To find events in both river discharge data and snow water equivalent data for the Rio Grande, New Mexico, moving averages (short term average over long term average, STA/LTA) are used to create an event detector. Additionally, a cross-correlation is performed on the data from the United States Geological Survey and the National Resource Conservation Service (1981-2010) to see the relationship between the river discharge data and the snow water equivalent data. Using lag times calculated from the cross correlation, the difference in peak times from discharge and snowmelt indicates a shift in snowmelt to earlier in the spring. Because of the high correlation between snow water equivalent and river discharge, plus the results of the cross-correlation, it is found that the peak in snowmelt is occurring earlier in the year and that the time lag between peak snowpack and peak river discharge is decreasing, meaning that the snowpack is generally melting faster. One implication resulting from hydrologic changes such as these is the adaptation of aquatic and terrestrial species depending on the system for survival.
INTRODUCTION

New Mexico, like many regions that rely on snowpack for water supply, is particularly vulnerable to climate change and the possibility of peak snowmelt occurring earlier in the year (Serreze et al. 1999). In the western United States, for example, approximately 50-70% of the precipitation falls as snow, and spring/early summer snowmelt runoff accounts for approximately 50-80% of the total annual runoff for snowmelt-dominated basins (Serreze et al. 1999), thus creating a need to study the interaction of spring snowmelt runoff and river discharge (Adams and Comrie 1997, Pelletier and Turcotte 1997, and Peterson et al. 2000). April 1st is typically used as the date of maximum snow accumulation (Maurer et al. 2007 and Graf 2006) but models show that a 1º - 2ºC temperature increase could yield a 10 - 15 day peak shift of discharge (IPCC 2007). A 3ºC temperature increase (under mid to high greenhouse gas emission scenarios) could also shift peak streamflow by 30 days (IPCC 2007). Additionally, snow has a high albedo, so there will be a positive feedback of polar amplification of global warming if there is a decrease in yearly snowpack (Mote 2003). Warmer winters yield less snowpack due to precipitation falling as rain rather than snow, and additionally move the peak snowmelt runoff to an earlier date. If the peak of spring snowmelt shifts to be earlier in the year, water managers will be faced with an annual challenge of storing winter precipitation for use later in the year (Stewart 2009).

Previous studies have found hydroclimatological changes in the last 50 years in the western United States. The changes are evident in the timing of spring runoff (Stewart 2009), in the fraction of rain versus snow (Knowles et al. 2006), in the amount of water contained in the snow (Mote 2003), in the fraction of annual streamflow throughout the year (e.g. Hidalgo et al. 2009 found that the March fraction of annual streamflow has increased while the April – July fraction of annual streamflow has decreased), and in climate-sensitive biological variables (Cayan et al. 2001). It is thought that these changes are related mainly to temperature increases as they affect snowmelt-dominated basins in ways predicted in response to warming (Mote 2003), and it is suspected that the warming trends causing the changes are in part due to anthropogenic effects. Recent studies have shown that snowpack volumes and snowmelt runoff have varied with climate on many temporal and spatial levels (Zhang et al. 2007), but recent global surface temperature increases due to anthropogenic greenhouse gas emissions are now well recognized along with their potential implications on the hydrologic system (IPCC 2007).

In this study, two methods are used to quantify spring snowpack and snowmelt. The first method is an averaging method that uses the STA/LTA algorithm, which is a short time running average divided by a long time running average. This method can be used to pick events out of time series data based on a threshold level (Wong et al. 2010). The use of this moving time-averaging method is necessary for a quantitative analysis of river discharge and SWE data because it mathematically calculates ratios. Additionally, with an event detector such as this, we can make proper decisions regarding water quality linked to events of interest and other management concerns (Cayan et al. 2001). The second method is a cross-correlation to analyze the similarities and differences between snowpack and spring snowmelt runoff. Lag times between peaks are also calculated in order to determine if the timing of snowmelt runoff has shifted in the last 30 years, how sensitive the Rio Grande Catchment is to mountain snowpack runoff, and if the lag time between peak snowpack and peak river discharge has also shifted in the last 30 years.
BACKGROUND

