

# Geostatistical modelling of compositional variability across granitoid complexes: Its relevance to petrogenetic interpretations and specification of parent-rock properties in sediment-generation models



Gert Jan Weltje\*, Bram Paredis

KU Leuven, Department of Earth and Environmental Sciences, Celestijnenlaan 200E, 3001 Leuven, Belgium

## ARTICLE INFO

**Keywords:**  
 Multivariate geostatistics  
 Granites  
 Modal composition  
 QAPF maps  
 Log-ratio transformation  
 Principal components analysis  
 Optimization  
 Sampling  
 Petrogenesis  
 Fractionation  
 Mixing

## ABSTRACT

Sediment-generation models need an accurate specification of the fundamental properties of parent rocks and their variability at the scale of first-order (mono-lithologic) drainage basins. Georeferenced point-count data of five extensively surveyed plutons were extracted from the literature to examine the spatial variability of modal composition in granitoids. Point-count data must be considered inherently noisy from the point of view of geospatial analysis because sampled areas (thin sections) are small compared to the average crystal size of granitoids. Geostatistical modelling of such data is further complicated by the fact that crucial information on short-range variability is unavailable because sampling was carried out according to more or less regular patterns to achieve an equal density of data coverage across plutons. Geostatistical modelling of compositional data was carried out by transforming the data to centred log-ratios and calculating their Principal Components (PCs). The PC scores were used as input for a geostatistical workflow based on Ordinary Kriging, coupled with cross-validation and stochastic simulation to assess the predictive capabilities of the models. Sets of omnidirectional exponential variogram models with fixed range and sill, and variable nugget were used for each pluton. The local neighbourhood (search radius) for geostatistical interpolation was set equal to the range. Variogram modelling was formulated as an optimisation problem aimed at estimating the nugget for which the mean squared cross-validation (prediction) error reaches its minimum. The best model of each pluton was selected from all possible combinations of models obtained from the PCs. Results of a permutation test indicate that the residuals of observed and predicted compositions do not exhibit significant cross-covariance within the search radius adopted and may be interpreted as stationary random errors, which suggests that little may be gained by application of co-kriging. Random sampling from the geostatistical models indicates that up to several hundreds of specimens must be analysed to successfully predict the area-weighted mean modal composition of plutons and its spatial covariance structure as depicted in single-component and QAPF (Streckeisen) maps. An illustration of the internal consistency of the log-ratio approach and petrogenetic models is provided by the analysis of the compositional pattern in one of the granite complexes, which can be explained by the combined effects of mixing and fractionation. The volumes sampled by bulk geochemical analysis are equivalent to the areas of thin sections, and geochemical data of coarse-crystalline rocks are thus subject to the same limitations as modal analyses. The combined data-reduction and geostatistical modelling strategy outlined in this study is expected to be particularly efficient for the modelling of such high-dimensional data. The advent of quick non-destructive measurement techniques such as XRF and NIR is expected to play a crucial role in future attempts at rigorous quantitative mapping of lithosomes. If sediment generation can be simulated, the uncertainties associated with the initial conditions (area-weighted means of parent-rock properties) can be propagated all the way through to uncertainties of predicted sediment properties.

## 1. Introduction

Lithology, climate and tectonics - modulated by vegetation - exert a primary control on sediment generation in upland areas of sediment-

routing systems (Graham et al., 1986; Dickinson, 1988; Kirkby and Cox, 1995; Weltje et al., 1998; Forte et al., 2016; Allen, 2017). In sedimentary provenance studies, considerable attention has been given to constraining these boundary conditions (climatic-physiographic

\* Corresponding author.

E-mail address: [gertjan.weltje@kuleuven.be](mailto:gertjan.weltje@kuleuven.be) (G.J. Weltje).

regimes in source areas of sediments) from the modal composition of siliciclastic detritus (Basu, 1985; Suttner and Dutta, 1986; Graham et al., 1986; Dickinson, 1988; Grantham and Velbel, 1988; Ingersoll, 1990; Palomares and Arribas, 1993; Johnsson, 1993; Weltje et al., 1998). Sediment-generation studies attempt the opposite, i.e., to predict the characteristics of sediments produced in source areas from specified rock types and supplied to basins under given environmental conditions (Ibbeken and Schleyer, 1991; Basu, 2003; Weltje and von Eynatten, 2004; Heins and Kairo, 2007; Weltje, 2012). The basic requirement for successful sediment-generation modelling is a firm grip on the initial conditions of the sediment-forming process, i.e., the fundamental properties of parent rocks and their variability at the scale of first-order (mono-lithologic) drainage basins (Heins, 1993, 1995; Caracciolo et al., 2012; von Eynatten et al., 2012, 2016; Weltje et al., 2018; Paredis et al., 2020). In this study, we will attempt to shed light on the spatial variability of modal composition in granitoids through multivariate geostatistical modelling. Granitoids were selected because they represent a ubiquitous and relatively well-studied rock type and are a major producer of quartzose-to-feldspathic sand.

The importance of systematically surveying compositional and textural properties of granitoids to understand their spatial variability and assess the robustness of petrogenetic inferences was pointed out by Whitten (1959), who stated that: “Qualitative variations of plutonic rocks have often been described, but in most descriptive and petrogenetic studies of granite masses, little or no attention has been given to the bulk quantitative variation. This is somewhat surprising because in numerous instances a series of rocks collected astride a granitic contact has been assumed to be typical and has been subjected to detailed examination including chemical analysis. Such results are then used directly as the basis for petrogenetic conclusions. Subsequently, other authorities abstract such data from the literature and use them for correlation and comparison with other results.”

To capture the full compositional-textural variability of granitoids, we should at least like to know in which way the modal composition of the rock, the crystal-size probability distributions of its minerals, and the spatial arrangement of crystals vary simultaneously in space (Weltje et al., 2018; Paredis et al., 2020). But this is far too ambitious given available data: full characterization of the joint compositional-textural properties across entire plutons cannot be attempted at this stage. In this study, we will model the spatial variability of the petrographic (modal) composition of five granitoid complexes from legacy data acquired between 1955 and 1983 (Whitten, 1955, 1957, 1959, 1961; Taubeneck, 1957; Dawson and Whitten, 1962; Wadsworth, 1968; Vistelius et al., 1983), and extrapolate our inferences to mapping using state-of-the-art data-processing technology. Granitoids lend themselves to a quantitative petrographic analysis by means of optical microscopy because of their comparatively large average crystal size. The popularity of quantitative petrographic analysis of granitoids in the 1950s and 1960s reflects the available data-acquisition and processing technology of the time, i.e., the development of the point counter and the statistical basis of modal analysis, respectively (Chayes, 1949, 1956).

**Table 1**

Database used in this study. Codes of plutons as used in the text are listed within brackets, as are the UTM zones in which the plutons are located. n = number of control points, N = (average) number of points counted in thin section. Mineral classes: Q: quartz, P: plagioclase, K: potassium feldspar, Others: the sum of all minerals other than Q, P, K. (Others includes: A: accessories, Am: amphibole, Ap: apatite, B: biotite, C: carbonates, CI: colour index (sum of dark minerals), E: epidote, H: hornblende, M: mica, Mag: magnetite, O: opaque minerals, Ort: orthite, Pen: penninite, Px: pyroxene, S: sphene/titanite, Z: zircon).

Name (code)	Location (UTM zone)	Reference	n	N	Mineral classes
Bald Mountain batholith (BM)	Oregon, USA (11 T)	Taubeneck (1957)	130	2000	Q, P, K, B, O, A, H
Cornelia pluton (CP)	Arizona, USA (12S)	Wadsworth (1968)	67	1850	Q, P, K, B, H, Px, CI
Donegal granite (DO)	Donegal, Ireland (29 U)	Whitten (1955) Whitten (1957) Whitten (1959) Whitten (1961)	13 51 71 65	13 51 71 65	Q, P, K, B, M, H, Mag, Ap, Ort, Pen, Z, S, CI Q, P, K, B, H, other Q, P, K, CI Q, P, K, CI, H, S, E
Lacorne - La Motte - Preissac Complex (LC)	Québec, Canada (17 U)	Total (this study)	200	2050	Q, P, K, Others
Meech Lake pluton (ML)	Québec, Canada (18 T)	Dawson and Whitten (1962) Vistelius et al. (1983)	167 143	1500 1000	Q, P, K, CI Q, P, K, O, Am, Px, S, M, C

Petrographic modal analysis of igneous rocks is a rare phenomenon nowadays as the acquisition of geochemical data has become the method of choice in petrology.

Quantitative petrographic analysis of igneous and metamorphic rocks in the geochemical era has been mainly focused on the joint variability of modal composition and crystal-size probability distributions of individual minerals, based on the concept of the CSD (“crystal size distribution”), which actually represents a crystal size-volume density function (Marsh, 1988; Higgins, 2000, 2006; Paredis et al., 2020). To our knowledge, georeferenced data sets of CSDs comparable in size to the legacy data sets analysed in this study are not available. Georeferenced data sets referring to the spatial arrangement of crystals (relative interface frequencies) are available for two out of the five plutons studied in this contribution (Wadsworth, 1975; Vistelius et al., 1983). They fall outside of the scope of this study and will not be discussed here.

Whitten (1959) rightly pointed out that not much was known about the representativeness of specimens sampled from heterogeneous rock bodies, nor about the number of specimens required to characterize a given property within specified limits of uncertainty (accuracy and precision). Owing to the popularity of geochemical methods in petrology, the number of systematic quantitative descriptions of the petrographic composition of lithosomes has barely increased over the past decades, and the same goes for our knowledge of its spatial variability. The design of sampling schemes for characterizing heterogeneous rock bodies represents a conundrum because an adequate sampling scheme resulting in a representative data set of a particular rock body can only be designed based on prior knowledge about the spatial variability of its properties (Griffiths, 1967). One way out of this intricate problem is to construct geostatistical models of spatial variability from extensive data sets. Sampling from these models provides the opportunity to examine the variability of the statistics obtained from real-world sampling campaigns and may permit generalizations about sampling design to be made. In the final part of this paper, we explore the connection between the log-ratio-based approach to geostatistical analysis adopted in this study and exploratory petrogenetic modelling and propose a protocol for systematic mapping of petrogenetically constrained components based on non-destructive (hyper)spectral techniques.

## 2. Material

### 2.1. Pre-processing

Legacy data (modal compositions and accompanying sample locations) of five extensively surveyed plutons were extracted from the literature. The metadata of the plutons are listed in Table 1. We will refer to these five plutons through the abbreviations BM, CP, DO, LC, and ML. Modal data were digitized from tables and spatial coordinates were assigned to control points using figures (maps with sample locations) and/or local coordinates reported in tables. All local coordinate

systems in the database were matched to the EPSG:4326 (WGS 84) coordinate reference system and transformed to UTM coordinates (in meters) for geospatial analysis. The UTM zones of each pluton are listed in Table 1. The modal data were subjected to a quality check: all entries should be non-negative and their sum should equal 100% (within round-off error). The modal data were recalculated to the original number of points counted (raw point-count data) to replace occasional zero values in the raw point-count data by 2 points. This ensured positivity, honoured the sample size (total count length), and avoided artefacts in the data. The volume percentages of all mineral classes can be found in the Online Supplement (Table 1 lists the (average) count length of each data set). Modal compositions of the five plutons have been reported in different formats, which required some recalculation to harmonize the database. Apart from the three major constituents, i.e., quartz, plagioclase, and alkali feldspar (Q, P, and K, respectively), a range of other minerals has been reported. For the purpose of inter-comparison of the five data sets, we decided to analyse their lowest common denominator, which implies amalgamation of all minerals apart from Q, P, and K in a group termed Others (all other minerals). In addition, we analysed the lowest common denominator of BM and CP, which comprises six variables. For these two plutons, the group Others was split into B (biotite), H (hornblende), and a rest group that was labelled A (accessories) in BM, and Px (pyroxenes) in CP, according to their dominant components.

## 2.2. The data set

Fig. 1 depicts the modal compositions of the control points of all five plutons in the ternary QAP diagram, a subspace of the QAPF diagram (Streckeisen, 1974, 1980) in which Q and F are mutually exclusive. None of our specimens contained any F, and therefore all samples fall within the QAP triangle. Class K in our data sets is equivalent to A in the QAPF scheme (although strictly speaking, class A includes nearly pure albite, i.e., members of class P with An < 10%). The Q-A-P sub-composition was obtained by normalizing the Q, K, and P values to unity.

Fig. 1 shows that the database represents a wide spectrum of granitoids. The spatial distribution of control points within the five plutons is shown in Fig. 2. Each control point has been colour-coded according to its QAPF class. The outlines of the plutons as given in the original references have been added to enable comparisons with the geostatistical models to be constructed.

## 3. Methods

### 3.1. Compositional geostatistics

The analysis of compositional data has made a big leap forward since the time of the above systematic petrographic surveys. The value of systematically surveying rock bodies can only be brought out by subjecting such data to appropriate statistical analysis, which was not yet possible at the time. The breakthrough in this field is attributable to Aitchison (1982, 1986), who introduced the idea of log-ratio transformations to provide an internally consistent statistical framework for the analysis of compositional data. As a result of his work, statistical analysis of compositional data is now firmly embedded in the mainstream of multivariate statistics. The first attempts to extend the concepts of Aitchison to georeferenced compositional data (summarized by Pawlowsky-Glahn and Olea, 2004) were soon followed by more advanced geostatistical models (Tolosana-Delgado and van den Boogaart, 2013; Pawlowsky-Glahn and Egozcue, 2016; McKinley et al., 2016). Geostatistical analysis of compositional data has now reached a fully mature stage (Tolosana-Delgado et al., 2019). The algorithms for spatial interpolation and simulation of compositional data developed thus far treat the problem as a case of isotopic co-kriging (Wackernagel, 1994, 2003), which refers to the multivariate interpolation of quantities which are available at all control points. In co-kriging, the covariance structure of the data set is used as a source of information to improve the quality of predictions relative to methods which treat the components of a multivariate data set independently. A drawback of the current approach to geostatistical modelling of compositional data is

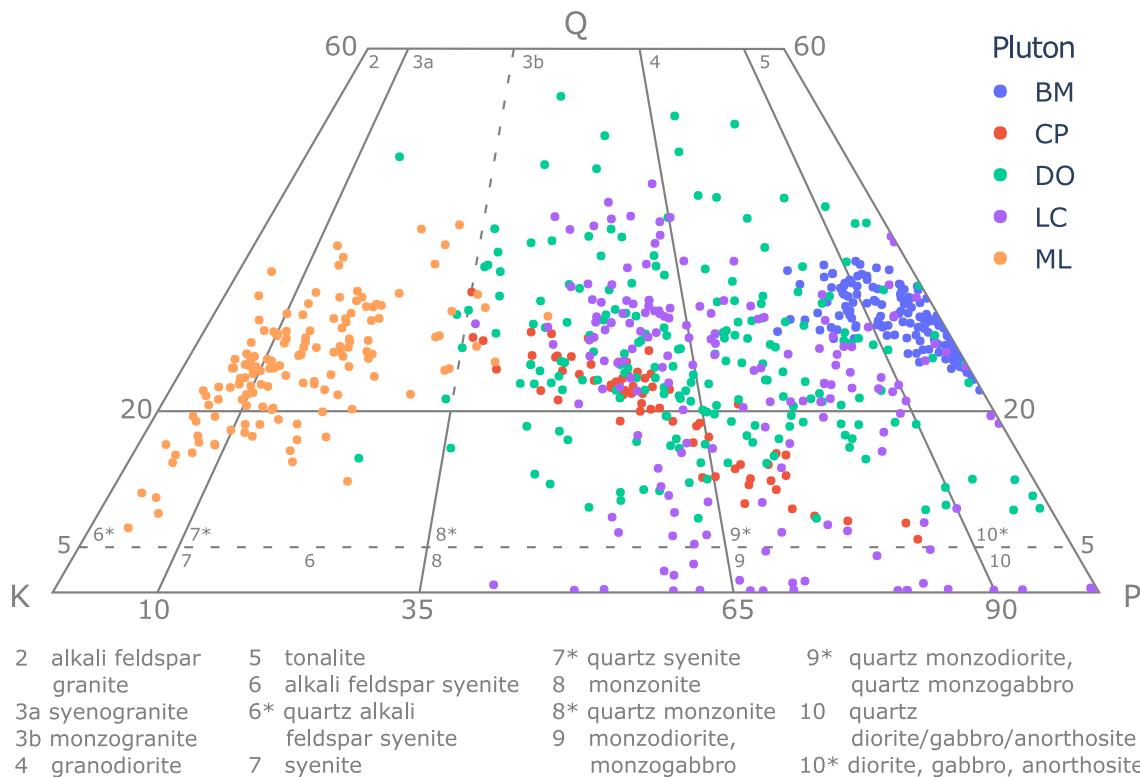
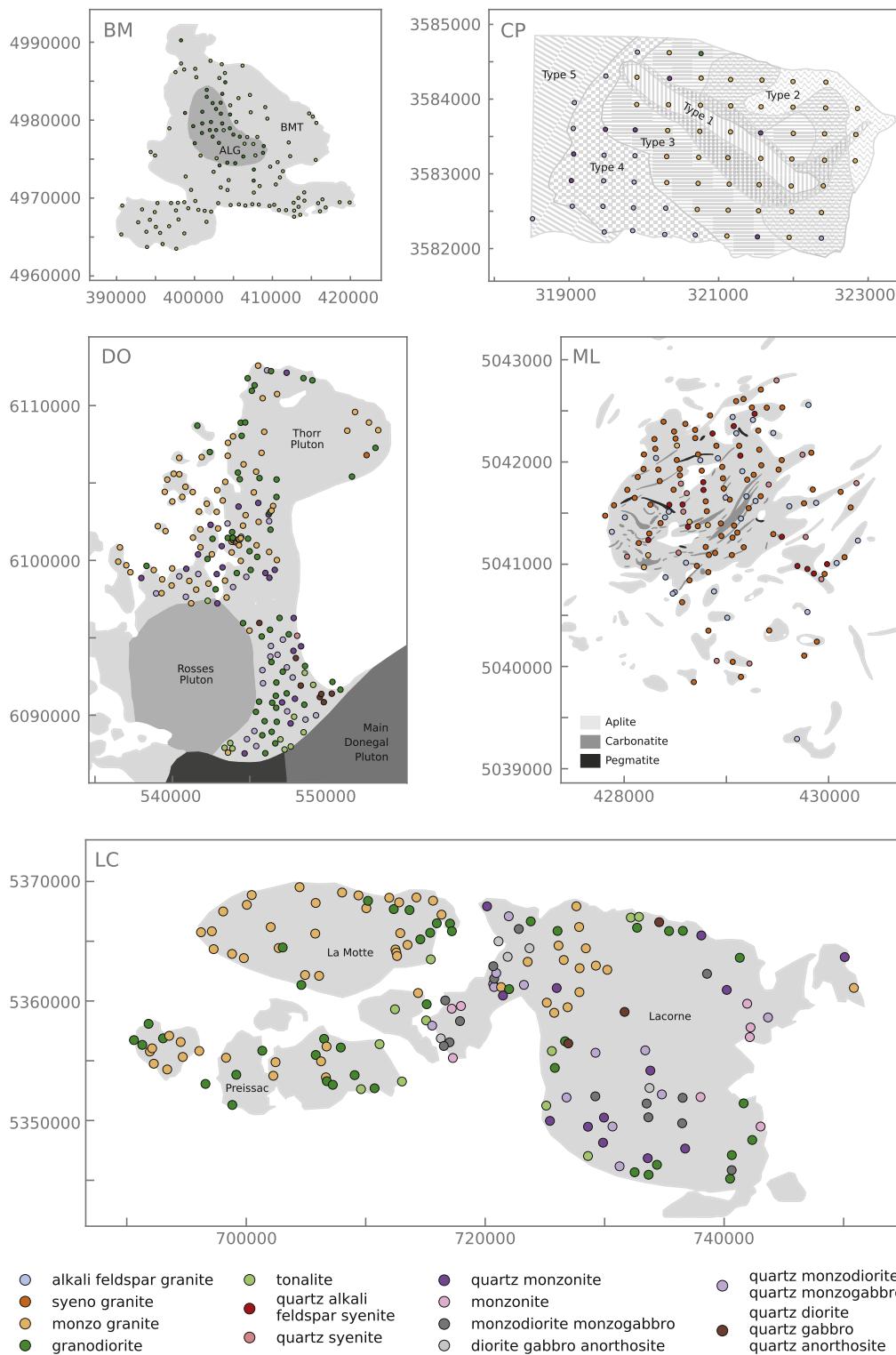


Fig. 1. Ternary subspace of the QAPF diagram (Streckeisen, 1974, 1980) with control points of the five plutons.



