



Precision and accuracy of modal analysis methods for clastic deposits and rocks: A statistical and numerical modeling approach

Pierre-Simon Ross*, Bernard Giroux*, and Benjamin Latutrie†,*

Institut National de la Recherche Scientifique, 490 de la Couronne, Québec, Québec G1K 9A9, Canada

ABSTRACT

Quantifying the proportions of certain components in rocks and deposits (modal analysis or componentry) is important in earth sciences. Relevant methods for cross-sections (two-dimensional exposures) of clastic rocks include point counts or line counts. The accuracy of these methods has been supposed to be good in the literature but not necessarily verified empirically. Natural materials are inappropriate for assessing accuracy because the true proportions of each component are unknown. The precision of modal analysis methods has traditionally been evaluated from statistical models (primarily the normal approximation to the binomial distribution) but again rarely verified in practice because it is also extremely difficult to obtain different slices through the same material at outcrop scale. Here we create a set of numerical models of red and blue spheres with different proportions and sizes and cut 60 slices through the models, on which we perform point counts and line counts. We show that both of these methods are indeed able to retrieve the correct volumetric proportions of components, on average, when enough fragments are counted or intersected. As already known, precision is controlled by component abundance and the number of points counted or clasts intersected. However, we show that other important factors include differences between slices, which are relevant for our unequal-size models, and the proportion of voids, matrix, and/or cement in the rock. We present empirical precision charts for clast counts and line counts

based on our models and make recommendations for future field studies.

INTRODUCTION

Importance and Basic Principles of Modal Analysis

Quantifying the composition of clastic rocks and deposits is of fundamental importance in earth sciences. For example, the mineralogical composition of a sandstone can indicate its provenance (Dickinson and Suczek, 1979; Ingersoll et al., 1984; Weltje, 2002), a microfossil assemblage can be used for paleoenvironmental interpretations (Patterson and Fishbein, 1989; Fatela and Taborda, 2002), and the composition of a pyroclastic rock or deposit can help in deciphering various eruptive processes (Houghton and Smith, 1993; Latutrie and Ross, 2020a, 2020b). Various terms are used—for example, “componentry analysis” in volcanology, “sandstone composition analysis” or “modal analysis” in sedimentology. “Modal analysis” will be employed in the rest of this article for simplicity.

The ultimate aim of modal analysis measurements is to reliably determine the proportions of various types of components such as clasts, minerals, or fossils in a volume of rock or a deposit, perhaps an entire lithofacies. It is not possible to characterize all of the grains individually, so instead, one or several smaller—hopefully representative—domains are selected, such as a 1 m² surface in a rock face or outcrop, a hand-sized sample, a

certain size fraction from a sieved deposit, or a thin section for analysis. Even then, there are generally still too many grains to count them all, so modal analysis methods such as point counting or line counting subsample the chosen material for consideration of only a predetermined number of grains (Delesse, 1848; Rosiwal, 1898; Fleet, 1926; Chayes, 1944, 1945; van der Plas and Tobi, 1965; Patterson and Fishbein, 1989). If a cross-section (two-dimensional [2-D] exposure) through the material is involved, such as with a rock face or a thin section, this implies a change from three dimensions (3-D; a volume) to 2-D (a surface). Line counting further reduces the dimensions to one, and point counting to zero. The Delesse principle states that the volume fraction of a component will be the same as the area fraction of that component on a representative slice, and so on down to the proportion of points counted (Higgins, 2006), but this applies only if enough particles are considered and if the material is homogeneous.

Problem Statement

Getting a reliable modal analysis result therefore means counting or measuring enough particles, but how much is enough to answer the scientific question being asked? Besides, is the studied slice representative in the first place? These questions have been considered since at least the 1940s, but at the moment, various research groups use different count numbers, partly because of different applications but also because of different traditions within research fields or research groups. Some arbitrary numbers are sometimes being employed (e.g., 300 clasts or points; Ingersoll et al., 1984; Patterson and Fishbein, 1989).

Pierre-Simon Ross <https://orcid.org/0000-0002-5302-698X>

*E-mails: rossps@ete.inrs.ca, Bernard.Giroux@ete.inrs.ca, Benjamin.LATUTRIE@uca.fr

†Now at Laboratoire Magmas et Volcans, Université Clermont Auvergne, 63000 Clermont-Ferrand, France

Aims of the Current Study

A reexamination of these questions is due, and our contribution uses both a statistical approach, as in many previous studies (with some improvements), and a novel numerical modeling approach that can simultaneously take into account multiple sources of variation. We consider two modal analysis methods involving cross-sections: point counts and line counts.

The reliability or uncertainty of modal analysis measurements has two independent dimensions: precision and accuracy. These dimensions need to be better studied for modal analysis methods. Previous studies on this topic have been based largely on statistics (e.g., van der Plas and Tobi, 1965; Howarth, 1998), but such an approach does not typically address accuracy. Statistical approaches can predict some sources of variability between counts, but might not include all sources of uncertainty that contribute to precision, such as:

- Is the cross-section or sample representative?
- What is the effect of particles of unequal sizes?
- What is the impact of the matrix and cement?

In this paper, we construct a set of numerical models of red and blue spheres with different proportions and sizes. The models serve as a basis for point counts and line counts using different numbers of points and lines. We show that these methods can indeed be accurate, and we create empirical precision charts for point counts and line counts. Recommendations for carefully designing modal analysis studies are proposed based on those charts.

■ BACKGROUND

Point Counts

Point counting is normally done with a regular grid, and particles on this grid are successively assigned by an expert to component bins until the predetermined total number of points is reached. Point counts provide the modal fraction of the components within a sample (e.g., van der Plas and Tobi, 1965; Underwood, 1970; Patterson and Fishbein,

1989) without needing stereological conversion, although some authors have questioned whether point counting is appropriate for components of unequal sizes (Howard, 1993).

Point counts can be performed in the field (Blatt, 1992; Ross and White, 2006, 2012; Carrapa and DeCelles, 2008; Latutrie and Ross, 2020a, 2020b). Field point counts on coarse clastic rocks are commonly acquired using string nets of 1 m² (Fig. 1B; Ross and White, 2006, 2012; Latutrie and Ross, 2020a, 2020b). Different meshes yield different number of points, for example 100 points for a 10 cm mesh or 400 points for a 5 cm mesh over 1 m². In many studies, each component greater than or equal to 4 mm (in some cases 2 mm in older studies) found below a point in the grid is counted into bins (component classes), whereas smaller components are called “matrix.” There may also be a “cement” category.

Petrographic point counts can be done using either images and specialized software (Fig. 1C) or mechanical point-counting devices attached to petrographic microscopes (Shand, 1916; Chayes, 1956; Galehouse, 1971; Douce and Johnston, 1991; Roduit, 2007; van Otterloo et al., 2013; Stamper et al., 2014; Bélanger and Ross, 2018). Petrographic point counts are classically used to quantify components smaller than 4 mm or 2 mm, so are a natural complement to field point counts.

Point counts are used in volcanology (e.g., Mastin et al., 2004; Ross and White, 2006, 2012; Bélanger and Ross, 2018; Latutrie and Ross, 2020a, 2020b), petrology (e.g., Giddings, 1986; Douce and Johnston, 1991; Stamper et al., 2014), paleontology (e.g., Retallack, 1994; Baarli et al., 2014), palynology and other paleo-environmental studies (e.g., Clark, 1982; Patterson et al., 1987; Magyar et al., 2010), as well as in sedimentology and soil science (e.g., Dickinson and Suczek, 1979; Weltje, 2002; Carrapa and DeCelles, 2008; McKinley et al., 2012).

Line Counts

Line counts as defined here have been used in volcanology (e.g., Lefebvre, 2013; Lefebvre et al., 2013; Bélanger and Ross, 2018; Latutrie and Ross, 2020b) and other disciplines (e.g., Wentworth, 1923;

Galehouse, 1971; Campbell and Galehouse, 1991). In volcanology, line counts have recently been employed mainly as a field method on horizontal to vertical outcrops of volcanoclastic rocks (Figs. 1D, 1E; Lefebvre et al., 2013; Bélanger and Ross, 2018; Latutrie and Ross, 2020b). A tape measure is placed on the outcrop, typically across 1 m, and each clast longer than a minimum length is classified and measured. Specifically, the intersection length of each fragment is measured. Some workers have used only one 1 m line (Lefebvre et al., 2013; Bélanger and Ross, 2018) whereas others recently summed up three 1 m lines separated vertically by 50 cm each to “cover” an area of 1 m³ (Latutrie and Ross, 2020b). In theory, it would be possible to add more lines within the square meter and/or use longer lines to increase the data quality. The sum of the intersection lengths of each component divided by the total length of the tape gives the proportion of this component. As with field point counts, the minimum size of clasts is commonly taken as 4 mm because smaller components are difficult to classify visually (e.g., Lefebvre et al., 2013; Bélanger and Ross, 2018; Latutrie and Ross, 2020b). Clasts smaller than 4 mm and cement are not explicitly measured, so they fall in a “matrix + cement” category, which brings the total of each count to 100%. In field line counts, the lines are generally parallel to each other, placed in one orientation only. This may introduce biases in the data if the rock has clasts with a preferred orientation or if the rock is anisotropic in some other way. However, this issue is beyond the scope of our study.

Line counts—measuring intersection lengths of fragments or minerals along lines—can also be done on thin sections, and this is known as the Rosiwal or Rosiwal-Shand method (Rosiwal, 1898). However, use of the Rosiwal-Shand method declined in the mid-20th century, when it was progressively replaced by petrographic point counting (Galehouse, 1971).

To our knowledge, no study yet empirically confirms that line counts based on intersection lengths can accurately quantify a volumetric proportion (modal fraction) without corrections. It is also unclear what the precision of this method is, depending on the number of lines counted (or the number of objects intersected) per site or sample.

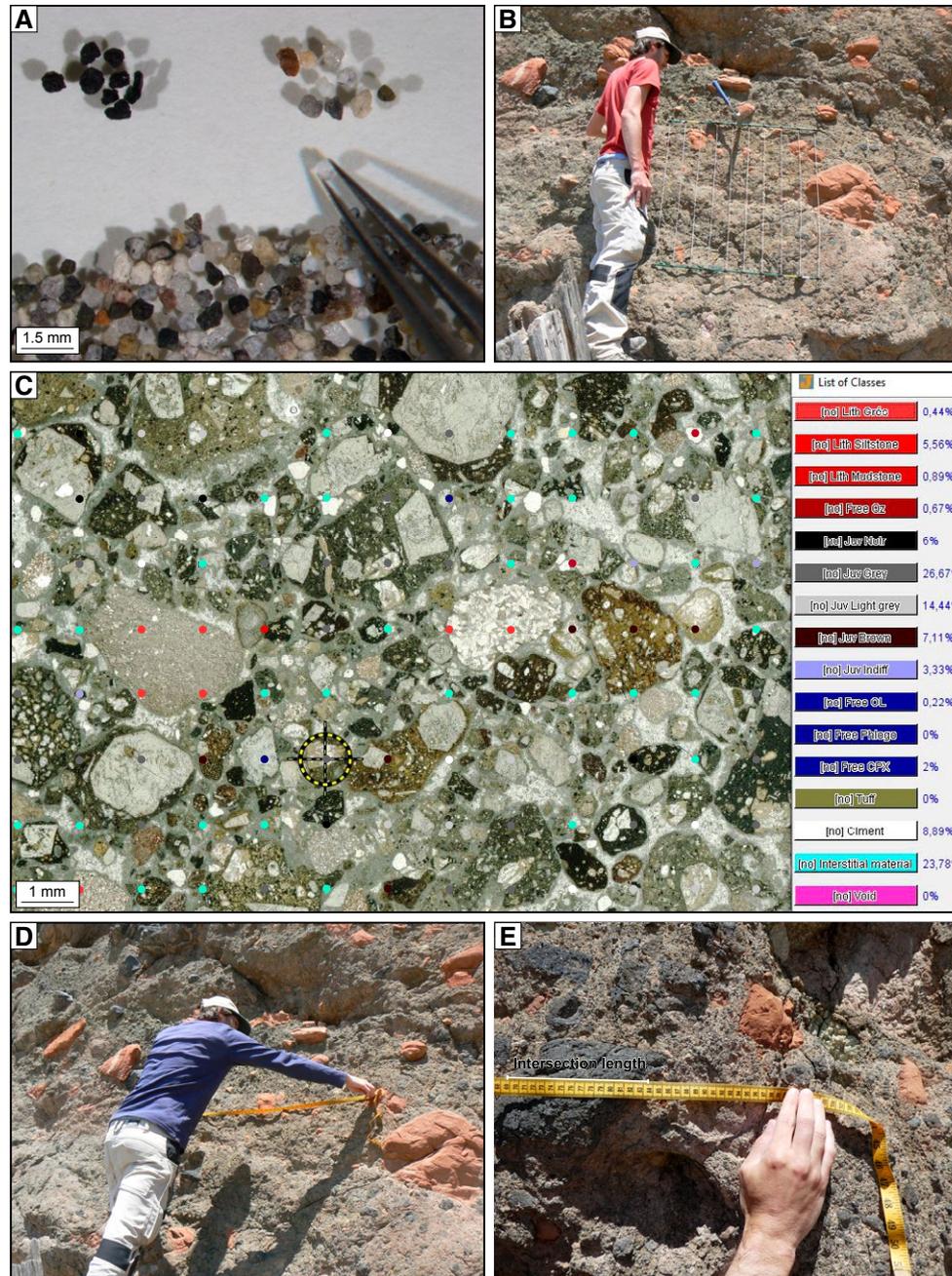


Figure 1. Examples of modal analysis methods used in the field and in the laboratory. (A) Particle counting under a binocular microscope considering two categories of grains, juvenile (left) and lithic (right) (1ϕ particles from a phreatomagmatic deposit of the Ukinrek east maar volcano, Alaska, USA; image courtesy of Pier Paolo Comida). (B) Field point count using a grid of 1 m^2 with 10 cm spacing between strings and 10 cm between black dots on the strings (100 points). (C) Petrographic point count on a high-definition scan of a thin section with the JMicroVision 1.2.7 software (Roduit, 2007) (colored point locations have been enlarged for better visibility). (D–E) Field line counts with 1-m-long measuring tape. The intersection length of each clast along the tape is recorded.

