PEAKFLOW PREDICTION USING AN ANTECEDENT PRECIPITATION INDEX IN SMALL FORESTED WATERSHEDS OF THE NORTHERN CALIFORNIA COAST RANGE

by

Gregg Bousfield

A Thesis

Presented to

The Faculty of Humboldt State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

In Natural Resources: Watershed Management

April, 2008

ABSTRACT

PEAKFLOW PREDICTION USING AN ANTECEDENT PRECIPITATION INDEX IN SMALL FORESTED WATERSHEDS OF THE NORTHERN CALIFORNIA COAST RANGE

Gregg Bousfield

The vast majority of small watersheds in Northwest California lack stream gage information. Understanding the high flow behavior of these watersheds is crucial for guiding resource managers in project planning. The purpose of this thesis was to develop a predictive relationship between precipitation and peakflow of streams draining small forested watersheds of the Northern California Coast Range. An antecedent precipitation index approach was developed for this purpose.

The five selected watersheds are covered by coastal coniferous forests with drainage areas ranging from 0.4 to 34 km^2. Streamflow and precipitation data from the South Fork of Caspar Creek was used to create the calibration model. Data from the North Fork of Caspar Creek, Hennington Creek, Little Lost Man Creek, and Freshwater Creek were used for independent model testing.

The calibration linear regression model, predicting peakflow as a function of peak antecedent precipitation index, resulted in a r^2 of 0.83 and a residual standard error of 1.20 L s^{-1} ha^{-1}. When peakflow was predicted, using precipitation data from test watersheds, the results were fair to poor with average absolute prediction errors ranging from 28.6 to 66.3 percent. When the ten largest peakflows were predicted separately, the average absolute prediction errors were significantly lower at 10.2 to 44.9 percent. The

model was positively biased at all test watersheds except Freshwater Creek. The root mean square error was within 15 percent of the calibration residual standard error at all test watersheds except Little Lost Man Creek.

The variability in prediction accuracy could be explained by changing unit-discharge relationships, heterogeneous lithologies, different cumulative land management effects, and spatial variation in precipitation intensity. Prediction errors were the greatest for the smallest peakflows, which may be due to greater variation in interception rates during small rainfall events. The antecedent precipitation index approach outlined in this study is best suited for predicting larger rather than smaller peakflow events that may be influenced more by factors other than short-term rainfall history.

ACKNOWLEDGEMENTS

I would like to thank Dr. E. George Robison, Jack Lewis, Dr. Andre Lehre, Dr. Andrew Stubblefield, and Dr. Gary Hendrickson for their extensive and helpful comments during the thesis review process. My committee chair, Dr. E. George Robison, was responsible for the subject and direction of this thesis. A special thanks to Jack Lewis for his statistical expertise and intimate knowledge of Caspar Creek Experimental Watershed. Thanks to Randy Klein for streamflow and precipitation data recorded at the Little Lost Man Creek gaging station. Thanks to Clark Fenton for streamflow and precipitation data recorded at the Freshwater Creek gaging station. I would also like to thank my true love who I met at the beginning of graduate school. She helped me become a better person and was supportive throughout the entire thesis process. My parents were very influential in the continuation of higher education.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF APPENDICES

INTRODUCTION

The prediction of streamflow in response to precipitation is a recurring theme in watershed management. Methodologies used to create rainfall-runoff models differ in both complexity and data requirements. Modeling strategies range from physically based to empirical. Physically based models use theoretical equations to simulate all runoff generation processes. Empirical models rely on statistical relationships between precipitation inputs and streamflow outputs. Most rainfall-runoff models are not purely physically based or empirical but lie somewhere in-between (Brooks et al. 1997).

Coefficients are required to adjust equations found in physical models due to the stochastic nature of hydrologic processes (Haan 2002). The majority of coefficients are derived using statistical techniques from experimental lab data. For example, infiltration rate coefficients are developed for different soil types by measuring dye wetting front movement rates on soil blocks in a lab. Even cultivated soils will show extreme variability in infiltration rates across the wetting front (Beven 2001). For these reasons, physically based models often have high costs and computational demands.

Empirical or black-box models rely on statistical relationships with little regard to the inherent physical processes. Black-box models require recalibration when applied to different climatic and geologic environments since they are strongly influenced by data. Black-box models are good for re-sizing stream crossing culverts on vast parcels of Federal lands where little data exists and economic incentives are low (Piehl et al. 1988, Cafferata et al. 2004). Simplicity and low cost are the strengths of black-box models.

Rainfall-runoff modeling remains an important tool in watershed management, although there continues to be a lack of simple modeling approaches to estimate peakflows in small forested catchments. Peakflow prediction in these watersheds is crucial for designing bridges, culverts, or channel habitat restoration structures. Unfortunately, natural resource practitioners may only have precipitation data available. Using an antecedent precipitation index (API) as the key variable for streamflow prediction has shown promise in environments with low data availability (Fedora 1987, Beschta 1990).

API was originally conceived to represent current soil moisture conditions in models predicting storm volume (Betson et al 1969, Kohler and Linsley 1951, Lee and Bray 1969, Sittner et al 1969). The universal form of an API equation is as follows:

$$API_t = API_{t-1} C + P_{\Delta t} \qquad (1)$$

where API_t is API at time t, $P_{\Delta t}$ is the precipitation occurring between times t-1 and t, and C is the recession coefficient. The theory of API is that earlier precipitation should have less influence on present streamflow response than recent precipitation. The recession coefficient represents the "memory" of a particular watershed by decaying the effect of accumulated rainfall at each time step.

A long-term API reflects seasonal moisture conditions while a short-term API reflects the most recent rainfall intensity governing peakflow response. The determination of the recession coefficient dictates whether a particular API decays rapidly or slowly. Besides *a priori* estimates, recession coefficients have been determined through optimization techniques (Moreda et al. 2006, Reid and Lewis 2007)

and physical parameters (Beschta 1990, Fedora 1987, Smakhtin and Masse 2000, Ziemer and Albright 1987).

Fedora (1987) developed an API methodology to predict storm hydrographs in small forested catchments of the Oregon Coast Range. His API was assumed to decay at the average rate of storm hydrograph recession. The relatively small watersheds used in the study had steep recession limbs resulting in a short-term API. Fedora's method resulted in average absolute peakflow and storm volume errors of 14.8 and 14.2 percent, respectively.

Beschta (1990) tested Fedora's methodology in tropical environments using data from a small catchment and a large river basin. Peakflow simulation of the four largest storms from the small catchment resulted in an average absolute error of 14 percent compared to 15.4 percent using a physically based model (Shade 1984). Peak stage of the three largest flood events was predicted with an average absolute error of 14.8 percent. Fedora's method may be widely applicable when the model has been locally re-calibrated. However, Beschta's study is the only published independent test of the methodology.

An API model was recently developed to detect changes in peakflows following experimental clearcut harvesting in the North Fork of Caspar Creek (Reid and Lewis 2007). Three different API components were used in a non-linear model ($r^2 = 0.84$) predicting daily peakflow. The components were assumed to represent quick, subsurface, and groundwater flow. Each component had different recession coefficients derived through optimization with quickflow having the fastest decay and groundwater flow the

slowest. When compared with Fedora (1987) and Beschta (1990), the API approach developed by Reid and Lewis (2007) was relatively complex.

The purpose of this study was to develop a simple API approach for modeling peakflow in small forested watersheds located on the Humboldt and Mendocino County Coasts of California. The recession coefficient used in this study was derived following the methodology developed by Fedora (1987). The API model created in this study was solely for peakflow prediction, unlike Fedora's (1987) method of simulating continuous hydrographs for both peakflow and stormflow volume prediction. The research questions were as follows:

1. Can streamflow and precipitation data from the South Fork of Caspar Creek consistently and accurately predict peakflow as a linear function of peak API?
2. Will an antecedent flow rate threshold improve model precision and accuracy?
3. Can the model consistently and accurately predict peakflow elsewhere in the Northern California Coast Range?
4. Does the model predict larger peakflows more accurately than smaller peakflows?

MATERIALS AND METHODS

Data Sources

The following criteria were used to select watersheds for API model development: forested watershed within 25 km of the Pacific Ocean, rain-dominated, drainage area less than 50 km^2, rain gage located within 5 km of the watershed centroid, gaging stations maintained and calibrated on a regular basis, streamflow and precipitation data resolution of one hour or finer, and five or more years of concurrent streamflow and precipitation data.

The distance from the Pacific Ocean was important to keep the analysis focused on coastal watersheds. Rain-dominated watersheds were sought to minimize the influence of snowmelt on streamflow generation. Small watersheds were necessary to study systems with less groundwater and channel routing influences (Gomi et al. 2002). Precipitation gages near the watershed centroid should better estimate average rainfall for the entire watershed. Poor stage-discharge relationships can have an error of 20 percent or more, which makes accurate gages a necessity (Rantz 1982). One hour or finer precipitation data is required since runoff in small watersheds responds rapidly to rainfall inputs (Beven 2001).

Gaging stations on the North and South Forks of Caspar Creek, Little Lost Man Creek, and Freshwater Creek met the criteria. There are other gaged watersheds in the region, but they lack a nearby rain gage or the data are only available at a daily time step. The South Fork of Caspar Creek was chosen as the calibration watershed due to its

moderate size and accurate data set. There was also a lack of forest harvesting at the South Fork of Caspar Creek during the period of concurrent streamflow and precipitation data. Table 1 compares basic gaging station characteristics.

All watersheds are dominated by mixed redwood (*Sequoia sempervirens*) and Douglas-fir (*Pseudotsuga menziesii*) forest. Soils are derived from the Franciscan geologic formation. The Franciscan formation contains a variety of lithologies, creating heterogeneous soils across the landscape (Woiska 1981). The Freshwater Creek watershed also contains the Yager and Wildcat formations, which are more consolidated than the Franciscan (Glass 2003). Figure 1 shows the relative location of the selected watersheds. Individual watershed maps are located in Appendix A through C.

Data Quality

The stream gaging stations have similar equipment, but different control structures. Unlike the other selected watersheds, Caspar Creek Experimental Watershed uses flumes and weirs for artificial control. Artificial control structures have empirically derived stage-discharge relationships that are relatively accurate (5 to 10 percent) and stable. The Freshwater Creek and Little Lost Man Creek gage sites are natural channels that aggrade and degrade though time.

Gage sites without artificial control require routine stage-discharge re-calibration. Randy Klein, the primary hydrologist at Redwood National Park, does not have confidence in peakflows above 3.0 L s^{-1} ha^{-1} at the Little Lost Man Creek gage site after the 1997 water year due to a lack of rating curve measurements and changes in control

Table 1 Gaging station characteristics.

	Calibration Watershed	Test Watersheds			
	South Fork of Caspar Creek	Hennington	North Fork of Caspar Creek	Freshwater Creek	Little Lost Man Creek
Distance from Pacific Ocean, km	6	7	6	15	5
Elevation Range, m	50 - 330	130 - 320	85 -320	25 - 850	60 - 650
Drainage Area, km^2	4.2	0.4	4.7	34	9.1
Rain Gage to Watershed Centroid, km	2	0.8* / 2**	1.5* / 1.5**	5	3
Years of Concurrent Streamflow and Precipitation	18	18	18	6	5

* North Fork Caspar Creek (N408) tipping bucket rain gage.
** North Fork Caspar Creek (N620) tipping bucket rain gage.

