

Field estimate of paleoseismic slip on a normal fault using the Schmidt hammer and terrestrial lidar: methods and application to the Hebgen fault (Montana, USA)

Alexander Tye¹ and Timothy Stahl²

¹Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI 48104

²Department of Geological Sciences, University of Canterbury, Christchurch, New Zealand

Abstract

The weathering characteristics of bedrock fault scarps provide relative age constraints that can be used to determine fault displacements. Here, we report Schmidt hammer rebound values (R-values) for a limestone fault scarp that was last exposed in the 1959 M_w 7.3 Hebgen Lake, Montana earthquake. Results show that some R-value indices, related to the difference between minimum and maximum R-values in repeated impacts at a point, increase upward along the scarp, which we propose is due to progressive exposure of the scarp in earthquakes. An objective method is developed for fitting slip histories to the Schmidt hammer data and produces the best model fit (using the Bayesian Information Criterion) of 3 earthquakes with single event displacements of ≥ 1.20 m, 3.75 m, and c. 4.80 m. The same fitting method is also applied to new terrestrial lidar data of the scarp, though the lidar results may be more

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1002/esp.4403](https://doi.org/10.1002/esp.4403)

influenced by macro-scale structure of the outcrop than by differential weathering. We suggest the use of this fitting procedure to define single event displacements on other bedrock fault scarps using other dating techniques. Our preliminary findings demonstrate that the Schmidt hammer, combined with other methods, may provide useful constraints on single event displacements on exposed bedrock fault scarps.

Introduction

Exposed rock surfaces weather via physical and chemical processes that change rock mass and surface properties through time (c.f. Birkeland and Noller, 2000). Depending on the lithology and climate conditions, this weathering usually degrades the mechanical integrity of the rock and increases surface roughness through time (Benedict, 1985; Crook and Gillespie, 1986; Maizels, 1989; McCarroll, 1991; Stewart, 1996). Relative exposure age dating techniques were developed on the premise that these weathering-related changes can be quantified and used to compare the ages of rocks and landforms, and are usually supported empirically via calibration on surfaces of known age (Colman and Dethier, 1986, and references therein). Relative and calibrated exposure dating techniques applied to rock surfaces have been used to date moraines (Matthews and Shakesby, 1984; Winkler, 2005), rock glaciers (Laustela et al., 2003; Aa et al., 2007), snow avalanche ramparts (Matthews et al., 2015), bedrock surfaces (Gupta et al., 2009; Matthews and Owen, 2010), fluvial terraces (Stahl et al., 2013), patterned ground (Winkler et al., 2016), and erratic boulder trains (Darvill et al., 2015), among others. They have also been used in the field of paleoseismology to date coseismic rock avalanche deposits (e.g., Bull, 1996 et al., 1994), correlate deformed river terraces (Stahl et al., 2016), and identify displacement patterns on bedrock normal fault scarps (Stewart, 1996; Tucker et al., 2011; He et al., 2016).

Bedrock scarps can be progressively exposed in repeated earthquakes along a normal-slip fault (Bosi et al, 1993). If the scarp face is preserved over several earthquake cycles, as is commonly the case in indurated limestones, several episodes of displacement may be preserved along the face. Previous researchers have performed terrestrial cosmogenic nuclide dating (TCND) transects along scarp faces, using ^{36}Cl in limestone, to date the timing of earthquakes and demarcate single event displacements (SEDs) that scale with earthquake magnitudes (e.g., Wells and Coppersmith, 1994). Some have observed that the different “slip patches” have distinguishing weathering characteristics that can be used to delineate SEDs independently of absolute ages (Stewart, 1996), or that could assist in choosing sample locations for TCND and/or the final interpretation of ages (Zreda and Noller, 1998). Terrestrial lidar scanning (TLS) has been used to investigate the microtopographic expression of these slip patches (Giaccio et al., 2003; Wei et al., 2013; Wiatr et al., 2015; He et al., 2016).

Here, we test the applicability of using TLS and the Schmidt hammer to quantitatively distinguish weathering zones and slip patches on the Hebgen fault in Montana, USA. The interpretation of this scarp as being related to progressive exposure in earthquakes has been questioned due to discrepant paleoseismic histories in nearby records (e.g., Pierce et al., 2000). Therefore, we propose that our approach could be used to clarify the exposure history of this scarp by assessing variations in time-dependent rock weathering properties along the scarp face and estimate the SEDs of past earthquakes. In our analysis, we use curve-fitting techniques

that can be used to assess the appropriate form of the weathering index versus scarp height curve, which could be used to objectively identify SEDs using other relative-age dating techniques.

Background

Significance of single-event displacements in paleoseismology

Determining the surface displacements of past earthquakes on faults is an essential component of paleoseismology. Fault SEDs scale with the magnitude of an earthquake (e.g., Wells and Coppersmith, 1994), are used to define fault avoidance zones in seismic hazard assessments (e.g. Boncio et al., 2012; Villamor et al., 2012), and can be used to estimate average earthquake recurrence intervals if long-term fault slip rates are available. Paleoseismic trenching can be used to determine SEDs where a fault has ruptured through unconsolidated sediments, but is not typically feasible across bedrock scarps. TCND along normal-slip bedrock scarps can be used to directly date and measure displacements of multiple paleo-earthquakes (Benedetti et al., 2002; Benedetti and van der Woerd, 2014; Carcaillet et al., 2008; Mitchell et al., 2001; Zreda and Noller, 1998). Accurate reconstructions of slip histories in these studies usually rely on several closely-spaced samples from the scarp face, numerical modeling (e.g., Schlagenhauf et al. 2010), and/or secondary scarp weathering evidence to corroborate the interpretation of modeled ages.

Geomorphic framework for relative dating of SEDs

Surface-rupturing normal faults accumulate displacement episodically in earthquakes or via continuous creep. In the absence of near-surface creep and significant erosion, bedrock fault scarps can record several meters of cumulative displacement over many earthquake cycles. Under ideal circumstances, each earthquake brings a previously unexposed, relatively unweathered, slip patch up to the surface and raises the previously exposed patches from older earthquakes higher along the scarp. Thus, from the bottom to the top of a normal fault scarp (i.e., moving up-dip), there should be progressively more weathered slip patches. This framework assumes that (i) that fault colluvium does not bury the newly exposed slip patch inhibiting subaerial weathering, and (ii) no external processes, besides weathering, significantly alter the scarp after it is exposed (e.g., stream modification of the base of the scarp, removal of rock mass via slope processes, biological weathering due to lichen or vegetation cover, remineralization of calcite lower on limestone scarps) (e.g., Kastelic et al., 2017). These assumptions are not always valid and limit the applicability of all scarp dating techniques at some sites.

Numerous studies have outlined methods for identifying SEDs independently of cosmogenic exposure ages on bedrock scarps and typically focus on differing degrees of rock weathering. In the first TCND study of a bedrock scarp, Zreda and Noller (1998) used qualitative metrics of rock weathering (e.g., roughness, preservation of slickensides, pitting, and discloration) as

corroborative evidence of their preferred slip history. Measurements of surface roughness using a micro-roughness meter (Stewart, 1996) and TLS (Giaccio et al., 2003; Wei et al., 2013; Wiatr et al., 2015), pit depth measurements (Tucker et al., 2011), image analysis (Giaccio et al., 2003), and geochemical signature (Carcaillet et al., 2008; Manighetti et al., 2010; Mouslopoulo et al., 2011) have yielded promising results that can assist in interpreting scarp exposure and weathering histories. Other relative dating techniques that measure rock weathering grade are useful in reconnaissance and assessment of bedrock fault scarps (Table 1).

The Schmidt hammer in rock weathering and relative age-dating studies

The Schmidt hammer tests rock hardness by measuring the rebound of a spring-loaded piston in a controlled impact against the rock surface, yielding a rebound value (R-value) (Day, 1980). The technique has been used in geomorphology for over 40 years (Day and Goudie, 1977; Goudie, 2006) and has been applied to rapidly assess rock strength in the field and as a relative exposure age dating tool to distinguish landforms of different ages (Goudie et al., 2006; Shakesby et al., 2011). The latter operates on the assumption that rock surface compressional strength and/or mechanical integrity decreases with time. Calibration with rock surfaces of known age allows this assumption to be tested and in some cases, enables researchers to establish chronofunctions that can be used to date rock surfaces or landforms over 10^2 - 10^4 year timescales (e.g., Stahl et al., 2013).

In order to use the Schmidt hammer as a relative dating tool, controls must be emplaced to ensure that exposure age is the dominant, if not the only, variable that affects R-values.

Previous researchers have investigated the relationship between R-value and time-independent variables that affect R-values like lithology (e.g., Goudie, 2006; T **örök et al., 2007), sa**

dimensions (Sumner and Nel, 2002; Aydin 2009; Demirdag et al., 2009), number of samples (Niedzielski et al., 2009), operator bias (Shakesby et al., 2006), instrument degradation (McCarroll 1987), biological weathering (Matthews and Owen, 2008), moisture content (Sumner and Nel, 2002), and spatially-variable weathering rates (Stahl et al., 2013). If these variables are known and adequately controlled, the Schmidt hammer can be used to distinguish rock surfaces with different exposure ages (Shakesby et al., 2006; Shakesby et al., 2011; Stahl et al., 2013).