Other Methods of Modal Analysis

There is a third major modal analysis method in use in various research fields, i.e., counting loose particles (e.g., Buzas, 1990; Houghton and Smith, 1993; Rebertus and Buol, 1989; Fatela and Taborda, 2002; Go et al., 2017), commonly under the binocular microscope (Fig. 1A) and commonly after sieving, but our numerical models are not suited to study this application.

A fourth class of methods is component area measures, involving image analysis, but this requires that components be easily distinguishable based on color or grayscale or the use of some other automated segmentation methods. Image segmentation is commonly difficult to perform on complex natural materials. In this paper, we use component area measures only for comparison with point counting or line counting results.

Accuracy and Precision

When making scientific measurements, accuracy is related to systematic errors, whereas precision is related to random errors (Taylor, 1997). Valid measurements are both accurate and precise. In simple terms, accuracy is the ability of a method to give correct results on average (Fig. 2A). It can be calculated

as the difference between the measured value and the true value, commonly normalized to the true value, i.e., the relative difference, in percent. The more accurate a series of measurements is, the smaller the systematic bias. Modal analysis methods are generally implied to be accurate in the literature, but this has been little studied, and we specifically test this assumption below for point counts and line counts.

Precision is the variability between different measurements on the same material. Precision is commonly taken as one or two absolute or relative standard deviations of a series of measurements (Taylor, 1997). Relative standard deviation is simply the standard deviation divided by the mean measurement. In general, if data are normally distributed, then about two-thirds of measurements (68.27%) should fall within one standard deviation on either side of the mean, and 95.45% should fall within two standard deviations (e.g., Davis, 2002; Fig. 2A).

Counting Error as a Measure of Precision

In a modal analysis context, there are many reasons for different measurements on the same material to yield different results. For example, a thin section subject to point counting may not be representative of the specimen from which it was cut, i.e., the material is heterogeneous at this scale.

Yet traditionally, investigators have assumed that the thin section or area investigated is representative (e.g., van der Plas and Tobi, 1965) and have instead focused on the “counting error.” The latter is the variability of results due to the fact that modal analysis methods do not take into account all of the fragments or minerals in the thin section, rock surface, or size fraction but only a subset of them to save time. Intuitively, this counting error would be larger for small numbers of points counted and for low-abundance components. If certain assumptions are met, counting error can be predicted based on statistical theory and models. Supplemental File S1¹ reviews the approaches that have been taken so far in the literature to predict counting error, encompassing topics such as the binomial distribution and confidence intervals for a binomial proportion. The chart by van der Plas and Tobi (1965) is probably the best known (Fig. 2B).

Apart from issues reported in File S1, another potential shortcoming of the van der Plas and Tobi (1965) chart, when applied to point counts on cross-sections, is that it assumes that the cross-section being studied (e.g., a 1 m² area of a rock outcrop or a single thin section) is representative of the whole volume of material to be characterized. While that may be more or less correct for well-sorted homogeneous fine-grained sandstones, it may not be for, say, poorly sorted volcanoclastic rocks where

Geosphere

Precision and Accuracy of Modal Analysis Methods for Clastic Deposits and Rocks: a Statistical and Numerical Modeling Approach

Pierre-Simon Ross¹, Bernard Gosses¹, Benjamin Lattard^{1,2}

¹ Institut national de la recherche scientifique, 490 de la Concession, Québec (QC), G1K 9A9, Canada

² Now at Université Clermont Auvergne (UCA), Laboratoire Magmas et Volcans, 63000 Clermont-Ferrand, France

* Corresponding author: p.ross@nrc.ca

Supplemental File 1

A REVIEW OF SOME PREVIOUS STUDIES AND STATISTICAL CONCEPTS

The Binomial Distribution to Statistically Model Point Counts

Point counts have often been studied using statistics, but we have not seen an equivalent treatment for line counts, and we do not attempt one here. In the paper, we only use statistical approaches to study point counts.

The binomial distribution describes a Bernoulli process, in which there are two possible outcomes for each trial: success or failure. For example, we randomly draw balls from a bag which contains blue and red balls, and drawing a blue ball is a “success,” with probability p , which is to count the proportion of blue balls. The selected ball is placed back in the bag and the balls are mixed after each draw, so each draw is independent. The number of trials, N , is fixed and known in advance, and the number of successes in discrete, X : (i) a rock or deposit contains only two components, (ii) the point counting grid is larger than the largest component, and (iii) the rock or deposit is homogeneous. Then point counting is a Bernoulli sampling process (Nelson and Brockman, 1977; Weippl, 2002) and the binomial distribution applies¹.

The binomial distribution is a discrete function, unlike the normal distribution which is continuous. The binomial population mean is $\mu = Np$ and its population standard deviation is $\sigma = \sqrt{Np(1-p)}$, in terms of the number of successes (Taylor, 1997; Howarth, 1998). For example, suppose that we count $N = 400$ points in a homogeneous thin section which contains blue and red particles and the true proportion of blue particles is $p = 0.5$. If we count these 400 points once, we will find 4 blue particles, somewhere close to 200, but probably not

¹In some cases, the binomial distribution may not be a relevant statistical model for point counting, for example if the thin section, specimen or area under study is not randomly homogeneous, or if the observations are not statistically independent, or if the number of trials is not fixed. In such cases, the binomial distribution may not be applicable. However, the following discussion will assume that the binomial distribution applies, a generalized version of the binomial distribution, where more than two components are present in the multivariate distribution (Howarth, 1998; Weippl, 2002).

¹Supplemental Material. File 1: Review of some previous studies and statistical concepts. File 2: Figures for people with color blindness (color vision deficiency). Please visit <https://doi.org/10.1130/GEOS.S.14448111> to access the supplemental material, and contact editing@geosociety.org with any questions.

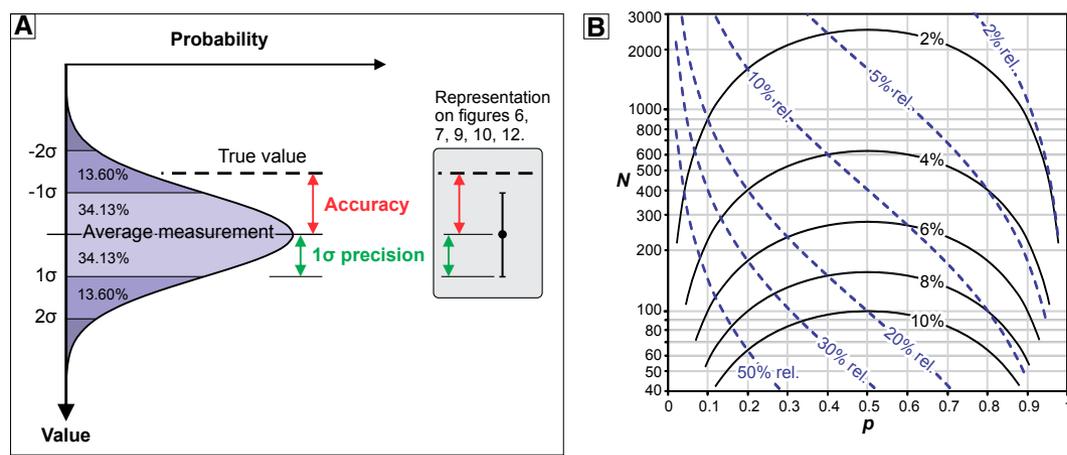


Figure 2. Illustration of some statistical concepts and previous “error charts”. (A) There are two types of measurement uncertainty: accuracy and precision. Accuracy is the ability of a method to provide the correct or accepted value, on average. Repeating measurements on the same material will yield variable results, as shown schematically by the bell curve. The relative standard deviation of this distribution of measurements is a measure of precision. The percentage values are the proportions of measurements that should fall in the shaded areas if the data are normally distributed. The way these concepts are illustrated in the rest of the paper, with a dashed line (true value), a dot (average measurement), and error bars, is shown on the right. (B) Point counting “error chart” similar to the one of van der Plas and Tobi (1965), plotting two-standard-deviation values (absolute or relative [rel.]) as a function of the number of clasts counted (N) and the proportion of the component of interest (p).

heterogeneity may be present. In other words, the counting error discussed above may not be the only source of variation between measurements that we wish to capture with precision.

Finally, the van der Plas and Tobi (1965) chart assumes that all counted points fall on relevant constituents rather than in voids, cement, or irresolvable material. If component proportions are recalculated on a 100% basis after exclusion of the void, cement, or irresolvable proportions, then fewer points are actually taken into account than the total number of points counted, and the van der Plas and Tobi (1965) chart applied to the total count would underestimate the variability of the results.

METHODS

Natural materials are inappropriate for verification of the accuracy of modal analysis methods because the “true” proportions of each component are not known. It is also extremely difficult to obtain different slices through the same material at outcrop scale to assess the slicing effect and check the precision of different methods. Therefore, we use numerical models.

3-D Model Creation

We build >1 m³ cubic packs of spheres of known characteristics. Sphere-pack models are generated with Yade software, an open-source framework for discrete numerical models based on the discrete element method (Šmilauer et al., 2015). Sphere packs are obtained after mimicking a gravitational depositional process. First, a low-density cloud of solid spheres is created inside a vertically elongated box, the initial position of the individual spheres being randomly determined. At this stage, the spheres can be thought to be “floating” in air. The spheres are then allowed to fall and interact with each other as well as with the faces of the box until static equilibrium is nearly reached (perfect equilibrium is never reached in practice because of finite precision computation). Linear elastic-plastic interaction with friction between the

bodies is allowed (Cundall and Strack, 1979), but the sphere and box material are defined as elastic so no permanent deformation occurs in the process. Ultimately, the spheres accumulated at the bottom of the box constitute a random pack much denser than the initial cloud. The size of the box and initial number of spheres are chosen so that the resulting pack is more or less cubic, with dimensions >1 m in all directions. The final 1 m³ pack is extracted from the center of the deposited pack to avoid any edge effects. The modeled volume is large enough to contain one extra sphere on every side of the 1 m³ pack. For example, for 20-cm-diameter spheres (10 cm radius), the box is 1.4 m across. The coordinates of the center and the radius of each sphere are stored in a Visualization Toolkit (VTK) file (Schroeder et al., 2006) for further processing and visualization.

In all cases, two populations of spheres are included in the models, i.e., red spheres and blue spheres, which represent different components. In the nine basic “equal-size” models, all spheres have a 1 cm radius (2 cm diameter) but the proportion of blue spheres varies from 0.1 to 0.9 (or 10% to 90%) in 0.1 (or 10%) increments (e.g., Fig. 3; File S2 [footnote 1]). We also have a 0.01 (or 1%) model to show the effect of very low proportions. It should be noted that due to the random nature of the model-building process, the proportions of blue spheres are not exactly 0.1, 0.2, etc., and vary by a small fraction of a percent from the intended proportion. For example, the 0.1 model actually contains a blue sphere proportion of 0.1003, as further explained below.

Components of equal sizes are a special case, so we also build unequal-sizes models. Four additional unequal-size scenarios are considered involving red spheres of 1 cm radius with blue spheres of increasing radii: 2 cm, 3 cm, 5 cm, and 10 cm (Fig. 4; File S2 [footnote 1]). This is done to check the influence of components of different average sizes on modal analysis results (a natural analogue would be a pyroclastic fall deposit with large pumice and much smaller dense lithics). For the 2-cm-, 3-cm-, and 5-cm-radii cases, the volumetric proportions of red and blue spheres vary from 0.1 to 0.9 in 0.1 increments; we do not have 0.01 models for

unequal sizes. With spheres of different radii, the procedure to create the initial “floating” cloud in Yade is based on the curve of the cumulative particle size distribution, i.e., the relative amount, by count, of particles present according to size. To build this curve, the volumetric proportion must be converted to a relative number of spheres for each component (numerical proportion). It can be shown, after some algebra, that the numerical proportion of blue spheres, n_b , is:

$$n_b = \frac{1}{1 + \frac{(1 - p_b) v_b}{p_b v_r}}, \quad (1)$$

where v_b and v_r are, respectively, the volume of a blue sphere and a red sphere, and p_b is the volumetric proportion of blue spheres. The numerical proportion of red spheres is then $n_r = 1 - n_b$.

Spheres are generated randomly, and their size is determined from the cumulative distribution curve. For each sphere, this is achieved by picking a value between 0 and 1, projecting this value horizontally from the ordinate axis to the curve, and then vertically from the curve to the corresponding size on the abscissa axis (Fig. 5A). It should be pointed out that for two-radii models, the theoretical cumulative distribution curve has two steps, each step located at the respective radius values. In practice, however, the steps must be approximated by sharp increases, as shown in Fig. 5A, to avoid an infinite slope. A consequence of the finite slope is that the actual radius of the spheres is allowed to deviate slightly from the target values. This is illustrated in Fig. 5B, which shows the radius histograms for all spheres in the case of 1 cm and 2 cm radii modeled in this study. It can be seen that the radius values are evenly distributed and very close to the target values (deviation of <0.5%). In all cases, the true value of the radius is stored and used for calculating the true proportions of each population (detailed below).

