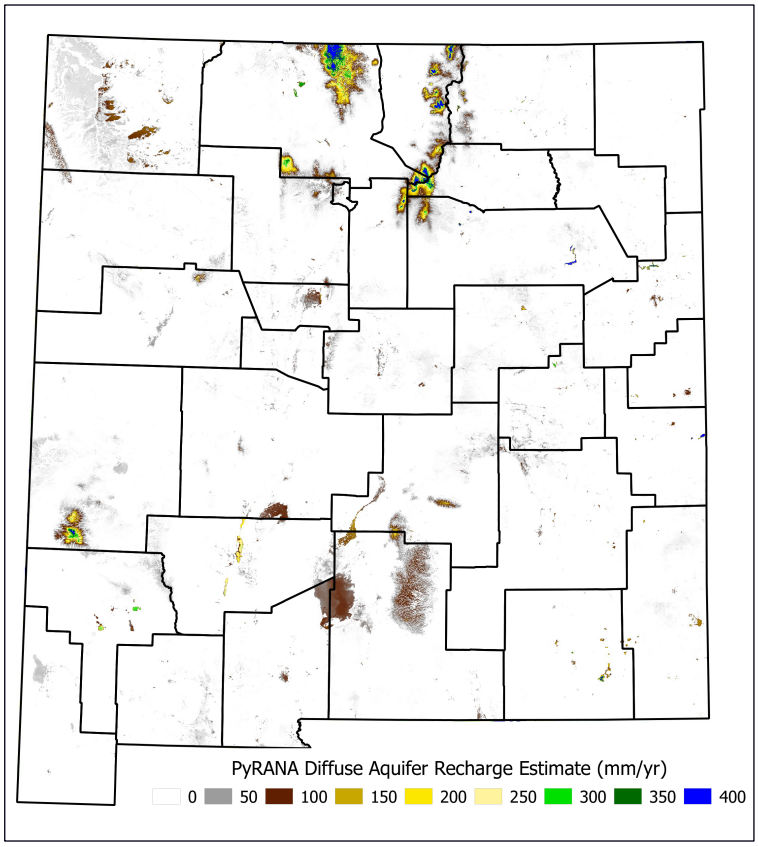
May 2020

ESTIMATING SOIL WATER HOLDING CAPACITY AND RUNOFF IN NEW MEXICO TO IMPROVE MODELED RECHARGE RATES

NM WRRI Technical Completion Report No. 384

Daniel Cadol
Gabriel Parrish
Sarah Reuter
Talon Newton
Fred M. Phillips
Jan M.H. Hendrickx



New Mexico Water Resources Research Institute
New Mexico State University
MSC 3167, P.O. Box 30001
Las Cruces, New Mexico 88003-0001
(575) 646-4337 email: nmwrri@nmsu.edu

ESTIMATING SOIL WATER HOLDING CAPACITY AND RUNOFF IN NEW MEXICO TO IMPROVE MODELED RECHARGE RATES

By
Daniel Cadol
Associate Professor of Hydrology
Department of Earth & Environmental Science
New Mexico Tech

Gabriel Parrish
MS Graduate Student Hydrology Program
Department of Earth & Environmental Science
New Mexico Tech

Sarah Reuter
MS Graduate Student Hydrology Program
Department of Earth & Environmental Science
New Mexico Tech

Talon Newton
Hydrologist
New Mexico Bureau of Mines and Geology
New Mexico Tech

Fred M. Phillips
Emeritus Professor of Hydrology
Department of Earth & Environmental Science
New Mexico Tech

Jan M.H. Hendrickx
Emeritus Professor of Hydrology
Department of Earth & Environmental Science
New Mexico Tech

TECHNICAL COMPLETION REPORT
Account Number EQ01963
Technical Completion Report #384

May 2020
New Mexico Water Resources Research Institute.

The research on which this report is based was funded by the State of New Mexico Legislature NMWRRI2018 through the New Mexico Water Resources Research Institute as part of the Statewide Water Assessment.

DISCLAIMER

The New Mexico Water Resources Research Institute and **affiliated institutions** make no warranties, express or implied, as to the use of the information obtained from this data product. All information included with this product is provided without warranty or any representation of accuracy and timeliness of completeness. Users should be aware that changes may have occurred since this data set was collected and that some parts of these data may no longer represent actual conditions. This information may be updated without notification. Users should not use these data for critical applications without a full awareness of its limitations. This product is for informational purposes only and may not be suitable for legal, engineering, or surveying purposes. The New Mexico Water Resources Research Institute and **affiliated institutions** shall not be liable for any activity involving these data, installation, fitness of the data for a particular purpose, its use, or analyses results.

ACKNOWLEDGEMENTS

We gratefully acknowledge programming and code version management support from Jake Ross of the New Mexico Bureau of Geology and Mineral Resources. He has been very patient in training us in the use of Python and Github. This work also builds on the efforts of previous MS students who guided the development of the various models used here. David Ketchum developed the Evapotranspiration and Recharge Model (ETRM), Peter ReVelle developed the Gridded Atmospheric Data downscalinG Evapotranspiration Tools (GADGET), and Esther Xu expanded ETRM into the Python Recharge Assessment for New Mexico Aquifers (PyRANA) model.

ABSTRACT

This report summarizes efforts to improve groundwater recharge estimates using the Python Recharge Assessment for New Mexico Aquifers (PyRANA) model by developing new methods to parameterize soil-water holding capacity (SWHC, also called total available water – TAW) and rainfall-runoff relationships in mountainous portions of northern New Mexico. The PyRANA model is a soil-water-balance model that runs on a daily time step and at a 250 m grid resolution across the state of New Mexico, tracking precipitation inflows and evapotranspiration (ET), runoff, and deep percolation outflows at each grid cell. Parameters that partition rainfall into runoff and infiltration, and the amount of water storage available in each cell (i.e., SWHC) were identified as the most poorly constrained of the highly influential model parameters in previous research efforts.

The SWHC parameterization effort explored a depletion-tracking method, which uses independent estimates of precipitation and ET to track soil-water content through time and assumes that the vegetation root zone is naturally adjusted such that SWHC meets the maximum water demand over a medium time range. In this effort, we found that two of the best available monthly ET products, the operational Simplified Surface Energy Balance model (SSEBop) and the Priestly-Taylor-based Jet Propulsion Laboratory product (PT-JPL), were not in adequate agreement with precipitation estimates from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) for use in this exercise. In an effort to parsimoniously correct the ET biases, we scaled the two ET products so that long-term ET matched long-term precipitation from PRSIM. This led to plausible SWHC estimates in natural upland landscapes (non-agricultural and non-riparian) using the scaled PT-JPL, but not using the scaled SSEBop. In an effort to use an ET estimate that is limited by water availability as well as energy availability, we also explored a recursive and iterative approach that employed the ET estimate from PyRANA itself in the depletion-tracking SWHC estimation method. This led to plausible values of SWHC, but spatial patterns that are unverifiable at this time. Finally, we explored an ecosystem-based approach to map SWHC, using landcover classification and a vegetation greenness index (normalized difference vegetation index – NDVI). This approach would require extensive field validation to determine appropriate relationships between NDVI and SWHC for unique vegetation types. But our preliminary exploration yielded promising maps of SWHC that are plausible in both magnitude and spatial variation.

The rainfall-runoff parameterization effort first explored potential biases in the PRISM precipitation dataset. Earlier research had found that PRISM accurately models total seasonal rainfall, but it under-predicts the magnitude of high-rainfall days and over-predicts the number of days with small rainfall amounts during monsoon season in the Walnut Gulch Experimental Watershed (WGEW) in south-east Arizona. In this report, we found a similar, but slightly less pronounced, pattern of bias in the Sangre de Cristo mountains of northern New Mexico. In both WGEW and northern New Mexico we only observed a bias during the monsoon season; winter storms have clear agreement between PRISM and local rain gauges. Next, we used runoff data from small (~ 1 km^2) gauged watersheds in the Jemez and Sangre de Cristo mountains to develop a multiple linear regression to predict runoff using rainfall magnitude and intensity. We found that unlike the experience with WGEW, intensity was not a significant predictor of runoff. Instead, the daily rainfall amount alone was an adequate predictor of runoff for use in PyRANA.

TABLE OF CONTENTS

LIST OF FIGURES

OVERVIEW

Motivation

In most basins in New Mexico, groundwater resources are being utilized faster than they are recharged (Rinehart et al., 2015; Rinehart et al., 2016; Rinehart and Mamer, 2017). In order to better define a sustainable usage rate, water managers need information on current rates of aquifer recharge. However, estimating recharge is extremely difficult because it is a small component of the hydrological budget at any given point. In New Mexico, the hydrologic budget is dominated by precipitation and evapotranspiration (ET), leaving a small residual of these two large terms to account for both runoff and aquifer recharge. Small errors in estimates of ET or precipitation will lead to large errors in estimates of runoff and recharge (Gee and Hillel, 1988; Hendrickx and Walker, 1997; Kearns and Hendrickx, 1998). Previous research conducted as part of the NM WRRI Statewide Water Assessment determined that PRISM (Parameter-elevation Regressions on Independent Slopes Model (Daly et al., 2008; Daly and Bryant, 2013) is an adequate daily precipitation product, at 800 m resolution, in determination of annual precipitation depths, but no adequate product for spatially continuous ET was found (Schmugge et al., 2015). In order to achieve the goal of estimating recharge, development of a new ET model was therefore determined to be a requisite first step, leading to the combined ET and Recharge Model (ETRM) (Ketchum, 2016) and its successor, the Python Recharge Assessment for New Mexico Aquifers (PyRANA) model (Xu, 2018).

The PyRANA model is a soil-water-balance model that runs on a daily time step and at a 250 m grid resolution. Each cell tracks the depletion of water in the root zone of the soil, with "depletion" referring to extractions of water from the soil-water reservoir that deplete it below the maximum holding capacity of the soil. Tracking the depletion involves accumulating water inflows from PRISM precipitation and outflows to deep percolation, ET, and runoff. Partitioning of precipitation into these three outflows is the core function of PyRANA. It should be noted that there are no lateral flows in PyRANA; runoff is considered lost to the soil, and we have not yet incorporated attempts to track it through the channel network. Hence, there is no run-on of water from neighboring cells, no subsurface water inflows, and no water sources other than precipitation. This limits application of the model somewhat, excluding its use in riparian zones and irrigated agriculture, but it makes the calculations tractable and efficient. PyRANA can model thirteen years of daily data for the entire state of New Mexico in six hours, running on a basic desktop computer. Another important result is that PyRANA reports only diffuse recharge, that is, recharge through the soil. Focused recharge, for example through channel beds and playas, must be calculated separately, such as by estimating a percent of the modeled runoff that recharges aquifers rather than being transpired by riparian vegetation or evaporated from lakes.

Deep percolation

The first of the three water-budget outflows mentioned above, deep percolation, occurs when the soil-water holding capacity (SWHC) of the soil is exceeded in any given day's water budget. At these times, the depletion goes to zero, and all excess water is modeled to percolate below the root zone. Conceptually, this water is assumed eventually to recharge the underlying aquifer. The term SWHC is synonymous with total available water (TAW) and many other similar terms (see Parrish et al., 2017 and Parrish, 2020, for a complete discussion). This deep percolation is the target outflow of the PyRANA model, but it depends on the accurate calculation of the other two outflows, ET and runoff.

