

# Methods in Ecology and Evolution

## The CV is dead, long live the CV!

Journal:	<i>Methods in Ecology and Evolution</i>
Manuscript ID	MEE-23-03-146.R1
Wiley - Manuscript Type:	Research Article
Keywords:	Kvålseth's coefficient of variation, Masting metric, MASTREE+, Pearson's coefficient of variation, Scale invariant statistics of variability
Abstract:	<p>1. Biology has an increasing need to reconsider the tools used to assess the variability of measurements, in addition to their central tendency. More than 100 years after Pearson's publication, most biologists still use the "good old" Pearson's coefficient of variation, PCV, despite its documented flaws such as sensitivity to excess zero values and/or irrelevant low mean values, which may compromise its use in some biological applications.</p> <p>2. A new statistic was developed in 2017 by Kvålseth, KCV, which is easy to implement. Unlike PCV, KCV is bounded (between 0 and 1), and it can be computed from PCV, ensuring backward compatibility with past studies. In addition to simulated data, we used the recent MASTREE+ database comprising the time series of the fruiting dynamics of perennial plants worldwide to compare the properties of PCV and KCV.</p> <p>3. Using as a benchmark the loose hump-shaped relationship between the interannual variability of fruiting and latitude, KCV led to significant increase in statistical power as it required almost half as many time series as PCV to detect the relationship. Perhaps most importantly, simulated data showed that KCV allows huge reductions in the length of time series required to estimate the population true variability, saving more than half the duration of long-term monitoring if fruiting fluctuations are very large, which is common in perennial plant species. Compared to the widely used PCV, KCV has great accuracy for estimating and analysing variability in biology, while strongly increasing statistical power.</p> <p>4. Selecting appropriate tools to assess the variability of measurements is crucial, particularly where the variability is of primary biological interest. Using Kvålseth's KCV is a promising avenue to circumvent the well-known issues of the former Pearson's PCV, its properties remain to be explored in other fields of biology, for purposes other than purely statistical ones (e.g. estimating heritability or evolvability of traits).</p>

SCHOLARONE™  
Manuscripts

# 1 The CV is dead, long live the CV!

2

3 Jean R. Lobry<sup>1</sup>,

4 Marie-Claude Bel-Venner<sup>1</sup>,

5 Michał Bogdziewicz<sup>2</sup>,

6 Andrew Hacket-Pain<sup>3</sup>,

7 and Samuel Venner<sup>1</sup>

8

9 Correspondence author: [jean.lobry@univ-lyon1.fr](mailto:jean.lobry@univ-lyon1.fr)

10 Running headline: the CV is dead, long live the CV!

11

1 <sup>1</sup>Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, Laboratoire de Biométrie et Biologie Evolutive, UMR5558, Villeurbanne, France

2 <sup>2</sup>Forest Biology Center, Faculty of Biology, Adam Mickiewicz University, Ul. Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland

3 <sup>3</sup>Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool, UK

## 12 Abstract

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

1. Biology has an increasing need to reconsider the tools used to assess the variability of measurements, in addition to their central tendency. More than 100 years after Pearson's publication, most biologists still use the "good old" Pearson's coefficient of variation,  $^P\text{CV}$ , despite its documented flaws such as sensitivity to excess zero values and/or irrelevant low mean values, which may compromise its use in some biological applications.
2. A new statistic was developed in 2017 by Kvålseth,  $^K\text{CV}$ , which is easy to implement. Unlike  $^P\text{CV}$ ,  $^K\text{CV}$  is bounded (between 0 and 1), and it can be computed from  $^P\text{CV}$ , ensuring backward compatibility with past studies. In addition to simulated data, we used the recent MASTREE+ database comprising the time series of the fruiting dynamics of perennial plants worldwide to compare the properties of  $^P\text{CV}$  and  $^K\text{CV}$ .
3. Using as a benchmark the loose hump-shaped relationship between the interannual variability of fruiting and latitude,  $^K\text{CV}$  led to significant increase in statistical power as it required almost half as many time series as  $^P\text{CV}$  to detect the relationship. Perhaps most importantly, simulated data showed that  $^K\text{CV}$  allows huge reductions in the length of time series required to estimate the population true variability, saving more than half the duration of long-term monitoring if fruiting fluctuations are very large, which is common in perennial plant species. Compared to the widely used  $^P\text{CV}$ ,  $^K\text{CV}$  has great accuracy for estimating and analysing variability in biology, while strongly increasing statistical power.
4. Selecting appropriate tools to assess the variability of measurements is crucial, particularly where the variability is of primary biological interest. Using Kvålseth's  $^K\text{CV}$  is a promising avenue to circumvent the well-known issues of the former Pearson'  $^P\text{CV}$ , its properties remain to be explored in other fields of biology, for purposes other than purely statistical ones (e.g. estimating heritability or evolvability of traits).

## 39 Keywords

- 40 Kvålseth's coefficient of variation
- 41 Masting
- 42 MASTREE+
- 43 Pearson's coefficient of variation
- 44 Scale invariant statistics of variability

## 45 Introduction

46 Biology has an increasing need to reconsider the tools used to assess the variability of  
47 measurements, in addition to their central tendency. This is particularly important in the fields  
48 of ecology and evolution, especially in the context of ongoing global change. For example, it  
49 is necessary to properly quantify the variability of population abundance in order to compare  
50 population dynamics and assess extinction risk. The dynamics and evolution of populations  
51 also depend strongly on the degree of variability of the environment or of individual  
52 phenotypes, which must be carefully assessed. There is general agreement that the  
53 appropriate statistic to estimate variability has to be scaled to the mean to facilitate  
54 comparisons (Gaston & McArdle, 1994; Inchausti & Halley, 2002; McArdle & Gaston, 1995;  
55 Pélabon et al., 2020; Pimm, 1991). In this respect, Pearson's coefficient of variation,  ${}^P\text{CV}$   
56 (Pearson, 1896), is the statistic almost exclusively used in biological studies to date.  ${}^P\text{CV}$  is  
57 computed from a series  $\mathbf{x}$  of  $n$  non-negative elements as the sample standard deviation to  
58 the sample mean ratio. Note that  ${}^P\text{CV}$  is unaffected by the order of elements in  $\mathbf{x}$  so that if  
59 you are interested in this aspect you should use order-dependent statistics (see for example  
60 Bogdziewicz et al., 2023) and that there is a loss of information if your interest is the  
61 variance-mean relationship (see for instance figure 2 in Pélabon et al., 2020).

