Geophysical Research Letters

Supporting Information for

Smartrock transport during snowmelt floods: Discharge controls on rest scaling from seconds to seasons

Kealie L. G. Pretzlav[1,*], Joel P. L. Johnson [1,†] , D. Nathan Bradley[2]

[1] Dept. of Geological Sciences, The University of Texas at Austin, Austin, TX, USA

[2] U.S. Bureau of Reclamation, Denver, CO, USA, ORCID 0000-0003-3062-9494

*Now at Balance Hydrologics, Berkeley, CA, USA, ORCID 0000-0002-9917-7206

[†]Corresponding author, joelj@jsg.utexas.edu, ORCID 0000-0001-6286-9949

Contents of this file

Text S1: Comparison of methods and uncertainties for calculating rest-time scaling exponents

Text S2: Description of scaling break determination, based on methodology of Clauset et al. (1999)

Text S3: Description of the theoretical framework for determining diffusion scaling from rest time scaling of Weeks et. al (1996) and Weeks and Swinney (1998)

Figures S1 to S4

Text S1: Comparison of methods and uncertainties for calculating rest-time scaling exponents

In addition to the methods reported in Clauset et al. (2009) and Nuyts (2010), another way to calculate power-law exponents uses the slope of linear regressions to log-transformed distributions and exceedance probabilities. For example, the equation $P(t_d > t) = kt^{-\alpha}$ becomes $\log(P(t_d > t)) = -\alpha \log(t) + \log(k)$, and a simple linear regression can be used to find slope α and *k* from the intercept. It is established that this method can result in exponent values that appear to visually match plotted data but can be inaccurate and biased (e.g., Clauset et al., 2009). Nonetheless, since this method is straightforward and widely used, as a methodological comparison we also calculate α in this manner.

The log-transformed method gives α_1=0.38±0.0005 (R^2=0.997, p≈0, 7 s ≤ t_r ≤ 12.5 min), α_2=0.62±0.0009 (R^2=0.99, p≈0, 12.5 min ≤ t_r ≤ 12.3 hr), and α_3=1.33±0.05 (R^2=0.97, p≈0, t_r ≥ 12.3 hr). Uncertainties correspond to 95% confidence intervals on the regressions. Comparison to Table 1 shows that all of the exponents are similar, with α_2 and α_3 within uncertainty bounds of the other method. The empirical uncertainties from the Clauset et al. (2009) method (using the estimator of Nuyts (2010)) are much larger than regression uncertainties from the log-transformations. In particular, the ratio of Clauset et al. (2009) to log-transformed confidence intervals are 40, 55, and 6 for α_1, α_2 and α_3 respectively.

Finally, for considering uncertainties, we note that Clauset et al. (2009) say that their "experience suggests that $n \gtrsim 50$ is a reasonable rule of thumb for extracting reliable [tail exponent] parameter estimates", where *n* is the number of points fit by the power law. Estimates in previous work based on signficantly fewer points than this may be relatively inaccurate.

Text S2: Description of scaling break determination, based on methodology of Clauset et al. (1999)

The method we use to determine rest duration scaling breaks is based on Clauset et al. (2009). Their method finds a cutoff value (in our case, a rest duration) at which the cumulative exceedance probability of larger values best follows a power-law distribution. The goodness-of-fit between power-law and data is measured using the classic Kolmogorov-Smirnov statistic *KS*:

$$KS = max\left(\left|S_{(x)} - P_{(x)}\right|\right) \; for \; x > x_{min} \qquad \text{(S1)}$$

where x is some variable (in our case, rest durations t_d), $S_{(x)}$ is the cumulative distribution function (CDF) of the observations of x greater than a given cutoff value x_{min}, and $P_{(x)}$ is the CDF of the power-law fit to $S_{(x)}$. Lower *KS* values mean that a given x_{min} allows

the power-law to fit the tail of the distribution better. It is also worth noting that the use of cumulative distribution functions means that considering more data points tends to result in a better match between $S_{(x)}$ and $P_{(x)}$, because smaller data sets give more statistical fluctuations (Clauset et al., 2009). Thus, this approach balances the goodness of power-law fit and fitting as many points as possible. Clauset et al. (2009) apply this method to synthetic exceedance probability data and show that it works accurately to identify break points in power-law scaling.

