours of each day. Daily PET from Gridmet was applied over the remaining 24.

Soil hydraulic properties were obtained from the POLARIS data set (Chaney et al., 2019) that derived 30-m probabilistic soil properties from terrain and soils data in the National Cooperative Soil Survey Soil Geographic Database (SSURGO) and gridded SSURGO database (Soil Survey Staff, 2014). Additionally, the SSURGO depth to restrictive horizon was rasterized onto the POLARIS grid and used to define the soil profile maximum depth within a given POLARIS soil layer. POLARIS consists of six soil layers with boundaries at 5, 15, 30, 60, 100 and 200 cm (Table A5.1).

Table A5-1. Soil texture and hydraulic properties extracted from the Polaris dataset at each field study site.

	Depth	Sand	Silt	Clay	Bulk density	θ_r	θ_s	α	n	K_s
	(cm)	(%)	(%)	(%)	(g/cm^3)	(m^3/m^3)	(m^3/m^3)	(1/cm)	(–)	(cm/h)
Savanna										
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.724
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.669
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.735
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.915
Woodland										
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.636
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.582
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.639
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.827
Grassland										
L1	0–5	26.5	37.3	26.5	1.23	0.080	0.536	0.108	1.35	0.709
L2	5–15	24.0	33.5	29.5	1.29	0.079	0.513	0.109	1.35	0.653
L3	15–30	16.5	30.4	42.5	1.29	0.100	0.513	0.079	1.28	0.706
L4	30–60	8.0	32.0	57.5	1.29	0.119	0.513	0.055	1.25	0.899

Hydraulic soil properties are described by the modified Mualem-van Genuchten model (Vogel et al., 2000) in which,

$$S_e = \frac{\theta-\theta_r}{\theta_s-\theta_r} = \begin{cases} [1 + |\alpha h|^n]^{-m} & h < h_s \\ 1 & h \geq h_s \end{cases} \qquad \textit{Eq. 5.7}$$

and,

$$K(S_e) = K_s S_e^l \left[1 - \left(1 - {S_e}^{1/m}\right)^m\right]^2 \qquad \textit{Eq. 5.8}$$

in which S_e is the effective soil water content, θ_r and θ_s denote the residual and saturated soil water content respectively ($L^3\ L^{-3}$), α (L^{-1}) and n (–) are parameters that define the shape of the water retention function, K_s represents saturated hydraulic conductivity ($L\ T^{-1}$), l is the pore-connectivity parameter by Mualem (1976), and h_s <0 is the maximum pressure head allowed at the soil surface (i.e., the non-zero minimum capillary height) (Vogel et al., 2000). Empirical coefficients were used at their assumed values in which, $l = 0.5$ and $m = 1–1/n$.

To solve Equation 5.1, the soil profile boundary and initial state conditions must be defined as,

$$h(z,t) = h_i(z) \qquad \textit{at } t = 0 \qquad \textit{Eq. 5.9}$$

and,

$$-K\left(\frac{\partial h}{\partial z} + 1\right) = q_0(t) - \frac{\partial h}{\partial t} \qquad \textit{at } z = 0, \textit{ for } h_a \leq h \leq h_s \qquad \textit{Eq. 5.10}$$

- $h(0,t) = h_a$ for $h \leq h_a$
- $h(0,t) = h_s$ for $h > h_s$

in which $h_i(t)$ is the initial pressure head set to −1000 cm, $q_0(t)$ is the net infiltration rate (i.e., precipitation minus evaporation), and h_a signifies the minimum pressure head allowed at the soil surface set equal to −10,000 cm. Equation 5.8 describes the atmospheric boundary condition at the soil interface, which switches between a prescribed flux condition and a prescribed head condition depending on the prevailing pressure head at the surface. During infiltration, the flux ($q(t)$) cannot exceed the $K(h_s)$. Any additional flux is considered surface runoff. We specify h_s as −1 cm to improve convergence when solving Equation 5.1. Lastly, the bottom boundary can be prescribed as either a unit gradient or seepage-face condition. Conceptually, the former is equivalent to an infinite soil column, whereas the later must reach a prescribed h to initiate flux out of the profile. This is similar to a column experiment in which the exit is essentially at atmospheric conditions and saturation must be reached prior to any discharge. Impermeable or semi-permeable karst fractures are more appropriated simulated with a seepage-face boundary, set here to −10 cm.

6. Geophysics

Plot-scale geophysical measurements were collected using electromagnetic induction (EMI), ground-penetrating radar (GPR), and time-domain electromagnetic induction (TEM). Collection of EMI and GPR measurements was attempted at all EAA PET locations. At each site, a 100-m-long transect was marked using plastic stakes so it could be revisited. During each measurement, the depth to bedrock was measured using a 1-m extended drill bit. Soil water content (at 0–12 cm) was measured with a portable soil moisture probe (Hydrasense II, Campbell Scientific). Measurement dates include:

Date	EMI	GPR	TEM
April 14, 2016	X		
August 10, 2016	X		
September 8, 2016	X		
November 14–16, 2016	X	X	X
April 2017	X	X	
November 2017	X	X	
March 7, 2018	X		
March 22, 2018	X		
April 12, 2018	X		
April 19, 2018	X		

These efforts were largely failures. Electrical conductivity was generally very low (<20 mS/m) in both the soil and epikarst. For either EMI or GPR to image the subsurface, there must be dichotomy in materials. We found none, even after wet periods. Thus, the EMI inversion software had difficulties resolving material properties. The GPR reflections were very faint and of little value. We also found our validation (drilling and portable moisture meter) was insufficient. The drilling was laborious and ineffective. The drill was routinely stopped by what was possibly gravel. If the soil was dry, it stopped much shallower than when wet. The moisture sensor resolved soil moisture too shallow to be comparable to either EMI or GPR.

The results are presented here, but we found little use overall for the geophysical measurements.

6.1 Electromagnetic Induction (EMI)

A Dualem-2 sensor (Dualem Instruments, Ontario, Canada) was used to measure bulk electrical conductivity from four dipoles ranging from 0.5 to 3.0 m in depth. Soil-water storage, soil mapping, and depth-to-bedrock are critical components for estimating and scaling the storage of infiltrated waters in the shallow vadose zone. Using time-lapse EMI, we can map changes in electrical properties of the shallow subsurface to better understand the spatial variability of transient soil moisture and static soil properties (Abdu et al., 2007; Robinson et al., 2008; Robinson et al., 2012).

The Dualem-21S operates in a narrow band around a frequency of 9.0 kHz (Figure A6-1a). A schematic of the sensor shows dual-geometry receivers at 1 and 2 m from the transmitting coil (Figure A6-1b). This configuration allows for simultaneous sounding of four depths of conductivity and the ability to analyze subsurface layering in the top few meters of the earth. All data were inverted using the EMTOMO software (Santos et al., 2010; Monteiro Santos et al., 2011).

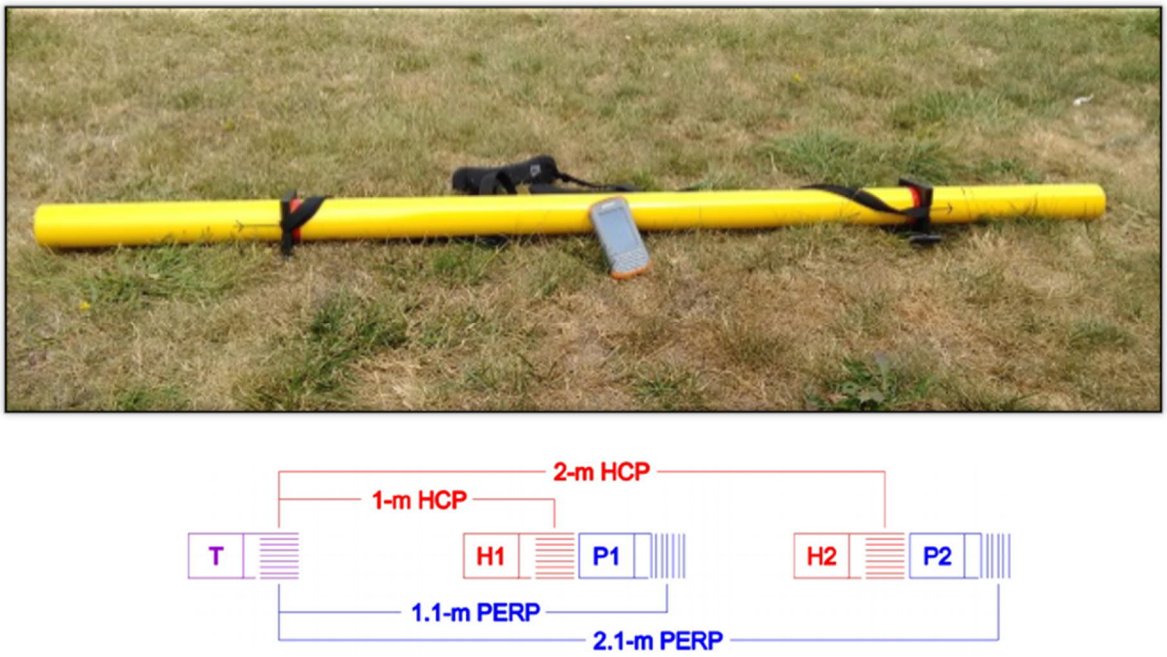
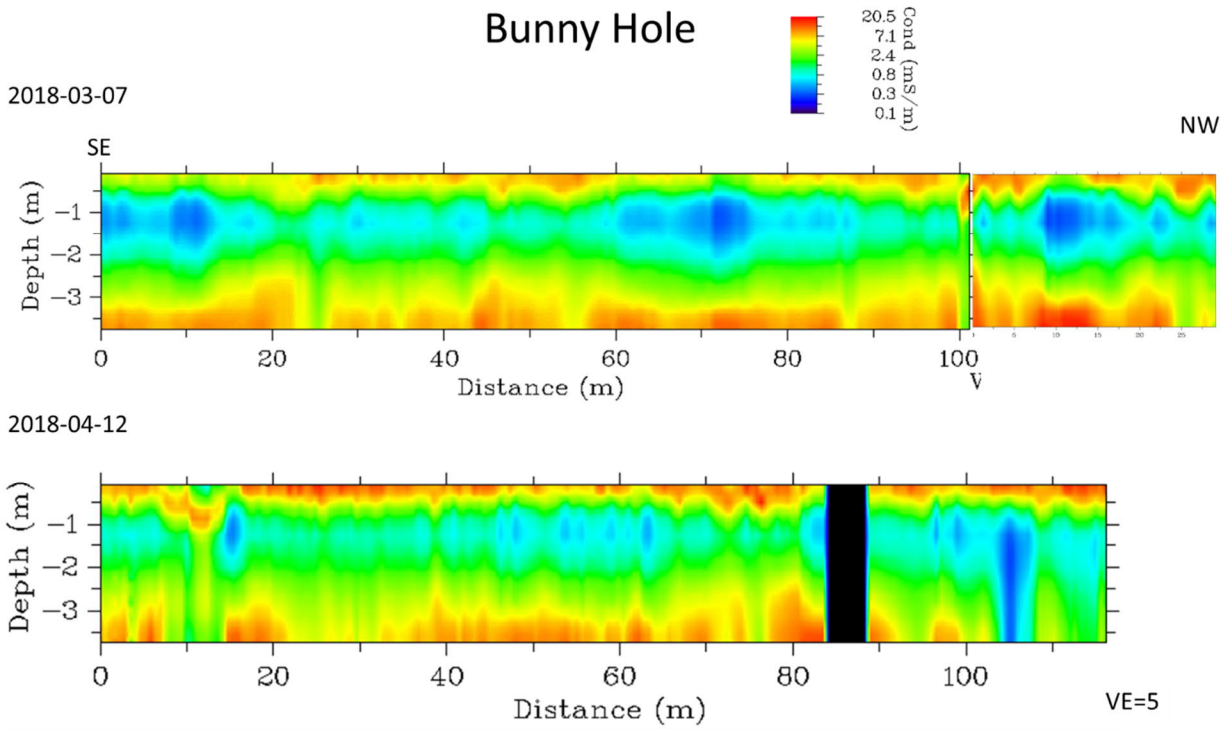


