Received: 4 May 2021 | Revised: 29 August 2021 | Accepted: 12 October 2021

DOI: 10.1002/hyp.14412

RESEARCH ARTICLE

WILEY

Effects of hydrologic variability and remedial actions on first flush and metal loading from streams draining the Silverton caldera, 1992–2014

Tanya N. Petach[1] | Robert L. Runkel[2] | Rory M. Cowie[3] | Diane M. McKnight[1]

[1]Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, Colorado, USA

[2]U.S. Geological Survey, Denver, Colorado, USA

[3]Alpine Water Resources, Silverton, Colorado, USA

Correspondence
Tanya N. Petach, Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, CO, USA.
Email: tanya.petach@colorado.edu

Funding information
U.S. Geological Survey; National Science Foundation, Grant/Award Number: DGE 1650115

Abstract

This study examined water quality in the upper Animas River watershed, a mined watershed that gained notoriety following the 2015 Gold King mine release of acid mine drainage to downstream communities. Water-quality data were used to evaluate trends in metal concentrations and loads over a two-decade period. Selected sites included three sites on tributary streams and one main-stem site on the Animas River downstream from the tributary confluences. During the study period, metal concentrations and loads varied seasonally and annually because of hydrologic variability and remedial actions designed to ameliorate the effects of acid mine drainage. Water-quality data were divided into two periods based on the timing of remedial activities in the watershed. The first period includes active water treatment, surface reclamation and installation of bulkheads in adits; the second period includes the decade following these activities. Water-quality data were used to estimate annual and monthly zinc loads using the Adjusted Maximum Likelihood Method (using LOADEST software) and U.S. Geological Survey streamflow data. This study presents one of the first applications of LOADEST focused on metal loads. Monthly flow-weighted concentrations were analysed using a Mann-Kendall trend test to determine the direction, magnitude, and significance of temporal trends in zinc loading in any given month and using *t*-test comparisons between the two periods. Zinc loads estimated for the Animas River below the tributaries indicate decreased zinc loading during the rising limb of the hydrograph in the second period, perhaps reflecting a reduction of snowmelt-derived zinc load following surface reclamation activities. In contrast, base-flow zinc loading increased at the main-stem site, perhaps because of the cessation of water treatment in tributary streams. Flow weighting of monthly load estimates yielded increased statistical significance and enabled more nuanced differentiation between the effects of hydrologic variability and remedial activities on zinc loading.

KEYWORDS
abandoned mines, acid mine drainage, LOADEST, remediation, water quality

1 | INTRODUCTION

The generation of acid rock drainage (ARD) affects natural waters in sulphide-mineral rich regions worldwide. Hard-rock mines in the United States have left a legacy of more than 200 000 abandoned or inactive mines (U.S. Environmental Protection Agency, 1997), many of which are in headwater regions of the Rocky Mountains (Riebsame et al., 1997). These mine discharge waters are often affected by ARD and are characterized by low pH and high concentrations of metals. ARD is predominantly produced by the oxidation of sulphide-rich minerals (e.g. pyrite) that typically occurs in a series of multi-step reactions and is catalysed by acidophilic bacteria such as *Thiobacillus ferrooxidans*, commonly found naturally in ARD (Singer & Stumm, 1970).

Mining-affected surface water can have elevated acidity and metal concentrations in which fish and insects cannot survive (Besser & Brumbaugh, 2007; Besser & Leib, 2007). In some locations where organisms from lower trophic levels are able to survive, the riparian birds become affected by the uptake of metals through the food chain (Larison et al., 2000). Elevated environmental metal concentrations also negatively affect beaver, raccoon, otter and muskrat populations (Ganoe, 2019; Wren, 1984). Additionally, aluminium and iron oxide deposits on streambeds can be toxic to aquatic organisms (Niyogi et al., 2002), including rainbow trout, (Todd et al., 2006) microbes, algae and macroinvertebrates (McKnight & Feder, 1984). Metal concentrations in ARD affected streams are tied to patterns in discharge, and can be highly temporally variable (August et al., 2002; Brooks et al., 1998; Nordstrom, 2009; Shaw et al., 2020).

Metal concentrations in streams and rivers are controlled by a variety of hydrologic processes. In streams with hydrographs dominated by snowmelt, metal concentrations can spike in the early spring as snowmelt infiltrates mine dumps, tailings and soils that have been unsaturated during winter, promoting the generation of ARD and remobilizing evaporative salts (Brooks et al., 1998). This "first flush" of metals from the watershed typically occurs during the initial portion of the rising limb of the hydrograph, before peak streamflow. Following the first-flush period, streamflow rises throughout spring snowmelt and metal concentrations typically reach their lowest concentrations at peak snowmelt. An increase of metal concentration is often observed on the falling limb of the hydrograph (Brooks et al., 2001), and maximum concentrations are reached when streamflow returns to base-flow conditions (base-flow). In addition to these snowmelt-driven events, monsoons can cause spikes in concentration, which are exacerbated by dry antecedent conditions that allow for build-up of soluble salts in soils and tailings (Runkel et al., 2016). Although many surface-water ARD sources are influenced by local hydrologic conditions, draining mine adits are a notable exception. Most adits discharge groundwater at relatively constant rates throughout the entire year (August et al., 2002; Church et al., 2007) and are less affected by either first- or post-rainfall flush phenomena.

Numerable remediation strategies have been employed in mining-affected sites (Johnson & Hallberg, 2005; Walton-Day, 2003) with varying complexity and efficacy. Surface reclamation includes the removal of waste rock and mill tailings, diversion of runoff away from mine sites, revegetation, and erosion control. Because surface reclamation targets periods of elevated runoff, it primarily affects water quality during the first flush and rainfall runoff. Other remedial strategies, including active treatment and bulkhead emplacement, target draining mine tunnels that contribute metals and acidity to the receiving streams throughout the year. Active treatment typically involves raising the pH of ARD through a base addition (e.g. lime) that initiates a series of pH-dependent reactions that decrease metal solubility (Johnson & Hallberg, 2005; Runkel et al., 2012). Although active treatment can be effective at improving water quality throughout the year, treatment systems require continuous operation with ongoing costs and maintenance and are often considered to be end-of-pipe solutions (Skousen et al., 1998; Wireman & Stover, 2011). Bulkheads, dam-like structures that impede the flow from draining mine tunnels, are considered a low-cost alternative to active treatment, but their efficacy is unclear (Walton-Day et al., 2021). Following the emplacement of a bulkhead, the water table behind the bulkhead rises, flooding mine workings and limiting oxygen exposure of pyrite, which can slow down ARD generation. Bulkheads can also redirect ARD to new surface locations; sometimes shunting ARD affected water from one drainage to a neighbouring stream (Walton-Day et al., 2021; Walton-Day & Mills, 2015).

Evaluating the long-term efficacy of remediation strategies is complicated by hydrologic variability in streamflow, timing of snowmelt, and frequency of rainfall events. Furthermore, periods of high flow vary in magnitude and timing between years. Large hydrologic variability can lead to different preferential flow paths and different levels of exposure to source areas. This extensive hydrologic variability leads to variable metal loads (concentration time's streamflow) from year to year, simply because of the direct relation between load and streamflow (Walton-Day et al., 2021). In above average snowpack years, a large watershed surface area is in contact with water and may produce ARD, whereas the opposite is true in dry years (Runkel et al., 2009). The amount of flushing is also important; snowpack can cover the same area in dry and wet years, but less runoff in dry years causes reduced contact between mine tailings and runoff water. Hydrologic variability complicates inter-year comparisons of water-quality data and may obfuscate remediation outcomes (Runkel et al., 2009). Untangling the differential effects of hydrologic variability and remedial actions can be complex but is essential for evaluating the effects of remedial actions on water quality.

Monitoring of water quality in mined watersheds typically involves the collection of discrete samples at stream gages. Discrete estimates of metal load can then be determined as the simple product of concentration and streamflow. Discrete data are often sparse because of infrequent sampling, which can complicate inferences about metal behaviour during varied hydrologic conditions. Load-estimation methods (Aulenbach et al., 2016) provide a means to extend the observations by utilizing the observed relations between load and streamflow. Regression-based methods such as LOADEST (Load Estimator; Runkel et al., 2004), for example, use observations of load to calibrate a regression model that expresses metal load as a function of streamflow and time. The regression equation is then used with the continuous measurements of streamflow to generate estimates of daily, monthly and annual load.

Load estimation techniques are frequently used to quantify nutrient loads and most LOADEST applications to date are focused on nitrate, phosphorus, dissolved organic carbon or sediment load (e.g. Drake et al., 2021; additional applications at https://water.usgs.gov/software/loadest/apps/). LOADEST applications involving metals and/or mining-affected watersheds are virtually non-existent, possibly because of the reactive nature of many metals. The recent publication by Rossi et al. (2021), uses LOADEST to estimate sulphate loads from the El Indio mining district in northern Chile. Attempts to estimate arsenic, copper, and iron loads were unsuccessful, however, and this may be caused by the reactive nature of these constituents (i.e., the effects of pH-dependent reactions on concentration result in poor correlations between load and streamflow). Additional metals/mining applications include the work of Donato (2006) and Shrestha et al. (2020). The research presented herein represents one of the first LOADEST applications focused on metal loading.

One challenge in addressing water-quality assessment and quantification is discerning between hydrologic variation and remediation-based effects. This study quantifies changes in water quality over time within a mined and remediated watershed in southwestern Colorado, the upper Animas River watershed. Zinc is the analyte of interest in this study because it is not subject to pH-dependent reactions nor sorption processes in this watershed (Schemel et al., 2007) and is therefore a relatively conservative analyte in this system. Numerous remediation strategies were employed in this watershed including surface reclamation, active treatment, and bulkhead implementation, which provide a long-term (20-year) window into the relative effects of different remediation strategies. Specific objectives of this study are to: (1) document changes in zinc concentrations and loads over a two-decade period; (2) quantify the hydrologic effects on zinc concentration and load, including the effects of first flush during snowmelt; and (3) ascertain the degree to which the observed changes in concentration and load are attributable to different remedial activities within the watershed. This study compiled 23 years of water-quality data in the ARD affected and remediated upper Animas River watershed and daily streamflow measurements from four nearby U.S. Geological Survey (USGS) gages. The computer code LOADEST was used to estimate monthly and annual zinc loads, and flow-weighting techniques were implemented to distinguish changes between hydrologic variability and remediation efforts.