62

63  ${}^P\text{CV}$  is commonly used as a convenient *dimensionless* statistics of variability, for instance in  
64 repeatability experiments it facilitates between-laboratories comparisons as they may use  
65 different units. Another interest is when there are widely different means between groups, for  
66 example in finance to compare the variability of securities in a stock exchange. In biology, it  
67 is perhaps the scale invariance property of  ${}^P\text{CV}$ ,

$${}^P\text{CV}(\lambda\mathbf{x}) = {}^P\text{CV}(\mathbf{x})$$

68

69 where  $\lambda$  is a strictly positive constant, that may explain its success. Intuitively, this statistic  
70 fits well with our need to assess the same *relative* variability in  $\mathbf{x} = (1, 2, 3)$  as in  $\mathbf{x} = (100,$

71 200, 300).

72

73 Non-negative series is a common situation since all *extensive* variables (e.g. mass, length,  
74 surface, volume, meristic variables such as seed counts) belong to this class. Calculating  
75 variability in the case of *intensive* variables (e.g. speed, pressure, density, temperature)  
76 could be more tricky, for example the same set of temperature data (Kimber, 1991) gives  
77 different  $^{\text{P}}\text{CV}$  values when expressed in degree Celsius or in degree Fahrenheit (Eisenhauer,  
78 1993) because  $^{\text{P}}\text{CV}$  is not invariant by translation. Scale types that are meaningless for  $^{\text{P}}\text{CV}$   
79 are given in table 1 in Pélabon et al., 2020. Even if the scale type is appropriate, there are  
80 still other well-known issues with  $^{\text{P}}\text{CV}$  (Gaston & McArdle, 1994; Kvålseth, 2017; Lewontin,  
81 1966; McArdle et al., 1990; McArdle & Gaston, 1995; Pélabon & Hansen, 2008; Silveira &  
82 Siqueira, 2022) such as its sensitivity to outliers and the fact that it is strongly affected by  
83 small variation in the mean, or errors in the estimation of the mean.

84 A few attempts have been made to find alternatives to  $^{\text{P}}\text{CV}$ : Lewontin (1966) proposed to  
85 work with the standard deviation of log-transformed values which can easily be mobilised in  
86 the context of allometric studies, where sample values are *strictly* positive. While this does  
87 not work, however, for counts including zero values, correcting the problem by using an  
88 arbitrary constant  $\alpha$  to enforce positivity in  $\log(\alpha + x_i)$  is inappropriate because the scale  
89 invariance property is lost in the process. The proportional variability statistic,  $\text{PV}$ , was  
90 proposed (Heath, 2006; Heath & Borowski, 2013) to address these challenges, but itself has  
91 major weaknesses (see table 1).

Dataset number	Ten-year dataset										PV	<sup>P</sup> CV	<sup>K</sup> CV	
	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$t_7$	$t_8$	$t_9$	$t_{10}$				
1	5	5	5	5	5	5	5	5	5	5	1000	0.20	2.86	0.94
2	995	995	995	995	995	995	995	995	995	995	0	0.20	0.33	0.32
3	1	2	3	4	5	6	7	8	9	1000	0.60	2.86	0.94	
4	0	0	0	0	0	0	0	0	0	1000	0.20	3.00	0.95	

92

93 Table 1. Comparison of PV (Heath, 2006), <sup>P</sup>CV (Pearson, 1896) and <sup>K</sup>CV (Kvålseth, 2017)

94 values on the same time series. The first issue is that the same PV values are obtained for

95 time series composed mainly of very low values and including seldom high values and those

96 mirror series mostly composed of high values with seldom low values (sets 1 and 2). The

97 second issue is that very different PV values are obtained for time series that are nearly

98 identical from a biological perspective (sets 3 and 4). The differences between sets 3 and 4

99 are minute or even meaningless yet commonly encountered as they may arise due to

100 sampling fluctuations. In these case studies, both <sup>P</sup>CV and <sup>K</sup>CV are sensitive to meaningful

101 differences and are insensitive to artifactual differences.



102 Recently, a new coefficient of variation has been proposed (Kvålseth, 2017), called hereafter  
103 Kvålseth's coefficient of variation,  ${}^K\text{CV}$ , which has gone largely unnoticed by biologists.  ${}^K\text{CV}$   
104 is as easy to compute as  ${}^P\text{CV}$ , since it is the sample standard deviation divided by the  
105 square root of the mean of squared values. What is more,  ${}^K\text{CV}$  can be seen as a variance  
106 stabilisation transformation of  ${}^P\text{CV}$ :

107

$${}^K\text{CV} = \sqrt{\frac{{}^P\text{CV}^2}{1 + {}^P\text{CV}^2}}$$

108

109

110 This relationship allows us to compute  ${}^K\text{CV}$  from formerly reported  ${}^P\text{CV}$  values, even if the  
111 original dataset is no longer available. This relationship also shows that when  ${}^P\text{CV}$  tends to  
112 infinity,  ${}^K\text{CV}$  is still bounded below 1. The other advantages of  ${}^K\text{CV}$  over  ${}^P\text{CV}$  are theoretically  
113 demonstrated in Kvålseth's paper. For instance  ${}^K\text{CV}$  can be used with a signed ratio type  
114 scale mixing positive and negative values since it is not undefined, unlike  ${}^P\text{CV}$ , when the  
115 mean is 0, and, at least for someone familiar with multivariate analyses, there is a nice  
116 geometrical interpretation of  ${}^K\text{CV}$  in terms of Euclidean distances in  $\mathbb{R}^n$ .

