

# HydroBio

Advanced Remote Sensing

CONTRACT ADMINISTRATION

2012 AUG 21 PM 2:50



## GROUND WATER RECHARGE MODEL: CALIBRATION AND VALIDATION FOR REMOTELY-SENSED DUAL COEFFICIENT (RDC) GMA8 ET ESTIMATION

TWDB CONTRACT NO. 1004831114

FOR: THE TEXAS WATER DEVELOPMENT BOARD  
AUGUST 9, 2012



ALEXANDRA KIRK, DAVID GROENEVELD, DAVE BARZ, JOANNA HERRING, AND SCOTT MURRAY



### Final Report

# HydroBio

Advanced Remote Sensing



## **GROUND WATER RECHARGE MODEL: CALIBRATION AND VALIDATION FOR REMOTELY-SENSED DUAL COEFFICIENT (RDC) GMA8 ET ESTIMATION**

**TWDB CONTRACT No. 1004831114**

FOR: THE TEXAS WATER DEVELOPMENT BOARD  
*AUGUST 9, 2012*



ALEXANDRA KIRK, DAVID GROENEVELD, DAVE BARZ, JOANNA HERRING, AND SCOTT MURRAY





# HydroBio

Advanced Remote Sensing

## **Ground Water Recharge Model: Calibration and Validation for Remotely-Sensed Dual Coefficient (RDC) GMA8 ET Estimation**

TWDB Contract No. 1004831114

Prepared For: The Texas Water Development Board  
August 9, 2012

Prepared by: Alexandra Kirk, David Groeneveld, Dave Barz, Joanna Herring, and Scott Murray



# TABLE OF CONTENTS

<b>ABSTRACT</b> .....	<b>2</b>
<b>1. GMA8 SITE DESCRIPTION</b> .....	<b>2</b>
<b>2. RELATIONAL GROUNDWATER RECHARGE MODEL CONCEPT</b> .....	<b>3</b>
<b>3. EXECUTIVE SUMMARY</b> .....	<b>4</b>
<b>3. ACKNOWLEDGEMENTS</b> .....	<b>8</b>
<b>A1. DRIVING MODEL VARIABLES</b> .....	<b>A2</b>
<b>A2. MODELING PRECIPITATION</b> .....	<b>A2</b>
A2.1. 50-YEAR PRECIPITATION RECORD .....	A3
A2.2. PRISM VERSUS RAINFALL KRIGED FROM TEXAS DIGITAL CLIMATE ATLAS.....	A3
<b>A3. MODELING ET<sub>0</sub></b> .....	<b>A5</b>
A3.1 RECENT METEOROLOGIC RECORD .....	A6
A3.2 OPTIMIZATION OF PRECIPITATION-ET <sub>0</sub> MODEL .....	A7
A3.3. MODEL IMPLEMENTATION & TESTING.....	A9
<b>A4. SUMMARY</b> .....	<b>A10</b>
<b>B. SITE CHARACTERIZATION: GAGED WATERSHEDS AND SOIL PERMEABILITY</b> .....	<b>B2</b>
<b>B1. PERMEABILITY DATASET</b> .....	<b>B2</b>
B1.1 ANALYSIS OF SOILS DATA IN STATSGO .....	B3
B1.2 REDUCTION OF MULTI-LAYER SOIL DATABASE TO A SINGLE REPRESENTATIVE LAYER .....	B4
B1.3 AWP FOR SHALLOW SOILS AND/OR HIGHLY VARIABLE PERMEABILITY .....	B5
B1.4 PROOFING STATSGO.....	B6
B1.4.1 ANALYSIS OF STATSGO DATA—TRINITY AQUIFER .....	B7
B1.4.2. ANALYSIS OF STATSGO DATA—EDWARDS-TRINITY (PLATEAU).....	B8
B1.4.3. ANALYSIS OF STATSGO DATA—EDWARDS BALCONES FAULT ZONE .....	B9
B1.4.4 ANALYSIS OF STATSGO DATA—NACATOCH AQUIFER .....	B10
B.1.5. ANALYSIS OF STATSGO DATA—WOODBINE AQUIFER.....	B11
B.1.6. ANALYSIS OF STATSGO DATA—BLOSSOM AQUIFER .....	B12
B.1.7. ANALYSIS OF STATSGO DATA—PALEOZOIC AQUIFERS.....	B13
<b>B2. INDIVIDUAL WATERSHED Q DATA, THE BASIS FOR CALIBRATING FOR PERM</b> .....	<b>B14</b>
B2.1. SELECTING REPRESENTATIVE WATERSHEDS .....	B14
B2.2. PROCESSING INDIVIDUAL WATERSHED DATA.....	B15
<b>B3. SUMMARY</b> .....	<b>B18</b>
<b>C. CALIBRATING R<sub>INFIL</sub>: SPATIAL EXTRAPOLATION OF DISCHARGE</b> .....	<b>C2</b>
<b>C1. ESTIMATING R<sub>INFIL</sub>—WHAT CONSTITUTES “SURFACE WATER” LOSS</b> .....	<b>C2</b>
<b>C2. CODIFYING Q RELATIONSHIPS</b> .....	<b>C4</b>
C2.1. LINEARIZATION OF Q DATA .....	C5
C2.2. Q RELATIONSHIPS DEFINED BY PPT/ET <sub>0</sub> .....	C7
<b>C3. MULTIPLE LINEAR REGRESSION OF Q</b> .....	<b>C8</b>
C3.1 MULTIPLE REGRESSION ANALYSIS OF Q FOR UNIQUE DISCHARGE REGIONS.....	C9
C3.2. ASSEMBLING THE Q REGRESSION .....	C10
C3.3. EVALUATION OF Q THROUGH CONFIRMATION FOR EACH REGION .....	C12
<b>C4. SUMMARY</b> .....	<b>C12</b>

<b>D. REMOTELY-SENSED DUAL COEFFICIENT (RDC) ET ESTIMATION FOR GMA8.....</b>	<b>D2</b>
<b>D1. INTRODUCTION TO THE RDC METHOD .....</b>	<b>D2</b>
<b>D2. MAKING MODIS NDVI EQUIVALENT TO LANDSAT TM5 NDVI .....</b>	<b>D4</b>
<b>D3. PROCESSING MODIS IMAGES FOR ESTIMATION OF ET.....</b>	<b>D6</b>
<b>D4. SUMMARY.....</b>	<b>D8</b>
<b>E. RDC PARAMETERS—VALIDATION AND CALIBRATION.....</b>	<b>E2</b>
<b>E.1 SITE DESCRIPTIONS AND ANALYSIS PROCEDURES .....</b>	<b>E3</b>
<b>E.2 ETA BY VEGETATION CLASS.....</b>	<b>E5</b>
E.2.1 WOODY VEGETATION—THE FORESTED ETA CLASS.....	E7
E.2.2 HERBACEOUS VEGETATION—THE GRASSLAND ET CLASS .....	E8
E.2.3 CROPS ET CLASS .....	E9
<b>E3. SUMMARY .....</b>	<b>E10</b>
<b>F. APPLICATION OF ETA ESTIMATION .....</b>	<b>F2</b>
<b>F1. SELECTION OF REPRESENTATIVE ETA MODEL YEARS.....</b>	<b>F3</b>
<b>F2. MODIS NDVI* PROCESSING AND ETA QUARTERLY INPUT VARIABLES .....</b>	<b>F4</b>
<b>F3. GAP ETA CLASSES AND THEIR APPLICATION .....</b>	<b>F5</b>
<b>F4. QUARTERLY AND ANNUAL ETA FOR REPRESENTATIVE WET AND DRY YEARS.....</b>	<b>F6</b>
<b>F5. INTERMEDIATE YEAR ETA BY SLOPE BETWEEN WET AND DRY YEARS ETA.....</b>	<b>F8</b>
<b>F6. LINEAR RELATIONSHIP OF ETA TO PRECIPITATION STORED PER PIXEL.....</b>	<b>F11</b>
<b>F7. SUMMARY .....</b>	<b>F11</b>
<b>G. MODELING URBANIZED AREAS .....</b>	<b>G2</b>
<b>G1. HISTORIC URBAN BOUNDARIES .....</b>	<b>G2</b>
<b>G2. NDVI* RATIO.....</b>	<b>G5</b>
<b>G3. DYNAMIC URBAN Q.....</b>	<b>G6</b>
<b>G4. DYNAMIC URBAN ETA.....</b>	<b>G7</b>
<b>G5. SUMMARY.....</b>	<b>G9</b>
<b>H1. MATING ALL MODEL DATABASES .....</b>	<b>H2</b>
<b>H2. HIGH PPT / ET<sub>0</sub> GWR ADJUSTMENTS .....</b>	<b>H4</b>
H2.2 THEORETICAL BEHAVIOR OF GWR UNDER HIGH PPT/ET <sub>0</sub> .....	H5
H2.2 HIGH PPT/ET <sub>0</sub> ADJUSTMENTS .....	H6
H.2.2 EXTRAPOLATING GWR PREDICTIONS FOR VERY YEARS WITH HIGH PPT/ET .....	H7
<b>H3. FINAL GWR ESTIMATES .....</b>	<b>H9</b>
H3.1 GWR PLATEAU AND BLENDED MODEL RESULTS .....	H9
H3.2. ETA CALCULATED AS A RESIDUAL FROM PLATEAU AND BLENDED MODELS .....	H10
H3.3. COMPARISON OF GWR RESULTS AND PREVIOUSLY PUBLISHED RECHARGE ESTIMATES .....	H11
H3.4. COMPARISON OF GWR BEHAVIOR IN HYDROBIO AND HARDEN MODELS .....	H15
<b>H5. SUMMARY.....</b>	<b>H16</b>
<b>I1. LIMITATIONS AND RECOMMENDATIONS.....</b>	<b>I2</b>
<b>I1. LIMITATIONS OF THE GWR MODEL .....</b>	<b>I2</b>

I1.1 DATA RESTRICTIONS.....	I2
I1.2. BROAD CHARACTERIZATION:.....	I3
<b>I2. GENERAL RECOMMENDATIONS FOR GWR MODELING .....</b>	<b>I3</b>
<b>I3. RECOMMENDATIONS FOR FUTURE GWR MODELING IN GMA8.....</b>	<b>I5</b>
<b>I4. RECOMMENDATIONS FOR USING GWR ESTIMATES IN FUTURE MODELS.....</b>	<b>I5</b>
<b>J. APPENDIX 1: REFERENCES .....</b>	<b>J2</b>
<b>K. APPENDIX 2: DELIVERABLES.....</b>	<b>K2</b>
<b>L. APPENDIX 3: REVIEW OF MODEL CALCULATIONS .....</b>	<b>L2</b>
<b>L1. REFERENCE EVAPOTRANSPIRATION (ET<sub>0</sub>) ESTIMATION.....</b>	<b>L3</b>
<b>L2. DISCHARGE (Q) ESTIMATION.....</b>	<b>L3</b>
<b>L3. INFILTRATED RAINFALL (R<sub>INFIL</sub>) ESTIMATION.....</b>	<b>L5</b>
<b>L4. ANNUAL EVAPOTRANSPIRATION ESTIMATION (ETA) .....</b>	<b>L6</b>
<b>L5. ANNUAL GWR ESTIMATION: GWR PLATEAU AND GWR BLENDED.....</b>	<b>L7</b>
<b>L6. DELIVERED MODEL VARIABLES .....</b>	<b>L9</b>
<b>M. APPENDIX 4: DRAFT REPORT COMMENTS AND RESPONSES.....</b>	<b>M2</b>



# INTRODUCTION



Woodbine Aquifer Outcrop



Brazos River Alluvium



Trinity Aquifer Outcrop



## Abstract

*HydroBio developed a model to predict annual groundwater recharge (GWr) across approximately 50,000 square miles of North Central Texas at square-mile resolution in annual time-steps through 50 years (1960-2009). The model is defined by precipitation and evaporation-driven variables and measured permeability that were calibrated separately. GWr was calculated as a residual from annual precipitation and water disposition from stream discharge and landscape evapotranspiration (ETa). All model variables were calculated and manipulated as rasters so that the data are preserved spatially.*

*The Remotely-sensed Dual Coefficient method was used to calculate ETa across the region using MODIS data for wet and dry conditions. Because all model inputs were measured data from weather stations, stream gages and land cover and permeability that were linked in separate calibration efforts, the resulting GWr values are expected to be correct relationally. The GWr outputs can be varied by scaling for fine adjustment for future application within regional scale modeling while calibration offers a better understanding of how to adjust the parameters.*

## 1. GMA8 Site Description

The study area for GMA8, shown in Figure 1, comprises an area 34,964 square miles within the GMA8 boundary and a total of 48,909 square miles for the combined GMA8 extent inclusive of the zone identified for the GMA8 boundary (Figure 1).

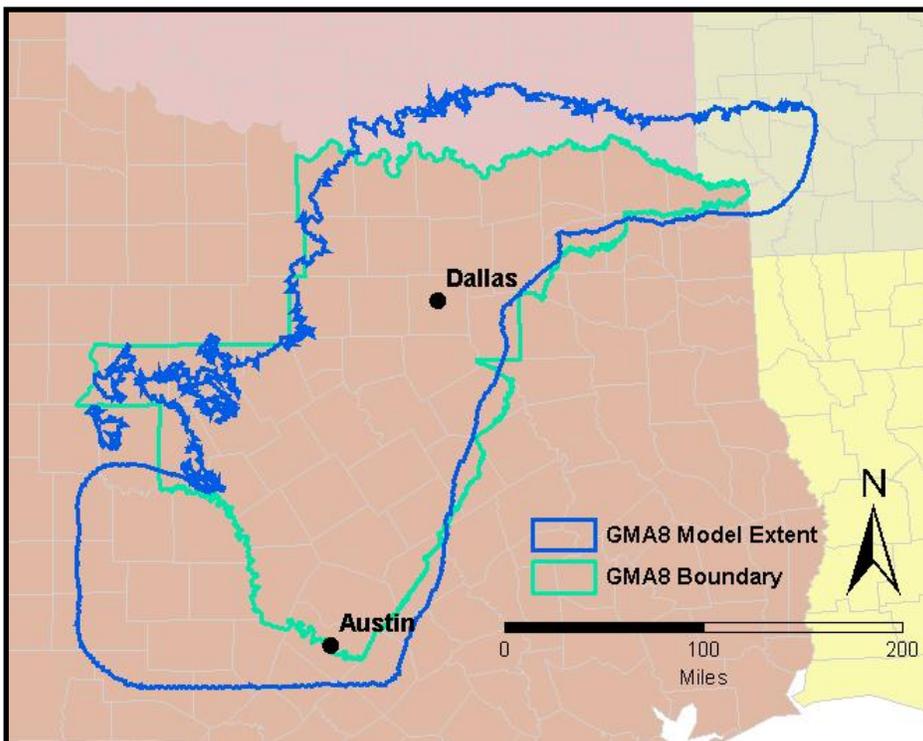


Figure 1. Map of the GMA8 Project area. The red outline is the original GMA8 boundary. The gray outline is the GMA8 project area that includes buffer areas.

## 2. Relational Groundwater Recharge Model Concept

The structure of this GWr model is different from other groundwater models because it was assembled entirely from weather inputs, stream gage data, land cover data, remote sensing-based estimates of ETa, and published values of measured permeability. Absent are model variable adjustments that are often made for matching water tables in groundwater models. It is anticipated that this independence will be of value when this work is combined with more traditional modeling approaches.

**Table. Abbreviations used in this report.**

<b>Table of Commonly Used Abbreviations</b>	
<b>AWP</b>	Intermediate synthesis of the multiple layers of STATSGO soil data that have been reduced, by averaging, to a single numeric value to represent soil permeability.
<b>BF</b>	Base flow. Subsurface component of discharge estimated using USGS streamflow partitioning program PART.
<b>EOS</b>	Earth observation satellites such as Landsat TM and MODIS.
<b>ET</b>	Evapotranspiration.
<b>ET<sub>0</sub></b>	Reference evapotranspiration. Reported in in/yr unless otherwise noted.
<b>ETa</b>	Annual evapotranspiration. Reported in in/yr unless otherwise noted.
<b>GAM</b>	Groundwater Availability Modeling Program operated by TWDB.
<b>GAP</b>	USGS survey to detect "gaps" in conservation coverage combined multiple sources of GIS data that classifies land cover nationwide.
<b>GIS</b>	Geographic information system.
<b>GMA8</b>	Groundwater Management Area 8; defined by TWDB.
<b>GRD</b>	Groundwater Resources Division of TWDB.
<b>GWr</b>	Groundwater recharge is the residual of Rinfil minus ETa. Reported in in/yr unless otherwise noted.
<b>K<sub>cb</sub></b>	Component of ET derived the water use of a crop canopy.
<b>K<sub>e</sub></b>	Component of ET from wetted surfaces, including soil and water bodies.
<b>Landsat TM</b>	Landsat Satellite Thematic Mapper.
<b>MODIS</b>	Moderate Resolution Imaging Spectroradiometer.

<b>NCRS</b>	Natural Resources Conservation Service.
<b>NDVI, NDVI*</b>	Normalized difference vegetation index used to quantify density of plant growth from NIR and SWIR. NDVI* is a calibrated form NDVI that is stretched from zero to one representing bare ground to maximally vegetated surfaces.
<b>NIR</b>	Near infrared.
<b>PART</b>	USGS stream partitioning program estimates base flow record in inches from gaged daily discharge and watershed area.
<b>Perm</b>	Permeability. Final value used to represent infiltration capacity of soil within model calculations. Reported in in/hr.
<b>Ppt</b>	Precipitation. Reported in in/yr unless otherwise noted.
<b>Q</b>	Discharge: surface water drained from recharge system. Reported in in/yr unless otherwise noted.
<b>RDC</b>	Remotely-sensed dual coefficient. ET estimation technique calibrated to land cover classes with ET flux data and satellite imagery developed by HydroBio.
<b>RefET</b>	Reference evapotranspiration estimation program available through the University of Idaho.
<b>R<sub>infil</sub></b>	Infiltrated rainfall; the residual of precipitation minus discharge. Reported in in/yr unless otherwise noted.
<b>RO</b>	Runoff. Residual of discharge minus base flow.
<b>STATSGO</b>	State Soil Geographic Database. Permeability data is reported in cm/hr of each standard soil layer and is derived from laboratory analysis of the mapped soil unit.
<b>SWIR</b>	Short-wave infrared band.
<b>TWDB</b>	Texas Water Development Board.
<b>USGS</b>	United States Geological Survey.
<b>WSS</b>	Web Soil Survey. Online soil database operated by NCRS.

### 3. Executive Summary

A groundwater recharge (GWr) model for a 50-year period, 1960 to 2009, was developed for Texas Groundwater Management Area 8 (GMA8) through solving a simple water balance that yielded GWr as a residual from calibrated inputs. The simplified water balance approach is appropriate because the GMA8 GWr model was formulated for one square mile grid cells within annual time steps. On a landscape level, the opportunity for GWr to occur is directly dependent upon precipitation less the water that drains leaves the landscape as drainage or evapotranspiration. This is expressed as:

$$\text{GWr} = \text{Ppt} - \text{Q} - \text{ETa}$$

Where Ppt is precipitation, Q is total discharge and ETa is annual evapotranspiration.

This water balance approach was spatially discrete with the variables measured, interpolated and estimated by pixels within rasters. Rasters are continuous fields of numbers that can be combined mathematically to solve the equation to yield GWr for all locations across the GMA8 region. The data input for solving GWr was calibrated prior to model assembly. Ppt and Q calibration occurred in discrete locations where measurements were obtained with values that were extrapolated and interpolated to larger areas. Ppt was calibrated from annual totals for stations located throughout the GMA8 region. Ppt was then interpolated across the entire region by kriging for each of the 50 years.

Q was calibrated for representative watersheds that were free from disturbances (reservoirs, irrigation diversions, and urbanization) and that had sufficient periods of measurement. Q was then extrapolated to the entire GMA8 project region using published maps of permeability. Satellite data were used to calibrate for ETa using a vegetation index to scale against ET<sub>0</sub> (ASCE Penman Monteith grass coefficient) that was kriged across the GMA8 region. These variables were calculated as rasters for each of the 50 years examined in the model.

Two weather variables were necessary for calibration in the GWr model, ET<sub>0</sub>, and precipitation. Ppt is the driving variable that determines the magnitude of recharge; when annual Ppt is low, no water is available for GWr. At high annual Ppt, GWr is expected to also be high. ET<sub>0</sub> was found to moderate Ppt for calculation of Q rendering the relationship quasi-linear and, therefore, more easily managed mathematically. An extensive record of precipitation was available for the project area.

The meteorological record sufficient for spatial calculations of ET<sub>0</sub> was relatively short—few stations collected enough data for the calculation and those that did only covered about a decade. Calibrating ET<sub>0</sub> to precipitation was chosen as the method for estimation of ET<sub>0</sub> across the project area for the full 50-year record. Quarterly ET<sub>0</sub> data were analyzed and annual ET<sub>0</sub> data were selected for use in model calibration. Annual Ppt was spatially interpolated by kriging to produce raster inputs for calculation of ET<sub>0</sub>. A regression relationship between Ppt and ET<sub>0</sub> was developed and kriged to model the 50-year ET<sub>0</sub> record.

Q within the GMA8 GWr model was characterized through analysis of gaged watersheds that were calibrated to soil permeability. The STATSGO soil database, NRCS Web Soil Survey, and review of literature describing permeability and infiltration capacity of individual aquifers within the GMA8 model boundaries were used to assemble and validate a permeability raster layer (Perm) to serve as the basis for dynamic estimation of Q across the GWr model region. Although the USGS gaged record included 400 records, a suite of only 40 gages was selected to represent discharge conditions across GMA8. This suite was based on criteria of data continuity, length of record, location, size and freedom from disturbance. Data from the selected representative watersheds were

collected from USGS stream gages and mated with Perm, Ppt, and  $ET_0$  data in preparation for calibration of  $Q$  for extrapolation to the model region.

To estimate the amount of infiltrated rainfall for GWR predictions, a technique was developed to estimate how much of the incoming Ppt drained from the watershed—thus, unavailable for aquifer recharge. This drainage was quantified as total gaged discharge  $Q$ , which was analyzed in 40 representative watersheds. The relationship between  $Q$  and Ppt was simplified and linearized by scaling the driving variable Ppt by  $ET_0$ —a ration that better expressed the opportunity for GWR across the model region than Ppt, only. The slope of the linear relationship between  $Q$  and  $Ppt/ET_0$  ( $b_Q$ ) provided a way to spatially codify  $Q$  behavior as a raster. Perm and 50-year  $Ppt/ET_0$  averages were also interpolated as rasters. A multiple linear regression model using  $Ppt/ET_0$ ,  $b_Q$ , and Perm were used to estimate  $Q$ . These raster calculations provided spatial estimates of  $Q$ . The  $Q$  predictions resulting from these analyses accurately modeled the  $Q$  observed in gaged data when spot checked for individual watershed.

Data from MODIS imagery (241m pixels) were used to estimate  $ET_a$  across the project region. To prepare for this estimation, Landsat TM data (30m pixels) were first calibrated to measured  $ET_a$  and  $ET_0$  within published data sets. For estimation of  $ET_a$ , a stretched version of NDVI was used,  $NDVI^*$ . Calculation of  $NDVI^*$  required the selection of a zero vegetation cover location on the MODIS imagery. These calibration targets were bare areas located within large quarries. To model the range of  $ET_a$  observed in GMA8, record wet (2007) and dry (2005) years were chosen to spatially represent the endpoints of  $ET_a$  across the varying climate conditions of GMA8.

MODIS NDVI was made equivalent to Landsat TM NDVI using a mathematical transformation. Transformation of MODIS enabled the use of the  $ET_a$  calibration at a much finer scale with Landsat TM data.  $ET_a$  was calculated for vegetation by MODIS Pixel using the Remotely-sensed Dual Coefficient (RDC) method that employed  $ET_0$ ,  $NDVI^*$  and fitted functions for three vegetation cover types: grassland, woodland and cropland. For estimation of  $ET_a$  using the mapped vegetation cover type relationships, GAP data across the GMA8 region were reclassified into these three simple cover types.

Urbanization greatly affects GWR through impermeable building footprints, roadways and parking lots that limit the area through which water can recharge. During the 50-year span of the GWR model period, the extent of the urbanized areas within the GMA8 model region has expanded significantly (e.g., Dallas-Fort Worth, Waco, Austin, etc.). Urban expansion was represented in the model for four time steps, 1960-1977 (from TDWR Land Cover Maps), 1978-1989 (interpolated contour), 1989-1999 (GAP Analysis from 1999-2001 Landsat Imagery), 2000-2009 (from 2007 MODIS Imagery). These steps were held constant with no interpolation made for gradual annual changes.

The effect upon GWR from expansion of urban areas was modeled using two key GWR variables:  $Q$  was increased with increasing area of impervious surfaces such as pavement and buildings, and water loss through evapotranspiration ( $ET_a$ ) was decreased when vegetation was replaced by urban hardscape. A metric for estimating the degree of

urbanization through changes in Q and ETa was developed that used the ratio of average NDVI\* in the model countryside to the NDVI\* measured over each urbanized pixel as a decimal fraction. This fraction was used to boost Q based on the degree of urbanization present in a pixel at a given year. ETa in the model automatically decreased because NDVI\* within the urbanized zone decreased. No adjustment was made for interception losses.

After the completion of calibration of the driving variables, these were assembled for calculation of GWr across the GMA8 project area. Under conditions of extreme Ppt/ET<sub>0</sub>, GWr exhibited an incorrect downward parabolic limb that indicated that the calibration for ETa through the entire range of dry to wet years was not adequately represented—under conditions of high Ppt/ET<sub>0</sub>, ETa decreases because the evaporative driving force, represented by ET<sub>0</sub> decreases due to cloud cover, high humidity and cooler temperatures. Calibration by modeling across the entire range of Ppt/ET<sub>0</sub> would correct for this, however, recalibration of ETa was judged to be beyond the capability conferred by the available data sets--MODIS data were available only for a decade and so ETa could not be calibrated for extreme values of Ppt/ET<sub>0</sub> because they were not available for all locations in the project area within the period of the MODIS record.

The GWr final model adjustment to overcome the inadequate ETa calibration was based upon accepting the relationship for GWr for lower levels of Ppt/ ET<sub>0</sub> and then changing the downward parabolic portion of GWr by replacing these aberrant values to follow a curve shape for diminishing returns. The diminishing returns curve is appropriate for GWr expected because Q increases proportionately during high levels Ppt/ET<sub>0</sub> while ETa decreases. GWr must respond as a diminishing returns curve because quasi-Hortonian overland flow occurs during longer periods of saturated soils when wet cycles generate higher surface runoff. As more water is received through precipitation, less water is available proportionately for GWr.

Two models for predicting GWr under extreme Ppt/ET<sub>0</sub> conditions were applied for pixels that experienced a downward parabolic limb. The first model, “Plateau”, held the upper limb of GWr constant at the measured parabolic peak value. The second model, “Blended”, allowed a small increase above the “Plateau” commensurate with the expected shape of a diminishing returns curve. The Plateau and Blended models for GWr, corresponded in magnitude but demonstrated some subtle differences in spatial distribution.

This approach to modeling GWr is completely new and provides a valuable understanding of the controlling processes and their spatial distribution. Because the GWr estimates are spatially calibrated to independent variables, including measured properties of Perm, land cover, ET<sub>0</sub>, Ppt, and Q, the model results are controlled by the actual environment as accurately as they could be represented. Spatial variation in GWr estimates corresponded to the varied environmental conditions within the project—demonstrating the relational accuracy of the model.

The relational properties of the GWr results are an important consideration. Within any calibration attempt, the resulting calibration often carries systematic error inferred by the data that were used. Although we do not know where such biases occurred, the data and resulting GWr calculations are relationally correct—a property that lends the model inputs and outputs for scaling regionally. Scaling can be as simple as applying a decimal multiplier to scale GWr up or down to balance with results from groundwater modeling working from the traditional perspective of balancing inputs to achieve water levels.

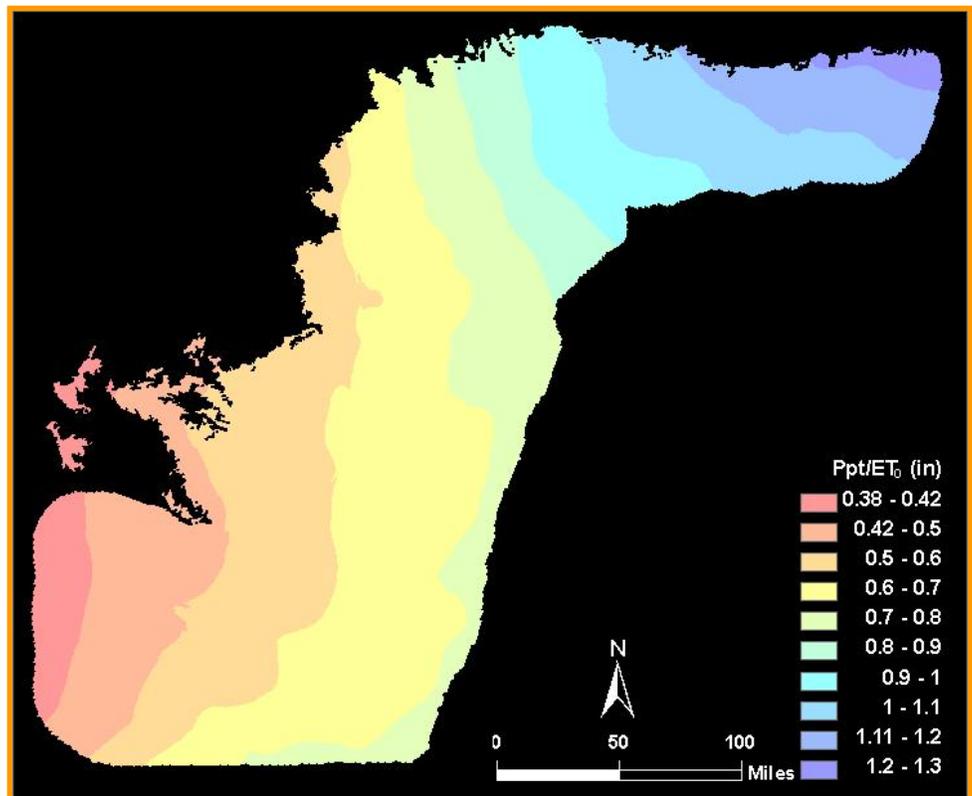
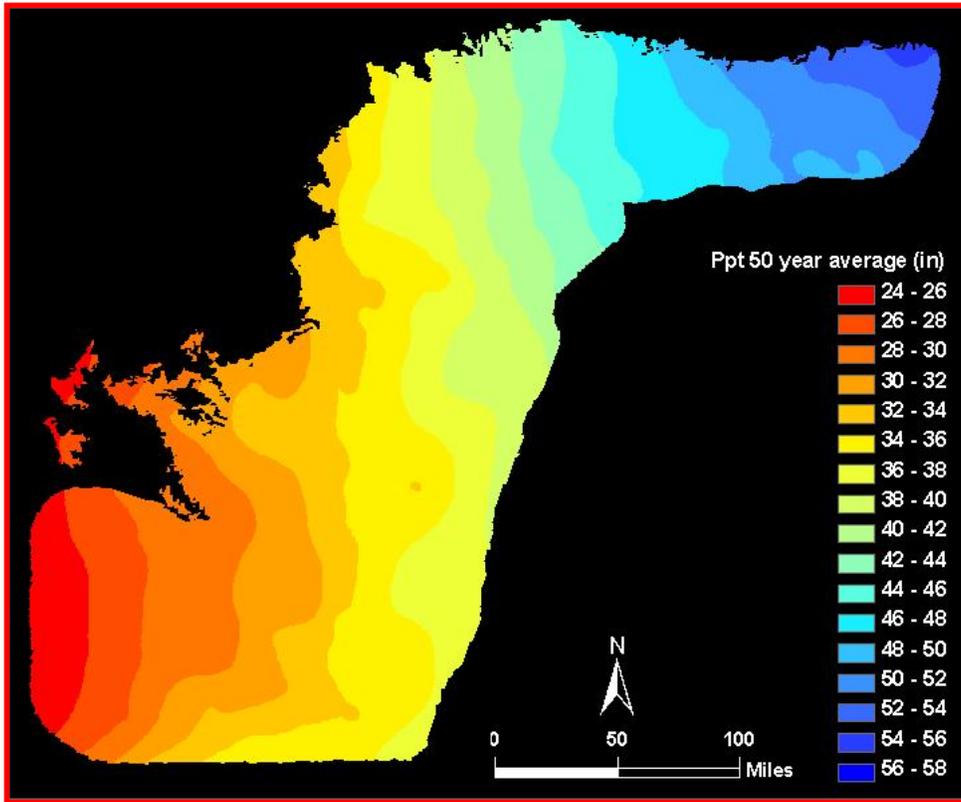
The GWr prediction for GMA8 was an effort wholly independent of traditional groundwater modeling methods because it calibrated the parameters of the surface water balance to then calculate what proportion remained for potential recharge. This approach should be valuable for understanding the relative magnitude and interrelationship of the parameters used for calculation of GWr. This is valuable feedback so that these parameters can be represented correctly in other groundwater modeling. For example, the shape of ETa from dry to wet annual weather was formerly an unknown that was predicted during final model adjustments. Likewise, the quasi-linearity of Q vs. Ppt/ETo is an advancement of knowledge that can now be used in future TWDB groundwater modeling.

### **3. Acknowledgements**

Insight and feedback from several individuals at the Texas Water Development Board were instrumental in the modeling and writing process. We would like to thank Cindy Ridgeway, Marius Jigmond, Jerry Shi, and Roberto Anaya of the Groundwater Availability Modeling department of TWDB. We would also like to thank Bill Hutchison, as director of the Groundwater Resources Division, for his constructive observations early in the modeling process.



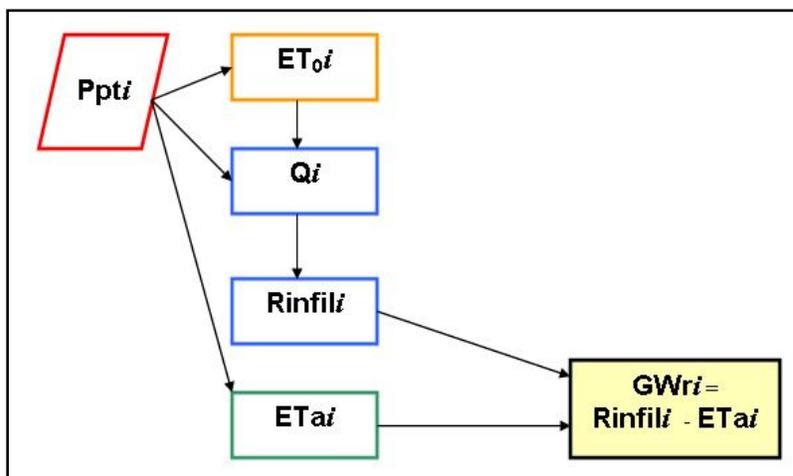
# A. CALIBRATING DRIVING MODEL VARIABLES: PRECIPITATION AND ET<sub>0</sub>





## A1. Driving Model Variables

Two key model inputs drive the 50 year annual groundwater recharge model, annual total precipitation (Ppt) and annual total  $ET_0$ , all input variables being annual values. Groundwater recharge (GWr) estimation was guided by four linear equations (Figure A1). The slope and intercept of each equation were developed using the relationship between precipitation and the appropriate model variable, including ASCE Penman Monteith grass reference evapotranspiration ( $ET_0$ ), stream discharge (Q), the proportion of Ppt that infiltrated the ground ( $R_{infil}$ ), and evapotranspiration of native vegetation and crops ( $ETa$ ). These variables allow GWr to be estimated for a given year from precipitation as a single input.  $ET_0$  was used as a driving variable for both  $ETa$  and Q estimation. An  $ET_0$  model was used to generate estimates of this driving variable based upon precipitation as an input since 50 years was desired for the model, but weather variables for calculation of  $ET_0$  were only available across the region for the past decade to decade and a half.



**Figure A1. GWr estimation for a given year “i” is guided by four Ppt-driven linear equations.**

- 1)  $ET_{0i} = a + b * Ppt_i$
- 2)  $Q_i = a + b * (Ppt / ET_{0i})$
- 3)  $R_{infil}_i = Ppt_i - Q_i$
- 4)  $ETa_i = a + b * Ppt_i$

Spatial representation of Ppt and  $ET_0$  for the model was generated from point data and geostatistical interpolation and extrapolation (kriging). Logically, both Ppt and  $ET_0$  vary gradually from station to station and so, geostatistical methods were chosen for interpolating the spatial variability between stations. Spatial models of the 50-year precipitation record and a spatial model for estimating  $ET_0$  were developed to represent both of these driving variables in the GWr model. For the  $ET_0$  model, the stations that gather sufficient input for calculation of a Ref-ET-based  $ET_0$  were used as points to estimate annual  $ET_0$  as a function of precipitation. Thus, through correlation, annual precipitation was also used as a proxy for annual  $ET_0$ .

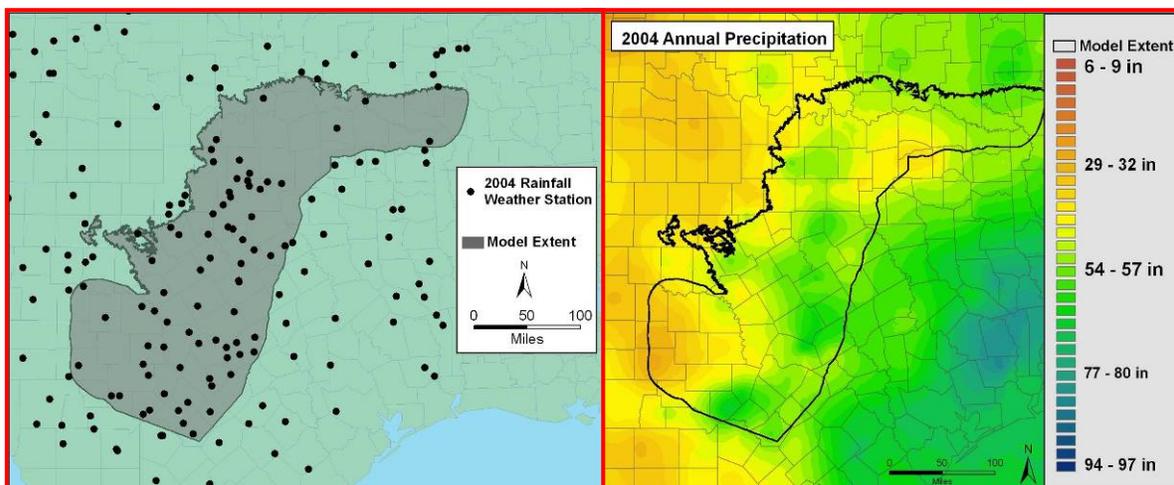
## A2. Modeling Precipitation

The GWr model was formulated to be dynamic from year to year across a 50-year calibration period. Weather stations that gather Ppt are point peppered across the GMA8 region and so the data density of this variable is much better than other more complex

variables, such as  $ET_0$ . The 50-year period was judged to be sufficient to capture the variability that can occur within the region.

### A2.1. 50-Year Precipitation Record

Tabulated precipitation data from The Digital Climate Atlas (DCA) of Texas was provided by the TWDB (TWDB, 2009). Two to five hundred weather stations reported hourly precipitation in the project region each year. Quarterly, bi-annual, and annual sums were calculated for each year—these values were used to populate annual point databases. A 50-year point-precipitation record was compiled covering the project region at these time resolutions. Point values were interpolated spatially by kriging (applied with exponential semivariogram, 12-point variable radius, and 0.25 mile<sup>2</sup> grid cells as a standard approach (Figure A2).



**Figure A2. Map of weather station locations and resulting kriged precipitation raster. Point measurements were used to interpolate precipitation values across the project region.**

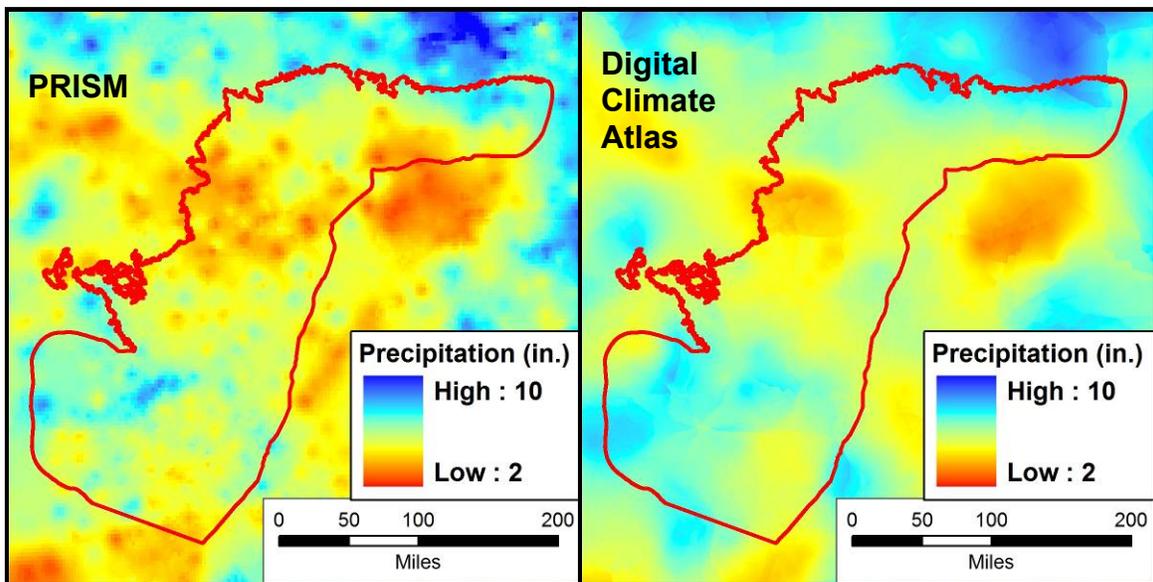
### A2.2. PRISM versus Rainfall Kriged from Texas Digital Climate Atlas

PRISM provides spatial data similar to the kriged data generated from the Texas Digital Climate Atlas (DCA: OSU, 2009 and TWDB, 2009). For PRISM, weather data from stations across the country were collected together and then spatially interpolated by kriging. An added influence in the spatially continuous PRISM model is an algorithm tied to a Digital Elevation Model (DEM) to account for increased precipitation in mountainous terrain. This is important for filling data gaps in areas often overlooked by city and airport weather stations—often such areas are mountainous. Spatial interpolation in mountainous areas is a mechanism to adjust for the effects of elevation change as opposed simply to distance from a station that is the intent of kriging calculations in the absence of elevation input.

For the GMA8 project, changes in elevation tend to be gradual and are generally insufficient to highly influence spatial distribution of precipitation. The Ozarks, well north of the project boundary, are the closest geographic feature to present an increase in precipitation that could be appropriately modeled with PRISM's elevation algorithm. Another potential drawback of the PRISM data is the 'blackbox' nature of the data set having raw data that was not accessible and station lists and data completeness that were not available for examination. Apprehensive of datasets that cannot be proofed, a decision was made to produce precipitation surfaces where the output can be directly checked against the input. PRISM and kriged DCA were compared for the same period.

A drawback with PRISM data shows up in the comparison of the outputs of PRISM and DCA. The PRISM data fluctuated over short distances in a “shotgun” manner, having numerous points in close proximity with similar values that stand out as extremes from the surrounding area. The precipitation values around these points quickly drop off or ramp up only to display opposing characteristics at a nearby point.

The kriged relationship that was developed using DCA was not perfect either, but was felt to more closely model average conditions across the project area rather than present a series of disconnected points. The kriged for the DCA data averaged the surrounding points and allowed for the building of a smooth surface between the data points. This reduced the point island effect that was prevalent in the PRISM data. Kriged DCA, instead, created larger swathes of gentle change in precipitation rather than a continuous fluctuating surface (Figure A3).



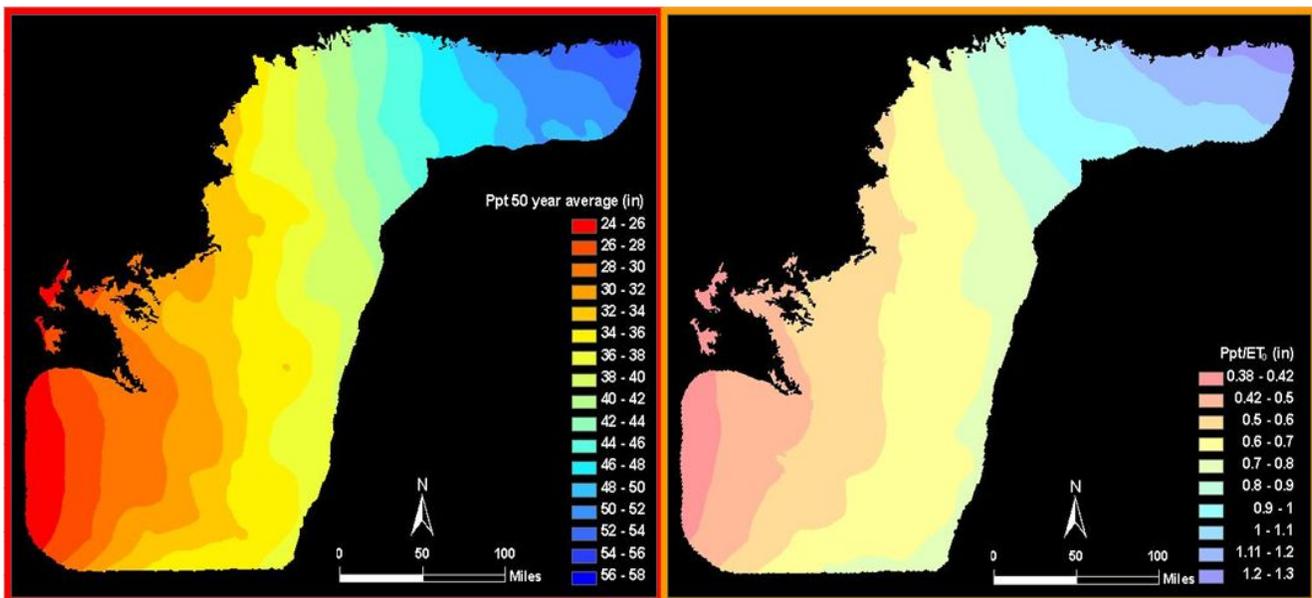
**Figure A3. Comparison of PRISM and kriged Digital Climate Atlas precipitation surfaces. PRISM data from the second quarter of 2005 contains many “point islands” in the distribution (left). Digital Climate Atlas data kriged for the same quarter; kriging produces a smoother transition between data points (right).**

The difference in the distribution shown in Figure A3 was heightened by the resolution of the data sets. PRISM data covers continental United States in 2.5 mile pixels (6.25 square miles), and so requires a resolution that would allow the data set to be manageable across very large regions. A scale of 0.5 mile pixels (0.25 square miles) was used for the kriged DCA rasters thus providing a 25x greater resolution than PRISM data.

In summary, the choice to use DCA and a standard kriging approach was based on a series of issues that made the PRISM data a less ideal choice for the GMA8 project. PRISM data are widely available but the model inputs remain unknown. The special elevation algorithms, unique to PRISM, are likely not required across this region of Texas. Finally the “point island” distribution and larger sized pixels of the PRISM data are less ideal for the GMA8 model scale. These factors combined to make the DCA data that is assessable, finer resolution, and exhibits smoother kriged results a better choice for GMA8 modeling.

### A3. Modeling $ET_0$

Although Ppt is the major input variable for the GWr model,  $ET_0$  is also an important weather variable because it moderates the water introduced by Ppt to more accurately reflect the degree of aridity or humidity in the system. For example, Figure A4 presents a 50-year kriged of Ppt and a 50-year kriged of Ppt/  $ET_0$ . Since both Ppt and  $ET_0$  change inversely from the southwest to the northeast across the project area, they enhance the effects of aridity to the west and humidity to the east. Both variables are required for providing competent spatial distribution of GWr.

