

Article

# A Quantitative Approach to Evaluate Changes in Hydrologic Conditions of Headwater Streams: A Case Study of Restoration and Recovery Following Longwall Mine Subsidence

Joshua M. Silvis <sup>1,\*</sup> , Brian C. Benson <sup>1</sup>, Michael L. Shema <sup>2</sup> and Mark R. Haibach <sup>2</sup>

<sup>1</sup> CONSOL Energy, Inc., Canonsburg, PA 15317, USA

<sup>2</sup> Civil & Environmental Consultants, Inc., Pittsburgh, PA 15205, USA

\* Correspondence: JoshuaSilvis@consolenergy.com; Tel.: +1-724-416-8311

Received: 13 May 2019; Accepted: 5 August 2019; Published: 8 August 2019



**Abstract:** Mine subsidence can induce streambed ruptures that pirate surface water from a stream. Current understanding of the effects of longwall mining on streams lacks rigorous analytical approaches to detect hydrologic effects and does not consider the efficacy of streambed restoration techniques to address flow disruption. CONSOL Pennsylvania Coal Company, LLC (CPCC) collected and analyzed pre-mining, post-mining, and post-restoration stream discharge and flow duration data from 51.9 km of streams overlying its Bailey Mine to define pre-mining flow variability, detect post-mining changes, and evaluate post-intervention flow recovery. The primary intervention method for restoring stream flow was bedrock permeation grouting. Pre-mine and post-mine baseflow data were compared using both parametric and non-parametric hypothesis testing, which yielded similar results. An environmental flow assessment method for headwater streams using probabilistic risk assessment and correlation analysis of change threshold criteria was developed to differentiate hydrologic change as induced by mine subsidence or explained by natural variability. The method is objective, pragmatic, and statistically delimited.

**Keywords:** mine subsidence; environmental flow; headwater stream; flow restoration; fracture grouting

## 1. Introduction

Scientific knowledge and theory pertaining to environmental flows advanced noticeably in the past few decades. This resulted in the conceptual development of various environmental flow assessment methodologies of river flows [1,2]. In the 1970s and 1980s, the water surface profile method [3,4], based on channel characteristics, and the Montana method [5], based on streamflow characteristics, were developed. The range in variability approach was developed in the 1990s to evaluate statistical variables of streamflow regarded as important indicators of hydrologic alteration [6]. More recently, the ecological limits of hydrologic alteration were developed to assess the flow alteration and ecological response between anthropogenic alterations and social, economic, and ecological preferences [1,7].

These existing environmental flow assessment methods were developed at the regional scale to balance ecosystem management with societal objectives for river management, which have the potential to negatively affect a river ecosystem by influencing the natural variations in flow [1,7,8]. These methods demonstrate that the degree of hydrologic change affects existing hydro-ecological relationships with the greatest effects on river ecosystems being associated with severe hydrologic alteration [6].

Theoretical and conceptual knowledge related to the impact of longwall mine subsidence of streams [9–11] and groundwater [12–14] existed since the early 1980s. Hydrogeologic studies were completed concomitantly to engineering studies examining mine subsidence theory of overburden movement and deformation [15–17] as it relates to local stress-relief fracture systems [18–20]. Existing environmental flow assessment methods do not, however, provide a proper tool to assess natural flows in headwater streams or evaluate flow in streams which underwent restoration to correct post-mine streamflow disruption.

CONSOL Pennsylvania Coal Company, LLC (CPCC) used intervention measures, such as streamflow augmentation and streambed grouting, to restore post-mine streamflow disruptions since 2006. The effects of longwall mine subsidence on streams in Pennsylvania are regulated under the Pennsylvania Clean Streams Law, the Mine Subsidence Act, and environmental regulations adopted pursuant to these and other laws, including Chapters 86, 87, 89, 93, and 105 of Title 25 of the Pennsylvania Code. The Pennsylvania Department of Environmental Protection (PADEP) also uses its October 2005 Technical Guidance Document No. 563-2000-655, entitled “Surface Water Protection Underground Bituminous Coal Mining Operations” (referred to herein as the Streams TGD) to interpret the various regulatory requirements applicable to longwall mining [21].

The Streams TGD provides guidance on subsidence-induced flow loss that may occur as the result of physical changes to the streambed, such as bedrock heaves and fractures, and subsidence-induced inverts that may create pooling in low-gradient streams and associated floodplain wetlands. The Streams TGD specifies hydrologic and biological monitoring frequency and reporting requirements and methods to compare pre-mine and post-mine biological indices to determine if restoration achieved established biological recovery end points. However, the Streams TGD does not prescribe methods to compare pre-mine and post-mine hydrologic conditions or contain specific hydrologic recovery end points. Throughout this paper, the term post-mine is used to reference both post-mine and post-restoration conditions.

Empirical data specific to the Streams TGD substantially improved CPCC’s understanding of the implications of longwall mine subsidence on local surface water flow regimes. In compliance with the Streams TGD, CPCC collected and analyzed hydrologic monitoring data to develop a hydrologic environmental flow method associated with streams undermined by longwall mining at its Bailey Mine located in Greene County, Pennsylvania, United States of America (USA).

## 2. Materials and Methods

### 2.1. Site Description and Hydrogeologic Setting

The Bailey Mine is a longwall mine operation producing thermal and high-vol metallurgical coal from the Pittsburgh coal seam (Figure 1). Annual production at the Bailey Mine averaged 11.16 million metric tons per year during the past three years (2016 to 2018). Extraction of coal beneath streams within the study area began fourth quarter, 2004 and ended first quarter, 2012.

A longwall mine operation is a highly productive and mechanized underground full extraction technique designed to induce controlled subsidence. The operation consists of a shearer with two rotating cutting drums and a series of self-advancing hydraulic roof supports that protect the working face. As a longwall panel is retreated, the overburden above the extracted coal seam progressively subsides, resulting in dynamic movement of the overburden and ground surface. The dimensions of the longwall panels within the study area are 2 m high (approximate extraction height), 4570 m long and 335 to 457 m wide, and are oriented south (S) 67° east (E).

The Bailey Mine is located in the Washington Hills section of the Appalachian Plateau physiographic province. The terrain is characterized by broad ridges, deeply incised headwater streams, and flat open valley bottom areas adjacent to low-gradient (<2%) streams. Land use is predominantly forest. Surface elevations of streams in the study area range from 286.5 m to 402.3 m above mean sea level.

Near-surface geology within the study area consists primarily of cyclothemic deposits of sandstones, siltstones, shales, limestones, clays, and coal beds of the Dunkard Group of late Pennsylvanian to early Permian Age. The exposed strata traverse the Waynesburg Sandstone unit of the Lower Waynesburg Formation to the Upper Rush Run Sandstone unit of the Tenmile Member, Greene Formation. Overburden thickness between the Pittsburgh seam and the streambeds within the study area ranges from 94.5 m to 256.0 m.

