Response to comments on the manuscript entitled ‘Impact of Noise on Landscapes and Metrics Generated with Stream Power Models’

M. J. Morris, G. G. Roberts

Introduction

We thank both reviewers for their assessment of the manuscript. We respond to all of the reviewers’ comments (in bold) below.

Response to comments from RC1 (25 June 2025)

“The authors conduct a thorough exploration and review of the influence of noise on landscape evolution. The manuscript is well-written and easy to follow; I particularly appreciate the explanations they provide throughout the paper to provide the reader with enough context to understand some of the complex material they cover. The motivation for the paper is clear and well justified. The figures illustrate their points well; I wish more papers utilized the style and clarity of Figure 3, which is worth a thousand words (or more). I think the authors use a thoughtful approach when choosing the different types of noise and model configurations. They cover different types of noise by choosing the end-member models of blue and red noise. Within these end-members are white noise, which is commonly used by landscape evolution researchers. The different types of model configurations (i.e., square, domal, escarpment) also represent the range of configurations most modelers use. I found the results and discussion about how noise affects the landscape metrics useful and thought-provoking. The metrics are typically straightforward when used on numerical results and harder to interpret for natural landscapes. By incorporating noise and multiple ensembles in their numerical experiments, the authors provide a great framework for interpreting the uncertainty seen in natural landscapes. I agree with the authors’ recommendation to use a more probabilistic/ensemble approach to landscape evolution modeling. I support this paper for publication in its current form, but would appreciate it if the authors incorporated some of the feedback below.”
We thank reviewer RC1 for their encouragement and the comments that follow, which we respond to below.

“Line 107 – 100 × 100 cells. Do you think this is a large enough grid to capture enough complexity in the drainage network? By eye, it looks as though the maximum stream order in the drainage networks is around 3 or 4.”
This concern is probably most straightforwardly addressed by re-running the ensemble of simulations with many grids of different resolution and spatial scales, which we think is largely beyond the scope of this paper for reasons we discuss below. Nonetheless, we have performed a few preliminary tests to explore this concern with results shown in Figure 1 of this document. Before examining those results we note that a single simulation with cell sizes of 1×1 km and 100 Myr of modelled time takes approximately 90 seconds to run using a desktop computer with an Intel i7-6700 CPU and 64 gigabytes of RAM for a 100 × 100 km grid, increasing the grid size to 400 × 400 km grid requires approximately 17 minutes of runtime (ensuring that the CFL stability conditions are satisfied). Given these practical considerations and our goal to develop statistical insight we think that experimentation using a 100 × 100 km grid is reasonable but readily acknowledge that

experimentation incorporating other scales would be valuable future work.

The additional tests we performed used a 400 × 400 km grid, again with cell resolution of 1 × 1 km, and were designed to reproduce the tests shown in Figure 6 of the main manuscript. Due to the change in number of grid cells, the inserted noise must be different between the two sets of experiments. Figure 1 of this manuscript shows the complexity of calculated drainage planforms on the larger grid presented with the same formatting as Figure 6 of the main manuscript. The drainage networks on the larger grids have streams with orders > 4, as expected. The networks have broadly similar arrangements to those on the smaller grid, in that there are a few large basins and flow pathways are broadly similar, or at least appear to be no more dissimilar than planforms generated with different arrangements of noise. The specific arrangements of the planforms, and to some extent their complexity, depend on the arrangement of inserted noise, the boundary conditions and uplift.

Establishing the impact of noise on the distribution of stream orders, including at different scales (e.g., at larger spatial scales or with finer grid resolution), would be valuable future work. Comparing the spectral power of flow pathways might also be useful for generating deeper insight into (and quantification of) the role of noise in generating complexity or simplicity in drainage planforms (e.g., Lipp and Roberts, 2021).

"Figure 1 – Do you expect your result to change much if you opt to not fill in the sinks?"
We have addressed this comment by regenerating model ensembles 1, 4, and 7, excluding the initial sink-filling step. Landscapes generated with and without sink filling are different of course (see Figure 2 of this document). Omitting the filling of initial sinks prevents the early formation of continuous flowlines from the interior to the exterior of the domain, and thus regions of internal drainage are produced. Due to the absence of initial sink-filling, early networks tend to be shorter than those in landscapes with filled initial sinks. A consequence of not filling sinks is that it takes longer for rivers to reach the centre of the domain (see Figure 2 of this document). However, after 100 Myr of evolution, there appears to be little difference between sink-filled or unfilled landscapes. Thus, in consideration with the following additional figures, we suggest that the statistical results presented in the main manuscript are largely unaffected by sink-filling.

Figure 3 shows drainage probabilities, identical to Figure S6a-c within the Supplementary Information, with the modification that the landscapes have been generated without an initial sink-filling step. Comparing the two demonstrates that neglecting this step produces, in general, drainage networks which are more direct than those which have been produced after an initial filling of sinks. However, the remaining patterns of drainage probability (i.e., broadly equal probabilities to four edges, low probabilities within a central cross) are consistent with those in Figure S6 and Figure 9 of the main manuscript.

Figure 4 of this document summarises the variability of drainage basins areas, akin to Figure 8 of the main manuscript. It was also produced using the 100 simulations without sink-filling. The results described in the main manuscript are also observed here. The same is true of Figure 5, which summarises metrics for the 'square' landscapes generated with no initial sink-filling (cf. Figure 12 of the main manuscript).