Figure 1b in the main paper shows *KS* calculated by setting x_{min} to each t_d value. The lowest *KS* value occurred at 12.5 minutes, with similarly low *KS* values over the range 11.8-23.6 minutes. Following Clauset et al. (2009), this duration is interpreted as a scaling break. In other words, the power-law fit to t_d>12.5 minutes gives the best match (as measured by *KS*) for the overall dataset. In addition, Figure 1b shows local minima in *KS* at ≈7 s and ≈12 hours, which we also interpret as scaling breaks. Importantly, these scaling breaks are supported by using the same *KS* analysis but cropping the data set at different scaling breaks. For example, when only considering t_d <≈12.5 minutes, the strongest scaling break is found at ≈7 s. Similarly, only considering t_d >≈12.5 minutes results preserves the *KS* minimum (i.e., interpreted scaling break) at ≈12 hours. As described in the main text, *KS* minima >12.3 hours in Figure 1b predict scaling exponents (α) that are within uncertainty of our α_3 estimate. Therefore, our data at these longer rest times are not sufficient to identify whether additional scaling breaks might be found over even longer timescales. Our data are useful because they cover orders of magnitude of rest times within one internally consistent data set. Future work that is able to fill a rest time data gap between ≈300 hours and ≈10 years will be required to understand diffusion over these timescales.

We emphasize that our methodology assumes the rest data are power-law distributed and that multiple different power-law scaling regimes occur in the data. Our method is designed to find scaling breaks for which power laws best match segments of the rest time data. The fact that scaling exponents are statistically significantly different for the different scaling regimes supports the interpretation that these scaling breaks are valid and meaningful. Nonetheless, other statistical distributions may also be able to fit the data robustly. In particular, we also find that power laws, Truncated Pareto, and Exponentially-tempered Pareto distributions can all fit longer rests in our data with high degrees of goodness-of-fit. We include multiple interpretations of diffusion for the longest rest times because we do not believe that our statistical analyses are sufficient to robustly distinguish whether the transition to less heavy tails at the longest rest times in our data are more consistent with power-law scaling, truncation effects, or tempered scaling.

Text S3: Description of the theoretical framework for determining diffusion scaling from rest time scaling of Weeks et. al (1996) and Weeks and Swinney (1998)

This section provides a summary of the previously published theoretical framework for application of anomalous diffusion theory to model the dispersion of bedload tracers (Weeks et al., 1996; Weeks and Swinney, 1998).

Dispersion of bedload is considered normal (e.g. Fickian) when the variance of particle displacements (σ_x^2) scales linearly with time (t):

$$\sigma_x^2 = \langle (X_i - \langle X \rangle)^2 \rangle \propto t^{\gamma} \quad \text{(s1)}$$

where X_i is the cumulative displacement distance of a given ith tracer and $\langle\ \rangle$ denotes the ensemble average for all tracers in the population. Thus, $\langle X \rangle$ represents the average displacement distance of the tracers over time, t. Dispersion is normal when the scaling exponent, γ, is equal to 1, and variance therefore grows linearly with time (i.e. $\sigma_x^2 \propto t^1$). When $\gamma \neq 1$, dispersion is considered "anomalous" (i.e. not normal/Fickian). Subdiffusion occurs when $\gamma < 1$, indicating that σ_x^2 grows more slowly with time than it would for normal dispersion. Conversely, dispersion is superdiffusive when $\gamma > 1$, and the variance of displacements increase more quickly with time compared with normal dispersion.

