

1 **Assessing the sensitivity and robustness of the Living Planet Index through simulated**
2 **population dynamics: strengths, stability, and challenges**

3 Cristian A. Cruz-Rodríguez^{1,2}, Gaëlle Mével³, Janaína A. Serrano^{2,4}, Shuaishuai Li^{2,4}, Jessica
4 Currie⁵, Maria Isabel Arce-Plata^{1,2}, Sarah Ravoth⁶, Timothée Poisot^{1,2}, Louise McRae^{7,8}, Robin
5 Freeman^{7,8}, Valentina Marconi^{7,8}, Sandra Emry⁶ and David AGA Hunt⁴

6 ¹ Université de Montréal, Biological Sciences, 1375 Ave. Thérèse-Lavoie-Roux, Montréal, CA,
7 H2V 0B3.

8 ² Quebec Centre for Biodiversity Science, 1205 Dr. Penfield Ave, Montreal, CA, H3A 1B1.

9 ³ McGill University, Department of Geography, 805 Sherbrooke St W, Montreal, CA, H3A 0B9.

10 ⁴ McGill University, Department of Biology, 1205 Dr. Penfield Ave, Montreal, CA, H3A 1B1.

11 ⁵ WWF-Canada, 410 Adelaide Street West, Toronto, CA, M5V 1S8.

12 ⁶ University of British Columbia, Department of Zoology, 6270 University Blvd 4200, CA, V6T
13 1Z4.

14 ⁷ Institute of Zoology, Zoological Society of London, Outer Circle, London, UK, NW1 4RY.

15 ⁸ Research Department of Genetics, Evolution and Environment (GEE), Division of Biosciences,
16 Faculty of Life Sciences, University College London, Gower St, Bloomsbury, London WC1E
17 6BT, UK.

18 Corresponding author: Cristian A. Cruz-Rodríguez cruzrodriguezcristian@gmail.com

19 **Abstract**

20 Understanding population change through time is crucial for effective conservation of biodiversity.
21 The Living Planet Index (LPI) is a key indicator for tracking global species abundance trends under
22 the Kunming-Montreal Global Biodiversity Framework, offering a picture of population change

23 over time. However, the sensitivity of the index to zero values or to the number of missing values
24 in time series has not been fully explored. Using simulations, we assessed how missing data and
25 zero values influence the index and associated trends. We found that the LPI method is informative
26 with complete datasets, as missing data raises variation but only results in minor deviations from
27 baseline trends. In contrast, the presence and distribution of zeros within the underlying data
28 substantially influence the results and produce significantly lower trends. These findings highlight
29 the need to develop complementary approaches for evaluating how data heterogeneity influences
30 LPI trends, towards a comprehensive understanding and better-informed decisions on biodiversity
31 change.

32 **1. Introduction**

33 The Living Planet Index (LPI) is considered a measure of the state of the world's biodiversity
34 (Ledger et al., 2023). It is based on combined population trends across vertebrate species (Collen
35 et al., 2009) and has become a leading worldwide indicator to discuss global tendencies in nature
36 towards international biodiversity goals ([CBD/COP/DEC/15/4](#), [CBD/COP/DEC/15/5](#)). The LPI
37 uses the largest repository of temporal vertebrate population abundance data, the *Living Planet*
38 *Database* (LPD). This database contains vertebrate population datasets extracted from scientific
39 papers, online databases, government sources, and expert reports (McRae et al., 2025, 2024; Collen
40 et al., 2009). Despite thorough processing, the data remains heterogeneous because the global
41 population time series vary greatly in both temporal resolution and duration (Deinet et al., 2024;
42 Toszogyova et al., 2024). Additionally, many series contain missing data, zeros, and irregular or
43 sparse sampling intervals (Deinet et al., 2024).

44 This data heterogeneity is a challenge for the LPI. The index should handle zeros effectively, as
45 their presence could reflect a true population loss or just temporary detection problems, especially

46 when a single-species population is zero (McRae et al., 2025; Korner-Nievergelt et al., 2022). It
47 also needs to be resilient against varying levels of available data, ensuring that missing data, zeros,
48 and irregular or sparse sampling intervals do not create a trend that does not reflect the state of
49 global biodiversity (Collen et al., 2009). Therefore, it is essential for robust interpretation of LPI
50 trends to identify and comprehend how these sources of variation could affect it.

51 Opinions vary regarding how the LPI responds to these variations. (McRae et al., 2025) argue that
52 using a comprehensive global dataset of species population time series reduces bias caused by
53 missing data. However, (Toszogyova et al., 2024; Troudet et al., 2017) show that uneven
54 availability of population data tends to focus on specific taxa and regions, resulting in substantial
55 taxonomic and spatial biases. Conversely, (McRae et al., 2017) suggest that applying weights that
56 balance groups (e.g., birds versus reptiles) and regions (e.g., North America versus Africa), across
57 various system configurations, helps provide a more uniform representation of global biodiversity,
58 and (Currie et al., 2026) contend that explicit testing of weighting and other treatments are critical
59 for the transparency of the methods used in the LPI to monitor long-term biodiversity trends.
60 Nevertheless, Buschke et al. (2021) show that the index is highly sensitive to fluctuations in
61 population data. They demonstrate that even with stable values on average, the LPI can show a
62 significant decline due to its calculation method. This indicates a potential bias toward reflecting
63 losses, which may lead to an overestimation of biodiversity loss (Buschke et al., 2021).

64 The potential impacts of the data heterogeneity on the robustness of the LPI have been evaluated
65 through the exploration of different methodological choices for index calculation. (Currie et al.,
66 2026) used the Canadian population data to evaluate how different treatments of zeros, changes in
67 confidence intervals and the uncertainty of the index, the effect of the time series length and the
68 number of data points required, different weighting, and the impact of the baseline on the results.

69 (Dove et al., 2023) developed a model of trend reliability using simulated datasets as stand-ins for
70 indicator datasets and a distance measure to quantify reliability by comparing partially sampled to
71 fully sampled trends. (Marconi et al., 2021) used Canadian population data to assess the taxonomic
72 and geographic coverage, as well as the quality of the data in the index. (Collen et al., 2009)
73 investigated how the use of generalized additive models provides information that enables us to
74 evaluate potential bias in the existing data and to explore the feasibility of data disaggregation
75 based on taxonomic and geographic variations. However, without a baseline that includes
76 complete observations, non-zero values, and regular, dense sampling intervals, the effects of these
77 evaluations can't be compared to determine which approach offers the best way to reduce the
78 potential bias from data heterogeneity.

79 Simulated data provide an opportunity to create a dataset that can be compared with the
80 heterogeneity of the real data. The use of synthetic data provides a high level of control, offering
81 the best-case scenario to assess whether correct estimates in an ecological system can be inferred
82 from real data (Zurell et al., 2010). It also provides the opportunity to evaluate and compare
83 different properties of our estimators and to evaluate their behaviour and performance (DiRenzo
84 et al., 2023; Lotterhos et al., 2022). In this way, the simulation of stable populations and combining
85 them with empirical data permits us to estimate contributions of missing data or zeros in the LPI
86 result. Here, we use simulated populations for which we can establish the actual trend to jointly
87 evaluate the impact of method and data structure on the sensitivity and robustness of the LPI. We
88 examine whether data removal, presence of zeros, and trend variation influence the simulated LPI
89 trend. Additionally, we analyze the public LPD to see how its current structure of missing data and
90 zeros affects the simulated LPI. Our goal is to assess how the method and data structure reacts to
91 these disturbances.

