**Regression relationship of groundwater irrigation water use to precipitation and ET for Rattlesnake-Zenith depletion response zone A within GMD5 groundwater model grid, with supporting documentation in Appendix**

Sam Perkins, David Barfield, Chris Beightel, David Engelhaupt, Jeff Lanterman and Ginger Pugh

KDA-DWR, Jan 9, 2018

In response to a request by GMD 5 made at a meeting on October 30, 2017, KDA-DWR has developed a regression model to estimate groundwater irrigation pumping as a function of reference evapotranspiration (ET) and precipitation (P), effectively the crop irrigation requirement (CIR).

The focus of this work is on the area where GMD5 groundwater model simulations show a 10 pct or greater depletion response in Rattlesnake C streamflow at the Zenith gage (Zone A). For this area, regression relationships have been found with r-squared values of 95 percent, based on 2000-2016 data. Similar relationships with high r-squared values have been found for pumping in GMD3 and GMD4, although these are not discussed here. These relationships may help implement and evaluate agreed upon pumping reductions by taking into account the variation of pumping demand with ET and P.

**Introduction**

The regression model has the form f(ET,P) = c1ET + c2P, which expresses expected pumping (f) either as a volume (acre-feet) or a volume per unit area irrigated (inches). Groundwater irrigation data are provided by WRIS database queries for years 2000-2016. ET and P are based on monthly spatial datasets of precipitation and temperature provided by Oregon State University’s PRISM Climate Group and downloaded from <http://www.prism.oregonstate.edu/>. These are “the most widely used spatial climate data sets in the United States,” according to “The PRISM Climate and Weather System – An Introduction” by Daly and Bryant (2013). The temperature data (average minimum and maximum daily temperature for each month, tmin and tmax) are used to calculate Hargreaves and Samani (1985, H-S) reference ET as outlined by Snyder and Eching (2005) for monthly time steps. The H-S method performs well among reduced-dataset methods of calculating ET in comparative studies by Xu and Singh (2002) and by Shahidian et al. (2012).

The above regression model can be compared with how CIR is calculated annually for the Republican River Compact (McKusick, 2003) as follows. The coefficient c1 corresponds to the product of two factors, denoted cp and cc, where cp is a calibration coefficient for each month that adjusts H-S ET for a short reference grass to match Penman ET calculated for a tall grass at specific weather stations. cc is a crop coefficient that varies with growing degree days over the year and is calculated according to Part B of Food and Agriculture Organization of the United Nations Guidelines for computing crop water requirements (FAO-56) by Allen et al. (1998). The second coefficient, c2, is negative and corresponds to the infiltrating fraction of rainfall. c2 is represented by a polynomial that varies with monthly rainfall. The second term in the above equation represents the reduction of crop irrigation requirement by effective rainfall.

**Methods and procedures**

Annual water use queries of the WRIS database were developed and run by Ginger Pugh and David Engelhaupt for GMD areas 3, 4 and 5 for 2000-2016. The queries included reported pumping and irrigated area for both active and inactive water rights over this period. Initial regression analysis was based on a query that was run on 4/18/2017 and subsequently updated on 12/6/2017. Regressions that have been updated show negligible change associated with the water use query update. A comparison of queries for Zone A and GMD5 is in the backup Excel file, sheet GMD5\_Summary\_2000-2016\_1206, range N1:X18 for years 2000-2016.

Monthly PRISM datasets were downloaded for precipitation and temperature for years 1999-2017. Included in zipfiles of PRISM data for each year are ASCII text files (with file extension .asc) for a grid of 4 km cells for the contiguous United States. Temperature data include daily minimum and maximum temperature averaged over each month (tmin and tmax). Fortran programs were developed to read these files and write grid files corresponding to local regions including GMD3, GMD5 and RRCA groundwater model domains. At this time, PRISM data for 2017 is stable for Jan-Apr, provisional for May-Oct, and represented in Nov-Dec by current monthly normal (1981-2010). For each month, the conversion programs look first for a file with stable data, then for provisional data if a stable version is not available, and then for a file with normal data if neither a stable or provisional version is available.

Evapotranspiration is calculated as reference ET according to Hargreaves-Samani (1985) as outlined by Snyder and Eching (2005) for monthly time steps, based on the PRISM monthly tmin and tmax datasets.

