

Figure S1: The optimal advection model n value in the 30m resolution models based on the Copernicus 1-arc second DEM (a) is shifted slightly higher compared with the lower resolution model, (optimal value $n=1.36$ compared with $n=1.28$ in the optimal 90m resolution model). Meanwhile, the optimal A_c (b) is lower ($A_c=0.02$ km² here vs 0.05 km² in the low resolution model). The different resolution models are still largely consistent in spite of these differences, and give a similar maximum NSE of 0.47 (c) compared with 0.48 for the 90m model.

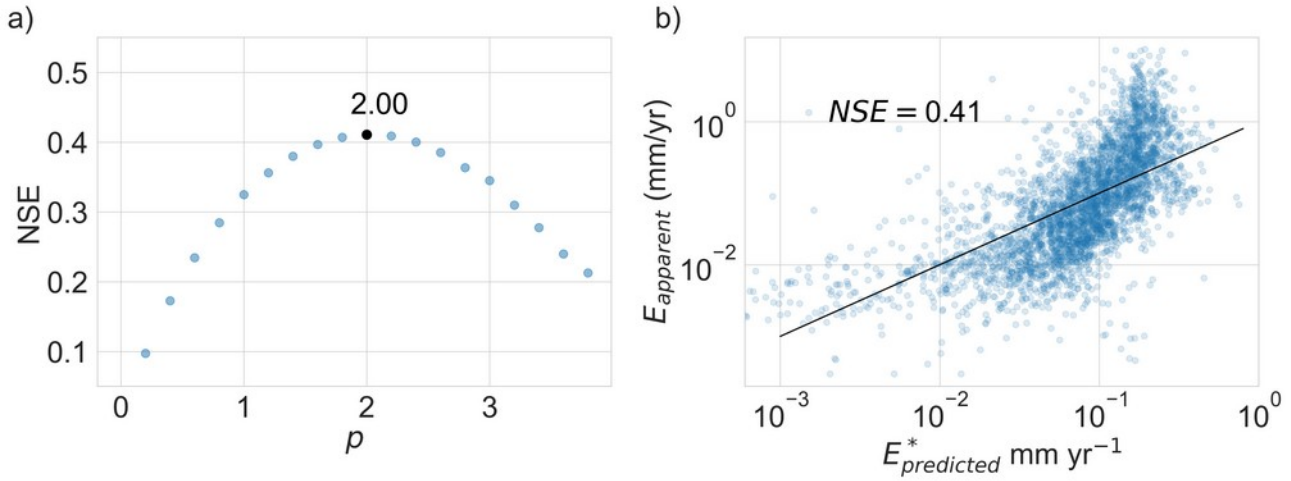


Figure S2: The optimised diffusion exponent in the 30m models of $p=2.0$ is consistent with that of the lower resolution model of $p=2.0$. However, the optimal fit is worse at $NSE = 0.41$ (compared with $NSE=0.51$ for SRTM 90). This difference may arise in part because the model based on the SRTM data appears better able to capture the variability of erosion rates (compare panel b to Figure 3b of the main text). This may suggest that the higher scatter of the SRTM surface elevation data, (whether real or due to noise) serves as somewhat of a proxy for surface roughness in steeper / rapidly eroding terranes.

