Rational and background for updating discharge data from Small Watersheds at the HJ Andrews Experimental Forest

Don Henshaw, Sherri Johnson, Steve Wondzell 24 August 2018
Revised 23 October 2018

Rating curves are used to calculate stream discharge from stage height at stream gages in the HJ Andrews Experimental Forest. The Andrews Forest hydrology field team has been regularly collecting rating calibration points for the past 20+ years. Most of the points collected over these years have been at relatively low flow periods when v-notches are installed or removed, although a small collection of points have been taken at higher flows.

Previous updates include rating points collected from 1996 to 2002 that were used to improve rating curves for some watersheds (WS 8, 9, 10, Mack), or used to construct rating curves (WS 6, 7, v-notch weirs), or simply collected as a means to continually check and verify the accuracy of existing curves (WS 1, 2, 3, and 6&7 H-flumes). Since 2002, updates have not occurred to rating curves and calibration points have only been used to check the accuracy of existing rating curves.

Over the spring and summer of 2018, all of the rating curves have been evaluated with respect to these check calibration points. While most of the rating curves appear to match the new calibration points well, there are issues which require attention or further investigation. In particular, issues with the original rating curves and v-notch transitions for WS 1, 2, 3 had been observed but never previously addressed. Updates to WS 8 and WS 10 are also being considered. This document explains the planned process to update the rating curves and discharge data. Rating curves for WS 6, 7, 9, Mack and for all WS when v-notches are on do not show a need for adjustment at this time but checking rating curves is a constant process. Because we have few recent calibration points at higher flows in a number of the watersheds, we need to systematically be collecting calibration points in the future.

This document describes proposed changes to rating curves based on field measurements and analyses of existing data. **If you have comments or concerns, please send them to Steve or me by Sept 21.**

Other updates will need to occur. We still need to address the precision of the discharge data based on the accuracy and repeatability of field measurements and calculations of rating curves. However, we have not worked through that methodology for calculating magnitude and sources of error that change throughout the annual hydrologic cycle, and how to display the changes in precision in the database.

Background information on rating curves:

A rating curve is developed for every weir (flume, v-notch, fish ladder) within the Andrews Forest. Typically a rating curve follows a power-law relationship between stage height and discharge. This relationship is linearized by log transforming both stage height and discharge and using a simple linear regression, or a piece-wise simple linear regression that each operate over a specified water depth (also called stage height) interval. The power-law for of the equation and its log transformed version are shown below:

$$Y = a\,b^{x} \quad \text{or} \quad \ln y = \ln a + b \ln x$$

(where y= flow as cfs and x = stage height in ft.

Each rating curve is customized for the specific weir by fitting the curve to a set of calibration points. Calibration points are collected to build a relationship between stage height (either measured using a floating hook gage in a stilling well or directly measured with a ruler in the flume) and stream discharge (calculated from velocity measurements times known areas, dye dilution, or through volumetric measurement).

WS 1, 2, 3 main flume rating curves:
Early calculations of discharge at WS 1, 2, 3 (1953-1958) were made using velocity measurements taken with a velocity-head rod multiplied times area of that measurement. Stage height measurements were also collected in the flume. Tables of rating curves were developed based on these measurement of discharge and stage height. After original hook gages were installed and properly adjusted, simultaneous flume and hook gage measurements (1957-1964) were used to generate calibration points using the hook gage- stage height reading and discharge determined through these original rating tables from the corresponding flume reading. These measured calibration points served as input for calculating the original rating curves (~1966) and have continued to be used (through WY 2017) at WS 1, 2 and 3.

Differences between new measurements and original rating curves can occur for several reasons: updated methods for measuring velocity and discharge are more precise compared to velocity-head rod, general settling of the weir that either changed the slope of the weir channel or its elevation relative to the hook gage, a change in roughness of the weir, or changes in shape of the stream channel and/or roughness of the upstream channel, streambed, and banks that change location and amplitude of the standing waves that have been observed within the flumes.

Issues for WS 1, WS 2 and WS 3:
WS 2: Calibration points have been collected WY 1996 to 2017. A new rating curve fit to these new calibration points show differences from original curve at lower stage heights and discharge. The new rating curve that includes these updated points shows that total annual yield would decrease by about 2%. Discharge calculated using new calibration points and rating curve will show less abrupt shifts when v-notch is installed or removed.