Evapotranspiration

The second outflow, ET, is calculated in PyRANA using the dual crop-coefficient method (Allan et al., 2005a; Allan, 2011). This method relies on first calculating a reference ET (ET_{ref}), which for our method is the ET that would result from the given weather conditions for a well-watered, healthy alfalfa crop (Allan et al., 1998; ASCE-EWRI, 2005). This is similar to some definitions of the commonly used term 'potential ET', although other dissimilar definitions are widespread, so we avoid its use here. We utilized a daily reference ET product that was generated by the Gridded Atmospheric Data downscalinG Evapotranspiration Tools (GADGET) (Revelle, 2017) at the native resolution of PyRANA. GADGET takes in daily radiation and atmospheric data from NLDAS and METDATA (Abatzoglou, 2013) and downscales them, while accurately representing the influence of topography and elevation. In particular, GADGET calculates insolation and shading adjustments due to topographic effects on direct, diffuse, and reflected radiation (Revelle, 2017).

The dual crop coefficient method then takes this ET_{ref}, which is the ET from healthy alfalfa, and adjusts it to account for different types of vegetation cover, and for stress due to limited soil moisture, using the equation

$$ET_a = ET_{ref} K_{cb} K_s \quad [1]$$

where ET_a is actual ET, K_{cb} is the crop coefficient, and K_s is the stress coefficient. The crop coefficient is estimated using the normalized difference vegetation index (NDVI) on the day being calculated, based on observations from the MODIS satellites. NDVI is a commonly used satellite remote sensing-based proxy for some combination of vegetation density, health, and vigor. Gaps due to limited satellite coverage (satellite overpasses occur approximately every other day) and cloud cover are filled by interpolation. NDVI typically ranges from 0.1-0.7 in natural land cover, and K_{cb} is calculated as

$$K_{cb} = 1.25[NDVI] \quad [2]$$

as proposed by previous researchers (Rafn et al., 2008; Appendix G in Jensen and Allen, 2016). The stress coefficient, K_s, ranges from 0 (for totally dry soil) to 1 (for unstressed vegetation). When soil moisture is greater than ~ 75% of its capacity, vegetation is expected to transpire at its reference rate (Manabe, 1969) and $K_s = 1$. This threshold can also be defined in a complementary way, instead describing the fraction of SWHC that a crop can extract from the root zone without suffering water stress (Allan et al., 1998). This maximum unstressed water extraction, p, varies with crop type, and for natural vegetation we use $p = 0.4$, following recommendations of the United Nations Food and Agriculture Organization and the American Society of Civil Engineers (Allan et al., 1998; ASCE-EWRI, 2005). As soil moisture drops below this level (i.e., as D/SWHC > 0.4, where D is the soil-moisture depletion), then vegetation begins to respond to water stress by closing stomata and reducing transpiration below the reference rate. Following Allan et al. (2005a), we assume a linear decrease in K_s from 1 to 0 as D/SWHC increases from p to 1. Additionally, PyRANA estimates evaporation from the soil surface, which is distinct from the transpiration calculated above, using the methods described in Allan et al. (2005a, 2005b) and Allan (2011) and described in previous technical reports (Hendrickx et al., 2016).

As can be seen from the above discussion, the estimated value of SWHC for each grid cell will have a first-order influence on the estimated K_s, and therefore the calculated ET_a. Accurate estimation of SWHC is therefore critical for producing accurate estimates of deep percolation (i.e., groundwater recharge). Previous versions of PyRANA used SWHC values from Natural Resources Conservation Service (NRCS) soil maps and soil property databases. More precisely, SWHC was calculated as the difference between volumetric soil moisture at field capacity and at the permanent wilting point (θ_{fc} - θ_{wp}) multiplied by the rooting depth, all three of which are estimated in the NRCS soils databases. However, the NRCS soil maps are created at a much coarser scale than the 250 m grid used in PyRANA, based on widely separated sampling points, and are not intended for use over such small extents (as described by a warning on the NRCS web data viewer). Rooting depth can vary greatly within a single mapped soil taxonomic unit. Since the beginning of this modeling effort, we have been searching for more suitable estimates of SWHC. The first chapter of this report, SWHC Calibration, explores three independent approaches to generating a high-resolution map of SWHC across New Mexico.

Runoff

The third outflow, runoff, was originally estimated as the amount of water falling at intensities exceeding the saturated hydraulic conductivity of the soil (K_{sat}). In this procedure, the K_{sat} of each cell was downloaded from the NRCS soils databases STATSGO and SSURGO, and intensity was calculated assuming a 6-hour storm duration (for the full daily water input from PRISM) during the monsoon season (July-Sept), and a 12-hour storm duration during the rest of the year. There are issues with both assumptions. First, as previously described, the NRCS soil maps are not suitable for use at such a detailed scale. Furthermore, the hydraulic conductivity and related parameters are calculated from estimated regression functions using percent sand, silt, and clay of the soil as inputs. In reality, soil texture is not uniform across the extent of a mapped unit, and again, it is inappropriate to use these soil data at a 250 m resolution. Second, the assumption that daily precipitation occurs steadily over six hours in the monsoon season is clearly a gross simplification.

To address these weaknesses in runoff estimation, MS student Esther Xu analyzed rainfall and runoff data from the Walnut Gulch Experimental Watershed in southeast Arizona (Xu, 2018), using Walnut Gulch rain-gauge data that were not part of the PRISM network. First, she found that the annual totals agreed with the gauge record at Walnut Gulch, but that the distribution was subtly different. PRISM missed the largest rainfall events and overestimated the number of days with small amounts of rainfall. In response, a PRISM precipitation adjustment was implemented in PyRANA. Second, she found that the intensity of rainfall had a moderate but significant correlation with the daily total amount of rainfall, with greater totals falling at higher intensities on average. Finally, she found that runoff from small, high-order watersheds (< 2 km^2) could be predicted based on the daily amount and intensity of rainfall measured at rain gauges within the watershed. Therefore, we implemented a regression-based runoff module in PyRANA. For each day, an intensity is estimated based on the probability density associated with the daily rainfall total, and this intensity and the total amount are used to predict runoff. Different regression models were fit for the monsoon season and the remainder of the year.

This runoff calibration is expected to be appropriate for low-elevation rangelands in southern and central New Mexico because it was calibrated at Walnut Gulch, which is

climatologically similar. However, we expect a different calibration will be needed for mountainous areas and northern New Mexico. The second chapter of this report, Rainfall-Runoff Relationships, presents an analysis of rainfall and runoff from study watersheds in the Santa Fe River basin and in Valles Caldera National Preserve. Additionally, rain gauge versus PRISM precipitation analyses were conducted using northern New Mexico data, including from a weather station at Taos, NM. These new relationships are used to implement new predictions of runoff in those portions of New Mexico for which the Walnut Gulch data are inappropriate due to climatological differences.

Report organization

For clarity and ease of communication, we have divided this report into two chapters: SWHC Calibration, and Rainfall-Runoff Relationships. The two chapters of this report present two distinct improvements to the PyRANA recharge model, and address what we, the model developers, perceive to be the two most significant sources of uncertainty in prior versions of the model. In the process of addressing these uncertainties, we have: 1) reviewed existing remote-sensing ET products and found them incompatible with the PRISM precipitation product for the purpose of soil water depletion tracking, 2) generated new estimates of SWHC, a fundamental soil property, for the state of New Mexico, and 3) developed simple new expressions describing watershed hydrology in the mountainous areas of northern and central New Mexico.

Chapter I: SWHC Calibration

Introduction

To model soil-water content through time, it is vital to quantify the fluxes, such as precipitation, evapotranspiration (ET), runoff, and deep percolation at appropriate time scales. It is also vital to know the amount of storage available, that is, the total available water (TAW), also known as the soil-water holding capacity (SWHC). Yet SWHC is rarely measured, perhaps due to the difficulty of measuring it at an adequate spatial resolution. The extreme heterogeneity of soils and rooting depth means that direct physical quantification is only possible for unrealistically simple scenarios.

The zone of soil and weathered bedrock through which vegetation extends roots, and thus maintains the potential connection between infiltrated water and the atmosphere, can vary greatly in its capacity to hold water. Higher rates of drainage (i.e., in circumstances of high infiltration and low field capacity) will tend to promote deep percolation at the expense of transpiration. Even when no vegetation is present, evaporation from the surface soil varies, as pore size and structure affect the ability of water to move upward from the deeper soil layers via capillary action or vapor diffusion. And where vegetation does extract water, from soil pores into roots, the tension the plant can exert and the spatial distribution of the root network influence the minimum soil water content that can be achieved. For all these reasons, an upscaled view of SWHC is likely to be more useful for soil-water-balance models, and parameterizing SWHC by calibration is more likely to be successful than parameterizing SWHC by direct measurement of soil properties at discrete points. A purely measurement approach would require tens or even hundreds of soil pits to be dug to obtain a representative value for a single 250 m pixel. We propose that calibration, for example trying a range of SWHC values in the model and selecting the value that leads to the best agreement of some key flux such as ET with independent observations, is the more efficient approach, with some degree of field-verification via soil pits.

In the PyRANA model, and the dual crop-coefficient method of ET estimation that it employs, SWHC affects results in two main ways. First, it determines the amount of storage that must be filled before percolation below the root zone is initiated and aquifer recharge can begin. A very large SWHC, due either to deep rooting depths or abundant storage between field capacity tension and the maximum tension of the local vegetation, will rarely be filled, meaning very little water will be displaced below the lowermost roots and continue to the water table. Second, ET is expected and modeled to respond to relative moisture (i.e., % of capacity) rather than absolute moisture (i.e., mm of water). So, if SWHC is very large, for example 1000 mm, a depletion of only 50 mm will not lead to significant drought stress and transpiration will remain high. In contrast, if SWHC is small, for example 100 mm, then a depletion of 50 mm may lead to major reductions in ET as the vegetation cover responds to the loss of water resources. However, if the soil is beginning from a nearly dry state, then a 50 mm rainfall event absorbing into a 1000 mm SWHC soil will hardly increase the water content enough to increase the ET, whereas a 50 mm rainfall event absorbing into a 100 mm SWHC soil will increase the water content enough to raise ET rates suddenly.

Thus, if SWHC is overestimated, we may expect recharge to be undercalculated and ET variation through time to be excessively subdued. In contrast, an underestimated SWHC could lead to exaggerated swings in ET following rainfall events and an overestimated recharge. The

differences in ET are likely to balance out over a long time period, but the recharge errors would accumulate and represent a major problem for the intended use of PyRANA.

As described above, in the earliest versions of the PyRANA model, SWHC was extracted from NRCS soil maps and their associated soil-property databases. Because these soil-property values were selected and vetted by soil scientists, this had the advantage of preventing any patently unrealistic values from being used. However, it also led to a surprisingly narrow range of SWHCs being mapped, some of which may be unrealistically small. It also produced a patchy, blocky texture in the SWHC map due to the large map units and the horizontal and vertical bounds to the individually published maps (Figure 1). In general, we had little confidence in any specific model cell's SWHC value. Therefore, we have made every effort to explore alternative means of developing the best available estimates of SWHC, and to define a defensible range of uncertainty for its values.