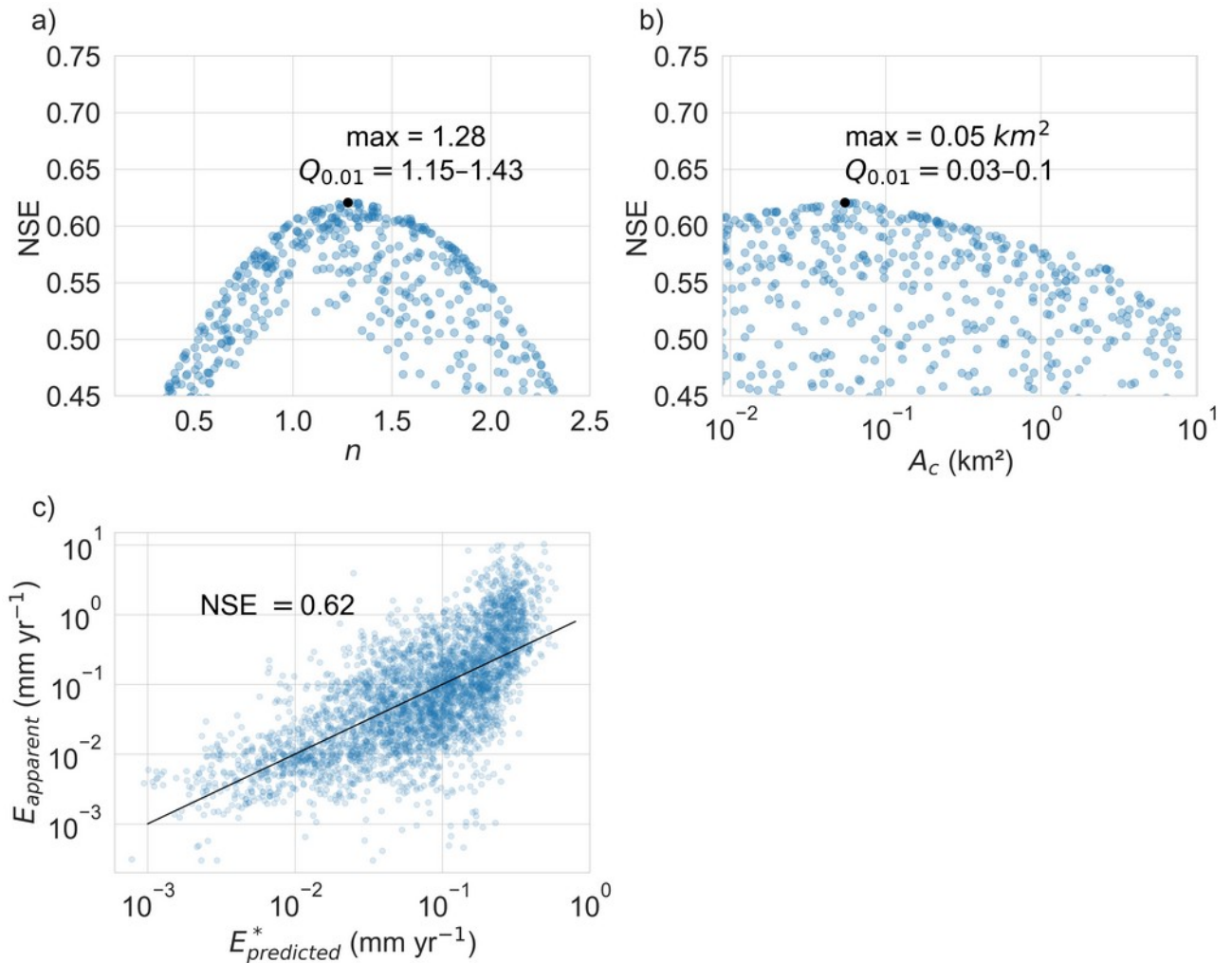


Figure S3: Example of the advection-only model run using a form of NSE that accounts for measurement error after Harmel and Smith (2007). The Model yields nearly the same optimized n (1.28) and A_c value (0.05 km^2) as the model run using NSE on the means from the main text ($n=1.28$, $A_c = 0.05$). The main difference is that the NSE metric is skewed upwards (c; $\text{NSE}=0.62$) because many $E_{predicted}^*$ estimates lie within the range of measurement error of $E_{apparent}$.