After downloading PRISM datasets, program PRISM\_to\_grid\_mo was used to map PRISM data (precipitation, tmin and tmax) onto corresponding GMD5 groundwater model grid cells. A variation on this program (PRISM\_to\_grid\_mo\_nper) was used to write summaries of annual and seasonal irrigation for each year 2000-2016 and each model grid cell to text files, and to optionally write a summary file of monthly data spatially averaged over a specified zone. A third program, ETref\_HS, was used to calculate reference ET according to Hargreaves-Samani (1985). Variations on these programs were used to optionally write summary files of monthly precipitation and ET spatially averaged over a specified zone. Fig. 1 is a plot of precipitation and ET for calendar months averaged over years 1999-2017. Fig. 2 is a similar plot with months lumped into groups Oct-Feb, Mar-May, Jun-Jul and Aug-Sep.

Supporting documentation of methods and procedures used to develop the PRISM-based precipitation and temperature datasets and to calculate reference ET is in the Appendix.

**Regression analysis in Excel spreadsheets**

The following regression relationships of reported water use to reference evapotranspiration( ET) and precipitation (P) were examined for Zone A in GMD5 using Microsoft Excel:

* u=f1(P) for annual precipitation P;
* u=f2(P) for precipitation over the irrigation season (May-September);
* u=f3(ET,P) for reference ET and precipitation over the irrigation season (May-Sep).
* u=f4(ET) for reference ET over the irrigation season (May-September);
* u=f5(ETi,Pi), where ETi and Pi are reference ET and precipitation over the three periods March-May, June-July and August-September.

Single-variable relationships to annual and seasonal precipitation were tested initially by superimposing a linear trend on a plot of the data for Zone A as well as for GMD5 and individual counties. For multivariate regression, Excel’s regression analysis tool was used. R-squared values and standard deviation of error for some of these relationships for Zone A are listed in Table 1.

Table 1. Summary of regression relationships of water use to precipitation (P) and reference ET: r-squared values and standard error by volume (KAF) and volume per unit area irrigated (inches).

Regressed variables U=f(ET,P) for 2000-2016: r-squared std error, s.e. (KAF) s.e. (in.)

Water use (af) vs. annual P 0.75 14.1 KAF 1.05

Water use (af) vs. May-Sep P 0.80 13.1 KAF 0.93

Water use (af or in) vs. May-Sep ET and P 0.86 10.7 KAF 0.75

Water use (af or in) vs. May-Sep ET 0.76 14.3 KAF 1.00

Water use (af or in) vs. ET and P, grouped months\* 0.95 6.5 KAF 0.46

 (\*) grouping: March-May, June-July, August-September

Regressions for water use in acre-feet and inches (12\*ac-ft/acres) show very similar r-squared values, usually within 1 percent. For example, these are 0.863 and 0.867, respectively, for the relationship of water use to ET and P over the irrigation season.

Comparison of the first two regressions shows that restricting precipitation to the nominal irrigation season of May-September has the effect of increasing the r-squared value from 75 to 80 percent and reducing the standard error from 14 to 13 KAF. For the third regression, inclusion of May-Sep ET increases r-squared from 80 to 87 percent, and reduces the standard error from 13 to 11 KAF.

Comparison of the second and fourth regressions listed above, against either P or ET, shows that ET appears to work nearly as well as P for estimating water use. Fig. 3 is a plot of ET vs. P for the irrigation season and shows an inverse relationship between the two estimators, so that one serves roughly as a surrogate for the other. However, inclusion of both ET and P in the regression model is justified by the increased r-squared value, reduced standard error and significance of coefficients for the third regression listed. The correlation of independent variables in a model is referred to as collinearity; a discussion of collinearity and its consequences can be found at <https://onlinecourses.science.psu.edu/stats501/node/346> .

For the fifth regression listed in Table 1 and identified as f5, monthly ET and P were lumped into the four groups Mar-May, Jun-Jul and Aug-Sep and Oct-Feb (Fig. 2). The group of months Oct-Feb was excluded from the regressions examined on the basis of regressions involving only precipitation that showed little contribution to the relationships. The remaining datasets were treated as six independent climatic variables for the purpose of the regression analysis.

Table 2 lists the coefficients and P-values for two versions of the regression model. Columns on the left are for water use in units of inches; columns on the right are for water use in units of acre-feet. The P-value is the probability that the corresponding term in the regression model has no effect. For more explanation of the P-value, see Frost (2013).

For the regression model identified as f5 (above), the y-intercept, or constant coefficient was set to zero, and March-May ET was excluded because of high corresponding P-values in regression models that included them. The remaining predictors are all significant at the 0.05 level.