"Line 118 – "100 MYR" – This seems excessive? With an uplift rate of 0.2 m/kyr, one relief unit (∼1200m) would take 6 MYR to erode through. Typically, landscapes reach a quasi-steady state around 2 relief units, but 100 MYR is ∼17 relief units. It might be useful to describe the model behavior (Line 416 to 423) in terms of relief units to relate the model parameters to the time."
Our motivation for using a model runtime of 100 Myr is to ensure that models can reach a topographic steady state ($\partial z/\partial t = 0$; or as close to steady state as can be expected). We agree that 100 Myr of modelled time could be regarded as being excessive for most models examined in this study if a user were looking to optimise computation time. For instance, as the histograms in Figures 5, S7 and S8 show, there is little change in the distribution of elevations of the models between 15 and 80 Myr. Moreover Figure 6 of this document—a

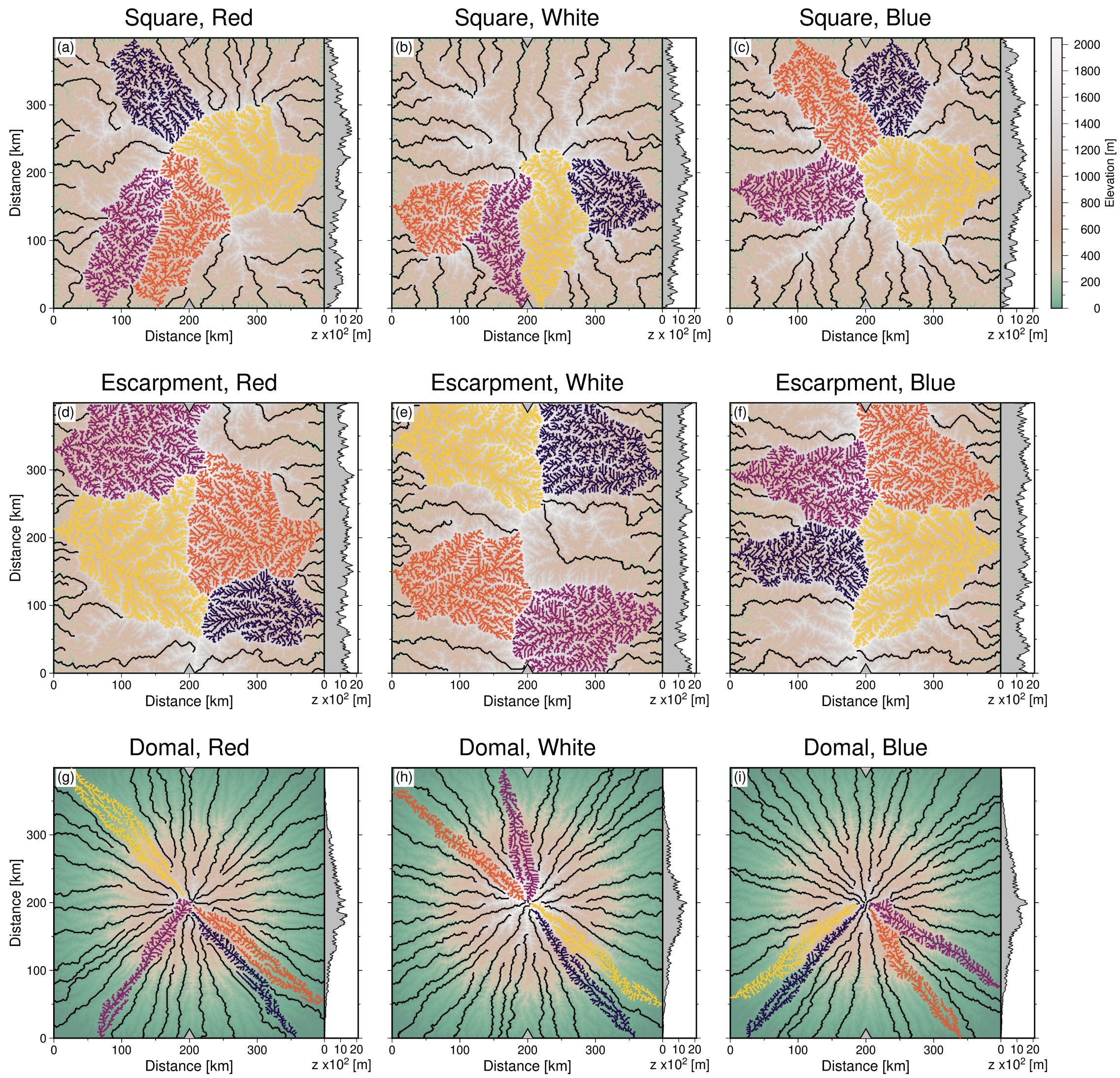

Figure 1: **Steady state landscapes generated with different initial conditions using a 400 × 400 km grid.** (a–c) Steady state 'square' landscapes generated with red, white and blue initial conditions, respectively. Planforms of channels in the four largest drainage networks in each landscape are emphasised by the coloured lines. Main (longest) channels of 40 largest basins shown in black lines, plus the four longest coloured lines. Adjacent topographic transects are between grey triangles at $x = 50$ km. (d–f) As above, for 'escarpment' landscapes. (g–i) As above, for 'domal' landscapes (cf. Figure 6 in the main manuscript).