Each year, the National Weather Service and the Soil Conservation Service issue monthly forecasts of the streamflow that can be expected during the main runoff period, April through July, for much of the western United States (Redmond and Koch 1991). These forecasts are mostly based on the existing snowpack but also on expected future precipitation. They are used extensively throughout the western United States to develop reservoir operation plans for flood control and water supply plans; in turn, decisions regarding the management of many agricultural operations are also based on expected water availability (Graf 2006 and Walton and Hunter 2009). The importance of these predictions indicates the economic implications of improved estimates of snowpack and spring runoff.

Redmond and Koch (1991) suggest that the confidence interval on a water supply forecast made very early in the water year (e.g. October or November) might be increased by as much as 15% due to climate change. It is for this reason that it is better to make predictions later in the water year. These late predictions may not give water managers sufficient time to make critical decisions regarding water use. Thus, it is important to accurately quantify the high correlation between snowpack and spring runoff (Krishna 2005).

Others have used various methods to quantify event data from hydrologic time series (e.g. Guralnik and Srivastava 1999, Smith et al. 1998, and Hamed 1998), but Norbiato et al. (2007), for instance, use the index variable method to do regional frequency analyses, allowing them to use data from nearby or similar sites to estimate quantiles of the underlying variable at their given sites. They find that attributing a single return period to a storm event is not realistic and that using a traditional rain gauge network can be too sparse to provide adequate sampling (radar data can give an advantage over actual rain gauge data (Delrieu et al. 1997 and Velasco-Forero et al. 2009)). Despite these findings, standard gauges can be used for the analysis presented here since we are studying events rather than entire volumes of precipitation for a given watershed.

Guralnik and Srivastava (1999) use a modeling approach to detect a change point by detecting the change of a model (or the parameters of the model) that describe the underlying data, and Smith et al. (1998) use wavelets to identify transient features to quantify the temporal variability of streamflow. Their study estimates precise locations of both stochastic and periodic events in time that revealed subtle structures not seen in time series data. Hamed (1998) then went on to use the Mann-Kendall test to detect trends in hydrological data. Since the test is not overly sensitive to outliers, it is accepted as a decent statistical method for hydrology.

Given the previous work, further analyses and correlations will be useful information for water management techniques. The magnitude of a given peak, plus its timing, are the two most important features of a hydrograph (Jain and Indurthy 2003), so those are the parameters that are analyzed here.

METHODS

In this analysis, we analyze National Resource Conservation Service (NRCS) SNOTEL (for SNOwpack TELemetry) data from the NRCS website (http://www.wcc.nrcs.usda.gov/snow/) at the Quemazon site in New Mexico. The site number is 708 and it is located in Los Alamos County, the latitude is 35 deg 55 min N, the longitude is 106 deg 24 min W, and the elevation is 9500 feet (Figure 1). The daily data, including snow water equivalent (SWE), precipitation, various air temperature quantities, and snow depth were downloaded from January 1981 through May 2010. The parameter of interest for this study is the SWE, measured in inches. To measure this
parameter, a measuring device called a snow pillow is positioned so that it can determine the water-
content of the snow covering. The working principle of the sensor is based on the detection of the 
hydrostatic pressure caused by the layer of snow (Peterson et al. 2000). This SNOTEL site is 
located northwest of the USGS stream gauge monitoring river discharge for this project (Figure 
1).