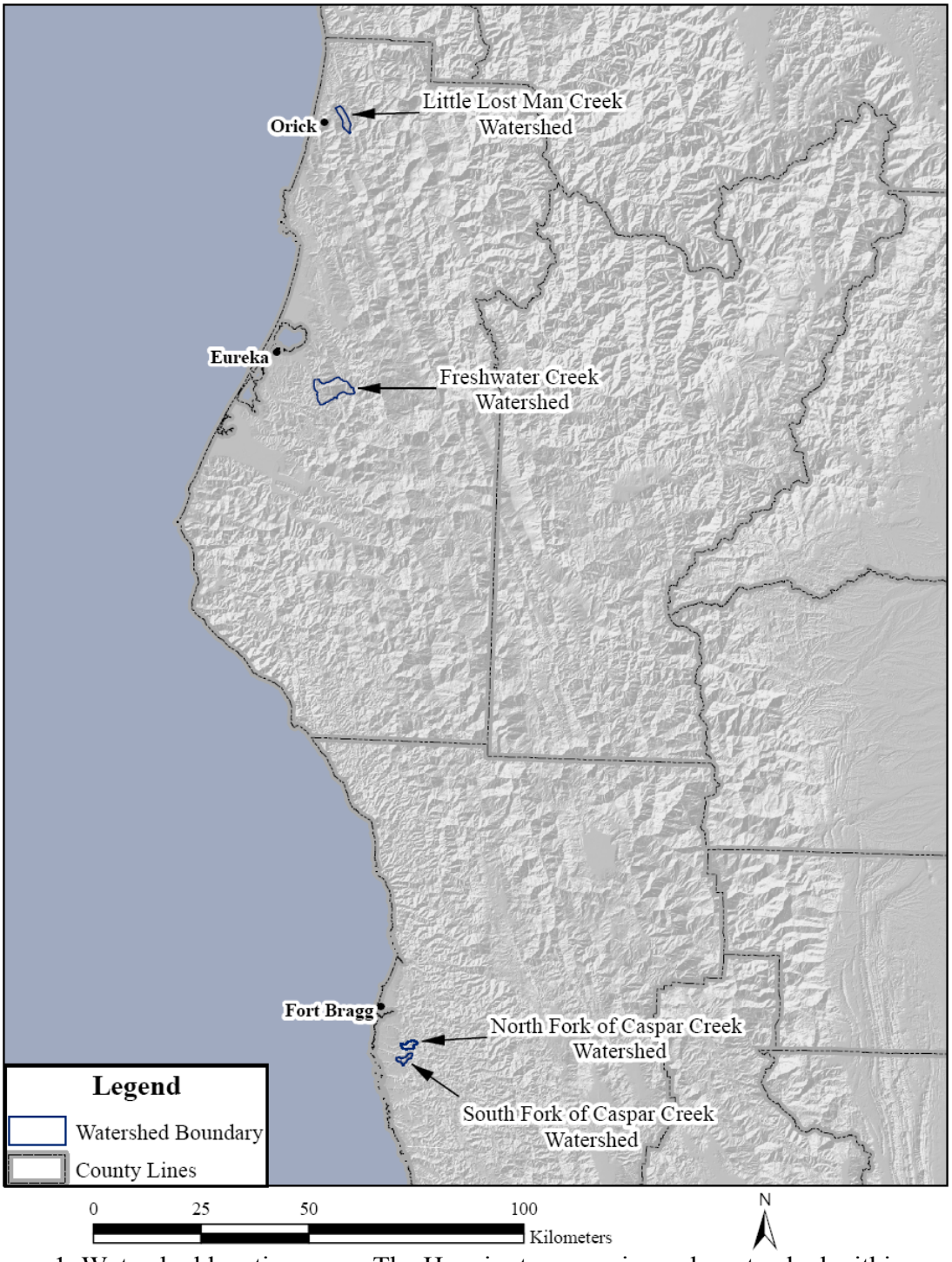


Figure 1 Watershed location map. The Hennington gage is a sub-watershed within the North Fork of Caspar Creek.

section geometry (Klein 2007, personal communication). Therefore, streamflow and precipitation data recorded at the Little Lost Man Creek gage site after 1997 was not used in this study.

Pressure transducers and tipping bucket rain gages have inherent error tolerances. All of the selected watersheds use similar pressure transducers to measure stage with an accuracy of 0.003 meters. Campbell Scientific tipping bucket gages are used at Little Lost Man Creek, Freshwater Creek, and the South Fork of Caspar Creek. The North Fork of Caspar Creek uses a Sierra Misco tipping bucket gage. Rain gage errors are five percent for intensities less than 8.0 cm hr^{-1} (Lewis 2007, personal communication).

API Model Development

The following steps were taken to develop the API model: frequency analysis, hydrograph recession analysis, API calculation, storm event analysis, and least squares regression modeling. Frequency analysis was undertaken to select events with peakflows whose return periods exceed one-year. The analysis used the annual maximum peakflows recorded at the South Fork of Caspar Creek from 1964 to 2004. The one-year peakflow (Q_1) was determined using the Log Pearson III method (Haan 2002). Selective harvesting that occurred during this period did not have a significant effect on annual maximum peakflows (Ziemer 1998).

Corresponding discharge hydrographs and rainfall hyetographs from the South Fork of Caspar Creek (1987 to 2004) were analyzed for their possible use in recession analysis. Recession analysis refers to the systematic observation of hydrograph recession

limbs in order to determine the average rate of discharge decline (Sujona et al. 2004). This analysis used recession limbs of peakflows exceeding Q_1 with data of fair or better quality. Hydrographs were eliminated if additional impulses of rainfall greater than 0.1 cm hr^{-1} or secondary peakflows occurred during the recession limb. These measures were taken to select recession limbs that best represent the recession characteristics of the South Fork of Caspar Creek to discrete rainfall events.

Recession limbs were defined as starting at the peak discharge and ending where Hewlett and Hibbert's (1967) 0.0055 L s^{-1} ha^{-1} baseflow separation line intersected the falling limb. Figure 2 provides an example of the recession limb selection process. Discharge from the selected recession limbs was plotted against discharge lagged by one hour. Following the methodology of Fedora (1987), the slope of the linear regression line was assigned to the recession coefficient in Equation 1.

Hourly time series' of API's were calculated using data from the S620 rain gage in South Fork of Caspar Creek (Equation 1). Calculations ran throughout the water year, since the rapidly decaying API of a prior event should have an insignificant influence after one or two days. For example, after rainfall ceases a recession coefficient of 0.90 will decay API to less than 10 percent of its peak value after 22 hours.

Matching hourly time series' of streamflows and API's from the South Fork of Caspar Creek, (1987 to 2004) were closely investigated. The following storm event attributes were investigated for peakflows exceeding Q_1: peakflow discharge rate, antecedent flow rate, peak API, data quality codes (Figure 3). Successive peakflows occurring on the same hydrograph had to be greater than 24 hours apart and recede by

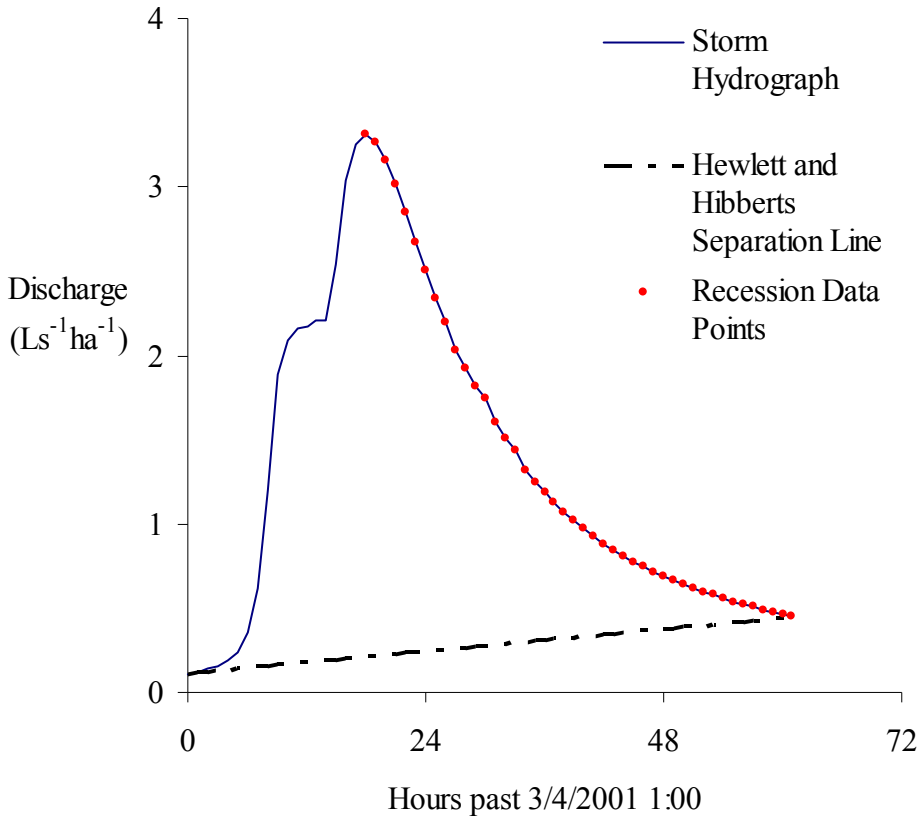


Figure 2 An example of a recession limb from a storm hydrograph recorded at the South Fork of Caspar Creek. Recession limbs began at the peakflow discharge and ended at the point where Hewlett and Hibbert's (1967) baseflow separation line intersects the hydrograph.

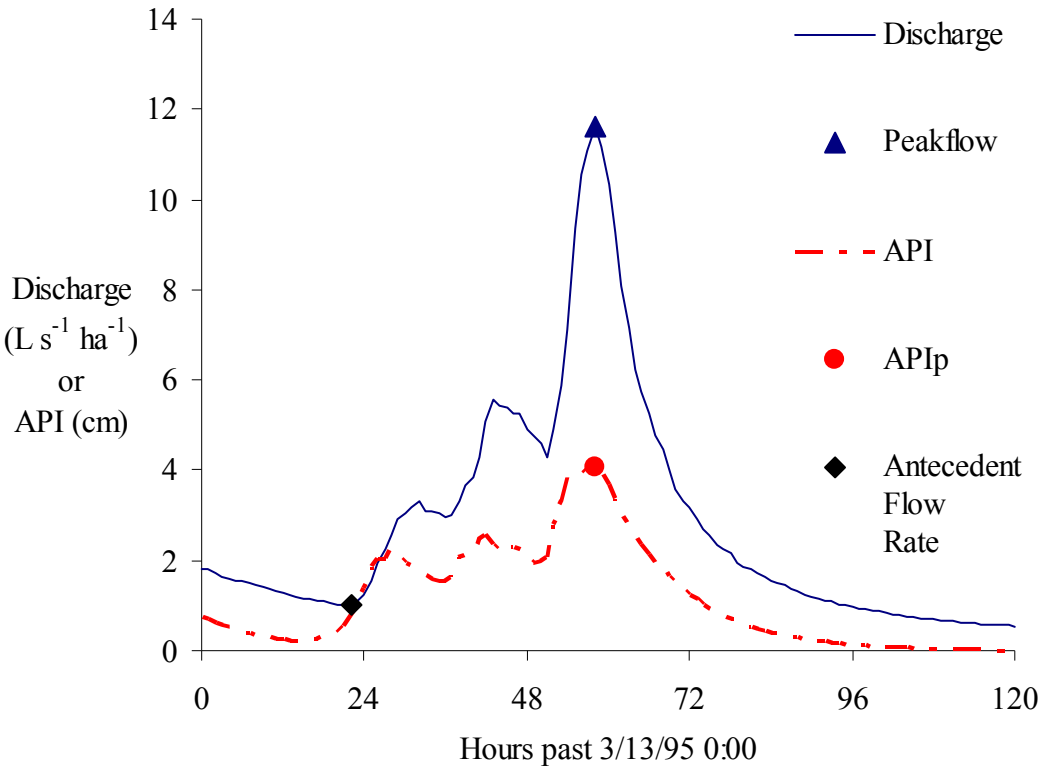


Figure 3 Hourly time series of discharge and API were plotted together to select corresponding peak API (API_p) values and peakflows exceeding Q_1. The previous peaks in the hydrograph were not recorded since they did not recede to less than half of their peak discharge and occurred within 24 hours of the largest peakflow.

50 percent of their peak discharge. Peakflows must rise to double the antecedent flow rate when they occur during hydrograph recession. These restrictions ensure relatively independent peak API values.

Data quality codes for both discharge and rainfall were investigated for each event. Storm events were excluded from this study when the codes indicated poor calibration, large gaps, or data reconstruction of either rainfall or discharge. Data quality codes for discharge were available only for Caspar Creek and Freshwater Creek. Caspar Creek was the only watershed with rain gage quality codes.

Fedora's (1987) API method revealed a different relationship for storm events occurring after extended periods without rain. It was thought that these "dry" events had a lower peakflow response due to a low water table and unsaturated soils. In this study, these conditions were investigated by recording a given storm event's antecedent flow rate. Scatterplots of peakflow as a function of peak API were studied to set an antecedent flow rate threshold that separated "wet" versus "dry" events.

The goal of least squares regression was to create a simple model of peakflow as a function of peak API. A data set must meet a set of assumptions in order to use least squares regression analysis for statistical inference. Since a best fit relationship for peakflow predictions was the main goal of this study, these assumptions were not strictly necessary, but were explored nonetheless. Outliers were first inspected using residual diagnostic techniques, since they can greatly influence the regression modeling results. Outliers could express missing independent variables or multiple populations (Haan 2002). Tests of normality ensured that the residuals were normally distributed.

Autocorrelation was tested using the Durbin-Watson statistic (Hintze 2004).

Independent Model Testing

API calculation and storm event analysis were repeated on the test watersheds. The one-year peakflow, hydrograph recession coefficient, and antecedent flow rate threshold were the same in the test and calibration watersheds. This was necessary to test the method as if rain gages were the only source of data available at the test watersheds. All restrictions applied to the calibration data set were also applied to data sets from the test watersheds for consistent evaluation of model performance.

Bias, precision, and accuracy were used to measure model prediction performance. Statistics used to calculate relative bias, precision, and accuracy were average prediction error, standard deviation of the prediction error, and average absolute prediction error, respectively (Walther and Moore 2005). The prediction error for each observation was calculated using the following equation (Green and Stephenson 1986):

$$E = (Q_p - Q_o / Q_o) * 100 \qquad (2)$$

where E is the prediction error, Q_p is the predicted peakflow in L s^{-1} ha^{-1} and Q_o is the observed peakflow in L s^{-1} ha^{-1}. The average and average absolute prediction error were calculated using the following equations (Green and Stephenson 1986):

$$E_m = (\sum E) / n \qquad (3)$$

$$E_a = (\sum | E |) / n \qquad (4)$$

where E_m is the average prediction error, E_a is the average absolute prediction error, and n is the sample size.