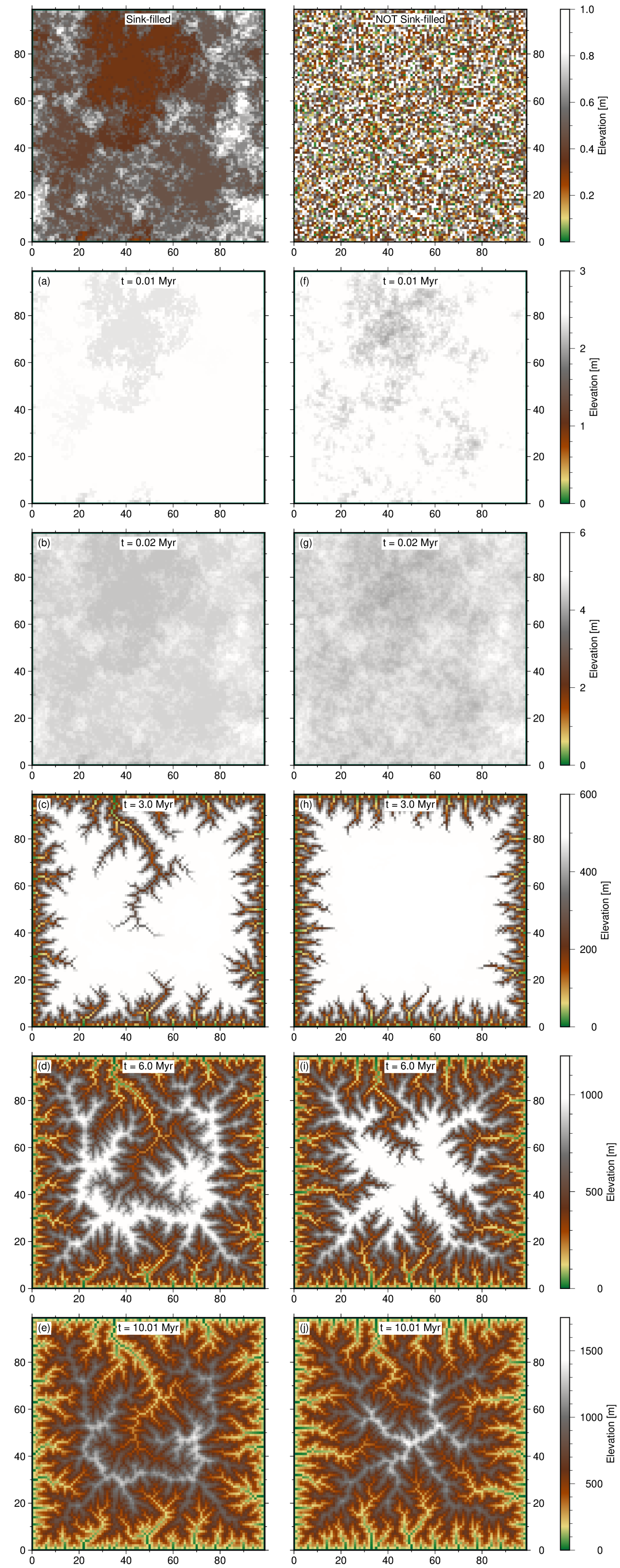

Figure 2: **Evolving landscapes with and without initial sink-filling.** Left column shows an initial condition which has been sink-filled, and a series of juvenile topographies. Right column shows the initial topography without sink-filling, and the resulting topographies. Note changing colour scales.

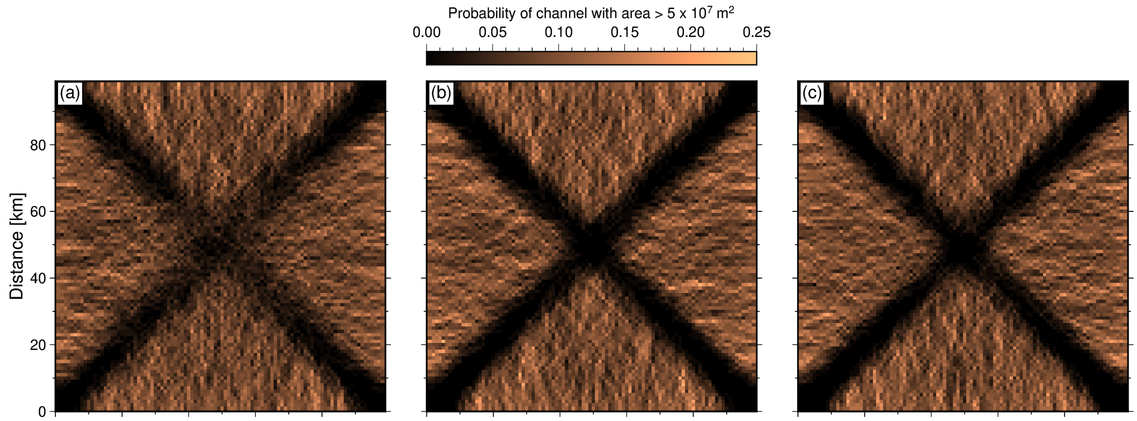

Figure 3: **Drainage probability maps for landscapes with no initial sink-filling. See body text for details.**

