

### Please cite the Published Version

Dwyer, GK, Cummings, CR, Rice, S , Lancaster, J, Downes, BJ, Slater, L and Lester, RE (2021) Using fractals to describe ecologically-relevant patterns in distributions of large rocks in streams. Water Resources Research, 57 (7). e2021WR029796 ISSN 0043-1397

DOI: https://doi.org/10.1029/2021WR029796

Publisher: Wiley/American Geophysical Union

Version: Accepted Version

Downloaded from: https://e-space.mmu.ac.uk/633885/

Usage rights: O In Copyright

Additional Information: This is the peer reviewed version of the following article: Dwyer, G. K., Cummings, C. R., Rice, S. P., Lancaster, J., Downes, B. J., Slater, L., & Lester, R. E. (2021). Using fractals to describe ecologically relevant patterns in distributions of large rocks in streams. Water Resources Research, 57, e2021WR029796, which has been published in final form at https://doi.org/10.1029/2021WR029796. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

**Data Access Statement:** To foster transparency, our data is available on Deakin University's research repository.

### Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

# Using fractals to describe ecologically-relevant patterns in distributions of large rocks in streams

- Authors: G. K. Dwyer<sup>1</sup>, C. R. Cummings<sup>1</sup>, S. P. Rice<sup>2</sup>, Jill Lancaster<sup>3</sup>, Barbara J. Downes<sup>3</sup>, L.
   Slater<sup>4</sup>, and Rebecca E. Lester<sup>1</sup>
- <sup>5</sup> <sup>1</sup>Centre for Regional and Rural Futures, Deakin University, Locked Bag 20000, Geelong,
- 6 Vic. 3220, Australia.
- <sup>2</sup> Department of Geography and Environment, Loughborough University, Loughborough,
   LE11 3TU, U.K.
- 9 <sup>3</sup> School of Geography, The University of Melbourne, Victoria, Australia
- <sup>4</sup>School of Geography and the Environment, University of Oxford, South Parks Road, Oxford
   OX1 3QY, UK
- 12
- 13 Corresponding author: Rebecca Lester (<u>Rebecca.Lester@deakin.edu.au</u>)
- 14 Key Points:
- Longitudinal patterns of emergent fluvial rocks in six streams in Scotland and
   Australia exhibited fractal behaviour (self-similarity).
- Fractal dimensions were related to development of bedform topography and the density and size of available bed materials in the streams.
- Fractal dimensions are a promising measure of physical complexity that enable
   comparisons across ecosystems, scales and linked disciplines like ecology and
   geomorphology.
- 22
- 23 Supplementary material references: [*Basharin*, 1959; *Miller*, 1955; *Seuront*, 2009; *Seuront*
- 24 *and Lagadeuc*, 1997; *Strobl*, 2005]

### 26 Abstract:

27 Measuring the physical complexity of habitats or ecological resources is often achieved using 28 system-specific methods that make comparisons across ecosystems difficult. One measure 29 that is applicable across multiple ecosystems and scales is the fractal dimension, which has 30 the benefit of generality as well as potential scale independence. This study evaluated the use 31 of box-counting and entropy fractal dimensions for characterising the complexity of emergent 32 rock distributions in six streams across Scotland and Australia. Emergent rocks (ER) are 33 important hydraulic features and ecological resources, including as oviposition sites for 34 aquatic insects and cover for fish. We complete fractal analysis on counts of ER in 5-m 35 segments along longitudinal stretches of the six streams. All six streams exhibited fractal 36 behaviour (self-similarity), suggesting that fractals can be used to measure the complexity of 37 longitudinal ER distributions in a way that is scale independent. Entropy was a superior measure due to its ability to differentiate among the six streams whereas box-counting could 38 39 not. Together, field results and numerical simulations showed that fractal dimensions of 40 emergent rock distributions were related to stream geomorphology. Well-developed 41 bedforms, like alternating pools and riffles had better organised emergent rocks because large 42 bed materials were more likely to be emergent in topographic highs. Streams with coarser 43 bed materials had more chaotic arrangements of emergent rocks because this increased the 44 general abundance of emergent rocks, making differentiation between topographic highs and 45 lows less distinctive. Fractal dimensions, therefore, can measure the complexity of river 46 systems in a way that is relevant to geomorphological and ecological processes.

### 47 **Plain language summary:**

48 Fractal dimensions are used to characterise the complexity of a wide range of patterns in 49 nature, from single objects (e.g. branched twigs) to whole environments (rainforests), and 50 learn where consistent patterns may occur. We measured the complexity of rock patterns 51 (specifically rocks that emerge above the water's surface) in six rivers from Scotland and 52 Australia using fractal dimensions. These rocks provide important habitat for plants, insects, and fish in rivers, and so are important to overall stream condition and functioning. Less 53 54 complex, more highly structured rock patterns (lower fractal dimensions) occurred in streams with smaller rocks, which had areas where emergent rocks were concentrated (riffles) and 55 56 many long pools without emergent rocks. These results suggest that fractal dimensions may 57 be a promising measure of complexity that can help us understand relationships between

- 58 physical characteristics of streams and their ecology. Fractal dimensions also allow
- 59 comparison of rock patterns in rivers with other habitats, such as shrubs in grasslands for
- 60 example. This may allow future research to explain patterns that are consistent across these
- 61 different ecosystems and so advance general ecological theory.

## 62 **Index terms:**

- 63 1. 1825 Geomorphology: fluvial (1625)
- 64 2. 0458 Limnology (1845, 4239, 4942)
- 65 3. 1856 River channels (0483, 0744)
- 66 4. 0439 Ecosystems, structure and dynamics (4815)
- 67 5. 0430 Computational methods and data processing

# 68 Keywords:

- 69 1. Ecological resources
- 70 2. Fluvial geomorphology
- 71 3. Large roughness elements
- 72 4. Habitat complexity
- 73 5. Spatial heterogeneity
  - 6. Information dimension
- 74 75

# 76 Abbreviations:

- 77 ER Emergent rocks
- 78 D<sub>X</sub> Fractal dimension, unspecified method
- 79 D<sub>B</sub> Box-counting fractal dimension
- 80 D<sub>E</sub> Entropy fractal dimension
- 81  $\delta$  Box size
- 82 LZS synthetic low-zero streams
- 83 HZS synthetic high-zero streams
- 84

### 85 **1. Introduction**

The physical structure of environments affects every aspect of ecosystem structure and 86 87 function [e.g. Cuthbert et al., 2019; Ossola et al., 2016]. Physical structure here refers to the three-dimensional configuration of living space: rocky coasts are more physically complex 88 89 than sandy beaches, forests are more physically complex than grassy plains, and so forth [Bell et al., 2012]. Increasing physical complexity creates a greater diversity of resources (e.g. 90 91 living spaces and food) and results in higher species diversity; this association holds true in 92 every ecosystem that has been tested [Barnes et al., 2013] and requires an explanation, but 93 quantifying physical complexity is difficult. Most researchers address this difficulty by 94 developing measures that are intrinsic to particular ecosystems, such as counts of cracks and crevices in rocks [Downes et al., 1998], wood loading in streams [Lester et al., 2007], or 95 96 tributaries across networks [Rice, 2017]. Such eclectic measures, while useful within a 97 system, cannot be applied across multiple ecosystems at different scales, thereby precluding 98 general tests of hypotheses and meta-analyses.

99 General methods for capturing physical complexity exist and one such measure, the fractal 100 dimension [Mandelbrot, 1967], is applicable across multiple ecosystems and scales. Fractal 101 dimensions lie between the well-understood dimensions of 1, 2 and 3 for a line, surface and 102 volume, respectively, and express the extent to which the space is filled. For example, a complex, wiggly line on a 2-dimensional plane (fractal dimension close to 2) would have a 103 104 larger fractal dimension than a straight line (fractal dimension would be 1 for a perfectly straight line). While natural systems commonly require more than a single exponent to 105 106 describe their dynamics (multiple scaling; multifractal), one compelling aspect of fractal dimensions is that many environments have the same fractal dimension over a range of 107 108 spatial scales, in which case they are termed self-similar over those scales [Mandelbrot, 109 1967]. Coastlines [Mandelbrot, 1967] and river networks [Mantilla et al., 2010] are two wellknown examples, i.e. their maps look the same regardless of the scale at which they are 110 111 mapped (self-similarity). Where environments are self-similar (fractal-like), fractal 112 dimensions can be used to measure complexity in a way that is scale independent and 113 transferable across ecosystem types.

114 Fractal dimensions and related functions have proven useful for describing characteristics of

- 115 river systems, from channel networks [Rinaldo et al., 1992; Rodríguez-Iturbe and Rinaldo,
- 116 2001; Tarboton, 1989; Yang and Shi, 2017], through riverbed topography [Sapozhnikov and
  - 4

117 Foufoula-Georgiou, 1996; Zhong et al., 2012] and planform sinuosity [e.g. Nikora, 1991] to grain-scale topography [e.g. Butler et al., 2001] in one, two and three dimensions. Fractal 118 119 descriptions have, in turn, offered new insights into fluvial processes, including the 120 generation of bed material fabrics and bedforms, sediment transport, hydraulic resistance and 121 hyporheic exchange [Aubeneau et al., 2015; Lee et al., 2020; Singh et al., 2011; Singh et al., 122 2009; Stewart et al., 2019]. Work using fractals to describe fluvial processes has extended 123 into descriptions of the areal arrangement of large clasts as pebble clusters [*Wu et al.*, 2018], 124 but has not been applied to regional (comparisons among streams) or local (comparisons 125 among habitat types within a stream) scales at which most theoretical ecological models apply [Wiens, 1989]. Most applications of fractals by ecologists have been over small spatial 126 127 scales [e.g. Jeffries, 1993; Warfe et al., 2008], which are less suited to understanding

128 populations.

129 An outstanding question is whether physical resources for insect populations, namely rocks

130 that protrude above the surface of the water (emergent rocks, ER) exhibit fractal

131 arrangements. Emergent rocks are an important resource for many ovipositing insects, which

rely on them for successful recruitment [*Lancaster et al.*, 2010]. They also serve as sites for

133 insect emergence [Petersen and Hildrew, 2003] and as substrates for bryophytes [Downes et

134 *al.*, 2003]. They create flow structures that provide resting, cover and feeding opportunities

135 for fish [*Hayes and Jowett*, 1994], trap drifting organic foodstuffs [*Hoover et al.*, 2006], and

136 create microscale habitat heterogeneity that affects the distribution of macroinvertebrates

137 [Bouckaert and Davis, 1998]. The spatial arrangement of ER has ecological implications. For

example, some caddisflies (*Ulmerochorema* spp.) lay eggs on ER surrounded by fast flow;

139 such ER are typically clumped in areas of high velocity (e.g. riffles) and this can lead to very

140 high local densities of caddisfly eggs and potentially of newly hatched larvae [Lancaster et

141 *al.*, 2003; *Lancaster et al.*, 2020].

142 ER also have important roles in stream geomorphology and hydraulics, including affecting

143 drag and shear stress [e.g. *Cooper et al.*, 2013], generating turbulent structures [e.g. *Lacey* 

144 *and Roy*, 2008], and influencing sediment entrainment, deposition and transport processes

145 [e.g. Monsalve et al., 2017; Papanicolaou et al., 2018]. The spatial distribution of ER reflects

146 patterns of sediment supply, dispersal, and sorting across multiple scales. The location of

- sediment supply points, the volume and size of sediments delivered to the channel, and
- subsequent sorting by particle size and shape are affected by factors including geology,
  - 5

149 geomorphological history and hydrometeorology. At local scales, large rocks are arranged

- 150 into structures via process-form feedbacks with the flow. The resulting bedforms are key
- 151 components in the definition of channel morphology, e.g., cascades, step-pools, plane beds
- and pool-riffle sequences [Montgomery and Buffington, 1997]. It is reasonable to expect ER
- to be more prevalent in riffles and similar bedforms (steps, cascades) that preferentially store
- 154 coarse sediments, and where the bed surface is elevated closer to the water surface, making
- emergence more likely. Particle characteristics may also affect the likelihood of emergence;
- 156 larger rocks should have a greater propensity to emerge, and rock shape (platy versus equant)
- 157 may influence emergence and sorting processes. At larger spatial scales, ER may be more
- 158 likely in zones of coarse sediment aggradation, for example at the upstream end of
- sedimentary links, downstream of significant tributaries [*Rice and Church*, 1998].

160 Beyond such simple expectations, there is limited understanding of how numbers of ER vary

161 in space or what controls their spatial arrangement. An ability to better describe and explain

162 the spatial organisation of ER is therefore key to improve our understanding of

163 geomorphological and ecological processes, such as flow resistance and species diversity. In

164 this work, we investigated the controls driving differences in fractal dimensions of ER using

165 numerical simulations of synthetic streams designed to manipulate key characteristics

166 independently. Specifically, we addressed three research questions:

- 167 (Q1) Are longitudinal ER distributions, at scales of up to a kilometre, self-similar
  168 (fractal)?
- 169 (Q2) Do fractal dimensions capture differences in ER distributions between rivers
   170 associated with differences in channel morphology and sediment characteristics?
- 171 (Q3) Using synthetic streams, which aspects of stream morphology are responsible for
  172 driving differences in fractal properties?

As fractal dimensions are influenced by the density of points (here ER) and their arrangement, see Figure 1, we hypothesised that channel morphology and sediment characteristics that increase the propensity for rocks to emerge (and hence increase the number of ER) or which describe the organisation of large rocks (into pool-riffle structure for example) would be related to fractal dimension. Larger rocks, particularly in steep, shallow streams, should be more likely to be emergent. In contrast, characteristics associated with bedform development, and hence the spatial organisation of ER, are hypothesised to be

- 180 negatively correlated with fractal dimension. Bedform development should be associated
- 181 with pool-riffle structure, long pools without ER or with low ER densities, riffles with high
- 182 ER densities, and lower stream slopes (see Supporting Information Table S1 for detailed
- 183 rationale of hypotheses).
- 184





186 Figure 1. How channel morphology and bed sediment characteristics are expected to influence fractal 187 dimensions. Fractal dimension is affected by the number of points and the arrangement of those points 188 in space. The fractal dimension of a single point is 0 and a line has the well-understood dimension of 189 1. ER in streams are analogous to points arranged along a line. Continuing to increase the number of 190 points or ER along the line will fill in the space until the points resemble a solid line, and hence this 191 will increase fractal dimension (top vs. bottom panels). If ER become spatially organized (i.e. into 192 clumps due to stream bedform topography), as opposed to a random distribution, then larger empty 193 spaces (i.e. pools) will occur along the line and fractal dimension will decrease (left vs. right panels). 194 How fractal dimension changes with the interaction between ER number, sediment characteristics, 195 and spatial organization is less clear.

### 196 **2.** Methods

197 We used field data from three streams each in Scotland and Australia to determine whether 198 patterns in ER counts in 5-m segments along ~1-km stream lengths were fractal-like. We then 199 tested the above hypotheses proposing how variables related to channel morphology and bed 200 sediment characteristics are correlated with fractal dimensions. The Scottish and Australian 201 sites have contrasting geomorphological histories and lithologies, and the streams vary in size 202 and slope, so the data set captures variations in channel morphology and particle 203 characteristics. Co-variance amongst variables used in hypothesis tests means that the 204 empirical work is necessarily constrained.

#### 205 2.1. Study systems

We investigated ER distributions from two sets of streams in SE Scotland and SE Australia 206 207 (Fig. 2; Fig. 3), which have been the focus of work on insect oviposition [Lancaster et al., 208 2010; Reich, 2004]. These regions have distinct bedrock geology, hydrometeorology and 209 geomorphological histories and vary in their reach-scale channel morphology sufficient to 210 provide a range of ER arrangements. These systems include three streams in SE Scotland 211 (Dye Water [Dye], Faseny Water [Faseny], and Kelphope Burn [Kelphope]) and three in SE Australia (Little River [Little], Snobs Creek [Snobs], and Steavenson River [Steavenson]). 212 213 The Scottish lithology is predominantly marine sedimentary (Silurian greywacke) [Davies et 214 al., 1986], whereas the Australian streams are underlain by volcanic complexes and marine 215 sediments [Marsden, 1973]. The hydrology of the Scottish system is quite flashy with short-216 lived floods [Lancaster, 2000], whereas the Australian system is less so. Mean annual rainfall 217 is ~830 mm in the area encompassing the Scottish streams (Scottish Environmental 218 Protection Agency, https://apps.sepa.org.uk/rainfall) and exceeds 1000 mm in the Australian 219 catchment (Australian Government Bureau of Meteorology, 220 http://www.bom.gov.au/climate/data/). The Scottish streams in the Lammermuir Hills have a 221 history of Pleistocene glaciation and post-glacial landscape adjustment. Each of the study reaches are located on floodplains of restricted width, set within convex hills with steep lower 222 223 slopes. Intermittent coupling to hillslopes, occasional bedrock outcrops and floodplain 224 erosion introduce some sediment to the streams but sediment supply is primarily from 225 upstream, headwater tributaries. The Australian streams, in the Goulburn River catchment of 226 Central Victoria, have not experienced glaciation and are set within a steeper, more 227 mountainous landscape. The reach settings are similar to the Scottish streams, with limited

228 floodplains, some points of hillslope coupling and occasional bedrock outcrops that affect channel orientation and recruit sediment. Woody debris is present in the Australian streams 229 230 but not in the Scottish streams. Using the Montgomery and Buffington [1997] classification, channel morphology differs among the streams (Table S2). Dye, Faseny, and Snobs are 231 232 dominated by plane bed morphology that tends toward cascades in places (steeper, smaller 233 depth to grain size ratio) with limited true riffles. Pools between plane bed sections are 234 common on Faseny, less common on Snobs and limited on Dye. Steavenson, Little and 235 Kelphope comprise mostly pool-riffle morphology, most prominent in Steavenson but with 236 increasing presence of plane bed sections in Little and Kelphope. Both regions have large 237 variations in the proportion of pools in the study reaches (Australia: 19, 34 and 38 % of the 238 stream length; Scotland: 14, 27 and 40 % of stream length bed (Fig. 3; Table S2). Channel 239 width varies among the streams, from Kelphope ( $2.5 \pm 0.5$  m; mean  $\pm$  SD), to Snobs ( $5.8 \pm$ 1.4), Faseny (5.8  $\pm$  1.7), Little (6.1  $\pm$  1.7), Dye (6.6  $\pm$  1.3), and Steavenson (9.5  $\pm$  2.6). 240



241

Figure 2. Scottish (circles) and Australian (triangles) study sites.

