1   **Clustered versus catastrophic global vertebrate declines**
2
3   Brian Leung[1,2*]
4   Anna L. Hargreaves[1]
5   Dan A. Greenberg[3]
6   Brian McGill[4]
7   Maria Dornelas[5]
8   Robin Freeman[6]
9
10   [1]Department of Biology, McGill University, Montreal, Quebec, Canada H3A 1B1
11   [2]School of Environment, McGill University, Montreal, Quebec, Canada, H3A 2A7
12   [3] Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia,
13   Canada V5A 1S6
14   [4]School of Biology and Ecology and Mitchell Center for Sustainability Solutions, University of
15   Maine, Orono, ME USA
16   [5] Centre for Biological Diversity, University of St Andrews, St Andrews KY16 9TH, UK
17   [6] Indicators and Assessments Unit, Institute of Zoology, Zoological Society of London, London,
18   NW1 4RY
19   *Corresponding author email: brian.leung2@mcgill.ca
20
21   **Summary**
22   Recent analyses have reported catastrophic global declines of vertebrate populations[1,2]. Yet
23   distilling many trends into a global mean index obscures variation that can inform conservation,
24   and can be sensitive to analytical decisions. For example, whereas earlier analyses estimated a
25   mean vertebrate decline of >50% since 1970 (Living Planet Index: LPI[2]), we find that this
26   estimate is driven by <3% of populations; excluding these extremely declining populations
27   switches the global trend to an increase. The sensitivity of global mean trends to outliers
28   suggests that more informative indices are needed. We propose an alternative approach,
29   identifying clusters of extreme decline (or increase) that differ statistically from the majority of
30   population trends. We show that, of LPI's 57 taxonomic-geographic systems, 16 systems
31   contain clusters of extreme decline (comprising ~1% of populations, occurring disproportionately
32   in larger animals) and 7 contain extreme increases (~0.4% of populations). The remaining
33   98.6% of populations across all systems showed no mean global trend. However, when
34   analyzed separately, three systems were declining strongly with high certainty (all Indo-Pacific),
35   and seven were declining strongly but with less certainty (mostly reptile-amphibian groups).
36   Accounting for extreme clusters fundamentally alters interpretation of global vertebrate trends
37   and should be used to help prioritize conservation effort.
38
39

Main

Rapid global change is threatening species across the globe[1]. Quantifying biodiversity trends is important to assess whether current investment is slowing or reversing declines, and identify regions and taxa of concern. While distilling disparate population trends into a single global index can focus attention on biodiversity trends[2,3,4], simple metrics can distort the full picture.

Estimates of global biodiversity trends vary depending on their data and mathematical model. The most apocalyptic garner extensive press coverage, even when based on controversial data (e.g., 'biological annihilation'[5] using trend estimates based largely on expert opinion; 'insect Armageddon' based on data disputed by the original collectors[6]). But even analyses of the best available data reach conflicting results. Analysis of a global dataset of abundance time-series estimated that on average vertebrate populations have declined by >50% since 1970 (Living Planet Index[2]:LPI), but other global analyses found that mean population size[7,8] and species richness[9,10] have remained stable over similar timeframes. Explanations for the discrepancies have been proposed[11,12,8,13], but not resolved.

One crucial consideration is that summary indices may be easily misinterpreted. Calculating the geometric mean across populations is the most common and straightforward approach, but is strongly influenced by extremes. To illustrate, imagine an ecosystem in which one population declined by 99%. Even if a second population increased 50 fold or 393 populations increased by 1% (i.e. a large net increase) a geometric mean would show a catastrophic 50% decline. Thus, a geometric mean decline of 50% could arise from significant, widespread loss is occurring across many populations (we term this the '*Catastrophic declines*' hypothesis), or from a few extremely-declining populations (we term this the '*Clustered declines*' hypothesis). Both scenarios involve important conservation issues, but suggest vastly different underlying problems and mitigation strategies[14], thus distinguishing between them is of real-world importance.

We derive a Bayesian hierarchical mixture (BHM) model to distinguish between the *Catastrophic* and *Clustered* decline hypotheses. The model statistically separates population trends into extreme declines, typical trends, and extreme increases (Fig 1), while accounting for time-series size, within population fluctuations, number of populations, and among-population variance. We test declines in abundance for >14 000 vertebrate populations (LPI[21]). We chose LPI data for its impressive scope, because the data and analytical details were publicly available, and because previous analyses of these data suggested widespread, global declines[2].