2 | DATA AND METHODS

This study aims to quantify changes in water quality in the upper Animas River watershed following multi-decade remediation efforts. Highly variable annual streamflow and sparse data collection complicate interpretation of water-quality results. Discrete concentration and load data are supplemented by estimates of load provided by LOADEST. To alleviate variability derived from interannual streamflow variability, data comparisons were executed with: (1) discrete concentration and load data, (2) estimated loads, and (3) flow-weighted concentration estimates.

2.1 | Site description

The upper Animas River watershed is in southwestern Colorado near the town of Silverton. (Figure 1). Discovery of gold and other precious metals in the late 1800s led to the development and later abandonment of over 1500 hard-rock mines in the region (Jones, 2007). Both disseminated and point sources of ARD are present in this watershed (Kimball et al., 2002). Forty-eight legacy-mining sites within the upper Animas River watershed are currently listed under the U.S. Environmental Protection Agency Superfund program collectively known as the Bonita Peak Mining District. The superfund listing followed an event in 2015 when a soil plug was disturbed near the Gold King Mine entrance, and an estimated pulse of 11 million litres of metal-rich water were released to the Animas River (Gobla et al., 2015; Rodriguez-Freire et al., 2016). The volume of ARD released in the spill is produced by draining mines and natural sources every few days across the upper Animas River watershed, but the Gold King Mine spill reached the lower Animas River downstream from Cement Creek, and the river ran an orange colour, carrying the pulse of ARD into New Mexico, Navajo Nation, and Utah. Prior to the Gold King release, metal concentrations exceeded aquatic-life standards in many stream reaches because of mining activities combined with natural sources of ARD (Besser et al., 2007; Kimball et al., 2007). Although this study focuses on water quality pre-Gold King release, this event highlights the importance of mine reclamation and quantification of water quality in the upper Animas River watershed. Numerous remedial actions that targeted these mines took place prior to the Gold King release; this period prior to the release (1991–2014) is the focus of this research.

The upper Animas River watershed is comprised of the upper Animas River, Cement Creek, and Mineral Creek, which converge near the town of Silverton. Each of the three tributaries has a corresponding USGS stream gage near its confluence with the main stem of the Animas River (stream gages A68, CC48 and M34 are shown in Figure 1, Table S1). A fourth stream gage is located below the confluences of the tributaries and measures the combined streamflow of all three tributary streams (stream gage A72 is shown in Figure 1, Table S1). The A72 stream gage was established by the Colorado Water Quality Control Commission and later adopted as the superfund water-quality compliance point for the upper Animas River watershed.

Streamflow in the Animas River is affected by the Rocky Mountain weather patterns. The watershed is set in rugged alpine to sub-alpine terrain with elevations ranging from 2830 m to more than 4050 m. Mean annual precipitation ranges from $\sim$600 to 1000 mm/year (NRCS, 2020), and streamflow is dominated by snowmelt runoff that typically occurs between April and June. The

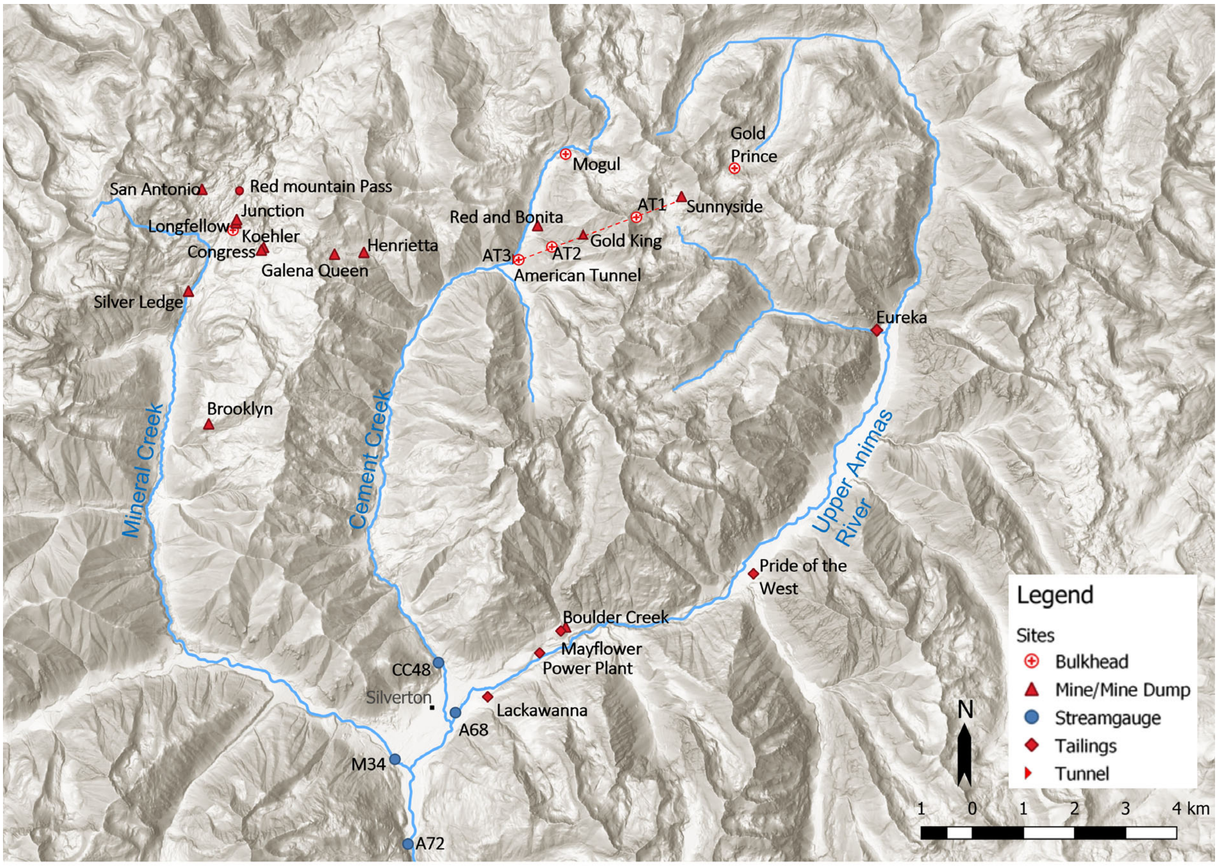

FIGURE 1 Map of the study region with highlighted locations of remediation activities

falling limb of the hydrograph (July–August) is punctuated by short-lived runoff spikes produced by summer monsoon events. Local geology of the upper Animas River watershed is a primary factor affecting the metal concentrations and pH of surface water in the study area (Church et al., 2007). The study area is located in heavily fractured and faulted Tertiary intrusive rocks that form the Silverton Caldera (Kimball et al., 2007). Hydrothermal alteration and historical mining activities compound ARD processes and influence acidity; stream pH of the study area ranges from 2.35 to 8.49 (Jones, 2007).

More than a century of mining activity resulted in miles of underground workings in the watershed. These mine workings provide preferential flow paths for groundwater that has reacted with mineralized rock to produce ARD (Bove et al., 2007; Church et al., 2007). Large volumes of waste rock and mill tailings result in a large surface area of pyrite that can oxidize and produce ARD. Following the closure of the last mine (Sunnyside Mine) and mill (Mayflower Mill) in 1991, the upper Animas River watershed underwent numerous remediation efforts. Nine bulkheads were installed by the Sunnyside Mining Company in the Bonita Peak area following the mine closure. Additional bulkheads were added later. A full overview of these efforts is provided by Finger et al. (2007); an abridged set of remedial actions for the three sub-watersheds is presented in Table 1 and discussed in the following paragraphs.

The largest mine in the Silverton area, the Sunnyside Mine, has workings that span the Cement Creek and upper Animas River watersheds. The American Tunnel was connected to the Sunnyside Mine workings in 1962 and served as the primary haulage and drainage tunnel thereafter. Mine drainage exiting the tunnel was treated from 1978 to 2003 using conventional lime treatment, which increased pH and lowered metal concentrations in the water, before discharging to Cement Creek (Walton-Day et al., 2021). The active conventional lime treatment was ceased in 2004 in accordance with a consent decree issued by the State of Colorado (Bonita Peak Community Advisory Group, 2016). While the treatment was in operation, the Sunnyside Mine company installed three bulkheads that led to lower outflow from the American Tunnel (Table 1). The outflow from several upgradient mines increased following the American Tunnel bulkheads installation (Walton-Day et al., 2020). This outflow led to a decision to install the bulkhead at the Mogul Mine in 2003. Additional mines and mill sites are located within the Cement Creek watershed, and these sites have also been subject to various forms of surface reclamation (Table 1; Finger et al., 2007).

The upper Animas River drains the area above stream gage A68 and contains numerous mines and mills, including the Gold Prince Mine in the headwaters and several mill sites along the river between Eureka and Silverton (Figure 1). Remedial actions in the upper Animas River have targeted surface reclamation efforts, and no substantial

TABLE 1 Overview of major remediation history in the study area

Basin	Type	Date	Site
Cement Creek	Active treatment	1978–2003	Lime treatment of American Tunnel
	Surface reclamation	1995	American Tunnel waste dump
	Bulkhead	1996	American Tunnel, AT #1
	Surface reclamation	1998, 2001	Galena Queen (Prospect Gulch)
	Bulkhead	2001	American Tunnel AT #2
	Bulkhead	2002	American Tunnel, AT #3
	Bulkhead	2003	Mogul Mine
	Surface reclamation	2004	Henrietta Mine (Prospect Gulch)
Mineral Creek	Surface reclamation	1996–1997	Longfellow Mine, Junction Mine, Koehler Tunnel
	Bulkhead	2003	Koehler Tunnel (re-grouted 2010)
	Surface reclamation	2004	Brooklyn Mine
	Surface reclamation	2010	Silver Ledge Mine
Upper Animas	Surface reclamation	1996–1997	Eureka and Pride of the West Mills, Boulder Creek Floodplain
	Surface reclamation	1997	Gold Prince Mine (Placer Gulch)
	Surface reclamation	1999–2003	Mayflower Mill (tailings and ponds)
	Surface reclamation	2000	Lackawanna Mill
	Surface reclamation	2003	Power Plant Tailings

attempts to install active treatment or bulkhead-draining mines have occurred to date. Although the Gold Prince Mine was a dry adit, a bulkhead was installed in anticipation of the water table rising after the installation of the American Tunnel (AT) bulkheads (Morris, written communication, March 8, 2021).