- 244 2.2. (Q1) Are longitudinal ER distributions, at scales of up to a kilometre, self-similar
  245 (fractal)?
- 246 Longitudinal profiles of ER counts were acquired to test whether patterns in ER are fractallike and consequently whether fractal dimensions reflect geomorphological features, 247 248 including measures of rock shape and channel morphology. ER of the Scottish streams were originally surveyed by Lancaster et al. [2010] and the Australian streams were surveyed 249 250 during the Austral summer (December - February) of 2016/17. Over the study length (685 -251 1000 m) of each stream, ER were counted within contiguous 5-m segments to describe the 252 spatial distribution of patterns along the stream lengths. 5-m corresponds roughly to the average channel width of the streams. ER were defined as any rock protruding above the 253 254 water surface, with a maximum b-dimension (width, perpendicular to longest axis) [Gordon 255 et al., 2004] of at least 5 cm and in at least 5 cm of water based on those typically used for 256 oviposition by many aquatic insects. These surveys were carried out at or near summer base 257 flow.



Figure 3. Longitudinal emergent rock (ER) distributions with examples of ER density in riffles and
pools. (a) Coloured bars illustrate 5-m segments denoted as pools (blue) and riffles-like segments
(black; inclusive of true riffles, step-pool, and plane bed; see section 2.3) for each of the six streams.
Unclassified segments (grey bars) were not surveyed for channel morphology. These were included in
the calculation of fractal dimensions but were excluded from simulations of the synthetic streams.
Photographs illustrate ER abundance in (b) pools and (c) riffle-like sections. Arrows indicate the
location of each photo.

266 Several methods are available for calculating fractal dimensions. Box counting is the default method, which has been used to measure complexity of forest understory and canopies 267 268 [Denny and Nielsen, 2017], rocky shores [Meager and Schlacher, 2013], and aquatic plant 269 habitat [Ferreiro et al., 2011], for example. Although wide use of the box-counting method 270 makes it well placed for comparisons among physical environments, it is a binary method 271 calculated by counting the number of occupied boxes (N<sub>B</sub>) and thus the underlying data are 272 simplified to presence/absence, which results in a loss of information [Halley et al., 2004]. 273 An alternative method, the information dimension, is based on the information science parameter, entropy, and involves determining the sum of the proportion of total ER ( $N_E$ ) 274 within each box, with weighting and correction for bias [Basharin, 1959; Miller, 1955]. This 275 276 method retains an estimate of the relative proportion of the variable of interest in the 277 calculation (here ER density), thus retaining more of the underlying data and potentially providing better explanatory power for systems where relative amounts of a resource are 278 279 important [Halley et al., 2004]. Because entropy is not commonly used for calculating fractal 280 dimensions, we apply both to allow comparison with other studies.

281 Box-counting  $(D_B)$  and entropy  $(D_E)$  fractal dimensions were calculated using methods 282 prescribed by Seuront [2009]. Both methods involve dividing the stream into a set of nested 283 equal-sized boxes, of size  $\delta$ . All possible values of  $\delta$  were used, i.e. 5, 10, 15, ..., 1000 m, however, the data set was reduced to include only unique values of  $N_X$  (i.e. values of  $\delta$  which 284 285 did not result in a change in the values of N<sub>X</sub> were excluded for either method, N<sub>B</sub> or N<sub>E</sub>). 286  $Log_e(\delta)$  versus  $log_e(N_X)$  plots provide an estimate of the fractal dimensions from the absolute 287 slope, and linearity of these plots indicates the possible presence of self-similarity (i.e. fractal 288 structure) (Fig. 4). For our datasets, D<sub>B</sub> will span a range of 0 (a single point) to 1 (a solid 289 line); however, D<sub>E</sub> retains a measure of relativity and so we might expect values greater than 290 1.

To avoid making assumptions of linearity in the log-log plots, we followed *Seuront's* [2009] three-step procedure to detect fractal-like properties in natural patterns. This ensures that only patterns that are fractal are described as such. These steps included the: 1)  $R^2$  – SSR [sum of squared residuals] Procedure, 2) the Zero-Slope Procedure and 3) the Compensated-Slope Procedure. *Seuront* [2009] specified that data sets could be considered fractal-like if they satisfied any two criteria of the three-step procedure. Reported estimates of the fractal dimensions (D<sub>x</sub>; unspecified method) in the results were taken from the Compensated-Slope

- 298 Procedure because this estimate is more robust to random non-fractal structure (i.e. artefacts
- of the data). All computations were undertaken using packages ggpmisc, stats, lmodel2 and
- 300 plyr in the open-source software R [*R Core Team*, 2019]. A more detailed description of
- 301 calculating fractal dimensions using *Seuront's* [2009] three-step procedure (Text S1; Fig. S1)
- 302 and the corresponding R script (Script S1) can be found in Supporting Information.





305 Figure 4. Calculating fractal dimensions of emergent rock (ER) distributions. To calculate fractal 306 dimensions, longitudinal counts of ER in 5-m segments along a 1-km stream stretch are observed at 307 different box sizes ( $\delta$ ). The minimum box size in fractal dimensions calculations corresponds to the 5-308 m segments ( $\delta = 5$ ). The box-counting fractal dimension (D<sub>B</sub>) and the entropy fractal dimension (D<sub>E</sub>) 309 methods use the presence (a) or proportion (b) of ER in each box, respectively. For all possible box 310 sizes (only six values of  $\delta$  are illustrated), the box-counting method calculates N<sub>B</sub> from the sum of boxes containing ER, whereas the entropy method calculates N<sub>E</sub> from the sum of the proportions of 311 312 ER within each box. Significant linearity in  $log(N_B)$  (c) and  $log(N_E)$  (d) plots with changing  $log(\delta)$ indicates the possible presence of a fractal structure. The absolute slope of this line can be used as an 313

 $\begin{array}{ll} \text{314} & \text{estimate of fractal dimension (here, $D_B$ and $D_E$ are 0.917 and 0.934, respectively). Equations are \\ \text{315} & \text{expressed in log-transformed units.} \end{array}$ 

316 2.3. (Q2) Do fractal dimensions capture differences in ER distributions between rivers 317 associated with differences in channel morphology and sediment characteristics? 318 As we confirm below, all six rivers are self-similar, which means their fractal dimensions 319 capture scale-independent aspects of how ER are distributed in these streams. Next, we tested 320 how fractal dimensions of ER vary with geomorphological variables related to bed 321 topography, because this affects the propensity of ER to be emergent and structured in space 322 (e.g. the strength of alternating pool-riffle development), and with the size and shape of bed 323 materials, because these affect the abundance of ER (Fig. 1). We conducted surveys to 324 provide measures of these characteristics.

325 Channel morphology surveys were carried out in Sept. 2018 in Scotland and Feb. 2018 in 326 Australia when discharge was at or near summer base flow. Surveys of morphology and ER 327 were aligned to ensure spatial compatibility. Morphological surveys were carried out on a reduced section of the ER survey length (525 - 940 m; Fig. 3), which was visually classified 328 329 for morphology at 2-3 m intervals using Montgomery and Buffington [1997] typology, as 330 described above and in Table S2. The primary purpose was to characterise the distribution 331 and arrangement of morphological units that are more or less likely to contain ER. We 332 therefore used a simplified binary morphological classification of pools and 'riffle-like', 333 lumping together true riffles, plane beds and step-pools as sections where the flow is 334 relatively shallow and bed materials are more likely to be exposed. For simplicity, we refer to 335 these sections as riffles and pools hereafter.

336 To document bed and water surface topography, a longitudinal survey was completed using a

dumpy level, with measurements at each interval along the thalweg. Channel width was

measured during the ER surveys at intervals of every 10 m for Scottish streams and every

339 20 m for Australian streams. The latter were measured at a coarser scale due to an observed

340 lack in variance. The shape and size of ER and submerged rocks were also surveyed.

341 Dimensions of ER and submerged rocks were measured at the top, middle and bottom of the

survey length using Wolman counts of 100 rocks at each location [*Wolman*, 1954]. We

343 measured a (length), b (width), and c-axes (thickness) of each rock to the nearest 5 mm

344 [Gordon et al., 2004; Rice and Church, 1996]. We characterised both ER and a random

345 sample of submerged rocks to determine whether the attributes of the ER themselves (which

346 were used to calculate the fractal dimensions) or the characteristics of the available bed

materials, which are better represented by the submerged rocks, are related to fractaldimensions.

349 Given the exploratory nature of the work, we calculated a suite of relevant variables from 350 these field data (Table S1; Table S3) and tested for correlation with fractal dimensions. For morphology, variables included stream slope, mean riffle and pool lengths (calculated 351 relative to the study length) and the number of pool-riffle transitions. For grain size we used 352 353 the mean of b and c axis and for shape we calculated mean values of equancy (c-axis/a-axis) 354 and flatness (c-axis/b-axis) ratios [Blott and Pye, 2008]. Some of these measures were 355 correlated with each other, particularly measures of sediment size and shape (Table S4). Moreover, the expectation of finding more ER in riffles than pools simplifies what is often a 356 357 complex pattern and is probably unrealistic in some streams. Our field observations 358 confirmed that, on average, riffles contained more ER than pools but also that most pools also 359 contained some ER in varying amounts. We therefore included several additional metrics 360 (Table S3) in the correlation tests that were intended to accommodate some of this uncertainty. First, we calculated the mean ER density in each study reach to capture 361 362 differences in the relative abundance of ER between the streams. Second, we removed stream morphology altogether by dividing the reaches into sections with and without ER and 363 364 calculating metrics for those units, including the number of segments without ER (Table S3). 365 These allow us to ask the fundamental question of whether fractal dimensions are related to 366 patterns of ER presence and absence, irrespective of whether that presence or absence is related to the pools and riffles identified in the field surveys. 367

368 Correlation tests (Pearson's product moment coefficient) were used to determine whether 369 fractal dimension is able to capture between-stream differences in ER associated with 370 differences in channel morphology and sediment characteristics, differences in the density of 371 ER in pools and riffles, and differences in the pattern of reaches with and without ER. These 372 were one-tailed correlation tests because, as explained in the Introduction, fractal dimension can only increase with numbers of ER, and can only decrease with an increase in the spatial 373 374 structure of ER. Relationships in the opposite directions are either mathematically 375 impossible, or at least improbable for the range of variable values likely to be encountered for 376 ER in rivers.

# 377 2.4. (Q3) Using synthetic streams, which aspects of stream morphology are responsible for 378 driving differences in fractal properties?

379 Multiple variables affecting fractal dimension were correlated with each other, hence we 380 created synthetic streams in a numerical model. These synthetic streams allowed some stream 381 characteristics to be manipulated individually and with greater replication, to test how the 382 characteristics influence fractal dimensions. Simulations were carried out in R based on 383 empirical measurements of morphological characteristics and ER numbers (for example see 384 Supporting Information Fig. S3), although some empirical data could not be simulated (e.g. 385 rock shape and size). Synthetic stream stretches were created by first randomly drawing riffle 386 and pool lengths from log-normal distributions. Riffle and pool lengths were alternated for 387 the length of the synthetic stream (1000 m). ER counts were then randomly drawn from the 388 negative-binomial distributions and assigned to each 5-m segment of the riffle and pool lengths. Log-normal distributions of riffle and pool lengths were produced from the observed 389 390 lengths of pools and riffles, which were fit to separate log-normal distributions. Log-normal distributions were chosen because riffle and pool lengths are continuous, positive (> 0) and 391 392 skewed with few high values. ER counts in riffle and pool segments were fit to separate 393 negative-binomial distributions. Negative-binomial distributions were chosen because these 394 data are discrete, nonnegative ( $\geq 0$ ), and skewed (i.e. suitable for few high ER counts). As the 395 ER distributions of the study streams fell into two distinct groups (those with few vs. many 396 segments without ER), separate negative-binomial distributions were produced from these 397 two groups to capture these differences (see Supporting Information Text S2 and Fig. S2 for 398 illustration). Low-zero streams (few segments containing zero ER) were simulated from the 399 ER count distributions of Dye, Little, Snobs, and Steavenson. High-zero streams (many 400 segments containing zero ER) were simulated from the ER count distributions of Faseny and 401 Kelphope.

To determine the effect of stream characteristics on fractal dimension, pool length, riffle length, and ER density were varied either individually or in combination and replicated 20 times. To disentangle the pattern of segments without ER from the total number of segments without ER, an additional set of synthetic streams was produced by alternating riffles with ER (randomly drawn from the negative-binomial distributions) and pools without ER of the same lengths. The lengths of the riffles and pools were not drawn using random processes but were systematically set to 14 different lengths, including: 5 (every second 5-m segment has

409 no ER), 10 (two 5-m segments with ER, two 5-m segments without ER), 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, and 500 m. For lengths that are factors of 1000 (the length of 410 411 the stream), the pattern of segments without ER will change, but the number of segments 412 without ER remains constant at 100 [500 m]. Twenty replicate synthetic streams were 413 produced for each of the 14 riffles and pool lengths. A total of 1360 synthetic streams were 414 simulated using both the random and systematic processes. For a more detailed description of 415 these simulations and the R script, see Supporting Information Text S2 and Script S2, 416 respectively. As a first test to see whether the ER distributions of low and high zero synthetic 417 streams produced systematically different values of D<sub>E</sub>, t-tests were used. To test for 418 differences in the variation of D<sub>E</sub> and variation in the number of ER between the two 419 distributions, modified signed-likelihood ratio (M-SLR) tests were used [Krishnamoorthy and 420 Lee, 2014]. M-SLR tests were implemented using the R package CVEQUALITY [Marwick and Krishnamoorthy, 2018]. Two-tailed linear regression analyses were performed to test for 421 422 relations between fractal dimensions and channel morphology characteristics. The slopes of lines show not only the direction of relationship with fractal dimension but also the relative 423 424 effect of each independent variable; steep slopes signal greater change in fractal dimension 425 than shallow slopes.

### 426 **3. Results**

# 427 3.1. (Q1) Are longitudinal ER distributions, at scales of up to a kilometre, self-similar 428 (fractal)?

429 The ER distributions of all six streams were fractal-like using both the box-counting and 430 entropy methods. All streams satisfied at least two of the three tests developed by Seuront 431 [2009]; this is sufficient to illustrate fractal behaviour (Table S5). Further, coefficients of 432 determination for all streams ranged from 0.99 to 1.0 for both the box counting and entropy 433 fractal dimensions (Fig S4). Fractal dimensions varied among the six streams and with the method of calculation (Fig. 5). There was very little discrimination among the streams using 434 the box-counting method (Table S5; Fig. 5a): four streams (Steavenson, Dye, Snobs and 435 436 Little) had the same value for D<sub>B</sub> and this value did not differ from the fractal dimension of a straight line (1.00), because very few segments ( $\leq 2$ ; equivalent to  $\leq 10$  m) in these streams 437 contained zero ER. Of the remaining streams, D<sub>B</sub> was 0.97 for Faseny (although statistically 438 439 the log-log slope was not significantly different from 1; Fig. S4) and 0.90 for Kelphope. 440 Significant breaks in the box-counting log-log slopes were found for all streams except Snobs

441 and Steavenson (Fig. S5).

442 The entropy fractal dimension (D<sub>E</sub>) largely matched the rank order of streams using the boxcounting method, but it discriminated numerically between all six streams. Values of D<sub>E</sub> 443 444 ranged from 0.91 to 1.05 and the rank order of streams from lowest to highest D<sub>E</sub> was Kelphope (0.91), Faseny (0.93), Steavenson (0.98), Dye (1.00), Little (1.02), and Snobs 445 446 (1.05) (Fig. 4b). This produced two groups, whereby log-log slopes for Kelphope and Faseny were significantly different from that of Steavenson, Little and Snobs (Fig. S4). Dye was only 447 significantly different to Kelphope. The log-log slopes for all streams except Dye were 448 significantly different from 1 (Fig. S4). Significant breaks in log-log slopes were found for all 449 450 streams, the position of these varied among the streams and was significantly correlated with 451 pool length and pool-riffle ER density ratio (Fig. S5). Due to its superior ability to 452 differentiate among streams, we consider D<sub>E</sub> the better method of measuring the complexity of ER distributions in our six streams. Therefore, analysis for questions Q2 and Q3 used only 453 454 the entropy fractal dimension (box-counting results can be found in Supporting Information 455 Table S6, Table S7 and Table S8).

456 3.2. (Q2) Do fractal dimensions capture differences in ER distributions between rivers 457 associated with differences in channel morphology and sediment characteristics? 458 Entropy fractal dimensions correlated well with some stream characteristics. D<sub>E</sub> was 459 negatively correlated with three stream characteristics that relate to the spatial arrangement of 460 ER: mean proportional pool length, the number of segments without ER and the maximum 461 length without ER (Table 1). D<sub>E</sub> was positively correlated with slope and ER density, which 462 relate to the number of ER. D<sub>E</sub> was also positively related to sediment characteristics of submerged rocks: the c-axis length, and equancy ratio. Given the small sample size (n = 6), 463 464 these results should be interpreted with caution given prospective lack of statistical power for some tests. More critically, some of the independent measures of stream characteristics are 465 466 correlated with each other, as well as with fractal dimension, which creates uncertainty about 467 cause and effect (Supporting Information Table S4).







4/1 Keipnope). Bars represent the presence (a) or proportion (b) of ER in each 5-m segment (the

472 minimum box size used in fractal dimensions calculations) and fill colour and the order of the streams

reflects the box counting fractal dimensions (DB for a) or the entropy fractal dimensions (DE for b) ofthe ER distributions.

475 Table 1. Study streams: Correlation tests (one-tailed) of associations between various stream
 476 characteristics and entropy fractal dimensions.