We first examined whether the previous mean decline estimate of >50%[2] was sensitive to extreme populations: robust declines would support the *Catastrophe* hypothesis; high sensitivity to few populations would support the *Cluster* hypothesis (Fig 1). We then applied our BHM model to assess the evidence for *Catastrophic* vs. C*lustered* declines globally and by region, taxonomy. Finally, we explore two additional conservation issues. First, we test whether declines occur disproportionately in larger animals (big animals tend to have lower reproductive rates), which might release small animals from predation[18]. Second, previous analyses often excluded time-series with few datapoints[10,12,19], but small time-series make up most of available data. We test the effects of their exclusion[20].

Results

Sensitivity of Geometric Mean index to extreme populations

The geometric mean underlying the LPI analysis was highly sensitive to extreme populations. Excluding only the 2.4% most strongly declining populations (354 of 14700 populations) reversed the estimate of global vertebrate trends from >50% loss to slightly positive growth (Fig 2). Similarly, excluding 2.4% of the most strongly increasing populations strengthened the mean decline to 71%. High sensitivity suggests that extreme populations are disproportionately affecting global trend estimates, such that clusters of extreme population decline should be considered explicitly.

Evidence for *Clustered* declines

Among LPI's 57 domain-realm-taxon systems, 16 systems contained clusters of extreme decline, and 8 contained clusters of extreme growth (3 systems are repeated, i.e. had both clusters of extreme decline and growth; Fig 3, SI Table S2). Together, clusters of extreme decline accounted for only 1% of populations across systems (2% of populations in the 16 systems in which they occurred). The mean population trend for extreme declining clusters across the 16 systems was $\theta_2 = -3.94$, or approximately 98% loss per year, and deviated substantially from the mean trend of the primary cluster in those systems. Clusters of extreme growth accounted for 0.4% of populations across systems (2.4% in the 8 systems where they occurred), with $\theta_2 = 3.51$, ie an explosive 33x growth per year (Fig 3, SI Table S2).

Extreme clusters showed some taxonomic and geographic patterns. The largest cluster of extreme declines was in arctic marine mammals, accounting for 7.6% of populations in that system. However, mammal systems generally had the fewest clusters of extreme decline [19% of 16 systems], followed by reptile/amphibian systems [21% of 14 systems], whereas bird and fish systems had more clusters of extreme declines [31% of 16 and 45% of 11 systems, respectively; Fig 3]. Clusters of extreme decline occurred throughout the world with half occurring in marine realms, whereas extreme increases occurred more in temperate regions or terrestrial realms (Fig 3).

Extreme population trends occurred predominantly in small time-series. Excluding time-series with <10 points removed all but two extreme clusters, but also removed 52% of the data (see SI Table S3). The higher frequency of extreme trends among small time-series was also apparent in the raw data (Figure 4). Thus the decision of whether to include small time-series will have large effects on the resulting estimates of global trends.

Body size was related to population trends. Larger species had three times more extreme declines than increases (15 vs 5 clusters of extreme decline vs extreme increase). Comparatively, smaller species had half as many (8) extreme declining and disproportionately more (7) extreme increasing clusters (SI Table S4). While size-specific models included fewer populations, especially for smaller species, the number of clusters was not uniformly lower (as might be expected given a reduction in power), so the differential occurrence of declining versus increasing extreme clusters suggests large animals are more vulnerable to extreme declines.

128

Evidence for *Catastrophic* declines

In contrast to the extreme clusters, the primary clusters accounted for the vast majority (98.6%) of populations across the 57 LPI systems. The overall growth rate of primary clusters was close to zero: θ_1 = -0.00035, corresponding to ~1.7% loss over 50 years, given a constant rate across populations and time (Fig 5). Also, in contrast to extreme clusters, primary cluster trends were robust to time-series size, as excluding series with <10 data points yielded a similar overall global trend (θ_1 = +0.0043) (Extended data Figure E3).

While the global BHM model reveals considerably more nuance than a geometric mean index, analyzing across systems still masked important patterns. When systems were analyzed separately (SI Table S2), primary population clusters were strongly declining (θ_1 < -0.015) with high certainty (95% credible intervals not overlapping zero) in three systems, all in the Indo-Pacific Realm (freshwater mammals, freshwater birds, terrestrial birds; red distributions, Fig 3). This suggests this region has the highest risk of system-wide decline and should be a conservation priority. In contrast, the primary cluster was increasing with high certainty in seven systems, six of which were in temperate regions. Seven additional systems had strongly declining primary population clusters but with less certainty (95% credible intervals overlapped zero), four of which were amphibian/reptile groups. 14 systems showed strong but low-certainty increases, with no obvious taxonomic nor geographic patterns (Fig 3).