117 Here, we highlight the interest of  ${}^K\text{CV}$  on the basis of a practical case study and of  
118 simulations used to compare the gain in *statistical* power or in the sampling effort associated  
119 with the use of  ${}^K\text{CV}$  vs  ${}^P\text{CV}$ . For that purpose, we used annual seed production in perennial  
120 plant populations as a case study. These populations show diverse fruiting dynamics,  
121 ranging from nearly constant annual production, through extreme interannual variation  
122 (masting), to semelparity in some species such as the mainland Chinese bamboo  
123 *Phyllostachys bambusoides* with its seeding cycle of about 120 years (Janzen, 1976). This  
124 may represent the greatest known variation ever recorded among biological variables in  
125 terrestrial ecosystems, providing an ideal example of the challenges with measuring  
126 variability.

## 127 Material and Methods

128

129 The demonstration of the statistical power gain associated with  $^K\text{CV}$  rather than  $^P\text{CV}$  in a  
130 biological context is carried out in two complementary steps, one based on quantified  
131 biological data in the field and the other on simulated data from true parameters known *a*  
132 *priori*.

133 In the first step, we used the numerous time series describing fruiting dynamics quantified at  
134 the scale of perennial plant populations and species, and in various localities around the  
135 world. Data recently made available in MASTREE+ (Hackett-Pain et al., 2022) offer a great  
136 opportunity to compare the behaviour of  $^P\text{CV}$  and  $^K\text{CV}$  because the series cover a very wide  
137 range of variability. This is a *libre* database available under a CC-BY-4.0 licence. We used  
138 the initial (2022-03-03) version. Quantitative time series with at least 12 documented values  
139 were selected ( $n = 1433$  time series). From this database we describe the relationship  
140 between  $^P\text{CV}$  and  $^K\text{CV}$  and then analyse the gain in statistical power associated with using  
141  $^K\text{CV}$  (compared to  $^P\text{CV}$ ) using a test to detect a previously published relationship between  
142 the degree of variability in fruiting and latitude (Pearse et al., 2020). To do this, we sub-  
143 sample the MASTREE database by randomly drawing time series. For each sub-sample  
144 size, we simulate 10,000 independent tests (either with  $^K\text{CV}$  or with  $^P\text{CV}$ ) and determine the  
145 proportion of tests that detected a significant ( $p < 0.05$ ) quadratic relationship between CVs  
146 and latitude. The power gain of using  $^K\text{CV}$  instead of  $^P\text{CV}$  is quantified by the difference  
147 between the sub-sample size needed by each statistic to detect a significant relationship in  
148 95% of the tests .

149

150 In a second step, we use a simulation experiment based on a lognormal distribution to  
151 generate the fruiting dynamics over a longer or shorter time series. The use of lognormal  
152 distribution has two advantages: (i) it allows the generation of fruiting dynamics consistent  
153 with observations, (ii) it requires the use of only one parameter. Once this parameter is fixed,

154 the true  $^P\text{CV}$  and  $^K\text{CV}$  are known (Kvålseth, 2017). From these simulation experiments, the  
155 gain in statistical power associated with the use of  $^K\text{CV}$  can be estimated by the savings in  
156 sampling effort (number of years saved) to estimate the true CVs with a chosen degree of  
157 precision. To illustrate the approach, we initially used  $\text{sdlog} = 1.010768$  in the  $\text{rlnorm}()$   
158 function, which corresponds to variability at the boundary between the "large" and "very  
159 large" ranges for  $^K\text{CV}$ , typical of masting studies, and the theoretical time series have true  
160  $^P\text{CV}$  and  $^K\text{CV}$  of 1.33 and 0.8 respectively. For each length of time series, 10,000 replicates  
161 were sampled and statistics were calculated on the *same* time series. Zero-inflated  
162 distributions were simulated by forcing a given fraction of the smallest values to zero. Then  
163 we generalised the procedure by using lognormal distributions to generate true  $^K\text{CV}$  ranging  
164 from 0.4 to 0.95 (in steps of 0.25).

165 All computations were done under the R statistical software (R Core Team, 2013).  
166 Nonparametric confidence intervals for statistics were computed with the boot package  
167 (Canty & Ripley, 2021; Davison, 1997) using the adjusted bootstrap percentile (BCa) method  
168 (Efron, 1987) and 9999 replicates. The R code to reproduce the analyses is available in the  
169 file *CVisDead.zip* at [pbil.univ-lyon1.fr/R/donnees/](http://pbil.univ-lyon1.fr/R/donnees/) in the form of an RMarkdown document  
170 (Allaire et al., 2020; Xie et al., 2018, 2020) compiled with knitr (Xie, 2014, 2015, 2020).

