



Pollen counts statistics and its relevance to precision

Paul Comtois^{1*}, Purificacion Alcazar^{1,2} & Daniel Néron¹

¹Laboratoire d'aérobiologie, département de géographie, Université de Montréal CP 6128, Montréal, Québec, Canada, H3C 3J7; ²Division de Botanica, Departamento de Biología Vegetal y Ecología, Facultad de Ciencias, Universidad de Córdoba, Spain

(*author for correspondence, e-mail: comtoisp@ere.umontreal.ca; fax: +1 514 3438004)

(Received 3 December 1996; accepted in final form 6 November 1998)

Key words: error, field, longitudinal, pollen count, precision, statistics, transverse

Abstract

In the day to day management of pollen counts from aerobiological samples of national networks, only a small proportion (usually from 12 to 15%) of the daily microscope slide is read. It is generally believed that, otherwise, too much time will be spent reading slides for a minimal increase in precision. Different networks use different slide sampling methods (longitudinal, transverse or at random) and a different number of lines are routinely read. However, the topic of the precision of the different pollen count strategies has seldom been the object of serious investigation. In this study, the precision of different sampling methods of 12 pollen types was investigated by: **a)** counting pollen grains over the whole slide (3 slides per taxa), **b)** spatially (i.e. microscope field per microscope field) recording over the 3120 fields found at 400× the location of each pollen grain, **c)** sub-sampling, by macro procedures, this population by selecting a number of transverse (1 to 48) or longitudinal (1 to 20) lines, or a number of random fields (90 to 2340), so that between 0.96 to 46.15% (transverse), 3 to 66.6% (longitudinal) or 3 to 75% (random) of the whole slide was artificially counted. Between nine and twelve procedures were built per reading strategy. The error found is much higher than what is normally believed, and it was significantly correlated with the abundance of a pollen taxa on the sampled slide. It is only with a total count over 1000 (corresponding to a concentration of above 500 m⁻³) that the mean error of 4 longitudinal lines (or 13.3% of the slide), the standard protocol of both the Italian Association of Aerobiology (AIA) and the Spanish Aerobiology Network (REA), was always below 30%.

“En compulsant les travaux de micrographie atmosphérique, je n'ai rien vu qui méritait le nom de statistique microbique. Au contraire, certains auteurs, guidés jusqu'alors par les seules vues de l'esprit, ont dédaigneusement traité les germes aériens de négligeables; d'autres, sur la foi d'expériences mal conduites, ont été effrayés de leur nombre.”

P. Miquel, 1883

of minute floating aeroplankton and of the gigantic atmosphere, makes it that the population we are investigating is invisible to the naked eye. A hundred years ago what was believed to be aerobiology was a matter of philosophy (see the above quotation), nowadays, it is a matter of what, and above all, of how we are sampling. Sampling here must be understood in its broadest sense. It not only includes the statistical principles behind sampling only a small part of the atmosphere, but also of observing only a portion of the sampled organisms.

Statistics exist, and probabilities used in aerobiology, as in other sciences, because we never know the whole biological population of a certain volume

1. Introduction

Aerobiology is a sampling science. Indeed, its interdisciplinary location, at the junction of the studies

of air. We only estimate it. And as for any estimation, it is tainted with error. So approximation of some sort cannot be excluded from our investigation. Our knowledge of the aerobiological content is based on partial observation and it is only through inference that we can reach conclusions. These conclusions therefore have limits.

2. Sources of errors in aerobiology

“si l’on néglige cette précaution on obtient des chiffres sans grande signification et dont l’utilité est dès lors bien contestable.”

P. Miquel, 1883

We must first distinguish between bias and error. They both can be seen as a sort of imprecision, but they are fundamentally different. Errors are random, i.e. they arise by chance in our sampling protocol. Statistical theory can be used to evaluate errors, and to take them in consideration in any comparative analysis between samples. They only represent the uncertainties of the natural world. Conversely, bias does not appear by chance. It is usually unidirectional (systematic). Bias restricts any generalizations and comparisons, and therefore should be controlled at any cost. It arises as a result of a failure to recognize and prevent the effects of variables other than those which are under investigation.

Errors will also come from “instrument reading” errors. In the particular case of aerobiology, they are represented by the counting errors. These can be divided in two subsets: the human error of counting and identifying the pollen grains (the so called technician error or variability) and the statistical errors attached to the fact that, as for the air sampling, the slide reading is also a sampling, this time of a surface, and that the concentration obtained at the end is an approximation of the true value. These probably constitute the largest proportion of non-biased errors found in aerobiological results. The slide sampling will be the subject of this paper.