Each primary cluster also contained variation among populations. In the 10 systems with significant (red) or non-significant (orange) mean declines (θ_1 < -0.015), 87% of the individual populations showed strong declines (Fig 5). These 10 systems accounted for ~20% of total global vertebrate populations, but ~61% of strong declines. The multimodality observed in Fig 5, was an outcome of aggregating unimodal primary clusters across systems, and suggests heterogeneous stressor levels among systems (i.e., similar principles as those causing extreme clusters within systems). The remaining ~11% of strongly declining populations were distributed across 47 of 57 systems; it is unclear whether they represent a deviation from natural dynamics expected in any naturally variable system.

Primary cluster trends were related to body size, but not as predicted. Compared to overall patterns, for larger animals the same systems showed significant declines and increases, but two additional temperate systems showed significant increases (Extended data Fig E4; SI Table S4). Smaller species also appeared to decline more than larger species; there were 27 systems where smaller species had more negative growth rates than larger species, versus 18 where the reverse was true. However, analyses of smaller species were based on a substantially fewer populations and trends were generally not significant (SI Table S4), so patterns remain tentative.

Discussion

By re-analysing the most comprehensive dataset of global wildlife population trends available, we show that previously estimated global declines are driven by a few extremely declining populations. Removing only 2.4% of declining populations reversed the estimated global trends from >50% mean decline since 1970 to slightly positive growth. Our BHM model revealed that

4

172 clusters of extreme decline are widespread and occur disproportionately in larger species, and
173 that a few clusters of extreme increase also exist and occur disproportionately in smaller
174 species. This is consistent with previous arguments of "trophic downgrading"[18].
175
176 Clusters of extreme declines were largely due to small time-series. However, neither random
177 sampling error nor "saw tooth" population dynamics (wherein ultimately stable populations
178 experience sudden declines followed by gradual increases) can fully explain this association
179 (see Supplementary information for full discussion). Additional explanations are needed.
180 Extreme trends could reflect transient populations that naturally leave or enter a survey area[22],
181 which could represent natural dynamics. Alternatively, researchers may stop sampling after
182 populations become (close to) extirpated, although the converse has also been suggested[23]. A
183 third possibility is that some regions experience both lower sampling effort and greater declines,
184 such that poor data correlates with factors linked to vulnerability, like lower national wealth or
185 conservation investment. Understanding why small time-series contain so many extreme
186 declines is particularly important given that studies that did not find widespread declines often
187 excluded short time-series[e.g., 7,10,12], potentially reconciling divergent findings among studies.
188
189 Once extreme clusters were statistically separated, no global trend remained across typical
190 populations (i.e. primary clusters; 98.6% of populations). However, aggregating systems into
191 one global trend hid important variation. Three systems, all in the Indo-Pacific, showed
192 widespread vertebrate declines across typical populations. Moreover, among typical populations
193 smaller species may be faring worse than larger ones. While tentative given lower sample sizes
194 and high uncertainty, this trend is contrary to common conservation assumptions and so merits
195 additional research.
196
197 Our results emphasize an important point: biodiversity trends within and across regions and
198 taxa are highly disparate. This likely reflects differences in both susceptibility and exposure to
199 anthropogenic environmental change[24, 25, 26]. Unravelling this variation is imperative to
200 understand where biodiversity is most threatened[27] and which conservation actions promote
201 stability or recovery. A productive global conversation about conservation requires that both
202 scientists and media pay more attention to variation and resist the temptation of simple
203 summary indices.
204
205 Shifting the message from ubiquitous catastrophe to foci of concern, also touches on human
206 psychology. Continual negative and guilt-ridden messaging can cause despair, denial and
207 inaction[28,29]. If everything is declining everywhere, despite the expansion of conservation
208 measures in recent decades, it would be easy to lose hope. Our results identify regions that
209 need urgent action to ameliorate widespread biodiversity declines, but also many systems that
210 appear to be generally stable or improving, and thus reason to hope that our actions can make
211 a difference.
212 **References**
213

214   1. IUCN. The IUCN Red List of Threatened Species. Version 2019-3.
215   <http://www.iucnredlist.org>. (2019).
216   2. WWF. Living Planet Report 2018: Aiming higher (eds. Grooten N & Almond REA). WWF,
217   Gland, Switzerland. (2018).

3. Rosenberg, K. V et al. Decline of the North American avifauna. *Science* **366**, 120 – 124 (2019).

4. Sánchez-Bayo, F. & Wyckhuys, K. A. G. Worldwide decline of the entomofauna: A review of its drivers. *Biol. Conserv.* **232** 8–27 (2019).

5. Ceballos, G., Ehrlich, P. R. & Dirzo, R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6089–E6096 (2017).

6. Willig, M. R. et al. Populations are not declining and food webs are not collapsing at the Luquillo Experimental Forest. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 12143–12144 (2019).

7. Daskalova, G. N., Myers-Smith, I. H. & Godlee, J. L. All is not decline across global vertebrate populations. Preprint at https://www.biorxiv.org/content/10.1101/272898v5. (2018).