Fig. 4 plots estimated vs. reported use (ac-ft); estimated use is based on the coefficients on the righthand side of the above table for 2000-2016. Fig. 5 superimposes estimated use for 1999-2017 with reported use for 2000-2016. Estimated use appears to agree fairly uniformly with reported use over years 2000-2016.

Fig. 5 is a time series plot of reported irrigation (inches, 2000-2016) and estimated use as a function of precipitation and ET for the volumetric irrigation regression model with coefficients listed in Table 2.

Fig. 6 is the same as Fig. 5, but shown with precipitation and ET for the irrigation season, which reference the axis on the right.

Table 2. Coefficients and p-values for water use estimate based on water use query of 12/6/2017 vs. ET and P for March-May, Jun-Jul and Aug-Sep 2000-2016.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| GMD5 Zone A | gwIrr(in) |   | gwIrr(af) |   |
| *Predictors* | *Coeffs* | *P-value* | *Coeffs* | *P-value* |
| Constant | 0 | #N/A | 0 | #N/A |
| Jun-Jul.et | 0.655528 | 0.005791 | 9315.527 | 0.006235 |
| Aug-Sep.et | 1.051469 | 0.001215 | 14360.8 | 0.001775 |
| Mar-May.ppt | -0.17021 | 0.000353 | -2430.15 | 0.000375 |
| Jun-Jul.ppt | -0.34909 | 2.97E-05 | -4808.31 | 4.46E-05 |
| Aug-Sep.ppt | -0.1969 | 0.045007 | -2981.13 | 0.036358 |
|  | r-square | s.e. (in.) | r-square | s.e. (AF) |
|  | 0.9513 | 0.455 | 0.949 | 6,548 |

**Summary and conclusions**

Regression relationships of groundwater irrigation water use to ET and P for Zone A began with annual precipitation with r-squared value of 0.75 and a standard error of 14.1 KAF. Restricting precipitation to the nominal irrigation season May-September improved the r-squared value to 0.80 and reduced the standard error to 13.1 KAF. Including both ET and P for the irrigation season further improved the r-squared value to 0.86 and reduced the standard error to 10.7 KAF. By simply regrouping the monthly ET and P data into the periods of March-May, June-July and August-September, a water use estimator based on five of the resulting six independent variables was found to have an r-squared value of 0.95 with the standard error reduced to 6.5 KAF. Coefficients for this relationship are listed in Table 2, both as a volume of water (acre-feet) and volume per unit area irrigated (inches). This relationship may be useful as a tool for estimating water use within the range of ET and P for years 2000-2016.

**References**

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: 1998, *Crop Evapotranspiration – Guidelines for*

*Computing Crop Water Requirements*, Food and Agriculture Organization of the United Nations [FAO] Irrigation and Drainage Paper 56 (FAO-56), FAO, 1998, ISBN 92-5-104219-5. <http://www.fao.org/docrep/X0490E/X0490E00.htm>

Daly, Christopher and Kirk Bryant (June 2013). The PRISM Climate and Weather System – An Introduction. PDF file at <http://www.prism.oregonstate.edu/documents/PRISM_history_jun2013.pdf>

Frost, Jim, 2013. How to interpret regression analysis results: P-values and coefficients. The Minitab Blog, <http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients>

Hargreaves, George H. and Richard G. Allen, 2003. History and Evaluation of Hargreaves Evapotranspiration Equation. Journal of Irrigation and Drainage Engineering, Vol. 129, No. 1, February 1, 2003. <http://onlinecalc.sdsu.edu/onlinehargreaves.pdf>

Hargreaves, G.H. and Z.A. Samani, 1985. Reference crop evapotranspiration from temperature. Transaction of ASAE 1(2):96-99.

McKusick, Vincent L., 2003. Final Report of the Special Master with Certificate of Adoption of RRCA Groundwater Model, p. 25, Crop Irrigation Requirements. PDF file at <http://www.republicanrivercompact.org/v12p/SMReportOnGWModel.pdf>

Samani, Zohrab (undated). Estimating Solar Radiation and Evapotranspiration Using Minimum Climatological Data (Hargreaves-Samani equation). New Mexico State U., Las Cruces. PDF file at