2.1. Error estimation in pollen analysis and aeropalynology

Reliability of estimation is also an important preoccupation in the theory of pollen analysis. According to Rull (1987), “pollen counting is the last step of a chain of successive approximations, where each stage represents a subsampling and produces a decrease in

the accuracy of the ultimate representation.” Certain statistical conditions must be satisfied for an adequate counting, and these were intensively studied in palynology from Barkeley (1934) to Regal and Cushing (1979). Reliability of concentrations calculations and estimation of the diversity required counts of 300–400 (Rull, op. cit.). The statistical determination of contemporary (surface) pollen counts requires 20 samples and counts of 250 according to Hill (1996).

In aerobiology, Pedersen and Moseholm (1993) have studied what they called the precision of the daily pollen count. However, their real interest was in the sources of variations external to the actual count, such as counters or trap variations, for which they found a relative uncertainty greater than 50%. These authors did not seem to realize that their variance component variations were only valid for a 12 transverse lines reading protocol. Kapyła and Penttinen (1981) have evaluated different counting procedures. These authors did not provide population count total for all their samples, which makes it difficult to reach precision estimation. However, they found a significant difference between two methods of estimation (transverse vs random fields). Tormo Molina et al. (1996) have also studied the actual slide sampling. They found a significant difference (of 7%) between 4 longitudinal traverses, and a stabilization of diversity after a count of 1000.

2.2. Error perception by aerobiologists

The percentage of error that is acceptable in aerobiology, like in any other science, is not a statistical decision. It is a personal/collective decision, related to significance. It is we, as aerobiologists, who can, and have, to decide what precision we need, and by consequence, what error will be acceptable. In order to have a gross idea of the percentage of error that seemed acceptable to professional aerobiologists, a small (not random) survey was done, to collect spontaneous opinions on the subject.

Results are presented in Figure 1. The distribution of answers is almost of normal kurtosis, but is skewed to the right. Average (mean) acceptable error was 10%, but the most frequent answer (mode) was 5%. The range extended to 30%, but the variance was small (0.006). There is therefore an expectation of high precision from aerobiologists, as regards to airborne pollen counts.

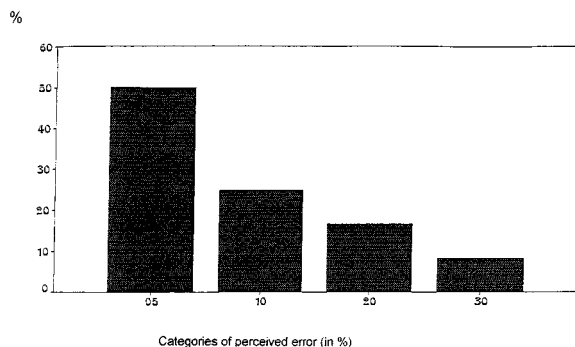


Figure 1. Pollen slide count error (in % categories) as perceived by a group of aerobiologists.

3. Methodology

In order to know the precision of any estimation (ϵ) arising from sampling, the population (the statistical universe ν) must be known. The sampling error will be the difference between ϵ and ν , while the estimation error will be related to σ_x . We therefore proceeded to the counting of the exact total of pollen found on 48×14 mm sampled Melinex tapes, for 12 taxa (*Acer*, *Ambrosia*, *Artemisia*, *Betula*, *Chenopodiaceae-Amaranthus* (Cheno.-Am.), *Cupressaceae*, *Fraxinus*, *Poaceae*, *Pinus*, *Plantago*, *Quercus*, and *Urticaceae*), and the same procedure was repeated 3 times per taxa for a total of 36 analysed slides. In addition, the exact location of microscope fields of each pollen grain counted was recorded. For our Wild-Leitz Laborlux microscope, there were 3120 fields at $400\times$, divided into 104 vertical transects and 30 horizontal lines. A single analyst has done the whole counting.

The records of the exact location of pollen grains on each slide were then transferred in an *Excel*TM worksheet, and macro procedures were built to subsample these 'populations'. Three types of reading protocols were studied: horizontal lines (longitudinal in regard to time), vertical lines (transverse in regard to time) and random fields (without any time constraint). A certain number of lines (longitudinal reading), columns (transverse reading) or random fields were then 'read' (Tables 1, 2, and 3). Longitudinally, from 1 to 20 lines were read, transversely, from 1 to 48 columns, and at random, from 90 to 2340 fields (Figure 2). Lines and columns were chosen in each protocol so that equal spacing between them was respected. Accordingly, from 3 to 66.6%, from 0.96 to 46.15%, and from 3 to 75% respectively of the slides were read. The macro procedures then summed the number of pollen grains that were read and calculated