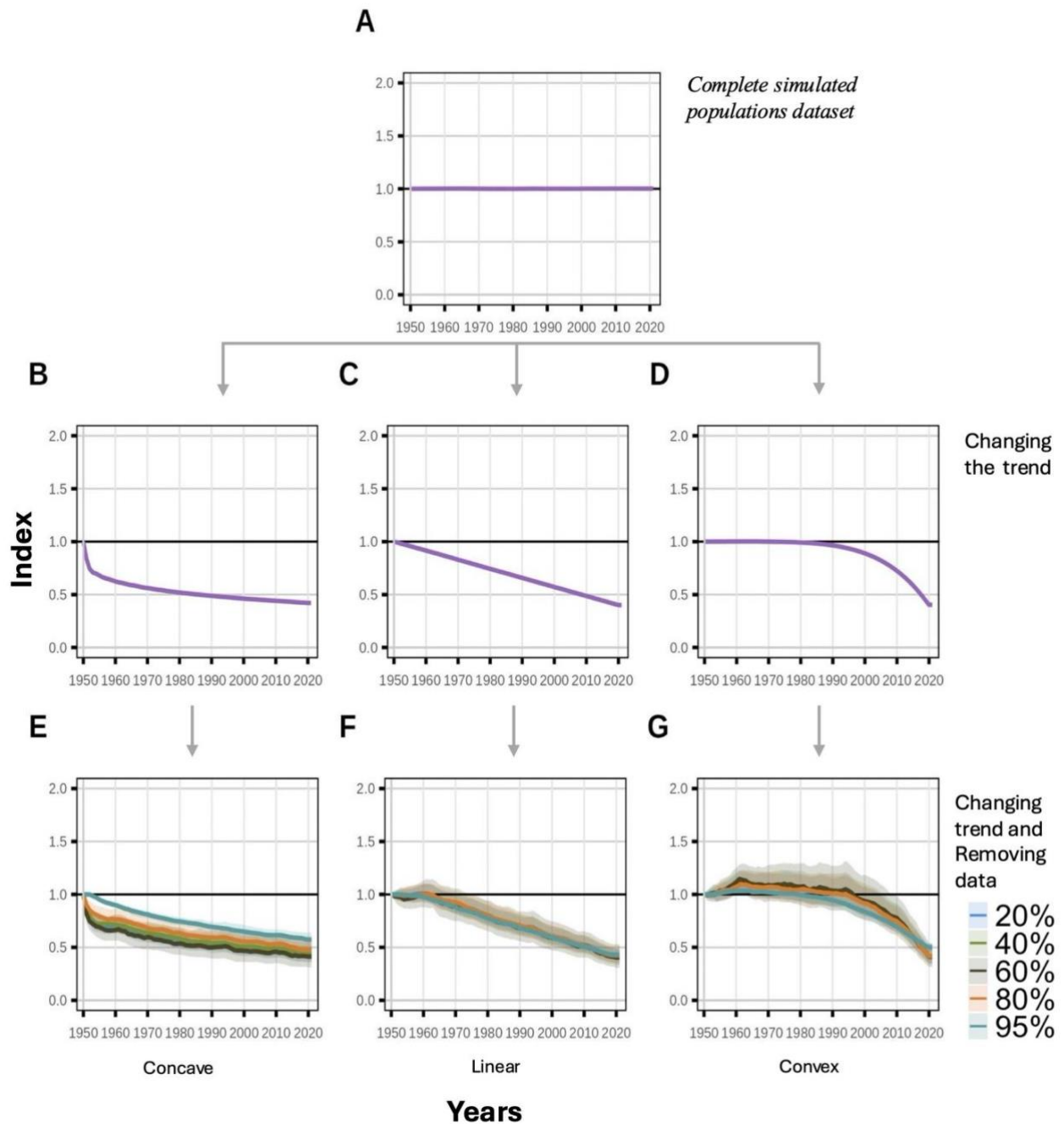
93 2. Results

94 *The LPI detects biodiversity trends on stable simulated data.*

95 When the LPI method is applied to simulated population data, it captures population tendency. A
96 similar pattern was also observed despite the intentionally introduced variations. The index starts
97 at 1 in the first year and then fluctuates, either increasing or decreasing relative to this baseline
98 (McRae et al., 2017). Using a *complete simulated population dataset*, the values across time were
99 close to one (1.0), a number which shows little variation (Fig 1A) and then stability in the
100 populations used. Also, the minimum and maximum coefficient variations are almost identical to
101 the index values, generating a flat line. (Fig 1A). Similarly, inclusion of the three decreasing trend
102 types (concave, convex and linear) produced a consistent tendency line that reflected the three
103 types of trends included, following the pattern of the input data without artificial fluctuations (Fig
104 1B-D). Next, we found that the absence of data (i.e., missing values in the temporal sequences)
105 increased trend variability and could modify the magnitude, but did not alter the direction of the
106 overall trend.

107 For all the trends analyzed, removing the data produced the expected trend, although there was a
108 noticeable increase in variability due to the gaps created by the missing data (Fig. 1E–G). The
109 20%, 40%, 60%, and 80% removal scenarios capture the same population trend, highlighting the
110 robustness to data removal, as only the most extreme 95% removal scenario showed a distinct
111 tendency (Fig. 1E–G). The greatest variability was seen in the convex trend (Fig. 1F), where the
112 tendency showed some increases and decreases over different years, but overall, during the
113 analyzed period, the trend remained convex. By contrast, the linear decreasing trend showed a

114 consistently linear decline (Fig. 1F), and the concave trend exhibited only minor changes (Fig.
 115 1G).



116 **Figure 1.** Resulting population trends using simulations for different scenarios used in the study.
 117 A) Results obtained by applying the LPI to the *simulated populations dataset* while accounting for
 118 sampling error; Results obtained after applying B) concave, C) linear, and D) convex trends using
 119 the full simulated dataset; and E–G) Median values obtained randomly removing 20%, 40%, 60%,

120 80%, and 95% of the simulated data under E) a concave, F) a linear, and G) a convex trend after
121 50 interactions. The values used in the E, F and G obey the median obtained after all of the
122 iterations.

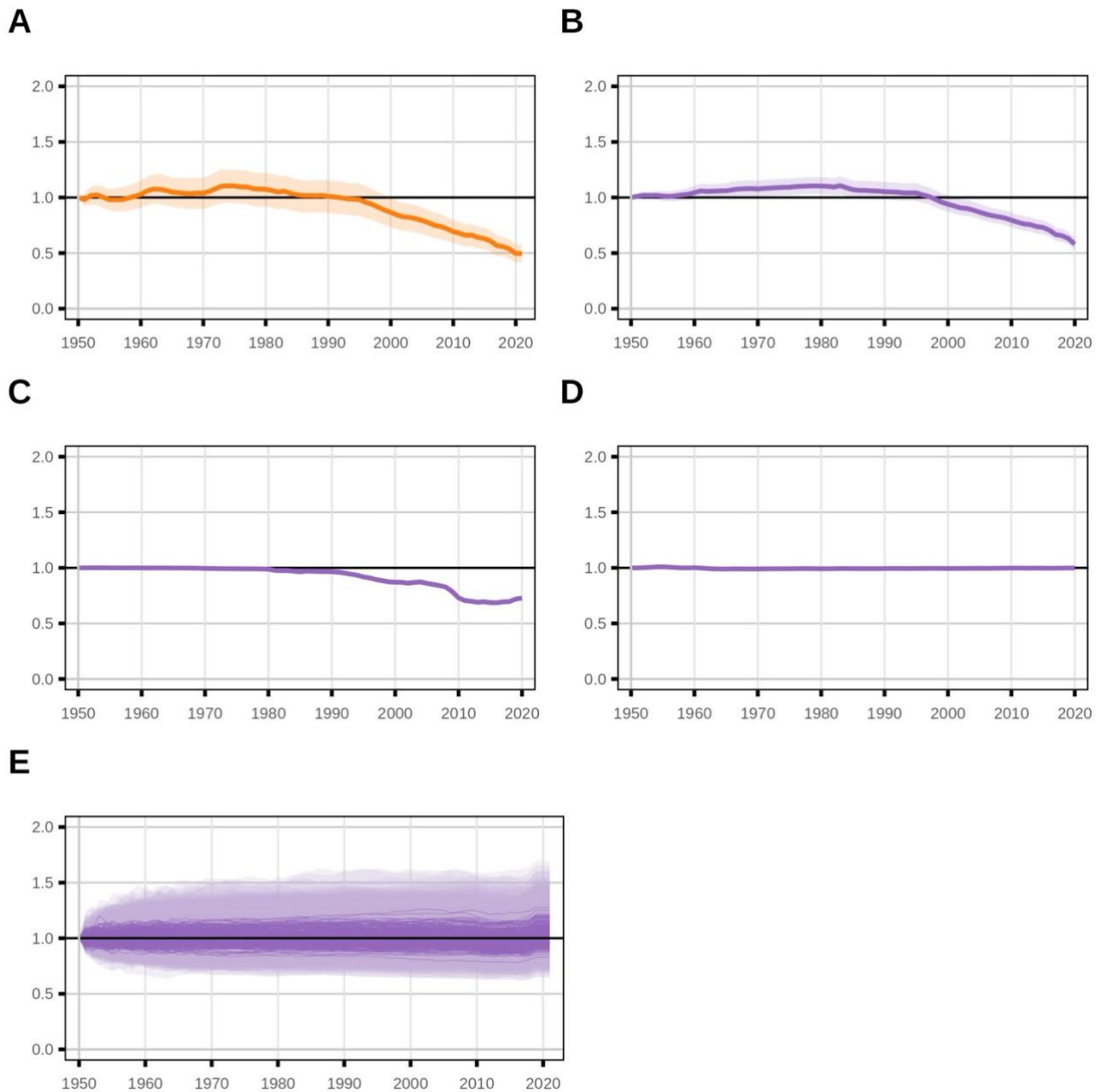
123 *The LPI shows a consistent decrease when using the Living Planet Database.*