In Mineral Creek, mining activity was concentrated in the headwaters near Red Mountain Pass, where breccia-pipe chimney deposits were mined for silver, lead, and copper (Bove et al., 2007). Remedial actions within this tributary include surface reclamation efforts and the installation of a bulkhead in the Koehler Tunnel in 2003 (Table 1). Discharge from the Koehler Tunnel was the largest source of zinc in the upper Animas River watershed prior to bulkhead construction (Kimball et al., 2007).

2.2 | Data aggregation

Daily streamflow data for the four USGS stream gages during the study period (1992–2014) were obtained through the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019). Because of the limited discharge data in 1991 and 1994, these years were removed from analyses.

Water-quality data at the same four sites were obtained from the Water Quality Portal (WQP) (National Water Quality Monitoring Council, 2019) and the U.S. Environmental Protection Agency (EPA) (2019). Additional data from peer-reviewed studies at the four sites were included to augment the database (Bove et al., 2000; Kimball et al., 2002). The compiled database includes dissolved and total metal concentration, anions, conductivity, pH and streamflow. Zinc concentration data and field parameter data were taken at infrequent (monthly to sub-monthly) intervals. Streamflow data have a temporal resolution of approximately 15 min between measurements. A detailed description of quality-assurance procedures used in this study can be found in the Supplementary materials (S2); Suspect data were eliminated from the database using protocols adapted from Mast (2018).

The compiled water-quality data include analytical results for dissolved zinc concentration. Dissolved zinc was chosen as the parameter of interest for this study for several reasons. First, zinc is nominally conservative because it is not subject to pH-dependent reactions because the pH is generally below levels at which zinc precipitates. Previous studies in this watershed found that dissolved iron, aluminium, copper and zinc are removed through sorption (Paschke et al., 2005). However, removal of zinc through sorption is lower than other metals, likely because of the higher pH range over which it sorbs to iron and aluminium precipitates (Schemel et al., 2007). Zinc travels conservatively from headwater sources to the monitoring locations used in this study and serves as a good indicator of ARD sources within the watershed. Second, zinc is toxic to aquatic life, and base-flow zinc concentrations exceed aquatic-life standards over most of the study area. Zinc is therefore a target of past and future remedial actions (other metals of concern that are monitored include aluminium, cadmium, iron, copper and lead). Third, zinc concentrations are highly correlated with streamflow, whereas other, more reactive constituents are correlated with streamflow and pH. This correlation between concentration and streamflow is needed for successful application of load estimation methods.

2.3 | Hydrograph divisions and definition of study periods

Data analysis was conducted by the dividing the year into three hydrologic periods based on the snowpack hydrograph: rising limb of the hydrograph, falling limb of the hydrograph, and base-flow. The rising limb was defined by calculating a 30-year average number of days between a doubling of base-flow and the peak flow for the Animas River below Silverton (A72). From 1991–2019, the average number of days between an initial spring doubling of base-flow to peak flow is 54 days. For any given year, the rising limb period is defined as the 54 days prior to that year's peak snowmelt runoff. The falling limb is defined as the 90 days following peak flow, and base-flow is defined as the remainder of dates. Although rigid day-length definitions of rising and falling limb do not represent the entirety of wetter years' snowmelt, they facilitate inter-year comparisons during these hydrologic periods.

Study periods were defined by remediation and post-remediation years. Water years 1992–2004 were analysed together as most remediation projects and all bulkhead installation occurred during this period. Water years 2004–2014 were analysed as a second period corresponding to the cessation of active treatment and decreased remedial activity (Table 1).

2.4 | Observed Streamflow, concentration, and load

Welch's t-tests were used to assess differences among mean zinc concentration, load, and streamflow between the two periods; Mann-Kendall trend tests were used to examine trends in monthly flow-weighted concentration during the entire study period. Welch's *t*-test is designed to compare means of sample populations with unequal sample distribution variance that has been optimized for minimizing type I error (Welch, 1938, 1947, 1951). Single- and double-sided Welch's *t*-tests were employed in this study. Single-sided tests were used for data with notable differences between periods; double-sided tests were used for ambiguous comparisons. Mann-Kendall trend tests were used to determine the strength of monotonic trends in monthly flow weighted zinc concentration (Kendall, 1948; Mann, 1945). Mann-Kendall trend tests are commonly used for climactic and hydrologic time-series trend analysis as the test is derived from a rank correlation test and is modified to include time order (Hamed, 2008).

2.5 | Estimated loads

The compiled dataset includes daily streamflow values as well as irregular and numerous concentration data (Section 3.1). Monthly and annual load were estimated using the Adjusted Maximum Likelihood Method (AMLE), which is known to provide estimates of load with a low level of bias (Cohn, 1988; Cohn, 1995). The AMLE method is implemented within LOADEST software package (LOADESTimator: Runkel et al., 2004), Within LOADEST, observed loads are used to calibrate several regression equations that express load as a function of streamflow and time (Supplementary Materials, S3). The regression equation used for load estimation was then selected from the set of calibrated models based on the Akaike Information Criterion (AIC) (Runkel et al., 2004). Daily values of observed streamflow were then used with the selected model to provide daily load estimates, and average loads were calculated for each month and year.

Observed data were aggregated into 3-year moving windows prior to LOADEST calibration such that the LOADEST model was fit to 3 years of data and then used to calculate loads for 1 year at a time. This moving-window approach is modified from other long-term load estimation applications (Aulenbach, 2006; Aulenbach et al., 2016; Yochum, 2000), and uses a shorter moving window because of both short timescale changes in the watershed and a shorter period of interest. This 3-year moving window approach accommodates for changing conditions in the watershed associated with remediation (i.e. the form of the regression equation and its associated coefficients are updated through time to reflect changes in the load-streamflow relation that result from remedial actions) while still aggregating enough input data to yield reliable estimates. Under the moving window approach, three estimates of load were developed for any given year. The first estimate utilized calibration data (observed concentrations and streamflow) from the current year and the two previous years; the second estimate used calibration data from the preceding year, the current year, and the following year; and the third estimate used calibration data from the current year and the two following years. Load estimates presented herein are median values from the three estimates.

Calibration results from six of the 270 LOADEST runs yielded estimates of load bias (Bp) in excess of 25%, indicating the potential overestimation of zinc loads. Load estimates from these six models were not used in the analysis based on guidance from Runkel (2013). Summary statistics and bias diagnostics for the remaining models include R^2 (35%–99%, median = 92%), Bp (−3.8–24%, median = −0.4%), and the Nash Sutcliffe Efficiency Index (E; 0.42–0.99, median = 0.90). These results indicate that the estimated loads are generally unbiased (Bp mean near 0.0; loads are neither over- or underestimated) and that there is a good fit between the observed data and the selected regression model (E mean near 1.0) (Runkel, 2013). The form of the selected regression equation and summary statistics describing the accuracy of the estimated loads for each dataset are presented in the Supplementary Materials (S10, S11) and the dataset is available in the Consortium of Universities for the Advancement of Hydrologic Science Hydroshare collaboration tool (Petach et al., 2021).

Monthly and annual load estimates were flow weighted to account for variability in streamflow. Flow weighted concentration estimates were developed by dividing the estimated load by the average streamflow value observed during the period of interest (e.g. to determine monthly flow weighted concentration for July 2000, the estimated load for July 2000 was divided by the average daily streamflow value for July 2000). Trends in flow-weighted concentrations were then analysed using the Mann-Kendall procedure discussed in Section 2.4.

3 | RESULTS

3.1 | Streamflow

Streamflow at each of the four stream gages used in this study was variable. Seasonally, streamflow ranged over an order of magnitude, and annual snowpack variability led to large ranges in peak streamflow and total net water volume. In the upper Animas watershed, 2002 was the lowest water year and 1997 was the highest water year during the two-decade study period. Peak runoff at A72 only reached 12.5 m^3/s in 2002 but was 5.2 times higher at 64.6 m^3/s in 1997. Total water volume was 1.1×10^8 m^3 in 2002 but reached 3.6×10^8 m^3 in 1997 (3.4 times greater than 2002). (Figure 2).

3.2 | Observed concentration

The concentrations of dissolved zinc at the four sites varied based on time of year, streamflow and remediation activities, which predominantly occur during the first period (Figure 2). Figure 3 shows these same data plotted by day of year with the average daily streamflow over the period of record. In Figure 3, shifts in concentrations between the first period (shown in red) and the second period (shown in blue) are especially evident at stream gages CC48 and M34 but less notable at stream gage A68.

Data from the stream gage A68, upper Animas River, show that dissolved zinc concentrations are generally greatest during the onset of spring runoff. This first flush signal is evident in the first (1992–2003) and second (2004–2014) periods. Dissolved zinc concentrations peak at 1900 μg/L during the first flush at the onset of spring runoff in 1999 with minimum around 51 μg/L during low flow in 2014. The trends and magnitudes of these data are similar for the first and the second period).

The CC48, Cement Creek, data show that dissolved zinc concentrations are generally greatest during base-flow. This pattern is evident in the first and second periods, although it is much more pronounced during the second period. Dissolved zinc concentrations peak at 3060 μg/L during low streamflow months in the second period in 2007; the minimum of dissolved zinc data is 116 μg/L and occurs during the first period in 1995. The magnitude of this pattern is less during the first period and greater during the second.