Annual rainfall recorded by the National Weather Service in Waynesburg, PA (WYNP1/NWS; mean ( $M$ ) = 110.9 cm,  $SD$  = 19.8 cm) varied across the study period (2004–2017) from a low of 88.0 cm (2014) to a high of 158.1 cm (2011). The 30-year (1986–2015) average annual rainfall was 103.8 cm.

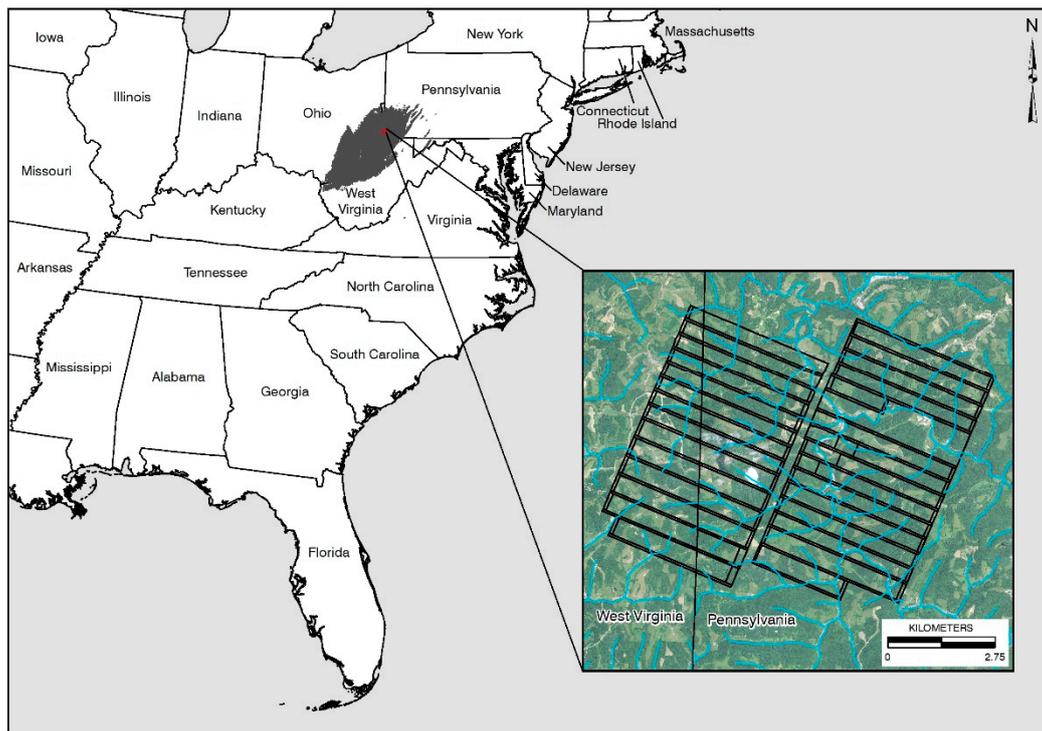


Figure 1. Map of the study area.

## 2.2. Streambed Restoration Techniques

The fundamental goal of stream restoration in the Commonwealth of Pennsylvania is to minimize any adverse effects that may be induced by mine subsidence and to ensure that coal mine activity permits are designed to protect and maintain the existing and designated uses of perennial and intermittent streams.

Mine subsidence can induce streambed ruptures, such as bedrock heaves and fractures, which pirate surface water from the streambed. Streamflow augmentation is used as a proactive flow management technique to bridge the time between the first observed incident of mine subsidence-induced flow loss and streamflow loss restoration. Streamflow augmentation involves the delivery and discharge of fresh water to the stream to restore and maintain streamflow, thereby supporting the existing biological communities. During the streamflow augmentation period, CPCC performs a pre-mitigation survey to identify the location(s) and extent of flow loss restoration that is necessary.

Flow loss restoration seeks to seal surface and/or near-surface bedrock fractures that may be the cause of flow loss. Bedrock permeation grouting is a streambed restoration procedure that injects a grout into the rock to reduce or eliminate the loss of surface water. CPCC initially used a blanket grouting procedure where boreholes were drilled in a single pass along the entire length of stream. A stable, cement-based grout was injected into the boreholes to fill fractures or voids. Communication of

grout to the streambed surface through adjacent injection boreholes was commonly observed, and full closure of fractures could not be confirmed.

CPCC revised its original technique in 2010 and since then used a performance-based closure-style grouting procedure. This technique involves drilling a primary set of boreholes, at approximately 3-m centers, to a depth of 1.8 m, which is the practical limit of manually operated equipment. A cement-, bentonite-, and water-based grout is injected into each borehole using a high-shear, high-speed colloidal mixer until the desired pressure is achieved. An inflatable packer, charged with compressed air, promotes confinement and prevents the return of grout to the surface. A custom manifold directs grout to a single or suite of injection boreholes for simultaneous injection, and recirculates unused grout back to the mixer to reduce waste. Grout density, viscosity, and compression strength are evaluated using standard testing methods, and field jar samples are collected and tested to ensure <2% shrinkage after curing. Drilling and injection of additional sets of boreholes at incrementally closer spacing is completed as necessary to achieve full closure.

This revised technique confirms closure (or determines the capacity of the streambed to convey water) by monitoring injection pressures and grout volumes and collecting streamflow measurements following each series (primary, secondary, etc.) of grouting. The current approach also minimizes disturbance of the streambed by targeting only reaches exhibiting ground deformation or streamflow reduction identified during the pre-mitigation stream survey. To date, CPCC grouted approximately 14.97 km of streambed in the study area.

### 2.3. Hydrologic Monitoring

CPCC analyzed two stream characteristics for this study: stream discharge, measured at fixed surface water monitoring locations (sample sites), and stream flow duration or intermittency, determined by observations and documentation of the longitudinal extent of flow along the stream channel.

This study involved collecting and analyzing 12,356 individual stream discharge measurements at 88 sample sites and recording flow observations along 29 streams (51.9 km) classified as warm water and trout-stocked fisheries. Stream discharge was measured using the United States Geological Survey (USGS) midsection current meter method [22]. Timed volumetric measurement was also used if water depths were too shallow to obtain accurate readings with a velocity meter. Stream flow observations included documenting the extent of flowing and non-flowing reaches (flow loss), flow origination points, ruptures and fractures of bedrock or bedload, pooling, discoloration caused by mineral leaching, and other observable conditions.

Monthly data and streambed observations, as recommended in the Streams TGD, were collected using TerraSync software (Version 5.86, Trimble Inc., Sunnyvale, CA, USA) installed on handheld Trimble global positioning system (GPS) units. These field data were post-processed using Pathfinder Office. Custom Microsoft Excel scripts expedited quality control prior to the data being uploaded to a propriety Microsoft Access database. Streambed observations, such as flow loss, were downloaded into ArcGIS to be measured and evaluated for trends.

Monthly sampling frequency, which is specified by the Streams TGD, does not enable the use of analysis techniques that require continuously monitored data, such as predicting low flow duration or peak flow magnitude. These approaches are valid for extreme event analysis and rooted in applications associated with drought and flood risk. For an analysis of change in hydrologic response, the use of base flow data provides a more balanced and representative distribution, which is more suitable to the objectives of this study.