For many tracer studies, measurements of dispersion have been conducted when only the starting and final resting positions of tracers is known. When the distribution of rest times and hop times or transport distances are also known, application of the theory of Weeks et al. (1996) and Weeks and Swinney (1998) can be used to determine if anomalous dispersion is either superdiffusive or subdiffusive. This framework is based upon modeling symmetric and asymmetric random walks and represents the dispersion pattern of individual hops rather than only the cumulative, event-based transport distances. It is worth noting that hop times and travel distances of individual hops are used somewhat interchangeably in their work, resulting in a built-in assumption of a linear correlation which may not be appropriate for all situations.

Figure S1a-c shows the phase diagrams from Weeks et al. (1996) and Weeks and Swinney (1998), which define the dispersive behavior (γ) as a function of the power law scaling exponents for hop and rest times (μ and ν) for symmetric, asymmetric, and strongly asymmetric random walks. Three of the four phases are normal, superdiffusive, and subdiffusive dispersion defined above. The fourth phase is called ballistic, and is defined by $\gamma = 2$. However, the ballistic phase does not apply to the strongly asymmetric case, which is what one would expect for mountain rivers where bedload moves overwhelmingly downstream.

From Weeks et al. (1996), if μ and ν are the exponents representing the asymptotic power law decay of the hop and rest time probability density functions (PDFs), then:

$$P_H(t_h) \sim t_h^{-\mu} \quad \text{(s2)}$$

and

$$P_R(t_r) \sim t_r^{-\nu} \tag{s3}$$

are the PDFs for hop and rest times distributions, respectively, as $t \to \infty$. Under this framework, dispersion is normal when both of the tails of the rest or hop time distributions scale as a power law with PDF scaling exponent μ or υ greater than 3 (Figure S1a-c). Notably, an exponential distribution, like is found in our study for hop times, can be thought of as having a power law scaling exponent that is very large, approaching ∞.

Our study, as well as Olinde and Johnson (2015) and Martin et al. (2012), used the exceedance probability, $P(X_i > x)$, to characterize dispersion of tracers. The exceedance probability is equivalent to the Complimentary Cumulative Density Function (CCDF) or $1 - CDF$, which allows power law functions to be more easily fit to distribution tails. Converting the power law scaling exponent from PDF units to CDF units requires integration of the power law function such that:

$$C_H(t_h) \sim t_h^{-(\mu-1)} \tag{s4}$$

and

$$C_R(t_r) \sim t_r^{-(\nu-1)}. \tag{s5}$$

As a result, we define CDF power law function scaling exponents such that:

$$C_H(t_h) \sim t_h^{-\beta}, \qquad \mu = \beta + 1 \tag{s6}$$

and

$$C_R(t_r) \sim t_r^{-\alpha}, \qquad \nu = \alpha + 1 \tag{s7}$$

where α and β are the power law decay exponents for the rest and hop durations, respectively, such that:

$$P(t_h > t) \propto t^{-\alpha} \tag{s8}$$

and

$$P(t_r > t) \propto t^{-\beta} \tag{s9}$$

Figure S1d-f shows our adaptation of the phase diagrams of Weeks et al., *[1996]* and *[1998]*, modified for use with the CDF of a power law function. Assuming that the distribution of hop times is approximately exponential, we can assume $\beta \to \infty$ and read the right side of each phase diagram for calculating dispersion coefficients. We also

assume that bedload transport is not well represented as symmetric as particles typically move in the downstream direction. Notably, under the assumption of exponentially distributed hop times, there is no difference between modeling dispersion as either an asymmetric or strongly asymmetric random walk (Figures S1e-f). We therefore get a function for γ under asymmetric random walks with an exponential hop time distribution:

$$\gamma = \begin{cases} 2\beta, & 0 < \beta \leq 1 \\ 3 - \beta, & 1 < \beta < 2 \end{cases} \quad \text{(s10)}$$

This equation is the same as equation (4) in the paper which accompanies this supplementary material.