124 Calculating the LPI using the *Living Planet Database* and without applying diversity weighting
125 (all population time series contributing equally to the index, regardless of taxonomic, realm or
126 geographic region) and represent all populations as distinct species—to ensure that no populations
127 are aggregated or averaged into a single species— projects a decrease in population abundances.
128 The index shows fluctuations that dipped and rose by very small amounts, remaining around 1.0,
129 rising slightly between 1960 and 1985 (Fig 2A). Afterward, the index experiences a steady decline.
130 The coefficient of variation is unstable before 1960, increasing between 1950 and 1962, but then
131 remains stable afterward (Fig 2A). These results were verified both by retaining the duplicate
132 records and by removing them, and the results remained the same, which demonstrates the
133 robustness of the workflow used

134 *The LPI produces a trend without tendencies despite missing data, but is affected by both the
135 presence and position of zeros.*

136 We observed that the addition of zeros in the *complete simulated population dataset* distorts the
137 tendency generated by the LPI, but that the tendencies are unaffected by missing data alone (Fig.
138 2). When we include in the *complete simulated population dataset* both the same zeros and missing
139 data present in the *LPD* (Fig. 2B), the resulting trend was different from the trend observed with
140 our stable simulations (Fig. 1A) and to the empirical data (Fig. 2A). The tendency obtained (Fig.
141 2B) showed an increase from 1960, reached its highest point in 1984, then declined until 2016,
142 when it began to rise again. Next, attempting to separate the effects of zeros and missing data, we

143 observed that removing the data at the same position where it is missing in the *LPD* yielded results
144 identical to the trend observed when we used the *complete simulated population dataset* (Fig 1A
145 and Fig 2D). However, only adding zeros to the dataset at the same positions in the *LPD* caused a
146 decline in the trend. Under this scenario, the trend initially matched that of the completed simulated
147 dataset (Fig. 2C) but diverged later in the series. After 1980, the trend began to drop; there was a
148 more noticeable decline in 1990, it remained relatively stable in 2000, then sharply fell in 2005.
149 Interestingly, after 2011, the trend reversed and started increasing, peaking in 2020 at a level
150 similar to 2005 (Fig. 2C). Finally, when we relocated the zeros in the *empirical population data*
151 with the same annual proportions there are in the *LPD* but across 300 different permutations (while
152 maintaining missing data at the same position), the resulting trend became unpredictable, and the
153 variation was not constant (Fig. 2E).



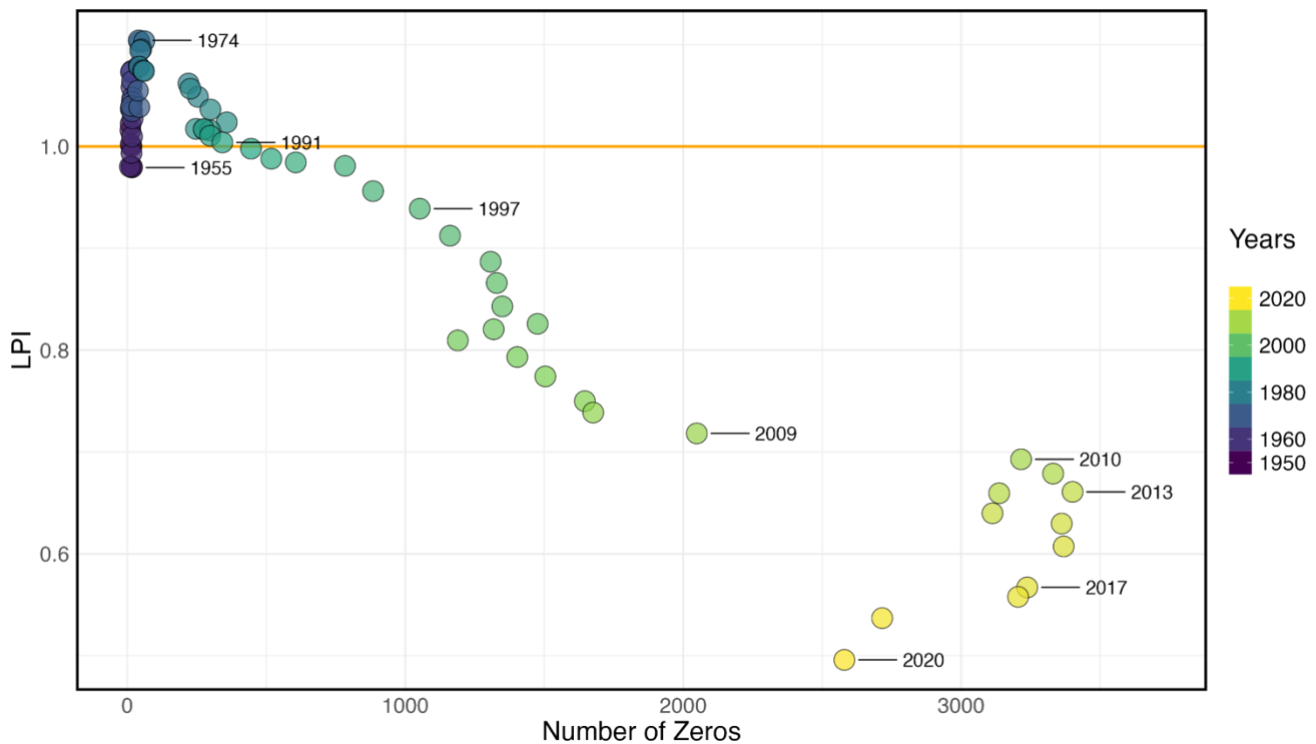
154 **Figure 2.** Implications of missing data and the presence of zeros in the *LPD* were assessed using
 155 a complete simulated population dataset with four approaches. A) Result to calculate the LPI using
 156 the *LPD* without including weights and representing all populations as distinct species. B)
 157 Complete simulated dataset with data removal at the same locations as missing values in the *LPD*,
 158 and including zeros at the same positions where they appear sin the *LPD* (see Methods for more

159 information). C) Zeros were added to the *complete simulated population dataset* at the same
160 positions where they appear in the *LPD*. D) Data were removed from the full simulated dataset at
161 the same positions that are missing in the *LPD*. E) The same number of zeros was added to the
162 *LPD*, but they were randomly distributed across 300 different permutations. Note that panels B
163 and C also have confidence intervals, as in the other figures, but all simulation results span a
164 narrower range, making the intervals less visible.

165 *The presence of zeros is a crucial structural feature of long-term time series.*

166 Comparing the number of zeros with the LPI trend reveals that data continuity likely influences
167 the observed results. As shown in Figure 3 and Figure S1, the index was most stable (around 1.0)
168 during the earliest period of documentation, which was characterized by a consistently low number
169 of zeros. After achieving a maximum value index in 1974 (Fig. 3), the LPI began a progressive
170 decline, remaining just below 1.0 until 1991, but with a progressive increase in the number of zeros
171 (Fig. 3, Fig. S1).