The M34, Mineral Creek, data show that dissolved zinc concentrations are generally greatest during the onset of spring runoff. This first flush signal is evident in the first and second periods, although zinc concentrations are greater in the first period as compared with the second throughout the year. Dissolved zinc concentrations peak at 732 μg/L during the first flush at the onset of spring runoff in 1993; the minimum of dissolved zinc data is 51 μg/L and occurs

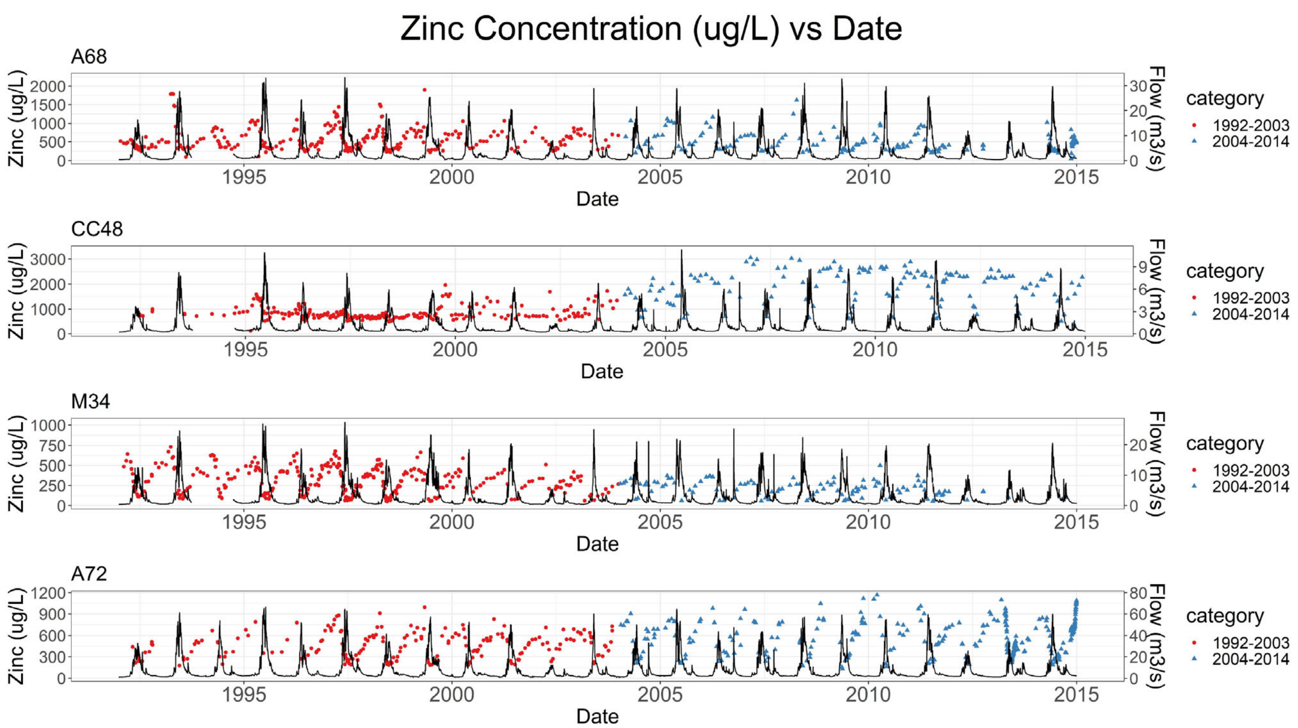

FIGURE 2 Streamflow (solid lines) and dissolved zinc concentrations at the four USGS stream gages. Red symbols are zinc concentrations during 1991–2003; blue symbols are zinc concentrations during 2004–2014. Date markings indicate January 1 of the year. From top to bottom, plots represent data from stream gages: A68, CC48, M34, and A72

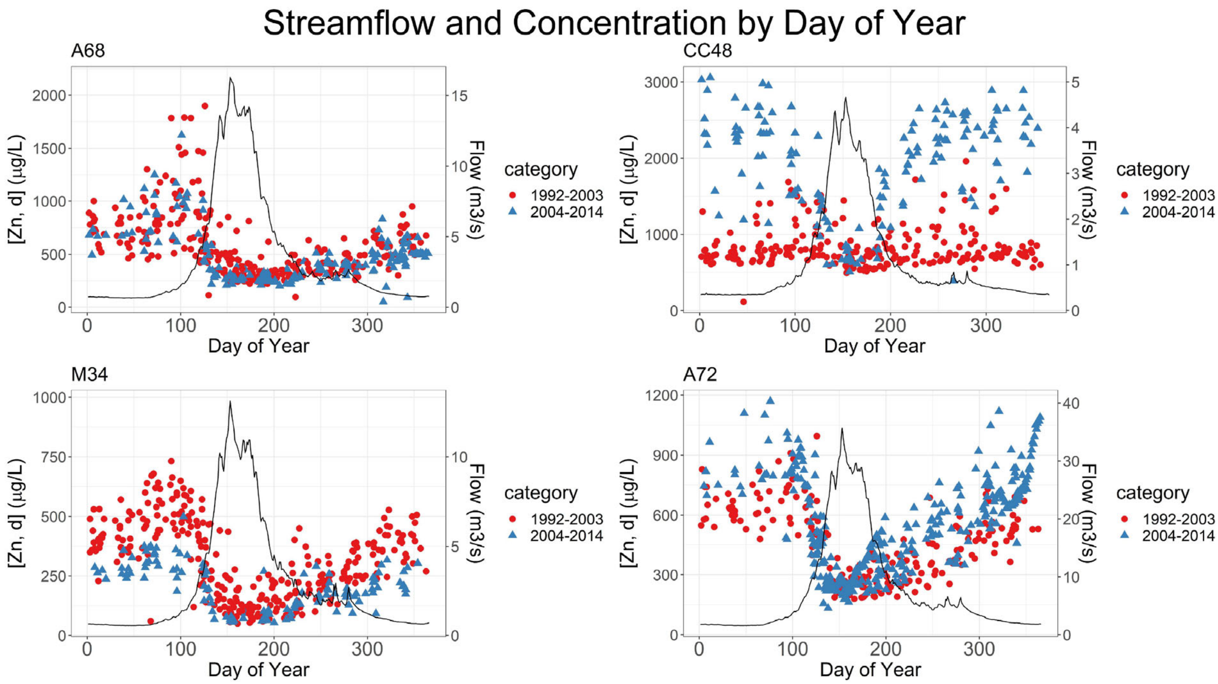

FIGURE 3 Observed zinc concentrations versus day of year with average daily discharge over period of record. Flow lines represent average daily streamflow over the entire study period (1992–2014). Red concentration measurements represent data from the first period; blue concentration measurements are from the second period. From top to bottom, plots represent data from stream gages: A68, CC48, M34, and A72

during the period of highest streamflow in 1998. The trends and magnitudes of these data are similar for the first and the second period.

The A72, Animas River below Silverton, data show that dissolved zinc concentrations are generally greatest during the second period. This signal is consistent with the pattern observed at CC48. Dissolved zinc concentrations peak at 1170 μg/L during the first flush at the onset of spring runoff in 2010; the minimum of dissolved zinc data is 133 μg/L and occurs during the period of highest streamflow in 2005. Zinc concentrations are slightly greater in the second period as compared with the first except during peak streamflow. Concentration measurements at A72 do not record a strong first flush response.

Comparison of average dissolved zinc concentrations for rising, falling, and base-flow portions of the hydrograph show different trends across the first and second periods for all gages (Figure 4).

Dissolved zinc concentrations at A68 demonstrate minor decreases during all three hydrologic periods in the second period as compared with the first (Figure 4). Similarly, as shown by the Welch's *t*-test, dissolved zinc concentrations at M34 were significantly less during the later second period as compared with the earlier period during all three hydrologic periods. In contrast, at CC48, dissolved zinc concentrations were significantly less (as determined by the Welch's *t*-test) during the first period as compared with the second period during all three hydrologic periods. These trends may indicate that during the second period, dissolved zinc levels are increasing at CC48.

Dissolved zinc concentrations at A72, Animas River below Silverton, were significantly less in the first period as compared with the second period during the falling-limb and base-flow hydrologic periods. No significant difference exists between the two averages during the rising limb phase, although all other differences between averages are statistically significant (Supplementary Materials S3).

3.3 | Observed load

The greatest zinc loads at all sites were observed (Figure 5) in times with high streamflow, likely because of the large volume of water discharged during spring snowmelt. At A68, zinc loads are statistically less in the second period during rising and falling limb hydrologic periods as compared to the first period. At M34, zinc loads significantly decreased from the first period to the second during all hydrologic periods. At CC48, zinc loads significantly increased from the first period to the second for base-flow conditions. At A72, zinc load in the second period as compared with the first significantly increased during base-flow but decreased during the rising limb. For p-value reporting on statistical tests, see (Supplementary Materials S4).

3.4 | Estimated load

Annual zinc loads, estimated using LOADEST, indicate a weak trend of decreasing zinc load at A68 and M34 and increasing zinc load at CC48. At A72, a relatively sporadic trend is observed with no net slope. Annual load estimates and flow-weighted annual concentrations estimates are reported in the Supplementary Materials S7, S8).