Characteristic	Expected	r	р	
	direction			
Stream morphology				
Median depth	Negative	0.791	0.969	
Stream slope	Positive	0.737	0.047	
Mean proportional pool length	Negative	-0.870	0.012	
Mean proportional riffle length	Positive	0.054	0.460	
Number of pool-riffle transitions	Negative	0.241	0.677	
Bed sediment				
SR (submerged rock) axis B	Positive	0.653	0.080	
SR axis C	Positive	0.950	0.002	
SR equancy ratio (C/A)	Positive	0.765	0.038	
SR flatness ratio (C/B)	Negative	0.755	0.958	
ER density measures				
Mean ER density over entire site	Positive	0.917	0.005	
Pool-riffle ER density ratio	Negative	0.541	0.866	
Segments without ER				
Number of segments without ER	Negative	-0.832	0.020	
Maximum longth without EP	Nogotivo	0.805	0.000	

477 Maximum length without ER Negative -0.895 0.008 Note. Bold text indicates p < 0.05 in the direction expected for the one-tailed correlation tests; DF = 4

477 Note: Bold text indicates p < 0.05 in the direction expected for the one-tanted conclusion tests, D1 = 4</li>
 478 for all tests. Tests involving sediment size and shape of ER were all non-significant and are reported
 479 in Table S6.

480	Table 2. Synthetic streams:	Linear regression tests (	(two-tailed) for relationshi	ps between various
		8	(	<b>F</b>

481 stream characteristics and entropy fractal dimensions, for low-zero streams (LZS) and high-zero

482 streams (HZS) in different simulations. For each simulation pool and riffle lengths, and the number of

483 ER in pools and riffle segments were held constant except for the manipulated variables.

Independent variable	Manipulated variable(s)	ER dist.	DF	Slope	t	р	R <sup>2</sup>	Fig.
Stream morphology								
Mean proportional pool	De al 1 an athr	LZS	78	- <b>8.9</b> ×10 <sup>-5</sup>	3.48	<0.001	0.13	71
length	Pool lengths	HZS	78	-7.8×10 <sup>-4</sup>	4.84	<0.001	0.23	70
Mean proportional pool	Pool lengths,	LZS	77	-3.8×10 <sup>-3</sup>	15.5	<0.001	0.76	74
length	riffle ER	HZS	77	-4.6×10 <sup>-3</sup>	8.46	<0.001	0.48	7 <b>u</b>
Mean proportional riffle	<b>Piffle lengths</b>	LZS	77	$8.2 \times 10^{-6}$	0.88	0.38	< 0.01	NΛ
length	Kille lenguis	HZS	77	1.3×10 <sup>-6</sup>	1.88	0.06	0.04	INA
ER density measures								
ED donsity	Riffle ER,	LZS	78	$-1.3 \times 10^{-6}$	0.70	0.49	< 0.01	NΛ
EK delisity	pool ER	HZS	78	7.1×10 <sup>-5</sup>	0.83	0.41	< 0.01	INA
Mean density of riffle FR	Riffle FR	LZS	78	-1.1×10 <sup>-4</sup>	12.3	<0.001	0.66	79
Weah density of Three EK	KIIIIC LK	HZS	78	-4.2×10 <sup>-4</sup>	5.02	<0.001	0.24	/ a
Mean density of riffle FR	Pool lengths,	LZS	77	$-3.2 \times 10^{-4}$	22.4	<0.001	0.87	7c
Weath density of Three Erc	riffle ER	HZS	77	$-1.9 \times 10^{-3}$	17.1	<0.001	0.79	70
Pool-riffle ER density	Riffle ER	LZS	78	0.075	11.9	<0.001	0.64	NA
ratio		HZS	78	0.093	4.97	< 0.001	0.24	1 11 1
Pool-riffle ER density	Pool lengths,	LZS	77	0.19	13.3	< 0.001	0.69	NA
ratio	riffle ER	HZS	77	0.38	9.71	<0.001	0.55	
Segments without ER								
Number of segments	Pool lengths	HZS	78	$-1.2 \times 10^{-3}$	6.18	<0.001	0.33	6a
without ER	-							
Maximum length without	Pool lengths	HZS	78	-1.8×10 <sup>-3</sup>	1.29	0.20	0.02	6b
ER								
Number of segments	۸11*	LZS	278	-3.1×10 <sup>-3</sup>	7.91	<0.001	0.18	60
without ER		HZS	278	$-3.9 \times 10^{-3}$	9.12	<0.001	0.23	UC

484 *Note.* t-tests indicate whether slopes differ from zero, significant tests are shown in bold text. See

485 supporting material for full description of simulations. \*Alternating riffles with ER (randomly drawn

486 from the negative-binomial distributions) and pools without ER of the same lengths.

# 487 3.3. (Q3) Using synthetic streams, which aspects of stream morphology are responsible for 488 driving differences in fractal properties?

489 Consistent with the study streams, D<sub>E</sub> of synthetic stream stretches was associated with 490 stream characteristics that relate to the spatial arrangement of ER, namely the number of 491 segments without ER, the maximum length without ER, and pool length. Without 492 manipulating stream characteristics (pool length, riffle length, or ER density), D<sub>E</sub> was 493 significantly different for simulations using the two ER distributions, i.e. low vs. high zero 494 streams (t = 34.2, p < 0.001). This suggests that the number of segments without ER influence D<sub>E</sub>, however the number of ER in pools and riffles could also contribute to this 495 result. Variation in  $D_E$  was significantly larger for the high-zero streams (Mean  $D_E \pm SD$ ; 496 497  $0.93 \pm 0.012$ ) compared with the low-zero streams ( $1.02 \pm 0.002$ ; M-SLR = 39.03, p < 498 0.001). By definition the number of segments without ER and the maximum length without 499 ER differed between the two ER distributions. However, as these simulations mimic the 500 natural systems, they also differed significantly in ER density (t = 69.8, p < 0.001) and 501 variance (M-SLR = 9.13, p = 0.003). Interestingly, the difference in ER variance was in the 502 reverse direction to the variation in D<sub>E</sub>, with the low-zero streams having higher ER variance 503 (Mean ER/m  $\pm$  SD; 10.90  $\pm$  0.51) than the high-zero streams (1.94  $\pm$  0.26). D<sub>E</sub> of simulated stream stretches was also negatively associated with the number of segments without ER for 504 505 high-zero streams with pool lengths manipulated to influence the number of segments 506 without ER (Table 2, Fig. 6a). The maximum length of segments without ER was not related 507 to  $D_E$  for these simulations (Table 2, Fig. 6b), which contradicts the results from the study 508 streams. However, systematic introduction of pools without ER resulted in a clear non-linear 509 pattern of D<sub>E</sub> with the maximum length of segments without ER (Fig. 6d). For these 510 simulations, minimum  $D_E$  resulted from pools (and riffles) of 50 – 75 m (10-15 segments) in 511 length and maximum  $D_E$  was observed with pools of 350 m (70 segments) in length.  $D_E$  was 512 also negatively associated with the length of pools, even when almost all segments contained 513 ER, as seen by manipulating pool length of low-zero streams (Table 2). The same trend was 514 seen for high-zero streams however, this increased pool length and the number of segments 515 without ER concurrently.

516 Relating to the number of ER, only ER density could be manipulated in simulations to create 517 certainty about cause and effect, whereas depth, slope and sediment characteristics could not

be simulated. Unlike the study streams,  $D_E$  was not correlated to overall ER density of

519 simulated stream stretches, as tested by manipulating the number of ER in pools and riffles in

520 combination (Table 2; Fig. 7).  $D_E$  decreased with ER density when only the number of ER in

521 riffles was manipulated (Table 2; Fig. 7a), however this negative relationship conflicts with

522 our expectations and the results from the study streams. This manipulation concurrently

523 created a gradient in the pool-riffle ER density ratio, which may have produced the

524 unexpected relationship. This is supported as further exaggeration of pool-riffle structure by

525 manipulations of ER density and pool length together, such that the highest ER densities were

526 paired with the longest pools creating the strongest relative difference (Fig 5c, d), increased

527 the magnitude of the slopes of  $D_E$  for both high-zero streams and low-zero streams. Taken

528 together, these results demonstrate that the spatial arrangement of ER (pool/riffle structure)

529 has a stronger effect on  $D_E$  than ER density.



Row 1) Manipulation of pool lengths



531 Figure 6. Using synthetic stream stretches to determine the influence of segments without ER on  $D_E$ . 532 (a-b) The number of segments without ER was altered by manipulating pool lengths of high-zero 533 streams (HZS: red triangles).  $D_E$  was associated with the number of segments without ER (a) but not 534 the maximum length of segments without ER (b). Low-zero streams (LZS) were not included in this 535 analysis as increasing pool lengths did not largely influence the number of segments without ER. (c-536 d) Further simulations whereby segments without ER were included by alternating pools without ER 537 and riffles with ER (using LZS (grey crosses) and HZS riffle distributions) for the length of the site 538 (1000 m). Pools and riffles were of length 5 (every second 5-m segment has no ER), 10 (two 5-m 539 segments with ER, two 5-m segments without ER), 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 540 and 500 m. In accord with the first set of simulations, increasing the number of segments without ER 541 resulted in a linear decline in  $D_{\rm E}(\mathbf{c})$ , however, the length of segments without ER resulted in a clear 542 non-linear relationship with  $D_E(\mathbf{d})$ . Loess curves are fit to non-linear relationships to illustrate 543 patterns. See Table 2 for summary of statistical tests.



Row 1) Manipulation of riffle ER density and pool length independently



545 Figure 7. Using synthetic stream stretches to determine the influence of pool-riffle structure on 546 entropy ( $D_E$ ).  $D_E$  is influenced by pool-riffle structure when described by the relative difference in ER 547 density between pools and riffles and the relative length of pools and riffles. Simulations used the two 548 ER distributions: low-zero streams, LZS (grey crosses) and high-zero streams, HZS (red triangles) 549 with manipulations of (a) ER density in riffles or (b) pool length, independently of each other, or (c-550 d) both ER density in riffles and pool length *in combination* (i.e. panels c and d are drawn from a 551 single set of simulations). ER density in riffles and pool length independently influenced LZS and 552 HZS simulations, however a greater effect (steeper slope) occurred when both stream characteristics 553 were manipulated. See Table 2 for summary of statistical tests.

### 554 **4. Discussion**

# 4.1. (Q1) Are longitudinal ER distributions, at scales of up to a kilometre, self-similar (fractal)?

557 Fractal dimensions can be used to measure the complexity of ER distributions in streams in a 558 way that is scale independent. Six streams exhibited fractal behaviour (self-similarity) 559 according to the criteria developed by *Seuront* [2009] for both the box-counting and entropy 560 methods. It seems likely that fractal dimensions may also measure complexity at scales smaller than the reach (e.g. single riffles) and beyond our reach lengths (regional scales). 561 However, we provide some evidence that our study streams exhibit multiple scaling 562 behaviour as discerned by breaks in the slope of the log-log plots, which may indicate that 563 different scaling regions (ranges of  $\delta$ ) are described by different fractal dimensions. The 564 position of the breaks varied among the streams and could be associated with the scale of 565 566 bedform spacing (Figure S5). This could also be an artifact of reaching the limits of the data 567 sets, and so further exploration is needed to determine whether a single or multiple 568 relationship is better linked to channel geomorphology. Because at least three orders of 569 magnitude are recommended for fractal analysis [Falconer, 1993] and our datasets confirm 570 fractal behaviour, even when these breakpoints are crossed, we will continue the discussion 571 with the box-counting and entropy fractal dimensions calculated using the entire dataset to 572 ensure the relevance of the fractal analysis.

573 Fractal dimensions varied among the six streams and with the method of calculation. The 574 box-counting method only differentiated between streams with many (1 stream), few (1 stream) and no (4 streams) segments without ER. This is problematic for comparison of 575 upland streams where the vast majority of segments in streams have at least one ER, which 576 results in  $D_B \sim 1$  (the value for a straight line). With a higher resolution, the entropy method 577 separated all six streams, indicating that the arrangement of ER varies among the streams. D<sub>E</sub> 578 579 ranged from 0.91 to 1.05, and while this is likely a narrow distribution out of a greater range 580 of fractal dimensions that are possible for real rivers, this range provides some opportunity to 581 relate ER arrangement to channel morphology and sediment characteristics using the study 582 streams. However, multicollinearity in the natural systems necessitates the use of synthetic 583 streams to isolate the causes of variation in D<sub>E</sub> among streams.

4.2. (Q2) Do fractal dimensions capture differences in ER distributions between rivers
associated with differences in channel morphology and sediment characteristics?

586 Our field data showed that entropy fractal dimensions correlated well with stream 587 characteristics that relate to the number and spatial arrangement of ER. Concerning the 588 number of ER, D<sub>E</sub> was related to ER density and channel morphology and sediment 589 characteristics that influence the propensity for large rocks to emerge. As expected, D<sub>E</sub> was 590 positively correlated with the c-axis and particle equancy (c/a) of the bed materials, as 591 characterised by random sampling of submerged sediments. Submerged rocks likely provide 592 a better characterization of the bed materials than ER (which were not correlated with  $D_E$ ), 593 which capture only a subset of the sediment distribution. At a given water depth, grains that 594 have a large c-axis and are more equant are more likely to be emergent because most grains 595 rest on their a-b plane, so that the c-axis then determines the elevation of the upper surface of the particle. Notwithstanding differences in water depth between streams, larger mean c-axis 596 597 and equancy are therefore likely to increase the abundance of ER. Indeed, our data showed a significant positive correlation of both c-axis and particle equancy with ER density. This 598 599 suggests that rock shape may influence the likelihood of emergence more than size (b-axis was not correlated with D<sub>E</sub>). However, there is also an expectation that c-axis varies with 600 601 channel slope, such that larger bed materials are more common in steeper channels (because 602 slope is generally adjusted to generate shear sufficient to transport the coarsest materials) and that as channel length (or drainage area) increases, so the maximum particle size declines due 603 604 to sorting and abrasion processes. As slope also influences bedform development causation 605 may be linked more closely to the spatial arrangement of ER.

606 Our expectation that the fractal dimension would be related to the spatial arrangement of ER 607 was supported by stream characteristics that identify the presence of pool and riffle structure. 608 These included negative correlations of  $D_E$  with the number and maximum length of segments without ER, mean proportional pool length, and slope. Segments without ER 609 610 provide the first evidence of pool-riffle structure however, at this coarse scale, pools are simplified to segments without ER, which rarely occurs. At a finer scale, all streams showed 611 612 some degree of pool-riffle structure, with riffles in each stream having higher ER density 613 compared to pools. This structure was greater for three streams (Kelphope, Faseny, and 614 Steavenson) where pools had fewer than half as many ER as riffles (<0.5:1; pool:riffle ER ratio); this resulted in the lowest values for D<sub>E</sub>. Accordingly, these streams have well-615

616 developed macro-scale bedforms like alternating pools and true riffles (e.g. Steavenson) or alternating pools and plane bed sections (e.g. Faseny). For the remaining streams (Snobs, 617 618 Little and Dye), the ratio of ER in pools and riffles was less pronounced (>0.7:1) and led to 619 the highest values of D<sub>E</sub>. Topographic bedform development is weak in these streams (e.g. 620 Dye and Snobs), and hence there is more chaotic, limited organisation of ER. In reaches 621 dominated by plane bed or with weak riffle-pool development, an absence of long pools 622 suggests that sediment storage dominates throughout, with little longitudinal topography. 623 This storage of sediment uniformly elevates the bed closer to the water surface and there is a 624 greater chance of rocks emerging. In contrast, where storage is organised into distinctive 625 topographic highs and lows, the propensity for large rocks to emerge is alternately higher and 626 lower. It is also notable that the steepest streams in each region (Dye, Snobs) are those with 627 the more irregular ER distributions whereas those on lower slopes (Faseny, Little, Steavenson) have better developed bedforms and ER organisation. This reflects the well-628 629 known association between channel steepness and bedform type, with the development of true riffle and pool sequences on slopes typically below 1 % [e.g. Buffington and 630 631 Montgomery, 2013].

632 The study streams provide evidence that larger, more-equant grain sizes are associated with 633 greater abundance but weaker longitudinal organisation of ER. This reinforces the suggestion that steeper, and also smaller catchments, are more likely to be associated with high D<sub>E</sub>. This 634 635 study is the first to investigate the fractal behaviour of ER in streams and so comparisons 636 with previous studies are limited; however, relationships with catchment characteristics align 637 with previous work on fractal dimensions of river networks, which are related to runoff and 638 sediment yield [Yang and Shi, 2017], flood frequency [Zhang et al., 2015], climate [Wang et 639 al., 2009], and tectonic forces [Shen et al., 2011]. Ultimately, the number of emergent rocks 640 and the presence of well-developed topographic bedforms reflects the interplay of many underlying geomorphological processes. This creates difficulties when attempting to identify 641 642 relationships between those variables and other constructs, such as fractal dimension, because variables are not independent and mechanisms are difficult to disentangle. Simulating streams 643 644 provides an elegant way to address these issues.

# 645 4.3. (Q3) Using synthetic streams, which aspects of stream morphology are responsible for 646 driving differences in fractal properties?

647 The simulated synthetic streams provided a clear basis upon which to hypothesise 648 mechanisms that underpin the relationship between any individual measure of stream 649 morphology and fractal dimension. They revealed a clear relationship of decreasing fractal 650 dimensions, and therefore stronger longitudinal organisation of ER, with increased pool-riffle 651 structure (Fig. 7). To illustrate this clearly, pool-riffle structure was exaggerated by 652 increasing pool lengths, the number of segments without ER in pools, and ER density in 653 riffles relative to pools (each manipulated independently from each other and in combination). The strongest effects were seen when these characteristics were manipulated in 654 655 combination, and changes were greater than the sum of the individual effects. This shows that 656 fractal dimensions capture the dynamics of multiple characteristics making them more useful than measuring any single characteristic. The use of simulations to investigate these 657 658 dynamics allows unhindered interpretation of results and have provided clear complementary support for the arguments made using the empirical analysis for a relationship between pool-659 660 riffle structure and fractal dimension in the study streams.