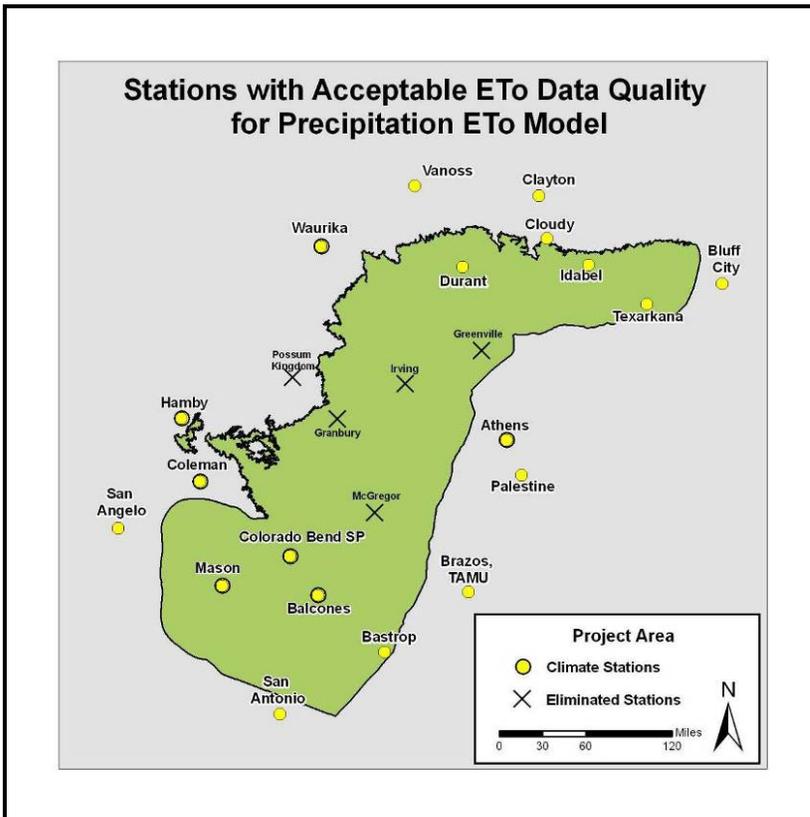


**Figure A4. 50-year spatial average Ppt shown on the left has a range that is about 2.4x while 50-year spatial average Ppt/ $ET_0$  on the right has a range of about 3.4x. Ppt/ $ET_0$  is a better indicator of the aridity-humidity differences across the project area than Ppt alone.**

### A3.1 Recent Meteorologic record

ET<sub>0</sub> is a critical driving variable in GWR calculations but the meteorological record has data sufficient for calculating ET<sub>0</sub> only in recent years (2000 to present). Because of the 50 years of record desired for the GWR model, the brevity of the ET<sub>0</sub> record necessitated a technique to predict this variable for model years prior to 2000. The recent meteorological record includes both precipitation and ET<sub>0</sub> parameters. Fortunately, the close relationship between annual ET<sub>0</sub> and annual precipitation permitted calibration to estimate ET<sub>0</sub> during earlier years.

Data were collected for all climate stations with sufficient annual records from Mesowest, Mesonet, and the Texas ET Network (Figure A5). The Ref-ET program (UID, 2010) was used to calculate reference evapotranspiration for stations that did not report ET<sub>0</sub> directly. REF-ET assimilates meteorological inputs from a suite of parameters to calculate a standard ASCE Penman-Monteith ET<sub>0</sub>. Quarterly, bi-annual, and annual ET<sub>0</sub> and precipitation were summed for each station.



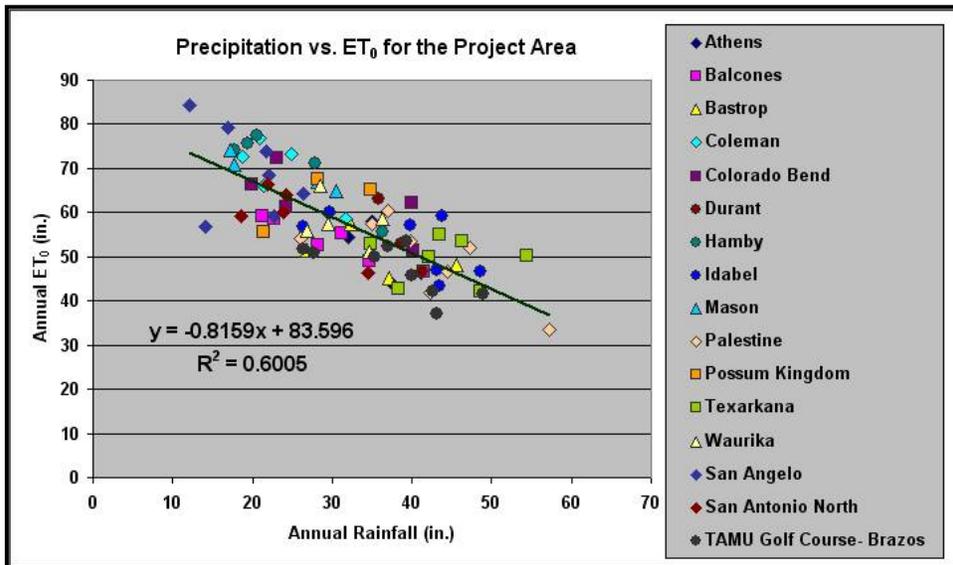
**Figure A5. Climate stations used to calibrate ET<sub>0</sub> to precipitation. Regression slopes and intercepts calculated at each station were kriged across the project region. Stations marked with an X were removed from the kriging model after noting that they did not fit well within the suite of other stations.**

Five stations were removed from the suite of potential stations as indicated in Figure A5. These stations were removed after (1) first observing that the data were recognizably different from the surrounding 19 stations, and (2) observing those stations on Google Earth™. The stations that were removed from further consideration had recognizable

problems associated with providing a  $ET_0$  that was truly representative of standard conditions. The problems were all related to easily observable issues within the fetch of the stations (within 50 to 100 m) that included parked cars, roadways and asphalt, and major obstructions such as trees and hedges.

### A3.2 Optimization of Precipitation- $ET_0$ Model

Initial observations of the annual precipitation and  $ET_0$  record revealed a clear linear relationship between the two variables (Figure A6.). To find the most sensitive and representative method to describe this relationship spatially across the project region, analyses began with quarterly data because they provided the largest data set with the highest spatial resolution. This analysis then proceeded to biannual and finally annual scales by applying the observations from the quarterly and bi-annual data to the interpretation. This progression took place because the few years available since the 2000 or later initiation of  $ET_0$  data collection represented a relatively sparse data set and biannual and quarterly time scales provided 2x and 4x the data magnitude.

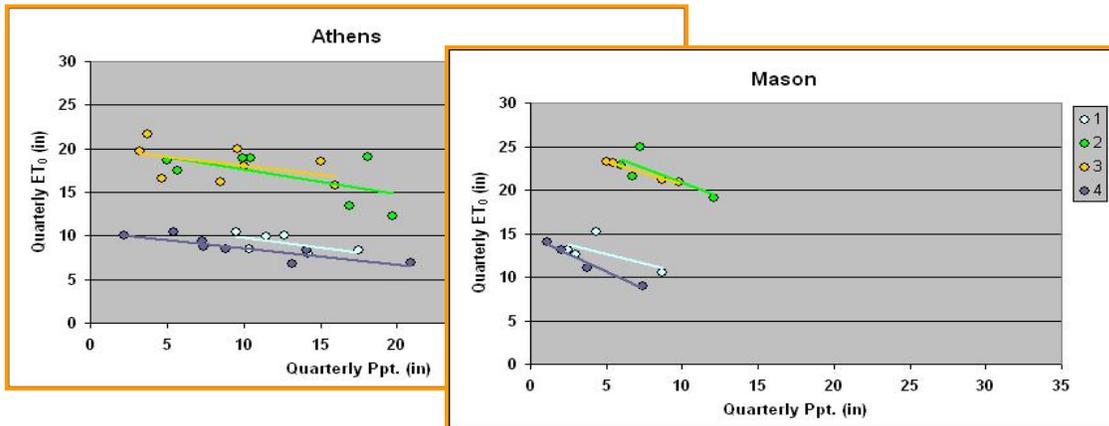


**Figure A6.** The overall annual Ppt- $ET_0$  trend is clear when data from across the project area are graphed together.

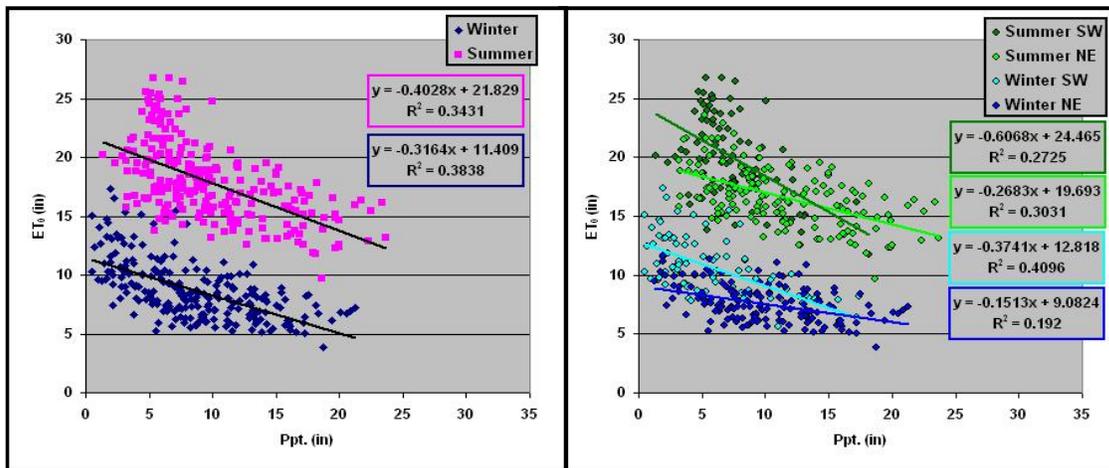
Quarterly relationships were best described by two regression relationships—for winter months (quarters one and four) and for summer months (quarters two and three). This result indicated that all  $ET_0$  values could be accurately described on a bi-annual basis (Figure A7.). The quarterly data were then lumped into project-wide winter and summer groups and these further demonstrated the clear seasonal difference in the precipitation- $ET_0$  relationship (Figure A8).

The results in Figure A7 and A8 are an important departure from biannual trends that are often used for calendar-based data. Instead of dividing the year into first half (January through June) and second half (July through December), the  $ET_0$  relationships show clearly that the year is best divided as winter (1<sup>st</sup> quarter and 4<sup>th</sup> quarter) and summer (2<sup>nd</sup>

quarter and 3<sup>rd</sup> quarter). While bi-modal trends are clear, the spread of data suggests a spatial influence that must be captured to preserve the integrity of the data structure (Figure A8). As described in numerous locations within this report, the method used to represent spatial data and to preserve spatial relationships was kriging the slope and intercepts of the individual Ppt-ET<sub>0</sub> regressions.



**Figure A7. Winter months (quarters 1 and 4) and summer months (quarters 2 and 3) share trends. Quarterly data can be described in terms of bi-annual trends. The regression slopes in the northeast (Athens) are notably shallower than those in the southwest (Mason) indicating a regional trend that exists within ET<sub>0</sub> data.**



**Figure A8. Quarterly data fits into bi-annual trends (left), but the spread of the data around the regression suggests spatial controls. When grouped into large NE and SW regions (right), systematic changes in slope are apparent: steep slopes in the dryer southwest, shallower slopes in the humid northeast.**

To model spatial variation in precipitation-ET<sub>0</sub> relationships, individual relationships were determined at each climate station and interpolated across the project area. Using observations from the quarterly data, linear regression were fitted into bi-annual and annual precipitation and ET<sub>0</sub> sums (Figure A9). The summation required for bi-annual

and annual data reduced the number of points available for fitting the relationships. Because of this data reduction, some stations in near proximity to one another, and with similar quarterly trends, were lumped in order to better define point-wise Ppt-ET<sub>0</sub> relationships. The annual evaluation provided curve fits that were no better (or worse) than those fitted for the quarterly and biannual relationships, leading to the conclusion that annual (calendar) data will perform well for capturing the spatial variability in ET<sub>0</sub> (Figure A9).

ET<sub>0</sub> was interpolated across the project region using the slope and intercept Ppt-ET<sub>0</sub> regression from each climate station or group of stations. Kriged rasters of slope and intercept of the Ppt-ET<sub>0</sub> regression were generated at quarterly and annual resolution using the chosen standard krigé parameters (exponential semivariogram, a variable 12 point radius, and half mile grid cells). These data were used to develop the Ppt/ET<sub>0</sub> relationship shown in Figure A4 that was subsequently used for all further modeling.

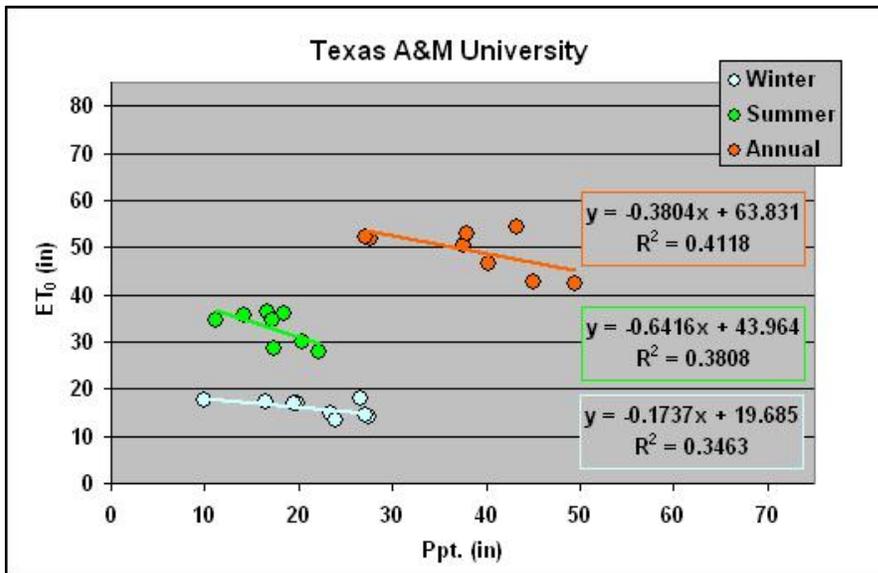


Figure A9. Annual fits performed as well or better than quarterly and biannual fits for each station. Biannual and annual Ppt-ET<sub>0</sub> relationships were fitted to individual stations in the project region.

### A3.3. Model Implementation & Testing

The intent for fitting the spatial ET<sub>0</sub> model was to (1) be able to estimate ET<sub>0</sub> as a factor of Ppt for years prior to when the ET<sub>0</sub>-capable weather records began, (2) to fit this relationship for recent years, since ET<sub>0</sub>-capable stations are sparse and have come and gone within this period, and (3) to provide the basis for estimation of ET<sub>0</sub> throughout the 50-year period desired for the model. Ppt rasters expressed as annual totals in inches and ET<sub>0</sub> slope and intercept rasters provided the means to estimate annual ET<sub>0</sub> from Ppt according to a standard linear relationship with slope (b) and intercept (a) as in the foregoing (Equation A1). Annual project-wide ET<sub>0</sub> was calculated for a given year by performing raster algebra with three raster inputs: (1) the annual precipitation for year “i”, (2) the ET<sub>0</sub> regression intercept, and (3) the ET<sub>0</sub> regression slope.

$$ET_{0i} = a + b * Ppt_i$$

Equation A1

ET<sub>0</sub> estimation from spatially interpolated data were tested against measured data by calculating ET<sub>0</sub> over a model watershed, Lampasas River, near a climate station reporting measured ET<sub>0</sub> variables, Colorado Bend. Bi-annual and annual estimates were calculated using precipitation and ET<sub>0</sub> data extracted from the spatially interpolated rasters for all years reported at the climate station. In the case of Lampasas, it appears that the annual regression may be a more accurate estimator of ET<sub>0</sub> than the bi-annual estimates which deviate more than the annual. Colorado Bend was chosen because of its proximity to Lampasas, but the point and watershed do not overlap—the two were expected to have comparable magnitudes of precipitation and ET<sub>0</sub>, but not exact matches (Table A1).

In summary, an extensive record of precipitation was available for the project area. The meteorological record sufficient for calculating ET<sub>0</sub>, however, was relatively short and few stations collected enough data for ET<sub>0</sub> calculation on those that did only covered the last ten years. Calibrating ET<sub>0</sub> to precipitation, however, allowed predictions of ET<sub>0</sub> to be made across the entire project area for the full precipitation record with the estimation calibrated spatially. Examining the data at quarterly and bi-annual intervals increased the amount of data, enhanced signal detection, and improved the robustness of estimates. Linear regression equations for the Ppt-ET<sub>0</sub> relationship were kriged across GMA8 and used to estimate annual ET<sub>0</sub> for the 50-year model record.

**Table A1. Testing Ppt-ET<sub>0</sub> model output. Ppt data extracted from rasters used to estimate ET<sub>0</sub> over the Lampasas watershed, using bi-annual and annual regression estimates extracted from rasters, is compared to the ET<sub>0</sub> calculated from measured meteorological data collected at nearby climate station Colorado Bend (the climate station does not fall within the watershed).**

Year	Bi-Annual Precipitation (in)		Ref-ET ET <sub>0</sub> at Colorado Bend (in)			Estimated ET <sub>0</sub> at Lampasas (in)		
	Winter	Summer	Winter	Summer	Annual	Winter	Summer	Annual
2006	12.2	16.9	19.7	40.5	60.2	20.2	38.6	60.7
2007	14.5	33.0	18.5	32.6	51.1	19.8	27.0	53.5
2008	7.5	14.6	22.5	43.8	66.3	21.0	40.3	63.5
2009	15.9	21.9	20.7	41.0	61.7	19.5	35.0	57.3

#### A4. Summary

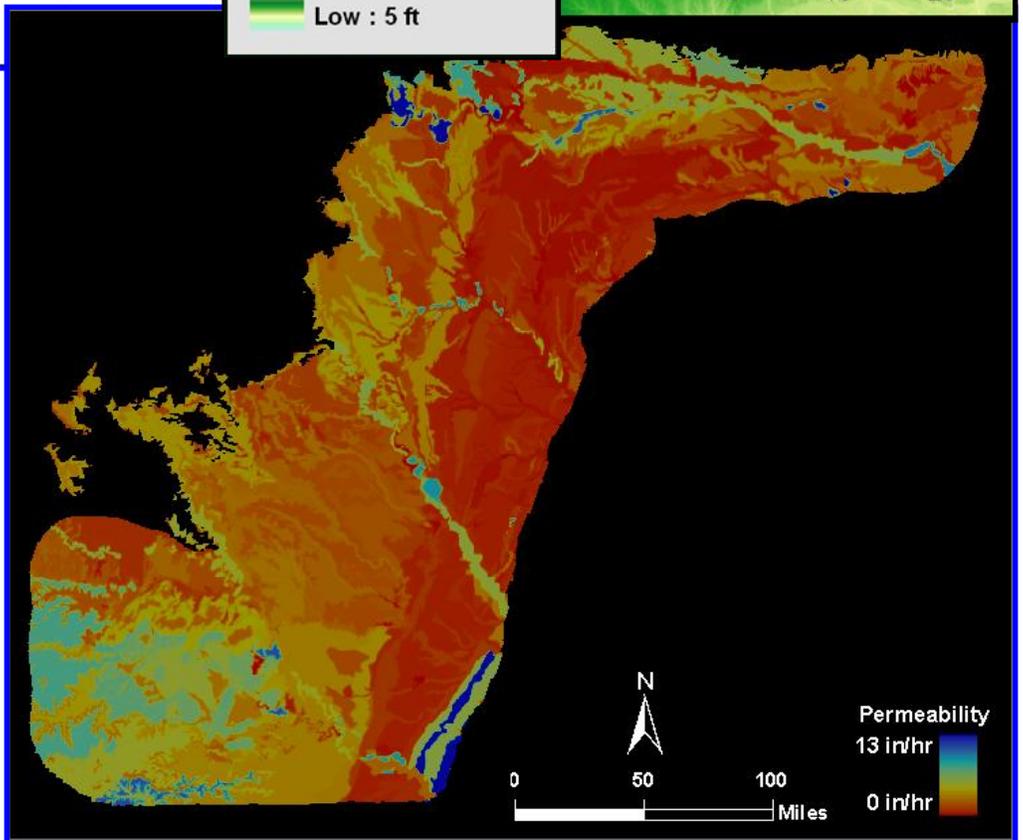
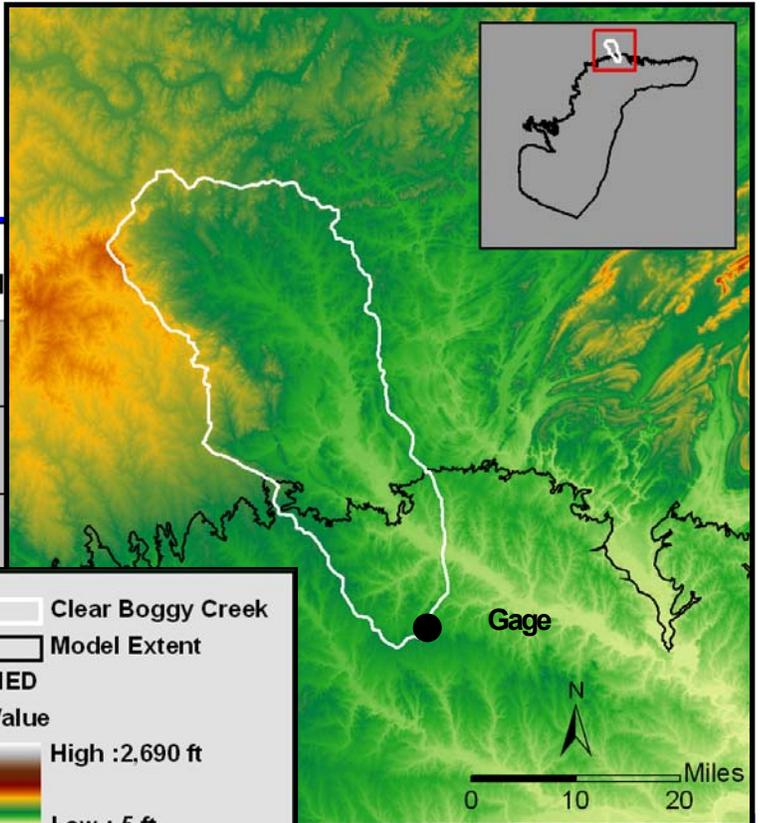
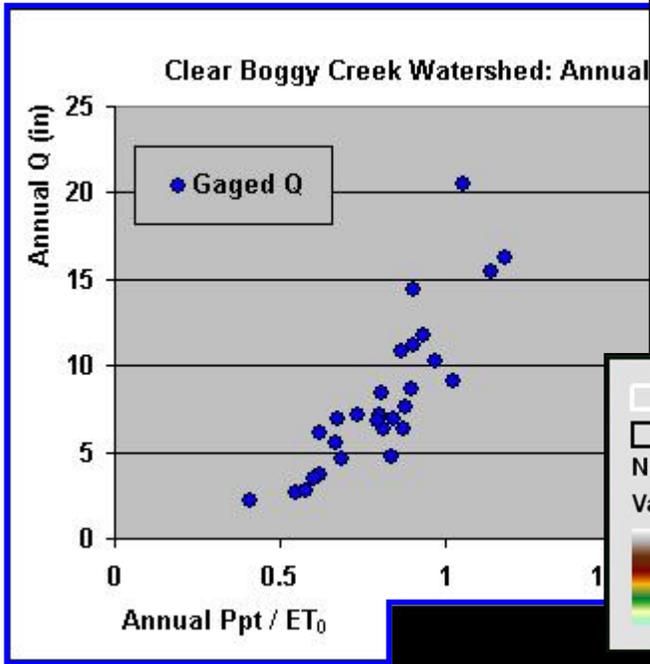
An extensive record of precipitation was available for the project area. The meteorological data sufficient for calculating ET<sub>0</sub>, however, was relatively small—few stations collect enough data for the calculation and those that do only cover the last ten years. Calibrating ET<sub>0</sub> to precipitation, however, allowed predictions of ET<sub>0</sub> to be made across the entire project area for the full record of precipitation. Quarterly and annual Ppt and ET<sub>0</sub> data were collected for model calibration. Quarterly and annual Ppt was spatially interpolated across the project area to produce raster GW<sub>r</sub> model inputs. A regression relationship between Ppt and ET<sub>0</sub> was developed and spatially interpolated in order to model a 50 year record of quarterly and annual ET<sub>0</sub>. Modeled quarterly and annual ET<sub>0</sub> were then spatially interpolated across the project area to use as intermediate GW<sub>r</sub> model inputs.

**Data provided from the work in Section A includes:**

- Annual kriged Ppt for 1960 – 2009 provided both as geodatabase and as individual rasters.
- Annual  $ET_0$  estimates for 1960 – 2009 provided both as a geodatabase and as individual rasters.
- Annual  $ET_0$  slope and intercept rasters in the model parameters geodatabase and as individual rasters.
- $ET_0$  calibration station locations point shapefile in the model parameters geodatabase and as a separate file.

# B. CALIBRATING $R_{INFIL}$ : SITE CHARACTERIZATION

## SOIL PERMEABILITY AND GAGED WATERSHEDS





## B. Site Characterization: Gaged Watersheds and Soil Permeability

Groundwater recharge (GWr) is the portion of precipitation that infiltrates the soil not lost through evapotranspiration. This is represented by a simple expression: GWr is the annual infiltrated rainfall ( $R_{infil}$ ) minus the water removed by annual evapotranspiration (ETa) (Equation B1). To model the first half of the GWr equation,  $R_{infil}$  was broken down into elements that could be quantified for the model. To better understand  $R_{infil}$ , two key parameters were analyzed for the model: (1) surface water discharge (Q) over representative gaged watersheds, and (2) soil permeability. Permeability functions as a proxy for the infiltration capacity of the soil, thus scaling the portion of rainfall permitted to infiltrate, and Q provides a measurement of the portion of rainfall that is not permitted to infiltrate for a given amount of rainfall (i.e., the portion that runs off from the surface that is then drained from the system).

$$GWr = R_{infil} - ETa$$

Equation B1

Permeability data were compiled for the model by evaluating the State Soil Geographic Database (STATSGO), the Natural Resources Conservation Service (NRCS) Web Soil Survey (WSS), and regional permeability factors described in literature concerning areas within the project region.

Modeling Q provided the framework for estimating  $R_{infil}$ , the residual of Ppt minus Q. Q data were compiled for the model using gaged watersheds in the project region and the Ppt and  $ET_0$  data described in Section A.  $R_{infil}$  was analyzed at the individual watershed scale to capture the Q-precipitation relationship.

This section will describe the data sources, processing techniques and evaluation of the Perm and Q data.

### B1. Permeability Dataset

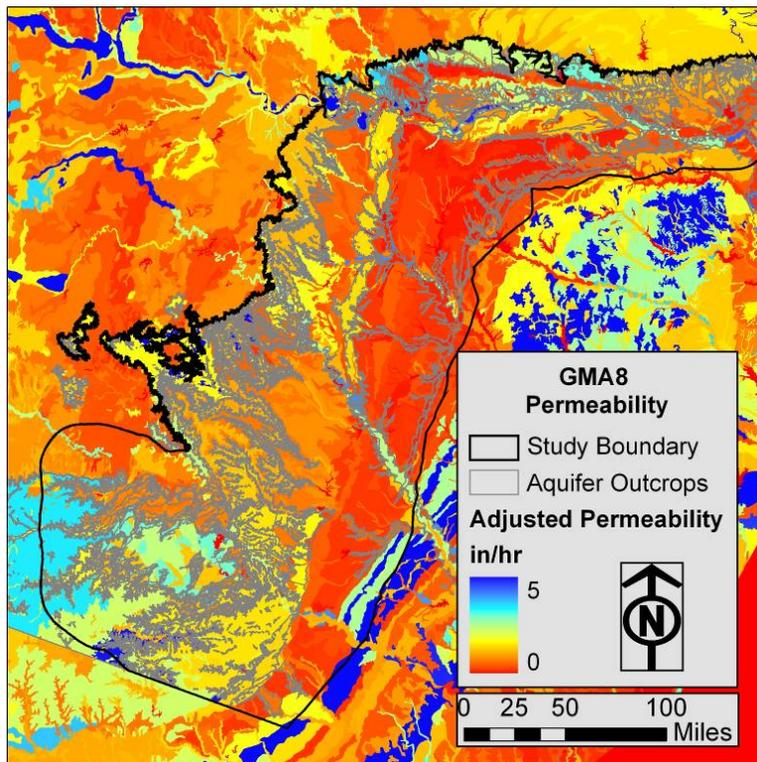
Quantifying the influence that surface properties have on GWr requires an understanding of both permeability and depth to bedrock. Because soils are the medium between precipitation and the aquifer itself, they are a major control for GWr. Mainly through the constraint of permeability, soils dictate the rate at which precipitation infiltrates the soil surface, the rate of discharge, the volume of water that is stored and available to plants, and the rate at which water moves through the profile either laterally or vertically. Permeability is a scalar value that was used to calibrate  $R_{infil}$  and to then apportion this calibration throughout the GMA8 region.

**Terminology Note:** Three terms are used to distinguish permeability data in this section. *Permeability* describes the raw numeric or descriptive data used to generate a model of the infiltration capacity of soils. *Average Weighted Permeability (AWP)* is the intermediate synthesis of the multiple layers of STATSGO soil data that have been reduced, by averaging, to a single numeric value to represent soil permeability. *Perm* is

the final value used to represent permeability within model calculations. Perm is equal to AWP across most of GMA8. Section B1 describes the process of collecting and evaluating permeability data. Subsections for each aquifer outcrop within the project region describe the validation process informed by a literature review and provide both methods and explanations for AWP adjustments made in select model regions.

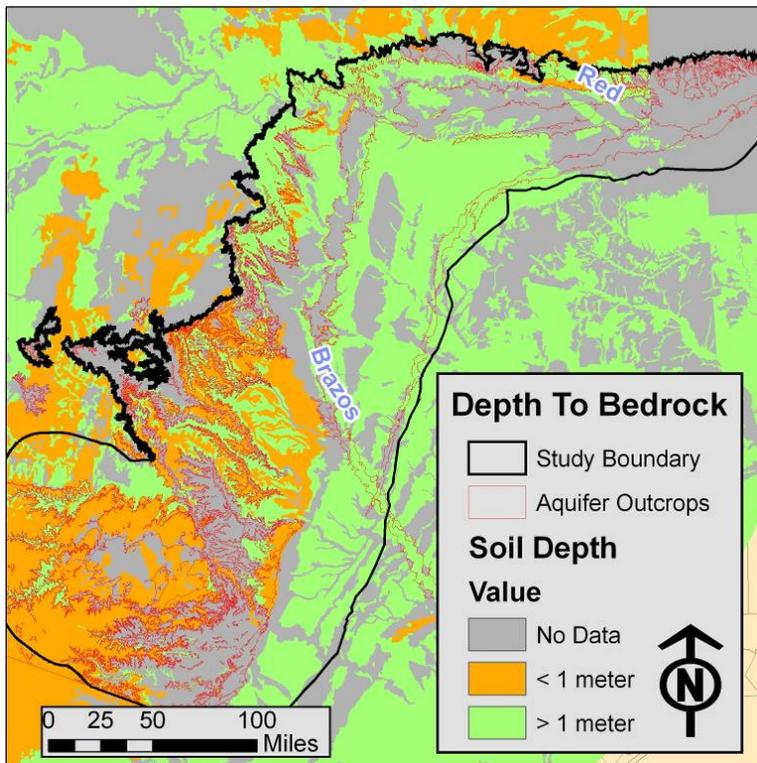
### B1.1 Analysis of Soils Data in STATSGO

Soil properties heavily influence the potential for recharge into outcropping aquifers because they influence how much water from precipitation runs off the landscape. STATSGO soil properties are generalized in a spatial soil database that defines, among other physical properties, permeability (Figure B1) and depth to bedrock (Figure B2) (PSU 2006). Because these data are what is available, STATSGO soil properties define the limit of our ability to assign parameters that influence recharge based on surface properties.



**Figure B1. Permeability rates were a major input to the HydroBio GWR model, and were used to represent surface properties. Permeability was a calibration parameter for stream discharge because such discharge is a factor of the water that is rejected from infiltration.**

Many methods to interpret the STATSGO data were proposed and tested to best represent recharge and discharge potential. STATSGO data were not entirely consistent for sampling and reporting depths. Soil properties were reported in 11 standard soil layer increments, from the uppermost 2 inches to 98 inches deep. As an example, many STATSGO components were only reported to 60 inches despite having deeper horizons. Due to data constraints and reasons to be discussed later, we chose to only examine the top 39 inches (1 meter) of the soil data presented in STATSGO.



**Figure B2. Depth to bedrock varies across the GMA8 study area. Soil depth influences recharge potential and was used to adjust permeability values. Soils less than one meter in depth were not adequately represented in the GWR model using simple AWP, and therefore were adjusted based on depth to bedrock. AWP is described below.**

### **B1.2 Reduction of Multi-Layer Soil Database to a Single Representative Layer**

A weighted average of the top one meter of the soil profile was chosen as the base soil metric to account for the limiting properties of the clayey B horizon and the influences of the surface layers. The permeability of each standard soil layer was weighted by thickness of the layer and averaged to both average and weight the permeability. The weighted permeability for all soil layers was represented as a single value that was given the name “average weighted permeability” (AWP). While this AWP method was determined to be reasonable for the majority of the soils across the GMA8 study area, unique regions required special attention and adjustments—regions where AWP did not sufficiently represent soil influences upon recharge in soils with highly variable permeability rates, and in soils shallower than one meter.

Soil analyses were limited to the top meter (39 inches) of the soil profile because it contained the soil horizons most important to rejecting / accepting infiltration and because STATSGO data were the most reliable within the top meter:

- The top one meter depth within STATSGO encompassed clayey B horizons that limit infiltration and retain the infiltrating water.
- The top of STATGO data had the fewest data gaps and questionable data values.
- Intake properties of the soil control the potential for recharge of aquifers, these properties are well defined by the top meter of soil.

- Initial calculation showed that 39 inches should be a sufficient depth to describe the potential for recharge:  $39 \cdot 0.3$  retention at field capacity = 11.7 inches retention. Retained water is then extracted from the soil by ETa throughout the year.

### **B1.3 AWP for Shallow Soils and/or Highly Variable Permeability**

Weighting for the effect of underlying aquifer bedrock was applied in the model for soils that were shallower than one meter. Where shallow soils overlay aquifer outcrops, flow into the aquifer could occur more readily. An assumption was made that the infiltration capacity through the underlying rock was non-limiting, for example the presence of cracks and joints would transmit water downward that could not be held within the shallow overlying soil. This calculation was made to represent conditions overlying aquifer outcrops that were believed to readily accept infiltration.

During the AWP calculation, STATSGO values for shallow rocky soils contained null values that did not enter the calculation for soil properties, only the layers with numeric values were represented in AWP. The STATSGO soil properties are relational (scalar) values that effectively treat all soils as having a full one meter thickness. In terms of transmitting water, however, the thinner the layer, the less the retention and the faster the travel time; factors that enhance recharge.

To account for the increased recharge potential in shallow soils, STATSGO depth to bedrock data were used to delineate soils shallower than 1 meter. Based on the knowledge that the shallow soils permit faster and more efficient recharge to the aquifer (Cooper 1990), HydroBio set out to adjust modeled permeability accordingly. All soil depths less than a meter to bedrock were adjusted to reflect higher actual recharge potential. As soil depth to bedrock decreases, the adjustment factor to permeability increases. Adjustments are based on Equation B2.

$$\text{Adjusted Permeability} = (1 \text{ meter} / \text{STATSGO Depth to Bedrock}) \times \text{AWP} \quad \text{Equation B2}$$

Adjustments to recharge potential were also made for soils with high permeability variability. Within the Nacatoch and Blossom aquifer outcrops, for example, very slowly permeable horizons exist within the first meter of soil below much higher permeability layers at the surface. Within these outcrops, recharge rates are dictated by the limiting clay horizons. Applying the AWP for highly variable soil profiles containing very slowly permeable layers artificially inflated recharge potential due to the coarse textured material at the surface, therefore the permeability of the most limiting layer was used in place of AWP to represent the recharge potential for the entire profile as dictated in these circumstances by the limiting clayey B horizon.

## **B1.4 Proofing STATSGO**

Although the STATSGO database provides highly useful information, it may contain errors. Suspect data, identified by searching for anomalous high or low values, were compared with soil surveys available through the Natural Resources Conservation Service (NRCS) online service in the form of Web Soil Survey (WSS) (NRCS 2011). WSS is an online collection of the most up-to-date soil surveys performed by NRCS. Soil data were organized as map units of discrete soil series or complexes of multiple soil series. Reported soil series properties included horizon formation, permeability, drainage, parent material, climate, water holding capacity, landform type and location, frequency of flooding, depth to water table, and depth to bedrock. Average permeability, permeability of the most limiting layer, and depth to bedrock were available through both STATSGO and WSS. Since Web Soil Survey contained the most up to date soil data and was less generalized, it was used to adjust raw STATSGO data where anomalous values were detected. Over aquifer outcrops, STATSGO and WSS data were evaluated using random sampling and specific sampling in areas with unusually high or low values; evaluation found good agreement between the two data sources.

In the calibration process, suspect STATSGO-derived permeability values were compared to permeability values reported on WSS. If discrepancies were discovered, the raw STATSGO permeability values were adjusted to reflect the smaller scale soils description in WSS. WSS was also used to cross check depth to bedrock values found using STATSGO. Table B1 summarizes instances of inaccurate STATSGO data detected and corrected using WSS.

The characteristics of the soil profiles described by STATSGO data and NCRS soil surveys are presented in the sub-sections that follow. These sections, presented by aquifer outcrop, explain the additional steps in processing STATSGO data to develop the permeability dataset, model Perm. In the reduction of the multi layer STATSGO database to a single representative unit, one of three different metrics presented at the beginning of this section was employed depending on the unique conditions within each outcrop. (1) AWP, the first metric, was applied across the entire project area before soils were analyzed by aquifer outcrop. After analysis, in some cases, AWP was replaced with one of the other two metrics: (2) AWP modified using equation B2 to adjust permeability upward based on depth-to-bedrock for shallow soils, or (3) the use of the most limiting soil layer from highly variable profiles in which the effects of restrictive layers such as clayey horizons that were not adequately reflected in AWP. The metrics used to represent permeability for each aquifer outcrop are summarized in Table B1 and detailed in the following sub-sections.

**Table B1. Overview of final metric used to represent permeability in the model for each aquifer outcrop within the model.**

Aquifer Outcrop	Perm Metric Used	Adjustments to STATSGO
Trinity	AWP, depth-to-bedrock applied to shallow soils in southern Trinity.	Unreasonably high permeability rates reported by STATSGO in Comanche, Erath, and Hood Counties were upwards of 6.5 in/hr—values only possible in alluvial soils, not the weathered shale and sandstone derived soils of this area. NCRS WSS measurements provided replacement permeability rates ranging from 0.5 - 0.75 in/hr.
Edwards-Trinity (Plateau)	AWP with depth-to-bedrock adjustment for shallow soils.	none
Edwards Balcones Fault Zone	AWP with depth-to-bedrock adjustment for shallow soils.	none
Nacatoch	Most limiting layer used for highly variable soil layers.	none
Woodbine	AWP.	none
Blossom	Most limiting layer used for highly variable soil layers.	none
Paleozoics	AWP with depth-to-bedrock adjustment for shallow soils.	none

#### **B1.4.1 Analysis of STATSGO Data—Trinity Aquifer**

The Trinity outcrop spans the entire western and northern border of the GMA8 project area, and displays a wide range of recharge values attributed to variability in precipitation, ET, lithology, and soil properties. Soils and permeability vary greatly between the northern extent of the aquifer in Oklahoma and Arkansas, and the shallow, calcic soils of the southern extent. The soil properties that influence recharge that were derived from STATSGO data are addressed in this section for each general portion of the Trinity Aquifer outcrop.

The northern extent of the Trinity outcrop, located north of the Red River, exhibits significantly different soil properties than the remainder of the outcrop. The alluvial valley along the Red River is dominated by sandy and loamy alluvium that promotes high

permeability and associated recharge rates. Moving north from the Red River, higher landform positions contain weathered sandstone and shale of the Ouachita Mountains and soils consist mainly of clay residuum from weathered shale. These sediments restrict permeability and recharge in contrast to the high permeability of the sandy and loamy alluvium along the river.

The portion of the Trinity outcrop extending from Tarrant County, west of Dallas, southward to Lampasas County is characterized by weathered shale and sandstone of at least 1 meter deep. This promotes moderately low permeability and recharge rates. The STATSGO data contained a number of major data flaws in Comanche, Erath and Hood counties that erroneously inflated permeability rates. STATSGO reported permeability rates upwards of 6.5 in/hr, which is possible only in coarse alluvial sediments, but not soils derived from fine sandstone and shale residuum. This was confirmed by permeability rates reported in soil survey data found on Web Soil Survey (NRCS 2011). Permeability rates were decreased from upwards of 6.5 in/hr, to values in the 0.5 – 0.75 in/hr range, as reported in the soil surveys.

Sediments overlying the Trinity aquifer outcrop are markedly different in the segment between Lampasas and Bexar Counties (San Antonio). STATSGO Soils here were generalized as shallow, well drained, moderately permeable limestone residuum. Areas that are shallower than 1 meter, the depth that Equation B2 considers, were adjusted to reflect the increase in recharge potential because the soils have relatively poor water holding capacity and are shallow. The Edwards Plateau portion of the Trinity aquifer outcrop is characterized by shallow, limestone dominated soils. The influences of these features and properties on recharge potential are discussed in the Edwards-Trinity Plateau Aquifer section.

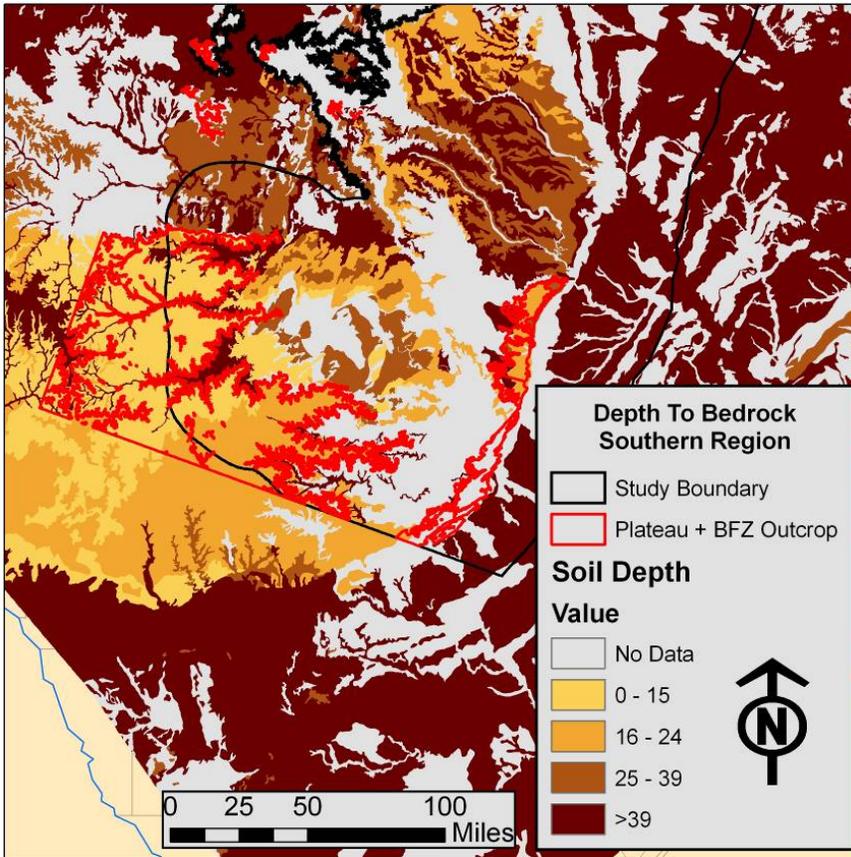
#### **B1.4.2. Analysis of STATSGO Data—Edwards-Trinity (Plateau)**

The Edwards-Trinity Plateau outcrop is unique within the GMA8 study area. The unique characteristics are due to soils that are relatively shallow, well drained and have poor water holding capacity. These properties are related to high recharge rates due to the absence of significant soil water retention. Soils occurring over the other aquifer outcrops in the GMA8 region are generally deep and provide a buffer between precipitation and recharge due to storage capacity. Water held in the vadose zone is also available to vegetation and is prevented from recharging underlying aquifers by ETa. Because the typical soils across GMA8 are deeper than one meter, the shallow soils in the Edwards-Trinity Plateau were treated differently in the GWr model.

The permeability term in the HydroBio GWr model is based on the average weighted permeability (AWP) of the top 39 inches of soil. The shallow soils of the Edwards-Trinity Plateau were not sufficiently represented using AWP so permeability rates were adjusted as a function of depth. The soils of the Edwards-Trinity Plateau are moderately permeable, but due to their shallow depth and low water retention, precipitation should travel relatively easily through the soil and into the bedrock outcrop. The outcrop is

commonly comprised of weathered limestone, which is not a limiting permeability layer, and therefore the soil profile was considered well drained.

Very small gaps in the STATSGO depth to bedrock data existed within the Edwards-Trinity Plateau outcrop, as seen in Figure B3. These areas were investigated using WSS and gaps were filled, where appropriate. Fortunately, most areas that had missing depth to bedrock data consisted of soils deeper than 39 inches and, therefore, did not require adjustment.



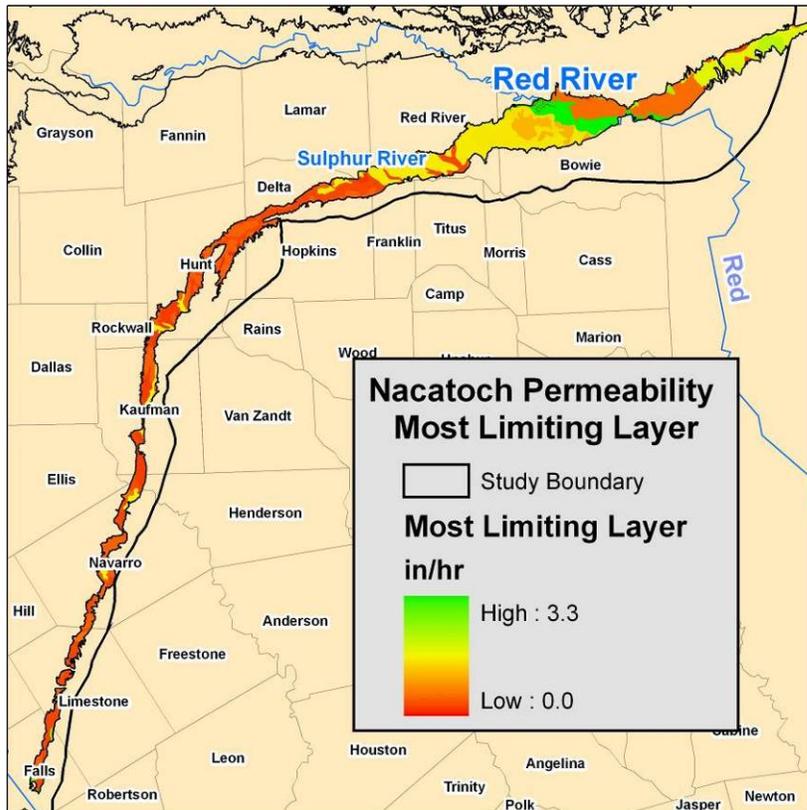
**Figure B3. Depth to bedrock is relatively shallow in the southern model region, which includes the Edwards-Trinity Plateau and Balcones Fault Zone. Shallow soils reject less precipitation to recharge, and therefore average permeability rates were adjusted as a function of depth.**

### **B1.4.3. Analysis of STATSGO Data—Edwards Balcones Fault Zone**

The soil and lithologic properties of the Edwards Balcones Fault Zone are similar to those of the Edwards-Trinity Plateau and were handled similarly. Recharge zones primarily consist of sinkholes, faults, and crevices, overlain by shallow, well-drained soils. Sinkholes and faults were generally associated with shallow soils, and permeability rates for shallow soils were adjusted upward because limestone bedrock generally provides no resistance to water intake. Equation B2 was applied to shallow soils over this aquifer outcrop, thereby enhancing Perm.

#### B1.4.4 Analysis of STATSGO Data—Nacatoch Aquifer

The literature suggests recharge rates over the Nacatoch Aquifer outcrop are limited due to very slowly permeable layers covered by coarser textures that would, otherwise, enhance infiltration (Ashworth, 1995). This coarse material skews the actual recharge potential when modeling the AWP of the upper meter of the soil profile. Analysis of the STATSGO data and most limiting layer of typical profiles across the outcrop confirmed the occurrence of very slowly permeable layers buried with significantly higher permeable sediments with the exception of an infiltration-accepting region in alluvial soil profiles in the northeastern portion of the outcrop (Figure B4).



**Figure B4. The most limiting permeability layer of the Nacatoch soils greatly influences recharge potential and was therefore used as the permeability input for the GWR model. In the northeast region of the outcrop the lowest permeability soil layer (the most-limiting layer) accepts moderate amounts of infiltration in some areas of river alluvium.**

The AWP across the outcrop averages 0.8 in/hr, which is relatively slow compared to the remainder of the GMA8 study area. However, the permeability of the most limiting layer averages 0.43 in/hr, including alluvial material associated with the Red River which is consistently permeable throughout the soil profile. The AWP for the soil profile over the Nacatoch outcrop was up to four times higher than the permeability of the most limiting layer due to the high variability in permeability rates within the profile. Due to the large difference between limiting layer permeability and the average weighted permeability, the limiting layer was used to dictate recharge rates. This rejects precipitation in areas where high variability in the profile obscured the affect of a very slowly permeable layer in the averaging procedure.

The Red River soils consist of sandy and loamy alluvium with high permeability rates comprising the majority of the roughly one third of the aquifer outcrop that is capable of accepting recharge. Another notable section of high permeability is the intersection of the outcrop with the Sulphur River in southern Red River County. Other such coarse alluvial soils exist along stream intersections within the outcrop and these locations are reflected in the STATSGO data with high permeability. STATSGO soil profiles consisting of sandy and loamy alluvium throughout the profile in the Red River area lacked infiltration-limiting layers in deeper profile positions, thus allowing recharge in these areas. Within the Red River vicinity, the lowest permeability soil layer did not limit infiltration and use of the lowest permeability layer as the model input Perm did not significantly differ from values derived from using AWP.