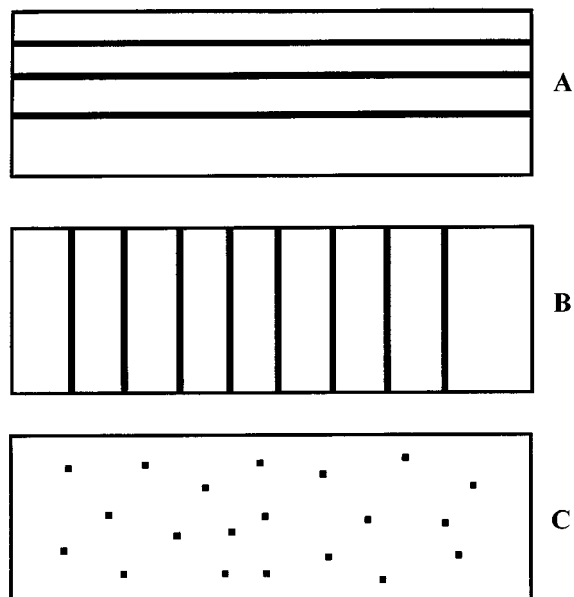


Figure 2. Slide sampling methodologies used: (A) longitudinal (horizontal) fields. In this example, 3 lines are represented. At 104 fields per line, this amounts to sampling 10% of the entire slide. (B) transverse (vertical) fields. In this example, 8 lines are represented. At 30 fields per line, this amounts to sampling 7.6% of the entire slide. (C) random fields. In this example, 20 cells are represented. At 3120 fields per slide, this amounts to sampling 0.6% of the entire slide.

the estimated total pollen concentration of the entire slide.

The sampling units of slides in aerobiology are microscope fields. The estimation error can therefore be calculated in the same manner whatever is the sampling protocol, because all reading procedures can be summarized as a number of microscope fields. In this manner, the field to field variation becomes a variance. Still, not all calculated variances will be equivalent. Only one procedure (randomization) sticks to the random field selection assumption. Part of the line to line variation will include the lateral difference in efficiency of the sampler (Tormo Molina et al., 1996), and most of the column to column variation will measure the magnitude of the diurnal variation of pollen emission. Nevertheless, as the number of selected lines or columns increases, a larger part of the variance can be ascribable to the normal variance between randomized elements. (This can be seen by looking at Figures 3 and 4, where the progression of the confidence interval is as expected under normal conditions.) An alternate solution would have been to calculate the variance between sets of lines (or columns). But, the spacing between them was not

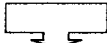
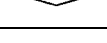
Table 1. Sampling protocol for longitudinal slide counting.

Number of lines	Sampled lines number	Number of cells	Observed %
1	15	104	3.33
2	10, 20	208	6.67
3	8, 15, 23	312	10
4	6, 12, 18, 24	416	13.33
5	5, 10, 15, 20, 25	520	16.67
6	4, 9, 13, 17, 21, 26	624	20
7	4, 8, 11, 15, 19, 23, 26	728	23.33
8	3, 7, 10, 13, 17, 20, 23, 27	832	26.67
9	3, 6, 9, 12, 15, 18, 21, 24, 27	936	30
10	3, 6, 8, 11, 14, 16, 19, 22, 25, 27	1040	33.33
14	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28	1456	46.67
20	1, 3, 4, 6, 7, 9, 10, 11, 13, 14, 16, 17, 19, 20, 21, 23, 26, 27, 29	2080	66.67

Table 2. Sampling protocol for transverse slide counting.

Number of columns	Sampled columns number	Number of Cells	Observed %
1	52	30	0.96
2	26, 78	60	1.92
3	17, 52, 87	90	2.88
4	13, 39, 65, 91	120	3.85
6	9, 26, 43, 61, 78, 95	180	5.77
8	7, 20, 33, 46, 69, 72, 85, 98	240	7.69
12	4, 13, 22, 30, 39, 48, 57, 65, 65, 74, 82, 91, 100	360	11.54
24	...	720	23.08
48	...	1440	46.15

Table 3. Sampling protocol for random slide counting.