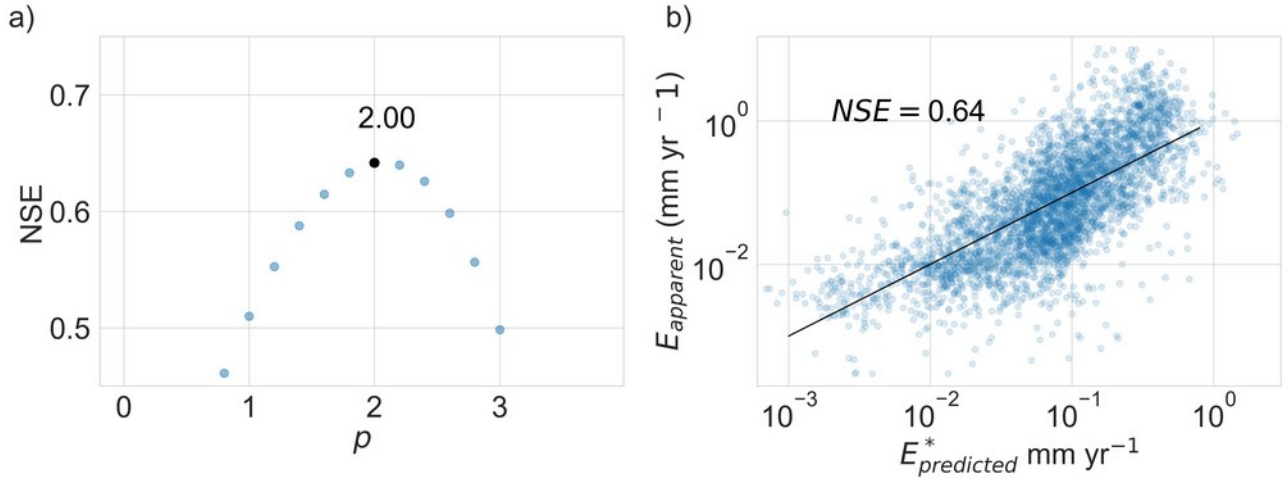


Figure S4: Example of the diffusion-only model run using a form of NSE that accounts for measurement error after Harmel and Smith (2007). The Model yields the same optimized p (2.0) as the model run using NSE on the means from the main text. The main difference is that the NSE metric is skewed upwards (b; $\text{NSE}=0.64$) because many $E_{predicted}^*$ estimates lie within the range of measurement error of $E_{apparent}$.