A small portion of the outcrop that has moderately permeable soils exists in central Bowie and Red River counties. These soils are not associated with alluvium, however, and are comprised of residuum from weathered shale and sandstone. In these areas as for the alluvium, the permeability through the soil profile did not vary significantly and the most limiting layer was notably different from the AWP. Application of the most limiting permeability layer for Perm values in the Nacatoch was chosen to allow limiting layers in the profile to hinder infiltration where present while allowing infiltration in areas of alluvium lacking infiltration-limiting soil layers.

#### **B.1.5. Analysis of STATSGO Data—Woodbine Aquifer**

The AWP of the top meter of soil across the Woodbine outcrop is generally high. AWP ranges from 0.3 in/hr in clayey soils and up to 5 in/hr in alluvial sediments. The average permeability over the entire outcrop is 1.42 in/hr, significantly higher than the surrounding landscape and outcrops. Another important factor in determining potential recharge from soil permeability is the capacity of the most limiting layer to transmit water. Across the Woodbine outcrop, even the most limiting soil layer is relatively more permeable than the soils of the surrounding landscapes and outcrops, suggesting that high recharge rates are possible.

Along the north-south stretch of the Woodbine aquifer outcrop, there is low variability in the permeability rates. Most of the variability occurs along the Red River within the east-west stretch of the outcrop where the most permeable soils occur due to the coarse alluvial material associated with the river. However, there is a large patch of low permeability soils at the northernmost point of the outcrop located in northwestern Bryan County, OK (Bryan is located above Grayson and Fannin Counties in Texas; Figure B5).

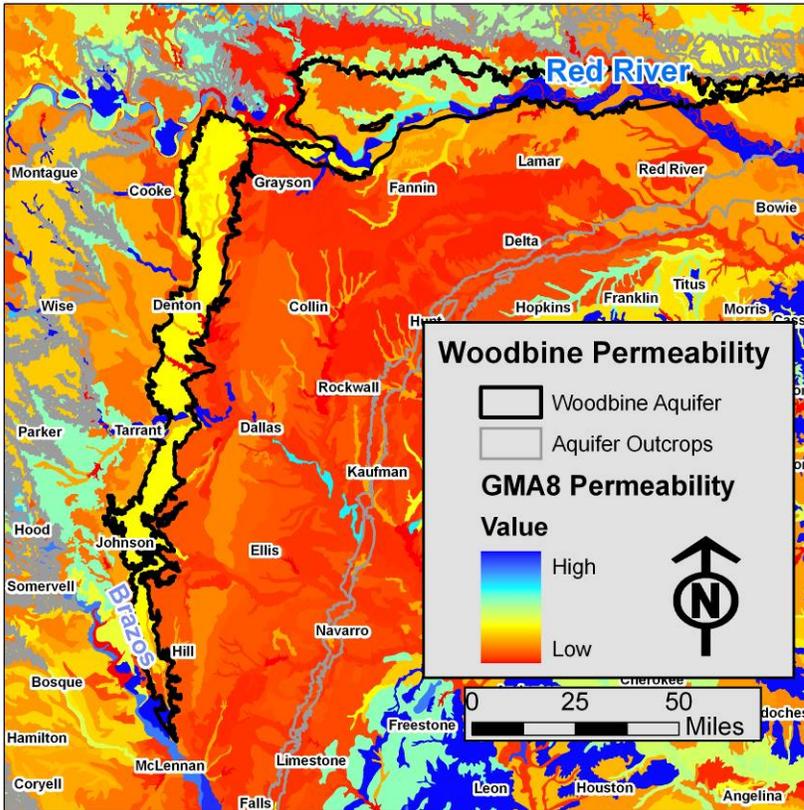


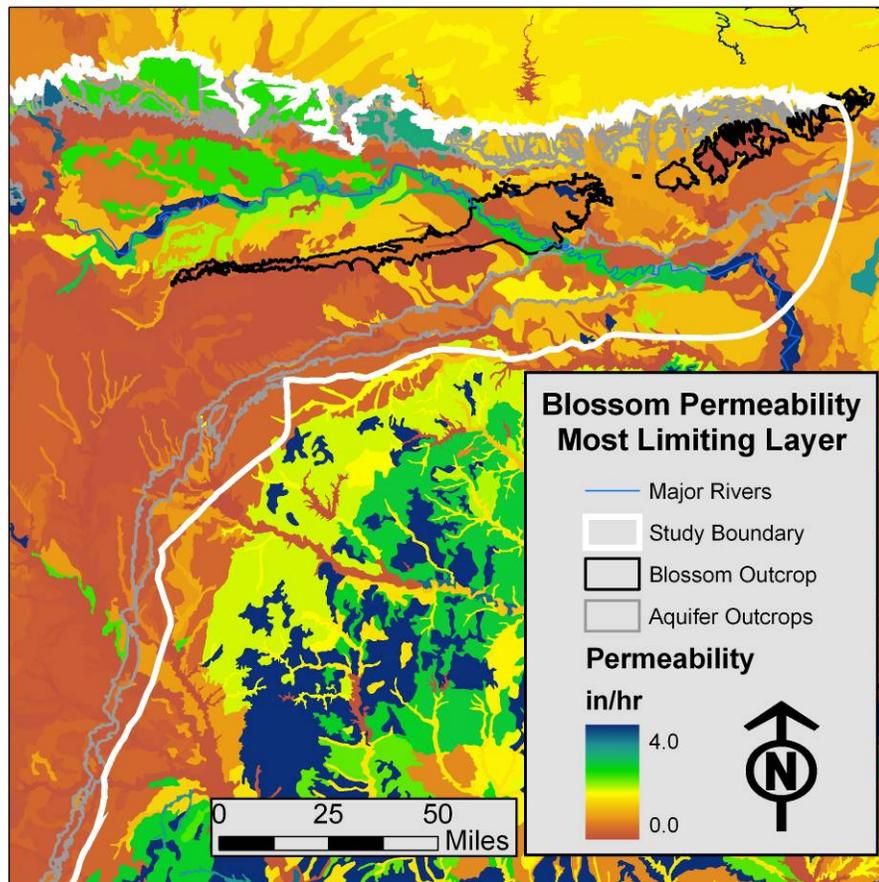
Figure B5. Perm rates of the soils overlaying the Woodbine Aquifer are significantly higher than surrounding soils.

### B.1.6. Analysis of STATSGO Data—Blossom Aquifer

Soils overlying the Blossom Aquifer closely resemble the properties of the soils overlying the Nacatoch Aquifer. Review of the STATSGO and soil survey data revealed that very low permeability horizons exist below relatively high permeability near the soil surface. Due to the occurrence of these limiting layers, the standard application of AWP for the top meter in the GWr model was not appropriate for this outcrop. Therefore, as applied for the Nacatoch outcrop, the most limiting layer approach was selected to represent the permeability input for the GWr model.

Like the Nacatoch outcrop, highly permeable soils also exist within the Blossom outcrop. These high permeability soils are mostly associated with coarse alluvium occurring in proximity to major streams. Around 10% of Blossom outcrop soils consist of Red River alluvium (Figure B6). Representing the soil properties with the most limiting layer did not greatly alter the recharge capacity of the alluvial sediments, since they are relatively evenly permeable throughout the top 39 inches of the soil.

**Figure B6. The most limiting layer was used to represent permeability in the Blossom aquifer outcrop due to the large difference between surface texture and fine textured subsurface soils.**



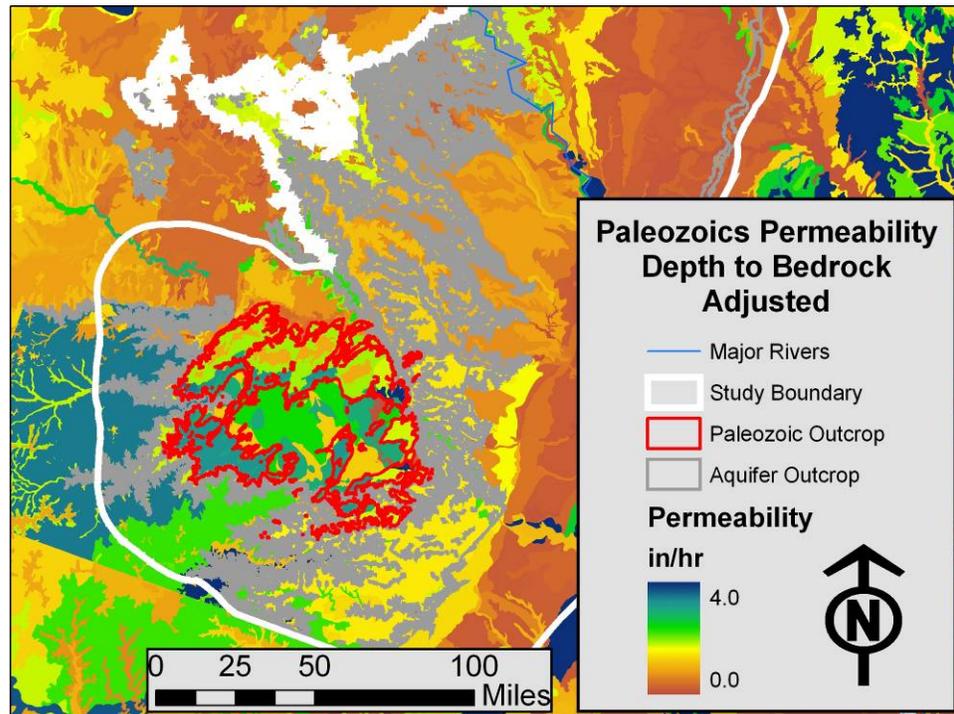
### **B.1.7. Analysis of STATSGO Data—Paleozoic Aquifers**

The soils overlaying the Paleozoic outcrops are shallow and contain high percentages of coarse material. Analysis of the STATSGO depth to bedrock data revealed extremely shallow, sandstone-derived soils to the southwest of the formation transitioning to deeper, limestone derived material to the north (Figure B7). After the analysis of soil survey data, we concluded that permeability rates should be adjusted upwards to account for the shallow soils and rapid recharge potential of the area. In particular, the southern half of the Paleozoic outcrops required permeability adjustment because most soils are less than 20 in, with some areas having only 4 inches of soil overlying bedrock. The soils with sandstone parent material are sandy/loamy at the surface and clayey in the subsurface. Normally the clay horizons would limit permeability; however these soils contain high percentages of coarse material such as gravel that provide pathways for infiltration though the profile.

Soil surveys also described the soils in both the northern and southern portions of the Paleozoic aquifers as well drained, which further supported the application of a depth to rock permeability adjustment for the outcrop. Water holding capacity was very low,

which is partly a consequence of shallow soils. The combination of well drained, shallow, and poor water retention of these soils suggested that potential recharge would be high in this region. If precipitation moves quickly through the thin layer of soil above bedrock, then that water has little potential to be disposed by plants or surface discharge. Due to these conditions, Equation B2, adjusting for depth to bedrock, was also applied to this region to better represent recharge potential and to maintain consistency in permeability adjustments made for similar soils occurring above the other aquifer outcrops.

**Figure B7.** The soils overlying the Paleozoic outcrops were adjusted upwards to reflect higher recharge potential from shallow soils.



## B2. Individual Watershed Q Data, the Basis for Calibrating for Perm

Gaged watersheds were analyzed to gain an understanding of Q conditions within small defined areas. Individual watersheds capture data representing physical relationships between model variables. If  $R_{infil}$  can be understood for a multiple small watersheds, this knowledge can be extrapolated for similar watersheds based on the underlying Perm as represented in STATSGO and regional climates affecting the humidity status of that location as typified by  $Ppt/ET_0$ .

### B2.1. Selecting Representative Watersheds

The individual watersheds used for  $R_{infil}$  calibration were chosen based on criteria to assure that each watershed would produce representative discharge measurements. To limit our data set to detect clear and interpretable relationships, we selected watersheds

most likely to represent the “pristine” conditions for Ppt, Perm, and Q. Those watersheds with significant disturbances such as urbanization or irrigation, likely to influence Q behavior, were removed. Thus, a representative watershed should provide a clear signal of the interaction between Ppt, Perm, and Q without the noise potentially introduced by disturbances to the natural conditions. Representative watersheds selected from the 400 USGS gages in the GMA8 region had:

- **Adequate discharge record** for use with USGS streamflow partitioning program (PART). 150 of the 400 gages had continuous daily discharge records. 80 of these had greater than 3 years of continuous data recorded during or after 1960.
- **Small, well-constrained watershed boundary** (1000 mi<sup>2</sup> or less). Forty of the pool of 80 gages that had adequate discharge records met the size criteria and were non-redundant (where multiple gages were present on a single stream, the gage with the most continuous data was selected).
- **No disturbances in the watershed.** A survey of the remaining 40 watersheds was conducted on images from Google Earth to remove watersheds with significant flow pattern disturbances, such as irrigation, upstream reservoirs, and urbanization.

## B2.2. Processing Individual Watershed Data

After reducing the USGS gages within GMA8 to a subset of representative watersheds, individual watershed boundaries were outlined using the National Elevation Dataset layer (USGS, Accessed 2010). A polygon shapefile was generated for each watershed. Daily discharge measurements were then downloaded and organized as text files (USGS, Accessed 2010). And finally, the USGS streamflow partitioning program (PART) was used to sum annual discharge and estimate annual base flow from daily discharge text files (USGS, Accessed 2010) (Figure B8). PART normalizes discharge measurements to the watershed area and reports discharge and base flow in linear units.

The discharge and base flow data were joined with the variables Perm, ET<sub>0</sub>, and Ppt. Watershed polygon shapefiles were used with the ArcMap zonal statistics tool to estimate the mean Perm, ET<sub>0</sub> equation variables, and annual Ppt for the 50 year model record within each watershed. A table was populated for each watershed combining all annual data (Ppt, measured Q, estimated BF) as well as static spatial data (slope and intercept of ET<sub>0</sub> equation, Perm). Table B1 summarizes the data for the 24 watersheds used in R<sub>infil</sub> calibration.

Initial evaluation of each watershed data table consisted of plotting annual Q measurements as a function of annual Ppt. During this evaluation, several potential watersheds were eliminated for being non-representative. Some watershed data showed high scatter which obscured the Q and Ppt relationship. Other watersheds showed variability that was clearly driven by factors other than Perm and Ppt (Figure B9). For

example, the Comal River gage was not useful for project-wide predictions of Q because nearly all of the flow gaged at Comal is derived from springs (Wahl, accessed 2011). The initial evaluation process reduced the representative watersheds to the 24 watersheds described in table B1.

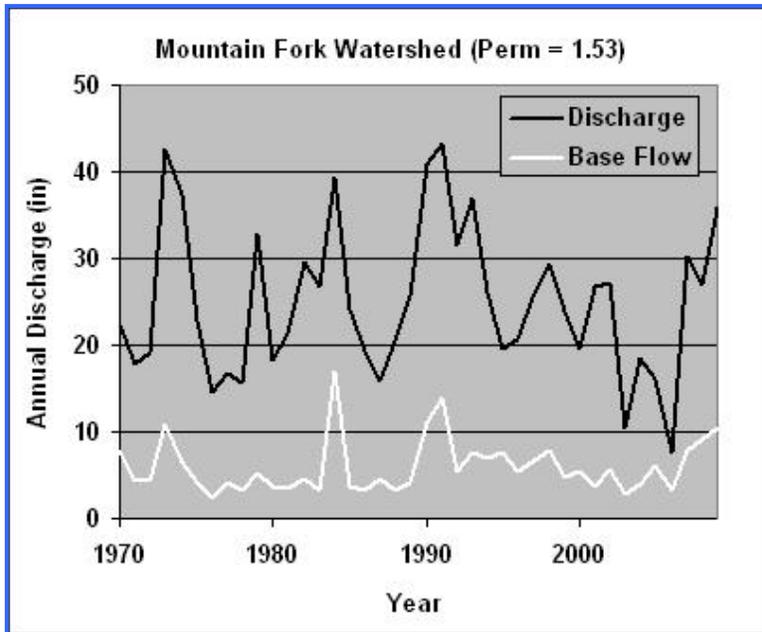


Figure B8. Example of USGS PART streamflow partitioning results for a moderate permeability basin. The program estimates base flow, in inches, from daily gaged discharge volume and watershed area.

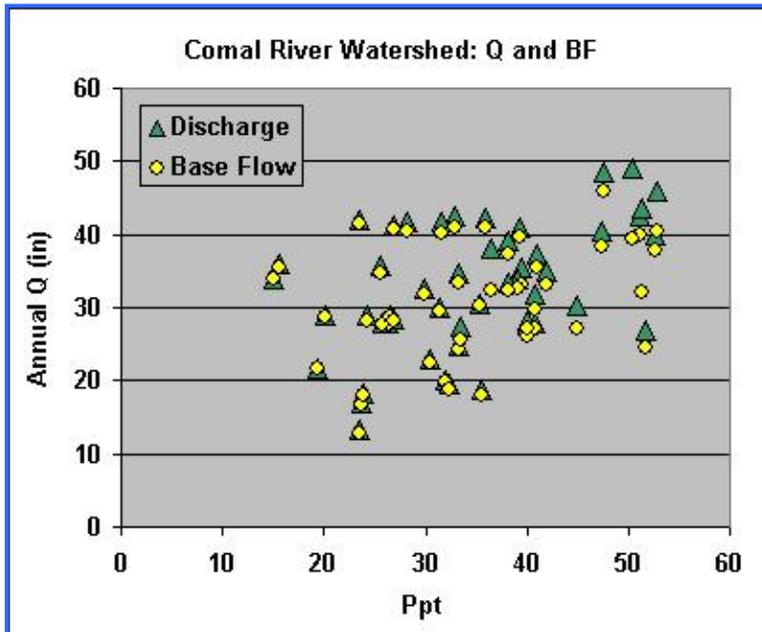


Figure B9. The measured Q and predicted BF record for the Comal River Watershed are not useful for calibrating  $R_{infil}$ . The Q record is invariant with precipitation and BF estimates are nearly equal to the measured Q. Nearly all of the flow gaged at Comal River is derived from springs (Wahl, accessed 2011). Comal is therefore not representative of the surface and climate conditions used to develop the  $R_{infil}$  model.

The gages presented in Table B2 were used to calibrate Q to annual climate conditions and to spatially extrapolate Q conditions beyond the representative boundaries. The temporal and spatial calibration process is described in the Section C: Spatial Extrapolation.

**Table B2. Overview of gaged watersheds used in Rinfil calibration. USGS number, name, drainage area, and discharge (Q) data were downloaded from USGS NWIS (National Water Information System). The Q record is the period over which continuous daily discharge data was available. Base flow (BF) was estimated using the USGS streamflow partitioning program PART. Permeability was averaged over each watershed from a raster adapted from the STATSGO soil database (described in Section B1. Precipitation (Ppt) was averaged over each watershed from rasters kriged from the Digital Climate Atlas point precipitation database.**

USGS Number	USGS Name	Drainage Area (mi <sup>2</sup> )	Q Record	Perm (in/hr)	Avg. Ppt in/yr	Q Avg. in/yr	BF Avg. in/yr
08150800	Beaver Ck nr Mason, TX	215	1964 - 2008	2.28	29.18	1.22	0.41
08065800	Bedias Ck nr Madisonville, TX	321	1968 - 2009	1.34	43.97	9.13	0.84
08086290	Big Sandy Ck abv Breckenridge, TX	280	1963 - 2009	0.71	27.97	1.37	0.06
08044000	Big Sandy Ck nr Bridgeport, TX	333	1960 - 1994, 2005 - 2009	1.14	33.47	2.72	0.83
07332500	Blue River near Blue, OK	476	1960 - 2009	0.95	41.27	9.33	3.31
07332600	Bois D Arc Ck nr Randolph, TX	72	1963 - 1984	0.24	41.07	10.3	1.82
07344486	Brushy Ck at Scroggins, TX	23	1978 - 2003	2.89	45.78	11.2	4.53
07335000	Clear Boggy Creek near Caney, OK	720	1988 - 2009	0.92	40.13	8.81	3.15
08139500	Deep Ck nr Mercury, TX	44	1960 - 1972	0.88	25.49	1.39	0.19
07315200	E Fk Little Wichita Rv nr Henrietta, TX	178	1964 - 2009	0.99	31.64	2.21	0.10
08109800	E Yegua Ck nr Dime Box, TX	244	1963 - 2009	6.57	37.43	3.72	0.98
08094500	Green Ck nr Alexander, TX	46	1960 - 1972	0.86	30.32	1.73	0.49
08058500	Honey Ck nr McKinney, TX	39	1960 - 1972	0.26	36.75	7.12	3.14
08103800	Lampasas Rv nr Kempner, TX	818	1963 - 2009	1.16	30.66	2.80	1.36
08158700	Onion Ck nr Driftwood, TX	124	1980 - 2009	1.48	35.37	5.80	4.02
08172400	Plum Ck at Lockhart, TX	112	1960 - 2009	0.32	35.45	5.89	1.79

### **B3. Summary**

The GMA8 model site was characterized for the GAM model through an evaluation of soil permeability and gaged watersheds within the model boundaries. The STATSGO soil database, the NRCS Web Soil Survey, and a review of the literature describing permeability and infiltration capacity of individual aquifers within the GMA8 model boundaries were used to assemble and validate a raster layer to represent soil permeability for the GWr model. USGS stream gages were selected to represent discharge conditions across GMA8 based on data continuity, watershed disturbance, location, and size criteria. Data from the selected representative watersheds was collected from USGS stream gages and mated with Perm, Ppt, and  $ET_0$  data in preparation for model Q calibration.

#### **Data provided with this report pertaining to Perm:**

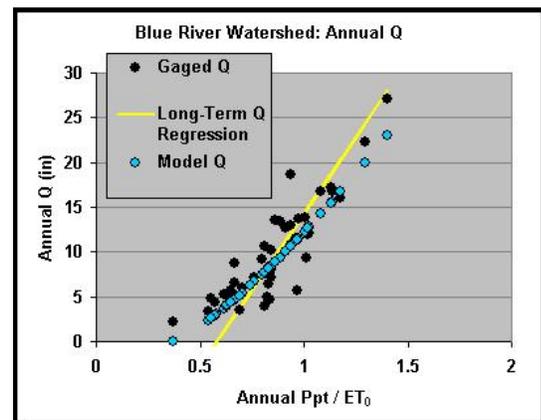
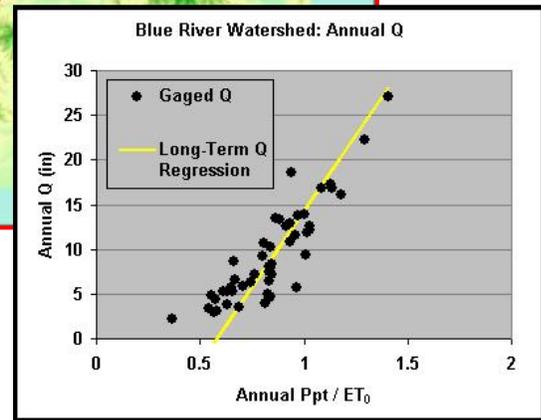
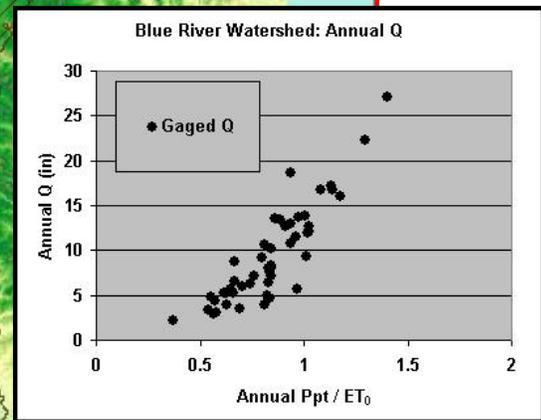
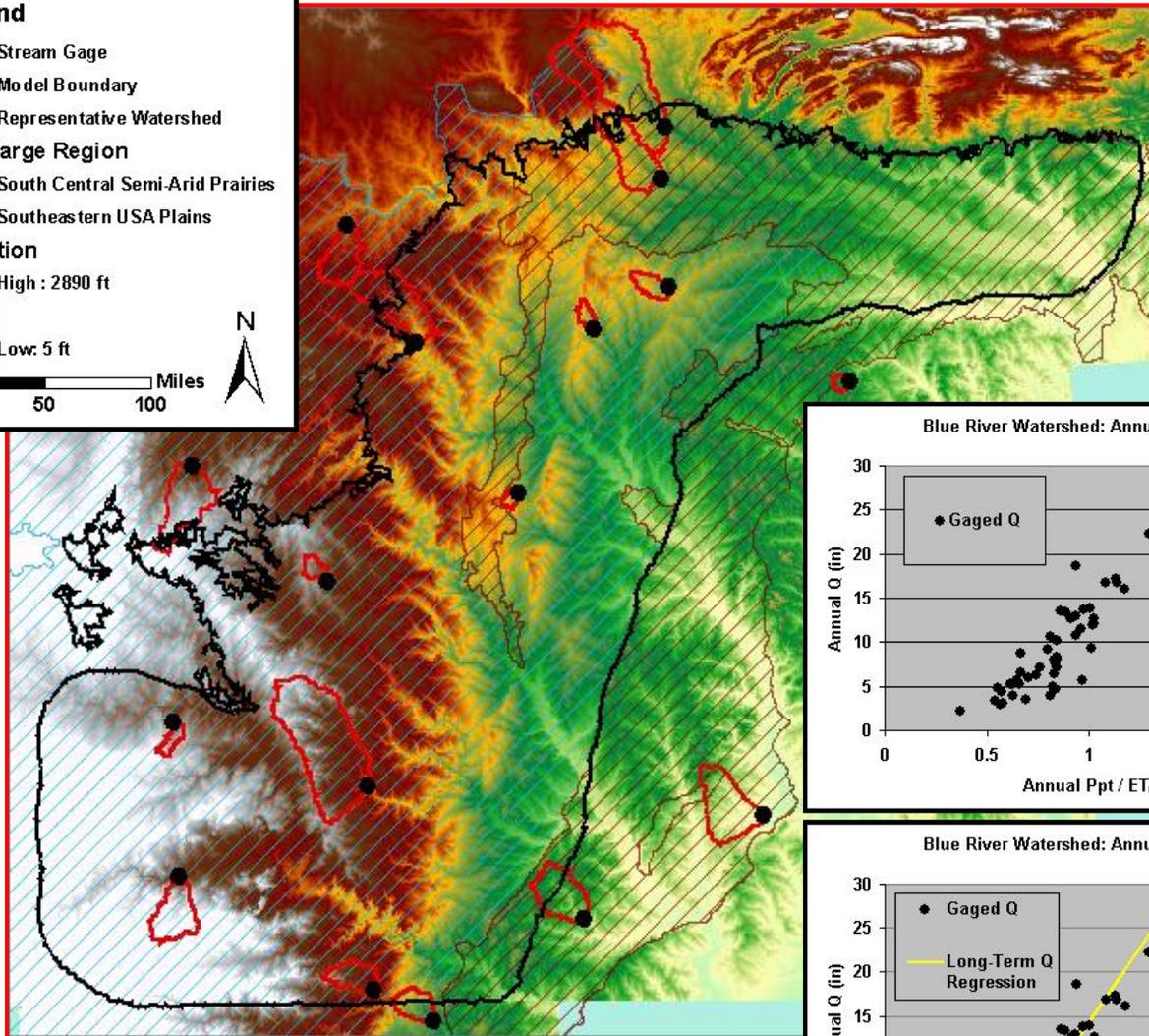
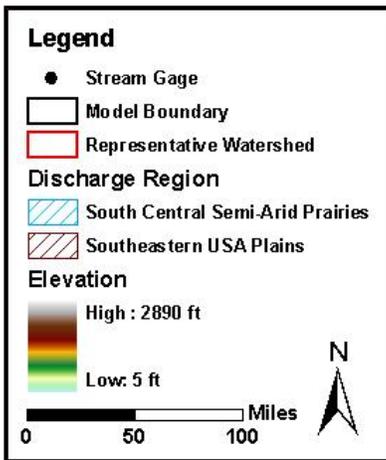
- Permeability rate (in/hr) raster both in the model parameters geodatabase and as an individual raster.

#### **Data provided with this report pertaining to Q:**

- Shapefiles of calibration watersheds for USGS gage locations and model parameters geodatabase.
- Shapefiles of calibration watershed boundaries and model parameters geodatabase.
- Calibration watershed gage data including quarterly and annual Q, Ppt, and  $ET_0$  in an Excel spreadsheet, Individual Watershed Analysis.
- Q estimates for 1960 – 2009 provided both as a model variable geodatabase and as individual rasters.



# C. CALIBRATING $R_{INFIL}$ AND SPATIAL EXTRAPOLATION OF Q





## C. Calibrating $R_{infil}$ : Spatial Extrapolation of Discharge

In Sections A and B we described the acquisition of the data for, and preparation of,  $R_{infil}$ . The concept of the variable  $R_{infil}$  parameter was developed to represent the total precipitation (Ppt) accepted by the soil and available to evapotranspiration over the year (ETa) in order to model groundwater recharge (GWr) at the coarse temporal and spatial resolution of the groundwater availability model (annual time-steps and square mile grid cells). The driving variable for the hydrologic processes governing GWr over the project region is precipitation (Ppt). The amount of Ppt that enters the soil system, infiltrated rainfall ( $R_{infil}$ ) is the critical parameter to estimate. This simplified water balance reduces the recharge system to three variables that explain GWr (Equation C1).

$$GWr = R_{infil} - ETa$$

Equation C1

In this section we will present the assembly of a precipitation-driven spatial model of  $R_{infil}$ . Modeled Q provides the framework for estimating  $R_{infil}$ , the residual of Ppt minus surface water loss (Equation C2). Q was chosen to represent surface water loss and was analyzed at the individual watershed scale to capture the Q-Ppt relationship. After defining Q-Ppt relationships for representative watersheds, we used the physical relationships between Q, Perm and climate (represented as the 50-year average Ppt/ET<sub>0</sub> relationship) to spatially extrapolate Q estimates across the GMA8 region. The following sections detail the procedure used to model  $R_{infil}$ .

$$R_{infil} = Ppt - \text{Surface Water Loss}$$

Equation C2

### C1. Estimating $R_{infil}$ —What Constitutes “Surface Water” Loss

Two estimates of surface water drainage were evaluated when calibrating  $R_{infil}$  as a means of sensitivity testing for the ultimate GWr results: measured total discharge (Q) and runoff (RO) estimated from Q (note “tested and accepted” for Equations C3 and C4). Q was collected directly from the USGS NWIS in daily discharge measurements. Base flow (BF) was estimated from daily discharge data using the USGS streamflow partitioning program PART (USGS, 2011). After assembling Q data and estimating BF, RO was calculated as the residual of Q minus BF (Equation C5).

$$R_{infil} = Ppt - RO$$

- Tested Formulation -

Equation C3

$$R_{infil} = Ppt - Q$$

- Accepted Formulation -

Equation C4

$$RO = Q - BF$$

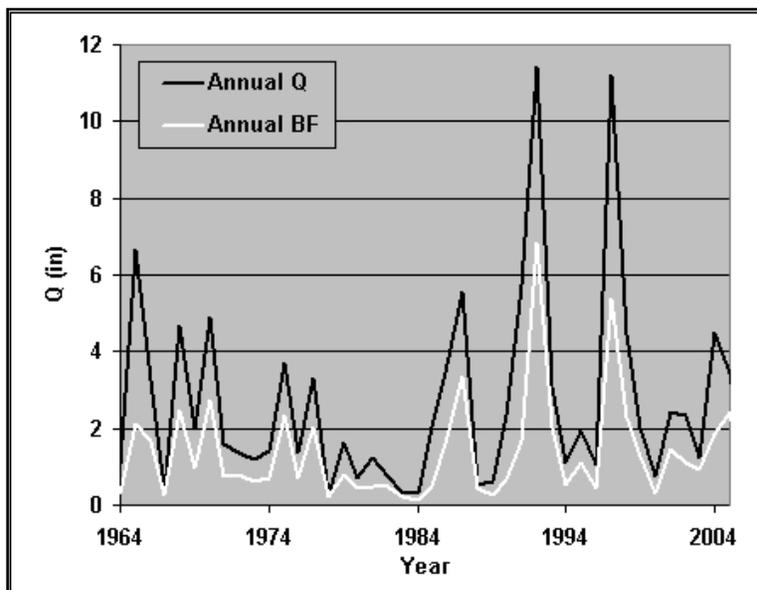
Equation C5

The ultimate intent for calibration of  $R_{infil}$  was to estimate the fraction of precipitation that infiltrates the soil and is available to recharge the aquifers across the project area. Two choices were examined for calculation of  $R_{infil}$ , represented by Equation C3 and C4. These formulations treat  $R_{infil}$  as the residual of RO from Ppt (RO being the amount

rejected by the landscape from storms), or  $Q$ , total discharge (including base flow). These methods were examined using calibration data from gaged watersheds. As indicated in these equations, the first representation was tested and rejected—the representation for  $R_{infil}$  used in the GWr model was calculated as the residual of Ppt and  $Q$ .

BF may be used as an estimate of minimum aquifer recharge given the assumption that BF is discharge from the aquifer. This combines an additional assumption that the portion of water contributing to bank storage cycles through the system: this portion infiltrates the alluvium during storms and exfiltrates afterward. While this may be true, it is not relevant at the temporal and spatial scale of the GWr model. Water that temporarily enters and exits the shallow aquifer falls below the detection limit of the GWr model that provides only coarse outputs in square miles across annual time steps. While the choice for calculation of  $R_{infil}$  using  $Q$  is obvious (Equation C4), use of RO in Equation C3 for the calculation was examined for comparison.

RO was calculated as a residual from  $Q$  determined by the PART model that partitions  $Q$  into BF and RO. These calculations are particularly sensitive to BF and how it is determined. PART sensitivity to storm events is apparent in the sharp increases and decreases present in annual BF estimates (Figure C1). “Flashy” BF predictions may introduce error to GWr model estimates. Extremely high BF predictions due to inflation of BF during storm events and lagging drainage of bank storage are suspected to lead to low RO estimates in representative watersheds (for example, Figure C2). Underestimates of surface water loss induced by quantifying the loss with biased RO lead to low estimates of drainage from the system. Consequently, low estimates of drainage from the system unrealistically exaggerate residual GWr predictions in comparison to estimates of recharge available in the literature.



**Figure C1.** In this example from Lampasas River watershed, base flow estimates produced by USGS PART surge with precipitation increase, and do not correctly represent the steady minimum groundwater discharge component needed to calibrate a RO-based GWr model.

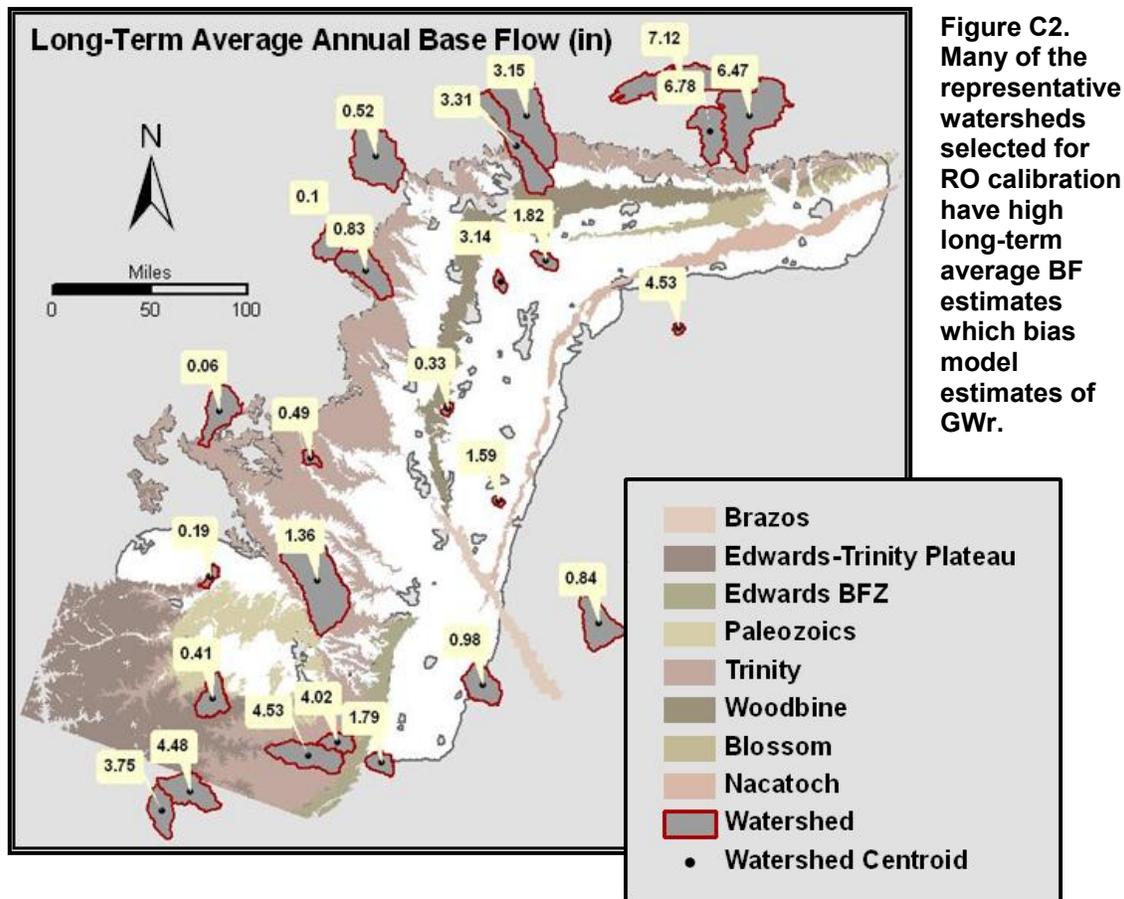


Figure C2. Many of the representative watersheds selected for RO calibration have high long-term average BF estimates which bias model estimates of GWr.

From the examination provided by applying PART, the use of BF and RO were rejected for calculations of GWr. For application of a coarse recharge model (years, miles), no further discrimination of the parameters for water balance beyond simple discharge from the watershed (Q) was made. Q was chosen because it is the water that is exported out as measurable flow.

## C2. Codifying Q Relationships

The procedure used to model  $R_{infil}$  utilized components that vary spatially (i.e. Ppt,  $ET_0$ , and Perm). All calculations were made as annual time steps because this was a simplifying step desired by the TWDB Groundwater Resources Division. The numbered overview below is described in greater detail within this and the following sections.

- (1) As a first step, the infiltration within discrete watersheds was quantified by modeling the Q drained from the Ppt received. This was determined by regressing annual calculated Q against annual Ppt.
- (2) The resulting Q-Ppt relationships were power functions that did not allow extrapolation across the landscape because the power functions were “touchy”.

- i.e., yielding values that quickly scaled out of the expected bounds for  $R_{infil}$  and that were not coherent across nearby watershed boundaries. What was needed was adjustment of the Q-Ppt to follow a functionally linear relationship.
- (3) Measured Ppt was scaled to reflect climatic variables that affect the amount of water available for  $R_{infil}$  (i.e. scaling for the linked effects  $ET_0$  that determine the disposition of Ppt to determine the residual Q in any annual period). This linearization was implemented by dividing Ppt by  $ET_0$  and then using this dynamic variable to drive Q predictions for the  $R_{infil}$  relationship—the resulting relationship between Q and the  $Ppt/ET_0$  ratio was functionally linear.  $Ppt/ET_0$  was chosen as the model input for Q calculations.
  - (4) After linearizing Q by  $Ppt/ET_0$  for each watershed, a linear regression was fit through the full Q record for individual watersheds. The slopes of the individual watershed Q regressions encoded the Q behavior—each controlled by permeability and spatial climate trends. These codified regression slopes were used to model Q for GMA8 using permeability and the spatial climate trends encapsulated in the 50-year average  $Ppt/ET_0$  for each model pixel. Such regressions varied from the SW (hotter, dryer) to the NE (wetter, cooler). The slope of the linear regression could then be used to spatially extrapolate Q behavior anywhere within the GMA8 region based on permeability and climate trends using multiple linear regressions (described in section C3).

### **C2.1. Linearization of Q Data**

Q prediction for the project area was calibrated spatially using individual watershed discharge analyses. For these analyses, gage location, watershed boundaries, Q and BF records, and zonal averages of annual Ppt and watershed permeability (Perm) were compiled into an Excel spreadsheet as described in Section B2.

To analyze individual watershed data, Q was plotted as a parameter dependent upon annual precipitation for regression to a best fit curve. The data clearly followed a non-linear trend (Figure C3).

To constrain problems introduced by modeling non-linear equations, the data were transformed by employing an inherent climatic relationship between precipitation and  $ET_0$ . Precipitation and  $ET_0$  vary inversely—as precipitation increases, humidity increases and  $ET_0$  decreases. Conversely as precipitation decreases, the air dries and  $ET_0$  increases. When precipitation is high in relation to  $ET_0$ , the ratio is stretched towards unity; when precipitation is low the ratio of the stretch is much smaller (Figure C4). The new plot of discharge versus the precipitation- $ET_0$  effectively linearized the precipitation-discharge relationship (Figures C3, lower graph).

Figure C3. Example transformation of non-linear Q-Ppt relationship using the dynamic variable  $ET_0$ . (A) Example watershed location and boundary. (B) Annual watershed Q is related to Ppt by a non-linear function. (C) Due to the inherent difficulty in modeling non-linear equations, a transformation using  $ET_0$  was used to functionally linearize the Q-Ppt relationship as Q vs.  $Ppt/ET_0$ . This transformation is depicted in figure C4.

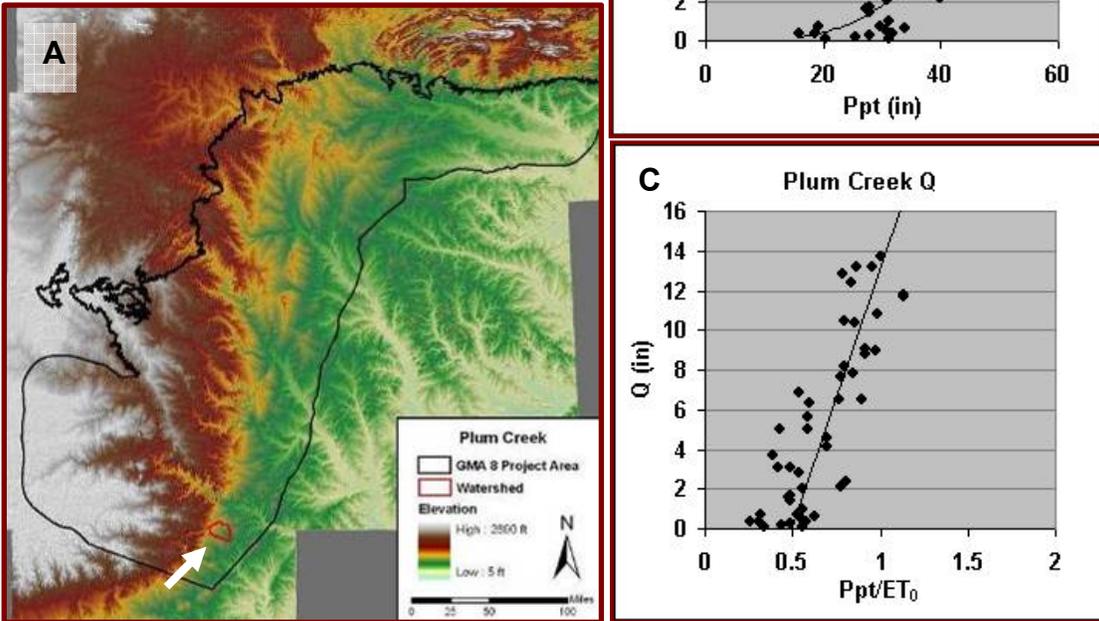
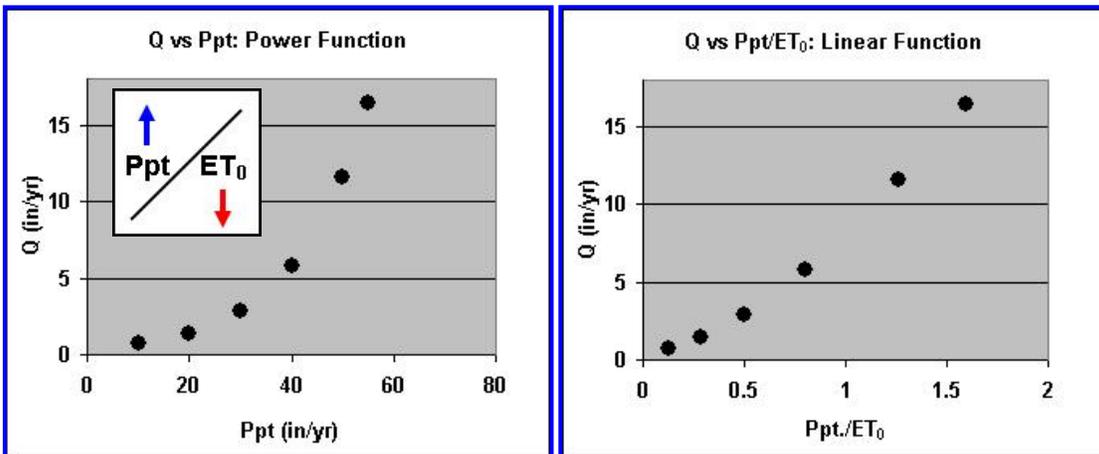


Figure C4. The logical relationship between Ppt and  $ET_0$  examined using synthetic data. As Ppt increases, humidity increases as well, therefore decreasing  $ET_0$  (left). The ratio of these two variables imposes a rightward stretching of the independent (x) variable and quasi linearization of the relationship between Q and the independent variable. Right: synthetic Ppt,  $ET_0$ , and Q developed from calibration data and simplified to illustrate the  $Ppt/ET_0$  transformation—while the data still exhibit a non-linear trend, the data can be adequately described by a linear regression.



## C2.2. Q Relationships Defined by Ppt/ET<sub>0</sub>

The first step in calibrating Q for the model was to analyze the Q versus Ppt/ET<sub>0</sub> relationship for each representative watershed. The annual sums of Q measurements were regressed against the annual Ppt/ET<sub>0</sub> for each watershed (Figure C5). The regression fit for each of the watersheds captured the long-term (LT) trends in Q (Equation C6). Test analysis of the Q regression results found that the Perm-Q relationships were preserved in the slope of the Q versus Ppt/ET<sub>0</sub> regression.

$$\text{Long-Term Watershed Q} = a_{\text{QLT}} + b_{\text{QLT}} * (\text{Ppt/ET}_{0\text{LT}}) \quad \text{Equation C6}$$

The differential Q slopes developed during regression analysis reflected differences in the underlying permeability—a highly permeable watershed, for example, resulted in a large amount of infiltration and a shallow Q slope while a relatively impermeable watershed rejected nearly all Ppt and resulted in a steep Q slope (Figure C6). Given the inherent Q-Perm relationship in the long-term Q regression slope, systematic Q behavior was used to link individual watershed analyses with the Perm dataset. The variability of the driving variable Ppt/ET<sub>0</sub> is clearly spatially controlled in the 50-year average and was also used in the spatial calibration of Q (Figure C7). The regression equation for each watershed was combined with the average Perm and average annual Ppt/ET<sub>0</sub> over that watershed to encapsulate the relationships between these variables for a project-wide regression analysis.