8. Dornelas, M. et al. A balance of winners and losers in the Anthropocene. *Ecol. Lett.* **22**, 847–854 (2019).

9. Vellend, M. et al. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19456–19459 (2013).

10. Dornelas, M. et al. Assemblage time series reveal biodiversity change but not systematic loss. *Science* **344**, 296–299 (2014).

11. Gonzalez, A. et al. Estimating local biodiversity change: A critique of papers claiming no net loss of local diversity. *Ecology* **97**, 1949–1960 (2016).

12. Leung, B., Greenberg, D. A. & Green, D. M. Trends in mean growth and stability in temperate vertebrate populations. *Divers. Distrib.* **23**, 1372–1380 (2017).

13. McGill, B. J., Dornelas, M., Gotelli, N. J. & Magurran, A. E. Fifteen forms of biodiversity trend in the anthropocene. *Trends Ecol. Evol.* **30**, 104–113 (2015).

14. Anderson, S. C., Branch, T. A., Cooper, A. B. & Dulvy, N. K. Black-swan events in animal populations. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 3252–3257 (2017).

15. Carrington, D. Humanity has wiped out 60% of animal populations since 1970, report finds. *The Guardian*. <https://www.theguardian.com/environment/2018/oct/30/humanity-wiped-out-animals-since-1970-major-report-finds>. (2018)

16. Chung, E. 60% of world's wildlife has been wiped out since 1970. <https://www.cbc.ca/news/technology/living-plant-wwf-2018-1.4882819>. *CBC News*. (2018)

17. Duncan, R. P., Boyer, A. G. & Blackburn, T. M. Magnitude and variation of prehistoric bird extinctions in the Pacific. Proc. Natl. Acad. Sci. U. S. A. **110**, 6436–6441 (2013).

18. Estes, J.A. et al. Trophic downgrading of planet Earth. Science **333**, 301-306 (2011).

19. Connors, B. M., Cooper, A. B., Peterman, R. M. & Dulvy, N. K. The false classification of extinction risk in noisy environments. *Proc. R. Soc. B Biol. Sci.* **281**, 20132935 (2014).

20. Hanks, E. M., Hooten, M. B., Baker, F. A. Reconciling multiple data sources to improve accuracy of large-scale prediction of forest disease incidence. *Ecol. Appl.* **21**, 1173–1188 (2011).

21. LPI. Living Planet Index database. <www.livingplanetindex.org/>. (2016).

22. Youngflesh, C. & Lynch, H. J. Black-swan events: Population crashes or temporary emigration? *Proc. Natl. Acad. Sci. U. S. A.* **114**, E8953–E8954 (2017).

23. Fournier, A. M. V., White, E. R. & Heard, S. B. Site-selection bias and apparent population declines in long-term studies. *Conserv. Biol.* **33**, 1370–1379 (2019).

263   24. Newbold, T. et al. Ecological traits affect the response of tropical forest bird species to
264   land-use intensity. *Proc. R. Soc. B Biol. Sci.* **280**, (2013).
265   25. Venter, O. et al. Sixteen years of change in the global terrestrial human footprint and
266   implications for biodiversity conservation. *Nat. Commun.* **7**, 1–11 (2016).
267   26. Allan, J. R. et al. Hotspots of human impact on threatened terrestrial vertebrates. *PLoS
268   Biol.* **17**, (2019).
269   27. Blowes, S. A. et al. The geography of biodiversity change in marine and terrestrial
270   assemblages. *Science* **366**, 339–345 (2019).
271   28. O'Neill, S. & Nicholson-Cole, S. "Fear won't do it" Visual and iconic representations. *Sci.
272   Commun.* **30**, 355–379 (2009).
273   29. Brennan, L. & Binney, W. Fear, guilt, and shame appeals in social marketing. *J. Bus.
274   Res.* **63**, 140–146 (2010).

275
276

Figure Legends

Figure 1. Stylized patterns of system-wide growth rates. Similar geometric mean population growth rates [$\log(N_{t+1}/N_t)$] can reflect contrasting systems. As a "null" model, systems can be stable (log growth rates centered at zero, middle panel). Deviations can occur in multiple ways. System can have most populations substantially declining ("*Catastrophic*" Hypothesis, top panel) or can have multiple clusters with the majority of populations showing a distribution of growth rates centered around zero, but a small cluster of populations experiencing extreme declines ("*Clustered*" Hypothesis, top-middle panel). Each has the same metric of mean decline (vertical red line = 1.5% annual decline, corresponding to a 50% loss over 50 years), even though most populations in the top-middle panel are stable. The converse can also happen, with a small cluster of extremely increasing populations, but otherwise a stable distribution (bottom-middle panel), or most populations increasing (bottom panel) (vertical blue line = 1.5% annual increase, corresponding to a doubling over 50 years).

