Sample Size for Bioindication of Pollution Level by the Method of Fluctuating Asymmetry

G Kozlov and M Pushkarev

Saint-Petersburg State Institute of Technology, 190013, 26 Moskovsky prospect, St. Petersburg, Russia

E-mail: malexpush@bk.ru

Abstract. Bioindication is a convenient tool for assessing the environmental situation and the level of pollution of natural and man-made landscapes. One of the most important indicator is the fluctuating asymmetry of plant organs, less common – organs of animals and humans. In 2019, we studied the fluctuating asymmetry of Pinus sylvestris L. located in the industrial zone (Yanino-1), as well as in city parks of St. Petersburg (Alexandrino, Sergievka) and Moscow (Sokolniki). The dependence of the fluctuation asymmetry on the sample size is studied, the volumes of samples are substantiated, an algorithm for processing experimental data is proposed, which allows increasing the accuracy of experiments, avoids false positive or false negative results.

1. Introduction

The relevance of the work is caused, on the one hand, by the lack of data on the fluctuating asymmetry of Pinus sylvestris L. growing in St. Petersburg, and, on the other hand, by the fact that most researchers neglect some aspects of the biological specificity of the indicator object under study.

Fluctuating asymmetry is widely used in various fields of science - this is how studies of fluctuating asymmetry of pine needles [1], birch leaves [2] are described to assess the quality of the environment; skulls of rodents [3] and birch leaves [4] are described to assess the effects of radioactive contamination in the Chernobyl zone; mammary glands [5] are described for the prognosis of cancer; human teeth [6] - in archeology. Moreover, the authors arbitrarily determine the sample size, which varies from units, tens [7], and hundreds to one and a half thousand [8] and more. Although, in most cases, the sample has a very significant size. With rare exceptions [9], it is not substantiated in any way. On the one hand, this circumstance can lead to redundancy of the experiment, or vice versa, the sample size will be insufficient, and a false positive or false negative result will occur. Studying the dependence of experimental values on the sample size made it possible to identify the causes of contradictions and, using the simplest transformations, to obtain well-interpreted results. We assume that ignoring the specificity of the dependence of the average values of sample indicators from sample size in biological and medical observations is one of the reasons for some unsuccessful attempts to apply fluctuating asymmetry in medicine [10] or radioecology [11], as well as the contradiction of the results of various studies [4, 11]. It is important to note that arbitrary (stereotypical) sample size is not an error or an incorrect approach. Its justification and transformation are one of those “style elements” in biological statistics that David L. Vaux speaks about [12].
2. Materials and methods
The study used the herbarium collected in obviously environmentally favourable and unfavourable locations in Moscow, St. Petersburg and the Leningrad Region (Russia). The degree of environmental well-being was determined on the basis of data from the state environmental report [13].

The measurements were made with a ruler with an accuracy of 0.5 mm. The fluctuation asymmetry was calculated according to the formula (1) [14] (the ratio of the difference between the lengths of the right and left needles of the paired needles to their average value):

\[ A_{f1} = 2 \cdot \frac{l_1 - l_2}{l_1 + l_2} \]  

where: \( l_1 \) - length of the left needle; \( l_2 \) - length of the right needle of paired needles.

3. Results and discussion
For study, the herbarium was selected from the following locations (points) with deliberately strong differences in environmental well-being (according to [11]): industrial zone (point I - Leningrad region, Yanino-1 village, waste processing complex SPb SUE “Zavod MPBO-2” - coordinates 59.939226, 30.600066), a city park near a busy road (point II - Moscow, Sokolniki park - coordinates 55.795333, 37.671889; point III - St. Petersburg Alexandrino park - coordinates 59.827989, 30.234273), and a park in the recreation area of the resort suburbs (point IV – Peterhof, Sergievka park, beach - coordinates 59.892370, 29.835802; point V - track - coordinates 59.892764, 29.835802).

For them, the dependences of the average value of fluctuating asymmetry on the sample size were plotted (Figure 1). The graphs show that for small sample sizes, fluctuations in the average value are very large, and even the volume of the 500 variants does not allow stabilizing the average value near the general one. Only with a sample size of more than \( n \geq 1700 \) in ecologically clean (park in the resort suburbs) and moderately polluted (city park) locations, the stabilization of the average value of fluctuating asymmetry near the general one is observed. In an ecologically “polluted” location (industrial zone), stabilization does not occur even with a sample size of \( n \geq 2000 \), which suggests that a larger number of variants is needed for reliable estimation of fluctuating asymmetries than are available in most publications.