Although the cumulative distribution curve approach works well for spheres as large as 5 cm, we found that for the case of the 1 cm versus 10 cm radii, the actual proportions deviate significantly from the 0.1–0.9 targets, e.g., by up to 0.07. This is likely due to the large ratio v_b/v_r , which causes n_b

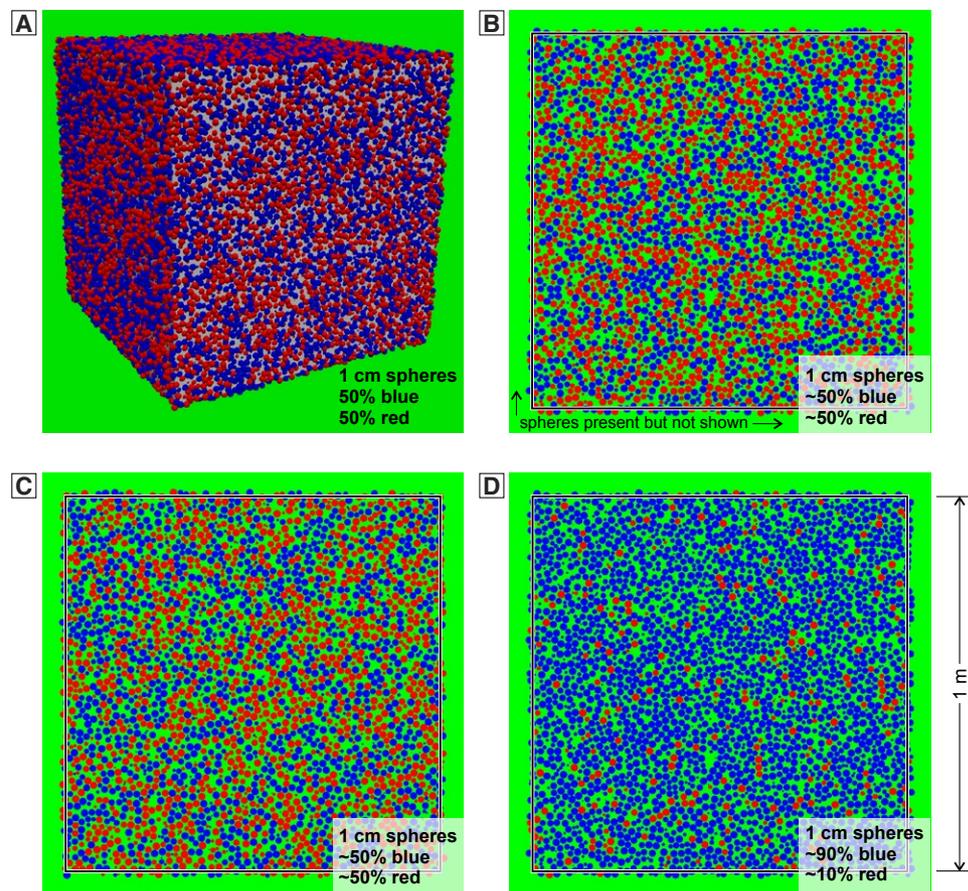


Figure 3. Equal-size sphere models with red and blue spheres of 1 cm radius. (A) Three-dimensional (3-D) model of 1 m³ comprising 50% red spheres and 50% blue spheres. (B–C) Horizontal and vertical slices through the center of the same 3-D model. (D) Horizontal slice through the middle of a different model, comprising 10% red spheres and 90% blue spheres. A similar figure is available in Supplemental File S2 (footnote 1) with gray backgrounds (for color-blind people).

to be very close to 0 and thus n_i to be close to 1. Given the dimensions of the problem, the overall number of blue spheres is too small for the population to statistically reflect the distribution curve, and numerous tests must be performed to obtain a satisfying value. Because the 1 cm versus 10 cm case is also quite computationally demanding (with each simulation running for more than three days on a multi-CPU server), only a single case is

presented, in which the red and blue spheres are of equal volumetric proportion.

Calculating the True Proportions of Each Component

For each 3-D model, the exact volumetric proportions of the two populations of spheres inside

the final 1 m³ cube must be computed to evaluate the accuracy of the point count and line count methods. Spheres completely inside the 1 m³ cube are separated by color, and the volume v of each sphere is added for each group, e.g., the total volume of blue spheres is $V_b = \sum v_b$. The proportion of blue spheres is then $p_b = \frac{V_b}{V_b + V_r}$. Spheres crossing

the faces, edges, and corners of the cube must be handled carefully to account only for the portion inside the cube. This is achieved by processing each sphere individually. If an intersection with a face, edge, or corner is detected, a test is performed to see if the intersection is happening along a single plane, leading to two spherical caps. In such a case, the analytical expression of the volume of the cap inside the cube is used. If the sphere is close to an edge or a corner and is intersected by two or three planes, the sphere is discretized with a tetrahedral mesh and the intersection of the sphere with the cube is computed numerically with the *PyMesh* library (Zhou, 2018). Meshes are generated with refinement order option equal to 5, leading to densely meshed objects. An example of a sphere close to an edge of the cube is illustrated in Figure 5C. The volume of the resulting mesh is then computed and added to the total volume of its corresponding group. Obtaining the total volume of blue and red spheres allows computing the proportion of voids in the 1 m³ cube, which is simply $(V_t - V_b - V_r) / V_t$, where $V_t = 1 \text{ m}^3$.

Component Area Measures, Point Counts, and Line Counts

Slices are extracted from the 3-D models in order to assess the precision and accuracy of the point counts and line counts. A total of 30 horizontal and 30 vertical slices are produced for each 3-D model, and for each slice, a 5000 × 5000 pixel RGB-encoded image is saved in a PNG file. The scripting capabilities of ParaView are used to generate the files automatically (Ahrens et al., 2005). The background between the particles is assigned a pure green color (RGB encoding of [0,1,0]), and

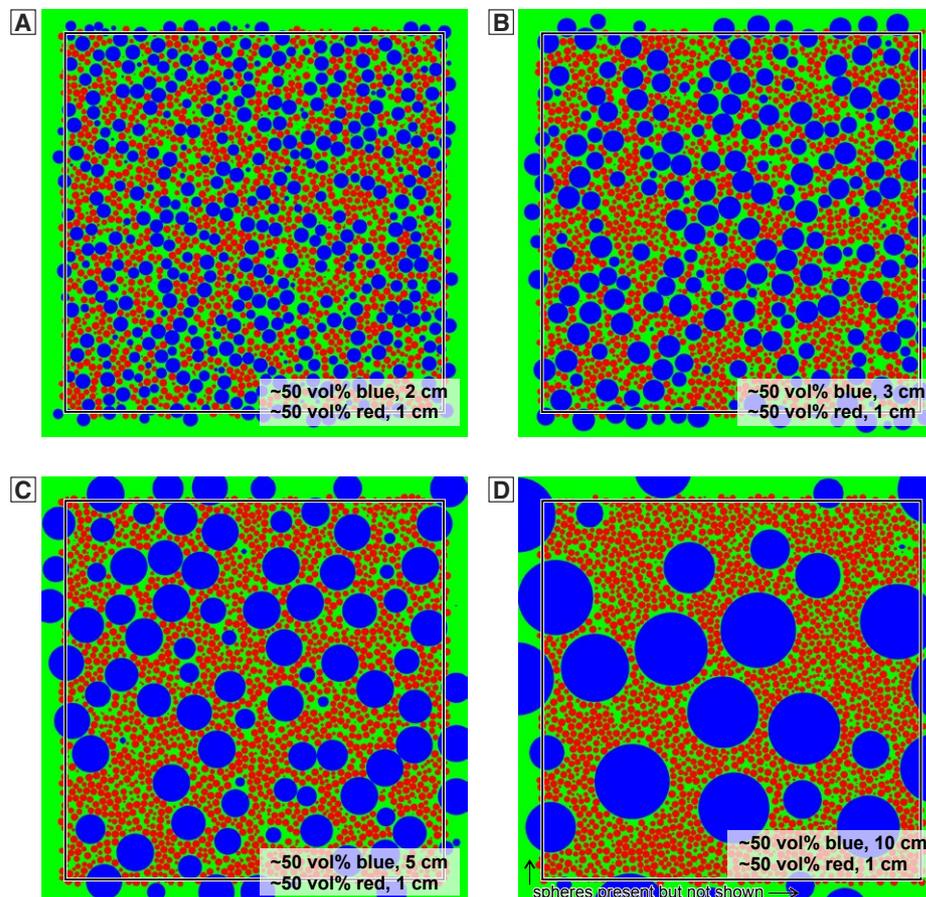


Figure 4. Horizontal slices through the middle of unequal-sphere models containing ~50% red spheres and ~50% blue spheres. (A) 1 cm radius red spheres and 2 cm radius blue spheres (1 + 2 cm). (B) 1 + 3 cm. (C) 1 + 5 cm. (D) 1 + 10 cm. The same figure is available in Supplemental File S2 (footnote 1) with gray backgrounds (for color-blind people).

the components are either pure red or pure blue (RGB encoding of [1,0,0] and [0,0,1] respectively). Lighting parameters are set to minimize shadowing in order to easily distinguish between colors. Examples of slices can be seen in Figures 3 and 4. Component area measures are obtained by simply retrieving the number of red and blue pixels on each slice, and the blue / (red + blue) ratio is calculated. The proportion of green pixels relative to other colors becomes the void fraction.

Point counting and line counting are implemented in Python, and the slices are processed in batches, which allows computing the mean proportion of red and blue spheres and the associated variance, for each considered scenario. For each slice, point counting is performed on a regular square grid of points covering a 1 m² area, with templates of 10 × 10, 15 × 15, 20 × 20, 30 × 30, and 50 × 50 points yielding 100, 225, 400, 900, and 2500 points, respectively. Note that these numbers of

points (N_{tot}) include a percentage of voids between the spheres. Line counting is performed with 1, 2, 3, 5, 7, and 10 lines, yielding different numbers of spheres intersected depending on their sizes (roughly 50–500 intersected objects for the models with red and blue spheres of 1 cm and 1 cm, respectively [notated hereafter as 1 + 1 cm]). Line counting is done by counting each pixel along 1-m-long parallel lines distributed equally over the slice. For example, lines are located at the bottom, middle, and top of the image when three lines are considered, similar to what would be done in the field. For both modal analysis methods, the presence of a blue sphere, a red sphere, or the background at any given pixel in the image is determined by the RGB channel with the highest value.

Calculating Accuracy and Precision

Each model has 60 slices. For the purpose of calculating accuracy and overall precision, each of these slices is point counted once at each grid spacing. For example, for the equal-size model with 30% blue spheres and 70% red ones, we have 60 point counts (one per slice) done with 100 points each (10 × 10 grid), 60 point counts done with 225 points each (15 × 15 grid), etc. Accuracy for a given number of points is calculated by averaging the point counted proportion of blue spheres (over the 60 slices) and comparing that with the true value. Similarly, the overall precision of the method for each model and for a certain number of points counted is the standard deviation of the point counts on the 60 slices through the model. This overall precision includes all sources of variability. The calculations are the same for line counts.

Calculating Counting Error for Point Counts

Counting error can be smaller than the total random error estimated by the overall precision just described. To isolate the counting error for the point counting method, we select a single slice through each model. We then make 60 repeated point counts on each of these unique slices for each

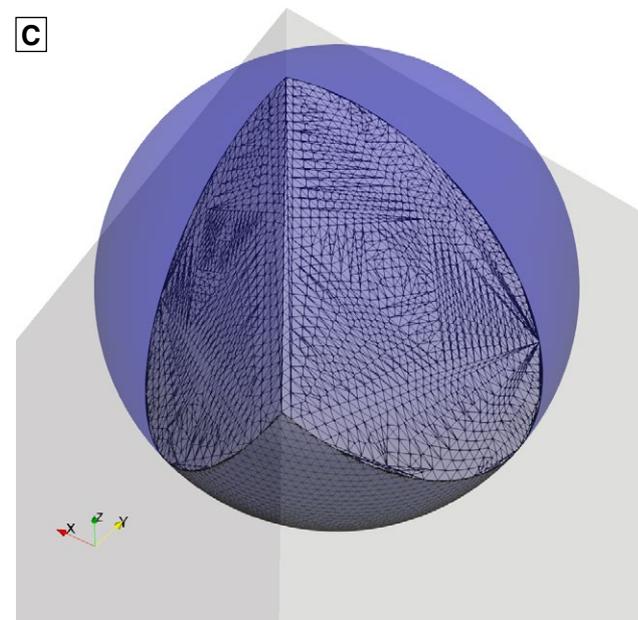
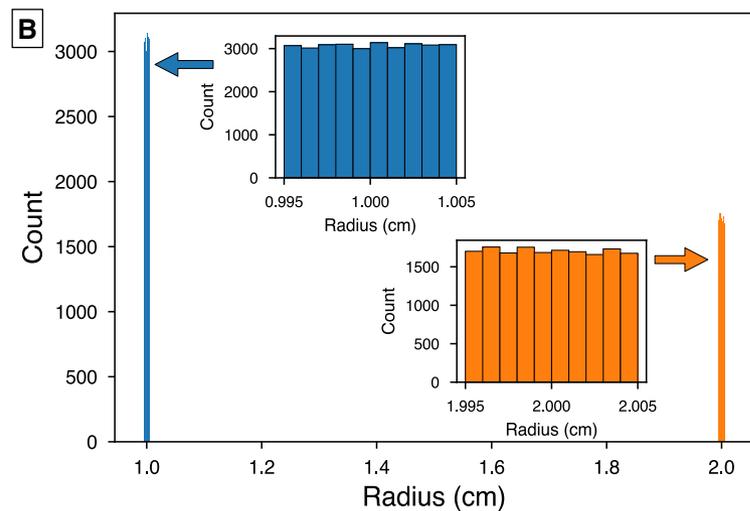
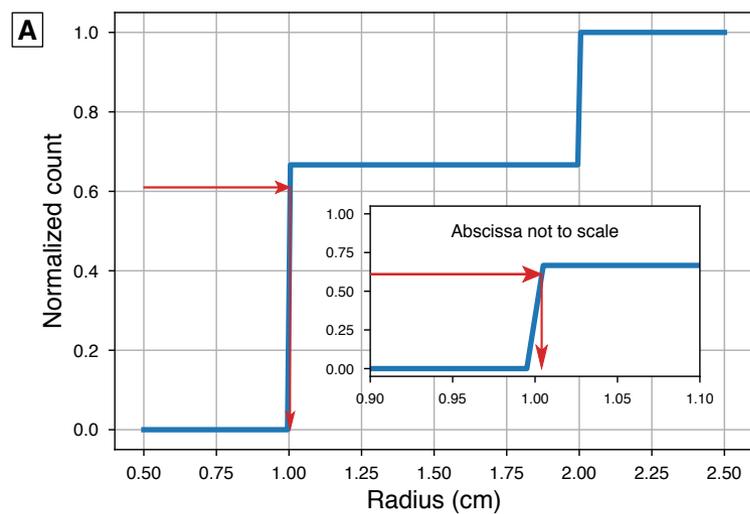


Figure 5. Information about the three-dimensional numerical models. (A) Example of cumulative particle-size distribution curve used to create the sphere populations. The radius of each individual sphere is chosen after drawing a number between 0 and 1 from a uniform distribution and projecting this number on the curve. (B) Example of histogram of the radius for a population of spheres resulting from the process described in A. (C) Illustration of the slicing of a sphere located on an edge of the cubic volume of interest. The volume of the fraction of sphere inside the cube is computed numerically after discretization with a tetrahedral mesh.

number of points investigated (100, 225, etc.). For convenience, the points are randomly placed each time on the slice, which should be equivalent to shifting the grid slightly each time. No such exercise was done for the line count method.