The difference between Schmidt hammer R-values in successive impacts at the same location increases with increasing weathering grade (Aydin and Basu, 2005; Nicholson, 2009; Matthews et al., 2016). Using this differencing approach allows detection of and control for small-scale lithologic, rock mass, or surface roughness variations that lead to time-independent variability in single impact R-values. This is partially due to a decrease in surface roughness after the first impact, and subsequent pulverization of the outer weathered zone, leading to higher R-values in successive impacts. Matthews et al. (2016) showed that this effect becomes insignificant after the fifth impact in a range of metamorphic and igneous lithologies.

Identification of bedrock fault scarp slip patches using microtopography

Microtopography data acquired using TLS have been combined with proxies for surface roughness to investigate fault processes (e.g. Sagy et al., 2007, Candela et al., 2009) and to investigate the weathering of bedrock fault scarps (Giaccio et al., 2003; Wei et al., 2013; Wiatr et al., 2015; He et al., 2016). Most of these approaches use fractal parameters as roughness proxies. Fractal parameter-based methods exploit the fact that on natural surfaces, the magnitude of relief is related to the interval over which it is measured by a power law (e.g. Ahnert, 1984). The exponent of this power law reflects how relief changes with scale. For instance, He et al. (2016) used the best-fit value of this exponent over the surface of a bedrock fault scarp to identify slip patches on limestone fault scarps. Wei et al. (2013) use fractal-based methods to show that on limestone fault scarps, weathering effects are typically expressed on scales of several centimeters and less.

Study site

The ideal study site for testing Schmidt hammer and TLS approaches should have (i) some evidence of multiple slip events exposed along a bedrock fault scarp, (ii) homogenous lithology along the scarp, and (iii) good scarp preservation with subaerial weathering being the predominant mode of post-exposure alteration. We selected a site along an exposed section of the Hebgen fault on the northern shore of Hebgen Lake in Montana (Fig. 1). The lake has an elevation of c. 2000 m above seal level. The region is cold and dry, with mean annual temperature of 2.8° C and mean annual precipitation of c. 500 mm. Mean monthly

temperatures range from -8 to 16° C; monthly precipitation is relatively constant at between 30-50 mm. The area of Hebgen Lake was periglacial during the last glacial maximum (LGM, ~15-18 ka BP Pinedale Glaciation) but not occupied by ice and may have been even drier than present day climate (Licciardi and Pierce, 2008). Terminal moraines and glacial till from the penultimate Bull Lake glaciation (~130-160 ka BP) indicate that the valley was occupied by ice in earlier Pleistocene glaciations (Licciardi and Pierce, 2008; Pierce, 2003).

Hebgen Lake is manmade, created by damming the Madison River in 1914, and is located within the intermontane Madison Valley. The region is currently undergoing NE-SW directed Basin and Range extension and is also at the western flank of Yellowstone hotspot-related deformation (Wicks et al., 2006). The 1959 M_w 7.3 Hebgen Lake earthquake produced surface rupture along the Hebgen and Red Canyon faults (Witkind et al., 1962) which led to NE-directed subsidence, lake-level rise on the northern shore of Hebgen Lake, and subsequent inundation of the old route US-287 (Fig. 1). The new road was established ~20 m higher, at the base of the c. 250 m-long outcrop considered in this study.

The limestone that comprises the scarp face is mapped as middle Cambrian Meagher Formation (Zreda and Noller, 1998; O'Neill and Christiansen, 2004). It is finely crystalline and appears massively bedded at the transect site though it is commonly thinly to medium bedded and oolitic in other areas (O'Neill and Christiansen, 2004). Fresh faces of the limestone appear grayish tan; weathered faces, in general, are light to dark gray, with some darker staining and

pitting evident on some faces. Meter-scale joints are present at some locations along the cliff exposure. No obvious variability in limestone sedimentology was observed along preserved sections of the scarp and no fault breccia or gouge was observed.

Our transect coincides with the study site of Zreda and Noller (1998), who used in-situ cosmogenic ^{36}Cl to obtain exposure ages of sections of the limestone scarp and model earthquake ages. The remainder of the roadside outcrop is unsuitable for absolute- or relative-age dating as the limestone scarp above the 1959 slip patch has been subject to rockfall and is therefore discontinuous up-dip. Zreda and Noller (1998) interpreted the lowest portion of the scarp at our site to have been exposed in 1959 M_w 7.3 Hebgen Lake earthquake, which caused up to c. 6 m vertical displacement along the fault. They inferred a total of six earthquakes at this site since 24 ka (four since 7 ka). Their data show that SEDs vary from c. 1-2 m per event, with 2.1 m of slip at the study site in the 1959 earthquake. Nearby trenching studies and mapping of Hecker et al. (2000; 2002) and Pierce et al. (2000) revealed similar >1.2-3.1 m SEDs, but found evidence of only one Holocene earthquake prior to the 1959 event (c. 1-3 ka) and one latest Pleistocene earthquake (c. 10-15 ka) (Schwartz et al., 2009). The nearest trench to the study site revealed a cumulative 2-3 m vertical displacement in the 1959 and penultimate events, and 5-6 m cumulative vertical displacement over the last three events (Schwartz et al., 2009). The discord between slip histories in TCND and paleoseismic trenching have led some to question the interpretation of event ages from the cosmogenic transect (e.g. Pierce et al. 2000) or even

that the scarp was rejuvenated in the 1959 earthquake. Regarding the latter question, however, Johnson (2017) has shown with airborne lidar that this trace did indeed rupture in the 1959 earthquake, making Zreda and Noller's (1998) interpretation of this site as a bedrock fault scarp sound.

Methods

Schmidt hammer framework and field protocol

We used a Proceq mechanical N-type Schmidt hammer with an impact energy of 2.207 N m. For each sample location spaced 0.25 m along the direction of slip on the scarp face (Fig. S1), we recorded both the first impact R-value and the subsequent 4 values at the same point of impact. This sample spacing is considered reasonable given the fault's dip-slip motion and all previously reported values of SEDs along the fault (Zreda and Noller, 1998; Hecker et al., 2000, 2002; Pierce et al., 2000; Hecker et al., 2009; Schwartz et al., 2009; Johnson, 2017), though centimeter-scale displacements in smaller surface rupturing earthquakes would be unlikely to be detected at this resolution.

The approach of collecting five impacts at the same location allowed us to assess a 'time-series' of R-values and examine various R-value metrics: first-impact (R_1), R-value range (ΔR), *I5 index (after Matthews et al., 2016), and *Imax index (this study). The latter three are defined below:

$$\Delta R = R_{\max} - R_{\min} \quad (\text{Eq. 1})$$

$$*I5 = 100 * (R_{u5} - R_1) / R_{u5} \quad (\text{Eq. 2; Matthews et al., 2016})$$

$$*I_{\max} = 100 * (R_{u_{\max}} - R_{\min}) / R_{u_{\max}} \quad (\text{Eq. 3})$$

where R_{\max} and R_{\min} are the highest and lowest R-values of the five successive impacts at a single location, respectively, R_{u_5} is the average 5th impact R-value of unweathered rock, and R_1 is the first impact at a single location, and $R_{u_{\max}}$ is the average of maximum R-values on unweathered rock. These unweathered rock R-values come from a single impact site for each distinct rock type and are assumed to be uniform for a given rock type.

For application to fault scarps, the area of scarp exposed in the most recent earthquake should be the least weathered, and thus yield the smallest values for all indices (Eq. 1 -3) (Fig. 2). In the absence of other modifying geomorphic processes (e.g., scarp renewal by landsliding or by stream abrasion) and complicating variables (non-uniform lithology or non-uniform weathering processes), R-value indices should increase in discrete steps along the scarp with increasing weathering grade (Fig. 2). Additionally, we propose that older slip patches should display larger variance in R-value indices (i.e., heteroscedasticity) due to small heterogeneities in weathering rates on older surfaces (Shakesby et al., 2011) (Fig. 2B).

During sampling, first return R-values were recorded if (i) the rock mass did not move or chip while sampling and (ii) the sound was resonant rather than hollow, the latter of which is indicative of a shallow discontinuity (e.g., Stahl et al., 2013). If these criteria were met, four additional R-values were recorded at the same location. Lichen were avoided and all

measurements were taken in dry conditions over the course of one day by a single operator to avoid operator bias. Sample locations were not abraded or otherwise prepared prior to impacts so as to preserve the original surface roughness. Sampling was limited to a single transect where the scarp was continuous along-dip; other transects along the outcrop were not feasible because slope failures have made identification of the original, continuous scarp face unreliable. Sample vertical locations were adjusted for fault dip (65°). R-values were not corrected for angle of incidence as all tests were conducted in the same orientation. The Schmidt hammer was calibrated on a test anvil before and after sampling to ensure no instrument degradation.