As for the values of fluctuating asymmetry and environmental pollution indicators, they are in good agreement with studies by other authors [8], official data [13] and our data [15] on heavy metal pollution of MSW treatment facilities in St. Petersburg.

Obviously, increasing the sample size is not always technically possible, and, in some cases, such as when using the parameters of the skeleton of animals, is not ethical. Therefore, to reduce the complexity, and in the case of the use of animals - for the ethics of research, it is necessary to use mathematical methods.

A consistent estimation tends to its general value as the sample size increases, which is shown by the graphs in Figure 1. Thus, we need to build the dependence using the available experimental data, straighten it, and determine the desired value by extrapolation.

The graphs in Figure 1 are not very suitable for transformations, since their type depends on which account one or another variant turns out to be. So, the first step to eliminate this effect is to rank the variants in descending order. A certain number of variants has a fluctuating asymmetry equal to zero. This is due to the fact that the differences between the needles in a pair are less than the measurement accuracy. Therefore, as long as the difference between the needles in a pair can be fixed by means of measuring, fluctuating asymmetry (in a ranked sample) will strive for a value equal to the ratio of the measurement accuracy to the upper limit of the average length of a pair of needles, and as soon as the measurement accuracy is insufficient to register it, it will abruptly take a zero value. The character of the dependence of nonzero values of fluctuating asymmetry is described by an exponential function.

If we plot the dependence of the average value of fluctuating asymmetry on a ranked sample, we find that the arithmetic average of fluctuating asymmetry asymptotically tends to its general value as the size of the sample for which the calculations are performed increases (Figure 2). Finding the value to which the quantity asymptotically tends is reduced to straightening the dependence (in the case of
an exponential function - using the inverse coordinates) and extrapolating it to the intersection with the corresponding coordinate axis (Figure 3). The Lineweaver-Burk equation is a classic example of using this technique in biology [16]. Performing the inverse transformation, we obtain the desired value.

**Figure 1.** Dependence of the average value of fluctuating asymmetry on the sample size.
The proposed algorithm is as follows:

1. Plot the dependence of the average value of fluctuating asymmetry ($\bar{A}_{f\ell}$) on the sample size ($n$) until it stabilizes around a certain constant value.
2. Rank the values (initial) of $A_{f\ell}$ in descending order.
3. Outline the number of points ($m$) by which the line will be built (we recommend $m=15÷20$).
4. Divide the sample into $m$ parts.
5. For each $n_i = n(N/m)_i$ (where $N$ – total sample size, and $i = 1, 2, ..., m$), find the average value of $(\bar{A}_{f\ell})_i$ from the ranked series of values of $A_{f\ell}$.
6. Find the inverse values of the obtained values of $(\bar{A}_{f\ell})_i$.
7. Rank them in descending order.
8. Build the dependence of the obtained values of $\frac{1}{\bar{A}_{f\ell}} = a + b \cdot n$.
9. Extrapolate the linear dependence on the coordinate axis of the inverse fluctuating asymmetry $\frac{1}{\bar{A}_{f\ell}} = a$.
10. Find the desired value of $\bar{A}_{f\ell} = \frac{1}{a}$.

![Figure 2](image)

* - values are calculated according to the equations in figure 3

**Figure 2.** Dependence of the average value of fluctuating asymmetry on the sample size ranked in descending order. Location - Sergievka park (track).

Having carried out the transformations in accordance with the above algorithm, we found that the best environmental situation from the considered locations is observed in Sergievka park in Peterhof. It is confirmed by the results of measurements of fluctuating asymmetries, while the sampling point near the site for building bonfires shows large values of fluctuating asymmetry (0.109) than the point where there are no local sources of air pollution (0.0070). Thus, weak sources of pollution in ecologically clean locations can give strong local effects. Therefore, when collecting herbarium, they must be avoided. In the city parks Sokolniki, Moscow, and Aleksandrino, St. Petersburg, the environmental situation, according to the measurement results, is almost identical - the fluctuating asymmetry is 0.0162 and 0.0159. In terms of fluctuating asymmetry (0.0335), the industrial zone territory (SPb SUE “Zavod MPBO-2”) corresponds to the zone with a high level of pollution with heavy metals [8], which is also caused by the presence of a large amount of heavy metals in MSW and MSW compost.
Figure 3. Dependences of the inverse values of fluctuating asymmetry on the sample size. Location - SPb SUE “Zavod MPBO-2”.

