Application of Solar-Atmospheric Connections Towards Improved Forecasts of the Animas River and other Streams in the Western US

Michael Wallace
Hydrologist, MW&A
Albuquerque, NM
mwa@abeqas.com

ABSTRACT

Considerations of connections between modern solar cycles, Hadley and Walker circulation patterns, along with streamflow characteristics of several mid latitude, high altitude watersheds have pointed to the potential for improved multi-annual to sub-decadal forecasting of streamflows in targeted locations. Such conditions appear to apply to watersheds of the southern Rocky Mountains of the western United States.

In this exploration, correlations and linear regressions were developed for key sequential features that have been provisionally identified. The resulting forecasting approach focused on streamflow time series for prime candidate streams of the Southern Rocky Mountains, including the Animas and Pecos Rivers. A third stream, the Gila River in the Mogollon Mountains was included in the comparative study because of its marginal candidacy with respect to geopotential height and latitude.

The forecasts based upon the new regression method were the most accurate of all featured methods, for the series considered under a five-year trailing average. Through a sequence of solar, trade wind and streamflow cross-regression exercises, forecasts for the Animas and Pecos Rivers were advanced as far as six years into the future.

INTRODUCTION

This paper is associated with two oral presentations and a poster by this author at the 1st and 2nd Annual Conferences on *Environmental Conditions of the Animas and San Juan Watersheds with Emphasis on Gold King Mine and Other Mine Waste Issues*, sponsored by the New Mexico Water Resources Research Institute in 2016 and 2017. Material here also summarizes a paper submitted to the *Hydrological Sciences Journal* which is in a closing stage of review (Wallace, 2018). Accordingly, extensive text from that paper has been condensed and consolidated by the same author to produce this complementary product.

Many hydrologists must at some point address the challenges of forecasting the long-term future states of streams associated with their study domains. To date most researchers have relied primarily upon stochastic approaches, which depend on records of the history of the subject stream system for estimating that system's future. Accordingly, the nominally predictive solutions are based on autocorrelation (AC) and/or autoregression moving average (ARMA) strategies, such as described in Wei (2006).

These strategies are typically parametric, given that they are often geared to develop best estimates of the first several moments (mean, variance, skew) for a given time series. Moreover such parametric results are typically cast into "climatological" products, including the mean flow for a given season or month, or as estimates of extreme events such as a "probable maximum flood." A more desirable type of forecast would for example project the arrival of a period of drought, with a certain lead time, span, and intensity across a given region, hopefully with acceptable accuracy for planning purposes.

Accordingly, if independent and reliable precursors were available over multi-season to multi-decadal scale time frames, then hydrologists would likely employ them in other regression based approaches towards improved forecasting. Some independent precursors to regional climatologic atmospheric moisture patterns have been described. These generally are limited to temperatures and/or pressure differentials associated with locations or regions within the Pacific Ocean, including the widely cited El Niño pattern. This sea surface temperature time series for a stationary location within the eastern equatorial Pacific Ocean is often lumped into a broader compendium of patterns termed the El Niño Southern Oscillation Index (ENSO) (NOAA 2017).

To date, the ENSO suite of precursors have not proven to be consistently reliable for hydrologic forecasting beyond a few months lead span. Deterministic numerical global circulation models (GCMs) have also been frequently deployed towards these same forecast objectives but none so far have been able to accurately simulate extended periods of integrated climate for the globe or any sub-region, at least with regard to hydroclimatology.

In the goal of improving hydrologic forecasting skill, researchers continue to experiment with the adoption of various combinations of selected precursors. As opposed to the conventional adoption of precursors that are primarily temperature or pressure based, this paper focuses on those indexes that exemplify high masses of atmospheric moisture. Moreover, this work contemplates the possibility that solar radiant forcing can drive the underlying precursor parameters.