RESULTS

The different numerical models allow us to study the slicing effect and quantify the influence of variable component proportions, sphere size, and number of points or lines counted (or objects intersected) on the accuracy and precision of point counts and line counts. The equal-size sphere models represent the best-case scenario to evaluate the performance of these modal analysis methods. In nature, the components may have unequal sizes, and this could potentially deteriorate the performance of the modal analysis methods; this is examined with the unequal-size models.

Point Counting Accuracy

We start with the point counting results for the equal-size (1 + 1 cm) models (Fig. 6). Recall that we average the results over 60 slices in each 3-D model. Generally, for N_{tot} (the total number of points including those in voids) between 100 and 2500 and p between 0.1 and 0.9, the point count method is accurate in this test. This means that the average measured proportions of blue spheres (colored dots, Fig. 6) are very close to the true volumetric proportion in the models (dashed lines on upper panel of Fig. 6 or “zero error” lines in other panels). Average absolute errors range from positive to negative and are always within 1% or 0.01 (positive or negative) for individual models (middle panel of Fig. 6), and the average relative error of all models combined ($p = 0.1-0.9$) is $\sim 0.0\%$, i.e., there is no systematic bias (lower panel of Fig. 6). The exception is the $p = 0.01$ model, where the results are very variable in relative terms, but counting several thousand points would likely generate accurate results.

We now consider the point counting results for spheres of unequal sizes, with proportions of

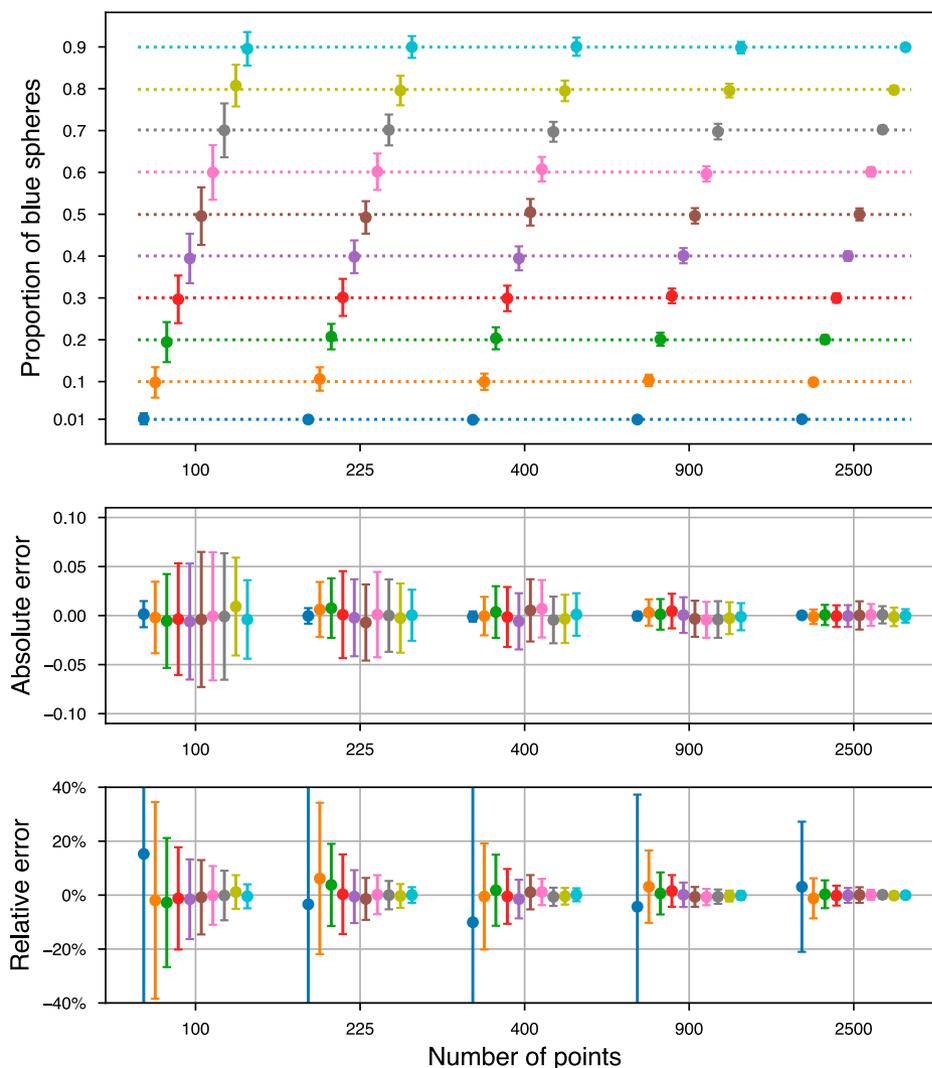


Figure 6. Point count results for 60 slices of the equal-size (1 + 1 cm) sphere models for 10 proportions of blue spheres, from 0.01 (1%, dark blue symbols) to 0.9 (90%, cyan symbols). All 60 slices are counted once. From top to bottom, diagrams display the counted proportions of blue spheres, the absolute error, and the relative error as a function of the total number of points counted (from 100 to 2500 points). In the upper diagram, the colored dots represent the average counted proportion of blue spheres, and the dashed lines represent the true proportion. Error bars are one standard deviation (top, middle diagrams) or one relative standard deviation (bottom diagram), giving an indication of precision. Refer to the Figure 2A inset for how to interpret precision and accuracy from the results presented.

~50% blue and 50% red spheres, and the equal-size (1 + 1 cm) case for comparison (Fig. 7). Note that the true proportions of blue spheres are not exactly 50% in the unequal-size models because of the way the models are generated (see Methods). Other proportions were calculated but are not shown. In this test, the accuracy of point counting is still good even when the spheres are not of equal size. For example, for the 1 + 2 cm, 1 + 3 cm, and 1 + 5 cm models, regardless of the proportion of blue spheres, the average absolute error is still mostly within 1% (positive or negative). For each of these unequal-sphere series, the average absolute errors of all models combined is also 0.0%, i.e., there is no systematic bias. However, the 1 + 10 cm model shows average absolute errors between 0.5% and 1.9% for the ~50% blue~50% red scenarios, depending on the number of points counted. This average absolute error is always positive, i.e., the much larger blue spheres are slightly overestimated. So, in extreme cases of components of different sizes coexisting in the same sample, there might be a slight systematic bias.

Point Counting Precision

Overall Precision

The overall precision of the method is estimated by calculating the standard deviation of our measurements on 60 slices, shown by the error bars in Figures 6 and 7. Because we are comparing measurements on many slices here, overall precision includes the counting error and any slicing effects, as discussed below. As expected, overall precision improves with more points counted (Fig. 7). However, the standard deviations remain high for models with blue spheres of 5 and 10 cm radii, even at $N_{tot} = 2500$ points (Fig. 7).

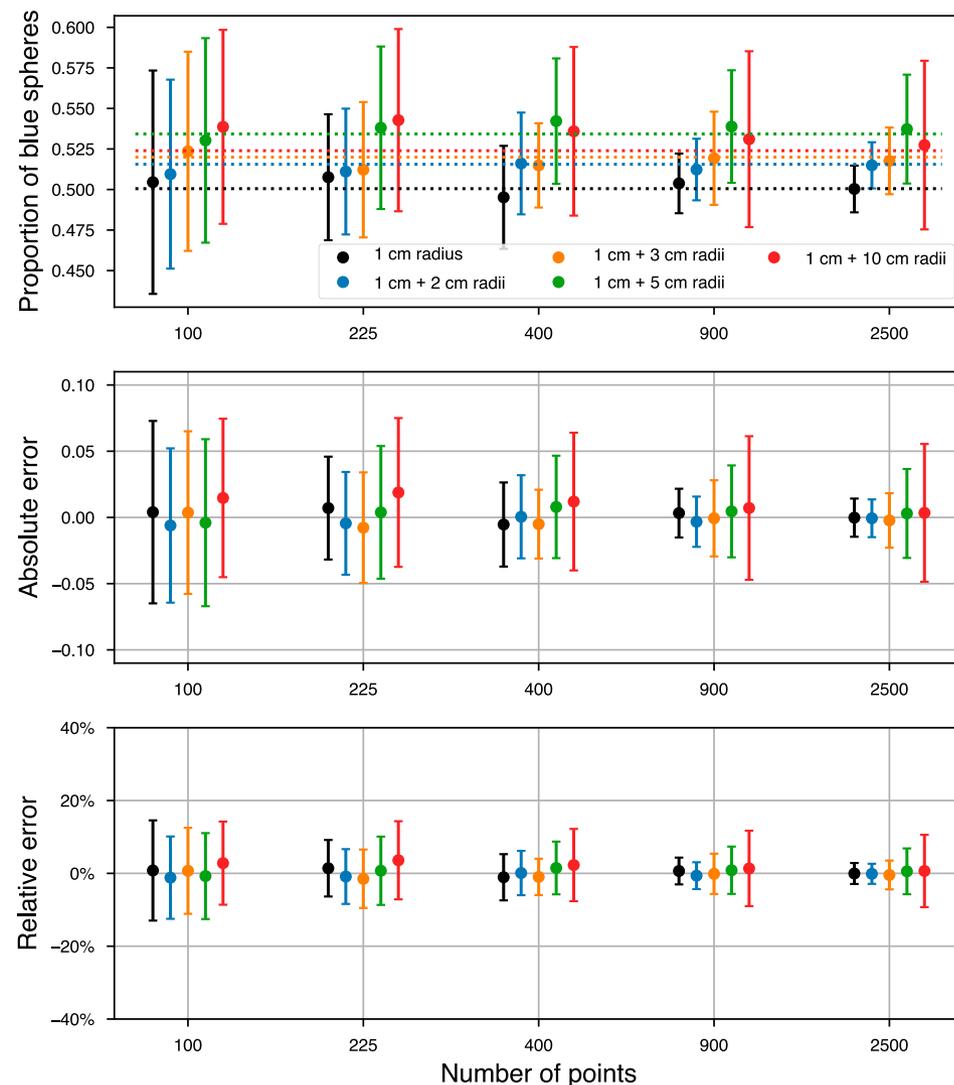


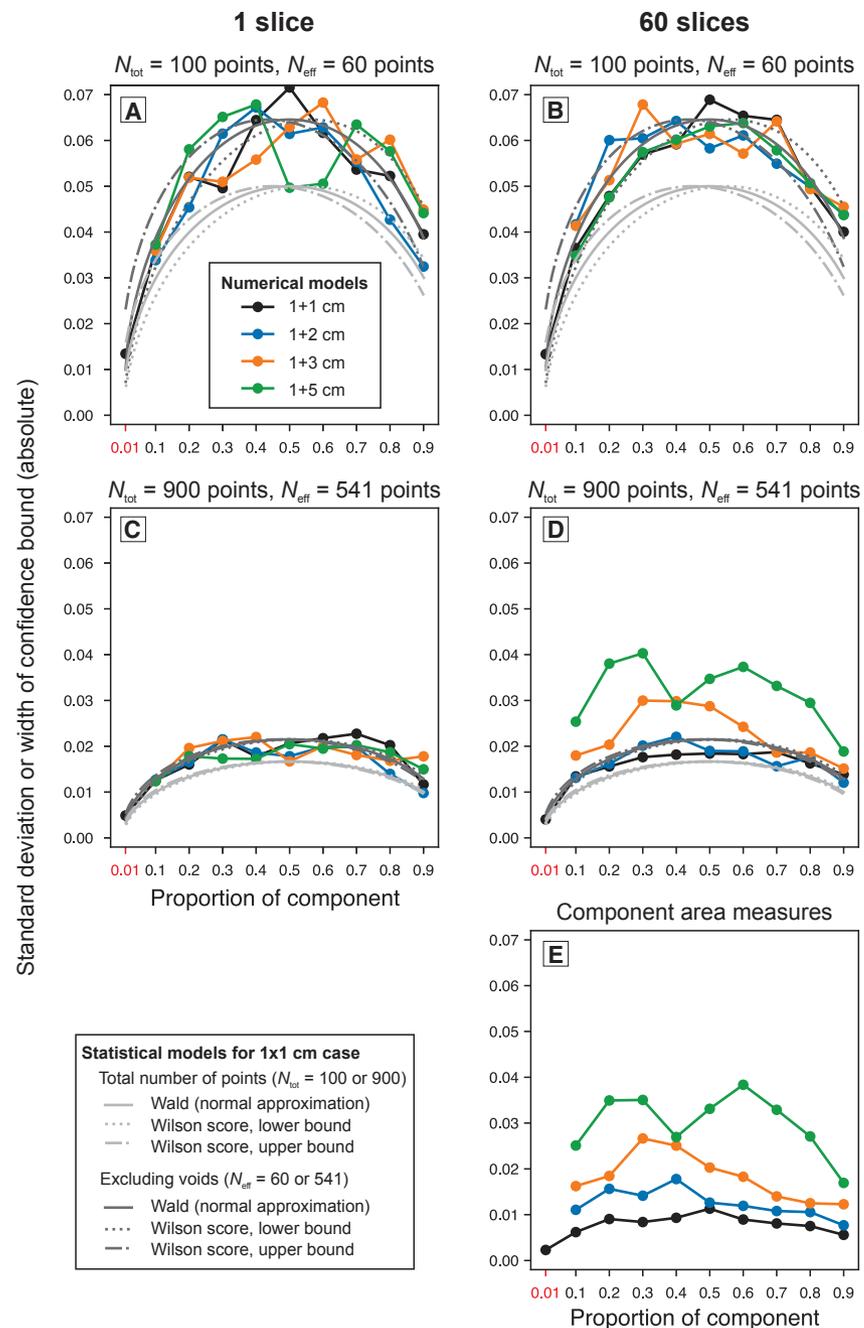
Figure 7. Comparison of the point count results for the equal-size (1 + 1 cm) and unequal-size sphere models (1 + 2 cm, 1 + 3 cm, 1 + 5 cm, and 1 + 10 cm) only for ~50% of blue spheres. Other aspects of the diagrams are as in Figure 6.