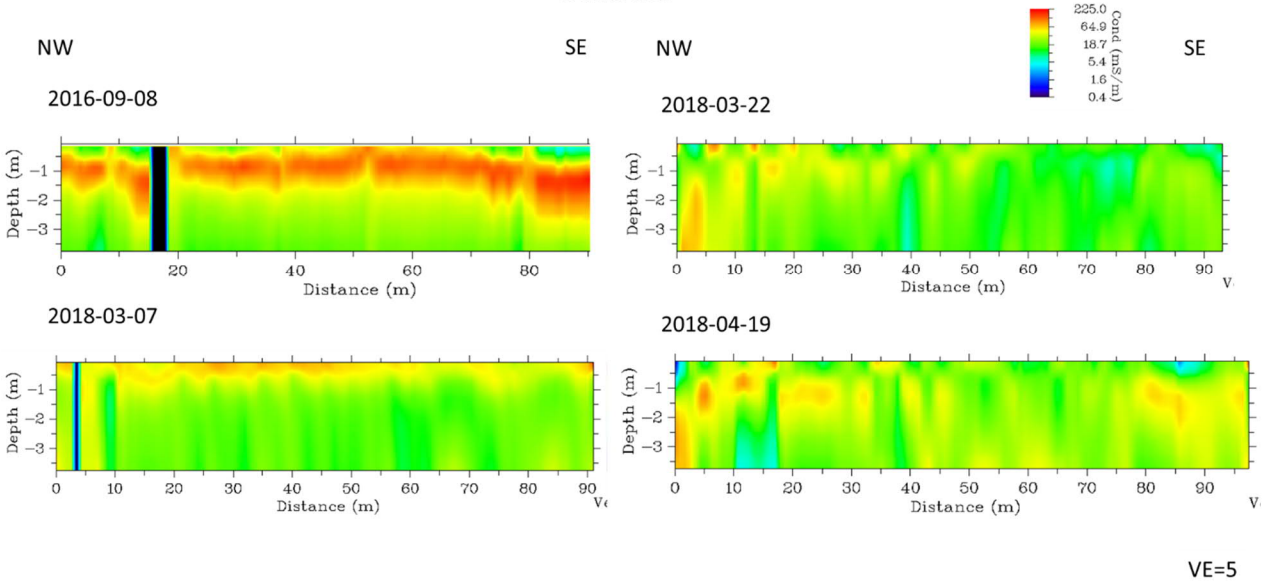
Figure A6-1: Schematic of the Dualem-21S sensor. It shows the transmitting coil (T), and four receiving coils—two of which are in a horizontal coplanar (HCP) (H1 & H2), whereas the other two are in a perpendicular (PERP) (P1 & P2) loop orientation with respect to the transmitting coil.

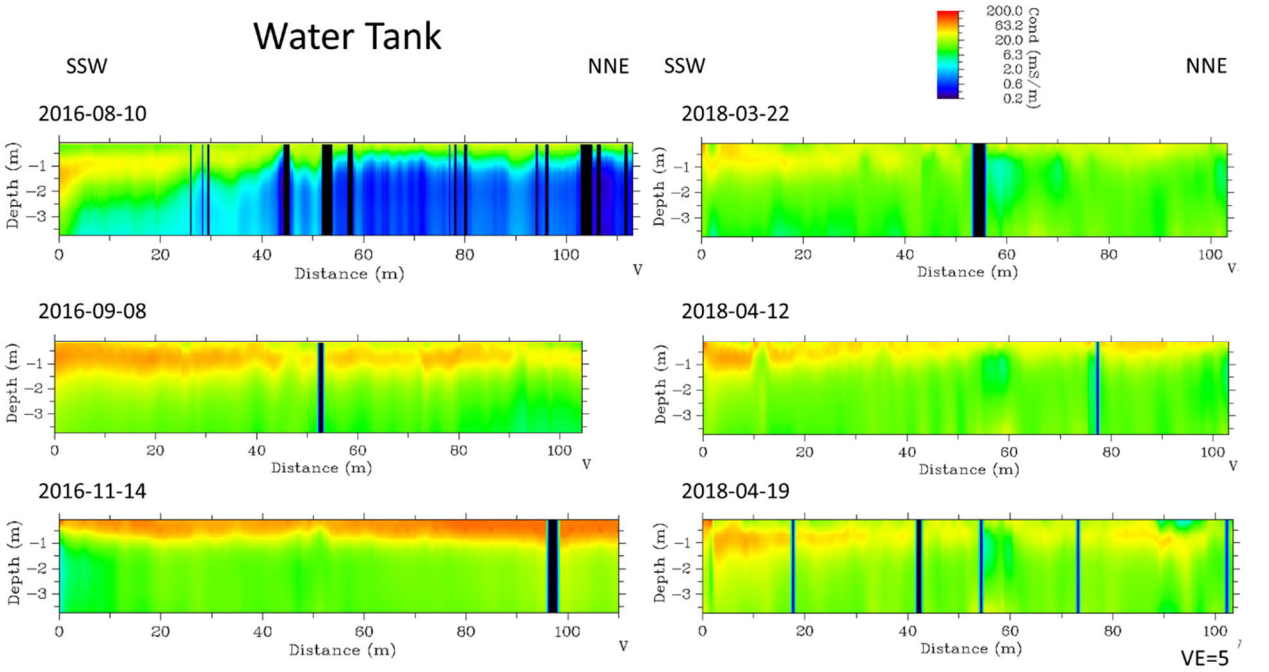
Below, inverted conductivity plots are presented by site.

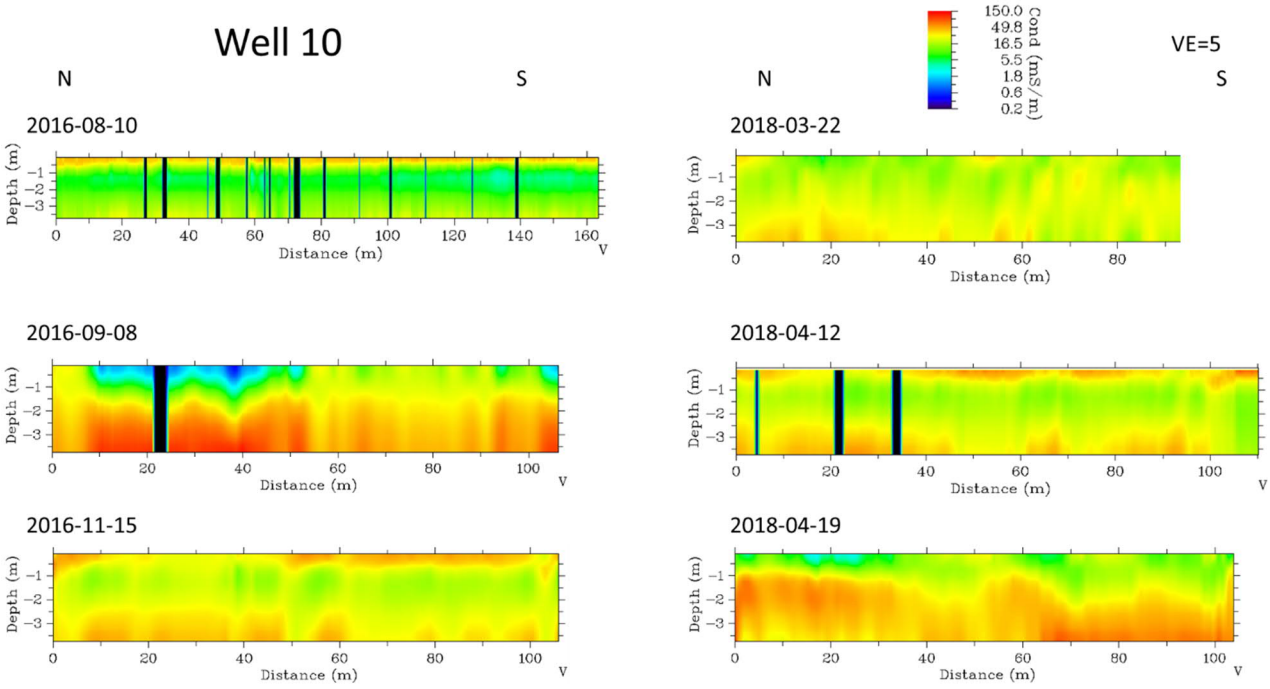
Bunny Hole

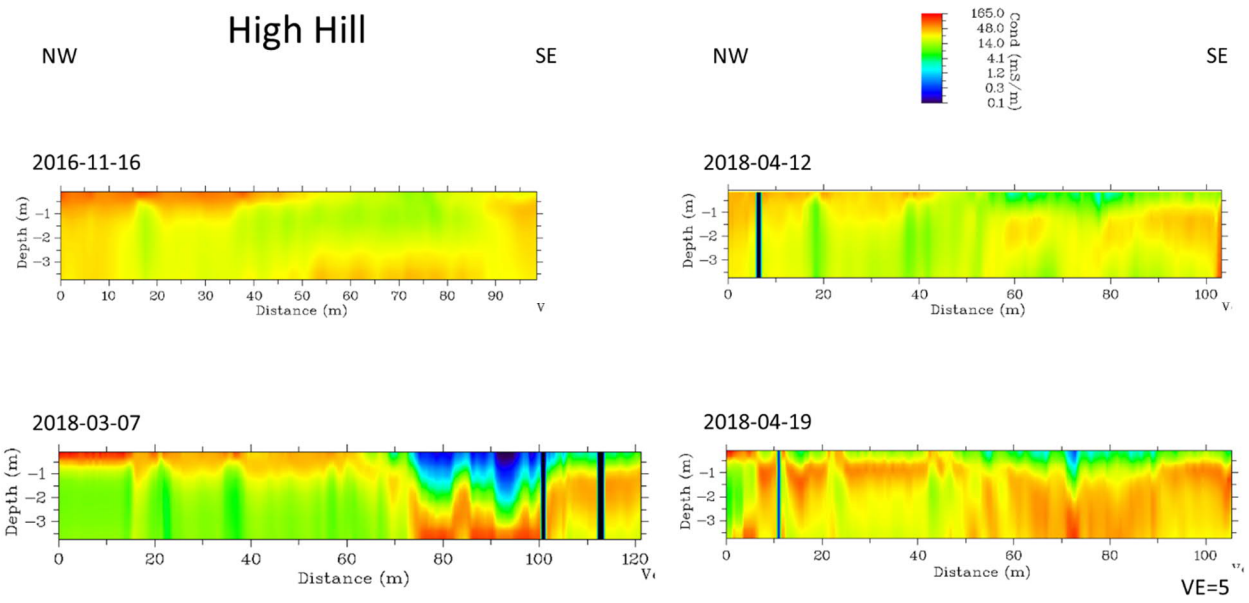


Salado




Water Tank
SSW
NNE
SSW
NNE
200.0
63.2
20.0
6.3
2.0
0.6
0.2
Cond (mS/m)
2016-08-10
2018-03-22
2016-09-08
2018-04-12
2016-11-14
2018-04-19
Depth (m)
Distance (m)
VE=5


Well 10
N
S
N
S
150.0
49.8
16.5
5.5
1.8
0.6
0.2
Cond (mS/m)
VE=5
2016-08-10
2018-03-22
2016-09-08
2018-04-12
2016-11-15
2018-04-19
Depth (m)
Distance (m)



6.2 Ground-penetrating Radar (GPR)

Ground-penetrating radar (GPR) with a 400-MHz antenna in constant offset was used to map soil thickness and common midpoint offset was used to accurately measure the depth to bedrock and soil dielectric permittivity (i.e., in situ, depth-integrated water content) at select locations. GPR uses an electromagnetic (EM) pulse to gather subsurface information (Knight, 2001). GPR data acquisition systems used the Subsurface Interface Radar-3000 System manufactured by Geophysical Survey Systems, Inc. (Figure A6-2).