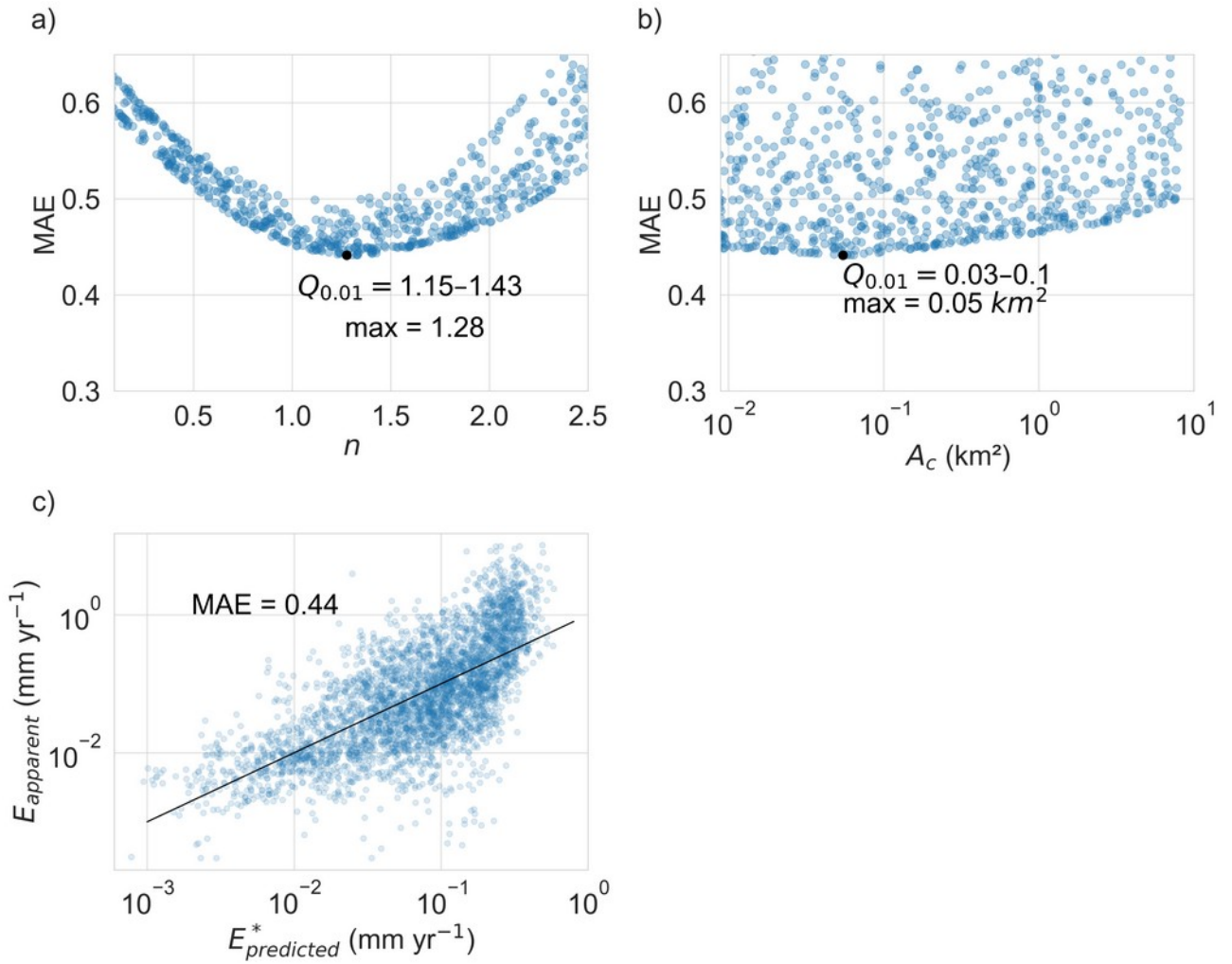


Figure S5: The optimal advection-only values, using MAE as a likelihood function a) $n=1.28$ and b) $A_c = 0.05 \text{ km}^2$ are identical to the results from models optimised using NSE of $n=1.28$ and $A_c=0.05$, respectively. The optimal model (c) has an average absolute error of 0.44 (log(mm yr⁻¹)).

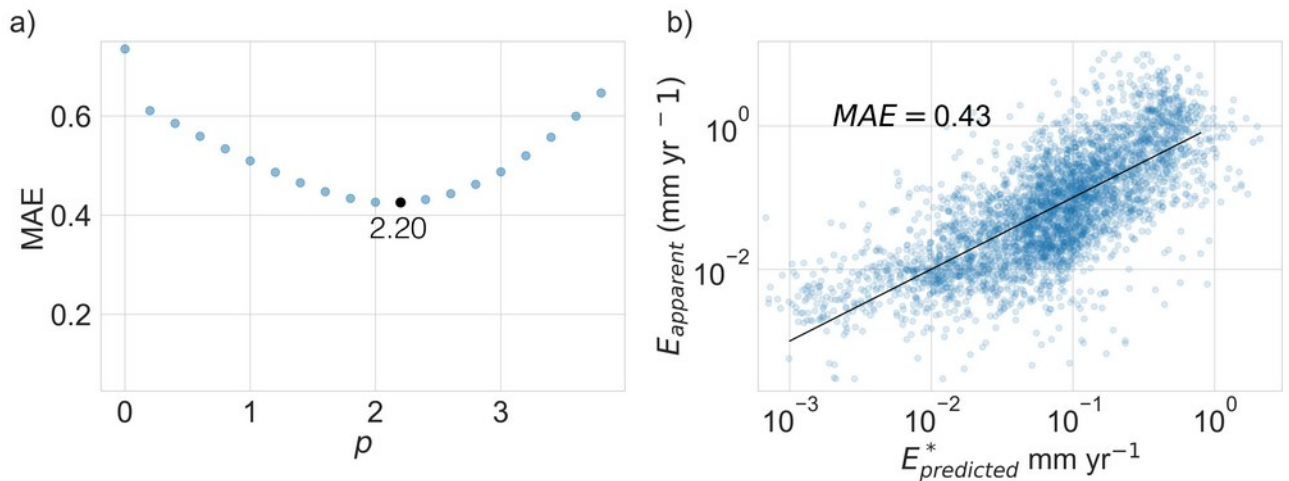


Figure S6: The optimal advection-only values, using MAE as a likelihood function a) $p=2.2$ is similar to the results from models optimised with NSE, $p=2.0$. The optimal model (c) has an average absolute error of 0.43 (log(mm yr⁻¹)).