Counting Error on One Slice versus Overall Precision

We now compare the absolute standard deviation obtained when performing many point counts on a single slice through the models, which is the counting error (Fig. 8, left), versus the absolute standard deviation obtained when acquiring a single point count on each of the 60 slices (Fig. 8, right), which is our overall precision from Figures 6 and 7.

The colored curves show the effect of changing the proportion of blue spheres in our models with equal and unequal sizes. When counting 100 points in total, the curves are convex, i.e., the absolute standard deviation tends to be largest for 50% blue spheres. The different models are generally similar to each other at 100 points, and there is no major difference between repeating measurements on one slice (Fig. 8A) versus measuring each of the 60 slices once (Fig. 8B), although the latter case shows less scatter between models and within a model as p changes. At 900 points counted in total, the standard deviation is much lower for the same models and proportions. But here the distinction between point counting a single slice several times (Fig. 8C) versus point counting 60 slices once each (Fig. 8D) is much clearer. With the 60 slices taken into account, increasing the size difference between spheres generally increases the standard deviation. Figure 8 is further discussed

Figure 8. Standard deviation of point counting results for numerical models (colored curves) showing a low count (100 points total, A and B) and a relatively high count (900 points total, C and D) as a function of component abundance (percentage of blue spheres). The left column (A, C) shows repeat counts on a single slice, whereas the right column (B, D) shows one count on each of the 60 slices through each model. These data are compared with 68.27% confidence bounds for a binomial proportion (gray curves; see Supplemental File S1 [text footnote 1] for details), based either on the total number of points counted (N_{tot}) or the effective number of points counted in spheres for the 1 + 1 cm model, excluding voids (N_{eff}). The bottom-right diagram (E) shows component area measures for the numerical models, i.e., counting all red versus blue pixels on each slice, for comparison. The horizontal scale is linear, except for the addition of a 0.01 proportion to the left.



below, including the component area measures and the statistical models.

Line Counting Accuracy

We now move on to line counting results, starting with the 1 + 1 cm models with various proportions of blue spheres (Fig. 9). Again, accuracy is judged by comparing the average measurements on 60 slices with the true values. Average absolute errors (position of colored dots in the middle panel of Fig. 9) are generally within 1% (positive or negative) for individual models, although with only one line measured, one scenario has an average absolute error of +2.4%. The average absolute error of all models combined (p between 0.1 and 0.9) is +0.1%, or 0.0% for between three and 10 lines, which suggests that there is essentially no systematic bias, i.e., line counting is accurate for the equal-size models with p between 0.1 and 0.9. Again, for $p = 0.01$, the results are quite variable, but counting enough lines would likely generate accurate results.

Next, we compare the equal- and unequal-size models for a proportion of blue spheres ~50% (Fig. 10). We also have data for other proportions (not shown). For unequal sizes, the accuracy of line counts is problematic at low numbers of lines (especially one or two lines; Fig. 10), which correspond to small numbers of intersected objects. For example, in the 1 + 2 cm models, using only one line yields average absolute errors that are always positive and range from +1.0% to +7.5%. In contrast, still in the 1 + 2 cm models, the average absolute error of all models is only +0.5% for three lines and goes to 0% when more lines are added. For other unequal sizes, the errors are also the worst for one or two lines but vary from positive to negative, and the method becomes accurate for higher numbers of lines.

Line Counting Precision

In 1 + 1 cm models, overall precision obviously improves with the number of lines counted (Fig. 9).

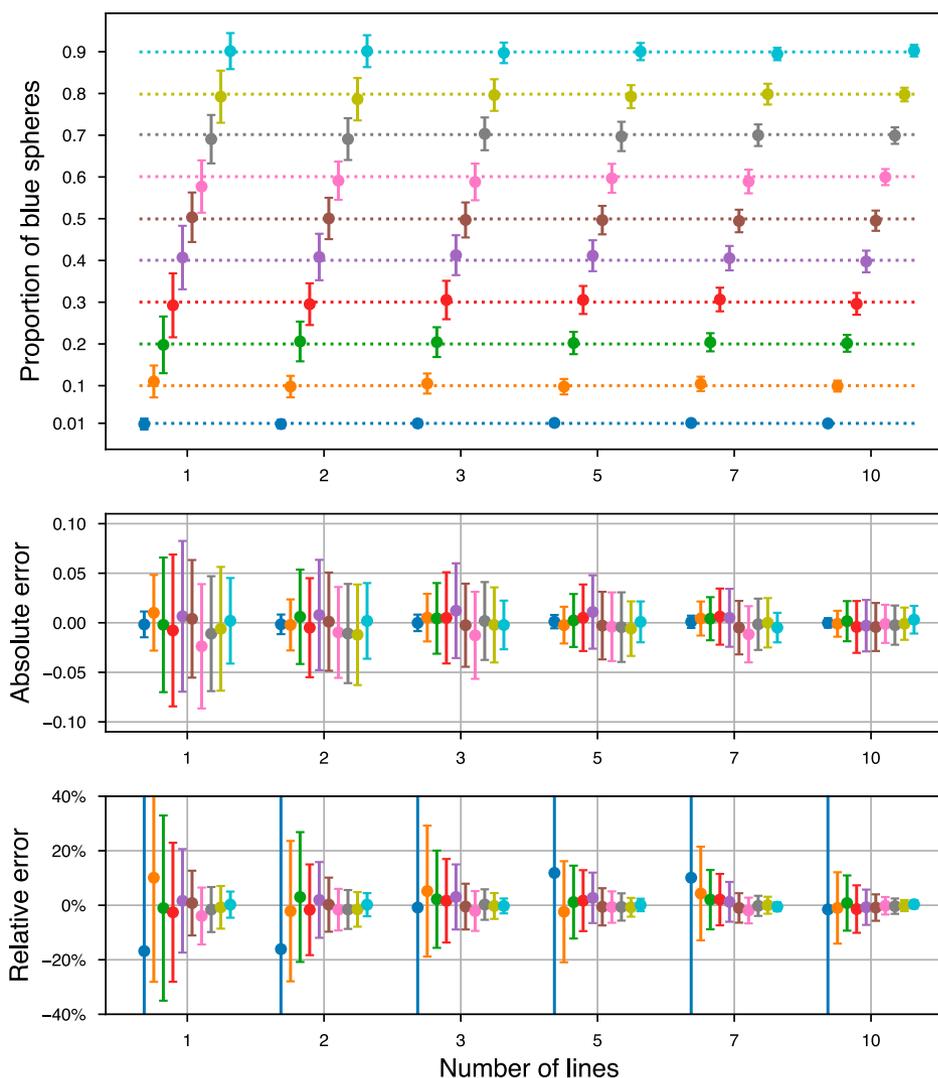


Figure 9. Line count results for the equal-size (1 + 1 cm) sphere models for 10 proportions of blue spheres as a function of the number of lines counted (from one to 10 lines). All 60 model slices are counted once each. Other aspects of the diagrams are as in Figure 6.

For a component with a 50% abundance, a standard deviation better than 5% absolute (10% relative) is obtained for three lines.

Overall precision is worse when the spheres are not of equal size. In fact, for a certain number of lines counted, the standard deviation increases systematically as the size difference between red and blue spheres increases (going from 1 + 1 cm to 1 + 10 cm) even for 5, 7, or 10 lines (Fig. 10). Figure 11 highlights the effect of the changing of proportions of blue spheres on the line count method for equal- and unequal-size sphere models, taking the 60 slices into account. Line count curves display the same general trend as the point count curves (Fig. 8). Increasing the size discrepancy between components leads to larger standard deviations. Figure 11 is further discussed below, including the contribution of the slicing effect to the precision of modal analysis methods based on cross-sections.

Void Proportion Accuracy

All the results presented so far show the proportions of red and blue spheres normalized to 100% spheres, ignoring voids. It is also interesting to check whether the void proportions (which in nature would correspond to a matrix and/or a cement ± voids) are correctly measured. There could be a systematic error on the void proportion due to the cut-section effect (Higgins, 2006). This is because it is unlikely that a certain sphere would be cut exactly at its greatest diameter, as visualized by looking at Figure 4C, where the blue circles have a range of sizes on the slice although the blue spheres are all essentially the same diameter in 3-D for this model. This means that for the average slice, the 2-D void proportion may be larger than the true volumetric proportion of voids. Note that this accuracy issue has no influence on the proportion of red versus blue spheres, but only applies to the void fraction.

Figure 12A shows that it is indeed the case that there is a small systematic bias in the slices for the void fraction. In this plot, the vertical axis is the relative difference between the measured 2-D void proportion on slices versus the true value in

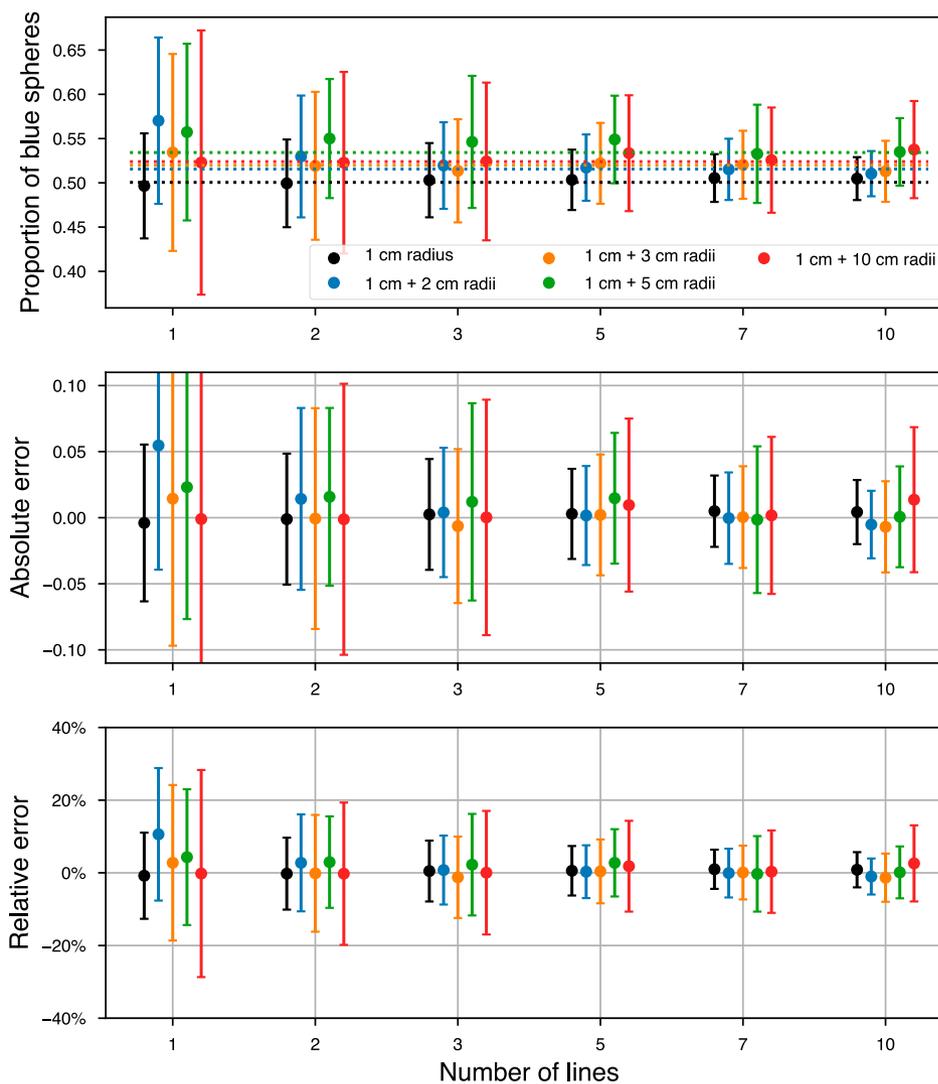


Figure 10. Comparison of the line count results for the equal-size (1 + 1 cm) and unequal-size sphere models (1 + 2 cm, 1 + 3 cm, 1 + 5 cm, and 1 + 10 cm) only for ~50% of blue spheres. All 60 model slices are counted once each. Other aspects of the diagrams are as in Figure 6.