**Fig. 2.** Spatial distribution of control points in the five plutons colour-coded according to QAPF class (for legend see Fig. 1). Outlines of the plutons according to geological maps in the original references (Table 1) have been added. The geological map of DO has been modified from Atherton and Ghani (2002).

that the algorithms are quite involved (Tolosana-Delgado et al., 2019) and therefore likely to deter non-specialists from using compositional geostatistics as a routine quantification tool.

In this contribution, we will investigate the usefulness of a trimmed-down heuristic approach to the multivariate geostatistical interpolation of “noisy” compositional data such as petrographic compositions. This approach follows in the footsteps of methods developed in

oceanography, meteorology, and physical geography that rely on the modelling of the spatial variability of (globally uncorrelated) principal components (PCs) of multivariate signals (Preisendorfer, 1988; Demšar et al., 2013). It should be noted that the term “globally uncorrelated” refers to the compositional space. It does not imply that the set of PCs is also spatially uncorrelated, i.e., that pairs of PCs are also uncorrelated at all distances between control points (zero cross-covariance). If the

PCs are spatially uncorrelated, co-kriging of PCs is equivalent to kriging of individual PCs (Wackernagel, 1994, 2003). Kriging of individual PCs significantly reduces computational effort because the multivariate geostatistical problem at hand may be solved by a series of univariate interpolations obtained by a standard algorithm known as Ordinary Kriging (OK) (Davis, 2002; Isaaks and Srivastava, 1989). Kriging of individual PCs may still be permissible in the presence of significant cross-covariance because accurate interpolation of individual PCs ensures that the geostatistical model will reproduce any significant cross-covariance after back-transformation and addition of the contributions of successive PCs to estimated compositions. To decide if any cross-covariance among the PCs has been faithfully reproduced by the interpolation algorithm, we adopt the rule that the spatial structure of the multivariate residuals of the geostatistical model obtained through cross-validation (the residual compositions) should be indistinguishable from stationary random errors with zero expectation.

Cross-validation is used to obtain the set of residual compositions and quantify the mean squared prediction error (MSPE) of the geostatistical models. The residual compositions are subjected to the following tests in order to determine whether our approach faithfully captures the spatial variability of petrographic composition: (1) a global test to determine whether the residual compositions are unbiased, (2) a permutation test to examine whether the cross-covariance functions of the PCs of the residuals differ significantly from zero. The models which passed these two tests were evaluated to select the best geostatistical model for each of the five plutons examined in this study, based on the trade-off between their MSPE values and the number of PCs involved (i.e., the rank of the model,  $r$ ). In cases where two or more models of a given pluton have approximately identical MSPE values, the model which employs the smallest number of PCs was preferred, following the principle of parsimony (traditionally referred to as Ockham's razor, and nowadays more often as KISS: keep it small and simple). The final set of geostatistical models may be used as a baseline for (1) estimating the proportions of the various rock types within plutons (cf. Streckeisen, 1974, 1980), (2) estimating the number of samples needed to adequately characterize the mean modal composition of plutons, and (3) assessing the intrinsic uncertainty of the initial conditions for sediment-generation modelling from the sizes of samples used to characterize the properties of granitoid parent rocks.

### 3.2. Data transformations and model formulation

A transformation from proportions (or percentages) to centred log-ratios was carried out to remove unit-sum and positivity constraints from the raw data. The centred log-ratio transformation of a  $k$ -part composition  $\mathbf{p}$  is defined as (Aitchison, 1982, 1986):

$$\mathbf{y} = \text{clr}(\mathbf{p}) = \left[ \ln\left(\frac{p_1}{g(\mathbf{p})}\right), \dots, \ln\left(\frac{p_j}{g(\mathbf{p})}\right), \dots, \ln\left(\frac{p_k}{g(\mathbf{p})}\right) \right],$$

$$g(\mathbf{p}) = \left( \prod_{j=1}^k p_j \right)^{1/k}$$

And the inverse transformation from centred log-ratios to compositions (proportions) is:

$$\mathbf{p} = \text{clr}^{-1}(\mathbf{y}) = \left[ \frac{e^{y_1}}{s(\mathbf{y})}, \dots, \frac{e^{y_j}}{s(\mathbf{y})}, \dots, \frac{e^{y_k}}{s(\mathbf{y})} \right],$$

$$s(\mathbf{y}) = \sum_{j=1}^k e^{y_j}$$

Each data set belonging to a given pluton is defined as a matrix  $\mathbf{P}$ , which represents the proportions of each petrographic class at each of the control points. In matrix notation, its clr-transformation is:

$$\mathbf{Q} = \text{clr}(\mathbf{P})$$

Each data set was centred by column-wise subtraction of the vector mean:

$$\mathbf{Q}_c = \mathbf{Q} - \bar{\mathbf{Q}}$$

The centred data were subjected to a Singular Value Decomposition (SVD):

$$\mathbf{Q}_c = \mathbf{U}\mathbf{S}\mathbf{V}^t$$

The matrix  $\mathbf{U}$  contains the left singular vectors, which represent coordinates in the space of a set of orthogonal reference vectors  $\mathbf{V}$ , the right singular vectors (superscript  $t$  indicates transpose). The matrix  $\mathbf{V}$  contains the relation between the left singular vectors and the centred composition  $\mathbf{Q}_c$ . The singular values in the diagonal matrix  $\mathbf{S}$  are the square roots of the corresponding variances of the PCs and act as scaling factors which may be combined in various ways with  $\mathbf{U}$  and/or  $\mathbf{V}$  (Aitchison and Greenacre, 2002). For this study, the left singular vectors ( $\mathbf{U}$ ) were column-wise multiplied by the corresponding singular values ( $\mathbf{S}$ ) to obtain a set of coordinates of the data points in the space of  $\mathbf{V}$ . The columns of the scaled matrix of left singular vectors will be referred to as the PC scores, and the right singular vectors as the PC loadings.

Each successive PC describes as much of the total variance of the data as possible. The use of SVD as a dimension-reduction technique relies on the fact that most of the signal will be present in the first (largest) PCs, whereas the last (smallest) PCs contain most of the noise. Eliminating the smallest PCs from the SVD by setting their corresponding singular values to zero should thus simultaneously reduce the noise and the number of variables to be considered.

The matrix of centred log-ratios  $\mathbf{Q}$  is singular because each element of the input matrix of raw compositional data ( $\mathbf{P}$ ) has been used in the clr-transformation. The linear dependence of one concentration on the sum of all others implies that this matrix has been overspecified, because we could eliminate one of the components and still perfectly reproduce the data, knowing that they sum to a constant. The redundancy in the clr-transformation also manifests itself as a constant-sum constraint (the coefficients of each clr-vector sum to zero). This sum-to-zero constraint implies that the clr-vectors are located in a subspace of the Euclidian space, a so-called hyperplane passing through the origin. The beauty of the SVD is that it automatically identifies this hyperplane and eliminates it from the data. And therefore, the number of PCs needed to reproduce the input data ( $f$ ) is always one less than the number of variables:  $f = k - 1$ . The SVD ensures that the clr-transformed data can be expressed in terms of a set of  $k - 1$  PC scores which are unconstrained and globally uncorrelated.

The augmented matrix of georeferenced compositions  $\mathbf{Z}$  is defined by assigning spatial coordinates  $\mathbf{X}$  to each set of PC scores:

$$\mathbf{Z} = (\mathbf{X} \mid \mathbf{US})$$

We intend to provide a statistically valid description of  $\mathbf{Z}$  in the area of interest, for instance as a set of values at arbitrary grid nodes ( $\mathbf{X}^*$ ). The geostatistical model  $\hat{\mathbf{Z}}$  represents the estimates of each PC score obtained by OK:

$$\hat{\mathbf{Z}} = (\mathbf{X}^* \mid \hat{\mathbf{US}})$$

Geostatistical interpolation requires a variogram model that describes the spatial continuity (auto-covariance) of each variable (PC) as a function of the distance between control points (the lag). The choice of the variogram model and the size of the local neighbourhood adopted for interpolation (the search radius) are crucial elements of geospatial analysis that will be discussed in detail below.

### 3.3. Model evaluation

All operations carried out on the raw data  $\mathbf{P}$  are invertible and permit the geostatistical model  $\hat{\mathbf{Z}}$  to be expressed in terms of  $\mathbf{Q}$ . The result of this transformation will be denoted  $\hat{\mathbf{Z}}^*$ :

$$\hat{\mathbf{Z}}^* = (\mathbf{X}^* \mid \hat{\mathbf{Q}})$$

The matrix  $\hat{\mathbf{Z}}^*$  consists of estimates of the composition obtained by back-transformation of the interpolated PC scores  $\hat{\mathbf{U}}$  evaluated at arbitrary grid nodes within the rock body:

$$\hat{\mathbf{Q}} = \hat{\mathbf{U}}\mathbf{S}\mathbf{V}^t + \bar{\mathbf{Q}}$$

Any geostatistical model based on interpolation of  $\mathbf{Z}$  is an approximation of rank  $r$ , where the rank is equal to the number of PCs used. This approximation relates to the data as follows:

$$\tilde{\mathbf{Q}}_r = \mathbf{Q} + \hat{\mathbf{E}}_r, \quad r = 1, 2, \dots, f - 1$$

In this paper, models will be referred to by their abbreviation (BM, CP, DO, LC, or ML), the number of variables in the data ( $k$ ), and their rank ( $r$ ). For example, CP(6,2) refers to a geostatistical model of the Cornelia Pluton of rank two, based on six variables.

Any full-rank geostatistical model ( $r = f$ ) is an exact interpolator and reproduces the data at each control point:

$$\hat{\mathbf{Q}}_f = \mathbf{Q}, \quad \hat{\mathbf{E}}_f = \mathbf{0}$$

This relation is a fundamental property of geostatistical algorithms such as OK and has no bearing on their predictive power. Assessment of the extent to which the model can predict the values measured at each of the control points is possible with cross-validation. Cross-validation is carried out by removing one control point from the set and predicting its composition from the remaining control points using the parameters of the geostatistical model to be tested. This procedure is repeated for each control point to give a matrix of predicted compositions of rank  $r$ , termed  $\tilde{\mathbf{Q}}_r$ . In order to evaluate the quality of the predictions, we examine:

$$\mathbf{E}_r = \mathbf{Q} - \tilde{\mathbf{Q}}_r, \quad r = 1, 2, \dots, f$$

In this equation  $\mathbf{E}_r$  represents the matrix of residual compositions. The rank of the models investigated in this study ranges from 1 to  $f$  because even full-rank models will have non-zero cross-validation errors:

$$\tilde{\mathbf{Q}}_f \neq \mathbf{Q}$$

If needed, we may apply the inverse of the centred log-ratio transformation to obtain predicted petrographic compositions in terms of proportions at arbitrary grid nodes:

$$\hat{\mathbf{P}} = \text{clr}^{-1}(\hat{\mathbf{Q}})$$

This operation is required to produce so-called “single-component” maps (cf. [McKinley et al., 2016](#)).

#### 3.4. Model testing

Throughout the geostatistical modelling, we adopted thresholds and confidence levels of 95%, and all statistical tests were conducted at a significance level of 5% ( $\alpha = 0.05$ ). The first test detects bias in the candidate geostatistical models. The mean and covariance matrix of the residual compositions are calculated to test if:

$$\bar{\mathbf{E}}_r = \mathbf{0}$$

here,  $\bar{\mathbf{E}}_r$  is the vector mean of the residual compositions obtained from cross-validation,  $\mathbf{E}_r$ . The vector mean may be evaluated using Hotelling's one-sample  $T^2$  test ([Davis, 2002](#)). Any model for which the  $p$ -value is less than  $\alpha$  is considered biased and must be rejected.

The candidate models were also subjected to a permutation test designed to detect if the cross-covariance functions of the PCs of the residual compositions contain any remaining signals (see [Appendix B](#)). The cross-covariance functions of the PC scores of  $\mathbf{E}_r$  were obtained from the SVD of  $\mathbf{E}_r$ . For this test, only the cross-covariance at lags smaller than the search radius of the geostatistical model has to be taken into account, because control points beyond this radius will not

be used for interpolation and any non-zero cross-covariance at lags exceeding the search radius is irrelevant. If the cross-covariance within the search radius is significantly different from zero, the residual compositions display systematic spatial patterns, and the spatial structure of the data has not been adequately captured by the geostatistical model. Any model for which the test statistic  $T^*$  is greater than its critical value corresponding to the chosen  $\alpha$  falls within the critical region and must be rejected.

Any model which passes the above statistical tests adequately captures the spatial structure of the data and is therefore considered admissible. Its residuals may be modelled as a stationary random error, whose variance is taken to be equal to the mean squared prediction error (MSPE) obtained by cross-validation.

#### 3.5. Variogram modelling

The legacy data analysed in this study do not represent ideal cases in which short-range variability can be directly estimated from replicate analyses or sets of samples taken at short-range within the data set to be geo-statistically modelled. A superposition of short-range and long-range sampling has been used in none of the campaigns. Instead, sampling has been more or less uniform in space, which implies that distances of control points to their nearest neighbours are in the range of hundreds of meters. The petrographic data sets present a particular challenge because of the mismatch in scale between the support (a few  $\text{cm}^2$ , the area covered by a thin-section analysis) and the spacing between control points. Under these conditions, variogram modelling becomes highly speculative owing to a lack of control on short-range variability.

In some cases, replicate analyses from the same population are available, from which a measure of inter-specimen variance (i.e., the experimental error) may be calculated (see [Appendix A](#)). However, the spatial scale associated with these variance estimates is poorly constrained. It most likely spans two orders of magnitude (from centimeter to meter-scale), and its characteristic scale may, therefore, be taken as a decimeter. The sampling (counting) error in clr-space ([Bloemsma and Weltje, 2015](#)) represents the irreducible uncertainty of individual specimens. We examined the relation between experimental and counting errors in three plutons as a basis for selecting the most appropriate class of variogram models to be used. The analysis of replicates reported in [Appendix A](#) indicates that the short-range variability increases rapidly away from lag zero, which implies that the direction coefficient of the variogram model at the origin ( $h = 0$ ) is large, and an exponential variogram seems an appropriate choice. We opted for a combination of an omnidirectional exponential variogram model combined with a nugget model for each PC, and thus followed the recommendations of [Goovaerts \(1997\)](#) who suggested that the theoretical variogram model should aim to capture only the major spatial features, to minimize the risk of fitting possibly spurious details of the empirical variogram (overfitting).

Another important parameter is the search radius adopted in the geostatistical model because only those samples falling within the search radius will be used to obtain local composition estimates by weighted averaging. The search radius also influences the result of the permutation test for detecting the presence of unresolved cross-covariance, because only sample pairs at lags smaller than the search radius have to be taken into account. Significant cross-covariance at greater lags does not influence the geostatistical model and is thus irrelevant. The range of possible search radii is constrained by the spatial arrangement of control points. An effective minimum search radius was calculated based on the restriction that the compositions of at least 95% of the control points can be predicted from at least three other control points. The maximum search radius equals the largest distance between control points in the data set. Setting the search radius equal to its maximum value implies that all control points will be used simultaneously for interpolation. If a pure nugget model is used in conjunction

with the maximum search radius, geostatistical interpolation returns the average composition of the data set at each grid node.

The sills of all variogram models were set equal to the squared singular values of the corresponding PCs, divided by  $n - 1$ . Because exponential variogram models do not have a sill at a finite lag distance, we defined the range of the variogram model as the distance at which the predicted variance reaches 95% of the sill (Isaaks and Srivastava, 1989). After an initial phase of extensive experimentation in which different variogram models and search radii were used, we decided to streamline the modelling by adopting a few restrictions. A single range was used in all variogram models of the PC scores from a single pluton. Furthermore, the search radius was taken equal to the range because the initial experiments showed that kriging predictions were insensitive to enlarging the search radius beyond this value, whereas the MSPE tends to increase for search radii smaller than the range.

### 3.6. Optimization and model selection

The above choice of candidate variogram models and search radii to be considered reflect a pragmatic and hopefully valid simplification of the geostatistical modelling problem to be solved. In this heuristic framework, the only unknown is the nugget to be assigned to the variogram of each PC. As the nugget cannot be predicted from theory and data of sufficient quality to describe the variance as a function of spatial scale at the short-range are unavailable, variogram modelling may be approached as an optimization problem. For each PC, we estimated the nugget by minimizing the MSPE obtained by cross-validation of the PC scores. The nugget was allowed to vary between zero and the sill of the variogram model.

A family of geostatistical models can be constructed from the set of PCs by combining them in all possible ways. Generally, one would expect distinct large-scale compositional trends within a pluton to be captured by the first few PCs, as the higher PCs describe a relatively small proportion of the total variance and the data are quite noisy (as explained in Appendix A). But it should be borne in mind that there is no compelling need for PC scores to correlate with spatial coordinates, and it is, therefore, conceivable that the first (largest) PCs do not even carry a significant spatial auto-covariance signal. This line of reasoning implies that all possible combinations of PCs should be evaluated.

The best geostatistical model of each data set was selected from the admissible models of each pluton, i.e., those whose residuals passed the two statistical tests outlined above. The operational definition of the best geostatistical model for each of the five plutons examined in this study was based on the trade-off between their MSPE values and the number of PCs used in their construction (their rank). In cases where two or more models of a given pluton had approximately identical MSPE values, the model with the smallest number of PCs (lowest rank) was preferred, following the principle of parsimony, consistent with our heuristic approach to solving the geostatistical problem at hand. The resulting set of selected models has been visualized and subjected to further analysis.

### 3.7. Visualization

Biplots were constructed to illustrate the properties of the geostatistical models in compositional (clr) space (Aitchison and Greenacre, 2002). Two types of biplots are in common use. Form biplots show the scores of the first two PCs ( $\mathbf{U}$ ) scaled by their singular values ( $\mathbf{S}$ ), as coordinates in the space of  $\mathbf{V}$ . The distances between data points in form biplots are a direct measure of their compositional dissimilarity. The form biplots presented in this study show the control points together with the OK predictions of a high-resolution grid (comprising between 100,000 and 200,000 nodes), colour coded according to QAPF class. The biplots were constructed using the control points only, and high-resolution grids of the OK surfaces were projected into the space of  $\mathbf{V}$  to

illustrate the relations between the systematic spatial components of the geostatistical models and the control points.