172 A prominent increase in the number of zeros started around 1980 (Fig. S1), surpassing 1,000
173 records by 1997 (Fig. 3). This number increased considerably during the subsequent decade with
174 an increase of 1,000 zeros by 2009, reflecting significant increase in its number in this period of
175 time (Fig. 3). Despite fluctuations between 2010 and 2018—a period that included the peak of
176 3,401 zeros in 2013—the overall trend shows that the index continued to decline (Fig. S1), even
177 as the total number of zeros decreased between 2017 and 2020 (Fig. 3). These observed patterns
178 were compared with the relationship between LPI values and the number of zeros through the 300
179 permutations. The results showed that when the number of zeros was below 1,500, the observed
180 values were similar to the simulations. However, once the count surpassed 1,500 numbers of zeros,
181 the results differed between the *LPD* data source and the simulations (Fig. S2).



182 **Figure 3.** Relationship between the number of zeros and LPI values between 1950 and 2020 using
 183 the LPD. Notice how the regularity of the number of zeros was constant before 2008, with the
 184 spread of values increasing as the years went by. In 2009, the frequency of zeros ceased to be
 185 regular, with variations in their number occurring over very short periods of time. Annotated years
 186 are shown for narrative purposes to highlight the main temporal trend.

187 3. Discussion

188 Using a *complete simulated population dataset*, representing different growth rates and a stable
 189 index trend over time, this study demonstrates that the LPI method generally preserves the
 190 underlying tendency despite missing data, although it increases its variability. Only when the
 191 proportion of missing data exceeded 95% (an extreme scenario) did the removal of data lead to a
 192 different tendency, underestimating the declining trend from the data. However, our analysis
 193 indicates that zeros in the *Living Planet Database (LPD)* —representing real absences in some
 194 populations— impact the results and influence trend outcomes based on where they occur.

195 Notably, including zeros from where they are found in the LPD in the simulated (stable)
196 populations resulted in discrepancies from expected trends, and were sufficient to create
197 decreasing trends in an otherwise stable population due to their temporal position within
198 population time series. Meanwhile, missing data on the LPD caused only minor variations when it
199 was included in the *complete simulated population dataset*. Overall, these findings underscore both
200 the importance of testing indicators with a baseline that can help identify limitations in the data
201 used for their calculation. Our results also call for increased effort in identifying the cause and
202 temporal position of zero values in time series that support the LPI calculation.

203 Using a *complete simulated population dataset* as a reference, our study reveals that the LPI
204 method is highly sensitive to the number and temporal positions of zeros present in the *LPD* across
205 all species considered. These zeros —when occurring in their observed temporal positions— have
206 a substantial impact, both qualitative and quantitative, on the final trend, and are sufficient to
207 explain the tendency observed in the LPD (shown in Fig. 2A). Similar to Toszogyova et al. (2024),
208 our analyses show that zeros can lead to significant drops or spikes in the index, when they are
209 temporally structured within population time series. Two factors likely explain this outcome:
210 firstly, the geometric mean, the method used to aggregate the LPI and similar multi-species indices,
211 is sensitive to the presence of zero values which presents a challenge to capturing change in rare
212 species or those hard to detect (Korner-Nievergelt et al 2022). Secondly, the zero values present
213 in the LPD are more prevalent in the second half of the time-series, with the increase in zero values
214 present corresponding with the decline in the simulated LPI (Fig. 2B and Fig. 2C).

215 To address such effect of zeros in a multi-species index, several alternatives have been proposed,
216 including treating zeros as missing (i.e. NA) (Currie et al., 2026; Marconi et al., 2021; Toszogyova
217 et al., 2024), adding a small constant such as 1% of the mean (Burns et al., 2018; Collen et al.,

218 2009), the minimum value, the number 1, or an extremely small value (e.g., 0.000001), as well as
219 conditional replacements where leading, middle, and trailing zeros are substituted with
220 combinations of NA and 1% of the mean (Currie et al., 2026; Collen et al., 2009). In one study,
221 these modifications result in different trends, causing a variation in the final outcome as well as in
222 the confidence intervals (Currie et al., 2026).

223 This issue is particularly challenging because, in ecological datasets, the presence of zeros can
224 originate from different processes. Zeros may represent true biological absences, but they may also
225 arise from imperfect detectability or sampling limitations, making it difficult to distinguish
226 ecological signal from observational noise. Multiple trailing zeros are more likely to represent a
227 local extinction, an important signal for a biodiversity indicator. Therefore, correctly classifying
228 zeros is crucial to avoid false zeros that could hinder real patterns or lead to misleading conclusions
229 about population change, a concern that has been analyzed in the literature (Martin et al., 2005).
230 Wenger & Freeman (2008), through practical examples, have investigated the explicit effect of
231 zeros, and demonstrate how the adoption of methods that explicitly model the sources of zero
232 observations improves the inference and strengthens robustness of ecological analyses. The *LPD*
233 and resulting LPI, however, does not currently incorporate this approach. Instead, the zeros are
234 transformed by adding 1% of the mean of the time-series to all values, in order to calculate rates
235 of change for the index on a log scale (Collen et al., 2009). As a result, the influence of zero values
236 on index trends appears to be driven more by the properties of the transformation than by the data
237 type itself, an important consideration when interpreting apparent declines in the LPI.

238 Similar to zero values, missing values have been identified as a potential source of bias in the LPI
239 due to the way they are incorporated into index calculations (Toszogyova et al., 2024). In practice,
240 the LPI accounts for missing values through methodological transformations designed to

241 standardize trajectories over time. Therefore, if a species is not observed in a particular year, values
242 are interpolated using different modelling approaches but only within the bounds of the first and
243 last year of the time-series (Collen et al., 2009). Our analysis shows that removing some of the
244 simulated data causes variations (Fig. 1D-E), but the overall trend stays consistent with the
245 baseline. The only exception was when we removed 95% of the data, at which point the LPI
246 maintained but started to underestimate the declining trend. Additionally, missing data in the *LPD*
247 does not alter the overall trend shape once added to the simulated stable populations (Fig. 2),
248 indicating a limited impact on the overall outcomes alone. However, our results suggest that
249 missing data do alter the impact that zeros have on the index (Fig. 2B), as in our example, the
250 index trend shifts to a more positive, then a more negative trend compared to the stable baseline.
251 Previous works have shown that random fluctuations naturally distort LPI trends, often biasing
252 toward decline, and should be evaluated against null-model baselines (Buschke et al., 2021).
253 Although these random fluctuations were not explicitly included in our analysis, the persistence of
254 minor variations when we analyze the effect of missing data supports that the index can maintain
255 the tendency despite missing data.

256 Addressing these challenges arising from zeros and missing data requires models with clearly
257 defined baselines to evaluate how the index responds to heterogeneous data and to ensure that
258 observed patterns reflect genuine population trends rather than methodological artifacts. Zeros
259 play a critical role in calculating the Living Planet Index (LPI) because their position can
260 substantially influence estimated trends, even when population time series are complete and show
261 no apparent declines. It is therefore essential to determine whether zeros represent true
262 observations or arise from missed observations before including them in the index. Although the
263 LPI can be informative when data are sparse, limited or discontinuous time series reduce accuracy

264 and increase trend variability, highlighting the need for more continuous monitoring to improve
265 robustness. Validation using both simulated and well-characterized ecological datasets is
266 particularly important, as it can help identify sources of variation, reduce arbitrary treatments of
267 zeros, and reveal data gaps that affect trend detection. While both the observed temporal trend and
268 the final value of the index may reflect the influence of zeros or near-zero values, the LPI still
269 indicates a downward trajectory after zeros are removed and replaced by missing data. We have
270 demonstrated the value of using simulated data to better understand how the LPI responds to zero
271 values and missing data, both of which are pervasive features of population timeseries data, and
272 more likely to increase as the timespan and the taxonomic diversity of the dataset to which the LPI
273 is applied increase.

274 Long-term research is a cornerstone of ecological studies (Kuebbing et al., 2018; Lindenmayer et
275 al., 2012), as it can capture long-term changes in biodiversity that sporadic collections of samples
276 cannot. Nevertheless, as we saw with the effect of the zeros, continuing to accumulate data while
277 also accumulating zeros can drive the population trends down; counter-intuitively, increased
278 monitoring efforts bias our understanding of actual population trends. This suggests a need for
279 improving the long-term research that captures and analyzes data (Festa-Bianchet et al., 2017),
280 primarily by more adequately documenting the process that leads to a population size being
281 registered as a zero. Introducing regular monitoring programs using structured protocols can help
282 mitigate gaps in our data streams, ensuring a more comprehensive understanding of the systems
283 we're monitoring (Lindenmayer et al., 2012). By adopting a long-term perspective that balances
284 the need for frequent updates with the importance of uninterrupted data collection, we can better
285 appreciate the significance of zeros as a structural feature of time series data.