Clauset, A., C.R. Shalizi, and M.E.J. Newman (2009), "Power-law distributions in empirical data" *SIAM Review* 51(4), 661-703 (2009). (arXiv:0706.1062, doi:10.1137/070710111)

Martin, R.L., Jerolmack, D.J., and Schumer, R., 2012, The physical basis for anomalous diffusion in bed load transport: Journal of Geophysical Research, v. 117, no. F1, p. F01018, doi: 10.1029/2011JF002075.

Nuyts, J. (2010), Inference about the tail of a distribution: Improvement on the Hill estimator, Int. J. Math. Math. Sci., 2010, 924013, doi:10.1155/2010/924013.

Olinde, L., and Johnson, J.P.L., 2015, Using RFID and accelerometer-embedded tracers to measure probabilities of bed load transport, step lengths, and rest times in a mountain stream: Water Resources Research, v. 51, no. 9, p. 7572–7589, doi: 10.1002/2014WR016120.

Weeks, E.R., Urbach, J.S., and Swinney, H.L., 1996, Anomalous diffusion in asymmetric random walks with a quasi-geostrophic flow example: Physica D: Nonlinear Phenomena, v. 97, no. 1-3, p. 291–310, doi: 10.1016/0167-2789(96)00082-6.

Weeks, E., and Swinney, H., 1998, Anomalous diffusion resulting from strongly asymmetric random walks: Physical Review E, v. 57, no. 5, p. 4915–4920, doi: 10.1103/PhysRevE.57.4915.