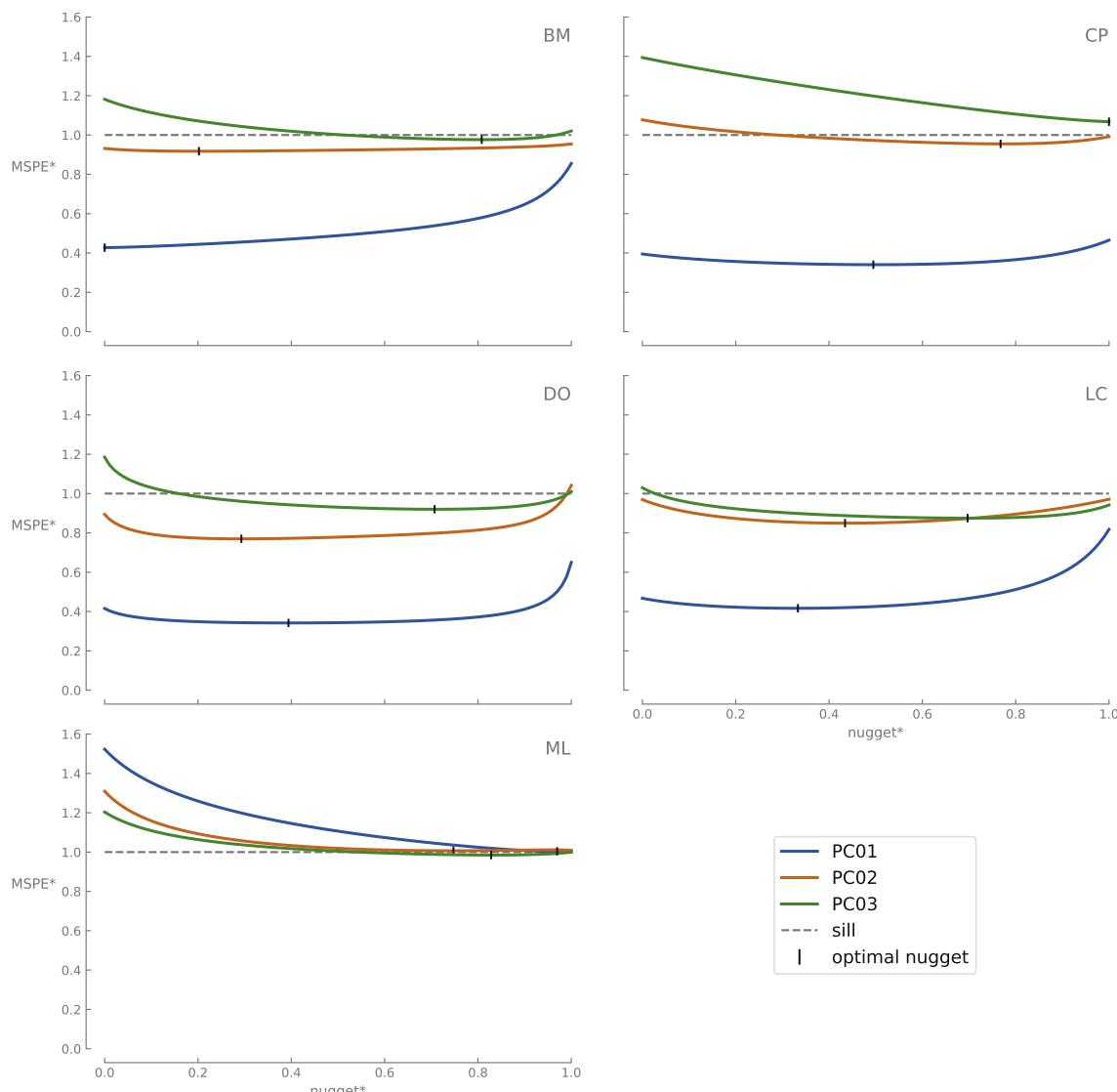
The relations between the original set of variables and the PCs are shown in the form of arrows (loadings). Loadings are most easily interpreted in covariance biplots, that show the scores of the first two PCs ( $\mathbf{U}$ ), as coordinates in the space of  $\mathbf{V}$  scaled by  $\mathbf{S}$ . The loadings in covariance biplots may be interpreted in terms of the distances (links) between arrowheads (Aitchison and Greenacre, 2002; van den Boogaart and Tolosana-Delgado, 2013). If the link between two arrowheads is very small, the corresponding log-ratio is nearly constant. If three (or more) arrowheads fall on a straight line, the corresponding sub-composition may be one-dimensional. i.e., the variables are perfectly (negatively and positively) correlated. Two orthogonal links strongly suggest that the two (sets of) log-ratios are statistically independent. The extent to which these conclusions apply depends on the proportion of the variance described by the first two PCs: if this is less than ~95%, each of these must be checked by further analysis.

The main spatial trends in the modal composition of each pluton have been visualized in the form of contour maps of the four petrographic classes constructed from the high-resolution OK grids (so-called “single-component” maps). An additional set of maps was prepared from the high-resolution OK grids to visualize the spatial distribution of QAPF classes within each pluton. Because the spatial arrangement of control points differs from one data set to another, there will be regions of the grid where no estimate can be obtained because they fall outside of the minimum search radius. Such regions of the grids were automatically blanked. The outlines of the geostatistical models displayed on the maps thus reflect the data density of the surveys and not necessarily the spatial extent of the plutons in outcrop.

### 3.8. Sampling

Samples were taken from the geostatistical models by randomly selecting points (coordinate pairs) falling within each pluton. The composition at each randomly selected coordinate pair was simulated as follows. For each PC participating in the geostatistical model, the local value was obtained by OK. The OK prediction represents the local mean of a Gaussian distribution whose variance was taken equal to the largest of two measures of uncertainty: the nugget and the MSPE. For PCs which do not participate in the model, the mean of the Gaussian distribution is equal to zero, and the variance equal to its squared singular value divided by  $n - 1$  (i.e., the sill). The predicted composition at each point was obtained by independent random draws from these Gaussians to produce a set of PC scores, followed by application of the inverse transformation discussed above. The result of this operation is a virtual specimen from the pluton. Note that this pointwise simulation is not a geostatistical simulation sensu stricto. It is distinctly different from true geostatistical simulation strategies such as Sequential Gaussian Simulation (SGS), in which simulated samples obtained from randomly chosen locations are added to the data and used in subsequent kriging operations to obtain a probabilistic realisation of the entire field (Johnson, 1987; Ripley, 1987). Our straightforward sampling strategy implies that the variance of the model will be slightly overestimated. This prevents us from providing an overly optimistic view of the minimum requirements for an adequate sampling programme.

The number of rock specimens needed to estimate the “true” (area-weighted) mean composition of a pluton as obtained from a high-resolution OK grid was determined by taking a large number of spatially random samples of various sizes from the models of each pluton according to the above procedure. The degree to which the mean composition of a sample deviates from the “true” value was expressed in terms of the mean squared deviation (MSD; the variance relative to the area-weighted mean value). For each sample size, 100 simulations were carried out. The sample sizes needed to approximate the “true” average



**Fig. 3.** Nugget optimization for each pluton (data sets with four variables). The nugget and the MSPE were normalized by the sills of the variograms to obtain nugget\* and MSPE\*.

with a precision equal to point counts of various sizes ( $N$ ) were extracted from the simulations to provide a guideline for evaluating the sampling design used in the five case studies.

## 4. Results

### 4.1. Geostatistical modelling

Because a large number of geostatistical models has been produced within the framework of this study, we cannot possibly show and tabulate all of them. Instead, we have made a selection of results to illustrate the modelling process and its most relevant outcomes, which will serve as a basis for discussion of the main results and the heuristics used.

The results of the variogram modelling of the homogenized data sets ( $k = 4$ ) are shown in Fig. 3. It appears that the predictive power of the geostatistical models does not depend strongly on the choice of nugget value, as demonstrated by the lack of a well-defined minimum of the curves (broad valleys), which have been normalized to the values of the sills for visual comparison. Fig. 3 shows that the nugget values which minimize the MSPE are typically in the range of 50% of the values of

the sills, which underscores the large proportion of noise in the modal data.

Table 2 lists the parameters of the best-fit variograms for all models. An initial estimate of the rank of the geostatistical models can be made by comparing the sills to the MSPEs of the PCs. Only PCs whose MSPE is smaller than the sill will contribute to explaining spatial trends in the plutons. The likely rank of the models is also mirrored in the cumulative proportion of the variance (CPV) described by the PCs if we assume that any PCs that describe less than ~5% of the total variability may be discarded. For the homogenized data sets comprising four variables, both lines of evidence suggest that geostatistical models of rank 1 are appropriate for BM and CP, whereas DO and LC are best represented by models of rank 2. The above reasoning does not apply to ML, as none of the PCs of this pluton seems to have any predictive power (sills and MSPEs are roughly equal), even though the CPV suggests that a model of rank 2 would be needed to represent its compositional variability. Fig. 3 shows that only a pure nugget model of the third PC might be capable of extracting some spatial information from the data. The variogram models based on the six-part compositions of BM and CP suggest that their geostatistical models should be of rank 3 and 2, respectively.

**Table 2**

Variogram modelling. Data set ( $k$  = number of variables), CPV = cumulative proportion of variance described by the PCs. For each pluton/data set, sill were set equal to the variance of PC scores, a single range was visually estimated, the search radius was set equal to the range, and the nugget was obtained by minimizing the MSPE of each set of scores (see Fig. 3). PCs whose MSPEs exceed their sill do not contribute to resolving the spatial structure of the plutons. See text for discussion.

Data set ( $k$ )	PC	CPV	Sill	Range [m]	Nugget	MSPE
BM(4)	1	0.972	1.661	12,000	0.000	0.710
	2	0.996	0.040	12,000	0.007	0.036
	3	1.000	0.008	12,000	0.006	0.007
CP(4)	1	0.934	0.575	1800	0.284	0.196
	2	0.980	0.028	1800	0.021	0.027
	3	1.000	0.012	1800	0.012	0.012
DO(4)	1	0.739	1.315	12,000	0.518	0.450
	2	0.949	0.375	12,000	0.113	0.288
	3	1.000	0.090	12,000	0.064	0.083
LC(4)	1	0.682	3.295	15,000	1.098	1.345
	2	0.974	1.412	15,000	0.599	1.199
	3	1.000	0.126	15,000	0.088	0.110
ML(4)	1	0.713	0.653	1500	0.633	0.655
	2	0.939	0.208	1500	0.159	0.208
	3	1.000	0.056	1500	0.044	0.055
BM(6)	1	0.846	1.928	12,000	0.000	0.827
	2	0.920	0.170	12,000	0.113	0.159
	3	0.977	0.129	12,000	0.027	0.107
CP(6)	4	0.997	0.044	12,000	0.044	0.044
	5	1.000	0.008	12,000	0.006	0.008
	1	0.494	2.201	1800	1.245	1.274
CP(6)	2	0.870	1.678	1800	1.034	1.199
	3	0.991	0.539	1800	0.447	0.532
	4	0.998	0.031	1800	0.019	0.031
	5	1.000	0.010	1800	0.010	0.011

**Table 3** summarizes the results of all possible models of the homogenized data, as well as a selection of models for BM and CP based on the six-part compositions. All models passed Hotelling's one-sample  $T^2$  test, indicating that the vector mean of the residual compositions is not significantly different from zero at a significance level  $\alpha = 0.05$  and all models are globally unbiased. Hence, the MSPE values in relation to the ranks of all permissible models ( $T^* < 1$ ) may be used to select the best model for each data set.

For BM(4,.) and CP(4,.), the choice is between models of rank 1 and 2. The gain in predictive power obtained by raising their ranks from 1 to 2 is negligible because in both cases, the MSPE values are reduced by only 0.4%. The reason we consider this negligible is that the MSPE, which represents a measure of the uncertainty, is itself an estimate and thus also subject to uncertainty. The uncertainty of the MSPE stems from the particular layout of control points. It is conceivable that leaving out one control point, or for that matter, even the decision to cut one of the thin sections from a different part of a hand specimen is already sufficient to provoke such a small change in the MSPE. Hence, BM(4,1) and CP(4,1) appear to be the best choices. Similar reasoning leads to the conclusion that DO(4,2) and LC(4,2) are the best models for these two plutons, because the gain in predictive power obtained by raising their ranks from 2 to 3 reduces their MSPE values by only 0.8% and 0.6%, respectively. For three out of these four models, the variance of the OK surface, which describes the spatial component of the model, is larger than the MSPE. Only for LC, this is not the case, indicating that this model is noisier than that of the other three. ML is the odd one out, as it appears that a geostatistical model based only on the third PC gives the smallest MSPE, but this value is still so much larger than the spatial component of the model (the variance of the OK surface) that the latter can be safely ignored. This implies that geostatistical modelling of ML has no added value whatsoever, and the only justifiable model for this pluton would be ML(4,0), i.e., a model employing a spatially constant composition with superimposed noise.

Leaving aside the pathological case of ML, the results of model

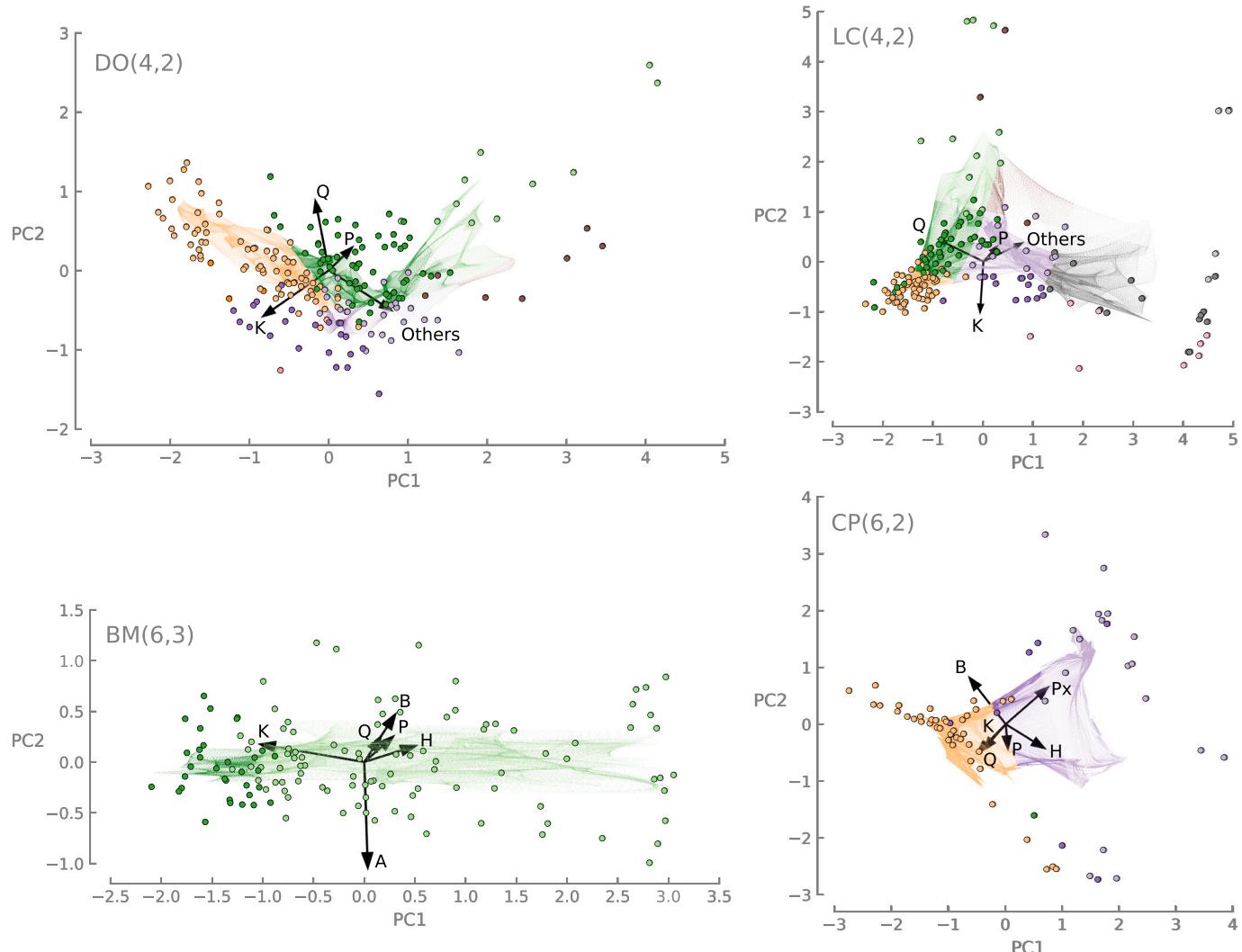
**Table 3**

Model selection. VAR(OK) = variance of kriging surface;  $T^*$  = test statistic for significant cross-covariance of residual compositions at  $\alpha = 0.05$ . Only models with  $T^* < 1$  are admissible. Selected models BM(4,1), CP(4,1), DO(4,2), LC(4,2), ML(4,1), BM(6,3), and CP(6,2) are marked in bold face. See text for discussion.

Model	r	PCs	VAR(OK)	MSPE	$T^*$
BM(4,.)	1	<b>1</b>	<b>1.240</b>	<b>0.757</b>	<b>0.56</b>
	1	2	0.015	1.692	0.98
	1	3	0.000	1.695	1.39
	2	12	1.255	0.754	0.74
	2	13	1.240	0.757	0.53
	2	23	0.015	1.692	0.82
	3	123	1.255	0.754	0.70
	CP(4,.)	1	<b>0.336</b>	<b>0.236</b>	<b>0.40</b>
	1	2	0.003	0.605	0.43
	1	3	0.000	0.607	0.50
	2	12	0.338	0.235	0.45
	2	13	0.336	0.237	0.39
DO(4,.)	2	23	0.003	0.606	0.39
	3	123	0.338	0.236	0.45
	1	1	0.874	0.912	1.52
	1	2	0.129	1.687	1.06
	1	3	0.004	1.764	1.26
	2	<b>12</b>	<b>1.003</b>	<b>0.828</b>	<b>0.93</b>
	2	23	0.133	1.680	0.82
	3	123	1.007	0.821	0.93
LC(4,.)	1	1	1.699	2.902	1.09
	1	2	0.344	4.601	1.31
	1	3	0.017	4.790	0.86
	2	<b>12</b>	<b>2.043</b>	<b>2.698</b>	<b>0.95</b>
	2	13	1.716	2.887	1.02
	2	23	0.361	4.586	0.94
	3	123	2.060	2.683	0.96
ML(4,.)	1	1	0.004	0.917	0.82
	1	2	0.007	0.913	0.84
	1	<b>3</b>	<b>0.003</b>	<b>0.909</b>	<b>0.85</b>
	2	12	0.011	0.920	0.84
	2	13	0.007	0.916	0.85
	2	23	0.010	0.912	0.87
	3	123	0.014	0.919	0.86
BM(6,.)	1	1	1.402	1.175	0.90
	2	12	1.419	1.166	0.94
	3	<b>123</b>	<b>1.462</b>	<b>1.145</b>	<b>0.99</b>
	4	1234	1.463	1.145	0.98
	5	12345	1.463	1.145	0.97
CP(6,.)	1	1	0.886	3.499	1.16
	2	<b>12</b>	<b>1.343</b>	<b>3.045</b>	<b>0.81</b>
	3	123	1.372	3.045	0.80
	4	1234	1.375	3.046	0.79
	5	12345	1.375	3.047	0.79

selection according to the MSPE thus confirm the inferences from the variogram modelling (Table 2). It also indicates that only cumulative models have to be taken into account, i.e., models based on the first PC or the first two PCs (e.g. BM and CP), or models based on the first two or three PCs (e.g. DO and LC). This was confirmed by running all possible models based on the six-part compositions of BM and CP (31 for each of these data sets). Therefore, only the five cumulative models for BM and CP have been listed in Table 3. According to the model-selection rules set out above, the best models are BM(6,3) and CP(6,2), again confirming the inferences from the variogram modelling (Table 2). Raising  $k$  from 4 to 6 results in a decrease of the signal-to-noise ratio of the models, indicating that inclusion of comparatively rare mineral classes that are noisier than major constituents is not necessarily a good idea.

Fig. 4 shows form biplots that depict the geostatistical models of the plutons in compositional space. Colour coding of data points and grid nodes according to the QAPF classification facilitates the comparison between the two. In all cases shown, the large nugget values (around 50% of the sill) result in a limited range of variability of the geostatistical models (the OK surfaces) relative to the noisy data.



**Fig. 4.** Form biplots of DO(4,2), LC(4,2), BM (6,3), and CP (6,2) showing the control points and the nodes of a high-resolution grid projected into the space of V, colour-coded according to QAPF class (see legend in Fig. 5). Arrows represent loadings of variables onto PCs.

Fig. 5 illustrates the geostatistical models of the five plutons based on the homogenized data sets comprising four variables in the form of QAPF maps, which depict the systematic spatial components of the models obtained from the OK surfaces. As expected, the model of ML does not show any spatial variation.

Figs. 6 and 7 highlight differences between alternative models of the same pluton in QAPF format. Binary difference maps are grey for grid nodes assigned to the same QAPF class, and black for grid nodes where predicted QAPF classes differ. The binary difference maps of DO and LC are based on models of rank 2 and 3, whereas those of BM and CP are based on the full data sets ( $k = 6$ ) and the homogenized data sets ( $k = 4$ ). The alternative QAPF models are also shown for comparison. Overall, differences are small considering the large amount of noise in the data, which indicates that the models are fairly robust.

