The Correlation Between Peak SWE and Baseflow in Snowmelt-Dominated Systems in Colorado

How are snowmelt and baseflow correlated

Introduction:

* Rationale (connection snow and baseflow, defining baseflow w prev lit)

In snow-dominated watersheds, snowmelt contributions to streamflow cause a peak in the hydrograph in spring with low flows returning in the fall. This paper defines baseflow as daily discharge falling between 50 and 90 percent probability of exceedance also sometimes referred to as low flows (Smakhtin, 2001). Baseflow originating as snowmelt sustains communities on Colorado’s Front Range during the winter months while water accumulates in the high elevation snowpack. In the spring, snowmelt influences these rivers and streams by adding significant quantities of water and increasing flows, but the influence of snowpack on baseflow in the fall and winter is less understood.

* Purpose

The traditional water year (WY, October 1st to September 30th) begins with low flows from a previous year’s snowpack. As the WY progresses, snow accumulates in the mountains and then begins to melt in the spring causing a peak in the hydrograph. After peak melt time, the hydrograph returns to low flows in the late summer and early fall. Thus, the WY contains two sets of baseflow that have been influenced by snowpack from two different years (Figure 1).

* Scope

There are several different methods that have been used to determine baseflow including baseflow separation, frequency analysis, and flow duration curves (FDC) (Smakhtin, 2001). Additionally, tracing the movement of stable isotopes is another way to measure baseflow. Researchers have used 15-day running minimum as a proxy for baseflow as well (Cooper et al., 2018; Godsey et al., 2013).

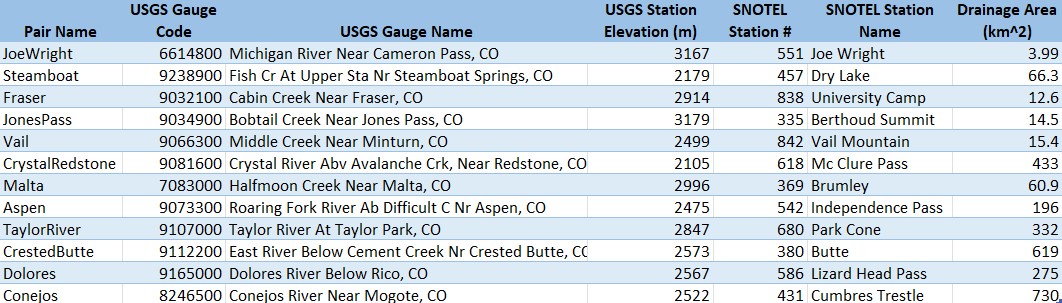
Similar research has been conducted in the Sierra Nevada Mountains of California using modeling to predict how summer low flows will be impacted by changes in peak SWE under a warming climate (Godsey et al., 2014). This research also suggests that there is a lag time or “memory” between when snow years deviating from average resulting in deviating baseflow.

To understand more about how water moves through these snow-dominated systems watersheds, our goals were to compute baseflow that occurs after snowmelt and correlate peak SWE and baseflow. Rather than using the traditional water year (WY), this was done using a melt year (MY) approach that identifies baseflow using that seasons melt. A yearly lag between peak SWE and baseflow correlation was implemented from 0 to 5 years. Additionally, we contrasted low-snow and high-snow years to correlate peak SWE and baseflow over subsequent years. Given trends in the data, we appraised the peak SWE versus baseflow correlation as a function of basin characteristics such as drainage area, elevation, land cover, and directional flow.

Data and Methodology:

To analyze the correlation between peak SWE and baseflow, discharge data from 13 USGS gauges in the Southern Rockies of Colorado were examined from October 1st, 1979 to September 30th, 2022. Stations with greater than ten percent missing data during the time period were excluded. The drainage basins that contain each stream gauge vary in latitude, elevation and area (Table 1). Each stream gauge was paired with an NRCS SNOTEL station within or just outside the drainage basin (Table 1). When the nearest SNOTEL station was lacking data, another nearby station was selected based on the clusters defined by Fassnacht and Derry (2010). Peak SWE data from the SNOTEL stations were used.

Table 1: the USGS and SNOTEL station pairs in order of northern-most latitude to southern-most.



In snow dominated watersheds, the hydrograph demonstrates a large peak occurring at the onset of melt that eventually levels off at lower values for most of the rest of the year. The traditional water year (WY, in the US October 1 through September 30) separates peak melt and baseflow into two different WYs. To reflect the hydrologic processes that are occurring in snow-dominated watersheds, we propose using a melt year (MY) beginning with the onset of snowmelt (the first deviation from baseflow) and ending with the onset of the following year’s snowmelt (Figure 1). The beginning of the MY was defined as the day at which the rate of increase in discharge changes.