Number of Cells	Procedure	Observed %
90		3
160		5
310	BASIC code for cells number generation	10
470	using a random numbers table	15
620	without replacement	20
780		25
870	generated cells number list	28
1030		33
1250	a single matrix is used for data pick-up	40
1560	for all species and slide specimens	50
2060		66
2340		75

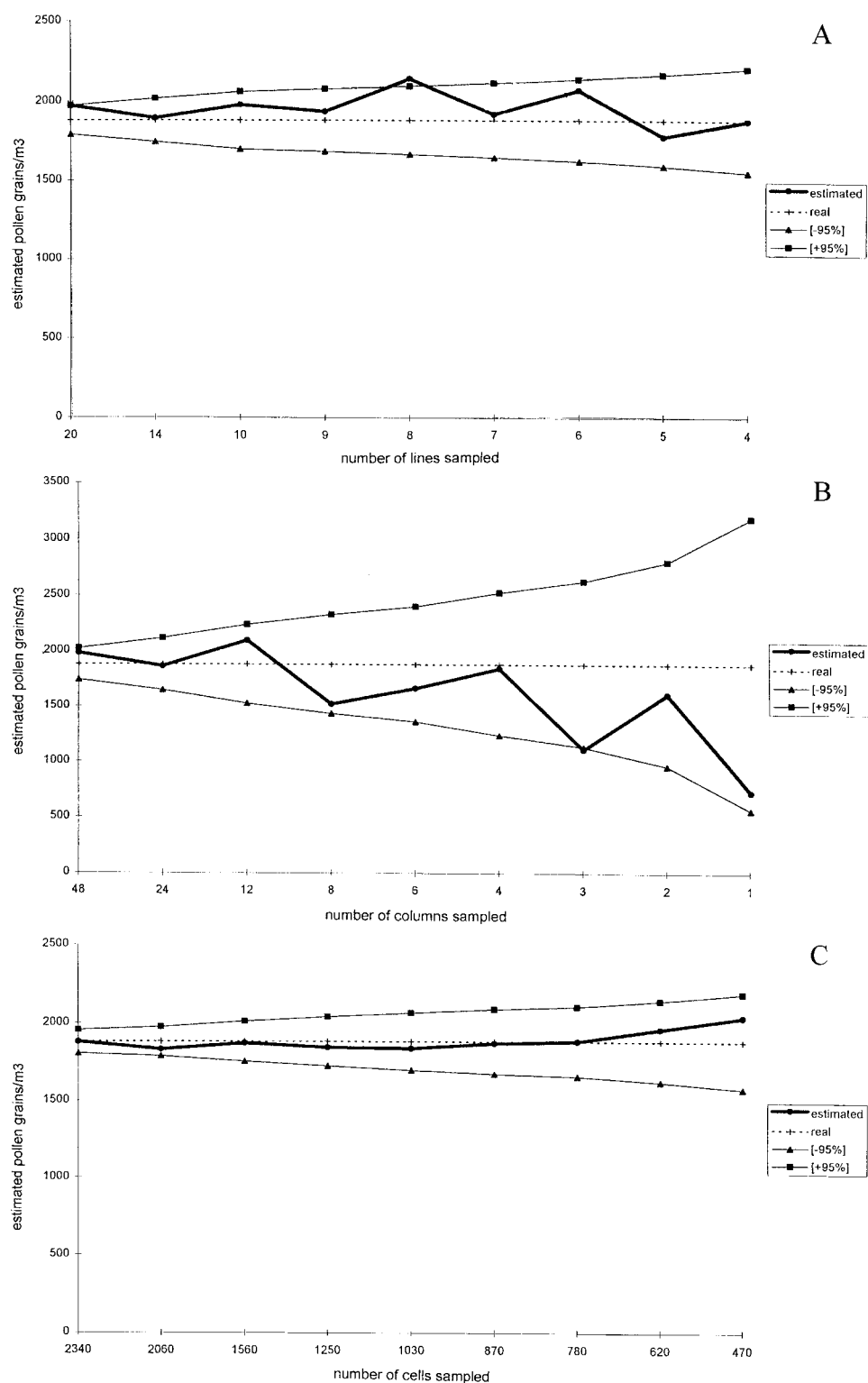


Figure 3. *Betula* pollen slide sampling error. Represented in each diagram are the estimated count (circles), the real count (dotted line), and the lower (triangles) and upper (squares) limits of the 95% confidence interval. (A) *Betula* longitudinal counts, (B) *Betula* transverse counts, (C) *Betula* random counts.