Additionally, we also analyze United States Geological Survey (USGS) stream gauge data from 
the USGS website (http://water.usgs.gov/) for the Rio Grande at Otowi Bridge site in New Mexico. 
The site number is USGS 08313000, the county is Santa Fe, the latitude is 35 deg 52 min 28.2 sec, 
the longitude is 106 deg 8 min 32.8 sec, and the elevation is 5488 feet (Figure 1). The daily data, 
including river discharge, were downloaded from January 1981 through May 2010. Historically, 
the Otowi gauge used several different types of instruments, but the technology today is the 
following: The stream channel cross section is divided into numerous vertical subsections and in 
each subsection the area is obtained by measuring the width and depth of the subsection, and the 
water velocity is determined using a current meter. The discharge in each subsection is computed 
by multiplying the subsection area by the measured velocity. The total discharge is then computed 
by summing the discharge of each subsection. This site is located southeast of the NRCS SNOTEL 
site used for this project (Figure 1).

Figure 1. New Mexico map depicting the NRCS SNOTEL site at Quemazon and the USGS stream 
gauge site at the Otowi Bridge.
STA/LTA

The first step of the moving average analysis is to calculate short term averages (STA) and long term averages (LTA). Let $x_i$ be the time series representing either river discharge or SWE data. Let the number of points in a short-term window be $n_s$ and the number of points in a long-term window be $n_l$ with $n_l > n_s$. The average values in the short term and the long-term windows preceding the index $i$ are as follows:

$$
STA = \frac{1}{n_s} \sum_{j=i-n_s}^{i} x_j^2
$$

(1)

$$
LTA = \frac{1}{n_l} \sum_{j=i-n_l}^{i} x_j^2
$$

(2)

If $j \leq 0$, set $x_j = (x_1 + x_2)/2$ (Wong et al. 2010). We can then define the STA/LTA ratio:

$$
\frac{STA}{LTA} = r_i
$$

(3)

For both the river discharge and the SWE calculations, $n_s = 100$ and $n_l = 400$. These values are determined by optimizing the length of the data set with the length of the yearly discharge and SWE event peaks, respectively. The $n_s$ and $n_l$ values for both data sets are the same: both data sets are daily and of the same number of years. To calculate the STA/LTA, the “filter” function is utilized in Matlab. This function filters the original time series data by the $n_s$ and $n_l$ values, giving a smoothed, averaged value. The STA/LTA is then plotted for each original data point of the time series data. A threshold value is also set in order to make a quantitative cut-off point of significant events (Wong et al. 2010). Based on the moving average, the threshold value is set at $r_i = 1$ for discharge and $r_i = 1.5$ for SWE. This is determined through careful exploration of average values among the years (i.e. percentages of points falling above and below the cut-off).

Cross-Correlation

To perform the cross-correlation, the following procedure is employed for both the discharge data and the SWE data. To find an “average year,” each annual peak is shifted to an arbitrary time scale. This allows us to normalize the data for lag time. For example, if we simply averaged each block of 365 days, there would be some years with an earlier peak and some years with a later peak. These off-years would not be included in the average and thus would not be fully represented. The goal is to see the peak of data without initially looking at the timing of the event. Once each year’s maximum peak event is overlaid on the other peaks from the 30-year data set, one average year can be determined. This information is plotted as both a 2D and a 3D plot.

The Matlab function “xcorr” is used to shift each year’s peak of data through the following sequence:

$$
r(d) = \frac{\sum_i [x(i) - mx] \cdot [y(i-d) - my]}{\sqrt{\sum_i (x(i) - mx)^2} \cdot \sqrt{\sum_i (y(i-d) - my)^2}}
$$

(4)
where \( x(i) \) and \( y(i) \) are the two data sets being compared, \( mx \) and \( my \) are the respective means, and \( d \) is the delay. If we assume \( mx \) and \( my \) to be zero, we have the following:

\[
\begin{align*}
\sum_{i} \left( x(i) \star (y(i-d)) \right) \\
\sqrt{\sum_{i} (x(i))^2} \sqrt{\sum_{i} (y(i-d))^2}
\end{align*}
\]

(5)

To figure the lag times associated with the correlations, a series of delays \( d \) are run through the above equation and the maximum value is determined. This is equal to the lag time. This function computes raw correlations without normalization, so a separate process is then employed to normalize the calculations. To do this, a cross correlation is performed between the average year and each separate year, the average year with itself, and each separate year with itself. The aforementioned 3 cross-correlations are then compared. This allows us to determine the normalized correlation between each year and the average year previously determined. These calculations also allow us to determine the lag time between each year’s peak and the average year’s peak. The correlations are then plotted as time series to see the correlation of each year’s peak to the average year’s peak. Additionally, the lag times are plotted as a time series to see the time difference between each year’s peak and the average year’s peak (with the mean removed to normalize the data).