WS 3: Calibration points have been taken WY 1996 to 2017. A rating curve fit to these more recent calibration check points shows that discharge is comparatively lower or higher in separate sections in the lower portion of the curve. Overall, total annual yield would decrease by approximately 3% using the new curves.

WS 1: Calibration points were collected WY 1996 to 2017 and have been used to rebuild the rating curve for WS 1. This new rating curve shows a distinct change in discharge from the original rating equations. Initial calculations using the new curves show that total annual yield is 18 to 24% less than previous total annual yield estimated using the current (original) rating curves. The new equations also provide more consistent discharge values during transitions of putting on and taking off the v-notch weir. V-notch rating curves were developed since the initiation of v-notch measurements (1998). Updated rating curves will help smooth (but not entirely remove) shifts in calculated discharge that occur on v-notch installation/removal.

Proposed plans for updating discharge data at WS 1, 2, 3:
Recalculate discharge from these three watersheds using the new rating curves. For WS 2 and WS 3, replace existing data beginning with WY 1999 through WY 2016, and continue forward using these new curves. Save and provide the existing discharge data (WY 1999 – WY 2016) online as a separate entity in HF004. [For WS 1, recalculate discharge for the period from WY 1962 to the end of WY 2016 with the new curves and continue forward with these new curves. *Note: WS1 was only replaced from WY 1999 through WY 2016*] Save and provide previous discharge data at 5-minute resolution and place these data online as a separate entity in HF004.

Rationale:
Calibration points taken from 1998 to 2017 appear to be internally consistent over this period. The v-notches in these small watersheds were first installed during summer of WY 1999. Recalculation of discharge using new rating curves beginning in WY 1999 will provide consistency from this point forward. Rerunning data through the new curve beginning in WY 1999 would create a closer association of the main flume and v-notch rating curves, as both sets of rating points have been collected within the same time interval. The gaging station at WS3 was destroyed in the February 1996 flood and was rebuilt and a new hook gage installed in September 1998. V-notches were installed at all three of these watersheds in the summer of 1999. Rerunning these data beginning in WY 1999 would directly follow these significant changes. Additionally, the discharge jump caused when the v-notch is installed and replaced each year will be improved in most years.

There is a significant mismatch in WS 1 between current rating points and the original rating curve. There is evidence that 30% of this change (the difference in discharge from the original equation to the new equation) occurred around 1998. Calibration points appear to be consistent (linear regression on log-transformed data) since 1998. However, rerunning calculations of discharge beginning in WY 1999 (October 1998) will lead to a major shift in discharge starting 1998.

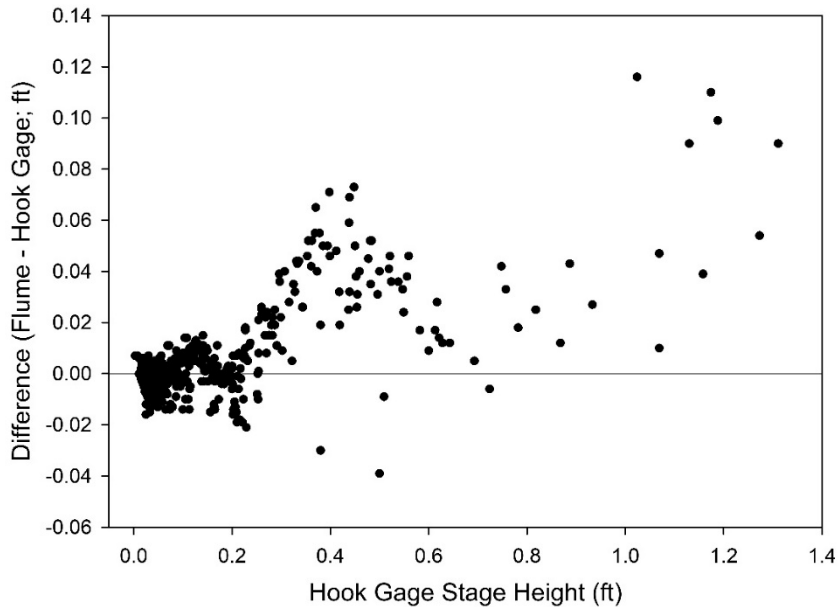
WS01 Gage: The Search for What Went Wrong Between 1962 and 1999