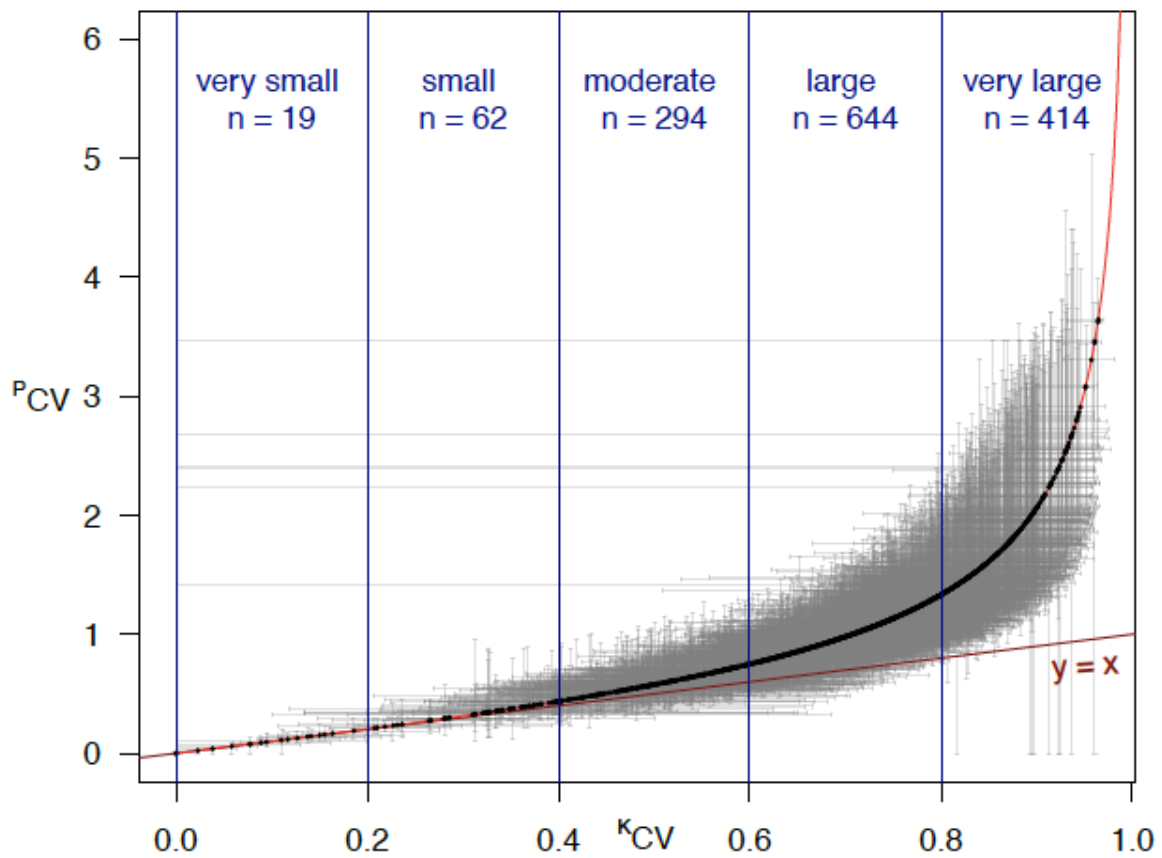
## 171 Results and discussion

### 172 Comparison of $^P\text{CV}$ and $^K\text{CV}$ general properties based on true 173 datasets

174

175 Paired calculations of  $^P\text{CV}$  and  $^K\text{CV}$  over a large dataset of field time series of fruiting  
176 dynamics by perennial plant species show that they are essentially the same up to moderate  
177 variability range, but for greater variability  $^P\text{CV}$  tends to stretch values to infinity. This is a  
178 common situation in masting studies since 74% of time series in MASTREE+ are in the large  
179 -or very large- variability range (Fig. 1). The  $^K\text{CV}$  estimates are accurate enough (with  
180 confidence intervals for  $^K\text{CV}$  typically  $\pm 0.1$ ) to consider as relevant the 5-class categorization  
181 of the [0,1] range values proposed by Kvålseth for verbal interpretation. The consistency of  
182 results when switching from  $^P\text{CV}$  to  $^K\text{CV}$  is ensured by their monotonic relationship, for  
183 instance all non-parametric rank based tests are equivalent since ranks are preserved. The  
184 “too many zeros issue”, common in masting studies, is solved neither by  $^P\text{CV}$  nor by  $^K\text{CV}$  but  
185 it can be at least detected by the confidence intervals including zero, meaning that the  
186 corresponding data set does not allow us to reject the null hypothesis “ $H_0 : \text{CV} = 0$ ”.

187



188

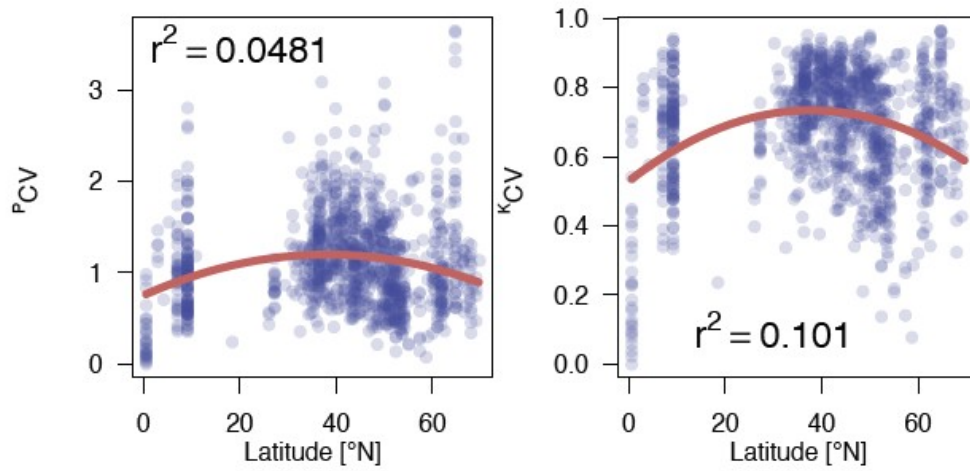
189 Figure 1. Comparison of  $^P\text{CV}$  and  $^K\text{CV}$  statistics for 1433 masting time-series with at least 12  
 190 observations from MASTREE+ (Hackett-Pain et al., 2022). The grey lines are the 95%  
 191 bootstrap confidence interval (Efron, 1987). The vertical blue lines are the boundaries of  
 192 Kvålseth's ranges for verbal interpretation of variability. The red curve is the theoretical  
 193 relationship ( $y^2 = x^2 / (1 - x^2)$ ) between  $^P\text{CV}$  and  $^K\text{CV}$ .

## 194 Comparison of $^P$ CV and $^K$ CV power with actual data

195

196 The level of variability in the population-scale fruit production has been recently examined  
197 over a large range of plant species and spatial scale in the Northern Hemisphere (Pearse et  
198 al., 2020) and the time-series variability exhibits a loose hump-shaped relationship with  
199 latitude (Fig. 2). This is a perfect benchmark because the small part of total variability  
200 accounted for by the model ( $r^2 = 0.0481$ ) requires a lot of data to get a significant  
201 relationship. With  $^P$ CV the relationship is questionable because there is an  
202 overrepresentation of data in the intermediate latitude range [35°-55°] likely including by  
203 chance most of the outliers (anomalously high  $^P$ CV values), which could be responsible for  
204 an artificial quadratic relationship. In this case, using  $^K$ CV the hump-shaped relationship is  
205 much more convincing because its values are bounded between 0 and 1 so that no heavy  
206 tailed distributions are possible. In this way,  $^K$ CV is similar to using a log scale when dealing  
207 with highly skewed data, but avoids the need for data transformations. In this case the  
208 advantage of a bounded statistic is obvious, preventing highly skewed distributions for  $^P$ CV  
209 values (Fig. 1) and helps, using Kelly's words (Kelly, 2023), in "fighting the urge to put things  
210 in bins".