286 We recommend determining the nature of the zero values in the empirical LPI to inform
287 appropriate treatments. Because the trend reported by the LPI responds, in addition to genuine
288 changes in species abundance, to the temporal position of zeros in the time series, we encourage
289 users of the method to carefully distinguish between actual zeros (the population is reliably
290 absent/locally extinct at this time point) from artifactual zeros (used in place of missing data).
291 Further research is needed to assess how the diversity of species or populations included affects
292 index performance, particularly to determine minimum data requirements across spatial scales,
293 biogeographic realms, time series coverage, and taxonomic groups.

294 **5. Methods**

295 We used a population-simulation matrix to assess the Living Planet Index's sensitivity to the
296 presence and position of zeros and to missing data (NA's). The matrix represents a dataset showing
297 the hypothetical number of individuals in a not-real population across different generations. For
298 methodological purposes, we consider each generation as 1 year. Then, we applied controlled
299 perturbations to the data quality to see if the LPI maintained the initial trend. Deviations in
300 indicator outcomes from the population simulations were used to assess robustness. The complete
301 workflow can be observed in Figure 3.

302 *Living Planet Index calculation using the Living Planet Database*

303 We used the latest available public version (www.livingplanetindex.org/data_portal) of the *Living*
304 *Planet Database (LPD)*, the largest repository of vertebrate population abundance data, to create
305 a trend based on genuine data. This public version contains data from 35,996 populations of 5,177
306 species, sourced from scientific papers, online databases, government, and expert reports (Ledger,
307 2023). As our aim was to evaluate the LPI using baseline simulated data, we included the full
308 number of years reported in the *LPD*. We followed the published LPI method for calculating an

309 index (Collen et al., 2009). With comparative purposes, we represent all populations as distinct
310 species, to ensure that no populations are aggregated or averaged into a single species. Also, as
311 our goal wasn't to study the impact of data availability, we didn't include the weighting that is used
312 for the global LPI, which assigns weights to populations based on the regions, realms, and taxa
313 where they were monitored (McRae et al., 2017).

314 *Simulation overview*

315 We used simulated population data to assess the sensitivity of the Living Planet Index (LPI)
316 method and data structure. The matrix shows a *complete simulated population dataset* which
317 reflects a stable population, indicating that the simulated populations remain consistent across
318 generations. Using the *complete simulated population dataset*, we applied controlled perturbations
319 to the data quality to see if the LPI maintained the similar tendency. Deviations in indicator
320 outcomes from the population simulation were used to assess robustness. The complete workflow
321 can be observed in Figure 3.

322 *Population Simulated Data:* To simulate each population, we used the discrete logistic growth
323 equation proposed by Richel in 1954 (Equation 1). It models density-dependent regulation in
324 populations with discrete generations by linking the population size at one time step to the next,
325 based on the intrinsic growth rate and carrying capacity (Ricker, 1954).

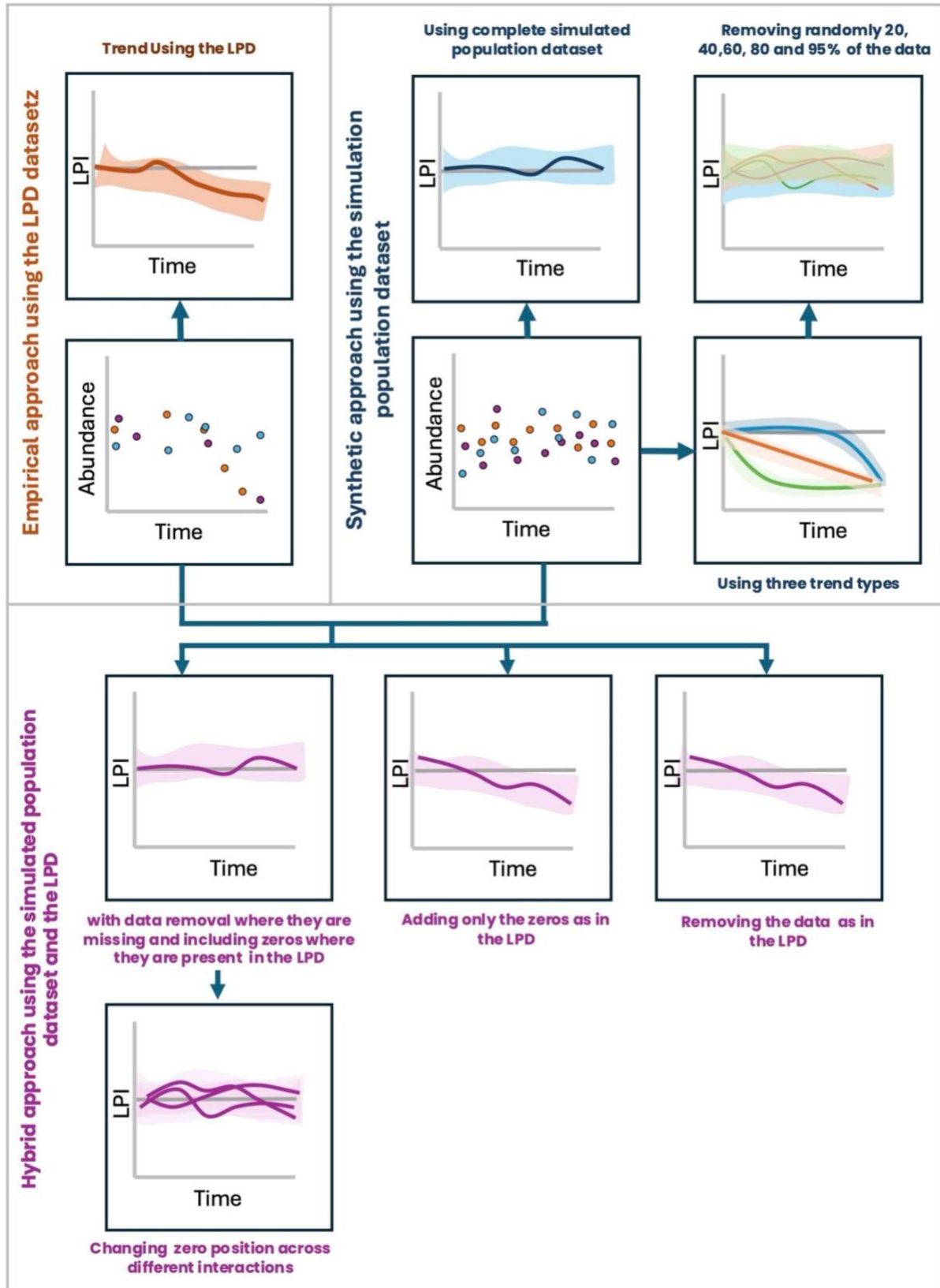
$$326 \quad N_{t+1} = N_t \cdot e^{r \cdot (1 - \frac{N_t}{K})} \quad (1)$$

327 Where N_t represents the population size at time t ; r the intrinsic rate of natural increase; K is the
328 carrying capacity, representing the maximum sustainable population size the environment can
329 support. Finally, e it is the base of the natural logarithm (≈ 2.718). Simulations were run on 50,000
330 independent hypothetical populations over 500 generations.

331 The initial population sizes simulated were randomly selected, ranging from a dozen to thousands,
332 representing various population sizes. The number of generations was chosen to ensure enough
333 time for populations to reach equilibrium or show stochastic fluctuations. The intrinsic growth rate
334 values ranged (r) from 0.05 to 0.45, reflecting the minimum and maximum growth rates
335 documented in the literature for vertebrates (Pardo & Dulvy, 2022; Dillingham et al., 2016). To
336 set the carrying capacity (K), we used the minimum value based on the Elk population's ($\sim <10$ per
337 herd in equilibrium) (Koetke et al., 2020), and the maximum estimate derived from Salmon
338 ($\sim 71,844$ per delta) (Hall et al., 2023). To account for environmental differences among
339 populations, these parameters were varied across replicate populations, allowing each to have its
340 own distinct growth rates and carrying capacities (Dennis & Ponciano, 2014; Gotelli & Gotelli,
341 2008). To obtain a *complete simulated population dataset* comparable with the *empirical*
342 *population data*, we randomly select the number of populations and generations of our simulated
343 dataset, matching the LPD dataset's structure in terms of population units (rows) and years
344 (columns).