Figure 2. Effect of extreme populations on global growth index. Removing a small fraction of extreme populations strongly influences the geometric growth index, using the LPI dataset. Each line represents a different number of populations removed, ranging from no removals (red line: keeping all 14,700 populations, showing a >50% mean decline), to removing 356 populations (yellow line: <2.4% of populations removed switches the global trend from negative to positive). A geometric growth index of one indicates no change (dashed horizontal black line).

Figure 3. Population trends by taxonomic groups and realms. Top panel shows the terrestrial realm, Middle Panel shows the freshwater realm, and bottom Panel shows the marine realm. Red and blue asterisks denote occurrence of extreme declining clusters (16 systems) and increasing clusters (8 systems), respectively. Distributions show the primary cluster in each system (Red = significant declines, Blue = significant increases, Orange = strong non-significant declines, Green = strong non-significant increases, Yellow = weak changes).

Figure 4. Effect of time series size. Number of data points in time series versus the mean log of the geometric mean growth rate.

Figure 5. Populations in the primary clusters across all systems, after removal of extreme clusters. The primary cluster of each system is unimodal, but because systems are experiencing decline (or growth) heterogeneously, plotting distributions across systems shows multimodality. Histograms show significantly declining systems (red), strongly but not significantly declining systems (orange), and weak changes or increases (yellow). Vertical lines show thresholds for strongly declining (-0.015) and strongly increasing (+0.015) growth rates, corresponding to ~50% loss or a doubling (over 50 years), respectively. Distributions of primary clusters were calculated based on the mean and standard deviations from the hierarchical model, and using the system specific weights to adjust for species richness.

318 **Methods**

319 <u>Dataset</u>

320

321 The publically available LPI data includes 15241 vertebrate populations from 3510 species[21].
322 When a species contained both finer resolution estimates within a country (2593 entries) and a
323 country-wide aggregate, we excluded the country-wide aggregate (537 entries), yielding 14700
324 populations. LPI groups species into 57 systems defined by a combination of habitat domain
325 (terrestrial, freshwater, marine), biogeographic realm (terrestrial/freshwater realms =
326 Afrotropical, Nearctic, Neotropical, Palearctic, Indo-Pacific; marine = Arctic, Atlantic north
327 temperate, Atlantic tropical/sub-tropical, Pacific north temperate, Indo-Pacific tropical/sub-
328 tropical, South-temperate/Antarctic), and taxonomic grouping (Fish=Actinopterygii,
329 Elasmobranchii, Holocephali, Myxini, Chondrichthyes, Sarcopterygii, Cephalaspidomorphi;
330 Birds=Aves, Mammals=Mammalia, Herps = Amphibia, Reptilia) (Extended data Figure E5-8).
331

332 To analyze the effect of body size, we obtained information on each taxonomic group. Given the
333 diversity of vertebrate groups in this data set, and the different conventions across groups, we
334 used different measures of body size for each taxonomic Class based on data availability. For
335 birds (N = 1397), mammals (N = 534), and reptiles (Squamata, N = 132; Testudines, N = 44;
336 and Crocodylia, N = 16) we used estimates of species' mass, in grams, collated in an extensive
337 comparative data set[30]. When mass data were missing for a species (N = 14 birds; N = 1
338 mammal; N = 25 reptiles), we estimated body mass as the geometric mean of available mass
339 estimates for species in that genus. For fishes (Chondrichthyes, Osteichthyes, Agnatha; N =
340 1211), estimates of mass were scarce for most species, so we instead used estimates of total
341 length or standard length, in centimetres, both of which were extracted from FishBase[31] using
342 the *rfishbase* R package[32]. These length estimates are an imperfect proxy for size (in terms of
343 mass) given the variability in body plans across groups, but given the large amount of variation
344 across these groups it suffices as a way to broadly categorize species into distinct size classes.
345 For amphibians, we used estimates of snout-vent length (SVL, in mm) as our proxy for body
346 size, as this is the most widely available metric of size across species. Data on SVL for
347 amphibian species (N = 175) was extracted from a comprehensive ecological trait dataset:
348 AmphiBio[33].

349 <u>Sensitivity of the "Geometric" indices to extreme population trends</u>

350

351 The LPI analysis was based on a geometric mean approach, calculated by summing across
352 logged growth rates[34]. We recreated the geometric-mean based analyses (see Supplementary
353 Information 1a for full details and model formulation), and examined the sensitivity of the global
354 estimate to extreme populations. We ordered populations and sequentially removed the largest
355 observed decline, determining the effect of each removal on the global estimate of biodiversity
356 loss. Low sensitivity would indicate that many/most populations are declining, supporting the
357 *Catastrophic* decline hypothesis. High sensitivity, i.e. if removal of relatively few populations
358 switched the strongly negative global trend to neutral or positive, would support the *Clustered*

359 declines hypothesis. For balance, we also examined sensitivity to sequential removal of the
360 greatest increasing populations.