To then see the correlation and the lag times between each SWE peak and the corresponding river discharge peak, a cross correlation is performed between each seasonal peak (i.e. each SWE peak versus its corresponding discharge peak). This allows us to see how, exactly, each year’s SWE data relates to the following spring’s river discharge. We then plot a time series graph for both the correlations and the lag times. Finally, to see how the lag times and the correlations compare among both the SWE peaks and the river discharge peaks, a plot is created to compare the correlations of each SWE year to the average SWE year versus the correlations of each river discharge year to the average river discharge year. Additionally, another plot is created to compare the lag times of each SWE year from the average SWE year versus the lag times of each river discharge year from the average river discharge year.

RESULTS

STA/LTA values are plotted for both the Rio Grande discharge data and the SWE data. These are presented in Figure 2. The ratio is on the vertical axis and time is on the horizontal axis. The threshold values are also plotted on each graph. The threshold for a significant peak on the discharge STA/LTA is 1 while the threshold value for the significant peak on the SWE STA/LTA is 1.5.

To view the average discharge year with each individual year peak, a plot is created with the shifted yearly peaks, as mentioned in the methods section. Figure 3 shows the average discharge peak with each of the other years’ discharge peaks shifted to a generic time scale. The importance of the plot is the discharge peak value, not the time that the peak occurs. Figure 3 also shows, as a surface plot, each year’s peak discharge shifted to the generic time scale. Discharge values are shown through a color bar as well as on a third axis. Trends over the 30-year time period are apparent.
To view the average SWE year with each individual year peak, a plot is created with the shifted yearly peaks, as mentioned in the methods section. Figure 4 shows the average SWE peak with each of the other years’ SWE peaks shifted to a generic time scale. Again, the importance of the plot is the SWE peak value, not the time that the peak occurs. Figure 4 also shows, as a surface plot, each year’s peak SWE shifted to the generic time scale. SWE values are shown through a color bar as well as on a third axis. Trends over the 30-year time period can be seen.

Several other relationships are plotted in Figures 5 and 6: the correlation of each year’s discharge peak with the average discharge peak, the time difference of each year’s discharge peak with the average discharge peak, the correlation of each year’s SWE peak with the average SWE peak, and the time difference of each year’s SWE peak with the average SWE peak.

The correlations are then calculated and plotted for each SWE peak and its corresponding discharge peak (Figure 7) where the normalized correlations are plotted on the vertical axis and the years are plotted on the horizontal axis. The lag times are then calculated and plotted for each SWE peak and its corresponding discharge peak (Figure 7) where the lag times are plotted on the vertical axis and the years are plotted on the horizontal axis.

To see the correlations between each year and the average year of both the discharge data and the SWE data, a plot is created with the normalized SWE correlations on the vertical axis and the normalized discharge correlations on the horizontal axis (Figure 8). To see how the lag times between each year and the average year of both the discharge data and the SWE data, another plot is created with the normalized SWE lag times on the vertical axis and the normalized discharge lag times on the horizontal axis (Figure 8).

**DISCUSSION**

In this study, two methods are used to quantify spring snowpack and snowmelt. The first method is an averaging method that uses the STA/LTA algorithm. This method is used to pick events out of time series data based on a threshold level (Wong et al. 2010). Interannual variability is clearly visible in the time series data of both river discharge and SWE so the STA/LTA method allowed us to “smooth” the data into detecting events of significance. Figure 2 shows that there was a significant amount of discharge in most years, but there were a few dry years within the 30-year data set (e.g. 1988 and 1996). The drought in 1988 was part of a drought that covered 36% of the United States and had a large effect on crop production (Riebsame et al. 2001), while the drought in 1996 covered much of Texas, New Mexico, California, Nevada, Utah, Colorado, Oklahoma, and Kansas. The depletion of ground water during both of these years affected the ecosystem for many years (Reuters 1996).