An absolute measure of model accuracy compared the residual standard error (RSE) of the calibration model to the root mean square error (RMSE) of the predicted regression line. The only difference between these two terms is that the sum of the squared residuals is divided by n-2 in the RSE compared to n in the RMSE. The n-2 is used for the calibration model to account for the information used up in estimating the slope and intercept. Model fit was evaluated using the r^2 from the regression of observed versus predicted peakflows. Model fit was also evaluated by testing whether the slope was significantly different from one and the intercept was significantly different from zero (95 percent confidence).

RESULTS

API Model Development

Forty-one annual maximum peakflows were recorded for South Fork Casper Creek with a mean and standard deviation of 10.3 and 5.08 L s^{-1} ha^{-1}, respectively. The largest peakflow on record had a maximum discharge rate of 21.5 L s^{-1} ha^{-1}. All peakflows exceeding Q_1 (2.0 L s^{-1} ha^{-1}) were investigated for their use in hydrograph recession and storm event analysis.

Nineteen recession limbs over the 18 years of record (1987 to 2004) for South Fork of Casper Creek met the stated requirements for hydrograph recession analysis. The associated peakflows had a mean and standard deviation of 6.35 and 5.13 L s^{-1} ha^{-1}, respectively. Segments exceeding 7.5 L s^{-1} ha^{-1} were removed from five recession limbs, since they accounted for 2.5 percent of the discharge observations. This may be explained by an unusually rapid recession following the largest peakflows. Peakflow generation with a greater proportion of saturation overland flow may explain the rapid recession. A regression of discharge lagged by one-hour for 758 discharge observations from the 19 recession limbs is shown in Figure 4. The slope of the linear regression line (0.91) was assigned as the API recession coefficient.

With the estimated recession coefficient of 0.91, API decayed by 90 percent in 26 hours. The time between peakflow events averaged 15 days, but varied from one to 135 days. Only one storm event occurred within 26 hours of a prior event. Peak API would have been reduced by 14 percent if the API time series were reset to zero between these

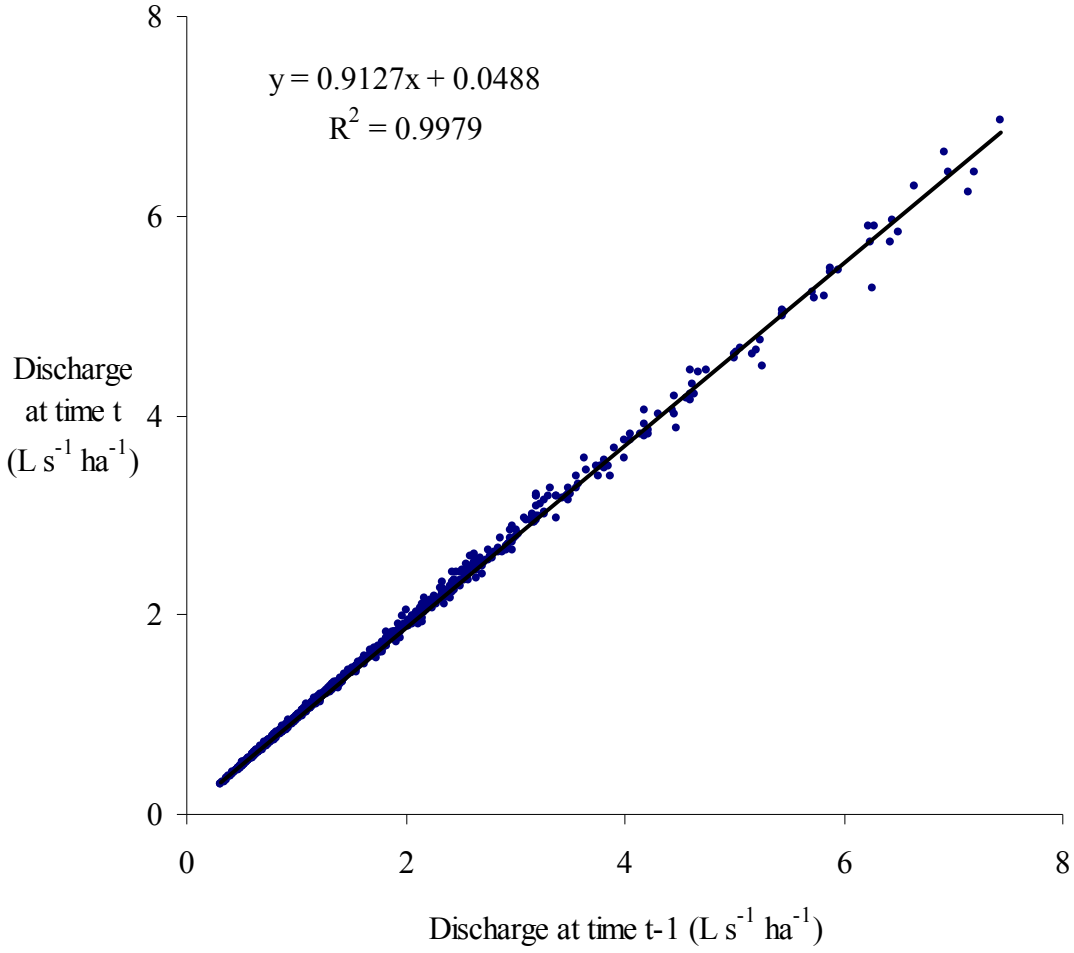


Figure 4 One-hour lag plot of hourly discharge from the South Fork of Caspar Creek.

two events. All peak API's were assumed to be independent since preceding observations had little to no influence on subsequent observations.

A data set of 71 storm events was initially analyzed. A scatterplot of peakflow as a function of peak API had a r^2 equal to 0.60 with a RSE of 2.13 L s^{-1} ha^{-1} (Figure 5). A subset of peakflows, with antecedent flow rates below 0.1 L s^{-1} ha^{-1}, was substantially smaller for a given peak API. Therefore, an antecedent flow rate threshold was set to remove these 12 "dry" events from the original data set. The remaining 59 events had an average peakflow of 5.67 L s^{-1} ha^{-1} and an average peak API of 3.14 cm (Table 2).

A visual inspection of peakflow as a function of peak API reveals a positive relationship. Residual diagnostics indicated that the largest peakflow, which occurred on March 24, 1999, was an outlier (Appendix D through H). Field notes on March 24, 1999 indicate that the V-notch weir was submerged by 0.5 feet (Lewis 2007, personal communication). Average event rainfall agreed to within 10 percent, and one-hour maximum rainfall agreed to within 15 percent at the three Caspar Creek tipping bucket gages. Yet the peakflow recorded at the North Fork of Caspar Creek had a 35 percent lower unit-area discharge rate than that of the South Fork. The March 24 1999 event was removed due to this large deviation in peakflow coupled with the residual diagnostic results.

All residual tests indicated the assumptions of normality were reasonable (alpha = 0.05). The Modified Levene test showed that the residual variance was not constant. Least squares regression analysis was continued regardless of this failure since a best fit for peakflow prediction was the main goal of this study. The Durbin-Watson

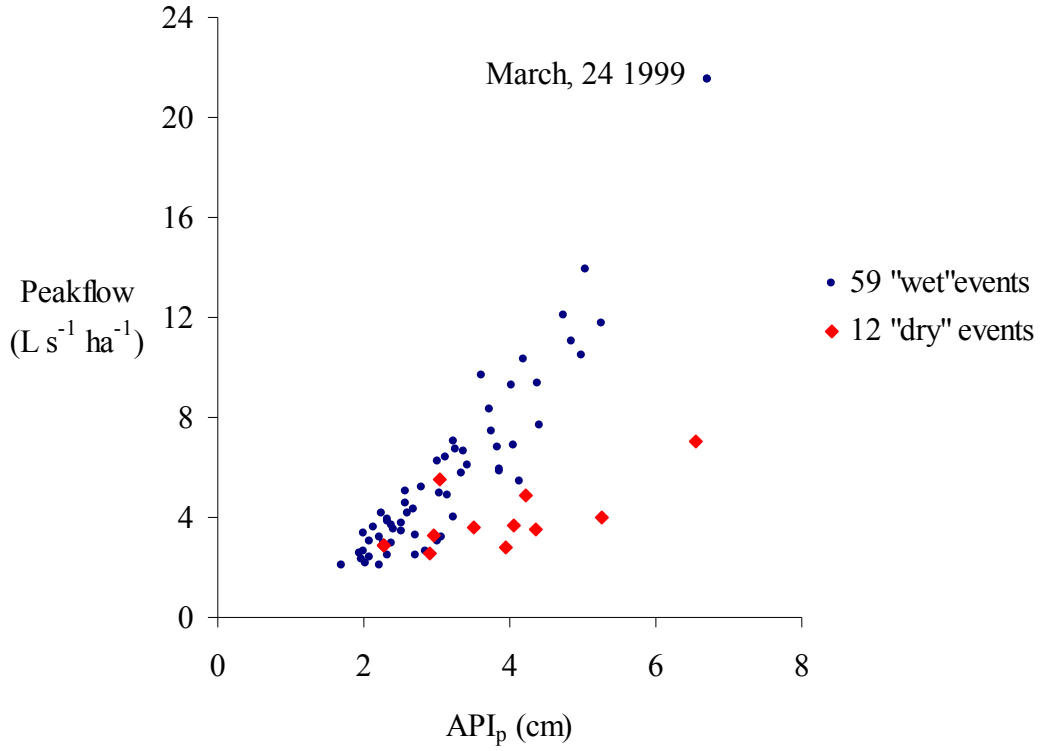


Figure 5 Scatterplot of the 71 selected events with the twelve "dry" events labeled. Storm events were considered "dry" when their antecedent flow rate was below 0.1 L s^{-1} ha^{-1}. The largest event was recorded on March 24 1999.

Table 2 Summary statistics after twelve "dry" storms out of 71 selected events were removed. Storm events were considered "dry" when their antecedent flow rate was below 0.1 L s^{-1} ha^{-1}.

	Peakflow, L s^{-1} ha^{-1}		Antecedent Flow Rate, L s^{-1} ha^{-1}		API_p, cm	
n	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
59	5.67	3.56	0.54	0.46	3.14	1.02

test for autocorrelation confirmed that peak API values were independent of one another. All tests of regression assumptions are summarized in Appendix I. Summary of the final storm event statistics are listed in Table 3.

The regression model used to predict peakflow as a function of peak API is:

$$Q_p = -3.52 + 2.90 * (API_p) \quad (5)$$

where Q_p is predicted peakflow in L s^{-1} ha^{-1} and API_p is peak API in cm. The r^2 was equal to 0.83 with a RSE of 1.20 L s^{-1} ha^{-1}. The slope term was highly significant ($p < 0.0001$). Figure 6 shows the least squares regression line along with the upper and lower 95 percent Working-Hotelling simultaneous confidence bands (Hintze 2004). These are the confidence bands for all possible values of peak API along the regression line. Additional regression statistics are located in Appendix J. The resulting model may only be applicable for peak API within a range of 1.71 to 5.25 cm. Peak API must be greater than 1.21 cm since lower values will result in negative predicted peakflows.

Independent Model Testing

The results of the API Calculation and Storm Event Analysis on the test watersheds are summarized in Table 4. The North Fork of Caspar Creek had the most observations, while Freshwater Creek had the fewest. The North Fork of Caspar Creek had the largest mean peakflow and peak API, while Freshwater Creek had the smallest.

Peakflow was initially predicted twice at the North Fork of Caspar Creek and Hennington since two rain gages were available. The N408 tipping bucket rain gage was

Table 3 Summary statistics after the March 24, 1999 outlier was removed.