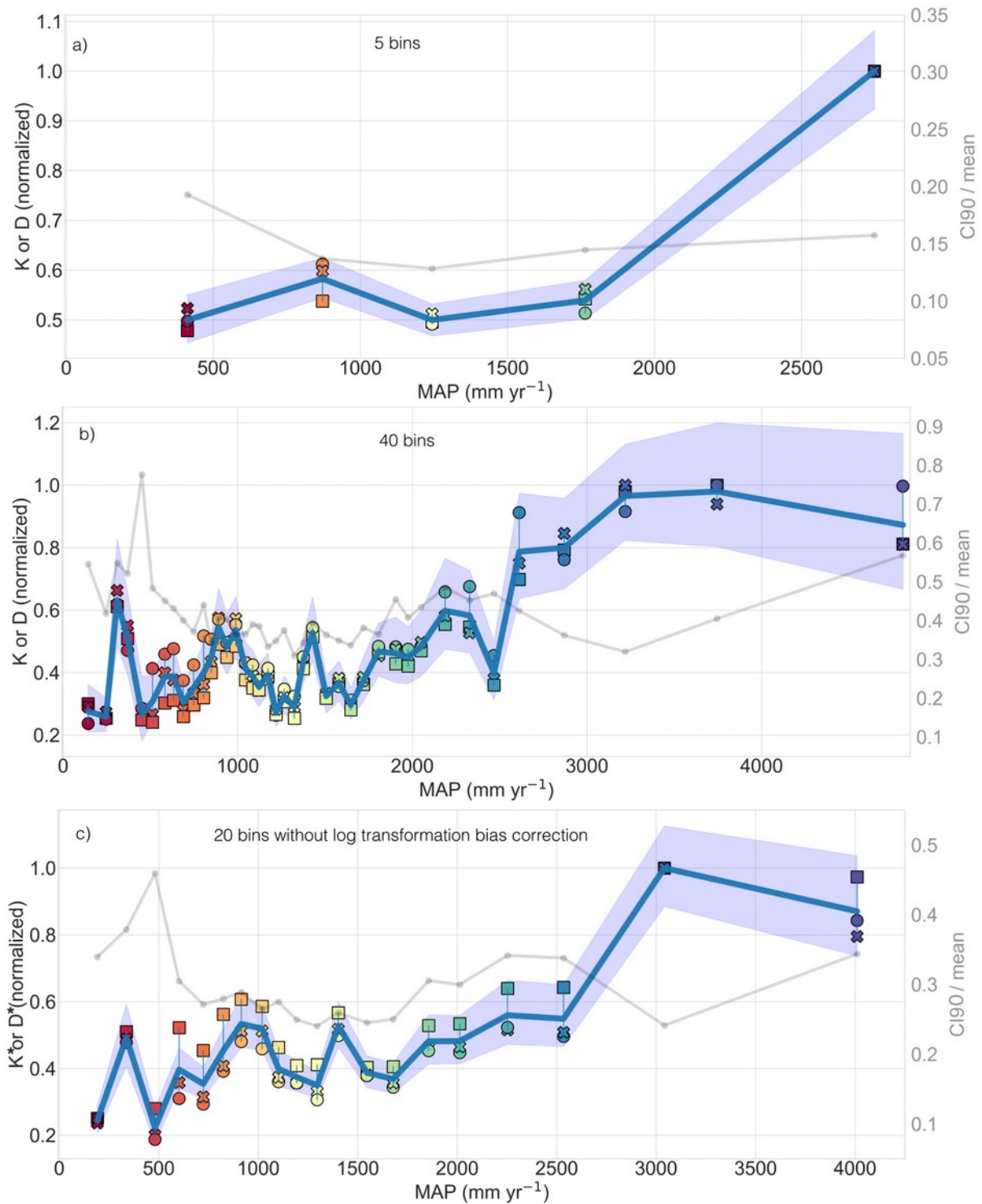


Figure S7: Exploration of binning on the relationships between (K or D) or (K^* or D^*) and MAP. a) K determined using only 5 MAP bins. b) K determined using 40 MAP bins, and c) non-log transformed K (K^*) determined using 20 MAP bins. Refer to Figure 5 of the main text for a full description of figure elements.