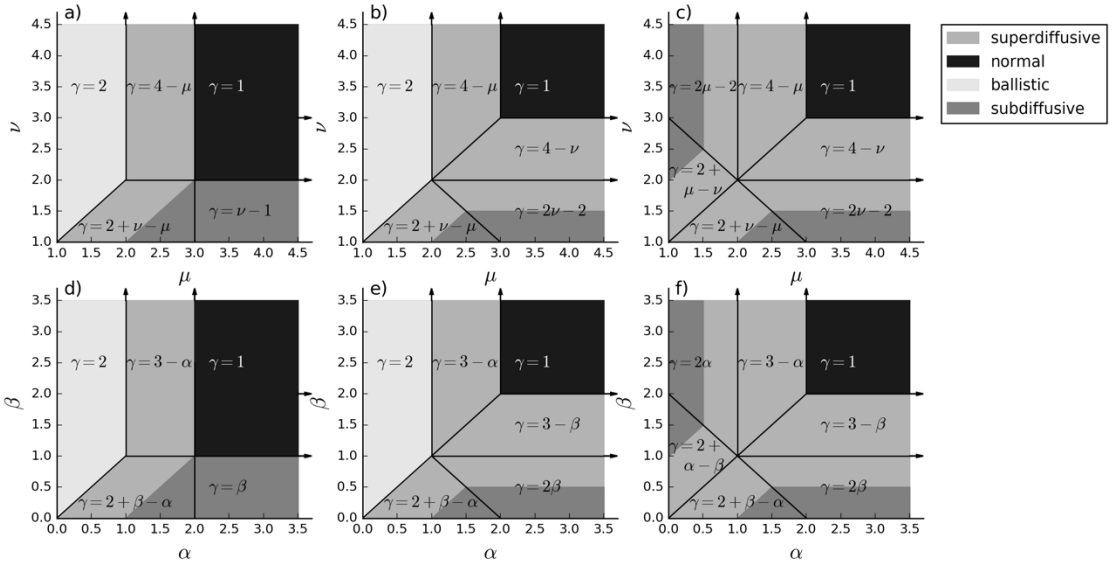

Figure S1. Modified from Weeks et al. (1996) and Weeks and Swinney (1998). Phase diagrams for variance of a) symmetric, b) asymmetric, and c) strongly asymmetric random walks for the power law decay of the PDF of rest times ($\boldsymbol{\upsilon}$) and hop times ($\boldsymbol{\mu}$). Phase diagrams for variance of d) symmetric, e) asymmetric, and f) strongly asymmetric random walks for the power law decay of the CCDF of rest times ($\boldsymbol{\alpha}$) and hop times ($\boldsymbol{\beta}$). The equations on each plot use values of power law scaling exponents on the x and y axes to determine the value of $\boldsymbol{\gamma}$, which determines whether dispersion is normal ($\boldsymbol{\gamma = 1}$), superdiffusive ($\boldsymbol{\gamma} > 1$), subdiffusive ($\boldsymbol{\gamma < 1}$), or ballistic ($\boldsymbol{\gamma}$ = 2), indicated by the shading. The phase diagram for the CCDF and the strongly asymmetrical case (f) was used for this analysis.