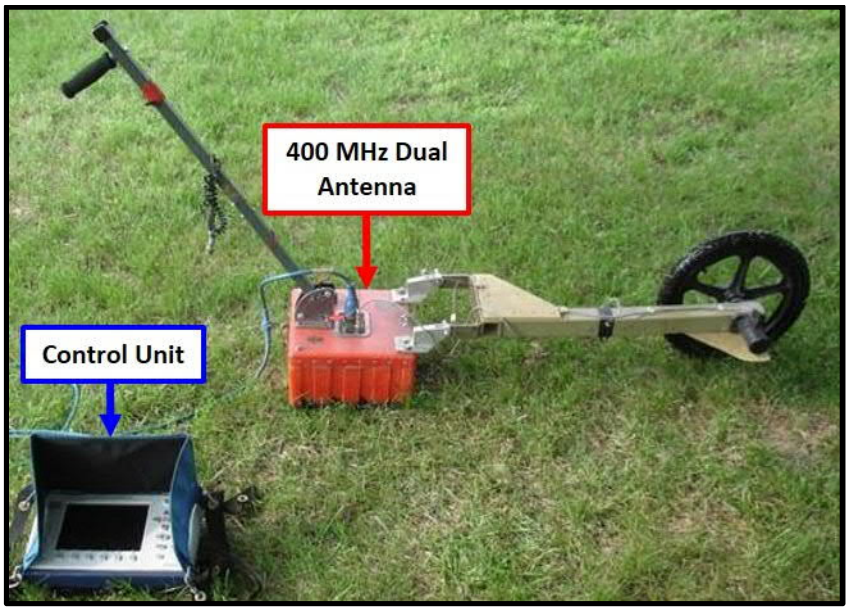


Figure A6-2: Setup of the GSSI SIR-3000 system with 400-MHz dual antenna and survey wheel used for this study.

GPR consists of transmitting and receiver antennae and a control unit that handles collection and setup parameters, data storage, and display. When the GPR is run along a survey line, EM pulses are radiated toward the ground by the transmission antenna (Figure A6-3). When the GPR encounters a point target or an interface with contrasting EM properties, part of the radar energy is reflected back to the receiver antenna, which then passes the radar energy to a recorder for the data to be displayed on the control unit. The time the pulse takes to travel down to a point target or an interface and return to the surface is known as the two-way travel time, which varies based on the propagation velocities within the subsurface materials as a function of their moisture content and resulting EM properties.

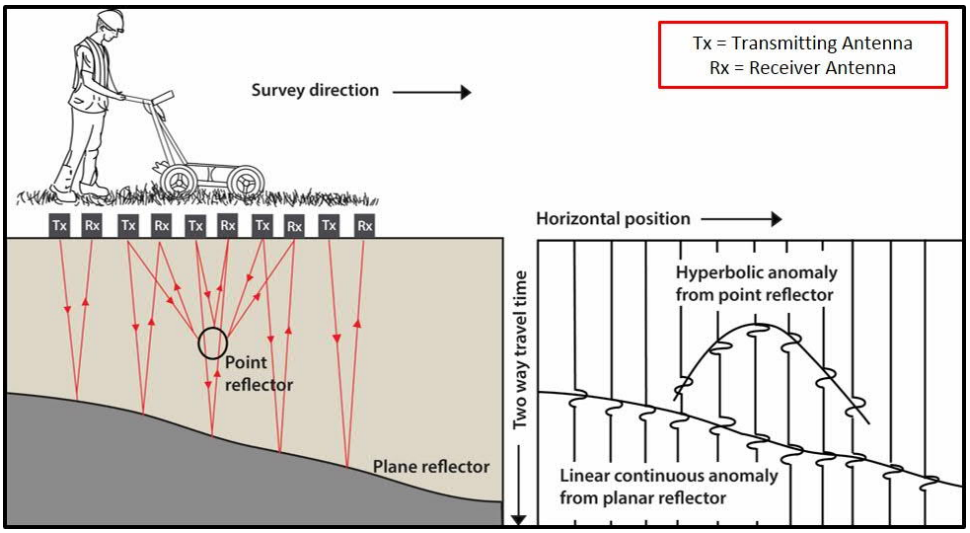


Figure A6-3: Working principle of the GPR (http://scantech.ie/scantech-about-gpr.html).

In fall 2016, GPR surveys were conducted during a relatively dry period (and following no major storm events) at four locations within Camp Bullis: Salado, High Hill, Acan, and Bunny Hole. Because site conditions have a strong control on data acquisition, overall open and relatively flat areas were targeted to minimize potential anthropogenic and topographic clutters. However, minor surficial dip changes and vegetation were encountered in places, adding to the noise in the data. At each site, GPR data was collected in both directions between the end points of each profile to ensure repeatability. The SIR-3000 system by GSSI was operated with a monostatic 400-MHz antenna. The antenna was chosen to balance the relatively shallow targeted depth of investigation and data resolution. The system was run in continuous mode and was configured for a sampling interval time window of 50 ns with 1024 samples per trace. Section lengths were 100 m with a trace spacing of 0.02 m. Considering the dielectric constant in limestone (the dominant subsurface lithology) in dry conditions is ~8 and GPR signal velocity is ~0.12 m/ns, subsurface penetration depths achieved in the field were estimated to be in the ~3.6 m range.

The GPR data collected along 2D profiles in the field at Camp Bullis was post-processed using matGPR software (v 3.1). The following steps were taken to process the raw data from each 2D GPR profile:

- Step 1: Move time-zero to adjust the vertical position of the surface reflection.
- Step 2: Trim time window to discard the late arrivals when they represented noisy data.
- Step 3: Remove DC component (arithmetic mean) from each trace of the GPR profile.

- Step 4: Equalize traces by making the sum of the absolute values of all samples in a trace the same for all traces.
- Step 5: Apply standard automatic gain control (AGC) to adjust gain as a function of the RMS amplitude computed over a sliding time window of 15 ns.
- Step 6: Remove global background trace, also known as stacking the data. This removes the background trace from the data to enhance coherent signal (horizontal banding) and reduces random noise (radar signal received from the subsurface).
- Step 7: Remove the foreground trace using a sliding window of width 15 to suppress dipping features.

See Figures A6-4 and A6-5 for examples of outcomes.

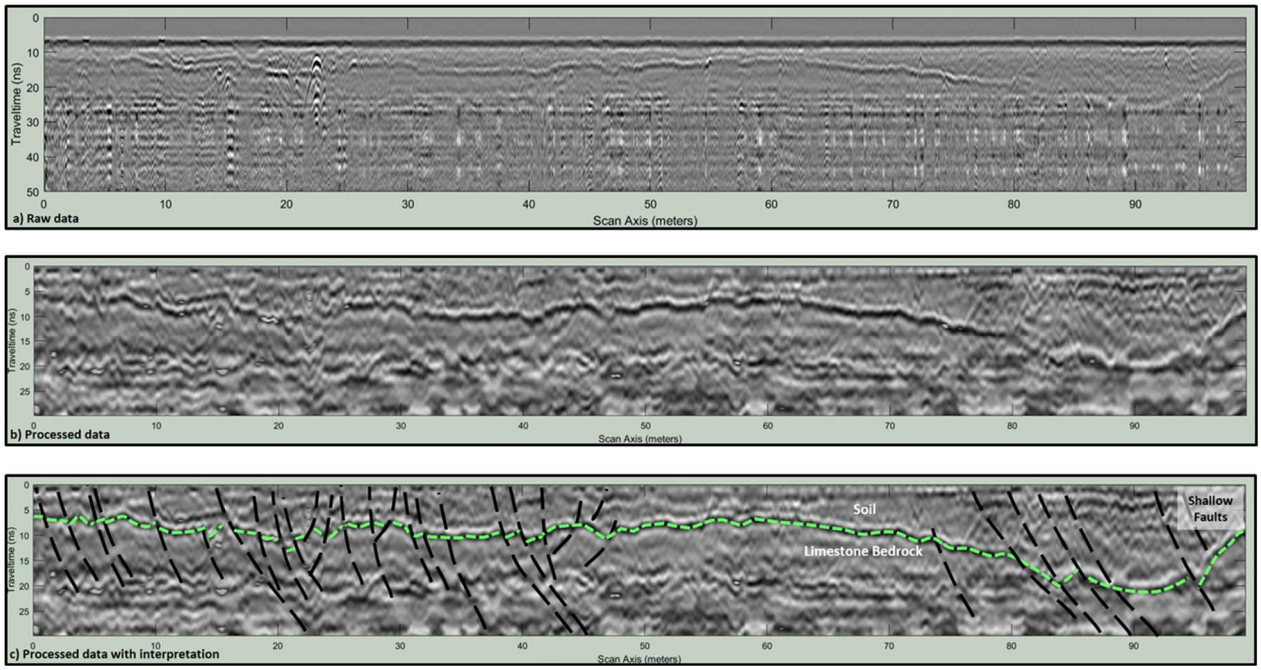


Figure A6-4: A good example showing a 2D GPR profile from the Salado location at Camp Bullis. Raw GPR data (top) was collected in the field, processed (middle), and interpreted. The geological interpretation overlain on the processed data shows a prominent undulating GPR reflector (in green). Based on the local geology, the GPR facies above the green interface is interpreted as soil, whereas the one below is interpreted as limestone bedrock (Woodruff et al., 2008). Offsets in reflectors are interpreted as shallow faults (black dotted lines), which are numerous along the profile.

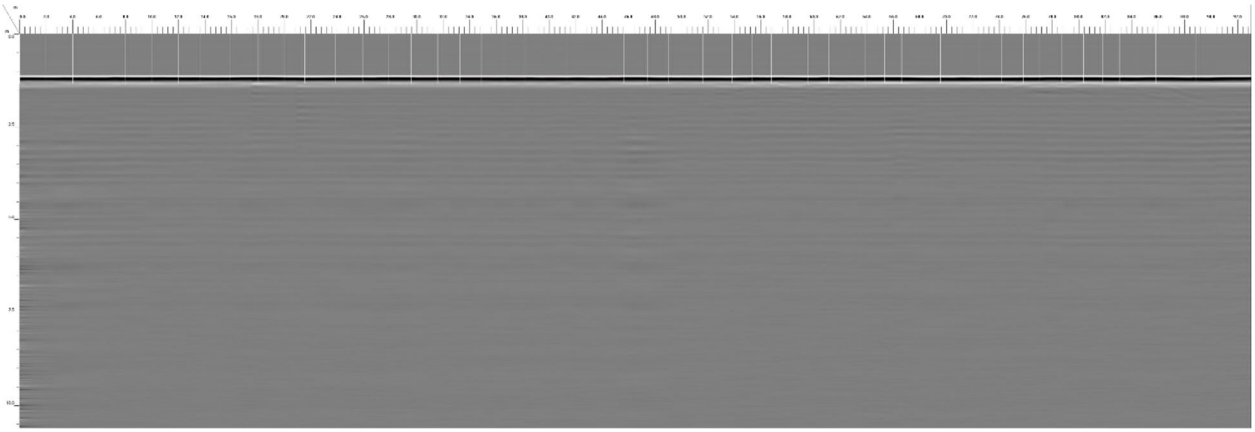


Figure A6-5: A more typical poor example at ACAN in which no reflectors are visible.

6.3 Time Domain Electromagnetic Induction (TEM)

EMI is a frequency-based EMI method, whereas TEM uses a static loop to record the time decay of an induced EM Signal. TEM transmits up to 3 amps into a 40-by-40-m transmitter loop with the receiver coil in the center (Figure A6-6). Measurements are sampled with 1 MHz and up to 200 gates to create soundings to ~100 m. We used a WalkTEM system from ABEM. Six sounds were collected in November 2016 (Figure A6-7). We present two cross sections trending east–west (Figure A6-8) and north–south (Figure A6-9).