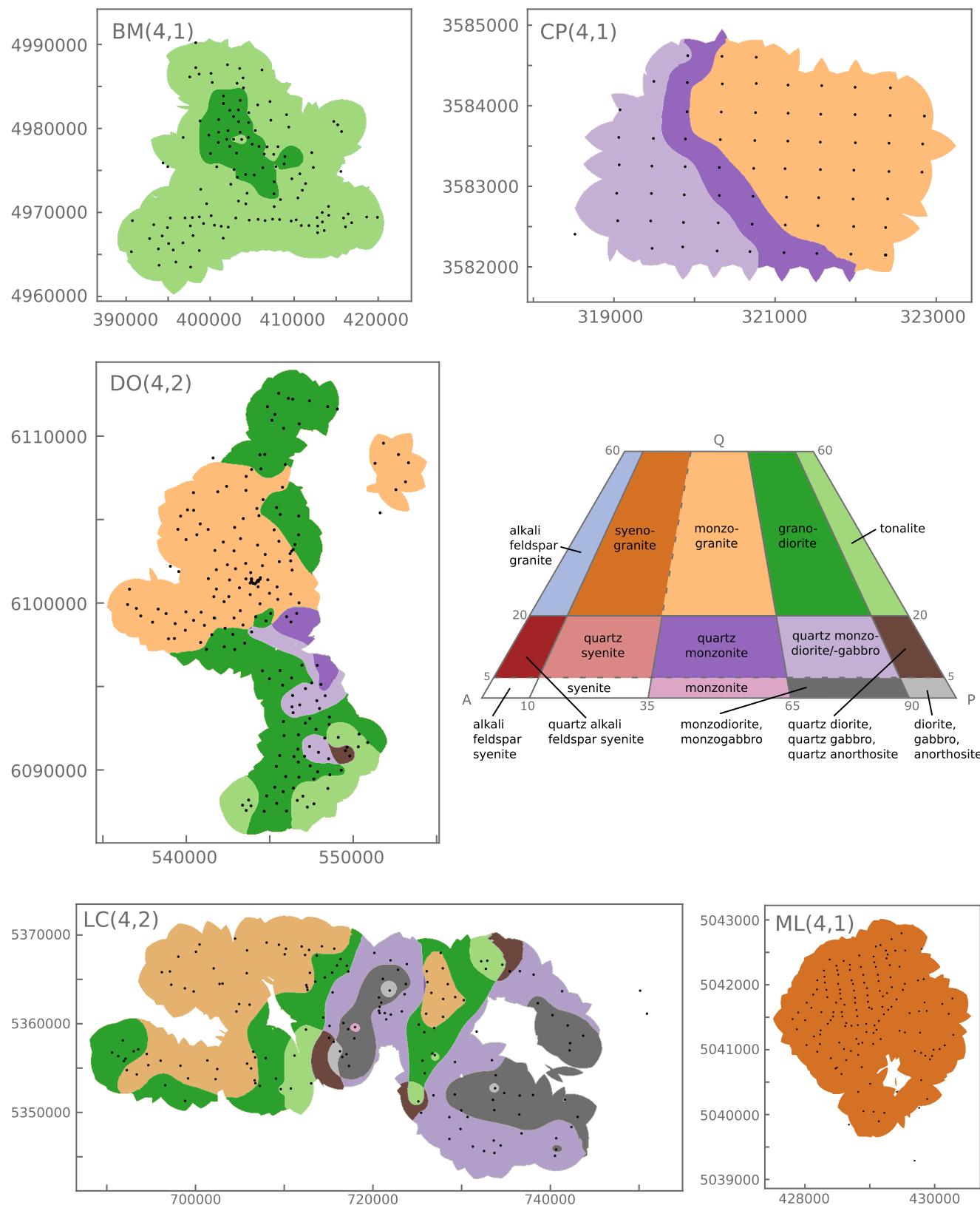
Even though the QAPF maps provide an attractive way to visualize the results of the geostatistical modelling, it should be borne in mind that they represent a discrete (and arbitrary) classification based on the ternary sub-composition Q-K-P, that does not bring out all the details of the models. The full output of the geostatistical modelling is shown in Figs. 8 and 9 in the form of so-called single-component maps of DO(4,2) and LC(4,2), respectively. These figures provide a detailed view of spatial compositional variation and permit visualization of robust trends within the plutons from which the overprint of short-range heterogeneity (“noise”) has been removed. The two plutons selected for

display are quite heterogeneous, which likely reflects their polygenetic nature.

#### 4.2. Fractionation patterns

Fig. 10 shows the loadings of the four variables of the DO and LC plutons in covariance biplots, which sheds light on the apparent correlations among the closed-sum variables in the single-component maps of Figs. 8 and 9. The arrowheads of Q, P, and Others fall almost perfectly on a straight line (link) in both biplots, which strongly suggests that the ternary sub-compositions Q-P-Others are 1-dimensional (collinear) in both plutons.

To illustrate the inferences drawn from the covariance biplots, we subjected both data sets of sub-compositions Q-P-Others to a clr-transformation. We centred each matrix by subtracting its column means and then subjected it to an SVD. The two PCs of pluton DO describe 88% and 12% of the total variance, respectively, which is not a spectacular result. However, the two PCs of pluton LC describe 96% and 4% of the total variance, respectively, confirming our inference based on the covariance biplot that we are dealing with a 1-dimensional (collinear) sub-composition resulting from fractionation. We will explore the implications of this result in the section on exploratory petrogenetic modelling below.

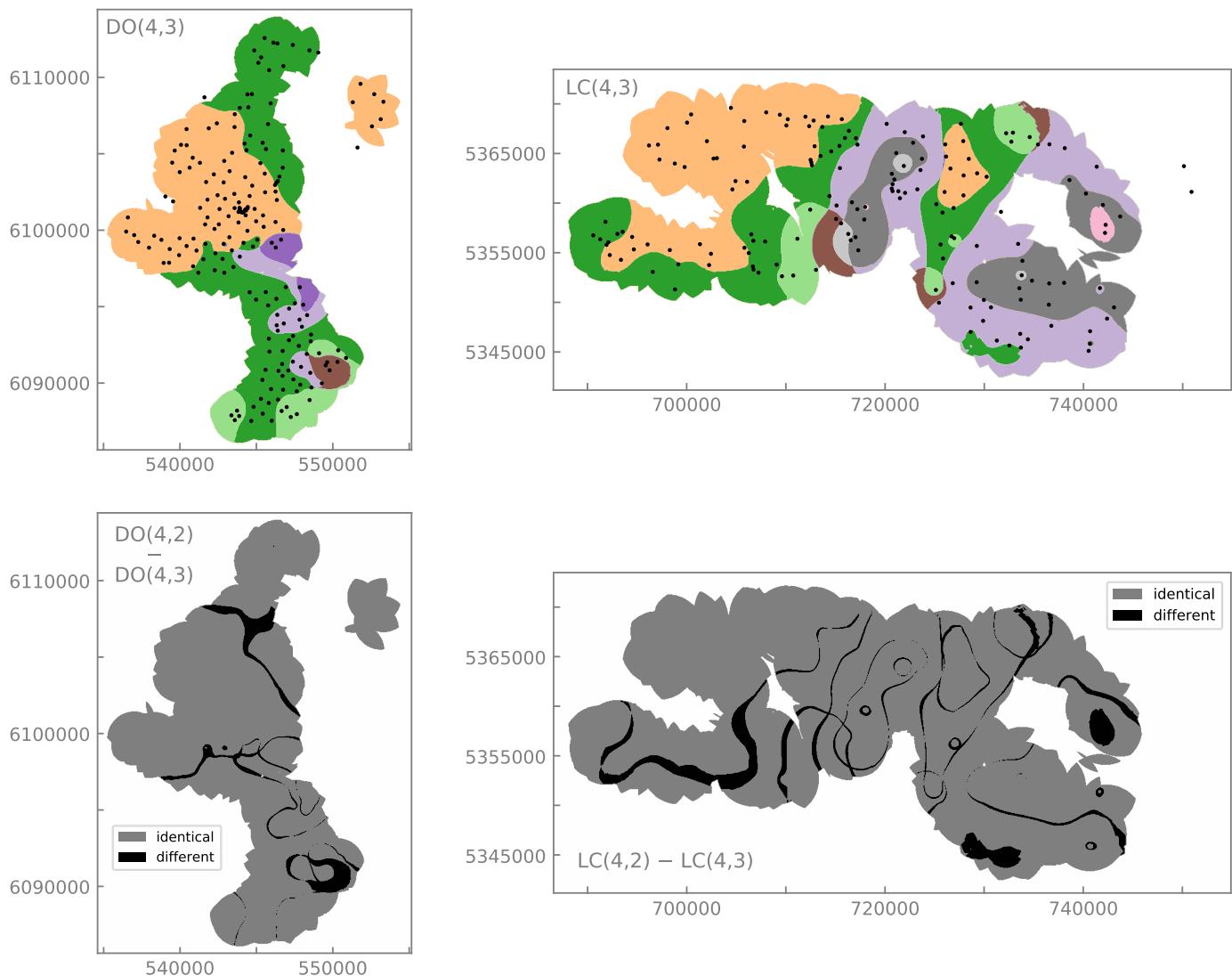


**Fig. 5.** Spatial distribution of QAPF classes within BM(4,1), CP(4,1), DO(4,2), LC(4,2), and ML(4,1).

#### 4.3. Sampling

The sample size needed for an adequate approximation of the area-weighted (“true”) mean composition of each pluton is illustrated with

two figures. In Fig. 11, the empirical variograms, the exponential variogram models, and a series of 100 simulated empirical variograms based on the same sample size as the actual number of control points are shown for BM(4,1). Even though the shapes of the empirical



**Fig. 6.** Comparison of models with different ranks. Spatial distribution of QAPF classes within DO(4,3) and LC(4,3), shown together with binary difference maps relative to corresponding models of Fig. 5.

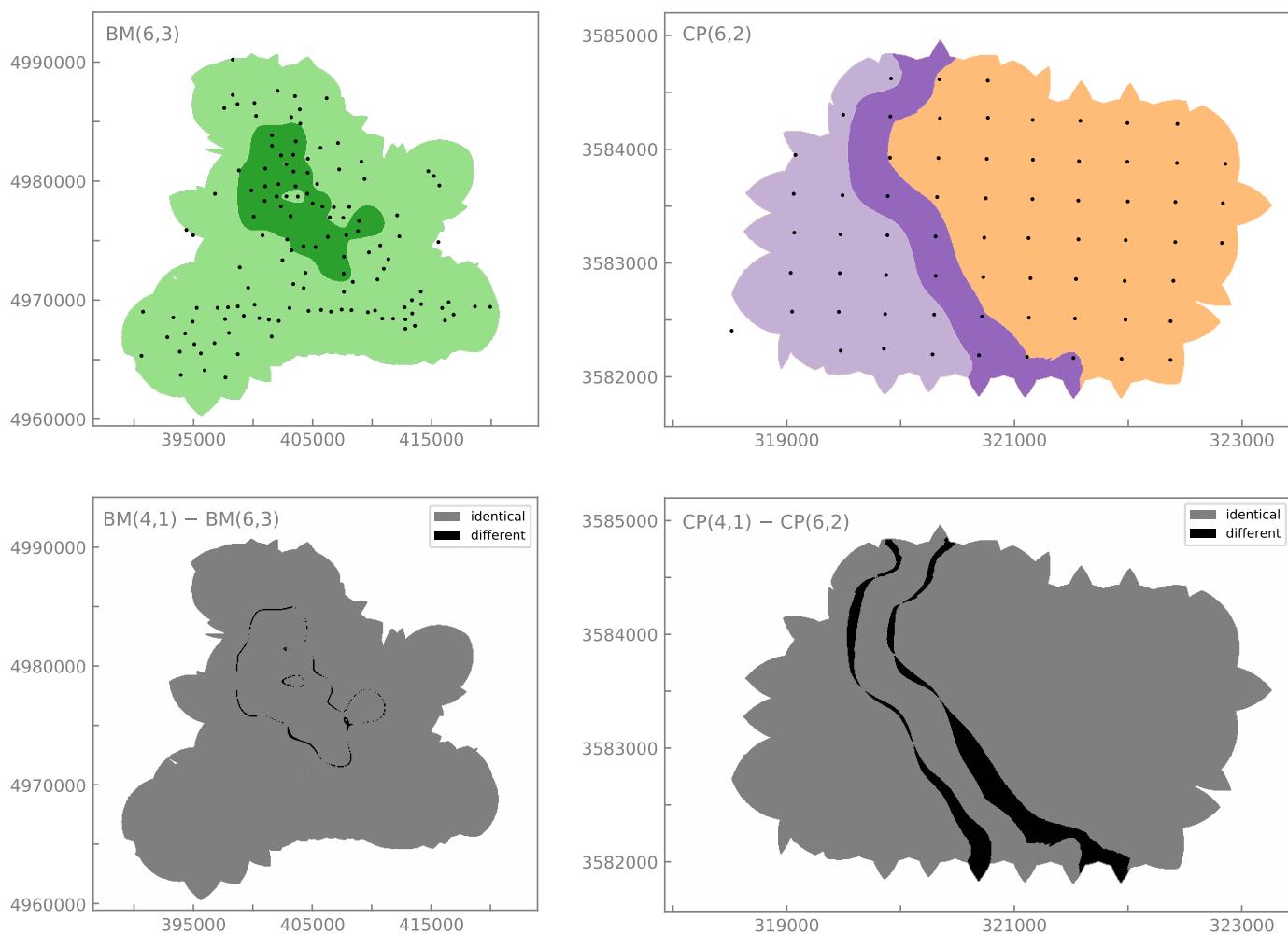
variograms are fairly well reproduced, it is clear that our simulation strategy overestimates the variance of the data, and as explained above, this results in a conservative estimate of the required sample size. A full overview of simulated random sampling from the models shown in Fig. 5 is given in Fig. 12. For each of the five plutons (seven data sets), the sample size needed to adequately estimate the area-weighted mean composition of the plutons has been calculated for three hypothetical values of the count length  $N$ . The operational definition of an adequate approximation is an estimate whose mean squared deviation (MSD) from the area-weighted mean is equal to the multinomial counting error in clr-space (MCE). Table 4 summarizes the results of the simulations and also lists the simulation results of BM and CP based on the six-part compositions.

## 5. Discussion

### 5.1. QAPF maps

To verify the geostatistical models of the five plutons, the QAPF maps of Fig. 5 were compared to the geological maps published in the original references (Fig. 2). It is worth noting that the QAPF classification (Streckeisen, 1974, 1980) had not been published when the original maps were made (except for ML). For the Bald Mountain

batholith (BM), the interpolation corresponds very well to of Taubeneck (1957), showing almost exclusively tonalite on the QAPF map except for the middle part of the batholith where the Anthony Lake granodiorite is present. Although the QAPF classes in of Wadsworth (1968) deviate from those of Streckeisen (1974, 1980), the geological map of the Cornelia pluton (CP) is well represented by the geostatistical model. Finer details are lost, however, since control points are located on a grid, and no textural information is present in the data. The model of the Donegal granite (DO) is more difficult to compare to the original geological maps since Whitten (1955, 1957, 1959, 1961) does not refer to a QAPF-type classification. Most of the monzogranite, granodiorite, and tonalite areas in the interpolations correspond to lithologies reported for the Thorr granite (Stevenson et al., 2008). The map of DO has been modified from Atherton and Ghani (2002). The granitic complex of Lacorne, La Motte and Preissac (LC) on the QAPF maps compares well to in Dawson and Whitten (1962). Lithologies other than granitoids are also present in this complex but were not captured by the QAPF interpolation since all control points are located within the boundaries of the granitic complex. The aplites of the Meech Lake pluton (ML) are present as syenogranite, alkali feldspar granite, and quartz alkali feldspar syenite on the map of control points (Fig. 2). Their occurrence coincides with of Vistelius et al. (1983), but lithologies other than aplite are not represented in the interpolations. This can be



**Fig. 7.** Comparison of models with different numbers of variables. Spatial distribution of QAPF classes within BM(6,3) and CP(6,2), shown together with binary difference maps relative to corresponding models of Fig. 5.

attributed to the fact that nearly all control points are located within the aplite bodies. From the geological description in Vistelius et al. (1983), it is clear that ML does display small-scale variability, but the sampling scheme does not permit any of this to be resolved. The difference maps (Figs. 6 and 7) clearly show where additional control points should be acquired if the objective of the geostatistical modelling exercise is to produce a reliable QAPF map: additional samples should be taken in areas where the black bands at the transition from one QAPF class to another are particularly wide, as in CP, or the eastern part of LC.

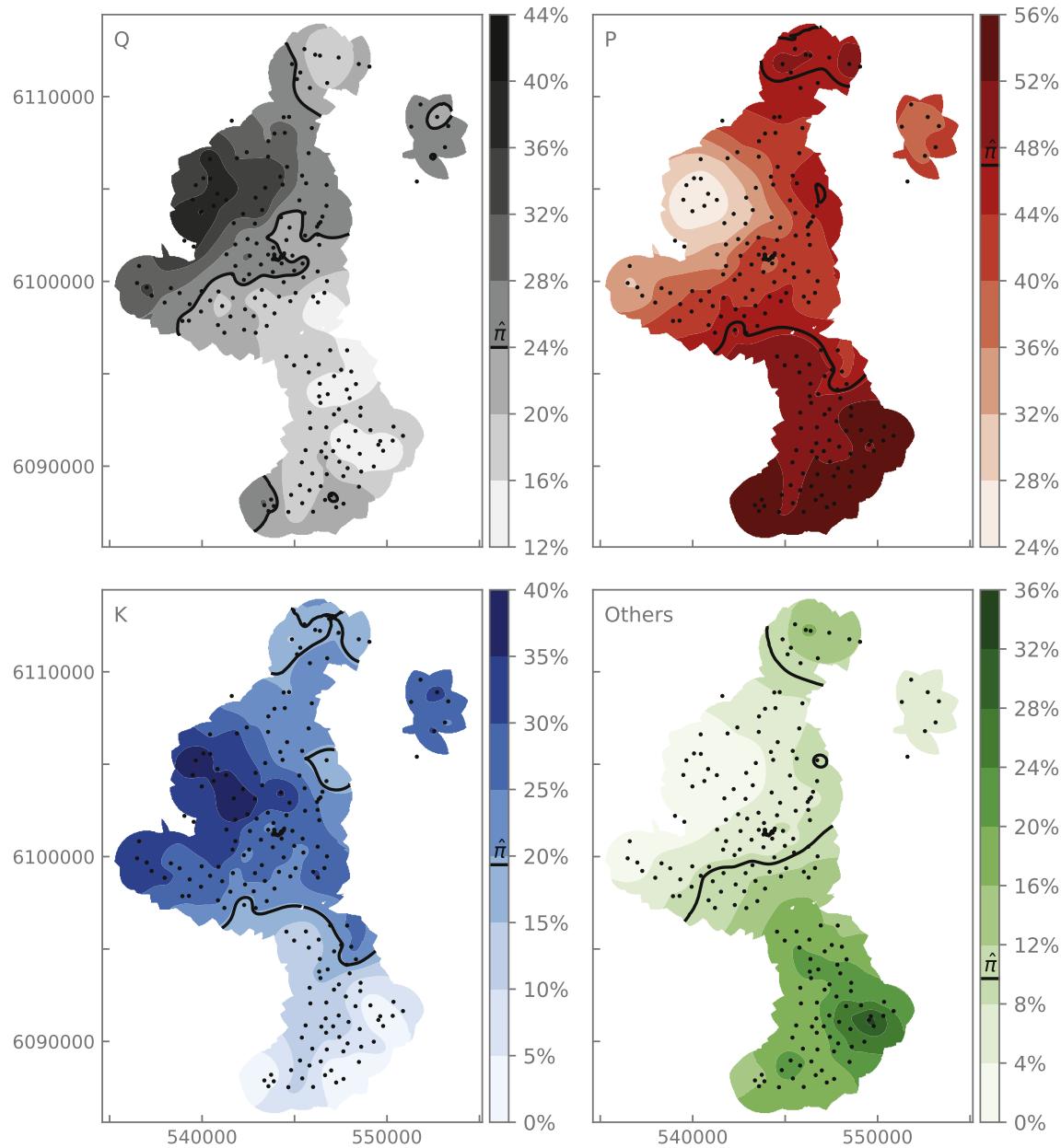
## 5.2. Single-component maps

The single-component maps of mineral-class percentages represent the output of the geostatistical models in their most detailed form and are thus a convenient illustration of the spatial compositional gradients within each pluton. Besides, the single-component format is much easier to digest than maps of PC scores of clr-transformed variables. The maps shown in Figs. 8 and 9 have been derived in the clr space. They were subsequently transformed back to “concentration space” and thus do not suffer from the problems associated with conventional single-component maps sensu McKinley et al. (2016), which are constructed by univariate interpolation of raw compositional variables (columns of  $\mathbf{P}$ ) with all the problems this entails. Consequently, conventional single-component maps violate the basic requirements of a valid statistical description of compositional data. Their values are not constrained to

be positive and the sum of values of the individual components predicted at each grid node is not equal to unity (or 100%). Conventional single-component maps, which are still commonly produced from geochemical surveys, are thus inadmissible and should be abandoned.