We also defined and recorded an Adjusted Geologic Strength Index (AGSI) (e.g., Marinos et al. 2005) to semi-quantitatively describe variations in surface weathering and ensure no major variability in the structural integrity of the limestone scarp. In doing so, we redefined some criteria from the original GSI 'surface quality' scale specifically for quantifying the surface weathering of limestone fault scarps: more weathered surfaces are rougher, more pitted, more stained/discolored, and have degraded and poorly preserved slickensides (Stewart, 1996; Zreda and Noller, 1998).

Terrestrial lidar scanning (TLS) and roughness analysis

To supplement our Schmidt hammer data, we acquired TLS microtopography on a portion of the bedrock fault scarp at Hebgen Lake at a resolution of 1-2 mm using a Riegl VZ-2000. The

scan area coincides with the scarp above the 1959 slip patch. Plants and artifacts were removed from the dataset manually. Coordinates were transformed such that one coordinate dimension represented the distance from the measured fault surface to the best-fit plane for the entire scarp. As a measure of surface roughness at the scale of interest, we calculated the variance of this fault plane normal distance in a square moving window of size ~ 20 cm x 20 cm. The mechanics of our moving window procedure are identical to those in He et al. (2016) though the method of analysis of the data within each window is different than their fractal-based method. Given the restricted scale over which weathering effects are topographically expressed (Wei et al., 2013), the topographic variance within a 20 cm window should capture the scale over which the desired signal should exist. Wiatr et al. (2015) use a similar window size.

Curve-fitting and SED determination

A determination of event-by-event SEDs depends on (i) assessing whether R-value indices and surface roughness are adequately characterized by discrete steps along the scarp (e.g., Fig. 2) and (ii) identifying the best-fit locations of those steps along the scarp. If (i) is assumed, then (ii) needs to be conducted in such a way that avoids bias from user input. We developed a maximum likelihood approach that determines the number of steps (i.e., number of earthquakes) and their locations (SEDs) for the observed data. For between $n=1$ and $n=6$ steps (full range of possible steps given previously reported earthquake histories), a model fit is generated for every possible location of the step(s) along the scarp. The mean (μ) and standard

deviation (σ) of observed indices are calculated over all of those steps and the output is the model that maximizes the likelihood function for the index relative to μ and σ . SEDs are then determined by the width of the steps along the x-axis. Maximum likelihood models were calculated for different functions using the natural logarithm of the R-value index, since a natural logarithmic transformation should reduce the heteroscedasticity inherent in the indices (Fig. 2). All errors were assumed to be normally-distributed on the natural logarithmic scale.

The same fitting method was also conducted on the TLS-derived surface roughness using microtopographic variance. We use the natural log transform of variance because it eliminates heteroscedasticity and is consistent with previous approaches to the same data (e.g. Sagy et al., 2007; Candela et al., 2009; Wei et al., 2013; He et al., 2016). Our moving window approach results in a two dimensional grid of topographic variance measurements that covers the measured portion of the fault scarp. To analyze the topographic variance as a function of scarp height, we use the mean log variance of each row of the grid (rows are perpendicular to fault dip). In addition, we use the standard deviation of each row as an independent estimate of uncertainty.

Model selection

If a stepwise fit is not assumed, then the model selection method must be able to accommodate the different functional forms of all candidate models. One such method is the

Bayesian Information Criterion (BIC; Schwarz, 1978), which quantifies the goodness-of-fit with penalties imposed for additional model parameters:

$$BIC = \ln(n) k - 2 \ln(\hat{L}) \quad \text{Equation 4}$$

where k =number of free parameters to be estimated, n =sample size, and \hat{L} is the maximized value of the log likelihood function of the model. Lower BIC values indicate more preferable models. While the stepwise model fits our and others' geologic interpretation of the site as a scarp exposed in episodic earthquakes, we used the BIC to determine if this model is statistically superior to linear and power law fits, and to determine the optimal number of steps for the stepwise model (Supplemental Information).

For calculating the BIC, linear and power law models have three parameters each: two corresponding to the function parameters (slope and intercept, and constant and exponent, respectively) and one corresponding to the standard deviation of the model residuals (Burnham and Anderson, 2004). Stepwise models have two parameters (for no steps, y =constant), plus two more for each step. These parameters are the constant y value of each step (e.g., R-value metric or topographic variance), the x values of the break points between steps, and the standard deviation of the model residuals. Linear and power law maximum likelihood parameters were determined by pattern search (Hooke and Jeeves, 1961), whereas stepwise maximum likelihood models were determined using brute force.

Results

Adjusted Geologic Strength Index

AGSI values generally decrease with increasing scarp height (Fig. 3) and permit discrimination of three loosely-defined weathering zones over the c. 10 m scarp height. The limestone was observed to have only small variations in structural integrity along the scarp, ranging from *Blocky* to *Intact/Massive* (Fig. 3D). Larger variations were observed in surface quality of the limestone (top axis, Fig. 3D). Zone I (AGSI=73-83, c. 0-1.25 m vertical) corresponds to the 1959 slip surface, which has well-preserved slickensides on a smooth, planar and unstained surface (Fig. 3C). Zone II (AGSI=60-75; c. 1.25-5 m vertical) is characterized by a slight decrease in structural integrity (*Very Blocky* to *Blocky*) but marked decrease in smoothness, slickenside preservation, and increase in discoloration/staining (Fig. 3A). Zone III (AGSI=65-75; c. 5-10 m vertical) has the highest structural integrity of the three zones (*Intact/Massive*) but is again characterized by a decrease in smoothness and slickenside preservation (Fig. 3B) and an increase in discoloration and pitting (Fig. 3A) from Zone II. We demarcated one additional zone Zone IV (AGSI=30-60; >10 m vertical) above the transect that is marked by a decrease in both structural integrity and surface quality, but it is not clear that this zone is a continuation of the planar fault scarp from below (van der Woerd et al., 2000)(Fig. 3A and D) and is thus not included in further analyses.

Schmidt hammer R-values

R-value minima (1st or 2nd impact in ~90% of sample locations, Fig. 4) show an overall decreasing trend with increasing scarp height, while R-value maxima (4th or 5th impact in ~80% of sample locations, Fig. 4) show less variability with scarp height. From one sample location to the next (i.e., every 0.25 m along the scarp) there can be significant variability in R-values. These variations typically affect all of the impacts at a given location, causing some alignment of 'peaks' and 'troughs' in the different impact series (Fig. 4). Thus, ΔR remains relatively unchanged over short length scales despite the large fluctuations in individual readings. R-value range (ΔR) and $*I$ indices (Eqn. 2, 3) increase with scarp height (Fig. 5). For our application, the ΔR metric is preferred over the $*I_5$ and $*I_{max}$. The $*I$ metrics (Matthews et al., 2016 and this study) rely on an R-value measurement of unweathered rock, which is then assumed to be constant for that rock type. The entire Hebgen Lake fault scarp is limestone, and lithologic variations that might be expected to affect R-values were not observed during visits to the site. Thus, the use of $*I$ metrics would assume a single, characteristic unweathered R-value for the entire dataset, which conflicts with our data: R-values for repeated impacts at different sites on the 1959 slip patch (Zone I in Figs. 3, 4) converge on values that differ by up to ~10, more than a third of the total spread in R-values observed across our dataset. This difference suggests that, at the Hebgen Lake site, non-weathering factors besides general rock type have a significant effect on the R-values of unweathered rock, invalidating the assumption

of *I metrics. In contrast, the ΔR values for these impact sites are remarkably consistent (Figs. 4, 5). As such, we conduct further analysis only on ΔR .

Model comparison of ΔR

The maximum likelihood models for linear, power law, and stepwise relationships between ΔR and scarp height are shown in Figure 6. The values of BIC indicate that a stepwise model with two break points is preferred, corresponding to three distinct patches of ΔR moving up the scarp (Table 2, Fig. 6). It can be seen that the support from BIC for the stepwise model with two breaks over the linear model and stepwise models with more breaks is not overwhelming; however, under the classification scheme of Kass and Raftery (1995), the preference of the two-step model over the linear and three-step models should be considered 'positive', and is 'very strong' when compared to power law and one-step models (Table 2).

TLS-derived surface roughness

The TLS data above the 1959 slip patch assist in interpreting the AGSI values and Schmidt hammer data in that part of the transect. The surface roughness (i.e. topographic variance) decreases as scarp height increases (Fig. 7). It can be seen that there is significant scatter to the topographic variance data. High and low variance values are not randomly distributed on the fault scarp, but rather form numerous clusters (Fig. 7B). Natural log transformed topographic variance values can be seen to be homoscedastic (Fig. 7C). Fits of log transformed topographic variance versus scarp height were conducted on a subset of the total scan, using the same set of functional forms as for R-value range. BIC values indicate a slight preference for a stepwise

model with one break over an exponential function (Table 3, Fig. 7C shows the one break step function and the power law function).