345 *Scenarios*

346 *Observed error:* To evaluate the impact of measurement uncertainty on observed population
347 variability, an extra layer of randomness was added to the simulated data to simulate sampling
348 error (Dennis et al., 2006). This was achieved by randomly modifying population values to account
349 for detection or counting errors, with a 0.1 probability to introduce more sampling variation. This
350 was done by randomly adjusting population values to emulate inaccuracies or imperfect detection
351 in our simulated population data. Incorporating observation error allows evaluation of how
352 sampling variability can alter perceived population trends and indices, even when underlying
353 dynamics remain stable.



355 **Figure 4.** Workflow of the approach used in this study. The empirical approach refers to the
356 calculation of the trend using the data available at the *LPD*. The synthetic approach refers to the
357 trends generated using the *complete simulated population dataset*. The results include the types of
358 trends and also remove 20%, 40%, 60%, 80%, and 95% of the dataset at random. Finally, the
359 hybrid approach represents the use of the *complete simulated population dataset*, but with data
360 removed as missing according to the structure in the *LPD*, adding zeros and changing their
361 position across interactions per row.

362 *Modifying the trend:* We generated three hypothetical population decline trajectories —concave,
363 linear, and convex— and subsequently applied them to the *complete simulated population dataset*
364 to reflect the diversity of trend types observed in nature (fig. 4). These trajectories were used
365 because even random fluctuations and non-linear trends (such as concave or convex shapes) can
366 influence the behaviour of overall biodiversity indices (Buschke et al., 2021). This approach
367 allowed us to test if changes in population dynamics affect the stability and interpretation of the
368 simulated results.

369 *Missing data:* Missing observations (NA values) in long-term population data can affect trend
370 detection and create bias interference (Onkelinx et al., 2017), as well as reduce the statistical power
371 and hamper inference (Łopucki et al., 2022). For that reason, we randomly removed 20%, 40%,
372 60%, 80%, and 95% of the data points for the whole dataset (fig. 4). This allowed us to examine
373 how data loss might bias the tendency captured by the LPI. We conducted these analyses using the
374 same trend as in the last step (Buschke et al., 2021) to assess whether combining different trends
375 with missing data could lead to unforeseen scenarios that the LPI couldn't detect. Finally, we
376 conducted 50 iterations for each combination to evaluate the robustness of the results with respect

377 to the locations of the data points removed. We reported the median, minimum, and maximum
378 index values from all iterations (fig. 4).

379 *Incorporating null and zero values in the simulated data:* We utilized the *LDP* structure and
380 incorporated it into our *complete simulated population dataset* to examine whether the structure
381 of the empirical dataset possessed particular features that affect the trend (McRae et al., 2025,
382 2017). First, we eliminated data points at identical time–series positions that were absent in the
383 empirical dataset per row, ensuring the simulated gaps mirrored the real-world missingness
384 pattern. Second, per row, we replaced values at those same positions with zeros to simulate cases
385 where monitoring records are true zero observations or non-detections (fig. 4).

386 After we evaluated how the distribution of missing or zero values could affect trend detection, we
387 performed 300 permutations as follows. First, we replace the zeros and remove the data from the
388 same cells that are in the LPD in the iterations in the *complete simulated population dataset*. After,
389 we shuffled the position of the zeros randomly within the time period covered by each time series,
390 but we maintained fixed cells with missing data in the matrix with the *complete simulated*
391 *population dataset*. With the resulting matrix, we applied the LPI method to calculate an overall
392 trend. We repeated this process 300 times. In each iteration, the zeros were shuffled per row.

393 All data analysis was performed in R (version 4.3, (R Core Team, 2023), using the main packages:
394 *data.table* (v1.16.2, (Barrett et al., 2024) , *future.apply* (v1.11.3, (Bengtsson, 2021), and *future*
395 (v1.40.0, (Bengtsson, 2021) for efficient data handling and parallel computation; *ggplot2* (v3.5.2,
396 (Wickham, 2016), *patchwork* (v1.2.0, (Pedersen, 2024), and *RColorBrewer* (v1.1-3, (Neuwirth,
397 2022) for data visualization; and *missMethods* (v0.4.0, (Rockel, 2022) and *tidyverse* (v2.0.0,
398 (Wickham et al., 2019) for data preprocessing. The Living Planet Index was calculated using the
399 *rlpi* package (v0.1.0, (Freeman et al., 2024).

400 **5. Data availability**

401 The study relies on the *Living Planet Database (LPD)*, which is publicly available through the
402 official Living Planet Index data portal: https://www.livingplanetindex.org/data_portal. All
403 datasets used in the analyses can be accessed and downloaded from this resource.

404 **6. Code availability**

405 All code used for the population simulations and the analytical routines described in this study is
406 openly available on GitHub at the following repository:
407 https://github.com/crcruZR/LPI_Simulations/. The repository includes scripts, documentation, and
408 instructions to reproduce the results.

409 **7. Acknowledgments**

410 This manuscript results from a working group supported by the Living Data Project (LDP), an
411 initiative of the Canadian Institute of Ecology and Evolution funded through a Collaborative
412 Research and Training Experience (CREATE) grant from the Natural Sciences and Engineering
413 Research Council of Canada (NSERC). We also thank the reviewers for their valuable feedback
414 during the development of this document. We acknowledge that this study was partially conducted
415 on land within the traditional unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki,
416 Mohawk, Huron-Wendat and Omàmiwininiwak nations.

417 **8. Author contributions**

418 C.C.R., G.M., J.S., S.L J.C., V.M., R.F., L.M., M.J., S.R., S.E. and D.A.H. conceived the study.
419 C.C.R., G.M., T.P. and D.A.H. developed the conceptual framework. C.C.R. designed and
420 implemented the population simulations and performed the analyses. S.L. and R.F. contributed to
421 model validation and checks. C.C.R. and T.P. prepared figures and visualizations. J.C. V.M., S.E.,

422 and D.A.H. acquired funding. C.C.R. wrote the first draft of the manuscript. All authors
423 contributed to manuscript revision and approved the final submitted version

424 **9. Competing interests**

425 The authors' institutions did not affect the study design, data analysis, interpretation of results, or
426 the decision to publish. All authors declare no competing interests.

427 **10. References**

428 Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Hocking, T., & Schwendinger,

429 B. (2024). *data.table: Extension of `data.frame`*. [https://CRAN.R-](https://CRAN.R-project.org/package=data.table)

430 [project.org/package=data.table](https://CRAN.R-project.org/package=data.table)

431 Bengtsson, H. (2021). A Unifying Framework for Parallel and Distributed Processing in R using

432 Futures. *The R Journal*, *13*(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>

433 Burns, F., Eaton, M. A., Hayhow, D. B., Outhwaite, C. L., Al Fulaij, N., August, T. A., Boughey,

434 K. L., Brereton, T., Brown, A., Bullock, D. J., Gent, T., Haysom, K. A., Isaac, N. J. B.,

435 Johns, D. G., Macadam, C. R., Mathews, F., Noble, D. G., Powney, G. D., Sims, D. W.,

436 ... Gregory, R. D. (2018). An assessment of the state of nature in the United Kingdom: A

437 review of findings, methods and impact. *Ecological Indicators*, *94*, 226–236.

438 <https://doi.org/10.1016/j.ecolind.2018.06.033>

439 Buschke, F. T., Hagan, J. G., Santini, L., & Coetzee, B. W. T. (2021). Random population

440 fluctuations bias the Living Planet Index. *Nature Ecology & Evolution*, *5*(8), 1145–1152.