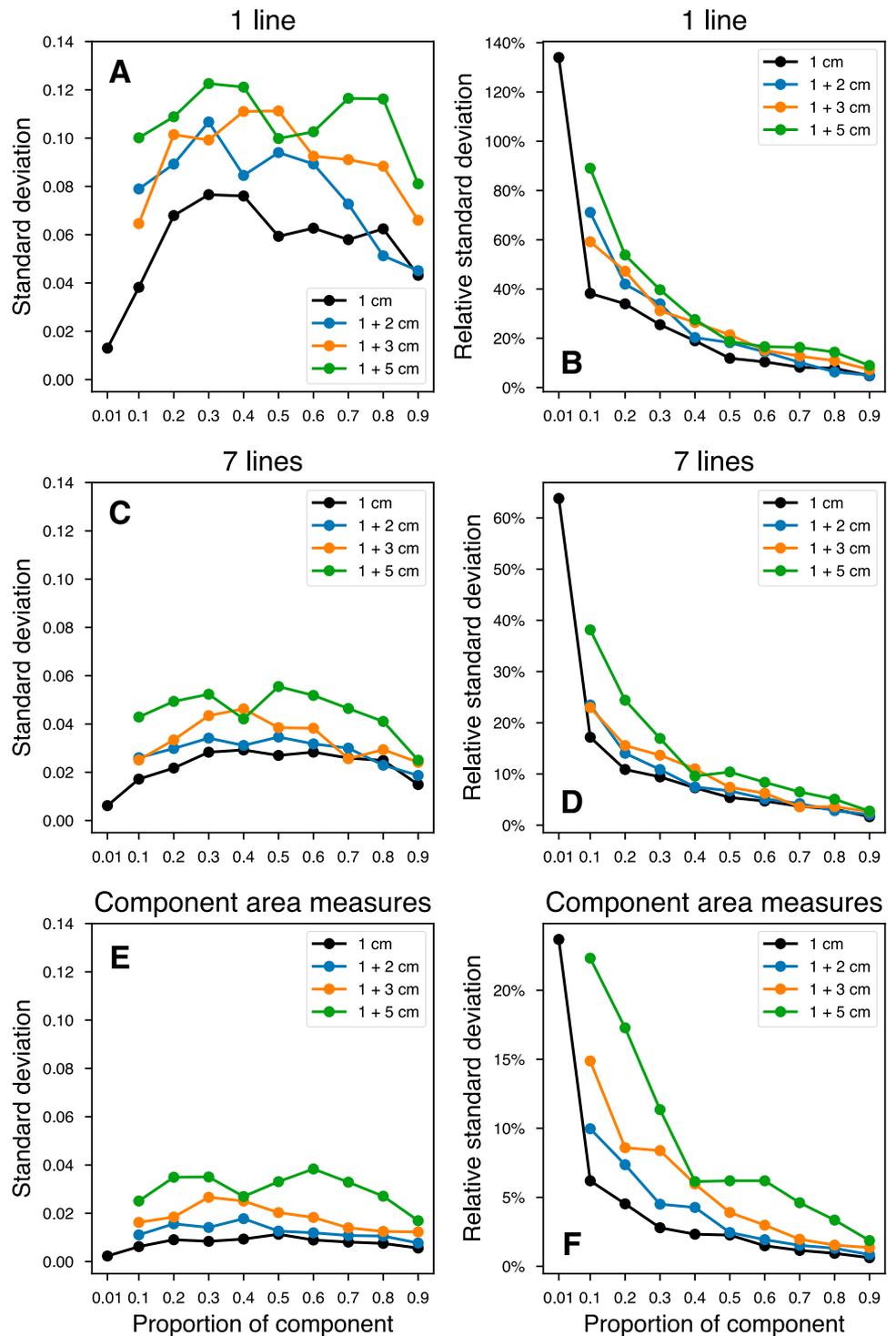


Figure 11. (A–D) Line counting standard deviation and relative standard deviation diagrams for a very low number of lines (one line, top) and a relatively high number of lines (seven lines, middle) as a function of component abundance (percentage of blue spheres). Both equal size and unequal size sphere models are illustrated. All 60 model slices are counted once each. (E–F) Component area measures, i.e., counting all red versus blue pixels on each slice, for comparison.

3-D. The horizontal axis is the true void proportion, which is maximum in the 1 + 1 cm model and minimal in the 1 + 10 cm model because small spheres can fit between larger spheres in the unequal-size models. When components are the same size (1 + 1 cm models) or no more than two times different (1 + 2 cm models), the systematic error due to slicing is almost negligible (<1%). The error is higher when components are very different in size, with the worst systematic error, ~2.5% relative (and the highest standard deviation, ~7.5% relative), for the 1 + 10 cm sphere model.

In theory, because the slices slightly overestimate the proportion of voids, the point counts and line counts—which are of course based on the slices—should also have the same issue. In practice, this is only clearly recognizable for 2500 points counted in total (Fig. 12B), where the pattern of always-positive systematic error, increasing with models of unequal sizes, is the same as in the slices (Fig. 12A). The magnitude of these systematic errors is also the same at 2500 points compared to the slices, for example, ~0.7% relative for the 1 + 1 cm model, or 2.2% relative for the 1 + 10 cm model. For most other point counts and line counts, the systematic error ranges from positive to negative without a discernable pattern. And in general, this systematic error is insignificant compared with the random error represented by the standard deviation, so the systematic error on the void proportion can be safely ignored in most cases.

DISCUSSION

Accuracy of Point Counts and Line Counts

Our study shows that point count and line count methods are both able to provide the correct proportion of a component, on average, i.e., they are both accurate without any corrections if the number of points or lines (i.e., objects intersected by lines) is high enough (Figs. 6, 7, 9, 10). This is true for both equal- and unequal-size models. Point counts are even accurate at low *N* for models featuring equal-size spheres (Fig. 6). There are situations when the accuracy of point counts and line counts decreases

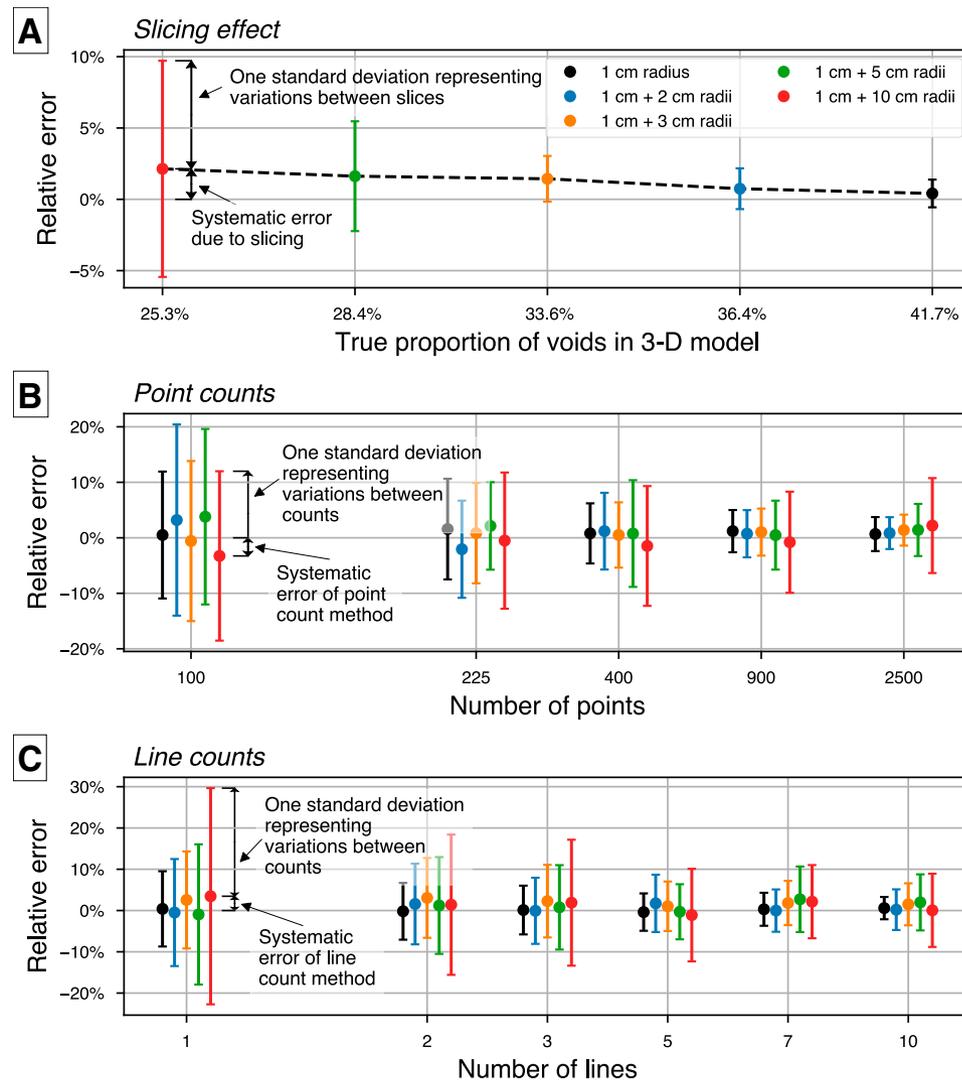


Figure 12. Relative errors of void quantification for the models with equal proportions of blue versus red spheres. (A) Slicing effect highlighted by comparing the average proportion of void (green pixels on slices) on the 60 slices with the true value of the void present in the three-dimensional (3-D) models; if there is no difference, the relative error is zero. (B) Proportion of void found with the point count method. (C) Proportion of void found with the line count method.

(low counts or large difference in component sizes). However, the main concern is identifying the value of N that would achieve the necessary precision required to answer the scientific question being investigated, which would also require high enough counts and would likely take care of any accuracy issues. In this discussion, we therefore focus on precision, first comparing our numerical model-derived point count standard deviations to confidence bounds from statistical models. We then make recommendations for field modal analysis studies based on new precision charts for point and line counts.

Point Counting Precision: Numerical Models versus Statistics

Figure 8 compares the standard deviation of our point counting results on spheres of equal and unequal sizes with 68.27% confidence bounds based on two statistical models representing binomial proportions (see File S1 [footnote 1] for details). The first statistical model is the normal approximation to the binomial distribution as used by van der Plas and Tobi (1965), also known as a Wald interval, in which the symmetrical confidence bounds correspond with one standard deviation. The second estimator is the Wilson score method, which gives a lower and an upper bound. These are calculated both for the total number of points N_{tot} and the effective number of points N_{eff} where voids are excluded, based on the average number of points falling on spheres (components) for the 1 + 1 cm radius models.

Again, the left side of Figure 8 displays repeat counts on a single slice, i.e., the counting error. This is specifically what the statistical models are supposed to represent (e.g., van der Plas and Tobi, 1965). For 100 points counted in total, the absolute standard deviation of the numerical models is as much as 0.02 units higher than the width of confidence bounds predicted by statistical models based on N_{tot} but fits well with those based on N_{eff} (Fig. 8A). This good correspondence supports the relevance of our numerical models and confirms that the binomial distribution can be used to theoretically

predict the variability of point counts in the form of “counting error”. However, charts such as those of van der Plas and Tobi (1965), or ours shown below, should not be used with N_{tot} but instead with N_{eff} (where voids, matrix, and cement are not included in N), otherwise the stated counting error would be too low.

Now we consider the equal-size (1 + 1 cm) models at 900 points total, still for one slice only (Fig. 8C). Again, there is a good correspondence with the statistically predicted confidence bounds using N_{eff} . The unequal-size numerical models are not very different from the equal-size models on this plot, and the standard deviation of a series of counts on a single slice again corresponds with the counting error.

However, for the 60 slices at 900 points, the standard deviation increases systematically as the size of the blue spheres increases relative to that of the red spheres, i.e., from the 1 + 2 cm to the 1 + 5 cm models, also progressively diverging more and more from statistical predictions (Fig. 8D). Therefore, there is a new source of variation between counts beyond the counting error. According to the Delesse principle, one slice can represent the whole volume, but this only works if (1) the volume is large relative to the size of the clasts, and (2) the distribution of the components is homogeneous. Our unequal-size models are meant to be homogeneous by design, but the blue spheres in the 1 + 10 cm models are 20 cm in diameter, one-fifth of the length of the side of the square (Fig. 4D). Therefore, individual slices through such a model are unlikely to contain the same surface proportion of red versus blue spheres as the volumetric proportion in the overall model. In other words, there are important modal variations between slices, as illustrated by the component area measures (Fig. 8E). Counting even a very large number of points on a single slice (in nature, an outcrop or a thin section) would not remove this effect, if present. For example, with the 1 + 5 cm models (green lines), when counting $N_{tot} = 900$ points, once per slice on 60 slices, the standard deviation is largely explained by the difference between slices (Fig. 8D). In contrast, when counting only $N_{tot} = 100$ points, the standard deviation is much larger and dominated

by the effect of the low number of points counted (counting error) (Fig. 8B).

Our choice of a 1 m³ numerical model with such large blue spheres is justified by our observations of maar-diatreme volcanoes, where we have conducted point counting and line counting within 1 m² areas of lapilli tuff and tuff breccia, i.e., very coarse heterogeneous volcanoclastic rocks, commonly with outsized clasts (Latutrie and Ross, 2020b). It is sometimes impossible to find a truly representative area to place the 1 m² net on such an outcrop. For example, in the photo of Figure 1B, moving the net upward or to the right by 0.5 m would increase the proportion of large orange clasts significantly. A reasonable solution here would be to study an area much larger than 1 m² by progressively moving the same net over adjacent areas, as discussed in the next section; this is partly similar to studying several slices through our numerical models instead of just one. In short, natural geological materials can be even more heterogeneous than our 1 + 10 cm models in terms of average component size and also in terms of compositional variability in 3-D. This would increase the variability of modal analysis measurements even for relatively large numbers of points counted or objects intersected within a 1 m² area, and we capture this effect at least partly with our unequal-size models. When heterogeneity is suspected, practitioners will want to add data for different sites or samples, and Vermeesch (2018) discusses statistical methods to distinguish the effects of counting error versus true compositional variability in a series of samples.