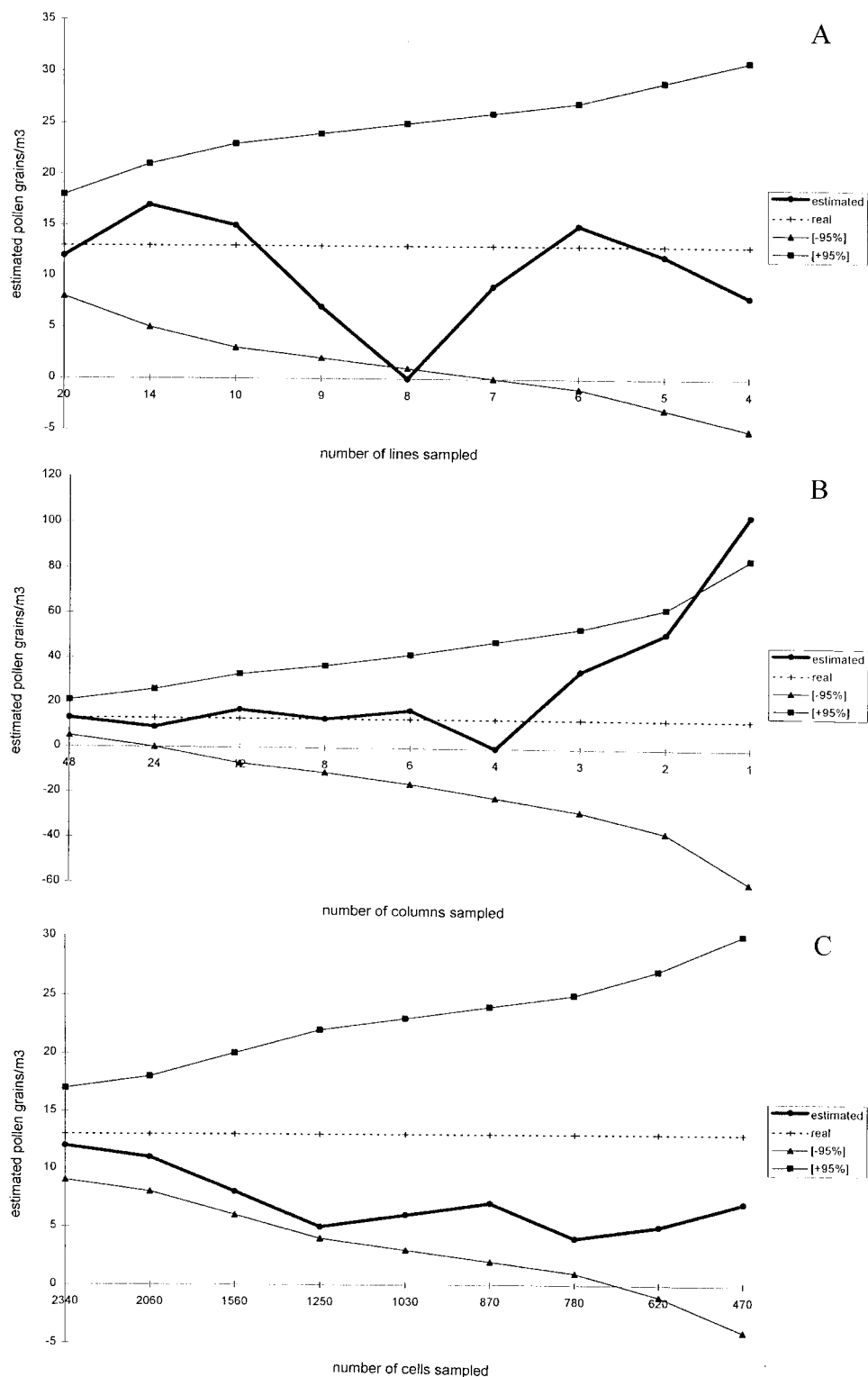


Figure 4. *Urticaceae* pollen slide sampling error. Represented in each diagram are the estimated count (circles), the real count (dotted line), and the lower (triangles) and upper (squares) limits of the 95% confidence interval. (A) *Urticaceae* longitudinal counts, (B) *Urticaceae* transverse counts, (C) *Urticaceae* random counts.

randomized but systematic, so the same problem arises again.

Knowing the sampling effort (the ratio between the number of fields read (n) and the total number of fields (N)), this particularity allows us to calculate the estimation error. The estimation error, for a sample, can be written as a confidence interval, under the form of:

$$X \pm Z_{\alpha/2} \sigma_x$$

where X represents the estimated concentration, Z the area (or probability) under the Normal curve, expressed in units of standard deviations, α the error, and therefore $(1 - \alpha)$ the confidence level in %, and σ_x the sample standard error. For our study, we have elected a 95% confidence percentage, and therefore the interval becomes, for a population:

$$X \pm 1.96 S_{qRt}(1 - n/N)(\sigma_{\frac{1}{n}}^2)$$

Since the calculation of this error takes into consideration the sampling effort (n/N) and the sample size (n), the confidence interval will decrease in width with any increase of the counting effort. The sampling effort is taken into consideration as an exhaustivity factor $(1 - n/N)$ because our sampling strategy is without replacement. This estimation error allows the comparison of results whatever is the sampling method used.

4. Results

“Il existe, dans le cas qui nous occupe, une limite d’approximation et un nombre de chiffres négligeables. Dans mes analyses de poussières, cette approximation varie de 1/10 à 1/15 du nombre moyen de bactéries”.