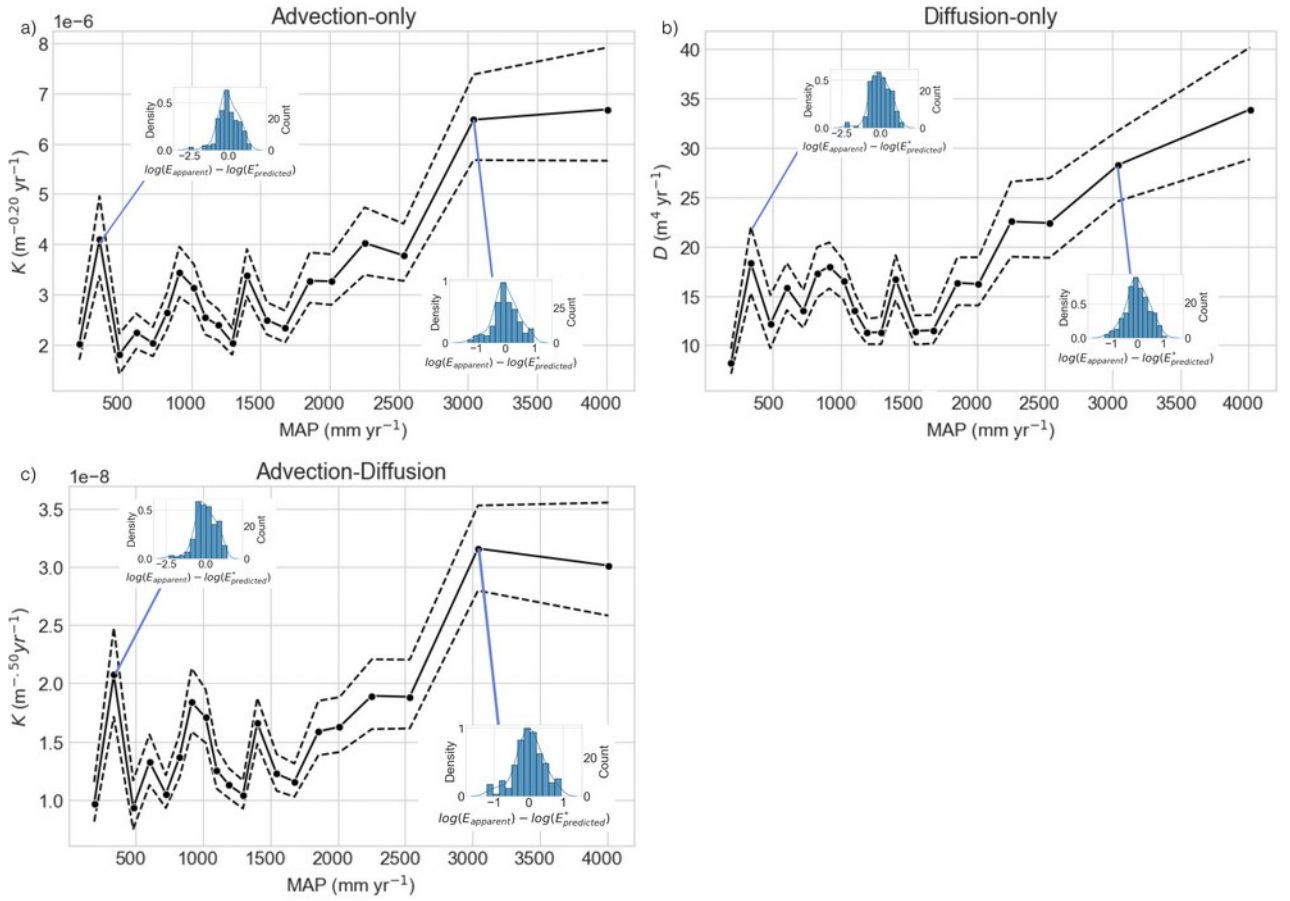


Figure S8: Corresponding Figure for Figure 5 of the main text, showing the 90% confidence intervals surrounding optimized K values within different climate bins and residuals for each bin. CI was determined by calculating K within each bin using (nonparametric) bootstrapped values of E_{apparent} for 1000 iterations. We also show residuals and corresponding high scatter for select bins, but with a generally log-normal appearance for most bins. a): Advection only with $n = 1.28$, $A_c = 0.05$; b): Diffusion-only with $p = 2.0$; c): Advection-Diffusion with $n = 2.26$, $A_c = 0.03$, $D/K = 1.79 \cdot 10^6$.