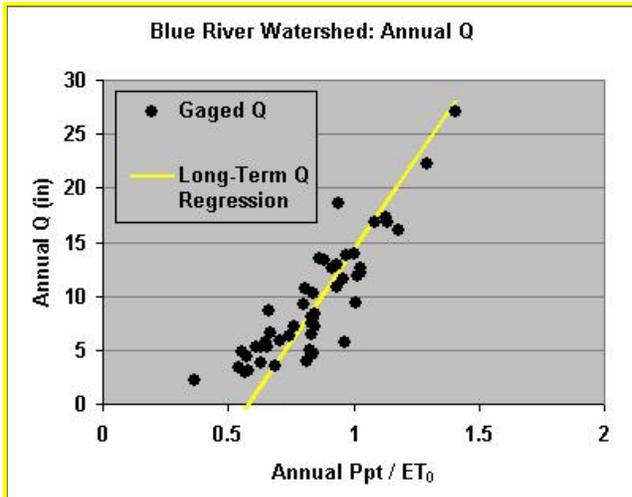
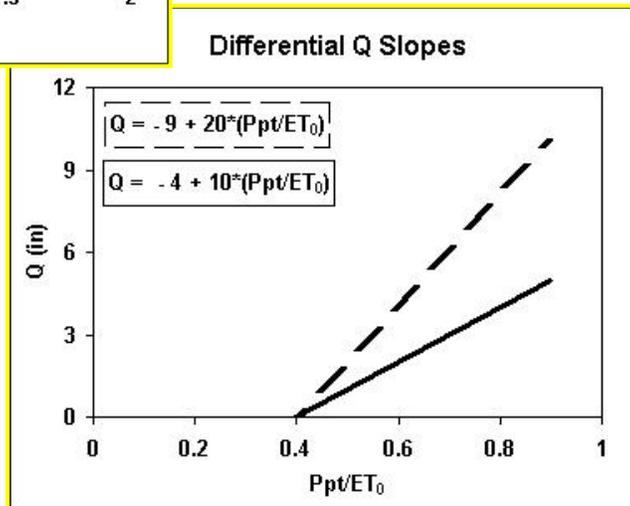
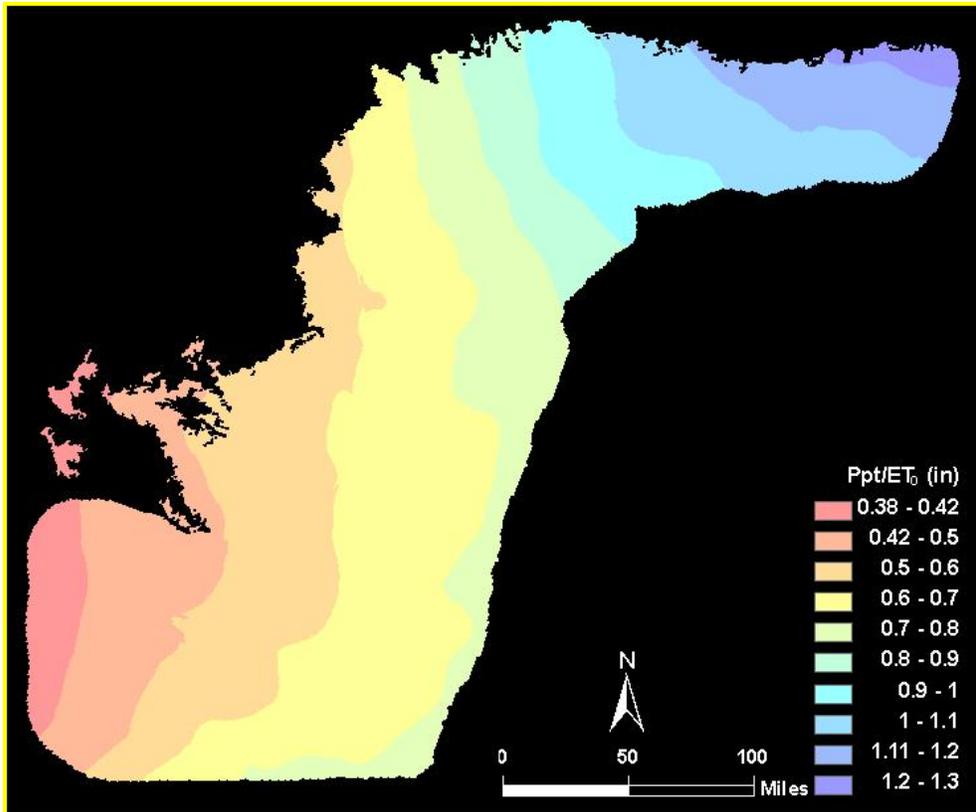


Figure C5. Example of gaged discharge, long term Q regression (Equation C4). The relationship of gaged Q to Ppt was linearized using the Ppt/ET<sub>0</sub> transformation and the regression equation was recorded for each representative watershed.

Figure C6 Perm-controlled and Ppt/ET<sub>0</sub>-driven Q demonstrates the relationship needed to model Q. A steep slope (dashed line) indicates that more incoming rainfall goes to Q than a gentle slope—a condition that occurs with low Perm soils. The solid line is a soil with high Perm that produces less Q, hence, boosting R<sub>infil</sub>, the residual of Ppt minus Q.





**Figure C7.** The 50-year average Ppt/ET<sub>0</sub> ratio illustrates the gradient in climatic conditions across the project area from the NE to the SW that varies by a factor of 3.4 across the project region.

### C3. Multiple Linear Regression of Q

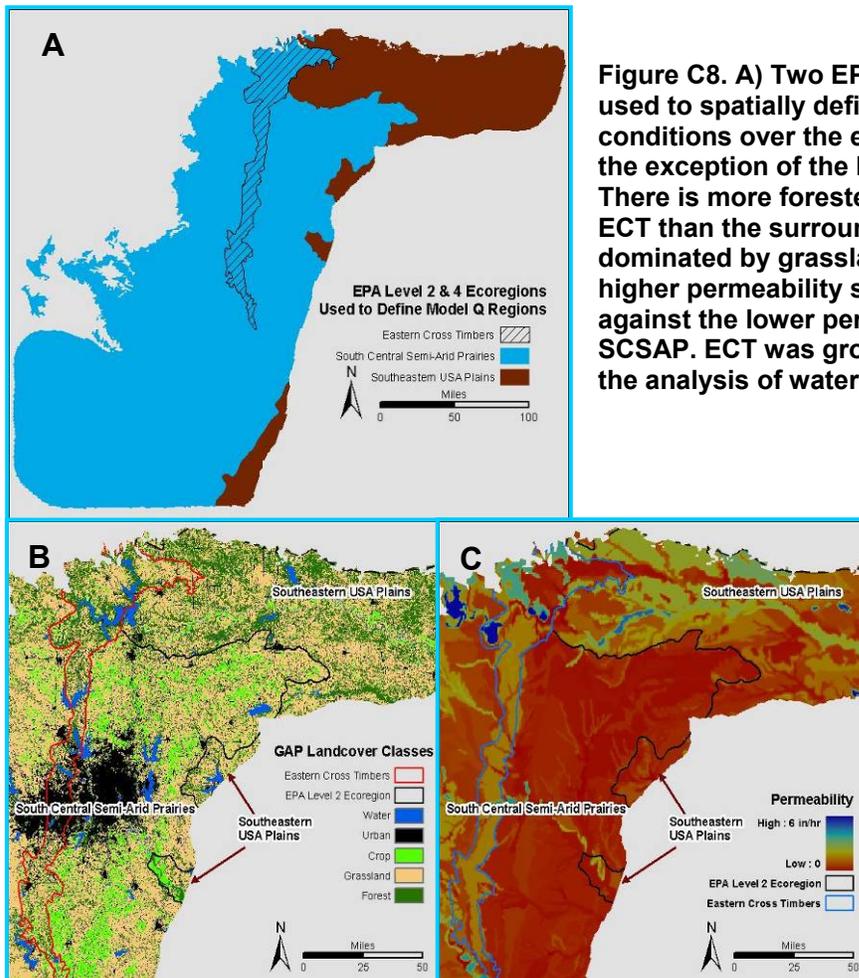
After analyzing the data for each watershed individually, we compared Q data for individual watersheds to spatially-controlled variables in order to extrapolate Q across the project. We tested two techniques for relating Q data to the spatial variables Perm and Ppt/ET<sub>0</sub>: (A) lumping all annual Q, ET<sub>0</sub>, and Ppt as well as spatial Perm data into a multiple regression analysis, and (B) using a representative set of long-term average values from each watershed to relate individual watershed data to the project area. In preparation for the (B) technique, the regression equation for each watershed was combined with the average Perm and average annual Ppt/ET<sub>0</sub> over each watershed to encapsulate the relationships between these variables for a project-wide regression analysis.

Initially, we tried grouping annual values of Q, Ppt/ET<sub>0</sub>, and Perm from all watersheds for a multiple regression analysis. The resulting regression relationships were weak due to the high variability in timing of Ppt, ET<sub>0</sub>, and Q in the annual data. Next, to enhance the signal in the relationship through removal of the annual variability, we reduced the Q

versus  $Ppt/ET_0$  relationship for each watershed using Equation C6 and the 50-year averages. The signal for  $Q$  versus  $Ppt/ET_0$  was encoded by the slope,  $b_{QLT}$ , of the regression relationship for each watershed—an example line is portrayed in yellow in Figure C5. Thus encoded by watershed, the slope  $b_{QLT}$  was analyzed through multiple linear regression using  $b_{QLT}$  as the dependent variable predicted by two independent variables, Perm and average  $Ppt/ET_0$ .

### C3.1 Multiple Regression Analysis of $Q$ for Unique Discharge Regions

Though representative of broad regions, the  $Q$  from all watersheds could not be well explained by one regression model so spatial subsets were evaluated. To consider spatial groupings, watersheds were divided based on their location within the EPA Level Two Ecoregions (EPA, accessed 2011). Unique conditions of  $Q$  across the model were described by two spatial groups: the South Central Semi-Arid Prairies (SCSAP) in the east and northeast of the project region and the Southeastern USA Plains (SEP) for the remaining majority of the project region excepting a sub-region of SCAP: Eastern Cross Timbers (Figure C8, A). The Eastern Cross Timbers sub-region lies within SCSAP, but the sub-region's floristic and Perm characteristics made the region more similar, in terms of the  $Q$  model, to SEP (Figure C8, B and C). For the purpose of estimating  $Q$ , the Eastern Cross Timbers (ECT) region was grouped with SEP. Variability in watershed  $Q$  within each region was captured by a unique regression for each region.



### C3.2. Assembling the Q Regression

In order to arrive at the equation to solve for  $R_{infil}$  from Ppt and Q (Equation C4), the variable Q must first be defined. The dynamic relationship between Q, Ppt,  $ET_0$ , and Perm was captured using multiple regression analysis. The equations to estimate Q parameters revealed by the regression models indicated dynamic changes in the unique conditions at each pixel for each model year. The dynamic Q parameters for each model year were then used to solve a simple linear regression to calculate annual Q. The following steps were followed for deriving the Q parameters for each model year. The Q parameters were then used to estimate Q based on annual Ppt/ $ET_0$ .

1) Watershed data were compiled in a table for each discharge region according to four variables:

- Long-term discharge equation slope:  $b_{QLT}$ .
- Long-term discharge equation x-intercept: **x-intercept** $_{QLT}$ .
- Long term average **Ppt/ $ET_{0LT}$** .
- Average watershed **Perm**.

2) Multiple linear regression analysis was performed for the dependent variable  $b_{QLT}$  and independent variables Perm and Ppt/ $ET_{0LT}$  for each discharge region. Regional coefficients,  $b_{climate}$  and  $b_{Perm}$ , and constant $_b$  shown in Equation C7 were defined in this process. The constants in Table C1 and Equation C7 were used to estimate  $b_{Qi}$  for each model year “i”—this unique slope for each model year was used in equation C10 to solve for annual Q for year “i”.

**Table C1. Q Multiple Regression Constants**

	Region 1 (SCSAP)	Region 2 (SEP)
$b_{climate}$	50.65	4.00
$b_{Perm}$	- 3.41	- 1.50
<b>Constant<math>_b</math></b>	- 4.39	26.00

$$b_{Qi} = ( b_{climate} * Ppt_i / ET_{0i} ) + ( b_{Perm} * Perm ) + Constant_b$$

**Equation C7**

3) Non-linear regression analysis was performed for the dependent variable x-intercept $_{QLT}$  against Ppt/ $ET_{0LT}$  for each discharge region. Regional coefficient  $b_{x-int}$  and constant $_x$  shown in Equation C8 were defined. After both  $b_{Qi}$  and x-intercept $_{Qi}$  were defined,  $a_{Qi}$  was calculated using Equation C9, the resulting intercept,  $a_{Qi}$ , was used in equation C10 to solve for annual Q.

**Table C2. Q Regression Constants**

	Region 1 (SCSAP)	Region 2 (SEP)
$b_{x-int}$	0.2983	0.2000
$Constant_x$	0.5974	0.5752

$$x\text{-intercept}_{Qi} = b_{x-int} * \ln( Ppt_i / ET_{0i} ) + Constant_x \quad \text{Equation C8}$$

$$a_{Qi} = - b_{Qi} * x\text{-intercept}_{Qi} \quad \text{Equation C9}$$

4) All regional coefficients and constants were incorporated into the model script. Equations C7 through C9 were solved for each year “i,” and these solutions were then used to estimate annual discharge  $Q_i$  with Equation C10.

$$Q_i = a_{Qi} + b_{Qi} * ( Ppt / ET_{0i} ) \quad \text{Equation C10}$$

5) Dynamic adjustments to Q predictions over urbanized areas were applied as described in Section G: Modeling Urbanized Areas.

6) The final estimate of  $Q_i$  was capped at the annual precipitation using equation C11 and negative Q estimates were replaced with zeros using Equation C12.

$$\begin{aligned} \text{If } Q_i > Ppt_i, Q_i &= Ppt_i \\ \text{If } Q_i \leq Ppt_i, Q_i & \end{aligned} \quad \text{Equation C11}$$

$$\begin{aligned} \text{If } Q_i < 0, Q_i &= 0 \\ \text{If } Q_i \geq 0, Q_i & \end{aligned} \quad \text{Equation C12}$$

The predictions made using these equations trace the non-linear behavior observed in gaged Q (Figures C10 and C11). Predicted Q agrees well with gaged Q (Figure C9).

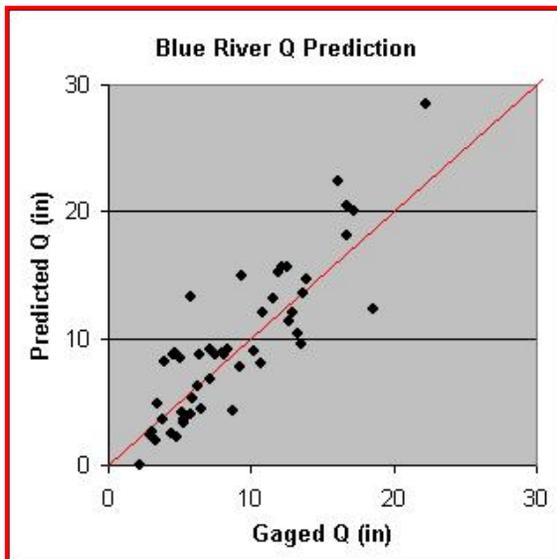
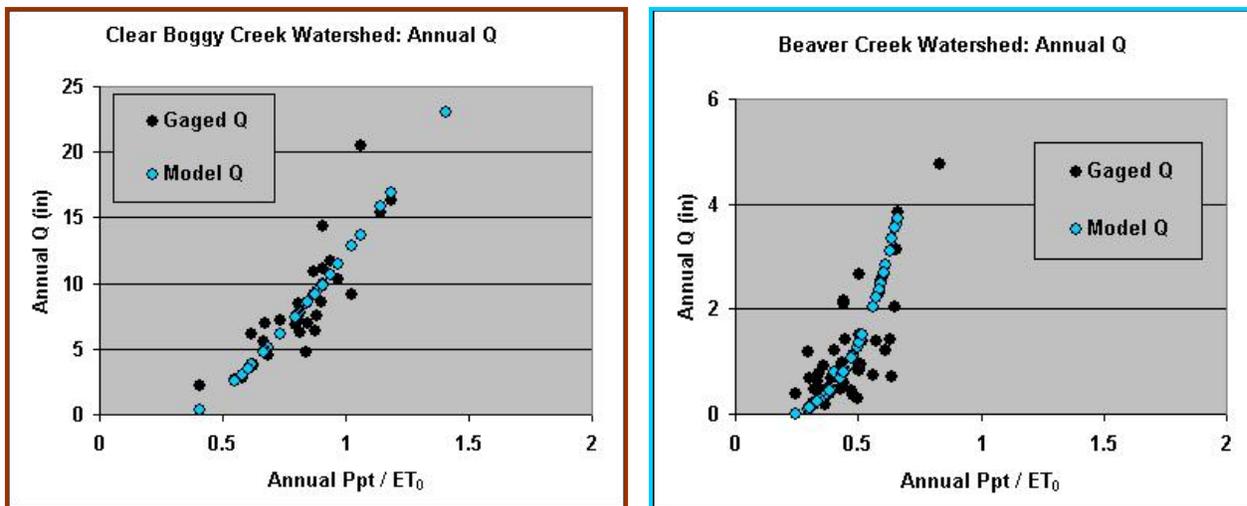


Figure C9 Q predicted with the new multiple regression model agrees with Q measured at stream gages. The example shown is the Blue River drainage in the far north of the project region.

### C3.3. Evaluation of Q through Confirmation for Each Region

To verify the results of the multiple regression analysis, the technique was applied to individual watersheds from both the two model regions, SCSAP and the SEP. The Perm coefficient, climate coefficient, and constant developed for each region were used to model Q according to Perm and the annual Ppt/ET<sub>0</sub> ratio for each year of a watershed's Q record. Figure C10 illustrates the agreement between modeled and measured Q at stream gages for both discharge regions of the model.

**Figure C10. Q modeled using multiple-regression derived coefficients compared to gaged Q for each discharge region of the model. Left: Q predictions and measurements for a watershed in the SEP region. Right: Q predictions and measurements for a watershed in the SCSAP region. In both regions, predicted Q follows the trend observed in the original gaged data.**



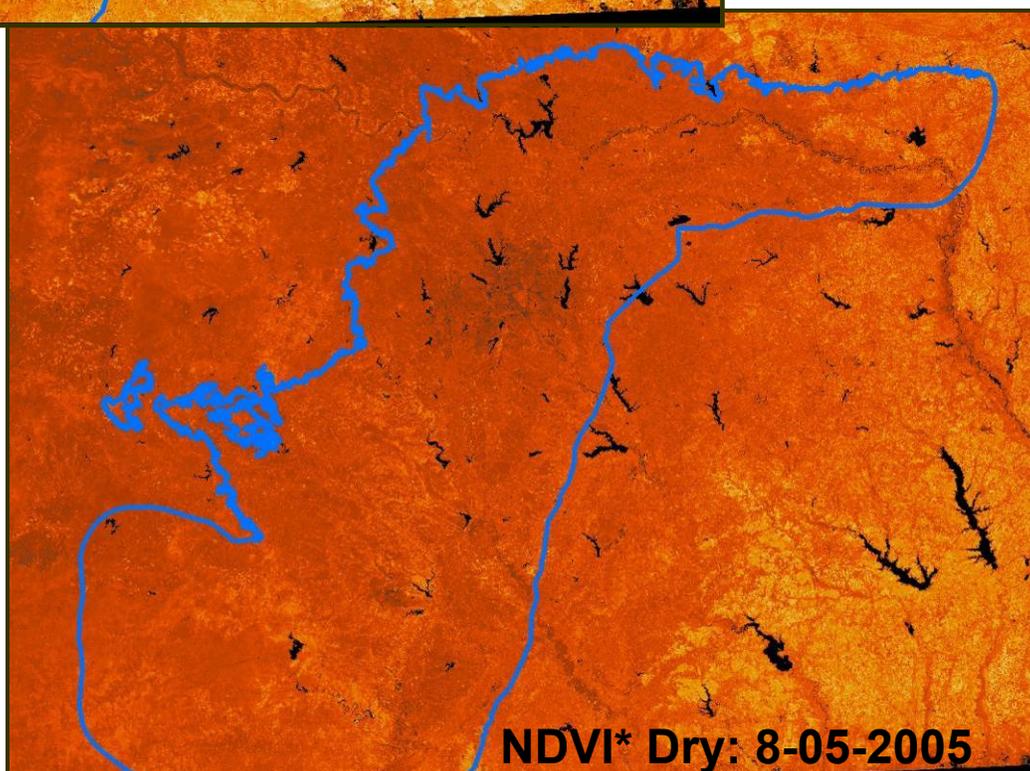
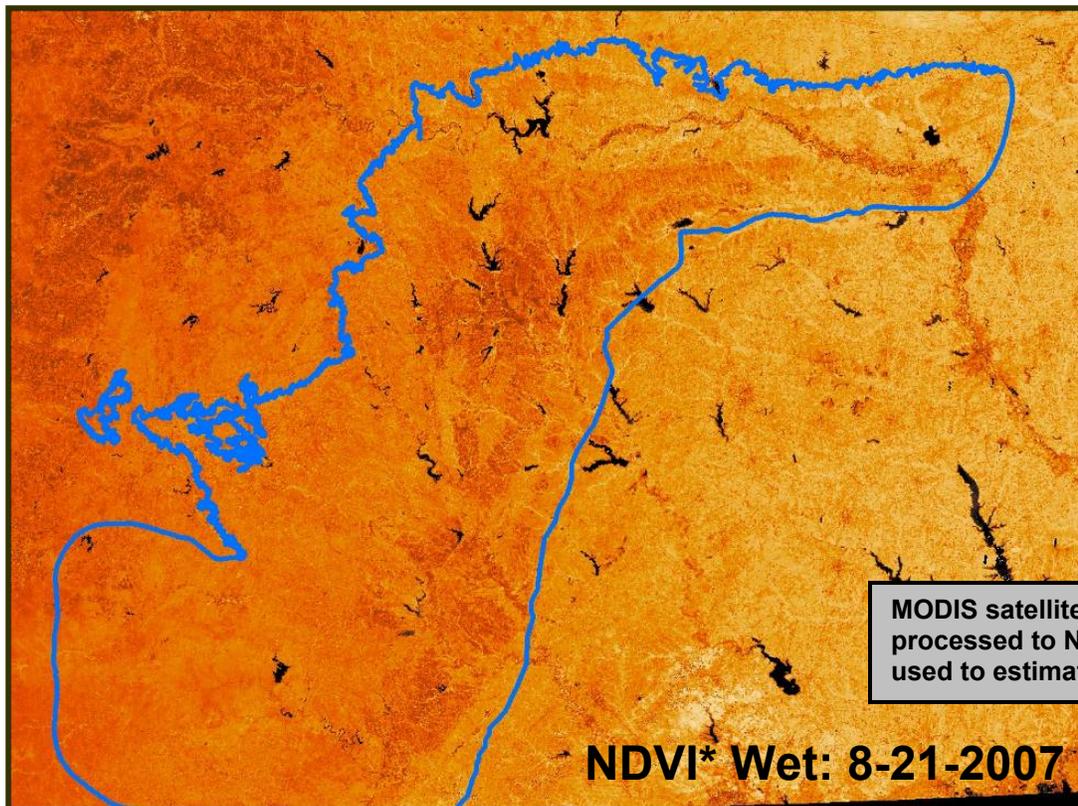
### C4. Summary

To estimate the annual depth of infiltrated rainfall ( $R_{infil}$ ) for groundwater recharge (GWR) predictions, a technique was developed to estimate how much of the incoming precipitation (Ppt) was lost at the surface. Surface water loss was quantified as total gaged discharge (Q), which was analyzed at representative watersheds across GMA8. The relationship between Q and Ppt was simplified and linearized by scaling the driving variable Ppt with ET<sub>0</sub>. The slope of the linear relationship between Q and Ppt/ET<sub>0</sub>,  $b_Q$ , provided a way to codify Q behavior spatially based on soil permeability (Perm) and climate variables (50-year Ppt/ET<sub>0</sub> average). A multiple linear regression model was developed to predict Q based on the spatially-codified relationships for each of two model regions. The Q predictions resulting from the multiple linear regression analyses accurately modeled the Q observed in gaged data.

**Data provided from the work in this section include:**

- Calibration watershed data including quarterly and annual Ppt,  $ET_0$ , and gaged Q in an Excel spreadsheet, “Individual Watershed Analysis”.
- Q estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- $R_{infil}$  estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- Unique Q regions boundaries defined in raster provided in a geodatabase of model parameters and as a separate raster.

## D. REMOTELY-SENSED DUAL COEFFICIENT (RDC) ET ESTIMATION FOR GMA8





## **D. Remotely-Sensed Dual Coefficient (RDC) ET Estimation for GMA8**

Groundwater recharge cannot be measured directly but must be inferred as a residual from rainfall after surface processes of discharge and evapotranspiration (ET<sub>a</sub>) have been subtracted. This section presents validation of calculations from the Remotely-Sensed Dual Coefficient (RDC) evapotranspiration estimation method developed by HydroBio. The technique was designed for use with Landsat Thematic Mapper 5 (TM5) and other moderate resolution (20-30 meter) Earth Observing Satellite (EOS) data. For application to the comparatively large GMA8 project area, synoptic scale satellite data from Moderate Resolution Imaging Spectroradiometer (MODIS) was applied.

The methods used to estimate ET<sub>a</sub> in the GMA8 region relied upon an internally calibrated vegetation index that was chosen to scale plant activity. Through fitted mathematical relationships to represent broad classes of vegetation, described in Section F, this index enabled generation of spatial estimates of annual ET<sub>a</sub> for a dry year (2005) and a wet year (2007). The description of the calibration and validation analyses for ET<sub>a</sub> for three types of land cover—crops, woody vegetation and grassland vegetation, are presented in Appendix 2. Within the project area, woody vegetation is primarily forested while grassland vegetation is primarily savannah (dominated by grass but with shrubs and trees).

The intent with the application of RDC calculations of ET<sub>a</sub> was intended to provide close estimates of the proportion of precipitation that was lost from the landscape. Inherent in this estimate were errors that were not measurable within RDC calibration and application. The intention for ET<sub>a</sub> estimation for the GMA8 project was to provide an estimate that was robust relationally, meaning that inherent errors in the ET<sub>a</sub> estimation are shared equally across the landscape. Such relational robustness enables the use of scaling for sensitivity analysis against other regional groundwater model components. For example, for cross checking regional groundwater models, ET<sub>a</sub> can be varied by decimal fractions up or down to enhance or retard GWR in the GWR model

### **D1. Introduction to the RDC Method**

The RDC method uses estimates of actual ET for a crop “n” (ET<sub>a<sub>n</sub></sub>) scaled by reference ET (ET<sub>0</sub>) and canopy and soil factors for that crop over a given time interval. RDC employs an adaptation of the normalized difference vegetation index (NDVI) that was proven to be more accurate than the other published vegetation indices for an evaluation that used actual satellite data to predict known hydro-ecologic processes (Baugh and Groeneveld, 2006). This index, NDVI\*, was employed in RDC to estimate ET<sub>a</sub>. The following series of papers were used as the basis for RDC:

- Testing vegetation indices from EOS data that most accurately capture the hydrologic signal—NDVI\*: (Baugh and Groeneveld, 2006).

- Testing methods to process EOS data to enhance the hydrologic signal in vegetation (NDVI\*): (Groeneveld and Baugh, 2007)
- Developing a method for EOS estimation of ETa using NDVI\* in use by three western states and Bureau of Reclamation (Groeneveld et al., 2007).
- Fitting ETa/groundwater-drawdown curves with the single-scene NDVI\* method for a site with declining water table (Groeneveld, 2008).

RDC is based upon a simple model used in agriculture, the dual coefficient method, used for scaling the ETa of crop as a function of reference ET (Allen et al., 1998). The dual coefficient method provides one-dimensional estimates of ETa (as depth at each location) and ETa for any crop “n” is often calibrated within temporal lookup table values (Equation D1). ETa<sub>n</sub>, for any area occupied by a crop n, results from summation of multiple time steps (t) representing all n crop development through the season. Summation of ETa<sub>n</sub> in any grid cell yields ETa (Equation D1).

$$ETa_{tn} = \Sigma ( Kcb \cdot ET_0 + Ke \cdot ET_0 )_{tn} \quad \text{Equation D2}$$

**Where Kcb and Ke are lookup-table ET scaling factors that represent the canopy (Kcb) and the lumped soil surface (Ke), and ET<sub>0</sub> represents the grass reference ET calculated by the ASCE Penman Monteith method (Allen, undated).**

The RDC method uses NDVI\* as an estimator of Kcb with correlation statistics used to solve for Ke. Ke can be a fitted as a constant or an equation to enhance NDVI\* so that when multiplied by ET<sub>0</sub>, yields an unbiased estimator of measured ETa. The RDC approach was developed prior to the GMA8 project and so, was validated and further calibrated to yield ETa for application within GMA8. Data were acquired from the AmeriFlux network that measures, among other parameters, the input for ETa and ET<sub>0</sub> calculation for many sites across the United States (Heilman 2005 – 2006; Katul, 2002-2005; Oechel, 2005 – 2006; Verma 2003 – 2006).

For estimation of ETa in the study area, numerous EOS images were analyzed that were assembled as image stacks that formed a data cube—space as x and y, with the z dimension being NDVI\* and with numerous geocorrected images incorporated to represent time wise snapshots stacked one over the other. The concept of a cube arises from the fact that the geoposition, determined by x and y dimensions remain in the position through dataset.

Validation/calibration analyses for RDC to yield unbiased estimation of ETa in GMA8 used measured ET data from the AmeriFlux network that were paired with Landsat TM5 data. The large project area of the GMA8 GWr model required the use of relatively coarse resolution MODIS data. RDC was developed using Landsat TM data because the 30m pixels enable comparison at the same approximate scale of the flux measurement locations (within a radius of about 3 - 4 pixels or 90 - 120 meters (98-131 yards). Rather than attempt calibration with MODIS, the finer resolution for TM data was used because precise spatial control was necessary for pairing AmeriFlux measured ETa with EOS data for deriving solutions to Equation D1. MODIS, with 231 meter pixels, was far too coarse.

On the other hand, Landsat TM data were judged unsuitable for ET estimation across the large project area because they are far too data rich, require about 9 images to cover the project region, would bog computations for one square mile grid system and have images acquired only at 16-day repeat intervals.

The 16-day repeat interval of TM combined with a high potential for cloud cover over the project area was demonstrated to cause data skips through large critical periods that were unacceptable for project application. With its daily repeat images and coarser scale, MODIS was the best data platform for remote-sensing estimation of ETa across the project area. Use of MODIS to estimate ETa using relationships that were first calibrated with TM required that the MODIS data were made equivalent to Landsat TM.

## **D2. Making MODIS NDVI Equivalent to Landsat TM5 NDVI**

RDC was calibrated using Landsat TM data. MODIS and Landsat TM do not have equivalent band sensitivities for red and near infrared bands used for assessing NDVI. NDVI\* analysis was made so that raw NDVI from MODIS could be transformed to be equivalent to TM5. For calibration both a Landsat TM5 image (Path/Row 28/29) and a MODIS image (H9/V5) covering the Austin/Dallas region of Central Texas were chosen from the same day (10-01-2010). This snapshot was selected because the region of interest was clear of clouds and lacked haze and represented optimal conditions for remote sensing analysis.

NDVI, normalized difference vegetation index, was calculated from the data provided by MODIS and Landsat TM. The relationship governing NDVI is presented in Equation D2.

$$\text{NDVI}_i = (\text{NIR} - \text{red}) / (\text{NIR} + \text{red}) \qquad \text{Equation D2}$$

**Where  $i$  represents the  $i$ th pixel and red and NIR (near infrared) are broad bands found on most EOS platforms.**

Though the nominal pixel dimension is 250 meter for MODIS, these data actually come as 231 meter pixels for the Texas region. MODIS data are delivered in reflectance while Landsat TM5 data require correction to reflectance using equations in Irish (undated) with the published gains and offsets for Landsat TM5. The TM5 data were rescaled to 231 m using the cubic convolution algorithm in ENVI version 4.8. The paired 231 m pixels from the MODIS and the rescaled TM were then sub-sampled to the one pixel overlying the centroid of a 5 square kilometer grid as a means of limiting the number of samples. Sub-sampling enabled constraining the sample size to be sufficient for a robust statistical fit, while remaining manageable for curve fitting and outlier removal by creating a sample size of 899 paired values. The paired data were graphed and a preliminary relationship was fitted using linear regression.

Every geocorrection naturally contains at least one-half pixel spatial uncertainty. For that reason, these paired data may contain bits of open water (zero NDVI) mixed into selected

pixels. This inclusion would tend to create scatter both above and below an expected line for comparison of the two EOS data sets. Such values can be highly influential on regression since they tend to lie at the extreme end of the data distribution. For this reason, the suite of 899 paired values were evaluated for outliers and 20 values were removed from the data cloud (2.2%) in order to provide a relationship of the highest precision (Figure D1).

The analysis for processing MODIS to yield TM5-equivalent NDVI\* was accomplished by first using GAP data (USGS, Accessed 2010) to define the location of large water bodies within the GMA8 region. Water body locations were used to screen out all pixels that contained water bodies since the reflectance from clear deep water (lacking significant algae or entrained sediments) is zero, or slightly negative; values that would skew the results. Additionally, when entrained particles or algae are present in water bodies, the backscattering of light that occurs is different in red and near infrared bands that are used to calculate NDVI (Groeneveld and Barz, 2010); this can greatly affect the results that are obtained. Each pixel is the functional average of the area within and the many of water-affected pixels would skew the distribution of NDVI values in the analysis. Clouds were also identified spectrally and then masked out of the images.

An area between Austin and Dallas was chosen as the area for this investigation. NDVI was calculated for water- and cloud-masked pixel values. The collected masked NDVI values were displayed as cumulative distribution functions (CDFs) during the two years chosen for analysis.

The resulting regression relationship in Figure D1 was used to yield MODIS NDVI values that are equivalent to Landsat TM5 values. Scatter in the data cloud of Figure D1 is due to geopositional uncertainty inherent in these data plus spatial uncertainty due to the cubic convolution method for pairing data. Cubic convolution simply selects the centroid of a pixel when combining pixel values for scaling TM data to the size of a MODIS pixel.

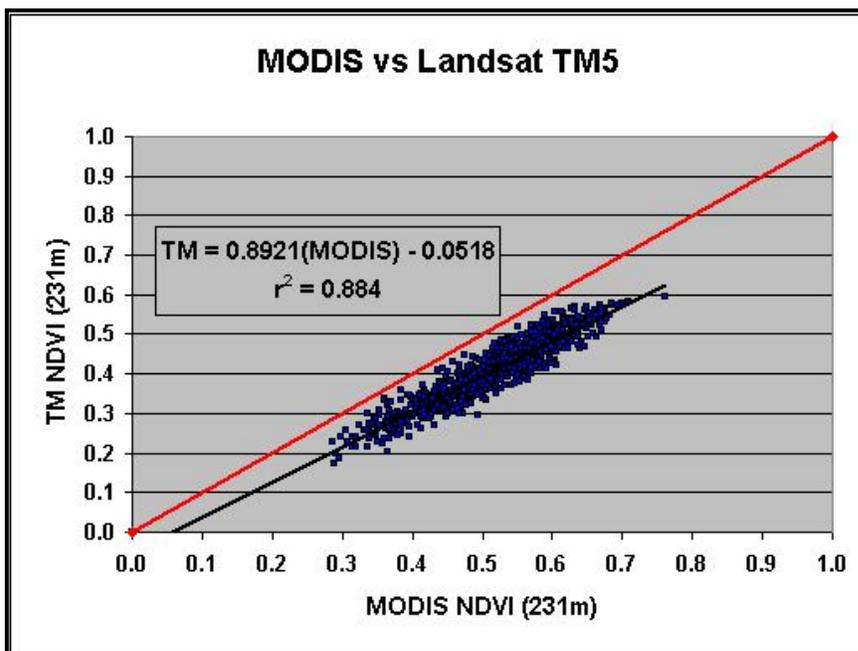


Figure D1. Graph of the data and the equation to yield MODIS NDVI that is equivalent to TM5 NDVI. The calibrated equation transforms MODIS data to fit the one-to-one line (red) so that the TM-derived calibration curves for ETa can be used with the MODIS data.

### D3. Processing MODIS Images for Estimation of ET

NDVI\*, a derivative of NDVI, is a stretched version of NDVI that codes vegetation density from zero to one in scale against  $ET_0$  to run the RDC calculations of  $ET_a$ . To calculate NDVI\*, two values are needed,  $NDVI_0$  and  $NDVI_s$ , that represent NDVI for vegetation cover at zero and at maximum expression, respectively (Equation D3). Calculation of NDVI\* is necessary because even bare soils often present an NDVI response that is greater than one, often about 0.10, and this is further altered because atmospheric aerosols differentially scatter the red and NIR bands.

$$NDVI^* = (NDVI_i - NDVI_0) / (NDVI_s - NDVI_0) \quad \text{Equation D3}$$

**Where  $NDVI_0$  is the value at zero vegetation cover and  $NDVI_s$  is the value at NDVI saturation, a theoretic peak vegetation canopy response.**

When choosing  $NDVI_0$  values for TM data, a number of methods can be used including regression of the lower limb of the NDVI CDF as in Baugh and Groeneveld (2006) and Groeneveld and Baugh (2007). Such methods work for TM5 data, because, at 30 meters, the pixels are small enough to statistically incorporate all or some bare ground, while the lowest limb of the curve represents Gaussian relationships for pixel mixes with water, parking lots, roadways and buildings. The regression predicts  $NDVI_0$  and cuts off the Gaussian portion of the distribution.

To establish  $NDVI_0$  for MODIS pixels, at 231 meters, regression is not a reliable method and another method to select  $NDVI_0$  for a known target that was devoid of vegetation was used. This measured value of  $NDVI_0$  was then used for the entire MODIS scene. Because of the size of the MODIS pixels, the bare target was chosen carefully by first looking at low non-water values of NDVI. A portion of a quarry near Georgetown, north of Austin (Figure D2) was selected on the imagery and then confirmed by Google Earth. This location yielded three pixels that were centered within low NDVI values, thus avoiding mixed edge pixels and issues with geocorrection. Unresolved geospatial error generally incorporates about one half pixel of spatial uncertainty for any EOS data.

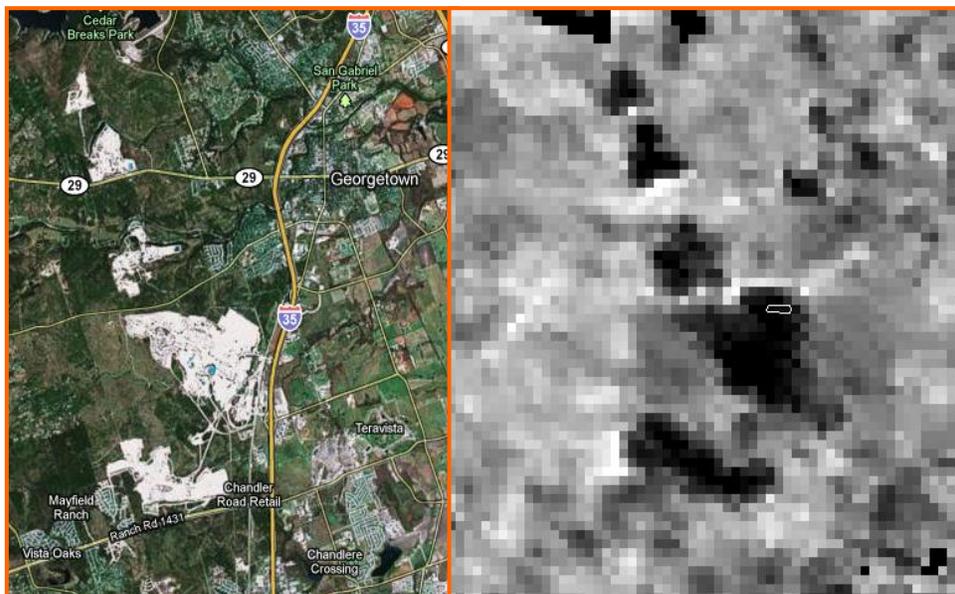
Table D1 presents the extracted  $NDVI_0$  values for the pixels located in Figure D2. Two years worth of data were evaluated, 2005 and 2007; these years were chosen to represent record dry and wet years across the GMA8 region. A number of images were eliminated from these tables because of issues with clouds within the region of the target area. These data show little variability; hence, choosing the bare quarry values for  $NDVI_0$  was anticipated to work satisfactorily. The overall average values highlighted in these tables were used for images that were cloudy over the  $NDVI_0$  target area. Likewise, since the MODIS eight day cloud-free product potentially contains a mosaic spanning all eight days, the low variability of the quarry data ensures that very little error will be introduced into the analysis due to different dates for a multi-day mosaic of MODIS data. Also, since the “cloud free” product often isn’t cloud free, using the overall averages to compensate for missing quarry values introduced little additional error over that induced by the mosaic of different days within each eight day snapshot of MODIS.

The years 2005 and 2007 were chosen for evaluation of spatial (i.e., per pixel) ETa for recent record dry and wet years, respectively. All intermediate years within the precipitation record were interpolated within these two extremes. The MODIS scenes for 2005 and 2007 were first processed to NDVI, and then converted to TM5 NDVI equivalency according to the relationship in Figure D1. For calculation of the TM5-equivalent NDVI\*, NDVI<sub>0</sub> was chosen as the average quarry value in Tables D1, expressed as TM5 NDVI equivalent. The value of NDVI<sub>5</sub> used an established value of 0.86, that has been found in other HydroBio work to yield a robust representation of NDVI\* with a minimum of uncertainty when working with Landsat TM data.

**Table D1. Day of year and MODIS NDVI<sub>5</sub> values for the three pixels indicated in Figure D1.**

2005	Record Dry Year				Average	2007	Record Wet Year				Average
	Low	Med	High	Average			Low	Med	High	Average	
9	0.100	0.102	0.103	0.1017	1	0.103	0.107	0.122	0.1107		
65	0.090	0.098	0.112	0.1000	25	0.093	0.094	0.115	0.1007		
73	0.087	0.104	0.156	0.1156	33	0.095	0.097	0.100	0.0973		
89	0.070	0.088	0.094	0.0840	41	0.098	0.106	0.109	0.1043		
113	0.071	0.083	0.086	0.0800	49	0.088	0.093	0.105	0.0953		
137	0.090	0.091	0.091	0.0907	65	0.091	0.091	0.095	0.0923		
169	0.137	0.137	0.141	0.1383	89	0.079	0.080	0.097	0.0853		
233	0.095	0.109	0.114	0.1060	97	0.134	0.141	0.147	0.1406		
265	0.081	0.082	0.109	0.0907	105	0.098	0.105	0.110	0.1043		
281	0.093	0.093	0.114	0.1000	113	0.098	0.103	0.112	0.1043		
297	0.101	0.105	0.107	0.1043	161	0.086	0.087	0.104	0.0923		
313	0.092	0.098	0.132	0.1073	217	0.145	0.171	0.172	0.1626		
329	0.086	0.089	0.094	0.0897	265	0.094	0.102	0.106	0.1007		
337	0.101	0.105	0.107	0.1043	281	0.084	0.090	0.091	0.0883		
353	0.069	0.074	0.077	0.0733	297	0.079	0.089	0.127	0.0983		
361	0.073	0.098	0.103	0.0912	313	0.102	0.103	0.109	0.1047		
<b>Average</b>	<b>0.0897</b>	<b>0.0972</b>	<b>0.1088</b>	<b>0.0986</b>	<b>Average</b>	<b>0.0973</b>	<b>0.1031</b>	<b>0.1175</b>	<b>0.1055</b>		

**Figure D2. Location of target for bare areas to calibrate NDVI<sub>0</sub> for MODIS. The small area indicated in the NDVI MODIS grayscale at right was used for extraction of the three pixels whose centroids fell within the area.**



## D4. Summary

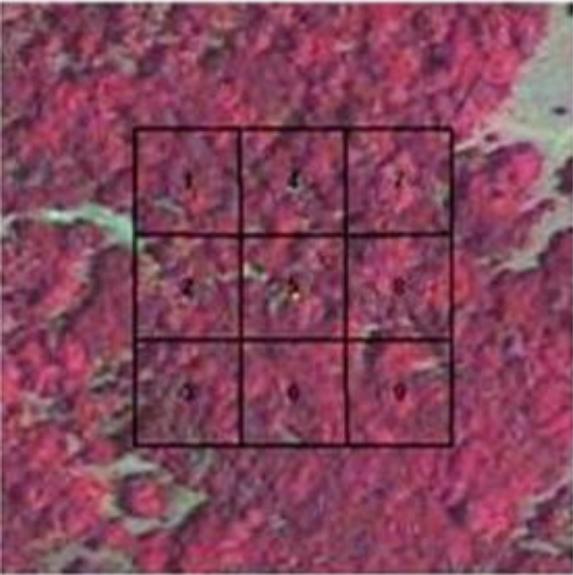
Data from MODIS imagery were used for estimation of ETa. From published data sets, Landsat TM data were used for calibration to estimate ETa. MODIS NDVI can be made equivalent to Landsat NDVI using a mathematical transformation. Transformation of MODIS enables the use of the ETa calibration methods. For estimation of ETa, a stretched version of NDVI is used, NDVI\*. NDVI\* requires the selection of a zero vegetation cover location on the MODIS for its calculation. Calibration targets, bare areas devoid of vegetation, were chosen to enable calculation of NDVI\* from MODIS data. To model the range of ETa observed in GMA8, record wet (2007) and dry (2005) years were chosen to spatially represent the endpoints of the spectrum of ETa across the varying climate conditions of GMA8.

### **Data provided from the work in this section include:**

- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters.
- MODIS raw and intermediate calibration data.
- MODIS-derived quarterly NDVI\* for 2005 and 2007.



# E. CALIBRATING ETA: RDC PARAMETERS





## E. RDC Parameters—Validation and Calibration

Section D provided an overview of ET<sub>a</sub> calculated by the RDC method. This section describes the culmination of techniques and data used to calibrate RDC parameters for calculating ET<sub>a</sub> across GMA8 region before they were applied to MODIS data.

The RDC method was originally developed using data sets recorded in drier climates than represented by the GMA8 region. For this reason, the RDC analysis for GMA8 was conducted in two parts: first, validation to ensure that the approach was correct for the GMA8 climate and then calibration, to ensure that the relationships were correct for the range of climatic conditions and vegetation in GMA8 region. The study sites chosen for the validation/calibration effort contained a variety of climates and vegetation. Inherent in this wide range of subject sites was the intent to determine the robustness of the relationships to estimate ET<sub>a</sub> correctly across the highly variable climate and vegetation of the region.

Landsat TM images were used to estimate ET<sub>a</sub> and were then calibrated to ET<sub>a</sub> measured at flux towers. This approach correctly matched the scale of ET<sub>a</sub> flux (across about 100 - 200 meters) with the scale of multiple TM pixels. The relationships developed from these analyses were later applied to the MODIS data for estimation of ET<sub>a</sub> over the GMA8 region.

For adaptation to use with MODIS, an unbiased estimator of K<sub>cb</sub> (defined in Equation E1) was chosen as the product of NDVI\* derived per pixel and ET<sub>0</sub> derived spatially according to point-wise measurements and extrapolated geostatistically (Groeneveld, 2008). K<sub>e</sub>, and any unresolved error from the K<sub>cb</sub> estimate, were then solved by fitting appropriate transforms to yield total ET<sub>a</sub> estimates equivalent to the measured ET<sub>a</sub>. The transformation to estimate the K<sub>e</sub> component was fitted for three broad cover types.

$$ET_{a_{tn}} = \Sigma ( K_{cb} \cdot ET_0 + K_e \cdot ET_0 )_{tn} \quad \text{Equation E3}$$

**Where K<sub>cb</sub> and K<sub>e</sub> are lookup-table ET scaling factors that represent the canopy (K<sub>cb</sub>) and the lumped soil surface (K<sub>e</sub>), and ET<sub>0</sub> represents the grass reference ET calculated by the ASCE Penman Monteith method (Allen, undated).**

ET<sub>a</sub> for Texas vegetation fitted within three natural cover types: (1) woody perennial vegetation, such as forests and shrublands, (2) non-woody perennial vegetation such as grasslands and savannahs, and (3) cultivated crops, both irrigated and non-irrigated. Each of these classes was applied to the MODIS NDVI\* estimate that had first been transformed to TM5 equivalency as described in Section D.

Within the GMA8 project area, in addition to the three major ET<sub>a</sub> classes, there are urban and water bodies that were not calibrated separately because no specific data were available for these cover types. ET<sub>a</sub> for urban areas was chosen, a 50:50 combination of woody and non-woody classes because urban landscaping in Texas generally contains a

combination of trees, shrubs and lawn. A set factor was used to estimate evaporation from water bodies.

## E.1 Site Descriptions and Analysis Procedures

Validation and calibration used measured ETa data. A series of twelve ET Flux towers in five locations were used for the analysis that covered various vegetation and crop types (Figure E1). The towers were located in highly variable climatic and vegetation regions across the United States, allowing for a robust analysis and calibration of RDC through a range of climate zones.

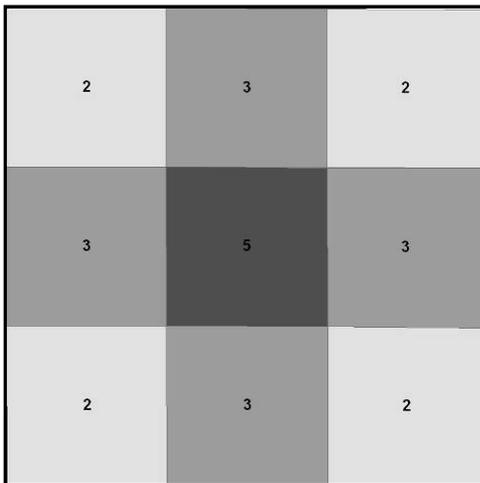


**Figure E1. Google Earth image showing AmeriFlux locations used for RDC analysis.**

The AmeriFlux network, established in 1996, is a group of 120 eddy covariance towers that monitor long-term variations of carbon dioxide, water vapor and energy exchange from various ecosystems in the western hemisphere (AmeriFlux, 2010). AmeriFlux plays a key role in the quantification and comprehension of the processes that regulate and affect variations in water and carbon fluxes. Data collected from these research sites are combined into consistent, quality-assured datasets by the Oak Ridge National Laboratory, Tennessee and are available at <http://public.ornl.gov/ameriflux>.

To accurately compare ET Flux measurements to the RDC method, weighted pixel grids were created around each tower (Figure E2). This approach can be shown to reduce error

in calculating ETa induced by air flow toward the flux tower location across radii with variable cover. Data were analyzed graphically as seven-day averages of measured daily ET flux centered on the day of the satellite overpass. Daily average ASCE Penman Monteith  $ET_0$  was calculated with the RefET software (Allen, undated) for each day within the seven day period at each site (the date of the overpass, plus and minus 3 days). Ground-truth ETa were then compared to the product of the spatially-weighted NDVI\* measured during the overpass and the coincidental seven-day average  $ET_0$ . This product was compared to the average ETa for the seven days of the paired RDC and flux data. The entire analysis consisted of 34 site/year combinations with 214 paired, seven-day ETa points developed from a total of 87 images. The sites, location, rainfall and crops for the individual towers are shown in Table E1.



