

# Unsupervised machine learning with petrological database ApolloBasaltDB reveals complexity in lunar basalt major element oxide and mineral distribution patterns

Kim A. Cone <sup>a,\*</sup>, Richard M. Palin <sup>a,b</sup>, Kamini Singha <sup>a</sup>

<sup>a</sup> Geology and Geological Engineering Department, Colorado School of Mines, 1516 Illinois St., Golden, CO 80401, USA

<sup>b</sup> Department of Earth Sciences, University of Oxford, 3 South Parks Road, Oxford OX1 3AN, UK



## ARTICLE INFO

### Keywords:

Moon, surface

Volcanism

Data reduction techniques

## ABSTRACT

Diversity of lunar basalt characteristics is partly a consequence of lunar mantle heterogeneity. Although the cumulate mantle overturn hypothesis is the current standard model invoked to explain mantle asymmetries of unknown length scale in both compositional and geometrical space, successful petrological modeling of this mixing event requires a specific set of parameters not currently agreed upon. In contrast, surface basalt patterns may yield clues to both localized and nearside lunar interior structure.

Using two multidimensional data analysis approaches – principal component analysis (PCA) and K-means cluster analysis (KCA) – we report the patterns produced from basalt characteristics over changing spatial scales, from intra-site to inter-site to nearside. The data are sourced from a newly developed, self-contained database of lunar basalt characteristics (ApolloBasaltDB), which includes major element oxides, mineral modes, ages, and textures for petrological and statistical modeling. Through the simultaneous considerations of multiple basalt characteristics contained in the database, we find that terrestrial-based basalt classifications cannot adequately describe the complex and overlapping distribution patterns of major element oxides and mineral modes that define multiple distinct basalt groupings over multidimensional space. These patterns provide opportunities for alternative lunar basalt classification schemes. Our analyses suggest that  $\text{Al}_2\text{O}_3$  volumetric content is more diverse inside the Procellarum KREEP Terrane rift boundary versus content for older Apollo samples in close proximity to the eastern arm of the same rift boundary. Northernmost basalt samples show increased pyroxene diversity. Easternmost sites suggest anti-correlations in modal ilmenite and plagioclase, based on major element oxide PCA biplots, while nearside analyses of either major element oxides or mineral modes similarly suggest plagioclase (and  $\text{Al}_2\text{O}_3$ ) diversity comes at the expense of ilmenite (and  $\text{TiO}_2$ ) diversity. There is evidence to suggest that approximate mineral content can be extracted from major element oxide data based on correlative patterns between major element oxide PCA biplots and mineral mode PCA biplots. These patterns have implications for remote sensing missions in that onboard data manipulation may provide lithologic basalt vectors of interest.

## 1. Introduction

Lunar basalts exhibit a wide range of geochemical, petrological, and textural characteristics that each relate to pre-, syn-, and post-eruptive conditions. These characteristics are well-documented to vary within and between Apollo mission sites (Zussman, 1972; Rhodes and Hubbard, 1973; Papike et al., 1976; Lindstrom and Haskin, 1978; Longhi, 1992; Ryder and Schuraytz, 2001). Inter-site comparisons among returned samples frequently display variations in major and trace element

content, mineral mode, texture, and isotopic age, which suggest localized compositional, thermal, and/or process-driven variations in the mantle source-region during melt generation, ascent, and extrusion. In this context, changes in lunar mantle states must reflect changes in 3D mantle compositional geometry. Key aspects of how this compositional distribution varies within the lunar mantle are intimately linked to forms of the cumulate mantle overturn event. This event reflects the end-stage of the lunar magma ocean (LMO) solidification that produced a negatively-buoyant upper layer, consequently forcing end-stage mantle

\* Corresponding author at: Geology and Geological Engineering Department, Colorado School of Mines, 1516 Illinois St., Golden, CO 80401, USA.

E-mail addresses: [kcone@mines.edu](mailto:kcone@mines.edu) (K.A. Cone), [richard.palin@earth.ox.ac.uk](mailto:richard.palin@earth.ox.ac.uk) (R.M. Palin), [ksingha@mines.edu](mailto:ksingha@mines.edu) (K. Singha).

mixing from an initially molten Moon (Elkins-Tanton et al., 2002, 2011; Elardo et al., 2011; Li et al., 2019), although the overturn event may have occurred before complete LMO solidification (Boukaré et al., 2018). The lunar interior was already stratified before complete solidification of the LMO ~100–150 Myr after the Moon-forming impact, with denser and early-formed Mg-rich olivine having preferentially accumulated in the lower mantle (Elkins-Tanton et al., 2011). During continued cooling, orthopyroxene precipitation increasingly characterized the upper mantle. At ~75 to 80% crystallization of the LMO, earlier-formed Ca-rich plagioclase floated to the surface to form a primary anorthositic highland crust (Taylor and Jakes, 1974; Elardo et al., 2011; Elkins-Tanton et al., 2011), the average global thickness of which has been determined to range from 34 to 50 km (Khan et al., 2000; Lognonné et al., 2003; Wieczorek et al., 2013). A secondary mafic crust was emplaced only after the formation of the primary anorthositic crust through both endogenic volcanism and impact-induced crustal thinning and heat generation (Jolliff et al., 2006). This secondary crust is restricted almost entirely to the lunar nearside in the form of lunar maria (i.e. basaltic “seas”).