Discussion

Use of ΔR as preferred Schmidt hammer metric

On length scales greater than 0.25 m, the change in ΔR is predominantly controlled by decreasing R_{\min} (Fig. 4). Values of R_{\max} remain approximately the same (60 ± 3 , 1σ) over the scarp and are interpreted to reflect the underlying structural integrity of the rock mass. It is thought that R_{\min} varies with surface conditions, like ~ 1 cm scale surface roughness, (top row of AGSI, Fig. 3) and R_{\max} remains comparatively stable with structural integrity (left column of AGSI, Fig. 3). The small drop in R_{\max} across weathering Zone II compared to Zone III (Fig. 4) is reflected by the slight decrease in AGSI structural integrity of Zone II (Fig. 3). Because of the covariance of R_{\max} and R_{\min} values (Fig. 4), ΔR is a more stable metric for weathering duration than a single-impact R-value. The reduced noise of ΔR relative to single impact R-values is important for model selection. Conducting multiple impacts at a single location also yields important information on the underlying conditions of seemingly homogeneous bedrock not obtained by taking a single impact at each location.

There is much greater scatter in both *I metrics than in ΔR (Fig. 5). This can be attributed to two separate causes. For the *I5 value, the fifth impact does not necessarily yield the highest R-value in our tests, and therefore it is not the best indicator of unweathered rock. This contrasts

with the findings of Matthews et al. (2016) and could be due to their use in relatively competent rock compared to limestone. $R_{u_{max}}$ might be considered to be a better indicator of unweathered rock in limestone, and therefore $R_{u_{max}}$ should yield a more useful relationship with exposure age. However, as discussed above, the range of $R_{u_{max}}$ values within the 1959 slip patch, from which $R_{u_{max}}$ is calculated, vary significantly. This could be due to non-weathering related variables within the rock and make ΔR are more suitable measure of weathering degree.

Apparent steps and SEDs of the Hebgen fault

A step pattern of ΔR is expected on a fault scarp that has been progressively exposed in surface rupturing earthquakes (Fig. 2). The width of the steps (change in scarp height, Δx) should define the SEDs in these earthquakes. Our preferred model for the Hebgen fault distinguishes the area of scarp exposed in the 1959 Hebgen Lake earthquake and two other slip patches (Fig. 6). The 1959 slip patch extends from 0.00–1.25 m vertical; a second section extends from c. 1.25–5.00 m; and a third section is interpreted from c. 5.00–9.80 m. We note that the vertical displacement for the 1959 patch is a minimum because our transect did not extend to the base of the 1959 scarp (which was observed along strike). If interpreted as vertical displacements, these ranges translate to >1.25, 3.75, and 4.80 m vertical slip in past earthquakes. It is unknown if the length of the final step (4.8 m) represents a larger SED or the failure of the Schmidt hammer to discriminate two or more older earthquakes in this age range due to the high

variance (Figs. 5, 6). In any case, these displacements are considered reasonable given the maximum displacements reported along all faults in the 1959 earthquake (c. 3-6 m; Witkind et al., 1962; Schwartz et al., 2009) and the range reported in trenching investigations along the fault (1-3 m).

Like the Schmidt hammer results, TLS-derived topographic variance results from TLS above the 1959 slip patch are best fit with stepwise function (Fig. 7). It should be noted that the stepwise relationship is only barely preferred over a power law relationship according to the BIC values (Table 3), which is due to the high scatter. Some warping is observed between the reliable scarp height values of the TLS data and the height estimates of the Schmidt hammer data, which are estimated using tape measurements and the measured dip of the scarp. However, scarp features such as plants and the track of a past TCND transect provide tiepoints between the TLS data, Schmidt hammer data, and AGSI interpretations. These tiepoints reveal that the location of the step in topographic variance corresponds to within ~ 1 m with the boundary between Zone II and Zone III in AGSI (Figs. 3 and 4), and the upper break revealed by the R-values (Fig. 6).

The TLS data is somewhat counterintuitive, however, in that it shows topographic variance *decreasing* towards the upper part of the scarp. This means that slip patches exposed by older earthquakes have smoother surfaces, at the scales investigated, than those exposed by younger earthquakes. The trend we observe is the opposite of the relationship between variance and scarp height observed by other studies on limestone normal fault scarps (Wei et al., 2013; Wiatr

et al., 2015) for length scales similar to our ~20 cm window. In addition to the results shown in Figure 7, we performed the variance analysis with a ~5 cm moving window, which produced nearly identical results to the ~20 cm window. In another analysis, we attempted to compensate for some non-planar topography observed on the scarp at small spatial scales with a local best-fit plane within each window. This procedure was undertaken using both ~5 cm and ~20 cm moving windows, and the relationship between variance and scarp height was virtually identical to that depicted in Figure 7 in both cases.

We interpret the unexpected relationship between topographic variance and scarp height that we observe as a product of jointing in the lower part of the scarp (Fig. 3). This is mirrored by the lower value of AGSI Structural Integrity for Zone II than for Zone III in Figure 3. Thus, our results suggest that variations in macro-structural features such as joints may obscure the effects of weathering in roughness analyses of fault scarps. Our findings from the Hebgen scarp suggest no relationship between the development of macro-scale structural features such as joints and the age of a paleoseismic slip patch (Fig. 4), which implies that topographic variance due to jointing is not useful for identifying slip patches. It is possible that a method for filtering out features like joints could mitigate this problem.

Reconciling SEDs with other records

Our results based on Schmidt hammer are at odds with the single event displacements previously interpreted for this specific site from TCND (Zreda and Noller, 1998). We propose

that the interpretation of more events with smaller displacements delineated by Zreda and Noller (1998) could be due to the methods they used to group samples. Setting aside the six sample groupings based on weathering characteristics presented in Zreda and Noller (1998), we fit stepwise exponential functions to the ^{36}Cl apparent ages using a procedure similar to the stepwise fitting presented above (Fig. 8; Supplemental Information). The objective of this fitting is not to redefine the absolute ages of events, as could be accomplished with full elemental analysis of the samples and well-established, robust modelling routines for determining the exposure histories of faults (Schlagenhauf et al., 2010). Rather, we aim to find the locations and number of events for which there are statistical support. Each stepwise piece can be expressed as

$$y = y_0 + Ae^{\lambda x}, x_{lb} < x \leq x_{ub} \quad \text{Equation 5}$$

where y is the predicted cosmogenic apparent age, y_0 is the age offset of the current step, A is the exponential coefficient, λ is the attenuation length, and x_{lb} and x_{ub} are the lower and upper bounds of x for the current step. The age offset y_0 is not a standard term of the exponential function, but is necessary here and reflects the uniform exposure to cosmic ray spallation that a given slip patch undergoes once it is exhumed. Details of the parameters and assumptions used in this fitting are listed in the Supplementary Information. Figure 8 shows our model fits to Zreda and Noller (1998) data, revealing that the data only support a model of the scarp that includes 3 slip patches (using BIC, shown in in Fig. 8). The locations of the slip patch boundaries

inferred from the Schmidt hammer and cosmogenic exposure age data are somewhat offset from one another, though we stress that the coregistration between the two datasets is not robust.

While further testing is required to resolve the debate between paleoseismic trenching and TCND of the Hebgen Fault, we consider a three event history, as evinced in the Schmidt hammer data, reanalysis of cosmogenic exposure ages, and late Pleistocene record of events in trenches, to be a plausible outcome. The differences in displacements between trenching and the study site here are not unexpected due to along-strike differences in fault kinematics and distributed (i.e., off-fault) deformation in unconsolidated sediments.

There are two alternative interpretations that our data cannot resolve. The first is that both trenching and the Schmidt hammer miss smaller events that are recorded by the original cosmogenic exposure ages of Zreda and Noller (1998). The second is that progressive exposure at this scarp could be assisted by non-tectonic, continuous or diffusive lowering of the colluvial cover at the base of the scarp. Kastelic et al. (2017) observed exposure rates due to non-tectonic erosion or deposition of colluvial material at the base of normal faults in Italy, and concluded that exposure rates could outpace fault slip by orders of magnitude. If this was the only process operating on the Hebgen fault, the linear or power-law models for progressive exposure (Fig. 6) could be the most appropriate curve fits. While our BIC models show favorability for the step-function over these other fits, and there is good evidence for discrete

displacements, we consider that the non-tectonic contribution of scarp exposure could lead to the large amount of noise for older events and potentially limit the applicability of using the Schmidt hammer beyond the last ~ 2 events.

Practical considerations of Schmidt hammer methodology

This study was limited to a single transect along the Hebgen fault, and so further testing should be conducted on other faults in different climates to confirm the utility of the Schmidt hammer in these studies. The technique seems most promising for distinguishing between MRE and penultimate event slip patches, as Stewart (1996) observed using measurements of surface roughness on limestone faults. In fact, we suggest that the main attribute of the rock weathering being measured by the Schmidt hammer is the variable cm-scale surface roughness of different slip patches. Because R-values are relatively simple to obtain in the field, the Schmidt hammer could be used as an auxiliary dataset in conjunction with other techniques (Table 1).

Other recommendations for future use are discussed below. Where possible, R-values should be collected in two dimensions along the scarp with their locations surveyed using a string grid or laser-based ranging. This approach would allow for more rigorous spatial and statistical analysis of the data. As with any measurement of paleoseismic slip at a point, care should be taken in applying the derived value of SED (or set of SEDs) to other positions along the fault, as these are likely to vary along-strike and between earthquakes. If a bedrock scarp is exposed in

several places along strike, the Schmidt hammer may be a useful tool for providing rapid, preliminary assessment of this slip variability.