The sun is indisputably the primary driver of our weather. However explicit and physically consistent correlations between longer term surface climate patterns and solar cycles, including the 11 year cycle (Schwabe 1844), have yet to be fully verified or widely documented. In contrast to this contemporary lack of evidence, older studies from the 19th and early 20th centuries produced extensive products that purported to show high correlations between solar cycles and climatic features, including for example global temperature estimates (Koppen 1914) and cyclone frequencies in the Indian Ocean (Meldrum 1885). Notably, the apparent persistence of these synchronous correlations had faded by the middle of the 20th century (Hoyt and Schatten 1997).

Ultimately new connections between the solar irradiance and climate signatures were identified by researchers such as Labitzke and Van Loon (1995) in the upper atmosphere by the late 20th century. However the establishment of lower atmospheric and near surface moisture and temperature correlations to SSNs has remained elusive to date. Many researchers have nonetheless recognized the potential to connect SSNs to surface water hydrology. This was considered a productive study area because streamflow records are commonly understood to represent an integrated signature of climate over their watershed footprint.

In addressing this objective, the recent work of Wallace (2018) included the development of a conceptual model in which the global hydrosphere undergoes energy changes over time and within its Hadley and Walker circulatory limbs as a result in part of slight changes in the total solar irradiance (TSI) of approximately 0.17 W/m^2 on average. The TSI can be represented as an index in numerous variations, including the published monthly Sunspot Numbers (SSN) as archived at the Royal Observatory of Belgium (WDC-SILSO).

From this work, new correlations between solar cycles and atmospheric parameters have been found over the Western Equatorial Pacific (WEP) region. That work identified for the WEP region that solar cycles express pervasive lagged correlations to multiple components of the atmosphere. These correlations were identified across that footprint from the surface via trade winds (TWWP), through the core via latent heat, and up to the top of the atmosphere (TOA) via outgoing longwave radiation.

As featured in that study by several maps of the geopotential height, the overlying weight of the full atmosphere drops significantly over major high altitude land masses in middle latitudes, and those indentations are routinely indicative of enhanced atmospheric moisture. The study accordingly extended the exploration to streams originating from such regions. The initial results were suggestive that significant lagged correlations between those streams and the TWWP could be identified.

Given the lagged correlation between the solar signature and the atmospheric parameters overlying the WEP, it appeared possible to apply a cross-regression moving average (CRMA) approach towards their prediction. Moreover, given the lagged correlation between those atmospheric parameters and the streamflow signatures within the Southern Rocky Mountains (SRM), it also appeared possible to apply a CRMA approach towards streamflow prediction. In essence, these two correlative pairings represent the potential for an accurate, two step CRMA analysis connecting solar cycles to streamflow records.

APPROACH AND METHODOLOGY

Three SRM streamflow records were selected for the CRMA analysis as identified in Table 1 and the map of Figure 1. The three SRM gages are represented in the map by the solid blue dots. Within that domain the Animas is the northernmost gage, the Pecos is the central gage to the east and the Gila gage is at the southern end of the cluster. The positions of the selected streamflow gages with respect to the geopotential height depressions, which can also be identified in the figure and that were developed from the routine satellite ERA-Interim resource were expected to impact the performance of this forecasting approach.

The first featured site is the Pecos near Pecos, New Mexico gage (PnP). It captures water from a catchment which reaches over 4 km in elevation and resides within the noted geopotential indentation at a middle latitude (approximately 37 degrees N). This gage lies largely upstream of any significant human operations such as reservoirs and irrigation. The Pecos River drains ultimately to the Gulf of Mexico within the greater Atlantic Ocean. The second featured site is the Animas River gage near Farmington, New Mexico. The Animas gage shares nearly all of the same characteristics as the Pecos. However its flow rates are roughly 5 to 10 times higher than those of the Pecos, and it drains ultimately to the Upper Colorado River, which reaches the Pacific Ocean.