Although end-member models representing both 1) continuous fractional crystallization and 2) equilibrium crystallization followed by fractional crystallization have produced similar stratification scenarios (Elardo et al., 2011 and references therein), the composition, orientation, and the length scale of mantle heterogeneity remain unconstrained. The cumulate mantle overturn event introduced both horizontal and vertical complexity to this already heterogeneous mantle through a gravitationally induced mixing event of indeterminate length scale. A late-stage layer enriched in incompatible elements including potassium (K), rare earth elements (REE), and phosphorus (P) (referred to as KREEP) between the anorthositic highland crust and the upper mantle formed as a consequence of LMO crystallization, and it is the dense  $TiO_2$ -bearing ilmenite in the KREEP layer that is thought to have been the primary driver for the solid-state overturn. Although Rayleigh-Taylor instability downwellings from the cumulate overturn event have been modeled (Li et al., 2019), the compositional and spatial extent to which this incompatible-rich layer is variably mixed with the underlying mantle is unknown. Assuming that lunar basalts are representative of mantle sources with geochemical characteristics established near the completion of the cumulate mantle overturn, basalt characteristics should reflect localized mantle conditions at the time of extrusion, degree of crustal contamination, post-extrusion processes, and cumulate mantle overturn geometry.

In this study, we present ApolloBasaltDB (ApolloBasaltDB v1.0, currently hosted at the Colorado School of Mines, Colorado, U.S.A.) – a newly developed petrological database containing geochemical, geochronological, mineralogical, and textural information for returned mare basalts from the Apollo 11, 12, 14, 15, 16, and 17 missions (Fig. 1). This database includes parent sample and subsample numbers, several early classification schemes, major element oxide concentrations, isotopic ages, analytical methods, reported mineral volumes, and general textures that can serve as crystallization-speed proxies. We demonstrate the potential use of the database predominantly through machine learning to infer localized-to-nearside patterns of lunar basalt characteristics. Such databases are increasingly useful as part of the data-driven movement to extract and predict larger and previously unrecognized spatio-temporal patterns of geochemical, petrological, and process-driven behavior across the solar system as well as solely on Earth. Terrestrial database equivalents hosting hundreds of thousands of similar analyses of basalt (e.g. EarthChem: <http://www.earthchem.org/>) have allowed for large-scale information processing, leading to benchmark studies that have quantified the extent of secular variations in major, minor, and trace element geochemistry and extend to geo-dynamics between the Archean and the present day (Keller and Schoene, 2012; Furnes et al., 2014; Condé et al., 2016; Ganne and Feng, 2017; Palin and Dyck, 2018). The information contained in ApolloBasaltDB is thus expected to be of interest to researchers actively investigating the

geological evolution of the Earth-Moon system or volcanism on other rocky planets in our solar system (e.g. Nimmo and McKenzie, 1998; Thomas et al., 2014; Wade et al., 2017).

## 2. Materials: context and construction of ApolloBasaltDB

Each Apollo mission employed a textural and/or geochemical classification framework reflecting dominant site-specific trends in sample characteristics. When sample trends sharply varied between successive missions, a new or adjusted classification framework was created. Here, we provide brief geological overviews of the Apollo sites pertaining to ApolloBasaltDB sample characteristics, the types and amounts of basaltic samples collected at each, and the classification schemes employed. Database construction and content are also explored.

### 2.1. Apollo sample and site overviews

#### 2.1.1. Apollo 11: Mare Tranquillitatis

The Apollo 11 landing site is situated in the southwestern region of a multi-flow mare plain in the ~800 km diameter, pre-Nectarian Tranquillitatis basin (Fig. 1) (Sutton and Schaber, 1971). The samples were collected ~400 m west of the 180-m diameter West crater (Jerde et al., 1994; Hiesinger and Head, 2006). The basin contains ejecta (including Highland materials) from younger basins and mare flows dated at c. 3.80–3.35 Ga (Rajmon et al., 1999). Approximately 22 kg of soil and rock samples were returned, including two 2-cm diameter core tubes.