Taken together, the field and simulated components of this study have helped to disentangle 661 662 the geomorphological processes that may generate variation in physical complexity. Fractal dimensions captured well-known patterns that arise from sediment sorting, abrasion, and 663 664 storage processes and are associated with longitudinal fluvial gradients. Upstream, steeper slopes, less well-developed bedform topography, and greater abundances of large rocks lead 665 666 to an irregular organisation of ER and high fractal dimension. Downstream, the development 667 of true riffle and pool sequences on low slopes and lower abundances of large rocks result in 668 stronger longitudinal organisation of ER and low fractal dimension. Variation in D<sub>E</sub> values 669 between streams relates to aspects of stream morphology and sediment character. As such, the entropy fractal dimension is a promising measure of physical complexity that captures 670 671 differences in ER distributions and organisation driven by geomorphological processes. D<sub>E</sub> is, 672 therefore, a useful metric in both geomorphological and ecological studies, which frequently 673 rely on measurements of geomorphological characteristics to explain ecological patterns.

674 4.4. Fractal dimensions across ecosystems

675 Landscape complexity generates patchiness in environmental conditions (e.g. flow resistance,

676 turbulence, etc.) and resources (e.g. living spaces and food) that drive ecological processes29

30

[*Huffaker*, 1958]. This patchiness can facilitate or impede dispersal of organisms, abundance
and persistence of species, and interactions among species. Entropy, describing the
complexity of the physical landscape, is likely to be related to these ecological processes with
the added benefit of being comparable across disciplines and potentially across subject
matter. The entropy fractal dimension may therefore be useful for bridging the gap between
ecology and geomorphology, enabling general questions of assembly to be tested across
ecosystem types.

684 Self-similarity allows fractal dimension metrics to be transferred across ecosystem types even 685 when they are measured at different scales. While we have shown that the entropy fractal dimension may be a useful method to describe the landscape complexity of streams, the box-686 687 counting method is still more commonly used across many disciplines. Using a common method allows integration of ecosystem patterns and broad-scale hypotheses tests across 688 689 landscapes scales, and so corresponding results using box counting are presented in the 690 Supporting information (Table S6, Table S7 and Table S8). While integrating patterns among 691 ecosystems was not the purpose of this study, it is interesting to note that the complexity and 692 lack of structure of ER habitats ( $D_B = 0.92 - 1.00$ ) seen here is generally greater than that 693 reported for other systems where D has been measured in one dimension, including evergreen 694 forest canopy (0.78 - 0.95), deciduous forest canopy (0.69 - 0.95), understory shrubs (0.70 -695 0.81), grassland shrubs (0.61) and grassland grasses (0.80) [Denny and Nielsen, 2017; 696 Ritchie, 2009] (slopes of log-log plots are reported here to provide consistency with other papers). How this relates to ecological response variables in these habitats (i.e. species 697 698 diversity, dispersal) is an interesting avenue for further research. Ultimately, determining 699 whether fractal dimensions provide a meaningful description of physical landscapes across 700 scales and locations will be contingent on the overarching goals in any attempt to integrate 701 ecosystem patterns more broadly.

Despite its promise, calculating fractal dimensions to measure landscape complexity has its
challenges [*Halley et al.*, 2004]. These start with identifying an appropriate measure of fractal
dimension but also include choosing the range of scales over which to calculate fractal
dimensions, i.e. the smallest and largest scales of observation (in this study, 5 m segments
and 685 – 1000 m river lengths, respectively). We identified the largest scale for each stream
individually, recognising that these would then vary across streams, and used the associated
fractal dimension for each for comparative purposes. Alternative choices included applying a

- single set of delta values ( $\delta$ ) for comparison across all streams, using the scale at which the
- 710 maximum (or minimum) entropy value was calculated, among others. These different choices
- 711 have a material effect on the resultant fractal dimensions and the ranking of streams and,
- therefore, are likely to affect the outcome of comparative studies. Very little published
- 713 literature assists with these choices, despite their impact on the final fractal dimensions
- calculated. Thus, fractal dimensions should be applied thoughtfully, and additional guidance
- is needed to ensure that ecologists and geomorphologists apply the techniques in a
- 716 mathematically robust manner.

### 717 4.5. Conclusions:

718 Here, the entropy fractal dimension was a meaningful measure of the complexity of ER 719 distributions, whereas the box-counting method was less useful for comparisons among 720 upland streams where most segments had at least one ER. The entropy fractal dimension was principally driven by the development of well-defined bed topography, for example in the 721 722 form of pool-riffle sequences, as this affected longitudinal patterns of ER distribution and, to some degree, by rock size as this affected the propensity for ER to be abundant irrespective 723 724 of bedform topography. Due to this ability to reflect the physical characteristics of the environment, the entropy fractal dimension shows great potential to measure the complexity 725 726 of river systems in a way that is relevant to ecological processes, provided it is calculated 727 consistently across the systems of interest.

### 728 **5.** Acknowledgments:

- 729 Thank you, Vicky Mak and Korbinian Strimmer, for expert advice on calculating fractal
- dimensions. For advice on earlier versions of the manuscript, we thank Galen Holt. For
- assistance collecting field data, we thank Alena Glaister, Galen Holt, Richard Mason and
- Harry Sanders. Thank you also to the various landowners for allowing access to Scottish field
- sites and to Peter Grant for facilitating access to Snobs Ck. Australian field work was carried
- out in conjunction with a Research Permit (No. 10007855) under the National Parks Act
- 735 (Australia), from the Department of Environment, Land, Water and Planning (Victoria).

### 736 **6. Funding**

- 737 This research was supported by data that were originally funded by a grant from the Natural
- 738 Environment Research Council, UK (NE/E004946/1) and a Discovery grant from the
- Australian Research Council (DP 160102262).

### 740 **7. Data statement**

To foster transparency, our data is available on Deakin University's research repository.

### 742 8. References

Aubeneau, A. F., R. L. Martin, D. Bolster, R. Schumer, D. Jerolmack, and A. Packman (2015),

744 Fractal patterns in riverbed morphology produce fractal scaling of water storage times,

- 745 Geophysical Research Letters, 42(13), 5309-5315.
- 746 Barnes, J. B., I. P. Vaughan, and S. J. Ormerod (2013), Reappraising the effects of habitat
- structure on river macroinvertebrates, *Freshwater Biology*, *58*(10), 2154-2167.
- 748 Basharin, G. P. (1959), On a statistical estimate for the entropy of a sequence of
- independent random variables, *Theory of Probability & Its Applications*, 4(3), 333-336.
- 750 Bell, S. S., E. D. McCoy, and H. R. Mushinsky (2012), Habitat structure: the physical
- 751 *arrangement of objects in space,* Springer Science & Business Media.
- Blott, S. J., and K. Pye (2008), Particle shape: a review and new methods of characterization
  and classification, *Sedimentology*, 55(1), 31-63.
- Bouckaert, F. W., and J. Davis (1998), Microflow regimes and the distribution of
- 755 macroinvertebrates around stream boulders, *Freshwater Biology*, 40(1), 77-86.
- 756 Buffington, J., and D. Montgomery (2013), Geomorphic classification of rivers, in *Treatise on*
- 757 *Geomorphology; Fluvial Geomorphology*, edited by J. Shroder and E. Wohl, pp. 730-767,
- 758 Academic Press, San Diego, CA.
- 759 Butler, J. B., S. N. Lane, and J. H. Chandler (2001), Characterization of the structure of river-
- 760 bed gravels using two-dimensional fractal analysis, *Mathematical Geology*, *33*(3), 301-330.
- 761 Cooper, J. R., J. Aberle, K. Koll, and S. J. Tait (2013), Influence of relative submergence on
- spatial variance and form-induced stress of gravel-bed flows, *Water Resources Research*,
  49(9), 5765-5777.
- 764 Cuthbert, R. N., T. Dalu, R. J. Wasserman, A. Callaghan, O. L. Weyl, and J. T. Dick (2019),
- Using functional responses to quantify notonectid predatory impacts across increasingly
   complex environments, *Acta Oecologica*, *95*, 116-119.
- 767 Davies, A., A. McAdam, and I. Cameron (1986), Geology of the Dunbar District. Memoir for
  768 Sheet 33E and Part of Sheet 41 (Scotland).
- 769 Denny, C. K., and S. E. Nielsen (2017), Spatial heterogeneity of the forest canopy scales with
- the heterogeneity of an understory shrub based on fractal analysis, *Forests*, 8(5), 146-163.
- 771 Downes, B. J., T. J. Entwisle, and P. Reich (2003), Effects of flow regulation on disturbance
- frequencies and in-channel bryophytes and macroalgae in some upland streams, *River*
- 773 *Research Applications*, 19(1), 27-42.
- 774 Downes, B. J., P. S. Lake, E. Schreiber, and A. Glaister (1998), Habitat structure and
- regulation of local species diversity in a stony, upland stream, *Ecological Monographs*, 68(2),237-257.
- Falconer, K. (1993), *Fractal geometry: mathematical foundations and applications*, John
  Wiley & Sons.
- 779 Ferreiro, N., C. Feijoó, A. Giorgi, and L. Leggieri (2011), Effects of macrophyte heterogeneity
- and food availability on structural parameters of the macroinvertebrate community in a
- 781 Pampean stream, *Hydrobiologia*, 664(1), 199-211.
- 782 Gordon, N. D., T. A. McMahon, B. L. Finlayson, C. J. Gippel, and R. J. Nathan (2004), Stream
- 783 *Hydrology: an Introduction for Ecologists,* 2nd ed., John Wiley and Sons, West Sussex.
- Halley, J., S. Hartley, A. Kallimanis, W. Kunin, J. Lennon, and S. Sgardelis (2004), Uses and
- abuses of fractal methodology in ecology, *Ecology Letters*, 7(3), 254-271.

- Hayes, J., and I. Jowett (1994), Microhabitat models of large drift-feeding brown trout in
- three New Zealand rivers, North American Journal of Fisheries Management, 14(4), 710-725.
- Hoover, T. M., J. S. Richardson, and N. Yonemitsu (2006), Flow-substrate interactions create
- and mediate leaf litter resource patches in streams, *Freshwater Biology*, *51*(3), 435-447.
- 790 Huffaker, C. (1958), Experimental studies on predation: dispersion factors and predator-prey
- 791 oscillations, *Hilgardia*, 27(14), 343-383.
- Jeffries, M. (1993), Invertebrate colonization of artificial pondweeds of differing fractal
- 793 dimension, *Oikos*, *67*, 142-148.
- 794 Krishnamoorthy, K., and M. Lee (2014), Improved tests for the equality of normal
- 795 coefficients of variation, *Computational Statistics*, *29*(1-2), 215-232.
- The spatial characterization of turbulence around large
- roughness elements in a gravel-bed river, *Geomorphology*, *102*(3-4), 542-553.
- Lancaster, J. (2000), Geometric scaling of microhabitat patches and their efficacy as refugia
   during disturbance, *Journal of Animal Ecology*, *69*(3), 442-457.
- Lancaster, J., B. J. Downes, and P. Reich (2003), Linking landscape patterns of resource
- distribution with models of aggregation in ovipositing stream insects, *Journal of Animal Ecology*, 72(6), 969-978.
- 803 Lancaster, J., B. J. Downes, and A. Arnold (2010), Environmental constraints on oviposition
- 804 limit egg supply of a stream insect at multiple scales, *Oecologia*, *163*(2), 373-384.
- Lancaster, J., B. J. Downes, R. E. Lester, and S. P. Rice (2020), Avoidance and aggregation
- create consistent egg distribution patterns of congeneric caddisflies across spatially variable
   oviposition landscapes, *Oecologia*, *192*(2), 375-389.
- Lee, A., A. F. Aubeneau, and M. B. Cardenas (2020), The sensitivity of hyporheic exchange to
- 809 fractal properties of riverbeds, *Water Resources Research*, *56*(5), e2019WR026560.
- Lester, R. E., W. Wright, and M. Jones-Lennon (2007), Does adding wood to agricultural
- 811 streams enhance biodiversity? An experimental approach, Marine and Freshwater Research,
- 812 *58*(8), 687-698.
- 813 Mandelbrot, B. (1967), How long is the coast of Britain? Statistical self-similarity and
- 814 fractional dimension, *Science*, *156*(3775), 636-638.
- 815 Mantilla, R., B. M. Troutman, and V. K. Gupta (2010), Testing statistical self-similarity in the
- 816 topology of river networks, *Journal of Geophysical Research: Earth Surface*, 115(F3).
- 817 Marsden, M. A. H. (1973), Palaeozoic evolution of east-central Victoria: Regional Guide to
- 818 Victorian Geology, edited by M. McAndrew and M. A. H. Marsden, pp. 175-201.
- 819 Marwick, B., and K. Krishnamoorthy (2018), cvequality: Tests for the Equality of Coefficients
- 820 of Variation from Multiple Groups, *R software package version 0.1, 3*.
- 821 Meager, J. J., and T. A. Schlacher (2013), New metric of microhabitat complexity predicts
- species richness on a rocky shore, *Marine Ecology*, *34*(4), 484-491.
- 823 Miller, G. (1955), Note on the bias of information estimates, in *Information Theory in*
- 824 Psychology: Problems and Methods, edited by H. Quastler, pp. 95–100, The Free Press,
- 825 Glencoe.
- 826 Monsalve, A., E. M. Yager, and M. W. Schmeeckle (2017), Effects of bed forms and large
- 827 protruding grains on near-bed flow hydraulics in low relative submergence conditions,
- S28 Journal of Geophysical Research: Earth Surface, 122(10), 1845-1866.
- 829 Montgomery, D. R., and J. M. Buffington (1997), Channel-reach morphology in mountain
- drainage basins, *Geological Society of America Bulletin*, 109(5), 596-611.

- Nikora, V. I. (1991), Fractal structures of river plan forms, *Water Resources Research*, 27(6),
  1327-1333.
- 833 Ossola, A., A. K. Hahs, M. A. Nash, and S. J. Livesley (2016), Habitat complexity enhances
- comminution and decomposition processes in urban ecosystems, *Ecosystems*, *19*(5), 927941.
- Papanicolaou, A., A. G. Tsakiris, M. A. Wyssmann, and C. M. Kramer (2018), Boulder array
- effects on bedload pulses and depositional patches, *Journal of Geophysical Research: Earth Surface*, 123(11), 2925-2953.
- 839 Petersen, I., and A. G. Hildrew (2003), Emergence of *Leuctra nigra* (Plecoptera) from a 840 southern English stream, *Archiv für Hydrobiologie*, *158*(2), 185-195.
- 841 R Core Team (2019), R: A language and environment for statistical computing
- 842 Reich, P. (2004), Patterns of composition and abundance in macroinvertebrate egg masses
- 843 from temperate Australian streams, *Marine and Freshwater Research*, 55(1), 39-56.
- Rice, S. P. (2017), Tributary connectivity, confluence aggradation and network biodiversity, *Geomorphology*, *277*, 6-16.
- 846 Rice, S. P., and M. Church (1996), Bed material texture in low order streams on the Queen
- 847 Charlotte Islands, British Columbia, *Earth Surface Processes Landforms*, *21*(1), 1-18.
- 848 Rice, S. P., and M. Church (1998), Grain size along two gravel-bed rivers: statistical variation,
- spatial pattern and sedimentary links, *Earth Surface Processes Landforms: The Journal of the British Geomorphological Group*, *23*(4), 345-363.
- 851 Rinaldo, A., I. Rodriguez-Iturbe, R. Rigon, R. L. Bras, E. Ijjasz-Vasquez, and A. Marani (1992),
- 852 Minimum energy and fractal structures of drainage networks, 28(9), 2183-2195.
- 853 Ritchie, M. E. (2009), Scale, heterogeneity, and the structure and diversity of ecological
- 854 *communities,* Princeton University Press.
- Rodríguez-Iturbe, I., and A. Rinaldo (2001), *Fractal river basins: chance and self-organization*,
  Cambridge University Press.
- 857 Sapozhnikov, V., and E. Foufoula-Georgiou (1996), Self-affinity in braided rivers, *Water*
- 858 *Resources Research*, *32*(5), 1429-1439.
- 859 Seuront, L. (2009), *Fractals and multifractals in ecology and aquatic science*, CRC Press.
- 860 Seuront, L., and Y. Lagadeuc (1997), Characterisation of space-time variability in stratified
- and mixed coastal waters (Baie des Chaleurs, Québec, Canada): application of fractal theory,
   *Marine Ecology Progress Series*, 159, 81-95.
- 863 Shen, X., L. Zou, G. Zhang, N. Su, W. Wu, and S. Yang (2011), Fractal characteristics of the
- 864 main channel of Yellow River and its relation to regional tectonic evolution, *Geomorphology*, 865 *127*(1-2), 64-70.
- 866 Singh, A., S. Lanzoni, P. R. Wilcock, and E. Foufoula-Georgiou (2011), Multiscale statistical
- characterization of migrating bed forms in gravel and sand bed rivers, *Water Resources Research*, 47(12).
- 869 Singh, A., K. Fienberg, D. J. Jerolmack, J. Marr, and E. Foufoula-Georgiou (2009),
- 870 Experimental evidence for statistical scaling and intermittency in sediment transport rates,
- 871 Journal of Geophysical Research: Earth Surface, 114(F1).
- 872 Stewart, M. T., S. M. Cameron, V. I. Nikora, A. Zampiron, and I. Marusic (2019), Hydraulic
- 873 resistance in open-channel flows over self-affine rough beds, *Journal of Hydraulic Research*,
- 874 *57*(2), 183-196.
- 875 Strobl, C. (2005), Variable selection in classification trees based on imprecise probabilities,
- 876 *ISIPTA*, *5*, 340-348.
- 877 Tarboton, D. G. (1989), The analysis of river basins and channel networks using digital
- 878 terrain data, Massachusetts Institute of Technology.
- Wang, B., F. Tian, and H. Hu (2009), Relationship between fractal dimension of river
- networks and their climates, *Journal of Tsinghua University Science Technology*, 49(12),
  1948-1953.
- 882 Warfe, D., L. Barmuta, and S. Wotherspoon (2008), Quantifying habitat structure: surface
- convolution and living space for species in complex environments, *Oikos*, *117*(12), 17641773.
- Wiens, J. A. (1989), Spatial scaling in ecology, *Functional Ecology*, *3*(4), 385-397.
- Wolman, M. G. (1954), A method of sampling coarse river-bed material, *EOS, Transactions American Geophysical Union*, *35*(6), 951-956.
- 888 Wu, F.-C., C.-K. Wang, and G.-H. Huang (2018), Delineation of gravel-bed clusters via
- 889 factorial kriging, *Geomorphology*, 308, 161-174.
- 890 Yang, H., and C. Shi (2017), The fractal characteristics of drainage networks and erosion
- 891 evolution stages of ten kongduis in the upper reaches of the Yellow River, *Journal of*
- 892 *Resources and Ecology*, 8(2), 165-173.
- 293 Zhang, S., Y. Guo, and Z. Wang (2015), Correlation between flood frequency and
- 894 geomorphologic complexity of rivers network–a case study of Hangzhou China, *Journal of*
- 895 *hydrology*, *527*, 113-118.
- 896 Zhong, L., F. Zeng, and G. Xu (2012), Comparison of fractal dimension calculation methods
- 897 for channel bed profiles, *Procedia Engineering*, 28, 252-257.