<http://www.zohrabsamani.com/research_material/files/Hargreaves-samani.pdf>

Shahidian, Shakib, Ricardo Serralheiro, João Serrano, José Teixeira, Naim Haie and Francisco Santos, 2012. Hargreaves and Other Reduced-Set Methods for Calculating Evapotranspiration. Chapter 4 of 23 in Evapotranspiration - Remote Sensing and Modeling, Edited by [Ayse Irmak](https://www.intechopen.com/books/editor/evapotranspiration-remote-sensing-and-modeling), ISBN 978-953-307-808-3, 526 pages, Publisher: InTech, Chapters published January 18, 2012. <https://www.intechopen.com/books/evapotranspiration-remote-sensing-and-modeling>

Snyder, R.L. and S. Eching, 2002. Penman-Monteith daily (24-hour) and Hargreaves-Samani Equations for Estimating Reference Evapotranspiration from Monthly Data. UC-Davis.

Monthly reference ET: <http://biomet.ucdavis.edu/Evapotranspiration/PMmonXLS/PMmon.htm>

Daily reference ET: <http://biomet.ucdavis.edu/Evapotranspiration/PMdayXLS/PMday.htm>

Xu, C-Y and Singh, 2002. Cross Comparison of Empirical Equations for Calculating Potential Evapotranspiration with Data from Switzerland. *Water Resources Management* **16:** 197–219, 2002. Downloaded from <http://folk.uio.no/chongyux/papers_SCI/WARM_8.pdf>

**Figures**

Figs. 1-6 are from the Microsoft Excel file PRISM\_4season\_precip\_et\_Hargreaves-Samani\_for\_Zone\_A\_1999-2017\_2018\_0109.xlsx that is provided as backup. Locations of figures in this file are as follows.

Fig. 1. ZoneA\_monthly\_ppt\_et!V3 Mean Monthly ET and precipitation (1999-2017)

Fig. 2. ZoneA\_monthly\_ppt\_et!V28 Mean monthly ET and P grouped into four periods

Fig. 3. ZoneA\_monthly\_ppt\_et!CC3 ET vs. P for irrigation season May-September

Fig. 4. ZoneA\_gwIrr(af)\_vs\_ppt\_et!I86 Estimated vs. reported irrigation for f5(ET,P)

Fig. 5. ZoneA\_gwIrr(af)\_vs\_ppt\_et!O86 Time series of estimated f5(ET,P) and reported irrigation

Fig. 6. ZoneA\_gwIrr(af)\_vs\_ppt\_et!W86 Same as Fig. 5 with ET and P superimposed



Fig. 1.



Fig. 2.



Fig. 3.



Fig. 4.



Fig. 5.



Fig. 6.

**Appendix. Documentation of procedures to map monthly PRISM climatological data onto GMD5 model grid and to calculate reference ET according to Hargreaves-Samani (1985).**

This Appendix provides supporting documentation of procedures and methods behind the precipitation and ET datasets that were used to develop regression relationships to irrigation water use. These include:

* Programs (or apps) PRISM\_to\_grid\_mo and PRISM\_to\_grid\_mo\_nper: procedures to map PRISM monthly precipitation and temperature datasets at the 4 km scale onto the GMD5, GMD3 and RRCA groundwater model grids at each cell. (These, or variations, should also be applicable for mapping PRISM data onto PLSS sections). A variation, PRISM\_to\_grid\_mo\_nper\_zone, optionally writes a text file of monthly precipitation spatially averaged over a specified zone.
* Program ETref\_HS: Calculation of monthly reference ET according to Hargreaves-Samani (1985), based on PRISM tmin and tmax data (average daily minimum and maximum temperature for each month). A variation, ETref\_HS\_zone, optionally writes a text file of monthly ET spatially averaged over a specified zone.

**Precipitation and temperature derived from PRISM data for the GMD5 groundwater model grid:**

Precipitation and temperature are derived from PRISM datasets that were downloaded from <http://www.prism.oregonstate.edu> as zipfiles. The zipfiles contain space-delimited text files with extension ‘.asc’, which were copied into the folder PRISM. Datasets are stable except for the most recent available six months of data, which are provisional. File names indicate this status with either ‘stable’ or ‘provisional’ as part of the name, along with grid scale and format version. The folder PRISM contains monthly and annual files for 1995-2017 for precip (ppt), tmax, tmean and tmin, where tmax and tmin are max and min daily temperatures averaged over the month or year. The Hargreaves ET calculations use all three. Tmean can either be taken from the tmean data files or calculated as an average of tmin and tmax.