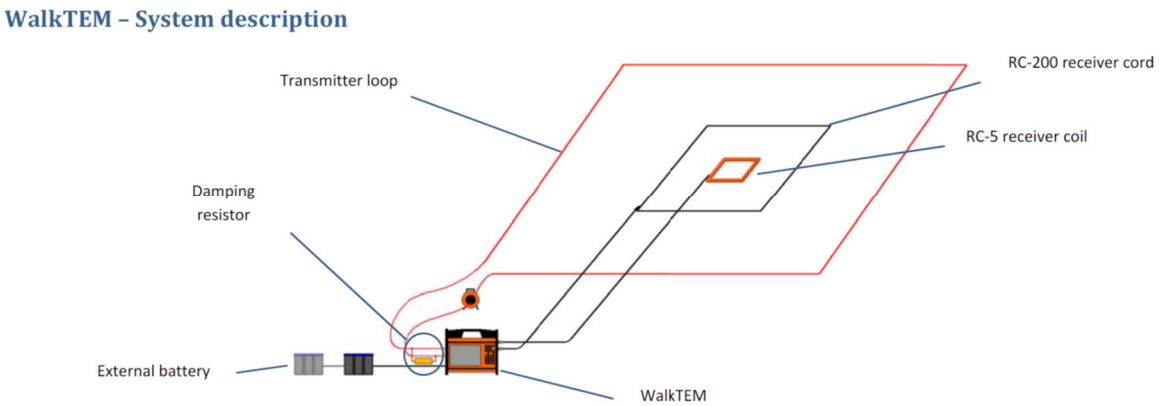


Figure A6-6: Basic system configuration, consisting of a transmitter loop, one or two receiver coils, an external battery with a series of dampeners, and the WalkTEM instrument.

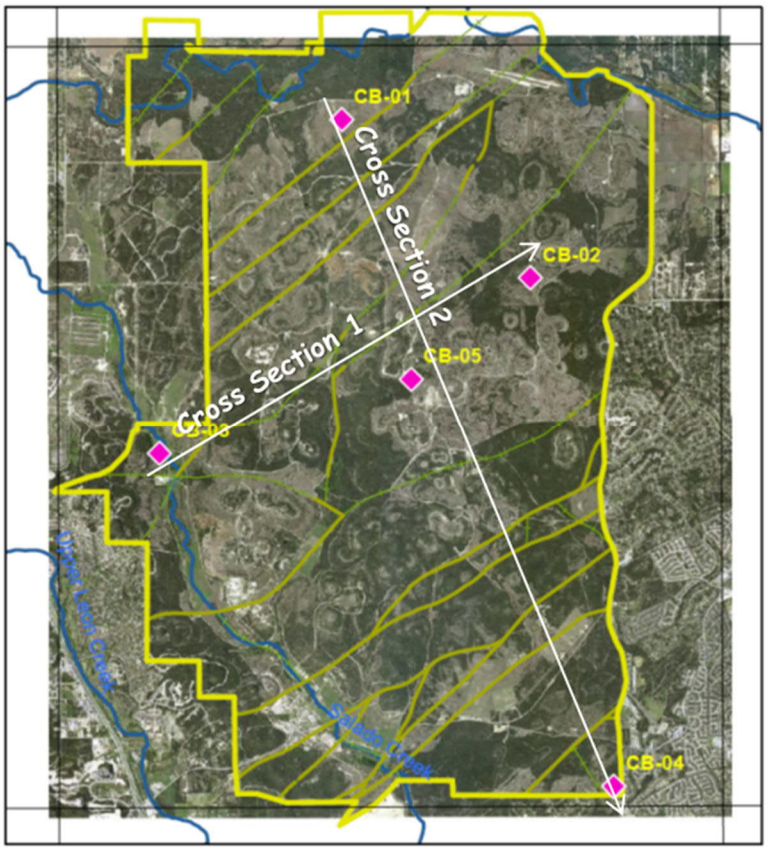


Figure A6-7: TEM measurement locations on Camp Bullis.

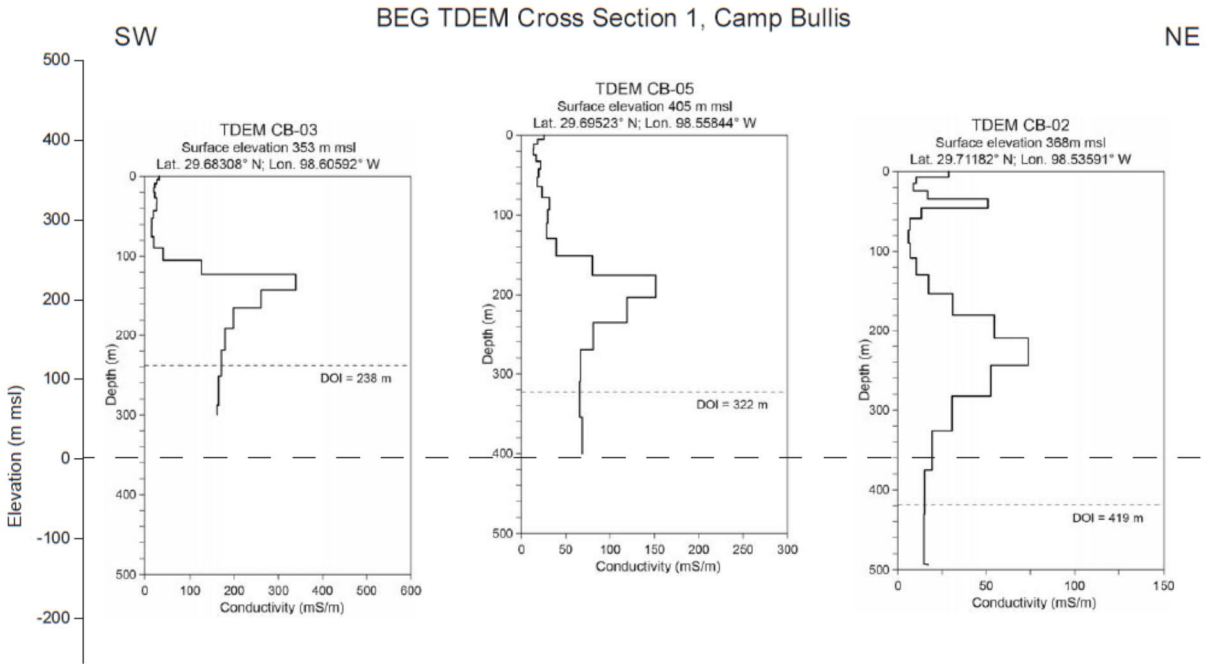


Figure A6-8: TEM sounding along cross section #1 from west to east.

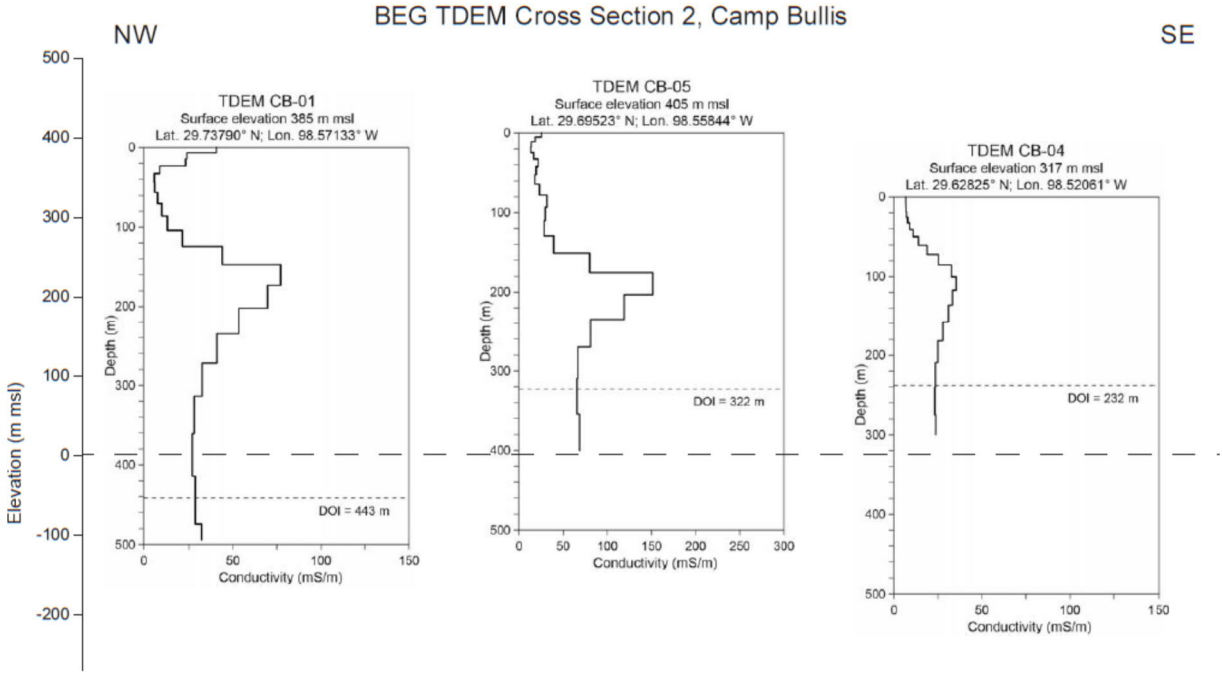


Figure A6-9: TEM sounding along cross section #2 from north to south.

7. Results of Statistical Analyses, using SigmaPlot (v13)

Linear Regression Wednesday, September 16, 2020, 3:38:58 PM

Data source: Data 1 in statistics.JNB

HYDRUS ET W = -26.085 + (1.460 * EC ET W)

N = 24

R = 0.897 Rsqr = 0.805 Adj Rsqr = 0.796

Standard Error of Estimate = 13.445

	Coefficient	Std. Error	t	P	Std. Coeff.
Constant	-26.085	9.717	-2.684	0.014	
EC ET W	1.460	0.153	9.519	<0.001	0.897

Analysis of Variance:

	DF	SS	MS	F	P
Regression	1	16379.302	16379.302	90.609	<0.001
Residual	22	3976.929	180.770		
Total	23	20356.231	885.054		

Normality Test (Shapiro-Wilk) Failed (P = 0.015)

Power of performed test with alpha = 0.050: 1.000

Linear Regression Friday, September 25, 2020, 3:00:51 PM

Data source: Data 1 in statistics.JNB

HYDRUS ET S = 28.553 + (0.583 * EC ET S)

N = 27

R = 0.525 Rsqr = 0.276 Adj Rsqr = 0.247

Standard Error of Estimate = 27.039

	Coefficient	Std. Error	t	P	Std. Coeff.
Constant	28.553	11.972	2.385	0.025	
EC ET S	0.583	0.189	3.085	0.005	0.525

Analysis of Variance:

	DF	SS	MS	F	P
Regression	1	6956.067	6956.067	9.515	0.005
Residual	25	18277.215	731.089		
Total	26	25233.282	970.511		

Normality Test (Shapiro-Wilk) Passed (P = 0.884)

Power of performed test with alpha = 0.050: 0.815

Linear Regression Wednesday, September 16, 2020, 3:39:45 PM

Data source: Data 1 in statistics.JNB

HYDRUS ET G = 7.670 + (0.643 * EC ET G)

N = 8

R = 0.794 Rsqr = 0.631 Adj Rsqr = 0.569

Standard Error of Estimate = 14.354

	Coefficient	Std. Error	t	P	Std. Coeff.
Constant	7.670	13.231	0.580	0.583	
EC ET G	0.643	0.201	3.200	0.019	0.794