	Peakflow, $L\ s^{-1}\ ha^{-1}$		Antecedent Flow Rate, $L\ s^{-1}\ ha^{-1}$		API_p, cm	
n	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
58	5.40	2.90	0.54	0.47	3.08	0.91

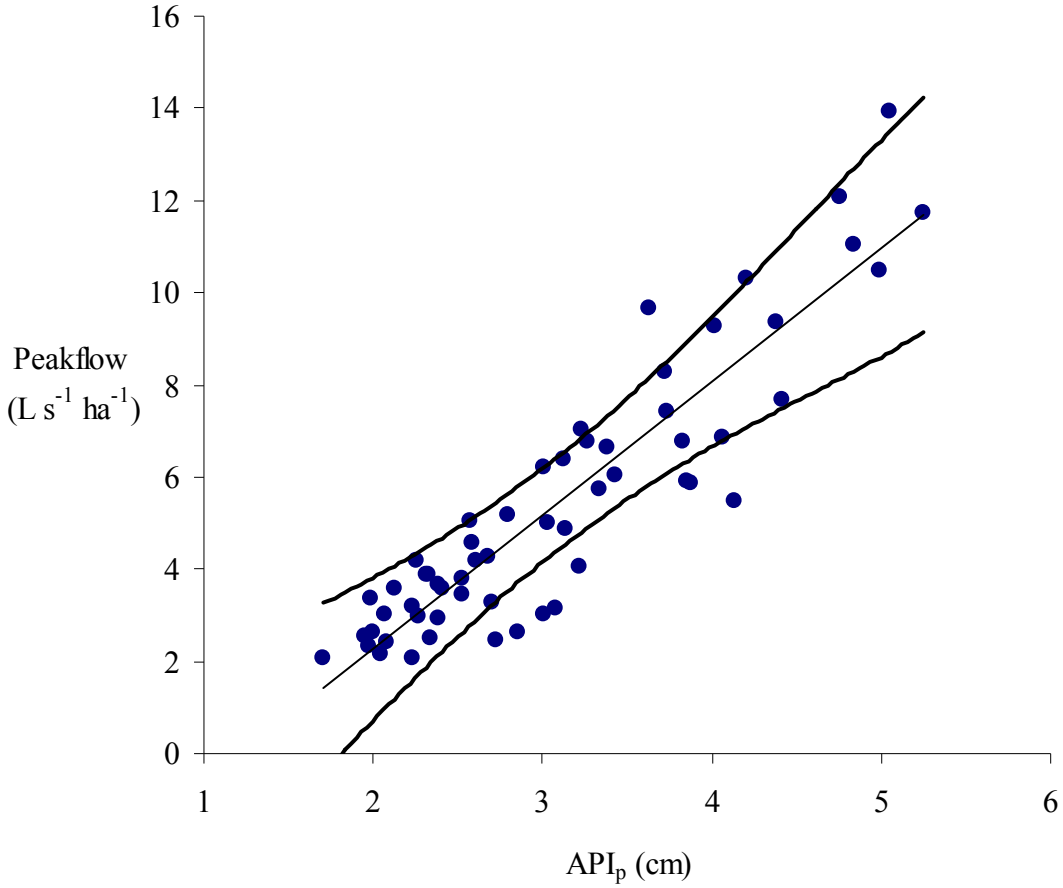


Figure 6 Linear regression line fitted to the 58 selected events along with the upper and lower 95 percent Working-Hotelling confidence bands (bold lines).

Table 4 Summary statistics for selected storm events from the test watersheds.

Test Gaging Station	n	Peakflow, $L\ s^{-1}\ ha^{-1}$		Antecedent Flow Rate, $L\ s^{-1}\ ha^{-1}$		API_p, cm	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Hennington	39	5.09	2.72	1.22	1.01	3.24	0.98
North Fork of Caspar Creek	49	5.70	2.98	0.79	0.55	3.49	0.97
Little Lost Man Creek	35	4.82	4.95	0.74	0.59	3.46	1.44
Freshwater Creek	32	4.77	3.20	0.69	0.33	2.47	0.89

retained for analysis with Hennington since it produced the best results. Similarly, the N620 tipping bucket rain gage was retained for analysis with the North Fork of Caspar Creek.

Figures 7 through 10 show the prediction error (Equation 2) for each storm event at the test watersheds. All test watersheds showed a decrease in prediction error with increase in peakflow. Unlike the other test watersheds, the majority of the peakflows were under predicted at Freshwater Creek. Little Lost Man Creek had the largest over prediction with almost a third of the errors exceeding 100 percent. Eighty percent of the prediction errors ranged from -50 to 50 percent at all test watersheds, except Little Lost Man Creek.

Bias, precision, and accuracy are summarized in Table 5. The model was positively biased at all test watersheds except Freshwater Creek. Little Lost Man Creek had the lowest precision at 54.2 percent compared to Hennington at 31.5 percent. Little Lost Man Creek had the lowest accuracy at 66.3 percent compared to Hennington at 28.6 percent.

Bias, precision, and accuracy for the ten largest peakflows are summarized in Table 6. The model was positively biased for the ten largest peakflows at all test watersheds except Freshwater Creek. Precision ranged from 12.7 percent at the North Fork of Caspar Creek to 42.8 percent at Little Lost Man Creek. Accuracy ranged from 10.2 percent at the North Fork of Caspar Creek to 44.9 percent at Little Lost Man Creek.

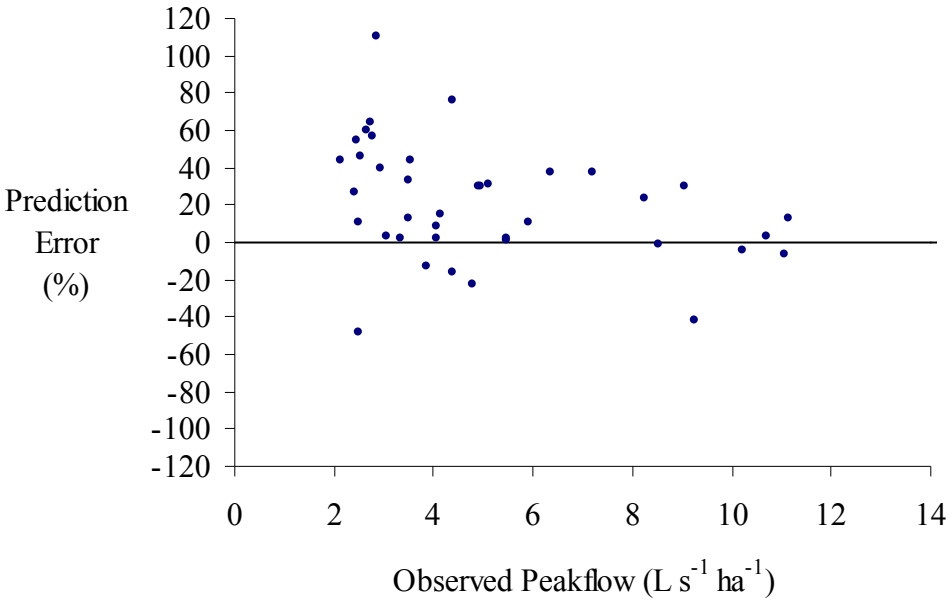


Figure 7 Model prediction errors at the Hennington test watershed show a decrease in variability as peakflows increase in magnitude.

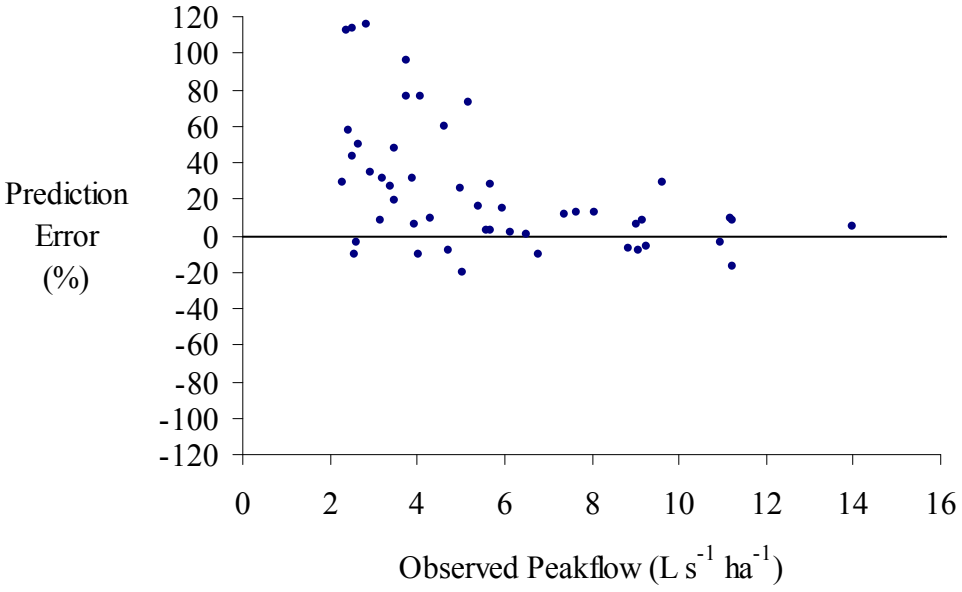


Figure 8 Model prediction errors at the North Fork of Caspar Creek shows a decrease in variability as peakflows increase in magnitude.

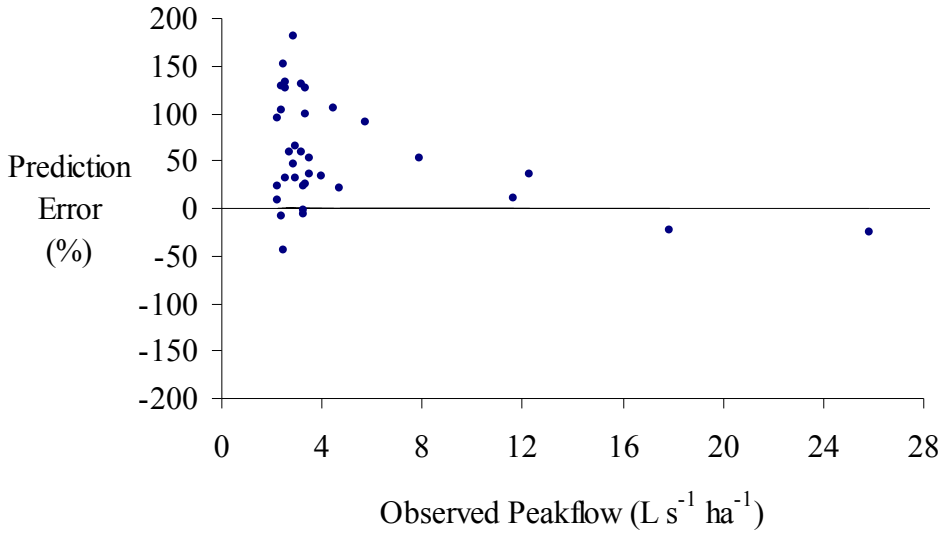


Figure 9 Model prediction errors at Little Lost Man Creek show a decrease in variability as peak flows increase in magnitude.

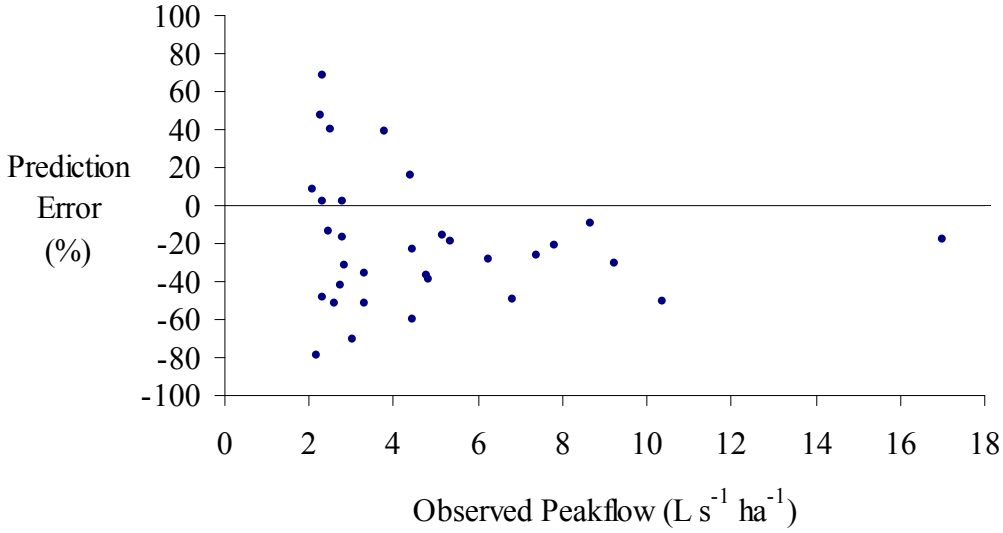


Figure 10 Model prediction errors at Freshwater Creek show a decrease in variability as peakflows increase in magnitude.

Table 5 Bias, precision and accuracy of predicted peakflows at the test watersheds.

		Bias	Precision	Accuracy
		E		
Station	n	Average	Standard Deviation	E_a
		(%)		
Hennington	39	20.7	31.5	28.6
North Fork of Caspar Creek	49	24.5	35.1	29.0
Little Lost Man Creek	35	62.4	54.2	66.3
Freshwater Creek	32	-20.1	34.4	34.0

Table 6 Bias, precision and accuracy for the ten largest peakflows at the test watersheds.

	Observed Peakflow		Bias	Precision	Accuracy
			E		
	Mean	Standard Deviation	Average	Standard Deviation	E_a
Station	L s^{-1} ha^{-1}	L s^{-1} ha^{-1}		(%)	
Hennington	10.2	1.38	4.79	20.2	15.3
North Fork of Caspar Creek	10.5	1.57	3.11	12.7	10.2
Little Lost Man Creek	9.80	7.30	35.5	42.8	44.9
Freshwater Creek	8.50	3.45	-24.6	15.8	24.6

Figure 11 through 14 show regressions of observed versus predicted peakflow at the test watersheds. These contrast with Figures 7 through 10 by showing absolute rather than percentage error. Most peaks were over predicted at the test watersheds except Freshwater Creek. Only the two largest peakflows were under predicted at Little Lost Man Creek. An exponential relationship was observed in Figure 13. This suggests a non-linear relationship between peakflow and peak API at Little Lost Man Creek.

Table 7 lists the least squares regression statistics of the observed versus predicted from the test watersheds. The slope terms were not different from zero and the intercept terms were not different from one (alpha = 0.05). The North Fork of Caspar Creek had the strongest correlation ($r^2 = 0.82$). Hennington and the North Fork of Caspar Creek had the lowest RMSE at 1.27 and 1.26, respectively. Little Lost Man Creek had the lowest r^2 and highest RMSE due to a non-linear relationship.