The sensitivity of detecting potential mine-subsidence effects is improved by removing the confounding effect of extreme weather events. To identify and remove the effect of these events, a Microsoft Excel script identified data collected within 48 h of an event that recorded a rainfall of 1.1 cm (daily mean storm event across the 30-year precipitation record) or greater. The parameters used to identify data influenced by extreme weather event flows were selected a priori to strike a balance between excluding event-generated variance and preserving sufficient data to ensure an

adequate sample size. The script also identified and removed data collected during periods of USGS-declared droughts.

The duration of pre-mine monitoring ( $M = 3.2$  years,  $SD = 1.1$  years) was nearly equivalent to the duration of post-mine monitoring ( $M = 3.4$  years,  $SD = 1.2$  years). The comparison of pre-mine to post-mine data over several years reduced the influence of annual precipitation differences. The rolling three-year average precipitation ( $M = 111.3$  cm,  $SD = 10.7$  cm) was less variable than the annual average precipitation ( $M = 110.9$  cm,  $SD = 19.8$  cm), which supported a more consistent comparison by controlling the influence of precipitation differences between the two periods.

#### 2.4. Statistical Analysis

The distribution of the pre-mine and post-mine stream discharge data for each sample site was positively skewed and heavily tailed. A Shapiro–Wilk goodness-of-fit test using ProUCL 5.1 concluded that data for a majority of the sample sites (86 of 88) were non-normally distributed. Non-normal distributions can affect sample statistics, which are calculated to infer characteristics associated with measures of central tendency, spread, and symmetry [23].

Nonparametric statistics or the use of data transformation and parametric analysis of the transform of the data are typically used to analyze non-normally distributed data. The principal advantage of nonparametric statistics is that they are invariant to measurement scale and can be used to test population parameters when the assumption of normality cannot be met [23]. Because nonparametric tests can be less sensitive, a larger difference is needed before the null hypothesis can be rejected. When the assumptions of parametric methods are met, the nonparametric counterpart needs a larger sample size to overcome the loss of information to obtain the same result; therefore, the nonparametric test can also be less efficient [24].

Parametric tests require symmetry, linearity, and homoscedasticity to interpret data effectively. The log transformation is a traditional approach to handle positively skewed data and can produce these characteristics [23]. When data transformation successfully produces a normal or nearly normal distribution, the parametric test is a better estimator of the difference between the two groups of data. Parametric procedures can also enable regression analysis, which is nearly impossible to complete using nonparametric methods [23].

To determine if any difference existed between the distribution of pre-mine and post-mine stream discharge data, parametric and nonparametric statistical hypothesis testing were performed. Statistical significance was determined by comparing the attained  $p$ -value to an *a priori* criterion of 0.05 and fit to a two-tailed test. To determine if long-term systematic trends existed, analysis was performed on a comparison of the entire pre-mine and post-mine periods; therefore, a seasonality analysis was not performed.

The parametric two-sample  $t$ -test and  $z$ -test were used to examine differences of base-10 log transformation of the discharge data (log base flow) between the 88 pre-mine and post-mine discharge sample sites. Because the log of zero is undefined, pre-mine and post-mine zero streamflow events were assigned a value of 0.38 L/min (0.1 gallon/min). The  $z$ -test was used on the log base flow data collected at 80 of the 88 sample sites, as these datasets met the sample size assumption. The two-sample  $t$ -test was used for the remaining eight sample sites. The appropriate  $t$ -test formula was determined by the results of the  $F$ -test, as there are two options for the  $t$ -test discriminated by equal or unequal variances.

An equivalent nonparametric test for independent samples was used to determine any similar differences. The Wilcoxon rank sum test is a robust technique that is used to determine statistical significance of the difference in magnitude between the ranks of the pre-mine and post-mine stream discharge data.

While statistical significance relates to whether an effect exists, practical significance refers to the magnitude of the effect and its potential for field-based application. To determine the practical significance of identified relationships, three methods were used to test the strength of these relationships: linear regression, probabilistic risk assessment, and correlation analysis of change.

#### 2.4.1. Linear Regression

To compare the cumulative pre-mine and post-mine stream discharge datasets, a simple linear regression of stream discharge and drainage area was completed for pre-mine and post-mine discharge data. The empirical relationship between stream discharge and drainage area was expressed as a power function in the form  $Q = bA^c$ , where  $b$  is a proportionality constant and  $c$  is an exponent that scales discharge to drainage area [25]. A Pearson's  $r$  correlation using the Fisher  $z$  transformation solution was used to determine the degree of association between average stream discharge and drainage area. These two linear regression lines were compared to determine mining impacts.

#### 2.4.2. Probabilistic Risk Assessment

To predict the likelihood that changes in stream discharge are likely or unlikely to occur, probabilistic risk assessment (PRA), involving the use of likelihood and a measure of the magnitude of potential consequences should changes occur, was completed. The likelihood metric (level of probability) was derived from the nonparametric hypothesis test results, specifically differentiating the results into statistically significant or non-significant events. The magnitude of potential consequence (MPC) was calculated as  $((\text{post-mine log average } Q - \text{pre-mine log average } Q) / \text{pre-mine log average } Q) \times 100$ .

#### 2.4.3. Correlation Analysis of Change

To estimate the threshold criteria of stream flow extent and duration associated with likely hydrologic alteration, a correlation analysis of stream discharge and stream flow observation data was completed. Hydrologic data associated with intermittent flow conditions were used to investigate changes to the extent (spatial variation) and duration (temporal variation) of flow to determine the direction and magnitude of change between the pre-mine and post-mine periods.

Change in intermittent flow condition is directional; a regressive change indicates an increase to the extent and duration of a stream's intermittency that is trending toward a drier streambed condition, whereas a progressive change indicates a decrease to the extent and duration of a stream's intermittency trending toward a continuous flowing condition. It is reasonable to expect a regressive change in intermittent flow condition where stream discharge decreases and a progressive change where stream discharge increases. It is also logical to state this relationship in terms of magnitude, where the degree of difference in intermittent flow condition is dependent on changes in stream discharge.

#### 2.4.4. Environmental Flow Assessment

Hydrologic-based environmental flow assessment methodologies are the most frequently used methodologies used to make environmental flow recommendations [8]. The Montana method, or a derivation of the original, is routinely used in 16 US states or Canadian provinces in North America and is now the most commonly applied hydrologic environmental flow assessment methodology worldwide [8].

Mimicking the design of the Montana method, the degree of hydrologic alteration measured at sample sites located at the mouth of the 29 streams within the study area was examined to determine whether downstream use impairment was likely. This approach aggregates all drainage features within a given watershed into a single representative sample, allowing multivariate analysis of the changes in the discharge measured at the mouth and changes in intermittency across the entire monitored length of the streams.