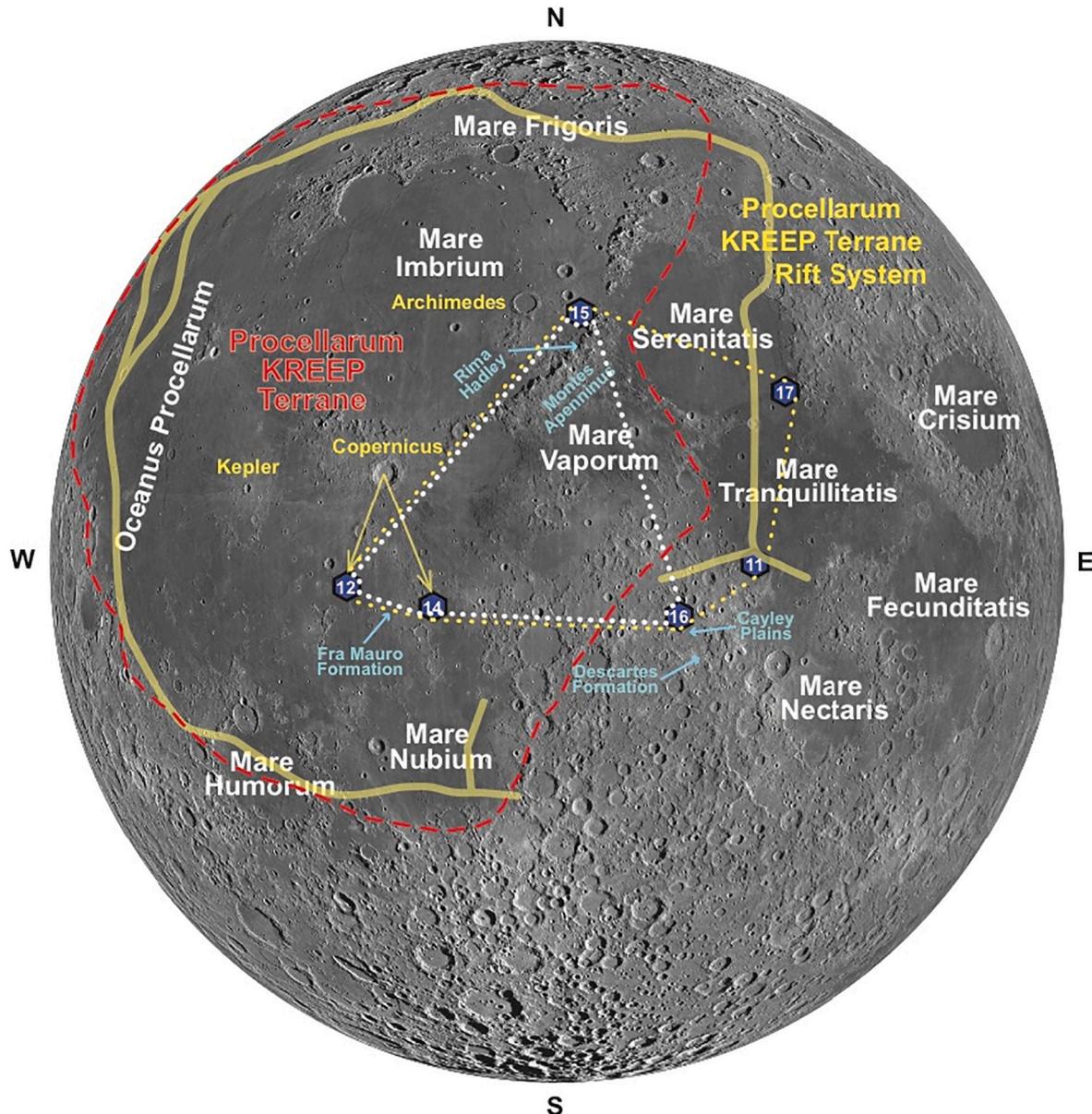
Basalts were primarily high-Ti/low-Al and were initially categorized by texture into four groups: Type A (fine-crystalline and vesicular), Type B (medium-crystalline, microgabbroic, or microdoleritic), Type C (brecciated), and Type D (fines - particles <1 cm) with later classification schemes reflecting a greater spectrum of textures (Zussman, 1972). Relations between texture, geochemistry, and age of crystallization were since established with a focus on trace element concentrations: Type A basalts were re-defined as high-K/high-REE, Type B as low-K/low-REE (further discriminated by their age of formation: B1 =  $3.67 \pm 0.02$  Ga, B2 =  $3.85 \pm 0.02$  Ga, and B3 =  $3.71 \pm 0.02$ ), and Type D as low-K/high-REE (Beatty and Albee, 1978). Type A basalts were identified as being the youngest of all samples collected, forming at  $3.59 \pm 0.02$  Ga (Jerde et al., 1994).

#### 2.1.2. Apollo 12: Oceanus Procellarum

The Apollo 12 site in the southeastern portion of Oceanus Procellarum is dominated by young mare basalts (c. 3.3–3.1 Ga) and minor outcrops of non-mare material derived from the nearby older Fra Mauro formation (Hiesinger and Head, 2006). Soil samples contain fragments younger than c. 3 Ga (parent soil 12,002). Ejecta rays from Copernicus and other craters likely contaminate the site with ejecta (Pieters and Wilhelms, 1985). The first KREEP samples were discovered at the site (parent sample 12,013), often in the form of ropy glasses that also contained KREEP basalt clasts (Wilhelms, 1984, 1987). Approximately 34 kg of rock and soil were returned, predominantly mare basalts, but which also included noritic and troctolitic varieties. Basalt samples here are typically richer in olivine than those collected from the Apollo 11 site (Wood et al., 1971). The general classification scheme for basalts from the Apollo 12 site comprises four petrological groups: olivine basalt, pigeonite basalt, ilmenite basalt, and feldspathic basalt (Neal et al., 1994). Aside from ilmenite-rich samples, most are low-Ti/low-Al/low-K.