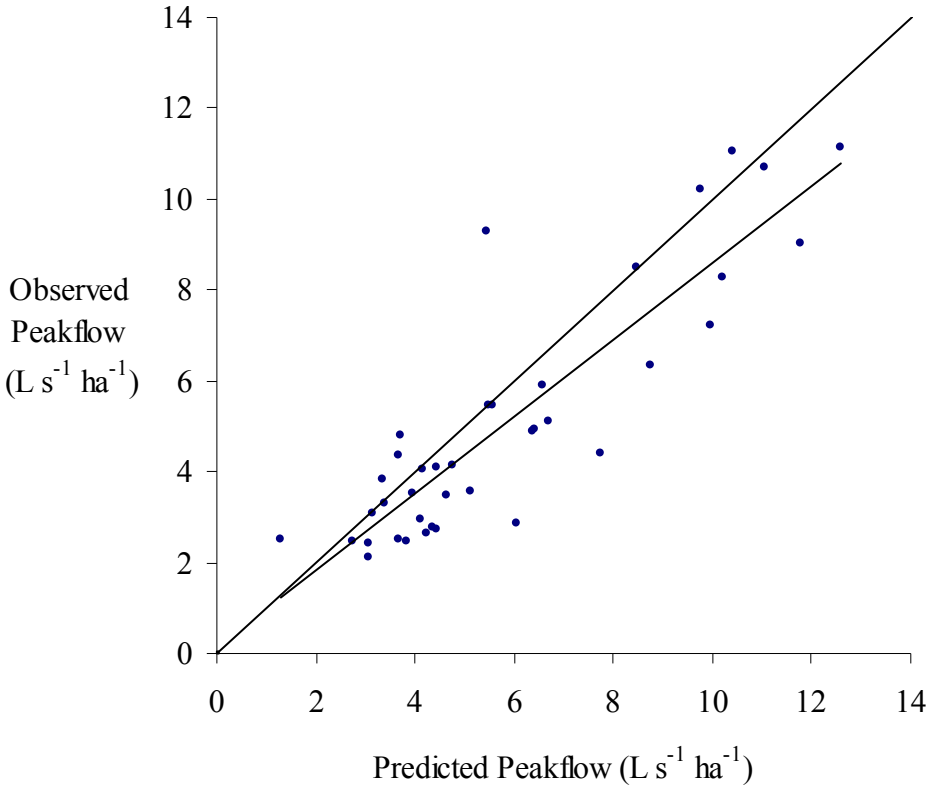


Figure 11 Observed versus predicted peakflow of the 39 events selected from Hennington. The one to one line of perfect agreement is displayed to compare with the linear regression line.

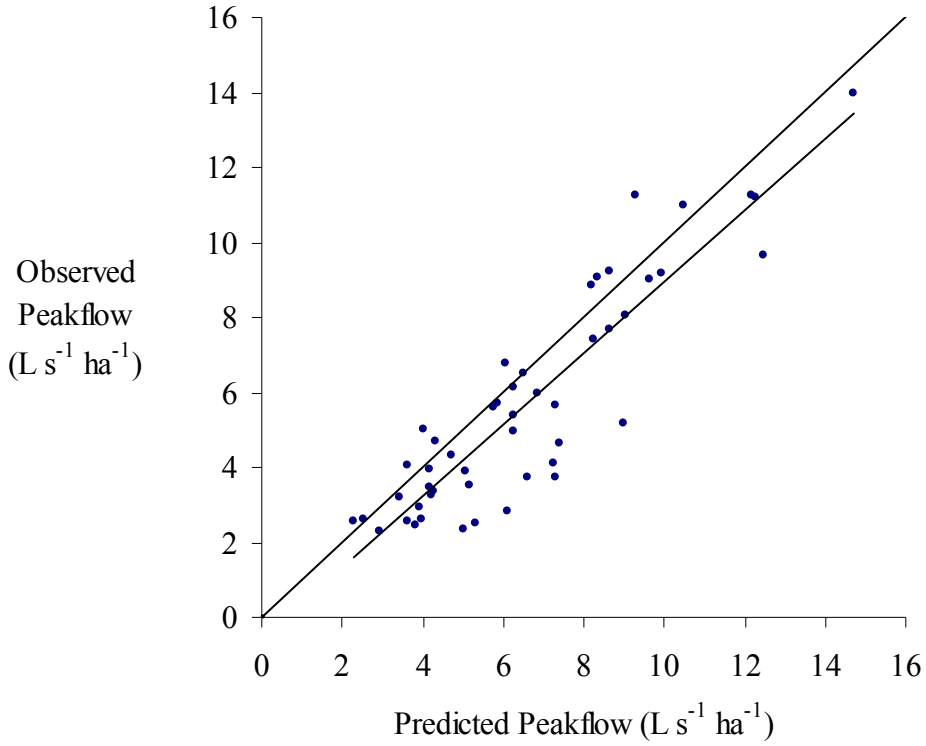


Figure 12 Observed versus predicted peakflow of the 49 events selected from the North Fork Caspar Creek. The one to one line of perfect agreement is displayed to compare with the linear regression line.

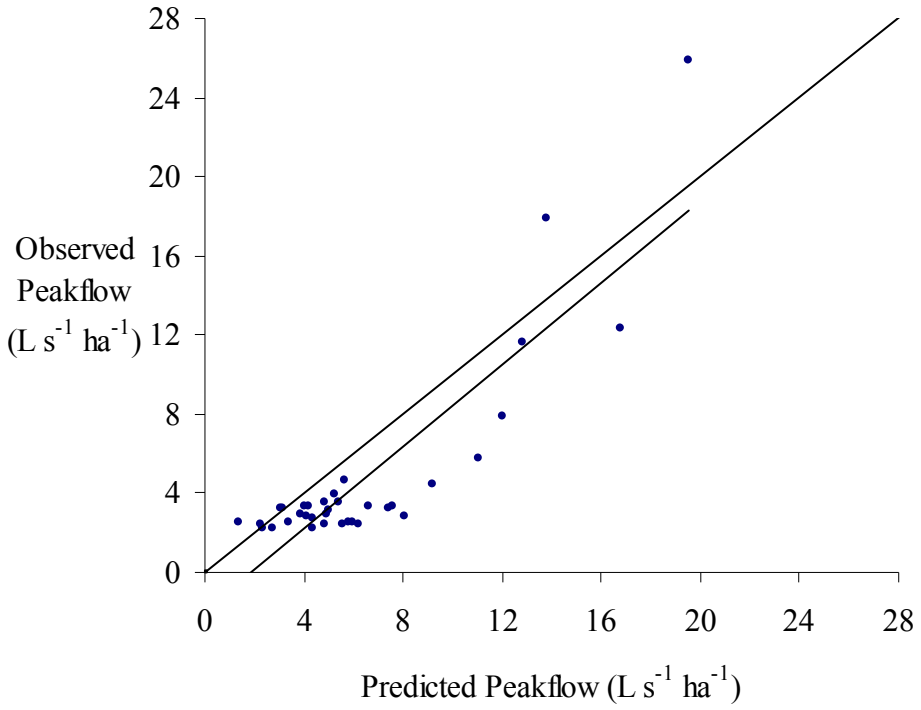


Figure 13 Observed versus predicted peakflow of the 35 events selected from Little Lost Man Creek. The one to one line of perfect agreement is displayed to compare with the linear regression line.

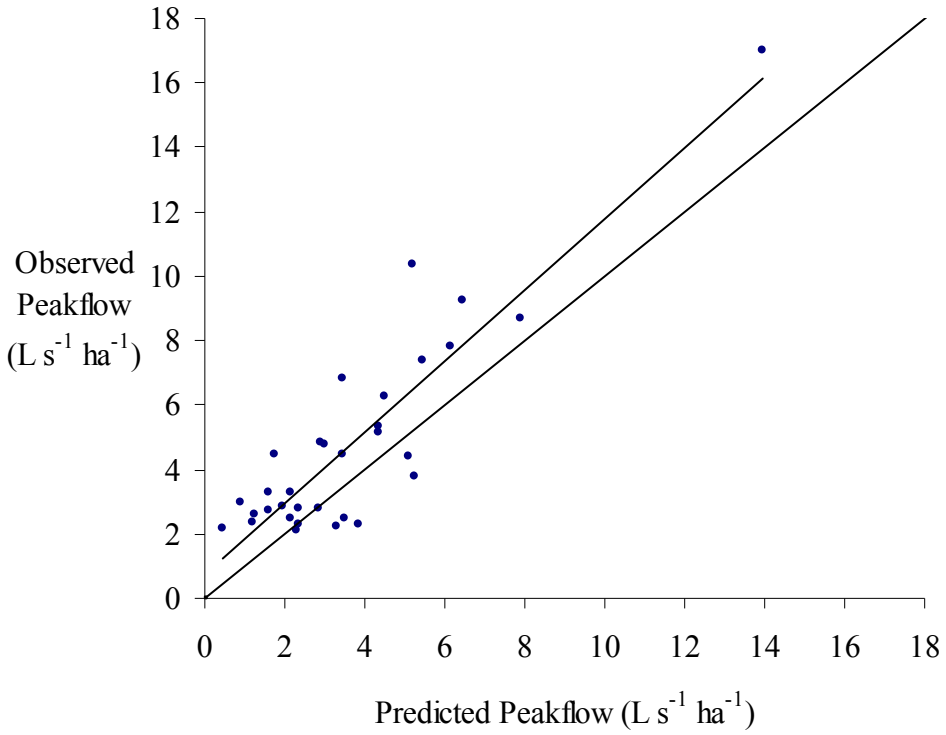


Figure 14 Observed versus predicted peakflow of the 32 events selected from Freshwater Creek. The one to one line of perfect agreement is displayed to compare with the linear regression line.

Table 7 Least squares regression statistics of the observed versus predicted peakflow from the test watersheds.

Station	n	Slope	Intercept	r^2	RMSE
Hennington	39	0.85	0.13	0.78	1.27
North Fork of Caspar Creek	49	0.95	-0.58	0.82	1.26
Little Lost Man Creek	35	1.04	-1.92	0.76	2.38
Freshwater Creek	32	1.11	0.73	0.80	1.42

DISCUSSION

A positive correlation exists between peakflow and peak API at the calibration watershed. The intercept is negative making the model only valid for peak API's above 1.21 cm. Variability in the relationship between peakflow and peak API was characterized by several measures. The r^2 value indicated that peak API explained 83 percent of the variability in peakflow. The residual standard error was 21 percent of the average peakflow. Nineteen percent of the observations fell outside of the confidence bands. Variability can be attributed to a simple linear regression model being used to predict complex non-linear rainfall-runoff processes. These processes, which include rainfall intensity, interception, evapotranspiration, soil hydraulic conductivity, pipeflow, and local saturation overland flow, vary spatially and temporally over a watershed throughout a storm event (Beven 2001).

The relationship between peakflow and peak API showed that "dry" events with antecedent flow rates below 0.1 L s^{-1} ha^{-1} produced substantially smaller peakflows for a given peak API. The calibration model had a 28 percent lower r^2 and a 44 percent higher RSE prior to the removal of the twelve "dry" events. The results are similar to those of Fedora (1987) in that a recession coefficient based on hydrograph recession analysis caused peak API to decay so quickly that long-term antecedent moisture conditions were not properly addressed.

The muted streamflow response with low antecedent flow was most likely due to soil moisture and shallow groundwater deficits occurring after prolonged periods of

drought. Greater antecedent flow indicates higher soil moisture and an elevated water table, creating a larger saturation overland flow response to rain. However, exploratory multiple regression analysis revealed that antecedent flow rate was not a reliable variable throughout the range of peakflows analyzed in this study.

Antecedent flow rate was not related to peakflow or peak API, but proved a reliable threshold indicator of catchment wetness. Lynch and Corbett (1982) explored the relationship between antecedent flow rate, antecedent soil moisture and hydrograph parameters. Antecedent soil moisture was a steep function of antecedent flow rate that flattened to a slope of zero above 0.05 L s^{-1} ha^{-1}, which is relatively close to the threshold set in this study. The small watersheds in this study, like those studied by Lynch and Corbett (1982), have relatively "flashy" and more ephemeral streamflow response than larger watersheds due to less groundwater interaction in holding and releasing flows.

Both consistent under or over prediction at the test watersheds may be due to variability in unit-area discharge relationships. Unit-area discharge had less variability in watersheds larger than 10 km^2 in drainage area (Robinson et al. 1995). Ziemer and Rice (1990) found that mean flow path had a significant positive association with lag-time and an insignificant negative association with unit-area discharge of progressively larger sub-watersheds within the North Fork of Caspar Creek. These results indicate that hillslope processes strongly control streamflow response in the North Fork of Caspar Creek.