The single-component maps of DO (Fig. 8) and LC (Fig. 9) permit some refinement of the QAPF maps (Fig. 5). For ease of interpretation, the area-weighted mean composition has been displayed in these maps as a contour line representing the corresponding concentrations, which permits an apparently straightforward subdivision of both complexes into areas of different composition. The single-component maps of both plutons seem to indicate a strong positive correlation between the concentrations of Q and K, as well as between those of P and Others. DO displays the strongest compositional contrasts between its north-western-central part, which is largely made up of monzogranite, and the southern- and northern-most parts, which are made up of granodiorite and tonalite. Similarly, LC appears to be made up of four subareas. Its westernmost part mainly consists of monzogranite and granodiorite and its easternmost part of quartz monzonite and quartz monzodiorite. The central area of the LC complex may be subdivided into two parts: a western-central area that is similar to the eastern part, and an eastern-central area that is similar to the western part.

A note of caution is in order, however, because the inverse transformation from clr-space to concentration space needed to display the model in the form of single-component maps re-introduces the constant-sum constraint on compositions. It is therefore impossible to infer petrogenetic relationships from single-component maps (or cross-plots



**Fig. 8.** Single-component maps of DO(4,2). Black contour corresponds to the concentration values obtained from inverse clr-transformation of the area-weighted mean composition ( $\hat{\pi}$ ).

of concentrations, for that matter). Unfortunately, this is still common practice because many researchers are unaware of the limitations to interpretation imposed by the nature of compositional data. These limitations apply equally to cases in which single-component maps have been produced by inverse transformation of a statistically sound model, as in Figs. 8 and 9. The endemic (and easily avoidable) problems of conventional petrological data analysis based on concentrations will be brought into sharper focus in the next section.

### 5.3. Exploratory petrogenetic modelling

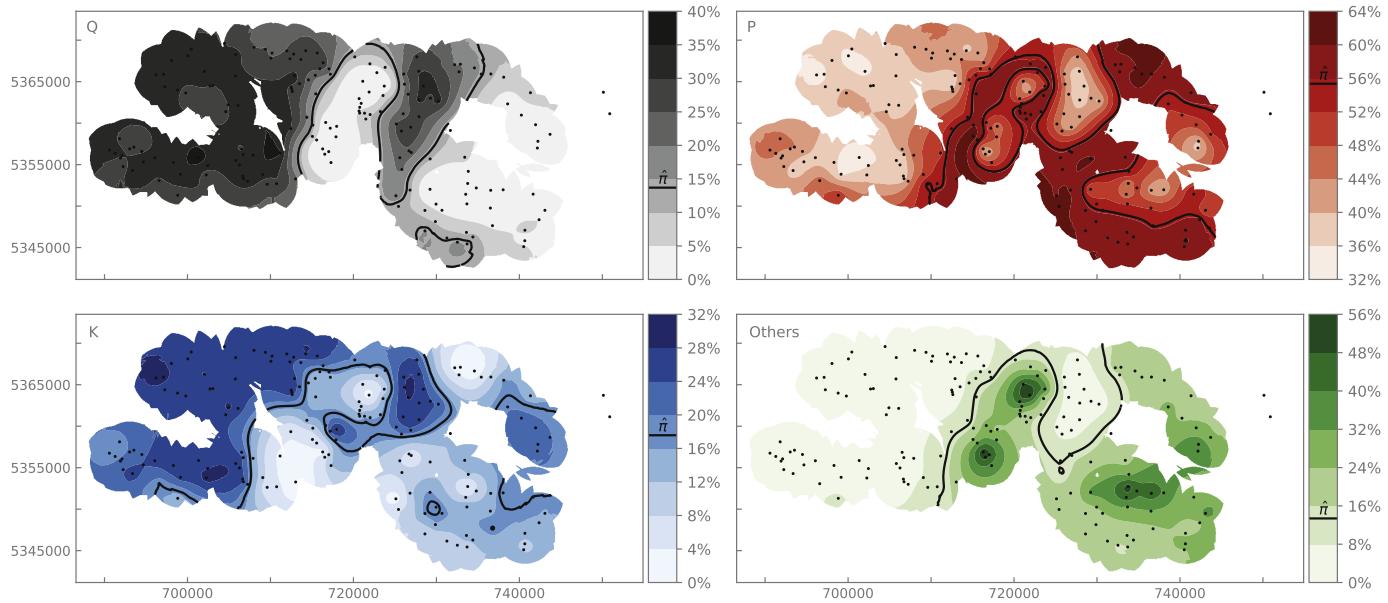
To illustrate how the log-ratio approach to compositional data analysis may aid in setting up a petrogenetic model, we return to the analysis of the Q-P-Others sub-composition of pluton LC (Fig. 10d). If we define the matrix  $\mathbf{R}$  as the clr-transformed set of ternary sub-compositions, SVD permits each row vector of this matrix to be written as:

$$\mathbf{r}_j = \bar{\mathbf{r}} + f_j \mathbf{v}_1 + g_j \mathbf{v}_2, \quad j = 1, 2, \dots, n$$

where  $\mathbf{r}_j$  is the  $j^{\text{th}}$  row of  $\mathbf{R}$ ,  $\bar{\mathbf{r}}$  is the vector of column means of  $\mathbf{R}$ . The vectors  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are the first and second eigenvectors (PCs) of the centred matrix  $\mathbf{R}$ , respectively, whereas  $f_j$  and  $g_j$  are the scaled PC scores of the  $j^{\text{th}}$  row of the centred matrix  $\mathbf{R}$ .

Because the last term on the right-hand side describes a very small part of the total variance (4%), we will set all  $g_j$  to zero, to obtain a 1-D approximation  $\hat{\mathbf{R}}$  that can be generated from  $f$  alone (using the clr mean  $\bar{\mathbf{r}}$  and the first eigenvector  $\mathbf{v}_1$ ). The resulting equation represents a so-called linear compositional process model (von Eynatten et al., 2003).

The fractionation pattern of the ternary sub-composition Q-P-Others may be understood as follows: for  $f = 0$ , we obtain the average sub-composition of pluton LC (i.e.,  $\bar{\mathbf{r}}$ ); a positive value of  $f$  indicates enrichment of P and Others, accompanied by depletion of Q; a negative value of  $f$  indicates depletion of P and Others, accompanied by enrichment of Q. The petrological significance of  $f$  may be further clarified by an appropriate scaling procedure. The pattern of compositional



**Fig. 9.** Single-component maps of LC(4,2). Black contour corresponds to the concentration values obtained from inverse clr-transformation of the area-weighted mean composition ( $\hat{\pi}$ ).

variability about the ternary clr mean  $\bar{r}$  is defined as the product of  $f$  and the first eigenvector,  $v_1$ . This term is unaffected if we multiply  $f$  with some constant, and divide the coefficients of the first eigenvector by the same constant. We may thus approximate the eigenvector coefficients by integers:

$$f \begin{bmatrix} -0.7717 \\ 0.1547 \\ 0.6169 \end{bmatrix} = F \begin{bmatrix} -5.001 \\ 1.003 \\ 3.998 \end{bmatrix} \cong F \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix}$$

where  $F = 0.1543 \cdot f$

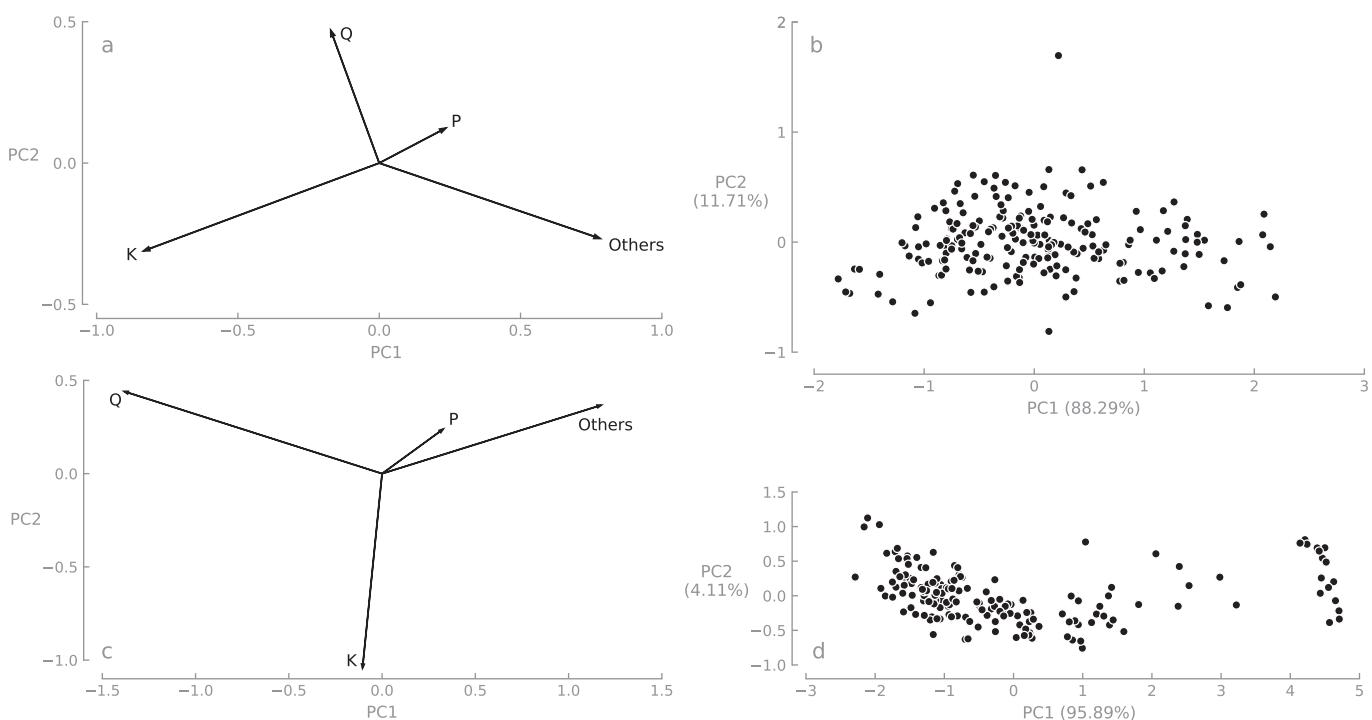
If we carry out the same exercise for the second eigenvector of the sub-composition Q-P-Others, we obtain:

$$g \begin{bmatrix} -0.2668 \\ 0.8017 \\ -0.5348 \end{bmatrix} = G \begin{bmatrix} -0.998 \\ 3.000 \\ -2.001 \end{bmatrix} \cong G \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$$

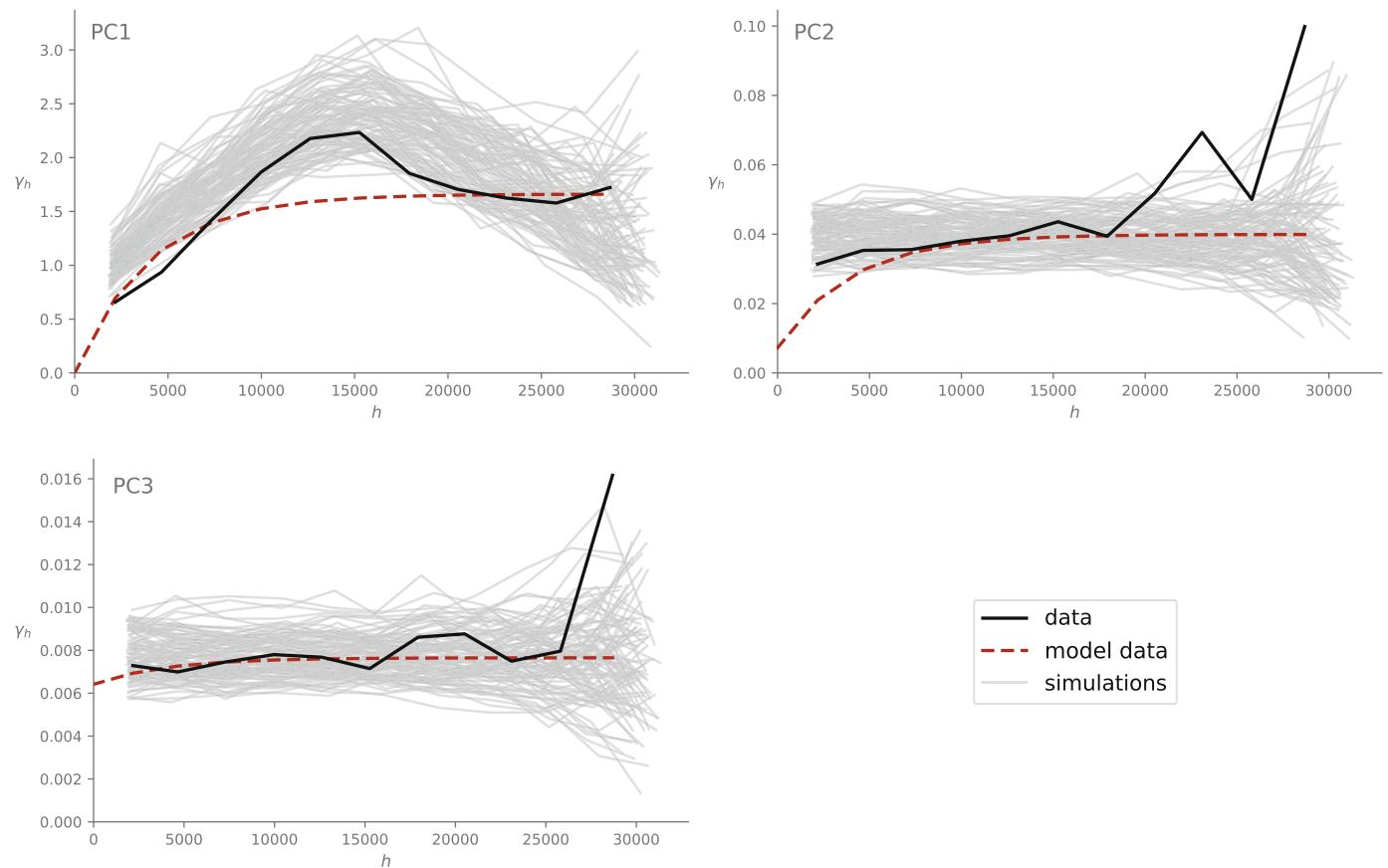
where  $G = 0.2672 \cdot g$

Because the Q-P-Others sub-compositions represent a 1-D set, this term must be equal to zero and thus acts as a constraint. In terms of the raw compositional variables (volume proportions), the constraint equals:

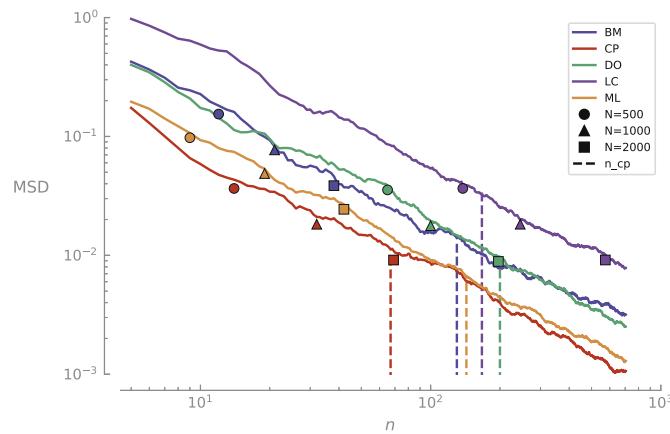
$$\ln\left(\frac{P^3}{Q \cdot Oth^2}\right) = c \cong 4.2$$



**Fig. 10.** covariance biplots of the full compositions and form biplots of the Q-P-Others sub-compositions of plutons DO (a, b) and LC (c, d).



**Fig. 11.** Empirical variograms, variogram models, and 100 simulated empirical variograms obtained by random sampling from BM(4,1) using the actual value of  $n = 130$ . The empirical variograms are fairly well reproduced, but the conservative simulation strategy overestimates the variance.



**Fig. 12.** Random sampling of plutons for models with four variables. The mean squared deviation (MSD) of sample means relative to the actual (area-weighted) mean composition of the pluton is plotted against the sample size ( $n$ ) needed to bring the MSD within the bounds of the counting error for different values of count size ( $N$ ). Vertical dashed lines mark the actual number of control points ( $n$ ) of the five data sets.

This condition represents the equilibrium state of the fractionation process. The term between brackets is of the same mathematical form as equilibrium reaction constants in chemistry (Aitchison, 1999).