361 _Catastrophic_ vs. _Clustered_ declines approach (BHM model)

362

363 We developed an approach to separate extreme population clusters whose growth or decline
364 statistically deviated from typical population trends, such that a small number of extreme
365 populations would no longer mask trends of the majority of populations (Fig 1). While some
366 summarization is needed to understand global trends, heterogeneous growth rates and
367 potentially multimodal distributions could be expected, given multiple stressors with diverse
368 effects, and differences in species vulnerabilities. We used a Bayesian Hierarchical Mixture
369 (BHM) model as our statistical architecture, as it has several desirable properties: 1) can
370 represent the "null" model and assess deviations from it; 2) allows testing for both negative and
371 positive extremes (sometimes both existed in the same system); 3) quantifies the magnitude
372 and proportion of those extremes; 4) provides a coherent way to separate extreme populations
373 from the majority of populations (the primary cluster), enabling tests of the Clustered and
374 Catastrophic declines hypotheses; 5) provides a measure of uncertainty as a direct outcome of
375 analysis (via the posterior distribution); and 6) accounts for population fluctuations and adjusts
376 for the number of data points in time-series.

377
378 First, we specify the "null" model. Even in a system with no overall trend, we expect stochastic
379 fluctuations in population size. We also expect some populations to be increasing or decreasing
380 during any time interval, given complex, real-world ecological dynamics. Thus, the null model
381 should include among-population heterogeneity, and therefore consists of a distribution of
382 growth rates (Fig 1 middle panel). Statistical deviations from this null model could be caused by
383 a shift in the overall distribution, whereby a system-wide mean growth <0 (i.e., decline) could
384 indicate risk to the entire system, and would support the _Catastrophic declines_ hypothesis (Fig 1
385 top panel). Alternatively, statistical deviation from the null model could be caused by a few
386 populations experiencing extreme declines, consistent with the _Clustered declines_ hypothesis
387 (Fig 1 mid-top panel).

388
389 To specify our model, we begin with a standard Bayesian hierarchical formulation (i.e., it does
390 not yet contain mixtures of distributions). We define θ, τ as the system-wide mean and variance,
391 respectively, of log-growth rates across all populations in the system (i.e., hyperparameters in
392 Bayesian terminology). θ, τ determine the distribution of log population trends (μ_i), and define
393 the properties of the overall system. However, within population dynamics are also occurring,
394 and the log growth rate for population i at time t are modelled as a population trend (μ_i) and
395 within population fluctuations (σ) (See Supplementary Information 1b for full details and model
396 formulation).

397
398 Using a standard Bayesian hierarchical model, we can test the _Catastrophe_ hypothesis by
399 determining the probability that the system-wide mean θ is <0. Testing the _Cluster_ hypothesis
400 however requires a mixture model to assess the evidence for the occurrence of clusters. Thus,
401 we define K is the number of clusters in the mixture, f_k is the fraction of populations in the kth
402 cluster, and θ, τ and f denote the vectors of parameters for the K clusters.

