

1 Response to Reviewer 1

We thank reviewer 1 for the thoughtful comments. We have considered many of their comments and applied we think that the updated manuscript addresses many of them.

Comment

The reviewer's main concern is with equation (2), from which many of the main results on $f_A(A)$ follow. It is not clear how the authors obtained it and what assumptions are involved. In principle, the ensemble average of A , given l , is $\langle A \rangle = \int A(f_l(l|A)f_A(A))/(f_l(l))dA$. So, the ensemble average of A should contain information about the entire joint probability distribution. How does it reduce to equation (2)?

Response

Response: It is true that Hack's Law is a statement of the mean of the conditional distribution,

$$\mu(A|l) = \int_0^{\infty} Af(A|l)dA. \quad (1)$$

However, because $f_l(l)$ doesn't appear anywhere in this expression, Hack's Law does not depend on the joint distribution - only on the conditional distribution. The joint distribution is important for our goal of understanding flow routing and distributions of hydraulic variables, but it is not necessary for understanding Hack's Law.

We suspect that the reviewer is questioning how the probabilistic representation of the mean of the conditional distribution $f_A(A|l)$ yields Hack's law the specific form of Hack's Law written in (2). Hack's Law is an empirical scaling observation which we state in the beginning of this paragraph. Equation 2 is simply a restatement of (1) with l as the independent variable and so it is an empirical result. One of our goals is to demonstrate a reasoning for the nonlinear form of Hack's Law and the value of parameters ϕ and m .

Comment

Is ϕ really a constant?

Response

We now refer to ϕ as a dimensional coefficient because, as the reviewer suggests, ϕ can change depending on the class of network one is considering.

1.0.1 Comment

The reviewer suggests that we clarify the new elements of this work as compared to *Dodds and Rothman* (2000).

1.1 Response

We agree that much of this work builds on the reasoning from *Dodds and Rothman (2000)* and we have added language to more clearly acknowledge what components are new and which build on previous work. In particular, early on in section 2, we state that our primary contribution here is to contribute towards a formal understanding of the moments of the Hack distributions, whereas *Dodds and Rothman (2000)* developed the forms of the distributions.

Comment

The reviewer strongly suggests dividing section 2 and 3 in subsections.

Response

For section 2 we have divided into 3 sections: Geometry, Area, Length. For section 3 we have divided into 2 sections: Hydraulic distributions and Sampling.

Minor Comments

We have accepted the minor comments.

2 Response to Reviewer 2

We thank the reviewer for a careful reading and critique of our manuscript. The reviewer correctly points out some issues that may be altered to improve a revised draft. In our view, the reviewer's major contribution here highlights that we should clarify the reasoning for our initial and boundary conditions that determine the probability function for watershed width. Further they suggest that we clarify that our theory is based on a continuous random variable, whereas the Scheidegger model has discrete random variables.

Comment

Starting from Eq (5), this equation does not read as it meant "a unity probability with zero variance". It should be expressed as an atom of probability at $w = 1$ and zero for other w . Second, should it be $w = 0$ instead of $w = 1$? Right above this paragraph, it is stated that watersheds are closed at $w = 0$; thus they should also initiate at $w = 0$.

Response

We have changed language in the manuscript to highlight that the value of the Scheidegger network is to inform the probabilistic elements of constructing watersheds and the probability functions of geometric variables. Therefore, the discrete nature of the Scheidegger model is only to guide the mathematics for a continuous version. We agree with the reviewer that the initial condition should be represented as a dirac function and have made that change.

We disagree that the initial condition should be $w = 0$ for the following reason. It is true that at $s = 0$, the watershed has a width $w = 0$. However, in order for the watershed to exist, it necessarily must widen to a width of $w = r$ at $s = 1$. This is where our initial condition applies and we have clarified this in the text

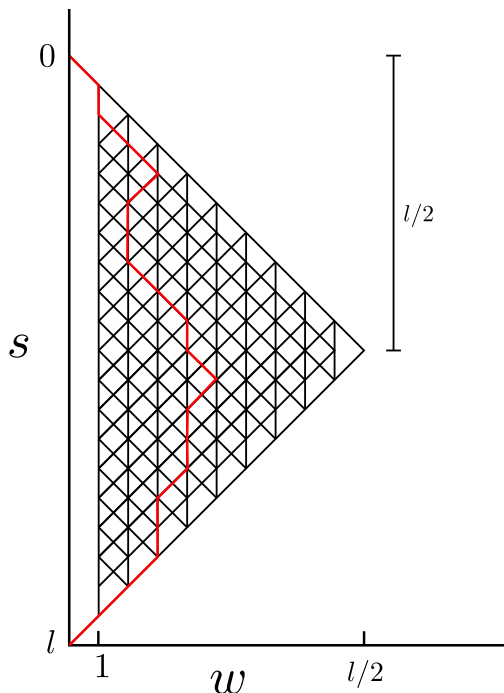


Figure 1: Conceptual diagram showing all possible paths of the width function for a watershed of length, l . The red line is one realization. The ensemble of paths is symmetrical about $l/2$ and the boundary condition is illustrated by no paths reaching $w = 0$ before l .