#### 2.1.3. Apollo 14: Fra Mauro formation

The Fra Mauro formation is a highland valley ~1100 m west of the 340-m diameter Cone crater (Swann et al., 1971). The site is located ~500 km south of the southern edge of the Imbrium basin (Wilhelms and McCauley, 1971) and likely contains ejecta contamination from Copernicus crater situated to the north (Pieters and Wilhelms, 1985). Although located in the Highlands, the Fra Mauro formation lies within



**Fig. 1.** Apollo landing site locations (numbered) and major geological features of the lunar near side. Background image from the Goddard Space Flight Center, based on a mosaic of NASA Lunar Reconnaissance Orbiter images. Smooth, dark gray regions are mare basalts. The red dashed line represents the approximate boundary of a 3.5 ppm Th surface domain, after Jolliff et al. (2000). The solid yellow line traces the exterior location of the Procellarum KREEP Terrane rift system, as determined by Andrews-Hanna et al. (2014). The white dotted triangular region outlines the surface region covered by Apollo seismometers. The yellow dotted line approximates the boundaries of the Apollo missions discussed herein. Arrows leading away from Copernicus crater indicate impact contamination at sites 12 and 14. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the Procellarum KREEP Terrane (PKT), resulting in several samples and fragments in parent breccias rich in KREEP components (Dickinson et al., 1985). Approximately 42 kg of rock and soil samples were returned, and most mare basalts occurred as breccia clasts (Gancarz et al., 1971). The dominant basalt type is low-Ti and high-Al (Shervais et al., 1985). Sample 14321 (a basalt-bearing impact-melt/crystalline matrix breccia) contains a clast that has suggested a possible terrestrial origin (Bellucci et al., 2019).

#### 2.1.4. Apollo 15: Hadley-Apennines

The Apollo 15 landing site is situated at the eastern edge of the Mare Imbrium, between the western foot of the impact-induced Apennine Front and to the east of Hadley Rille (Wilhelms and McCauley, 1971; Spudis and Ryder, 1985). Among the ~77 kg of rock and soil samples collected were anorthosites, including “Genesis Rock” (a large early-formed fragment of primary lunar crust) (Steele and Smith, 1971), and a large basalt sample ~9.6 kg (parent sample 15555) (Rhodes and Hubbard, 1973). The Apollo 15 mission recovered Mg-rich pyroclastic green glasses, which were subsequently shown to have been derived from depths of ~400 km within the lunar mantle (Ridley et al., 1973; Delano, 1979). Mare basalts were mostly categorized as olivine-normative or quartz-normative types (the latter sometimes alternately referred to as pigeonite basalts), although some feldspathic KREEP basalts with crystallization ages of c. 3.9 Ga (parent rake sample 15382) were also collected (Dymek, 1986).

#### 2.1.5. Apollo 16: Descartes formation

Apollo 16 returned relatively few mare basalt samples, as the landing site was situated in the highlands on the Cayley Plains, adjacent to the Descartes Formation (Hodges et al., 1973; Andre and El-Baz, 1982; Spudis, 1984). The above units were initially assumed to be characterized by volcanics due to smooth topography, but the surface was dominated by breccias. Ejecta from Mare Nectaris exists at the site (McGetchin et al., 1973). A total of 96 kg of rock and soil samples were returned from the Apollo 16 site, with the few basaltic members being mostly aluminous and KREEP-rich (Wänke et al., 1973).

#### 2.1.6. Apollo 17: Taurus-Littrow

The Apollo 17 landing site is located in the Taurus-Littrow valley, on the southeastern edge of Mare Serenitatis, bounded by the Montes Taurus (Pieters et al., 1973; Haase et al., 2012). Nearly 111 kg of rocks and soils were returned, including noritic (78236) and troctolitic (76535) samples, multiple types of pyroclastic glass beads, and basalts (Longhi et al., 1974; Rhodes et al., 1976). Basalts were mostly high-Ti/low-Al/low-K and were originally classified on a textural basis by Brown et al. (1975) into three types: Type 1A = porphyritic olivine- and ilmenite-bearing basalt; Type 1B = olivine-poor and slower-cooled basalt; and Type II = olivine-absent, low-Mg, ilmenite basalt. Type 1A are the most common and one occurrence of a KREEP basalt was discovered in a breccia sample (72275) (Shih et al., 1992).