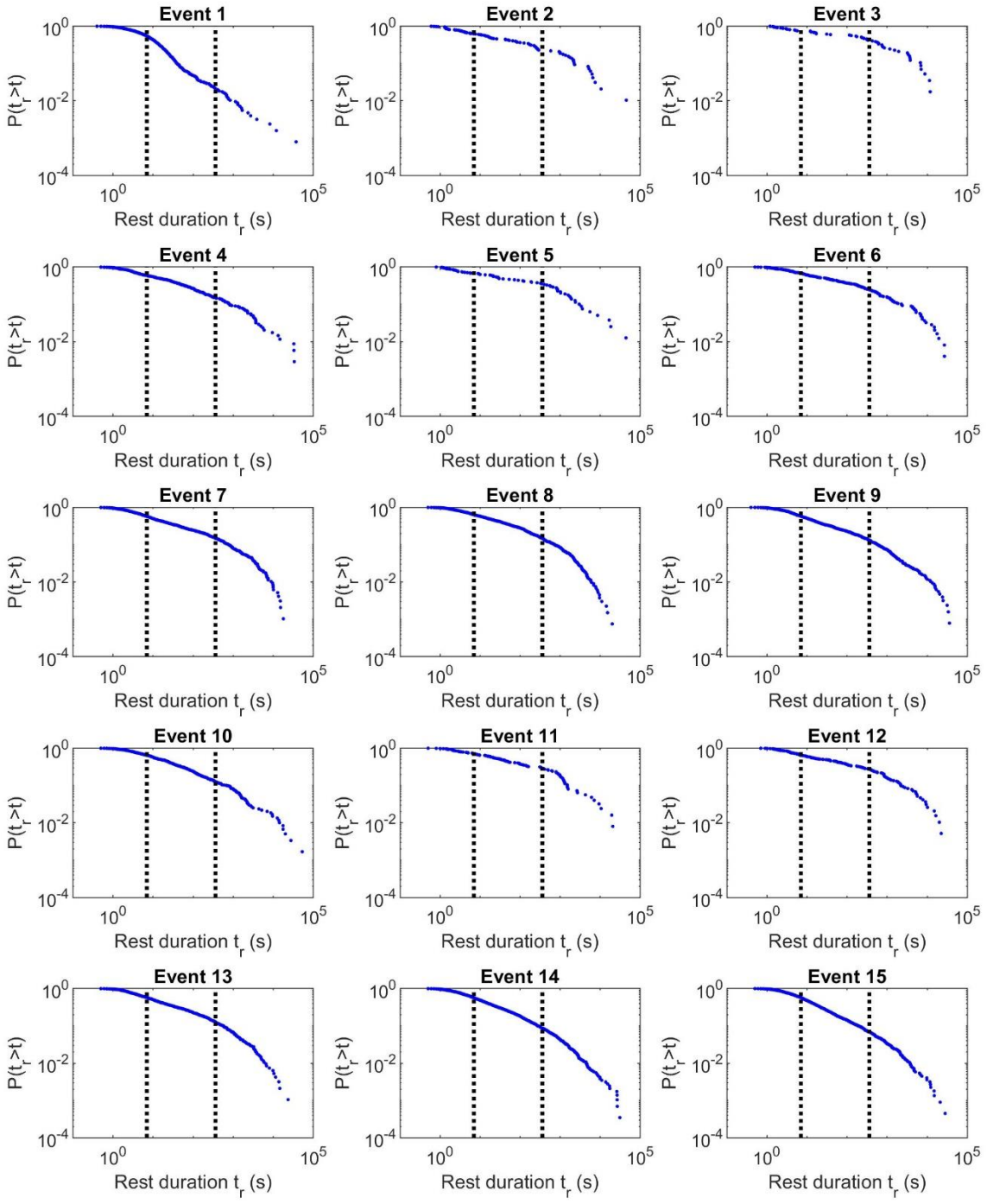

Figure S2. Rest time distributions for Halfmoon Creek daily event hydrographs 1-15. Vertical dotted lines indicate 7 second and 12 minute cutoff times used for α fits.