Comment

The reviewer has issue with our boundary conditions, which we state are fixed boundary conditions at $w = 0$. They suggest that given that it is possible for a watershed to close, that there is finite probability that $w = 0$ and therefore the boundary condition cannot be $f_w(0) = 0$.

Response

We understand the reviewers comment and how the previous manuscript led readers to that conclusion. However, we are confident in our boundary condition and have added language and a figure to explain why. Consider the ensemble of watersheds of length l . The width function, $w(s)$ can take any random walk over the domain $[0, l]$, but it must begin and end at $w = 0$. At $s = 1$ the variance of widths of the ensemble of watersheds of length l is 0. When $s = l/2$ the variance is at a maximum. While $s > l/2$ and as $s \rightarrow l$ the variance must begin to decline back towards 0. Therefore, the evolution of $f_w(w, s)$ from 0 to $l/2$ is mirrored by the evolution from $l/2$ to l . Our boundary condition only applies to the $s \leq l/2$, or the ‘growing’ part of the watershed. The mirroring about $l/2$ then allows for us to know the form of the distribution over all $s < l$. The calculation of μ_A (Equation 10) involves multiplying the integration of μ_w by 2 to account for this.

Comment

The reviewer is under the impression that we have fixed l .

Response

This is, in fact, the case for describing the conditional distribution $f_A(A|l)$. The response above addresses this issue and illustrates how a diffusion equation is appropriate for the description of $f_w(w, s)$. We consider l as a random variable for constructing the entire

Comment

The reviewer suggests that we remove semi-colons when referring to the spatial variable in the Fokker-Planck like equations.

Response

We have replaced these with commas. However, we have also removed most references to a Fokker-Planck-like equation to avoid confusion.

Comment

The reviewer got the impression that at some point we were equating Scheidegger models and optimal criticality networks.

Response

We are uncertain about where the reviewer got this impression. We refer to OCNs in two places in the paper. The first instance, we say “Other network classes exist including optimal criticality networks...” We think that this clearly states that they differ from Scheidegger networks.

The other location that we refer to OCNs is in the discussion where we highlight the slight mismatch in form of the probability functions of area for real topography and a Scheidegger network. We have added language to make sure that readers do not think that we suggest Scheidegger and OCN are the same:

“However, because those networks are not amenable to the type of theory developed above because they lack the clarity in rules for links and nodes of the network. The Scheidegger model serves as a guide to inform probability distributions and provide a basic reasoning for nonlinear relationships.”

Comment

Unclear how equation 11 is obtained, is it empirical?

Response

This result is semi-empirical. We expect that the variance of A scales as $\sigma^2 l^3$. The presence of 6 in the denominator is not theoretically derived. If one could formally explain this, then the problem would be complete. We have added:

“We emphasize that this is a semi-empirical result that warrants a stronger theoretical solution”

Comment

Rill flow length needs to be defined in the intro

Response

We have changed the appearance of “rill flow length” and now only have “watershed length” which is defined in the intro.

Comment

Clarify how rills are “efficient”.

2.0.1 Response

We have removed the word “efficient”

Comment

Unclear: “second we ask if the particular arrangement of the rills focuses flow such that it leads to a nonlinear sediment yield relationship...”

Response

We have reworded this to: “Second, we ask if a well defined network of rills focuses flow such that it leads to a nonlinear sediment yield relationship”

Comment

OCN: Optimal Channel Networks?

Response

YES, we have changed.

Comment

Line 75- “As such, . . .”: Is this referring to OCN? What is the “constraint”? OCN has a “clear rule” to construct networks.

Response

Yes it is referring to OCN, we think this is sufficiently clear. The constraint is that it minimizes the energy expenditure as stated above. We disagree that OCN has a clear rule for constructing networks that leads to clear probabilistic insights. That OCN minimize energy and satisfy continuity equations does not lead to particularly clear conclusions regarding the construction of links and nodes. In contrast, the Scheidegger model is constructed by a set of uniformly spaced paths that take simple random walks in the cross slope dimension. Second, uniform drainage density is maintained so that when two rills meet, another is formed. These lead directly to a graphical representation of the network in a way that the rule for OCN does not. We do not think that further clarification benefits this manuscript as the focus is not on OCN.