To complete our description of the pattern of compositional variability in pluton LC we need to take into account the alkali feldspar content (volume proportion). A straightforward approach is to define the log-ratio:

$$m = \ln\left(\frac{K}{1 - K}\right)$$

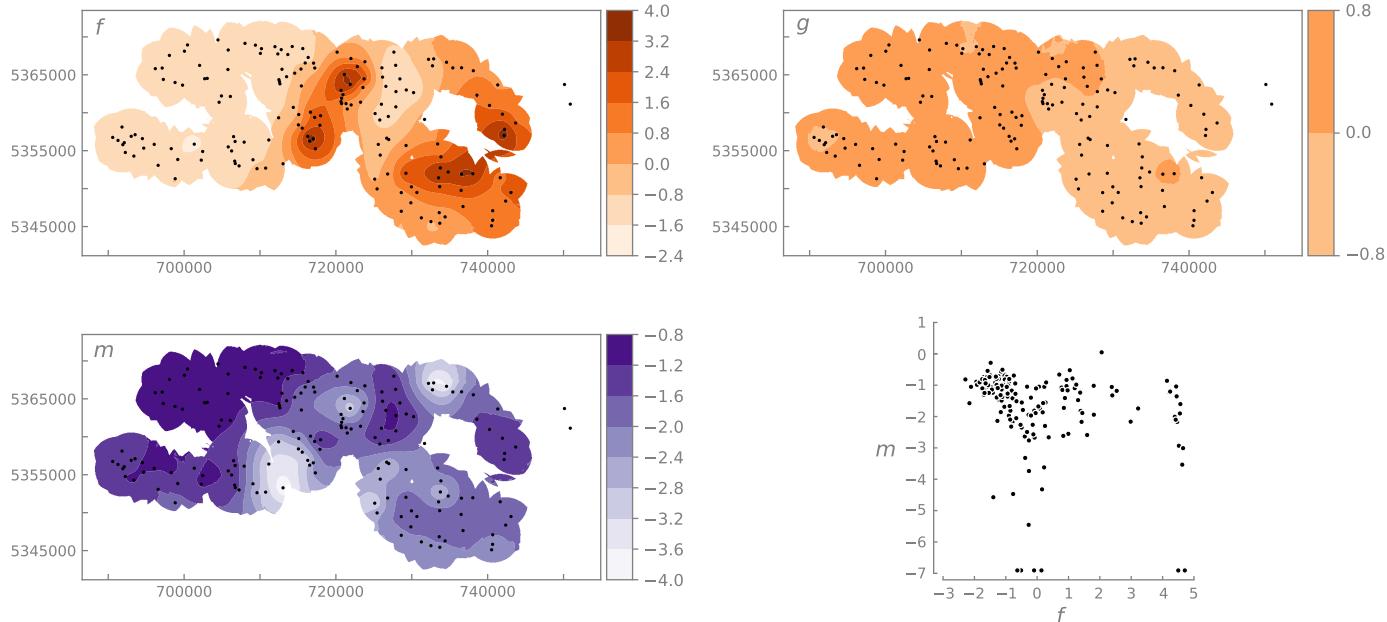
where  $1 - K = Q + P + Oth$

The log-ratio  $m$  represents the admixture of K-feldspar to the Q-P-Others sub-composition in the form of an unconstrained variable. The variability of the four-part composition of pluton LC may thus be captured with only two variables:  $f$  and  $m$ . The geostatistical models of these two variables are shown in Fig. 13. For the sake of completeness, the geostatistical model of  $g$  is shown on the same colour scale as  $f$ . The spatial component of  $g$  displays an E-W trend. However, as its contribution to the overall pattern is very small, this trend falls entirely within the noise (MSPE). The relation between  $f$  and  $m$  has been clarified in the cross-plot of Fig. 13. It is clear that  $f$  and  $m$  are uncorrelated, and that a few data points have anomalously low values of  $m$ . We now understand why the four-part composition of pluton LC can be adequately approximated by a model of rank 2 because Q, P and Others are intimately linked and K varies independently of these three. In petrogenetic terms, the structure of compositional variation in pluton LC seems to result from a combination of fractionation among Q, P, and Others (mafic minerals), with a random admixture of K. However, the spatial pattern of  $m$  is quite distinct and indicates that the admixture of K is not spatially random (Fig. 13), but varies from relative high values in the WNW to relatively low values in the ESE. The four-fold geographic subdivision of pluton LC that was apparent from the single-component maps (Fig. 9) is also visible in the spatial structure of  $f$ : the western-central and eastern parts of pluton LC are characterized by enrichment of P and Others, accompanied by depletion of Q, relative to the other two sectors of the pluton. The above interpretation of the compositional structure within the LC granite complex is purely data-driven and has been made without taking into account any petrological

**Table 4**

Simulation results and evaluation of experimental design. Selected models are indicated by pluto code, with number of variables and rank in brackets. VAR(OK) = variance of kriging surface; MSPE = mean squared prediction error; TVAR(s) = total variance from simulations; TVAR(d) total variance of data; Noise [%] = percentage of noise in model;  $n(N = \dots)$  = required sample size for hypothetical values of  $N = 500, 1000, 2000$ , respectively;  $n$  = number of control points;  $N$  = number of points counted in thin section; MCE( $N$ ) = multinomial counting error in clr-units based on area-weighted mean and actual  $N$ ; MSD( $n$ ) = mean squared deviation from area-weighted mean for actual  $n$ ; Diagnosis: A = adequate, O = over-sampled, U = under-sampled; Remarks: approximate sample size  $n^*$  required to meet the objective for actual  $N$ .

Model	BM(4,1)	CP(4,1)	DO(4,2)	LC(4,2)	ML(4,1)	BM(6,3)	CP(6,2)
PCs	1	1	12	12	3	123	12
VAR(OK)	1.240	0.336	1.003	2.043	0.003	1.462	1.343
MSPE	0.757	0.236	0.828	2.698	0.909	1.145	3.045
TVAR(s)	2.007	0.724	2.086	5.344	0.927	2.676	4.989
TVAR(d)	1.709	0.615	1.780	4.834	0.916	2.278	4.459
Noise [%]	38	54	52	62	> 99	45	73
$n(N = 500)$	14	16	65	138	9	5	5
$n(N = 1000)$	21	37	100	245	19	10	8
$n(N = 2000)$	39	75	197	574	42	17	20
$n$	130	67	200	167	143	130	67
$N$	2000	1850	2050	1500	1000	2000	1850
MCE( $N$ )	0.038	0.010	0.009	0.012	0.049	0.128	0.246
MSD( $n$ )	0.014	0.011	0.010	0.032	0.007	0.019	0.084
Diagnosis	O	A	A	U	O	O	O
Remarks	$n^* = 40$	–	–	$n^* = 400$	$n^* = 20$	$n^* = 20$	$n^* = 20$



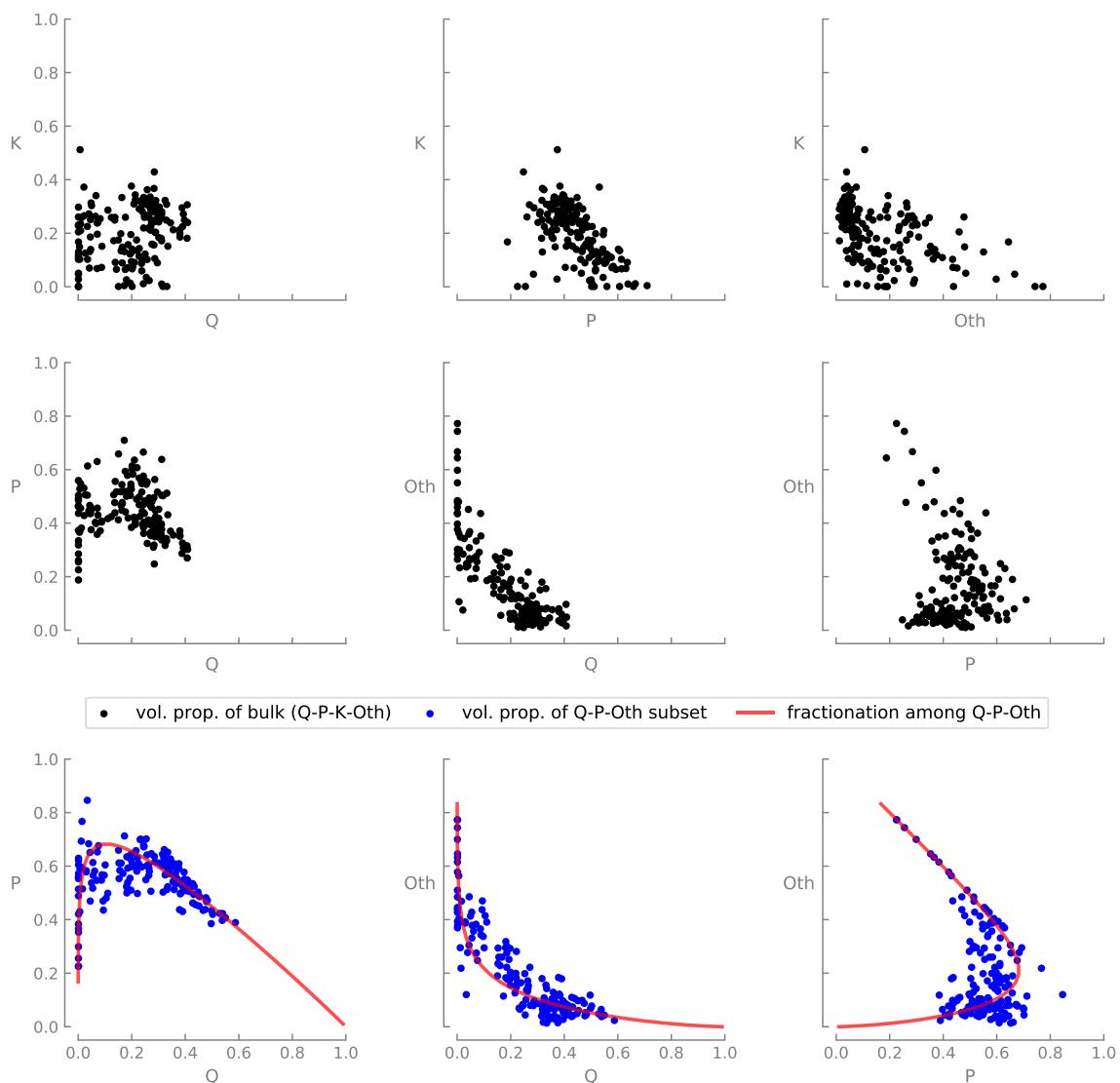
**Fig. 13.** Geostatistical models of  $f$ ,  $m$ , and  $g$  and cross-plot of  $f$  and  $m$ .  $f$  is interpreted as a fractionation pattern among Q, P, and Others (mafic minerals),  $g$  represents the residuals of the 1-D approximation, and  $m$  represents the admixture of K-feldspar to the fractionated ternary sub-composition. See text for a discussion of results.

information. The possibility to automatically extract fractionation and mixing patterns from the data permits testing of competing petrogenetic hypotheses on the origin of spatial heterogeneity within granitoids, and by extension, other types of compositional data.

To conclude this section we will regress to the universe of conventional petrological data analysis of raw compositional variables in light of the petrogenetic inferences presented above. The upper two rows of Fig. 14 show all possible cross-plots of the raw compositional variables (volume proportions). Four out of the six cross-plots seem to display patterns: the three that make up the Q-P-Others sub-composition, and the P-K cross-plot as well. All patterns are quite noisy, because of the interference between “real” (process-based) interdependence and the closed-sum effect. This is most apparent in the cross-plots of Q-K and Others. Based on the cross-plots shown in the upper two rows of Fig. 14, it would be hard to unambiguously identify a sub-composition that could be explained by fractionation. The only way to detect such a sub-composition would be to carry out some kind of systematic search, e.g.

first leave out one component and then recalculate the remaining three to unit sum. With this approach, one would produce 12 more cross-plots. Then the same would have to be done for all two-part sub-compositions, which adds another 6 cross-plots to the list. It will be quite difficult to identify the three-part sub-composition Q-P-Others from these 18 cross-plots, as each shows only two out of the three components involved in fractionation.

Let us assume for the sake of argument that this daunting task has been accomplished, the Q-P-Other sub-composition has been identified, and we have produced the three cross-plots shown at the bottom of Fig. 14. Then the next problem appears: modelling the highly non-linear patterns visible in these plots would be challenging, especially taking into account that the sum of Q, P, and Others would have to be equal to unity everywhere along the three fractionation curves (or the fractionation plane if we venture into 3-D space), negative values of Q, P, and Others are not permissible, and we would like some equilibrium condition to apply to the system as well (even though there would be no



**Fig. 14.** Top and middle row: conventional cross-plots of raw compositional variables (volume proportions) of pluton LC. Bottom row: Cross-plots of ternary Q-P-Others sub-composition expressed as volume proportions recalculated to unit sum with superimposed fractionation model.

obvious choice for such a criterion based on the data in the format shown here).

To be able to routinely tackle such tasks one could (and probably should) turn to a log-ratio based analysis. Although the projections of the fractionation model into the bivariate concentration space of the ternary sub-composition Q-P-Others (shown in the three cross-plots at the bottom of Fig. 14) appear as highly non-linear (hyperbolic) curves, they appear as a simple linear function in the log-ratio universe. Such linear representations of fractionation among multiple components can be routinely identified in and extracted from biplots. The linear fractionation model in clr-space automatically obeys the unit-sum constraint after inverse clr-transformation and fulfills a uniquely specified equilibrium condition. Given these glaring advantages, it is a mystery why log-ratio-based analysis (which has been around for more than 30 years) is still not widely considered as the default approach to exploratory modelling of petrogenetic relationships.

#### 5.4. Precision and sample size

Evaluation of the data-acquisition schemes used to sample the five plutons has been made possible by adopting an operational definition of the required precision with which we aim to estimate their area-

weighted mean compositions. We wish to stress that the (fairly modest) objective of estimating the area-weighted mean is not necessarily identical to the ideal objective of mapping, which should aim at modelling the covariance structure of the composition of entire lithosomes, but it serves as a useful frame of reference for discussion. The required precision has been defined in terms of the irreducible counting error, which is a function of  $\hat{\pi}$ , the estimated area-weighted mean composition, and  $N$ , the number of points counted in thin section (Bloemsma and Weltje, 2015) as discussed in Appendix A. Random sampling from the plutons yields an estimate of the number of control points (sample size  $n$ ) needed to achieve the objective. The results of this exercise are summarized in Table 4. If the number of points counted (the precision) is in agreement with the sample size, the experimental design is considered adequate for achieving the objective. If not, we conclude that the pluton has been either over- or under-sampled. For instance, in the case of LC, which has been under-sampled according to the above criterion, adopting a lower precision by setting  $N = 600$  would result in a balanced data-acquisition scheme (as can be derived from inversion of the equation of the clr-space variance of count data). The alternative is to maintain the high precision associated with  $N = 1500$ , and in that case, the number of control points should be increased to  $n = 400$ . In the case of BM(4,1), BM(6,3), CP(6,2), and ML(4,1), which have been

over-sampled, a significant reduction of  $n$  would still permit us to achieve the objective.

The most important result emerging from the random sampling is that the number of points counted is proportional to the size of the sample needed to estimate the area-weighted mean. Roughly speaking, doubling the count length  $N$  also doubles  $n$ , which indicates that the total effort increases by a factor of four. Because of this effort, which translates directly into time and cost, one may wonder how to define the desired precision and achieve the objective with a minimum effort. The problem was eloquently summarized by Griffiths (1967), who stated that: “*The adoption of elaboration in experimental technique with concomitant reduction in the number of experiments and, in contrast, an increase in the number of experiments without elaboration of technique are important alternatives in achieving any degree of precision. The choice of alternative should be given very careful consideration, and the final decision as to which is the most efficient procedure will be a function of the objective, time, and cost involved in the experimental program*” (Griffiths, 1967, p. 9).

In order to define the desired precision, we should take into account the size (area) corresponding to the measurement of a single specimen (thin section), commonly referred to as “support” in geostatistics, in relation to the scale of heterogeneity of the material under investigation (cf. McCanta et al., 2017). In our case, the smallest level of heterogeneity is determined by the crystal size of the granitoids. The most efficient setup for point counting corresponds to the case in which successive observations are statistically independent. This may be achieved by adopting a grid size that is large enough to prevent the same crystal from being counted twice. Standard thin sections cover an area of a few  $\text{cm}^2$ , and given the large crystal sizes of granitoids, this implies that no more than a few hundred points can be counted without seriously violating this requirement. The implied justification for counting much larger numbers of points (Table 1) can be traced back to the work of Chayes (1956), who advocated count lengths of 1500 up to 2000 because it seemed to bring a high degree of precision to the analysis. However, excessively large values of  $N$  as recommended by Chayes (1956) and used in the above studies are justifiable only if the objective is to characterize the specimen under investigation with a high degree of precision (cf. Chayes, 1951; Chayes and Fairbairn, 1951).

In the vast majority of geological studies (including the surveys examined in this paper), the target population is not the specimen but the lithosome as a whole (Griffiths, 1967). Increasing analytical precision does not bring any more insight into the large-scale variability of lithosomes in cases where the spread among the properties of those specimens (the heterogeneity of the lithosome) is larger than the precision with which each specimen has been characterized. As discussed by Weltje (2004), maximizing  $N$  carries the undesirable implication that even minute differences between adjacent specimens are likely to be considered statistically significant and thus require some form of explanation. Let us assume that a set of specimens from a lithosome has been measured by two different techniques, one with a low precision in the range of the heterogeneity and another with much higher precision. The low-precision technique does not permit any statistically significant differences to be detected among the specimens because of its large intrinsic uncertainty, whereas the high-precision technique allows us to conclude that each specimen is statistically distinct. At first glance, this seems to suggest we have gained something by applying the latter. However, to achieve our objective, we need to characterize the properties of the lithosome under investigation by averaging the sets of specimens and calculating their variances. This gives more or less the same results with both techniques. The source of the variance is different, however. In the case of the low-precision technique, it is controlled by the intrinsic uncertainty, whereas in the case of the high-precision technique, it is determined by the heterogeneity. Taking into account the efforts associated with both approaches, the choice of alternative is not particularly difficult.

The usefulness of adopting values of  $N$  which are much smaller than those actually used for acquiring the data sets re-analysed in this study is clearly brought out by the large nugget values of the variogram models which minimize the MSPE, as well as by the comparatively large values of the MSPE itself. Both are much larger than the precision associated with the determination of the composition of individual specimens. Most likely, a study of the small-scale heterogeneity of granitoids (on the scale of centimeters to tens of meters) would permit us to conclude that much smaller values of  $N$  are sufficient to achieve the objective with less effort (cf. Weltje, 2004). Alternatively,  $n$  might be increased substantially while keeping the total effort of data acquisition constant (although this would increase the cost of sample preparation). Enlarging  $n$  increases the probability that the sample is representative of the population (lithosome), and this is the only way to improve geospatial models of the properties of lithosomes.

### 5.5. Implications for sediment-generation studies

A data set of granitoids and their sandy detritus acquired by Heins (1992, 1993, 1995) was expanded and re-analysed in two recent studies (Weltje et al., 2018; Paredis et al., 2020). If the results of the geostatistical simulations reported above are taken as a guideline, the plutons in the data set of Heins (1992) are likely to have been under-sampled, with typical values of  $n \cong 10$ . The notion that a high degree of uncertainty is associated with the estimates of parent-rock properties based on such small samples poses certain limitations on predictions of sediment properties. Insight into the uncertainties of predicted sediment properties is needed to interpret discrepancies between predictions and observations in a probabilistic sense. If sediment generation can be simulated, the uncertainties associated with the initial conditions (estimates of fundamental parent-rock properties sensu Paredis et al., 2020) can be propagated all the way through to uncertainties of predicted sediment properties. Quantification of the uncertainties associated with predicted and observed sediment properties will permit us to evaluate the performance of sediment-generation models and provide suggestions for their improvement.

### 5.6. High-dimensional data sets

As discussed above, the quality of geostatistical models is largely governed by the knowledge of the short-range variability of the data, which is much larger than the analytical (counting) error. Essentially the same conclusions would have been drawn if geochemical data had been used instead of point-count data because both have comparable support (analysed area or volume) and thus are equally representative (McCanta et al., 2017). As discussed above, this applies regardless of the precision of the analytical technique used. The number of rock-forming minerals is usually much smaller than the number of chemical elements that can be measured routinely in bulk geochemical analysis. Interpolation of PCs instead of the original variables will, therefore, be more effective with high-dimensional geochemical data ( $k \geq 30$ ), because the underlying pattern rarely has more than  $\sim 10$  dimensions ( $r \leq 10$ ), since specific linear combinations of chemical elements form the building blocks of minerals, and specific combinations of minerals fractionate to form collinear sub-compositions (as illustrated for the LC pluton).

Given the noisiness of the modal data, it is not surprising that the rank of the geostatistical models never exceeds three. In theory, the highest rank of a model is equal to the number of detectable rock-forming minerals, but this number is never reached owing to practical limitations on the representativeness of individual specimens (noise) and systematic fractionation effects. Dimension reduction of compositional data thus seems a useful and even necessary pre-processing step. Not only does it reduce the number of variables to be interpolated, but it also decreases the noise in the data which is concentrated in the highest PCs that drop out of the modelling at an early stage. Reduction

of the multivariate geostatistical modelling problem to a series of univariate problems which may be solved by means of OK implies that any PC which does not contribute to explaining the spatial variability of the data can be identified in the initial stage of variogram modelling because its MSPE is larger than the sill of the variogram model. The possibility to assess the geospatial information content of PCs, coupled with the notion that only cumulative models need to be taken into account, thus reduces the number of candidate models to be evaluated *a priori*.

### 5.7. Recommended experimental design

The geostatistical models permit the total variability of modal composition to be decomposed into a systematic part (the variance of the OK surface) and a random error (the MSPE representing the unexplained variance), as shown in Table 4. This unexplained variance, which makes up around 50% of the total variance, represents the superposition of small-scale heterogeneity that cannot be resolved by the model based on a lack of control on short-range variability, and actual noise resulting from the fact that the ratio of crystal size to specimen size is large, which inflates the sampling error relative to the ideal case in which this ratio is very small and each thin section would have been representative (McCanta et al., 2017).