Fortran programs were developed to extract subsets of the PRISM data for the GMD5 model grid or others, including those for GMD3 and RRCA. Monthly and annual versions of the PRISM precip and temperature data for each year and month were written to subfolder GMD5. The files in this folder are all text files and mostly grid files, one per month or year with 165 rows and 335 columns (space delimited, 335f10.2 format).

**Programs PRISM\_to\_grid\_mo and PRISM\_to\_grid\_mo\_nper:**

Files with extensions ppt, tmax, tavg and tmin were all written by program PRISM\_to\_grid\_mo as shown in the batch file run\_PRISM\_to\_grid\_mo\_GMD5.bat in subfolder ‘bat’. Source code and executables are in subfolder ‘bin’. The program uses the available monthly PRISM data. For each year and month of the period specified, it first looks for a file with stable data to read; if that fails, it looks for provisional data. If that also fails, it reads a file with normal (1981-2010) data for the month.

For example, as of November 30, 2017, GMD5 model grid files for precipitation and temperature can be assembled for 2017 that include four months of stable data Jan-Apr, six months of provisional data May-Oct and two months of normal data Nov-Dec, thereby providing provisional data for the 2017 irrigation season May-Sep.

For each run of PRISM\_to\_grid\_mo listed in the batch file, the program reads a parameter file (ext. par, redirected keyboard input). The parameter file specifies the following on lines 1-6:

Line 1: parameter filename.

Line 2: georeferencing mapfile path and name. The mapfile contains one record for each model grid cell, and includes fields that identify PRISM grid cells (row, column) corresponding to GMD5 model grid cells for each cell. The mapfile also includes (latitude, longitude) coordinates for each model cell center and other fields. [Analogous mapfiles georeference the RRCA and GMD3 model grids.]

Line 3: output folder where the GMD5 grid files are to be written.

Line 4: Number of rows and columns of the GMD5 grid.

Line 5: Beginning and ending years of data to be mapped onto model gridfiles.

Line 6: File extension of the PRISM data to be read. This file extension must be one of the following: 'ppt', 'tmin', 'tmax', 'tmean'; otherwise, the program will let you know and stop.

Program PRISM\_to\_grid\_mo\_nper is similar but applies only to precipitation data, and writes annual and seasonal summary files in which rows correspond to model nodes (165x335, or 60300 records) and columns correspond to years 1999-2016 (first column is node number).

For precipitation, the above example for 2017 was run as follows from the PRISM folder:

bin\PRISM\_to\_grid\_mo\_nper < mapfiles\PRISM\_to\_GMD5\_2017nper.par

> mapfiles\PRISM\_to\_GMD5\_2017nper.log

**Program ETref\_HS:**

This program, demonstrated by the last line of the batch file, uses the monthly temperature files for tmin and tmax to calculate monthly reference ET according to the Hargreaves and Samani method (1985). as described at <http://biomet.ucdavis.edu/Evapotranspiration/PMmonXLS/PMmon.htm> by Snyder and Eching (2002, 2006). The source file ETref\_HS.f90 also provides this link to the monthly calculation procedure, which the source file implements (hopefully correctly, subject to additional checking). This program writes grid files with the extension ‘eth’ in the same format and to the same folder as the precip and tmin GMD5-grid files.

Program ETref\_HS was written in double precision to avoid unnecessary numerical problems. It also writes seasonal and annual summary text files, analogous to those written by PRISM\_to\_grid\_mo\_nper (see above). The “seasons” are defined in the parameter file etref\_HS\_GMD5\_1999-2016.par, line 8, with labels on line 9 (3 seasons: Jan-Apr, May-Sep and Oct-Dec). Annual file is 1999-2016.eth; Seasonal file for May-Sep is 1999-2016May-Sep.eth

**Spatial averaging of precipitation and ET data**

Precipitation and ET annual and seasonal data were spatially averaged by county and for all of GMD5 and Zone A in Excel files. These averages can be updated for the grouping of monthly data for the four seasons October-February, March-May, June-July and August-September for the purpose of applying regressions by county. This process would be similar to that used for the above three seasons for precipitation data [file PRISM\_precip\_1999-2017\_3seasons\_GMD5\_grid.xlsx] and ET data [file PRISM\_et\_1999-2016Qtr\_GMD5\_grid.xlsx].

These files contain sheet gridZones, which associates GMD5 model grid cells with various zones (basin, county, township, GMD and response Zone A), and is used in calculation of averages by county, GMD5 and Zone A for annual and seasonal precipitation or ET.