On each check date, water levels in the gage are recorded – in two ways. Five measurements of the water level in the flume are taken using a ruler. These 5 measurements are averaged to give the observed depth of water in the flume. Then, the water level in the stilling well is recorded from observing the hook gage. The hook gage was originally calibrated to the flume, and surveyed from a fixed bench mark. The hook gage and bench mark are surveyed annually. If they drift, then the hook gage is reset to the bench mark and the stage heights collected over the time interval from the current survey check to the previous check are corrected using a linear interpolation (i.e., assuming that the hook gage slowly and consistently drifted out of calibration over the course of the entire period). The original calibration and these annual surveys should keep the two measurements – ruler in the flume versus hook gage – closely aligned. That is, the relationship between the hook gage and flume should be stationary.

We can use that relationship to test whether there has been unaccounted drift in either the elevation of the flume or the hook gage. Changes in this relationship would appear as a consistent bias in the measured stage height.

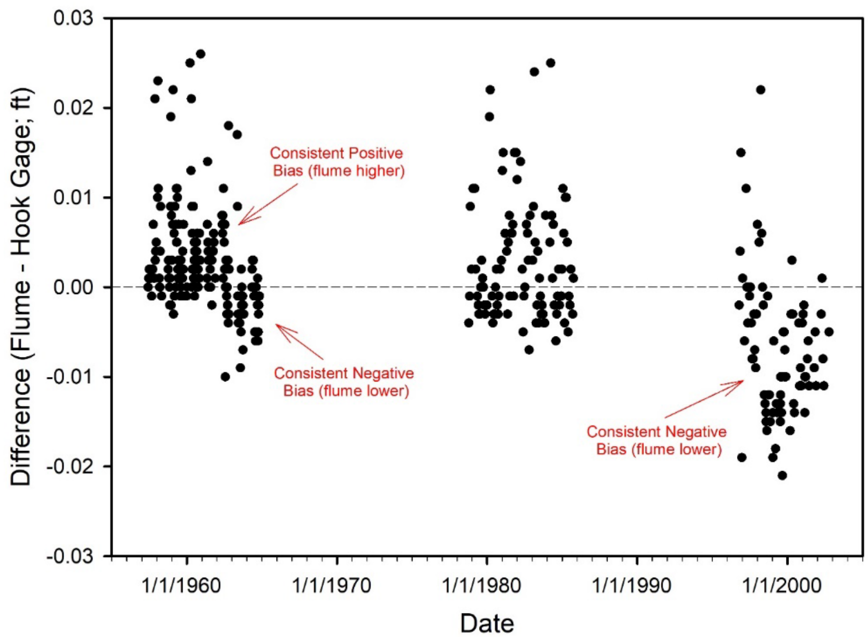
Note also that the relationship between the flume and hook gage appears biased at high discharge, with the water level measured with a ruler in the flume substantially higher than the measurement using the hook gage (Figure 1). This is not surprising. At high velocities, the water "piles up" along the leading edge of any obstruction, including a ruler.

Figure 1: Relationship between the accuracy of the water depth measured in the flume (expressed as the difference between the flume depth and the hook gage) versus stage height in the flume. Note – measurements made with the V-notch weirs in place have been removed.



Because of bias at higher flows, to test the possibility of a drift in the elevations of the flume versus the hook gage, we only want to examine the low flow data. Using the lower flow data (excluding all points with stage height > 0.300 ft), we then look to see how well the flume and hook gage measurements agree (Figure 2).

Figure 2: The relationship between the water depth in the flume versus the water level measured from the hook gage in the stilling well adjacent to the flume.



While these flume and hook gage measurements have been made at each watershed check (at least every 3 weeks) for the duration of the record, we have only transcribed data from the field check forms for three periods shown above: A) 1957-1964, B) 1978-1985, and C) 1996-2002.

Also note that any measurements made at high flow (i.e., stage height measured off the hook gage > 0.300 ft) have been deleted from this analysis.