### 3. Results

#### 3.1. Hypothesis Testing

Parametric two-sample  $t$ -tests and  $z$ -tests collectively showed that 47 tests accepted the  $t$ -test and  $z$ -test hypothesis of no difference, while 41 tests rejected the hypothesis of no difference (Supplementary

Materials Tables S1–S3). Of the 41 tests that indicated a statistically significant difference between the pre-mine and post-mine data, the post-mine discharge was greater than the pre-mine discharge at 21 monitoring locations and less than the pre-mine discharge at 20 sample sites.

The nonparametric Wilcoxon rank sum test collectively revealed that, of the 88 sample sites, 49 tests accepted the hypothesis of no difference, while 39 tests rejected the hypothesis of no difference (Supplementary Materials Table S4). Of the 39 tests that indicated a statistically significant difference between the pre-mine and post-mine data, 21 tests concluded that the post-mine discharge increased, and 18 tests concluded that the post-mine discharge decreased.

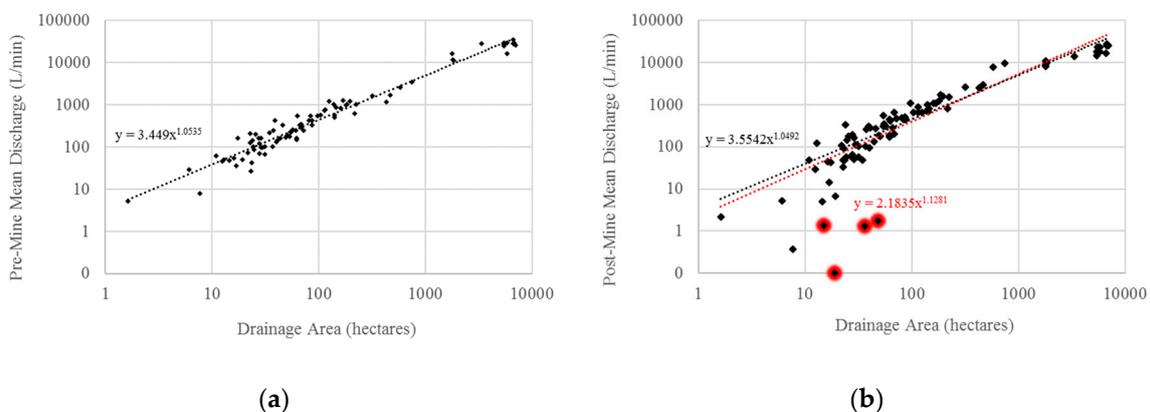
For nearly normal data produced by log transformation, the nonparametric and parametric tests often give the same result [23]. The two test procedures produced the same hypothesis decisions for 78 of 88 (89%) of the sample sites. These similar results provided a high degree of confidence that conclusions based on changes to the distribution of discharge, specifically differences in the log base flow data, were accurate and could be used for correlation and regression analysis of trends.

### 3.2. Trend Analysis

The above statistical analysis suggested several data trends, which required further analysis to determine their applicability for developing an environmental flow assessment methodology.

#### 3.2.1. Linear Regression

Linear regression analysis indicated a very strong positive relationship existed between the dependence of the mean discharge on the drainage area for both pre-mine ( $r = 0.978$ ,  $n = 88$ ,  $p = 0.013$ ) and post-mine ( $r = 0.936$ ,  $n = 88$ ,  $p = 0.041$ ) datasets, as shown in Figure 2.



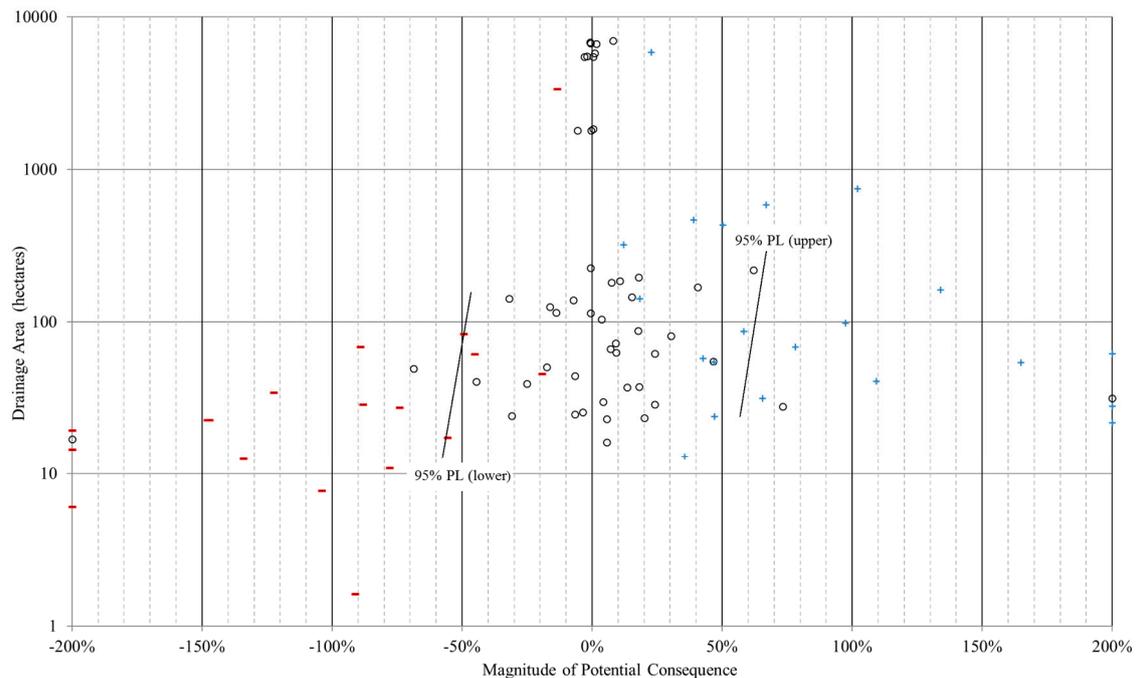
**Figure 2.** Regression trendlines showing drainage area versus pre-mine (a) and post-mine (b) mean discharge for sample sites in the study area. Pre-mine regression line is a best-fit power function:  $Q = 3.449A^{1.054}$ ,  $r^2 = 0.957$ . Post-mine regression (red line) is a best-fit power function:  $Q = 2.184A^{1.128}$ ,  $r^2 = 0.876$ . Optimized post-mine regression (black line) is a best-fit power function:  $Q = 3.554A^{1.049}$ ,  $r^2 = 0.926$ .

The post-mine relationship differs from the pre-mine discharge with respect to the proportionality constant ( $b = 2.184$ ) and the scaling exponent ( $c = 1.128$ ), which suggests the composite post-mine discharge-to-drainage area ratio decreased. To account for this, a univariate search method was used to exclude sample sites with the greatest measured reduction in mean discharge until the proportionality constant and scaling exponent were similar to those determined by regression analysis of the pre-mine data. This approach may be somewhat unorthodox; however, the results ( $r = 0.962$ ,  $n = 84$ ,  $p = 0.025$ ), proportionality constant ( $b = 3.554$ ), and scaling exponent ( $c = 1.049$ ), as shown in Figure 2b, suggested the reduced, post-mine discharges measured at four sample sites (red highlighted points) accounted for the net difference between pre-mine and post-mine discharge-to-drainage area relationships.