Unlike the other test watersheds, the API model was negatively biased for Freshwater Creek. One would expect the API model to be biased to over predict, instead of under predict at Freshwater Creek, since channel roughness and bank storage should

increase lag-time and flatten peakflow response in larger watersheds (Gomi et al. 2002). A combination of clearcut and selective harvesting from 1989 to 1999 removed roughly 82 percent of the timber volume above the stream gage (Glass 2003). Reid and Lewis (2007) indicated a 29 percent increase in rainfall that reaches the forest floor after clearcut timber harvesting. The under prediction of peakflows at Freshwater Creek is most likely due to lower interception and evapotranspiration rates.

The consistent over prediction at the other three watersheds could be due to skid trails in the South Fork of Caspar Creek. Soil compaction due to legacy skid trails could cause overland flow, which artificially extends the natural drainage system. An overland flow component may not have been captured in this API methodology. This phenomenon is less prominent in the North Fork of Caspar Creek since cable yarding produced less soil compaction when compared to selective tractor yarding (Ziemer 1998).

Fedora (1987) analyzed the largest annual events, which resulted in only six to 20 events from his study watersheds. Beschta's (1990) test of Fedora's methodology only looked at four peakflows and three flood events. In contrast, my study looked at every peakflow exceeding Q_1, which resulted in 32 to 58 events from the selected watersheds. Accuracy for peakflow prediction ranged from 10.4 to 30.4 percent in Fedora's (1987) study and 14 to 14.8 percent in Beschta's (1990) study compared to 28.6 to 66.3 percent in my study. Higher variability was expected in my study because the data set represents peakflow response over a wider range of rainfall intensities, amounts, and antecedent soil moisture conditions.

The API model predicted peakflows at Hennington better than the other test watersheds with 28.6 percent accuracy. The North Fork of Caspar Creek was equally accurate at 29.0 percent. Predicted peakflow at Freshwater Creek was 32 percent more accurate than at Little Lost Man Creek. As expected, the RMSE for prediction in the test watersheds exceeded the calibration RSE. Prediction at the North Fork of Caspar Creek (RMSE = 1.26 L s^{-1} ha^{-1}) and Hennington (RMSE = 1.27 L s^{-1} ha^{-1}) was only slightly less accurate than in the calibration watershed, South Fork of Caspar Creek (RSE = 1.20 L s^{-1} ha^{-1}). Little Lost Man Creeks RMSE was 98 percent greater than the calibration RSE. Freshwater Creek had a RMSE 18 percent greater than the calibration RSE, which was surprisingly better than Little Lost Man Creek.

The regression of observed versus predicted peakflow at Little Lost Man Creek revealed a positive exponential transition from larger to smaller peaks. This suggests that the linear relationship used in this study was not adequate for peakflow prediction at Little Lost Man Creek. An exponential relationship between peakflow and peak API should increase the predictive capability at Little Lost Man Creek. Although not explored in this study, an exponential transformation of peak API may be useful to increase prediction power in future applications of this methodology.

The South Fork of Caspar Creek may not truly represent the processes that control streamflow generation at the test watersheds. Errors in peakflow prediction could be due to localized geologic and pedologic variability. The South Fork of Caspar Creek may have greater connectivity in soil macropores and pipes, creating a faster response and generating larger peakflows. Even though the watersheds have relatively similar

geology, heterogeneous lithology could restrict preferential flow paths. The geological formations in the Oregon Coast Range watersheds used by Fedora (1987) may not have as much variability in localized lithology as in the Northern California Coast Range.

Rainfall variability over a given watershed is very hard to quantify unless a dense network of rain gages is present. Rain gages are sparse throughout the Northern California Coast Range, although Caspar Creek Experimental Watershed is an exception. Rainfall intensity can vary significantly within one km (Singh 1997). Individual storms could have errors in rainfall measurements up to 75 percent due to the effects of wind and location (Dingman 2002). Due to orographic influences on rainfall amounts and intensities, rain gages misrepresent a watershed's actual mean rainfall. Erroneous rainfall data may have been used to calculate peak API at the other test watersheds, since only Caspar Creek had rain gage error codes.

The different gaging station control structures could have also influenced model performance. Without artificial control, the location of a gaging station can greatly affect the accuracy and consistency of streamflow measurements. None of the stream gages in this study met all of the criteria for optimal stream gage location (Rantz 1982). It is very hard to find a location in these small watersheds where the stream course is straight for 100 m upstream and downstream. Stage data quality was not available to remove erroneous data at Little Lost Man Creek.

When the ten largest peakflows were analyzed separately, the API model had a higher accuracy of 10.2 to 44.9 percent. The average accuracy of the predicted peakflows at the test watersheds was increased by 40 percent. The test of the API

methodology, like the findings of Fedora (1987) and Beschta (1990), revealed that the largest peakflows on record had the lowest errors. These are promising results for flood prediction since the largest peakflows in this study had return periods which ranged from a 4-year to a 10-year event.

Better prediction of these large events was most likely due to a simplification of physical processes once the soils are saturated and macropores reach their maximum flow rate (Ziemer and Lisle 1998). This may also be explained by decreased variability of interception rates as peakflows increased in magnitude (Link et al. 2004, Pypker et al. 2005, Reid and Lewis 2007). Smaller events could have greater variability in the interactions between the processes that control streamflow generation. These interactions were not addressed in this study.

LITERATURE CITED

Beschta, R.L. 1990. Peakflow estimation using an antecedent precipitation index (API) model in tropical environments. Pages 128-137 in R.R. Ziemer, C.L. O'loughlin, and L.S. Hamilton, editors. Research needs and applications to reduce sedimentation and erosion in tropical steeplands. International Association of Hydrologic Sciences. Publication 192. Washington, D.C.

Betson, R.P., R.L. Tucker, and F.M. Haller. 1969. Using analytical methods to develop a surface runoff model. Water Resources Research 5: 103-111.

Beven, K. J. 2001. Rainfall-runoff modeling: the primer. John Wiley & Sons, Ltd., New York, New York.

Brooks, K.N., P.F. Ffolliott, H.M. Gregersen, and L.F. DeBano. 1997. Hydrology and the management of watersheds. Iowa State University Press, Ames, Iowa.

Cafferata, P., T. Spittler, M. Wopat, G. Bundros, and S. Flanagan. 2004. Designing watercourse crossings for passage of the 100-year flood flows, wood, and sediment. California Department of Forestry and Fire Protection, California Forestry Report No. 1. Sacramento, California

Dingman, S.L. 2002. Physical hydrology. Prentice Hall, Upper Saddle River, New Jersey.

Fedora, M. A. 1987. Simulation of storm runoff in the Oregon Coast Range. Bureau of Land Management Technical Note 378. Denver, Colorado.

Glass, D. 2003. Freshwater Creek watershed analysis cumulative effects assessment. Report of Watershed Professionals Network to Pacific Lumber Company (PALCO), Scotia, California.

Gomi, T., R.C. Sidle, and J.S. Richardson. 2002. Understanding processes and downstream linkages of headwater systems. BioScience 52: 905-916.

Green, I.R.A. and D. Stephenson. 1986. Criteria for comparison of single event models. Hydrological Sciences 31: 395-411.

Haan, C. T. 2002. Statistical methods in hydrology. Iowa State University Press, Ames, Iowa.

Hewlett, J. D. and A. R. Hibbert. 1967. Factors affecting the response of small watersheds to precipitation in humid areas. Pages 275-290, in W.S. Sapper and H.W. Lull, editors. International Symposium on Forest Hydrology. Pergamon Press, New York, New York.

Hintze, J. 2004. NCSS and PASS. Number cruncher statistical systems. Kaysville, Utah. Available at www.ncss.com. Contacted on May 20, 2007.

Kohler, M. A., and R. K. Linsley. 1951. Predicting the runoff from storm rainfall. United States Weather Bureau Research Paper 34. Washington, D.C.

Lee, J. and D. I. Bray. 1969. The estimation of runoff from rainfall for New Brunswick watersheds. Journal of Hydrology 9: 427-437.

Link, T.E., M. Unsworth and D. Marks 2004. The dynamics of rainfall interception by a seasonal temperate rainforest. Agricultural and Forest Meteorology 124: 171-191.

Lynch, J. A. and E. S. Corbett. 1982. Relationship of antecedent flow rate to storm hydrograph components. Pages 73-77, in A.I. Johnson and R.A. Clark, editors. International Symposium on Hydrometeorology June 13-17, 1982, Denver, Colorado. American Water Resources Association.

Moreda, F., V. Koren, Z. Zhang, S. Reed, and M. Smith. 2006. Parameterization of distributed hydrological models: learning from experiences of lumped modeling. Journal of Hydrology 320: 218-237.

Piehl, N.E., M.R. Pyles, and R.L. Beschta. 1988. Flow capacity of culverts of the Oregon Coast Range forest roads. Water Resources Bulletin 24: 631-637.

Pypker, T.G., B.J. Bond T.E. Link, D. Marks, and M.H. Unsworth. 2005. The importance of canopy structure in controlling the interception loss of rainfall: Examples from a young and an old-growth Douglas-fir forest. Agricultural and Forest Meteorology 130: 113-129

Rantz, S.E. 1982. Measurement and computation of streamflow: Vol. 1 Measurement of stage and discharge. United States Geological Survey, Water Supply Paper 2175. Washington, D.C.

Reid, L.M. and J. Lewis. 2007. Rates and implications of rainfall interception in a coastal redwood forest. Pages 107-117 in R.B. Standiford, G.A. Giusti, Y. Valachovic, W.J. Zielinski, and M.J. Furniss, editors. Proceedings of the redwood region forest science symposium: What does the future hold? United States Department of Agriculture, Forest Service, Pacific Southwest Research Station, General Technical Report 194. Albany, California.

Robinson, J. S., M. Sivapalan, and J. D. Snell. 1995. On the relative role of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments. Water Resources Research 31: 3089-3101.

Shade, P.J. 1984. Hydrology and sediment transport, Moanalua Valley, Hawaii. United States Geological Survey, Water Resources Investigation Report 84-4156. Honolulu, Hawaii.

Singh, V.P. 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrological Processes 11: 1649-1669.

Sittner, W. T., C. E. Schauss, and J. C. Monroe. 1969. Continuous hydrograph synthesis with an API-type model. Water Resources Research 5: 1007-1022.

Smakhtin, V. Y. and B. Masse. 2000. Continuous daily hydrograph simulation using duration curves of a precipitation index. Hydrological Processes 14: 1083-1100.

Sujono, J., S. Shikasho, and H. Hiramatsu. 2004. A comparison of techniques for hydrograph recession analysis. Hydrological Processes 18: 403-413.

Velleman, P.F. and R.E. Welsch. 1981. Efficient computing of regression diagnostics. The American Statistician 35: 234-242.

Walther, B.A. and J.L. Moore. 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography 28: 815-829.

Woiska, E. P. 1981. Hydrologic properties of one major and two minor soil series of the Coast Ranges of Northern California. Master of Science Thesis, Natural Resources, Humboldt State University, Arcata, California.

Ziemer, R. R. 1998. Flooding and Stormflows. Pages 15-24, in R. R. Ziemer, editor. Proceedings of the conference on coastal watersheds: The Caspar Creek story. United States Department of Agriculture, Forest Service, Pacific Southwest Research Station, General Technical Report 168. Albany, California.

Ziemer, R. R. and J.S. Albright. 1987. Subsurface pipeflow dynamics of north-coastal California swale systems. Pages 71-80, in R. L. Beschta, T. Blinn, G. E. Grant, G. G. Ice, and F. J. Swanson, editors. Erosion and sedimentation in the Pacific Rim. International Association of Hydrological Sciences. Publication 165. Washington, D.C.

Ziemer, R.R. and T. E. Lisle. 1998. Hydrology. Pages 43-68, in R.J. Naiman and R.E. Bilby, editors. River ecology and management: lessons from the Pacific Coastal Ecoregion. Springer-Verlag, New York, New York.

Ziemer, R. R. and R. M. Rice. 1990. Tracking rainfall impulses through progressively larger drainage basins in steep forested terrain. Pages 413-420, in H. Lang and A. Musy, editors. Hydrology in mountainous regions. 1 – Hydrological measurements; the water cycle. International Association of Hydrological Sciences. Publication 193. Washington, D.C.