Practical Recommendations: Point Counting

We use the numerical models to construct new simple-to-use “error charts” that show the overall precision, i.e., the combined counting error and heterogeneity effect (when present). We show one standard deviation (absolute or relative) for the equal-size (1 + 1 cm) end member on the left and the 1 + 5 cm unequal size end member on the right (Fig. 13). Note that the van der Plas and Tobi (1965) charts were using 2σ precision instead, representing only the counting error, because they assumed

a homogeneous material. The horizontal axis is N_{eff} , the number of points within components excluding voids. This is more relevant than N_{tot} , the total number of points counted, which is model specific (or sample specific) and a function of the proportion of voids, matrix, and cement. Dash-dotted horizontal lines represent variability between slices, as represented by the standard deviation of the component area measures: at high-enough numbers of points, the variability of results is entirely explained by the slicing effect, especially in the 1 + 5 cm models, and counting more points would not improve overall precision.

To use these error charts in the field or laboratory, geologists should first visually estimate the proportion of the different components present in the deposit or rock, and whether the components have different average sizes (or are somewhat heterogeneously distributed). Then the error charts will provide the N_{eff} required to reach a certain overall precision. The required level of precision depends on the scientific question being asked.

If the abundance of a geologically important component is low, more points need to be counted to achieve the same precision. For example, based on the 1 + 1 cm models, the quantification of a component with a proportion of 0.1 (or 10%) would require a N_{eff} of ~1000 to obtain results with a relative standard deviation of ~10%. At the other end of the spectrum, components with abundances of 50% or more only need a N_{eff} of 100 to reach the same precision (Fig. 13, bottom left). Acquiring more than 1000 points within components per site would be extremely time consuming in the field, requiring, for example, at least 10 juxtaposed 1 m² nets with a 10 cm spacing between strings. Therefore, a relative standard deviation larger than 10% must probably be tolerated for low-abundance components in field applications. For the 1 + 5 cm models, it is clear that if components have very unequal sizes on average, a relative standard deviation of 10% is unattainable if the larger-sized component proportion is 30% or less.

Taking everything into account, for field cases where the main components of interest are relatively abundant, a N_{eff} between 200 and 300 is a reasonable compromise between time and

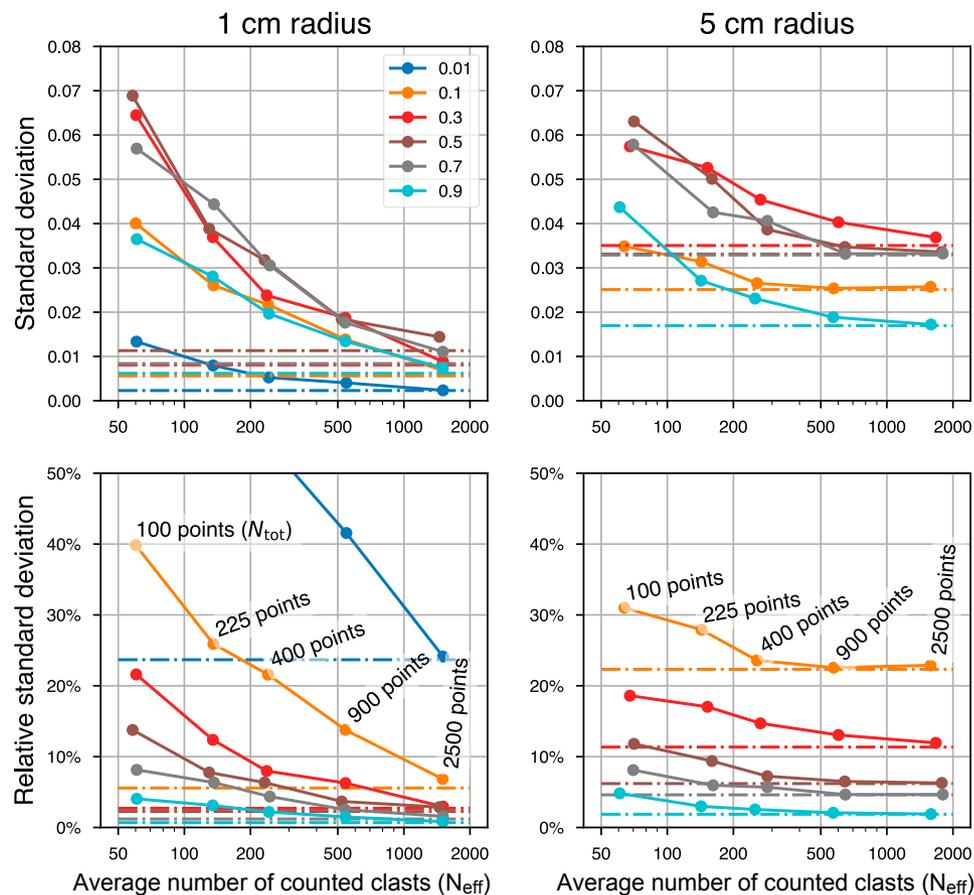


Figure 13. Point counting “error chart” based on the equal-size (1 + 1 cm) sphere models (left) and the 1 + 5 cm models (right). This shows the overall precision, which incorporates both the counting error and the slicing effect N_{tot} and N_{eff} are the total and effective number of points counted, respectively. Numbers from 0.01 to 0.9 are the proportion of blue spheres in the models, or in nature, the proportion of the component of interest. The dash-dot lines represent variability between slices, as represented by the standard deviation of the component area measures.

precision. If the proportion of voids, matrix, and cement in the materials investigated is similar to that of our models, this corresponds to a N_{tot} of ~400 points. This could be acquired using a 1 m² net with a 5 cm spacing at one place, or using a 1 m² net with a 10 cm mesh at four contiguous places (to cover 4 m²), depending on grain size and heterogeneity. Although a N_{tot} of ~400 points is much

higher than used in many previous field studies, users should remember that quantification with low precision can help confirm visual impressions and give a general idea of proportions, but not much more. If different lithofacies are to be compared, the quantitative data should be of sufficient precision and representativeness to allow statistical testing of hypotheses (see Davis, 2002). If the proportion

of voids, matrix, and cement is much more than 40%, for example, when point counting some lapilli tuffs in the field, then N_{tot} would need to increase to obtain the same N_{eff} .

For petrographic point counting, the traditional guideline of counting 300–400 points still seems relevant for high-abundance components, but this should be for points within components (N_{eff}), not total points including voids, matrix, and cement (N_{tot}). If low-abundance components are of interest, a N_{eff} of 1000 or more may be needed to achieve reasonable precision. In nature, rocks are made of many components, not just two. Therefore, the way to use Figure 13 for petrographic point counting would be to visually estimate the proportion of the least-abundant component of interest, select an acceptable absolute or relative standard deviation, and read N_{eff} from the horizontal axis. For example, suppose that an equigranular crystalline rock is made up of abundant olivine, clinopyroxene, plagioclase, and ~1% oxides (with no voids, groundmass, matrix, or cement), and we are very interested in quantifying these oxides. We use the 0.01 curve for the 1x1 cm models, and from the bottom left plot in Figure 13, we can see that counting a N_{eff} of 1000 points would yield a relative standard deviation slightly greater than 30%. If this is too high, we could use 1500 points.

Practical Recommendations: Line Counting

Similar error charts are presented for line counts, again for the 1 + 1 cm and 1 + 5 cm cases with the number of intersected objects displayed in the horizontal axis (instead of the number of lines) (Fig. 14). Counting three lines, as has been done by some authors in the field recently (Latutrie and Ross, 2020b), is obviously an improvement to counting one line only, but it still implies intersecting only ~150 clasts if the natural material under study looks like the 1 + 1 cm models, and commonly less than 100 clasts in a material similar to the 1 + 5 cm models. Intersecting 200–300 clasts as per the suggestion for point counting above requires approximately five lines in the 1 + 1 cm models. In the 1 + 5 cm models, intersecting over 200 clasts

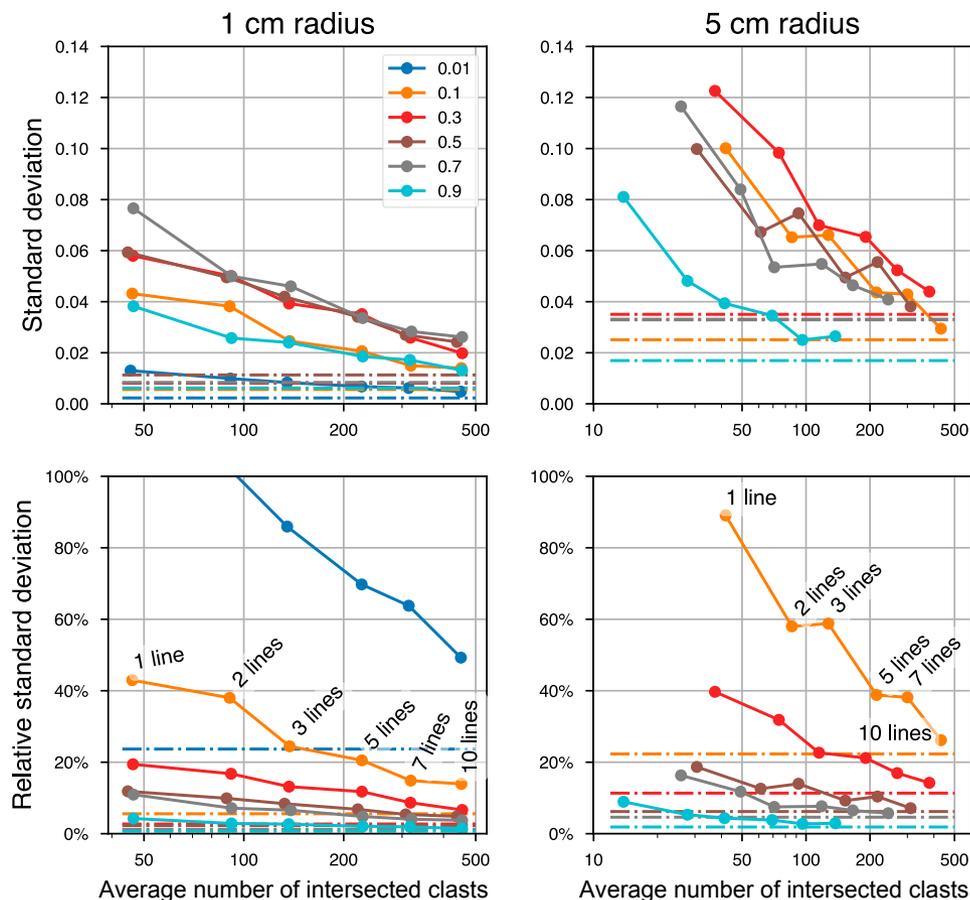


Figure 14. Line counting error chart based on the equal-size (1 + 1 cm) sphere models (left) and the 1 + 5 cm models (right). This shows the overall precision, which incorporates both the counting error and the slicing effect. Numbers from 0.01 to 0.9 are the proportion of blue spheres in the models, or in nature, the proportion of the component of interest. The dash-dot lines represent variability between slices, as represented by the standard deviation of the component area measures.

requires between five and >10 lines depending on the proportion of the component of interest, and this may not be practical.

Line counts and point counts can potentially achieve the same precision if the same number of clasts are counted or intersected. Further, line counts are convenient on vertical rock faces where it may be easier to steadily hold a 1-m-long tape

measure than a large string net. However, in the presence of large clasts or in poorly sorted deposits that may contain significant matrix and/or cement, it takes many lines to reach a high enough number of intersected clasts, especially if they are of unequal size. A further practical issue with field line counts is that each clast intersection length must be measured and then the lengths totaled, which is

a time-consuming exercise. It is therefore faster to achieve a given precision with point counts.

CONCLUSIONS

In this paper, we have considered two major modal analysis methods performed on cross-sections: point counts and line counts. Our numerical modeling based on different proportions and sizes of blue spheres versus red spheres demonstrates that both point and line counts can be *accurate*. They are both potentially able to provide the correct abundance of a component, on average, without any corrections, if the number of points or lines (i.e., objects intersected by lines) is high enough, even if components are of unequal size. To our knowledge, this had not been empirically demonstrated before for line counts, or for both methods using components of unequal sizes. One implication is that if enough data of sufficient quality are acquired by both methods on the same outcrops or other types of surfaces, the average results should be the same, therefore data from both methods can be directly compared (there is no systematic bias).

In practice, the main preoccupation of the scientist is likely to control the *precision* of modal analysis data, which would also require high enough counts and should take care of any accuracy issues. Previously published “error” (precision) charts for the point count method were based on statistical theory (including the normal approximation to the bimodal distribution) and ignored the effects of voids, matrix, and cement, the effects of components of unequal sizes, as well as variability between slices. New error charts showing all sources of variation have been proposed for point counts and line counts and can be used to design field or laboratory modal analysis studies. Achieving a given overall precision, expressed as the relative standard deviation, is more difficult for low-abundance components. Therefore, in practice, the effective number of points counted or the number of objects intersected by lines should be chosen as a function of the least abundant component of *interest*, but not necessarily the least abundant component overall, because modal analysis data

should be fit for purpose. Although point counts and line counts can achieve comparable precision if the same number of objects are counted or intersected, in practice this takes more time to achieve with line counts.

Our models so far are all based on spheres, each component having a single (unimodal) size or nearly so. Follow-up studies should examine the effects, on the accuracy and precision of point and line count methods, of factors such as:

- the morphology of components (e.g., spheres, cubes, ellipsoids);
- any anisotropy in the deposit;
- the size distribution of each component;
- the grid spacing relative to particle size; and
- adding extra components.

ACKNOWLEDGMENTS

This project was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants to PSR (RGPIN-2015-06782) and to BG (RGPIN-2017-06215). Michael Higgins kindly reviewed a pre-submission version of the manuscript, which helped in clarifying some ideas. We thank John P. Hogan, an anonymous reviewer, and Associate Editor Michael Williams for their constructive reviews and comments.