### 2.2. ApolloBasaltDB: structure and content

ApolloBasaltDB is organized by Apollo mission, with data for each contained in a separate worksheet in an Excel file. A summary of basic characteristics and general observations for each landing site is on the first worksheet (Intro\_Apollo\_Basalts). The second worksheet (Intro\_ABDB) describes database headers and their significance, such as the definition of low-K and high-K mare basalts, which varies between published studies. Here, we employ the 0.2 wt% cutoff for defining K content in a descriptive manner (Jerde et al., 1994); however, as this is

not consistently applied in the literature, the reader is encouraged to refer to the original source for further clarification. Cutoffs for qualitative descriptors for varying Al and Ti concentrations are also defined on the Intro\_ABDB worksheet, alongside mineral abbreviations. The subsequent six worksheets contain lunar basalt characteristics for Apollo missions 11, 12, 14, 15, 16, and 17. Data include sample and sub-sample numbers, general sample classifications and parent types, major element oxides, isotopic ages, coarse textural descriptions, and mineral modes.

Source references for major element oxides were initially extracted from Clive Neal’s 2008 unpublished Mare Basalt database, housed at the University of Notre Dame, Indiana (<https://www3.nd.edu/~cneal/Lunar-L/>). These references were reviewed for data transcription accuracy and any discrepancies noted appear in ApolloBasaltDB. Additional major element oxide, mineral mode, texture, and age data were subsequently added and reflect data from as recent as 2018. Additional data were compiled from 1) the NASA Lunar Sample Compendium (<https://curator.jsc.nasa.gov/lunar/lsc/>); 2) the Mineralogical Society of America’s *New Views of the Moon*; 3) The Lunar and Planetary Institute’s Basaltic Volcanism on the Terrestrial Planets (<http://ads.harvard.edu/books/bvtp/>); 4) conference abstracts; 5) articles published in scientific journals up to 2018; and 6) the Lunar and Planetary Institute’s online Apollo resources (<https://www.lpi.usra.edu/lunar/missions/apollo/>). Two worksheets of references finalize the database: one duplicated from Neal’s Mare Basalt database and a second documenting additional data added to ApolloBasaltDB. Specific sources for some data appearing on worksheets within ApolloBasaltDB are referenced therein.

To maximize ease of data comparisons between different missions, each worksheet provides similar basalt characteristics in a consistent format and using similar terms, when possible. When available, each sample entry (i.e. each row) includes (1) a parent sample number, (2) the nature of the analyzed sample relative to the parent sample (e.g., soil particle, rake sample, breccia clast, or basalt fragment), (3) a brief textural description, (4) a mission-specific geochemical classification, (5) a subsample number, (6) an assignment according to the three basic geochemical classifications employed for mare basalts (Ti, Al and K groups), (7) a reference for geochemical analysis, (8) isotopic age and method of determination, (9) reference for isotopic age, (10) major element oxides in weight percent, (11) notes on oxide entries, (12) minerals proportions, most commonly in volume percent, (13) reference for mineralogical data, and (14) related notes. An isotopic age for a given sample/subsample does not necessarily correspond to a specific analysis from a specific row, but if so, the age source-reference and the analytical source-reference is identical. In total, there are 330 row entries for Apollo 11 samples, 278 for Apollo 12, 285 for Apollo 14, 686 for Apollo 15, 129 for Apollo 16, and 389 for Apollo 17, comprising 2097 entries for all missions.

A single parent sample number can contain multiple entries when multiple analyses have been performed on the same sample/subsample by different research groups. The results of each analysis are then listed on separate rows. Mineral modes corresponding to a particular sample/subsample may or may not be listed in a separate row from major element oxides. In the event that mineral modes and major element oxides reflect the same sample/subsample, the reader is referred to the related “Reference” columns to determine if the analyses are derived from the same study. Isotopic ages for subsample analyses from a single larger basalt rock are often listed in adjacent rows at the beginning of a given parent sample number, as these analyses are assumed to reflect homogeneity within a larger basalt rock sample.

A complete analysis of data within ApolloBasaltDB is beyond the scope of this study, although examples of the potential for multidimensional geological interpretations are given below, in part focusing

**Table 1**

Variances for principal components 1 and 2 for site-specific major element oxide analyses and strongest/weakest major element oxide co-trends; df = diversity factor.

Mission	n	PC1 (%)	PC2 (%)	(PC1 + PC2) %	(PC1-PC2) %	df	Maximum co-trending oxides (PC1)	Anti-trend along PC1
A11	78	44	19.3	63.3	24.7	2.6	CaO/Al <sub>2</sub> O <sub>3</sub> ; FeO/TiO <sub>2</sub>	pairs anticorrelate
A12	56	55.7	22.1	77.8	33.6	2.3	CaO/Al <sub>2</sub> O <sub>3</sub>	MgO
A14	49	55.3	14.9	70.2	40.4	1.7	CaO/Al <sub>2</sub> O <sub>3</sub> ; MgO/FeO	pairs anticorrelate
A15	255	53.3	20.9	74.2	32.4	2.3	Al <sub>2</sub> O <sub>3</sub> /NaO	FeO
A16	7	57.1	26.7	83.8	30.4	2.8	CaO/Al <sub>2</sub> O <sub>3</sub> ; MgO/FeO	pairs anticorrelate
A17	187	44.8	29.2	74.0	15.6	4.7	Al <sub>2</sub> O <sub>3</sub> /K <sub>2</sub> O	FeO

on the new data types that succeed the 2008 Neal database structure. These data additions comprise mineral volumes, isotopic ages, and textures, which serve to enhance understanding of lunar basalt volcanism outside of strictly geochemical approaches.