898



Water Resources Research

Supporting Information for

## Using fractals to describe ecologically-relevant patterns in distributions of large rocks in streams

G. K. Dwyer<sup>1</sup>, C. R. Cummings<sup>1</sup>, S. P. Rice<sup>2</sup>, J. Lancaster<sup>3</sup>, B. J. Downes<sup>3</sup>, L. Slater<sup>4</sup>, and R. E Lester<sup>1</sup>

1Centre for Regional and Rural Futures, Deakin University, Locked Bag 20000, Geelong, Vic. 3220, Australia.

2 Department of Geography, Loughborough University, Loughborough, LE11 3TU, U.K.

3 School of Geography, The University of Melbourne, Victoria, Australia

4School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

### Contents of this file

Text S1 - Calculating Fractal Dimensions

- Text S2 Simulating Emergent Rocks in Streams
- Figure S1 Four Methods of Calculating Fractal Dimensions
- Figure S2 Frequency Distributions of ER
- Figure S<sub>3</sub> Example of Four Synthetic Stream Stretches
- Figure S4 Log-log plots of the box counting dimension and the entropy dimension for ER counts in the six study streams.
- Figure S5 Exploring Multiple Scaling.
- Table S1 Rationale for directional hypotheses.
- Table S2 Proportions of Channel Morphologies in the Six Study Streams
- Table S3 Geographical and Physical Features of the Study Streams along the Survey Lengths
- Table S4 Correlation Tests among Variables Describing Stream Characteristics
- Table S5 Fractal Dimension Results using the Box-Counting and Entropy Methods.
- Table S6 Correlation Tests of the Study Stream Characteristics with Fractal Dimensions
- Table S7 Linear Regression Tests of  $Log(\delta)$  versus  $Log(N_X)$
- Table S8 Linear regression of the Stream Characteristics with Fractal Dimensions using Synthetic Stream Stretches

### Additional Supporting Information (Files uploaded separately)

Captions for Script S1 - R scripts for calculating fractal dimensions Captions for Script S2 - R scripts for simulating synthetic stream stretches

## Text S1.

## **Calculating fractal dimensions**

The calculation of the fractal dimensions involved the following five steps:

- 1) Calculate N<sub>X</sub> for all possible values of  $\delta_i$ ;
- 2) Log-transform linear regression of  $\delta$  and N<sub>x</sub>;
- 3)  $R^2 SSR$  Procedure;
- 4) Zero-slope Procedure;
- 5) Compensated-Slope Procedure

# Step 1: Calculation of N<sub>X</sub> for all possible values of $\delta$

All possible values of  $\delta$  were identified. This is simply a sequence of values from the smallest scale of observation (here, 5 m segments) to the largest scale of observation (here, the study sites ranged from 685 – 1000 m) increasing in increments of the smallest scale. For example, the box sizes used for Dye are as follows:

δ: 5, 10, 15, ..., 990, 995, 1000.

For each value of  $\delta$ , the dataset was divided into boxes of size  $\delta$  to calculate N<sub>x</sub>. As the data collected for each stream is 1-dimensional, the stream section is divided in only that one direction, along the length of the site. For the box-counting method, N<sub>B</sub> is calculated as the number of occupied boxes. Irregularities in the ER distributions of a size smaller than  $\delta$  are disregarded in box sizes greater than  $\delta$ (Seuront, 2009). For example, sections without rocks that were 10 m in length (2 segments) would be ignored at values of delta greater than 15 m (3 segments). As a result, the values of N<sub>B</sub> using the boxcounting method on the Dye dataset, which does not contain any segments without ER, are as follows:

Dye values of N<sub>B</sub>: 200, 100, 67, ..., 2, 2, 1.

For the entropy method, N<sub>E</sub> is calculated using the proportions of the total ER, following weighting and correction for bias in the estimate of entropy as per Miller (1955) and Basharin (1959). The weighted estimate (*H*; theoretical Shannon entropy) was calculated as:

$$H = -\sum p \, \times \log_e p$$

where *p* is the proportion of total ER in each box. The correction for bias in the estimate of entropy as per Miller (1955) and Basharin (1959) was achieved according to the following equation:

$$\log N_E = \propto H + \frac{(s-1)}{2N} \times \log_2 e$$

where s is sample size (number of boxes) and N is the total number of ER. The constant  $\log_2(e)$  has a value of 1.442695.  $\propto$  indicates that we neglect terms of the order of  $O\left(\frac{1}{N^2}\right)$ ; this correction provides a decent approximation of the true entropy value, but only for sufficiently large sample sizes (Strobl, 2005). Further, the correction is negligible for N to infinity (Strobl, 2005). For example, the values of N<sub>E</sub> using the entropy method on the Dye dataset are as follows:

Log N<sub>E</sub>: 4.97, 4.31, 3.92, ..., 0.00, 0.00, 0.00.

The list of  $\delta$  and associated N<sub>x</sub> was condensed to remove duplicate and inconsequential values of N<sub>x</sub>. Using the Dye dataset for instance,  $\delta$  values of 100, 105, and 110 all result in N<sub>B</sub> of 10 and  $\delta$  values between 500 and 995 all result in an N<sub>B</sub> of 2. Further using the entropy method,  $\delta$  values were removed where the resulting value for log N<sub>E</sub> equalled 0, as this indicates that the entire data set fit into fewer than two boxes. The final condensed list (of  $\delta$  and N<sub>x</sub>) is used in the remaining four steps to provide an estimate of fractal dimensions (D<sub>x</sub>) including Seuront's (2009) three-step procedure to confirm the presence of fractal-like properties in the ER distributions.

## Step 2: Log-transformed linear regression

Simple linear regression of  $log(\delta)$  and  $log(N_x)$  provides a first estimate of the fractal dimension. The slope of the linear regression estimated over the values of  $\delta$  provides an estimate of the scaling exponent (D<sub>x</sub>) (Fig. S1A)

# Step 3: R<sup>2</sup> – SSR Procedure

The R<sup>2</sup> – SSR Procedure estimates D<sub>x</sub> by using the values of  $\delta$  which simultaneously maximised the coefficient of determination (R<sup>2</sup>) and minimized the total sum of the squared residuals (SSR) (Seuront & Lagadeuc, 1997) (Fig. S1B). This was achieved by looping over the list of  $\delta$  and N<sub>x</sub> with a 'regression window'. Similar to the 'boxes' used to loop over the ER counts to calculate N<sub>x</sub>, the regression window looped over the list of  $\delta$  and N<sub>x</sub> to allow calculation of the R<sup>2</sup> and SSR for each window. The range of window sizes used here ranged from a minimum of six data points to a maximum that contained the entire list. The R<sup>2</sup> – SSR criterion was only satisfied if the largest R<sup>2</sup> value and smallest SSR value were produced from the same regression window (Seuront & Lagadeuc, 1997). An estimate of D<sub>x</sub> was then calculated from the slope of the points contained in the regression window that satisfied the R<sup>2</sup> – SSR criterion. This process ensures that artefacts of linearity in the log-log plot (e.g. due to power-law relationships) are avoided.

## Step 4: Zero-slope Procedure

The Zero-slope Procedure provides an estimate of  $D_x$  from the intersection of a fitted line of zero-slope with the axis describing the derivative of  $log(N_x)$  with respect to  $log(\delta)$  (Fig. 1C). Again a sliding regression window is used to determine whether the slope of the derivative of  $log(N_x)$  with respect to  $log(\delta)$  for any regression window is significantly different from zero. To satisfy the criterion of the Zero-Slope Procedure, the slope of all windows must not be significantly different from zero. An estimate of  $D_x$  is calculated from the intercept, using the largest window where the slope was not significantly different to zero. This procedure has a disadvantage in that enhanced noise is generated by taking the first derivative of any linear trend which causes problems for standard statistical procedures; however this is overcome in the following compensated-slope procedure.

# Step 5: Compensated-Slope Procedure

The Compensated-Slope Procedure estimates  $D_x$  from a range of compensated exponents. The compensated exponent which best estimates  $D_x$  is determined by the plot of log [ $\delta^C \times \delta^{-D_x}$ ] versus log ( $\delta$ ), where the slope is closest to zero (Fig. 1D). The estimate of  $D_x$  used for this procedure is produced from the linear regression of log( $\delta$ ) and log( $N_x$ ) in step 2. We used 101 values of C that ranged from o to  $D_x+1$ , to allow an excess of values extending into the next dimension. The value of C whereby the slope was closest to zero was used as the  $D_x$  estimate in the in the current paper, as plateau behaviour in this relationship is a manifestation of scaling and therefore should not be the result of random non-fractal structure (i.e. an artefact of the data). The Compensated-Slope criterion was only satisfied if the slope was not significantly different from zero.

As stated by Seuront (2009), fractal-like behaviour is only confirmed if the criteria of two of the three test procedures described above are satisfied. R scripts for the calculation of the fractal dimensions are found in Supporting Information Script S1.

## References

- Basharin, G. P. (1959). On a statistical estimate for the entropy of a sequence of independent random variables. *Theory of Probability & Its Applications*, 4(3), 333-336.
- Miller, G. (1955). Note on the bias of information estimates. . Glencoe, IL: Free Press.
- Seuront, L. (2009). Fractals and multifractals in ecology and aquatic science: CRC Press.
- Seuront, L., & Lagadeuc, Y. (1997). Characterisation of space-time variability in stratified and mixed coastal waters (Baie des Chaleurs, Québec, Canada): application of fractal theory. *Marine Ecology Progress Series*, 159, 81-95.
- Strobl, C. (2005). Variable Selection in Classification Trees Based on Imprecise Probabilities. Paper presented at the ISIPTA.

## Text S<sub>2</sub>.

## Simulating emergent rocks in streams

Synthetic stream stretches were simulated using base R (R Foundation for Statistical Computing, Vienna, Austria) using the characteristics observed in the six study streams. These characteristics included: 1) site length, 2) feature (riffle and pool) lengths and 3) the number of ER in the 5-m segments.

To produce the synthetic stream stretches: First, a sequence of feature lengths was produced by alternately selecting riffle and pools lengths (X<sub>rL</sub> and X<sub>pL</sub>) from log-normal distributions, calculated based on the study stream data, until the cumulative length exceeded stream length (e.g. a sequence may include 3pL, 7rL, 10pL, 21rL, 4pL, 22rL, 3pL, 7rL, 6pL, 7rL..., indicating that the first feature was a pool of length 3m and the next feature was a riffle of length 7 m, and so on). The parameters of these distributions were  $\mu$  = 1.372 and 2.049, and  $\sigma$  = 0.571 and 0.829 for pools and riffles, respectively. Next, the sequence of riffle and pool lengths was translated into a binary sequence, with each digit (r or p) representing a single 5-m segment (e.g. for the above example, the binary sequence would begin with p, p, p, r, r, r, r, r, r, r, r, and so on). Finally, digits (r and p) of the binary sequence were replaced with values selected from negative binomial distributions based on the ER numbers in either riffles or pools of the study streams. These values represent the number of ER in each 5-m segment (e.g. 36<sub>PER</sub>, 11<sub>PER</sub>, 2per, 71rer, 48rer, 10rer, 99rer, 74rer, 16rer, 30rer..., indicates that the first 5-m segment in the example sequence above was a pool with 36 ER, the next segment was also part of the same pool but with 11 ER, given that the first pool extended for three segments in the stream length sequence, and so on). Two riffle/pool sets of ER distributions were used for the simulations, one fitted from the low-zero streams (low-zero streams; including Kelphope and Faseny) and one from the high-zero streams (high-zero streams; including Dye, Steavenson, Snobs, and Little) (Fig. S2). The parameters of the ER distributions for the low-zero streams were: size = 1.28 and 1.64, and  $\mu$  = 43.2 and 60.1 for riffles and pools, and for the high-zero streams: size = 0.266 and 0.638 and  $\mu$  = 5.46 and 14.8 for riffles and pools.

To determine the effect of each stream characteristic on fractal dimension, each characteristic was manipulated individually, leading to a total of 640 synthetic streams (see Fig. S<sub>3</sub> for an example). A default of 1000 m was used for the length of synthetic stream stretches. However, to test for any influence of site length on fractal dimensions (D<sub>x</sub>) site length 'variants' of 685, 765, 845, 920, and 1000 m were used, which spans the observed range. Variants of feature lengths were produced using multipliers (1, 2, 4, and 8) on riffle and pool lengths separately. Variants of ER density were produced using multipliers (1, 2, 4, and 8) on riffle ER counts separately, and pool and riffle ER counts simultaneously. Pool length and riffle ER density were also manipulated simultaneously (using multipliers 1, 2, 4, and 8 on both; Fig. S<sub>3</sub>). Further simulations were produce by adding systematic features with segments containing zero ER (including features of length 1 [every second segment has zero ER], 2 [two segments with ER, two segments with zero ER], 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90 and 100), Each set of simulations included 10-20 replicates of each variant; i.e. for the set of simulations for which site length was manipulated, 5 variants (685, 765, 845, 920, and 1000 m) were produced, with 10 replicates of each, equating to a total of 50 synthetic stream stretches.

R scripts for simulating synthetic stream stretches are found in Supporting Information (Script S2).



**Figure S1.** Four methods of calculating fractal dimensions using the box-counting method with the Dye dataset for example. (a) Log-transformed linear regression of  $\delta$  and N<sub>B</sub>. (b) The R<sup>2</sup> – SSR Procedure. Each point shows R<sup>2</sup> and the SSR for each regression window. Here, the R<sup>2</sup> – SSR criterion were not met as the regression window with the highest R<sup>2</sup> (red shaded point) and the regression window with the lowest SSR (blue shaded point) were not one and the same. (c) The Zero-slope Procedure. Here the derivative of log(N<sub>B</sub>) is plotted against log( $\delta$ ). The red solid line gives the best estimate of D<sub>x</sub> and the blue dashed line provide confidence intervals based on the largest and smallest plausible values of D<sub>x</sub> for these data (d) The Compensated-Slope Procedure. Here log [ $\delta^C \times \delta^{-D_X}$ ] is plotted against log( $\delta$ ). The compensated exponent which best estimates D<sub>x</sub> is where the slope is closest to zero (blue filled points). Methods **b** – **d** are required for Seuront's (2009) three-step procedure to confirm the presence of fractal-like properties.







**Figure S3.** Example of four synthetic stream stretches where pool lengths and riffle ER density have been manipulated simultaneously using multipliers (M) of 1, 2, 4, or 8. Segments of pools (blue) and riffles (black) are 5 m in length and span a 1000 m stream stretch.



**Figure S4. (a)** Log-log plots and **(b)** a comparison of the corresponding absolute slopes for the box counting dimension (left columns) and the entropy dimension (right columns) for ER counts in the six study streams. Significant differences among streams are signified with lower-case letters (e.g. for the box-counting method Kelphope was significantly different to all other streams) and were determined using post hoc analysis tests using the 'lstrends' r package. Asterisks signify a significant difference

from a slope of one, tests of which were executed using two-tailed linear regression analyses with an offset term fixed at -1. Equations are presented in log-transformed units.



**Figure S5.** Exploring multiple scaling. **(a)** Log-log plots for the box counting dimension (left columns) and the entropy dimension (right columns) for ER counts in the six study streams. Data were divided at breakpoints that were identified using the r package 'segmented' that identifies where the linear relation changes using bootstrap restarting. Significant differences between the small- (blue lines and equations) and large-scale (red lines and equations) slopes for each stream were determined using post

hoc analysis tests using the 'lstrends' r package. Equations are presented in log-transformed units. Breakpoints were significantly related to pool-riffle structure as seen with the relation between (b) pool lengths and (c) pool-riffle ER density ratio. Breakpoints were not related to channel width or stream length.

Characteristic (Expected direction)	Propensity for rocks to emerge and thence the number of ER	Bedform development and thence the spatial organisation of ER
Stream morphology		
Median depth (Negative)	Rocks are less likely to be emergent in deeper topography (fewer ER)	Depth increases downstream and is associated with better bedform development
Stream slope (Positive)	Greater slopes are associated with larger bed materials (more ER)	Slope decreases downstream and is associated with better bedform development
Mean proportional pool length (Negative)	Longer pool length implies greater occurrence of deeper topography (fewer ER)	Longer pool lengths imply stronger organisation of bed topography
Mean proportional riffle length (Positive)	Longer riffle-like length implies greater occurrence of shallow topography (more ER)	ΝΑ
Number of pool-riffle transitions (Negative)	ΝΑ	The number of pool-riffle transitions implies greater bed form development and structure
Bed sediment		
Axis B (Positive)	For a given water depth, larger bed materials are more likely to be emergent (more ER)	Larger bed materials are more common in steeper channels. Slope decreases downstream and is associated with reduced bedform development
Axis C (Positive)	Larger c-axis increases elevation above the average bed surface because most grains rest on their a-b plane (more ER)	Larger bed materials are more common in steeper channels. Slope decreases downstream and is associated with reduced bedform development
Equancy ratio (C/A) (Positive)	Greater equancy increases elevation above the average bed surface (more ER)	ΝΑ
Flatness ratio (C/B) (Negative)	Greater flatness decreases elevation above the average bed surface (fewer ER)	ΝΑ
ER density measures		
Mean ER density (Positive)	ER density is directly related to the number of ER	ΝΑ
Pool-riffle ER density ratio (Negative)	ΝΑ	Pool-riffle ER density ratio is directly related to spatial organisation of ER
Segments without ER		
Number of segments without ER (Negative)	Number of segments without ER is directly related to the number of segments with ER	Coarse measure of the number of pools, implies strong organisation of bed topography
Maximum length without ER (Negative)	Maximum length without ER is directly related to the number of segments with ER	Coarse measure of longest pool, implies strong organisation of bed topography

	Scotla	nd		Australia					
	Dye	Faseny	Kelphope	Little	Snobs	Steavenson			
Pools	0.14	0.40	0.27	0.34	0.19	0.38			
Riffle-like	o.86	0.60	0.73	0.66	0.81	0.62			
True riffle	0.00	0.04	0.40	0.41	0.07	0.50			
Plane Bed	o.86	0.56	0.33	0.25	0.65	0.12			
Step Pool	0.00	0.00	0.00	0.00	0.09	0.00			

**Table S2.** Proportions of channel morphologies in the six study streams.

*Note*. Simplified morphology into pools and 'riffle-like', provide a high-level distinction between deeper pools and those sections where the flow is relatively shallow and bed materials are more likely to be exposed. The 'riffle-like' category is the sum of true riffles, plane beds, step-pools. See Methods, section 2.3.