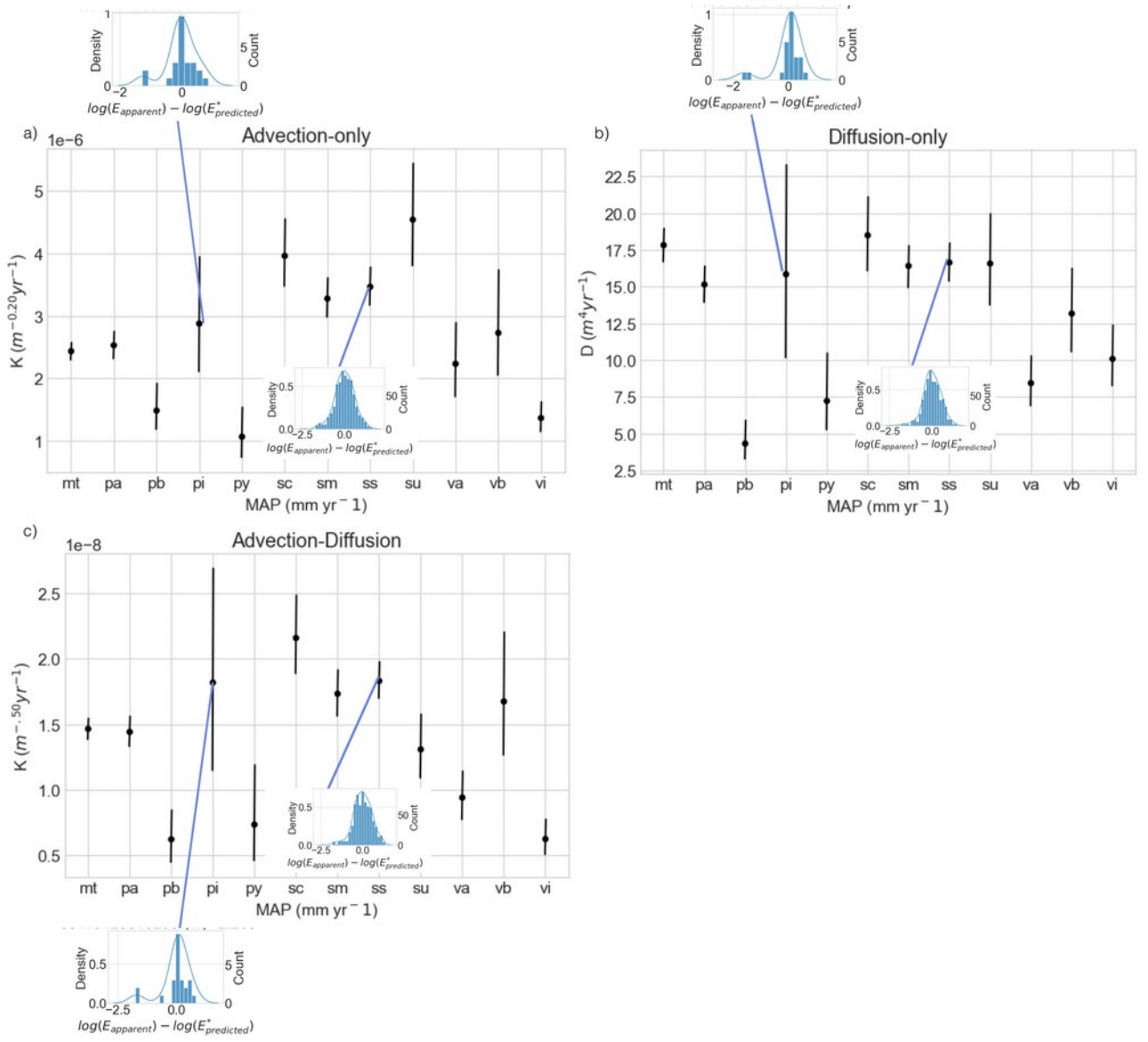


Figure S9: Corresponding Figure for Figure 6 of the main text, showing the 90% confidence intervals surrounding optimized K values within different climate bins and (inset) residuals select bins. CI was determined by calculating K within each using bootstrapped values of E_{apparent} for 1000 iterations. We also show residuals and corresponding high scatter for select bins, but with a generally log-normal appearance for most bins. The exception to this may be bins with lower numbers of samples (e.g., Plutonic Intrusive), which also have high uncertainty. a: Advection only with $n = 1.28$, $A_c = 0.05$; b: Diffusion-only with $p = 2.0$; c: Advection-Diffusion with $n=2.26$, $A_c=0.03$, $D/K = 1.79 \cdot 10^6$.