403
404   To test the *Clustered* hypothesis we modeled three clusters: a primary cluster, corresponding to
405   the typical trend; a negative extreme cluster; and a positive extreme cluster (Figure 1). Although
406   our main interest was in the mechanisms behind apparent global population declines (i.e.
407   *Catastrophic* vs. *Clustered* hypotheses), we also assayed positive extreme clusters so that
408   analyses were not biased to find only negative population trends. We considered four cluster
409   combinations: 1) a single distribution; 2) a primary distribution and a negative extreme
410   distribution; 3) a primary distribution and a positive extreme distribution; or 4) a primary
411   distribution and both positive and negative extreme distributions (Figure 1). For referencing
412   purposes, we denote $k=1$ as the primary cluster, $k=2$ as the negative extreme cluster, and $k=3$
413   as the positive extreme cluster. Note that reality need not be bi(or tri)-modal, but exploring
414   generalities in trends necessitates some aggregation. Nonetheless, the extreme clusters
415   identified by the mixture model could contain multiple extreme modes in the data (or even result
416   from a skewed distribution). With any of these deviations, model selection (see below) would
417   still choose the mixture model as explaining the data better than a single normal distribution.
418   (See Supplementary Information 1c for full details and model formulation).
419
420   We used the (lowest) Deviance Information Criteria (DIC) value to select the mixture model with
421   the strongest statistical evidence[35]. The *Catastrophic* declines hypothesis would be supported
422   by a mean decline of the primary population cluster ($\theta_1 < 0$ and credible intervals did not
423   overlap zero), and would be particularly severe if the mean θ_1 was also strongly negative (e.g.,
424     $\theta_1 = -0.015$ would correspond to >50% loss over 50 years). The *Clustered* declines hypothesis
425   would be supported if the DIC selected a mixture with a negative extreme cluster (combinations
426   2 or 4 above). Note that the *Catastrophic* and *Clustered declines* hypotheses are not mutually
427   exclusive, as a system could have both a negative extreme cluster and declining primary
428   cluster. A large fraction of populations in the negative extreme cluster (f_2) could also be
429   interpreted as widespread *Catastrophic* declines, but this did not occur in our results. Although
430   our hypotheses focus on understanding declining trends, our model will also detect increases in
431   abundances.
432
433   To estimate model parameters we used Bayesian analyses and the Markov Chain Monte Carlo
434   algorithm, which simultaneously estimated uncertainty. For each Bayesian analysis, we ran 3
435   chains, each with 10000 iterations (3000 used for warm-up). Convergence was determined
436   using Rhat ≈ 1. Values for all parameters across all systems ranged from (0.999 < Rhat <
437   1.005). Bayesian analyses were conducted using the STAN language[36], and processed and
438   analyzed in R[37].
439
440   Additionally, we explored the theoretical behavior of each model, including the
441   the geometric mean model, in the presence of clustered declines (see Supplementary
442   Information 1d, 2a), and our *Catastrophic* and *Clustered declines* approach given our selection
443   of priors, application of constraints, and other modeling choices; these simulation analyses
444   showed that our approach yielded appropriate theoretic behavior (see Supplementary
445   information 1e, 2b, Extended data Figure E1). Finally, we conducted sensitivity analyses and
446   showed that results were robust to modelling choices (see Supplementary Information 2c, Table
447   S1, Extended data Figure E2).
448

449

Application of the *Catastrophe* and *Clusters* approach to LPI data

We tested for extreme clusters in each of LPI's 57 domain-realm-taxon systems, by choosing the mixture model with the lowest DIC value. We also examined the number of populations in each cluster, as a fraction of the total number of populations, scaled using the LPI system-specific weightings[38] (see Supplementary Information 1f for more details).

Next, we examined evidence for the *Catastrophic declines* hypothesis in each system by assaying for negative mean growth rates in the primary cluster (θ_1). We defined "high certainty" of decline (or increase) as 95% credible intervals not overlapping zero, and "strong" decline as $\theta_1 < -0.015$, corresponding to a ~50% decline if it persisted for 50 years ($\theta_1 > 0.015$ was used for a strong positive relations, corresponding to a doubling over 50 years).

We assessed the effect of small time-series on both extreme clusters and trends in primary clusters, by omitting all data with fewer than 10 points, as has often been done in other studies[e.g.,12]. These small time-series accounted for 52% of the population estimates (7110 populations remained in the analysis).

Finally, we examined whether trends differed between large versus small-bodied animals. Within each Class (but with Agnatha lumped with Osteichthyes), we scaled body size as standard deviations on the natural log-scale – thereby creating an index of relative species size within a taxonomic group. In two cases, we separated out different groups within a Class that had relatively distinct body plans that would influence this size scaling. We scaled size within the Superorder Batoidea (Rajiformes, Myliobatiformes, and Torpediniformes) and separately scaled size for the rest of the Chondrichthyans (Selachimorpha and Holocephali). For the amphibians, we separated out the Orders Caudata and Anura and scaled size within each of these groups. For each taxonomic group we scaled body size and separated species into larger-than-average (henceforth 'larger') versus smaller-than-average (henceforth 'smaller') species. This yielded 9596 populations from 1765 larger species, and 5103 populations from 1745 smaller species. We then reran the BHM model for larger animals and again for smaller animals. Body sizes were divided unevenly among habitat domains and realms; 12 domain-realm-taxon systems contained ≤1 smaller species so were excluded from the small-animal model.

Methods References

30. Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., Ernest, S. K. M., An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles. Ecology **96**, 3109 (2015).
31. Froese, R. & Pauly, D. FishBase. World Wide Web electronic publication. version (12/2019) (2019). Available at: www.fishbase.org.
32. Boettiger, C., Lang, D. T. & Wainwright, P. C. rfishbase: exploring, manipulating and visualizing FishBase data from R. J. Fish Biol. **81**, 2030–2039 (2012).