4. Conclusions
To assess the fluctuating asymmetry in ecologically clean locations (city parks), it is necessary to use a sample size of at least 1600-1700, in ecologically polluted locations - more than 2000 replicates. An algorithm is proposed that allows one to more accurately calculate the values of fluctuating asymmetry, including from samples that are insufficient to stabilize the average value near the general value. In the study of fluctuating asymmetries, local sources of air pollution should be avoided.

Acknowledgments
The work was performed as part of the state assignment of Ministry of Science and Higher Education of the Russian Federation (785.00. X6019).

References
[1] Kozlov M V and Niemela P (1999) Difference in Needle Length - A New and Objective Indicator of Pollution Impact on Scots Pine (Pinus sylvestris). in Sheppard L J, Cape J N (eds) Forest Growth Responses to the Pollution Climate of the 21st Century (Springer, Dordrecht) https://doi.org/10.1007/978-94-017-1578-2_31
[2] Kozlov M, Wilsey B, Koricheva J, and Haukioja E (1996) Journal of Applied Ecology 33(6), 1489-1495. https://doi.org/10.2307/2404787
[3] Oleksyk T K, Novak J M, Purdue J R, Gashchak S P, and Smith M H (2004) *Journal of Environmental Radioactivity* **73**(1), 1-20 https://doi.org/10.1016/j.jenvrad.2003.07.001

[4] Møller A P, and Møller A P (1998) *Oikos* **81**(3) 444 https://doi.org/10.2307/2546765

[5] Scutt D, Lancaster G A, and Manning J T (2006) *Breast Cancer Research* **8**(2) R14 https://doi.org/10.1186/bcr1388

[6] Khudaverdyan A Yu, Yengibaryan A A, Vardanyan Sh A, Karalyan Z A, and Matevosyan R Sh (2014) Comparison of fluctuation asymmetry of the dental system with physiological stress indicators. *Biological Journal of Armenia* **66**(2) 12-18 (In Russian)

[7] Ialkovskaia L E, Fominykh M A, Mukhacheva S V, Davydova Iu A, and Borodin A V (2016) *Ekologiya* **3** 213-220 (In Russian) https://doi.org/10.7868/s0367059716030185

[8] Chudzinska E, Pawlaczyk E M, Celinski K, and Diatta J (2013) *Water and Environment Journal* **28**(4) 533-539 https://doi.org/10.1111/wej.12068

[9] Trubyanov A B and Glotov N V (2010) Fluctuating asymmetry: variation trait and correlation left - right. *Doklady AN* **431**(2) 283-285 (In Russian)

[10] Żelaźniewicz A and Pawłowski B (2015) *American Journal of Human Biology* **27**(5) 690-696 https://doi.org/10.1002/ajhb.22716

[11] Kashparova E, Levchuk S, Morozova V, and Kashparov V (2020) *Journal of Environmental Radioactivity* **211** 105731 https://doi.org/10.1016/j.jenvrad.2018.05.015

[12] Vaux D L (2012) *Nature* **492**(7428) 180-181 https://doi.org/10.1038/492180a

[13] Serebrickij I A (ed.) (2019) Report on the environmental situation in St. Petersburg in 2018 (Ltd "Sezam-print", Saint Petersburg) (In Russian) https://www.gov.spb.ru/static/writable/ckeditor/uploads/2019/08/12/42/doklad_za_2018_EKOLOGIA2019.pdf

[14] Palmer A R and Strobeck C (2003) Fluctuating asymmetry analyses revisited. Developmental instability: causes and consequences, 279-319.

[15] Pyankova E D, Ivahnyuk G K, and Kozlov G V (2008) Study of the effect of the Yuzhny municipal solid waste landfill on groundwater pollution. *Zhurnal prikladnoj himii* **81**(1), 169b-169. (In Russian)

[16] Lineweaver H and Burk D (1934) *Journal of the American Chemical Society* **56**(3) 658-666 https://doi.org/10.1021/ja01318a036