211



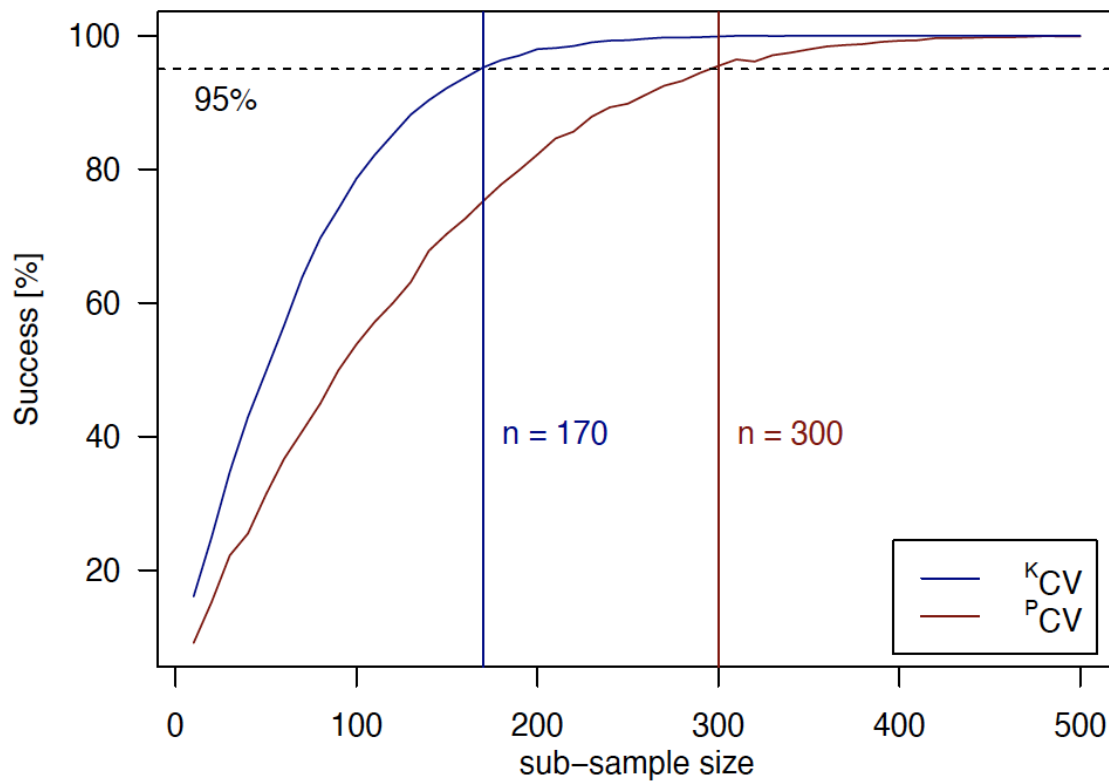
212

213 Figure 2. Quantifying the relationship between variability and latitude using the two methods  
214 of calculating CV. Points show a subset of 1138 time-series from the Northern Hemisphere  
215 showing the relationship between  $^P$ CV (left) or  $^K$ CV (right) and latitude. The red line is the  
216 quadratic fit that minimises the sum of squared residuals.

217 Based on these data and from sub-sampling simulation, we found a massive gain in  
218 statistical power when using  $^K\text{CV}$  instead of  $^P\text{CV}$  (Fig. 3): we may save about 40% of the  
219 sampling effort to reach a significant result. The advantage of using the  $^K\text{CV}$  instead of  $^P\text{CV}$   
220 is worth considering, given the logistical difficulties in long-term field monitoring of seed  
221 production, which pose a major obstacle to progress in the field (Clark et al., 2021; Koenig,  
222 2021).



223



224

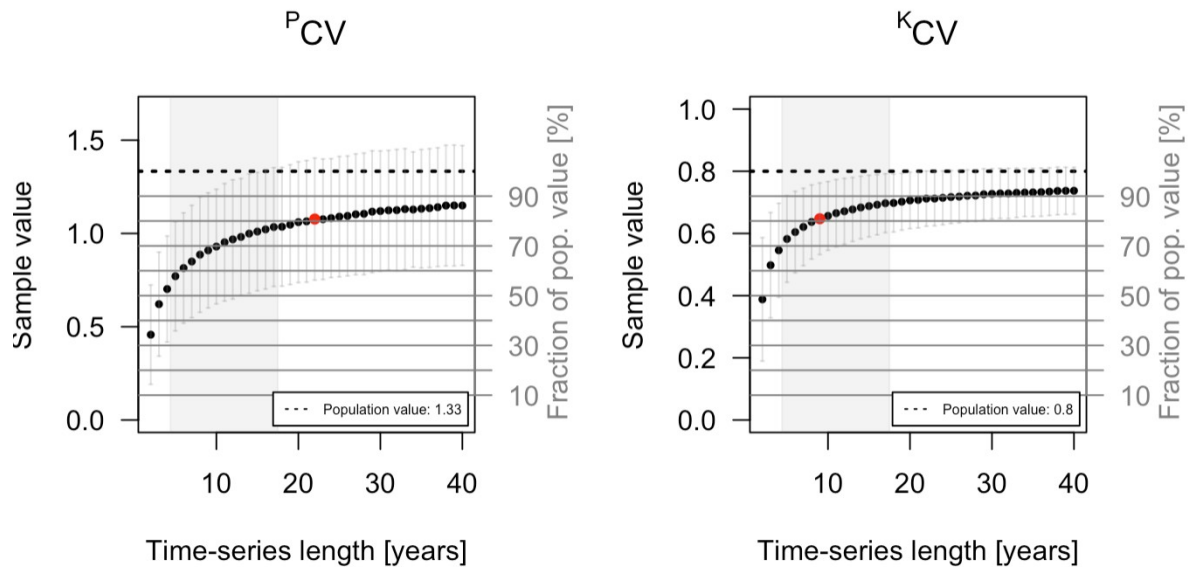
225 Figure 3. Sub-sampling simulation showing how  $^k\text{CV}$  dramatically reduces the number of  
 226 samples required to detect a significant relationship between CV and latitude, as shown in  
 227 Fig. 2. Lines show the percentage of simulations for a given sub-sample size that produce a  
 228 significant ( $p < 0.05$ ) quadratic relationship between CVs and latitude, based on 10,000  
 229 replicates for each sub-sample size. Sub-samples were randomly selected from the 1138  
 230 MASTREE+ time-series. Detecting the relationship using  $^k\text{CV}$  saved 43% of the sampling  
 231 effort as compared with  $^P\text{CV}$ .