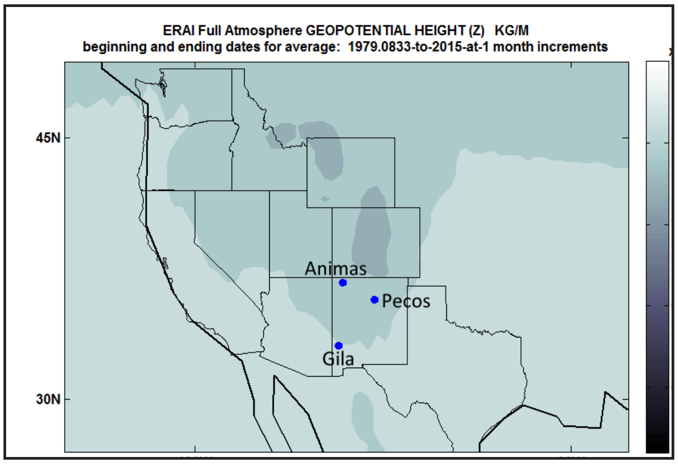

Figure 1. Geopotential height contours across the SRM and stream gages.

Table 1. Stream Gages Utilized.

River, State and Country	Streamflow Gage ID	Decimal Longitude	Decimal Latitude	drainage area (km^2)	Elevation of highest edge of catchment (km amsl)	Years of continuous records available
Pecos River	PnP	254.31730	35.70835	490	4.0	86
Animas River	Animas	251.79825	36.72250	3522	4.2	85
Gila River	GnG	251.46261	33.06150	4828	2.7	88

Sources:
USGS 08378500 Pecos River near Pecos, NM accessed online at http://www.usgs.gov/water/
USGS 09364500 Animas River at Farmington, NM accessed online at http://www.usgs.gov/water/
USGS 09430500 Gila River near Gila, NM accessed online at http://www.usgs.gov/water/

The third site, the Gila River near Gila, New Mexico was chosen primarily due to its marginal qualities with respect to the candidate criteria. As noted, the other two catchments reach above 4 km in elevation but the Gila's watershed only reaches to slightly below 3 km in elevation. Moreover, the Gila watershed is further to the south of the middle latitude target (at approximately 33 degrees N). The Gila also drains ultimately to the Pacific Ocean.

An exploration of the autocorrelations for each time series is helpful in targeting some of the most useful moving averages and lags for the forecasting exercises. Figure 2 outlines a series of autocorrelation functions (ACFs) associated with the SSN and TWWP indexes along with the selected streams. The maximum lag considered was 11 years, specifically because that is the primary period of the SSN index. Many notable features of this chart stand out. For example, the ACFs of the Sun and the Animas River are congruent for the most part. Moreover, to a greater or lesser extent, each ACF shows a cyclostationary behavior, with a dip to low or negative autocorrelations near the five-year lag followed by a return to higher and more positive ACFs near the 11-year lag.

The proposed method is subsequently shown to produce a promising yet variable degree of forecasting span and accuracy improvement across the study region. The first step of the process developed and applied the lagged correlations between each stream gage and the satellite era TWWP for a five-year trailing averages (5yta) streamflow forecast with a lead time of three years into the future.

As part of this initial step, those correlations are applied to the customary linear regression solution as Equation 1.

$$y = Ax + B \qquad (1)$$

Where:

x is the independent variable (TWWP for the first CRMA regression step or simply the appropriate streamflow value for the ARMA case),

y is the dependent variable (streamflow for the first regression step),

A is the dimensionless regression coefficient calculated from a linear fit of a scatter plot between historical values of y and x, and

B is the intercept (the value of y when x is equal to zero).