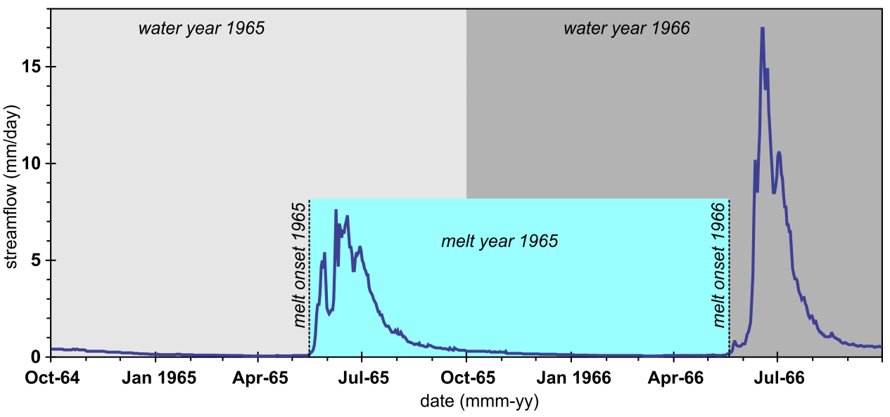


Figure 1: the above hydrograph shows the traditional water year (WY) denoted by the grayed boxes and the melt year (MY) in light blue.

To find the onset of melt, the cumulative sum of the runoff was calculated for each water year. Using the cumulative sum, the onset of melt was selected as the first day when the daily slope was K times greater than the baseflow slope. The constant K was determined to be 6 after testing other values in order to ensure that an onset of melt date was selected for a reasonable time frame. The baseflow slope was calculated between January 1 and March 31.

Once the MY was identified, we ranked and found the probability of exceedance for each discharge value. Flow duration curves (FDC) were constructed by plotting the discharge values against the probability of exceedance for each value per year (Figure 2). The discharge values that fell within 50-90% probability were deemed to be baseflow (Smakhtin, 2001)(Figure 2). Then, we directly correlated peak SWE and the average baseflow values from the same MY. We also compared average baseflow to the previous years’ peak SWE using a 5, 4, 3, 2, 1, and 0-year lag. For example, MY 1980 average baseflow was compared to peak SWE from MY 1975-1980.

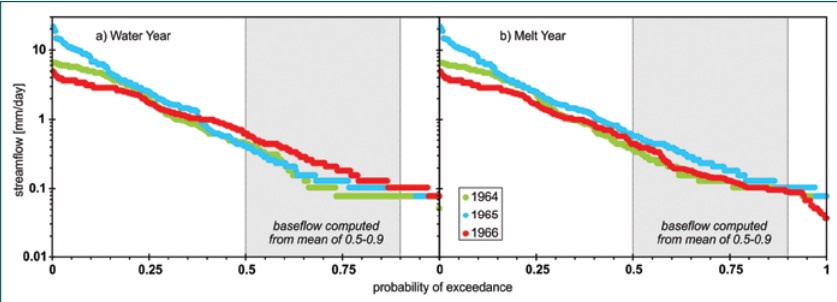


Figure 2: the above figure depicts a sample flow duration curve (FDC) with baseflow isolated in gray. The WY is displayed on the left and the MY is displayed on the right.

The peak SWE values were sorted into low, average, and high snow years using a histogram separating the bins by 0.5 standard deviations from the mean. A multivariate regression was run on average baseflow values with the intercept being a 0–5-year lag to peak SWE. This was calculated for low, high and all snow years. Covariance was calculated for the same parameters as well. Finally, the standard deviation of the average baseflow was calculated for each station pair.

Basin characteristics like drainage area, elevation, and latitude were then examined with respect to the results from the correlation between peak SWE and baseflow as well as the contrast between 0-5-year lags.

Results:

The correlation between the MY and WY is low at the Fraser stations with R = 0.30. This was consistent across basins. This is expected because the WY baseflow values can occur during two different calendar years when calculating baseflow using FDC (Figure 1). MY baseflow values occur during one calendar year because melt begins the year.

The correlation between the MY average baseflow and peak SWE is positive when the values are from the same year (Figure 4). This trend also occur with a lag of 3 years (Figure 4). The correlation across the basins becomes less consistent with lag 4 and lag 5 (Figure 4). After splitting the times series into low and high snow years, the correlation changed. The high snow years tended to show a high correlation across most lags (Figure 4). However, the low snow years had much more variability across the lags (Figure 4).

The covariance analysis showed similar results with greatly consistency amongst stations for all and high snow years, and greater positive trends in lags 0-3. Again, the low snow years were inconsistent, more negative, and generally lower.

The results of the multivariate regressions were inconclusive. The correlation coefficients and p-values were not significant across most basins.

Basins that were in similar spatial locations tended to have similar correlation and covariance values when correlating all years and high snow years. The low snow years varied enough that there did not appear to be any basin characteristic in common.

Chart, scatter chart

Description automatically generated

Figure 3: The above figure shows the correlation between peak SWE and average MY baseflow from the same year at the Fraser stations.

A picture containing graphical user interface

Description automatically generated

Figure 4: The above figure depicts the correlations between peak SWE and average MY baseflow across and all, high, and low snow years for all stations with lags 0-5.