232 *In silico* comparison of <sup>P</sup>CV and <sup>K</sup>CV power

233

234 Using simulated data allows us to study the impact of sampling effort (the length of time-  
235 series) on the sampling fluctuation of statistics values, from both a central tendency and a  
236 dispersion point of view (Fig. 4). The dispersion of sampling fluctuations decreases with  
237 sampling effort with <sup>K</sup>CV but is almost unchanged with <sup>P</sup>CV, an undesirable property.  
238 Examining the central tendencies, the convergence to the true value is faster with <sup>K</sup>CV than  
239 with <sup>P</sup>CV (Fig. 4). For example, reaching 80% of the true population value requires 22 years  
240 with <sup>P</sup>CV while it takes only 9 years with <sup>K</sup>CV, corresponding to a 13-year gain (*i.e.* more than  
241 50% saved years). A similar gain was observed with zero-inflated time-series (not shown). At  
242 the expense of no extra cost, with the same dataset, we are always closer to the true  
243 population value with <sup>K</sup>CV. Crucially in the case of masting analyses, this enables substantial  
244 reduction in the number of years of monitoring needed prior to accurately measuring the  
245 intensity of masting (Fig. 4).

246



247

248

249

250

251

252

253

254

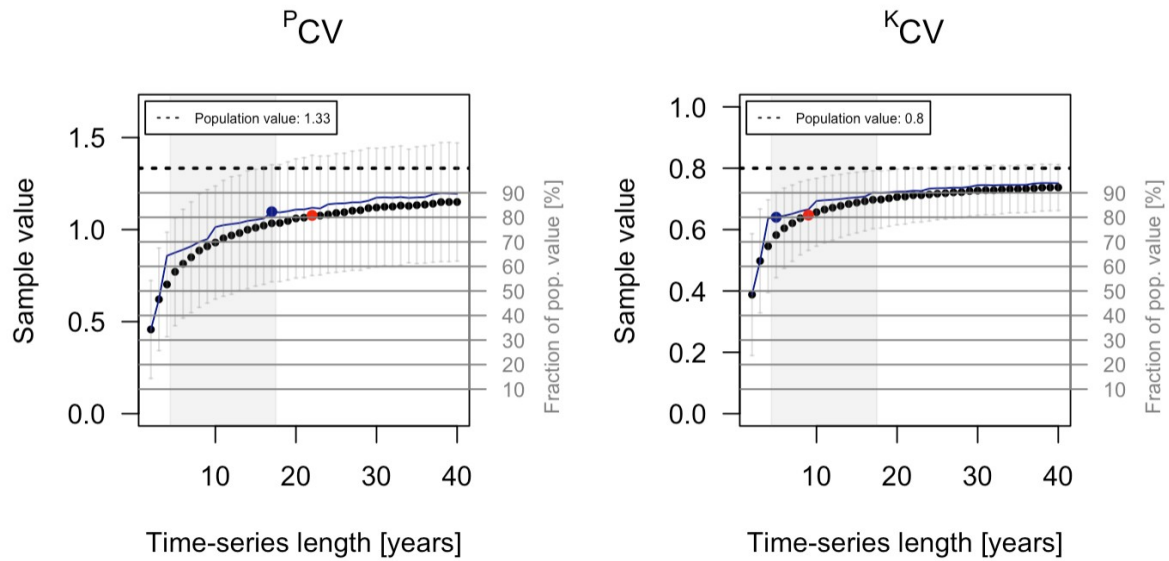
255

256

257

Figure 4. Simulation experiment using a log-normal distribution to demonstrate that while  $P_{CV}$  and  $K_{CV}$  both underestimate the true population value,  $K_{CV}$  converges more rapidly than  $P_{CV}$ , reducing the number of years of observation required to estimate its value. The dotted lines are the true population values for  $P_{CV}$  and  $K_{CV}$ . The x-axis scale is representative of the length of the masting series available in MASTREE+, whose median is 10 years, and 50% of the time-series are between 4 and 17 years (indicated by the grey shading). The black point is at the mean and the bars represent plus or minus one standard deviation (not confidence interval for the mean) to illustrate the dispersion of the sample statistics. The red point indicates the time-series length where 80% of the true population value is reached.