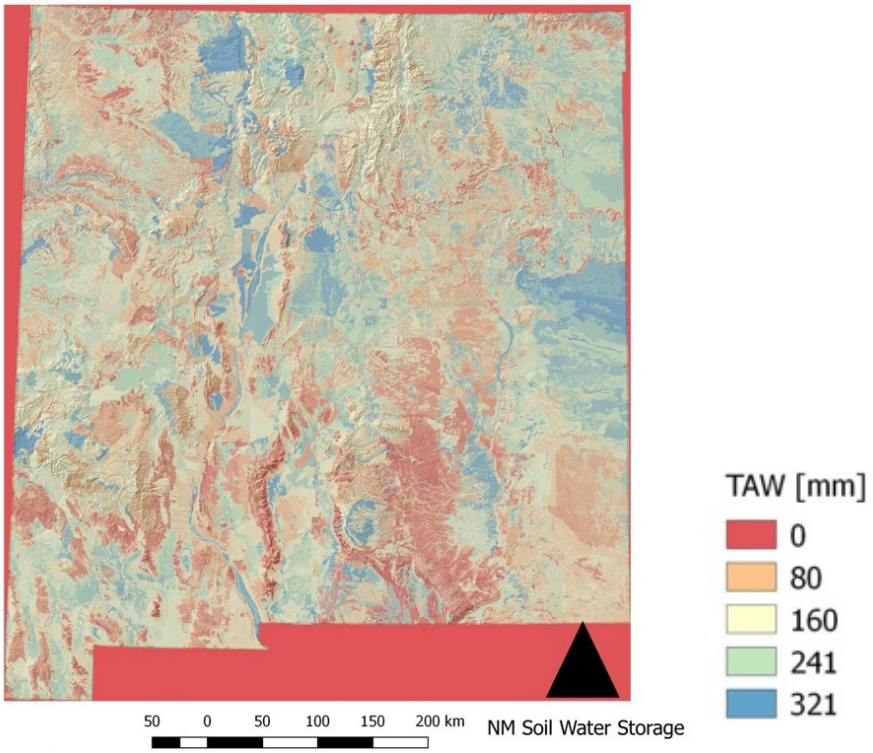


Figure 1. Soil-water holding capacity (SWHC, also known as total available water – TAW) within the root zone, extracted from NRCS databases for New Mexico (Ketchum, 2016). An alternative version, with a color ramp consistent with later figures of SWHC, is presented in Figure 4.

The calibration, rather than measurement, of SWHC has been attempted previously, and we build from these efforts, especially those of Wang-Erlandsson et al. (2016) using a 'depletion-tracking' method. In this approach the cumulative depletion is tracked over a long period of time, summing monthly estimates of precipitation, which reduce the depletion, and monthly estimates of ET, which increase the depletion. Based on the assumption that vegetation, especially forest vegetation, will extend its roots deep enough to access an adequate store of water to survive the largest local drought (consistent with Budyko, 1956), the depletion-tracking method then finds the maximum depletion in the time series and assumes that this provides an

accurate estimate of the SWHC. In grasslands, droughts may periodically kill the vegetation, meaning the maximum depletion may at times exceed the SWHC. But if drought kills the grasses every ten years on average, we may infer that the ten-year-recurrence depletion will be an accurate estimate. We also note that desert shrubs have been documented to dry out the vadose zone for many meters below their maximum root depth (Sandvig and Phillips, 2006), thereby producing a SWHC greater than even multiple years of precipitation. In this case, the maximum depletion observed in a few decades of data will provide only a lower limit on the actual SWHC.

Methods

Depletion tracking to estimate SWHC

The first SWHC calibration we present here is based on the Wang-Erlandsson depletion tracking method (Wang-Erlandsson et al., 2016). We tracked depletions within the soil over the period 2000-2013, which is the time period for which we have full PRISM, GADGET, and remotely sensed ET data for the state. Following Wang-Erlandsson et al. (2016), we began with zero depletions, then each month added the monthly total ET and subtracted the monthly total precipitation. If this caused the depletion to become negative, it was reset to zero (since any negative depletions will become runoff or recharge). We also explored what would happen if we allowed depletions to go negative, but this became unnecessary when we adjusted the ET products to scale with precipitation as described below.

For the precipitation input of water (subtraction from depletion) we used downscaled monthly PRISM data. For the ET output of water (addition to depletion), we first used monthly Simplified Surface Energy Balance operational (SSEBop) ET estimates (Senay et al., 2011; 2013) based on MODIS data. Second, we used monthly ET estimates from a Priestly-Taylor-based product from the NASA Jet Propulsion Laboratory (PT-JPL) (Fisher et al., 2008). A detailed description of the physics and assumptions underlying each product is presented in Parrish (2020). In both cases, we found that there were many cells for which annual ET consistently exceeded annual precipitation from PRISM, leading to steadily rising depletions, often exceeding 5 m of soil-moisture deficit by the end of the 13-year analysis (or only ten years for PT-JPL due to data only being available from 2002 onward and excluding 2013 due to excessive data gaps). This clearly breaks with the conceptual model of Wang-Erlandsson, suggesting an error in either the precipitation or ET products. Because we found that the cumulative PRISM data agree well with independent rain gauges (see Chapter 2 of this report as well as Schmugge et al., 2016), we concluded that the ET remote-sensing products are biased high in these areas. We also found, for both ET products, cells in which depletions stayed at zero for the base method, or if we allowed depletions to go negative, had steadily declining depletions throughout the 13-year analysis period. It is possible that these cells could be areas that produce very large volumes of runoff or recharge, but the extreme behavior and broad extent of these areas, especially in low-elevation desert shrubland, suggests that these are more likely to be areas where the remote sensing ET products are biased low.

Because our initial analysis led us to question the validity of these ET products, and because they are unusable in the depletion-tracking method if they do not match the net precipitation in the long term, we elected to apply a linear adjustment. We forced the total 13-year ET (or 10-year ET for PT-JPL) at each grid cell (ET_{tot}) to equal the total 13-year (or 10-year) precipitation from PRISM for that cell ($PRISM_{tot}$), multiplying each monthly ET value by

the ratio $PRISM_{tot}/ET_{tot}$. Adjusting ET in this way means that the effect of runoff and recharge is missed in this calibration, but these fluxes are expected to be much smaller than ET, and missing their contribution to the cumulative depletion will minimally alter the estimated SWHC. And other than a small change to the SWHC, this omission in no way affects their proper calculation in the future running of the PyRANA model. With ET adjusted to be of the same long-term magnitude as PRISM precipitation, there was then no further need to allow negative depletions, and we could proceed with the conceptual model of Wang-Erlandsson.

This ET adjustment assumes that the physical basis for estimating ET and the relative variations in ET are valid for the remote-sensing methods, but that each remote-sensing product has a consistent linear bias. This is probably not strictly true for either of the ET products, but it gives maximum credence to the products while permitting them to be used with PRISM for depletion tracking. Previous work conducted as part of the New Mexico Statewide Water Assessment has suggested that PRISM is more reliable than any of the reviewed ET products, which included SSEB but not PT-JPL (Schmugge et al., 2016).

Recursive depletion tracking to estimate SWHC

From the beginning of its development, PyRANA has been motivated by a need to calculate a statewide ET product that is constrained by both the energy budget and the water budget. Because it uses the dual-crop-coefficient method to estimate ET, the model is unable to estimate ET that is significantly greater than precipitation in the long term. PyRANA is thus not subject to the ET-bias problem we found with SSEBop, and to a lesser extent with PT-JPL. Thus, we attempted to develop a procedure that could take advantage of PyRANA's well-constrained ET estimation in the depletion-tracking approach to calibrating the SWHC that is used in PyRANA itself. Clearly there are potential issues with circular calibration, but these are minimized by the non-linear response of ET to changes in SWHC and by the observation that errors in ET estimation caused by inaccurate SWHC values tend to balance in the long term.

In this method, we began by setting SWHC to a large value, for example, 1000 mm, across the entire state of New Mexico. Then, beginning with zero depletions, we ran the PyRANA model on a daily time step from 2000-2013 (Ketchum, 2016; Xu, 2018). Because the soils initially held a great deal of water, the first few years of the simulation showed a rapid drying of the soils, especially in vegetated areas. Following this, depletions rose and fell with seasonal and annual variations in precipitation and ET. We then selected the maximum and minimum depletion, excluding the first two years of spin-up, and used this range of depletions that reflects the depth of soil-moisture storage utilized by the vegetation community as the next estimate of SWHC. We then reran PyRANA with this new SWHC map, again starting with zero depletions. For most cells, there was initial dry-down as in the first iteration, but because the storage was smaller, this spin-up period completed more quickly. For the remainder of this second simulation, the depletion rose and fell as before, following climate trends, but usually in a more constrained range of values, because the storage was smaller. There began to be recharge events following large precipitation inputs, because in places the SWHC was beginning to be small enough to be filled and therefore to allow the onset of deep percolation. Again, we analyzed the time series in each cell for the maximum and minimum depletion and used this range as the SWHC for the next iteration of the model. The decrease in the SWHC in this

iteration was in almost all cells much less than the decrease following the first model run.

We repeated this process ten times. In almost all cells, the SWHC declined with each iteration until it reached a minimum value and there plateaued. The rate of decline from one iteration to the next declined as well, even before the plateau was reached. Most cells had reached a constant value by the seventh iteration, so for simplicity we used the SWHC map from this iteration as the final result for this method.

Vegetation-based SWHC estimation

Given the issues regarding independent ET products available for SWHC calibration, we also explored approaches for estimating SWHC using other observations. The most promising one that we pursued utilizes remote sensing observations of vegetation cover to map SWHC. This approach is likely to be superior to the soil-property-based method because satellite imagery has been used to produce detailed vegetation maps at high spatial resolution (Cihlar, 2000; Homer et al., 2004; Rollins, 2009), whereas soil-property maps are based on quite sparse data. We began with the assumption that each vegetation community or ecosystem will have a characteristic soil-moisture demand, and depending on the soil and subsurface hydrology, it will extend its root zone to meet that demand. We also assumed that not every patch of that community will fully meet its water needs every year, and that this failure (inadequate SWHC) will be manifested in reduced vegetation abundance and vigor, and increased plant spacing and bare soil exposure. These effects should be visible in long-term average NDVI values. Thus, SWHC should be a function of vegetation community and NDVI. Determining the exact form of this function and calibrating the subsequent parameters is beyond the scope of this exploration. Instead, we assumed a linear variation in SWHC from zero to a community-specific maximum as long-term average NDVI varies from a lower to an upper limit (see Parrish, 2020 for details).