This holistic approach is more representative of a cumulative effect and less descriptive of individual, stream-specific hydrologic change.

### 3.2.2. Probabilistic Risk Assessment

Analyzing the magnitude of hydrologic consequence measured at each sample site as it related to drainage area identified three statistically distinct groupings, as shown in Figure 3: no difference in discharge (o), an increase in post-mine discharge (+), and a decrease in post-mine discharge (-). The transitions between these groups are conspicuous and were statistically derived as prediction limits using the ordinary least squares (OLS) regression model.



**Figure 3.** Magnitude of potential consequence versus drainage area for sample sites in the study area. The lower 95% prediction limit is shown as an exponential trendline:  $y = 6.37 \times 10^6 e^{22.80x}$ . The upper 95% prediction limit is shown as an exponential trendline:  $y = 1.95 \times 10^{-5} e^{24.65x}$ .

The 95% prediction limit is a benchmark of risk that distinguishes between possible and likely hydrologic alteration based upon the statistical significance of the difference in the means of the log base flow data. The lower and upper 95% prediction limits were obtained from the sample sites ( $n = 36$ ) that accepted the null hypothesis of no difference using the nonparametric Wilcoxon rank sum test. Thirteen sample sites with drainage areas greater than 400 ha ( $n = 11$ ) or MPC greater than 200% ( $n = 2$ ) were excluded from the OLS regression analysis. An MPC that is less than the lower 95% prediction limit is likely to show a statistically significant decrease in the post-mine mean stream discharge. An MPC that is greater than the upper 95% prediction limit is likely to show a statistically significant increase in the post-mine mean stream discharge. An MPC within the 95% prediction interval is likely to show a statistically insignificant difference between the pre-mine and post-mine means of the log stream discharge data. The threshold criteria representing the lower and upper 95% prediction limits are listed in Table 1 and are applicable to streams with drainage areas between 4 and 400 ha.

Probabilistic risk assessment is a reliable (as indicated by the results of the comparison of two test procedures), accessible (requires common data), and simple (comparison of two values) method that can be used to determine the degree of change relative to a calculated gradient of hydrologic alteration.

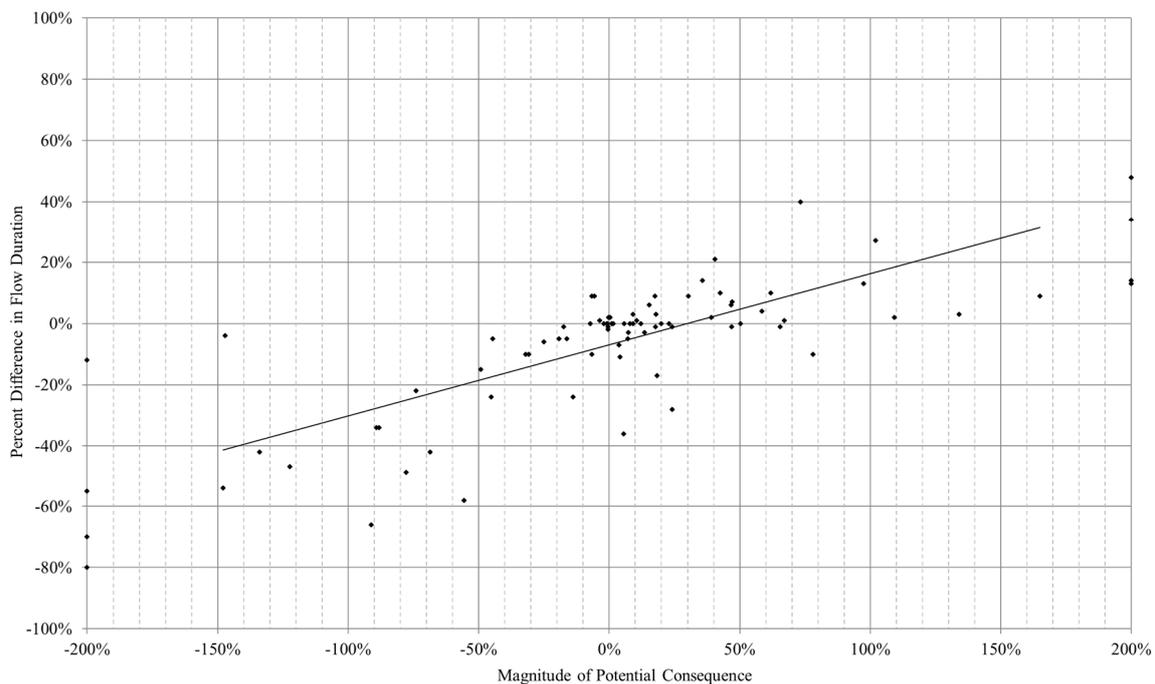
**Table 1.** Threshold criteria (magnitude of potential consequence, MPC) for likely mining effects on stream discharge as a function of drainage area (DA).

Drainage Area	Lower 95% Prediction Limit MPC = $\ln\left(\frac{DA}{6.37 \times 10^6}\right) / 22.80$	Upper 95% Prediction Limit MPC = $\ln\left(\frac{DA}{1.95 \times 10^{-5}}\right) / 24.65$
4 ha <sup>a</sup>	−63%	50%
20 ha	−56%	56%
40 ha	−52%	59%
200 ha	−45%	66%
400 ha <sup>a</sup>	−42%	68%

<sup>a</sup> The threshold criteria for 4 and 400 ha are extrapolated values of the 95% prediction limit trendlines.

### 3.2.3. Correlation Analysis of Change

The extent and duration of a stream's intermittency is commonly recognized as an important indicator of biological potential [6]. The relationship between the MPC and the difference in intermittent flow condition was explored through correlation analysis (Supplementary Materials Table S5). Twelve sample sites where the MPC was greater than 200% ( $n = 8$ ) and post-mine flow condition data were unavailable ( $n = 4$ ) were excluded from this analysis. For the remaining 76 sample sites, Figure 4 shows that changes observed in intermittent flow condition are dependent upon the MPC, with a strong, positive correlation existing between these paired data ( $r = 0.712$ ,  $n = 76$ ). The equation of the trendline is expressed as a linear function,  $y = 0.232x - 6.94$ , where  $y$ , the difference in flow duration, is dependent on  $x$ , the MPC.

**Figure 4.** Correlation between the magnitude of potential consequence and difference in flow duration for sample sites in the study area.