### 3. Methods

#### 3.1. Data preparation

All data are derived from the database. Multiple major element oxide, mineral mode, and age analyses of a single parent rock sample number, when that parent rock is classified as a “basalt rock” in the database (third column), are averaged to produce a single series of major element oxide, mineral mode, and age values; the assumption is that each subsample analysis represents more-or-less a homogenous portion of a homogeneous parent rock. Identical subsample numbers followed by a different letter (e.g. subsamples: 100A and 100B) represent two separate analyses of a single subsample (subsample number 100). In this case, the two analyses would be averaged together before any further averaging with other subsamples of a parent rock. Major element oxide data were not weighted, and no preference was assumed for analytical method. In the event a parent rock is classified as a type of breccia (polymict, regolith-matrix, rake-fragmental, impact-melt, or crystalline-matrix), each clast within the breccia is treated as a separate and distinct basalt fragment, and no averaging of characteristics is performed unless stated otherwise. Soil basalt fragments are all considered as separate and distinct basalt rocks (or rocklets), and their characteristics are not averaged together unless stated otherwise. Isotopic ages are typically averaged together, and no assumptions of age accuracy are made. Lunar basalts consist mostly of four minerals: pyroxene, plagioclase, olivine, and ilmenite, with other minerals rarely contributing >5% to the total volume of any given sample (Heiken et al., 1991). For the purpose of mineral mode analyses, all reported species of pyroxenes are amalgamated into a single pyroxene family. Reported opaque volumes are treated as ilmenite when ilmenite is not reported separately. Although spinels may constitute a minor fraction of the original opaques, ilmenite is the major opaque mineral in lunar basalts (Jolliff et al., 2006), and this trend is noted in ApolloBasaltDB. For parent sample 14321, any major element oxide analysis used for multidimensional analyses are considered in terms of “Dickenson Group Averages” (Dickinson et al., 1985), and the reader should take note that the recent reclassification of part of 14321 from a felsite clast suggests a potential terrestrial origin (Bellucci et al., 2019). Some reported major element oxide values have no place (zero place) accuracy, and the reader should refer to the database for individual entries. Reported descriptive textures have been converted to numeric values such that fine grained = 1, fine-medium grained = 1.5, medium grained = 2, medium-coarse grained = 2.5, and coarse grained = 3. Although this scheme does not produce a numerically continuous set of values, given that textural descriptions in the literature are not always in agreement but do appear to generally coincide, the goal here is to associate descriptive texture to loosely-defined numerical values for preliminary comparisons against other basalt characteristics. In the database and in this paper, Apollo missions may be abbreviated by Axx, where xx represents the mission number (e.g. A15 for Apollo 15).