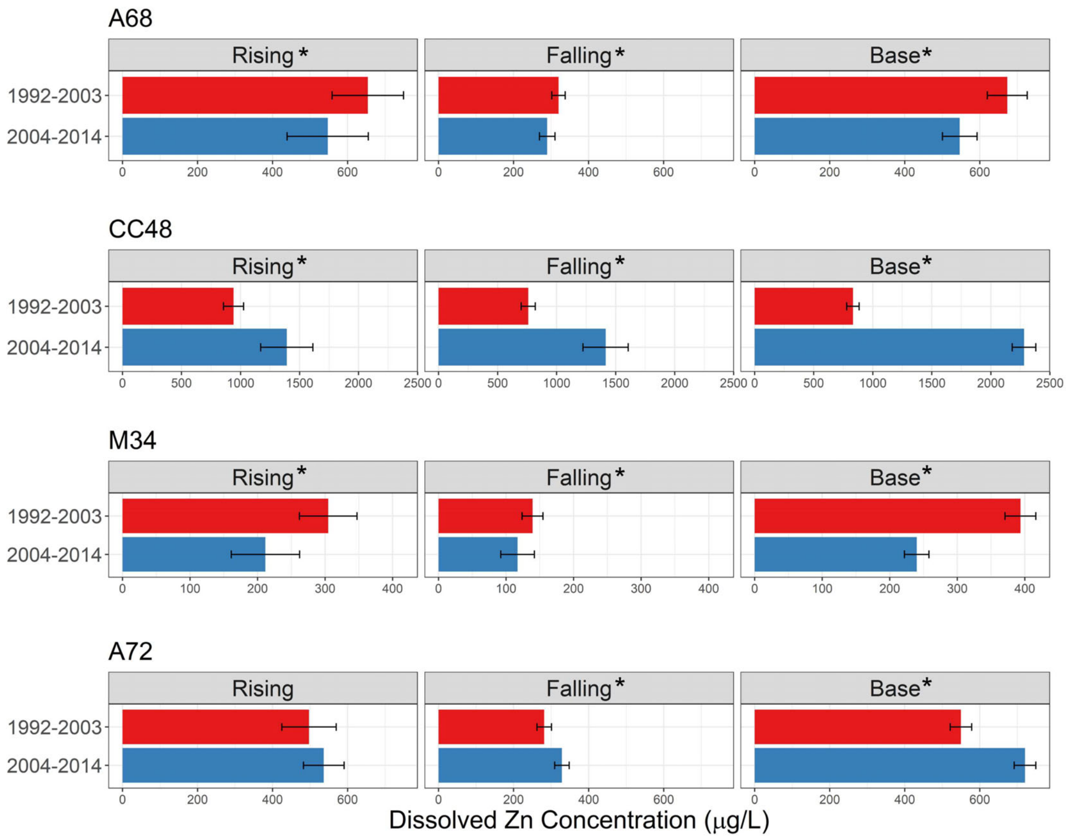

FIGURE 4 Concentration of dissolved zinc in hydrologic limbs. Average dissolved zinc concentration during each period of the hydrograph (rising limb, falling limb, and base-flow). Data are divided into two periods: 1991–2003 and 2004–2014. Error bars represent 95% confidence intervals on the average. Different magnitudes of the *x*-axes. Statistically significant differences are indicated by *

Monthly zinc loads plotted through time (Figure 6) show similar trends to annual zinc loads, but with exacerbated trends in the zinc load increase through time at CC48 and the zinc load decreases through time at A68 and M34. Time-dependent changes in monthly loads fluctuate at A68 and A72, but CC48 records a clear increase in base-flow zinc loading during the second period (Figure 6). At site M34, a subtle decline in base-flow zinc loads is observed along with the greatest spring peak loading during the first time period.

Flow-weighted monthly concentrations were analysed using a Mann-Kendall trend test to determine the direction, magnitude, and significance of changes in zinc loading through time in any given month (Figure 7). At A68 and M34, results of the Mann-Kendall trend test indicate decreasing zinc loading over time with larger magnitude slope changes in low streamflow months (Mann Kendall slopes <0). In contrast, results at CC48 indicate increasing zinc loading over time in all months, and results at A72 indicate increasing zinc loading during 10 of 12 months (Mann Kendall slopes >0). Larger magnitude slopes and higher statistical significance (lower *p*-values) were recorded during low streamflow months at these sites (Figure 7). Raw monthly load analyses using a Mann-Kendall trend test (not streamflow weighted concentration data) and a detailed streamflow analysis can be found in the (Supplementary Materials S5, S6, S9).

4 | DISCUSSION

4.1 | Zinc concentration and load in tributaries: decreases at M34 and A68 during certain periods and increases at CC48

Changes in zinc concentration and load were observed at all sites from the first (1992–20 003) to the second (2004–2014) time periods. Slight decreases in dissolved zinc concentration (Figure 4) and load (Figure 5) occurred at A68 between the first and second time periods, particularly during rising streamflow months. These results correspond with targeted surface remediation activities in the upper Animas River, which removed and capped waste rock and tailings prone to first flush pulses of dissolved zinc. For M34, results indicate decreased zinc concentrations (Figure 4) and loads (Figure 5) during base-flow and rising limb months. The bulkhead installation at the Koehler Tunnel may have contributed to decreased base-flow dissolved zinc concentrations. In addition, like A68, surface reclamation in the watershed is likely related to the diminished first flush signal. The significance of these changes is often greater in streamflow-weighted concentration data, which may be useful for future studies investigating similar water-quality trends.

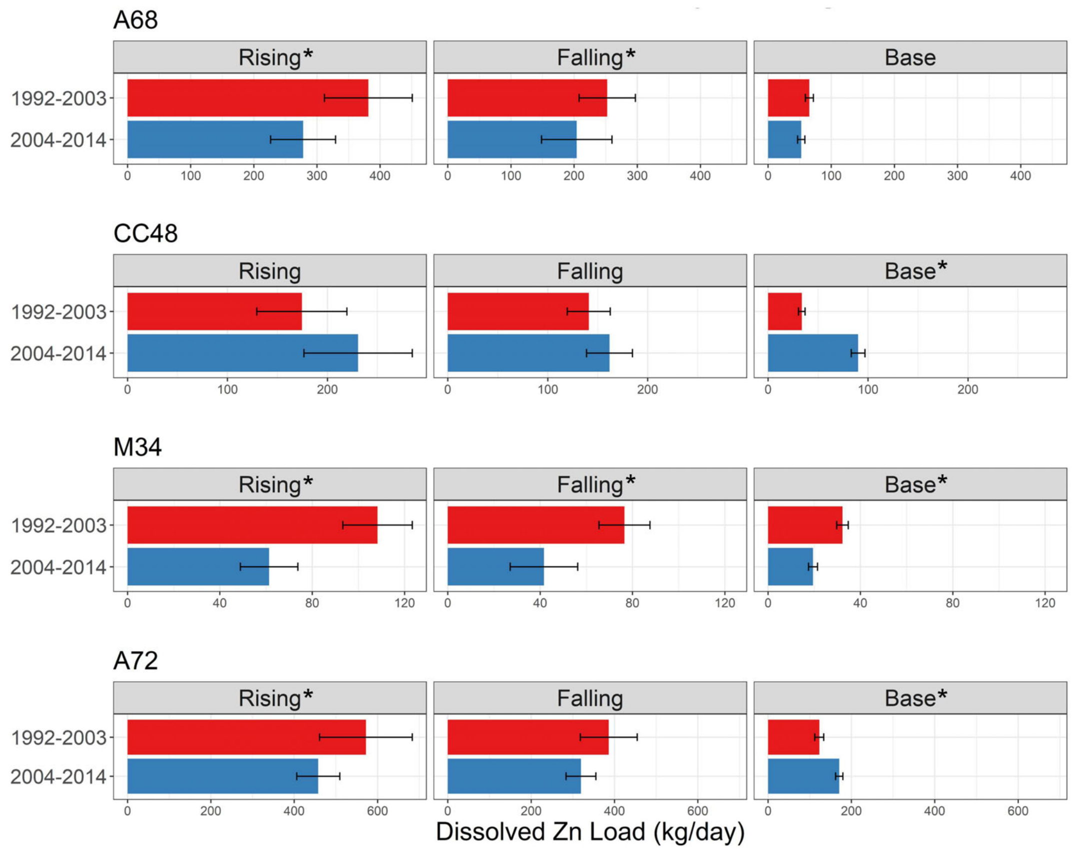

FIGURE 5 Dissolved zinc load in hydrologic limbs. Average dissolved zinc load during each period of the hydrograph (rising limb, falling limb, and base-flow). Data are divided into two periods: (1) 1992–2003 and (2) 2004–2014. The different *x*-axis scales. Error bars are the 95% confidence intervals on the average. Statistically significant differences are indicated by *

At CC48, results show increased dissolved zinc concentrations (Figure 4) and loads (Figure 5) during base-flow, rising limb, and falling limb months, and these increases may be attributable to the cessation of active treatment during the second period. The transition from water treatment to bulkheads in the Cement Creek headwaters was a multi-year event, however, and the full effects of these changes may not be evident during the current duration of study (generation of metals and acidity caused by pyrite oxidation is expected to decrease following bulkhead emplacement, but the full effects are not expected to be observed until the system re-equilibrates). Some bulk headed water may re-emerge and enter Cement Creek untreated, as evidenced by the increases in flow from the Mogul and Red and Bonita mines following emplacement of the American Tunnel bulkheads (Walton-Day et al., 2020). Longer-term investigations will help quantify the role of bulkheads given the complex mine workings and fractured-rock groundwater system that underlies the Cement Creek headwaters.

4.2 | Zinc concentration and load at the Animas River below Silverton (A72): decreases in spring and increases during base flow

The stream gage A72 is the farther downstream monitoring location and thus captures the broad-spectrum trends of the combined watershed scale. At A72, trends in zinc concentration and load vary depending on the time of year. Zinc concentration increased significantly from the first period to the second during base-flow and falling limb periods (Figure 4). Trends in zinc concentration were not significant during the rising limb. Zinc loads significantly increased between the two periods only during base-flow (Figure 5). Zinc loads decreased significantly between the two periods during rising limb months and increased significantly during base flow months. Differences between patterns observed in rising and falling limb conditions may be caused to hydrologic differences between the first and second time periods (Supplementary Materials S5, S6).

Results from the Mann-Kendall trend test (Figure 7) similarly show the greatest and most statistically significant increase in zinc load during low streamflow months. The Mann-Kendall trend test indicates increased zinc load through time during 10 of 12 months, although the rising and falling limb months have lower slope values and are less statistically significant than base-flow months. At the larger watershed scale of A72, decreasing zinc loads at M34 and A68 seem to be negated by increasing zinc loads from CC48. During high-flow months, A72 exhibits statistically significant negative slopes in changes in zinc load, perhaps driven by the decreasing zinc loads at M34 and A68. During base-flow months, zinc load trends have positive slopes, reflecting the increased zinc load inputs from CC48.