The threshold criteria for likely mining effects on the duration of a stream's intermittent flow condition are based on the correlation between the difference in pre-mine and post-mine log mean discharges and differences in flow duration. These threshold criteria are estimated using the trendline equation in Figure 4 and the lower and upper 95% prediction limits of the MPC (Table 1). These threshold criteria represent the maximum difference in a stream's intermittent flow condition that could potentially be explained as natural variation. A difference in flow duration that is less than

the lower 95% prediction limit is likely to show a regressive change in intermittent flow condition. A difference in flow duration that is greater than the upper 95% prediction limit is likely to show a progressive change in intermittent flow condition. The threshold criteria representing the lower and upper 95% prediction limits are listed in Table 2 and are applicable to streams with drainage areas between 4 and 400 ha.

**Table 2.** Threshold criteria for likely mining effects on the extent and duration of a stream's intermittent flow condition as a function of drainage area.

Drainage Area	Lower 95% Prediction Limit	Upper 95% Prediction Limit
4 ha <sup>a</sup>	−21%	5%
20 ha	−20%	6%
40 ha	−19%	7%
200 ha	−17%	8%
400 ha <sup>a</sup>	−17%	9%

<sup>a</sup> The threshold criteria for 4 and 400 ha are extrapolated values of the 95% prediction limit trendlines.

The threshold criteria estimated for flow duration were compared to the observed differences in flow duration at 70 sample sites, as 18 sample sites had drainage areas greater than 400 ha or their post-mine flow condition data were unavailable. Streamflow conditions at 38 of the sample sites indicated no change in flow duration as the observed differences did not exceed the threshold values. For the remaining 32 sample sites, regressive flow is likely at 17 sites and progressive flow is likely at 15 sites.

The threshold criteria estimated for extent of the flow condition were compared to the observed differences in the extent of the flow condition for 25 streams, as three streams had drainage areas greater than 400 hectares and flow condition data were unavailable at one additional stream. Streamflow conditions observed at 16 of the 25 streams, representing 37.15 km, maintained flow conditions consistent with pre-mine streamflow conditions, six streams exhibited regressive flow trends (573 m), and three streams exhibited progressive flow trends (274 m). The net change in the extent of the flow conditions (299 m), as determined by total length (37.15 km), was <1%.

Although correlation analysis could be a reliable tool for qualifying biological response, it may not be as statistically robust as probabilistic risk assessment for quantifying the magnitude of potential consequence.

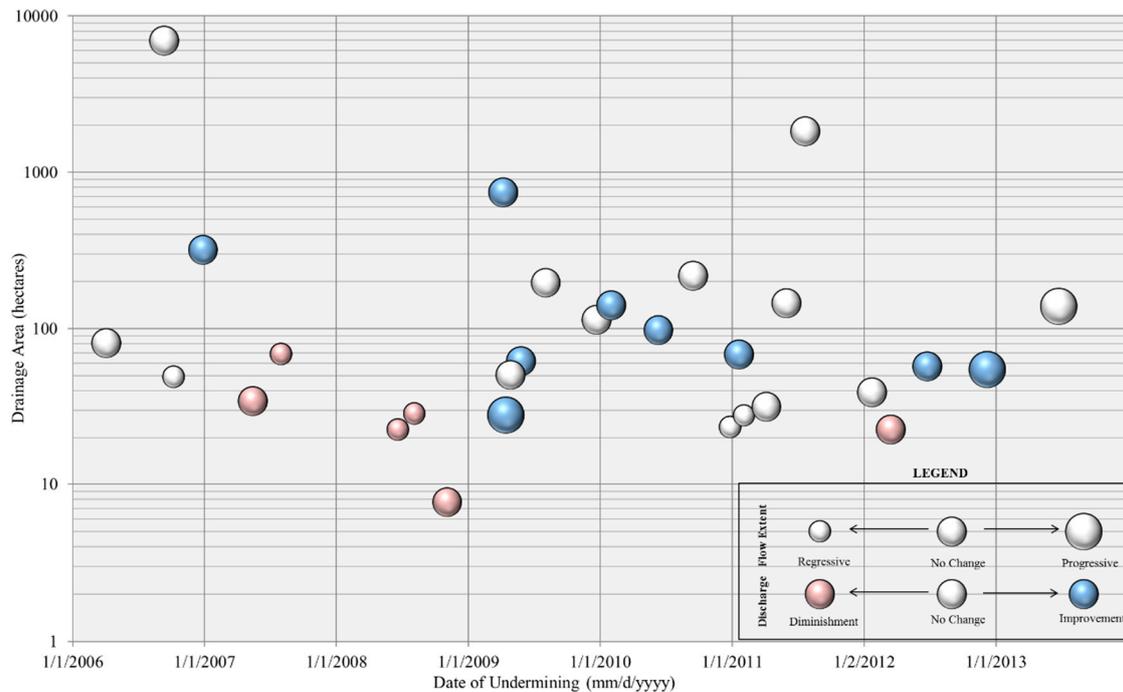
#### 4. Discussion

The conceptual foundation used to facilitate environmental flow assessment is based on classifying streams according to ecologically meaningful streamflow characteristics which are intuitively amenable to other research studies [1]. A risk-based approach can be initiated by using simple tools and readily available data, provided the hydrologic analysis generates information of sufficient certainty. As the need for prediction resolution increases, the environmental flow assessment method becomes more sophisticated and resource-intensive [1].

Existing large-scale environmental flow assessment methodologies qualified the various degrees of hydrologic alteration based upon the quantity of hydrologic change (Montana method) and the frequency, duration, and timing of hydrologic change (range in variability approach). The hydrologic methods divide the range of possible hydrologic alteration into three classes of equal range: 0% to 33% represents little or no alteration, 34% to 67% represents moderate alteration, and 68% to 100% represents a high degree of alteration [5,6]. Table 1 shows that the probabilistic risk assessment methodology produced threshold criteria results that are within the range of moderate hydrologic alteration as described by these methods. Table 1 results are also logical from a perspective of statistical variance, as these threshold criteria trend from higher values of acceptable alteration for headwater streams, which exhibit flashier flow characteristics as the baseflow is primarily derived from precipitation, to lower

values of acceptable alteration for larger order streams, which display relatively more uniform flow characteristics as the source of baseflow is sustained through groundwater influence.

The hydrologic response observed at 14 of the 29 streams indicated no difference between the pre-mine and post-mine distribution of stream discharge. For the remaining 15 streams, six streams had downward trends indicating diminished post-mine discharge, and nine streams had upward trends indicating improved post-mine discharge, as shown in Figure 5.