		Scotland		Australia					
	Kelphope	Faseny	Dye	Snobs	Little	Steavenson			
Latitude	N 55.78639	N 55.85111	N 55.81528	S 37.27382	S 37.34908	S 37.48389			
Longitude	W 2.78389	W 2.59194	W 2.57222	E 145.87713	E 145.75202	E 145.75242			
Altitude (m)	230	300	270	283	308	379			
Survey length (m)	1000	1000	1000	785	685	885			
Stream morphology									
Mean channel width (m)	2.5 ± 0.5	5.8 ± 1.7	6.6 ± 1.3	5.8 ± 1.4	6.1 ± 1.7	9.5 ± 2.6			
Median depth (m)	0.13	0.22	0.23	0.36	0.30	0.37			
Slope (%)	0.94	1.04	1.09	3.08	1.47	1.01			
Mean proportional riffle lengths	0.10 ± 0.06	0.05 ± 0.03	0.11 ± 0.13	0.09 ± 0.09	0.06 ± 0.03	0.05 ± 0.05			
Mean proportional pool lengths	0.04 ± 0.02	0.04 ± 0.02	0.02 ± 0.01	$0.02 \pm 0.01$	0.03 ± 0.02	0.03 ± 0.02			
Pool-riffle transitions	25	2	14	19	26	24			
Bed sediment									
ER axis A (mm)	191.5 ± 52.2	375.3 ± 153.9	371.6 ± 117.2	290.7 ± 152.7	256.4 ± 110.8	202.9 ± 106.3			
ER axis B (mm)	126.3 ± 33.6	241.2 ± 89.9	253.9 ± 74.7	203 ± 102.9	185.7 ± 74.4	147.8 ± 72.2			
ER axis C (mm)	76.9 ± 31.9	134.1 ± 73.3	122.1 ± 54.3	147.2 ± 83.5	121.2 ± 59.1	95.2 ± 57.1			
ER elongation ratio	0.68 ± 0.15	0.67 ± 0.17	0.71 ± 0.17	0.71 ± 0.13	0.74 ± 0.13	0.75 ± 0.13			
ER flatness ratio	0.61 ± 0.20	0.57 ± 0.24	0.50 ± 0.22	0.73 ± 0.16	0.66 ± 0.19	0.65 ± 0.19			
ER equancy ratio	0.41 ± 0.14	0.37 ± 0.16	0.34 ± 0.14	0.52 ± 0.14	0.48 ± 0.14	0.48 ± 0.15			
SR axis A (mm)	136.1 ± 43.8	202.2 ± 97.9	220.6 ± 103.8	187.1 ± 82.5	178.2 ± 79.1	160.4 ± 87.4			
SR axis B (mm)	93.7 ± 31.5	134.7 ± 67.4	152.3 ± 69.6	139.6 ± 62.6	133.8 ± 59.6	115.4 ± 63.5			
SR axis C (mm)	42.6 ± 19.1	58.6 ± 37.1	65 ± 32.3	94.6 ± 47.9	87.2 ± 40.3	73.7 ± 43.3			
SR elongation ratio	0.70 ± 0.15	0.68 ± 0.15	0.71 ± 0.15	0.76 ± 0.12	0.76 ± 0.12	0.73 ± 0.14			
SR flatness ratio	0.47 ± 0.19	0.44 ± 0.19	0.46 ± 0.19	0.68 ± 0.15	0.66 ± 0.16	0.65 ± 0.18			
SR equancy ratio	0.32 ± 0.12	0.30 ± 0.13	0.31 ± 0.12	0.51 ± 0.13	0.50 ± 0.13	0.47 ± 0.14			
ER density measures									
Density of ER (no. m <sup>-2</sup> )	0.31	0.57	1.01	2.31	2.44	1.27			
Pool-riffle ER density ratio	0.63	0.34	0.79	0.68	0.73	0.49			
Segments without ER									
Number of segments without ER	109	34	0	0	2	1			
Maximum length without ER (m)	50	25	<5	<5	5	5			

**Table S3.** Geographical and physical features of the study streams along the survey lengths. Where relevant, values are reported as mean ± SD. SR indicates submerged rocks.

**Table S4.** Correlation tests among variables describing physical features of the study streams. All tests are two-tailed with DF = 4. r-values significant at  $\alpha$  < 0.05 are shown in bold. Sets of cells in shaded grey describe similar aspects of the stream: stream bed morphology, sediment size and shape, and ER density. Cells shaded yellow indicate significant correlations among different aspects. SR indicates submerged rocks.

	Median depth	Slope	Mean length riffles	Mean length pools	Pool-riffle transitions	ER axis B	ER axis C	ER flatness	ER equancy	SR axis B	SR axis C	SR flatness	SR equancy	ER density	ER density ratio	редптептс милооц ЕК
Median depth																
Slope	0.549															
Mean proportional riffle length	-0.426	0.184														
Mean proportional pool length	-0.581	-0.593	-0.421													
Pool-riffle transitions	0.215	0.095	0.145	-0.146												
ER axis B	0.028	0.136	0.097	-0.397	-0.779											
ER axis C	0.437	0.662	-0.09	-0.505	-0.541	0.780										
ER flatness	0.621	0.718	-0.274	-0.141	0.499	-0.48	0.173									
ER equancy	0.734	0.66	-0.301	-0.252	0.603	-0.491	0.128	0.972								
SR axis B	0.362	0.358	0.082	-0.687	-0.487	0.909	0.849	-0.216	-0.164							
SR axis C	0.875	0.753	-0.221	-0.684	0.222	0.193	0.656	0.653	0.73	0.547						
SR flatness	0.862	0.614	-0.31	-0.451	0.602	-0.33	0.213	o.856	0.948	0.059	0.859					
SR equancy	0.847	0.632	-0.316	-0.441	0.585	-0.304	0.249	0.863	0.947	0.082	0.874	0.997		1		
ER density	0.753	0.687	-0.166	-0.608	0.396	0.071	0.53	0.662	0.744	0.444	0.958	o.867	0.892			
ER density ratio	0.018	0.297	0.712	-0.67	0.528	0.044	0.034	0.009	0.100	0.284	0.312	0.233	0.242	0.457		
Segments without ER	-0.823	-0.379	0.272	0.734	0.096	-0.519	-0.654	-0.155	-0.295	-0.788	-0.813	-0.548	-0.547	-0.694	-0.189	
Max length without ER	-0.813	-0.450	0.107	0.845	-0.009	-0.478	-0.621	-0.177	-0.324	-0.781	-0.831	-0.579	-0.575	-0.727	-0.347 <b>C</b>	0.982

-	5		5	• •			
		Scotlan	d	Australia			
	Dye	Faseny	Kelphope	Little	Snobs	Steavenson	
Box-Counting							
Log(δ)-log(N <sub>B</sub> ) slope	1.00	0.98	0.92	1.00	1.00	1.00	
R <sup>2</sup> – SSR	0.99	0.95	1.01	0.99	0.99	0.99	
Zero-Slope	1.03	1.02	0.97	1.03	1.02	1.02	
Compensated-Slope (D <sub>B</sub> )	1.00	0.97	0.90	1.00	1.00	1.00	
Entropy							
Log(δ)-log(N <sub>E</sub> ) slope	1.00	0.94	0.93	1.04	1.06	1.01	
R <sup>2</sup> – SSR	0.94	0.91	1.03	0.96	0.98	0.96	
Zero-Slope	0.98	0.98	1.19	1.11	1.10	0.99	
Compensated-Slope (D <sub>E</sub> )	1.00	0.93	0.91	1.02	1.05	0.98	

**Table S5.** Fractal Dimension Results using the Box-Counting and Entropy Methods.

Note. One estimate of fractal dimension can be produced from the absolute slope of  $log(\delta)$ - $log(N_x)$  linear regression, where significant linearity (here indicated with bold text) indicates the likely presence of a fractal structure. A more robust method of confirming the presence of fractal-like properties follows *Seuront's* [2009] three-step procedure, including the 1) R<sup>2</sup> – SSR, 2) Zero-Slope, and 3) Compensated-Slope procedures. Using this method, any stream must satisfy any two of the three criteria to be considered fractal [*Seuront*, 2009] (see Supporting Text S1). Bold text indicates that the data satisfy the criterion for the test in question. Plain text indicates that the data did not satisfy the criterion. As described in the methods above, Dx was selected using the Compensated-Slope Procedure, as indicated by the parentheses.

	Expected	Box-cou	nting	Entropy		
	direction	r	p	r	р	
Stream morphology						
Median depth	Negative	0.817	0.977	0.791	0.969	
Stream slope	Positive	0.365	0.238	0.737	0.047	
Mean proportional pool length	Negative	-0.709	0.057	-0.870	0.012	
Mean proportional riffle length	Positive	-0.301	0.719	0.054	0.460	
Number of pool-riffle transitions	Negative	-0.124	0.408	0.241	0.677	
Bed sediment						
ER axis B	Positive	0.791	0.030	0.653	0.080	
ER axis C	Positive	0.804	0.027	0.950	0.002	
ER equancy ratio (C/A)	Positive	0.534	0.137	0.765	0.038	
ER flatness ratio (C/B)	Negative	0.534	0.862	0.755	0.958	
SR axis B	Positive	0.532	0.139	0.291	0.288	
SR axis C	Positive	0.663	0.076	0.636	0.087	
SR equancy ratio (C/A)	Positive	0.283	0.294	0.582	0.113	
SR flatness ratio (C/B)	Negative	0.146	0.609	0.482	0.833	
ER density measures						
ER density over entire site	Positive	0.679	0.069	0.917	0.005	
Pool-riffle ER density ratio	Negative	0.157	0.617	0.541	o.866	
Segments without ER						
Number of segments without ER	Negative	-0.999	<0.001	-0.832	0.020	
Maximum length without ER	Negative	-0.974	0.001	-0.895	0.008	

**Table S6.** One-tailed correlation tests of the study stream characteristics with the box-counting and entropy fractal dimensions. DF = 4. Bold text indicates p < 0.05.

	В	lox-co	unting	Entropy					
River	F-statistic	DF	p-value	R²	F-statistic	DF	p-value	R²	
Dye	84177	25	<0.01	1.00	6263	24	<0.01	1.00	
Faseny	34778	25	<0.01	1.00	6208	25	<0.01	1.00	
Kelphope	3498	25	<0.01	0.99	2604	24	<0.01	0.99	
Little	47493	20	<0.01	1.00	2901	18	<0.01	0.99	
Snobs	63688	21	<0.01	1.00	4434	20	<0.01	1.00	
Steavenson	151983	23	<0.01	1.00	5399	22	<0.01	1.00	

**Table S7.** Linear regression tests of the  $log(\delta)$  versus  $log(N_X)$  for the box-counting  $(N_B)$  and entropy  $(N_E)$  methods. Linearity indicates the likely presence of fractal structure.

Indonendentverighte	Manipulated	ER	DE	Box counti	ng	Entropy					
independent variable	variable(s)	dist.	DF	Slope	t-value	p-value	<b>R</b> <sup>2^</sup>	Slope	t	р	R²
Stream morphology											
Mean proportional pool	Poollongths	LZS	78	-2.6×10 <sup>-8</sup>	0.08	0.94	<0.01	<b>-8.9×10</b> ⁻⁵	3.48	<0.001	0.13
length	Poorlengtils	HZS	78	–4.4×10 <sup>-4</sup>	5.60	<0.001	0.29	-7.8×10⁻⁴	4.84	<0.001	0.23
Mean proportional pool	Pool lengths,	LZS	77	4.3×10 <sup>-8</sup>	0.11	0.91	<0.01	-3.8×10 <sup>-3</sup>	15.5	<0.001	0.76
length	riffle ER	HZS	77	-3.5×10⁻⁴	4.29	<0.001	0.19	-4.6×10 <sup>-3</sup>	8.46	<0.001	0.48
Mean proportional riffle	Piffle longths	LZS	77	–7.3×10 <sup>-8</sup>	0.78	0.44	0.01	8.2×10 <sup>-6</sup>	0.88	0.38	<0.01
length	Rime lengths	HZS	77	1.4×10 <sup>-4</sup>	3.73	<0.001	0.15	1.3×10 <sup>-6</sup>	1.88	0.06	0.04
ER density measures											
	Riffle ER,	LZS	78	-3.3×10 <sup>-9</sup>	0.19	0.85	<0.01	-1.3×10 <sup>-6</sup>	0.70	0.49	<0.01
ER density	pool ER	HZS	, 78	-1.3×10 <sup>-5</sup>	0.49	0.63	<0.01	,1×10 <sup>-5</sup>	, 0.83	0.41	<0.01
		LZS	, 78	-4.1×10 <sup>-8</sup>	1.37	0.18	0.02	-1.1×10 <sup>-4</sup>	12.3	<0.001	o.66
Mean density of riffle ER	Riffle ER	HZS	, 78	-5.1×10 <sup>-5</sup>	1.26	0.21	0.02	-4.2×10 <sup>-4</sup>	5.02	<0.001	0.24
	Pool lengths,	LZS	77	-9.1×10 <sup>-9</sup>	0.30	0.77	<0.01	-3.2×10 <sup>-4</sup>	22.4	<0.001	0.87
Mean density of riffle ER	riffle ER	HZS	77	-8.9×10 <sup>-5</sup>	3.28	0.002	0.12	_ −1.9×10 <sup>-3</sup>	17.1	<0.001	0.79
Pool-riffle ER density		LZS	78	-2.0×10 <sup>-5</sup>	1.00	0.32	0.01	0.075	11.9	<0.001	0.64
ratio	RITTEER	HZS	78	-8.1×10 <sup>-3</sup>	0.90	0.37	0.01	0.093	4.97	<0.001	0.24
Pool-riffle ER density	Pool lengths,	LZS	77	–7.6×10⁻ <sup>7</sup>	0.04	0.97	<0.01	0.19	13.3	<0.001	0.69
ratio	riffle ER	HZS	77	3.1×10 <sup>-2</sup>	5.11	<0.001	0.25	0.38	9.71	<0.001	0.55
Segments without ER											
Number of segments without FR	Pool lengths	HZS	78	-7.8×10 <sup>-4</sup>	9.93	<0.001	0.03	- <b>1.2×10</b> <sup>-3</sup>	6.18	<0.001	0.33
Maximum length without ER	Pool lengths	HZS	78	-3.4×10 <sup>-3</sup>	5.70	<0.001	0.29	-1.8×10 <sup>-3</sup>	1.29	0.20	0.02

**Table S8.** Linear regression tests of the stream characteristics using synthetic stream stretches with the box-counting and entropy fractal dimensions. Entropy results are repeated for ease of comparison.

*Note.* t-tests indicate whether slopes differ from zero, significant tests are shown in bold text.

Script S1.