An obvious question is whether the models can be improved by adopting a different layout of control points. The data of the surveys analysed in this paper were all acquired by sampling on a grid or by using a stratified random approach (random sampling within each cell of a grid). As a result of this, there is almost no control over the relation between the variability and the distance between control points at the short-range. We, therefore, formulated the variogram modelling as an optimization problem, which resulted in large values of the nugget (typically around 50% of the sill). It is conceivable that the quality of geospatial models of lithosomes can be improved by adopting a different sampling design.

We recommend a multi-stage data-acquisition strategy. In the first stage, a low-density sampling campaign should be conducted to cover the entire area of the outcropping lithosome, coupled with several areas of higher sampling density to constrain short-range variability (the nugget or experimental error). Based on this first round of sampling an initial data set is produced for setting up a geostatistical model. Simulated sampling from this model may be used to estimate how many control points would be required. The additional control points needed are then sampled in a second campaign, followed by an update of the geostatistical model. In this second sampling campaign, particular attention should be paid to areas of the lithosome where spatial compositional gradients appear to be steep, as well as to areas around data points which have extremely large cross-validation errors. If necessary, the process is repeated a third time. This iterative approach allows us to escape from the conundrum that an adequate experimental design requires *a priori* knowledge of the pattern of spatial heterogeneity.

### 5.8. Future research

The insights that emerged from the geospatial modelling exercises presented in this paper are that (1) we should maximize sample size instead of the precision of our analytical techniques, (2) we should adopt a multi-scale approach to designing the layout of control points, (3) we should adopt an iterative (multi-stage) data-acquisition strategy, and (4) the proposed heuristic approach to geostatistical modelling of compositional data is expected to be particularly effective for high-dimensional geochemical data sets. To this, we should add the fact that our objective for evaluation of the data-acquisition schemes does not, in general, provide us with a guideline for the actual sample size needed to map entire lithosomes. Much larger samples, on the order of hundreds (up to perhaps thousands) of control points, are needed if the objective of estimating the area-weighted mean, which we adopted for the sake of

argument, is traded for the more realistic objective of estimating the spatial covariance structure of the composition of entire lithosomes. The latter objective may be formulated as petrogenetically constrained compositional mapping. An example of such mapping exercise has been given above for the LC granite complex.

Putting these insights into practice raises the question of the analytical effort involved. Taking very large samples from lithosomes which have to be processed in the lab to yield estimates of the properties of interest is a daunting task and the main obstacle to producing high-quality maps of lithosomes from which reliable petrogenetic inferences may be drawn. The advent of non-destructive techniques for rapid acquisition of compositional data by means of hand-held X-ray Fluorescence (XRF) and Near-Infrared Reflectance (NIR) analysers has the potential to change this. The hardware for applying these techniques in the field is available at a moderate cost, and consequently, the ability to collect large numbers of high-quality georeferenced measurements with user-specified support in a short time is now well within reach. Reliable algorithms based on the principles of compositional data analysis have been developed for XRF data acquired *in situ* without control on measurement geometry (Weltje and Tjallingii, 2007; Weltje et al., 2015; Henares et al., 2019). The performance of these algorithms is more than adequate for the mapping task at hand. The development of algorithms for quantifying NIR data is still at an early stage. From a theoretical point of view, there are no fundamental obstacles to the quantification of NIR data along the same lines as used for quantification of *in-situ* XRF measurements, because all spectroscopic data exhibit the same statistical behaviour in the clr-space (Bloemsma and Weltje, 2015). A particularly attractive aspect of NIR spectroscopy in the context of mapping is its application in remote sensing, which is complementary to the use of NIR in the field. The joint use of XRF and NIR is also worth considering, as the former provides information on the chemical composition whereas the latter is sensitive to mineralogical composition. Combining both sets of measurements will shed light on the relation between chemical and mineral composition, which may be exploited in end-member modelling of chemical data to obtain the likely number of mineral phases involved, as well as their chemical composition (Weltje, 1997; Tolosana-Delgado et al., 2011). Calibration of XRF and/or NIR data to properties of interest will be needed to turn “proxy” maps obtained by these techniques into maps of fundamental rock properties (cf. Henares et al., 2019; see also Paredis et al., 2020).

## 6. Conclusions

Geostatistical modelling of the modal composition of a series of five extensively surveyed granitoids shows that point-count data must be considered extremely noisy from the point of view of geospatial analysis because sampled areas are small compared to the average crystal size of granitoids and crucial information on short-range variability of modal composition is missing. The same applies to data acquired in geochemical surveys because the information content of control points is primarily limited by the volumes analysed (which are comparable to the areas of thin sections) relative to the size of the crystals that make up the rock and not by the precision of analytical technique. Simulations indicate that up to several hundreds of specimens must be analysed to successfully predict the mean (area-weighted) modal composition of plutons and its spatial covariance structure, as schematically depicted in the QAPF maps.

Spatial interpolation of compositional data is extremely useful in many fields of application, but a far from trivial task (McKinley et al., 2016). Current algorithms (Tolosana-Delgado et al., 2019) make use of co-kriging of the original variables in the data set. The alternative explored here is to apply dimension-reduction techniques (SVD) to the data, and model a much smaller number of PC scores using Ordinary Kriging. For reasons of tractability, we kept the approach outlined in this study as basic as possible (following the KISS principle). We did not include any further transformations of the PC scores, as advocated by

Tolosana-Delgado et al. (2019) for the original log-ratio transformed variables, and made use of standardized variogram models (exponential with fixed range and search radius) for each data set. The nugget values were estimated by minimizing the MSPE. Despite these restrictions, the results of the permutation test indicate that the residuals of the interpolations obtained with Ordinary Kriging of PC scores do not exhibit significant cross-covariance, which suggests that little may be gained by application of co-kriging. A comparative study of the two geostatistical methods applied to the same data sets is needed to shed more light on their strengths and weaknesses but falls outside the scope of this study.

In this study, we presented a concise overview of techniques available to systematically map the compositional properties of lithosomes. The ability to routinely identify petrogenetic structure resulting from combinations of fractionation and mixing within the data set emerged as a logical extension of the proposed geostatistical modelling approach, which is based on a combination of dimension reduction (SVD) of log-ratio-transformed data and conventional geostatistical interpolation (OK). The tools and algorithms needed for data-driven geospatial petrogenetic modelling are nearly all in place. Large georeferenced data sets of rock properties may be acquired by non-

destructive (hyper)spectral analysers (specifically XRF and NIR) in the field and combined with remotely sensed NIR data. As shown in this study, log-ratio based analysis is very well equipped to process such integrated data sets and achieve this objective.

## Declaration of Competing Interest

The authors have no competing interests.

## Acknowledgements

We gratefully acknowledge the perceptive comments of two anonymous reviewers that permitted us to significantly improve the manuscript. BP would like to acknowledge the PyKrig project (Murphy et al., 2020) which formed the basis of the in-house Python code developed for this paper. The research reported here was funded by KU Leuven. Supplementary materials (raw data, graphs, and maps) are available at <https://github.com/Complexio/Weltje-Paredis-2020>.

## Appendix A. Nugget model

### Experimental error

The experimental error may be estimated from sets of replicate measurements which must be specimens from the same population but do not have to form part of the spatially indexed data set to be geo-statistically analysed. The equation below is based on the definition of the variance at lag zero for a univariate case (i.e., each set of PC scores):

$$\gamma_r = \sum_{j=1}^{n_s} \sum_{i=1}^{n_j} w_j (x_{ij} - \bar{x}_j)^2$$

where  $\gamma_r$  is the variance estimated from replicates,  $n_s$  is the number of replicate sets,  $n_j$  is the number of specimens in the  $j^{\text{th}}$  set,  $x_{ij}$  is the  $i^{\text{th}}$  specimen of the  $j^{\text{th}}$  set, and  $\bar{x}_j$  is the mean of specimens of the  $j^{\text{th}}$  set. The weights assigned to each set,  $w_j$ , are given by:

$$w_j = \frac{n_j}{\sum_{j=1}^{n_s} n_j (n_j - 1)}$$

Sets of spatially indexed replicates were taken out of the data and replaced by a single composition obtained by averaging the set of replicates in the clr space. Each control point is therefore associated with a unique composition estimate in the data sets to be geostatistically analysed.

### Counting error

Point-count data are estimates of modal composition based on a sample of finite size. The counting error associated with such data is an irreducible sampling error that follows a multinomial distribution. The magnitude of the counting error can be determined from the representative composition (the clr-mean of the control points or the OK grid), the total number of points counted ( $N$ ), and the number of classes ( $k$ ), provided that successive observations on a counting grid are statistically independent and the specimen is homogeneous. Statistical independence is ensured by adopting a grid size that exceeds the size of the largest object likely to be encountered. Bloemsma and Weltje (2015) showed through error propagation that the total variance of multinomially distributed data (MDD) in the clr-space is given by:

$$\gamma_c = \text{VAR}\{\text{clr}(\text{MDD}(N, \pi_k))\} = \frac{1}{N} \sum_{j=1}^k \frac{1 - \pi_j}{\pi_j}$$

where the vector  $\pi = (\pi_1, \pi_2, \dots, \pi_j, \dots, \pi_k)$  represents the proportions in the population from which a sample of size  $N$  has been drawn. This vector is obtained by inverse clr-transformation of the clr mean. An alternative method for calculating the predicted clr-space variance of count data makes use of the Dirichlet distribution (Vermeesch, 2005; van den Boogaart and Tolosana-Delgado, 2013).

### Nugget estimation

Our objective is to provide an estimate of the nugget in variogram models based on the experimental error in thin-section analyses of granitoids (for the case  $k = 4$ ). Replicate sets are available for only two out of the five plutons (BM and CP), and therefore, analyses of the Mount Westerly (MW) granite (Chayes, 1951; Chayes and Fairbairn, 1951) were added to enlarge the database. The number of points counted varies among the three data sources (from 1500 up to 2500) and we cannot compare them directly. We, therefore, scaled the experimental error with the associated counting error by forming a ratio of the two variances:

$$\frac{\gamma_r}{\gamma_c} = e^h$$

The logarithm of this ratio ( $h$ ) follows an approximately normal distribution. We estimate its mean as:

$$\hat{h} = \frac{1}{n_s} \sum_{i=1}^{n_s} \ln\left(\frac{\gamma_{r,i}}{\gamma_c}\right)$$

If the sizes of the replicate sets are not identical, we have to replace the above by a weighted mean. Mean values of  $\hat{h}$  and their 95% confidence limits were calculated for each of the three plutons separately. The estimated nugget  $\hat{\gamma}_0$  is related to  $\hat{h}$  by:

$$\hat{\gamma}_0 = e^{\hat{h}} \gamma_c$$

The results of nugget estimation are shown in Table A1. The only pluton whose nugget can be well constrained is CP (based on 19 replicate sets). It is obvious that MW is less heterogeneous on the small scale than either BM and CP, as expected given the nature of the single slab of MW granite used by Chayes (1951) and Chayes and Fairbairn (1951), which was selected from a quarry based on its macroscopic (visual) homogeneity. The apparent homogeneity of the large specimen from which several thin sections were cut was considered a prerequisite for examining the experimental error associated with point counting. In light of this consideration, it is worth noting that the experimental error is still larger than the counting error even in this nearly ideal case.

A more satisfactory explanation may be given if the data-acquisition parameters are called into question. If fewer points had been counted in the thin sections cut from the slab of MW granite, the conclusion that the latter is indeed homogeneous might well have been drawn. This would have resolved the apparent contradiction between the point-counting results and the macroscopic (visual) assessment of its homogeneous nature. Application of the equation of Bloemsma and Weltje (2015) shows that this would have been the case if  $N \leq 1200$ . There are other good reasons for not counting so many points in thin section. This topic is more fully discussed in the section on precision and sample size in the body text.

In conclusion, the estimate  $\hat{\gamma}_0$  represents the variance among replicates sampled within a centimeter-to-meter range, whereas the value of  $\gamma_c$  represents the (irreducible) variance at the scale of a single thin section (a few centimeters). Even though these distances differ by roughly one order of magnitude, they may be considered negligible relative to the typical spacing of control points, which is in the range of hundreds of meters. Because  $\gamma_c < \hat{\gamma}_0$ , an obvious conclusion is that the variance increases very rapidly with the distance between control points, and the effective nugget of the corresponding variogram models may be larger still, given the likely greater variability at the meter-to-decrometer scale that cannot be assessed from the data at our disposal. An alternative explanation is that far too many points were counted by the operators who acquired the data sets at our disposal. This leads to the same conclusion, i.e., that the nugget must be quite large for all of the granites examined in this study.

Table A1

Nugget modelling.  $k$  = number of variables,  $n$  = number of replicate sets,  $N$  = (average) number of points counted in thin section,  $\gamma_c$  = multinomial counting error in clr-units,  $\gamma_0$  = experimental errors (estimates of the nugget), where L and U refer to the lower and upper limits of the 95% confidence interval.

	BM	CP	MW
$k$	4	4	4
$n$	4	19	1
$N$	2000	1850	1750
$\gamma_c$	0.029	0.010	0.013
$\gamma_0$ (L)	0.010	0.023	–
$\gamma_0$	0.043	0.035	0.020
$\gamma_0$ (U)	0.197	0.051	–

## Appendix B. Test for significant cross covariance

A fundamental requirement of any valid geostatistical model is that its residuals are stationary, i.e. they do not exhibit any significant spatial patterns. The objective of the test described below is to determine whether the residual compositions exhibit a spatial structure significantly different from random noise. If this turns out to be the case, we cannot accept the geostatistical model as a valid and unbiased representation of the data. If not, it may be concluded that the geostatistical model accurately captures the main spatial trends in the data. The magnitude of the stationary random error (variance) is estimated as the mean squared prediction error (MSPE) obtained from cross-validation.

The principle of a permutation test is to compare a test statistic calculated from the data to its distribution under the null hypothesis. The distribution under the null hypothesis is established utilizing stochastic simulation (randomization). In the case of the data tested in this study, we randomized the spatial coordinates  $x$  relative to the residual compositions. The number of unique ways in which one could distribute the residuals over the spatial coordinates of  $n$  control points equals  $n!$ , an extremely large number for the data sets analysed in this study which have values of  $n$  around 100. Any spatially randomized realisation of the data thus may be thought of as having been drawn with equal probability from a virtually infinite set of possibilities.

The power of permutation tests lies with their ability to provide meaningful estimates regardless of the multivariate distribution of the residual compositions and the spatial arrangement of the control points. Permutation tests are internally consistent provided that the test statistic is formulated in such a way that its value is a monotonic function of the likelihood of the data under the null hypothesis. If we use a test statistic whose magnitude is inversely proportional to its probability under the null hypothesis (e.g., zero indicates perfect agreement and large values indicate poor agreement), its critical value is defined as the  $100(1 - \alpha)$  percentile of the values obtained from the randomized data. In this test, we approximated the distribution of our test statistic under the null hypothesis by running 1000 simulations for each pluton.

The permutation test consists of comparing the value of the test statistic obtained from the actual data to its critical value. If the value of the test statistic is less than its critical value, we may conclude that there is nothing to suggest that the pattern is anything but random.

The test was set up as follows. The matrix of residual compositions was subjected to SVD to give three globally independent sets of PC scores. The cross-covariance functions of the three PCs were evaluated for all pairs of scores whose lag is smaller than the search radius employed in the geostatistical model. The values of the cross-covariance were placed in bins containing the same number of pairs. The number of bins was taken equal to the square root of the number of control points rounded to the nearest integer.

The test statistic for detection of significant cross-covariance among PCs was inspired by the *t*-test (Davis, 2002). We evaluate two measures of departure from randomness:

$$t_{ij}^{\text{avg}} = \sum_{h=1}^l \frac{|\bar{C}_{ijh}|}{w_{ijh}}$$

$$t_{ij}^{\text{max}} = \max_h \left\{ \frac{|\bar{C}_{ijh}|}{w_{ijh}} \right\}$$

here,  $t$  represents the sum and the maximum value, respectively, of the absolute values of the mean cross-covariances ( $\bar{C}$ ) between the scores of the  $i^{\text{th}}$  and the  $j^{\text{th}}$  PC, normalized to their standard deviations ( $w$ ) over the range of lag bins considered ( $h$ ). The upper limit of bin  $l$  is equal to the search radius. As the number of pairs is equal for all bins within a given data set, it is not necessary to divide the  $t$  values by the square root of sample size (as in a conventional *t*-test). The overall test statistic  $T^*$  is obtained in two steps. First, we sum the values of  $t$  over all cross-covariance functions:

$$T^{\text{avg}} = \sum_{i=1}^{r-1} \sum_{j=i+1}^r t_{ij}^{\text{avg}} (s_i^2 + s_j^2)$$

$$T^{\text{max}} = \sum_{i=1}^{r-1} \sum_{j=i+1}^r t_{ij}^{\text{max}} (s_i^2 + s_j^2)$$

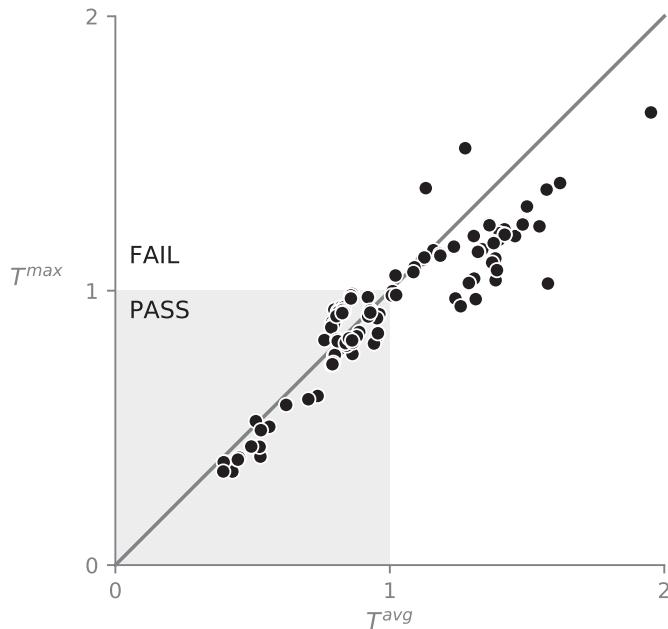
The term after  $t$  is a weight factor, which represents the proportion of the total variance described by the cross-covariance functions, as given by the sum of the squared singular values ( $s$ ). This weight factor was introduced to emphasize the cross-covariance functions which describe the strongest spatial patterns in the residual compositions. In the case of the residual compositions, the maximum value of  $r$  is equal to  $f = k - 1$ . In the case of the homogenized data, the two  $T$  values are calculated from the cross-covariance functions of three PCs ( $C_{12}$ ,  $C_{13}$ , and  $C_{23}$ ).

In order to test if the residual compositions display significant spatial structure, a test statistic was formed by evaluating the two  $T$  values under the null hypothesis from 1000 realisations of the spatially randomized data. The critical values  $T_c$  are defined as the  $100(1 - \alpha)$  percentiles of these statistics. The final test statistic  $T^*$  is taken as the largest value of the two ratios of  $T$  to their critical values:

$$T^* = \max \left\{ \frac{T^{\text{avg}}}{T_c^{\text{avg}}}, \frac{T^{\text{max}}}{T_c^{\text{max}}} \right\}$$

The result of the test is: PASS if  $T^* < 1$ , else FAIL (at  $\alpha = 0.05$ ).

The above test was also applied to the control points of each pluton to examine if the PCs of the data were spatially uncorrelated. For the homogenized data with  $k = 4$ , only two models are available: for  $r = 2$  only  $C_{12}$  must be evaluated, and for  $r = 3$ , all three cross-covariance functions have to be considered. Testing of the six-part compositions of BM and CP requires evaluation of up to 10 cross-covariance functions for  $r = 5$ . The results shown in Table B1 indicate that significant cross-covariance among PCs is present in all data sets, apart from CP(4) and ML(4).



**Fig. B1.** Results of all permutation tests for significant cross-covariances among PCs of data sets and residual compositions. Models and data sets plotting inside the shaded region pass the test (at  $\alpha = 0.05$ ). Most points plot below the line  $y = x$ , indicating that the test employing the average cross-covariance ( $T^{\text{avg}}$ ) is more sensitive than the one using the largest cross-covariance ( $T^{\text{max}}$ ). The actual test statistic  $T^*$  is defined as the largest of these two values.