**Figure E2. Grid and weighting factors used for comparison of NDVI\* to measured ETa. Rectangular grids were used where winds were strongly north-south; portions of the rectangle extending one pixel row above and below this square grid were weighted as 1.**

The workflow for the RDC-to-groundtruth comparison for each site was:

1. Download and organize the AmeriFlux ETa data.
2. Search for and download cloud-free Landsat TM5 scenes corresponding to the flux tower site and its period of operation.
3. Calculate reflectance for the Landsat TM5 data by equations in Irish (undated).
4. Calculate NDVI according to Equation D2.
5. Calculate NDVI\* according to Equation D3 following methods presented in Groeneveld and Baugh (2007).
6. Plot the flux tower site on the imagery in a geographic information system.
7. Plot the pixel locations as a grid to provide a weighting system for the NDVI\* values (Figure E2).
8. Adjust the grid for prevailing wind, if warranted.
9. Extract NDVI\* from cells in the weighted grid and calculate weighted NDVI\*.
10. Download weather data from the flux tower and calculate ASCE Penman Monteith  $ET_0$  using RefET (Allen, undated).
11. Plot the first-order ETa estimates ( $ET_0 \cdot NDVI^*$ ) against measured ETa.
12. Fit a transform so that the estimated ETa data becomes a competent predictor of measured ETa by class.

**Table E1. Sites used for calibration of RDC for use with MODIS.**

<b>Location Name</b>	<b>State</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Years</b>	<b>Average Annual Rainfall (in)</b>	<b>Vegetation</b>
<b>Freeman Ranch</b>	TX	29.940N	-97.990W	2005-2006	34-36	Woodland
<b>Freeman Ranch</b>	TX	29.9300N	-98.0100W	2005-2006	34-36	Grassland
<b>Ponca</b>	OK	36.7667N	-97.1333W	1998-1999	36-40	Winter Wheat
<b>Mead Site 1</b>	NE	41.1651N	-96.4766W	2003-2006	28-30	Maize
<b>Mead Site 2</b>	NE	41.1649N	-96.4701W	2003-2006	28-30	Maize/ Soybeans
<b>Mead Site 3</b>	NE	41.1797N	-96.4396W	2003-2006	28-30	Maize/ Soybeans
<b>Duke Open Field</b>	NC	35.9712N	-79.0934W	2005-2006	44-46	Grassland
<b>Duke Hardwood</b>	NC	35.9736N	-79.1004W	2002-2005	44-46	Mature Oak/Hickory
<b>Duke Pine Forest</b>	NC	35.9782N	-79.0942W	2002-2005	44-46	Loblolly Pines
<b>Sky Oaks New</b>	CA	33.3844N	-116.6403W	2005-2006	15-20	Chaparral
<b>Sky Oaks Young</b>	CA	33.3772N	116.6227W	2005-2006	15-20	Chaparral
<b>Sky Oaks Old</b>	CA	33.3739N	-116.6229W	2005-2006	15-20	Chaparral

## **E.2 ETa by Vegetation Class**

All classes of ETa, except open water, rise and fall dynamically with annual fluctuations in two controlling parameters:  $ET_0$ , the driving force for evaporation, and NDVI\*, a measure of vegetation vigor and water use. Examination of the first-order ETa curves for the individual flux sites indicated three natural groupings: (1) forested: vegetation dominated by woody species that shade the ground; (2) grassland: vegetation dominated by grasses, natural meadow or pasture (that may include sparse trees or shrubs—

savanna); and (3) crops. To apply RDC to develop ETa estimates, these classes were delineated spatially using USGS GAP data (described in section F).

The calibration analysis of the flux data was approached in three steps: (1) plotting the paired data for first-order RDC ETa estimates against the ground-truth measured ETa for each class, (2) calculating a linear first-order fit; and then (3) fitting a transform (if necessary) to correct the estimate so that the RDC-estimated ETa provided an unbiased predictor of the measured ETa. The simplified math for these steps is shown in Equations E2 and E3. Unlike the traditional dual-coefficient approach for estimation of ET expressed in Equation E1, this is a new approach that accommodates the effects of soil evaporation to now provide a new K from the function  $f(K)$  and an important improvement over the concept of an additive estimate for  $K_e \cdot ET_0$  that remains a static number or set tabular distribution through the entire growth cycle. This dynamic fitted variable was designated  $f(K)$ ; it is the correction for the lumped influences that cause the first-order RDC estimate of ETa to diverge from the measured ETa.

$$\text{First-order estimate:} \quad \hat{ETa} = ET_0 \cdot NDVI^* \quad \text{Equation E2}$$

Where  $\hat{\phantom{x}}$  designates estimate, and  $NDVI^*$  operates as a surrogate for  $K_{cb}$  for dual coefficient estimation of ETa.

$$\text{Transformed estimate:} \quad \hat{ETa} = f[NDVI^*] \cdot ET_0 \quad \text{Equation E3}$$

Where  $f(NDVI^*)$  is the fitted transformation of  $NDVI^*$  that includes the effects of  $K_e$  whose product with  $ET_0$  yields ETa (constituting a dynamic K factor).

The term “unbiased” is important for estimating ETa because of the natural variability that exists in the flux measurements and in the RDC estimates. Estimates of ETa that are not systematically either over or under the ground truth rates are unbiased although they contain scatter from estimation of  $ET_0$ , calculation of  $NDVI^*$ , and a suite of other issues that cause error in the actual flux measurement that constitute the ground truth for calibrating RDC. Because the physics, temporal variability, and spatial variability governing ET are highly complex—even the most careful measurements produce significant scatter. Although scatter must exist in these data due to the very nature of ETa and its measurement, this method approximates ETa in such a manner that the annual average of multiple and unbiased data points will define an accurate annual estimate of ETa.

The linear transformation for the suite of first-order estimates provided a lumped parameter correction to arrive at total ETa losses from the system, including canopy interception losses and soil surface evaporation. Examination of the distribution for the RDC-estimated and flux-measured ETa permits physically-based interpretation. It is evident that the character of the transformation is dominated by soil-surface evaporation because such transformation is necessary in such short-statured cover such as both cultivated crops and grassland classes, but is not necessary within forested vegetation that is self shading, hence, apparently greatly constraining soil surface evaporation. In some ways the transformation functions work as fitted  $K_e$  values, but they are more dynamic

than the additive way a  $K_e$  is generally used within the dual coefficient method. This is because the  $f(K)$  replacing  $K_e$  within this application is multiplicative and can vary with the magnitude of NDVI\*.

The remaining two ET class types require separate estimation techniques. ET<sub>a</sub> for urban vegetation is estimated simply as 0.5 grassland plus 0.5 forested rates. Evaporation from open water was estimated using the relationship from Allen et al. (1998) as  $1.05 \cdot ET_0$ .

### E.2.1 Woody Vegetation—the Forested ET<sub>a</sub> Class

Vegetation dominated by trees and shrubs is a complex category for estimation of ET because it includes sites that vary from hydric to mesic conditions within the project area. In comparison to crops, which have controlled canopy architecture and growth that follows predicted stages, woody vegetation can be found in virtually any stage. To provide a wide range of climatic conditions and vegetation, sites were chosen in shrub-dominated California chaparral, in woodland in Central Texas, and in an experimental forest run by Duke University in North Carolina. Together, these sites constituted 12 site/year combinations (Table E2) representing 74 separate data points.

**Table E2. Sites, dominant vegetation types, general locations, and number of years of operation for each site dominated by woody plants—the Forested ET class.**

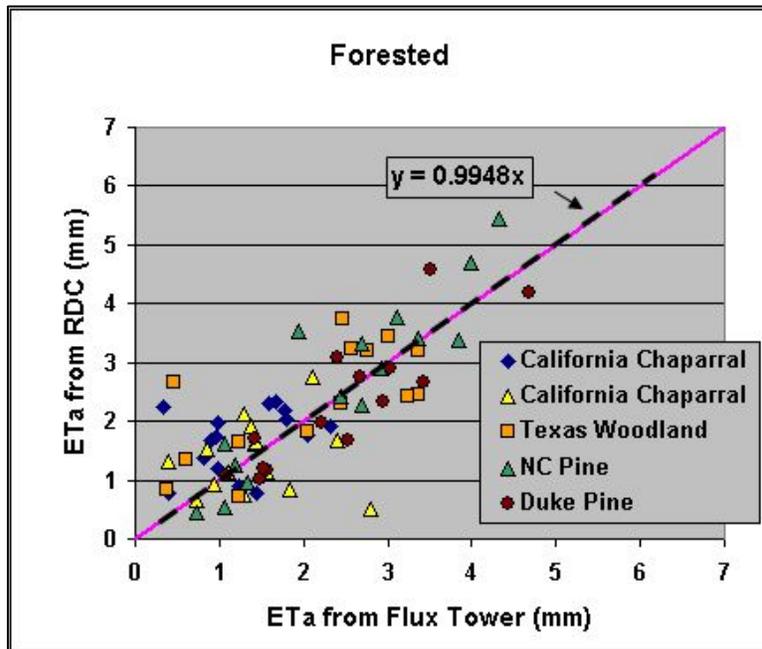
Sky Oaks New	Shrub-dominated chaparral	Southern California	2
Sky Oaks Young	Shrub-dominated chaparral	Southern California	2
Freeman Ranch Woodland	Live oak and cedar	Central Texas	2
Duke Hardwood Forest	Eastern hardwood forest	North Carolina	3
Duke Pine Forest	Eastern pine forest	North Carolina	3

The locations evaluated for this ET class include California chaparral, specifically because this vegetation type contains sclerophyllous evergreen species that are characterized by drought hardy woody plants that have persistent leaves with thick waxy coatings, leaves that are retained through the year and often for multiple years. The GMA8 region contains such vegetation in the form of forests of live oak.

The forest sites in North Carolina were included for several reasons, the first being that the deciduous and pine forests of the far northeast portions of the project area have many of the same species and characteristics as those in North Carolina. Another reason was to test whether pine forests have significantly different water use than hardwood forests when judged by their level of NDVI\*. This is important because the project area contains intermixes of both hardwood forests and pine forests in the northeast.

The Freeman Ranch data provided the only woodland data available within the study area, so this site was included. The location of the representative forested ET class site, however, was nearly immaterial for the RDC method. The input ET<sub>0</sub> is determined by local climatic inputs and adjusts estimates of ET<sub>a</sub> to accurately represent any location of interest for RDC estimates for a given class.

All first-order data points for woody vegetation were collected together in Figure E3. The first-order estimates fit well with the measured values; no separate calculation of  $K_e$  was necessary for estimation of the  $ET_a$  for vegetation dominated by woody species.



**Figure E3.** All first-order estimates of  $ET_a$  estimated by RDC plotted against their paired flux-tower measured values. The first-order estimate is a competent and unbiased estimator of  $ET_a$  without transformation.

### E.2.2 Herbaceous Vegetation—the Grassland ET Class

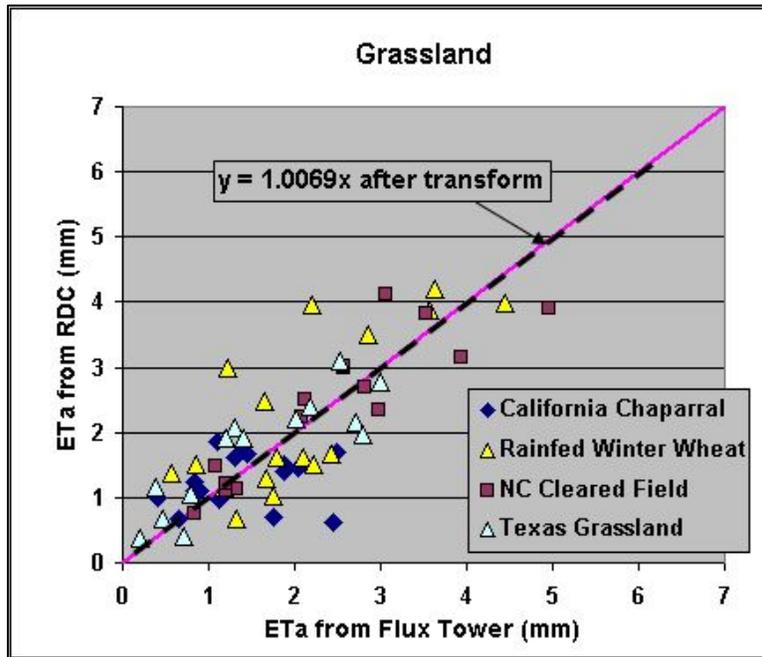
The Grassland class is herbaceous dominated; non-grass herbaceous species are a significant component of the class. Savanna vegetation is included in this class because even though trees may be present within savanna cover, the grass component dominates. For example, Sky Oaks Old is included within the Grassland class, even though it is within California chaparral—it is categorized as savanna and fits within the grassland data and its transform. The data for the grassy sites totaled 10 site/year combinations and 62 separate points.

**Table E3.** Sites, vegetation cover and number of years chosen for evaluating  $ET_a$  for grasslands.

Sky Oaks Old	Chaparral savannah	Southern California	2
Ponca wheat	Cultivated winter wheat	Northern Oklahoma	3
Duke open field	Successional cleared field	North Carolina	3
Freeman Ranch Grassland	Grassland	Central Texas	2

Winter wheat was included in the Grassland class even though it is a cultivated crop. Its water use behaved more like grasses than crops and the transformation for grasses worked reasonably well for winter wheat, as can be seen in Figure E4. Some of the

values, however, were elevated slightly above the 1:1 line. Because only one winter wheat site was evaluated, winter wheat will be dealt with in the Crop ET estimation class because of (1) the uncertainty due to data limitations—only one winter wheat site was available, (2) the areas planted as winter wheat change each year, thereby requiring classification for each year of interest, (3) specific cropping data is not available for most years of interest from the USDA (that are provided in a product known as the Cropland Data Layer), (4) winter wheat is not distinguished from other crops within the GAP data that is the basis for locating the classes for estimating ETa, and finally, (5) minimal error would accrue due to this decision.



**Figure E4.** Transformed first-order estimates of ETa for Grassland sites plotted against their paired flux-tower measured values. The transformation was  $Kc = (1.3 \cdot NDVI^*)$ .

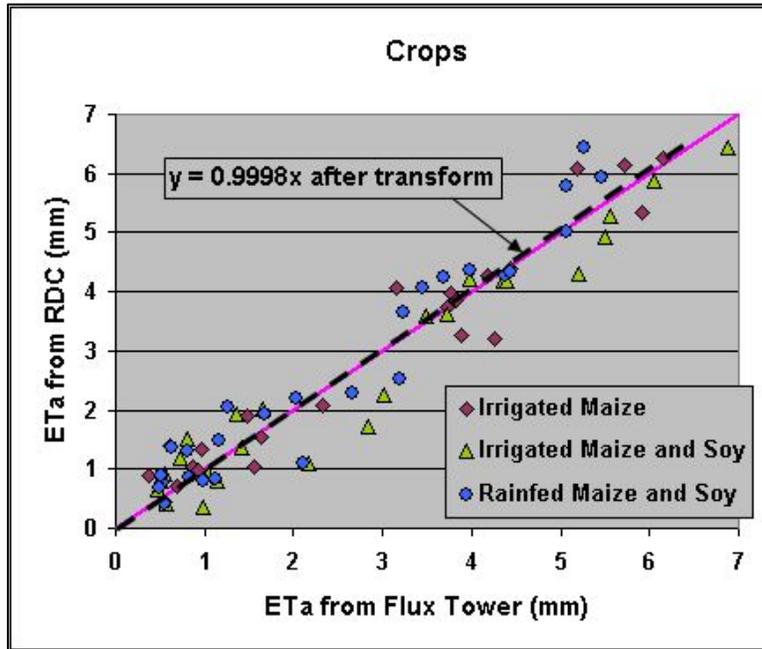
### E.2.3 Crops ET Class

This class consists of cultivated crops and, for application to the GMA8 project area, included winter wheat that was analyzed with Grassland as described above. The first-order estimates of the two crops that were evaluated, maize and soy, can be fit with a single transformation despite large differences in physiology, physiognomy, taxonomy, and growth pattern. Thus, for purposes of estimating ETa, the crop relationship that was developed from the flux data was used for all crops. This relationship is extremely robust because the maize and soy grown on Mead Sites 1 and 2 received supplemental irrigation, while Site 3 was entirely rain-fed: all sites fit on the same relationship.

**Table E4.** Sites, vegetation cover and number of years chosen for evaluating crop ETa.

Mead 1 Irrigated	Maize	East Central Nebraska	4
Mead 2 Irrigated	Maize and soybeans	East Central Nebraska	4
Mead 3 Rainfed	Maize and soybeans	East Central Nebraska	4

Even with large differences in crops and treatment, data from these three sites fit well together within one graph and were calibrated with a single transformation (Figure E5). These crop data were acquired in Mead Nebraska that has summertime humidity and temperatures like the GMA8 region and include 12 site/year combinations and include 78 data points.



**Figure E5. Transformed first-order estimates of ETa for crops, maize and soy, plotted against their paired flux-tower measured values. The three symbols represent three different flux towers. The crop transformation was:  $Kc = [0.3 + (0.765 \cdot NDVI^*)]$ .**

### E3. Summary

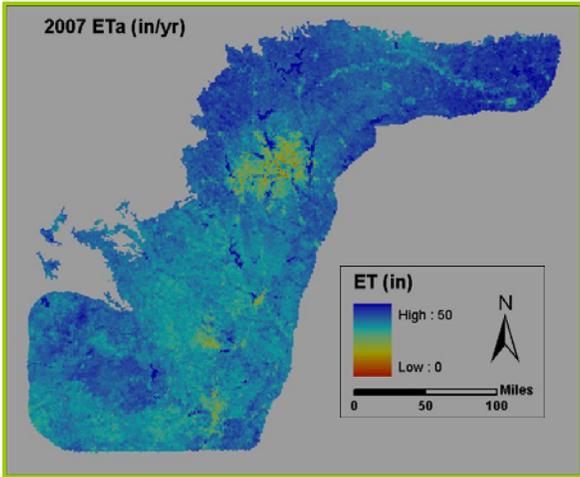
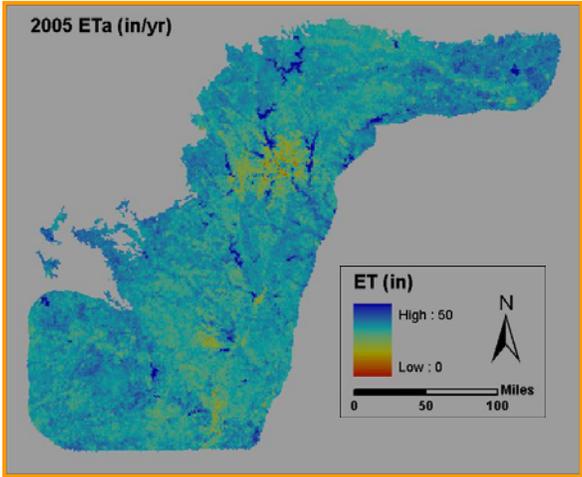
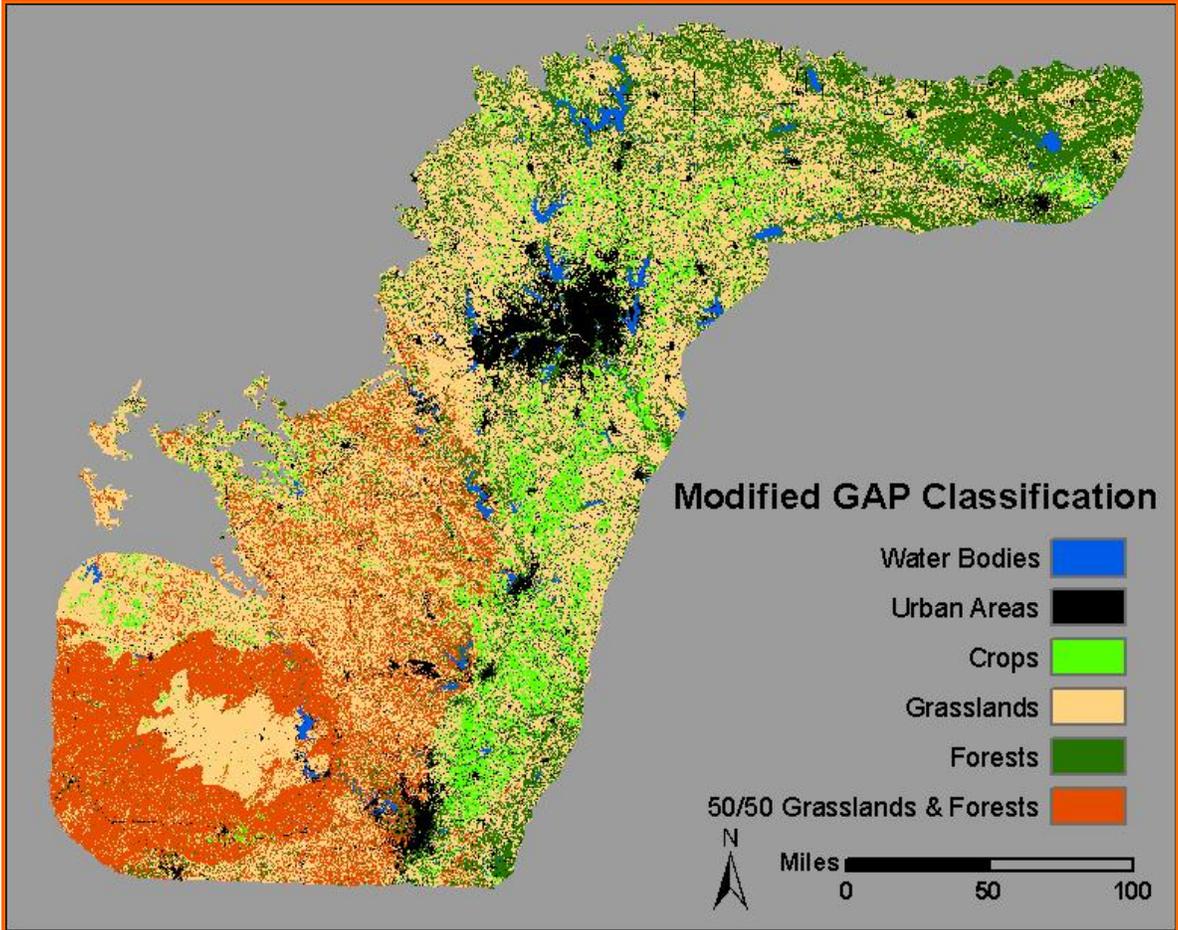
ETa was calculated for vegetation by MODIS Pixel using the RDC estimation that employed  $ET_0$ ,  $NDVI^*$  and fitted functions by vegetation cover type. Three cover types were chosen through calibration to typify the vegetation within GMA8, grassland, woodland and cropland. For estimation of ETa using the spatially correct vegetation cover type relationships, GAP data were then consulted to estimate the vegetation across the GMA8 region as described in Section F.

#### Data provided from the work in this section include:

- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters.
- MODIS raw and intermediate calibration data.
- MODIS-derived quarterly  $NDVI^*$  for 2005 and 2007.



# F. CALIBRATING ETA: SPATIAL EXTRAPOLATION





## F. Application of ETa Estimation

Because of scale and ease, MODIS 8-day cloud free products were chosen as the basis for ETa estimation. Given that MODIS was the remotely sensed data on which to base ETa estimates, calibration for variable conditions required selection of recent years. MODIS Terra which was used for this study was launched in December, 1999. The method for calibration was to choose the wettest and driest years during the MODIS record to set end members for expected ETa.

This section calibrates ETa for precipitation. Sections D and E described the method that was used for estimating ET that was developed using Landsat TM5 data, calibration for MODIS NDVI to yield Landsat TM5-equivalent NDVI, and calibration for NDVI<sub>o</sub> used to derive NDVI\* from the MODIS data. This section describes the application of the remotely sensed dual-coefficient (RDC) ETa estimation techniques within GMA8 to construct a spatially accurate precipitation-driven model.

Eight steps are summarized below that were required for development of the relationships that were needed to model ET. Steps 1 through 3 were calibration steps that used data subsets as described in Sections A, D and E. Steps 4 through 7 were conducted within the GMA8 Groundwater Availability Model (GAM) grid cells. The final step translated the newly developed ETa prediction into raster formats appropriate for use within the GWr model.

Overall Step 1. RDC method of ETa estimation was calibrated for GMA8 with Landsat TM5 using datasets of NDVI\* and AmeriFlux ET flux data.

Overall Step 2. Wet (2007) and dry (2005) years were chosen from the recent record to be coincident with the availability of MODIS data. These Ppt extremes were chosen to capture the expected range of ETa conditions across GMA8.

Overall Step 3. GAP data were used to classify ETa classes across the GMA8 region.

Overall Step 4. Spatially explicit NDVI\* from the MODIS data were paired with ET<sub>0</sub> to calculate ETa according to the GAP identified ETa class. These data were calculated as quarterly data that were then extracted to the GAM model grid.

Overall Step 5. Spatially explicit quarterly Ppt was summed to annual Ppt for the wet and dry years.

Overall Step 6. To provide an estimator for any Ppt year between the wet year and dry years, ETa and Ppt were used to derive a linear regression equation to estimate ETa from Ppt as  $ETa = a + (b * Ppt)$ . About 90% of the GAM grid cells behaved as expected by having a positive slope with higher ETa during the wet year. In other pixels, however, this produced a slope that was negative, due to a transient condition of cloudiness during the 2007 wet year. For such cells showing dry year ETa in excess of wet year, the ETa during the wet year was set equivalent to the

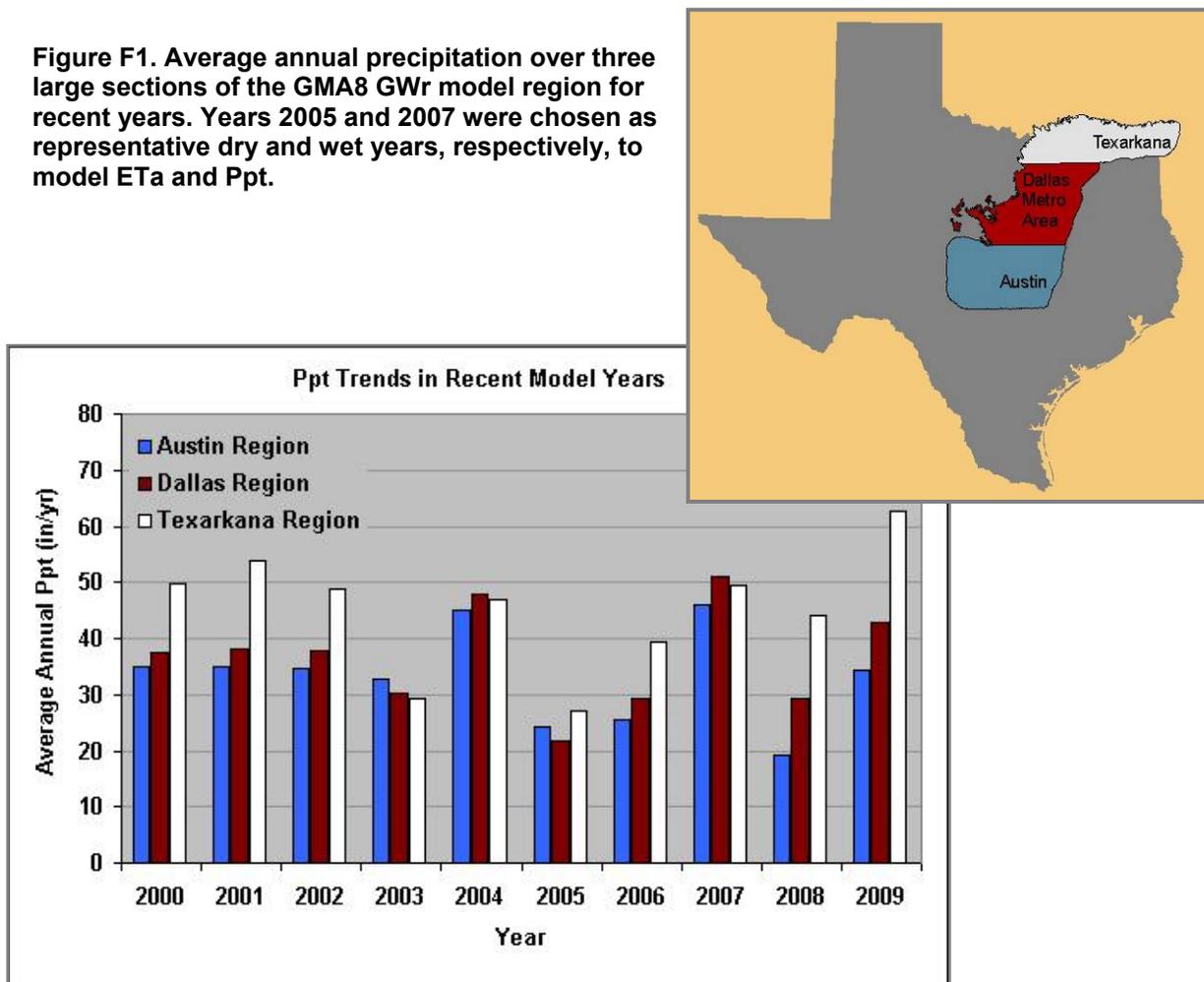
dry year 2005 wet year, effectively setting the slope as one and the intercept equal to the 2005 intercept for that cell.

Overall Step 7. The ETa-Ppt slope and intercept were exported as rasters to be input parameters in the GWR model.

### F1. Selection of Representative ETa Model Years

Two years were selected for calibration of ETa for wet and dry conditions. To represent the range of Ppt that occurs within the model area, a record dry year (2005) and a record wet year (2007) were chosen from recent weather record. Figure F1 shows the large-scale average Ppt across three major model regions from these Ppt years, 2005 and 2007. These were selected on the basis of consistently low and high Ppt across the project.

**Figure F1.** Average annual precipitation over three large sections of the GMA8 GWR model region for recent years. Years 2005 and 2007 were chosen as representative dry and wet years, respectively, to model ETa and Ppt.



## F2. MODIS NDVI\* Processing and ETa Quarterly Input Variables

MODIS eight-day composite images data were acquired for 2005 and 2007, the dry and wet focus years for ETa analysis. Two MODIS scenes were necessary to cover the GMA8 project region, 47 images were collected per year, per scene, totaling 188 scenes for the full analysis. A sensitivity analysis performed on these data found that the ETa within the quarters was not highly variable and this enabled a simplifying step for calculation of total ETa per quarter for each year.

To develop quarterly estimates of NDVI\*, raw MODIS scenes were evaluated, merged for simplicity, and processed to TM-equivalency before calculating and averaging quarterly NDVI\* using seven steps. Calculations in Steps 4-6 represent a variant of the method generally used for calculating NDVI\*.

Processing Step 1. Any remaining clouds or cloud shadows in the “cloud free” MODIS product were masked out and reclassified as no data. Even though the 8-day MODIS product was used, some clouds are still present in a few scenes.

Processing Step 2. The two MODIS scenes, H9 V5 and H10 V5, were merged into a single scene for each 8-day product and then clipped to the approximate outline of the project area to reduce the volume of data and the required processing steps.

Processing Step 3. NDVI was calculated for each merged image with Red equal to MODIS Band 1 and near Infra-red equal to Band 2.

Processing Step 4. The use of MODIS for the estimation was necessary because its scale enabled the estimation of ETa over the study area; however, the ETa relationships were developed using Landsat TM. This required that the data be calculated as TM equivalent. First however, all scenes were indexed by an offset calculation to correct for regional haze, thin clouds, or other issues. This was accomplished by measuring the NDVI values over a quarry devoid of vegetation cover. NDVI was first evaluated on one clear reference scene to determine the bare soil/rock background value of NDVI. The difference of NDVI values for this quarry area on all other images from this reference scene were then added or subtracted to all NDVI values on each MODIS image. This variant step enabled checking each MODIS image for comparability with the others.

Processing Step 5. A regression equation was applied to calculate values of MODIS NDVI that were equivalent to TM NDVI (Figure D1).

$$\text{TM NDVI} = 0.8921 \cdot \text{MODIS NDVI} - 0.0518.$$

Processing Step 6. The MODIS NDVI were transformed to NDVI\* values by applying stretching parameters  $\text{NDVI}_0$  and  $\text{NDVI}_S$ . Because of the indexing of Step 4, a set value for  $\text{NDVI}_0$  was used that represented the pixels with bare surfaces at the quarry (Figure D2). A standard  $\text{NDVI}_S$  (saturated NDVI) value was applied.

$$\text{NDVI}^* = (\text{TM equivalent MODIS NDVI} - .0705) / (0.8600 - 0.0705).$$

Processing Step 7. NDVI\* was averaged for each quarter to reduce the 47 eight-day MODIS composite images for each year to four rasters to represent NDVI\* for each quarter of the wet and dry years.

Processing Step 8. Quarterly NDVI\* data were extracted to the GAM grid that were paired with quarterly modeled ET<sub>0</sub> and quarterly interpolated Ppt (both described in Section A) were extracted for each GAM grid. Because of the geospatial complexity introduced by the GAM grid for these calculations, they were accomplished in spreadsheet software (Excel). Data in Table F1 summarize quarterly variability in NDVI\*, Ppt, and ET<sub>0</sub> across the three major regions of GMA8 depicted in Figure F1.

**Table F1. Quarterly variables used to calibrate ETa across GMA8. Regions are depicted in Figure F1.**

	Qtr	2005 Average			2007 Average		
		NDVI*	Ppt (in/Qtr)	ET <sub>0</sub> (in/Qtr)	NDVI*	Ppt (in/Qtr)	ET <sub>0</sub> (in/Qtr)
<b>Austin Region</b>	<b>Q1</b>	0.39	7.58	10.33	0.36	11.61	9.09
	<b>Q2</b>	0.50	6.44	20.45	0.54	19.18	14.45
	<b>Q3</b>	0.46	7.81	19.73	0.53	10.81	18.51
	<b>Q4</b>	0.35	2.45	11.01	0.38	3.38	10.81
<b>Dallas Region</b>	<b>Q1</b>	0.35	8.09	9.76	0.36	11.59	8.55
	<b>Q2</b>	0.52	5.17	20.63	0.56	22.35	13.45
	<b>Q3</b>	0.45	6.63	19.87	0.53	10.27	18.73
	<b>Q4</b>	0.28	1.65	10.65	0.36	5.50	9.74
<b>Texarkana Region</b>	<b>Q1</b>	0.38	9.25	8.16	0.40	9.62	7.98
	<b>Q2</b>	0.63	6.32	18.38	0.66	19.96	14.81
	<b>Q3</b>	0.58	9.85	17.96	0.63	11.11	17.79
	<b>Q4</b>	0.37	2.72	8.98	0.43	7.83	8.25

### F3. Gap ETa Classes and Their Application

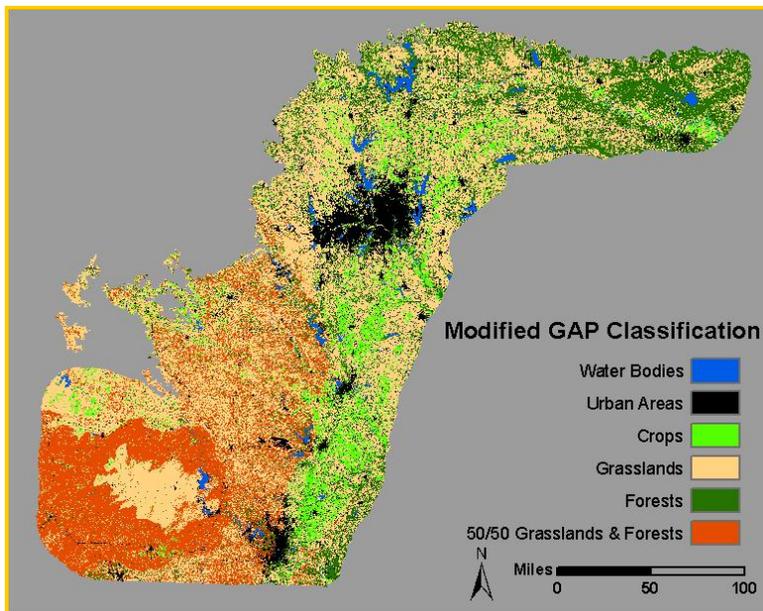
After calibrating ETa estimation to specific cover types in Section E, USGS GAP data were chosen as the basis for mapping to apply the five ETa classes. GAP data, mapped nationwide, included GIS layers of detailed ecological descriptions of land cover. The GAP survey combined multiple sources of data including digital elevation models and satellite imagery from multiple years and seasons that are integrated through a common classification system that classifies land cover nationwide (USGS, Accessed 2010).

To define the vegetation that influences ETa across the project region, each grid cell was described in terms of the GAP land cover it contained. The detailed ecological descriptions of the GAP analysis were used to spatially define the five classes of land cover in GMA8: water, urban, crop, grassland, and forest. These five groups, chosen for their relevance to ETa estimation, are compilations of similar GAP land cover classes. The grassland ETa class, for example, combines GAP analysis grasslands and prairies

with bare ground and ephemeral meadowlands because similarities permitted lumping into a common ETa class.

GAP classes that did not readily fit into one of the ETa groups were evaluated using satellite imagery provided on the USGS Land Cover Viewer assignments. Within the project region, all five land-cover classes were defined as well as a hybrid region detected over the Edwards Plateau—this region exhibited sparsely forested areas interspersed with grasslands and was modeled, like the urban areas, as one-half grassland and one-half forested. (Figure F2 and F3). A new raster was generated to spatially represent the extent of these five classes as well as a hybrid region of half forest and half grassland over the Edwards plateau.

The new raster for ETa classes maintained the GAP 30 meter resolution. These data, with a much higher resolution than the GAM one mile grid cells, were used to proportion the influences of different ET classes represented within each GAM grid cell. The area of individual pixel ETa classes were extracted from the ETa class raster for each GAM grid cell and used to create a decimal fraction to apply in ETa calculation. The decimal fraction of each ETa class was applied to the RDC calibrated  $f(K)$  for each ETa class. These values were then summed to form weighted average ETa within each grid cell.



**Figure F2. Five key ETa classes were mapped in the project region using the land cover data compiled by the USGS Gap (Gap Analysis Project).**

#### **F4. Quarterly and Annual ETa for Representative Wet and Dry Years**

Quarterly ETa was estimated for each cell in GMA8. The calculation was tailored to each individual cell by considering the proportions of land cover represented within the cell and was driven by quarterly  $ET_0$ . The decimal fractions of each ETa class were applied to estimate ETa according to Equation F1. The unique  $f(K)$ , statistically fitted to each of the ET land cover classes, was previously derived during RDC calibration in Section E.

Table F2 summarizes the project-wide results and Figure F4 illustrates the quarterly and annual ET estimates for the wet (2007) and dry (2005) calibration years.

$$ETa_i = \sum_{t=0} [f_k [NDVI^*]_i * \sum ET_{0i}]_t \quad \text{Equation F1}$$

Where *i* is the *i*th pixel, *k* is the *k*th function for the ET land cover class, NDVI\* is an average value for the quarter, designated *t* and summed for all four quarters of the year.



Figure F3. Examples of grassland (top), forested (middle), and blended grassland / forested land cover (bottom) (Google Earth, Accessed 2012). Due to the proportions of both grassland and forest areas over the Edwards Plateau (extent shown in Figure F2), ET estimates for the region were calculated as one-half grassland and one-half forest.

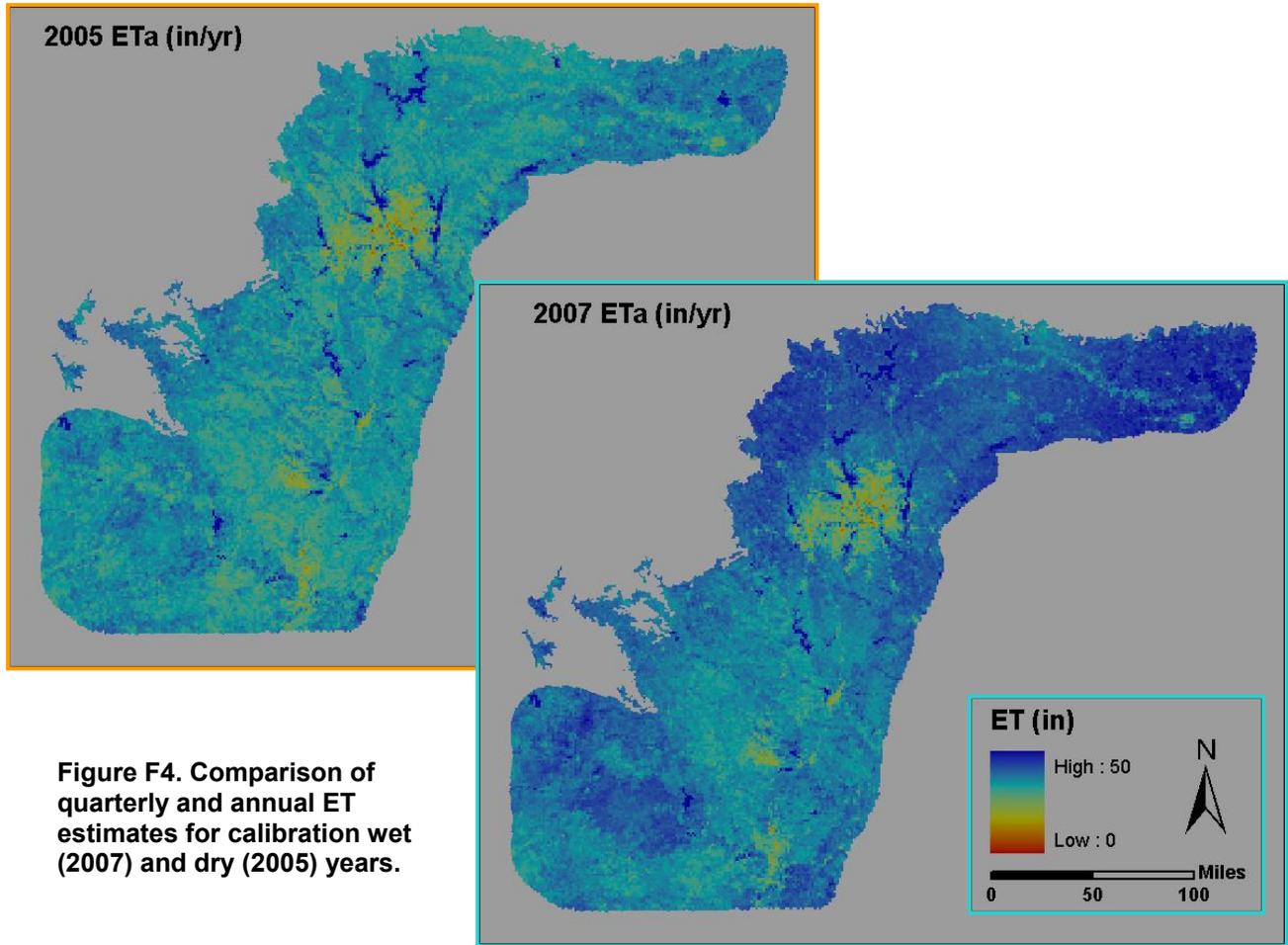
**Table F2. Project-wide averages of ETa driving variables and ETa.**

Average ET Parameters for Record Wet (2007) and Dry (2005) Years						
		Q1	Q2	Q3	Q4	Annual Sum (NDVI* is Annual Avg.)
Precipitation (in)	2005	8.25	5.98	7.97	2.26	24.46
	2007	11.04	20.42	10.76	5.36	47.59
ET <sub>0</sub> (in)	2005	9.53	19.94	19.28	10.31	59.05
	2007	8.60	14.22	18.36	9.72	50.91
NDVI* (f(x) transformed)	2005	0.38	0.54	0.49	0.34	0.44
	2007	0.37	0.59	0.57	0.39	0.48
ETa (in)	2005	4.59	13.38	11.84	4.48	34.29
	2007	10.30	13.38	12.99	4.90	41.57

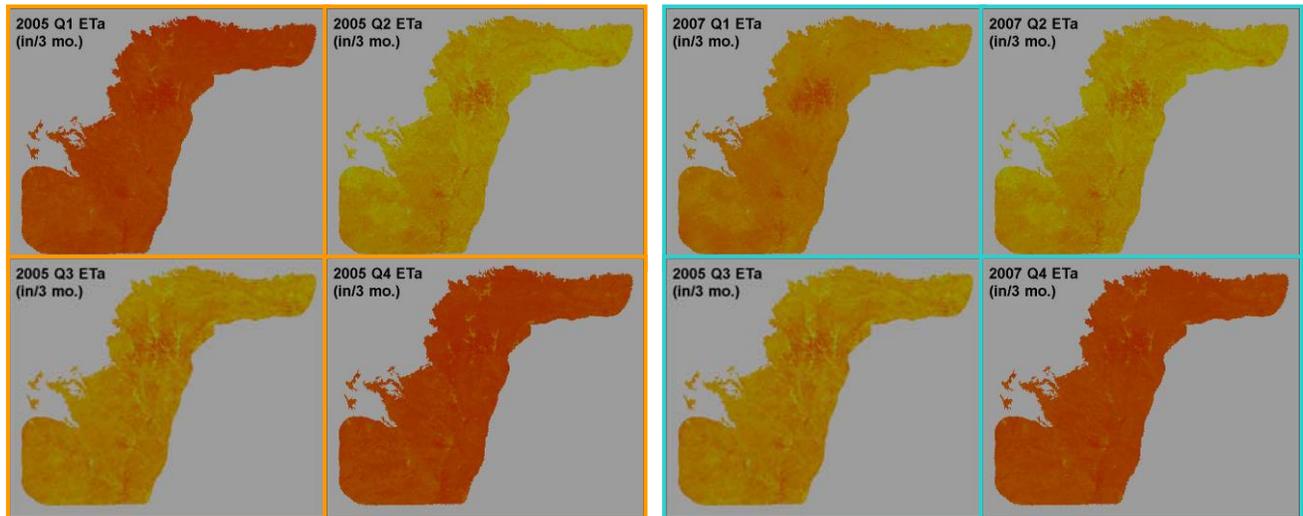
### **F5. Intermediate Year ETa by Slope between Wet and Dry Years ETa**

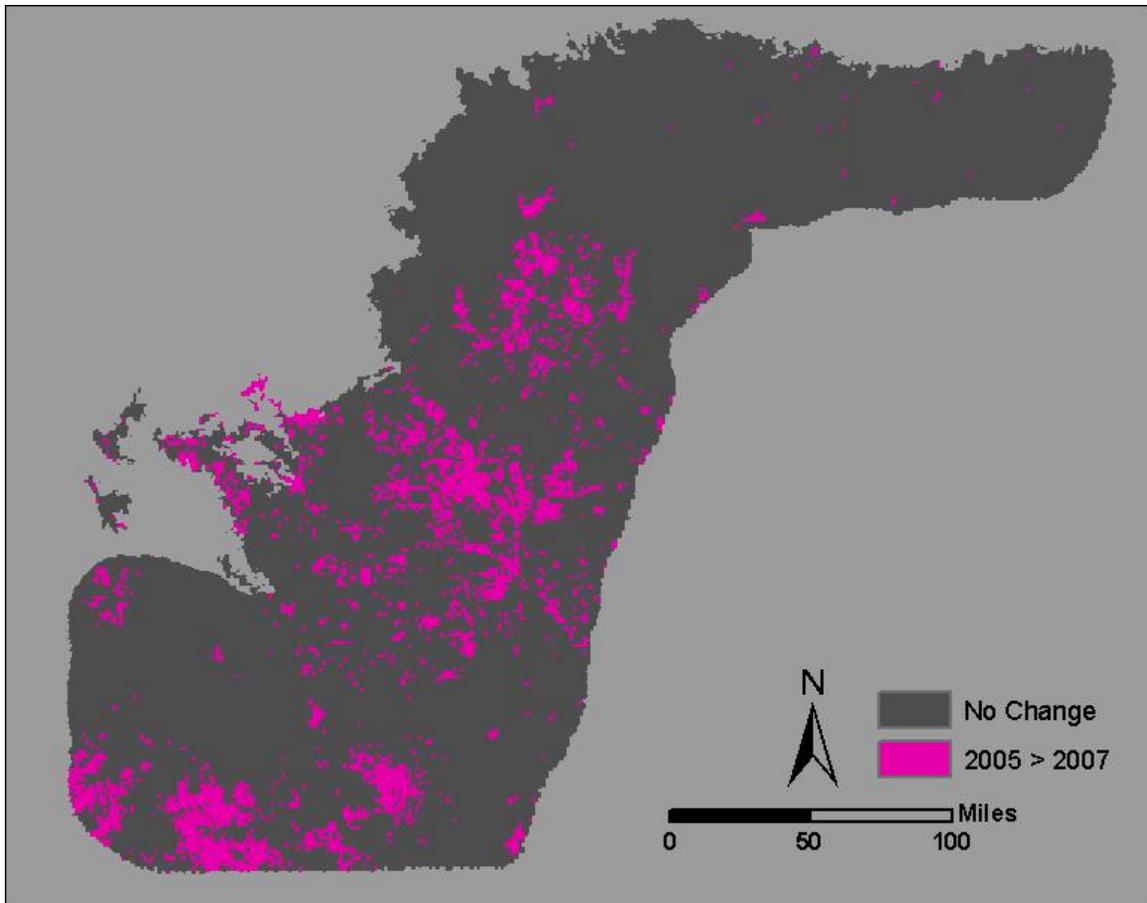
The end members for high and low precipitation were expected to yield a positive slope where the wet year ETa was higher than the dry year ETa. The concept for using slope and intercept to define the relationship for each pixel between wet and dry years was developed to provide ETa estimation as predicted by Ppt for years that are intermediate to the two end member years. This method permits encoding the ETa vs. Ppt relationship throughout the project region using only two rasters portraying slope and intercept.