To classify the entire state of New Mexico into vegetation community classes we used the LandFire pre-existing vegetation product (Rollins and Frame, 2006; Rollins, 2009). This is a trained classification that uses Landsat multispectral imagery to group each 30 m pixel into a vegetation class, of which 97 are present in New Mexico. These 97 classes are excessive for this exploratory exercise, so we grouped them into seven communities: bare ground, arid shrubland (mostly creosote), low elevation grassland and rangeland, piñon-juniper woodlands, ponderosa and mixed conifer forest, montane grassland and meadows, and riparian zones and agricultural lands. For each community, we selected an estimated maximum SWHC based on average rooting depths reported by Sandvig and Phillips (2006). These rooting depth values were: 0 m for bare ground, 3 m for creosote, 2 m for grassland, 4 m for piñon-juniper, 3 m for ponderosa forest, 2 m for montane grassland, and effectively unlimited SWHC for riparian zones and agricultural lands (PyRANA is not intended for use in such areas). For simplicity, these rooting depths were multiplied by a uniform field capacity to wilting point difference (θ_{fc} - θ_{wp}) of 0.2 to estimate SWHC. These maximum SWHC values should be subject to adjustment based on future research. Also note that PyRANA does parameterize total evaporable water (TEW, storage capacity in the evaporative layer) using the NRCS soils databases. So, although the rooting depth may be zero for cells classified as bare ground, the minimum allowed TAW equals TEW.

For each of these seven communities, we needed to identify NDVI thresholds to associate with zero SWHC and maximum SWHC. To do so, we extracted the long-term (2000-2013)

average NDVI for each 250 m grid cell classified as part of each community, and analyzed the probability distribution functions. Based on visual inspection, we selected the 5th percentile NDVI value to be the lower threshold, and the 75th percentile NDVI value to be the upper threshold. Cells within that community that had NDVI less than the lower threshold were given a SWHC of zero, characteristic of bare ground, while those with NDVI greater than the upper threshold were given the community's maximum SWHC. Cells with NDVI between the thresholds were given SWHC that scaled linearly between zero and the maximum.

Results

Depletion tracking to estimate SWHC

The comparison of long-term PRISM to SSEBop (2000-2013) and PT-JPL (2002-2013) shows major incompatibilities between the products, especially between PRISM and SSEBop (Figure 2). It is normal and expected for there to be much more ET than precipitation in riparian zones and irrigated agriculture. And it is also normal and expected for high alpine areas to have less ET than precipitation, since these are areas where runoff and aquifer recharge will be important components of the water budget. But both products show much greater ET in forested mountains than there is precipitation to supply that ET, especially in areas where shading is an important component of the energy budget, suggesting a possible source of error. And SSEBop estimates nearly zero ET in many arid rangelands of New Mexico, for example in the creosote community in the Chihuahuan desert. We know from eddy covariance flux tower data available in the AmeriFlux climate data network that ET in these areas is not in fact zero, and that annual ET approximates annual precipitation. Overall, PT-JPL does well in most parts of the state, especially the rangelands, but appears to overestimate ET in the mountains, which is the dominant locale where recharge occurs and where we particularly desire accurate SWHC estimates. In contrast, SSEBop does not perform well in any portions of the state, with the possible exception of riparian zones and irrigated agriculture (which we are unable to evaluate accurately), leading us to doubt its capability to assist in estimation of SWHC in natural vegetation.

Nonetheless, we conducted depletion tracking using both unadjusted and adjusted SSEBop and PT-JPL ET estimates. The resulting SWHC maps show wide variations across the state (Figure 3). Recall that maximum depletion is interpreted as a minimum SWHC estimate assuming the vegetation community does not die of drought stress. In the unadjusted maps (Figure 3A and 3C), riparian zones are shown as having extremely high SWHC because the imported water sustains more ET than precipitation alone can explain, leading to continually increasing calculated depletions. In most of the state, SSEBop estimates little to no ET ever, meaning the depletion never increases beyond a tiny value before rainfall refills the depletion. In forest-covered mountains, however, the uncorrected SSEBop estimates ET that far exceed PRISM precipitation (Figure 2) leading to unrealistically high SWHCs.

In the adjusted SSEBop depletion tracking SWHC estimate, many rangeland areas have very high maximum depletions. This is because SSEBop only estimates ET to occur on a very small number of days with all others being zero, and when these ET values are scaled up to match the total PRISM precipitation, all the ET piles on to these few days. The adjusted ET estimates for these few days are completely unrealistic (meaning our assumption that ET errors are uniformly biased is wrong), and these lead to a saw-toothed pattern in the cumulative

depletion through time: depletions are filled by precipitation over many months, then a single day of unrealistically high ET depletes the soil by a proportionally large amount, followed by slow refilling and eventual re-depletion. For those few cells with no ET ever in the 13-year record, the SWHC is zero, leading to speckled high and low (red and blue) cells in the SWHC map (Figure 3B).

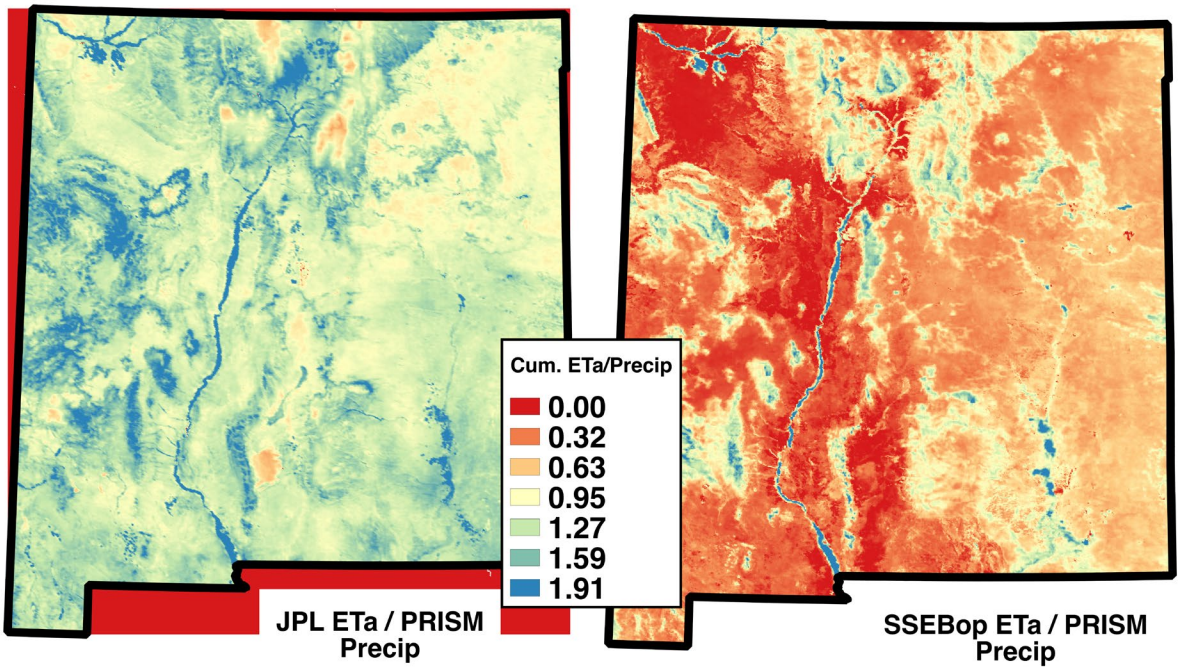


Figure 2. Ratio of long-term unadjusted ET_a to long-term PRISM precipitation for two remotely sensed ET products: PT-JPL and SSEBop.

The unadjusted PT-JPL ET product slightly overestimates ET in much of the lower elevation regions of the state, typically by less than 50% (Figure 2). This nonetheless leads to steadily increasing depletions that results in maximum depletions in excess of 1000 mm (Figure 3C). But in mountainous areas, the unadjusted PT-JPL ET is less than PRISM precipitation, leading to seasonal cycles in depletion and 'runoff' (depletion values that would have become negative had they not been reset to zero according to the method) that might be expected in these areas, especially during snowmelt. These values might reflect reality.

When the PT-JPL ET is adjusted to match total PRISM precipitation, however, these mountainous areas no longer produce runoff. Rather, due to the proportionally increased ET, they have larger seasonal swings in cumulative depletion (high depletion at the end of the summer season, low depletion at the end of winter and spring melt). This leads to larger maximum depletions than in the unadjusted version (Figure 3C and 3D), approaching and exceeding 1000 mm. Yet in the low-elevation areas, the linear adjustment of PT-JPL ET to match precipitation leads to very reasonable depletion cycle range and variability, which in turn produces plausible SWHC estimates. This is especially true in the eastern plains, where the ET_{PT_JPL}/P_{PRISM} ratio was between 0.8-1.2.

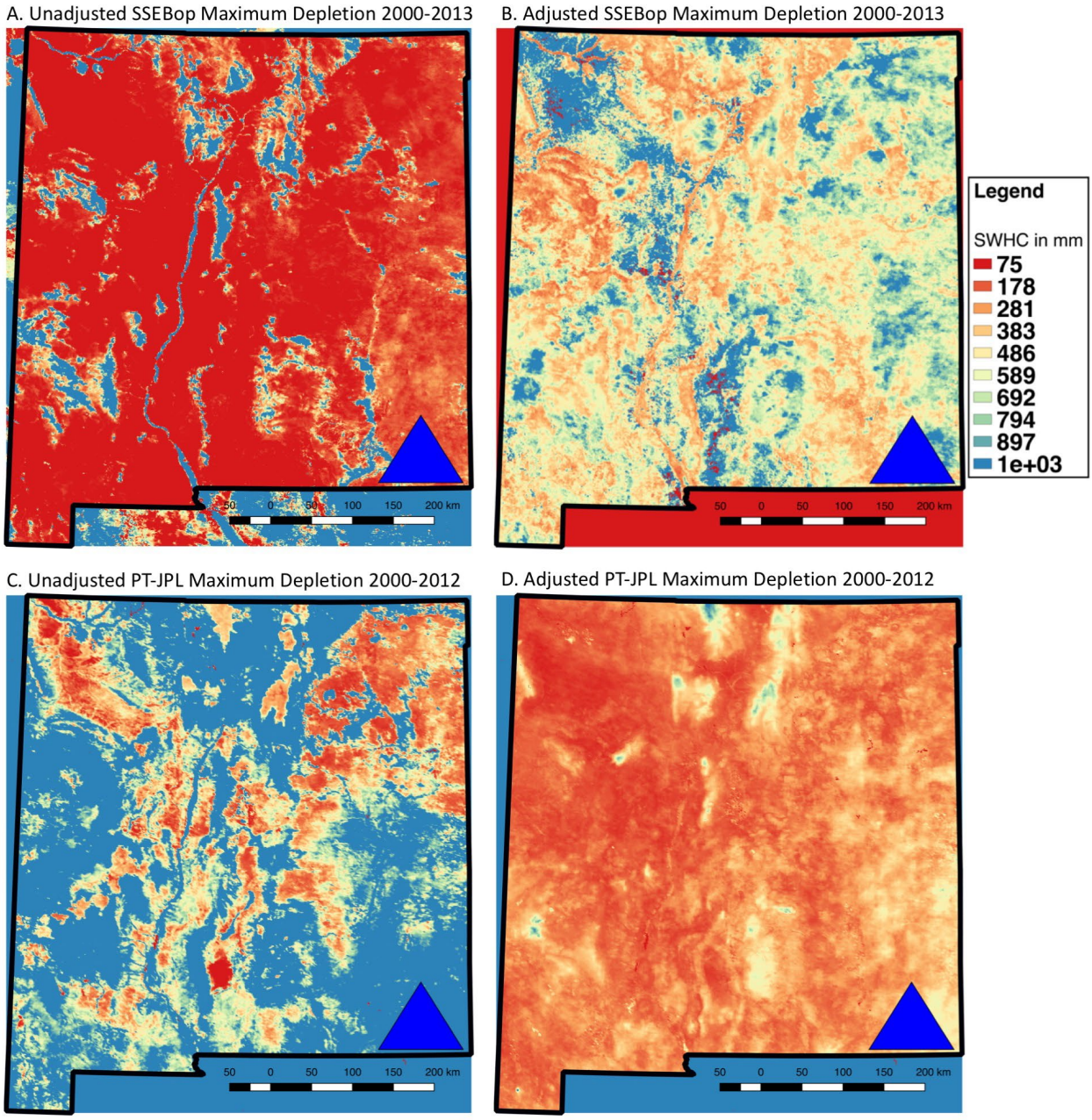


Figure 3. Maximum soil-water depletions after tracking precipitation inputs with PRISM and ET outputs with A) unadjusted SSEBop, B) SSEBop adjusted to match total PRISM precipitation, C) unadjusted PT-JPL, and D) PT-JPL adjusted to match total PRISM precipitation.

