Revisions made to Struble and Roering based on comments by Simon Mudd. *All line number references refer to numbering in original manuscript (as referenced by reviewer).*

<u>AE, Mudd:</u>

Overarching comments:

C_{HT} isn't that bad at E*=10: For D=0.003 m^2 yr^{-1}, perhaps not. But for lower diffusivities, it becomes more problematic. We've made adjustments to Figure 11A to emphasize this. We'd also refer you back to earlier changes we made to Figure 11B which demonstrate this too.

Color Figure 11A by E*: Unfortunately, that would simply result in a linear increase of E* parallel to erosion rate on the x axis. Plus, it would corner us even more into discussing a single diffusivity. Instead, we've opted to expand Figure 11A to include more than a single diffusivity to demonstrate that the deviation between expected and measured C_{HT} can occur at slower erosion rates and depends strongly on D.

Figure of C_{HT}/known C_{HT} as function of E*: That is basically what Figure 11A is now. It now allows us to more clearly interrogate D, as opposed to it being hidden within E*.

Sentence with major disagreement: We acknowledge that "likely" is a bit strong, so we've changed it to "may be." Again, though, D is hugely important. If we use a smaller value, closer to D=0.001 m^2 yr^{-1}, we end up seeing deviation between known and measured C_{HT} >10 percent at erosion rates of 0.04 mm yr^{-1}. Does that mean all of the measurements in Gabet et al. are underestimated? Absolutely not. But if you don't know the diffusivity *a priori*, then it is much harder to know how drastically C_{HT} is underestimated from measurements of C_{HT} and E. Unfortunately, for moving forward, this becomes circular ("you need to know diffusivity to accurately measure C_{HT} (or at least know which erosion rates are dangerous for underestimation) to estimate diffusivity"), but that's where we're at.

"I think what is supported by your paper is that there is a high risk of data artifacts at E*>10, and that could lead to a spurious E vs D relationship with an exponent less than 1. But I think your paper also shows that, given the erosion rates of sites in our paper, the square root relationship is *unlikely* to result from data artifacts alone. Do you disagree with this conclusion?"

Lots to unpack here. First off, we need to be clear here about what "data artifacts" are. We aren't simply discussing issues with grid spacing and data resolution. This is a geometric problem with the inherent shape of hillslopes. That is, if hillslopes have any kind of planarity near the hilltop, mathematical functions are unable to extract an accurate C_{HT} while *simultaneously* accounting for topographic roughness (both real landscape roughness and that introduced from *actual data artifacts*.). We think you know this, but it is a crucial point. We've made a few minor adjustments throughout the paper to hopefully above confusion about whether artifacts are of geometric origin or from data resolution.

To address your actual question though: per our previous comment, without knowing the diffusivity at the sites in Gabet et al. *a priori*, we really can't answer that question. It may be that the erosion rates at those sites are slow enough and diffusivity is sufficiently high to make the square root relationship unlikely to be spurious. It may also be, however, that diffusivity is lower than you think at some (doesn't have to be all) sites, resulting in an underestimation of C_{HT}, which then leads to an underestimation in D. So, we sadly can't answer your question, as our data only emphasize that D is important (I think we can all agree on that!).

Our reanalysis of your sites: Again, we agree that "likely" is strong and premature given our data, but the importance of D that our data demonstrate reinforces that the square root relationship, from our perspective, is not a clear observation clearly free from bias by planar hillslopes. Hopefully, though, this will just motivate future studies!

Line 51: Removed redundant phrase, "a nonlinear formulation implied…"

Line 175: Yes, that is technically correct. We have reordered material in this section to first introduce this property of convolutions, and then dive into the Ricker wavelet. This way, we have provided a general framework for utilizing the wavelet transform, and then we can discuss the specifics of DoG wavelets. It also avoids most of the confusion around the second derivatives, as we now discuss the convolution property as simply for taking derivatives and not just for calculating curvature.

Line 184: Yeah, that's right. See response to previous comment. Isn't that cool, though?!

Line 189: Yes, λ has dimensions (and so does *s*, since it is the standard deviation of the Gaussian (for the zeroeth order DoG)). To be honest, we've rarely (if ever) seen any papers report units of a wavelet basis function. This may be because most papers utilize wavelets in 1D on temporal data and consider them strictly in the frequency domain. That said, yes, if you follow the units, ψ should have units of $1/L^4$. Throughout the paper, we've added unit designation for each variable where appropriate.