441 <https://doi.org/10.1038/s41559-021-01494-0>

442 Collen, B., Loh, J., Whitmee, S., McRae, L., Amin, R., & Baillie, J. E. M. (2009). Monitoring

443 change in vertebrate abundance: The living planet index. *Conservation Biology: The*

444 *Journal of the Society for Conservation Biology*, 23(2), 317–327.
445 <https://doi.org/10.1111/j.1523-1739.2008.01117.x>

446 Currie, J., Ravoth, S. M., Marconi, V., McRae, L., Arce-Plata, M. I., Emry, S., Freeman, R.,
447 Jousse, M., Mével, G., Li, S., Cruz-Rodríguez, C. A., Hunt, D. A. G. A., Oppenheimer,
448 P., Gill, L., Serrano, J., & Deinet, S. (2026). Navigating methodological decisions:
449 Balancing rigor and data volume of the Canadian Living Planet Index. *Facets*, 11, 1–14.
450 <https://doi.org/10.1139/facets-2025-0112>

451 Deinet, S., Marconi, V., Freeman, R., Puleston, H., & McRae, L. (2024). *Living Planet Index*
452 *technical supplement 2024*.
453 [https://www.google.com/url?q=https://doi.org/10.17605/OSF.IO/6TE9H&sa=D&source=](https://www.google.com/url?q=https://doi.org/10.17605/OSF.IO/6TE9H&sa=D&source=docs&ust=1775937926621981&usg=AOvVaw0r8YXVJTCqeMkw73kRj_JN)
454 [docs&ust=1775937926621981&usg=AOvVaw0r8YXVJTCqeMkw73kRj_JN](https://www.google.com/url?q=https://doi.org/10.17605/OSF.IO/6TE9H&sa=D&source=docs&ust=1775937926621981&usg=AOvVaw0r8YXVJTCqeMkw73kRj_JN)

455 Dennis, B., & Ponciano, J. M. (2014). Density-dependent state-space model for population-
456 abundance data with unequal time intervals. *Ecology*, 95(8), 2069–2076.
457 <https://doi.org/10.1890/13-1486.1>

458 Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L., & Staples, D. F. (2006). Estimating
459 Density Dependence, Process Noise, and Observation Error. *Ecological Monographs*,
460 76(3), 323–341. [https://doi.org/10.1890/0012-](https://doi.org/10.1890/0012-9615(2006)76%255B323:EDDPNA%255D2.0.CO;2)
461 [9615\(2006\)76%255B323:EDDPNA%255D2.0.CO;2](https://doi.org/10.1890/0012-9615(2006)76%255B323:EDDPNA%255D2.0.CO;2)

462 Dillingham, P. W., Moore, J. E., Fletcher, D., Cortés, E., Curtis, K. A., James, K. C., & Lewison,
463 R. L. (2016). Improved estimation of intrinsic growth r_{max} for long-lived species:
464 Integrating matrix models and allometry. *Ecological Applications*, 26(1), 322–333.
465 <https://doi.org/10.1890/14-1990>

466 DiRenzo, G. V., Hanks, E., & Miller, D. A. W. (2023). A practical guide to understanding and
467 validating complex models using data simulations. *Methods in Ecology and Evolution*,
468 *14*(1), 203–217. <https://doi.org/10.1111/2041-210X.14030>

469 Dove, S., Böhm, M., Freeman, R., McRae, L., & Murrell, D. (2023). Quantifying reliability and
470 data deficiency in global vertebrate population trends using the Living Planet Index.
471 *Global Change Biology*. <https://doi.org/10.1111/gcb.16841>

472 Festa-Bianchet, M., Douhard, M., Gaillard, J.-M., & Pelletier, F. (2017). Successes and
473 challenges of long-term field studies of marked ungulates. *Journal of Mammalogy*, *98*(3),
474 612–620. <https://doi.org/10.1093/jmammal/gyw227>

475 Freeman, R., McRae, L., Deinet, S., Amin, R., & Collen, B. (2024). *rlpi: Tools for calculating*
476 *indices using the Living Planet Index method*. [https://github.com/Zoological-Society-of-](https://github.com/Zoological-Society-of-London/living_planet_index)
477 [London/living_planet_index](https://github.com/Zoological-Society-of-London/living_planet_index)

478 Gotelli, N. J., & Gotelli, N. J. (2008). *A Primer of Ecology* (Fourth Edition, Fourth Edition).
479 Oxford University Press.

480 Hall, J., Roni, P., Ross, K., Camp, M. J., Nuckols, J., & Ruffing, C. (2023). Estimating Juvenile
481 Salmon Estuarine Carrying Capacities to Support Restoration Planning and Evaluation.
482 *Estuaries and Coasts*, *46*(4), 1046–1066. <https://doi.org/10.1007/s12237-023-01185-y>

483 Koetke, L. J., Duarte, A., & Weckerly, F. W. (2020). Elk population dynamics when carrying
484 capacities vary within and among herds. *Scientific Reports*, *10*(1), 15956.
485 <https://doi.org/10.1038/s41598-020-72843-5>

486 Korner-Nievergelt, F., Strebel, N., Buckland, S. T., Freeman, R., Gregory, R. D., Guélat, J.,
487 Isaac, N. J. B., McRae, L., Roth, T., Schirmer, S., Soldaat, L. L., Voříšek, P., & Sattler,
488 T. (2022). Multi-species population indices for sets of species including rare,

489 disappearing or newly occurring species. *Ecological Indicators*, 140, 109005.
490 <https://doi.org/10.1016/j.ecolind.2022.109005>

491 Ledger, S. E. H. (2023). *Past, present, and future of the Living Planet Index*.
492 <https://doi.org/10.1038/s44185-023-00017-3>

493 Ledger, S. E. H., Loh, J., Almond, R., Böhm, M., Clements, C. F., Currie, J., Deinet, S.,
494 Galewski, T., Grooten, M., Jenkins, M., Marconi, V., Painter, B., Scott-Gatty, K., Young,
495 L., Hoffmann, M., Freeman, R., & McRae, L. (2023). Past, present, and future of the
496 Living Planet Index. *Npj Biodiversity*, 2(1), 12. [https://doi.org/10.1038/s44185-023-](https://doi.org/10.1038/s44185-023-00017-3)
497 [00017-3](https://doi.org/10.1038/s44185-023-00017-3)

498 Lotterhos, K. E., Fitzpatrick, M. C., & Blackmon, H. (2022). Simulation Tests of Methods in
499 Evolution, Ecology, and Systematics: Pitfalls, Progress, and Principles. *Annual Review of*
500 *Ecology, Evolution, and Systematics*, 53(1), 113–136. [https://doi.org/10.1146/annurev-](https://doi.org/10.1146/annurev-ecolsys-102320-093722)
501 [ecolsys-102320-093722](https://doi.org/10.1146/annurev-ecolsys-102320-093722)

502 Marconi, V., McRae, L., Müller, H., Currie, J., Whitmee, S., Gadallah, F. (ZuZu), & Freeman, R.
503 (2021). Population declines among Canadian vertebrates: But data of different quality
504 show diverging trends. *Ecological Indicators*, 130, 108022.
505 <https://doi.org/10.1016/j.ecolind.2021.108022>

506 Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre,
507 A. J., & Possingham, H. P. (2005). Zero tolerance ecology: Improving ecological
508 inference by modelling the source of zero observations. *Ecology Letters*, 8(11), 1235–
509 1246. <https://doi.org/10.1111/j.1461-0248.2005.00826.x>

510 McRae, L., Cornford, R., Marconi, V., Puleston, H., Ledger, S. E. H., Deinet, S., Oppenheimer,
511 P., Hoffmann, M., & Freeman, R. (2025). The utility of the Living Planet Index as a

512 policy tool and for measuring nature recovery. *Philosophical Transactions of the Royal*
513 *Society B: Biological Sciences*, 380(1917), 20230207.
514 <https://doi.org/10.1098/rstb.2023.0207>

515 McRae, L., Deinet, S., & Freeman, R. (2017). The Diversity-Weighted Living Planet Index:
516 Controlling for Taxonomic Bias in a Global Biodiversity Indicator. *PLOS ONE*, 12(1),
517 e0169156. <https://doi.org/10.1371/journal.pone.0169156>

518 McRae, L., Marconi, V., Deinet, S., Puleston, H., Living Planet Index Team, & Freeman, R.
519 (with Center For Open Science). (2024). *Living Planet Index technical supplement 2024*.
520 OSF. <https://osf.io/6te9h/>

521 Neuwirth, E. (2022). *RColorBrewer: ColorBrewer Palettes*. [https://CRAN.R-](https://CRAN.R-project.org/package=RColorBrewer)
522 [project.org/package=RColorBrewer](https://CRAN.R-project.org/package=RColorBrewer)

523 Onkelinx, T., Devos, K., & Quataert, P. (2017). Working with population totals in the presence
524 of missing data comparing imputation methods in terms of bias and precision. *Journal of*
525 *Ornithology*, 158(2), 603–615. <https://doi.org/10.1007/s10336-016-1404-9>