Recursive depletion tracking to estimate SWHC

The recursive depletion tracking method resulted in a SWHC map (Figure 4) that broadly agrees with the NRCS soils database SWHC map (Figure 1 and 4B). However, the precise locations of relatively high SWHCs are not perfectly consistent between the two. The recursive method finds high SWHCs in the mountainous areas, where large seasonal swings in depletion are estimated, as well as in the southern high plains. Note that because ET is limited by precipitation in the PyRANA calculations, there are no run-away depletions as in the SSEBop

and PT-JPL cumulative depletions. Rather, the range of depletions are a function of: 1) the seasonality of precipitation and ET (more winter precipitation leads to larger swings and larger SWHC estimates in order for the vegetation, mostly forests, to stay alive), and 2) the abundance of green vegetation (more vegetation leads to higher K_{cb} estimates and greater ET, assuming the soil-water is available). Hence, mountains with forests produce the largest SWHCs using this method. Irrigated and riparian areas do not stand out because the only water input in PyRANA is rainfall, and so the ET in these areas is capped in a way that does not represent the reality of water imports.

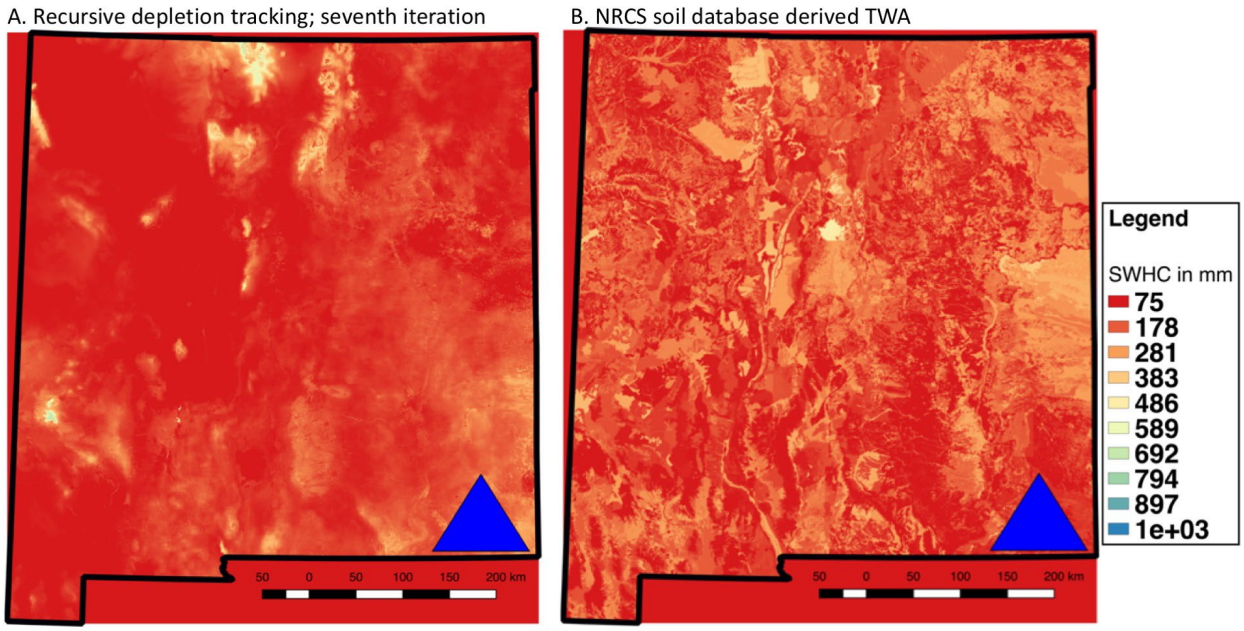


Figure 4. A) Range of soil-water depletions in the seventh recursive depletion tracking iteration using PRISM for water influx and PyRANA ET for water outflux. B) For comparison, the NRCS soil database SWHC map is presented at right, using the same color ramp. The seventh iteration is our preferred SWHC map for New Mexico using the recursive Wang-Erlandsson depletion tracking method, though future research and SWHC evaluation may prove that another number of iterations is more consistently reliable.

Vegetation-based SWHC estimation

The vegetation-based estimate of SWHC tended to produce larger values than the NRCS soil database or the recursive Wang-Erlandsson approach (Figure 5). Only the unadjusted PT-JPL depletion tracking and the adjusted SSEBop depletion tracking produced higher average values of SWHC. It captured the high effective SWHC of riparian and agricultural lands, and also estimated higher SWHCs on the eastern plains, where there are grasslands and shrublands with relatively high average NDVI values for the state. The low SWHC values in the San Juan basin, Rio Grande Rift, and southwestern lowlands of New Mexico reflect the relatively low average NDVI values of these areas. Similarly, the high SWHC values in the Sangre de Cristo, Gila, and Sacramento mountains (really all mountainous areas) reflect high NDVI of these forests compared to the same communities in the foothills, as well as the high estimated rooting depth of piñon-juniper (4 m) and ponderosa pine (3 m). The fringe of piñon-juniper around most mountain ranges is visible as a high SWHC area due to our assignment of the highest rooting depth to this vegetation community. Clearly, the parameters selected for use in this method

significantly affect the results, and further refinement and calibration of field measurements is required.

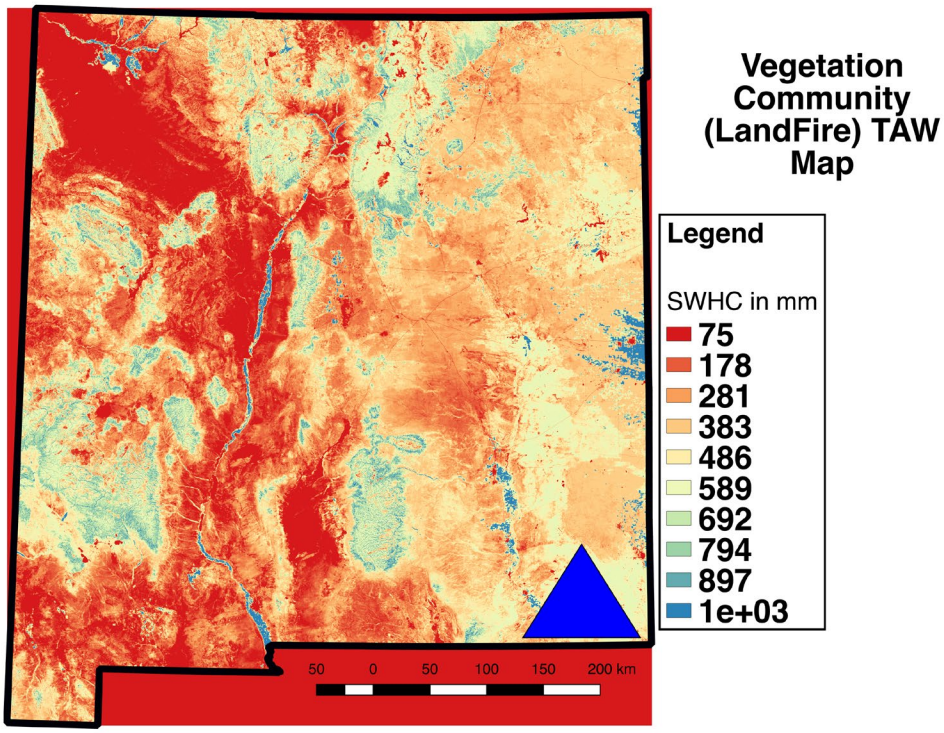


Figure 5. Estimated SWHC across New Mexico using the vegetation community and NDVI estimation method.

Discussion

This attempt to use depletion tracking to estimate SWHC has revealed weaknesses in the SSEBop estimates of ET in native vegetation land cover in New Mexico (Figure 2). The PT-JPL product performed better in all natural land covers, yet was still inadequate for use in the depletion tracking method for most areas. The exception appears to be mountainous areas, where unadjusted PT-JPL depletion tracking gave reasonable values, which were quite similar to those produced by the recursive depletion tracking using PyRANA ET. When PT-JPL was adjusted to match PRISM, the lowland SWHC estimates (mostly grassland and shrubland) became more spatially consistent (i.e., smoothed) and also more consistent with the soils database and recursive methods for these areas. In spite of the improvement in the lowlands, the SWHC estimates in the mountains were not improved by the adjustment to match PRISM. Rather, the estimates grew larger than expected, especially in the high alpine areas. Because most aquifer recharge occurs in these areas, an overestimation of SWHC in these regions would be especially problematic for accurate recharge modeling, and would lead to an underestimated recharge.

Though the PT-JPL depletion-tracking SWHC maps each have issues, they are an improvement over the SSEBop maps, which do not appear to be usable for this application. The strong overprediction of ET in the forested mountains (up to two times the long-term precipitation) and the strong underprediction in the low-elevation shrublands (the vast majority of days with zero estimated ET) may be due to the calibration of SSEBop in agricultural areas as well as more humid portions of the United States (Senay et al., 2013). Many of the areas where

both SSEBop and PT-JPL appear to overpredict ET is in shadowy canyons and slopes of mountains. This suggests that the topographic radiation corrections used in GADGET are vital for predicting ET and recharge in mountainous terrain. Though SSEBop may do well in riparian and agricultural areas, where PyRANA does not function, it was not useful for the purposes of this research project.

The use of PyRANA ET in a recursive manner to conduct depletion tracking was attempted because our application was clearly inappropriate for PT-JPL and SSEBop. However, in spite of the presumed increased ET accuracy of this approach, there are potential conceptual problems with using the model to calibrate itself. If we continue to repeatedly select the next iteration of SWHC based on the range of used soil storage, it will shrink until both complete drying and recharge occur. However, due to the stress coefficient, K_s, and its declining value as the depletion approaches the full SWHC, it is nearly impossible to completely dry the soil. Hence, the SWHC should continue to shrink by a small amount with each iteration. Yet this does not always occur in practice, primarily because of our method of ignoring runoff and recharge in our depletion tracking. Because infiltrated water can be less than precipitation during any particular day in the PyRANA model run, and in most cases really is slightly smaller, ET is able to deplete the soil fully. This full depletion and full filling lead to stabilization in the iterative SWHC.