#### 3.2. Unsupervised learning

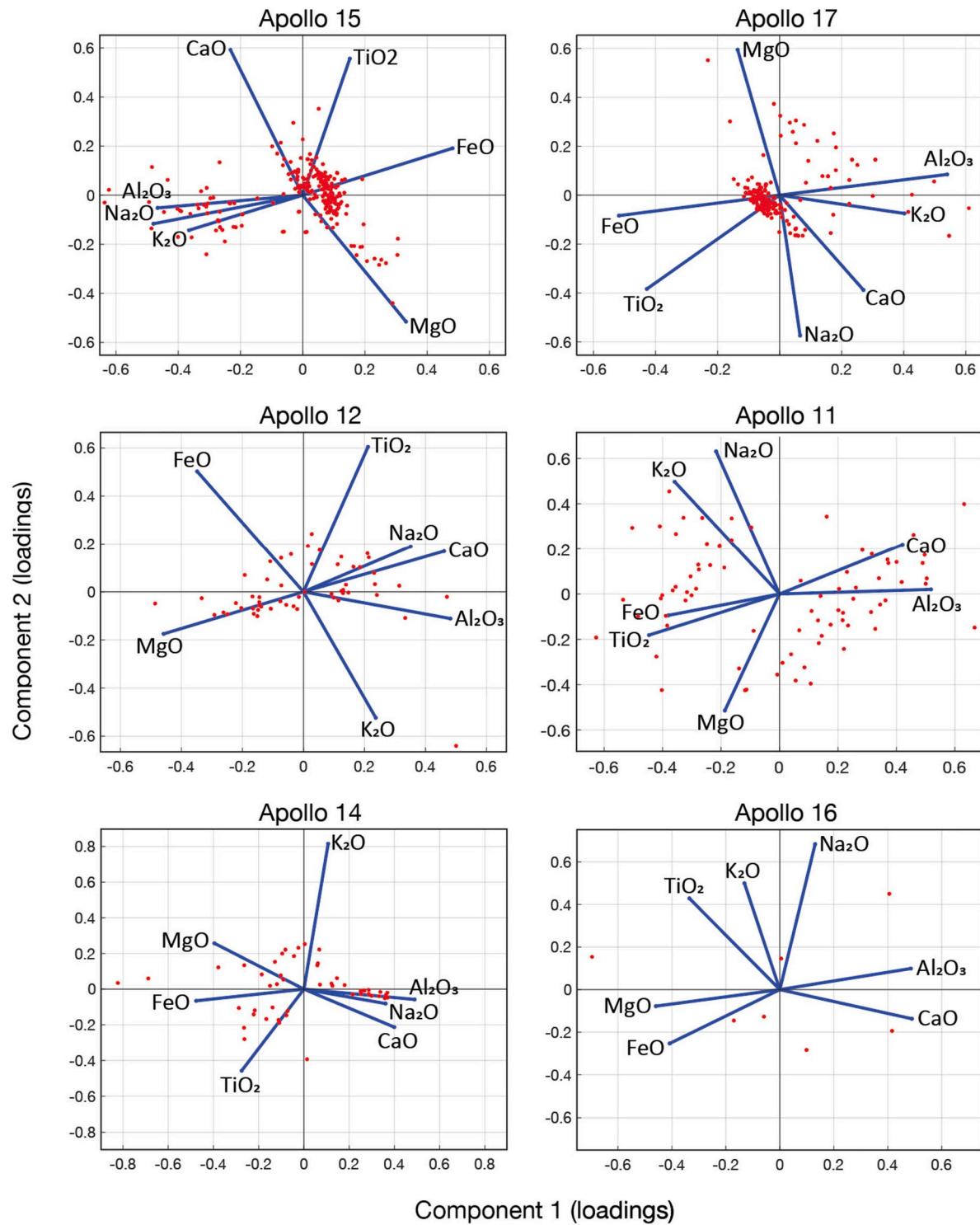
Two multivariate approaches are used here: principal component analysis (PCA) and K-means cluster analysis (KCA). PCA establishes significance among sample characteristics based on orthogonal and uncorrelated variance-based vectors, while KCA iteratively organizes (and reorganizes) sample characteristics and assigns data points to discretized clusters by minimizing the squared distance between data points and cluster means. Neither method requires Gaussian distributions, although such distributions would provide a more immediately robust basis from which conclusions may be drawn. However, if the underlying variable relationships in the data are meaningful, then non-Gaussian distributions may still provide results that reflect the true population (Dudzinski et al., 1975). In order to reveal potentially novel natural groupings, no inherent data patterns were assumed. Both methods are forms of unsupervised machine learning that are underutilized in petrological studies, although big data approaches have become increasingly popular in recent years for constraining terrestrial mineral occurrences (Hazen et al., 2015; Morrison et al., 2017). Both PCA and KCA have been employed in basalt and magma research (Butler, 1976; Allègre et al., 1995; Iwamori et al., 2017; Ueki and Iwamori, 2017) and provide non-traditional perspectives on the co-evolution of petrological traits.

Seven major element oxides are commonly reported from mare basalt geochemical analyses: TiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, FeO, MgO, CaO, Na<sub>2</sub>O, and K<sub>2</sub>O. Texture and age data were then added to major element oxide data for new analyses. However, because not all major element oxides have associated age and texture characteristics, only data containing all three types of entries per unique sample were kept. Mineral mode analyses included olivine (Ol), pyroxene (Px), plagioclase (Pl), and ilmenite (Ilm). Texture and age data were combined with mineral mode data for reanalysis. The addition of these two characteristics resulted in a reduced number of samples over nine variables in the case of major element oxide analysis (seven major element oxides, age, and texture) and over six variables in the case of mineral modal analysis (four minerals, age, and texture).