P. Miquel (op. cit.)

Results of counts corresponding to different protocols and to different sampling efforts can be represented as a curve on a diagram, where we find, on the ordinate, the estimated total (ϵ) of pollen grains on the slide; and, on the abscissa, the decreasing number of subsamples. Also represented, as a straight horizontal line with a null slope, is the real pollen total (ν), obtained by counting the whole slide; and the 95% confidence interval, as a cone of increasing width with diminishing sampling size. Please note that it is the confidence interval for the total slide count, not the confidence interval that would bound the estimated count at each point that is presented. Examples of

these diagrams are presented in Figure 3 (A–C) and 4 (A–C). *Betula* (Figure 3) and *Urticaceae* (Figure 4) were chosen to represent, respectively, examples of large pollen producer (average daily concentration of 1883 pollen grains m^{-3}) and low pollen producer (average daily concentration of 13 pollen grains m^{-3}).

Our precise pollen counts varied from 9 pollen grains per slide (*Chenopodiaceae-Amaranthus*) to 8396 pollen grains/slide (*Betula*) for an average of 710. Although there were only 3 replicates per taxa, the coefficients of variation were acceptable (mean of 0.54), and varied from 0.03 (*Plantago*) to 1.13 for *Betula*, the only value above 1.0 (Table 4). The estimation error, expressed in the form of a confidence interval, has high graphical value, but it is difficult to comprehend, unless it is expressed in relative values. This is why we have chosen to express it as percent of the real pollen total per slide. In this form, a mean error value of 98% was obtained for all our sampling (36 slides). Therefore, on average, the 95% confidence interval will have the same width as the pollen count. However, there was a large variation between slides. The minimum value was of 6% (*Betula*) and the maximum of 703% (*Cheno.-Am.*).

There was always a general trend of decreasing estimation error with increasing percentage of the slide read. When there was a departure from this trend, it was never for more than two values, and these exceptions were normally restricted to the first steps of the chain of sampling, i.e. when only 1 or 2 lines were read. The random fields strategy generally presented smaller initial errors when only a small area of the slide was read (< 470 fields), but, on the other hand, many times, especially when the total pollen number was below 50, the estimation error could not be calculated, because no pollen was found with this protocol, and therefore the procedure ended with a division by 0. So on one hand it can be more efficient, but on another hand, it can also lead to a dead end.

If we take as an example the standard procedure of 4 longitudinal lines (13.3% of fields), the mean estimation error (i.e. $Z_{\alpha/2}\sigma_x$) was of 63.3%, and varied between 10.5% and 474%. It is only with counts above 1000 (i.e. atmospheric concentration above 520 m^{-3} , our correction factor having a value of 0.52 at 400 $\times$), that it averages 16.9%. We can compare this with the equivalent 12 transverse lines (11.5% of fields) which has given us a mean estimation error, again for counts above 1000, of 14.8%; and with either 310 random fields (10%) or 470 random fields (15%), which has given us respectively an estimation error of 19.1% and

Table 4. Population count, coefficient of variation (C.V.) and count estimation error for 3 protocols for each studied taxa. Average of 3 slides are presented.

Taxa	Sampling protocols		Mean count estimation error		
	Mean population count	C.V.	4 Longitudinal lines (416 cells or 13.3%)	12 Transverses (360 cells or 11.5%)	310 Random cells (10% of total cells)
<i>Acer</i>	649	0.71	15%	11%	31%
<i>Ambrosia</i>	1177	0.66	14%	3%	7%
<i>Artemisia</i>	32	0.42	10%	15%	div./0
<i>Betula</i>	3665	1.13	7%	6%	5%
Cheno-Am.	15	0.37	96%	37%	24%
Cupressaceae	384	0.40	18%	24%	12%
<i>Fraxinus</i>	242	0.42	28%	19%	15%
<i>Pinus</i>	817	0.60	6%	6%	18%
<i>Plantago</i>	21	0.03	33%	21%	7%
<i>Poaceae</i>	294	0.36	16%	14%	7%
<i>Quercus</i>	1240	0.59	16%	10%	12%
Urticaceae	52	0.76	18%	30%	11%

15.3%. Therefore it is only with counts above 1000 that we reach, whatever the sampling strategy, the precision achieved by Miquel 100 years ago.

The estimated count error (i.e. $|\epsilon - \nu|$) was, on average for longitudinal sampling, of 23%, and varied between 0% and 168%, i.e. up to above 1.5 times the real total. For our 432 different samplings of the longitudinal type, 50.6% have given a negative error value (underestimation of the total) and 49.4% a positive value (overestimation of the total). The count error seems therefore to be totally random. For standard procedures (12 transverses, 4 longitudinal lines or 10% of random fields), the estimated count error is presented for each species (Table 4). Although the estimation error takes into consideration the field to field variance, there was a significant Pearson's correlation coefficient between the estimation error and the sampling error of 0.724. The mean estimated value is therefore the most important factor in determining both types of error.