This ignoring of runoff and recharge in the depletion tracking is a key weakness of this work. Yet we have no choice, because for the PT-JPL and SSEBop depletion tracking, there is no way to estimate these fluxes on a monthly basis. The recursive method could be modified to take them into account using PyRANA, but then it would produce an always declining SWHC estimate with each iteration. This perpetual decline could be addressed by picking an iteration to stop on, but we do not currently have adequate field estimates of SWHC to decide which iteration is best.

Of the methods explored in this report, we are most optimistic about the vegetation-based estimate, though it too requires more calibration data. The method produces plausible values, and is controlled by what we know to be the main control on SWHC – the vegetation itself. The method is easily adjustable. However, it is dependent on good vegetation community classification data. LandFire is not intended to be used at the 30 m scale at which it is produced, but rather at the small watershed scale or larger, as it is for use in fire behavior models (Rollins and Frame, 2006).

Furthermore, there is a challenge in defining SWHC in arid rangelands, especially with creosote. These shrubs pull such a great tension on the soil, that they have effectively dried out the sediment below them for many meters over the past several thousand years (Sandvig and Phillips, 2006). Thus, the volume of storage that must be filled before deep percolation can occur is extremely large, generally in excess of several years of precipitation. However, the amount of storage that is actually used, even in the wettest season, is quite small within our decade of data, and probably beyond. As a result, depletion tracking can never adequately represent this SWHC because: 1) it was depleted over time periods much greater than our analysis, and 2) the creosote does not currently use it, which breaks the assumptions of the Wang-Erlansson depletion-tracking method. Even our vegetation-based method underestimates SWHC in these areas,

because they have low NDVI compared to other shrublands throughout the state, so that they fall close to the lower NDVI threshold, which is based on the 5th percentile of NDVI for that community. This could be corrected by separating out creosote from other shrublands into its own community with its own characteristic rooting depth, or by adjusting the lower NDVI threshold. Future work is planned to pursue these avenues of research.

Future work

In addition to improving the vegetation-based SWHC estimation method, we are also exploring the use of SWHC optimization based on repeated PyRANA runs using a range of SWHC values. This method of optimization has already been explored at the point scale using eddy correlation flux tower data (Parrish, 2020). This process involves running PyRANA with a SWHC of, for example, 25 mm, 50 mm, 75 mm, and so on through 1000 mm. Then the output ET time series is compared against the flux tower ET time series. The SWHC that leads to the smallest disagreement (e.g., chi squared error) is the best SWHC. Future work will explore optimization of the model against spatially variable ET estimates from the mapping evapotranspiration with internalized calibration (METRIC) ET estimation method using Landsat imagery (Allen et al., 2007). METRIC ET estimates are labor intensive to produce at present, especially in mountainous areas where topographic corrections are needed. So as an intermediate step, we will explore the use of PT-JPL as a calibration data set. Although we know that it does not strictly agree with PRISM, the bias may be systematic enough that minimization of the ET disagreement will still result in the most accurate SWHC.

Chapter II: Rainfall-Runoff Relationships

Introduction

Modeling runoff based on rainfall has a long history (see Bevin, 2012), but typically this exercise takes place on a watershed-by-watershed basis. For PyRANA, we seek a method to estimate runoff from the land surface of a 250 m grid cell. Our approach is to base this on the foundation of plot-scale experiments, from which we know that both rainfall characteristics and soil characteristics play a role. Development of a multi-parameter equation to predict runoff based on both rainfall and soil would require calibration sites across the full range of soil types and hydroclimate regimes. Unfortunately, runoff data for the small, ephemeral streams necessary for this approach are very rare. The best data for this purpose are from the Walnut Gulch Experimental Watershed (WGEW) in southeast Arizona, which is located in the foothills of the San Pedro watershed, and on which both rainfall and runoff data have been collected for over 50 years. The WGEW is characterized by low-relief topography that is covered by shrubs and grasses with elevations ranging between 4,000 and 5,000 feet above sea level. The desert soils and monsoon-dominated rainfall in WGEW are typical of low elevations across southern and central New Mexico. Therefore, we elected to develop a rainfall-based equation to predict runoff, and assumed that soil properties are of secondary influence in areas where Hortonian overland flow dominates. Esther Xu developed the statistical precipitation-runoff correlations for WGEW, and developed the code to apply them in PyRANA (Xu, 2018). This approach was described in the Overview and Introduction section of this report.

Although this rainfall-runoff regression analysis at WGEW was successful, the specific coefficients for the regressions that relate precipitation to runoff are likely different at higher elevations in mountainous areas, where most recharge occurs. In part, this is because the runoff processes are different. While Hortonian overland flow is dominant in the desert rangelands, throughflow may be more common in forested areas (e.g., Dunne and Black, 1970), and this may lead to different rainfall-runoff relationships. In fact, the greater infiltration capacity of the forest soils may preclude overland flow, unless factors such as hydrophobicity are prevalent. To investigate this possibility, and to develop runoff equations specific to this contrasting landscape, statistical analyses similar to those done by Xu (2018) were conducted at three high-elevation, steep, forested catchments in northern New Mexico. These catchments represent the only adequately detailed paired rainfall and runoff measurements we could find for small (< 2 km^2) watersheds in New Mexico. Two basins are within the Santa Fe municipal watershed, and are tributary to the Santa Fe River just upstream of McClure Reservoir. They were studied by Amy C Lewis as part of a paired watershed study (Lewis, 2018) to investigate the water-yield effects of forest thinning and prescribed burns. The third basin is the La Jara Creek watershed in Valles Caldera National Preserve. Climate data were collected as part of the Critical Zone Observatory network, and runoff data were collected by an initiative of the National Park Service scientific staff at the Preserve.

The scarcity of multi-year stream-gauging data from small watersheds makes it unlikely that this technique can be extended to include soil properties in predicting runoff. But given the diversity of these three watersheds, we can begin to explore soil and physiographic factors that may be important, and to evaluate if they are of the same order of importance as rainfall intensity and total amount.

In addition to developing an empirical rainfall-runoff relationship, Xu (2018) observed that PRISM rainfall data did not perfectly agree with rain gauge data in Walnut Gulch. There was a single rain gauge near the town of Tombstone that was often used as a training gauge in the PRISM regression algorithm, but the gauges that we analyzed were never used as input gauges for the PRISM product. At these locations, PRISM fairly accurately estimated annual precipitation, but on a daily basis underestimated the largest rainfall events and generated many days of small rainfall for which the gauges showed no precipitation. This is consistent with PRISM effectively 'smearing' or spreading rainfall through time and space, as would be expected of such an interpolation-and-regression algorithm when gauges in the network are relatively sparse and convective storm size is much smaller than the gauge spacing. This is a problem because the rainfall-runoff regression showed a non-linear increase in runoff with rainfall amount, thus implying that missing the largest events and replacing them with small events would lead to an underestimation of runoff. To fix this issue, Xu proposed a linear adjustment to PRISM, which shifts the probability distribution of PRISM rainfall to match the probability distribution of the gauge record. Two such adjustments were developed, one for the monsoon season (July-September) and one for the rest of the year.

This PRISM adjustment was larger in the monsoon season, when precipitation is dominated by small convective storm cells that may hit one location with heavy rain while leaving a location only a few kilometers away completely dry. This short-spatial-scale variability in rainfall is especially hard for an interpolation product to describe accurately. During the winter, when larger frontal systems dominate precipitation, the difference between PRISM and the rain gauges at WGEW was negligible. Following Xu's work, these two PRISM adjustments (monsoon and non-monsoon) were applied throughout the state of New Mexico. However, it is not clear that PRISM needs to be adjusted in the more mountainous portions of the state. Monsoon precipitation is still a major source of moisture, but the spatial heterogeneity may be reduced enough to enable PRISM to provide an adequate representation. Therefore, a key goal of this research was to compare PRISM rainfall to daily gauge rainfall at the three study watersheds, as well as at the long-term precipitation gauge at Taos, NM.

Methods

Taos rainfall data

Rainfall data for the 2003-2010 time period from the Taos Village rain-gauge was downloaded from the NOAA National Weather Service website, and then compared to PRISM rainfall data for the grid cell that contained the Taos Village gauge. The monsoon season (July to September) and the non-monsoon season were analyzed separately. The analysis focused on the monsoon season data, because disagreements between PRISM and rain gauges had only previously been documented during this period of highly localized convective precipitation (Xu, 2018). Due to one-day event timing disagreements between PRISM and the gauges (most probably related to PRISM's practice of dividing days at noon UTC time), we then sorted both data sets from highest precipitation to lowest. We plotted the ordered gauge data on the x-axis and the ordered PRISM data on the y-axis. A simple linear regression was used to determine the relationship between PRISM and the gauge for the days on which rain occurred.

Santa Fe rainfall and runoff data

These data were collected from two small ephemeral watersheds that drain into the Santa Fe River (Lewis, 2018). Lewis (2018) placed three rain gauges on the ridge between the two basins she was studying and two flumes on the ends of each ephemeral stream. In this analysis, we only used the data she collected during the monsoon season, because the processes affecting runoff during winter and spring are much different due to snow and snowmelt. Due to tree cover and data gaps at the upper gauge, only the middle and lower gauges were deemed usable. The data from the two gauges were initially averaged and compared to PRISM, but there were some days when one gauge had significant rainfall and the other had none, and so averaging resulted in the storm appearing half as small as it probably was. After the two basins were analyzed separately, we found that the middle gauge had significantly lower total rainfall and more days with zero rainfall. Since it is more likely that a tipping bucket rain gauge will miss rain (jamming, clogging) than produce spurious rain (manual bucket tips), and considering that the lower gauge was located in the most open space, it was used for the analysis that followed. The relationship between gauge data and PRISM data was determined using the same method as described in the previous section.

Rainfall data were recorded every hour and runoff data were recorded every 15 minutes. Initially we attempted to sum up all rainfall and runoff for each individual day, but no relationship was visible between the two because the duration of rainfall and runoff events often exceed one day and there can be a large lag between initiation of rainfall and initiation of runoff. We elected to create and analyze hydrographs and hyetographs (rainfall time series) in tandem. From this we could clearly see that not all rainfall that causes runoff occurs the same day the runoff is generated. We manually identified associated rainfall and runoff events using a set of assumptions: 1) All rainfall that happened within 24 hours of the beginning of runoff contributed to the runoff generation process; 2) Any rainfall that occurred more than an hour after runoff stopped could not be associated with the runoff event preceding it; 3) Any runoff that occurred without additional rainfall was attributed to the same rainfall event that initiated the runoff; and 4) Rainfall intensity was calculated as the ratio of the total depth of rain that fell divided by the number of hours in which it fell, excluding hours without rain.