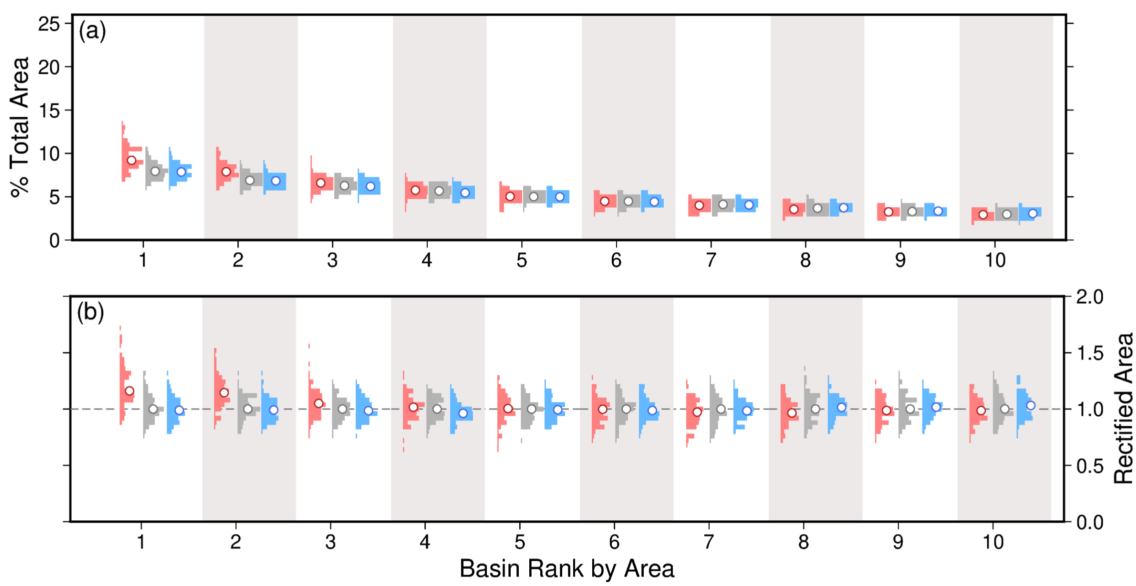

Figure 4: **Variability of drainage basin areas due to noise in landscapes without initial sink-filling. See body text for details.**

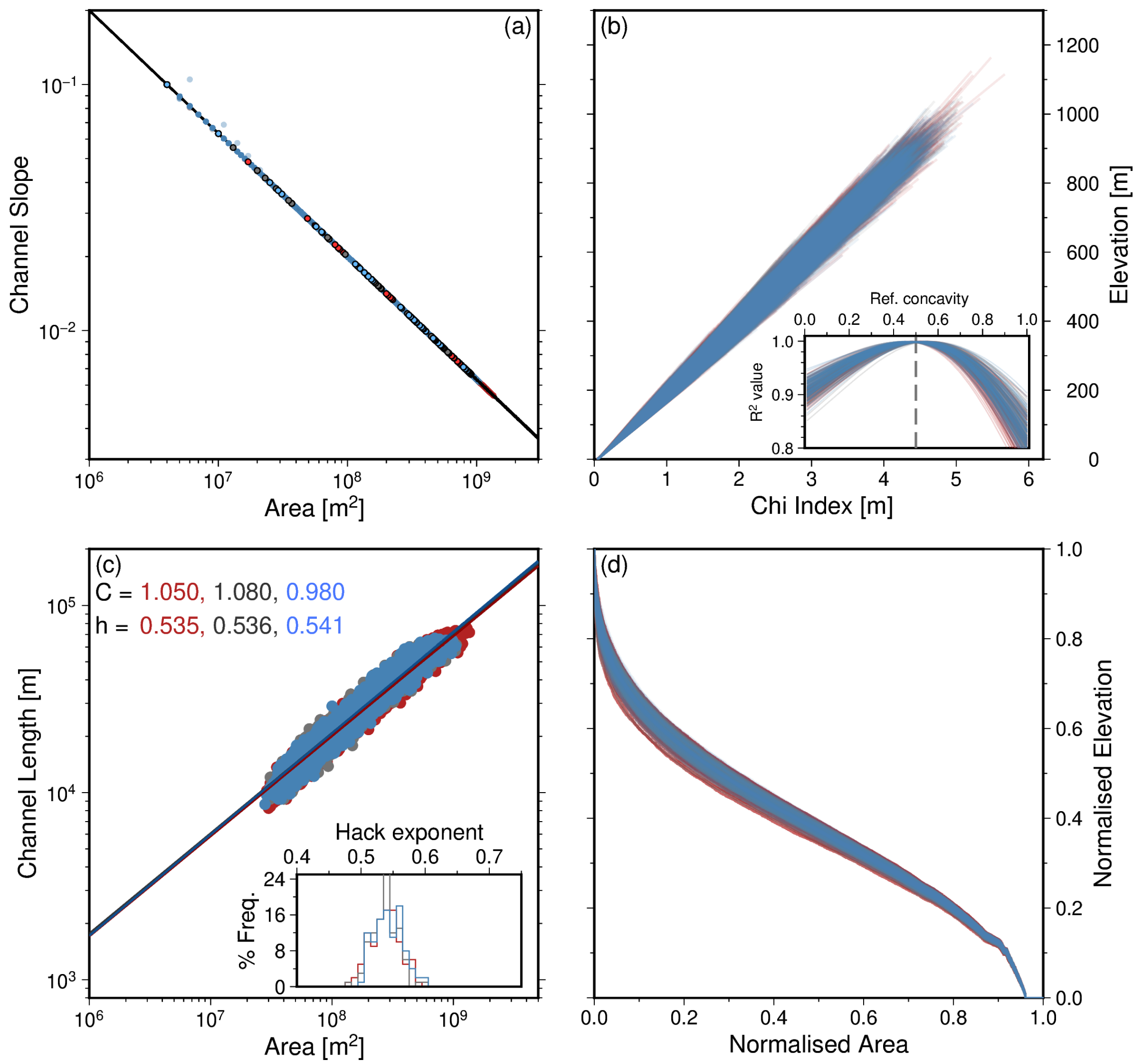

Figure 5: **Geomorphic metrics for 'square' landscapes with no initial sink-filling. See body text for details.**