Additional change at Lines ~185-190: To address Tyler Doane's continued concern about the way we are defining the two different λ, we have added some detail on how both Lashermes et al. (2007) and Torrence and Compo (1998) define λ. We've discussed this with Tyler, and our understanding of his concern is that because λ is different for the two definitions, then the wavelets are different and, therefore, the outputs aren't comparable. Granted, the wavelets are "different," but only in as much as two parabolas with a different coefficient are different. The functional form is the exact same; we simply use a different definition for the scale, which "stretches" the wavelet and considers topography with a different characteristic wavenumber. The outputs are certainly comparable, since we are calculating curvature for both. Otherwise, we wouldn't be able to compare to the output from the 2D polynomial either. We have added brief statements that clarify the methods by which Lashermes et al. and Torrence and Compo define λ. We personally find these definitions to be a bit superfluous and unnecessary to understand that λ dictates the smoothing scale of the wavelet (and we fear this new added explanation may confuse this understanding for some), but we acknowledge that some readers may appreciate this specificity.

Line 195: We do not use the wavelet to define the smoothing scale for the polynomial. We independently define the values of λ ahead of time, with the only stipulations being that $s>1$ for the wavelet (since the convolution can't be applied at scales smaller than the grid spacing; translates to $\lambda>5$ for 1-meter data) and that λ be odd to accommodate the window-size (diameter) requirement for the polynomial. Thus, we applied both the CWT and PFT for odd $\lambda>5$ (at irregular intervals for computations efficiency at larger λ; see Figure 3). We've added some text back at the start of section 3.1 to hopefully clarify that the PFT is applied for a range of λ

independently from the CWT. (We also clarify that λ for the PFT is the diameter of the smoothing window).

Lines 232-234: We adjusted language here to avoid confusion about hillslope length vs the "length" of the hilltops (how far *along* the hilltop we measure C_{HT}).

Line 239: Correct. We now specify this (that we use diameter here) more clearly in section 3.1. But you are correct in that Roering et al. report the value as a radius of 7.5 m.

Line 275: Added some sentences that give a bit more detail.

Lines 399-403: Thank you!

Line 474: Removed modest; though, we maintain the rest of that sentence.

Line 494: Unfortunately, no. It all depends on the curvature of the noise signal. For instance, the overestimation of curvature at modest smoothing scales ($\lambda \approx$7-15m) in some cases (Figure 10C, D, G, H, I) corresponds to locations where negative curvature with a higher magnitude than the underlying hillslope form biases the value to more negative values (some underestimations may be greater than they would be without noise due to a similar issue). An early version of the manuscript discussed this more, but we felt it distracted greatly from the more notable underestimation seen at most smoothing scales, and importantly, when no noise was added. Overall, this is more of a question regarding topographic noise, which for our purposes here, is poorly constrained and is the subject of future papers. We've added a statement to this effect.

Line 567: Figure 11A is diffusivity dependent, which is why we originally added Figure 11B. If the diffusivity is lower, then the erosion rate at which this deviation occurs is also lower. So, to help make this clearer and to avoid misinterpretation of Figure 11A as being the absolute erosion rate at which this problem manifests, we've adjusted Figure 11A to also include a range of diffusivities (0.001-0.003 m^2 yr^{-1}). This is certainly not an exhaustive examination of the range of diffusivities found in natural landscapes. But it demonstrates that if you visit a landscape where diffusivity is low, then you are *more likely* to underestimate C_{HT} at slower erosion rates. We acknowledge that given more work is needed to clarify this problem, our language is perhaps a bit strong. We've modified "likely" to "may be." We think this is very reasonable given the analysis we've completed.

Fig. 1: Added this statement.

Fig. 2: Font size adjusted.

Fig. 3: Adjusted font size (in this figure and others). Added L07 and TC98 definitions.

Fig. 5: Added L07 and TC98 to appropriate figures.

Fig. 11: Since we chose a range of E*, then yes, hillslope length is relevant. But if we want to utilize Equation 1 ($E = -\frac{\rho_s}{\rho_r} D C_{HT}$), then we want to isolate the C_{HT} term. The hillslopes are 100 m long. We have added that information in the same parenthetical where we specify the range of diffusivities.