Analysis of Variance:

	DF	SS	MS	F	P
Regression	1	2109.973	2109.973	10.240	0.019
Residual	6	1236.298	206.050		
Total	7	3346.271	478.039		

Normality Test (Shapiro-Wilk) Passed (P = 0.936)

Power of performed test with alpha = 0.050: 0.677

---

8. Supplemental Figures

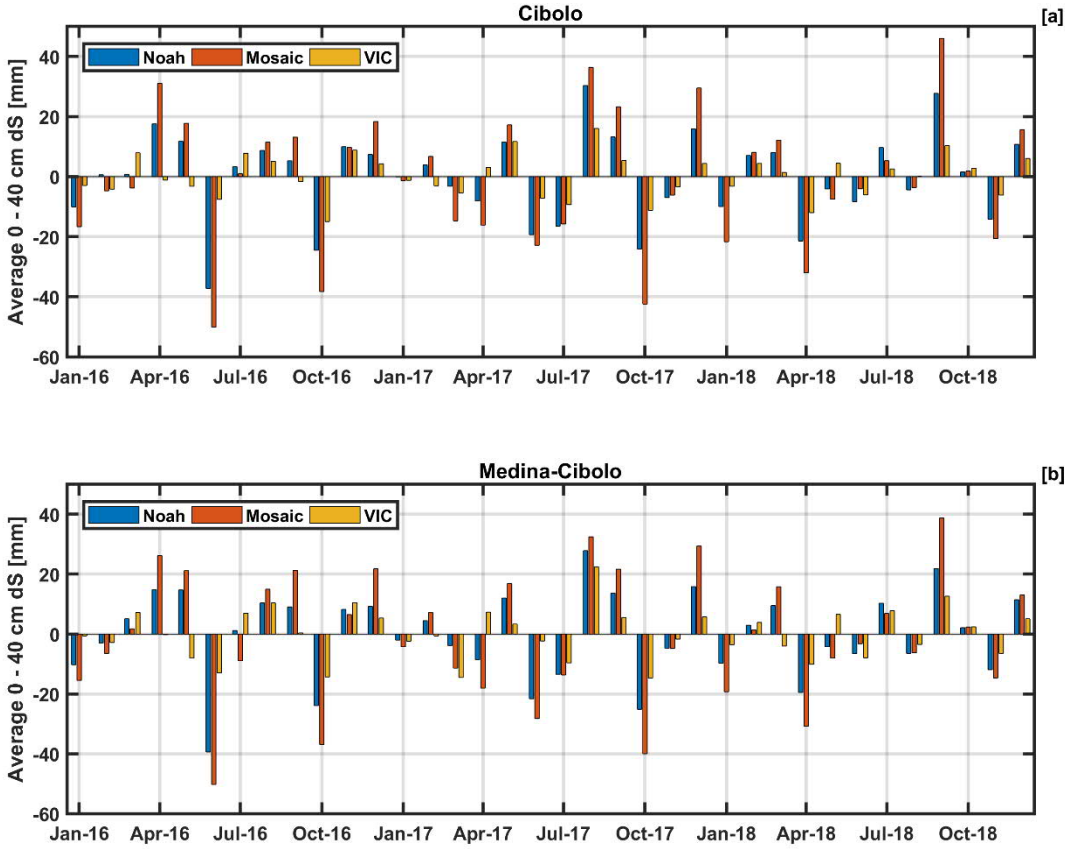


Figure SI1: Comparison of average monthly change in soil-water storage (dS) from the NLDAS-2 Noah, Mosaic, and VIC land-surface models for the [a] Cibolo and [b] Medina-Cibolo Basins.

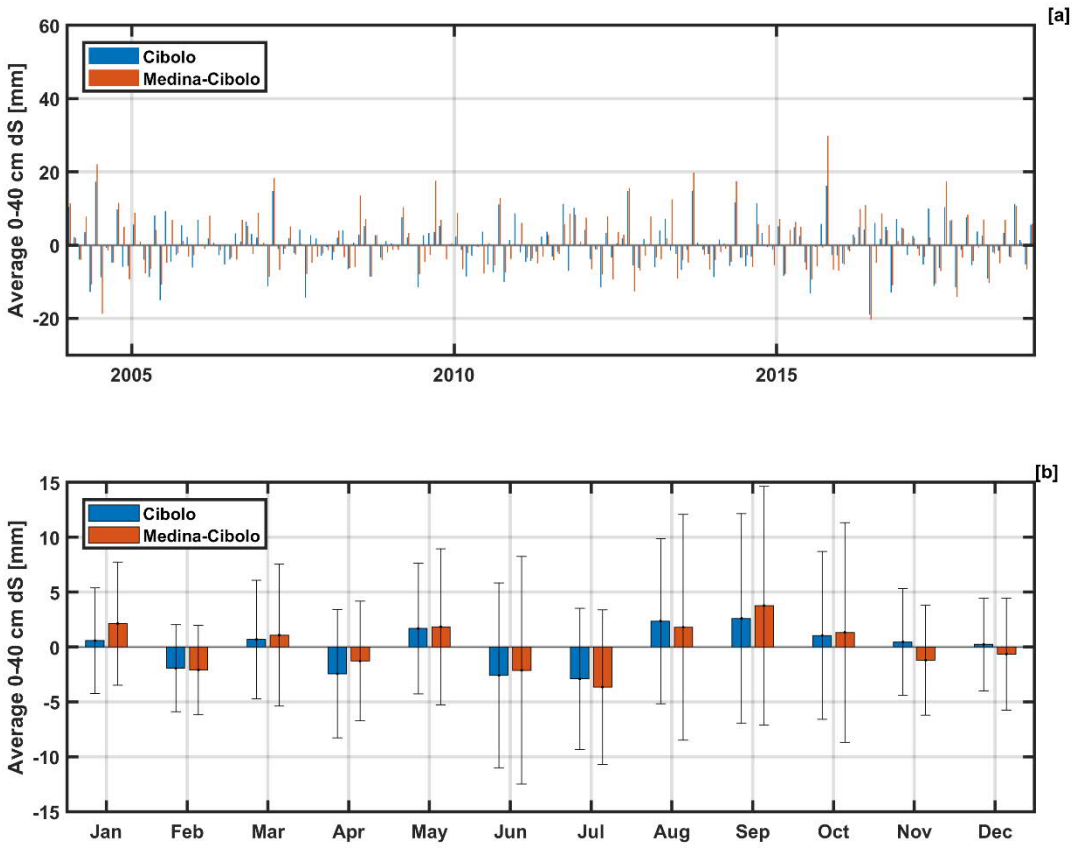


Figure SI2: Comparison of [a] monthly average and [b] monthly standard deviation of change in soil-water storage (dS) from NLDAS-2 Noah at 0–40 cm for the Cibolo and Medina-Cibolo Basins during the study period.

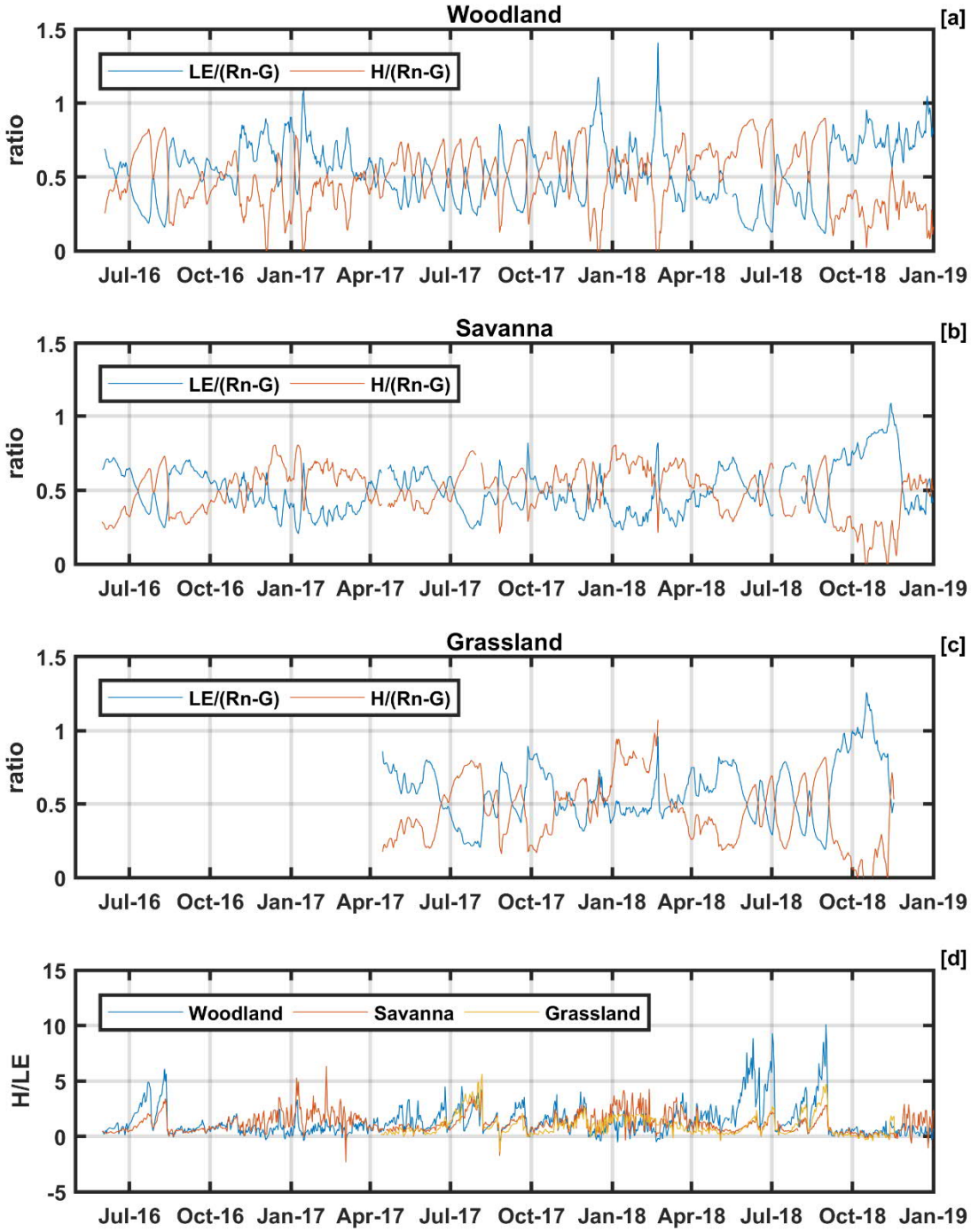


Figure SI3: Five-day running average of the ratios LE/(R_n – G) and H/(R_n – G) of the [a] Woodland, [b] Savanna, and [c] Grassland using the daily total fluxes.