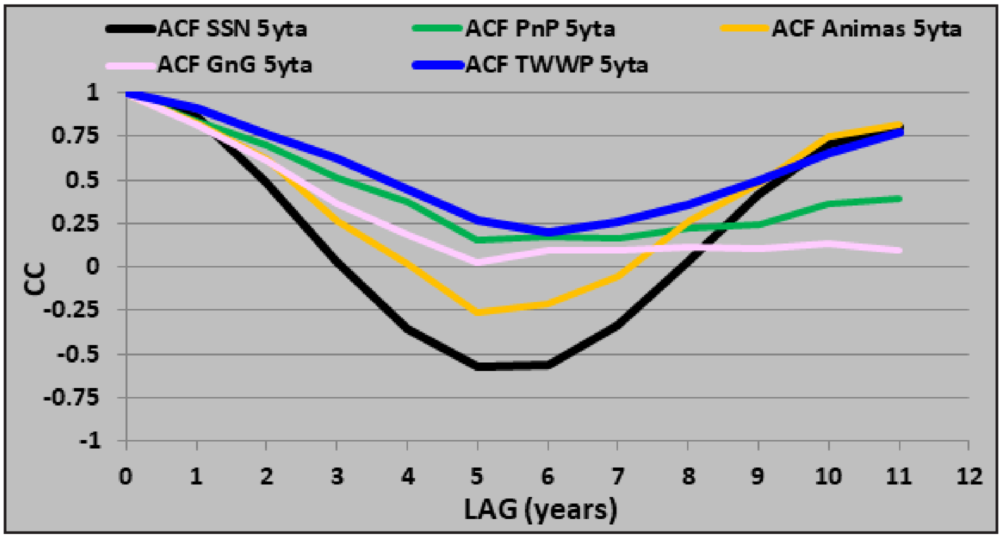

Figure 2. Autocorrelation functions (ACFs) for five-year trailing averages (5yta) of the SSN, TWWP and selected streamflow observation time series.

Accordingly, Pearson correlation coefficients were calculated between the TWWP and each SRM stream in the Wallace study (2018). This exercise demonstrated that for all streams, the TWWP correlations were consistently the highest. Moreover the TWWP correlations generally showed the highest significance scores. The resulting ARMA and CRMA calculations produced the dependent time series for the three SRM streamflow gages.

The second step of the analysis was largely identical to the first step in structure but was limited to extending the three year lead forecasts for the 5yta CRMA cases an additional three years. Moreover, it was limited to the two stream gages associated with the Pecos and Animas Rivers, as opposed to the more marginal Gila River case. This second step was motivated by the correlations indicated between the TWWP and solar cycles. It was additionally informed by the fact that solar cycles themselves are routinely (if not always accurately) forecast in advance by one or more years, as documented for example at the WDC-SILSO site. Accordingly the step entailed first obtaining a record of 5yta SSN values to the current time, and then adding the projected SSN values for the subsequent year. Next, the Pearson correlation coefficient was calculated between the SSN record and the TWWP for a two year lag. The resulting regression coefficients were calculated and those were used to forecast the 5yta TWWP for three additional years into the future.

These projected TWWP values were then utilized, again through the regression relation of Equation 1 and the coefficients corresponding to each stream, to predict the 5yta streamflows of the Pecos and Animas rivers for an additional three years into the future. For those specific cases therefore, streamflows were quantitatively predicted for a total of six years beyond the year 2016. The results of this final step were simply appended to the results from the previous step for the two relevant cases.