The two watersheds analyzed were part of a forest management experiment. One basin, which we designated the Treated basin, was mechanically thinned and had low-severity fire reintroduced. The other basin, which we designated the Control basin, was not thinned or burned. However, this was not the only difference between the basins. The Control basin was much steeper, with a relief ratio of 0.33 compared to 0.23 for the Treated basin. (Relief ratio is the maximum basin relief divided by the longest horizontal distance of the basin parallel to the main stream.) And while both basins had a geology dominated by granitic and gneissic basement rock, the Treated basin also has the Borrego fault zone running down its length, while the Control basin has large talus slopes that cover ~ 7% of its surface area. Finally, the Treated basin is very slightly larger, at 1.79 km^2, compared to 1.53 km^2 for the control basin.

La Jara rainfall and runoff data

The La Jara stream is much different than the streams in Santa Fe, because base flow is a significant contribution to total runoff, whereas the streams in Santa Fe are ephemeral. In addition to being perennial, La Jara drains Bandolier Tuff that was subsequently uplifted into the

resurgent dome of Redondo Peak in Valles Caldera National Preserve. The watershed is larger than those in the Santa Fe basin (3.57 km^2) and slightly less steep (relief ratio = 0.19). Because the water table is higher in relation to the stream bed in the La Jara basin, there was never a time when the stream bed was dry, complicating efforts to associate runoff with a particular rainfall event. To deal with this, the initial baseflow prior to a rainfall event was averaged with the final baseflow and then subtracted from each runoff measurement before the total amount of runoff was summed. This effort to separate baseflow is crude, but simple and easy to apply consistently. The baseflow after the runoff event had ended was difficult to determine in some cases because some events raised the water table, thus raising the post-event baseflow. Therefore, if there was not a clear return to the base flow present before the rainfall event, then the post-event baseflow was determined to be the discharge at which runoff leveled out and the cause of variation in stream stage was due to diurnal cycles, in which phreatophytes lower the water table during the day. It is difficult to distinguish exactly between baseflow and direct runoff purely by using a streamflow hydrograph, but the same methods were used by the same operator in an effort to remain consistent. Some runoff events did not have a chance to return to baseflow before another rain event began and so the interrupted events were combined with the events that followed into one large event spanning numerous days.

Results

PRISM versus gauge rainfall

At both the Santa Fe watershed rain gauge and the Taos rain gauge, PRISM had more days with reported rainfall than the gauges (Figures 6 and 7). But for the large events, the depth of rain reported by the gauges was greater than that of PRISM (Figures 6 and 7). This is consistent with previous research using the rain gauges at Walnut Gulch [Xu, 2018]. The two data sets are also consistent with each other in the relationship between gauge rainfall and PRISM rainfall that they predict. At Taos, PRISM = 0.73(Gauge) + 0.99, while at Santa Fe, PRISM = 0.74(Gauge) + 1.17. In summary, these equations suggest that any PRISM rainfall event that is less than 1 mm is probably a false event, and that PRISM rainfall estimates are on average ¾ of the locally measured rainfall, especially for the large events. Yet because of the many additional small events, the long-term total PRISM is consistent with the gauges. The biases appear consistent and easily correctable.

Rainfall-runoff relationships

The three watersheds had consistent behaviors in runoff production, and the trend between rainfall and runoff was quite linear after log-transforming the runoff data (Figure 8). Basin runoff, in mm, is predicted by $0.01e^{0.058P}$, where P is rainfall in mm. The Treated basin in Santa Fe produced less runoff than the other two basins, consistent with its lower steepness than the Control basin in Santa Fe. If topography were the only control, we would then expect La Jara to produce even less runoff. However, it is a perennial stream, and the response time to rainfall events was almost instantaneous (less than the one-hour resolution of our data). In contrast, both of the Santa Fe watersheds typically saw lags of five hours to a day between the first rainfall and initiation of flow at the basin outlet. The higher water table at La Jara means that less storage must be filled before flow occurs, and for small rainfall events the runoff is particularly high at this site (Figure 8). As rainfall events get larger, the enhanced runoff at La Jara is no longer evident, at least compared to the Control basin. Similarly, as rainfall event size increases, the difference between the Control and Treated basin runoff responses becomes less pronounced.

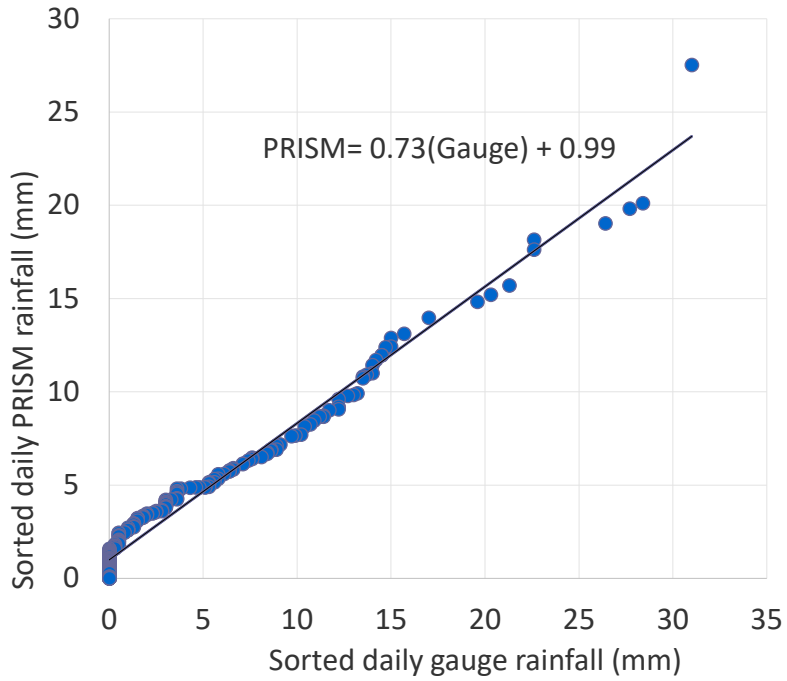


Figure 6. Daily PRISM rainfall vs. daily gauge rainfall for the Taos Village rain gauge for only the months July-September (monsoon season) in the years 2003-2010. Both data sets were sorted before being plotted against each other, to account for timing inconsistencies between the two.

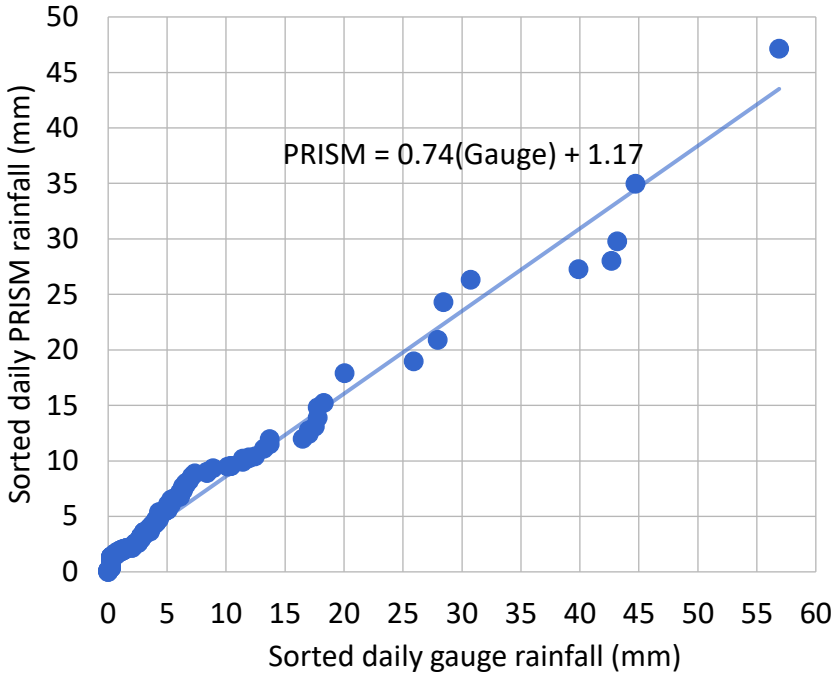


Figure 7. Daily PRISM rainfall vs. daily gauge rainfall for the lower Santa Fe paired basins rain gauge for only the months July-September (monsoon season) in the years 2009-2017. Both data sets were sorted before being plotted against each other, to account for timing inconsistencies between the two.

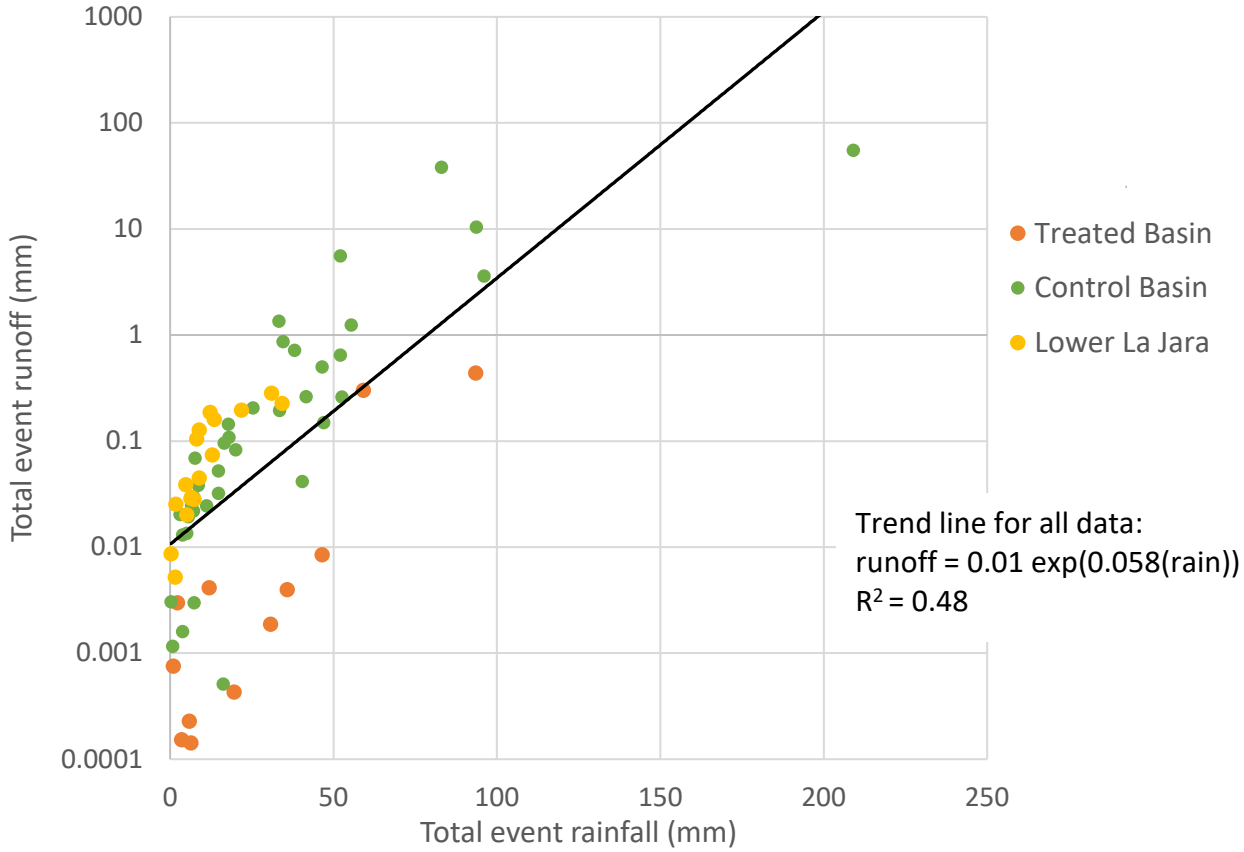


Figure 8. Rainfall-runoff relationships for three small watersheds in mountainous terrains of north-central New Mexico. Runoff is reported as a depth, that is, volume of runoff divided by basin area. The trend line (solid black) is a regression for all three data sets combined.