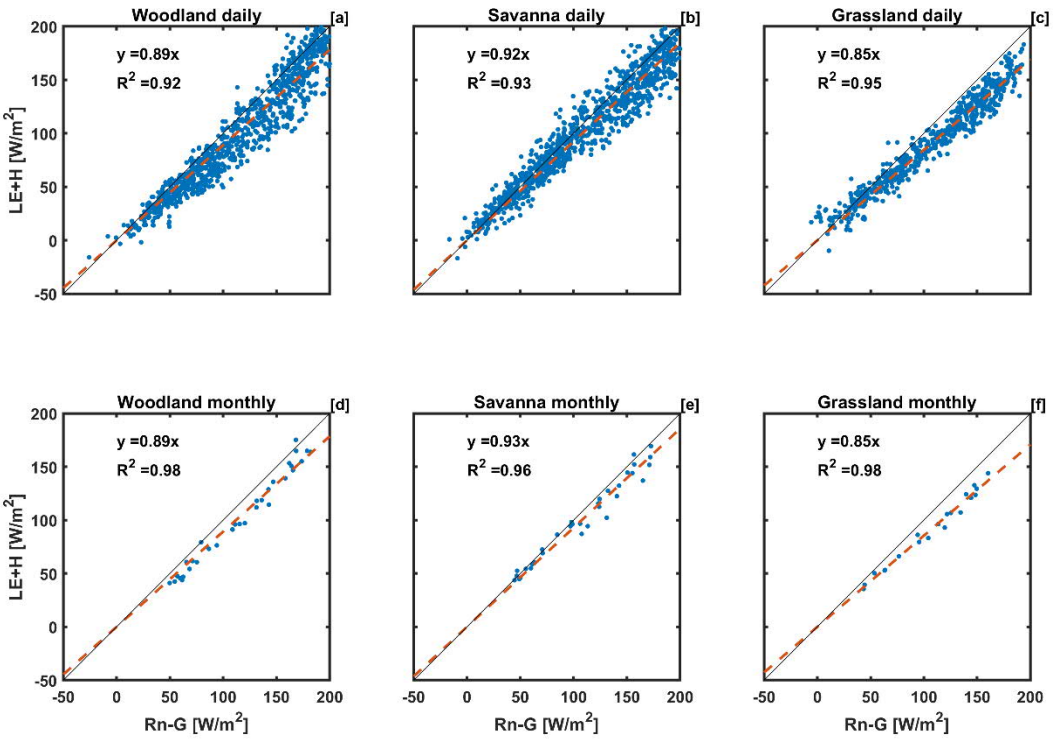


Figure SI4: Energy balance closure for [a–c] daily averages and [d–f] monthly averages for the [a and d] Woodland, [b and e] Savanna, and [c and f] Grassland.

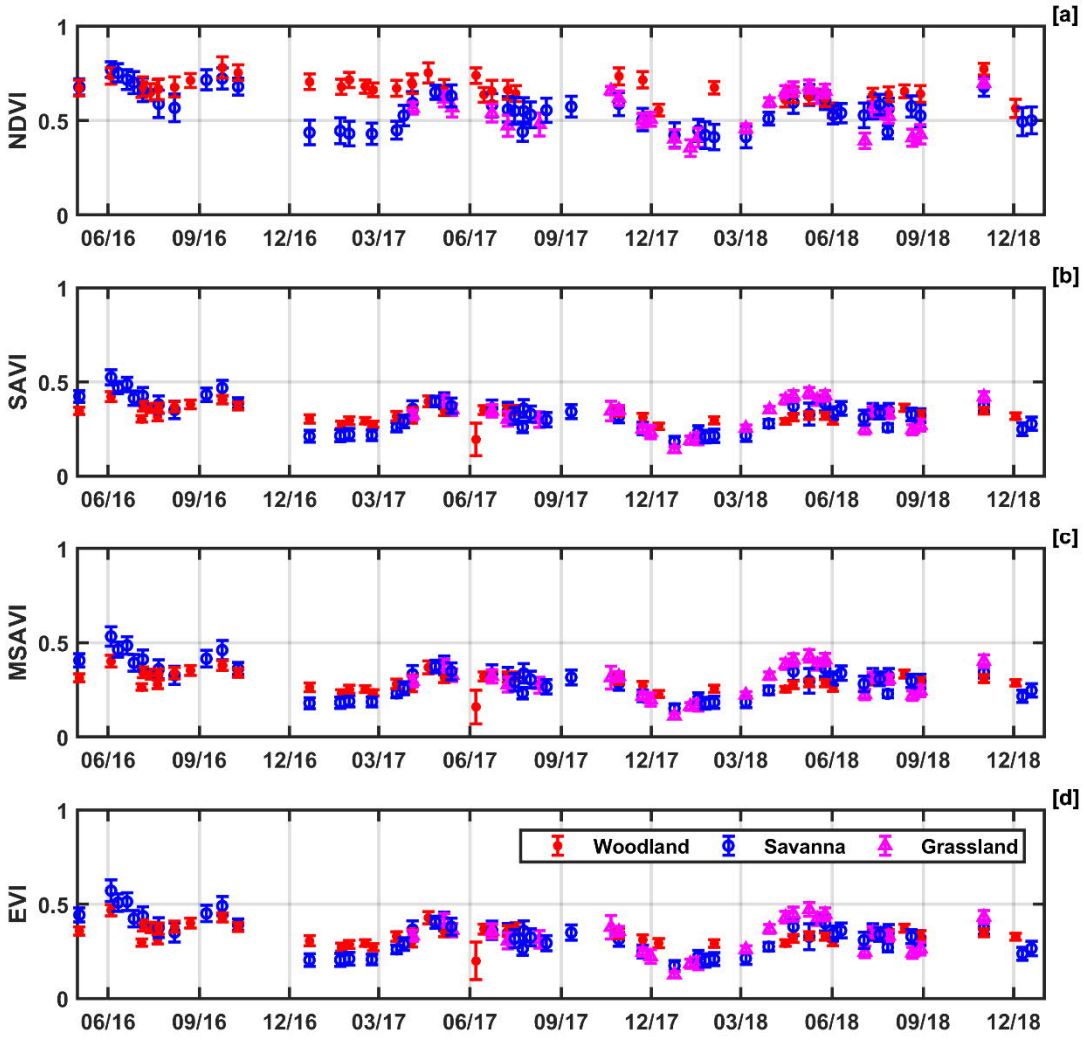


Figure SI5: Time series from Landsat for the vegetation indices ([a] normalized difference vegetation index (NDVI), [b] soil adjusted vegetation index (SAVI), [c] modified soil adjusted vegetation index [MSAVI], and [d] enhanced vegetation index [EVI]) for the three eddy covariance locations during the study period.

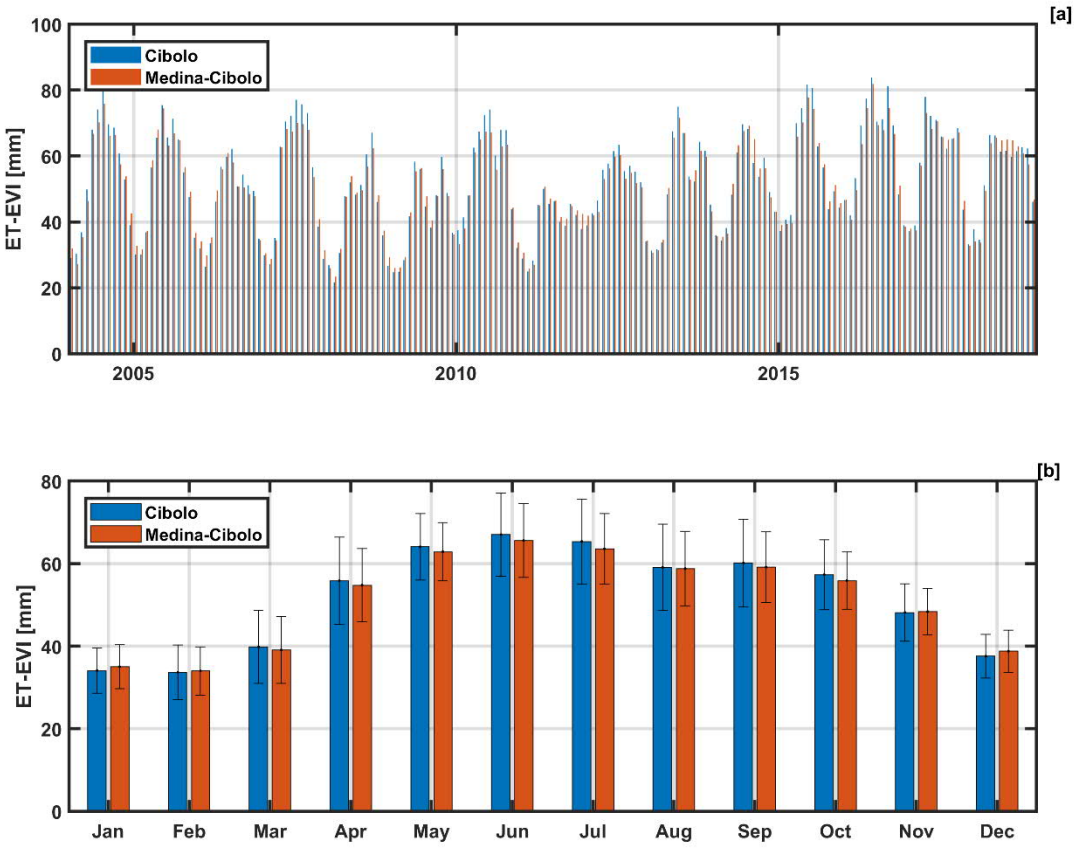


Figure SI6: Comparison of [a] average monthly and [b] monthly average and standard deviation of evapotranspiration (ET) from ET-EVI for the Cibolo and Medina-Cibolo Basins during the study period.

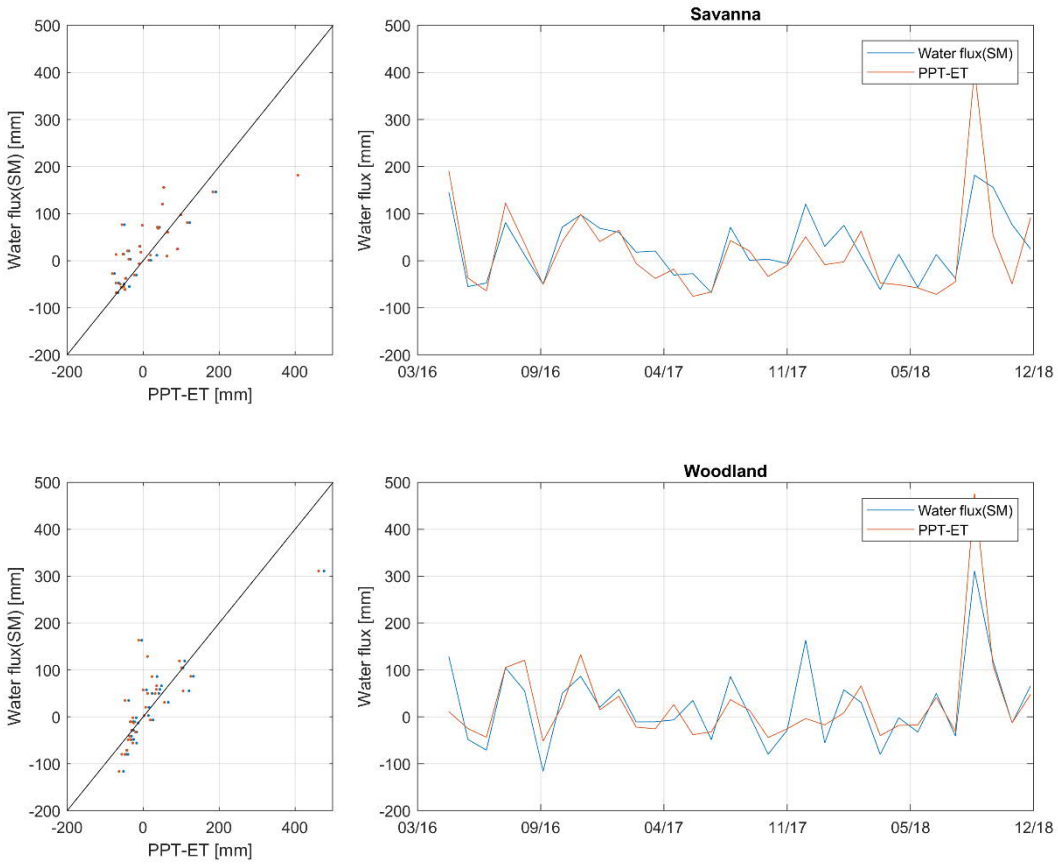


Figure SI7: Monthly water flux derived from soil moisture vs PPT-ET.