Although calibrating using record dry and wet years is a sound concept, when applied to large and highly variable regions such as GMA8, this method may provide results in some locations that are contrary to expected, i.e., when ETa increases during a dry year over that measured in a wet year. This response would not be expected in an arid environment; however, much of the GMA8 region is relatively mesic, moderate, not wet or dry. Under these conditions, a reversal of the expected relationship where ETa increases during a dry year can occur. A look at Table F2 provides the reason why this is so. When averaged across the entire GMA8 region, ETa during a very wet year was only about 21% higher than a very dry year. A large range is lacking and the majority of the difference between the two years occurred during the first quarter (January through March), the difference between years is not very great in terms of the calculated ET even though the Ppt is different. A reversal in ETa occurred due to cloud cover and higher humidity during the wet year, reducing ET<sub>0</sub> so that the ETa calculated from it was actually lower. This relationship likely does not concatenate through all years from dry to wet (i.e., as the years from dry to wet vary, a bit more rain would not be expected to have a slight decrease in ETa). To constrain the possible error arising from this problem, the values of ETa for the wet year were set equal to the dry year, hence, having a slope of 1. Figure F5 shows the locations where this reversal occurred and where slope was set to 1.



**Figure F4. Comparison of quarterly and annual ET estimates for calibration wet (2007) and dry (2005) years.**





**Figure F5. In some locations, the ETa calculated for the calibration dry year (2005) exceeded the ETa calculated for the calibration wet year (2007). The unexpected result for ETa in a dry year exceeding the ETa in the same location for a wet year was likely caused by cloud-cover suppression of  $ET_0$  in the wet year in a system without a great deal of difference in NDVI\* from year to year.**

To correct for underestimation of ETa, the 2007 estimate was replaced with the 2005 ETa estimate. This change was applied across 10.9% of the model area as shown in pink. Reservoirs were identified in this process since  $ET_0$  tends to be depressed in wet years, hence, ETa in wet years for reservoirs, calculated by a simple k factor for water bodies as  $1.05 * ET_0$ , is lower than in dry years. Reservoirs remained within these spatial calculations at this stage were later removed since the land-surface calculations for GWr for the majority of the region are inappropriate for application to reservoirs that may or may not have recharged depending upon regional water tables, intersections with aquifer outcrops and underlying permeability and the position and influence of poorly permeable layers of fine sediments that can build up.

## F6. Linear Relationship of ETa to Precipitation Stored Per Pixel

Quarterly ETa estimates were summed to represent annual ETa. The Ppt and ET estimates for the wet and dry years were then used as end members to fit simple linear relationships for ETa by precipitation. Each grid cell was fitted with a unique equation to derive annual ETa from precipitation—the equation can be used to both interpolate between the wettest and driest conditions in the model record and to extrapolate ETa beyond the precipitation range.

A logical precondition was applied to these data disallowing ETa to exceed annual Ppt because the GWr model contains no memory function to account for ETa in excess of Ppt. ETa was then estimated for each grid cell over the span of the 50-year model using Ppt as the driving variable and the two model parameter rasters representing the slope and intercept of Equation F2. Three steps were taken to estimate ETa:

1. Annual Ppt, ETa slope, and ETa intercept rasters were used to solve Equation F2.

$$ETa_i = a + b * (Ppt_i)$$

**Equation F2**

2. ETa in urbanized areas was adjusted for the model years before urbanization using the dynamic ETa adjustment process described in Section G: Modeling Urbanized Areas
3. ETa was capped at annual precipitation using Equation F3.

$$\begin{aligned} \text{If } ETa_i > Ppt_i, ETa_i &= Ppt_i \\ \text{If } ETa_i \leq Ppt_i, ETa_i & \end{aligned}$$

**Equation F3**

## F7. Summary

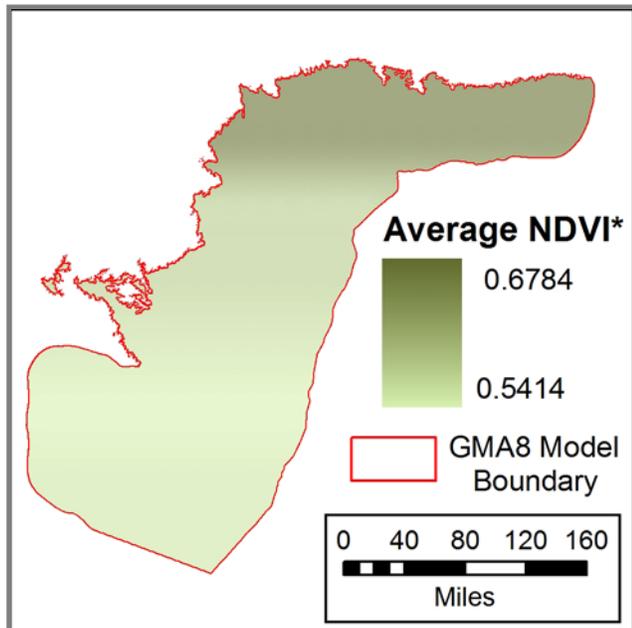
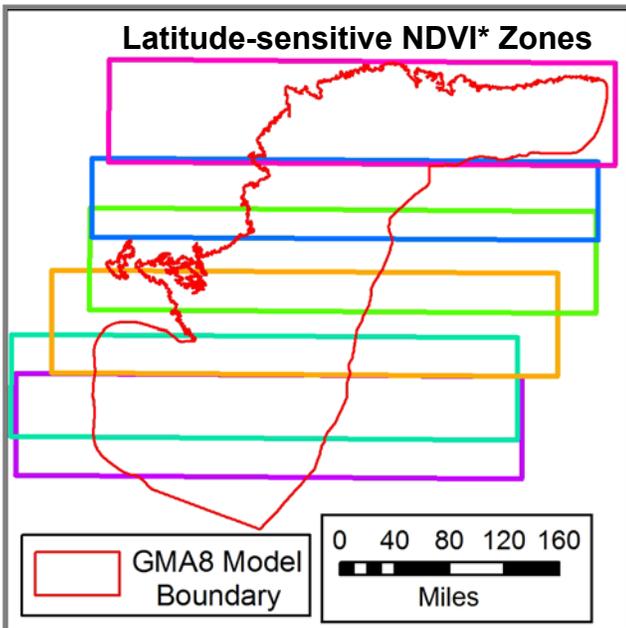
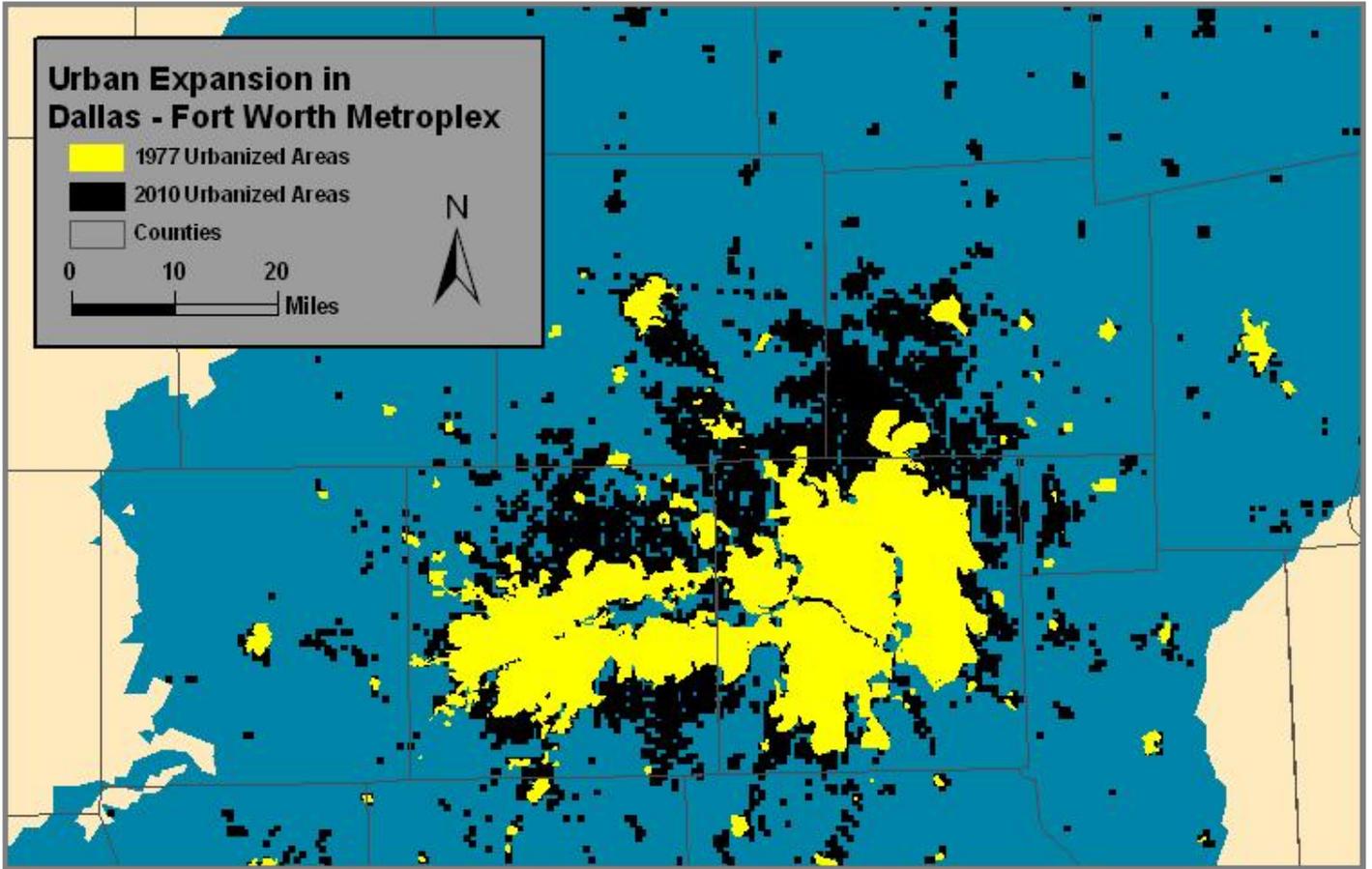
The application of remotely-sensed dual coefficient ETa estimation to GMA8 was used to estimate ETa, a key variable for ground water recharge estimation. HydroBio used the RDC method that was developed from Landsat TM5 for predicting ETa that used canopy greenness as a scalar against  $ET_0$ . RDC calibrated to three ETa classes of land cover, grassland, woodland and cropland, were modified for application to MODIS data. The modified RDC method was combined with USGS GAP for ETa prediction throughout the project area.

**Data provided from the work in this section include:**

- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters.
- MODIS raw and intermediate calibration data.
- MODIS-derived quarterly NDVI\* for 2005 and 2007.
- Modified GAP land cover classification raster provided in model parameters geodatabase and as an individual raster.



# G. MODELING URBANIZED AREAS





## G. Modeling Urbanized Areas

Urbanization over the 50-year span of the GMA8 groundwater recharge (GWr) model has altered the natural recharge conditions within significant areas of the GMA8 project area. Two critical components of GWr are altered by urbanization:

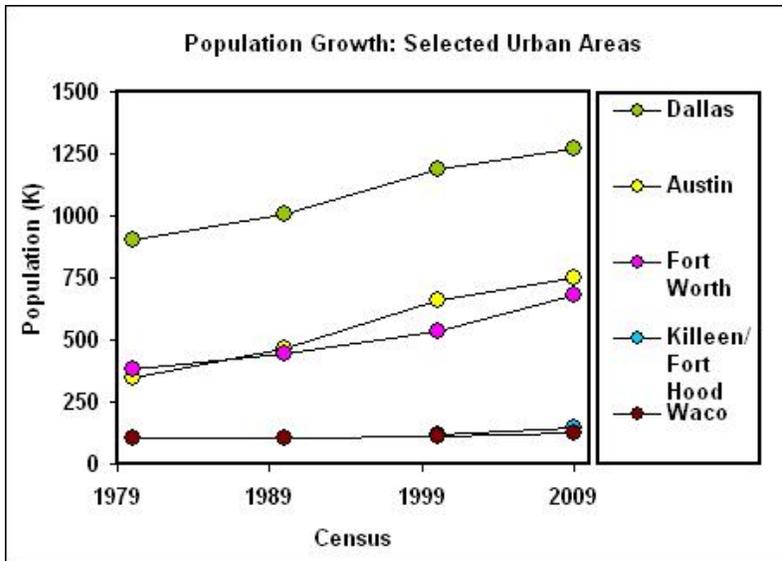
1. Discharge (Q) increases over urbanized areas due to changes in infiltration capacity with the proliferation of urban hardscape. This causes divergence from simple soil permeability (Perm) based estimates of Q.
2. ETa decreases over urbanized areas due to depressed NDVI, making the urbanized ETa calibration constructed from recent satellite imagery insufficient for representing ETa in areas before they were urbanized.

A two step approach was used that first identified those analysis grid cells with a significant enough portion of urban land use to affect Q and ETa within that cell. The second step used NDVI\*, the normalized difference vegetation index that has been corrected to yield a zero value for bare/unvegetated areas and unity for peak expected vegetation expression. The NDVI\* calculated for each urban grid cell was compared to a modeled average for the surrounding vegetated areas outside the urbanized zone. The ratio between these values represents a value of how much the urban hardscape had depressed the expected average NDVI\* signature.

Fractional NDVI\* was developed as a proxy for the degree of Q and ETa affecting urbanization within a model grid cell. The fraction represents a reduction in surfaces with the capacity for infiltration and ETa relative to the surrounding undisturbed landscape. This decline in NDVI\* relative to undisturbed areas is a proxy for how much of the area within each grid cell was covered by non-NDVI generating cover, such as roads, parking lots and building footprints that render the ground surface impermeable. For example if half of an area of interest is covered by non-NDVI\* generating cover, then the NDVI\* value will be about half of the expected level. Because the extent of urbanization has changed over the model calibration years, from 1960 to 2009, the urban extent was developed for four different time-steps in order to limit urban model adjustment to the extent of urbanization appropriate for each model year.

### G1. Historic Urban Boundaries

To model the urban changes throughout the model record, areas of population growth were identified and urban boundaries were determined using historical datasets. To identify the major urban areas that required dynamic urban boundaries, US Census data were evaluated. In the last 30 years, Austin, Dallas, and Fort Worth populations grew, while the Waco population remained fairly constant (Figure G1). Austin, Dallas, and Fort Worth were flagged for hand-analysis in the 1977 land-cover maps provided by TWDB.



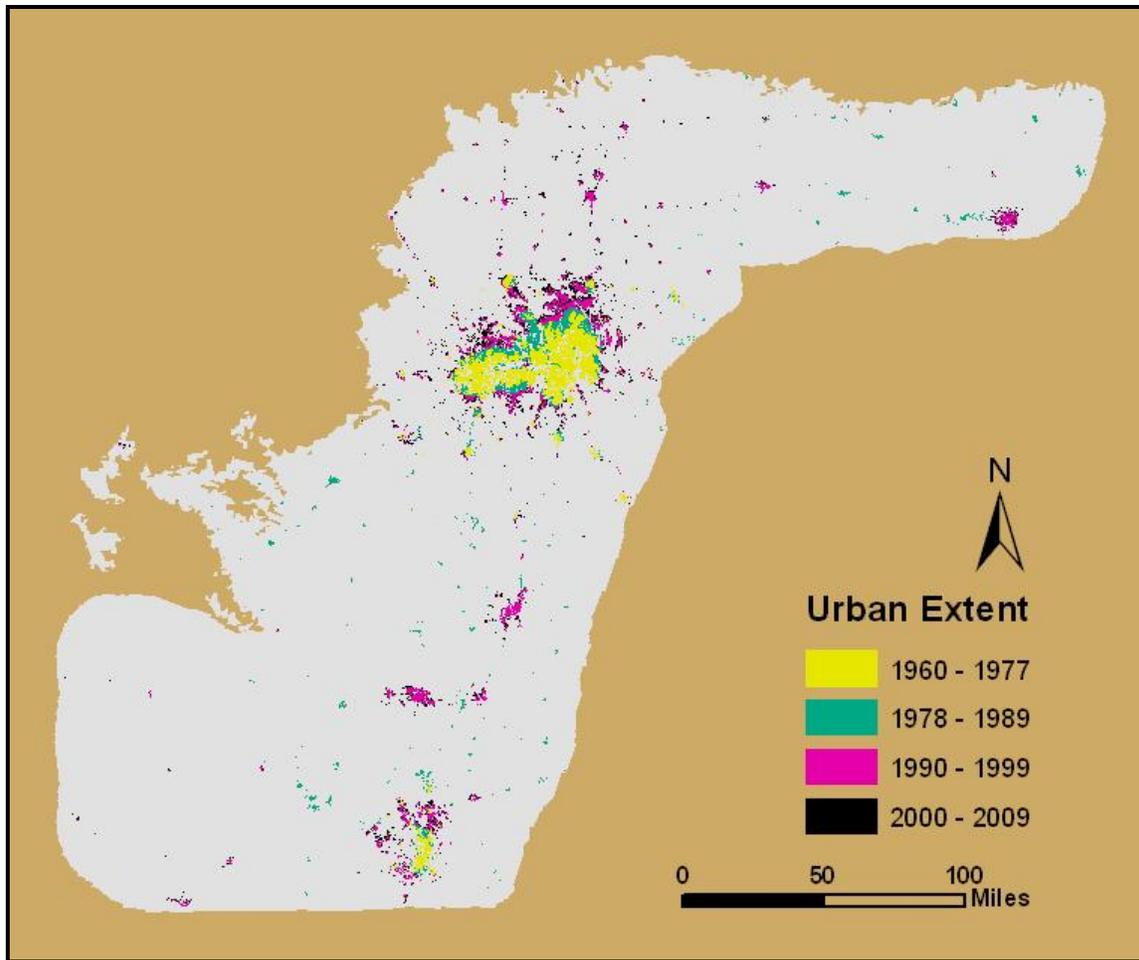
**Figure G1. Population growth in the major urban areas within the project area. Killeen / Fort Hood data was only collected for recent years.**

Given the low spatial and temporal resolution of the model (square-mile grid cells and annual time steps), a decade-scale was chosen to model urbanization. Three mapped urban extents and one interpolated boundary were used to capture the changes in urbanization over the duration of the model (Table G1).

**Table G1. Four time-steps were used to represent urbanization in the GWr model. Each contour represents the extent that urbanization has reached by the end of the stated time-step.**

<b>Time-Step</b>	<b>Source of Urban Extent</b>
1960 – 1977	Land Use / Land Cover Maps of Texas (TDWR)
1978 – 1989	Interpolated Contour Using ArcMap Spatial Analyst
1990 – 1999	GAP Analysis Data based on 1999 – 2001 Landsat Imagery
2000 – 2009	2007 MODIS Imagery

The urbanization represented in the GAP data and MODIS imagery was previously constrained for the model; shapefiles were generated from existing data. The urbanization represented in the Land Use / Land Cover Maps of Texas was defined by hand. Digitized copies of the maps containing Austin, Dallas, and Fort Worth were loaded into ArcMap and polygons were traced around the urban areas (TWDR, 1977). After combining these boundaries into a raster, a contour between the 1977 extent and the 1990 extent was interpolated using the Contour tool in the ArcMap Spatial Analyst, Surface Analysis toolbox. The urban extent at each of the four time-steps is shown in Figure G2.



**Figure G2. Urbanization over the model time-span necessitated dynamic spatial representation of urban expansion. Four decadal time-steps were chosen based on available data. These mapped and interpolated extents will be used to define when and where to apply urban corrections to GWr calculation.**

Combining the multiple data sources into a dynamic representation of urbanization in the model required three steps:

- 1.) All urban boundaries were defined by polygons in ArcGIS shapefiles;
- 2.) The shapefiles were combined and converted to an input ArcGIS GRID to develop a contour model; and,
- 3) The urban extent for the gap between documented urban boundaries was interpolated.

Final urban delineated zones for each time step were used to spatially identify where urban corrections were needed and for what model years.

## G2. NDVI\* Ratio

The urban NDVI\* ratio provided a decimal fraction representing the magnitude of urbanization's effect on GWr conditions. This NDVI\* ratio was used to enhance Q over areas where the Perm-Q relationship was obscured by urbanization and to increase ETa in model areas before urbanization reached the extent that depressed the 2007 NDVI\* (pre-urban areas).

The extent of the urban depression of NDVI\* values was examined in relation to three large urban areas located within the GMA8 project area; Dallas, Waco-Temple, and Austin, north to south respectively. The NDVI\* value for a MODIS satellite 241m pixel was reduced in areas where urban structures and infrastructure replace vegetation. The degree of difference between the 'urbanized' NDVI\* values and the average NDVI\* values within the surrounding undeveloped countryside were used as a scalar for the degree of urban development within each grid cell. Ratios ranged from 1 for areas of parks, golf courses, and wooded parcels that experienced no reduction on NDVI\* to close to zero representing extreme urban impact such as airports and downtown development. Table G2 shows some of the average NDVI\* differences found between the pixels within the city limits of Dallas, Waco, and Austin and the surrounding countryside. Urban Dallas had the lowest NDVI\* values, whereas Austin had a greener footprint. All urban classes had lower values than the surrounding countryside.

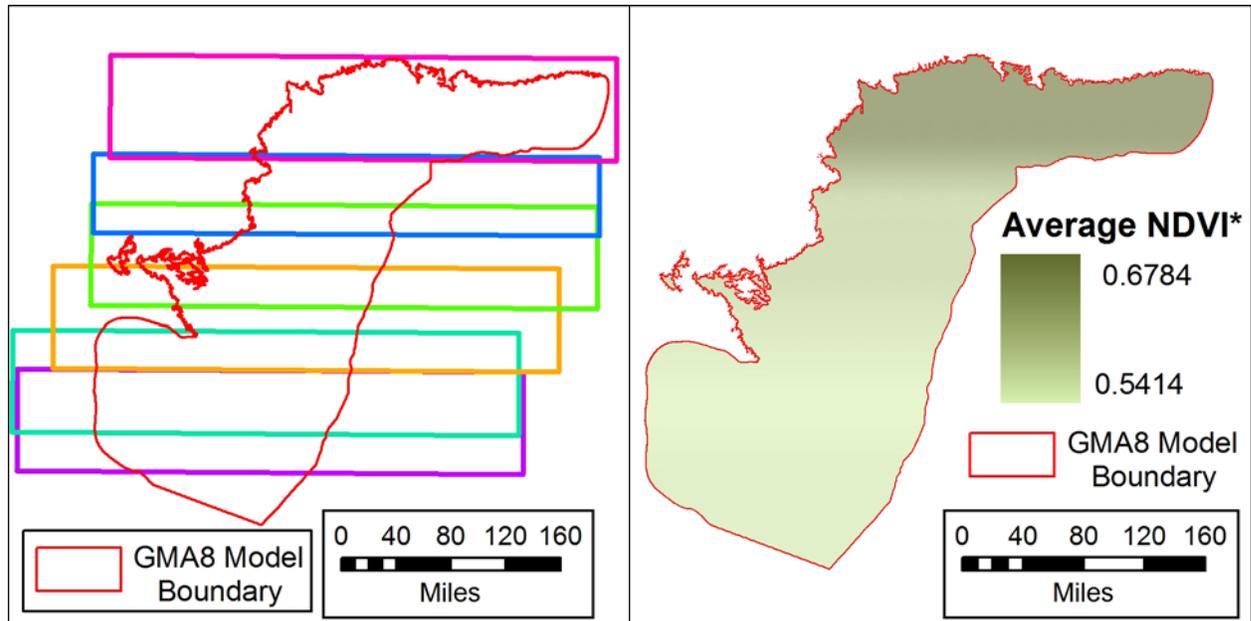
**Table G2. Urban NDVI\* averages compared to the surrounding countryside.**

City	NDVI* AVERAGES		
	Urban	Countryside	Ratio
Dallas	0.3903	0.6560	0.59
Waco-Temple	0.4046	0.5221	0.77
Austin	0.4635	0.5451	0.85
<b>3 Cities Average</b>	<b>0.4194</b>	<b>0.5744</b>	<b>0.73</b>

To model the NDVI\* consistently at a pixel level given that the non-urban vegetation showed systematic spatial variation, a method for determining average NDVI\* across the project area was devised. Having observed that average NDVI\* increases with increases in latitude, a zoned approach was taken. NDVI\* values were collected within linear zones as shown in Figure G3. Only NDVI\* from areas designated as non-urban by GAP and outside of the three urban districts were included to sample the background of climatically-determined regional NDVI\*. The average within each zone was converted to a grid of points for use in a kriging calculation. Where zones overlapped, the average of the two zones was used. This grid of points was then converted to a continuous grid layer through the use of kriging (Figure G4).

NDVI\* values from each cell with greater than 35 percent urban cover were indexed against the modeled average NDVI\* kriged layer. The difference found in each analysis cell from a summer image that was non-water limited (8/5/2007) was an indication of

development that has reduced a potential NDVI\* signal. This simplified approach allowed development of a rough urban impact index that was used to model the amount of impervious surface found in a pixel. This metric for the degree of urban impact was named the NDVI\* ratio.



**Figures G3 and G4. (G3) (left) Latitude-sensitive NDVI\* zones used to capture systematic spatial variation of NDVI\* outside of urban areas. (G4) Average NDVI\* layer used to represent the spatial average of NDVI\* in undisturbed areas to compare the depressed values observed in urban pixels.**

### **G3. Dynamic Urban Q**

The discharge model for GMA8 was built using precipitation (Ppt), reference ET ( $ET_0$ ), and Perm as its main inputs. This was unsuitable for urban areas with structures and infrastructure creating large areas of impervious surfaces that were not attributed in the soil data. Such urban areas generate additional Q and should, therefore, lower  $GW_r$ . Model  $GW_r$  predictions, however, were high due to the under-prediction of both Q and  $ET_a$ . To correct for this inaccuracy, factors were devised for application to urban areas to more correctly represent the expected increased runoff from impermeable surfaces. It was assumed that the reduction in NDVI\* was spatially coincident with areas of lowered permeability that rejects more infiltration and therefore leads to greater Q.

The NDVI\* ratio was determined on a cell-by-cell basis for grid cells that contained 35 percent or more urban area in the 2007 Landsat image. These data were stored in a raster to represent the degree of urban impact at the maximum urban extent for the model. The maximum extent corresponded to the 2000 – 2009 extent in Figure G2. For these years,

the Q enhancement of Equation G1 was applied to estimate Q over every cell using the NDVI\* ratio. To model Q for earlier years, the NDVI\* ratio was masked to each of the earlier urban extents. For example, in 1980, the maximum urban extent NDVI\* raster is masked to the 1978 – 1989 urban boundary and Q is only enhanced within this boundary. Example G1 provides a sample calculation using Equation G1. Table G3 lists the appropriate NDVI\* rasters to adjust Q estimates within the urban boundary for each model time-step.

**Example G1** If the NDVI\* ratio between urban and countryside NDVI\* was 0.4, the correction should calculate that 1 – 0.4, or 0.6 of the pixel was represented by structures that rendered the ground impermeable. Thus, this correction provided for infiltration through only 40 percent of the pixel while 60 percent contributed to increased runoff. The corrective equation for Q in this example is:

$$\text{Urban Pixel Q} = ( \text{Q calculated using Equation C10} * 0.4 ) + \text{Ppt} * 0.6$$

Equation G1 demonstrates the general form of this equation. Note that Urban NDVI Ratio<sub>k</sub> describes the NDVI\* raster from time-step “k” (the appropriate rasters for each model time-step are listed in Table G3).

$$\text{Urban Pixel Q} = ( \text{Q calculated using Equation C10} * \text{Urban NDVI Ratio}_k ) + \text{Ppt} * ( 1 - \text{Urban NDVI Star Ratio}_k ) \quad \text{Equation G1}$$

**Table G3. Urban NDVI\* ratio rasters used in model Q calculations.**

Time-Step “k”	NDVI* Ratio Raster (Urban NDVI Ratio <sub>k</sub> )
1960 – 1977	Q_Urban_NDVI_Ratio_1.tif
1978 – 1989	Q_Urban_NDVI_Ratio_2.tif
1990 – 1999	Q_Urban_NDVI_Ratio_3.tif
2000 – 2009	Q_Urban_NDVI_Ratio_4.tif

#### G4. Dynamic Urban ETa

GWr is calculated as a residual from precipitation less Q, less ETa. Hence, ETa values are important inputs for estimation of the correct GWr. ETa was spatially calibrated to precipitation (Ppt) for the model using recent satellite imagery (2005 and 2007), while the spatially explicit Ppt-ETa relationships thus determined were used in the model to predict ETa as a function of Ppt. Estimates using this relationship are referred to as 2005/2007 calibrated ETa. Within urban areas in recent satellite imagery, the non-NDVI\* generating cover depressed the expected NDVI\* signature. Depressed NDVI\* from urbanization, therefore, biases ETa calculations in these areas towards lower values that then generate high levels of GWr. A pixel that is urbanized in recent years may not, however, have been urbanized at an earlier date. ETa for these pixels for the model years before urbanization, referred to as pre-urban, will be too low due to the depressed NDVI\* bias in the model (Figure G5). This would result in an over-estimation of GWr.

To correct for the bias from urbanized ETa calibration for early pre-urban conditions, the NDVI\* ratio (urban cell/regional average) was used to boost ETa to match the average NDVI\* in the surrounding countryside. For example, for a pre-urban pixel with a NDVI\* ratio of 0.4, the ETa was boosted by an additive of  $1 - 0.4$ , or 60 percent of the original estimated ETa (Example G2, Equation G2).

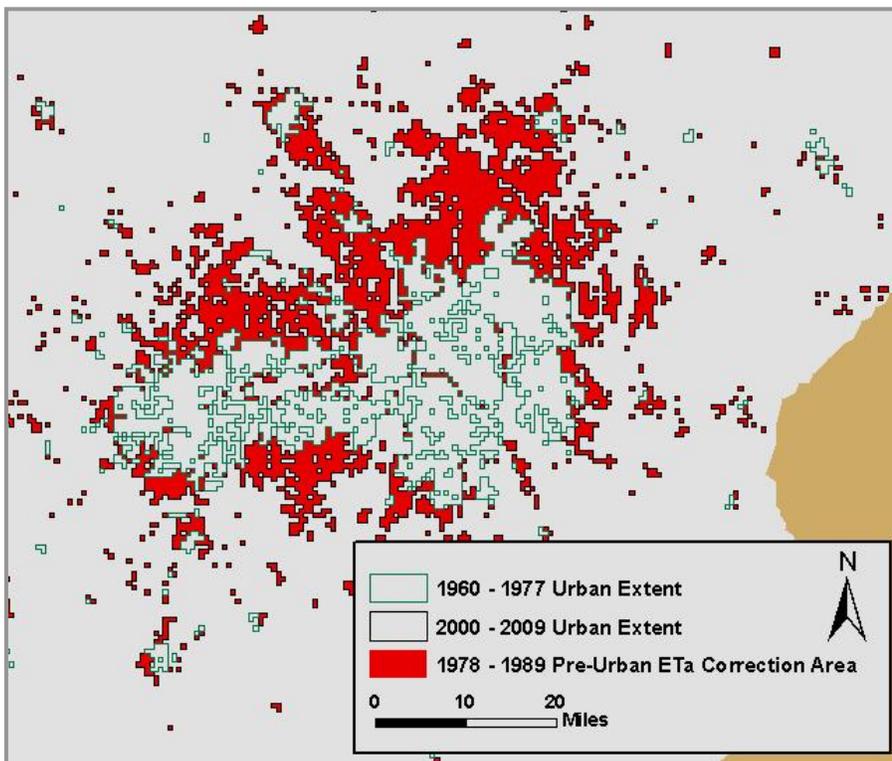
$$\text{Pre-Urban Pixel ETa} = (\text{ETa calculated using Equation F2}) + 0.6 * (\text{ETa calculated using Equation F2}) \quad \text{Example G2}$$

$$\text{Pre-Urban Pixel ETa} = (\text{ETa calculated using Equation F2}) + (1 - \text{Pre-Urban NDVI Ratio}_k) * (\text{ETa calculated using Equation F2}) \quad \text{Equation G2}$$

Note that Pre-Urban NDVI Ratio<sub>k</sub> describes the NDVI\* raster from time-step “k” (the appropriate rasters for each model time-step are listed in Table G4).

**Table G4. Pre-Urban NDVI\* ratio rasters used in model ETa calculations.**

Time-Step “k”	NDVI* Ratio Raster (Pre-Urban NDVI Ratio <sub>k</sub> )
1960 – 1977	ETa_Pre_Urban_NDVI_Ratio_1.tif
1978 – 1989	ETa_Pre_Urban_NDVI_Ratio_2.tif
1990 – 1999	ETa_Pre_Urban_NDVI_Ratio_3.tif
2000 – 2009	No Adjustment



**Figure G5. Correcting for depressed NDVI\* over pre-urban areas. Model ETa was calibrated the years 2005 and 2007—in these years NDVI\* was depressed over urban areas (2000 – 2009 Urban Extent) therefore lowering ETa estimates in these areas for the entire time span of the model. To correct ETa for the years before urbanization had reached the modern extent, a correction was applied over pre-urban areas. For example, for model years 1960 – 1977 the correction was applied to the area between the 1960 – 1977 extent and the modern urban extent (shown in red).**

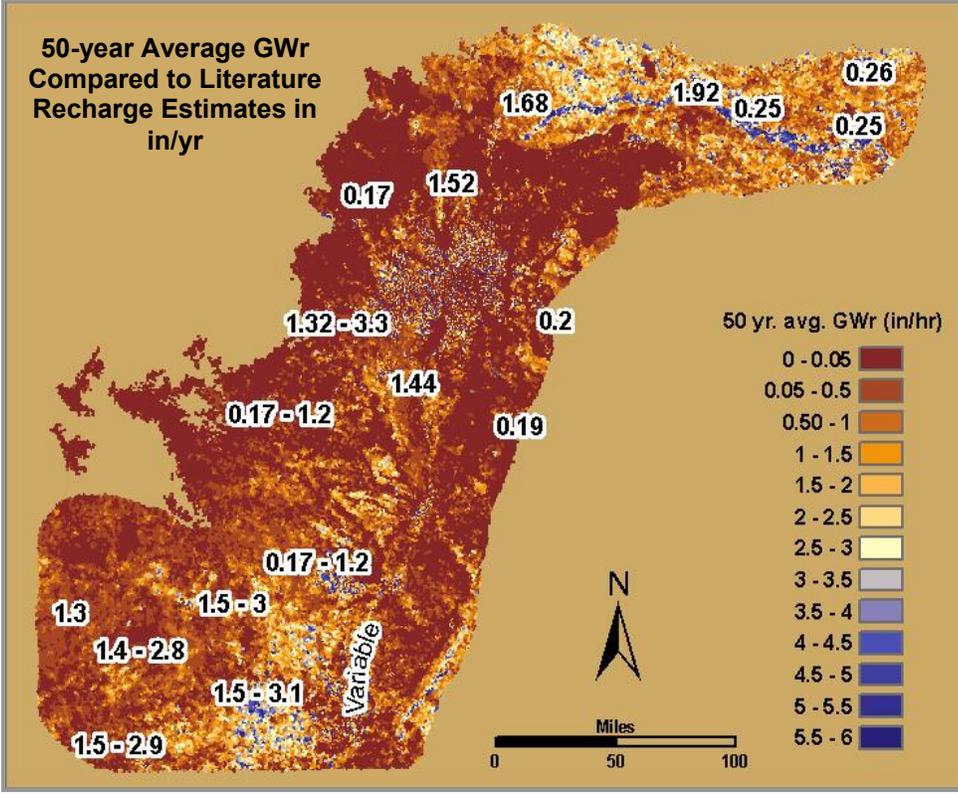
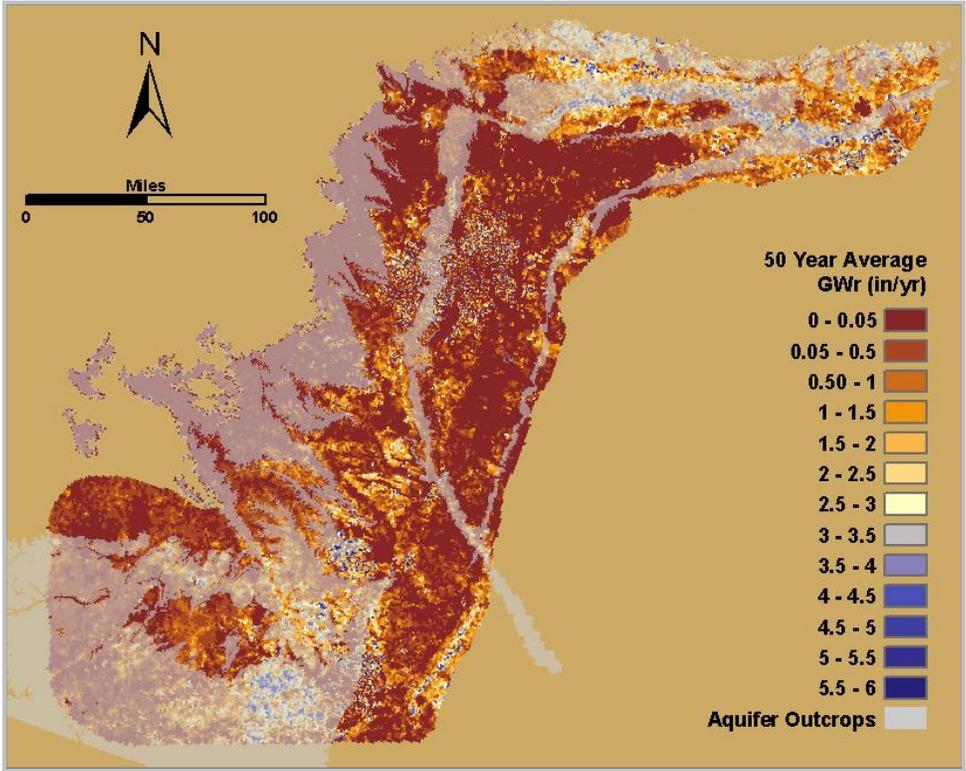
## G5. Summary

The groundwater recharge (GWr) model spans 1960 – 2009 and the extent of urbanization within the model area has changed significantly over this period. To represent this expansion in the model, urban boundaries were developed for four different time steps. The expansion of urban areas affects two GWr variables: (1) surface water loss measured as discharge (Q) is increased with the introduction of impervious surfaces such as pavement and buildings, and (2) water loss through evapotranspiration (ETa) is decreased when vegetation is replaced with urban hardscape. A metric developed using the ratio of the average NDVI\* from non-urban regional environments to the NDVI\* measured over an urbanized pixel was used to represent the degree of urbanization as a decimal fraction. This fraction was used to boost Q and decrease ETa based on the degree of urbanization present in a pixel at a given year.

### **Data provided from the work in this section include:**

- Dynamic urban boundary shapefiles in the model variables database and as individual files.
- NDVI\*-based Q adjustment rasters for each of the four urbanization time steps in the model parameters geodatabase and as individual rasters.
- NDVI\*-based ETa adjustment rasters for each of the four urbanization time steps in the model parameters geodatabase and as individual rasters.

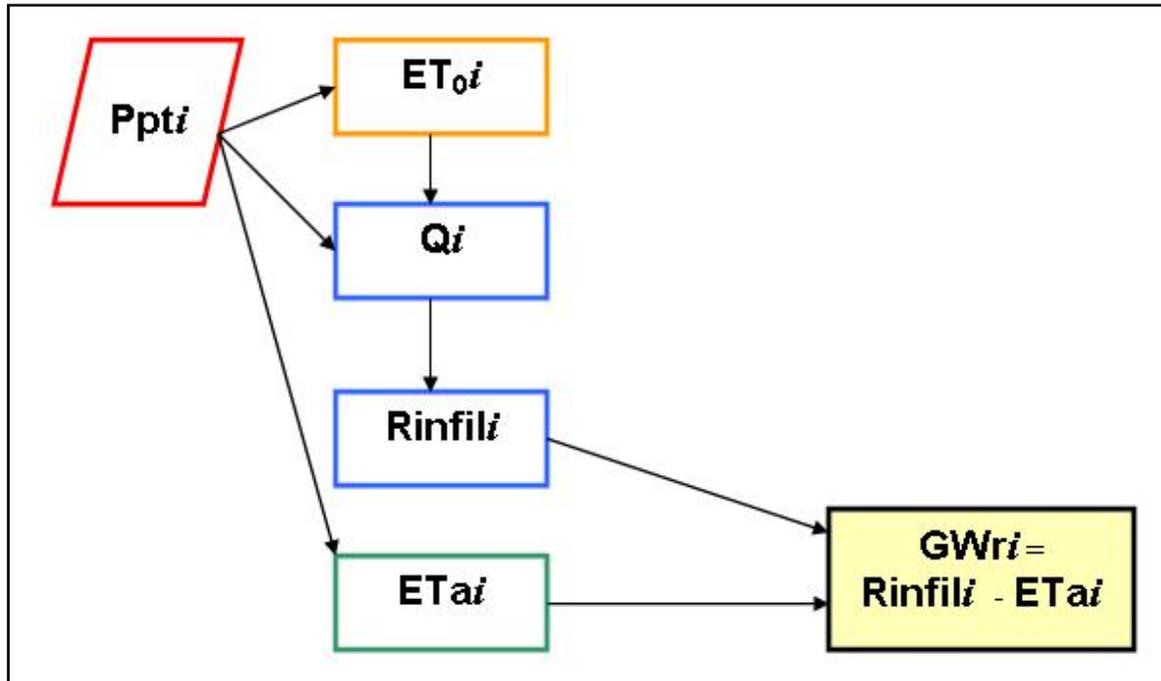
# H. INTEGRATING ALL COMPONENTS INTO RELATIONAL GWR MODEL





## H1. Mating all Model Databases

Sections A through G describe data collection, preparation and analysis used to develop the model relationships. Groundwater recharge (GWr), the residual of infiltrated rainfall ( $R_{infil}$ ) minus annual ET ( $ET_a$ ) was defined as a function of availability of precipitation ( $Ppt$ ) that is the driving variable (Figure H1).



**Figure H1. Flow of GWr estimation and structure of model relationships.**

The simple relational structure shown in Figure H1 was developed spatially using raster (image) data that allowed for calculations of variables stored by location as pixels. Geostatistical interpolation of variables measured at points was combined with regressed model relationships for pixels. For output of GWr, each pixel across the region represents 50 years of interpolated and modeled variables as well as representations of the spatially- and temporally-controlled relationships between model variables to enable relational model estimates within each raster cell. Table H1 summarizes the raster variable data, the math relationships that were used, and the flow of relational model calculations that culminated in the estimates of annual GWr for each cell (Equation H1).

$$GWr_i = R_{infil}_i - ET_{ai}$$

**Equation H1**

In this section GWr results are evaluated and final adjustments to GWr estimates are discussed. Final GWr results are also discussed in relation to previously published groundwater recharge models within the GMA8 model area.

**Table H1. Model variables, model regression equations and rasters associated with each variable, and the databases mated to model each variable. All calculations occurred at the pixel level in rasters. \*Dynamic urbanization rasters were used to apply minor adjustments to modeled Q and ETa in urban areas—these rasters are discussed in section G. Q, ETa, and GWr were each capped at Ppt.**

Major Model Components			Databases Used To Calibrate Model Components
Variable	Model Equation(s)	Raster Datasets	Calibration Database(s) / Tools
<b>Annual Precipitation (Ppt)</b>	N/A	50 annual kriged precipitation rasters.	Digital Climate Atlas: 50 years of point precipitation measurements (TWDB, 2009).
<b>Annual Reference Evapotranspiration (ET<sub>0</sub>)</b>	$ET_0 = a + (b * Ppt)$	(A) Spatially interpolated (kriged) rasters for intercept and slope of ET <sub>0</sub> -Ppt regression to estimate ET <sub>0</sub> from Ppt for each model year. (B) 50 annual ET <sub>0</sub> rasters estimated from annual Ppt.	Mesowest (University of Utah, Department of Atmospheric Sciences, 2010), Mesonet (Oklahoma Climatological Survey, 2008), TexasET (Texas A&M AgriLife Extension Service, 2010), ref-ET (University of Idaho, College of Agriculture and Life Sciences, 2010).
<b>Annual Discharge (Q)</b>	$Q = a_Q + [b_Q * (Ppt/ET_0)];$ $b_Q = (b_{climate} * Ppt/ET_0) + (b_{perm} * Perm) + Constant_b.$ $x\text{-intercept}_{Qi} = b_{x-int} * \ln(Ppt/ET_{0i}) + Constant_x.$ $a_Q = -b_{Qi} * x\text{-intercept}_{Qi}.$	50 years of Q rasters estimated from a multiple regression of Ppt/ET <sub>0</sub> and Perm.—each annual calculation inputs annual Ppt, ET <sub>0</sub> , and Perm rasters.	USGS gage data (2010), STATSGO (Penn State University ESSC, 2006) and NCRS WSS soil data (2011).
<b>Annual Infiltrated Rainfall (R<sub>infil</sub>)</b>	$R_{infil} = Ppt - Q$	50 annual Rinfil rasters estimated from annual Ppt and Q.	N/A
<b>Annual Evapotranspiration (ETa)</b>	$ETa = a + (b * Ppt)$	(A) Spatially extrapolated rasters for intercept and slope of ET <sub>0</sub> -ETa regression allow estimation of ETa from Ppt. (B) 50 annual ETa rasters estimated from annual Ppt.*	Ameriflux Network data (Heilman, J., 2005 - 2006; Katul and Oren, 2002 - 2005; Oechel, 2005 - 2006; Verma, 2003 - 2006), MODIS TM Imagery (NASA LP DAAC, 2005, 2007), USGS Gap Analysis Program (2010).
<b>Annual Ground Water Recharge (GWr)</b>	$GWr = R_{infil} - ETa$	(A) 50 annual GWr Plateau Model rasters and (B) 50 annual GWr Blended model rasters, both estimated from annual Ppt, R <sub>infil</sub> , and ETa.	N/A

## H2. High Ppt / ET<sub>0</sub> GWr Adjustments

As Ppt/ET<sub>0</sub> increases, the GWr model results indicated incipient GWr occurs after a Ppt/ET<sub>0</sub> threshold has been reached, then increasing as a diminishing returns curve in a non-linear fashion for the majority of Ppt/ET<sub>0</sub> conditions. At high Ppt/ET<sub>0</sub>, however, GWr predictions for many of the pixels across the GMA8 region showed a parabolic downswing. Q, ETa, and GWr logically increase with increasing Ppt/ET<sub>0</sub> for the majority of model Ppt conditions and so this behavior indicates that the GWr model was not yet correctly calibrated (Figure H2).

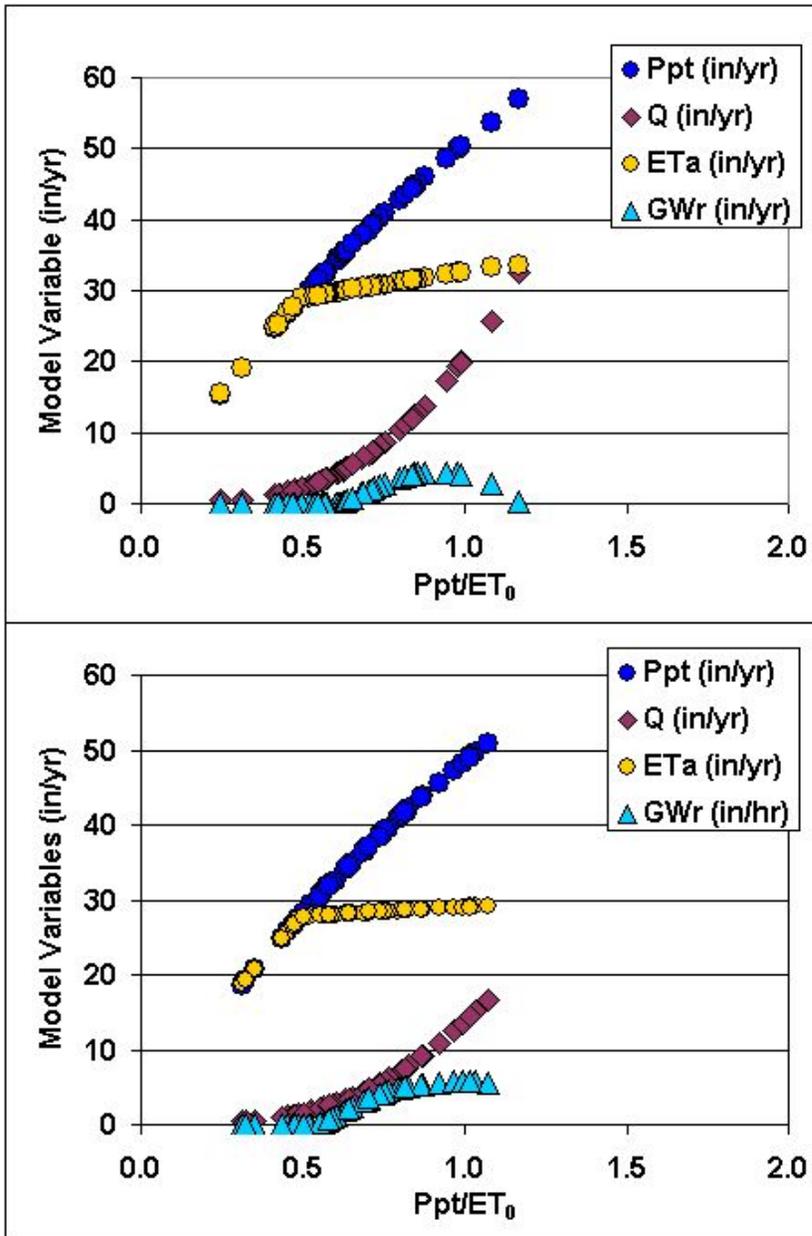


Figure H2. Q, ETa, and GWr results from nearly completed GWr model.