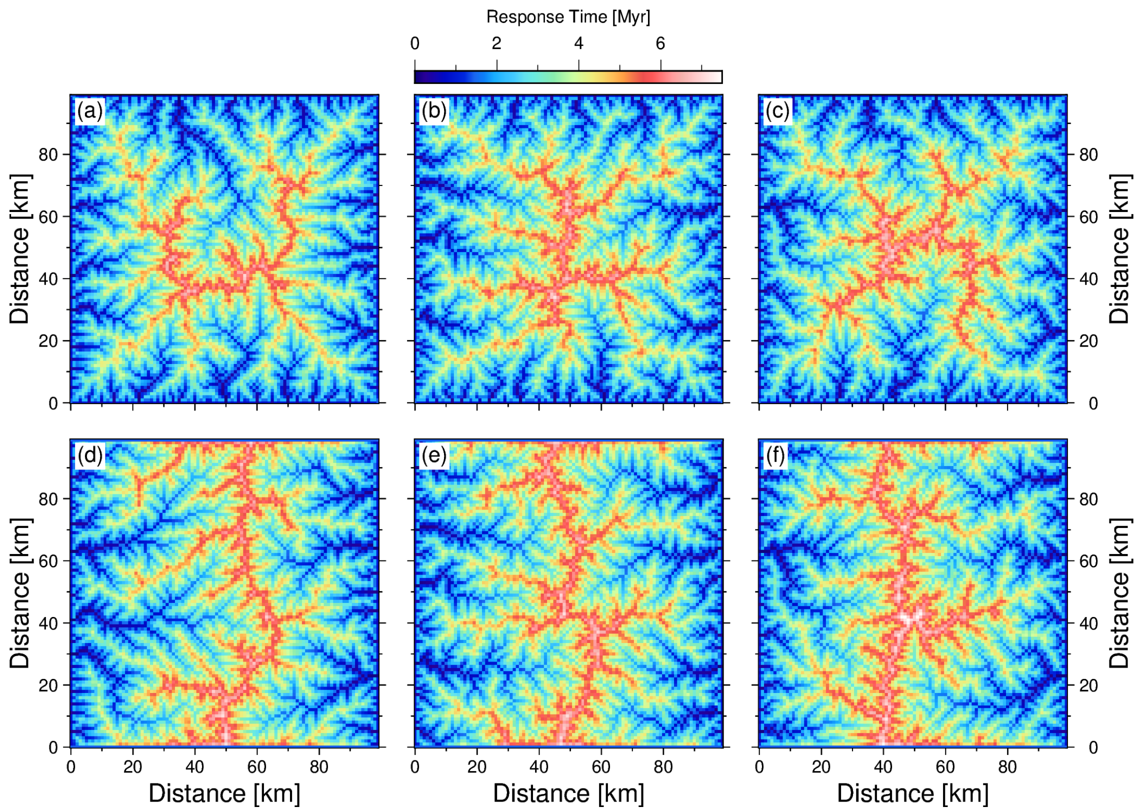

Figure 6: **Landscape response times, τ, for typical 'square' (a–c) and 'escarpment' (d–f) landscapes.**

map of response times for the landscapes shown in Figures 5a-f of the main manuscript calculated such that $\tau = \int_{\circ}^{x} \mathrm{d}x/KA(x)^m$, where x is streamwise distance—further emphasises the reviewer's statement regarding time to steady state. We note that we do not show response time maps for the domal landscapes here because of the spatially varying nature of uplift meaning assuming that tectonic signals propagate from the edge of the domain to fluvial headwaters is not reasonable. Nonetheless, we feel it is useful to run the models out to 100 Myr when assessing, for example, the impact of quenched and especially spatio-temporal noise.

"Line 306 – typo – "river being in a cell is calculated by DIVIDING the number of times""
Now corrected from 'divided' to 'dividing'.