In practice, we found that field investigations of steep bedrock scarps using the Schmidt hammer are likely to be limited by along strike alteration of the scarp and/or location of suitable sites to set up a rope or ladder to reach sample locations. Care should be taken to avoid areas of visible scarp modification by processes other than weathering. Operating the Schmidt hammer while suspended by rope is challenging and requires a second person to record values (for non-digitally recording hammers). Even so, it is not always possible to get reliable impacts at every planned sample location due to the geometry and relief of the fault scarp.

For future studies, our approach can be adapted to distinguish SEDs on fault scarps provided that (i) lithology does not change over the fault scarp, (ii) the fault scarp has not been significantly eroded in older events and can be traced as a single, continuous plane over the sample transect, (iii) there are no significant variations in weathering processes and rates through time (e.g., case hardening of older exposed sections), and (iv) non-tectonic exposure of the fault plane does not outpace the rate of exposure in earthquakes.

Faults with higher slip rates, smaller SEDs, and/or shorter recurrence intervals may not show enough inter-event ΔR variability to discriminate displacements (numerically or otherwise).

Similarly, faults with much slower slip rates and longer recurrence intervals may have uniform weathering characteristics and therefore ΔR over the exposed scarp. We anticipate that faults with several thousand-year recurrence intervals will yield enough variability in R-values to delineate weathering zones and SEDs. Additionally, with calibration on scarp sections of known age, chronofunctions could be developed that yield numerical paleo-earthquake ages from Schmidt hammer measurements.

Applicability of step function fitting to SED identification

The step function fitting presented in this paper requires little user input, incorporates the effects of heteroscedasticity, and requires few initial assumptions. For other paleoseismic data that are expected to show monotonically increasing, stepwise behavior, but do not have constraints on the absolute *location* or *number* of events, this approach is useful in identifying SEDs. Other approaches using regression trees or the sliding window Student's *t*-test of He et al. (2016) are promising, but require more user input and could introduce user bias in defining SEDs.

The assessment of models using the BIC should be used with an understanding of what the BIC represents. This BIC penalizes additional parameters in curve fitting, so it is ideal for avoiding over-fitting slip histories and interpreting too many earthquakes. However, for long-lived bedrock scarps with many small displacements, the BIC may prefer functions with only a few steps or continuous models over a 'true' but complex slip history, thus under-representing the

earthquake hazard of a fault. As such, user-discretion is still advisable for selecting the most appropriate model from information criteria.

Conclusions

The major conclusions of this preliminary study are outlined below:

- (1) We have outlined a field methodology and application of the Schmidt hammer for delineating single event displacements (SEDs) on bedrock fault scarps.
- (2) SEDs can be identified by fitting step-functions and other models to the R-value range over 5 impacts at a point (ΔR) and assessing the model fits using the Bayesian Information Criterion (BIC). We propose that this approach is applicable to other scarp dating techniques (Table 1).
- (3) For the Hebgen fault, the best-fitting model is a stepwise, three-event model. SEDs are generally consistent with several meter-scale displacements of the fault in three recent events as proposed by trenching studies and by reanalysis of cosmogenic exposure ages for the same scarp. The exact lengths of these displacements are not resolved, but they probably represent a maximum of c. 4.8 m.
- (4) The Schmidt hammer should be used in conjunction with other absolute and relative-age dating techniques to further explore its utility in assessing fault slip in paleo-earthquakes. The results presented here are promising and could facilitate rapid field assessment of SEDs along active normal faults.

Acknowledgements

This research was supported by NSF EAR 1451466. This material is based in part on TLS data services provided by the UNAVCO Facility with support from the National Science Foundation (NSF) and National Aeronautics and Space Administration (NASA) under NSF Cooperative Agreement No. EAR-0735156. We thank Chris Crosby and William Medwedeff for assistance in the field and are grateful to Kendra Murray and Eric Portenga for providing constructive reviews of early versions of this manuscript. We are also grateful to Nathan Niemi, Marin Clark, and Eric Hetland for insightful conversations about Schmidt hammers, field sites, and model selection.

References

- Aa AR, Sjøstad J, Sønstegaard E, Blikra LH. 2007. Chronology of Holocene rock-avalanche deposits based on Schmidt-hammer relative dating and dust stratigraphy in nearby bog deposits, Vora, inner Nordfjord, Norway. *The Holocene* 17: 955–964. DOI: 10.1177/0959683607082411
- Ahnert, F. (1984). Local relief and the height limits of mountain ranges. *American Journal of Science*, 284(9), 1035-1055.
- Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. Pages 267-281 in B. N. Petrov, and F. Csaki, (Eds.) *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.
- Aydin A. 2009. ISRM suggested method for determination of the Schmidt hammer rebound hardness: revised version. *International Journal of Rock Mechanics and Mining Sciences* 46:627–634.
- Aydin A, Basu A. 2005. The Schmidt hammer in rock material characterization. *Engineering Geology* 81: 1-14. DOI: <http://dx.doi.org/10.1016/j.enggeo.2005.06.006>.
- Benedetti L, Finkel R, Papanastassiou D, King G, Armijo R, Ryerson F, Farber D, Flerit F. 2002. Post-glacial slip history of the Sparta fault (Greece) determined by ³⁶Cl cosmogenic dating: Evidence for non-periodic earthquakes. *Geophysical Research Letters* 29: 87-1–87-4. DOI: 10.1029/2001GL014510.
- Benedetti LC, van der Woerd J. 2014. Cosmogenic Nuclide Dating of Earthquakes, Faults, and Toppled Blocks. *Elements* 10: 357-361. DOI: 10.2113/gselements.10.5.357.
- Benedict JB. 1985. Arapaho Pass, Glacial Geology and Archaeology at the Crest of the Colorado Front Range, Research Report no. 3. Center for Mountain Archaeology: Ward, CO; 197 pp.
- Birkeland, P. W. and Noller, J. S. 2000. Rock and Mineral Weathering. in *Quaternary Geochronology: Methods and Applications* (eds J. S. Noller, J. M. Sowers and W. R. Lettis), American Geophysical Union, Washington, D. C.. doi: 10.1029/RF004p0293.
- Boncio P, Galli P, Naso G, Pizzi A. 2012. Zoning surface rupture hazard along normal faults: Insight from the 2009 M w 6.3 L'Aquila, central Italy, earthquake and other global earthquakes. *Bulletin of the Seismological Society of America* 102: 918-935. DOI: 10.1785/0120100301.
- Bosi C, Galadini F, Messina P. 1993. Neotectonic significance of bedrock fault scarps: case studies from the Lazio-Abruzzi Apennines (central Italy). *Zeitschrift fur Geomorphologie, Supplementband* 94: 187-206.
- Bull WB. 1996. Prehistorical earthquakes on the Alpine fault, New Zealand. *Journal of Geophysical Research: Solid Earth* 101: 6037-6050. DOI: 10.1029/95JB03062.
- Burnham KP, Anderson DR. 2004. Multimodel Inference. *Sociological Methods & Research* 33: 261-304. DOI: doi:10.1177/0049124104268644.
- Candela T, Renard F, Bouchon M, Brouste A, Marsan D, Schmittbuhl J, Voisin C. 2009. Characterization of Fault Roughness at Various Scales: Implications of Three-Dimensional High Resolution Topography Measurements. *Pure and Applied Geophysics* 166, 1817-1851.
- Carcaillet J, Manighetti I, Chauvel C, Schlagenhauf A, Nicole JM. 2008. Identifying past earthquakes on an active normal fault (Magnola, Italy) from the chemical analysis of its exhumed carbonate fault plane. *Earth and Planetary Science Letters* 271: 145-158. DOI: <http://dx.doi.org/10.1016/j.epsl.2008.03.059>.