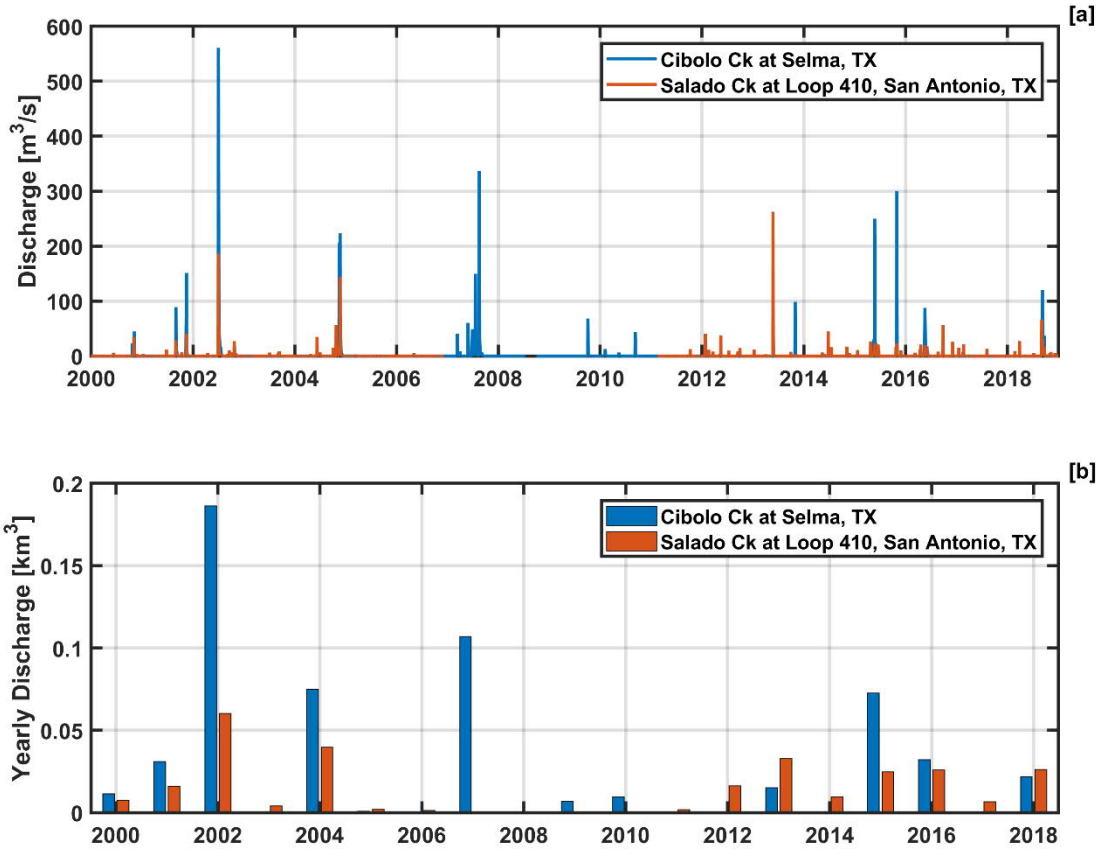


Figure SI8: [a] Daily and [b] yearly discharge at the outlet for the Cibolo and Medina-Cibolo Basins.

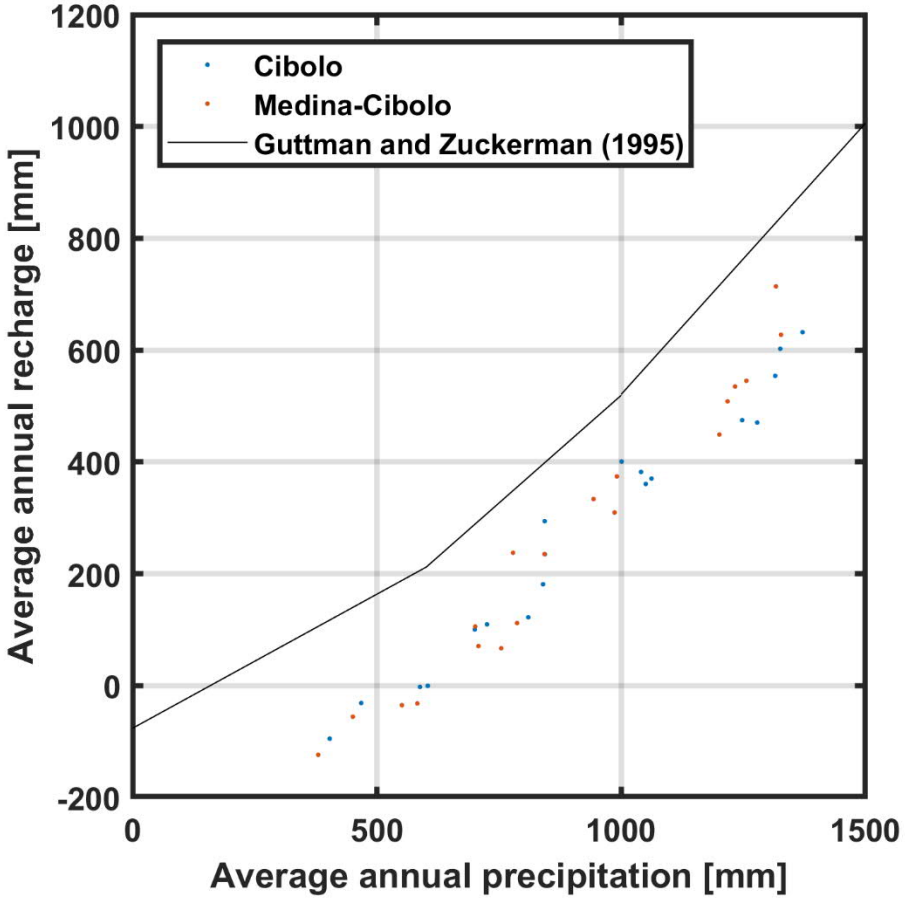


Figure SI9: Average annual recharge vs average annual precipitation of Cibolo and Medina-Cibolo Basins compared to the Guttman and Zuckerman (1995) model.

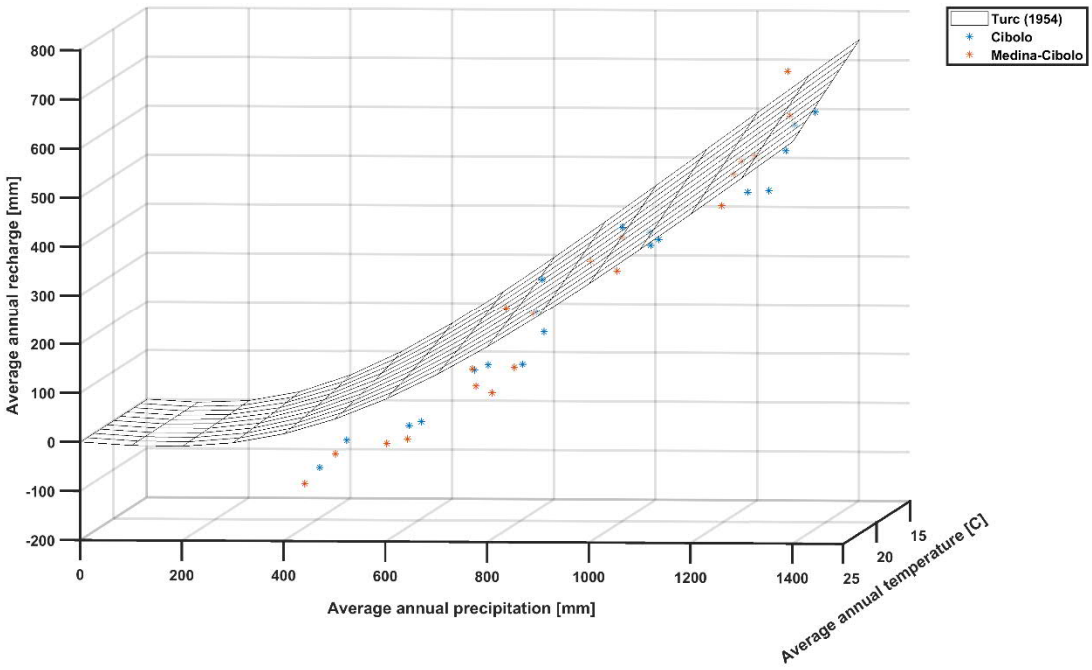


Figure SI10: Average annual recharge vs average annual precipitation vs average annual temperature of Cibolo and Medina-Cibolo Basins compared to the Turc (1954) model.

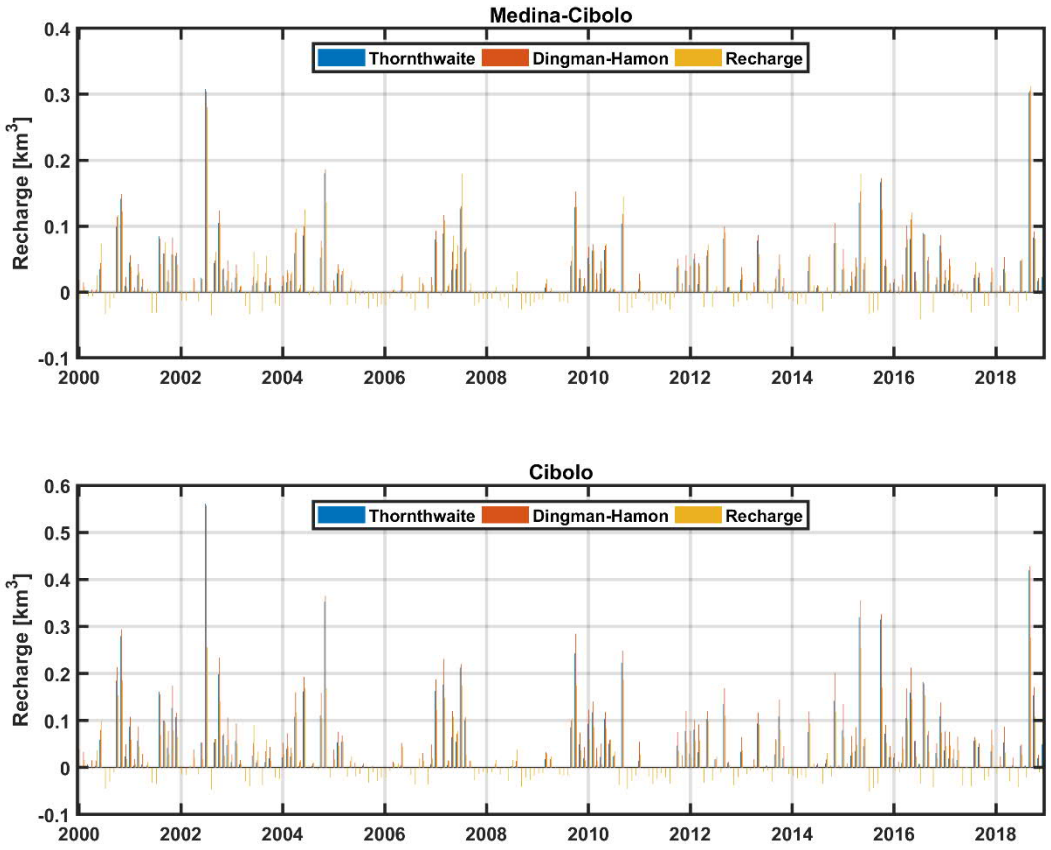


Figure SI11: Monthly recharge comparison between models in ESPERE and the recharge found in this study.

9. References

Abatzoglou, J.T. 2013. Development of gridded surface meteorological data for ecological applications and modelling. International Journal of Climatology, 33, 121–131. doi:10.1002/joc.3413.

Abdu, H., D.A. Robinson, and S.B. Jones. 2007. Comparing bulk soil electrical conductivity determination using the DUALEM-1S and EM38-DD electromagnetic induction instruments. Soil Science Society of America Journal, 71, 189–196.

Andreo, B., J. Vías, J. Durán, P. Jiménez, J. López-Geta, and F. Carrasco. 2008. Methodology for groundwater recharge assessment in carbonate aquifers: application to pilot sites in southern Spain. Hydrogeology Journal, 16, 911–925.