### R scripts for calculating fractal dimensions

### 

## IsInBox

```
IsInBoxBC <- function (IsInBoxdata, boxx, boxy)</pre>
ş
 IsInBoxdata <- subset(IsInBoxdata, IsInBoxdata[,1]>= boxx[1])
 IsInBoxdata <- subset(IsInBoxdata, IsInBoxdata[,1]<= boxx[2])</pre>
 if (nrow(IsInBoxdata)>o){
  for (i in 1:nrow(IsInBoxdata))
  { if(any((lslnBoxdata[i,2] >= boxy[1]) & (lslnBoxdata[i,2] <= boxy[2])))</pre>
  {return(TRUE)}} }
 return(FALSE)
}
## CrossingBox
CrossingBoxyBC <- function (Crossdata, boxx, boxy)
ş
 Crossdata <- subset(Crossdata, Crossdata[,1]>= boxx[1])
 Crossdata <- subset(Crossdata, Crossdata[,1]<= boxx[2])
 Add <- o
 nrow(Crossdata)
 if (nrow(Crossdata)>o){
 for (j in 1:nrow(Crossdata))
 { if(Crossdata[i,2] >= boxy[2]) {Add <- 1}}
 }
 return(Add)
}
CrossingBoxxBC <- function (Crossdata, boxx, boxy)
ş
 Crossdata <- subset(Crossdata, Crossdata[,2]>= boxy[1])
 Crossdata <- subset(Crossdata, Crossdata[,2]<= boxy[2])
```

```
Add <- o
 if (nrow(Crossdata)>o){
  for (k in 1:nrow(Crossdata))
  { if(Crossdata[k,1] >= boxx[2]) {Add <- 1}}
}
 return(Add)
}
CrossingBoxxyBC <- function (Crossdata, boxx, boxy)
ş
 Crossdata <- subset(Crossdata, Crossdata[,1]>= boxx[2])
 Crossdata <- subset(Crossdata, Crossdata[,2]>= boxy[2])
 Add <- o
 if(nrow(Crossdata) < 0) {Add <- 1}
 return(Add)
}
## BoxCount
BoxCount_Fun <- function (data, numcases, boxsize, data.range.x, data.range.y)
ş
 Ninternal <- o
 width <- boxsize
 height <- boxsize
 minboxx <- data.range.x[1]</pre>
 minboxy <- data.range.y[1]</pre>
 maxboxx <- data.range.x[2]</pre>
 maxboxy <- data.range.y[2]</pre>
 qh <- trunc(maxboxx/width)</pre>
 qv <- max(trunc(maxboxy/height),1)</pre>
 for (hor in 1:qh) {
  boxxlimits <- c((min(data[,1]) + (hor-1) * width),(min(data[,1]) + (hor) * width))
  for (ver in 1:qv) {
   boxylimits <- c((min(data[,2]) + (ver-1) * height), (min(data[,2]) + ver * height))</pre>
   if(IsInBoxBC(data, boxxlimits, boxylimits))
   {Ninternal <- Ninternal + 1}
   if(hor == qh) 
    Addx <- CrossingBoxxBC(data, boxxlimits, boxylimits)
    Ninternal <- Ninternal + Addx }
   if(ver == qv)
```

```
Addy <- CrossingBoxyBC(data, boxxlimits, boxylimits)
Ninternal <- Ninternal + Addy }
}
Addxy <- CrossingBoxxyBC(data, boxxlimits, boxylimits)
Ninternal <- Ninternal + Addxy
}
return(Ninternal)
}
```

```
## Run Box-counting functions
Function_Run_Box_Counting <- function(Delta = TRUE) {
    library(plyr)
    if("Rocks" %in% colnames(data1D)){
        # 1D data:
        data <- subset(data1D, Rocks > o)
        data <- uncount(data = data, weights = data$Rocks)
        data$Y..m. <- 1
        data$X..m. <- data$X_coord
        min_box_size <- 5
        decimal_places <- 0</pre>
```

### }else{

```
# 2D data:
data <- data1D
min_box_size <- 0.2
decimal_places <- 2
}</pre>
```

```
Rivers <-as.character(unique(data$River))
```

```
boxcount_results_Im1 <- array(o, dim = c(length(Rivers), 4))
boxcount_results_Im2 <- array(o, dim = c(length(Rivers), 4))
Delta2onwards_LIST <- list()
DELTAS <- data.frame(River = Rivers, new_delta = vector("numeric",length(Rivers)))
Snipping_values <- data.frame(River = Rivers, pointsremove =
vector("numeric",length(Rivers)))</pre>
```

```
for (g in 1:length(Rivers)) {
    data2D_format <- data[data$River==Rivers[g],]
    data2D_format <- subset(data2D_format, select = c(X..m.,Y..m.))
    n <- nrow(data2D_format)
    data.range.x <- range(data2D_format[1:n,1])
    data.range.y <- range(data2D_format[1:n,2])
    sizes <- round(seq(from = data.range.x[1], to = data.range.x[2], by = min_box_size), digits =
decimal_places)
    sizes <- sizes[sizes != 0]</pre>
```

```
delta <- length(sizes)</pre>
 neff <- trunc((n - 1)/((delta - 1)*5)) * ((delta - 1)*5) + 1
 Box <- array(o,c(delta,3))
 for (a in (1:delta)) {
  Box[a,1] <- sizes[a]
  Box[a,2] <- log(BoxCount_Fun(data = data2D_format, numcases = neff, boxsize = sizes[a],
data.range.x, data.range.y))
  Box[a,3] <- a
 }
 colnames(Box) <- c("delta", "log_boxcount", "old_delta")</pre>
 Box <- as.data.frame(Box[!duplicated(Box[,2]), ])</pre>
 Box <- as.data.frame(Box[1:nrow(Box)-1, ])</pre>
 Delta2onwards_LIST[[g]]<- Box
 DELTAS$new_delta[q] <- nrow(Box)</pre>
 Snipping_values$pointsremove[q] <- nrow(Box)</pre>
 Delta2onwardsdf <- as.data.frame(Box)
 Im <- Im(log_boxcount~log(delta), data=Delta2onwardsdf)</pre>
 boxcount_results_lm1[q,1:4] <- rbind(as.character(Rivers[q]),
as.numeric(summary(lm)$coefficients[2]), summary(lm)$coefficients[1],
summary(lm)$r.squared)
 Im2 <- Imodel2(Delta2onwardsdf$log_boxcount~log(Delta2onwardsdf$delta),
data=Delta2onwardsdf)
 lm2_intercept <- lm2$regression.results[3,2] # print intercept</pre>
 lm2_slope <- lm2$regression.results[3,3] # print slope</pre>
 lm2_rsquare <- lm2$rsquare # print R2</pre>
 boxcount_results_lm2[q,1:4] <- rbind(as.character(Rivers[q]), lm2_slope, lm2_intercept,
lm2_rsquare)
 rm(delta, g, data2D_format)
}
 names(Delta2onwards_LIST) <- Rivers
 colnames(boxcount_results_lm1) <- c("River", "lm1_slope", "lm1_Intercept", "R_squared")
 colnames(boxcount_results_Im2) <- c("River", "Im2_slope", "Im2_Intercept", "R_squared")
 boxcount_LM_results <- cbind(as.data.frame(boxcount_results_lm1),</pre>
as.data.frame(boxcount_results_lm2[,2:4]))
 # Output:
 List_boxcount <- list()
```

```
List_boxcount$boxcount_LM_results <- boxcount_LM_results
```

```
List_boxcount$Delta2onwards_LIST <- Delta2onwards_LIST
```

```
List_boxcount$DELTAS <- DELTAS
```

```
List_boxcount$Snipping_values <- Snipping_values
```

```
return(List_boxcount)
}
##Entropy Functions
###############
## IsInBox
IsInBoxE <- function (IsInBoxdata, boxx, boxy)</pre>
ş
 IsInBoxdata <- subset(IsInBoxdata, IsInBoxdata[,1]>= boxx[1])
 IsInBoxdata <- subset(IsInBoxdata, IsInBoxdata[,1]<= boxx[2])</pre>
 Add <- o
 if (nrow(IsInBoxdata)>o){
 for (i in 1:nrow(IsInBoxdata))
  { if(any((IsInBoxdata[i,2] >= boxy[1]) & (IsInBoxdata[i,2] <= boxy[2])))</pre>
  {Add <- Add + 1}} }
 Add
 return(Add)
}
## CrossingBox
CrossingBoxyE <- function (Crossdata, boxx, boxy)
{
 Crossdata <- subset(Crossdata, Crossdata[,1]>= boxx[1])
 Crossdata <- subset(Crossdata, Crossdata[,1]<= boxx[2])
 Add <- o
 nrow(Crossdata)
 if (nrow(Crossdata)>o){
 for (j in 1:nrow(Crossdata))
  { if(Crossdata[j,2] >= boxy[2]) {Add <- Add + 1}}
 }
 Add
 return(Add)
}
CrossingBoxxE<- function (Crossdata, boxx, boxy)
{
 Crossdata <- subset(Crossdata, Crossdata[,2]>= boxy[1])
 Crossdata <- subset(Crossdata, Crossdata[,2]<= boxy[2])
 Add <- o
```

```
if (nrow(Crossdata)>o){
 for (k in 1:nrow(Crossdata))
 { if(Crossdata[k,1] >= boxx[2]) {Add <- Add + 1}}
}
return(Add)
}
CrossingBoxxyE <- function (Crossdata, boxx, boxy)
ş
 Crossdata <- subset(Crossdata, Crossdata[,1]>= boxx[2])
 Crossdata <- subset(Crossdata, Crossdata[,2]>= boxy[2])
 Add <- o
 if(nrow(Crossdata) < 0) {Add <- Add + 1}
 return(Add)
}
## Entropy
Entropy_Fun <- function (data, numcases, boxsize, data.range.x, data.range.y)
{
 width <- boxsize
 height <- boxsize
 minboxx <- data.range.x[1]</pre>
 minboxy <- data.range.y[1]</pre>
 maxboxx <- data.range.x[2]</pre>
 maxboxy <- data.range.y[2]
 qh <- trunc(maxboxx/width)</pre>
 qv <- trunc(maxboxy/height)
 if(qv < 1) {qv <- 1}
 # Calculate the total number of rocks
 CountRocks <- array(o, dim = c(max(1,qh), max(1,qv)))
 for (hor in 1:qh) {
 boxxlimits <- c((min(data[,1]) + (hor-1) * width),(min(data[,1]) + (hor) * width)))
 boxxlimits
 for (ver in 1:qv) {
  boxylimits <- c((min(data[,2]) + (ver-1) * height), (min(data[,2]) + ver * height))</pre>
  boxylimits
  Addbase <- IsInBoxE(data, boxxlimits, boxylimits)
  CountRocks[hor,ver] <- Addbase
 }
 }
 CountRocks
```

sum(CountRocks)

```
# Calculate weighted pi
 pivalues <- array(o, dim = c(nrow(CountRocks),ncol(CountRocks)))
 pilogpivalues <- array(o, dim = c(nrow(CountRocks),ncol(CountRocks)))
 TotalRocks <- sum(CountRocks)
 for (m in 1:nrow(CountRocks))
 ş
 for (l in 1:ncol(CountRocks))
 ş
  if(CountRocks[m,l] > 0)
  Ł
   pivalues[m,l] <- CountRocks[m,l] / TotalRocks
   pilogpivalues[m,l] <- pivalues[m,l] * log(pivalues[m,l])</pre>
  }
 }
}
 Hdelta <- -sum(pilogpivalues)
 # Correct for bias in the estimate of entropy as per Basharin (1959)
 Samplesize <- length(CountRocks)</pre>
 Hdeltaunbiased <- Hdelta + ((Samplesize - 1)/(2 * TotalRocks)*(log2(exp(1))))
 return(Hdeltaunbiased)
}
## Run Entropy functions
Function_Run_Entropy <- function (Delta = TRUE) {</pre>
 library(plyr)
 if("Rocks" %in% colnames(data1D)){
 #1D data:
 data <- subset(data1D, Rocks > o)
 data <- uncount(data = data, weights = data$Rocks)</pre>
 data$Y..m. <- 1
 data$X..m. <- data$X_coord
 min_box_size <- 5
 decimal_places <- o
 }else{
 #2D data:
 data <- data1D
 min_box_size <- 0.2
 decimal_places <- 2
}
```

Rivers <-as.character(unique(data\$River))

```
Delta2onwards_LIST <- list()</pre>
DELTAS <- data.frame(River = Rivers, new_delta = vector("numeric",length(Rivers)))
Snipping_values <- data.frame(River = Rivers, pointsremove =
vector("numeric",length(Rivers)))
for (g in 1:length(Rivers)) {
 data2D_format <- data[data$River==Rivers[q],]</pre>
 data2D_format <- subset(data2D_format, select = c(X..m.,Y..m.))
 n <- nrow(data2D_format)
 data.range.x <- range(data2D_format[1:n,1])</pre>
 data.range.y <- range(data2D_format[1:n,2])</pre>
 sizes <- round(seq(from = data.range.x[1], to = data.range.x[2], by = min_box_size), digits =
decimal_places)
 sizes <- sizes[sizes != 0]</pre>
 delta <- length(sizes)</pre>
 neff <- trunc((n - 1)/((delta - 1)*5)) * ((delta - 1)*5) + 1
 Entropy <- array(o,c(delta,5))</pre>
 for (a in (1:delta)) { #a<-1
  Entropy[a,1] <- sizes[a]</pre>
  Entropy[a,5] <- Entropy_Fun(data = data2D_format, numcases = neff, boxsize = sizes[a],
data.range.x, data.range.y)
  Entropy[a,2] <- log(Entropy[a,5])</pre>
  Entropy[a,3] <- a
  Entropy[a,4] <- round(Entropy[a,5], digits = 1)</pre>
 }
 colnames(Entropy) <- c("delta", "log_Entropy", "old_delta", "Round_Entropy",
"Entropy_Unlogged")
 Entropy <- as.data.frame(Entropy[!duplicated(Entropy[,4]), ])</pre>
 Entropy <- as.data.frame(Entropy[1:nrow(Entropy)-1, ])</pre>
 Delta2onwards_LIST[[q]]<- Entropy
 DELTAS$new_delta[q] <- nrow(Entropy)</pre>
 Snipping_values$pointsremove[g] <- nrow(Entropy)
 rm(delta, g, data2D_format)
}
names(Delta2onwards_LIST) <- Rivers
# Ouputs:
List_Entropy <- list()
List_Entropy$Delta2onwards_LIST <- Delta2onwards_LIST
List_Entropy$DELTAS <- DELTAS
List_Entropy$Snipping_values <- Snipping_values
```

```
return(List_Entropy)
```

}

```
## Log-Log regression
Function_FDSlope <- function() {
```

```
Rivers <- unique(data1D$River)
```

FDSlope\_results <- data.frame(River = Rivers, Intercept= vector("numeric",length(Rivers)), Slope = vector("numeric",length(Rivers)), R2 = vector("numeric",length(Rivers))) #NULL

```
for (g in 1:length(Rivers)) {
    Delta2onwards <- as.data.frame(Delta2onwards_LIST[[g]])</pre>
```

```
# Log-Log Plot
```

```
Plot <- ggplot(Delta2onwards, aes(x= log(Delta2onwards$delta), y = Delta2onwards[,2])) +
geom_point(size = 3) +
xlab(expression(log(delta))) + ylab(expression(log(N["X"]))) +
ggtitle(as.character(Rivers[g]))+
theme_bw()+ theme(text = element_text(size=15), axis.text = element_text(size=15),
legend.position = "none", panel.grid = element_blank()) +
stat_smooth(method="lm", se=FALSE)+
stat_poly_eq(formula = y ~ x, aes(label = paste(..eq.label.., ..rr.label.., sep = "~~~")), parse =</pre>
```

```
TRUE, label.y = "top", label.x = "right")
```

```
print(Plot)
```

```
#Extract intercept, slope, and R-squared from lm()
river_lm <- lm(Delta2onwards[,2] ~ log(Delta2onwards$delta))
FDSlope_results$Intercept[g] <- summary(river_lm)$coefficients[1]
FDSlope_results$Slope[g] <- summary(river_lm)$coefficients[2]
FDSlope_results$R2[g] <- summary(river_lm)$r.squared
}
return(FDSlope_results)</pre>
```

```
}
```

## ## R2SSR

```
Function_R2SSR <- function() {

Rivers <- unique(data1D$River)

R2SSR_results <- data.frame(River = Rivers, R2SSR_FD = vector("numeric",length(Rivers)),

R2SSR_test_result = vector("logical",length(Rivers))) #NULL
```

```
for (g in 1:length(Rivers)) {
 delta <- DELTAS[q,2]
 Plotdataframe <- as.data.frame(Delta2onwards_LIST[[q]])
 windowsize <- 6 #smallest window size
 iterations <- o
 for (j in windowsize:(delta-1))
 ş
  iterations <- iterations + (delta - (j))
 }
 R_2SSR <- array(o, dim = c((iterations), 4))
 colnames(R2SSR) <- c("Rsquared", "SSR", "Slope", "Iterations")
 startlocation <- 1
 for (k in 1: iterations)
 ş
  Residualcalc <- array(o, dim = c(windowsize,6))
  # Standard major axis regression
  R2SSRIm <- Imodel2(Plotdataframe[startlocation:(windowsize+startlocation-1),2] ~
log(Plotdataframe[startlocation:(windowsize+startlocation-1),1]), data = Plotdataframe)
  # Calculate sum of squared residuals and R2
  SMAslope <- R2SSRIm$regression.results[3,3]
  SMAintercept <- R2SSRIm$regression.results[3,2]
  for (i in 1:nrow(Residualcalc))
  ş
   Residualcalc[i,1] <- log(Plotdataframe[(startlocation+i-1),1])
   Residualcalc[i,2] <- Plotdataframe[(startlocation+i-1),2]
   Residualcalc[i,3] <- SMAslope * Residualcalc[i,1] + SMAintercept
   Residualcalc[i,4] <- Residualcalc[i,3] - Residualcalc[i,2]
   Residualcalc[i,5] <- Residualcalc[i,4]<sup>2</sup>
   Residualcalc[i,6] <- Residualcalc[i,1]*Residualcalc[i,2]
  }
  SSRSMA <- sum(Residualcalc[,5])
  TotalSS <- sum(Residualcalc[,2]^2) - ((sum(Residualcalc[,2])^2)/nrow(Residualcalc))
  RegressionSS <- (((sum(Residualcalc[,6]))-
((sum(Residualcalc[,1])*sum(Residualcalc[,2]))/nrow(Residualcalc)))^2)/(sum(Residualcalc[,1]^2
)-((sum(Residualcalc[,1])<sup>2</sup>/nrow(Residualcalc))))
  R2SMA <- RegressionSS / TotalSS
  R2SSR[k,2] <- SSRSMA
  R2SSR[k,1] <- R2SMA
  R2SSR[k,3] <- SMAslope
  R2SSR[k,4] <- k+1
```

```
if (startlocation < ((delta)-(windowsize)))
  ş
   (startlocation <- startlocation + 1)
  } else
  Ł
   startlocation <- 1
   windowsize <- windowsize + 1
  }
 }
 # Output:
 R2SSR <- as.data.frame(R2SSR)
 R2SSR_test_result <- which.min(R2SSR[,2]) == which.max(R2SSR[,1])
 R2SSR_FD <- as.numeric(abs(R2SSR[which.min(R2SSR[,2]),3]))
 R2SSR_results[g,2] <- R2SSR_FD
 R2SSR_results[q,3] <- R2SSR_test_result
 ggplot(R2SSR, aes(x= SSR, y= Rsguared))+
  geom_point(size = 3, shape = 21) +
  geom_point(data = R_2SSR[R_2SSR_3R_3] alpha = 0.5, colour
= "red", size = 3) +
  geom_point(data = R2SSR[R2SSR$SSR == min(R2SSR$SSR),], alpha = 0.5, colour = "blue",
size = 3) +
  ggtitle(as.character(Rivers[g]))+
  ylab(expression(R^2)) + xlab("SSR") +
  theme_bw()+ theme(text = element_text(size=15), axis.text = element_text(size=15),
legend.position = "none", panel.grid = element_blank())
}
return(R2SSR_results)
}
# Function for the Zero Slope Procedure
## Zero Slope
Function_Zero_Slope <- function() {</pre>
Rivers <- unique(data1D$River)
Zero_Slope_results <- data.frame(River = Rivers, Optintercept =
vector("numeric",length(Rivers)), ZeroSlope_test_result = vector("logical",length(Rivers)),
ZeroSlope_Lower = vector("logical",length(Rivers)), ZeroSlope_Upper =
vector("logical",length(Rivers))) #NULL
for (q in 1:length(Rivers)) {
```

```
delta <- DELTAS$new_delta[g]
Plotdataframe <- as.data.frame(Delta2onwards_LIST[[q]][1:delta,])
```

```
Independentz <- log(Plotdataframe[,1])</pre>
 Dependentz <- Plotdataframe[,2]</pre>
 deriv <- function(x, y) diff(y) / diff(x)</pre>
 middle_pts <- function(x) x[-1] - diff(x) / 2
 first_d <- deriv(Independentz, Dependentz)</pre>
 Independentza <- middle_pts(Independentz)
 iterationz <- o
 minwindowsize <- 6
 for (j in minwindowsize:(delta-1)) # 6 is the minimum window size
 ş
  iterationz <- iterationz + (delta - (j+1))
 }
 Zeroslopez <- array(o, dim = c((iterationz), 4))
 colnames(Zeroslopez) <- c("F", "p", "windowsize", "startinglocation")</pre>
 startlocationz <- 1
 windowsizez <- minwindowsize
 for (k in 1:iterationz)
 Ł
  Residualcalcz <- array(o, dim = c(windowsizez, 6))
  NewDependentz <- first_d[startlocationz:(windowsizez+startlocationz-1)]
  NewIndependentz <- Independentza[startlocationz:(windowsizez+startlocationz-1)]
  # Standard major axis regression
  R2SSRImz <- Imodel2(NewDependentz ~ NewIndependentz)
  Zslopez <- R_2SSRIm_sregression.results[3,3]
  Zerointerceptz <- R2SSRImz$regression.results[3,2]
  for (i in 1:nrow(Residualcalcz))
  ş
   Residualcalcz[i,1] <- NewIndependentz[i]
   Residualcalcz[i,2] <- NewDependentz[i]
   Residualcalcz[i,3] <- Zslopez * Residualcalcz[i,1] + Zerointerceptz
   Residualcalcz[i,4] <- Residualcalcz[i,3] - Residualcalcz[i,2]
   Residualcalcz[i,5] <- Residualcalcz[i,4]<sup>2</sup>
   Residualcalcz[i,6] <- Residualcalcz[i,1]*Residualcalcz[i,2]
  }
  SSRSMAz <- sum(Residualcalcz[,5])
  TotalSSz <- sum(Residualcalcz[,2]^2) - ((sum(Residualcalcz[,2])^2)/nrow(Residualcalcz))
  RegressionSSz <- (((sum(Residualcalcz[,6]))-
((sum(Residualcalcz[,1])*sum(Residualcalcz[,2]))/nrow(Residualcalcz)))^2)/(sum(Residualcalcz[,
1]<sup>2</sup>)-((sum(Residualcalcz[,1])<sup>2</sup>/nrow(Residualcalcz))))
```