Both Figures 3 and 4 show the shifted yearly peaks with the average peaks on a generic time scale, for both river discharge and snow water equivalent, respectively. The importance of the plots is the peak values in the respective units, not the time that the peaks occurred. Due to the interannual variability of river discharge and SWE, the peaks appear diverse, which allows a further analysis of each peak’s correlation and time lag in relation to the average peak as seen in the figure. Figures 3 and 4 also show each year’s peak shifted to the generic time scale, which gives a nice visual representation of the variability over time. For instance, there are several large peaks earlier in the data sets that are not in the later part of the data sets. These results indicate a general decrease in both peak discharge and peak SWE over time.
In order to see the correlation of each year’s discharge peak with the average discharge peak, a time series plot is created (Figure 5), which shows that the correlations may have a subharmonic throughout time. There is a weak oscillation among the 30 years that would have to be further studied before any conclusions can be drawn. Each year’s discharge peak has a correlation of at least 0.7 with the average discharge peak, meaning that most years are similar to each other and thus the average year. Figure 5 also shows the difference in time between each individual year of discharge data with the average discharge curve. The lag times seem to fluctuate around the zero line, with a big, distinct drop in the late 80s. In the years that have a positive lag time, the peak in discharge is actually later than average, and in the years that have a negative lag time (e.g., 1987,
Coherence of river discharge and snow water equivalent, New Mexico, USA  Tichy

1988, and 1989), the peak in discharge is actually earlier than the average. Given the drought of 1988 (Riebsame et al. 1991), the discharge that year is smaller and earlier than other years in the data set.

Figure 6 presents the correlation of each year’s SWE peak with the average SWE peak, and shows that the correlations may have subharmonic throughout time (perhaps every 8 years). There is a weak oscillation among the 30 years that would have to be further studied to make a conclusion. Each year’s SWE peak has a correlation of at least 0.9 with the average SWE peak, meaning that the yearly SWE is quite consistent. Figure 6 also shows that during the beginning of the data set, there are positive lag times between the yearly SWE peaks and the average SWE peak, but over time, this value switches to be primarily negative. Having a negative lag time means that the SWE peaks

Figure 5. LEFT: Normalized correlation between each year’s discharge peak and the average discharge peak. RIGHT: Normalized lag time between each year’s discharge peak and the average discharge peak. All data are from the Rio Grande Otowi gauge.
are becoming earlier in the year. SWE peaks are being seen earlier in the year, showing evidence for the hypothesis that water will have to be stored for longer periods of time given this shift. If this trend continues, the biological organisms around the Rio Grande will be forced to change their reproductive cycles to match the river system, potentially having adverse affects while the species adapt.

Figure 7 shows the correlation between each winter’s SWE peak and each spring’s discharge peak, in an attempt to see the close relationship between the two. The correlations range from 0.3 to 0.9, meaning that there is a high correlation in some years but not in all. This leads us to the conclusion that there are multiple other factors affecting river discharge. The STA/LTA analysis also hints at this on/off relationship, but it is apparent here. The groundwater and surface water
relationship most likely has a high effect on this system (Wang et al. 2008), in addition to other anthropogenic forcings (Burkholder 1997).

In the same manner that the correlations between each winter’s SWE peak and the corresponding spring’s discharge peak are not uniformly strong, neither are the lag times between each winter’s SWE peak and the corresponding spring’s discharge peak. The time between the 2 peaks for each year ranges from 30 to 130 days (Figure 7). In 1990, there is an unusually high lag time between the winter’s peak in SWE and the spring’s river discharge (128 days), which could indicate a colder-than-normal spring and in 2001 there is an unusually low lag time between the winter’s peak in SWE and the spring’s river discharge (33 days), which could indicate a warmer-than-normal spring.