"Line 560 – Why did you decide to add noise that is 0 to 1 m instead of -0.5 to 0.5 m? While I think it is easy to follow when the noise is averaged, it essentially adds ∼0.5m per time step (25% of the uplift rate). It would be easier to make your points about uplift recovery on Figure 17g and 18g if it averaged around 0m. Is it another way to interpret Scenario B and C as a noisy constant spatial uplift rate from 0.2 to 0.25 m/kyr and a noisy spatiotemporal uplift rate from 0.2 to 0.25 m/kyr, respectively?"
We chose to add noise with elevations 0 to 1 m to avoid elevations becoming negative due to noise (e.g., early or late in the simulations when elevations can be low), which we think helps to keep the numerical modelling relatively simple, tending to avoid, for instance, formation of internal drainage and need for additional sink filling steps. As you say, noise added to Scenarios B and C could be interpreted as insertion of noisy uplift. Nonetheless, to address your comment, we now incorporate results within the Supplemental Information Figures S9–10 from a series of models in which quenched and spatio-temporal noise (Scenarios B and C),

were added within the range −0.5 to 0.5 m (see also hashed polygons in the revised Figures 17 and 18 of the main manuscript). The results from these models are similar to those with noise in the range 0 to 1 m. We note that (as expected) recovered uplift rates are centred on the true value of 0.2 m/kyr for uplift rates estimated from χ analyses when noise is quenched or spatio-temporal, and also for uplift rates estimated from slope-area analysis when spatio-temporal noise is present in the models. We discuss these results, and the difficulties with choosing appropriate mean values for noise, alongside existing results for Scenarios B and C in the revised Sections 3.2 and 3.3 of the main text.

"Figure 17b and 18b – It seems like the relief of the landscape is around ∼1250 m from the color scalebar. This also seems constant across Scenario A, B and C. Shouldn't the relief in Scenario B and C be higher because of the added noise component. Since 0.5 meters, on average, are being added every timestep, 10 kyr, the apparent uplift rate should be 0.25 m/kyr. If relief scales linearly with the uplift rate, then landscapes in Scenario B and C would be about 25% taller, right?"
Yes, the landscapes generated in Scenarios B (quenched noise) and C (spatio-temporal noise) have higher maximum elevations than their counterparts in Scenario A. We have amended the colour scale in Figures 17 and 18 to address the reviewer's concern.

"Line 613 – Which supplementary information figure are you referring to?"
We have clarified this line to: *"In fact, even 20 m of added noise can result in calculated uplift rates being unreliable (Figure 20b; Supplementary Information Figures S11-13)."*.

"Line 667 – I found this result so strange that it was not just offset by the mean value of the noise addition. I was excited to read about this in the discussion (mentioned on line 579 to 580), but I was a little disappointed by the explanation here. I think Scenario C had an offset because the apparent uplift rates varied around 0.25 m/kyr due to the spatiotemporal noise configuration. Over enough timesteps this should average to 0.25 m/kyr in most locations. In contrast, this averaging does not happen in Scenario B. Do you think this might have something to do with the disparity?"
We too found this result strange when it was first produced, and further considered its cause following your comment. As you say, the uplift rates calculated in scenario C appear to have a fairly simple explanation (i.e., spatio-temporal noise averaging to giving an additional 0.05 m/kyr of expected uplift, in addition to the 0.2 m/kyr uplift rate inserted at each time step at every non-boundary cell). In contrast, uplift rates calculated for scenario B, with quenched ('frozen') noise up to a maximum elevation of 1 m added at each time step (i.e., up to 0.1 m/kyr) are not centred on 0.25 m/kyr, and some are larger than 0.3 m/kyr. This result is perhaps confusing, given that the maximum possible 'uplift rate' imposed for any grid cell is 0.3 m/kyr (i.e. 0.2 m/kyr 'true' uplift rate + 0.1 m/kyr of noise). We suggest that these results are principally a consequence of the sensitivity of the y-intercept (from which uplift rates are calculated) generated from the linear regression of log(slope) - log(area) data. The presence of quenched noise in models tends to generate considerable spread in slope-area data, and thus in calculated intercepts (and uplift rates). We note the R^2 values for models with quenched noise ($0.985 \leq R^2 \leq 0.995$) are typically lower than those with just initial noise ($R^2 > 0.996$).

To further illustrate these results, Figure 7a–c of this document shows slope-area relations for four rivers and their tributaries (extracted from the four largest drainage basins) for three select models generated with quenched red, white or blue noise, and their associated linear regressions. Values of θ (gradient of regression lines) and U (derived from y-intercepts of regression lines) for those models and the other 291 (97 for each colour) generated with quenched noise are shown in panels d–f. The positive correlation between θ and U emphasises the spread in slope-area data especially for smaller areas and larger slopes when quenched noise is present. These results further emphasise the challenges associated with interpreting slope-area data in the presence of noise, which we expect to always be the case in the real world. We have modified the discussion