Blanken, P., T.A. Black, P. Yang, H. Neumann, Z. Nesic, R. Staebler, G. Den Hartog, M. Novak, and X. Lee. 1997. Energy balance and canopy conductance of a boreal aspen forest: partitioning overstory and understory components. Journal of Geophysical Research: Atmospheres, 102, 28,915–28,927.

Celia, M.A., E.T. Bouloutas, and R.L. Zarba. 1990. A general mass-conservative numerical solution for the unsaturated flow equation. Water Resources Research, 26, 1483–1496.

Chaney, N.W., B. Minasny, J.D. Herman, T.W. Nauman, C.W. Brungard, C.L.S. Morgan, A.B. McBratney, E.F. Wood, and Y. Yimam. 2019. POLARIS Soil Properties: 30-m Probabilistic Maps of Soil Properties Over the Contiguous United States. Water Resources Research, 55, 2916–2938. doi:10.1029/2018WR022797.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.L. Pan, V. Koren, Q.Y. Duan, M. Ek, and A. Betts. 1996. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. Journal of Geophysical Research: Atmospheres, 101, 7251–7268.

Dingman, S.L. 2002. *Physical Hydrology*. Prentice Hall: Upper Saddle River, NJ. 575 p.

Ek, M., K. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. Tarpley. 2003a. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. Journal of Geophysical Research: Atmospheres, *108*.

Ek, M.B., K.E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.D. Tarpley. 2003b. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. Journal Of Geophysical Research-Atmospheres, 108, 8851. doi:10.1029/2002jd003296.

Feddes, R.A., P.J. Kowalik, and H. Zaradny. 1978. Simulation of Field Water Use and Crop Yield. John Wiley and Sons: New York, NY. 189 p.

Foken, T., R. Leuning, S. R. Oncley, M. Mauder, and M. Aubinet. 2012a. Corrections and data quality control, in *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*, edited by M. Aubinet, T. Vesala and D. Papale. Springer, Netherlands. pp. 85–131.

Foken, T., R. Leuning, S. R. Oncley, M. Mauder, and M. Aubinet. 2012b. Corrections and data quality control. In *Eddy covariance*. Springer. pp. 85–131.

Friedl, M.A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang. 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114, 168–182, doi:10.1016/j.rse.2009.08.016.

Guttman, J., and H. Zuckerman. 1995. Flow model in the eastern basin of the Judea and Samaria hills, Tahal Consulting Engineers Ltd. Report 01/95/66. Tel Aviv, Israel. 33 p.

Horst, T. W., and D. H. Lenschow. 2009. Attenuation of Scalar Fluxes Measured with Spatially-displaced Sensors. Boundary-Layer Meteorology, 130, 275–300.

Huete, A. R. 1988. A soil-adjusted vegetation index (SAVI), *Remote sensing of environment*, *25*, 295-309.

Kaimal, J., S. Clifford, and R. Lataitis. 1989. Effect of finite sampling on atmospheric spectra. Boundary-Layer Meteorology, 47, 337–347.

Kemp, P.R., J.F. Reynolds, Y. Pachepsky, and J.-L. Chen. 1997. A comparative modeling study of soil water dynamics in a desert ecosystem. Water Resources Research, 33, 73–90. doi:10.1029/96wr03015.

Kessler, H. 1967. Water balance investigations in the karstic regions of Hungary, AIHS-UNESCO. *Act Coll Dubrovnik*, Paris, France. pp. 91–105.

Kljun, N., P. Calanca, M.W. Rotach, and H.P. Schmid. 2004. A simple parameterisation for flux footprint predictions. Boundary-Layer Meteorology, 112, 503–523.

Knight, R. 2001. Ground penetrating radar for environmental applications. Annual Review of Earth and Planetary Sciences, 29, 229–255.

Kormann, R., and F.X. Meixner. 2001. An analytical footprint model for non-neutral stratification. Boundary-Layer Meteorology, 99, 207–224.

Koster, R.D., and M.J. Suarez. 1992. Modeling the land surface boundary in climate models as a composite of independent vegetation stands. Journal of Geophysical Research: Atmospheres, 97, 2697–2715.

Lee, X. 1998. On micrometeorological observations of surface-air exchange over tall vegetation. Agricultural and Forest Meteorology, 91, 39–49.

Liang, X., D.P. Lettenmaier, E.F. Wood, and S.J. Burges. 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99, 14,415–14,428.

Liu, H.Q., and A. Huete.1995. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Transactions on Geoscience and Remote Sensing, 33, 457–465.

Loveland, T.R., and A.S. Belward. 1997. The IGBP-DIS global 1km land cover data set, DISCover: First results. International Journal of Remote Sensing, 18, 3289–3295. doi:10.1080/014311697217099.

Moncrieff, J., R. Clement, J. Finnigan, and T. Meyers. 2004. Averaging, detrending, and filtering of eddy covariance time series. In *Handbook of micrometeorology*. Springer. pp. 7–31.

Moncrieff, J.B., J. Massheder, H. De Bruin, J. Elbers, T. Friborg, B. Heusinkveld, P. Kabat, S. Scott, H. Soegaard, and A. Verhoef. 1997a. A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. Journal of Hydrology, 188, 589–611.

Moncrieff, J.B., J.M. Massheder, H. deBruin, J. Elbers, T. Friborg, B. Heusinkveld, P. Kabat, S. Scott, H. Soegaard, and A. Verhoef. 1997b. A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. Journal of Hydrology, 189, 589–611.

Monteiro Santos, F.A., J. Triantafilis, and K. Bruzgulis. 2011. A spatially constrained 1D inversion algorithm for quasi-3D conductivity imaging: Application to DUALEM-421 data collected in a riverine plain. Geophysics, 76, B43-B53, doi:10.1190/1.3537834.

Moore, C. J. 1986. Frequency-Response Corrections for Eddy-Correlation Systems. Boundary-Layer Meteorology, 37, 17–35.

Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resources Research, 12, 513–522.

Myneni, R.B., Y. Knyazikhin, and T. Park. 2015. MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006, edited, NASA EOSDIS Land Processes DAAC.

Papale, D., M. Reichstein, M. Aubinet, E. Canfora, C. Bernhofer, W. Kutsch, B. Longdoz, S. Rambal, R. Valentini, and T. Vesala. 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences, 3, 571–583.

Qi, J., A. Chehbouni, A. Huete, Y. Kerr, and S. Sorooshian. 1994. A modified soil adjusted vegetation index. Remote sensing of environment, 48, 119–126.

Reichstein, M., E. Falge, D. Baldocchi, D. Papale, M. Aubinet, P. Berbigier, C. Bernhofer, N. Buchmann, T. Gilmanov, and A. Granier. 2005. On the separation of net ecosystem exchange into assimilation

and ecosystem respiration: review and improved algorithm. Global Change Biology, 11, 1424–1439.

Ritchie, J. T. 1972. Model for predicting evaporation from a row crop with incomplete cover. Water Resources Research, 8, 1204–1213.

Robinson, D.A., H. Abdu, I. Lebron, and S.B. Jones. 2012. Imaging of hill-slope soil moisture wetting patterns in a semi-arid oak savanna catchment using time-lapse electromagnetic induction. Journal of Hydrology, 416, 39–49, doi:10.1016/j.jhydrol.2011.11.034.

Robinson, D.A., H. Abdu, S.B. Jones, M. Seyfried, I. Lebron, and R. Knight. 2008. Eco-Geophysical Imaging of Watershed-Scale Soil Patterns Links with Plant Community Spatial Patterns. Vadose Zone Journal, 7, 1132–1138, doi:10.2136/vzj2008.0101.

Rouse Jr, J.W., R. Haas, J. Schell, and D. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS.

Santos, F. A. M., J. Triantafilis, K. E. Bruzgulis, and J. A. E. Roe. 2010. Inversion of Multiconfiguration Electromagnetic (DUALEM-421) Profiling Data Using a One-Dimensional Laterally Constrained Algorithm. Vadose Zone Journal, 9, 117–125. doi:10.2136/vzj2009.0088.

Schotanus, P., F.T.M. Nieuwstadt, and H.A.R. Debruin. 1983. Temperature-Measurement with a Sonic Anemometer and Its Application to Heat and Moisture Fluxes. Boundary-Layer Meteorology, 26, 81–93.

Simunek, J., M.T. van Genuchten, and M. Sejna. 2012. Hydrus: Model Use, Calibration, and Validation. Transactions of the ASABE, 55, 1261–1274.

Simunek, J., M.T. van Genuchten, and M. Sejna. 2016. Recent Developments and Applications of the HYDRUS Computer Software Packages. Vadose Zone Journal, 15. doi:10.2136/vzj2016.04.0033.

Soil Survey Staff. 2014. Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United State, edited, United States Department of Agriculture, Natural Resources Conservation Service.

Tanner, C.B., and G.W. Thurtell. 1969. Anemoclinometer measurements of Reynolds stress and heat transport in the atmospheric surface layer. WI Final Oct 1965-Jan 1969. Univ. Wisconsin-Madison: Madison, 200 p.

Thornthwaite, C. W. (1948. An Approach toward a Rational Classification of Climate, Geographical Review, 38, 55–94.

Turc, L. 1954), Le bilan d'eau des sols: Relations entre les précipitations, l'évaporation et l'écoulement, *Annales Agronomiques*, *5*, 491–595.

Twine, T.E., W. Kustas, J. Norman, D. Cook, P. Houser, T. Meyers, J. Prueger, P. Starks, and M. Wesely. 2000. Correcting eddy-covariance flux underestimates over a grassland. Agricultural and Forest Meteorology, 103, 279–300.

van Dijk, A. 2002. Extension to 3D of "The effect of line averaging on scalar flux measurements with a sonic anemometer near the surface" by Kristensen and Fitzjarrald. Journal of atmospheric and oceanic technology, 19, 80–82.

Van Dijk, A., A. Moene, and H. De Bruin. 2004. The principles of surface flux physics: theory, practice and description of the ECPACK library. Internal Rep, 1, 99.

Vickers, D., and L. Mahrt. 1997. Quality control and flux sampling problems for tower and aircraft data. Journal of Atmospheric and Oceanic Technology, 14, 512–526.

Vogel, T., M.T. van Genuchten, and M. Cislerova. 2000. Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions. Advances in Water Resources, 24, 133–144.

Vrugt, J.A., J.W. Hopmans, and J. Simunek. 2001. Calibration of a two-dimensional root water uptake model. Soil Science Society of America Journal, 65, 1027–1037.doi:10.2136/sssaj2001.6541027x.

Webb, E.K., G.I. Pearman, and R. Leuning. 1980. Correction of flux measurements for density effects due to heat and water vapour transfer. Quarterly Journal of the Royal Meteorological Society, 106, 85–100.

Wilczak, J.M., S.P. Oncley, and S.A. Stage. 2001. Sonic Anemometer Tilt Correction Algorithms. Boundary-Layer Meteorology, 99, 127–150. doi:10.1023/a:1018966204465.

Wilson, K., A. Goldstein, E. Falge, M. Aubinet, D. Baldocchi, P. Berbigier, C. Bernhofer, R. Ceulemans, H. Dolman, and C. Field. 2002. Energy balance closure at FLUXNET sites. Agricultural and Forest Meteorology, 113, 223–243.

Wood, E.F., D. Lettenmaier, X. Liang, B. Nijssen, and S.W. Wetzel. 1997. Hydrological modeling of continental-scale basins. Annual Review of Earth and Planetary Sciences, 25, 279–300.

Xia, Y., J. Sheffield, M.B. Ek, J. Dong, N. Chaney, H. Wei, J. Meng, and E.F. Wood. 2014. Evaluation of multi-model simulated soil moisture in NLDAS-2. Journal of Hydrology, 512, 107–125.