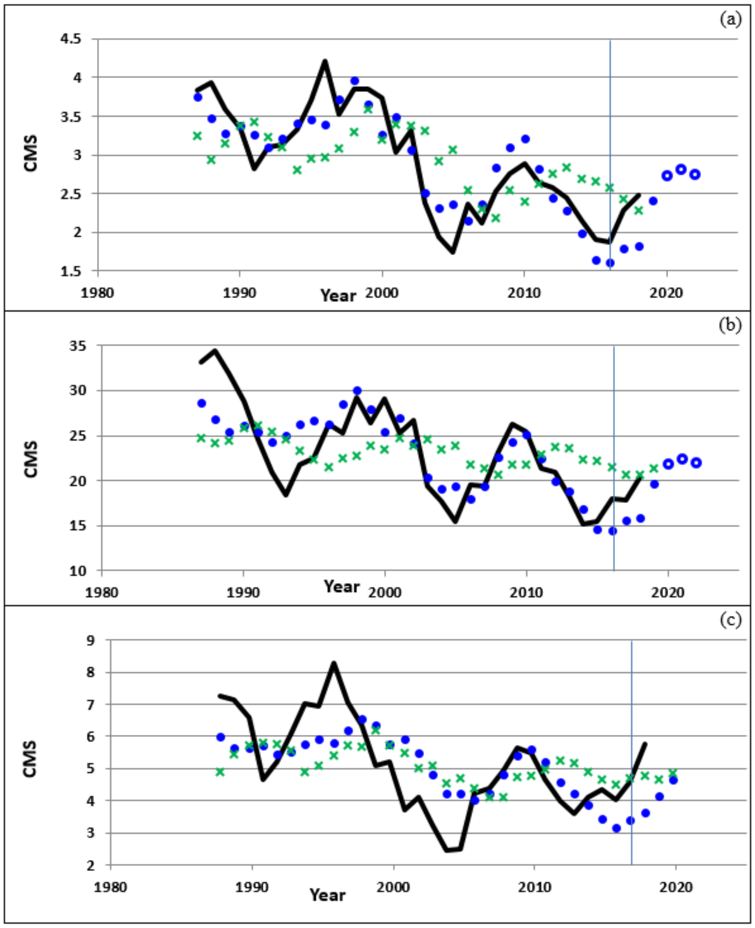

Figure 3. Comparisons of three year-lead predictions to observed values for the featured streams. (a) PnP (b) Animas (c) GnG. All values are for five-year trailing averages (5yta) in cubic meters per second (CMS).

RESULTS

As noted in the previous section, the results documented here include specific cases from application of the widely used Cross Regression and Auto Regression Moving Average (CRMA and ARMA) approaches towards time series predictive analyses. A graphical overlay of observations to the simulated time series is helpful if not essential as a first order assessment of the skill of each exercise. Figure 3 accordingly highlights these overlays for the 5yta cases for each of the three stream gages in this study. In this figure set, the solid black line represents the observed series and dotted lines represent the simulations. The green Xs define the ARMA solutions and the blue dots define the CRMA solutions. The thin vertical line for each subfigure defines the transition from the so-called training period to the test period. In other words, the regressions were developed based upon observation records through for the most part the end of the year 2015.

For most cases the TWWP results exhibit lower errors across the range of streamflow values. Naturally, the closer the fit to both mean and extreme flow values, the better the overall performance of the method. Also, most of the methods show nominally good performance at estimating mean values. Skill performances are routinely quantified through the common Root Mean Squared Error (RMSE). However a greater sensitivity in comparing results was desired and accordingly two Goodness of Fit (GOF) measures were adapted from common significance and hypothesis testing resources. Those measures were the well known Chi squared (χ2) and Kolmogorov Smirnov (K-S) tests. These quantified skill measures of the ARMA and CRMA regression calculations are featured in Table 2. Figure 3 and Table 2 therefore document that among the three approaches explored, the TWWP CRMA exercises produce the most accurate forecasts, particularly for the two streams located well within the main target domain identified.

SUMMARY AND CONCLUSIONS

Considerations of past research regarding solar cycles, Hadley and Walker circulation patterns, and streamflow characteristics of several mid latitude, high altitude watersheds have pointed to the potential for improved multi-annual to sub-decadal forecasting of streamflows in targeted locations. Such conditions appear to apply to the SRM of the Western United States.

In the development of this conclusion, sets of correlations were explored for key sequential features based upon previous published research and currently available solar cycle, and trade wind observations. Equivalent exercises were applied towards potential connections of some of those parameters to streamflow data sets for candidate streams of the SRM. A two staged cross-regression based forecasting approach (CRMA) was then applied to forecast streamflows up to six years into the future. The skill of the CRMA forecasts for the training period were compared to forecasts for the same stream sets via a conventional autocorrelation technique (ARMA).