Because of the unknown time to equilibrium following bulkhead installation, it is possible that the initial bulkhead installation

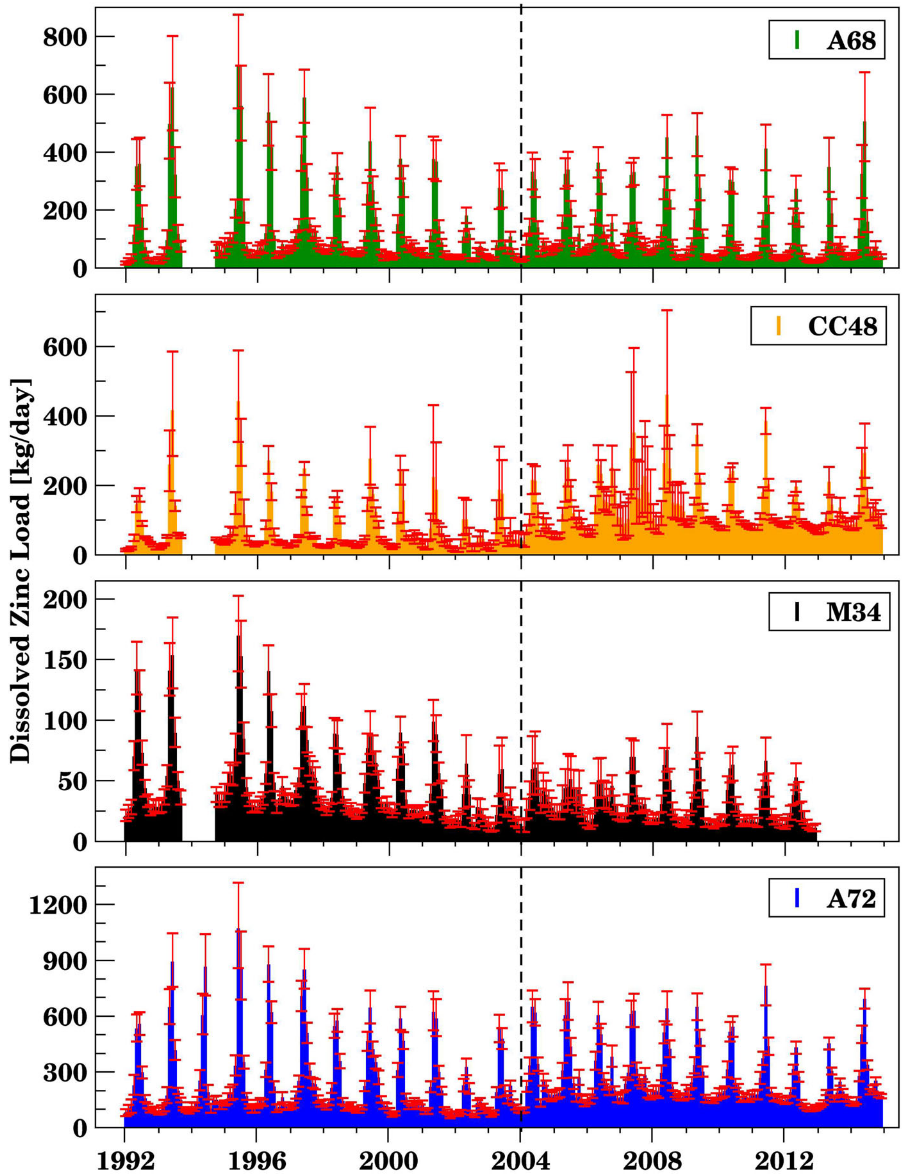

FIGURE 6 Monthly zinc loads estimated using LOADEST from each stream gage. Error bars indicate 95% confidence intervals on load estimates. The increase in base-flow troughs between peaks at CC48 and A72 following the 2004 dashed dividing line

redirected discharges into different areas in the watershed, which would influence loads in sub-watersheds more than the overall net change below the study area. Bulkhead installation dramatically changes local hydrology; following bulkhead closure, the rebound of the water table in surrounding host rocks occurs over long periods as compared with ceasing or starting treatment (Walton-Day & Mills, 2015). A sum of each watershed's contributing loads was calculated to compare net changes between the sum of the three sub-watersheds to the overall observed changes at A72 (Figure 8).

Notably, the highest zinc loads at A72 are all recorded in the first period, indicating that surface reclamation may have decreased peak loading at the watershed outlet (Figure 8). Although the peak zinc loads of 1993 and 1995 are driven in part by the large snowpack of those years, negative trends in flow-weighted concentrations during rising limb months (Figure 7) provide additional evidence that surface remediation may have influenced the decreasing peak annual load at A72.

Interpretation of results is facilitated by the separation of the study period into separate periods (1991–2003; 2004–2014). The binning of periods inevitably incorporates a range of runoff conditions, which potentially complicates interpretation. Although both periods in this study incorporate 1 year of drought (2002 and 2012 respectively) and notable wet years, there are subtle differences between the hydrology in each period (S5 and S6). The implementation of bulkheads in the first period influences subsurface flow paths such that a greater portion of upper Cement Creek source water is moved through the remaining open mines (Walton-Day et al., 2020). As mine workings tend to release water slower than snowmelt runoff, it is possible that hydrologic changes following bulkhead implementation drive elevated base-flow zinc levels and dampened peak streamflow.

Summed zinc loads from stream gages A68, CC48 and M34 are in strong agreement with loads at A72 (Figure 8). Both these summed and estimated loads for A72 indicate that zinc loads increased during

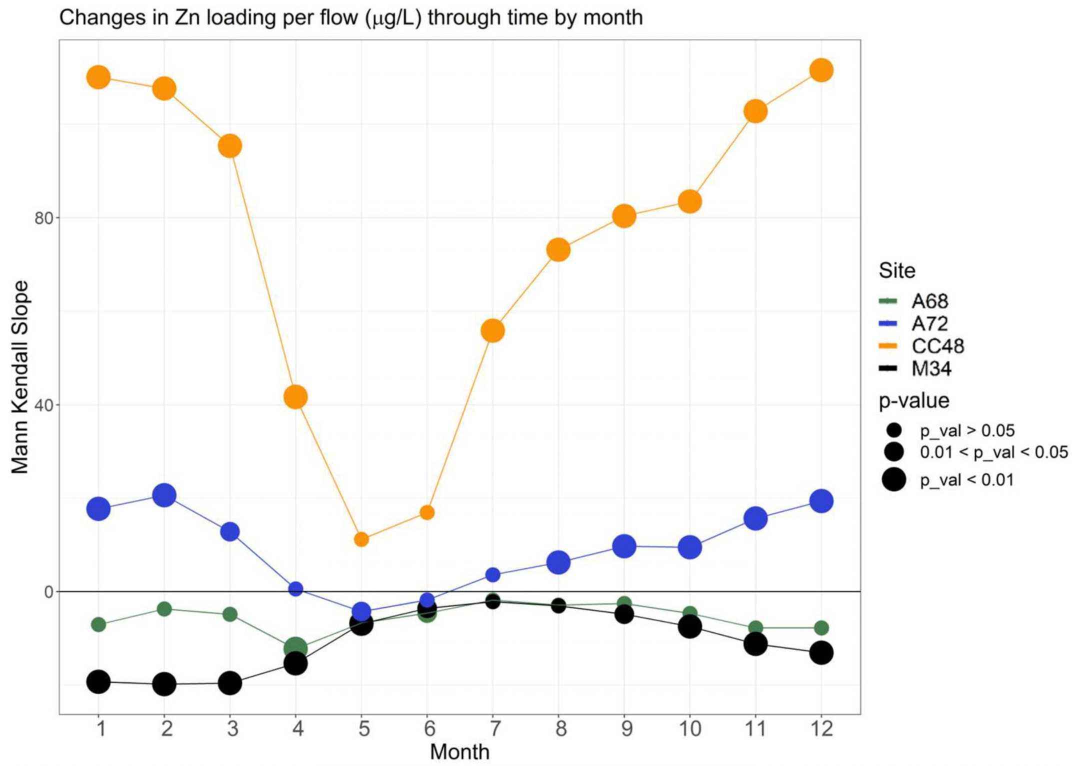

FIGURE 7 Trends in monthly flow weighted zinc concentration (μg/L) through time. Mann-Kendall trend test results for monthly flow weighted concentrations. *y*-axis values are slopes of flow weighted zinc trends through time plotted over month (μg/L/year). Larger points correspond with smaller *p*-values; data below the solid black line indicate negative slopes in zinc load (decreases in load over time) and data above the solid line indicate positive slopes (increases in load over time). Calculated slopes of flow weighted concentration trends over time (*y*-axis) is plotted for each month (*x*-axis)

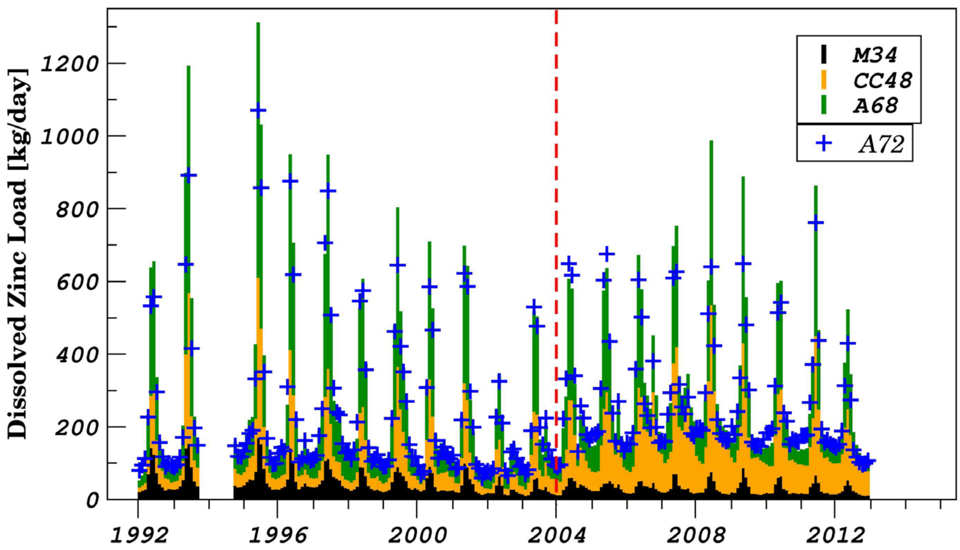

FIGURE 8 Monthly modelled zinc load at all four sites. Loads are plotted stacked on top of one another; blue plus signs indicate load estimates from A72

groundwater-dominated base-flow months and decreased with the combined groundwater and surface water inputs during high streamflow months. The increase in zinc loading at base-flow is notable because these loads are the result of high concentrations, rather than large flow values. High metal concentrations can be taxing for aquatic life.