Discussion

There was reasonably good consistency in runoff production among the three small mountainous watersheds studied here. For small rainfall events there was about two orders of magnitude range of runoff production, declining to one order of magnitude for larger rainfall events (> 50 mm). Given the major differences among the sites, we are satisfied with this relationship.

The basin characteristics that seem most important in promoting higher runoff are: bare rock (or talus), steep slopes, and perennial stream flow (which, based on the La Jara results, can overcome the effects of low slope and relatively permeable bedrock). The factors that reduce runoff are low slope and thick soils. Additionally, the Treated basin in Santa Fe was underlain by a fault system, which may have preferentially transported water out of the channel and into the regional aquifer.

The PRISM-gauge correction derived from the Walnut Gulch rain gauges is not dramatically different from those found here. In Walnut Gulch, the PRISM-to-gauge correction equation was Gauge = (PRISM – 1.61)/0.68, whereas here the rough average was Gauge = (PRISM – 1)/0.74. This suggests that PRISM does not underestimate large monsoon events by

quite as much in northern New Mexico compared to the south, but it is still a significant effect. Because it is so easily corrected, this seems to be prudent to implement in PyRANA.

The runoff relationship in Walnut Gulch was found to be highly dependent on rainfall intensity, consistent with a Hortonian overland flow conceptual model for runoff generation. However, in the mountainous basins, rainfall intensity was not a valuable predictor of runoff. Simply rainfall alone had nearly all the predictive power. This is consistent with runoff production being caused by processes other than Hortonian overland flow, such as throughflow and saturation overland flow. Given the relatively thick soils in the ponderosa and mixed conifer forests in the studied basins, especially when compared to the rangeland soils of southeast Arizona and Walnut Gulch, this finding makes good hydrologic sense.

Future Work

This new PRISM correction and rainfall-runoff relationship will be incorporated into PyRANA, to be applied in the northern and mountainous portions of the state. We will continue to use the Walnut Gulch derived relationships (Xu, 2018) for the southern and low-elevation portions of the state.

REFERENCES

Abatzoglou, J.T. (2013), Development of gridded surface meteorological data for ecological applications and modelling, *International Journal of Climatology*, 33(1), 121-131.

Allen, R.G. (2011), Skin layer evaporation to account for small precipitation events—An enhancement to the FAO-56 evaporation model, *Agric. Water Manage.*, 99(1), 8-18.

Allen, R.G., L.S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration: Guidelines for computing crop requirements. Irrigation and Drainage Paper No. 56., 300 pp., Food and Agricultural Organization of the United Nations, Rome, Italy.

Allen, R.G., L.S. Pereira, M. Smith, D. Raes, and J.L. Wright (2005a), FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions, *Journal of Irrigation and Drainage Engineering*, 131(1), 2-13.

Allen, R.G., W.O. Pruitt, D. Raes, M. Smith, and L.S. Pereira (2005b), Estimating evaporation from bare soil and the crop coefficient for the initial period using common soils information, *Journal of Irrigation and Drainage Engineering*, 131(1), 14-23.

Allen, R.G., M. Tasumi, and R. Trezza (2007), Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)-model, *Journal of Irrigation and Drainage Engineering*, 133, 380–394.

ASCE-EWRI (2005), The ASCE standardized reference evapotranspiration equation., 59 p. with six appendices, Environmental and Water Resources Institute, American Society of Civil Engineers, Reston, Virginia.

Beven, K.J. (2012), *Rainfall-Runoff Modelling: The Primer (2nd Edition)*, 457 pp., John Wiley & Sons, Ltd.

Budyko, M.I. (1956), *Heat balance of the earth's surface (in Russian)*, 255 pp., Gidrometeoizdat, Leningrad, USSR.

Cihlar, J. (2000), Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing 21*, 1093– 1114. doi:10.1080/014311600210092

Daly, C., and K. Bryant (2013), The PRISM Climate and Weather System - An Introduction. http://www.prism.oregonstate.edu/documents/PRISM_history_jun2013.pdf

Daly, C., M. Halbleib, J.I. Smith, W.P. Gibson, M.K. Doggett, G.H. Taylor, J. Curtis, and P.P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *International Journal of Climatology*, 28(15), 2031-2064.

Dunne, T., R. Black. (1970), An experimental investigation of runoff production in permeable soils, *Water Resources Res.*, 6(2), 478-490.

Fisher, J.B., K. Tu, and D.D. Baldocchi (2008), Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, *Remote Sens. Environ.*, 112(3), 901–919.

Gee, G.W., and D. Hillel (1988), Groundwater recharge in arid regions: review and critique of estimation methods, *Hydrol. Process.*, 2, 255-266.

Hendrickx, J.M.H., and G. Walker (1997), Recharge from precipitation, in *Recharge of phreatic aquifers in (semi)-arid areas*, edited by I. Simmers, pp. 19-114, Balkema, Rotterdam, The Netherlands.

Homer C, C. Huang, L. Yang, B. Wylie, and M. Coan. (2004). Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering and Remote Sensing,* 70, 829–840.

Jensen, M., and R. Allen (2016), *Evaporation, Evapotranspiration, and Irrigation Water Requirements*. Second Edition. ASCE Manual and Report No. 70, 554 pp., American Society of Civil Engineers.

Kearns, A.K., and J.M.H. Hendrickx (1998), *Temporal variability of diffuse groundwater recharge in New Mexico*, 43 pp., New Mexico Water Resources Research Institute Technical Completion Report No. 309, New Mexico State University, Las Cruces NM.

Ketchum, D.G. (2016), *High-Resolution Estimation of Groundwater Recharge for the Entire State of New Mexico Using a Soil-Water Balance Model*. Earth and Environmental Science Department, New Mexico Tech, Unpublished thesis, 125pp.

Lewis, A. (2018), Monitoring Effects of Wildfire Mitigation Treatments on Water Budget Components: A Paired Basin Study in the Santa Fe Watershed, New Mexico. New Mexico Bureau of Geology and Mineral Resources Bulletin-163, 52 pp.

Manabe, S. (1969), The atmospheric circulation and the hydrology of the earth's surface, *Monthly Weather Review*, 97(11), 739-774.

Parrish, G.E.L., J.M.H. Hendrickx, D. Cadol, B.T. Newton, (2017) Parametrization of Total Available Water (TAW) for Statewide Water Assessment in New Mexico. New Mexico Water Resources Research Institute Draft Final Progress Report, New Mexico State University, Las Cruces NM, https://nmwrri.nmsu.edu/wp-content/SWWA/Reports/Bawazir/Year3/technical_report_xxx_Parrish_Hendrickx_Cadol_Talon_5_jul_2017%20CM%20modified.pdf.

Parrish, G.E.L. (2020), *Parameterizing Total Available Water for New Mexico Soils*, Earth and Environmental Science Department, New Mexico Tech, Unpublished thesis, 150pp.

Rafn, E., B. Contor, and D. Ames (2008), Evaluation of a method for estimating irrigated crop-evapotranspiration coefficients from remotely sensed data in Idaho, *Journal of Irrigation and Drainage Engineering*, 134(6), 722-729.

ReVelle, P. (2017), *Evapotranspiration in mountain terrain – applying topographic-based energy constraints to evaluate the distribution of water fluxes and effect of vegetation cover change*. Earth and Environmental Science Department, New Mexico Tech, Unpublished thesis, 190pp.

Rinehart, A.J., E. Mamer, T. Kludt, B. Felix, C. Pokorny, and S. Timmons (2016), Groundwater Level and Storage Changes in Basin-Fill Aquifers in the Rio Grande Basin, New Mexico, 44 pp., New Mexico Water Resources Research Institute Final Progress Report, New Mexico State University, Las Cruces NM. https://nmwrri.nmsu.edu/wp-content/SWWA/Reports/Timmons/FinalReport/TimmonsNewtonGroundwaterlevelstoragechanges-RioGrandeRift-NM2016_LR.pdf.

Rinehart, A.J., and E. Mamer, (2017), Groundwater Storage Change in New Mexico Aquifers: Part 1. Method for Estimating Groundwater Storage Change in Variably Confined Aquifers in New Mexico. Part 2. Estimates for Groundwater Storage Change in the New Mexico High Plains Aquifer, 58 pp., New Mexico Water Resources Research Institute Technical Completion Report (in progress), New Mexico State University, Las Cruces NM.

Rinehart, A.J., S. Timmons, B. Felix, and C. Pokorny, (2015), Groundwater Level and Storage Changes – Regions of New Mexico, 42 pp., New Mexico Water Resources Research

Institute Final Progress Report, New Mexico State University, Las Cruces NM, https://nmwrri.nmsu.edu/wp-content/SWWA/Reports/Timmons/June%202015%20FINAL/report.pdf.

Rollins, M.G., C.K. Frame, tech. eds. (2006), *The LANDFIRE Prototype Project: Nationally Consistent and Locally Relevant Geospatial Data for Wildland Fire Management*. Gen. Tech. Rep. RMRS-GTR-175. Fort Collins: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 416 pp.

Rollins, M.G. (2009), LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire,* 18, 235-249.

Sandvig, R.M., F.M. Phillips. (2006), Ecohydrological controls on soil moisture fluxes in arid to semiarid vadose zones. *Water Resources Research,* 42, W08422. doi:10.1029/2005WR004644

Schmugge, T. J., S. Walker, and J.M.H. Hendrickx, (2015), Mapping statewide precipitation and evapotranspiration in New Mexico: Year One, New Mexico Water Resources Research Institute Final Progress Report, Las Cruces, NM, https://nmwrri.nmsu.edu/wp-content/SWWA/Reports/Hendrickx/June2015FINAL/Hendrickstechnical_report_Year_One_22_jul_2016_finalcmMODIFIED.pdf.

Senay, G.B., M.E. Budde, and J.P. Verdin, (2011), Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET: Validation with the METRIC model. *Agricultural Water Management*, 98(4), 606-618. doi: 10.1016/j.agwat.2010.10.014.

Senay, G.B., S. Bohms, R.K. Singh, P.H. Gowda, N.M. Velpuri, H. Alemu, and J.P. Verdin, (2013), Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *Journal of the American Water Resources Association*, 49(3), 577-591. doi: 10.1111/jawr.12057.

Wang-Erlandsson, L., W.G.M. Bastiaanssen, H. Gao, J. Jagermeyer, G.B. Senay, A.I.J.M. van Dijk, J.P. Guerschman, P.W. Keys, L.J. Gordon, and J.J.G. Savenije, (2016), Global root zone capacity from satellite-based evaporation: *Hydrol. Earth Syst. Sci*., 20, 1459-1481.

Xu, F. (2018), *Focused Recharge Estimation of New Mexico Using a Soil-Water-Balance Model PyRANA*. Earth and Environmental Science Department, New Mexico Tech, Unpublished thesis, 90pp.