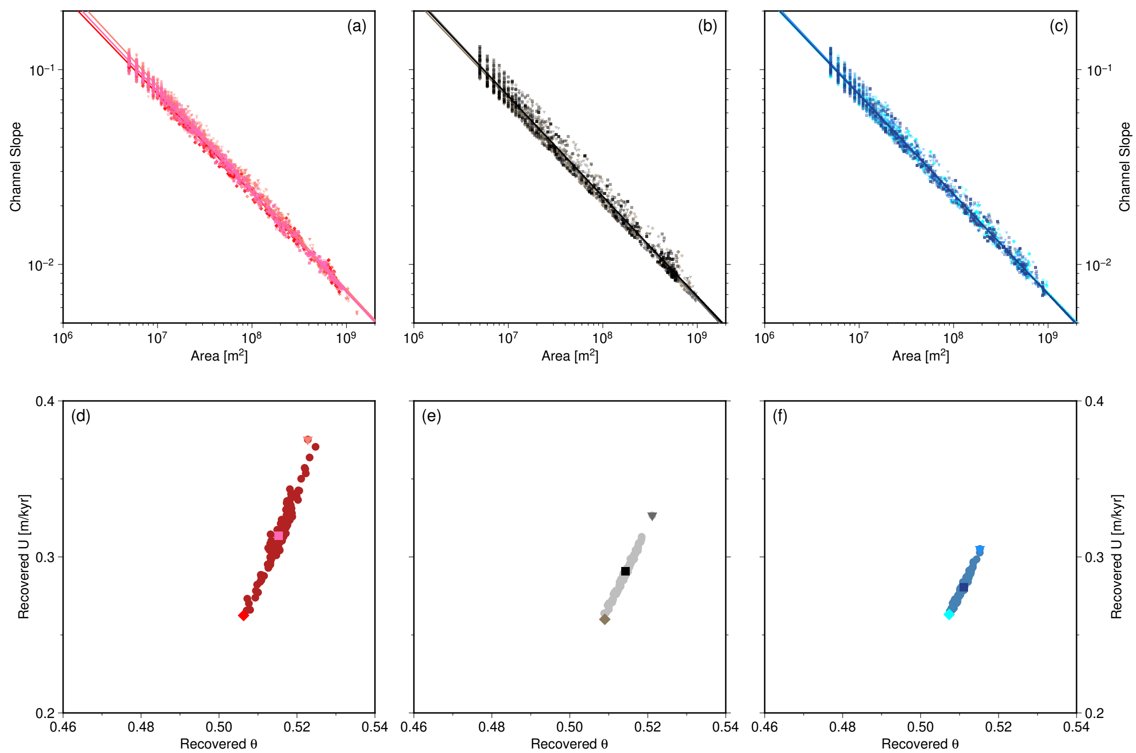

Figure 7: **Trade off between recovered parameters from slope-area data.** (a–c) Slope-area data from three selected simulations containing quenched ('frozen') red, white, or blue noise, respectively, where recovered uplift rate is low, intermediate, or high within the distribution of recovered rates. (d–f) Each panel shows the relationship between the value of θ and uplift rate, U, calculated from the best-fitting regression line of the 100 models containing quenched ('frozen') red, white, or blue noise respectively within Scenario B; i.e., the regressions shown within Figure 17f of the main manuscript. Data from selected models within panels (a)–(c) is highlighted with respective colours/symbology.

in the main manuscript to emphasise these results.

"Line 719-720 – Can you cite Kwang et al. 2021, who show how numerical models remove memory of initial conditions with lateral erosion."
Done.

Response to comments from RC2 (4 August 2025)

"This manuscript presents a diverse series of numerical experiments, designed to explore the influence of initial conditions (eg topographic noise) on topographic metrics extracted from DEMs generated using typical landscape evolution modelling approaches. The authors do an excellent job in drawing together a wide range of scenarios, ensembles and models into a coherent message. When I initially read the abstract I was worried that so many permutations would become unwieldy, but the clear and systematic organisation of the experiments and the discussion of their outcomes leads the reader well and arrives at clear, justified conclusions. The authors find that differing initial noise types and parameters can have a meaningful impact on the recovery of tectonic information using common approaches including chi and slope-area analysis is challenging. However they also demonstrate that the distributions of geomorphic properties (hypsometry, Hack relationships) remain stable across varying types and applications of noise into landscape evolution models. The manuscript concludes by making some clear recommendations about the need to generate ensembles of models as a way of quantifying the uncertainties inherent in landscape evolution modelling. Overall, I believe this is an excellent contribution which is a good fit for Esurf and will be of wide relevance to the landscape evolution modelling community, as well as the broader geomorphology community."
We thank the reviewer for their commentary and respond to their individual comments below.

"My only reservation with this work is in the choice to generate noise in the range 0-1m rather than centring the noise distribution around 0 for scenarios B and C. As you highlight in Section 4.4.1 this means that you are only ever adding elevation to your models, which feels like a poor representation of geomorphic processes which roughen landscapes. I would be reassured by some tests that demonstrate that this choice of constructive noise is not driving the results we see in scenarios B and C. Or alternatively, some more discussion about the use of noise within the discipline showing that standard practice is to lock noise to the range 0-1."
We chose to add noise with elevations 0 to 1 m to avoid the models having negative elevations (e.g. early on or late in their evolution when topography can be low), principally to avoid generation of internal drainage and the requirement of additional sink-filling steps. We agree of course that noisy geomorphic processes in reality could potentially increase or decrease elevations, we think it is generally unclear what appropriate mean elevations of noise might be. Nonetheless, following the reviewer's suggestion we ran an additional series of models in which quenched and spatio-temporal noise (Scenarios B and C), were added within the range -0.5 to 0.5 m. The results from these models are similar to those with noise in the range 0 to 1 m. We include these results within the revised Figures 17 & 18, and in the Supplemental Figures S9 & S10. We discuss them alongside existing results for Scenarios B and C in the revised Sections 3.2 and 3.3 of the main text.

"Something that I think could merit some additional discussion is the potential sensitivity of some of your metrics to the channel extraction threshold. In general we would expect that the most dynamic parts of a landscape are the first order basins and channels, so things like approximations of steady state in channels may vary based on the threshold used. I'm not asking for another parameter to be added to your models, but some reflection in the discussion could be interesting."
We think that the reviewer is referring to the threshold upstream drainage area used to define the upstream limit of the rivers analysed, which we assumed was 4 km^2 (uppermost 4 cells). In that sense we think that we are incorporating the relatively dynamic parts of the landscape being referred to here. We note that the assessments of steady state in our paper includes all cells in the landscape, we therefore think that we may have already satisfied the reviewer's concern and are unsure what additional material we should add to or discuss in the paper.