- Colman SM, Dethier DP. 1986. Rates of chemical weathering of rocks and minerals. Academic Press: Orlando; 603 p.
- Crook R, Gillespie AR. 1986. Weathering rates in granitic boulders measured by P-wave speeds. In Rates of Chemical Weathering of Rocks and Minerals, Colman SM, Dethier DP (eds). Academic Press: Orlando, FL; 395–417.
- Darvill CM, Bentley MJ, Stokes CR. 2015. Geomorphology and weathering characteristics of erratic boulder trains on Tierra del Fuego, southernmost South America: Implications for dating of glacial deposits. *Geomorphology* **228**: 382-397. DOI: <https://doi.org/10.1016/j.geomorph.2014.09.017>.
- Day MJ. 1980. Rock hardness: field assessment and geomorphic importance. *The Professional Geographer* **32**: 72-81. DOI: 10.1111/j.0033-0124.1980.00072.x.
- Day, MJ, Goudie AS. 1977. Field assessment of rock hardness using the Schmidt test hammer. *BGRG Technical Bulletin* **18**: 19–29.
- Demirdag S, Yavuz H, Altindag R. 2009. The effect of sample size on Schmidt rebound hardness value of rocks. *International Journal of Rock Mechanics and Mining Sciences* **46**: 725-730. DOI: <http://doi.org/10.1016/j.ijrmms.2008.09.004>.
- Doser DI. 1985. Source parameters and faulting processes of the 1959 Hebgen Lake, Montana, earthquake sequence. *Journal of Geophysical Research: Solid Earth* **90**: 4537-4555. DOI: 10.1029/JB090iB06p04537.
- Giaccio B, Galadini F, Sposato A, Messina P, Moro M, Zreda M, Cittadini A, Salvi S, Todero A. 2003. Image processing and roughness analysis of exposed bedrock fault planes as a tool for paleoseismological analysis: results from the Campo Felice fault (central Apennines, Italy). *Geomorphology* **49**: 281-301. DOI: [http://dx.doi.org/10.1016/S0169-555X\(02\)00191-5](http://dx.doi.org/10.1016/S0169-555X(02)00191-5).
- Goudie AS. 2006. The Schmidt Hammer in geomorphological research. *Progress in Physical Geography* **30**: 703-718. DOI: 10.1177/0309133306071954.
- Gupta V, Sharma R, Sah MP. 2009. An evaluation of surface hardness of natural and modified rocks using Schmidt hammer: study from northwestern Himalaya, India. *Geografiska Annaler* **91**: 179–188.
- He H, Wei Z, Densmore A. 2016. Quantitative morphology of bedrock fault surfaces and identification of paleo-earthquakes. *Tectonophysics* **693, Part A**: 22-31. DOI: <http://doi.org/10.1016/j.tecto.2016.09.032>.
- Hecker S, Stenner HD, Schwartz DP, and Hamilton JC. 2000. Paleoseismic results from the central part of the 1959 Hebgen fault rupture, Montana. *EOS Transactions, American Geophysical Union* **81**: Suppl., p. 1170.
- Hecker S, Stenner HD, Costa CH, Schwartz DP, and Hamilton JC. 2002. New paleoseismic results from the central part of the 1959 Hebgen fault rupture, Montana. *Geological Society of America Abstracts with Programs* **34**: p. 4.
- Marinos V, Marinos P, Hoek E. 2005. The geological Strength index: applications and limitations. *Bulletin of Engineering Geology and the Environment* **64**: 55-65.
- Matthews JA, Owen G. 2010. Schmidt hammer exposure-age dating: developing linear age-calibration curves using Holocene bedrock surfaces from the Jotunheimen-Jostedalbreen regions of southern Norway. *Boreas* **39**: 105–115. DOI: 10.1111/j.1502-3885.00107.x

- Matthews JA, Shakesby RA. 1984. The status of the 'Little Ice Age' in southern Norway: relative-age dating of neoglacial moraines with Schmidt hammer and lichenometry. *Boreas* **13**: 333–346.
- Hooke R, Jeeves TA. 1961. Direct Search Solution of Numerical and Statistical Problems. *Journal of the ACM* **8**: 212–229. DOI: 10.1145/321062.321069.
- Johnson, K.L. 2017. Applications of High Resolution Topography in Tectonic Geomorphology. Colorado School of Mines PhD Thesis: 138 p.
- Karaman K, Kesimal A. 2015. A comparative study of Schmidt hammer test methods for estimating the uniaxial compressive strength of rocks. *Bulletin of Engineering Geology and the Environment* **74**: 507–520. DOI: 10.1007/s10064-014-0617-5.
- Kass RE, Raftery AE. 1995. Bayes Factors. *Journal of the American Statistical Association* **90**: 773–795. DOI: 10.1080/01621459.1995.10476572.
- Kastelic V, Burrato P, Carafa MMC, Basili R. 2017. Repeated surveys reveal nontectonic exposure of supposedly active normal faults in the central Apennines, Italy. *Journal of Geophysical Research: Earth Surface* **122**: 114–129. DOI: 10.1002/2016JF003953.
- Laustela M, Egli M, Frauenfelder R, Kääb A, Maisch M, Haerberli W. 2003. Weathering rind measurements and relative age dating of rockglacier surfaces in crystalline regions of the Eastern Swiss Alps. Permafrost: Proceedings of the Eighth International Conference on Permafrost, July 2003, Phillips M, Springman S, Arenson L (eds), Zurich; 627–632. Licciardi JM, Maizels J. 1989. Differentiation of late Pleistocene terrace outwash deposits using geomorphic criteria: Tekapo valley, South Island, New Zealand. *New Zealand Journal of Geology and Geophysics* **32**: 225–241.
- Manighetti I, Boucher E, Chauvel C, Schlagenhauf A, Benedetti L. 2010. Rare earth elements record past earthquakes on exhumed limestone fault planes. *Terra Nova* **22**: 477–482. DOI: 10.1111/j.1365-3121.2010.00969.x.
- Matthews JA, Owen G. 2008. Endolithic lichens, rapid biological weathering and schmidt hammer R-values on recently exposed rock surfaces: storbreen glacier foreland, Jotunheimen, Norway. *Geografiska Annaler: Series A, Physical Geography* **90**: 287–297. DOI: 10.1111/j.1468-0459.2008.00346.x.
- Matthews JA, McEwen LJ, Owen G. 2015. Schmidt-hammer exposure-age dating (SHD) of snow-avalanche impact ramparts in southern Norway: approaches, results and implications for landform age, dynamics and development. *Earth Surface Processes and Landforms* **40**: 1705–1718. DOI: 10.1002/esp.3746.
- Matthews JA, Owen G, Winkler S, Vater AE, Wilson P, Mourne RW, Hill JL. 2016. A rock-surface microweathering index from Schmidt hammer R-values and its preliminary application to some common rock types in southern Norway. *Catena* **143**: 35–44. DOI: <https://doi.org/10.1016/j.catena.2016.03.018>.
- McCarroll D. 1987. The Schmidt Hammer in geomorphology: five sources of instrument error. *BGRG Technical Bulletin* **36**: 16–27.
- McCarroll D. 1991. The schmidt hammer, weathering and rock surface roughness. *Earth Surface Processes and Landforms* **16**: 477–480. DOI: 10.1002/esp.3290160510.

- Mitchell SG, Matmon A, Bierman PR, Enzel Y, Caffee M, Rizzo D. 2001. Displacement history of a limestone normal fault scarp, northern Israel, from cosmogenic ^{36}Cl . *Journal of Geophysical Research: Solid Earth* **106**: 4247-4264. DOI: 10.1029/2000JB900373
- Mouslopoulou V, Moraetis D, Fassoulas C. 2011. Identifying past earthquakes on carbonate faults: Advances and limitations of the 'Rare Earth Element' method based on analysis of the Spili Fault, Crete, Greece. *Earth and Planetary Science Letters* **309**: 45-55. DOI: <http://dx.doi.org/10.1016/j.epsl.2011.06.015>
- Nicholson DT. 2009. Holocene microweathering rates and processes on ice-eroded bedrock, Røldal area, Hardangervidda, southern Norway. *Geological Society of London Special Publications* **320**: 29-49. DOI: 10.1144/sp320.3.
- Niedzielski T, Migoń P, Placek A. 2009. A minimum sample size required from Schmidt hammer measurements. *Earth Surface Processes and Landforms* **34**: 1713-1725. DOI: 10.1002/esp.1851
- O'Neill JM, Christiansen RL. 2004. Geologic map of the Hebgen Lake quadrangle, Beaverhead, Madison, and Gallatin counties, Montana, Park and Teton counties, Wyoming, and Clark and Fremont counties, Idaho. In Scientific Investigations Map: Denver, CO.
- Ozbek A, Gul M. 2011. Variation of Schmidt hammer rebound values depending on bed thickness and discontinuity surfaces. *Scientific Research and Essays* **6**: 2201-2211
- Pierce KL. 2003. Pleistocene glaciations of the Rocky Mountains. In *Developments in Quaternary Sciences*. Elsevier; 63-76.
- Pierce KL. 2008. Cosmogenic exposure-age chronologies of Pinedale and Bull Lake glaciations in greater Yellowstone and the Teton Range, USA. *Quaternary Science Reviews* **27**: 814-831. DOI: <https://doi.org/10.1016/j.quascirev.2007.12.005>.
- Pierce KL, Lageson DR, Ruleman CA, Hintz RG. 2000. Holocene paleoseismology of Hebgen Lake normal fault, MT—The Cabin Creek site of the Hebgen Lake paleoseismology working group: *EOS Transactions, American Geophysical Union* **81**: Suppl., p. 1170.
- Sagy A, Brodsky EE, Axen GJ; Evolution of fault-surface roughness with slip. *Geology* **35** (3): 283–286. doi: <https://doi.org/10.1130/G23235A.1>
- Schlagenhauf A, Gaudemer Y, Benedetti L, Manighetti I, Palumbo L, Schimmelpfennig I, Finkel R, Pou K. 2010. Using in situ Chlorine-36 cosmnuclide to recover past earthquake histories on limestone normal fault scarps: a reappraisal of methodology and interpretations. *Geophysical Journal International* **182**: 36-72. DOI: 10.1111/j.1365-246X.2010.04622.x.
- Schwarz G. 1978. Estimating the Dimension of a Model. *Annals of Statistics* **6**: 461-464. DOI: 10.1214/aos/1176344136.
- Schwartz, D.P., Hecker, S., Stenner, H.D., Haller, K.M., Pierce, K., Lageson, D.R., Machette, M.N. 2009. The 1959 Hebgen Lake, Montana, surface rupture and record of late-Pleistocene-Holocene earthquakes. *Geological Society of America Abstracts with Programs* **41**: p. 53.
- Shakesby RA, Matthews JA, Karlén W, Los SO. 2011. The Schmidt hammer as a Holocene calibrated-age dating technique: Testing the form of the R-value-age relationship and defining the predicted-age errors. *The Holocene* **21**: 615-628. DOI: 10.1177/0959683610391322.
- Shakesby RA, Matthews JA, Owen G. 2006. The Schmidt hammer as a relative-age dating tool and its potential for calibrated-age dating in Holocene glaciated environments. *Quaternary Science Reviews* **25**: 2846-2867. DOI: <http://dx.doi.org/10.1016/j.quascirev.2006.07.011>.