PERSONAL COMMUNICATIONS

Klein, R. 2007. Personal Communications. Redwood National and State Parks, 1655 Heindon Road, Arcata, CA 95521

Lewis, J. 2007. Personal Communications. Redwood Sciences Laboratory, Pacific Southwest Research Station, USDA Forest Service, 1700 Bayview Drive, Arcata, CA 95521

LIST OF VARIABLES AND ACRONYMS

API = Antecedent precipitation index, cm

API_p = Peak API, cm

C = Recession coefficient, dimensionless

E = Prediction error for each observation, (%)

E_a = Average absolute prediction error, (%)

E_m = Average prediction error, (%)

$P_{\Delta t}$ = Precipitation occurring between times t-1 and t, cm

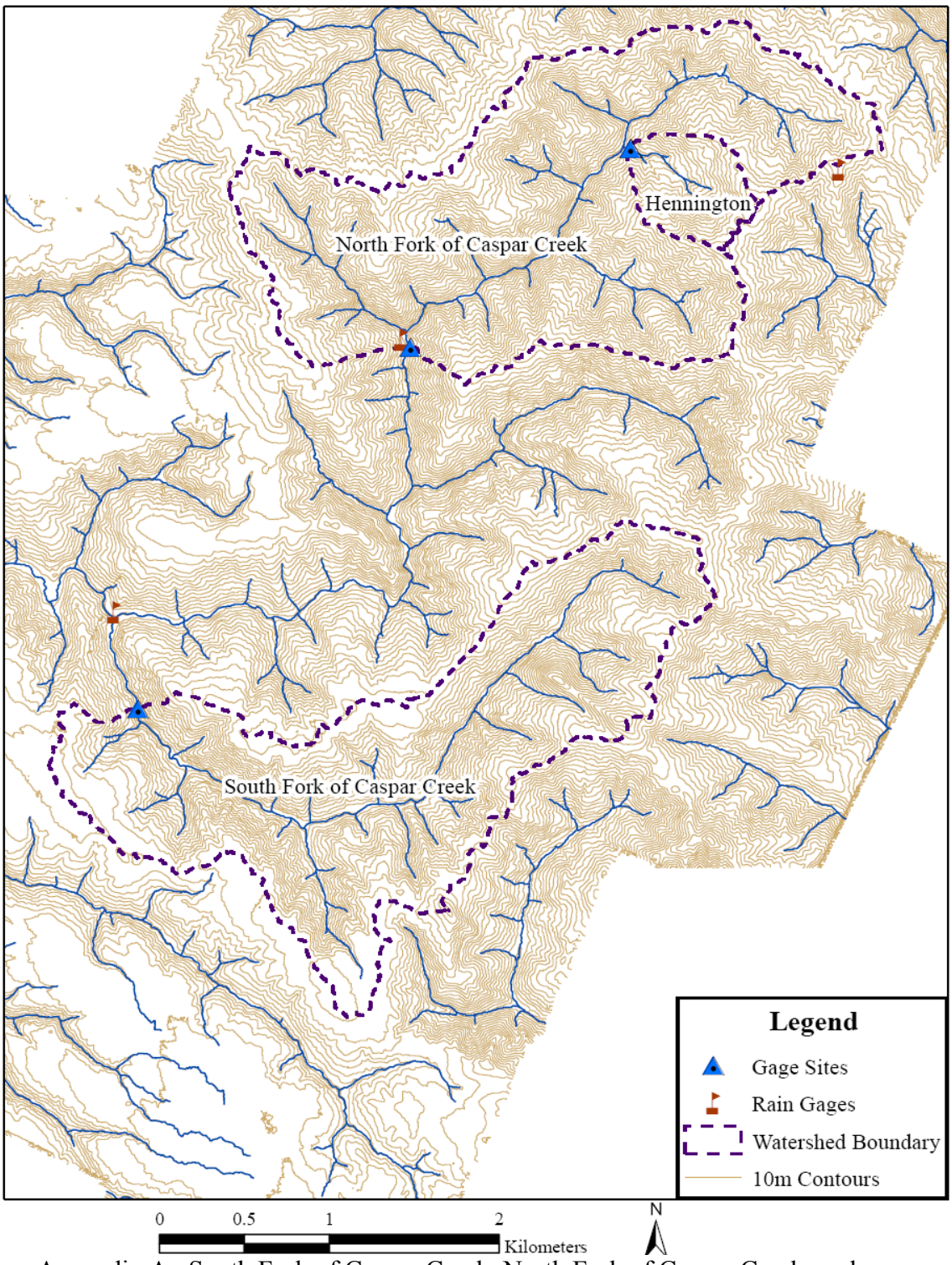
RMSE = Root mean square error, L s^{-1} ha^{-1}

RSE = Residual standard error, L s^{-1} ha^{-1}

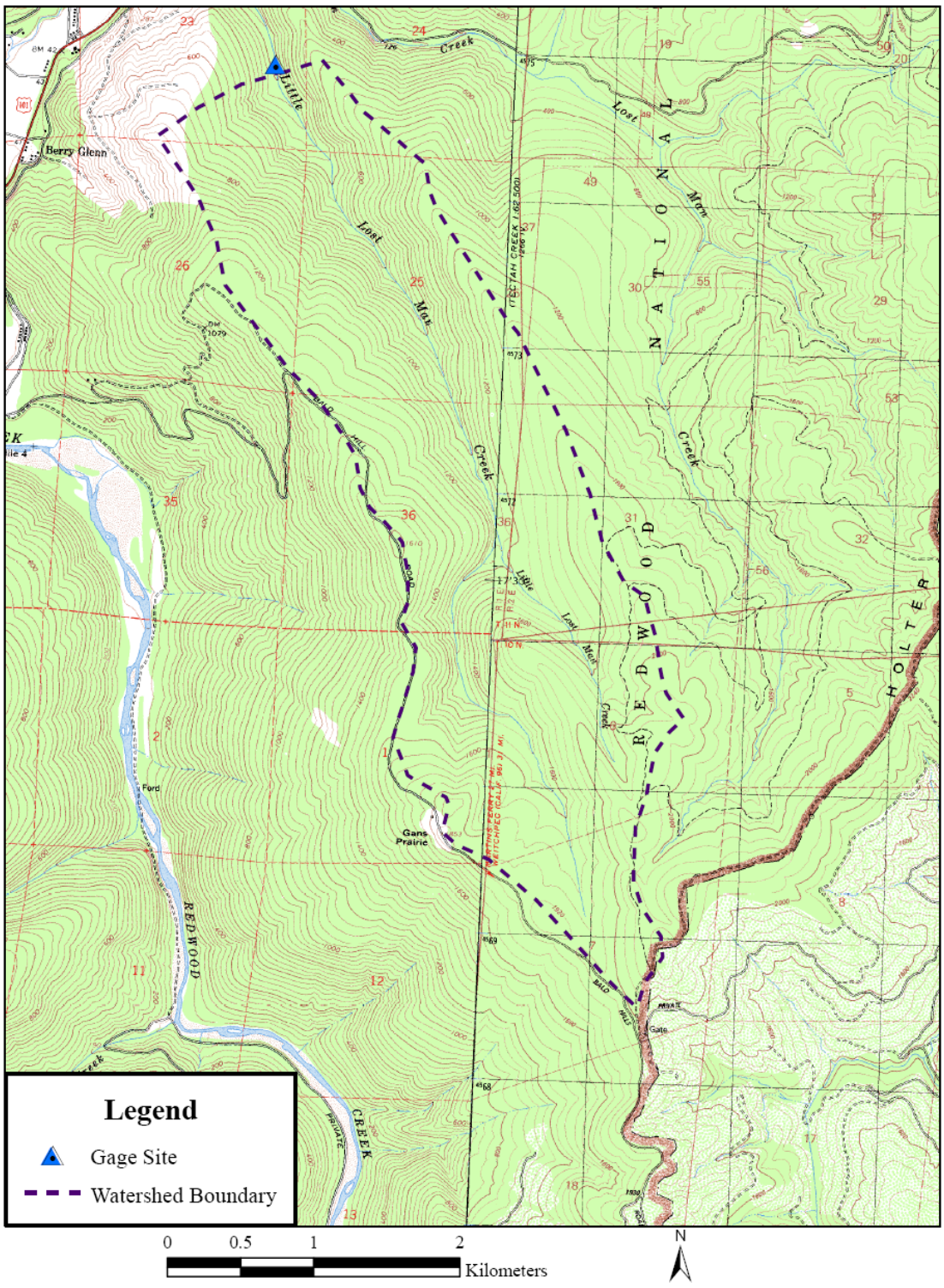
Q_1 = Peakflow with a return period of one-year equal to 2.0 L s^{-1} ha^{-1}

Q_o = Observed peakflow, L s^{-1} ha^{-1}

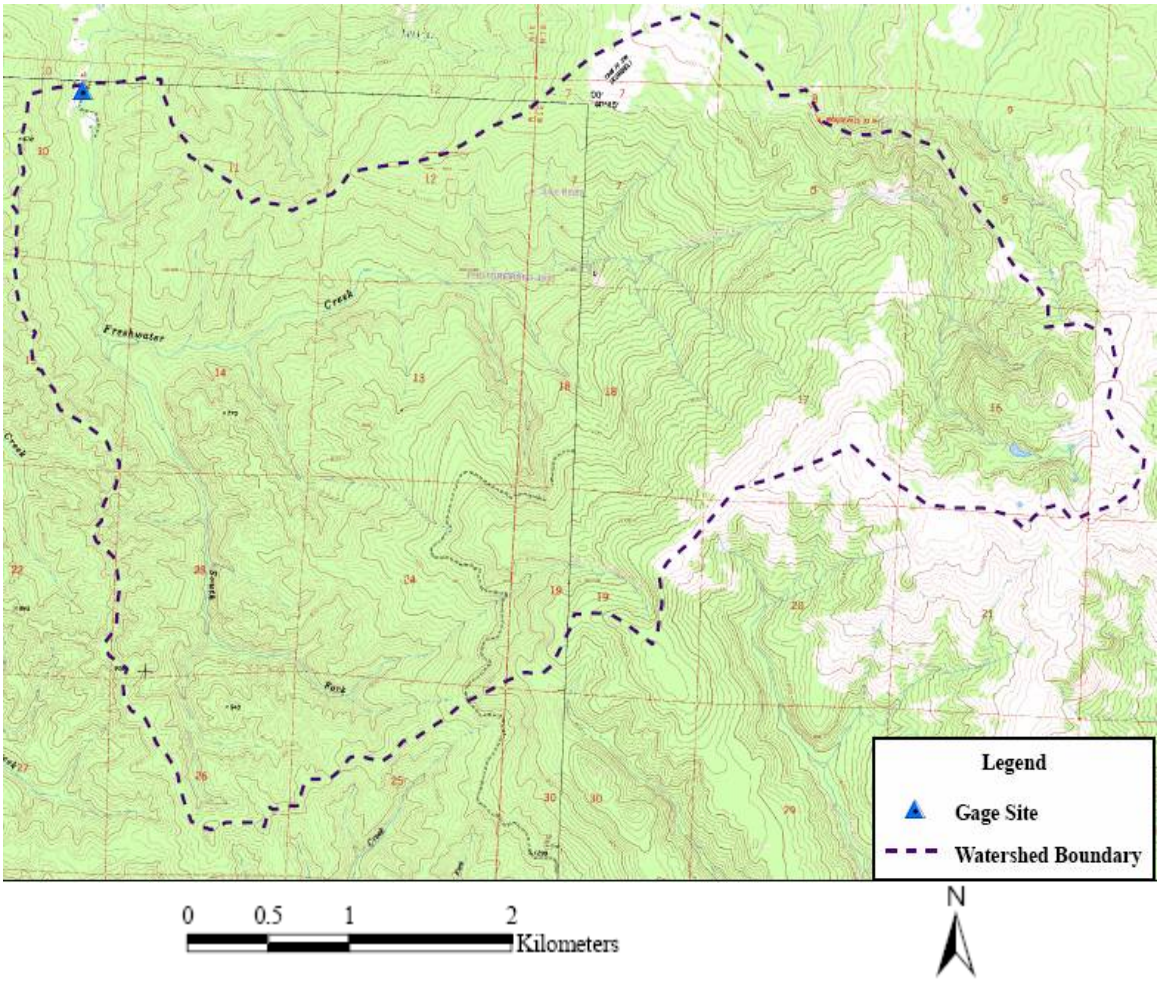
Q_p = Predicted peakflow, L s^{-1} ha^{-1}



Appendix A. South Fork of Caspar Creek, North Fork of Caspar Creek, and Hennington watershed boundaries with gage sites and rain gage locations



Appendix B. Little Lost Man Creek watershed boundary and gage site location.

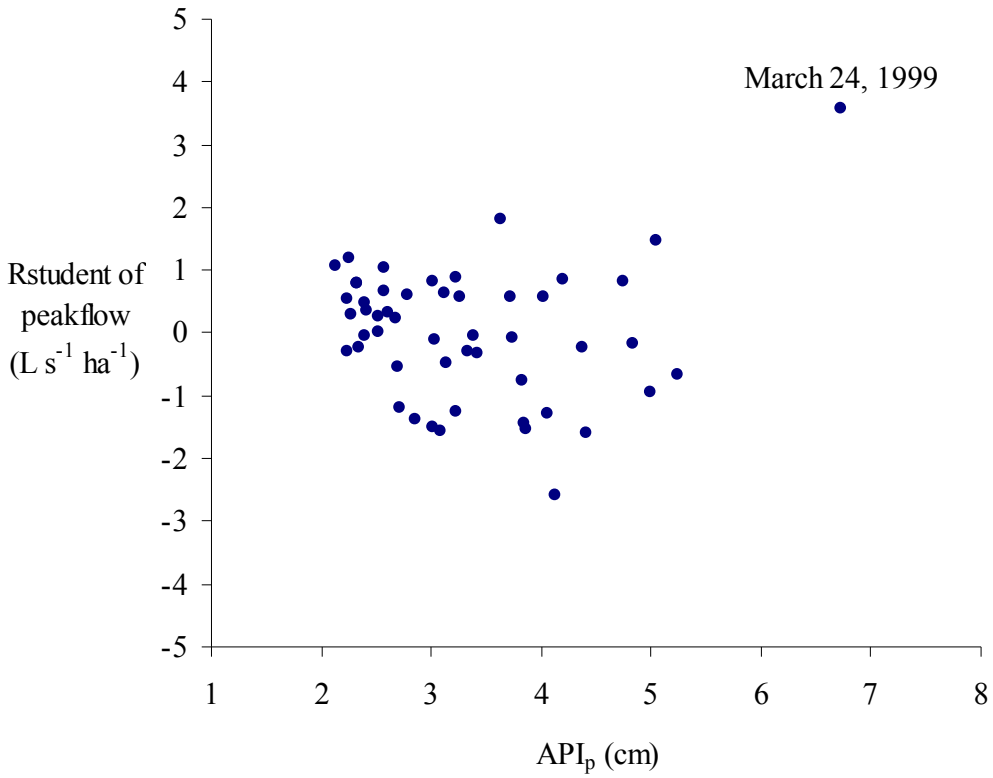


Appendix C. Freshwater Creek watershed boundary and gage site location.