Additionally, to see how the correlations compare among both the SWE peaks and the river discharge peaks, Figure 8 compares the correlations of each SWE year to the average SWE year versus the correlations of each river discharge year to the average river discharge year. This figure shows that, in general, when there is a high correlation between the SWE peak and the average SWE peak, there is also a high correlation between the yearly discharge peak and the average discharge peak. Although there are several years where there is a high SWE correlation with a low discharge correlation and several years where there is a high discharge correlation with a low SWE correlation, in general, most correlations are high for both SWE and discharge. The years with a high SWE correlation and a low discharge correlation (upper left hand corner of Figure 8) and the years with a high discharge correlation and a low SWE correlation (lower right hand corner of Figure 8) give us indication that other factors are at play in river discharge values. Rainfall and temperatures could have played a role with these outliers. It is interesting to see that none of the years had both a low SWE correlation and a low discharge correlation, indicating that the overarching factors of SWE and discharge are strongly related.

A comparison of the lag times of each SWE year from the average SWE year versus the lag times of each river discharge year from the average river discharge year (Figure 8) shows that some of the years in this 30 year data set have both a positive SWE lag time and a positive discharge lag time,
meaning that these years have peaks later than the average year for both SWE and discharge (Quadrant I). Just a few years have a positive SWE lag time and a negative discharge lag time, meaning that these years have a later-than-normal SWE peak but an earlier-than-normal discharge peak (Quadrant II). Still other years have negative lag times for both the SWE peak and the discharge peak, meaning that these years have an earlier-than-normal SWE peak and an earlier-than-normal discharge peak (Quadrant III). The final grouping of data has a positive discharge lag time and a negative SWE lag time, meaning that these years have an earlier-than-normal SWE peak and a later than normal discharge peak (Quadrant IV).

Quadrants I and III show the same-sided correlation between SWE and river discharge, meaning that the lag times are either both positive or both negative. If both lag times are positive, the peaks occur after the average peak, but if both lags are negative, the peaks occur before the average peak. With the SWE peak analysis, it is shown that the peak occurs after the average peak in the early part of the data set, but in recent years the peak occurs before the average peak. Quadrants II and IV show us that there are other factors at play in the system, such as groundwater and other anthropogenic forcings.

The analysis presented here allows us to answer several questions regarding snowpack and snowmelt but also raises others. It is found that the peak in snowmelt is occurring earlier in the year and that there is generally a high correlation between SWE and discharge. Future studies will need to analyze yearly data with an STA/LTA event detection. This will allow us to see events smaller than the annual cycle, pulling further information out of the data being collected. From this and other studies of the Rio Grande (e.g. Vivoni et al. 2009), proper management decisions can be made. Ideally, management will focus on restoring river processes, allowing the biological communities to naturally adjust to the system (Hanson et al. 2004).

ACKNOWLEDGMENTS

Carolyn Donnelly (U.S. Department of the Interior Bureau of Reclamation), Christian Gunning (University of New Mexico), Heidi Hopkins (University of New Mexico), Kelly Isaacson (Daniel B. Stephens & Associates), Etsuko Nonaka (Umeå University), Bob Sinsabaugh (University of New Mexico), and Helen Wearing (University of New Mexico) provided helpful feedback and are greatly appreciated for their comments on this text.

REFERENCES


Coherence of river discharge and snow water equivalent, New Mexico, USA

Tichy


Wong, J., L. Han, J. Bancroft, and R. Stewart, 2010. Automatic time-picking of first arrivals on noisy microseismic data. CREWES.


ADDRESS FOR CORRESPONDENCE
Jennifer L. Tichy
University of New Mexico
10932 Jicama Way SE
Albuquerque, NM 87123
jltichy@gmail.com