493   33. Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C. & Costa, G. C.
494   AmphiBIO, a global database for amphibian ecological traits. Sci. Data **4**, 170123
495   (2017).
496   34. Collen, B. et al. Monitoring Change in Vertebrate Abundance: the Living Planet Index.
497   *Conserv. Biol.* **23**, 317–327 (2009).
498   35. Gelman, A., Hwang, J., Vehtari, A. Understanding predictive information criteria for
499   Bayesian models. *Statistics and computing* **24**, 997-1016 (2014).
500   36. Carpenter, B. et al. Stan: A probabilistic programming language. *J. Stat. Softw.* **76**,
501   (2017).
502   37. R Core Team. R: A language and environment for statistical computing. R Foundation
503   for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>. (2016).
504   38. McRae, L., Deinet, S. & Freeman, R. The diversity-weighted living planet index:
505   Controlling for taxonomic bias in a global biodiversity indicator. *PLoS One* **12**, 1–20
506   (2017).

507 **Acknowledgements**
512

513 **Author contributions**
514 Authors are listed in order of their contributions. BL formulated the BHM model, conducted
515 analyses, and wrote the majority of the ms. AH discussed and clarified the ideas, and also
516 played a central role in writing. DG discussed and clarified the ideas, synthesized the data, and
517 contributed to writing. BM discussed and clarified the ideas, and commented on the ms. MD
518 discussed and clarified the ideas, commented on and improved presentation of the ms. RF
519 discussed and clarified the ideas, and provided insight into the LPI data and analyses.
520

521 There are no competing financial interests.
522

523 **Statement of data availability:**
524 Data can be obtained from the Living Planet Index database. <www.livingplanetindex.org/>.
525 (2016), AmphiBio database from
526 https://figshare.com/articles/Oliveira_et_al_AmphiBIO_v1/4644424>, fishbase database
527 <www.fishbase.org>, and lifehistory traits from
528 <https://doi.org/10.6084/m9.figshare.c.3308127.v1>
529

530 **Statement of code availability:**
531 Code for the BHM model is available at: https://doi.org/10.5281/zenodo.3901586
532

533 **Supplementary information** is available for this paper.
534

535 Correspondence and requests for materials should be addressed to brian.leung2@mcgill.ca
536

537

Extended Data Figure Headings

Figure E1. p-p plots showing that the posterior distributions for each estimated parameter are unbiased, and largely follow a 1:1 line for each hyper parameter (σ, τ) as well as the fraction in each cluster (f_1, $f_2=1-f_1$). The 1:1 line is the theoretic expectation, indicating that the true parameter value to fall below the 0.01 quantile 1% of the time, 0.02 quantile 2% of the time, etc.

Figure E2. The trends of the primary clusters (θ_1), for the main analysis (x-axis) versus the sensivity analysis (y-axis) for the threshold for extreme clusters (top panel) and the offset when N=0 was observed (bottom panel).

Figure E3. Each point represents a trend estimate for the primary cluster of a system, with the full dataset (x-axis) versus data excluding time series with <10 points (y-axis). The red dot was Freshwater Indo-Pacific Mammals, which was reduced from 22 populations (full) to 2 populations (only data >=10 points).

Figure E4. Mean trends of primary clusters across systems calculated using the Bayesian Hierarchical Mixture Model. Top) all species (14,700 populations). Middle) only large species (9596 populations). Bottom) only small species (5103 populations). The small species appear to be declining more than large species, although this needs to be interpreted with caution, as most primary distributions did not significantly deviate from zero for small species.

Figure E5. Histograms of observed growth rates and output of the Bayesian Hierarchical Mixture model for systems 1-16 (blue line – primary cluster, red line – extreme cluster(s) from model). Grey vertical lines show the range of observed values. In comparing model output to data: 1) the variation of the BHM primary cluster (blue line) is much lower than the raw data, because the BHM separates variation in among population trends from variation due to within population fluctuations. 2) The BHM model identifies evidence for extreme clusters in both (e.g., Terrestrial Indo-Pacific birds) or only one direction (e.g., Terrestrial Neotropical mammals), but not other apparent clusters (e.g., Terrestrial Indo-Pacific herps). The BHM integrates the magnitude of within population fluctuations, time-series sizes, number of populations, among-population variance, and the magnitude and frequency of the extreme populations in determining whether additional (extreme) clusters are needed to account for the observations.

Figure E6. Histograms of observed growth rates and output of the Bayesian Hierarchical Mixture model for each system 17-32 (blue line – primary cluster, red line – extreme cluster(s) from model). Grey vertical lines show the range of observed values.

Figure E7. Histograms of observed growth rates and output of the Bayesian Hierarchical Mixture model for each system 33-48 (blue line – primary cluster, red line – extreme cluster(s) from model). Grey vertical lines show the range of observed values.

Figure E8. Histograms of observed growth rates and output of the Bayesian Hierarchical Mixture model for each system 49-57 (blue line – primary cluster, red line – extreme cluster(s) from model). Grey vertical lines show the range of observed values.