**Fig. B1** summarizes the results of the permutations tests on all data sets and residual compositions. In general, the test detects more significant departures from randomness based on  $T^{avg}$  than on  $T^{max}$  (departure from the line  $y = x$  is highly significant). Note that both values have already been normalized to their critical values for the purpose of comparison, such that  $T^*$  is simply given by the largest of the two.

Table B1

Results of the permutation test indicate that significant cross-covariance is present among PCs in most of the data sets, except CP(4) and ML(4), for which all values of  $T^* < 1$  at  $\alpha = 0.05$ .

Data set	#PCs	$T^*$
BM(4)	2	1.58
	3	1.55
CP(4)	2	0.53
	3	0.53
DO(4)	2	1.39
	3	1.46
LC(4)	2	0.62
	3	1.24
ML(4)	2	0.52
	3	0.81
BM(6)	2	0.93
	3	1.57
	4	1.40
	5	1.42
CP(6)	2	1.95
	3	1.50
	4	1.62
	5	1.49

## References

- Aitchison, J., 1982. The statistical analysis of compositional data (with discussion). *J. R. Stat. Soc. Ser. B* 44, 139–177. <https://doi.org/10.1111/j.2517-6161.1982.tb01195.x>.
- Aitchison, J., 1986. *The Statistical Analysis of Compositional Data*. Chapman & Hall, London, pp. 416.
- Aitchison, J., 1999. Logratios and natural laws in compositional data analysis. *Math. Geol.* 31, 563–580. <https://doi.org/10.1023/A:1007568008032>.
- Aitchison, J., Greenacre, M.J., 2002. Biplots of compositional data. *Applied statistics*. *J. R. Stat. Soc. Ser. C* 51, 375–392. <https://doi.org/10.1111/1467-9876.00275>.
- Allen, P.A., 2017. *Sediment Routing Systems: The Fate of Sediment from Source to Sink*. Cambridge University Press<https://doi.org/10.1017/9781316135754>. 422 p.
- Atherton, M.P., Ghani, A.A., 2002. Slab breakoff: a model for Caledonian, late Granite syn-collisional magmatism in the orthotectonic (metamorphic) zone of Scotland and Donegal, Ireland. *Lithos* 62, 65–85. [https://doi.org/10.1016/S0024-4937\(02\)00111-1](https://doi.org/10.1016/S0024-4937(02)00111-1).
- Basu, A., 1985. Influence of climate and relief on compositions of sands released at source areas. In: Zuffa, G.G. (Ed.), *Provenance of Arenites*, NATO-ASI Series. D. Reidel, Dordrecht, pp. 1–18. [https://doi.org/10.1007/978-94-017-2809-6\\_1](https://doi.org/10.1007/978-94-017-2809-6_1).
- Basu, A., 2003. A perspective on quantitative provenance analysis. In: Valloni, R., Basu, A. (Eds.), *Quantitative Provenance Studies in Italy. Memorie Descrittive della Carta Geologica dell'Italia* 61, pp. 11–22.
- Bloemsma, M.R., Weltje, G.J., 2015. Reduced-rank approximations to spectroscopic and compositional data: a universal framework based on log-ratios and counting statistics. *Chemom. Intell. Lab. Syst.* 142, 206–218. <https://doi.org/10.1016/j.chemolab.2015.02.001>.
- Caracciolo, L., Tolosana-Delgado, R., Le Pera, E., von Eynatten, H., Arribas, J., Tarquini, S., 2012. Influence of granitoid textural parameters on sediment composition: Implications for sediment generation. *Sediment. Geol.* 280, 93–107. <https://doi.org/10.1016/j.sedgeo.2012.07.005>.
- Chayes, F., 1949. A simple point counter for thin-section analysis. *Am. Mineral.* 34, 1–11.
- Chayes, F., 1951. Modal analyses of the test rocks. In: Fairbairn, H.W. (Ed.), *A Cooperative Investigation of Precision and Accuracy in Chemical, Spectrochemical and Modal Analysis of Silicate Rocks*. United States Geological Survey Bulletin. 980, pp. 59–68.
- Chayes, F., 1956. *Petrographic Modal Analysis: An Elementary Statistical Appraisal*. Wiley, New York, pp. 113.
- Chayes, F., Fairbairn, H.W., 1951. A test of the precision of thin-section analysis by point counter. *Am. Mineral.* 36, 704–712.
- Davis, J.C., 2002. *Statistics and Data Analysis in Geology*, Third ed. Wiley, New York 656 p.
- Dawson, K.R., Whitten, E.H.T., 1962. The quantitative mineralogical composition and variation of the Lacorne, La Motte, and Preissac Granitic complex, Quebec, Canada. *J. Petrol.* 3, 1–37. <https://doi.org/10.1093/petrology/3.1.1>.
- Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A.S., McLoone, S., 2013. Principal Component Analysis on Spatial Data: an Overview. *Ann. Assoc. Am. Geogr.* 103, 106–128. <https://doi.org/10.1080/00045608.2012.689236>.
- Dickinson, W.R., 1988. Provenance and sediment dispersal in relation to paleotectonics and paleogeography of sedimentary basins. In: Kleinspehn, K.L., Paola, C. (Eds.), *New Perspectives in Basin Analysis*. Springer, New York, NY, pp. 3–25. [https://doi.org/10.1007/978-1-4612-3788-4\\_1](https://doi.org/10.1007/978-1-4612-3788-4_1).
- Forte, A.M., Yanites, B.J., Whipple, K.X., 2016. Complexities of landscape evolution during incision through layered stratigraphy with contrasts in rock strength. *Earth Surf. Process. Landf.* 41, 1736–1757. <https://doi.org/10.1002/esp.3947>.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, pp. 483.
- Graham, S.A., Tolson, R.B., DeCelles, P.G., Ingersoll, R.V., Bargar, E., Caldwell, M., Cavazza, W., Edwards, D.P., Follo, M.F., Handschy, J.F., Lemke, L., Moxon, I., Rice, R., Smith, G.A., White, J., 1986. Provenance modelling as a technique for analysing source terrane evolution and controls on foreland sedimentation. *Int. Assoc. Sedimentol. Spec. Publ.* 8, 425–436. <https://doi.org/10.1002/9781444303810.ch23>.
- Grantham, J.H., Velbel, M.A., 1988. The influence of climate and topography on rock-fragment abundance in modern fluvial sands of the southern Blue Ridge mountains, North Carolina. *J. Sediment. Petrol.* 58, 219–227. <https://doi.org/10.1306/212F8D5F-2B24-11D7-864800102C1865D>.
- Griffiths, J.C., 1967. *Scientific Method in Analysis of Sediments*. McGraw-Hill, New York, pp. 508.
- Heins, W.A., 1992. *The Effect of Climate and Topography on the Composition of Modern Plutoniclastic Sand*. 465 University of California, Los Angeles PhD thesis.
- Heins, W.A., 1993. Source rock texture versus climate and topography as controls on the composition of modern plutoniclastic sand. In: Geological Society of America Special Paper. 284. pp. 135–146. <https://doi.org/10.1130/SPE284-p135>.
- Heins, W.A., 1995. The use of mineral interfaces in sand-sized rock fragments to infer ancient climate. *Geol. Soc. Am. Bull.* 107, 113–125. [https://doi.org/10.1130/0016-7606\(1995\)107%3C0113: TUOMII%3E2.3.CO;2](https://doi.org/10.1130/0016-7606(1995)107%3C0113: TUOMII%3E2.3.CO;2).
- Heins, W.A., Kairo, S., 2007. Predicting sand character with integrated genetic analysis. In: Geological Society of America Special Paper 420, pp. 345–380. [https://doi.org/10.1130/2006.2420\(20](https://doi.org/10.1130/2006.2420(20).
- Haneraas, S., Donselaar, M.E., Bloemsma, M.R., Tjallingii, R., De Wijn, B., Weltje, G.J., 2019. Quantitative integration of sedimentological core descriptions and petrophysical data using high-resolution XRF core scans. *Mar. Pet. Geol.* 110, 450–462. <https://doi.org/10.1016/j.marpetgeo.2019.07.034>.
- Higgins, M.D., 2000. Measurement of crystal size distributions. *Am. Mineral.* 85, 1105–1116. <https://doi.org/10.2138/am-2000-8-901>.
- Higgins, M.D., 2006. *Quantitative Textural Measurements in Igneous and Metamorphic Petrology*. Cambridge University Press, Cambridge, pp. 270.
- Ibbeken, H., Schleyer, R., 1991. *Source and Sediment: A Case Study of Provenance and Mass Balance at an Active Plate Margin (Calabria, Southern Italy)*. Springer-Verlag, Berlin, pp. 286.
- Ingersoll, R.V., 1990. Actualistic sandstone petrofacies: discriminating modern and ancient source rocks. *Geology* 18, 733–736. [https://doi.org/10.1130/0091-7613\(1990\)018<0733:ASPDMA>2.3.CO;2](https://doi.org/10.1130/0091-7613(1990)018<0733:ASPDMA>2.3.CO;2).
- Isaaks, E.H., Srivastava, R.M., 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, pp. 592.
- Johnson, M., 1987. *Multivariate Statistical Simulation*. Wiley, New York, pp. 212.
- Johnsson, M.J., 1993. The system controlling the composition of clastic sediments. In: Geological Society of America Special Paper 284, pp. 1–19. <https://doi.org/10.1130/SPE284-p1>.
- Kirkby, M.J., Cox, N.J., 1995. A climatic index for soil erosion potential (CSEP) including

- seasonal and vegetation factors. *Catena* 25, 333–352. [https://doi.org/10.1016/0341-8162\(95\)00016-L](https://doi.org/10.1016/0341-8162(95)00016-L).
- Marsh, B.D., 1988. Crystal size distribution (CSD) in rocks and the kinetics and dynamics of crystallization. *Contrib. Mineral. Petrol.* 99, 277–291. <https://doi.org/10.1007/BF00375362>.
- McCanta, M.C., Dyar, M.D., Dobosh, P.A., 2017. Extracting bulk rock properties from microscale measurements: subsampling and analytical guidelines. *GSA Today* 27, 4–9. <https://doi.org/10.1130/GSATG290A.1>.
- McKinley, J., Hron, K., Grunsky, E., Reimann, C., de Caritat, P., Filzmoser, P., van den Boogaart, K.G., Tolosana-Delgado, R., 2016. The single component geochemical map: fact or fiction? *J. Geochem. Explor.* 162, 16–28. <https://doi.org/10.1016/j.gexplo.2015.12.005>.
- Murphy, B., Yurchak, R., Müller, S., 2020. GeoStat-Framework/PyKrig: v1.5.0 (Version v1.5.0). Zenodo. <https://doi.org/10.5281/zenodo.3739879>.
- Palomares, M., Arribas, J., 1993. Modern stream sands from compound crystalline sources: Composition and sand generation index. In: Geological Society of America Special Paper. 284, pp. 313–322. <https://doi.org/10.1130/SPE284-p313>.
- Paredis, B., Weltje, G.J., Heins, W.A., june 2020. Quantitative analysis of the joint variability of rock texture and composition in thin section: the Generalized Griffiths Descriptor and its application to sediment generation from granitoids. *Earth Sci. Rev.* 205, 103188. <https://doi.org/10.1016/j.earscirev.2020.103188>.
- Pawlowsky-Glahn, V., Egozcue, J.J., 2016. Spatial analysis of compositional data: a historical review. *J. Geochem. Explor.* 164, 28–32. <https://doi.org/10.1016/j.gexplo.2015.12.010>.
- Pawlowsky-Glahn, V., Olea, R.A., 2004. Geostatistical analysis of compositional data. In: *Studies in Mathematical Geology*. Oxford University Press, Oxford, pp. 204.
- Preisendorfer, R.W., 1988. *Principal Component Analysis in Meteorology and Oceanography. Developments in Atmospheric Science* 17. Elsevier, Amsterdam (425 p.).
- Ripley, B.D., 1987. *Stochastic Simulations*. Wiley, New York, pp. 214.
- Stevenson, C.T.E., Hutton, D.H.W., Price, A.R., 2008. The Travenagh Bay Granite and a new model for the emplacement of the Donegal Batholith. *Trans. R. Soc. Edinb. Earth Sci.* 97, 455–477. <https://doi.org/10.1017/S0263593300001565>.
- Streckeisen, A.L., 1974. Classification and nomenclature of plutonic rocks. Recommendations of the IUGS subcommission on the systematics of igneous rocks. *Geol. Rundsch.* 63, 773–785. <https://doi.org/10.1007/BF01820841>.
- Streckeisen, A.L., 1980. IUGS subcommission on the systematics of igneous rocks. Classification and nomenclature of volcanic rocks, lamprophyres, carbonatites and melilite Rocks. Recommendations and suggestions. *Geol. Rundsch.* 69, 194–207. <https://doi.org/10.1007/BF01869032>.
- Suttner, L.J., Dutta, P.K., 1986. Alluvial sandstone composition and paleoclimate: framework mineralogy. *J. Sediment. Petrol.* 56, 329–345. <https://doi.org/10.1306/212F8909-2B24-11D7-8648000102C1865D>.
- Taubeneck, W.H., 1957. Geology of the Elkhorn Mountains, Northeastern Oregon: Bald Mountain batholith. *Bull. Geol. Soc. Am.* 68, 181–238. [https://doi.org/10.1130/0016-7606\(1957\)68\[181:GOTEMN\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1957)68[181:GOTEMN]2.0.CO;2).
- van den Boogaart, K.G., Tolosana-Delgado, R., 2013. Analysing Compositional data with R. Springer, Heidelberg (269 p.) doi:<https://doi.org/10.1007/978-3-642-36809-7>.
- Tolosana-Delgado, R., van den Boogaart, K.G., 2013. Joint consistent mapping of high-dimensional geochemical surveys. *Math. Geosci.* 45, 983–1004. <https://doi.org/10.1007/s11004-013-9485-y>.
- Tolosana-Delgado, R., von Eynatten, H., Karius, V., 2011. Constructing modal mineralogy from geochemical composition: a geometric-Bayesian approach. *Comput. Geosci.* 37, 677–691. <https://doi.org/10.1016/j.cageo.2010.08.005>.
- Tolosana-Delgado, R., Mueller, U., van den Boogaart, K.G., 2019. Geostatistics for compositional data: an overview. *Math. Geosci.* 51, 485–526. <https://doi.org/10.1007/s11004-018-9769-3>.
- Vermeesch, P., 2005. Statistical uncertainty associated with histograms in the Earth sciences. *J. Geophys. Res.* 110. <https://doi.org/10.1029/2004JB003479>.
- Vistelius, A.B., Agterberg, F.P., Divi, S.R., Hogarth, D.D., 1983. A stochastic model for the crystallization and textural analysis of a fine-grained granite stock near Meech Lake, Gatineau Park. In: *Canadian Geological Survey Paper 81–21*, pp. 62.
- von Eynatten, H., Barcelo-Vidal, C., Pawlowsky-Glahn, V., 2003. Modeling compositional change: the example of chemical weathering of granitoid rocks. *Math. Geol.* 35, 231–251. <https://doi.org/10.1023/A:1023835513705>.
- von Eynatten, H., Tolosana-Delgado, R., Karius, V., 2012. Sediment generation in modern glacial settings: grain-size and source-rock control on sediment composition. *Sediment. Geol.* 280, 80–92. <https://doi.org/10.1016/j.sedgeo.2012.03.008>.
- von Eynatten, H., Tolosana-Delgado, R., Karius, V., Bachmann, K., Caracciolo, L., 2016. Sediment generation in humid Mediterranean setting: grain-size and source-rock control on sediment geochemistry and mineralogy (Sila Massif, Calabria). *Sediment. Geol.* 336, 68–80. <https://doi.org/10.1016/j.sedgeo.2015.10.008>.
- Wackernagel, H., 1994. Cokriging versus kriging in regionalized multivariate data analysis. *Geoderma* 62, 83–92. [https://doi.org/10.1016/0016-7061\(94\)90029-9](https://doi.org/10.1016/0016-7061(94)90029-9).
- Wackernagel, H., 2003. *Multivariate Geostatistics: An Introduction with Applications*. Springer, Berlin. <https://doi.org/10.1007/978-3-662-05294-5>. (388 p.).
- Wadsworth, W.B., 1968. The Cornelia Pluton, Ajo, Arizona. *Econ. Geol.* 63, 101–115. <https://doi.org/10.2113/gsecongeo.63.2.101>.
- Wadsworth, W.B., 1975. Petrogenetic significance of grain-transition probabilities, Cornelia Pluton, Ajo, Arizona. *Geol. Soc. Am. Memoir* 142, 257–282. <https://doi.org/10.1130/MEM142-p257>.
- Weltje, G.J., 1997. End-member modeling of compositional data: numerical-statistical algorithms for solving the explicit mixing problem. *Math. Geol.* 29, 503–549. <https://doi.org/10.1007/BF02775085>.
- Weltje, G.J., 2004. A quantitative approach to capturing the compositional variability of modern sands. *Sediment. Geol.* 171, 59–77. <https://doi.org/10.1016/j.sedgeo.2004.05.010>.
- Weltje, G.J., 2012. Quantitative models of sediment generation and provenance: state of the art and future developments. *Sediment. Geol.* 280, 4–20. <https://doi.org/10.1016/j.sedgeo.2012.03.010>.
- Weltje, G.J., Tjallingii, R., 2007. Calibration of XRF core scanners for quantitative geochemical logging of sediment cores: theory and application. *Earth Planet. Sci. Lett.* 274, 423–438. <https://doi.org/10.1016/j.epsl.2008.07.054>.
- Weltje, G.J., von Eynatten, H., 2004. Quantitative provenance analysis of sediments: review and outlook. *Sediment. Geol.* 171, 1–11. <https://doi.org/10.1016/j.sedgeo.2004.05.007>.
- Weltje, G.J., Meijer, X.D., de Boer, P.L., 1998. Stratigraphic inversion of siliciclastic basin fills: a note on the distinction between supply signals resulting from tectonic and climatic forcing. *Basin Res.* 10, 129–153. <https://doi.org/10.1046/j.1365-2117.1998.00057.x>.
- Weltje, G.J., Bloemsma, M.R., Tjallingii, R., Heslop, D., Röhl, U., Croudace, I.W., 2015. Prediction of geochemical composition from XRF-core-scanner data: A new multi-variate approach including automatic selection of calibration samples and quantification of uncertainties. In: Croudace, I., Rothwell, R. (Eds.), *Micro-XRF Studies of Sediment Cores. Developments in Paleoenvironmental Research* 17. Springer, Dordrecht, pp. 507–534. [https://doi.org/10.1007/978-94-017-9849-5\\_21](https://doi.org/10.1007/978-94-017-9849-5_21).
- Weltje, G.J., Paredis, B., Caracciolo, L., Heins, W.A., 2018. Quantitative analysis of crystal-interface frequencies in granitoids: implications for modelling of parent-rock texture and its influence on the properties of plutoniclastic sands. *Sediment. Geol.* 375, 72–85. <https://doi.org/10.1016/j.sedgeo.2018.01.004>.
- Whitten, E.H.T., 1955. The metasediments of Bunbeg (County Donegal), and their relationship to the surrounding granite. *Proc. Geol. Assoc.* 66, 51–67. [https://doi.org/10.1016/S0016-7878\(55\)80003-X](https://doi.org/10.1016/S0016-7878(55)80003-X).
- Whitten, E.H.T., 1957. The Gola Granite (Co. Donegal) and its regional setting. *Proc. Royal Irish Acad. Section B* 58, 245–292. <https://www.jstor.org/stable/20490925>.
- Whitten, E.H.T., 1959. Composition trends in a granite: Modal variation and ghost stratigraphy in part of the Donegal granite, Eire. *J. Geophys. Res.* 64, 835–848. <https://doi.org/10.1029/JZ064i007p00835>.
- Whitten, E.H.T., 1961. Quantitative areal modal analysis of granitic complexes. *Geol. Soc. Am. Bull.* 72, 1331–1360. [https://doi.org/10.1130/0016-7606\(1961\)72\[1331:QAMAOG\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1961)72[1331:QAMAOG]2.0.CO;2).