- Stahl T, Winkler S, Quigley M, Bebbington M, Duffy B, Duke D. 2013. Schmidt hammer exposure-age dating (SHD) of late Quaternary fluvial terraces in New Zealand. *Earth Surface Processes and Landforms* **38**: 1838-1850. DOI: 10.1002/esp.3427.
- Stahl T, Quigley MC, Bebbington MS. 2016. Tectonic geomorphology of the Fox Peak and Forest Creek Faults, South Canterbury, New Zealand: slip rates, segmentation and earthquake magnitudes. *New Zealand Journal of Geology and Geophysics* **59**: 568-591. DOI: 10.1080/00288306.2016.1212908.
- Stewart I. 1996. A rough guide to limestone fault scarps. *Journal of Structural Geology* **18**: 1259-1264. DOI: [http://dx.doi.org/10.1016/S0191-8141\(96\)00049-1](http://dx.doi.org/10.1016/S0191-8141(96)00049-1).
- Sugiura N. 1978. Further analysts of the data by akaike' s information criterion and the finite corrections. *Communications in Statistics - Theory and Methods* **7**: 13-26. DOI: 10.1080/03610927808827599.
- Sumner P, Nel W. 2002. The effect of rock moisture on Schmidt hammer rebound: tests on rock samples from Marion Island and South Africa. *Earth Surface Processes and Landforms* **27**: 1137-1142. DOI: 10.1002/esp.402.
- Török Á, Forgó LZ, Vogt T, Löbens S, Siegesmund S, Weiss T. 2007. The influence of lithology and pore-size distribution on the durability of acid volcanic tuffs, Hungary. *Geological Society of London Special Publications* **271**: 251-260. DOI: 10.1144/gsl.sp.2007.271.01.24.
- Tucker GE, McCoy SW, Whittaker AC, Roberts GP, Lancaster ST, Phillips R. 2011. Geomorphic significance of postglacial bedrock scarps on normal-fault footwalls. *Journal of Geophysical Research: Earth Surface* **116**: F01022. DOI: 10.1029/2010JF001861.
- van der Woerd J, Benedetti L, Caffee MW, Finkel R. 2000. Slip-rate and earthquake recurrence time on the Hebgen Lake fault (Montana)—Constraints from surface exposure dating of alluvial terraces and bedrock fault scarp: *EOS Transactions, American Geophysical Union* **81**: Suppl. p. 1160.
- Villamor P, Litchfield N, Barrell D, Van Dissen R, Hornblow S, Quigley M, Levick S, Ries W, Duffy B, Begg J, Townsend D, Stahl T, Bilderback E, Noble D, Furlong K, Grant H. 2012. Map of the 2010 Greendale Fault surface rupture, Canterbury, New Zealand: application to land use planning. *New Zealand Journal of Geology and Geophysics* **55**: 223-230. DOI: 10.1080/00288306.2012.680473.
- Wei Z, He H, Shi F. 2013. Weathering history of an exposed bedrock fault surface interpreted from its topography. *Journal of Structural Geology* **56**: 34-44. DOI: 10.1016/j.jsg.2013.08.008.
- Wells DL, Coppersmith KJ. 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* **84**: 974-1002
- Wicks CW, Thatcher W, Dzurisin D, Svarc J. 2006. Uplift, thermal unrest and magma intrusion at Yellowstone caldera. *Nature* **440**: 72. DOI: 10.1038/nature04507.
- Winkler S. 2005. The Schmidt hammer as a relative-age dating technique: potential limitations of its application on Holocene moraines in Mt. Cook National Park, Southern Alps, New Zealand. *New Zealand Journal of Geology and Geophysics* **48**: 105–116. DOI: 0028–8306/05/4801–0105.
- Winkler S, Matthews JA, Mourné RW, Wilson P. 2016. Schmidt-hammer exposure ages from periglacial patterned ground (sorted circles) in Jotunheimen, Norway, and their interpretative problems. *Geografiska Annaler: Series A, Physical Geography* **98**: 265-285. DOI: 10.1111/geoa.12134

- Witkind IJ, Myers WB, Hadley JB, Hamilton W, Fraser GD. 1962. Geologic features of the earthquake at Hebgen Lake, Montana, August 17, 1959. *Bulletin of the Seismological Society of America* **52**: 163-180.
- Zreda M, Noller JS. 1998. Ages of Prehistoric Earthquakes Revealed by Cosmogenic Chlorine-36 in a Bedrock Fault Scarp at Hebgen Lake. *Science* **282**: 1097-1099. DOI: 10.1126/science.282.5391.1097.

Figure 1: Location and fault trace map of the Hebgen and Red Canyon faults in southwestern Montana. Fault traces are from the USGS Quaternary Fault and Fold database.

Figure 2: Expected relationships between Schmidt hammer R-value, cosmogenic nuclide content, geologic strength index (GSI), surface roughness, scarp height, and relative earthquake age (e.g. EQ1, EQ2, etc.). (A) R-value range (ΔR) and $*I5$ index should increase in discrete steps along scarp height that correspond to single event displacements (SEDs) in paleo-earthquakes. (B) SEDs are defined by the distance along scarp height that ΔR are constant, with the expectation that older earthquakes show more scatter in ΔR .

Figure 3: (A) Photo of the scarp in this study and that of Zreda and Noller (1998) (location in inset). (B) Photo of slickensides near the top of the scarp in (A). Slickensides become more poorly preserved up-fault (C) Photo of scarp section produced in the 1959 Hebgen Lake earthquake. The section is marked by a smooth, striated plane with few weathering pits or staining. (D) Estimations of Adjusted Geologic Strength Index (AGSI) for the scarp. Surface conditions degrade up-scarp (lower values and shifted to the right on the top axis) while variability in structural integrity less pronounced.

Figure 4: Schmidt hammer R-value vs. scarp height for all impacts. Adjusted Geologic Strength Indices (AGSI) are shown for three weathering zones (I-III) and correspond to discrete changes in ΔR (the distance between lowest and highest R-values at each point, colored series).

Figure 5: Positive correlation of Schmidt hammer R-value range (ΔR), $*I5$, and $*I_{max}$ with scarp height. See text for discussion.

Figure 6: Model fitting to Schmidt hammer data in $\ln(\Delta R)$ space. Using a natural logarithmic ΔR scale reduces the heteroscedasticity observed in ΔR values simplifying model fitting and selection. Only the best-fitting step model (blue, Table 2) is shown here, which has a lower BIC than linear (red) and power law (yellow) model fits.

Figure 7: Microtopographic analysis from TLS data. (A) Hebgen lake fault scarp TLS dataset has been fit with a plane, and values indicate the distance of each point from the best-fit plane. A cosmogenic nuclide sample transect is visible and provides a common reference to Fig. 4. The 7.5m measurement indicates vertical height. (B) Variance of topography calculated over the fault scarp using a moving window. The slightly grayed out region is the full scan, and the fully opaque region is that used for the subsequent analysis. Inset shows mechanics of the moving window. The center of each cell corresponds to the center of one location of the moving window, which is actually twice as wide and twice as tall as the cells shown. Thus, adjacent cells in the main portion cover some of the same topography. (C) Mean and standard deviation topographic variance for each row of the moving window analysis. A step function fit (red) is barely preferred over a power law fit (green; see Table 3).

Figure 8: Correspondence of discrete breaks in Schmidt hammer data (grey filled circles) and apparent cosmogenic exposure ages of Zreda and Noller (1998). The scarp locations of the cosmogenic ages have

been shifted using the top of the 1959 slip patch as a common datum between the two datasets, though some uncertainty exists in the coregistration of the two datasets. The highly variable uncertainty in reported apparent cosmogenic ages results in some data having a much greater influence on the fit than other data and introduces uncertainty in the location of the steps.

Table 1: Overview of methods for distinguishing weathering zones on fault scarps.