- Even though the high discharge data have been removed (stage height > 0.300 ft), there are still several points when the flume measurement with the ruler was 0.010 to 0.030 ft higher than the hook gage. We suspect that there are "erroneous" readings.
- We should expect the errors to be small and somewhat evenly distributed above and below the zero line – as is seen in the middle cluster of points collected form the mid-1970s through the late 1980s. Further, the overall relationship should not change over time. So, while the data points collected between 1978 and 1985 are not perfectly evenly distributed, the general pattern and presence of high outliers should be consistent over time. That does not appear to be the case.
- The points collected in the 1950s and 1960s show two strange biases with the flume measurement consistently higher than the hook gage until the beginning of WY1962, after which the flume tends to read lower than the hook gage.
- By the late 1990s through the late 2000s, we see that the flume was consistently lower than the hook gage – and by a substantial amount (~0.010 feet) or 3 mm.
- Although a consistent bias of only 0.010 ft may seem like a very small amount, it is a substantial error when accounted for over the entire water year. Don Henshaw tested the influence of changing stage height by 0.010 ft and showed that an error of that magnitude would change total annual water yield by ~7%.
- We do not know why the flume and hook gage measurements should change and show a long-lasting bias. Some potential reasons:
  - The flume has settled over time, so while the hook gage and bench mark survey have been consistent, the calibration between the flume and hook gage has drifted. The data shown above seem to indicate that the change occurred rapidly, not from a slowly settling concrete slab. This explanation seems somewhat unlikely.
  - Settling of the bench mark. The bench mark is located on a concrete wing wall that rests on bedrock and prevents flow from bypassing the gage. The change in the flume – hook gage relationship does not appear to be consistent with a slowly settling concrete slab. This explanation seems somewhat unlikely.
  - Replacement of the hook gage after it was vandalized– during which time the survey to the bench mark was in error and that change has been fixed into the data.
  - Disturbance to the gage and flume, potentially including the flume, bench mark, and hook gage during a major flood.
  - An unspecified error made during the original or any of the annual survey checks on the gage.

Other Watersheds

WS 8: The H flume was replaced by trapezoidal flume in Oct 1987 (WY 1988) and a new hookgage installed. Twenty seven calibration points taken with a velocity meter and bucket points 1990-1996 were used to build the current rating curve. Checks made by plotting new calibration points on this rating curve are in general agreement, but the fuller set of rating points (1990-2017) shows improvement. An additional 44 points between 1997-2017 have been collected. These are a combination of low flow measurements with bucket catch and at higher flows, discharge calculated using velocity (from velocity meter) times area. When 71 points are used to build a new rating curve, there is a slight increase in in discharge, primarily at stage heights ranging from 0.25 to 0.40 ft. The new curve tends to increase total annual yield by ~3-4%.

Proposed plan for discharge data from WS 8 main flume: Rerun all data using the new rating curve beginning with WY 1988 through WY 2016, and continue forward using the new curve. Save and place the existing data (WY 1988 – WY 2016) online as a separate entity in HF004.

Rationale: The new curve uses more than double the number of calibration points than the original curve and improves the fit at the low end of the curve. All of the points from 1989 seem consistent overall and makes rerunning all of the data since this flume was constructed in 1988 the cleanest solution.
Caveat: There is still some uncertainty in the curve, particularly at stage heights ranging from 0.25 to 0.40 ft. range. This is likely the range contributing to the increases in calculated discharge. Collection of more rating points within this range would be desirable.

WS 10: 76 points based on the dye dilution points (collected 1975-1977) and velocity meter and bucket calibration points (collected 1996-2002) were used to build the rating curve. 28 additional check points match very well with the existing curve except at very low stage heights (<0.015 ft.). However, several new calibration points have been collected in this low range in recent years where none were available prior to 2002. The new curve uses the same equations as the original curve at stage heights above 0.037 ft.; Only the very low end range of calculated discharge will be changed. This has nearly no effect on total annual yield.
Proposed plan for discharge data from WS 10 main flume: Recalculate discharge using the new rating curve beginning with WY 1997 through WY 2016, and continue forward using the new curve. Save and place the existing data (WY 1997 – WY 2016) online as a separate entity in HF004.
Rationale: The new curve takes advantage of 28 additional calibration points taken at low levels to improve the low end fit of the curve. Recalculating discharge beginning in 1997 matches the year of the v-notch installation and will improve the transition from the main flume to v-notch rating curves. High flow discharge calculations will be unchanged.