3 Reviewer 3

We thank reviewer 3 for the meaningful review and suggestions on this manuscript. We believe that the manuscript is improved after following many of their suggestions.

Comment

... Since the primary goal is to demonstrate and understand the non-linearity of the length-sediment relation for a rilled surface, it would have been useful to have first discussed the relation for an unrilled surface...

3.0.1 Response

We agree with the reviewer and have added a paragraph dedicated towards addressing this. We review the work of Burch et al., 1986 which demonstrates nonlinear sediment yield-slope length relationship on planar surfaces. Their work suggests that the nonlinear relationship for planar and unrilled slopes approaches the least nonlinear relationship observed for rilled slopes. We suggest that this similarity is shared between planar surfaces and rill networks with linear rills. We then review the work of *McGuire et al., 2013*, which demonstrates that rill networks become increasingly dendritic with increasing rainfall detachment. This is the scenario that leads to more nonlinear sediment yield relationships.

Comment

It would be useful to make a little section for related literature, and especially to add a comprehensive summary of all of the relevant findings of Dodds and Rothman since they are employed/referred to so extensively.

Response

We have added a paragraph in the introduction summarizing these results and how we build on them.

Comment

The section on network geometry could be greatly improved for ease of understanding and readability by shortening and simplifying. It would be simpler to state Hacks law in deterministic form, state its inverse in deterministic form, and note that the goal is to derive its probabilistic representation from the theory of random walks in a Scheidegger network. The derivation of equation 7 as currently written, could then be simplified and disambiguated by

- noting that width is the difference between two normally distributed and independent RVs and must be positive;
- stating the ICs at $s = 0$, presumably as two rill sources separated by a distance $w = w_0$;
- noting that the distribution of the difference of two independent normal distributions $z_1 - z_2$ with means (μ_1, μ_2) and variances (σ_1, σ_2) is itself a normal distribution with mean $\mu_1 - \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$;
- noting that the first passage of a random walk with a normally distributed RV with diffusion coefficient D starting at $w = w_0$ is given by the Rayleigh distribution

since, as written, the description is not clear, equation (5) is confusing, while the brief discussion of the diffusion (6) and Fokker-Planck equations will not be helpful to many readers.

Response

We have largely taken these suggestions. We have clarified the initial conditions and added a figure that explains them and the boundary conditions. In our case, the Rayleigh distribution represents the distribution of watershed widths for $s \leq l/2$, which we have clarified in this case. We disagree though that $w(s)$ is a first passage problem because a watershed can return to the same value for w at many points along the watershed. We have; however, chosen to remove most references to the Fokker-Planck equation for simplicity.

Minor Comments

1. Line 12 understanding Hack's
2. Line 34 geometrical (not topological)
3. Line 169
4. Line 197 the this
5. Line 286 , for ;
6. Line 330 length(en)
7. Line 343 geometrical (not topological)
8. Line 358 distribution a

Figure 12B needs some attention. Its not clear what its meant to be (as seen by this reviewer.)

Response

We have accepted most of the minor comments that identify typos or improved language. Figure 12B is included to illustrate the results from flow routing. We have added the following language:

“This illustrates the results of numerical flow routing. From this result, we calculate exceedance probabilities that compare to theoretical distributions.”

Response to longer term improvements

We appreciate the reviewers suggestion regarding future developments and their support for further work. Indeed some form of what they mention is the subject of current work. We choose to keep the focus of this manuscript largely the same as in the previous manuscript. However, we respond to a few points here.

First, we address their suggestions for a second paper and uncertainty with regard to a Scheidegger model being suitable for understanding Hack’s Law. We understand that the reviewer thinks that the Scheidegger model is likely a good model, but not necessarily the best because two adjacent rills that define a ridge may not be the ones that meet to close it off. This is true, however, the observation that watershed divides take simple random walks and necessarily begin at $w = 1$ at $s = 1$ holds. In fact, the ridges form a network that is also a Scheidegger network and is simply the complement of the channels. We think that insofar as the Scheidegger model is a simplification of networks, it is a reasonable one for exploring Hack distributions and laws.

Second, we agree with the reviewer’s comment that our representation of channel widths is somewhat simplistic. One could reasonably make r_w a random variable that scales with Q and we think this would be a fruitful next step, but beyond the scope of this paper.