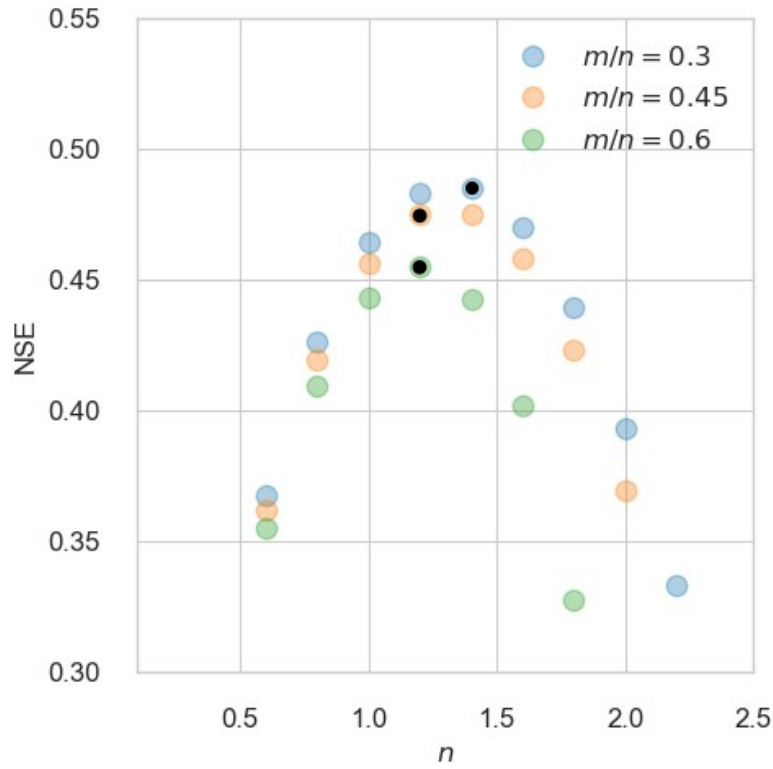


Figure S10: Model results for the advection-only model with free variable n using different m/n ratios. The best performance is the model run with $m/n = 0.3$ ($n=1.4$, NSE=0.48), slightly higher than the model run with $m/n=0.45$ ($n=1.2$, NSE=0.47) and higher than the model run with $m/n=0.6$ ($n=1.2$, NSE=0.45). However, our inability to optimise the global average of $m/n \sim 0.45$ suggests that perhaps our method cannot constrain this parameter while simultaneously solving for n , and instead it must be inferred directly from topographic morphometry (e.g. Gailleton et al., 2021).