4.3 | Load estimation method and flow weighted data

Successful application of LOADEST in the metals/mining setting is attributed to several factors. First, daily estimates of streamflow were readily available from the four stream gages in the Silverton area. This contrasts with mining areas in more remote locations, where long-term, continuous, high-quality streamflow data are often lacking. Second, this application benefited from a multi-decadal, multi-agency monitoring effort, and the rigorous analysis of hydrologic variability and remedial actions would not be possible without such a long-term effort. Finally, the nominally conservative behaviour of zinc resulted in a correlation between load and streamflow, which, coupled with frequent zinc concentration measurements, facilitated LOADEST application (Section 3.1).

Interpretation of time-dependent changes in zinc concentration and load at the four USGS stream gages is complicated by the compounding influence of three separate watersheds all of which experience large hydrologic variability spatially and temporally. Superimposed over these signals are numerous remedial activities using multiple approaches (surface reclamation, active treatment and bulkheads), each of which may affect water quality in distinct ways in response to hydrologic variation. Flow-weighted data can reduce hydrologic variability, which clarifies zinc trends more directly tied to remediation efforts. Using concentration-streamflow relations to model periods of missing data and flow weighting results by average streamflow to ameliorate confounding streamflow variability, statistically significant patterns help elucidate the effects of remedial action. These statistical patterns become clearer with flow-weighted data as statistical significance increased (p-values decreased) following flow weighting of data. Flow weighting of data is critical in comparing interannual ARD variations in the presence of hydrologic variability.

4.4 | Concluding remarks

This study presents an effective use of LOADEST to assess changes in a watershed where trends in data are obfuscated by interannual variability. The observed increase in zinc loading from the first time period to the second at CC48 persisted across all months. The spatial scale of this watershed is large, and it is possible that the system had not reached equilibrium in the decade following bulkhead installation. Improvements achieved with water treatment were not maintained following cessation of treatment and installation of bulkheads. Surface remediation upstream of A68 and M34 is correlated with decreased zinc load in some months.

Bulkhead structures internal to mine workings and at mine adits can affect mine water movement at large scales. It is even possible that bulkhead implementation may cause shifts across topographic watershed boundaries because of changes in groundwater flow directions (Walton-Day & Mills, 2015). Additional considerations when interpreting the increased zinc concentrations following bulkhead implementation include: (1) the timescale required to eliminate oxidation in saturated mine workings following bulkhead implementation, and (2) the time lag in observed changes following bulkhead implementation in a subsurface reservoir of this scale.

Although the time window examined in this study likely does not extend to hydrologic equilibrium following bulkhead installation, it provides a unique window into the watershed-scale response following the implementation of remediation techniques. Previous studies (Walton-Day et al., 2021; Walton-Day & Mills, 2015) focus on the stream-scale response, whereas this study attempts to take a larger, watershed-scale approach to the assessment of remedial actions. On the watershed scale, water-quality improvements from surface reclamation yielded decreased ARD during first flush and peak streamflow. Surface reclamation is not expected to improve water quality during low-flow periods, when groundwater is the primary source of streamflow. Watershed scale increases in zinc loading were quantified during base flow, and these increases may be attributable to the cessation of water treatment. Additional remediation techniques targeting base-flow periods may therefore warrant further investigation.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No DGE 1650115. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Additional support was provided by the U.S. Geological Survey's Toxic Substances Hydrology Program. We thank Peter Butler, Ian Bowen, Andrew Todd, Kirstin Brown, Katie Walton-Day, and Alisa Mast for their knowledge, data, and insights that proved invaluable to this study. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

DATA AVAILABILITY STATEMENT

All calibration data used in this study are publicly available from the USGS National Water Information System (U.S. Geological Survey, 2019), the EPA Bonita Peak Mining District Superfund Analytical Results (EPA, 2019) or other cited publications (Church et al., 2007; Walton-Day et al., 2020). LOADEST inputs and outputs are available on the Consortium of Universities for the Advancement of Hydrologic Science Hydroshare platform (Petach et al., 2021).

ORCID

Tanya N. Petach https://orcid.org/0000-0002-4109-1012

REFERENCES

August, E. E., McKnight, D. M., Hrncir, D. C., & Garhart, K. S. (2002). Seasonal variability of metals transport through a wetland impacted by mine drainage in the Rocky Mountains. *Environmental Science & Technology*, *36*(17), 3779–3786. https://doi.org/10.1021/es015629w

Aulenbach, B. T. (2006). Annual dissolved nitrite plus nitrate and total phosphorous loads for the Susquehanna, St. Lawrence, Mississippi-Atchafalaya, and Columbia River basins, 1968–2004. Open-File Report No. 2006–1087. https://doi.org/10.3133/ofr20061087

Aulenbach, B. T., Burns, D. A., Shanley, J. B., Yanai, R. D., Bae, K., Wild, A. D., Yang, Y., & Yi, D. (2016). Approaches to stream solute load estimation for solutes with varying dynamics from five diverse small watersheds. *Ecosphere*, *7*(6), e01298. https://doi.org/10.1002/ecs2.1298

Besser, J. M., & Brumbaugh, W. G. (2007). Status of stream biotic communities in relation to metal exposure. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 823–835. https://doi.org/10.3133/pp1651

Besser, J. M., Finger, S. E., & Church, S. E. (2007). Impacts of historical mining on aquatic ecosystems: An ecological risk assessment. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 87–106. https://doi.org/10.3133/pp1651

Besser, J. M., & Leib, K. J. (2007). Toxicity of metals in water and sediment to aquatic biota. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 839–846. https://doi.org/10.3133/pp1651

Bonita Peak Community Advisory Group. 2016. Gladstone Chronology from Silverton Standard. https://www.bonitapeakcag.org/wp-content/uploads/2019/10/Gladstone-Chronology-from-Silverton-Standard-reformatted.pdf

Bove, D. J., Mast, M. A., Dalton, J. B., Wright, W. G., Yager, D. B., & Church, S. E. (2007). Major styles of mineralization and hydrothermal alteration and related solid-and aqueous-phase geochemical signatures. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 161–230. https://doi.org/10.3133/pp1651

Bove, D. J., Mast, M. A., Wright, W. G., Verplanck, P. L., Meeker, G. P., & Yager, D. B.. 2000. Geologic control on acidic and metal-rich waters in the southeast Red Mountains area, near Silverton, Colorado. In Proceedings of the ICARD 2000 Meeting: Proceedings of the fifth international conference on acid rock drainage, Volume 1: Society for Mining, Metallurgy, and Exploration, pp. 523–535

Brooks, P. D., McKnight, D. M., & Bencala, K. E. (1998). Episodic increases in Zn concentrations associated with spring snowmelt in headwater catchments, Colorado. *Eos Transactions American Geophysical Union*, *79*(45), 545.

Brooks, P. D., McKnight, D. M., & Bencala, K. E. (2001). Annual maxima in Zn concentrations during spring snowmelt in streams impacted by mine drainage. *Environmental Geology*, *40*, 1447–1454. https://doi.org/10.1007/s002540100338

Church, S.E., von Guerard, P., and Finger, S.E., (Eds.) (2007). Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado. U.S. Geological Survey Professional Paper 1651, p. 1096. https://doi.org/10.3133/pp1651

Cohn, T. A. (1988). Adjusted maximum likelihood estimation of the moments of lognormal populations from type 1 censored samples. Open-File Report No. 88–350. US Geological Survey. https://doi.org/10.3133/ofr88350

Cohn, T. A. (1995). Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers. *Reviews of Geophysics*, *33*(S2), 1117–1123. https://doi.org/10.1029/95rg00292

Donato, M. M. (2006). Annual trace-metal load estimates and flow-weighted concentrations of cadmium, lead, and zinc in the Spokane River basin, Idaho and Washington, 1999-2004. U.S. Geological Survey Scientific Investigations Report 2006–5188, p. 38. https://pubs.usgs.gov/sir/2006/5188/

Drake, T. W., Hemingway, J. D., Kurek, M. R., Peucker-Ehrenbrink, B., Brown, K. A., Holmes, R. M., Galy, V., Moura, J. M. S., Mitsuya, M., Wassenaar, L. I., Six, J., & Spencer, R. G. M. (2021). The pulse of the Amazon: Fluxes of dissolved organic carbon, nutrients, and ions from the world's largest river. *Global Biogeochemical Cycles*, *35*, e2020GB006895. https://doi.org/10.1029/2020gb006895

Finger, S. E., Church, S. E., & von Guerard, P. (2007). Potential for successful ecological remediation, restoration, and monitoring. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 1065–1076. https://doi.org/10.3133/pp1651

Ganoe, L. S. (2019). Using a multi-faceted approach to assess ecological components affecting muskrat (Ondatra zibethicus) populations [unpublished master of science thesis], The Pennsylvania State University.

Gobla, M., Gemperline, C., & Stone, L. (2015). Technical evaluation of the gold king mine incident. Reclamation Managing Water in the West.