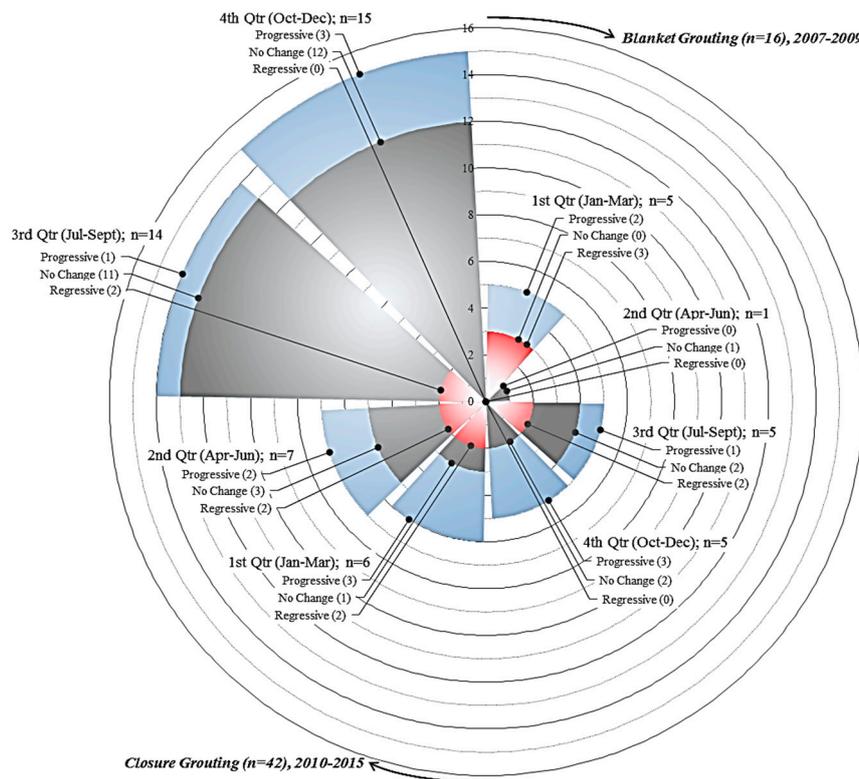


**Figure 5.** Nonparametric hypothesis testing and the extent of change in intermittent flow conditions as related to threshold criteria representative of lower and upper 95% prediction limits for the 29 streams within the study area.

Generally, it is suggested that streams within watersheds smaller than 80 ha experience dewatering longer than streams with greater drainage areas [10,11]. For this study, it does not appear that streams with watersheds smaller than 80 ha were more susceptible to hydrologic effects of mine subsidence; the difference is posited to be caused by a lack of restoration interventions within the previously studied streams. The streams that exhibited diminished post-mine streamflow conditions in this study ( $n = 6$ ) are located within watersheds smaller than 80 ha; however, seven streams within watersheds smaller than 80 ha indicated no difference, and five streams indicated improved post-mine streamflow conditions. The streams that exhibited diminished post-mine streamflow conditions did not affect downstream hydrologic conditions, as sample sites located immediately downstream of the confluence between the studied stream and its receiving stream indicated no difference between the pre-mine and post-mine distribution of discharge in all cases for which data were available.

Of the six streams identified with diminished streamflow conditions, five were mined between 2007 and 2009. This suggested a possible correlation between streambed grouting techniques and flow loss restoration. Figure 6 shows that, when analyzing all sample sites where flow loss restoration was performed ( $n = 58$ ), blanket grouting (performed from 2007 to 2009) was less effective than the performance-based closure-style grouting used since 2010.

Of the 16 samples sites where blanket grouting was performed (2007 through 2009), five sites exhibited diminished streamflow conditions, an incidence of 31%. Of the 42 sample sites where closure-style grouting was performed (since 2010), six sites exhibited diminished streamflow conditions, an incidence of 14%.



**Figure 6.** Change in the duration of intermittent flow condition as it relates to grouting technique and timing (quarter of year) at 58 sample sites.

Figure 6 also shows that flow loss restoration is less effective when grouting (blanket or closure-style) is performed in the first quarter (January–March) of the calendar year. Grouting was performed during the first quarter at 11 sample sites, five of which were identified with diminished streamflow conditions. This represents an incidence of 45%. The incidence of sample sites with diminished streamflow conditions when grouting was performed during the second, third, and fourth quarters was 25% (two of eight sample sites), 21% (four of 19 sample sites), and 0% (zero of 20 sample sites), respectively.

It is logical that flow loss restoration is less effective in the winter months and more effective in the summer and fall. Groundwater levels of unconfined aquifers follow a natural cyclic pattern of seasonal fluctuation. Typically, hydraulic head increases during the winter and spring in response to greater recharge rates, and decreases during the summer and fall in response to increased evapotranspiration rates. It is possible that the long-term effectiveness of flow loss restoration is affected when grout is injected into near-surface bedrock under a high hydraulic head condition, as this condition could affect the integrity of the grout through dilution or erosion prior to attaining a complete in situ cure.

## 5. Conclusions

Various approaches were examined to develop an environmental flow assessment method for headwater streams to evaluate the magnitude of likely hydrologic alteration and to determine the effectiveness of stream restoration techniques. These methods were derived from a common dataset and are differentiated by the technique used for statistical analysis. While most of the 88 monitoring sites and 29 streams showed no effect on flow or recovery following mining, four sample sites with significant hydrologic change were detected by all three methods: the univariate search method of linear regression, probabilistic risk assessment, and the flow duration parameter of correlation analysis. These four sample sites would also be classified as having a high degree of alteration, as determined by the Montana method and range in variability approach. Circumstantial evidence supports these

findings, as diminished streamflow observed at these four sites following flow loss restoration persisted longer than sample sites determined to be restored effectively.

Probabilistic risk assessment and correlation analysis of change are objective, pragmatic, and statistically delimited. However, hydrologic alteration, as determined by these environmental flow assessment methodologies, does not necessarily imply that ecological impact occurred, but rather that there is a greater likelihood of ecological change for a stream.

One issue confronting the study of environmental flows is the shortage of research documenting practical application, despite expanding literature on method development [2]. Empirical models that directly predict ecological response to various degrees of flow alteration are not readily available [1]. The establishment of a holistic methodology is an alternative approach via which environmental flow assessment can advance through the examination of multidisciplinary elements [8].

Additional study will be considered to investigate changes in species composition and abundance, biological metrics, habitat characteristics, and other applicable ecological indicators as they relate to various degrees of hydrologic alteration determined by the probabilistic risk assessment and correlation analysis of change criteria. The Streams TGD includes two standard biological sampling procedures to establish pre-mine conditions and to evaluate post-mine changes: (i) a field-based assessment of biological indicators, and (ii) a protocol for collecting, processing, and calculating an index of biotic integrity score using benthic macroinvertebrates. Biomonitoring locations are paired with over 50 sample sites used for this study. Future research will evaluate the hydro-ecological relationship of headwater streams and determine its application to compliance, permitting, and regulatory functions.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2306-5338/6/3/67/s1>: Table S1: Z-test results; Table S2: T-test results; Table S3: F-test results; Table S4: Wilcoxon rank sum test results; Table S5 Summary.