R2SMAz <- RegressionSSz / TotalSSz

```
RegressionMSz <- RegressionSSz / 1
ResidualMSz <- (TotalSSz - RegressionSSz)/(nrow(Residualcalcz)-2)
Zeroslopez[k,1] <- RegressionMSz/ResidualMSz
Zeroslopez[k,2] <- pf(Zeroslopez[k,1], df1=1, df2=(nrow(Residualcalcz)-2), lower.tail=FALSE)
Zeroslopez[k,3] <- windowsizez
Zeroslopez[k,4] <- startlocationz
if (startlocationz < ((delta-1)-(windowsizez)))
{
(startlocationz <- startlocationz + 1)
} else
{
startlocationz <- 1
windowsizez <- windowsizez + 1
}
```

```
}
```

*#* calculate the number of regression windows that have a slope significantly different from zero

```
length(Zeroslopez[,2])
sum(Zeroslopez[,2] < 0.05)
z_insig_results <- as.data.frame(subset(Zeroslopez, Zeroslopez[,2] > 0.05))
ZeroSlope_pass <- nrow(z_insig_results)>0
zeroslopez_largest <- z_insig_results[length(z_insig_results$windowsize),]</pre>
```

# Finding the confidence intervals for the ZERO slope lines for the window of interest by bootstrapping the entropy values

nruns <- 100 Critwindowsize <- 6 Critstartlocation <- 1

```
critdelta<- Critstartlocation + Critwindowsize -1
BootFDZ <- array(o,dim = c(nruns,1))</pre>
```

```
for (m in 1:nruns)
```

{ PlotdataClz <- array(o, dim = c(delta, 2)) colnames(PlotdataClz) <- c("delta", "entropy")

```
SelectedEntropy <- sample(1:nrow(Plotdataframe), delta, replace = TRUE)</pre>
```

```
for (j in 1:delta)
{
    PlotdataClz[j,2] <- Plotdataframe[SelectedEntropy[j],2]
}</pre>
```

```
PlotdataClz[,1] <- Plotdataframe[,1]
PlotdataClz[,2] <- sort(PlotdataClz[,2], decreasing = TRUE)
```

```
# calculate the zero slope
IndependentClz <- log(PlotdataClz[,1])
DependentClz <- PlotdataClz[,2]</pre>
```

```
first_dCl <- deriv(IndependentClz, DependentClz)
IndependentzCla <- middle_pts(IndependentClz)
NewDependentClz <- first_dCl
NewIndependentClz <- IndependentzCla
```

```
R2SSRImbootz <- Im(NewDependentClz ~ 1 + offset(o * NewIndependentClz))
BootFDZ[m] <- summary(R2SSRImbootz)$coefficients[1,1]
```

```
# find the 2.5th and 97.5th percentiles (95% confidence intervals)
BootFDZ <- sort(BootFDZ)
LowerScatterZ <- BootFDZ[nruns/40]
UpperScatterZ <- BootFDZ[39*nruns/40]</pre>
```

```
# Find the optimal intercept
OptDependentz <- first_d
OptIndependentz <- Independentza</pre>
```

```
R2SSRImoptz <- Im(OptDependentz ~ 1 + offset(o * OptIndependentz))
Optintercept <- summary(R2SSRImoptz)$coefficients[1,1]
OptSig <- summary(R2SSRImoptz)$coefficients[1,4]<0.05
```

```
zero_slope_data <- as.data.frame(OptIndependentz)
zero_slope_data$OptDependentz <- OptDependentz</pre>
```

```
# Outputs:
```

```
ggplot(zero_slope_data, aes(x= OptIndependentz, y= OptDependentz))+
geom_point(size = 3, shape = 21) +
geom_hline(yintercept = Optintercept, colour = "red") +
geom_hline(yintercept = LowerScatterZ, colour = "blue", linetype="dashed") +
geom_hline(yintercept = UpperScatterZ, colour = "blue", linetype="dashed") +
ylim(-10,2) +
ggtitle(as.character(Rivers[g]))+
xlab(expression(log(delta))) + ylab(expression(~italic(d)~log(N[B]))) +
theme_bw()+ theme(text = element_text(size=15), axis.text = element_text(size=15),
legend.position = "none", panel.grid = element_blank())
```

```
z_insig_results <- as.data.frame(subset(Zeroslopez, Zeroslopez[,2] > 0.05))
ZeroSlope_pass <- nrow(z_insig_results)>0
Zero_Slope_results[g,2] <- Optintercept
Zero_Slope_results[g,3] <- OptSig</pre>
```

```
Zero_Slope_results[q,4] <- LowerScatterZ
 Zero_Slope_results[g,5] <- UpperScatterZ
}
return(Zero_Slope_results)
}
# Function for the Compensated-slope Procedure
## Compensated-slope
Function_Compensated_Slope <- function() {</pre>
nCvalues <- 101
Rivers <- unique(data1D$River)
Compensated_Slope_results <- data.frame(River = Rivers, Cvalue =
vector("numeric",length(Rivers)), C_slope_MIn = vector("logical",length(Rivers)), C_slope_Max
= vector("logical",length(Rivers)), C_slope_MIn2 = vector("logical",length(Rivers)),
C_slope_Max2 = vector("logical",length(Rivers)))
for (g in 1:length(Rivers)) {
          <- DELTAS$new_delta[g]
 delta
                                         # delta <- 20
 Plotdataframe <- as.data.frame(Delta2onwards_LIST[[q]][1:delta,])
 FDvalue <- abs(FDSlope_results$Slope[q])
 Cvalues <- seq(o, FDvalue+1, length= nCvalues)
 # calculate constant k based on M(delta) = \log k \times DF \log delta
 CalcFD <- array(o, dim = c(nrow(Plotdataframe), 3))
 CalcFD[,1] <- Plotdataframe[,1]
 CalcFD[,2] <- Plotdataframe[,2]
 kconstant <- mean(exp(CalcFD[,2] + FDvalue * log(CalcFD[,1])))</pre>
 CalcFD[,3] <- (log(kconstant)- CalcFD[,2])/ log(CalcFD[,1])
 colnames(CalcFD) <-(c("delta", "entropy", "FD value"))</pre>
 iterationc <- o
 for (j in 6:(delta)) # 6 is the minimum window size
 ş
  iterationc <- iterationc + (delta - (j))
 }
 Compslopec <- array(o, dim = c((iterationc), 3, length(Cvalues)))
 colnames(Compslopec) <- c("F", "p", "cvalue")</pre>
 startlocationc <- 1
 windowsizec <- 6
 for (k in 1:iterationc)
 ş
  Residualcalcc <- array(o, dim = c(windowsizec, 6, length(Cvalues)))
  for (b in 1:length(Cvalues))
```
```
ş
   Independentc <- log(CalcFD[startlocationc:(windowsizec+startlocationc-1),1])
   Dependentc <- log((CalcFD[startlocationc:(windowsizec+startlocationc-1),1] ^ Cvalues[b]) *</pre>
(CalcFD[startlocationc:(windowsizec+startlocationc-1),1] ^ -
CalcFD[startlocationc:(windowsizec+startlocationc-1),3]))
   R2SSRImc <- Imodel2(Dependentc ~ Independentc)
   Cslopec <- R2SSRImc$regression.results[3,3]
   Compinterceptc <- R2SSRImc$regression.results[3,2]
   for (i in 1:nrow(Residualcalcc))
   ş
    Residualcalcc[i,1,b] <- Independentc[i]
    Residualcalcc[i,2,b] <- Dependentc[i]
    Residualcalcc[i,3,b] <- Cslopec * Residualcalcc[i,1,b] + Compinterceptc
    Residualcalcc[i,4,b] <- Residualcalcc[i,3,b] - Residualcalcc[i,2,b]
    Residualcalcc[i,5,b] <- Residualcalcc[i,4,b]<sup>2</sup>
    Residualcalcc[i,6,b] <- Residualcalcc[i,1,b]*Residualcalcc[i,2,b]
   }
   SSRSMAc <- sum(Residualcalcc[,5,b])
   TotalSSc <- sum(Residualcalcc[,2,b]^2) - ((sum(Residualcalcc[,2,b])^2)/nrow(Residualcalcc))
   RegressionSSc <- (((sum(Residualcalcc[,6,b]))-
((sum(Residualcalcc[,1,b])*sum(Residualcalcc[,2,b]))/nrow(Residualcalcc)))^2)/(sum(Residualca
lcc[,1,b]^2)-((sum(Residualcalcc[,1,b])^2/nrow(Residualcalcc))))
   R2SMAc <- RegressionSSc / TotalSSc
   R<sub>2</sub>SMAc
   RegressionMSc <- RegressionSSc / 1
   ResidualMSc <- (TotalSSc - RegressionSSc)/(nrow(Residualcalcc)-2)
   Compslopec[k,1,b] <- RegressionMSc/ResidualMSc
   Compslopec[k,2,b] <- pf(Compslopec[k,1,b], df1=1, df2=(nrow(Residualcalcc)-2),
lower.tail=FALSE)
   Compslopec[k,3,b] <- Cvalues[b]
  }
  if (startlocationc < (delta-(windowsizec)))
  ş
   (startlocationc <- startlocationc + 1)
  } else
  ş
   startlocationc <- 1
   windowsizec <- windowsizec + 1
  }
 }
 RevisedCompslopec <- array(o, dim = c(length(Cvalues),3))
 colnames(RevisedCompslopec) <- c("F", "p", "cslope")</pre>
```

```
Compslopec[,,b]
 for (xx in 1:length(Cvalues))
 ş
  RevisedCompslopec[xx,1] <- Compslopec[iterationc,1,xx]</pre>
  RevisedCompslopec[xx,2] <- Compslopec[iterationc,2,xx]
  RevisedCompslopec[xx,3] <- Compslopec[iterationc,3,xx]</pre>
 }
 FD_iteration <- which(RevisedCompslopec[,2] == max(RevisedCompslopec[,2]))
 # Plot
 par(mar = c(5.1, 4.5, 4.1, 1))
 plot(log(CalcFD[,1]), log((CalcFD[,1] ^ Cvalues[1]) * (CalcFD[,1] ^ -CalcFD[,3])), ylim = c(-8,10),
    xlab = expression(log(delta)),
    ylab = expression(paste("log(", delta^C, " ",x, " ", delta^-D[X],")")),
    main = as.character(Rivers[q]))
 for (i in 2:nCvalues)
 { points(log(CalcFD[,1]), log((CalcFD[,1] ^ Cvalues[(i)]) * (CalcFD[,1] ^ -CalcFD[,3])), col =
(i+1))
 }
 points(log(CalcFD[,1]), log((CalcFD[,1] ^ Cvalues[(FD_iteration)]) * (CalcFD[,1] ^ -CalcFD[,3])),
col = "black", pch = 19)
 Compensated_Slope_results[g,2] <- Cvalues[FD_iteration]
 Comp_slope_Cl_results <- data.frame()
 for (i in 1:nCvalues)
 ş
  Comp_slope_CI_results[i,1] <-Cvalues[i]
  Comp_slope_Cl_results[i,2] <- Compslopec[iterationc,2,i]
  Comp_slope_CI_results[i,3] <- Compslopec[1,2,i] <0.05
  Comp_slope_CI_results[i,4] <-length(which(Compslopec[,2,i] >= 0.05))
  Comp_slope_CI_results[i,5] <-length(which(Compslopec[,2,i] >=
0.05))/length(Compslopec[,2,i])
  Comp_slope_CI_results[i,6] <- Compslopec[iterationc,2,i]
 }
 colnames(Comp_slope_CI_results) <- c("Cvalues", "p", "Significant", "N_Significant",
"PER_Significant", "LAST_VALUE_SIG")
 Cl_Range<- Comp_slope_Cl_results[Comp_slope_Cl_results$PER_Significant > o,]
 CI_Range2<- Comp_slope_CI_results[Comp_slope_CI_results$LAST_VALUE_SIG > 0.05,]
 Compensated_Slope_results[q,3] <- min(Cl_Range[,1])
 Compensated_Slope_results[q,4] <- max(Cl_Range[,1])
 Compensated_Slope_results[g,5] <- min(Cl_Range2[,1])
 Compensated_Slope_results[q,6] <- max(Cl_Range2[,1])
}
return(Compensated_Slope_results)
}
```

Script S2.

## R scripts for simulating synthetic stream stretches

#Packages required: ############# library(dplyr) library(gpplot2)

```
# Function to generate Pool-Riffle sequence:
# Generates Riffle and Pool lengths using distrubutions of riffle and pool lengths from
# Australian and Scottish study streams.
# With ability to manipulate either riffle or pool lengths.
RP_sequence <- function(River_length = 1000,
          PL_meanlog = 1.3715, PL_sdlog = 0.5711,
          RL_meanlog = 2.0488, RL_sdlog = 0.8286,
          PL_multiplier = 1, RL_multiplier = 1
){
n <- River_length/5
P_lengths <- round(rlnorm(n = n, meanlog = PL_meanlog, sdlog = PL_sdlog)*PL_multiplier)
#POOLS #o
R_lengths <- round(rlnorm(n = n, meanlog = RL_meanlog, sdlog = RL_sdlog)*RL_multiplier)
#Riffles #1
#Translate these lengths in to sequences of 1s and os for riffle and pool segments:
```

```
PRo <- as.numeric()
for (j in 1:n) {
    PR1 <- c(rep(o, P_lengths[j]), rep(1,R_lengths[j])) #Pools = o, Riffles = 1
    PRo <- c(PRo,PR1)
}
```

PRo <- PRo[1:(River\_length/5)]

#Prepare output: RP\_10\_sequence\_data <- list() RP\_10\_sequence\_data\$Sequence <- PRo RP\_10\_sequence\_data\$P\_lengths <- P\_lengths RP\_10\_sequence\_data\$R\_lengths <- R\_lengths</pre>

```
RP_10_sequence_data }
```

# Function to simulate ER counts:

#Simulate ER values for the riffle and pool sequence using distributions from high-zero study rivers:

# With ability to maipulate either riffle or pool values

RP\_ER <- function(River\_length = 1000,

PR\_size = 1.2773, PR\_mu = 43.2381, RR\_size = 1.6393, RR\_mu = 60.1088, PR\_multiplier = 1, RR\_multiplier = 1

){

n <- River\_length/5

**#**Prepare output:

```
To<- as.data.frame(matrix(o, ncol = o, nrow = n))
To$X_coord <- seq(5, River_length, by = 5)
To$Pool_ER <- round(rnbinom(n = n, size = PR_size, mu = PR_mu)*PR_multiplier) #Pools = o
To$Riff_ER <- round(rnbinom(n = n, size = RR_size, mu = RR_mu)*RR_multiplier) #Riffles = 1
```

То }

```
data1D <- cbind(Temp_RP_sequence, Temp_RP_ER[1], Temp_RP_ER[2], Temp_RP_ER[3])
data1D <- mutate(data1D, ER = ifelse(Sequence == 0, Pool_ER, Riff_ER), RP = ifelse(Sequence
== 0, "P", "R"))
data1D <- data1D[, c(4,7,8)] # neat dataset
data1D$River <- "Synthetic_stream_01"</pre>
```

#Plot simulation: ggplot(data1D, aes(x = X\_coord, y = ER, fill = RP))+ geom\_bar(stat="identity", width = 5)