5. Discussion

“Cette précision tient du hasard, et l'on doit s'estimer heureux quand le second chiffre diffère seulement du premier d'une unité de l'ordre immédiatement inférieur”.

P. Miquel (op. cit.)

We have seen in the introduction that the perceived error by aerobiologists was on average 10%, and the

highest expectancy was 5%. The error found using the same standard protocol as those who have participated in our survey reveals a mean error of 23%. In fact, the count error never reaches, on average, 5%, even after 20 longitudinal lines (Figure 5). It is only after 10 lines (or 33.3% of the fields) that 10% is reached.

Moreover, the count error was highly correlated to the abundance of pollen grains on the slide, as it should be for any discrete sampling. The R^2 found between error% and sample size was 0.41. This relation was of course built around a negative slope (Figure 6). The relation was a negative exponential, and can be expressed as:

$$(\text{ERROR \%}) = 138.4 - 34.2 (\ln \text{COUNT}) + 2.3 (\ln \text{COUNT})^2.$$

The significance of the regression was high, with an α probability of 0.0002. Normally, with a 2 dimensional sampling, as it is the case for a slide sampling, there should not only be a correlation with abundance but also with the size of the particles. However, our selection of pollen types did not allow us to check this assumption.

Comparison of sampling strategies did not bring any significant difference between comparable slide sampled percentage (for example, the paired t statistic between 4 longitudinal and 12 transverse lines was of 1.331 associated to a p of 0.2100). However, on average, the random fields showed a smaller estimation error (13%) than the 12 transverses (16%), but the

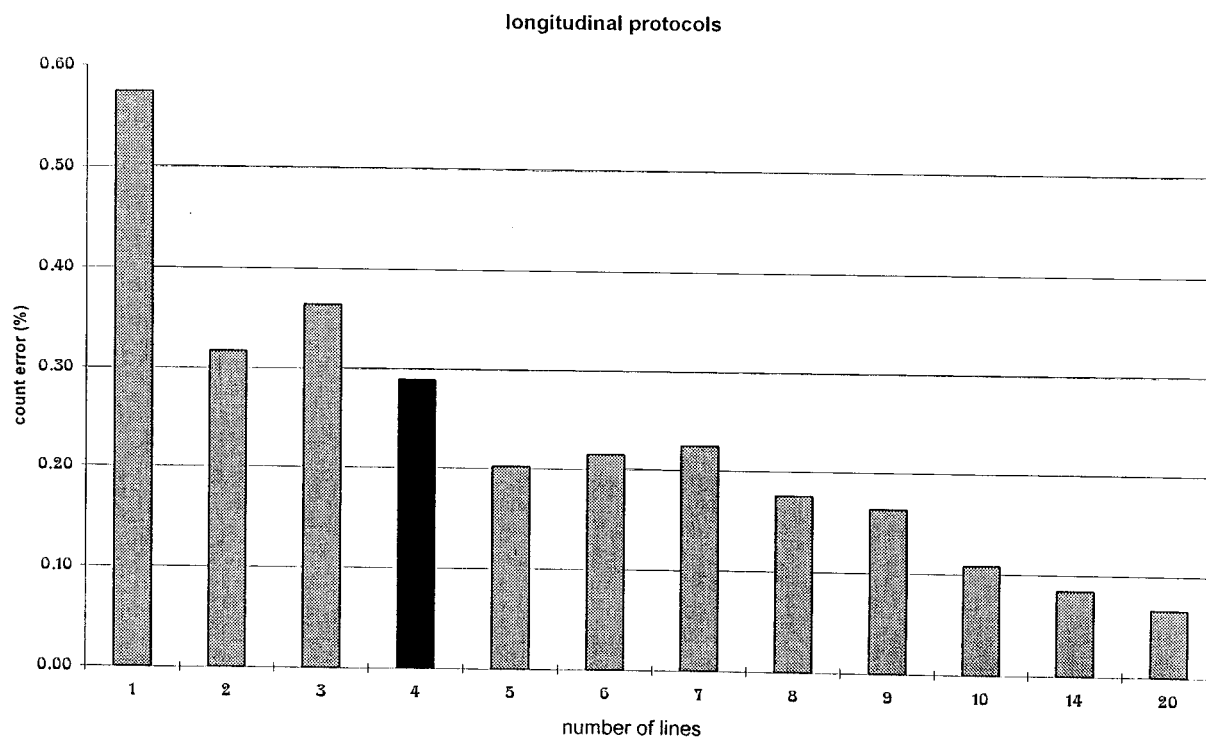


Figure 5. Bar chart of decreasing count error with the number of longitudinal lines read. Mean of all 36 studied slides. The standard 4 longitudinal lines is represented in black.