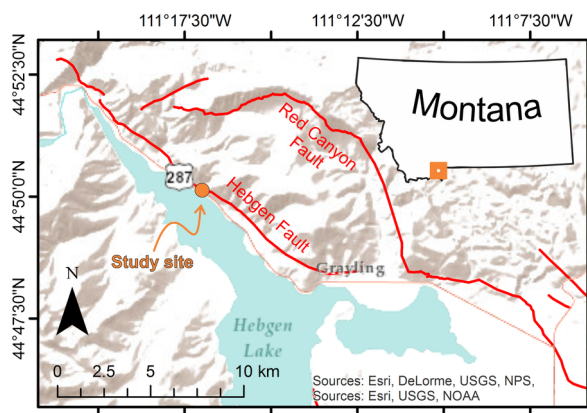
Method	Technique	Description/Utility	Practical and Theoretical Limitations	References
Surface roughness	Micro-roughness meter	Can distinguish most recent event in some circumstances	1D (transect-based), but can be applied in 2D if scarp is accessible	Stewart et al., 1996; Giaccio et al., 2003
	Terrestrial lidar	High-resolution, 3D topography of fault scarp; Can distinguish between Holocene-aged events	Equipment is expensive and required significant data processing	Wei et al., 2013; Wiatr et al., 2015; He et al., 2016; This study
	Pit Depth	Pit depths increase up-scarp	Requires manual measurement (transect-based), but can be applied in 2D	Tucker et al., 2011
Image analysis	Color	Identification of paleo bedrock-soil boundary	Some subjectivity in manipulation of color; Time-independent color changes	Giaccio et al., 2003
Geochemical	Rare Earth Element (REE) and Yttrium	Previous contact with soil enriches fault plane in REE prior to exposure, concentrations can be measured on exposed fault scarp	Requires rock sampling equipment and subsequent lab work	Carcaillet et al., 2008; Manighetti et al., 2010; Mouslopoulo et al., 2011)
Descriptive	Adjusted GSI	Incorporates numerous weathering properties of rocks	Some subjectivity in assigning GSI values; metrics depend on geomorphic framework	This study
Schmidt hammer	Rebound values	Incorporates numerous mechanical properties of rock into R-value measurement	1D (transect-based), but can be applied in 2D if scarp is accessible	This study

Table 2: BIC values for ΔR maximum likelihood models of various functional forms (low values indicate a better fit). The preferred model (Stepwise with two breaks) is bolded.

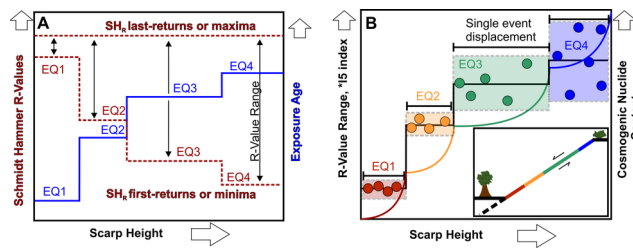
Model	BIC
Linear	49.14
Power law	63.95
Stepwise (1 break)	56.72
Stepwise (2 breaks)	46.77
Stepwise (3 breaks)	49.80

Table 3: BIC values for topographic log variance maximum likelihood models of various functional forms. The preferred model is bolded.

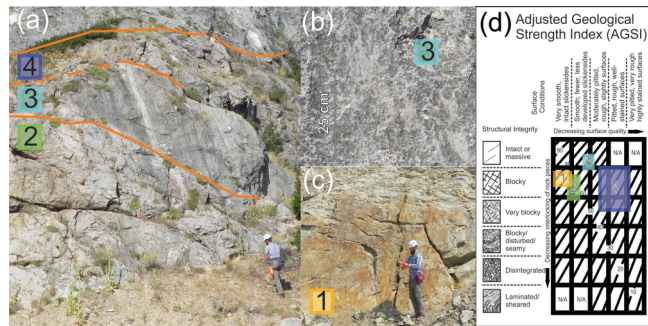
Model	BIC
Power law	179.89
Stepwise (1 break)	178.34
Stepwise (2 breaks)	185.64



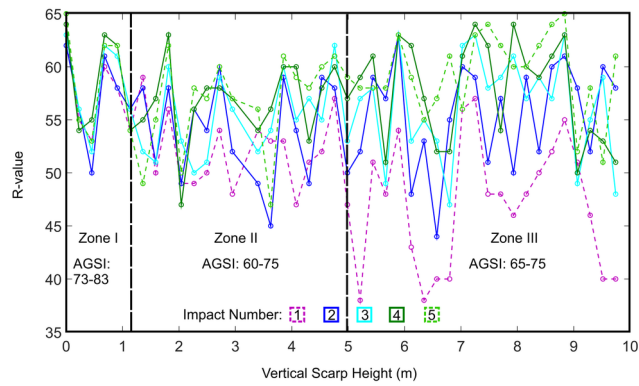
ESP_4403_F1.tif



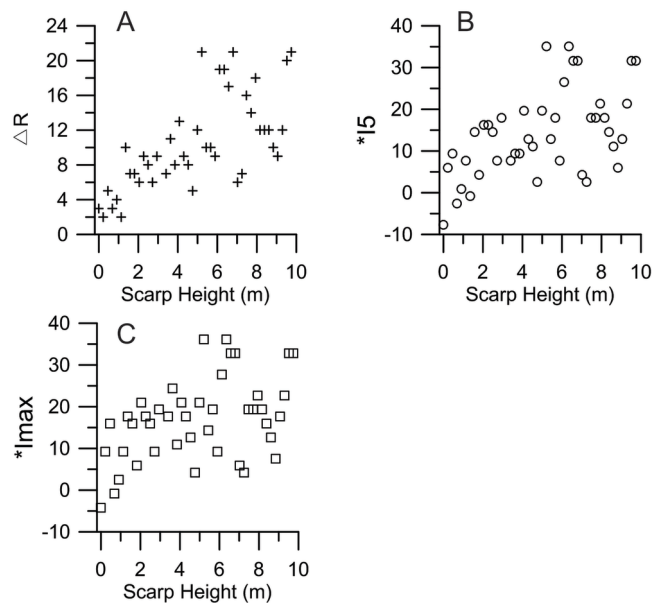
ESP_4403_F2.tif



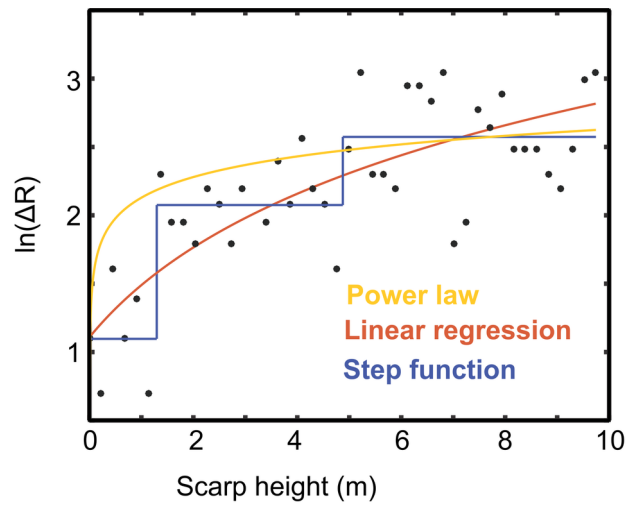
ESP_4403_F3.tif



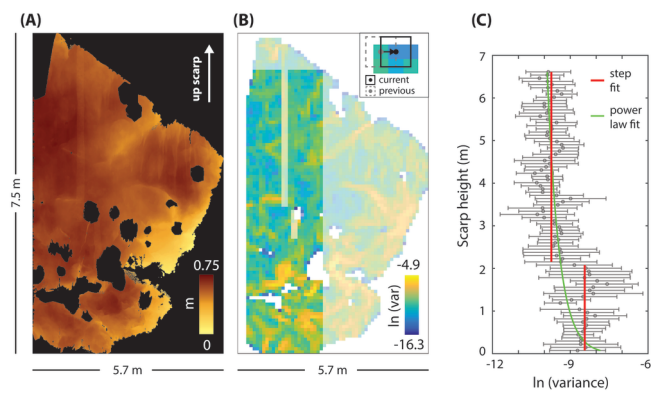
ESP_4403_F4.tif



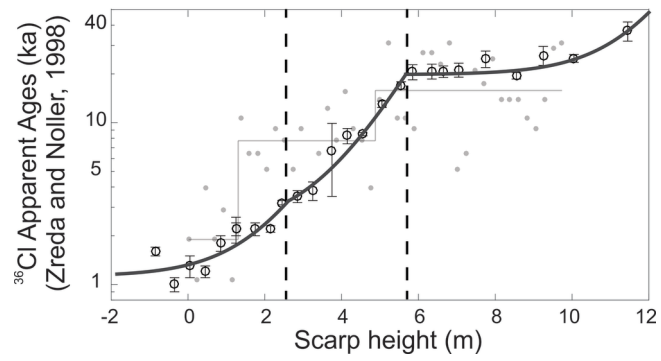
ESP_4403_F5.tif



ESP_4403_F6.tif



ESP_4403_F7.tif



ESP_4403_F8.tif