526 Pardo, S. A., & Dulvy, N. K. (2022). Body mass, temperature, and depth shape the maximum
527 intrinsic rate of population increase in sharks and rays. *Ecology and Evolution*, 12(11),
528 e9441. <https://doi.org/10.1002/ece3.9441>

529 Pedersen, T. (2024). *patchwork: The Composer of Plots*. [https://CRAN.R-](https://CRAN.R-project.org/package=patchwork)
530 [project.org/package=patchwork](https://CRAN.R-project.org/package=patchwork)

531 R Core Team. (2023). *R: A Language and Environment for Statistical Computing*.
532 <https://www.R-project.org/>

533 Ricker, W. (1954). Stock and Recruitment. *Journal of the Fisheries Research Board of Canada*,
534 11(5), 559–623. <https://doi.org/10.1139/f54-039>

535 Rockel, T. (2022). *missMethods: Methods for Missing Data*. [https://CRAN.R-](https://CRAN.R-project.org/package=missMethods)
536 [project.org/package=missMethods](https://CRAN.R-project.org/package=missMethods)

537 Toszogyova, A., Smyčka, J., & Storch, D. (2024). Mathematical biases in the calculation of the
538 Living Planet Index lead to overestimation of vertebrate population decline. *Nature*
539 *Communications*, 15(1), 5295. <https://doi.org/10.1038/s41467-024-49070-x>

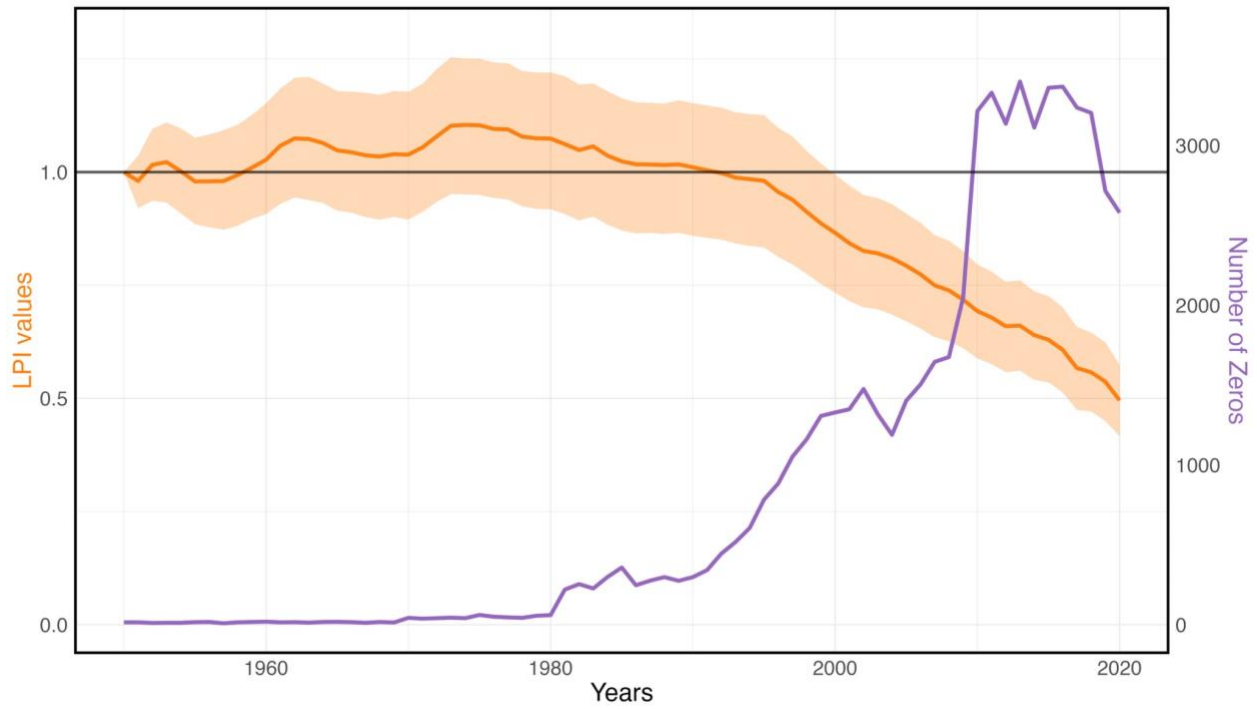
540 Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias
541 in biodiversity data and societal preferences. *Scientific Reports*, 7(1), 9132.
542 <https://doi.org/10.1038/s41598-017-09084-6>

543 Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
544 <https://ggplot2.tidyverse.org>

545 Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R.,
546 Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E.,
547 Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H.
548 (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
549 <https://doi.org/10.21105/joss.01686>

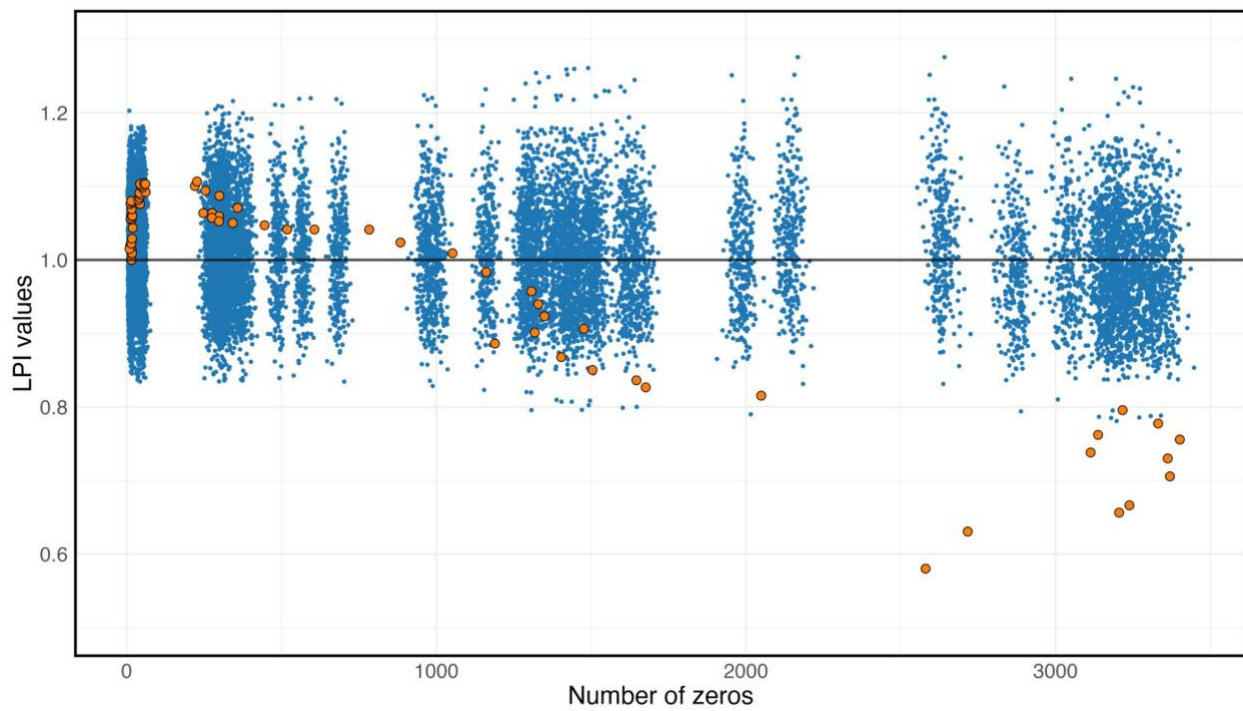
550 Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbaß, N.,
551 Pagel, J., Reineking, B., Schröder, B., & Grimm, V. (2010). The virtual ecologist
552 approach: Simulating data and observers. *Oikos*, 119(4), 622–635.
553 <https://doi.org/10.1111/j.1600-0706.2009.18284.x>
554

555 **Supplementary Material**



556

557 Figure S1. Relationship between the number of zeros and the LPI values across time (orange)
558 and the number of zeros per year (purple).



559

560 Figure S2: The relationship between the number of zeros and LPI values. Blue illustrates the
561 relation between the two metrics across 300 permutations, while orange dots reflect the
562 relationship using the LPD.