REFERENCES CITED

- Ahrens, J., Geveci, B., and Law, C., 2005, ParaView: An end-user tool for large data visualization, in Hansen, C.D., and Johnson, C.R., eds., *Visualization Handbook*: Amsterdam, Elsevier, p. 717–731, <https://doi.org/10.1016/B978-012387582-2/50038-1>.
- Baarli, B.G., Cachão, M., da Silva, C.M., Johnson, M.E., Mayoral, E.J., and Santos, A., 2014, A Middle Miocene carbonate embankment on an active volcanic slope: Ilhéu de Baixo, Madeira Archipelago, Eastern Atlantic: *Geological Journal*, v. 49, p. 90–106, <https://doi.org/10.1002/gj.2513>.
- Bélangier, C., and Ross, P.-S., 2018, Origin of nonbedded pyroclastic rocks in the Cathedral Cliff diatreme, Navajo volcanic field, New Mexico: *Bulletin of Volcanology*, v. 80, 61, <https://doi.org/10.1007/s00445-018-1234-0>.
- Blatt, H., 1992, *Sedimentary Petrology* (second edition): New York, Freeman, 514 p.
- Buzas, M.A., 1990, Another look at confidence limits for species proportions: *Journal of Paleontology*, v. 64, p. 842–843, <https://doi.org/10.1017/S00223360001903X>.
- Campbell, D., and Galehouse, J., 1991, Quantitative clinker microscopy with the light microscope: *Cement, Concrete and Aggregates*, v. 13, p. 94–96, <https://doi.org/10.1520/CCA10123J>.
- Carrapa, B., and DeCelles, P.G., 2008, Eocene exhumation and basin development in the Puna of northwestern Argentina: *Tectonics*, v. 27, TC1015, <https://doi.org/10.1029/2007TC002127>.

- Chayes, F.A., 1944, Petrographic analysis by fragment counting: Part 1, The counting error: *Economic Geology and the Bulletin of the Society of Economic Geologists*, v. 39, p. 484–505, <https://doi.org/10.2113/gsecongeo.39.7.484>.
- Chayes, F.A., 1945, Petrographic analysis by fragment counting: Part II, Precision of microsampling and the combined error of sampling and counting: *Economic Geology and the Bulletin of the Society of Economic Geologists*, v. 40, p. 517–525, <https://doi.org/10.2113/gsecongeo.40.8.517>.
- Chayes, F.A., 1956, *Petrographic Modal Analysis: An Elementary Statistical Appraisal*: New York, Wiley, 113 p.
- Clark, R.L., 1982, Point count estimation of charcoal in pollen preparations and thin sections of sediments: *Pollen et Spores*, v. 24, p. 523–535.
- Cundall, P.A., and Strack, O.D.L., 1979, A discrete numerical model for granular assemblies: *Geotechnique*, v. 29, p. 47–65, <https://doi.org/10.1680/geot.1979.29.1.47>.
- Davis, J.C., 2002, *Statistics and Data Analysis in Geology* (third edition): New York, John Wiley & Sons, 638 p.
- Delesse, A.E.O.J., 1848, Procédé mécanique pour déterminer la composition des roches: *Annales des Mines*, v. 13, p. 379–388.
- Dickinson, W.R., and Suczek, C.A., 1979, Plate tectonics and sandstone compositions: *American Association of Petroleum Geologists Bulletin*, v. 63, p. 2164–2182, <https://doi.org/10.1306/2F9188FB-16CE-11D7-8645000102C1865D>.
- Douce, A.E.P., and Johnston, A.D., 1991, Phase equilibria and melt productivity in the pelitic system: Implications for the origin of peraluminous granulites and aluminous granulites: *Contributions to Mineralogy and Petrology*, v. 107, p. 202–218, <https://doi.org/10.1007/BF00310707>.
- Fatela, F., and Taborda, R., 2002, Confidence limits of species proportions in microfossil assemblages: *Marine Micro-paleontology*, v. 45, p. 169–174, [https://doi.org/10.1016/S0377-8398\(02\)00021-X](https://doi.org/10.1016/S0377-8398(02)00021-X).
- Fleet, W.F., 1926, Petrological notes on the Old Red Sandstone of the West Midlands: *Geological Magazine*, v. 63, p. 505–516, <https://doi.org/10.1017/S0016756800085484>.
- Galehouse, J.S., 1971, Point counting, in Carver, R.E., ed., *Procedures in Sedimentary Petrology*: New York, Wiley Interscience, p. 385–407.
- Giddings, S.D., 1986, *Petrology, mineralogy, and geochemistry of the Goldlund gold deposit, northwestern Ontario* [M.S. thesis]: Grand Forks, University of North Dakota, 217 p.
- Go, S.Y., Kim, G.B., Jeong, J.O., and Sohn, Y.K., 2017, Diatreme evolution during the phreatomagmatic eruption of the Songgaksan tuff ring, Jeju Island, Korea: *Bulletin of Volcanology*, v. 79, 23, <https://doi.org/10.1007/s00445-017-1103-2>.
- Higgins, M.D., 2006, *Quantitative Textural Measurements in Igneous and Metamorphic Petrology*: Cambridge, UK, Cambridge University Press, 265 p., <https://doi.org/10.1017/CBO9780511535574>.
- Houghton, B.F., and Smith, R.T., 1993, Recycling of magmatic clasts during explosive eruptions: Estimating the true juvenile content of phreatomagmatic volcanic deposits: *Bulletin of Volcanology*, v. 55, p. 414–420, <https://doi.org/10.1007/BF00302001>.
- Howard, J.L., 1993, The statistics of counting clasts in rudites: A review, with examples from the upper Palaeogene of southern California, USA: *Sedimentology*, v. 40, p. 157–174, <https://doi.org/10.1111/j.1365-3091.1993.tb01759.x>.

- Howarth, R.J., 1998, Improved estimators of uncertainty in proportions, point-counting, and pass-fail test results: *American Journal of Science*, v. 298, p. 594–607, <https://doi.org/10.2475/ajs.298.7.594>.
- Ingersoll, R.V., Bulard, T.F., Ford, R.L., Grimm, J.P., Pickle, J.P., and Sares, S.W., 1984, The effect of grain size on detrital counts: A test of the Gazzi-Dickinson point counting method: *Journal of Sedimentary Petrology*, v. 54, p. 103–116, <https://doi.org/10.1306/212F83B9-2B24-11D7-8648000102C1865D>.
- Latutrie, B., and Ross, P.-S., 2020a, Phreatomagmatic vs magmatic eruptive styles in maar-diatremes: A case study at Twin Peaks, Hopi Buttes volcanic field, Navajo Nation, Arizona: *Bulletin of Volcanology*, v. 82, 28, <https://doi.org/10.1007/s00445-020-1365-y>.
- Latutrie, B., and Ross, P.-S., 2020b, What lithic clasts and lithic-rich facies can tell us about diatreme processes: An example at Round Butte, Hopi Buttes volcanic field, Navajo Nation, Arizona: *Journal of Volcanology and Geothermal Research*, v. 411, 107150, <https://doi.org/10.1016/j.jvolgeores.2020.107150>.
- Lefebvre, N.S., 2013, Volcanology of maar-diatreme volcanic vent complexes, Hopi Buttes Volcanic Field, Navajo Nation, Arizona, USA [Ph.D. thesis]: Dunedin, New Zealand, University of Otago, 269 p.
- Lefebvre, N.S., White, J.D.L., and Kjarsgaard, B.A., 2013, Unbedded diatreme deposits reveal maar-diatreme-forming eruptive processes: Standing Rocks West, Hopi Buttes, Navajo Nation, USA: *Bulletin of Volcanology*, v. 75, 739, <https://doi.org/10.1007/s00445-013-0739-9>.
- Magyari, E.K., Chapman, J.C., Passmore, D.G., Allen, J.R.M., Huntley, J.P., and Huntley, B., 2010, Holocene persistence of wooded steppe in the Great Hungarian Plain: *Journal of Biogeography*, v. 37, p. 915–935, <https://doi.org/10.1111/j.1365-2699.2009.02261.x>.
- Mastin, L.G., Christiansen, R.L., Thornber, C., Lowenstern, J., and Beeson, M., 2004, What makes hydromagmatic eruptions violent? Some insights from the Keanakākoʻi Ash, Kīlauea Volcano, Hawaiʻi: *Journal of Volcanology and Geothermal Research*, v. 137, p. 15–31, <https://doi.org/10.1016/j.jvolgeores.2004.05.015>.
- McKinley, J.M., Ruffell, A.H., and Worden, R.H., 2012, An integrated stratigraphic, petrophysical, geochemical and geostatistical approach to the understanding of burial diagenesis: Triassic Sherwood Sandstone Group, South Yorkshire, UK, in Morad, S., Ketzer, J.M., and De Ros, L.F., eds., *Linking Diagenesis to Sequence Stratigraphy: International Association of Sedimentologists Special Publication 45*, p. 231–256, <https://doi.org/10.1002/9781118485347.ch10>.
- Patterson, R.T., and Fishbein, E., 1989, Re-examination of the statistical methods used to determine the number of point counts needed for micropaleontological quantitative research: *Journal of Paleontology*, v. 63, p. 245–248, <https://doi.org/10.1017/S002233600019272>.
- Patterson, W.A., III, Edwards, K.J., and Maguire, D.J., 1987, Microscopic charcoal as a fossil indicator of fire: *Quaternary Science Reviews*, v. 6, p. 3–23, [https://doi.org/10.1016/0277-3791\(87\)90012-6](https://doi.org/10.1016/0277-3791(87)90012-6).
- Rebertus, R.A., and Buol, S.W., 1989, Influence of micaceous minerals on mineralogy class placement of loamy and sandy soils: *Soil Science Society of America Journal*, v. 53, p. 196–201, <https://doi.org/10.2136/sssaj1989.03615995005300010036x>.
- Retallack, G.J., 1994, Were the Ediacaran fossils lichens?: *Paleobiology*, v. 20, p. 523–544, <https://doi.org/10.1017/S0094837300012975>.
- Roduit, N., 2007, JMicroVision: Un logiciel d'analyse d'images pétrographiques polyvalent [Ph.D. thesis]: Geneva, University of Geneva, 116 p., <https://doi.org/10.13097/archive-ouverte/unige:468>.
- Rosival, A., 1898, Über geometrische Gesteinsanalysen: Ein einfacher Weg zur ziffer-mässigen Feststellung des Quantitätsverhältnisses der Mineralbestandtheile gemengter Gesteine: *Verhandlung der Kaiserlich-Königlichen Geologischen Reichsanstalt*, v. 1898, no. 5, p. 143–175.
- Ross, P.-S., and White, J.D.L., 2006, Debris jets in continental phreatomagmatic volcanoes: A field study of their subterranean deposits in the Coombs Hills vent complex, Antarctica: *Journal of Volcanology and Geothermal Research*, v. 149, p. 62–84, <https://doi.org/10.1016/j.jvolgeores.2005.06.007>.
- Ross, P.-S., and White, J.D.L., 2012, Quantification of vesicle characteristics in some diatreme-filling deposits, and the explosivity levels of magma-water interactions within diatremes: *Journal of Volcanology and Geothermal Research*, v. 245–246, p. 55–67, <https://doi.org/10.1016/j.jvolgeores.2012.07.006>.
- Schroeder, W., Martin, K., and Lorenzen, B., 2006, *The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics (4th edition)*: Kitware, 528 p.
- Shand, S.J., 1916, A recording micrometer for geometrical rock analysis: *The Journal of Geology*, v. 24, p. 394–404, <https://doi.org/10.1086/622346>.
- Šmilauer, V., Catalano, E., Chareyre, B., Dorofeenko, S., Duriez, J., Dyck, N., Eliáš, J., Er, B., Eulitz, A., Gladky, A., Guo, N., Jakob, C., Kneib, F., Kozicki, J., Marzougui, M., Maurin, R., Modenese, C., Scholtès, L., Sibille, L., Stránský, J., Sweißen, T., Thoeni, K., and Yuan, C., 2015, *Yade documentation (second edition)*: The Yade Project, 521 p., <https://doi.org/10.5281/zenodo.34073>.
- Stamper, C.C., Blundy, J.D., Arculus, R.J., and Melekhova, E., 2014, Petrology of plutonic xenoliths and volcanic rocks from Grenada, Lesser Antilles: *Journal of Petrology*, v. 55, p. 1353–1387, <https://doi.org/10.1093/petrology/egu027>.
- Taylor, J.R., 1997, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements (second edition)*: Sausalito, California, University Science Books, 327 p.
- Underwood, E.E., 1970, *Quantitative Stereology*: Reading, Massachusetts, Addison-Wesley, 274 p.
- van der Plas, L., and Tobi, A.C., 1965, A chart for judging the reliability of point counting results: *American Journal of Science*, v. 263, p. 87–90, <https://doi.org/10.2475/ajs.263.1.87>.
- van Otterloo, J., Cas, R.A.F., and Sheard, M.J., 2013, Eruption processes and deposit characteristics at the monogenetic Mt. Gambier Volcanic Complex, SE Australia: Implications for alternating magmatic and phreatomagmatic activity: *Bulletin of Volcanology*, v. 75, 737, <https://doi.org/10.1007/s00445-013-0737-y>.
- Vermeesch, P., 2018, Statistical models for point-counting data: *Earth and Planetary Science Letters*, v. 501, p. 112–118, <https://doi.org/10.1016/j.epsl.2018.08.019>.
- Weltje, G.J., 2002, Quantitative analysis of detrital modes: Statistically rigorous confidence regions in ternary diagrams and their use in sedimentary petrology: *Earth-Science Reviews*, v. 57, p. 211–253, [https://doi.org/10.1016/S0012-8252\(01\)00076-9](https://doi.org/10.1016/S0012-8252(01)00076-9).
- Wentworth, C.K., 1923, An improved recording micrometer for rock analysis: *The Journal of Geology*, v. 31, p. 228–232, <https://doi.org/10.1086/623009>.
- Zhou, Q., 2018, *PyMesh—Geometry processing library for Python*, v. 0.3 (installed 19 July 2020), <https://github.com/PyMesh/PyMesh>.