Table 2. Performance Results.

Forecast Case	Root Mean Squared Error (RMSE)	χ Squared ρ Value	Chi Test Significance at ρ <.1 ?	Kolmogorov Smirnov ρ Value	Auto-correlation of residuals
SSN 2 yr lead forecast of TWWP 5 yta	0.71	2.94E-03	N	0.6	0.581
TWWP 3 yr lead forecast of PnP 5 yta	0.3182	0.999	Y	0.652	0.054
AC 3 yr lead forecast of PnP 5 yta	0.6022	0.036	N	0.200	na
TWWP 3 yr lead forecast of Animas 5 yta	3.1279	0.438	Y	0.781	-0.123
AC 3 yr lead forecast of Animas 5 yta	5.0446	3.94E-14	N	0.011	na
TWWP 3 yr lead forecast of GnG 5 yta	1.1188	5.64E-04	N	0.211	0.089
AC 3 yr lead forecast of GnG 5 yta	1.3550	2.43E-04	N	0.109	na

"na" means not applicable to this exercise. "yta" means year trailing average. "AC" means autocorrelation. "PnP", "Animas", and "GnG" are streams as defined in text.

The training forecasts of flows in the Animas River, the Pecos River and the Gila River at the selected observation gages, based upon the CRMA approach were the most accurate of all featured methods. Given the three- to six- year forecast span, several years remain for subsequent review and evaluation. Should these forecasts continue to show high fidelity, they may point the way to a more routine, longer span and higher accuracy approach to streamflow and general drought forecasting to the benefit of all.

REFERENCES

Hoyt, D.V. and Schatten, K. H. 1997. The Role Of The Sun In Climate Change. Oxford University Press.

Koppen, W. 1914. Lufttemperaturen, Sonnenflecke und Vulkanausbruche. Meteorlogische Zeitung, 31, 305-328.

Labitzke, K. and Van Loon, H. 1995. Connection between the troposphere and stratosphere on a decadal scale. Tellus A, 47 A, 275-286. https://doi.org/10.1034/j.1600-0870.1995.t01-1-00008.x

Meldrum, C. 1885. On a supposed periodicity of the cyclones of the Indian Ocean south of the equator. British Association Report, 925-926. Nguyen, H., Hendon, H. H., Lim, E. P., Boschat, G., Mahoney, E. and Timbal, B. 2017. Variability of the extent of the Hadley circulation in the southern hemisphere: a regional perspective. Climate Dynamics printed online Springer-Verlag Berlin Heidelberg

Schwabe, A.N. 1844. Sonnen-Beobachtungen im Jahre. Astronomische Nachrichten, 21, 233.

Wallace, M.G., 2018, Application of newly identified solar – atmospheric connections towards improved forecasts of streamflows in the Western US. actively in review at the Hydrological Sciences Journal. Manuscript ID HSJ-2017-0119.R3

Wei, W.S. 2006. TIME SERIES ANALYSIS, Univariate and Multivariate Methods, Pearson Addison Weseley publishers Boston.

SOURCE DATA

Animas River Source: USGS 09364500Animas River at Farmington, NM Available from https://waterdata.usgs.gov /

ENSO including TWWP Sources: Available from http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml

NOAA (National Oceanic and Atmospheric Administration) National Climate Data Center (NCDC) 2017

https://www.esrl.noaa.gov/psd/cgi-bin/data/testdap/timeseries.pl

Gila River Source: USGS 09430500 Gila River near Gila, NM. Available from https://waterdata.usgs.gov

Pecos River Source: USGS 08378500 Pecos River near Pecos, NM Available from https://waterdata.usgs.gov

UCAR ERAI Source: Available from https://climatedataguide.ucar.edu/climate-data/era-interim

SSN Source: WDC-SILSO, Royal Observatory of Belgium, Brussels. Available from http://www.sidc.be/silso/