**Author Contributions:** J.M.S., conceptualization, methodology, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, and project administration; B.C.B., conceptualization, methodology, investigation, resources, data curation, writing—review and editing, and supervision; M.L.S., conceptualization, methodology, investigation, writing—review and editing, visualization, and project administration; M.R.H., conceptualization, methodology, writing—review and editing, visualization, and project administration.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors wish to thank the following people for their contributions to this paper: David Quatchak (editing), Dan Connell and Benjamin Blasingame (statistical data analysis), and Cynthia Cyprych (GIS support). The authors also acknowledge John Owsiany (pioneering CPCC's streamflow restoration concepts) and Burns Drilling and Excavating (restoration logistics, operation, and maintenance).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Poff, N.L.; Richter, B.D.; Arlington, A.H.; Bunn, S.E.; Naiman, R.J.; Kendy, E.; Acreman, M.; Apse, C.; Bledsoe, B.P.; Freeman, M.C.; et al. The ecological limits of hydrologic alteration (ELOHA): A new framework for developing regional environmental flow standards. *Freshw. Biol.* **2010**, *55*, 147–170. [[CrossRef](#)]
2. Tonkin, J.D.; Jahnig, S.C.; Haase, P. The Rise of Riverine Flow-ecology and Environmental Flow Research. *Environ. Process.* **2014**, *1*, 323–330. [[CrossRef](#)]
3. Davidian, J. *Computation of Water-Surface Profiles in Open Channels: U.S. Geological Survey Techniques of Water-Resources Investigations; Book 3*; U.S. Government Publishing Office: Washington, DC, USA, 1984; Chapter A15; p. 48.
4. Shearman, J.O. *Computer Applications for Step-Backwater and Floodway Analysis: U.S. Geological Survey Open-File Report 76-499*; U.S. Geological Survey: Reston, VA, USA, 1976; p. 103.
5. Tennant, D.L. *Instream Flow Regimens for Fish, Wildlife, Recreation and Related Environmental Resources*; U.S. Fish and Wildlife Service: Billings, MT, USA, 1975.
6. Richter, B.D.; Baumgartner, J.V.; Braun, D.P.; Powell, J. A Spatial Assessment of Hydrologic Alteration Within a River Network. *Regul. Rivers: Res. Mgmt.* **1998**, *14*, 329–340. [[CrossRef](#)]

7. Martin, D.M.; Labadie, J.W.; Poff, N.L. Incorporating social preferences into the ecological limits of hydrologic alteration (ELOHA): A case study in the Yampa-White River basin, Colorado. *Freshw. Biol.* **2015**, *60*, 1890–1900. [CrossRef]
8. Tharme, R.E. A global perspective on environmental flow assessment: Emerging trends in the development and application of environmental flow methodologies for rivers. *River Res. Appl.* **2003**, *19*, 397–441. [CrossRef]
9. Carver, L.; Rauch, H.W. Hydrogeologic Effects of Subsidence at a Longwall Mine in the Pittsburgh Coal Seam. In Proceedings of the Thirteenth International Conference on Ground Control in Mining, West Virginia University, Morgantown, WV, USA, 2–4 August 1994; College of Mineral and Energy Resources, West Virginia University: Morgantown, WV, USA, 1995; pp. 298–307.
10. Dixon, D.Y.; Rauch, H.W. The Impact of Three Longwall Coal Mines on Streamflow in the Appalachian Coal Field. In Proceedings of the Ninth International Conference on Ground Control in Mining, West Virginia University, Morgantown, WV, USA, 4–6 June 1990; College of Mineral and Energy Resources, West Virginia University: Morgantown, WV, USA, 1990; pp. 169–182.
11. Gill, D.R. Hydrogeologic Analysis of Streamflow in Relation to Underground Mining in Northern West Virginia. Unpublished Master's Thesis, Department of Geology and Geography, West Virginia University, Morgantown, WV, USA, 2000.
12. Hill, J.G.; Price, D.R. The Impact of Deep Mining on an Overlying Aquifer in Western Pennsylvania. *Groundw. Monit. Remediat.* **1983**, *3*, 138–143. [CrossRef]
13. Leavitt, B.R.; Gibbens, J.F. Effects of Longwall Coal Mining on Rural Water Supplies and Stress Relief Fracture Flow Systems. In Proceedings of the Third Workshop on Surface Subsidence Due to Underground Mining, West Virginia University, Morgantown, WV, USA, 1–4 June 1992; College of Mineral and Energy Resources, West Virginia University: Morgantown, WV, USA, 1992; pp. 228–236.
14. Walker, J.S. *Case Study of the Effects of Longwall Mining Induced Subsidence on Shallow Ground Water Sources in the Northern Appalachian Coalfield*; RI 9198; U.S. Bureau of Mines, U.S. Printing Office: Washington, DC, USA, 1988.
15. Hasenfus, G.J.; Johnson, K.L.; Su, D.W.H. A Hydrogeomechanical Study of Overburden Response to Longwall Mining. In Proceedings of the Seventh International Conference on Ground Control in Mining, West Virginia University, Morgantown, WV, USA, 3–5 August 1988; College of Mineral and Energy Resources, West Virginia University: Morgantown, WV, USA, 1988; pp. 149–162.
16. Kendorski, F.S. Effect of Full-Extraction Underground Mining on Ground and Surface Waters: A 25-Year Retrospective. In Proceedings of the Twenty-fifth International Conference on Ground Control in Mining, West Virginia University, Morgantown, WV, USA, 1–3 August 2006; College of Mineral and Energy Resources, West Virginia University: Morgantown, WV, USA, 2006.
17. Peng, S.S. *Surface Subsidence Engineering*; Society for Mining; Metallurgy and Exploration: Littleton, CO, USA, 1992.
18. Schubert, J.P. Fracture Flow of Ground Water in Coal Bearing Strata. In Proceedings of the Symposium on Surface Mining Hydrology, Sedimentology and Reclamation, Lexington, KY, USA, 1–5 December 1980.
19. Stoner, J.D. Probable Hydrologic Effects of Subsurface Mining. *Groundw. Monit. Remediat.* **1983**, *3*, 128–137. [CrossRef]
20. Wyrick, G.G.; Borchers, J.W. *Hydrologic Effects of Stress-Relief Fracturing in an Appalachian Valley*; United States Geological Survey Water-Supply Paper 2177; U.S. Department of the Interior, U.S. Printing Office: Washington, DC, USA, 1981.
21. Pennsylvania Department of Environmental Protection. Surface Water Protection—Underground Bituminous Coal Mining Operations. *Pennsylvania Bulletin* **2005**, *35*, 563–2000–655.
22. Rantz, S.E. *Measurement and Computation of Streamflow*; Measurement of Stage and Discharge; Geological Survey Water-Supply-Paper 2175; United States Geological Survey: Reston, VA, USA, 1982; Volume 1.
23. Helsel, D.R.; Hirsch, R.M. *Statistical Methods in Water Resources, Techniques of Water-Resources Investigations*; Book 4; United States Geological Survey: Reston, VA, USA, 2002; Chapter A3.
24. Bluman, A.G. *Elementary Statistics: A Step by Step Approach*, 6th ed.; McGraw-Hill: New York, NY, USA, 2007.
25. Whipple, K.X. Bedrock Rivers and the Geomorphology of Active Orogens. *Annu. Rev. Earth Planet. Sci.* **2004**, *32*, 151–185. [CrossRef]