"You ascribe the spike pattern observed in Figure 9 to the use of D8 flow routing and a square

grid. You could quickly test this by running the D-infinity landlab component on a bunch of your DEMs to see if that is enough to smooth out the spike across all noise types."
We thank the reviewer for their suggestion, we had considered this approach when producing the paper— it is surprisingly non-trivial to do so. It turns out that the D-infinity component is incompatible with a number of model components used within this study (e.g., `SinkFillerBarnes`, `FastscapeEroder`), and similar alternatives (`SharedStreamPower`, `SpaceLargeScalerEroder`). As we understand, it is not currently possible to perform the suggested tests, without major alterations to the fundamental model setup, which we think is beyond the scope of this contribution.

"A very minor point, but throughout the manuscript, you present and discuss the three colours of white noise in different orders. For example Figure 4 goes white, blue, red; Figure 6 goes red, white, blue; Section 2.2.2 goes white, red, blue. There's a lot of data being presented in this manuscript, so sticking to one consistent order will make it slightly easier for the reader to follow things."
We have modified the figures and text to be consistently in the order: red, white, blue (i.e., in order of increasing exponent p in Equation 4). The exception to this arrangement is in Section 2.2.2, where we feel an order of white, red, blue is more appropriate due to red and blue noise distributions being generated from white noise distributions, explained in the main manuscript.

"Line 46 - "noise to inserted" is missing a word."
Thank you! We have modified this sentence to "...noise to be inserted...".

"Line 86 - You discuss steady state in detail later (which I really enjoyed), but it would be good to add a citation here and a little more detail to outline what definition(s) you are using in this study."
Thank you. We now briefly define topographic steady state here, and point to the more fulsome explanation within Section 2.4 of the text. This now reads as *"...landscapes that are demonstrably at, or very close to, topographic steady state, which we define here as (practically) no elevation change in any grid cell between successive model time steps and explain further in Section 2.4 (see e.g., Willett & Brandon, 2002).*

"Line 113 - The units for K are missing their numerator."
Dimensional analysis of the stream power model shows that the prefactor K has units of $[1-2m][\mathrm{T}]^{-1}$, where m is the erosional exponent within the stream power model, and T is time. Within our modelling our choice to use a value of 0.5 for m results in K having units of $[\mathrm{T}]^{-1}$. Therefore in this case we believe the units of K are displayed correctly.

"Line 171 - A note to highlight how good it is to see the practice of recording or fixing seeds discussed!"
Thank you.

"Table 1 - It wasn't clear to me when reading the manuscript, what scenarios A, B and C refer to here. I think they are defined later in the manuscript, but that definition needs to come before their first use."
On reflection, it is not essential to refer to Scenarios A to C here, and we now omit their mention in this table caption. The first reference to these scenarios is therefore in Section 2.3, where we define them in full.

"Figure 8b - Missing values on y-axis."
The y-axis values are present on the right hand side of this panel, where we would like to retain them.

"Line 325 - "calculated from channel lengths and areas" This should read "and basin areas", so that there is no confusion about the calculation of channel areas."

Corrected.

"Line 354 - "a value often used in studies of real landscapes" - this needs a citation to at least one of these studies."
We now cite Whipple and Tucker (1999) and Roberts and White (2010).

"Figure 11g - The hack exponents for the largest basins reported here seem to sit considerably outside the norm. I did some work last year (https://doi.org/10.1029/2024GL111220) looking at Hack's exponents on a global scale only saw values that high when very odd things were happening. Is there a morphometric explanation for these values?"
The Hack exponent values shown in the old Figure 11g included those for the coloured lines, which were calculated using a non-linear regression of channel length-upstream basin area data sampled sequentially along 4 individual streams (using the `HackCalculator` component within Landlab). To avoid confusion we have now removed those results, and instead only show the results for the maximum length and areas of the rivers in each model (as is more commonly done). The values of Hack exponents obtained from those results are within the global range calculated for observed datasets, such as the work you refer to, which we now cite.

"It is great to see the code associated with this manuscript available online, and citable. I have gone through the code on github and it is well written and structured, and reflects the analyses described in the manuscript. To aid reproducibility it would also be helpful to record the package versions python versions you are using within a readme, in case future upgrades break things in your code."
Thank you for viewing associated code. Just to be clear—at the time of review there was code in a Zenodo repository, but no code available on Github associated with this manuscript. We have now made this available in a Github repository (link within main manuscript), and included a readme file with package versions.

References

Lipp, A. G. and Roberts, G. G.: Scale-Dependent Flow Directions of Rivers and the Importance of Subplate Support, Geophysical Research Letters, 48, e2020GL091 107, https://doi.org/10.1029/2020GL091107, 2021.

Roberts, G. G. and White, N.: Estimating uplift rate histories from river profiles using African examples, Journal of Geophysical Research: Solid Earth, 115, https://doi.org/10.1029/2009JB006692, 2010.

Whipple, K. X. and Tucker, G. E.: Dynamics of the stream-power river incision model: Implications for height limits of mountain ranges, landscape response timescales, and research needs, Journal of Geophysical Research: Solid Earth, 104, 17 661–17 674, https://doi.org/10.1029/1999JB900120, 1999.