Hamed, K. H. (2008). Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, *349*(3–4), 350–363. https://doi.org/10.1016/j.jhydrol.2007.11.009

Johnson, D. B., & Hallberg, K. B. (2005). Acid mine drainage remediation options: A review. *Science of the Total Environment*, *338*(1–2), 3–14. https://doi.org/10.1016/j.scitotenv.2004.09.002

Jones, W. R. (2007). History of mining and milling practices and production in San Juan county, Clorado, 1871-1991. Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado: US Geological Survey Professional Paper, Vol. *1651*, pp. 43–86. https://doi.org/10.3133/pp1651

Kendall, M. G. (1948). *Rank correlation methods*. Griffin. https://doi.org/10.2307/2333282

Kimball, B. A., Runkel, R. L., Walton-Day, K., & Bencala, K. E. (2002). Assessment of metal loads in watersheds affected by acid mine drainage by using tracer injection and synoptic sampling: Cement Creek, Colorado, USA. *Applied Geochemistry*, *17*(9), 1183–1207. https://doi.org/10.1016/s0883-2927(02)00017-3

Kimball, B. A., Walton-Day, K., & Runkel, R. L. (2007). Integrated investigations of environmental effects of historical mining in the Animas River watershed, San Juan County, Colorado. US Geological Survey Professional Paper, Vol. *1651*, p. 1096. https://doi.org/10.3133/pp1651

Larison, J. R., Likens, G. E., Fitzpatrick, J. W., & Crock, J. G. (2000). Cadmium toxicity among wildlife in the Colorado Rocky Mountains. *Nature*, *406*(6792), 181–183. https://doi.org/10.1038/35018068

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, *13*, 245–259. https://doi.org/10.2307/1907187

Mast, A. (2018). Estimating metal concentrations with regression analysis and water-quality surrogates at nine sites on the animas and San Juan Rivers, Colorado, New Mexico, and Utah. US Geologic Survey Scientific Investigations Report, Vol. *5116*. https://doi.org/10.3133/sir20185116

McKnight, D. M., & Feder, G. L. (1984). The ecological effect of acid conditions and precipitation of hydrous metal-oxides in a rocky-mountain stream. *Hydrobiologia*, *119*(2), 129–138. https://doi.org/10.1007/BF00011952

National Water Quality Monitoring Council (2019). Water Quality Portal. https://www.waterqualitydata.us

Niyogi, D. K., Lewis, W. M., Jr., & McKnight, D. M. (2002). Effects of stress from mine drainage on diversity, biomass, and function of primary producers in mountain streams. *Ecosystems*, *5*(6), 554–567. https://doi.org/10.1007/s10021-002-0182-9

Nordstrom, D. K. (2009). Acid rock drainage and climate change. *Journal of Geochemical Exploration*, *100*(2–3), 97–104. https://doi.org/10.1016/j.gexplo.2008.08.002

NRCS. (2020). Prism mean annual map: U.S. Department of Agriculture. http://www.wcc.nrcs.usda.gov/climate/prism.html

Paschke, S.S., Kimball, B.A., Runkel, R.L. (2005). Quantification and simulation of metal loading to the upper Animas River, Eureka to Silverton, San Juan County, Colorado, September 1997 and August 1998. US Geological Survey Scientific Investigation, Report, pp. 2005–5054. http://pubs.usgs.gov/sir/2005/5054.

Petach, T., R. Runkel, R. Cowie, D. McKnight (2021). Effects of Hydrologic Variability and Remedial Actions on First Flush and Metal Loading from Streams Draining the Silverton Caldera, 1992–2014 (LOADEST input/output files), HydroShare, http://www.hydroshare.org/resource/892eaec104ed4dfa8a3a41ee71f4b3a2

Riebsame, W. E., Gosnell, H., & Theobald, D. (Eds.). (1997). *Atlas of the New West: Portrait of a changing region*, W W Norton & Co Inc. https://doi.org/10.14714/cp30.664

Rodriguez-Freire, L., Avasarala, S., Ali, A. M. S., Agnew, D., Hoover, J. H., Artyushkova, K., Latta, D. E., Peterson, E. J., Lewis, J., Crossey, L. J., Brearley, A. J., & Cerrato, J. M. (2016). Post gold king mine spill investigation of metal stability in water and sediments of the Animas River watershed. *Environmental Science & Technology*, *50*(21), 11539–11548. https://doi.org/10.1021/acs.est.6b03092

Rossi, C., Oyarzún, J., Pastén, P., Runkel, R. L., Núñez, J., Duhalde, D., Maturana, H., Rojas, E., Arumí, J. L., Castillo, D., & Oyarzún, R. (2021). Assessment of a conservative mixing model for the evaluation of constituent behavior below river confluences, Elqui River basin, Chile. *River Research and Applications*, *37*, 967–978. https://doi.org/10.1002/rra.3823

Runkel, R.L., (2013). *Revisions to LOADEST*, U.S. Geological Survey. https://water.usgs.gov/software/loadest/doc/loadest_update.pdf

Runkel, R. L., Bencala, K. E., Kimball, B. A., Walton-Day, K., & Verplanck, P. L. (2009). A comparison of pre-and post-remediation water quality, Mineral Creek, Colorado. *Hydrological Processes*, *23*(23), 3319–3333. https://doi.org/10.1002/hyp.7427

Runkel, R. L., Crawford, C. G., & Cohn, T. A. (2004). Load estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers. U.S. Geological Survey Techniques and Methods Book Vol. *4*, Chapter A5, p. 69. https://doi.org/10.3133/tm4a5

Runkel, R. L., Kimball, B. A., Nimick, D. A., & Walton-Day, K. (2016). Effects of flow regime on metal concentrations and the attainment of water quality standards in a remediated stream reach, Butte, Montana. *Environmental Science & Technology*, *50*(23), 12641–12649. https://doi.org/10.1021/acs.est.6b03190

Runkel, R. L., Kimball, B. A., Walton-Day, K., Verplanck, P. L., & Broshears, R. E. (2012). Evaluating remedial alternatives for an acid mine drainage stream: A model post audit. *Environmental Science and Technology*, *46*(1), 340–347. https://doi.org/10.1021/es2038504

Schemel, L. E., Kimball, B. A., Runkel, R. L., & Cox, M. H. (2007). Formation of mixed Al-Fe colloidal sorbent and dissolved-colloidal partitioning of cu and Zn in the Cement Creek-Animas River confluence, Silverton, Colorado. *Applied Geochemistry*, *22*(7), 1467–1484.

Shaw, M., Yazbek, L., Singer, D., & Herndon, E. (2020). Seasonal mixing from intermittent flow drives concentration-discharge behaviour in a stream affected by coal mine drainage. *Hydrological Processes*, *34*(17), 3669–3682. https://doi.org/10.1002/hyp.13822

Shrestha, S., Gunawardana, S. K., Piman, T., & Babel, M. S. (2020). Assessment of the impact of climate change and mining activities on streamflow and selected metal's loading in the Chindwin River, Myanmar. *Environmental Research*, *181*, 108942. https://doi.org/10.1016/j.envres.2019.108942

Singer, P. C., & Stumm, W. (1970). Acidic mine drainage: The rate-determining step. *Science*, *167*(3921), 1121–1123. https://doi.org/10.1126/science.167.3921.1121

Skousen, J., Rose, A., Geidel, G., Foreman, J., Evans, R., & Hellier, W. (1998). *Handbook of technologies for avoidance and remediation of acid mine drainage* (Vol. 131). National Mine Land Reclamation Center.

Todd, A. S., McKnight, D. M., Jaros, C. L., & Marchitto, T. M. (2006). Effects of acid rock drainage on stocked rainbow trout (*Oncorhynchus mykiss*): An in-situ, caged fish experiment. *Environmental Monitoring Assessment*, *131*, 111–127. https://doi.org/10.1007/s10661-006-9382-7

U.S. Environmental Protection Agency. (1997). *EPA's National Hardrock Mining Framework*. U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2019. Bonita Peak Analytical Chemistry Results. Bonita Peak Superfund Site. https://epa.maps.arcgis.com/apps/MapJournal/index.html?appid=9ec10c38c47b435fb8f4bbe02b02b65d

U.S. Geological Survey. 2019. National Water Information System data available on the world wide web. USGS Water Data for the Nation. http://waterdata.usgs.gov/nwis/. https://doi.org/10.5066/F7P55KJN

Walton-Day, K. (2003). Geochemistry of active and passive treatment processes used to treat mine drainage. *Environmental Aspects of Mine Wastes*, *4*, 335–359. https://doi.org/10.1144/1467-7873/03-029

Walton-Day, K., Mast, M. A., & Runkel, R. L. (2021). Water-quality change following remediation using structural bulkheads in abandoned draining mines, upper Arkansas River and upper Animas River, Colorado USA. *Applied Geochemistry*, *127*, 104872. https://doi.org/10.1016/j.apgeochem.2021.104872

Walton-Day, K., & Mills, T. J. (2015). Hydrogeochemical effects of a bulkhead in the Dinero mine tunnel, sugar loaf mining district, near Leadville, Colorado. *Applied Geochemistry*, *62*, 61–74. https://doi.org/10.1016/j.apgeochem.2015.03.002

Walton-Day, K., Runkel, R.L., Mast, M.A, and Qi, S.L. (2020). Water-quality and discharge data from draining mine tunnels near Silverton, Colorado, 1988-2015: U.S. Geological Survey data release. https://doi.org/10.5066/P9FE667O

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, *29*(3/4), 350–362. https://doi.org/10.2307/2332010

Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, *34*(1/2), 28–35. https://doi.org/10.2307/2332510

Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, *38*(3/4), 330–336. https://doi.org/10.2307/2332579

Wireman, M., & Stover, B. (2011). Hard rock mining and water resources. *Groundwater*, *49*(3), 310–316. https://doi.org/10.1111/j.1745-6584.2011.00818.x

Wren, C. D. (1984). Distribution of metals in tissues of beaver, raccoon and otter from Ontario, Canada. *Science of the Total Environment*, *34*(1–2), 177–184. https://doi.org/10.1016/0048-9697(84)90049-4

Yochum, S. E. (2000). A revised load estimation procedure for the Susquehanna, Potomac, Patuxent, and Choptank Rivers. U.S. Geological Survey Water-Resources Investigations Report 2000–4265, Chapter iv, p. 49. https://doi.org/10.3133/wri004156

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Petach, T. N., Runkel, R. L., Cowie, R. M., & McKnight, D. M. (2021). Effects of hydrologic variability and remedial actions on first flush and metal loading from streams draining the Silverton caldera, 1992–2014. *Hydrological Processes*, *35*(11), e14412. https://doi.org/10.1002/hyp.14412