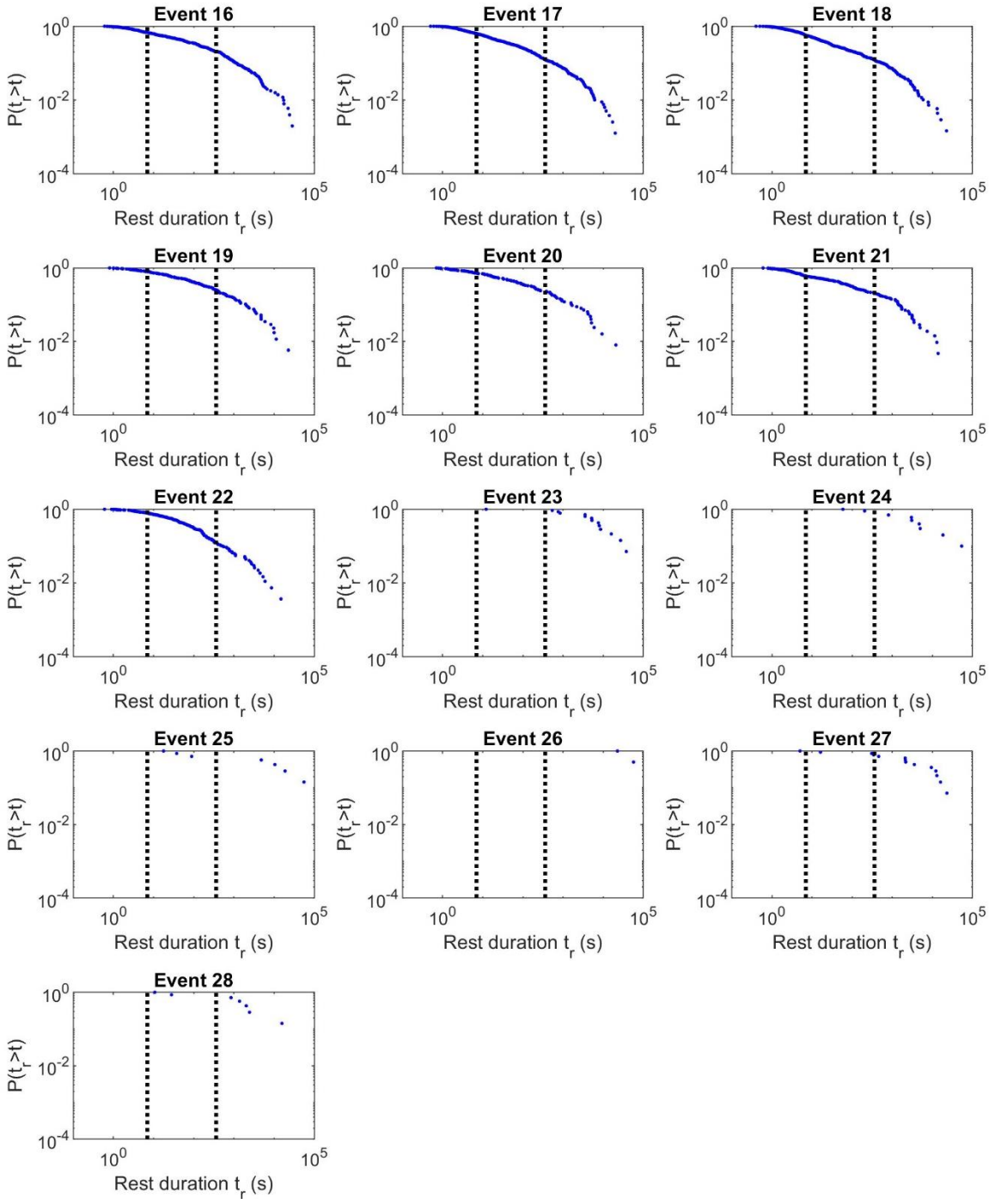

Figure S3. Rest time distributions for Halfmoon Creek daily event hydrographs 16-28. Vertical dotted lines indicate 7 second and 6 minute cutoff times used for α fits.

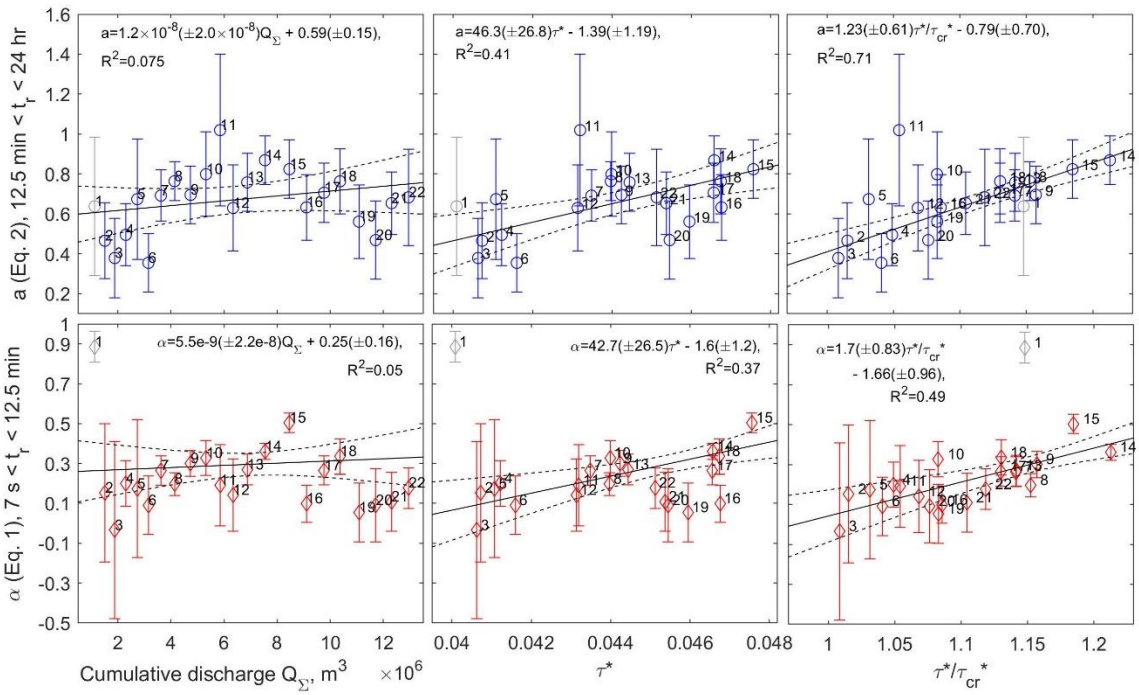

Figure S4. Plot of rest scaling exponents vs cumulative discharge, shear stress, and transport capacity, also showing numbers to identify hydrographs that correspond to data points. Note that the plots are in a rearranged order from Figure 2a-f, and a corresponds to a_{tr}.