258



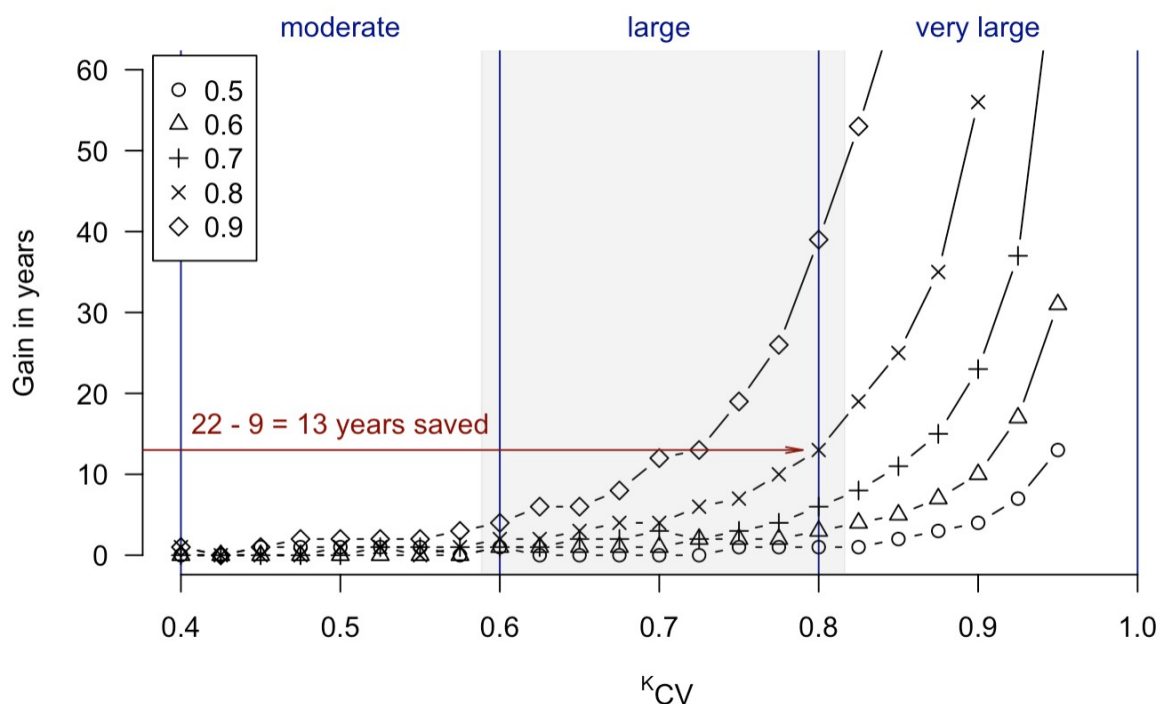
259

260 Not shown, for reviewers only. Same as Fig. 4 but the blue line is what is obtained with a  
 261 zero-inflated-log-normal distribution with a fraction of 15.15% of zero, corresponding to the  
 262 last quartile (25% of time-series in our subset of MASTREE+ have more than 15.15% of  
 263 zero). The blue point indicates the time-series length where 80% of the true population value  
 264 is obtained (17 years with  $P_{CV}$  and 5 years with  $K_{CV}$ , a gain of 12 years).

265 The amount of sampling effort saved when shifting from  ${}^P\text{CV}$  to  ${}^K\text{CV}$  was also found to  
266 increase along with the degree of variability in the data series (Fig. 5). For instance,  
267 considering that 80% of the true population value was reached, 13 years could theoretically  
268 be saved when  ${}^K\text{CV} = 0.8$ , 25 years when  ${}^K\text{CV} = 0.85$  and even 56 years when  ${}^K\text{CV} = 0.9$ . To  
269 summarise, whatever the length of the time series,  ${}^K\text{CV}$  always outcompetes  ${}^P\text{CV}$  and the  
270 reduction in the length of the time series allowed by  ${}^K\text{CV}$  increases along with the intrinsic  
271 variability level of the dataset.

272

273



274

275 Figure 5. The reduction in sampling effort (years) when switching from  $^P CV$  to  $^K CV$  to  
 276 measure variability of reproduction. Simulations plotted show the reduction in effort required  
 277 to reach, on average, a given fraction (0.5, 0.6, ..., 0.9 as indicated in the top-left box) of the  
 278 true population value as a function of the variability level ( $^K CV$ ). The red arrow highlights the  
 279 example shown in Fig. 4. The vertical blue lines are the boundaries of Kvålseth's ranges for  
 280 verbal interpretation. The grey area is the interquartile range for  $^K CV$  in our subset of  
 281 MASTREE+: 50% of quantitative time-series with at least 12 observations are between 0.59  
 282 and 0.82.

## 283 Conclusion

284

285 Kvålseth concludes his article by stating that “except for a long tradition of the use of  $^P$ CV,  
286 there appears to be no reason not to prefer the use of  $^K$ CV over  $^P$ CV”. The double negation  
287 in Kvålseth’s delicate wording appears to us as an understatement: *at least in studies*  
288 *devoted to understand the temporal or spatial variability of biological quantities, we do have*  
289 *good reasons to shift from  $^P$ CV to  $^K$ CV as a scale invariant statistic to properly quantify*  
290 *variability. Other applications of the  $^K$ CV deserve to be explored, such as in evolutionary*  
291 *biology where inferring the evolvability of a trait, its phenotypic plasticity, or its selective*  
292 *value relies on accurate, and still debated, measures of variability (Hansen et al., 2011;*  
293 *Houle, 1992; Houle et al., 2011; Pélabon et al., 2020). Nonetheless, while  $^K$ CV has a number*  
294 *of advantages for focal applications as presented in our paper, the choice of statistics will*  
295 *depend on the questions being asked.*

## 296 Acknowledgments

297 This work was funded by the ANR FOREPRO (ANR-19-CE32-0008) and the ECOFECT and  
298 RESPOND program of the Université de Lyon (ANR & UDL) and was supported by the  
299 Centre National de la Recherche Scientifique and the Université Claude Bernard Lyon 1,  
300 France. MB was supported by the European Union (ERC, ForestFuture, 101039066). Views  
301 and opinions expressed are however those of the authors only and do not necessarily reflect  
302 those of the European Union or the European Research Council. Neither the European  
303 Union nor the granting authority can be held responsible for them. *We thank the two*  
304 *anonymous reviewers for their constructive comments.*

## 305 Conflict of Interest

306 None

## 307 Author Contributions

308 All authors contributed critically to the drafts and gave final approval for publication. JLO  
309 coordinated the ideas from all authors and led the writing of the manuscript.



310 **Statement of where we intend to archive our data**

311 Not relevant: we used already published data from MASTREE+ (Hackett-Pain et al., 2022).

312

## 313 REFERENCES

314

315 Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng,  
316 J., Chang, W., & Iannone, R. (2020). *rmarkdown: Dynamic Documents for R*.

317 <https://github.com/rstudio/rmarkdown>

318 Bogdziewicz, M., Calama, R., Courbaud, B., Espelta, J. M., Hacket-Pain, A., Journé, V.,

319 Kunstler, G., Steele, M., Qiu, T., Zywiec, M., & Clark, J. S. (2023). How to measure  
320 mast seeding? *New Phytologist*, *239*, 830–838.