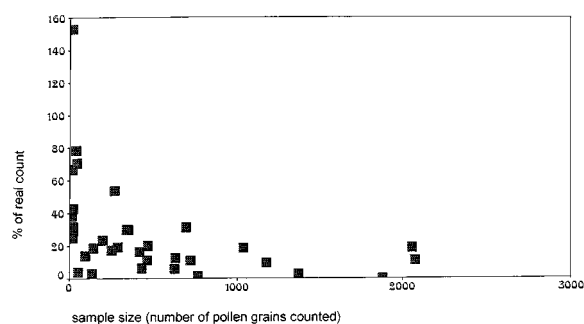


Figure 6. Scattergram of the slide total pollen sample size (X) and counting error (Y) expressed as % of the real count.

former procedure can sometimes lead to infinite error (division by zero). The 4 longitudinal lines procedure showed the largest average count estimation error (23%). This may indicate that the Hirst-type sampler is more prone to lateral variation of efficiency than expected, since, theoretically, complete longitudinal lines, should give better estimation of the slide total than transverses, which could miss pollination peaks, if they fall between 2 transverses.

6. Conclusion

statistics: n.pl. science of collecting and analysing significant numerical data.

The Oxford dictionary of current English.

Not all aerobiologists seem to realize that pollen slide counting is also a sampling, and that, by consequence, the pollen count obtained is impregnated with imprecision. In this paper we report a sampling error six times larger than the modal error expected by aerobiologists. To obtain the average expected error we would have to count a third of a daily slide. Only for concentrations above 500 pollen grains m^{-3} (counts above 1000 per slide at 400 $\times$) did the sampling error approach the expected one. The count error means that even if we could control all other extraneous factors, there will always be imprecision linked to the airborne pollen count, unless we count the whole slide. By chance alone, we could have obtained a count that would differ by some 30%.

There is no point in missing useful hourly information by counting pollen on randomly selected fields. The increase in precision is not significant, and although at low sampling efforts the mean error

was lower, there was the risk of not collecting at all individuals of rare species. The 12 hourly transverse sampling has given the lowest average error (as compared to the other standard, the 4 longitudinal lines sampling). However, these differences were not statistically significant.

We must agree on an acceptable sampling error (or precision) before we can determine an acceptable slide reading protocol (or sampling effort). Moreover, this decision must also be based on the expected use of data we are collecting. For allergy risks forecasts, an order of magnitude notation, such as the one usually given for standard microbiological water quality, would probably be sufficient. For regression analysis purposes, using meteorological data, both independent and dependent data sets need comparable errors. For biogeographical or dispersal studies, the highest possible precision should be our objective.

Pollen slide counting is a complex statistical entity. Although a relatively large error is linked to the pollen counts, it is probably not significant in terms of the interpretation we can (or should reach). It should however be pointed out that our sampling error will always be dependent on the size of the slide total and, therefore, no counting method can be satisfactory for all taxa all the time.

Acknowledgements

This work was started after a friendly discussion with Siegfried Jäger of the H.N.O. Klinik in Vienna. Many thanks for the impetus he has given to this project.

References

- Barkeley F.: 1934, The statistical theory of pollen analysis. *Ecology* **15**, 283–289.
- Hill T.R.: 1996, Statistical determination of sample size and contemporary pollen counts, Natal Drakensberg, South Africa. *Grana* **35**, 119–124.
- Kapyla and Penttinen: 1981, An evaluation of the microscopic counting methods of the tape in Hirst-Burkard pollen and spore trap. *Grana* **20**, 131–141.
- Miquel P.: 1883, *Les Organismes Vivants de l'Atmosphère*. Gauthier-Villars, Paris, 310 pp.
- Pedersen B.V. and Moseholm L.: 1993, Precision of the daily pollen count. Identifying sources of variation using variance component models. *Aerobiologia* **9**, 15–26.
- Regal R.R. and Cushing E.J.: 1979, Confidence intervals for absolute pollen counts. *Biometrics* **35**(3), 557–565.
- Rull V.: 1987, A note on pollen counting in palaeocology. *Pollen et Spores* **24**(4), 471–480.
- Tormo Molina R., Munoz Rodriguez A. and Silva Palacios I.: 1996, Sampling in aerobiology. Differences between traverses along the length of the slide in Hirst spore trap. *Aerobiologia* **12**, 161–166.