Chart

Description automatically generated

Figure 5: The above figure depicts the summary of the correlation values at the Fraser stations across lags and all, high and low snow years.

Discussion:

Using FDCs, baseflow computed as it occurs after snowmelt has a very low correlation to baseflow computed using the traditional WY across the study basins. This is to be expected given that the date range of the MY is significantly different from the WY.

Overall, there appears to be a slight positive correlation between peak SWE and baseflow in from the same MY (Figure 3). This is likely due to the relatively small basin size and snow-dominated nature of these systems.

The variability of correlation values within the low snow years is likely due to the fact that there is high interannual variability in peak SWE (Figure 6). High years could influence the lag effects of low snow years to the point that a correlation is not obvious in the low snow years. High snow years clearly have a lasting impact on the amount of water in the streams during the same year that the melt occurs, and for a few years after even if a low year follows (Figure 4). It is likely that this is also the reason that the covariance values followed these same general trends.

The multivariate regressions were not able to indicate any trends across basins. Basins that had a positive correlation coefficient or p-value were likely false positives due to the small number of years compared to the number of variables (lags). In the future, it could be useful to decrease the number of lags and only use lags 0 through 3. Increasing the total number of years in the time series could also improve results. Using just high snow or just low snow years will likely not provide a large enough sample size to run an accurate regression.

The similarities across spatially similar basins were likely due to similar latitudes, peak SWE values, and onset of melt timing. Diagram

Description automatically generated

Figure 6: The above figure depicts the interannual variability in peak SWE at the Fraser stations for the complete time series.

* Discuss how you could use land cover and orientation to group the watersheds.
  + Land cover:

A potential next step for this research is to group the watersheds by landcover and orientation to evaluate the impact that these traits might have on the correlation between peak SWE and baseflow. Common landcover types in this region include: herbaceous, evergreen forest, shrub/scrub, and perennial ice/snow. Depending on the type and quantity of vegetation in a basin, snowmelt could be absorbed or runoff differently.

* + Orientation:

The orientation of the basin could indicate

Conclusion:

In snow-dominated watersheds, snowmelt is the primary driver of increases in streamflow in the spring. Baseflow appears to be correlated to peak SWE especially during the same year. This trend can continue with subsequently lagged years during high-snow years. However, due to interannual variability in peak SWE in Colorado, the impact of low-snow years on this trend is variable. This is important to understand because high snow years can potentially mute the impacts of low-snow years on the watershed.

In the future, the 0 through 3 year lags should be further investigated. Also chemical ion trackers….?? Help Steven

This matters in the context of a changing climate and something about ski resorts.

Objectives:

1. Compute baseflow that occurs after snowmelt
2. Correlate peak SWE (snowpack amount) and baseflow directly and with 0-5-year lag
3. Contrast high and low snow years to correlate peak SWE and baseflow over subsequent years
4. Appraise the peak SWE versus baseflow correlation as a function of basin characteristics (drainage area, elevation, land cover, orientation)

Appendix:

Full station info? Coords?

Multivariate regression results?

Code? Link to bookdown/github?

References:

Miller, M. P., Susong, D. D., Shope, C. L., Heilweil, V. M., and Stolp, B. J. (2014), Continuous estimation of baseflow in snowmelt-dominated streams and rivers in the Upper Colorado River Basin: A chemical hydrograph separation approach, *Water Resour. Res.*, 50, 6986– 6999, doi:[10.1002/2013WR014939](https://doi.org/10.1002/2013WR014939).

NOTES: (save)

* Objectives

1. Identify baseflow occurring after melt
2. Correlation between peak SWE and baseflow
   1. High SWE years vs low SWE years
   2. Lag between peak SWE and baseflow
3. Identify trends between basin characteristics and correlation of peak SWE and baseflow

Methods outline:

1. Choose USGS streamflow and SNOTEL station pairs
2. Extract streamflow data
   1. Create WY column
3. Find onset of melt date for each WY
   1. Onset of melt occurs when cumulative baseflow – cumulative runoff > 10
      1. Calculate average Q from Nov1 to Mar15 of each WY (“average baseflow”)
      2. Create cumulative average baseflow column for each WY
      3. Create cumulative Q column
      4. Create difference column (cumulative baseflow – cumulative runoff)
      5. Select date, WY and Q on the first day when diff > 10 occurs (should be in April, May, June)
   2. End date of MY is start date of next MY – 1
4. Assign MY to each Q and date value based on range of start and end dates
5. Rank each day within MY by magnitude
6. Find probability of each rank within each MY
7. Extract Q values for 0.5-0.9 probability (Qbase for the MY)
8. Repeat 5-7 for WY to find WY Qbase
9. Correlate MY vs WY Qbase
10. Correlate MY vs peak SWE all years (0–5-year lags)
11. Separate low, medium, and high snow years
    1. Create histogram of peak SWE
    2. Use 0.5 std. dev. from median to separate 40 years into low, med or high
12. Run direct correlation for individual years from 0 to 5-year lags for all years, then low and high snow years
13. Run multivariate regression for all years from 0 to 5-year lags for all years, then low and high snow years