Upper graph:  
For this example 0.25 x 0.25 mi. pixel the GWr estimates increased to a peak and then diminished as ETa and Q estimates consumed the increasing Ppt/ET<sub>0</sub>.

Lower graph:  
The majority of the pixels across GMA8 had GWr that did not contain the downward curve of GWr at high Ppt/ET<sub>0</sub>.

The dogleg of the ETa was due to the truncation described in Section F6.

The hypothesis for the decrease in GWr at high Ppt/ET<sub>0</sub> is that the calibration for ETa is incomplete and should decrease at high values of Ppt/ET<sub>0</sub>.

The behavior of the GWR downswing at high values of  $Ppt/ET_0$  indicated that the model was not yet correctly calibrated. This lack of calibration was hypothesized to occur because  $ETA$  was not calibrated across all conditions, particularly under annual conditions that were extremely wet. The dogleg for the  $ETA$  was a logical precondition that was established in Section F6 is an example of the difficulty for calibration of  $ETA$ , however, it can be seen from Figure H2 that this lowermost  $ETA$  correction had little or no effect upon the GWR because it was below the threshold wetting conditions (as  $Ppt/ET_0$ ) to begin GWR accrual.

Because the apparently wettest conditions were used for calibration, extremely rainy years were not used for calibration data sets to establish  $ETA$  in some watersheds. This led to the underestimation of GWR for years that were highly wet. The divergence of GWR from the direct relationship with  $Ppt/ET_0$  is influenced by over prediction of  $ETA$  at high  $Ppt/ET_0$ . Hypothetically,  $ETA$  rates should decrease under the conditions that occur with high  $Ppt$  because  $ET_0$  decreases with the consequent increases in humidity and cloud cover during the wettest years. Because of limited  $ET_0$  data during recent years of coincident MODIS data, rather than recalibrating  $ETA$  for the entire range of conditions, we chose to adjust GWR directly using trends observed in model output. Two solutions were developed to remedy this issue; each uses the GWR estimations predicted by the original relational model to extrapolate GWR beyond a threshold  $Ppt/ET_0$  value as described in the next section.

## H2.2 Theoretical Behavior of GWR under High $Ppt/ET_0$

$Ppt/ET_0$  represents the relative opportunity for GWR to occur since this parameter represents the balance between the water supplied to the soil and near-surface vadose zone. As  $Ppt/ET_0$  increases, atmospheric humidity increases, temperatures are cooler, and cloud cover more frequent. These are all functions that reduce  $ET_0$  that is the driving force governing  $ETA$ .  $NDVI^*$ , the functional switch that scales  $ET_0$  for  $ETA$  disposition, was found to have relatively low inter-annual variability (Section F) in comparison to other parameters in the calculation of GWR, for example  $Q$  (Section C). Reduction in  $ET_0$  is a mechanism expected to decrease  $ETA$  at higher values of  $Ppt/ET_0$ .

As annual  $ET_0$  and  $ETA$  decline with increasing  $Ppt/ET_0$ , more water must be ported to runoff because the opportunities for the ground to accept recharge are less, i.e., as the degree of wetting increases, the water from precipitation will increasingly be disposed by runoff rather than soil uptake because the proportion of time that soils are at or near saturation increases. This concept is the same as Hortonian overland flow common to surface water hydrology (Bevin, 2004). The rapidly rising limb of  $Q$  with  $Ppt/ET_0$  illustrated in Figure H2, therefore, is conceptually correct. The parabolic downward response is incorrect and is simply an artifact arising from the lack of appropriate calibration for  $ETA$ .

## H2.2 High Ppt/ET<sub>0</sub> Adjustments

With the realization that full calibration of ET<sub>a</sub> through the entire relationship for ET<sub>a</sub> versus Ppt/ET<sub>0</sub> was beyond the reach of this investigation, another approach was adopted for constraining GWR to follow expected behavior. The adopted approach was based upon the robust assumption that the upper limb of GWR that shows parabolic decrease can be modeled conceptually. The GWR curves (Figure H2):

- show a minimal value of Ppt/ET<sub>0</sub> when GWR starts to occur that is both a reasonable and expected finding;
- may have a parabolic lower limb where GWR follows an expected shape of diminishing returns as would be expected given that, as wetter weather occurs, it enhances Q because wetted and saturated soils limit infiltration;
- may contain parabolas that peak and then diminish that are an artifact of incomplete calibration for ET<sub>a</sub>;
- are amenable to a method to constrain the upper limb of the GWR curve, therefore, compensating for incomplete calibration of ET<sub>a</sub>; and
- because a curve shape for diminishing returns is expected, this curve shape can be used as the conceptual curve to correct downward parabolic curve behavior.

A GWR model adjustment was applied only for high Ppt/ET<sub>0</sub> portions of the curves. The trigger point for when to apply GWR model corrections was defined when higher values of Ppt/ET<sub>0</sub> yielded values of GWR that were lower than the value at peak GWR. Exceedances of the parabolic threshold (peak GWR on the parabola) were identified within the data set by a set of appropriate queries.

Two approaches were used to constrain the GWR to prevent the decreasing values at higher values of Ppt/ET<sub>0</sub> (Figure H3). The first approach, called the Plateau Model, was to hold GWR constant once the maximum GWR was reached (Figure H3, shown in yellow). This adjustment can be regarded as extremely conservative as it uses the expected diminishing returns curve shape.

A second, more complex correction, also applied to high-Ppt/ET<sub>0</sub>, adjusted GWR using a linear regression through the minimum and maximum recharge (blue line) and incorporated the results from the plateau model. This second model, called the Blended Model, used regression of minimum-maximum GWR values that were “blended” with the Plateau model. This yielded a weighted average of the linear and plateau model that tethered the upper limb of recharge to the highest GWR previously predicted by the model while allowing moderate increase based on the Min/Max GWR Regression (Figure H3, GWR Blended Model shown in blue). The following sub-section describes how these adjustments were implemented in the model.

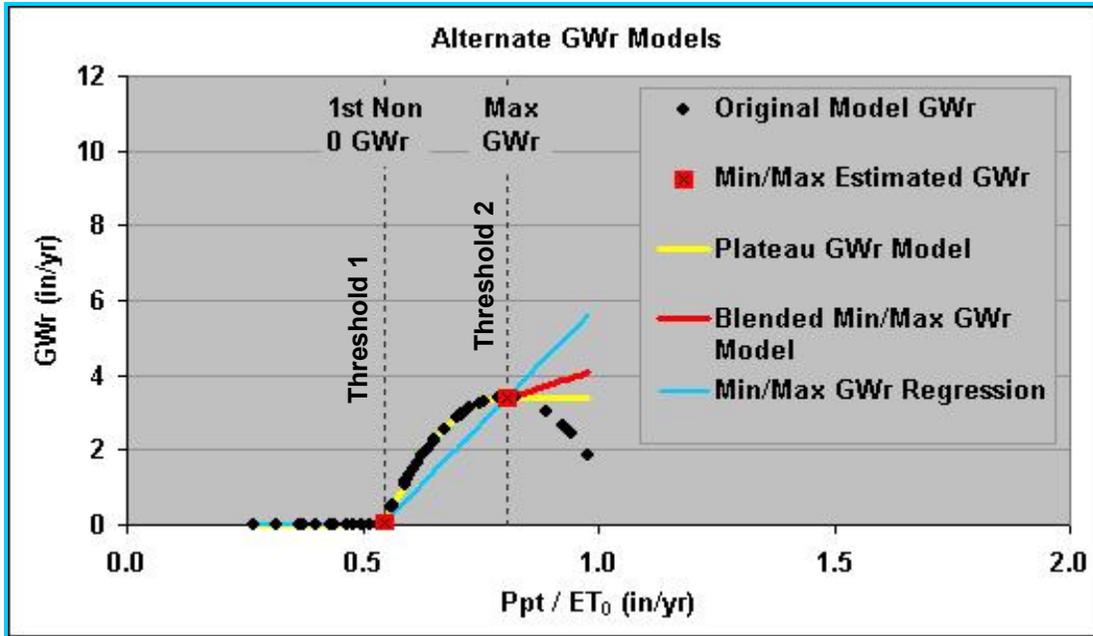


Figure H3. Exceedance of the  $Ppt/ET_0$  parabolic threshold caused the original model to incorrectly produce decreasing GWR estimates (black points). To correct this, two models were developed to predict GWR for high  $Ppt/ET_0$  using the Min non-zero value of GWR and the Max GWR (red boxes) paired with their associated  $Ppt/ET_0$  values. The two models were applied for years with  $Ppt/ET_0$  that exceeded the  $Ppt/ET_0$  corresponding to Max GWR.

## H.2.2 Extrapolating GWR Predictions for Very Years with High $Ppt/ET$

Relying on the accuracy of the relationships within the model structure, 50 years of GWR estimates were used to develop the data for extrapolating GWR as shown in Figure H3. For this operation we extracted four values from the 50-year GWR model to summarize the GWR relationship to  $Ppt/ET_0$  for each pixel in the model (Figure H3):

1. **Min GWR:** minimum non-zero GWR from 1960 – 2009.
2. **Threshold 1:**  $Ppt/ET_0$  corresponding to the minimum non-zero GWR (this point is the threshold in the  $Ppt/ET_0$  range at which recharge begins—below Threshold 1 no recharge occurs).
3. **Max GWR:** maximum estimated GWR from 1960 – 2009.
4. **Threshold 2:** The  $Ppt/ET_0$  corresponding to the maximum estimated GWR. Above this point GWR estimates deviated from expected.

These four values enabled extrapolation of GWR for very high  $Ppt/ET_0$  on a pixel-by-pixel basis. The minimum and maximum GWR and corresponding  $Ppt/ET_0$  defined a range of values characteristic to each pixel's unique recharge conditions. The majority of

the Ppt/ET<sub>0</sub> years, those below Ppt/ET<sub>0</sub> Threshold 2, required no adjustments for either the Plateau or the Blended GWr model because that portion of the GWr curve followed expected behavior (GWr increasing throughout the distribution for the leftward leg of the parabolic curve). The two models for correcting the parabolic exceedance problem were applied only when Ppt/ET<sub>0</sub> exceeded Threshold 2.

The Plateau Model was the more conservative of the two corrections. Above Threshold 2, GWr was held constant at Max GWr (Equation H3). This adjustment prevents GWr estimates from decreasing at high Ppt, but did not allow GWr to increase above the Max GWr predicted in the unmodified 50-year model (Figure H3, yellow line).

$$\begin{aligned} &\text{If } Ppt_i / ET_{0i} \leq \text{Threshold 2,} && \text{Equation H2} \\ &GWr_{\text{Plateau}i} = GWr_i \text{ calculated using Equation H1} \end{aligned}$$

$$\begin{aligned} &\text{If } Ppt_i / ET_{0i} > \text{Threshold 2,} \\ &GWr_{\text{Plateau}i} = \text{Max GWr} \end{aligned}$$

In the second, Blended Model, GWr, was developed to allow a moderate increase in GWr above Threshold 2 that followed more closely the expected shape for a diminishing returns curve. The blended GWr model combined the results of the Plateau Model with the Blended Model. The slope and intercept of the GWr<sub>Min/Max</sub> regression line for each pixel were defined for each model pixel in the form of Equation H4 (Figure H3, red line).

$$GWr_{\text{Min/Max}i} = a_{GWr} + ( b_{GWr} * Ppt_i / ET_{0i} ) \quad \text{Equation H4}$$

For the Blended Model, a weighted average between the predictions of the Plateau Model and the results of the Min/Max regression were used to calculate GWr for Ppt/ET<sub>0</sub> above Threshold 2. By using a weighted average, we allowed GWr estimates to increase with Ppt/ET<sub>0</sub> based on the Min/Max regression while moderating that growth with the Plateau model. Blended GWr was estimated using Equation H5 which inhibits GWr from rapidly diverging from the previously predicted maximum GWr.

$$\begin{aligned} &\text{If } Ppt_i / ET_{0i} \leq \text{Threshold 2,} && \text{Equation H5} \\ &GWr_{\text{Blended}i} = GWr_i \text{ calculated using Equation H1} \end{aligned}$$

$$\begin{aligned} &\text{If } Ppt_i / ET_{0i} > \text{Threshold 2,} \\ &GWr_{\text{Blended}i} = ( 0.7 * GWr_{\text{Plateau}i} ) + ( 0.3 * GWr_{\text{Min/Max}i} ) \end{aligned}$$

Both GWr Models, Plateau and Blended, were capped at annual precipitation and all occurrences of negative values were replaced with zeros. GWr predictions were not made over reservoirs, final GWr estimates were set to zero over reservoirs (Equations H6, H7, and H8).

$$\begin{aligned} &\text{If } GWr_{\text{Plateau or Blended}i} > Ppt_i, GWr_{\text{Plateau or Blended}i} = Ppt_i && \text{Equation H6} \\ &\text{If } GWr_{\text{Plateau or Blended}i} \leq Ppt_i, GWr_{\text{Plateau or Blended}i} \end{aligned}$$

$$\begin{aligned} &\text{If } GWr_{\text{Plateau or Blended}i} < 0, GWr_{\text{Plateau or Blended}i} = 0 && \text{Equation H7} \\ &\text{If } Qi \geq 0, GWr_{\text{Plateau or Blended}i} \end{aligned}$$

### H3. Final GWr Estimates

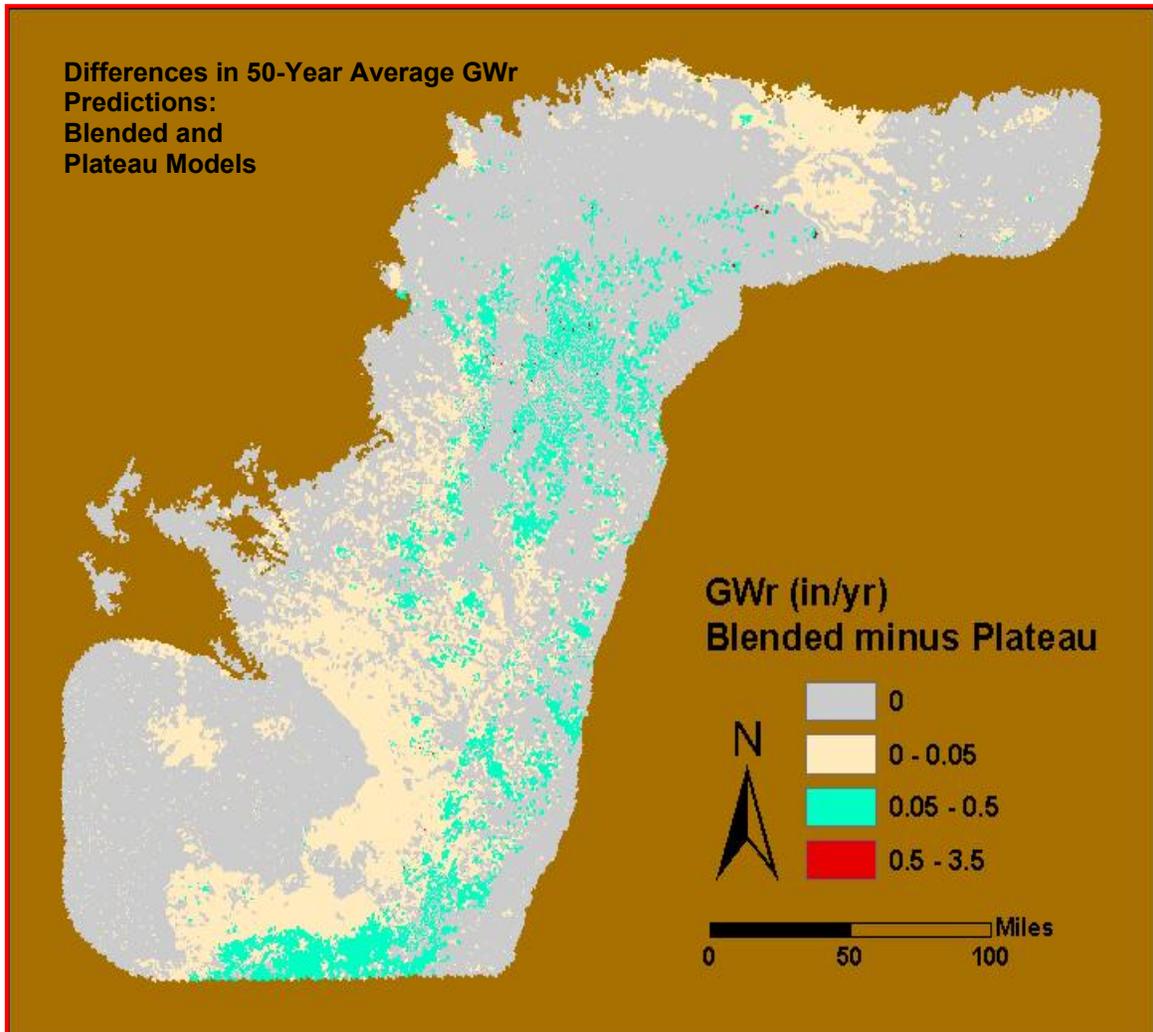
#### H3.1 GWr Plateau and Blended Model Results

To determine how widespread the effect of the high-Ppt/ET<sub>0</sub> GWr adjustments on GWr predictions was, we examined Ppt/ET<sub>0</sub> for a relatively dry year (2005), a relatively wet year (2007), for the 50-year average, and for the 50-year maximum. Table 1 presents the occurrence of pixel values exceeding Threshold 2 as a percentage of total model pixels. Pixels that exceed Threshold 2 are adjusted using either Plateau or Blended GWr adjustments. The portion of model cells adjusted is low for a dry year but increases notably for a very wet year. Under average annual Ppt/ET<sub>0</sub>, around 70 percent of model pixels did not exceed Threshold 2 and, therefore, remained unmodified.

**Table H2. Occurrence of Ppt/ET<sub>0</sub> above Threshold 2 presented as percentage of total model pixels. When Ppt/ET<sub>0</sub> for a given pixel exceeded that threshold, both the GWr Plateau and GWr blended model adjusted GWr predictions. No model adjustments were made at or below Threshold 2.**

	> Threshold 2 (model adjustment applied)	≤Threshold 2 (model adjustment <i>not</i> applied)	Notes
<b>2005 Ppt/ET<sub>0</sub> (dry year)</b>	1.93%	98.07%	During 2005, a dry year, only 1.9% of pixels were adjusted with Plateau and Blended GWr models.
<b>2007 Ppt/ET<sub>0</sub> (wet year)</b>	60.52%	39.48%	During 2007, an extremely wet year, 60.5% of pixels were adjusted in Plateau and Blended GWr models.
<b>50-year Average Ppt/ET<sub>0</sub></b>	29.63%	70.37%	29.6% of pixels exceed Threshold 2 for the average year and were adjusted by the Plateau and Blended models.
<b>50-year Max Ppt/ET<sub>0</sub></b>	67.54%	32.46%	32.5% of pixels never exceed Threshold 2 and remained unaffected by Plateau and Blended GWr adjustments.

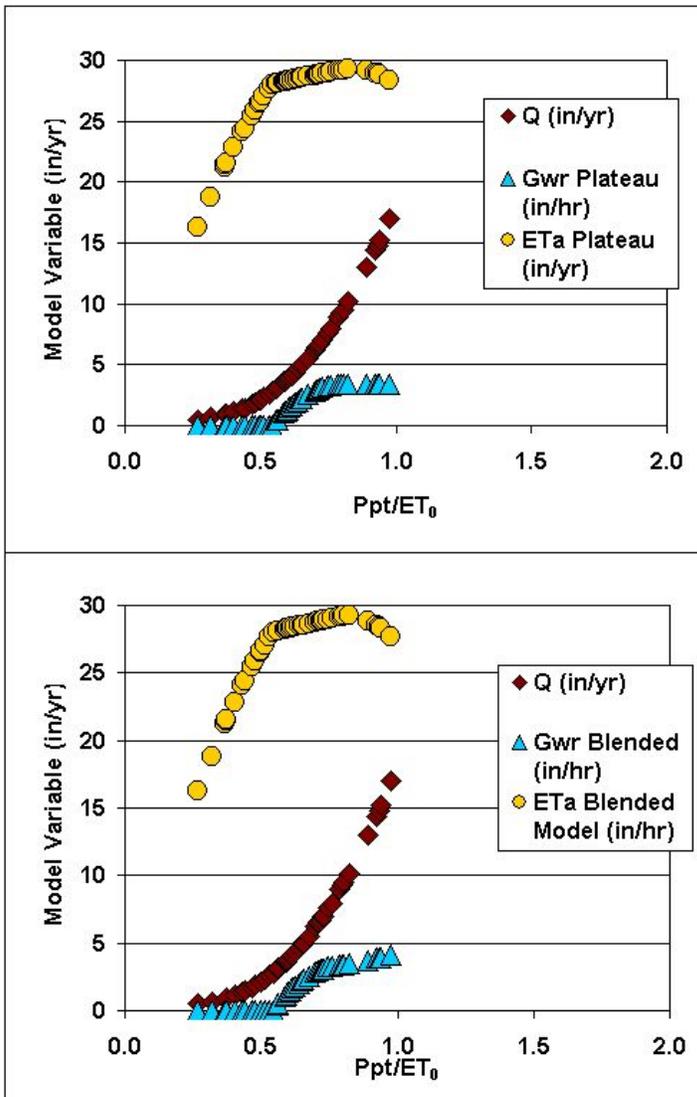
The differences between the Plateau and the Blended model were subtle and small. We evaluated the differences spatially by comparing the 50-year average GWr results of each model. While the difference between the models is small, the Blended GWr model predicts up to 0.5 inch per year over the Plateau model predictions in the areas shown in Figure H4.



**Figure H4.** Comparison of the 50-year average predictions of the Plateau and Blended GWR models. The majority of the difference between the two models was less than 0.5 in. A short study of the pixels with a difference greater than 0.5 in, shown in red, 15 were less than 1 in, and 8 pixels fell within the range of 1 in to 3.5 in (these 8 all fell within urban boundaries).

### H3.2. ETa Calculated as a Residual from Plateau and Blended Models

One of the benefits for constraining GWR using the two models chosen was the ability that this confers for understanding the shape of the ETa as a residual. Figure H5 presents the results for a single pixel that was constrained by Plateau and Blended Models. These outputs were used to calculate ETa as a residual to provide an advanced look at what ETa should look like were significant additional calibration accomplished over the full range of Ppt/ET<sub>0</sub>. Little difference was visually detectable in the ETa residual between the Plateau and Blended Model results.



**Figure H5. ETa estimates calculated as residuals after application of the Plateau and Balanced Model adjustments of GWr from**  
 $ETa = Ppt - Q - GWr$

**Upper Graph:**  
**Application of Plateau Model**

**Lower Graph:**  
**Application of Blended Model.**

### H3.3. Comparison of GWr Results and Previously Published Recharge Estimates

GWr results were compared to published GWr estimates. Figure H6 shows overall trends in the estimated recharge as controlled by the suite of spatially calibrated data that were used to construct the model—precipitation, climate variables, soil permeability, and land cover type. The incorporation of these individually calibrated variables in the GWr model produced reasonable estimates of GWr based on relationally accurate estimates of all GWr variables—Ppt,  $R_{infil}$ , and ETa.

We compared outputs from the Harden (2004) and HydroBio models on a per county basis. The Harden (2004) model was the basis for comparison because this was the most recent and extensive groundwater model available. Although calculated differently, Harden’s Recharge minus EVT (MODFLOW evapotranspiration) should estimate a

variable roughly equivalent to GWR—the resulting net recharge is the water that enters the aquifer. For the comparison, average annual recharge estimates over Trinity and Woodbine aquifer outcrops within each county were summed as total acre feet per year. Recharge data for the Harden county-by-county totals were provided by TWDB model runs for 1980 – 1999 (2011). Magnitudes of Harden recharge, Recharge minus EVT, and HydroBio’s GWR are shown in Figure H7.

In order to enable the most direct comparison, HydroBio GWR data in Figure H7 were generated for the same calibration period as the Harden Model (1980 - 1999). The relationships between the Harden and HydroBio models differed from one location to another. We collected the percentage of each land cover class over aquifer outcrop areas for each county to evaluate the differences between HydroBio and Harden results. Finally, each outcrop area was also evaluated in Google Earth to detect visible surface conditions that would affect GWR.

For counties where HydroBio’s GWR estimates were notably different from the modeled Harden Recharge minus EVT estimates, model variables and spatial controls were examined in order to determine which variables were responsible for driving HydroBio GWR results differing from the Harden results. Where notable differences existed between the two models, documented GWR-affecting variables were found to explain the differences.

When evaluations were concluded, four major factors arose as likely explanations for model disagreements (see Figure H6).

1. Outcrop areas in counties with large amounts of exposed limestone and low vegetation cover, including quarries, drill pads, and general land cover disturbances were associated with enhanced GWR (for example, Burnet, Coryell, Lampasas, Travis, and Tarrant). In these cases, HydroBio GWR was greater than Harden Recharge minus EVT. The higher values of GWR were responsive to special conditions of low vegetation cover that can foster GWR.
2. Outcrop areas with a large proportion of crops were associated with GWR estimates lower than Harden estimates (for example, see Comanche and Eastland in Figures). Crop water use and enhanced ETa from cultivated crops explain the lower GWR estimates.
3. Urbanized model areas were associated with GWR estimates higher than Harden estimates (for example, see Denton, Dallas, Tarrant, and Travis). The combination of land disturbance (exposed bedrock) and depressed ET over low-vegetated urban areas enhanced GWR in the HydroBio model.
4. GWR estimates over outcrops along the Red River were higher than Harden estimates (for example, Fannin, Lamar, and Red River). These areas showed enhanced Q and GWR due to high Ppt and rapid permeability through the alluvial soils in the HydroBio model.

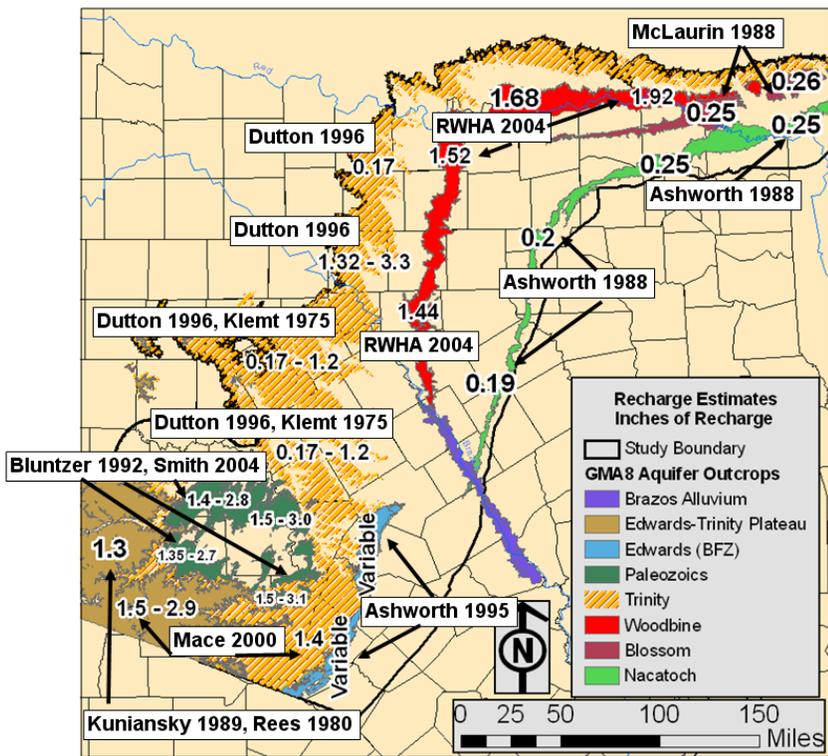
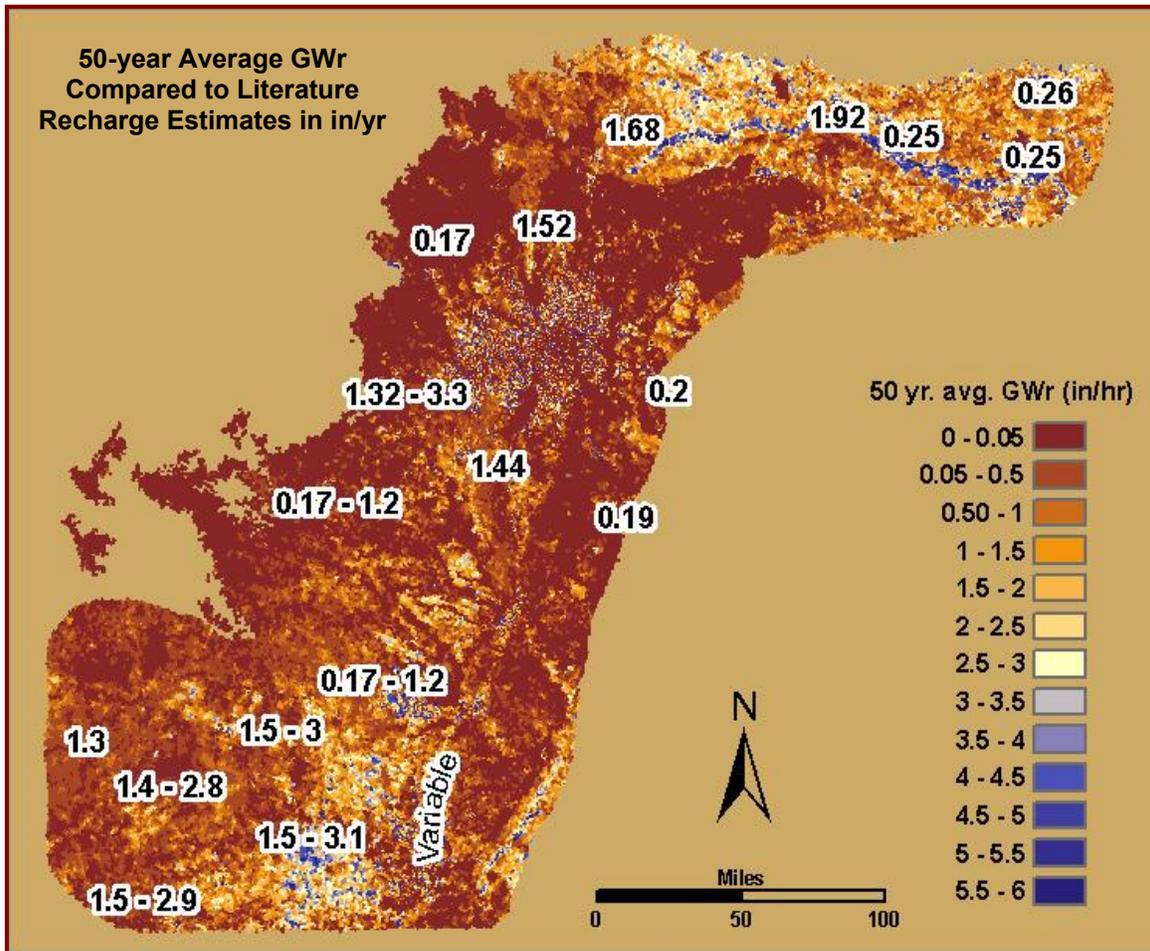
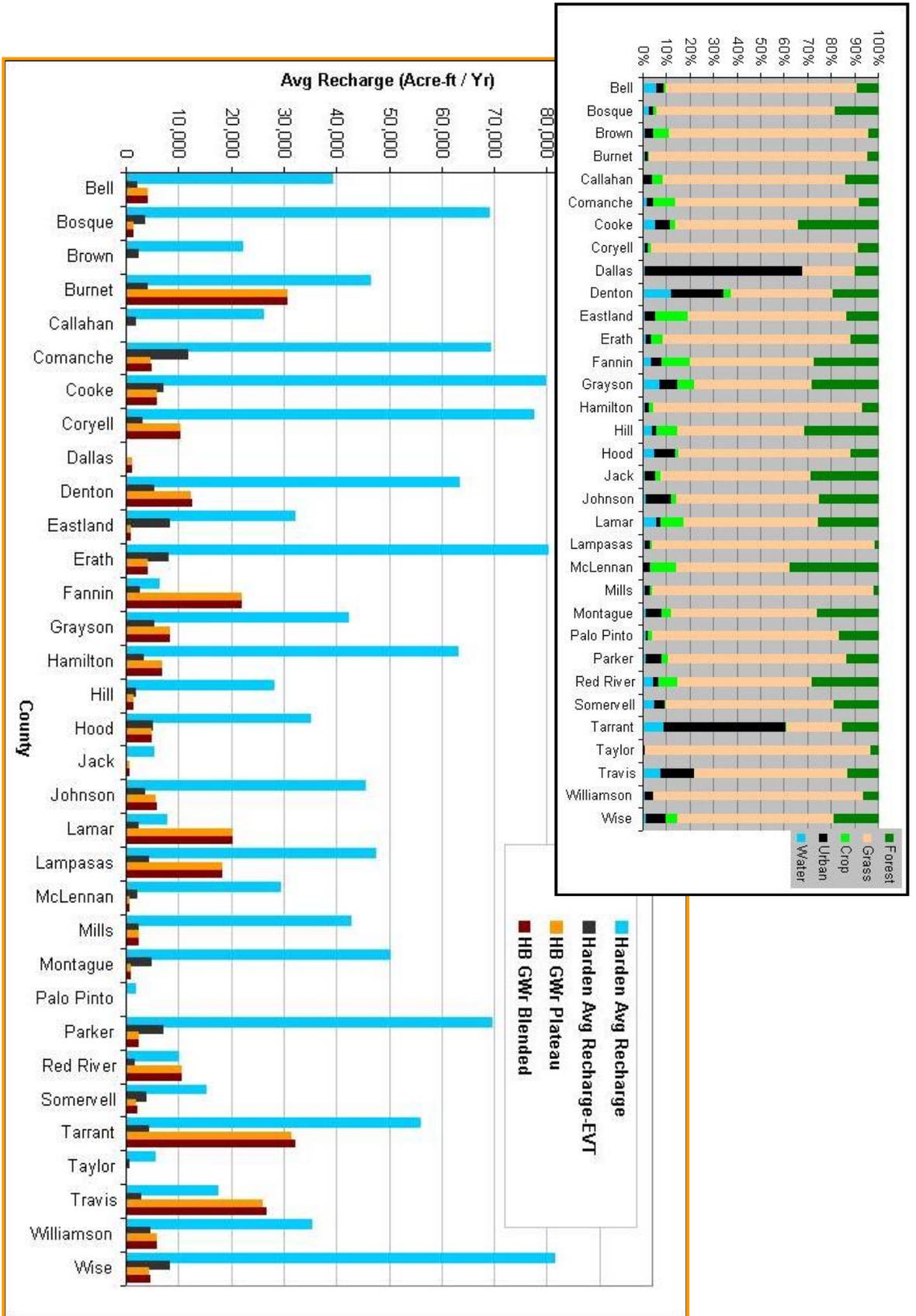


Figure H5. Top: Plateau GWR results fall within the range of estimates compiled in a literature review, Section C). Values on the map represent the approximate average or range of recharge estimates reported in the literature. (Blended GWR model not shown because differences between Plateau and Blended models are not detectable over the 50 year average at map scale.) Left: Citations for previously published groundwater recharge estimates.

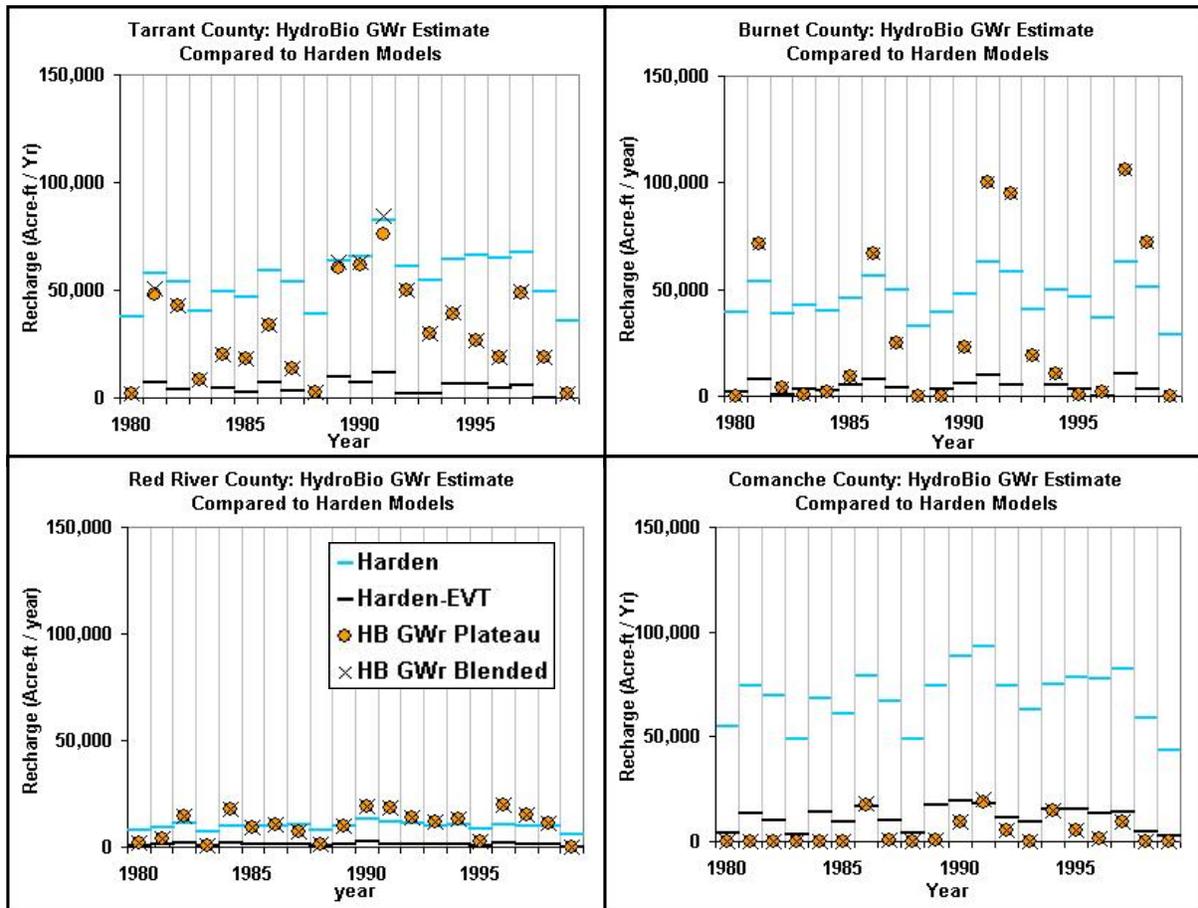
Figure H7. Top: Distribution of land cover classes by county. Bottom: Average recharge for the years 1980 – 1999 within the outcrops of each county. HydroBio averages are well below the Harden recharge estimates, but fluctuate above and below the Harden average estimates of recharge minus EVT. The differences between HydroBio's GWR Plateau and GWR Blended model are subtle.



### H3.4. Comparison of GWR Behavior in HydroBio and Harden Models

Because the HydroBio approach to estimation of GWR is radically different from the existing groundwater modeling in Texas, further comparisons were made to investigate the differences. Figure H8 provides a look at model behavior in a range of dry and wet years. The striking difference between the approaches is the relatively low inter-annual variability of the recharge predictions in Harden in comparison to the results from the HydroBio results that vary greatly depending upon the available precipitation. This section is not provided to call attention to either models shortcomings but simply to demonstrate the different results from the two efforts.

**Figure H8. HydroBio GWR estimates compared to Harden model outputs illustrate the effects of urbanization, exposed bedrock, river alluvium, and crop water upon GWR. Tarrant County (upper left) GWR estimates were notably higher than Harden Recharge minus EVT—this may be due to land cover disturbances and depressed ETa in an urban environment (alternatively this may also indicate the need for better calibration in urban areas). GWR in Burnet County (upper right) exceeds Harden Recharge minus EVT—the difference is attributable to exposed limestone bedrock accepting infiltration. Red River County (lower left) demonstrates disagreement between models associated with highly permeable river alluvium accepting infiltrated rainfall in the HydroBio model. In Comanche County (lower right) GWR is slightly lower than Harden Recharge minus EVT: crop water use is associated with the difference. The differences between the Plateau and Blended GWR models are indistinguishable at the annual recharge over outcrop by county scale in Burnet, Red River, and Comanche Counties. (Negative Harden Recharge minus EVT values not shown).**



## H5. Summary

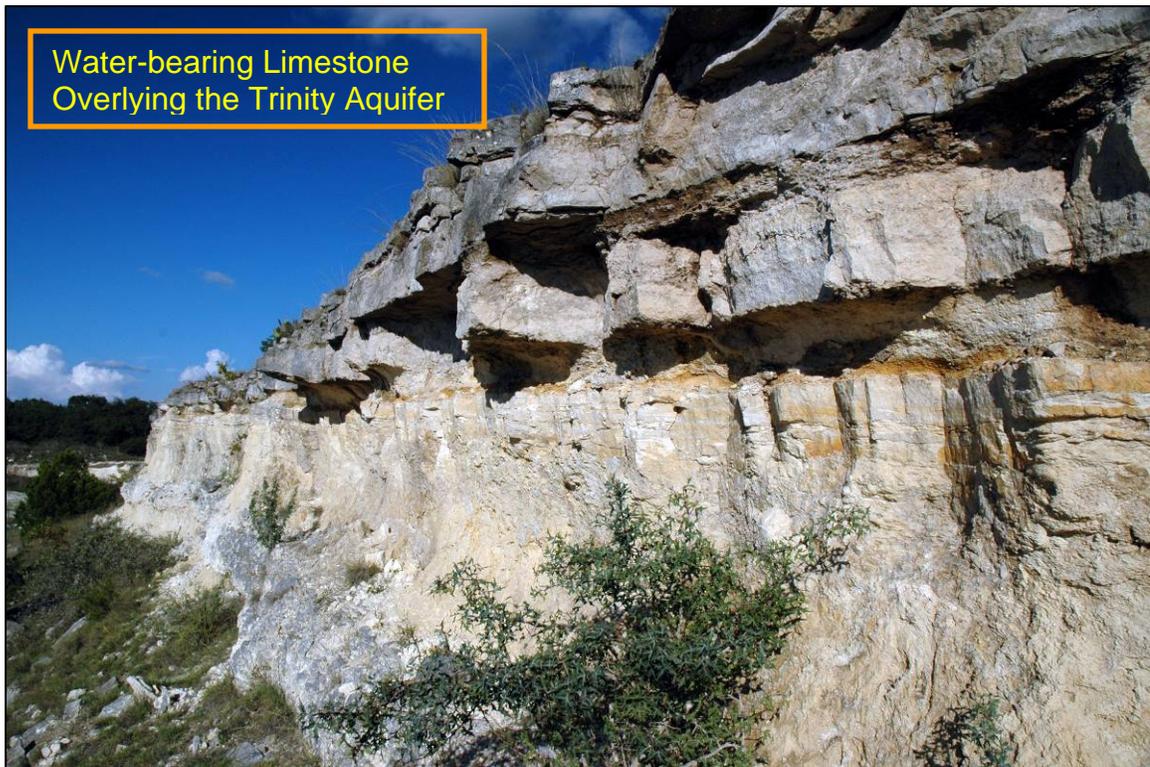
The model relationships between GWR driving variables were assembled to solve for GWR across the GMA8 project area. Model instability at very high precipitation, due to incomplete calibration of ETa over the entire possible range of wet to dry conditions necessitated the development of two alternate models for predicting GWR under extreme Ppt/ET<sub>0</sub> conditions, Plateau and Blended. The results of the two alternate models demonstrated subtle spatial differences. When compared to previously published recharge estimates, GWR Plateau and Blended model estimates corresponded in magnitude but demonstrated minor differences in spatial distribution. Because HydroBio GWR estimates are spatially calibrated to independent variables, including Perm and land cover type, the model results are controlled by the actual environment. Spatial variation in GWR estimates correspond to the varied environmental conditions within the project—demonstrating the relational accuracy of the model.

### Data provided from the work in this section include:

- GWR Plateau estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- GWR Blended estimates for 1960 – 2009 provided as a geodatabase and as individual rasters.
- Threshold 2 raster used in GWR Plateau and Blended calculations.
- Max GWR raster used in GWR Blended calculations.
- Slope and intercept rasters for GWR<sub>MinMax</sub> linear regression.
- Reservoir location raster.



## I. LIMITATIONS AND RECOMMENDATIONS





## **I1. Limitations and Recommendations**

This section should be reviewed in order to best use the groundwater recharge (GWr) model results. Here we provide an overview of the limitations introduced by data restrictions and the broad generalizations made at the model scale as well as recommendations based on insights from the modeling process. Recommendations are made for modeling GWr in general, for future GMA8 modeling projects, and for using the results of this GWr model in future groundwater availability models of GMA8.

### **I1. Limitations of the GWr Model**

In order to interpret and use the results of the GWr model, we provide the following description of the data restrictions and limitations.

#### **I1.1 Data Restrictions**

The GWr model was assembled from the most accurate characterization of each of the parameters possible. This necessitated required interpolation over model areas and/or calibration years with incomplete data coverage.

Evapotranspiration (ET) data were temporally limited so techniques were developed to estimate ET for the early model years. Reference evapotranspiration ( $ET_0$ ) was a critical variable for model calibration because it moderates both annual evapotranspiration ( $ET_a$ ) and discharge ( $Q$ ). Climate data sufficient for  $ET_0$  estimation, however, were only available for recent years in the model record. To model  $ET_0$  for the 50-year model record (1960 – 2009) precipitation (Ppt) was used to estimate  $ET_0$  for the complete model record (described in Section A3). Measured climate data for calculating  $ET_0$  on an annual basis for the model would be more accurate than the Ppt approximation used for the model, but these data was not available.

Discharge data were spatially limited within the model due to poor availability of undisturbed watersheds in the USGS stream gage archive. Many watersheds either had incomplete discharge records or had disturbances such as urbanization, reservoirs and irrigation diversions that could obscure the relationship between  $Q$  and the underlying soil properties. After eliminating incomplete and disturbed watersheds, 16 watersheds remained to represent 49,000 square miles. This necessitated extrapolation of these records throughout the model domain to simulate the many complex environments that give rise to GWr (the watershed selection process is described in Section B2).

## **11.2. Broad Characterization:**

Generalizations of the variables and processes controlling GWR were necessary to construct the model. Complex variables were distilled into single “lumped” variables. Spatial averaging was used to parse data into square-mile grid cells with data reported as single annual values. Some variables such as groundwater pumping were excluded in order to characterize the broad extent of the model. These generalizations are noted below for model users to consider during the interpretation of results and the use of GWR results in further modeling.

The square-mile grid cells and annual time steps used for the GMA8 model allowed for GWR estimation through lumped variables. As examples, Q was chosen as a lumped variable to represent aquifer discharge, bank storage, and overland flow with one value and ETa was represented as combined canopy interception, transpiration, and soil evaporation. While the lumped variable approach worked for the native countryside, for the highly variable recharge conditions in urbanized regions, such coarse resolution was not appropriate. The spatial variability of recharge conditions in urbanized areas was lost when complex variables were lumped over a square-mile.

Changes in urban extent over the 50-year model were captured only by the rough 10-year time-steps used to model urbanization --the decade time scale being the highest resolution possible to capture urban changes detectable at the square-mile scale. Hence, the GWR estimates within urbanized regions were generalized in a manner appropriate for assessing GWR over very large regions (e.g., many counties), but should not be used for higher resolution evaluation within urban regions (e.g., scale of a county or two).

This GMA8 GWR assessment did not consider pumping stresses on the aquifers across the region, so the parameters used in the model, without adjustment, cannot be represented dynamically in relation to changing heads induced by pumpage. For example, Q represents watershed discharge (water rejected by the aquifer) and is anticipated to vary with pumping stresses as the pumping makes space available in the aquifer for water intake, thereby inducing GWR and decreasing Q. The GMA8 GWR model depicts the relationships between model variables and the magnitude of  $P_{pt} / ET_0$ . These variables are appropriate for use as raw inputs within future modeling programs that consider the effects of pumping stresses. Such intensive, head-dependent evaluations are best made for smaller regions of GMA8, for example at the scale of a few counties rather than as an assessment over the full 49,000 square miles.