321 Canty, A., & Ripley, B. D. (2021). *boot: Bootstrap R (S-Plus) Functions*.

322 Clark, J. S., Andrus, R., Aubry-Kientz, M., Bergeron, Y., Bogdziewicz, M., Bragg, D. C.,

323 Brockway, D., Cleavitt, N. L., Cohen, S., & Courbaud, B. (2021). Continent-wide tree  
324 fecundity driven by indirect climate effects. *Nature Communications*, *12*(1), 1242.

325 Davison, D. V., A. C. Hinkley. (1997). *Bootstrap Methods and Their Applications*. Cambridge  
326 University Press.

327 Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical*  
328 *Association*, *82*(397), 171–185.

329 Eisenhauer, J. G. (1993). A measure of relative dispersion. *Teaching Statistics*, *15*(2), 37–  
330 39.

331 Gaston, K. J., & McArdle, B. H. (1994). The temporal variability of animal abundances:

332 Measures, methods and patterns. *Philosophical Transactions of the Royal Society of*  
333 *London. Series B: Biological Sciences*, *345*(1314), 335–358.

334 Hacket-Pain, A., Foest, J. J., Pearse, I. S., LaMontagne, J. M., Koenig, W. D., Vacchiano,  
335 G., Bogdziewicz, M., Caignard, T., Celebias, P., van Dormolen, J., Fernández-

336 Martínez, M., Moris, J. V., Palaghianu, C., Pesendorfer, M., Satake, A., Schermer, E.,  
337 Tanentzap, A. J., Thomas, P. A., Vecchio, D., ... Ascoli, D. (2022). MASTREE+:

338 Time-series of plant reproductive effort from six continents. *Global Change Biology*,  
339 *00*, 1–17.

340 Hansen, T. F., Pélabon, C., & Houle, D. (2011). Heritability is not evolvability. *Evolutionary*

- 341 *Biology*, 38, 258–277.
- 342 Heath, J. P. (2006). Quantifying temporal variability in population abundances. *Oikos*,  
343 115(3), 573–581.
- 344 Heath, J. P., & Borowski, P. (2013). Quantifying proportional variability. *PLoS One*, 8(12),  
345 e84074.
- 346 Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130,  
347 195–204.
- 348 Houle, D., Pélabon, C., Wagner, G. P., & Hansen, T. F. (2011). Measurement and meaning  
349 in biology. *The Quarterly Review of Biology*, 86, 3–34.
- 350 Inchausti, P., & Halley, J. (2002). The long-term temporal variability and spectral colour of  
351 animal populations. *Evolutionary Ecology Research*, 4, 1033–1048.
- 352 Janzen, D. H. (1976). Why bamboos wait so long to flower. *Annual Review of Ecology and*  
353 *Systematics*, 7, 347–391.
- 354 Kelly, D. (2023). Mast seeding: Study of oak mechanisms carries wider lessons. *Current*  
355 *Biology*, 33, R231–R233.
- 356 Kimber, A. (1991). What A Scorcher! An analysis of Some Temperature Data. *Teaching*  
357 *Statistics*, 13(2), 34–37.
- 358 Koenig, W. D. (2021). A brief history of masting research. *Philosophical Transactions of the*  
359 *Royal Society B*, 376(1839), 20200423.
- 360 Kvålseth, T. O. (2017). Coefficient of variation: The second-order alternative. *Journal of*  
361 *Applied Statistics*, 44, 402–415.
- 362 Lewontin, R. C. (1966). On the measurement of relative variability. *Systematic Zoology*,  
363 15(2), 141–142.
- 364 McArdle, B. H., & Gaston, K. J. (1995). The temporal variability of densities: Back to basics.  
365 *Oikos*, 165–171.
- 366 McArdle, B. H., Gaston, K. J., & Lawton, J. H. (1990). Variation in the size of animal  
367 populations: Patterns, problems and artefacts. *The Journal of Animal Ecology*, 59,  
368 439–454.

- 369 Pearse, I. S., LaMontagne, J. M., Lordon, M., Hipp, A. L., & Koenig, W. D. (2020).  
370 Biogeography and phylogeny of masting: Do global patterns fit functional  
371 hypotheses? *New Phytologist*, *227*, 1557–1567.
- 372 Pearson, K. (1896). VII. Mathematical contributions to the theory of evolution.—III.  
373 Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society*  
374 *of London. Series A*, *187*, 253–318.
- 375 Pélabon, C., & Hansen, T. F. (2008). On the adaptive accuracy of directional asymmetry in  
376 insect wing size. *Evolution*, *62*, 2855–2867.
- 377 Pélabon, C., Hilde, C. H., Einum, S., & Gamelon, M. (2020). On the Use of the Coefficient of  
378 Variation to Quantify and Compare Trait Variation. *Evolution Letters*, *4*, 180–188.
- 379 Pimm, S. L. (1991). *The balance of nature?: Ecological issues in the conservation of species*  
380 *and communities*. University of Chicago Press.
- 381 R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R  
382 Foundation for Statistical Computing.
- 383 Silveira, P. S. P., & Siqueira, J. O. (2022). Histogram lies about distribution shape and  
384 Pearson's coefficient of variation lies about relative variability. *Quantitative Methods*  
385 *for Psychology*, *18*, 91–111.
- 386 Xie, Y. (2014). knitr: A Comprehensive Tool for Reproducible Research in R. In V. Stodden,  
387 F. Leisch, & R. D. Peng (Eds.), *Implementing Reproducible Computational Research*.  
388 Chapman and Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>
- 389 Xie, Y. (2015). *Dynamic Documents with R and knitr* (2nd ed.). Chapman and Hall/CRC.  
390 <https://yihui.org/knitr/>
- 391 Xie, Y. (2020). *knitr: A General-Purpose Package for Dynamic Report Generation in R*.  
392 <https://yihui.org/knitr/>
- 393 Xie, Y., Allaire, J. J., & Golemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman  
394 and Hall/CRC. <https://bookdown.org/yihui/rmarkdown>
- 395 Xie, Y., Dervieux, C., & Riederer, E. (2020). *R Markdown Cookbook*. Chapman and  
396 Hall/CRC. <https://bookdown.org/yihui/rmarkdown-cookbook>