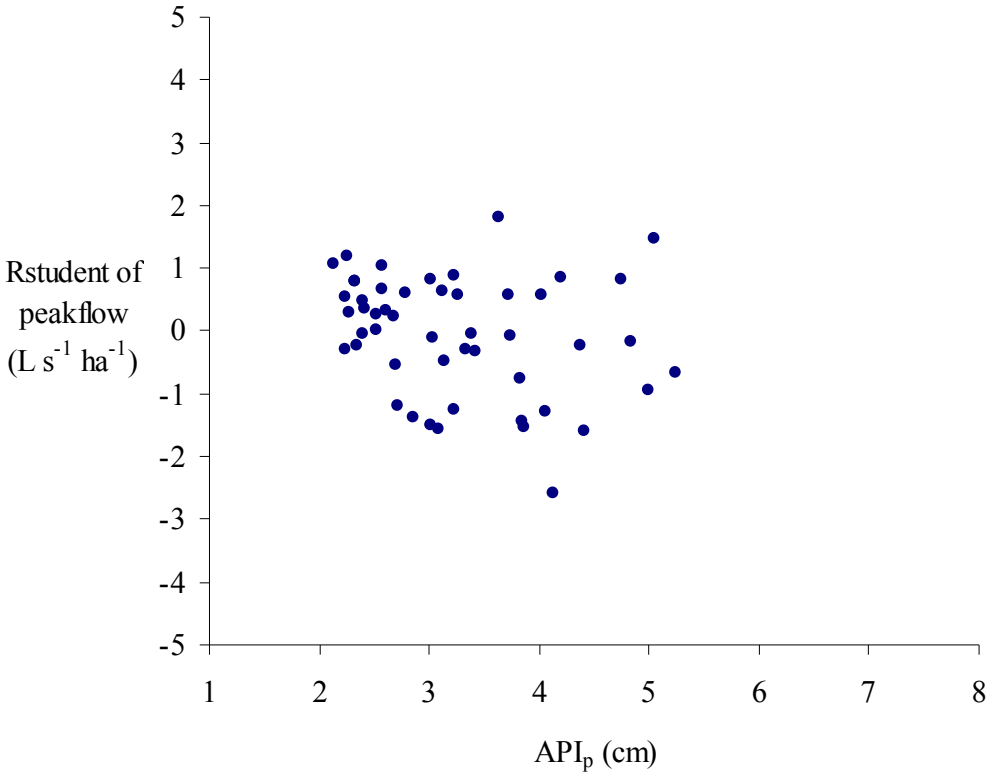
Appendix D. Outlier detection statistics from residual diagnostics before and after the March 24, 1999 event was removed. The March 24, 1999 event failed all four tests. The three remaining observations passed the DFFITS and Cook's D tests. Two out of the three were considered high leverage outliers based on Rstudent and Hat Diagonal, the other failed the Rstudent test. All statistics were calculated using NCSS (Hintze 2004).

	Initial Outlier Detection Statistics				
API_p	Residual	Rstudent*	DFFITS**	Cook's D***	Hat Diagonal****
4.13	-3.3886	-2.5746	-0.476	0.1016	0.0331
6.72	3.8943	3.5637	2.1085	1.7944	0.2593
	Outlier Detection Statistics after March 24, 1999 event removal				
API_p	Residual	Rstudent	DFFITS	Cook's D	Hat Diagonal
5.05	2.7111	2.3672	0.8255	0.3109	0.1084
3.63	2.684	2.2259	0.3476	0.0558	0.0238
4.13	-2.987	-2.5342	-0.5235	0.1231	0.0409

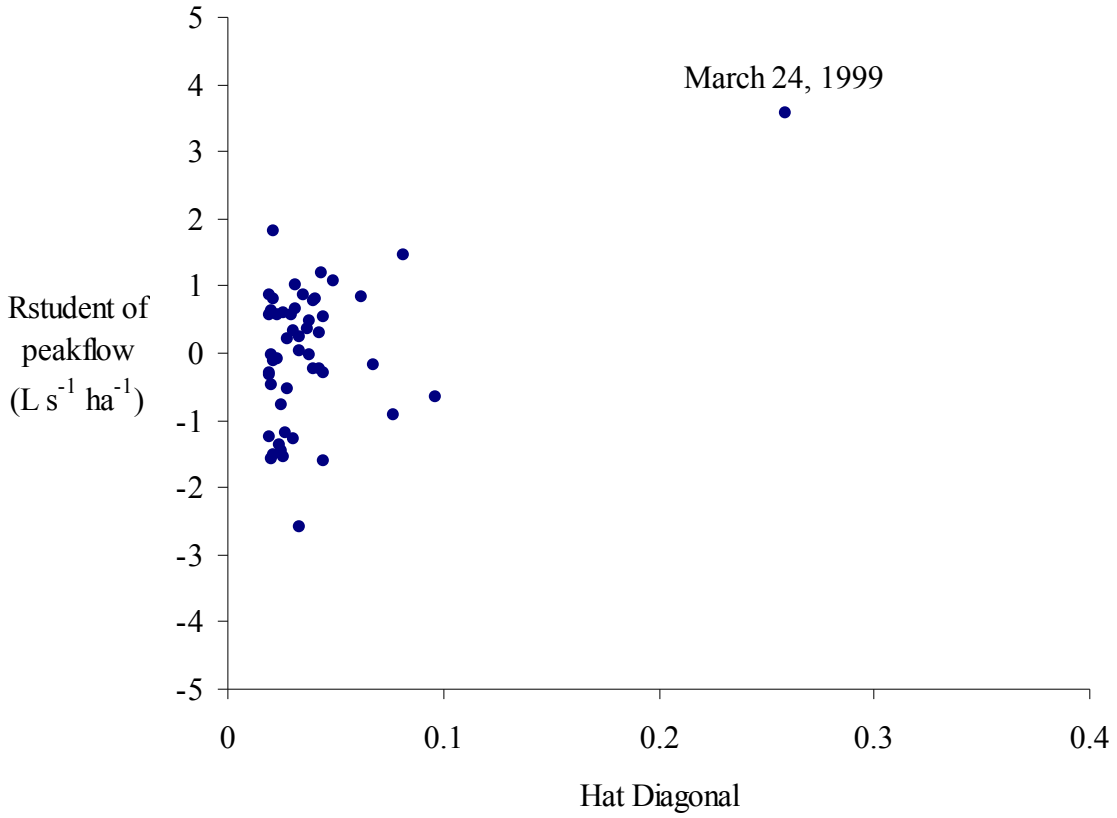
* An observation is considered an outlier if the absolute value of Rstudent (also known as the studentized deleted residuals) is greater than two (Hintze 2004).
** An observation is considered influential concerning prediction if the absolute value of DFFITS is greater than one. DFFITS measures the influence of a single observation on its fitted value (Velleman and Welsch 1981).
*** Cook's D values greater than one indicate that the observations have a large influence. It measures the influence of each observation on all fitted values (Velleman and Welsch 1981).
**** Hat Diagonal measures the remoteness of the observations in the X-space. Hat Diagonals greater than 2*degrees of freedom / n (2*2/59 = 0.068) are considered high-leverage observations. Leverage refers to the amount of influence a given observation has on the trend of the least squares regression estimate (Velleman and Welsch 1981).



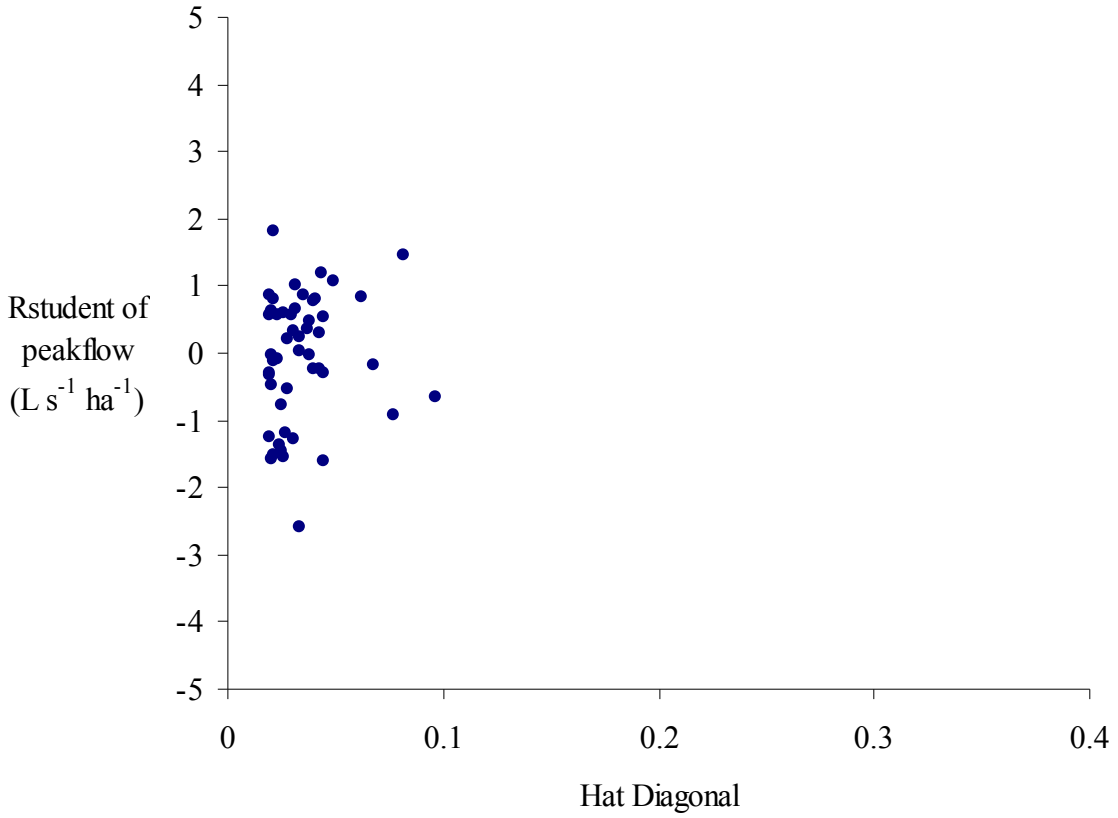
Appendix E. Rstudent as a function of peak API shows the March 24, 1999 event as an outlier.



Appendix F. Rstudent as a function of peak API with the March 24, 1999 event removed. Two observations with an absolute value of Rstudent greater than two remain. These observations were retained because they did not deviate significantly from the cloud.



Appendix G. Rstudent as a function of Hat diagonal indicates that the March 24, 1999 event was a high leverage observation.



Appendix H. Rstudent as a function of Hat diagonal with the March 24, 1999 event removed. One of the observations remaining was considered an outlier, four were considered high leverage, and one was considered a high leverage outlier. However, they all passed the DFFITS and Cook's D test unlike the March 24, 1999 event.

Appendix I. Tests of regression assumptions after the March 24, 1999 outlier was removed. The Modified Levene Test indicates a lack of constant residual variance. The other null hypotheses were not rejected at the 0.05 alpha level. The Durbin-Watson test indicated a lack of positive and negative autocorrelation (alpha = 0.05). All statistics were calculated using NCSS (Hintze 2004).

Do the residuals follow a normal distribution?	Test Value	Probability Level	Assumption Reasonable ($\alpha = 0.05$)
Shapiro Wilk	0.9736	0.235103	Yes
Anderson Darling	0.6689	0.080721	Yes
D'agnostino Skewness	-0.8441	0.398635	Yes
D'agnostino Kurtosis	0.572	0.5673	Yes
D'agnostino Omnibus	1.0397	0.594621	Yes
Constant residual variance?			
Modified Levene	10.7631	0.001785	No
Durbin-Watson test for lack of autocorrelation			
Positive	1.60	0.0626	Yes
Negative	1.60	0.9378	Yes

Appendix J. API model coefficients and related statistics.

Parameter	Intercept B(0)	Slope B(1)
Coefficients	-3.5222	2.8963
Lower 95% Confidence Limit	-4.6398	2.5482
Upper 95% Confidence Limit	-2.4046	3.2444
Standard Error	0.5579	0.1738
Standardized Coefficient	0.0000	0.9123
T statistic	-6.3134*	16.6668*

* Significant at 0.05 alpha ($p < 0.0001$)