## **12. General Recommendations for GWR Modeling**

The process of developing the GWR model broke new ground in the calibration and evaluation of recharge. The results and methods that were used in modeling can be applied in future modeling studies to help formulate parameters and processes for

assessing GWr in Texas. Below, we present insights on the use of precipitation measurements, stream gage data, and the ETa estimates provided with this model.

Although Ppt is the driving variable for GWr, the use of Ppt / ET<sub>0</sub> for assessing hydrologic processes was a significant improvement over raw Ppt measurements. Scaling Ppt with ET<sub>0</sub> effectively linearized Q, runoff and base flow regressions so that each could then be spatially extrapolated (see Section C2.1). Instead of using raw Ppt as the driving variable for GWr, future modeling should consider the use of Ppt / ET<sub>0</sub> as a variable indicative of available incoming rainfall that has been corrected for water losses driven by ET<sub>0</sub> conditions. This is especially appropriate for the coarse annual time step.

Total discharge, Q, was useful for assessing the proportion of incoming rainfall available for ETa and GWr. The discharge measured at stream gages can be used directly or it can be processed with programs such as USGS's streamflow partitioning program PART to separate gaged Q into two values, both runoff at the surface and sub-surface base flow (USGS, 2010). When the partitioned results were evaluated for use as individual variables the model, runoff as a measure of surface water loss and base flow as an estimate of minimum groundwater recharge, the results were questionable (see Section C1). The partitioning algorithm introduced errors and created uncertainty in model estimates. In particular, this uncertainty came from the interpretation of PART estimated "base flow". Because the interpretation of base flow estimates could not be applied universally for the model, neither base flow nor runoff was used for the model. Simplifying assumptions and choosing Q to represent that portion of Ppt rejected from recharge was a helpful simplification recommended for similar modeling efforts.

A landscape-scale study of ETa was beyond the scope of this project, but model ETa calibration provided insight into important details that should be considered when modeling this variable in the future. First, a landscape-scale model of ETa requires that the full range of ETa be represented in the calibration data, requiring imagery at a higher resolution and over a long enough time span to capture all variability (this would require additional processing time and the purchase of necessary imagery). ETa calibration was based upon a record wet and a record dry year, however, this failed to capture all regional variability as true extremes were not universally represented across GMA8 within the two calibration years. Second, the relationship between ETa with Ppt changes with higher magnitude precipitation. As Ppt approaches high values the potential for ET, as ET<sub>0</sub>, diminishes as a result of climatic feedback (cooler, more humid, less solar radiation) and the corresponding rate of ET should diminish concurrently. The last adjustments to the GWr model (Plateau and Blended GWr) were used to simulate this change in ET behavior at high Ppt by adjusting the GWr distribution to an expected diminishing returns curve shape at higher Ppt / ET<sub>0</sub>. Inverting either the Plateau or Blended GWr model, and estimating ETa as the residual of GWr minus infiltrated rainfall (R<sub>infil</sub>) confers an ETa curve through the entire range of Ppt / ET<sub>0</sub> to better represent the dynamic variation of ETa for future models (see Figure H5 in Section H3.2). This shape for ETa at the landscape level is a new contribution to the understanding of dynamic ETa and should be used to precondition the use of ETa in future studies of GWr within Texas.

### **13. Recommendations for Future GWR Modeling in GMA8**

The following projects are suggested based on what we have learned during model calibration and evaluation. The results of the relational GWR model provide a sound basis for more detailed and expanded studies in the following areas:

1. Incorporate a memory function for soil water to enable antecedent climate conditions to carry water over from a wet to a dry year as soil water storage so that the model can better represent GWR across multiple model years. This will require calibration and development at higher resolution than the GMA8 GWR model. A carryover can be effective for evaluating the changes in GWR from dryer and hotter conditions across GMA8 to represent climatic changes.
2. Evaluate the effect of bark beetle-induced cedar die-off on GWR. Recharge is directly affected by vegetation cover and large areas of cedar die-off occurred during 2011 in the Trinity Aquifer upland due to drought. Reduced cedar cover may affect GWR because cedar tend to be the climax species covering rocky uplands with very thin soils—zones of expected recharge. GWR increases driven by the removal of vegetation that intercepts Ppt and redirects it to ETa may be transitory or permanent through the feedback mechanism of climatic trends.
3. Reassess GWR within urbanized areas by conducting high-resolution studies of ETa and Q for urban areas. Imagery with a higher resolution than MODIS should be used to capture the high spatial variability in ETa over urbanized region for this GMA8 model (Landsat TM, with 30-m pixels and 25+ year archive would be ideal). Q in urban areas should be calibrated using watersheds that have become urbanized in recent years so that before and after conditions recorded at these gages can be used to model changes in Q while factoring the effects of urbanization. We did not use urbanized watersheds in model calibration because diversion and other disturbances obscured the relationship between Q and the underlying soil permeability.

### **14. Recommendations for Using GWR Estimates in Future Models**

The GWR model is anchored to the landscape-position-determined controlling variables by a strong relational structure. Each of the intermediate variables was carefully calibrated to measured data across the landscape and evaluation indicates that the resulting predictions of each variable closely model the calibration data (see, for example, section C3.3). By maintaining consistent relationships between all of the model variables and the underlying landscape, we have provided a product that can be easily manipulated using simple scaling or the addition or subtraction of a universal constant for the full project area. To maintain the relational integrity of the model it is important to apply adjustments universally for the project or area of interest within the project area so that the relative proportions of each variable are preserved.

When incorporating GWR values in future modeling efforts, GWR can be adjusted up or down universally to match expected or previously reported values. Figure I1 demonstrates simple manipulations for adjusting GWR and how each adjustment can vary across the range of GWR values. An additive (or subtractive) adjustment, for example, will affect results in the same increment across the entire range of GWR predictions. Subtraction will cause some values of GWR to become negative, an impossible result so if this test adjustment is performed all negative values should be converted to zeroes.

A scalar adjustment can be used to make small changes for low GWR estimates and relatively larger changes for high estimates. GWR can be adjusted upward using a constant to match the anticipated level of recharge over the western plateau. When this adjustment is made universally the change in the predictions across the remainder of the model will maintain proportion. In another scenario, GWR can be adjusted downward to minimize or eliminate GWR predictions over non-aquifer outcrop areas by applying a fractional scalar or subtracting the minimum value necessary to eliminate GWR over these areas. Thus the flexibility of the relational GWR results allows both the end-user and future modeler to “tweak” results to fit the GWR range necessary to match heads within traditional groundwater modeling while maintaining relationally-accurate values across GMA8.

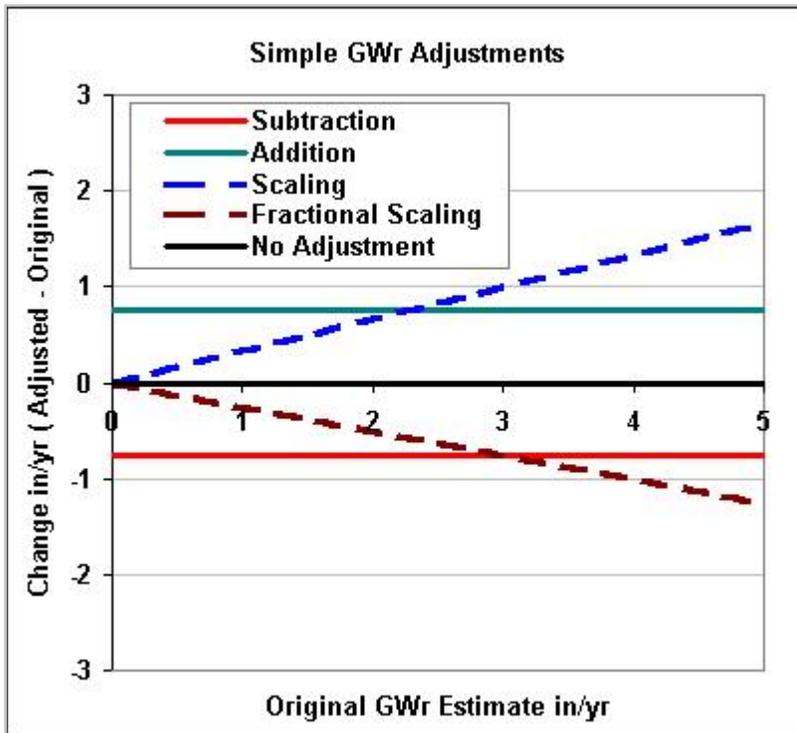


Figure I1. Changes resulting from simple GWR adjustments.



## J. APPENDIX 1: REFERENCES



High shrink-swell soils (laterites) of the Blackland Prairie caused wracking of the power lines in this photograph. Such tight, clay rich soils impede or prevent infiltration.



## J. Appendix 1: References

- Allen, R.B., Pereira, L.S., Raes, D., and Smith, M.S. 1998. Crop evapotranspiration (guidelines for computing crop water requirements). FAO Paper 56. 300 pp.
- Allen, R.G. Undated. REF-ET: Reference Evaporation Calculator for Support of ASCE Manual 70 (1990) and 2001 ASCE Standardizations and FAO Irrigation and Drainage Paper No. 56. <<http://www.kimberly.uidaho.edu/ref-et/>>.
- Anaya, R. 2004. Conceptual model for the Edwards-Trinity (Plateau) Aquifer System, Texas. Chapter 2: Aquifers of the Edwards Plateau. Texas Water Development Board. Report 360. 42p.
- Ashworth, J.B., Hopkins, J. 1995. Aquifers of Texas. Texas Water Development Board, Report 345. 69p
- Atmospheric Radiation Measurement (ARM) Program, August 2010. U.S. Department of Energy, <<http://www.arm.gov/>>.
- Baugh, W.M. and Groeneveld, D.P. 2006. Broadband vegetation index performance evaluated for a low-cover environment. *International Journal of Remote Sensing* 27:4715-4730.
- Beach, J.A., Ashworth, J.B., Davidson, T., Tang, C., Donnelly, A. 2006. Evaluation of Recharge for the Edwards-Trinity (Plateau) Aquifer. Prepared for Region F Water Planning Group by LBG-Guyton Associates.
- Bevin, K. 2004. Robert E. Horton's perceptual model of infiltration processes. *Hydrological Processes* 18:3447-3460.
- Bluntzer, R. L. 1992. Evaluation of the ground-water resources of the Paleozoic and Cretaceous aquifers in the Hill Country of Central Texas. Texas Water Development Board. Report 339. 161p.
- Cooper, J. D., Gardner, C.M.K., Mackenzie, N. 1990. Soil controls on recharge to aquifers. *Journal of Soil Science*, 41: 613–630.
- EPA. Accessed 2011. Ecoregions of North America Shapefile. <<http://www.epa.gov/wed/pages/ecoregions.htm>>.
- ESRI, 2005. ArcMap 9.1.
- Gibson, C. And K. Jung, U.S. Census Bureau, Population Division. February 2005. Working Paper No. 76 Historical census statistics on population totals by race,

1790 to 1990, and by Hispanic origin, 1970 to 1990, for large cities and other urban places in the United States.

Google Earth Accessed 2012.

Groeneveld, D.P. 2008. Remotely-sensed groundwater evapotranspiration from alkali scrub affected by declining water table. *Journal of Hydrology* 358: 294-303.

Groeneveld, D.P. and Baugh, W.M. 2007. Correcting satellite data to detect vegetation signal for eco-hydrologic analyses. *Journal of Hydrology* 344:135-145.

Groeneveld, D.P., Baugh, W.M., Sanderson, J.S., and Cooper, D.J. 2007. Annual groundwater evapotranspiration mapped from single satellite scenes. *Journal of Hydrology* 344:146-156.

Groeneveld, D.P., Watson, R.P. Barz, D.D., Silverman, J.B. and Baugh, W.M. in press. Assessment of two methods to monitor wetness to control dust emissions, Owens Lake, California. *International Journal of Remote Sensing*. In April or May issue, 2010.

Heilman, J. Principal Investigator. Undated. AmeriFlux Freeman Ranch Grassland US FR1 dataset. 2005 - 2006. Texas A&M University Department of Soil and Crop Sciences.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.freeman\\_grassland.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.freeman_grassland.01)>.

Heilman, J. Principal Investigator. Undated. AmeriFlux Freeman Ranch Woodland US FR3 dataset. 2005 - 2006. Texas A&M University Department of Soil and Crop Sciences.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.freeman\\_woodland.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.freeman_woodland.01)>.

Irish, R., Undated (last accessed April, 2009). Landsat 7 Science Data Users Handbook. Landsat Project Science Office, Goddard Space Flight Center.  
<<http://itpwww.gsfc.nasa.gov/IAS/handbook.html>>.

Katul, G., and Oren, R. Principal Investigators. Undated. AmeriFlux Duke Forest Hardwoods US-Dk2 dataset. 2002, 2003, 2005. Duke University School of the Environment. <[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.duke\\_hardwoods.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.duke_hardwoods.01)>.

Katul, G., and Oren, R. Principal Investigators. Undated. AmeriFlux Duke Forest Open Field US-Dk1 dataset. 2002, 2003, 2005. Duke University School of the Environment.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.duke\\_grassyfield.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.duke_grassyfield.01)>.



- Katul, G., and Oren, R. Principal Investigators. Undated. AmeriFlux Duke Forest Loblolly Pine US-Dk3 dataset. 2002, 2003, 2005. Duke University School of the Environment.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.duke\\_loblolly.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.duke_loblolly.01)>.
- Kuniansky, E. L., 1989, Precipitation, streamflow, and baseflow, in West-Central Texas, December 1974 through March 1977: U.S. Geological Survey Water-Resources Investigations Report 89-4208, 2 sheets.
- Mace, R.E., Anaya, R., 2004. Estimate of recharge to the Edwards (Balcones Fault Zone) and Edwards-Trinity (Plateau) aquifers in Kinney County, Texas
- Mace, R.E., Chowdhury, A.H., Anaya, R., and Way, S.C., 2000. Groundwater availability of the Trinity Aquifer, Hill Country Area, Texas: numerical simulations through 2050. Texas Water Development Board, Report 353, 119 p.
- McLaurin, C. 1988. Occurrence, availability, and chemical quality of ground water in the Blossom Sand Aquifer. Texas Water Development Board. Report 307. 22p.
- NASA Land Processes Distributed Active Archive Center (LP DAAC), 2005. MODIS MOD09Q1 8-day Composite Daily Reflectance. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota.
- NASA Land Processes Distributed Active Archive Center (LP DAAC), 2007. MODIS MOD09Q1 8-day Composite Daily Reflectance. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota.
- NRCS. 2011. Natural Resource Conservation Service. Web Soil Survey. US Department of Agriculture. <<http://websoilsurvey.nrcs.usda.gov>>. Last accessed August 2011.
- Oechel, W. Principal Investigator. Undated. AmeriFlux Sky Oaks New US-SO4 dataset. 2005 – 2006. San Diego State University Biology Department.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.sky\\_oaks.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.sky_oaks.01)>.
- Oechel, W. Principal Investigator. Undated. AmeriFlux Sky Oaks Old US-SO2 dataset. 2005 – 2006. San Diego State University Biology Department.  
<[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.sky\\_oaks\\_old.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.sky_oaks_old.01)>.
- Oechel, W. Principal Investigator. Undated. AmeriFlux Sky Oaks Young US-SO3 dataset. 2005 – 2006. San Diego State University Biology Department.

- [http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.sky\\_oaks\\_yn g.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.sky_oaks_yn g.01).
- Oklahoma Climatological Survey: Drought Monitoring Tools. Created 2008. Oklahoma Mesonet. [http://climate.mesonet.org/rainfall\\_update.html](http://climate.mesonet.org/rainfall_update.html).
- Oregon State University, PRISM Climate Group. Monthly Climate Data. Data accessed 2009. <http://www.prism.oregonstate.edu/>.
- Penn State University ESSC. 2006. STATSGO derived soil physical and hydraulic properties. [http://dbwww.essc.psu.edu/geotree/dbndx/tree/amer\\_n/us\\_48.html](http://dbwww.essc.psu.edu/geotree/dbndx/tree/amer_n/us_48.html) . Accessed November 2010.
- PRISM Climate Group. Created 2004. Oregon State University, <http://www.prismclimate.org/>.
- Rees, R., and Buckner, A. W., 1980. Occurrence and quality of ground water in the Edwards–Trinity (Plateau) aquifer in the Trans-Pecos Region of Texas: Texas Department of Water Resources Report 255, 41 p.
- RWHA. 2004. Northern Trinity / Woodbine aquifer groundwater availability model. Submitted to Texas Water Development Board. R.W. Harden & Associates, Inc
- RWHA. 2007. Northern Trinity/Woodbine GAM Assessment of Groundwater Use in the Northern Trinity Aquifer Due To Urban Growth and Barnett Shale Development. Submitted to Texas Water Development Board. R.W. Harden & Associates, Inc.
- Scanlon, B. R., Dutton, A. R., and Sophocleous, M., 2003, Groundwater recharge in Texas. Report to the Texas Water Development Board. The University of Texas at Austin, Bureau of Economic Geology
- Smith, R. 2004. Aquifers of the Edwards Plateau. Chapter 9: Paleozoic Aquifers of the Llano Uplift. Texas Water Development Board. Report 360. 20p.
- Texas A&M, Texas AgriLife Extension Service. Texas ET. Accessed 2010. <http://texaset.tamu.edu/>.
- Texas Department of Water Resources, 1977. Land Use / Land Cover Maps of Texas LP 62. Texas Department of Water Resources.
- Texas Water Development Board and Texas Agricultural Experiment Station of the Texas A&M University System. Digital Climate Atlas of Texas. Accessed 2009. <http://www.twdb.state.tx.us/gam/resources/resources.asp>.
- U.S. Census Bureau, FactFinder. Accessed October 2011. <http://factfinder.census.gov/>.

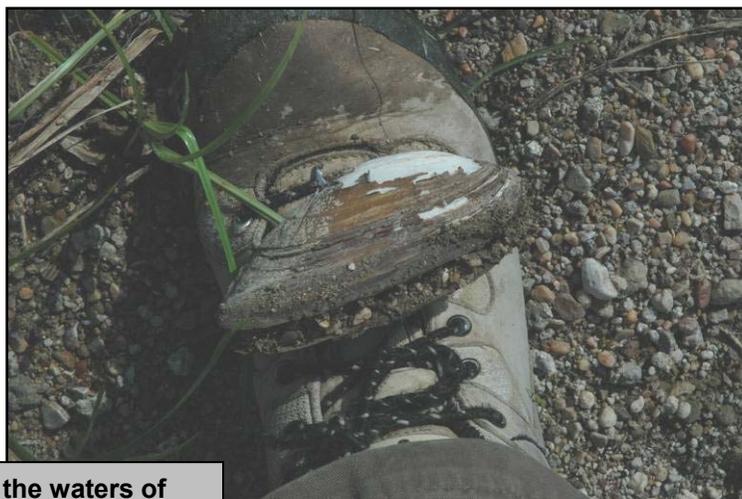
- University of Idaho, College of Agricultural and Life Sciences. REF-ET Version 3.1, accessed 2010. <<http://www.kimberly.uidaho.edu/ref-et/>>.
- University of Oklahoma, Oklahoma Climatological Survey. MesoNet. Accessed 2010. <<http://climate.mesonet.org>>.
- University of Utah, Department of Atmospheric Sciences. MesoWest. Data accessed 2010. <<http://mesowest.utah.edu/index.html>>.
- US Geological Survey, Gap Analysis Program (GAP). February 2010. National Land Cover, Version 1.
- USGS. Digital Elevation Model (DEM). Accessed 2010. <<http://seamless.usgs.gov>>.
- USGS. PART: Streamflow Partitioning Program. Accessed 2010. <<http://water.usgs.gov/ogw/part/>>.
- USGS. Stream Gage measurements. Accessed 2010. <<http://waterdata.usgs.gov/nwis>>.
- Verma, S. Principal Investigator. Undated. AmeriFlux Mead Irrigated Rotation US-Ne2. 2003 - 2006. University of Nebraska-Lincoln School of Natural Resources. <[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.mead\\_maize\\_soybean\\_irrigated.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.mead_maize_soybean_irrigated.01)>.
- Verma, S. Principal Investigator. Undated. AmeriFlux Mead Irrigated US-Ne1. 2003 – 2006. University of Nebraska-Lincoln School of Natural Resources. <[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.mead\\_maize.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.mead_maize.01)>.
- Verma, S. Principal Investigator. Undated. AmeriFlux Mead Rainfed US-Ne3. 2003 – 2006. University of Nebraska-Lincoln School of Natural Resources. <[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.mead\\_maize\\_soybean\\_rainfed.0](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.mead_maize_soybean_rainfed.0)>.
- Verma, S. Principal Investigator. Undated. AmeriFlux Ponca Winter Wheat US-Pon. 1998 - 1999. University of Nebraska-Lincoln School of Natural Resources. <[http://public.ornl.gov/ameriflux/Site\\_Info/siteInfo.cfm?KEYID=us.ponca\\_wheat.01](http://public.ornl.gov/ameriflux/Site_Info/siteInfo.cfm?KEYID=us.ponca_wheat.01)>.
- Wahl, K. L. and T.L. Wahl. Determining the Flow of Comal Springs at New Braunfels, Texas. 12 Apr 2011. Last accessed April 2011. <[http://www.usbr.gov/pmts/hydraulics\\_lab/twahl/bfi/texaswater95/comalsprings.html](http://www.usbr.gov/pmts/hydraulics_lab/twahl/bfi/texaswater95/comalsprings.html)>.



## K. APPENDIX 2: DELIVERABLES



The depositional discontinuity may have resulted from erosion due to early to mid 20<sup>th</sup> Century cotton farming in this Blackland Prairie photo.



Mussel shell found in the waters of the creek shown above.



## K. Appendix 2: Deliverables

### Section A

Data provided from the work in this section include:

- Annual kriged Ppt for 1960 – 2009 provided both as a geodatabase and as individual rasters. [GMA8\_1960\_annual\_ppt.tif]
- Annual  $ET_0$  estimates for 1960 – 2009 provided both as a geodatabase and as individual rasters. [etoi\_1960.tif]
- Annual  $ET_0$  slope and intercept rasters in the model parameters geodatabase and as individual rasters. [ETo\_Intercept\_a\_GAM.tif]
- $ET_0$  calibration station locations point shapefile in the model parameters geodatabase and as a separate file. [ET\_Stations\_Final.shp]

### Section B

Data provided from the work in this section include:

- Permeability rate (in/hr) raster both in the model parameters geodatabase and as an individual raster. [Permeability.tif]
- Calibration watershed USGS gage locations shapefile in model parameters geodatabase and as a separate shapefile. [USGS\_Gage\_Stations.shp]
- Calibration watershed boundaries shapefile in model parameters geodatabase and as a separate shapefile. [USGS\_Watershed\_boundaries.shp]
- Calibration watershed gage data including quarterly and annual Q, Ppt, and  $ET_0$  in Excel Individual Watershed Analysis spreadsheet. [Individual\_Watershed\_QAnalyses.xls]
- Q estimates for 1960 – 2009 provided both in a model variable geodatabase and as individual rasters. [q\_1960.tif]

### Section C

Data provided from the work in this section include:

- Calibration watershed USGS gage locations shapefile provided in the model parameters geodatabase and as a separate file. [USGS\_Gage\_Stations.shp]
- Calibration watershed boundaries shapefile provided in the model parameters geodatabase and as a separate file. [USGS\_Watershed\_boundaries.shp]
- Calibration watershed data including quarterly and annual Ppt,  $ET_0$ , and gaged Q in Excel Individual Watershed Analysis spreadsheet. [Individual\_Watershed\_QAnalyses.xls]
- Q estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [q\_1960.tif]
- $R_{infil}$  estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [rinfil\_1960.tif]
- Unique Q regions boundaries defined in raster provided in model parameters geodatabase and as a separate raster. [Discharge\_Regions.tif]

### Section D

Data provided from the work in this section include:



- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [etai\_1960.tif]
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters. [ETa\_Slope\_b.tif & ETa\_Intercept\_a.tif]
- MODIS raw and intermediate calibration data. [1\_Raw\_Files, 2\_Merged\_NDVI, & 3\_TM\_Adjusted file folders]
- MODIS-derived quarterly NDVI\* for 2005 and 2007. [4\_NDVIstar file folder]

### **Section E**

Data provided from the work in this section include:

- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [etai\_1960.tif]
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters. [ETa\_Slope\_b.tif & ETa\_Intercept\_a.tif]
- MODIS raw and intermediate calibration data. [1\_Raw\_Files, 2\_Merged\_NDVI, & 3\_TM\_Adjusted file folders]
- MODIS-derived quarterly NDVI\* for 2005 and 2007. [4\_NDVIstar file folder]

### **Section F**

Data provided from the work in this section include:

- ETa estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [etai\_1960.tif]
- ETa slope and intercept rasters provided as model parameters in a geodatabase and as individual rasters. [ETa\_Slope\_b.tif & ETa\_Intercept\_a.tif]
- MODIS raw and intermediate calibration data. [1\_Raw\_Files, 2\_Merged\_NDVI, & 3\_TM\_Adjusted file folders]
- MODIS-derived quarterly NDVI\* for 2005 and 2007. [4\_NDVIstar file folder]
- Modified GAP land cover classification raster provided in model parameters geodatabase and as an individual raster. [GAP\_ETa\_Classes.tif]

### **Section G**

Data provided from the work in this section include:

- Dynamic urban boundary shapefiles in the model variables database and as individual files. [Urban\_extent\_yr\_1977.shp]
- NDVI\*-based Q adjustment rasters for each of the four urbanization time steps in the model parameters geodatabase and as individual rasters. [Q\_Urban\_NDVI\_Ratio\_1.tif]
- NDVI\*-based ETa adjustment rasters for each of the four urbanization time steps in the model parameters geodatabase and as individual rasters. [ETa\_Pre\_Urban\_NDVI\_Ratio\_1.tif]

### **Section H**

Data provided from the work in this section include:

- GWR Plateau estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [gwr\_a\_1960.tif]

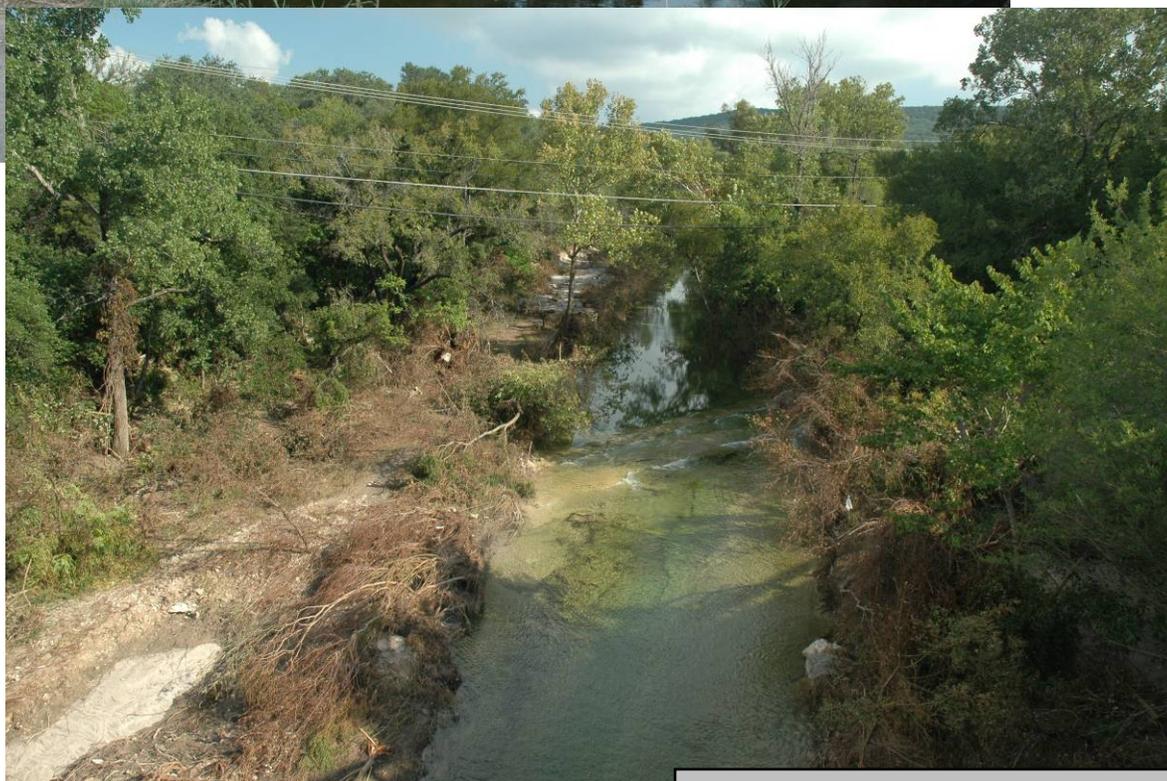
- GWr Blended estimates for 1960 – 2009 provided as a geodatabase and as individual rasters. [gwr\_c\_1960.tif]
- Threshold 2 raster used in GWr Plateau and Blended calculations. [Threshold\_2.tif]
- Max GWr raster used in GWr Blended calculations. [Max\_GWr.tif]
- Slope and intercept rasters for  $GWr_{MinMax}$  linear regression. [GWr\_MinMax\_Slope\_b.tif & GWr\_MinMax\_Intercept\_a.tif]
- Reservoir location raster. [Reservoir\_Mask.tif]



## L. APPENDIX 3: REVIEW OF MODEL CALCULATIONS



Depth gage on an Arizona crossing in the Trinity uplands. Sourced bedrock forms the bank and the gage has been battered during high flows.



Evidence of extreme discharge event following hurricane-derived high precipitation in Summer, 2011 over the Edwards-Trinity Aquifer region.



## L. Appendix 3: Review of Model Calculations

To estimate annual groundwater recharge (GWR), four intermediate variables were defined: reference evapotranspiration (ET<sub>0</sub>), discharge (Q), infiltrated rainfall (R<sub>infil</sub>), and annual evapotranspiration (ET<sub>a</sub>). These variables were calibrated to annual precipitation (Ppt) for the model and a sequence of regression equations was developed to estimate each variable—yielding the data necessary to estimate GWR. Figure L1 shows the flow of model calculations and the major model equations used (intermediate calculations are detailed in the following sub-sections). This section steps through the process of calculating each model variable from ET<sub>0</sub> through GWR. The purpose of this section is to explain the complete model calculation process including all equations and rasters used. Table L1 presents the variable naming conventions used in the model.

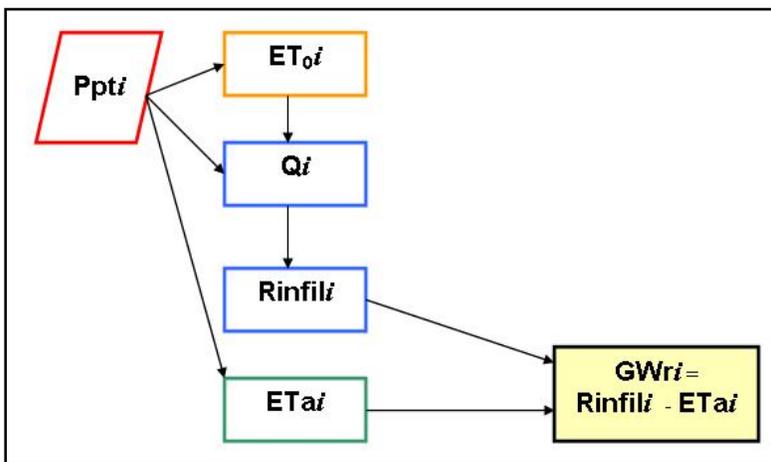


Figure L3. GWR estimation for a given year “i” is guided by four Ppt-driven linear equations:

- 1)  $ET_{0i} = a + b * Ppt_i$
- 2)  $Q_i = a + b * ( Ppt / ET_{0i} )$
- 3)  $R_{infil}^i = Ppt_i - Q_i$
- 4)  $ET_{ai} = a + b * Ppt_i$

Table L1. Overview of all variables used in model calculations.

Table of Commonly Used Abbreviations for Model Calculations

<b>ET<sub>0</sub></b>	Reference evapotranspiration. Reported in inches per year.
<b>ET<sub>a</sub></b>	Annual evapotranspiration. Reported in inches per year.
<b>GWR</b>	Groundwater recharge is the residual of infiltrated rainfall (R <sub>infil</sub> ) minus annual evapotranspiration (ET <sub>a</sub> ). Reported in inches per year.
<b>NDVI* Ratio</b>	A metric developed using the ratio of the average NDVI* from non-urban regional environments to the NDVI* measured over an urbanized pixel was used to represent the degree of urbanization as a decimal fraction. This fraction was used to boost Q (Urban NDVI* Ratio) and decrease ET <sub>a</sub> (Pre-Urban NDVI* Ratio) based on the degree of urbanization present in a pixel at a given year. NDVI* Ratio values are stored in individual rasters masked to appropriate urban / pre-urban boundaries for each model time-step.

<b>Perm</b>	Permeability. Final value used to represent infiltration capacity of soil within model calculations. Reported in inches per hour.
<b>Ppt</b>	Precipitation. Reported in inches per year.
<b>Q</b>	Discharge: surface water drained from recharge system. Reported in inches per year.
<b>R<sub>infil</sub></b>	Infiltrated rainfall; the residual of precipitation minus discharge. Reported in inches per year.

## L1. Reference Evapotranspiration (ET<sub>0</sub>) Estimation

L1.1. ET<sub>0</sub> for year “i” was calculated from Ppt for year “i” using Equation 1.

Input Rasters:

- Annual kriged precipitation (Ppt).
- ET<sub>0</sub> intercept (a).
- ET<sub>0</sub> slope (b).

$$ET_{0i} = a_{ET_0} + b_{ET_0} * Ppt_i$$

**Equation L1**

## L2. Discharge (Q) Estimation

L2.1. Annual slope (b) for Q estimation equation was calculated for year “i” using Ppt<sub>i</sub>, ET<sub>0i</sub>, and the constants presented in Table L2 to solve Equation L2. Q slope was calculated individually for each model year. Unique constants were used in the calculations for each of the two discharge regions in the model defined in the discharge regions raster.

Input Rasters:

- Annual kriged precipitation (Ppt).
- Annual ET<sub>0</sub> calculated using Equation L1.
- Discharge regions.
- Permeability.

**Table L2. Q multiple regression constants used to calculate slope of annual Q regression equation.**

	Region 1	Region 2
<b>b<sub>climate</sub></b>	50.65	4.00
<b>b<sub>Perm</sub></b>	- 3.41	- 1.50
<b>Constant<sub>b</sub></b>	- 4.39	26.00

$$b_{Qi} = ( b_{climate} * Ppt_i / ET_{0i} ) + ( b_{Perm} * Perm ) + Constant_b \quad \text{Equation L2}$$

L2.2. X-intercept of Q estimation equation was calculated for year “i” using Ppt<sub>i</sub>, ET<sub>0i</sub>, and the constants presented in Table L3 to solve Equation L3. Q x-intercept was calculated individually for each model year. Unique constants were used in the calculations for each of the two discharge regions in the model defined in the discharge regions raster.

Input Rasters:

- Annual kriged precipitation (Ppt).
- Annual ET<sub>0</sub> calculated using Equation 1.
- Discharge regions.

**Table L3 Q regression constants used to calculate x-intercept of annual Q regression equation.**

	Region 1	Region 2
<b>b<sub>x-int</sub></b>	0.2983	0.2000
<b>Constant<sub>x</sub></b>	0.5974	0.5752

$$x\text{-intercept}_{Qi} = b_{x-int} * \ln( Ppt_i / ET_{0i} ) + Constant_x \quad \text{Equation L3}$$

L2.3. Intercept (a) of Q estimation equation was calculated using slope and intercept for year “i” (calculated with Equations L2 and L3) to solve Equation L4. Q intercept was calculated individually for each model year.

$$a_{Qi} = - b_{Qi} * x\text{-intercept}_{Qi} \quad \text{Equation L4}$$

L2.4. The Q slope (b) and intercept (a) calculated in Equations L2 and L4 for year “i” were used with Ppt<sub>i</sub> and ET<sub>0i</sub> to solve Equation L5. This initial estimate of Q was refined using Equations L5 through L8, below, to produce a final estimate of Q<sub>i</sub>.

Input Rasters:

- Annual kriged precipitation (Ppt).
- Annual ET<sub>0</sub> calculated using Equation L1.

$$Q_i = a_{Qi} + b_{Qi} * ( Ppt_i / ET_{0i} ) \quad \text{Equation L5}$$

L2.5. Dynamic adjustments to Q predictions over urbanized areas were applied using the Q estimated with Equation L5, Ppt<sub>i</sub>, and the appropriate Urban NDVI\* ratio raster to solve Equation L6. The appropriate rasters for each model time-step are listed in Table L4. Note that “Urban NDVI Ratio<sub>k</sub>” describes the NDVI\* raster from time-step “k”. For each time-step the Urban NDVI Ratio raster was masked to the modeled urban extent and was populated with the value “1” in all other

locations—when “1” is input for Urban NDVI Ratio value the Q estimated using Equation L5 is unchanged.

**Table L4. Urban NDVI\* ratio rasters used in model Q calculations.**

Time-Step “k”	NDVI* Ratio Raster (Urban NDVI Ratio <sub>k</sub> )
1960 – 1977	Q_Urban_NDVI_Ratio_1.tif
1978 – 1989	Q_Urban_NDVI_Ratio_2.tif
1990 – 1999	Q_Urban_NDVI_Ratio_3.tif
2000 – 2009	Q_Urban_NDVI_Ratio_4.tif

$$Q = ( Q \text{ calculated using Equation 5} * \text{Urban NDVI Ratio}_k ) + Ppt_i * ( 1 - \text{Urban NDVI Ratio}_k ) \quad \text{Equation L6}$$

L2.6. Q was capped at the annual precipitation using Equation L7. This adjustment prevents the occurrence of negative infiltrated rainfall ( $R_{infil}$ ) estimates in Equation L9.

$$\text{If } Q_i > Ppt_i, \quad \text{Equation L7} \\ Q_i = Ppt_i$$

$$\text{If } Q_i \leq Ppt_i, \\ Q_i \text{ calculated using Equation L6}$$

L2.7. Negative Q estimates were replaced with zeros using Equation L8. This adjustment prevents illogical  $R_{infil}$  estimates in Equation L9—a negative estimate of Q subtracted from Ppt leads to a  $R_{infil}$  estimate greater than Ppt.

$$\text{If } Q_i < 0, \quad \text{Equation L8} \\ Q_i = 0$$

$$\text{If } Q_i \geq 0, \\ Q_i \text{ calculated using Equation L7}$$

### L3. Infiltrated Rainfall ( $R_{infil}$ ) Estimation

L3.1.  $R_{infil}$  for year “i” was calculated using  $Ppt_i$  and  $Q_i$  to solve Equation L9.

Input Rasters:

- Annual kriged precipitation (Ppt).
- Annual Q calculated using Equations L2 - L8.

$$R_{infil} = Ppt_i - Q_i \quad \text{Equation L9}$$



#### L4. Annual Evapotranspiration Estimation (ETa)

L4.1. ETa for year “*i*” calculated using  $Ppt_i$ , the ETa intercept, and the ETa slope to solve Equation 10. This initial estimate of ETa was refined using Equations L11 and L12, below, to produce a final estimate of  $ETa_i$ .

Input Rasters:

- Annual kriged precipitation ( $Ppt$ ).
- ETa intercept ( $a$ ).
- ETa slope ( $b$ ).

$$ETa_i = a_{ETa} + b_{ETa} * (Ppt_i) \quad \text{Equation L10}$$

L4.2. To correct for a bias in ETa calibration in modern urban areas, an adjustment was applied to ETa estimates for the years before the modern urban extent was reached. For pre-urban conditions, the NDVI\* ratio (urban cell/regional average) was used to boost ETa to match the average in the surrounding countryside. Note that “Pre-Urban NDVI Ratio<sub>*k*</sub>” describes the NDVI\* raster from time-step “*k*” (the appropriate rasters for each model time-step are listed in Table L5). For each time-step the Pre-Urban NDVI Ratio raster was masked to the modeled pre-urban extent and was populated with the value “1” in all other locations—when “1” is input for Urban NDVI Ratio value the ETa estimated using Equation L10 is unchanged.

**Table L5. Pre-Urban NDVI\* ratio rasters used in model ETa calculations.**

Time-Step “ <i>k</i> ”	NDVI* Ratio Raster (Pre-Urban NDVI Ratio <sub><i>k</i></sub> )
1960 – 1977	ETa_Pre_Urban_NDVI_Ratio_1.tif
1978 – 1989	ETa_Pre_Urban_NDVI_Ratio_2.tif
1990 – 1999	ETa_Pre_Urban_NDVI_Ratio_3.tif
2000 – 2009	No Adjustment

$$ETa = ( ETa \text{ calculated using Equation L10} ) + ( 1 - \text{Pre-Urban NDVI Ratio}_k ) * ( ETa \text{ calculated using Equation F2} ) \quad \text{Equation L11}$$



L4.3. ETa was capped at annual precipitation using Equation L12.

**If  $ET_{ai} > P_{pti}$ ,  
 $ET_{ai} = P_{pti}$**

**Equation L12**

**If  $ET_{ai} \leq P_{pti}$ ,  
ETa calculated using Equation L11.**

## **L5. Annual GWr Estimation: GWr Plateau and GWr Blended**

L5.1. GWr for year “i” was first estimated from  $R_{infil_i}$  and  $ET_{ai}$  using Equation L13. The initial estimate from Equation L13 was then refined to produce two separate GWr estimates: GWr Plateau and GWr Blended (described below).

Input Rasters:

- $R_{infil}$  estimated using Equation L9
- ETa estimated and refined using Equations L10 through L12.

**$GWr_i = R_{infil_i} - ET_{ai}$**

**Equation L13**

L5.2. GWr Plateau was estimated from  $GWr_i$  using two rasters: Threshold 2 and Max GWr. Threshold 2 represents a critical  $P_{pt}/ET_0$  value above which GWr is set equal to the maximum GWr for that pixel. This estimate was further refined using equations L15 through L17.

Input Rasters:

- GWr initial estimates calculated using Equation L13.
- Annual kriged Ppt.
- $ET_0$  estimated using Equation L1.
- Threshold 2.
- Max GWr.
- Reservoirs.

**If  $P_{pt} / ET_0 \leq \text{Threshold 2}$ ,  
 $GWr_{\text{Plateau}_i} = GWr_i$  calculated using Equation H1**

**Equation L14**

**If  $P_{pt} / ET_0 > \text{Threshold 2}$ ,  
 $GWr_{\text{Plateau}_i} = \text{Max GWr}$**

L5.2.1. GWr Plateau for year “i” was capped at  $P_{pti}$  using Equation L15.

**If  $GWr_{\text{Plateau}_i} > P_{pti}$ ,  
 $GWr_{\text{Plateau}_i} = P_{pti}$**

**Equation L15**

**If  $GWr_{\text{Plateau}_i} \leq P_{pti}$ ,  
 $GWr_{\text{Plateau}_i}$  calculated using Equation L14**



L5.2.2. Negative estimates of GWr Plateau for year “*i*” were replaced with zeros using Equation L16.

$$\begin{aligned} \text{If } GWr_{\text{Plateau}i} < 0, \\ GWr_{\text{Plateau}i} &= 0 \end{aligned} \qquad \text{Equation L16}$$

$$\begin{aligned} \text{If } GWr_{\text{Plateau}i} \geq 0, \\ GWr_{\text{Plateau}i} &\text{ calculated using Equation L15} \end{aligned}$$

L5.2.3. Final GWr Plateau estimates were set to zero over reservoirs using Equation L17.

$$\begin{aligned} \text{If Reservoir Raster Value} = 1, \\ GWr_{\text{Plateau}i} &= 0 \end{aligned} \qquad \text{Equation L17}$$

$$\begin{aligned} \text{If Reservoir Raster Value} = 0, \\ GWr_{\text{Plateau}i} &\text{ calculated using Equation L16} \end{aligned}$$

L5.3. GWr Blended uses both the GWr estimated from the Plateau model and an estimate generated using a linear regression through the minimum and maximum GWr at each pixel.

Input Rasters:

- GWr initial estimates calculated using Equation L13.
- GWr Plateau estimates calculated using Equation L14 and refined using Equations L15 – L17.
- Annual kriged Ppt.
- ET<sub>0</sub> estimated using Equation L1.
- Threshold 2.
- GWr Min/Max intercept (a).
- GWr Min/Max slope (b).
- Reservoirs.

L5.3.1. GWr Min/Max was estimated using Ppt<sub>*i*</sub>, ET<sub>0*i*</sub>, and the GWr Min/Max intercept and slope to solve equation L18. This estimate was further refined using equations L19 through L21.

$$GWr_{\text{Min/Max}i} = a_{\text{GWr}} + b_{\text{GWr}} * Ppt_i / ET_{0i} \qquad \text{Equation L18}$$

L5.3.2. GWr Min/Max for year “*i*” was capped at Ppt<sub>*i*</sub> using Equation L19.

$$\begin{aligned} \text{If } GWr_{\text{Min/Max}i} > Ppt_i, \\ GWr_{\text{Min/Max}i} &= Ppt_i \end{aligned} \qquad \text{Equation L19}$$

$$\begin{aligned} \text{If } GWr_{\text{Min/Max}i} \leq Ppt_i, \\ GWr_{\text{Min/Max}i} &\text{ calculated using Equation L18} \end{aligned}$$

L5.3.3. Negative estimates of GWR Min/Max for year “i” were replaced with zeros using Equation L20.

**If  $GWR_{Min/Max}^i < 0$ ,**  
 **$GWR_{Min/Max}^i = 0$**  **Equation L20**

**If  $GWR_{Min/Max}^i \geq 0$ ,**  
 **$GWR_{Min/Max}^i$  calculated using Equation L19**

L5.3.4. Final GWR Min/Max estimates were set to zero over reservoirs using Equation L21.

**If Reservoir Raster Value = 1,**  
 **$GWR_{Min/Max}^i = 0$**  **Equation L21**

**If Reservoir Raster Value = 0,**  
 **$GWR_{Min/Max}^i$  calculated using Equation L20**

L5.3.5. Final GWR Plateau and GWR Min/Max estimates were used to calculate GWR Blended for year “i” using equation L22.

**If  $Ppt / ET_0^i \leq \text{Threshold 2}$ ,**  
 **$GWR_{Blended}^i = GWR^i$  calculated using Equation H1** **Equation L22**

**If  $Ppt/ET_0^i > \text{Threshold 2}$ ,**  
 **$GWR_{Blended}^i = (0.7 * GWR_{Plateau}^i$  calculated using Equation L17) +**  
**(0.3 \*  $GWR_{Min/Max}^i$  calculated using Equation L21)**

## L6. Delivered Model Variables

The model variables described above were delivered as individual rasters and as geodatabases for the years 1960 – 2009. All variables are reported in inches per year.

- Kriged precipitation (Ppt).
- Reference evapotranspiration ( $ET_0$ ) calculated using equation L1.
- Discharge (Q) calculated using equations L3 – L8.
- Infiltrated rainfall ( $R_{infil}$ ) calculated using equation L9.
- Annual Evapotranspiration (ETa) calculated using equations L10 – L12.
- Plateau model groundwater recharge ( $GWR_{Plateau}$ ) calculated using equations L13 – L17.
- Blended model groundwater recharge ( $GWR_{Blended}$ ) calculated using equations L13 and L18 – 22.

## M. APPENDIX 4: DRAFT REPORT COMMENTS AND RESPONSES



Deeply incised channels like this one carry large amounts of runoff from the Blackland soil region.



M1

## M. Appendix 4: Draft Report Comments and Responses

1. TWDB Comment: Please remove the word draft from all headers and re-submit as final report.  
HydroBio Response: Draft removed from all headers for final submission of report.
2. TWDB Comment: Please adjust size of the pdf of report to less than 10MB for ease of downloads from web and due to size limitations for attaching report to emails (or break report into smaller components).  
HydroBio Response: Report images were compressed to meet size requirement.
3. TWDB Comment: Page C3: please consider re-writing statement “Q was chosen because it is the water that is exported out as measurable flow, hence unavailable to recharge groundwater”. Surface water/groundwater interactions due to streambed seepage are frequently considered a source of inflow/recharge to the aquifer system. Possibly eliminate “hence unavailable to recharge groundwater”.  
HydroBio Response: Removed “hence unavailable to recharge groundwater” from the end of statement. This qualifier was unnecessary to the explanation of modeling Q provided for the report.
4. TWDB Comment: Page C11, caption for Figure C10: for clarity please adjust reference to figures in the caption from “top” and “bottom” to “left” and “right”.  
HydroBio Response: Caption corrected.
5. TWDB Comment: Page E7: please update reference to figure in line 3 from F3 to E3 and adjust location of table caption for Table E3 to stay on the same page as the table.  
HydroBio Response: Reference corrected. Table E3 placement adjusted to fall on the same page as the caption.
6. TWDB Comment: Section F: please review references to figures and adjust as needed. For example figures appear on page F2 with a caption, “F2. ETa driving variables”.  
HydroBio Response: Figure on page F2 was unnecessary and was removed. Discussion in paragraph one relating to this figure was also unnecessary and removed. All other Section F figure and table references and captions verified and corrected as necessary.
7. TWDB Comment: Page H1: please update caption for Figure H3 to Figure H1.  
HydroBio Response: Caption corrected.