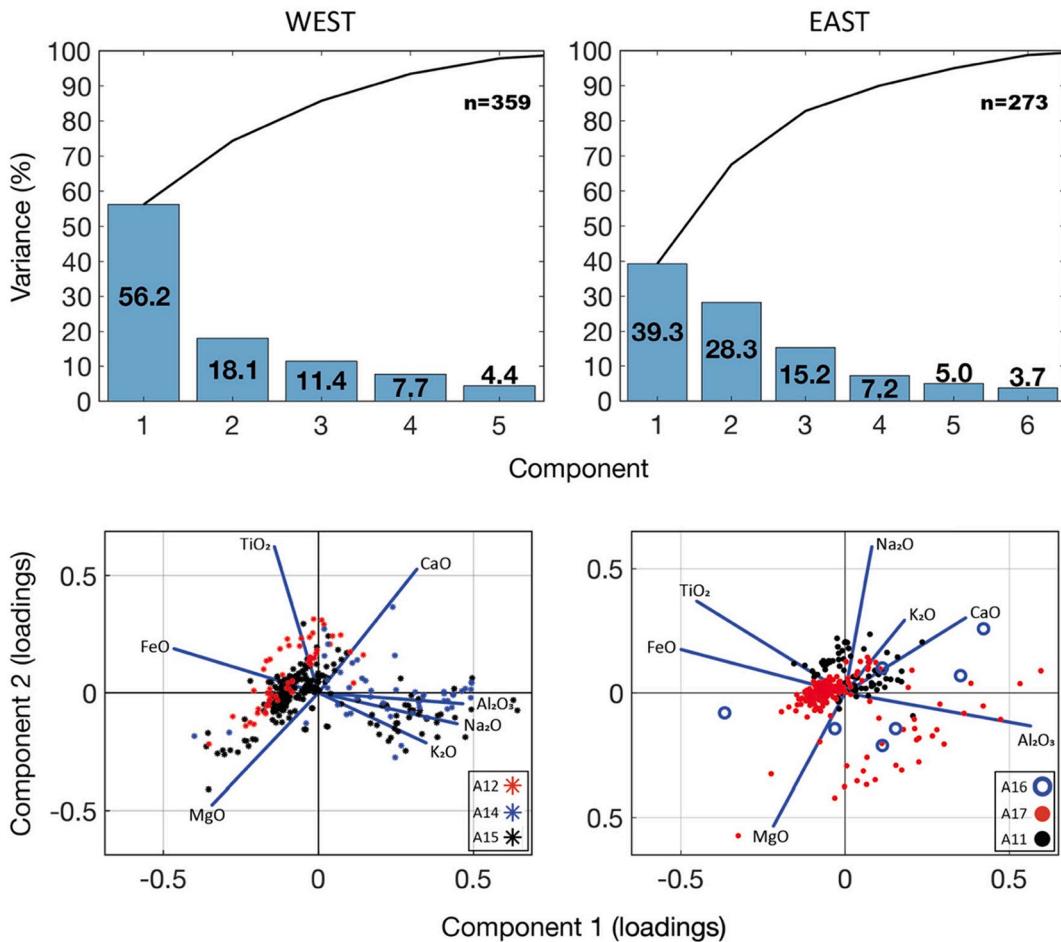
##### 3.2.1. Principal component analysis

PCA is a dimension-reduction approach that projects the original, multidimensional data onto a reduced number of summary indices, or variables. The newly established variables, or components, are orthogonally oriented to represent vectors of maximum uncorrelatedness. The first principal component (PC1) represents the greatest variance in the dataset and subsequent components for n variables (PC2, PC3, ..., PCn) represent increasingly reduced variance. The greater the inherent correlation between the original variables (in this case, basalt characteristics), either positively or negatively, then the greater the reduction in the number of original variables (i.e. fewer components). This technique projects the original data into the fewest number of dimensions that still account for the majority of data variance. Detailed descriptions of PCA for use in petrological studies are provided in LeMaitre (1982) and Albarède (1995).

PCA was performed on ApolloBasaltDB using singular value decomposition on the correlation matrix to produce principal component coefficients, or loading values. Columns of the coefficient matrix



**Fig. 2.** PCA biplots for unique lunar basalt analyses, reflecting the seven major element oxide weight percents and the variances shown in Supplementary Fig. 1. Apollo site 14 includes analyses for crystalline matrix breccia 14321, but only by the five Dickinson groups after [Dickinson et al. \(1985\)](#).



**Fig. 3.** Pareto diagrams (top row) and PCA biplots (bottom row) for major element oxide weight percents based on unique basalt sample analyses. West (Apollo 15, 12, and 14) and East (Apollo 17, 11, and 16) groups reflect proximity to the eastern PKT rift boundary (Fig. 1).

represent principal components PC1 to PC $n$ , from left to right, while rows represent the variables, where the maximum  $n$  is limited by the number of basalt characteristics considered. Loading values in the coefficient matrix range from  $-1$  to  $+1$  and indicate the relative strength and direction of the relationship between variables with respect to a given principal component. Standardization of data was applied, transforming all individual measurements into unit variance, zero-meanned datasets. Two principal components are displayed in PCA biplots, PC1 and PC2, which encompass the majority of variance from all Pareto variance plots. Three data groupings were analyzed: 1) site-specific, 2) West and East groups (West = Apollo sites 15, 12, 14; East = Apollo sites 17, 11, and 16), and 3) nearside (all unique sample data combined, or West + East). West and East were defined by site location with respect to the far east, vertically oriented PKT boundary (Fig. 1).

The difference between PC1 and PC2 approximates a variance flatness value where greater values indicate increased distribution flatness. The sum of PC1 and PC2 approximates the sum degree of variance in context of emphasizing only the first two principal components. Therefore,  $(PC1 + PC2)/(PC1 - PC2)$  approximates a diversity factor (df) about PC1 and PC2, where larger values indicate relatively decreasing diversity. When paired with pronounced major element oxide (or mineral mode) trends, diversity factors along with loading values should provide insight into how pronounced trends are.

### 3.2.2. K-means cluster analysis

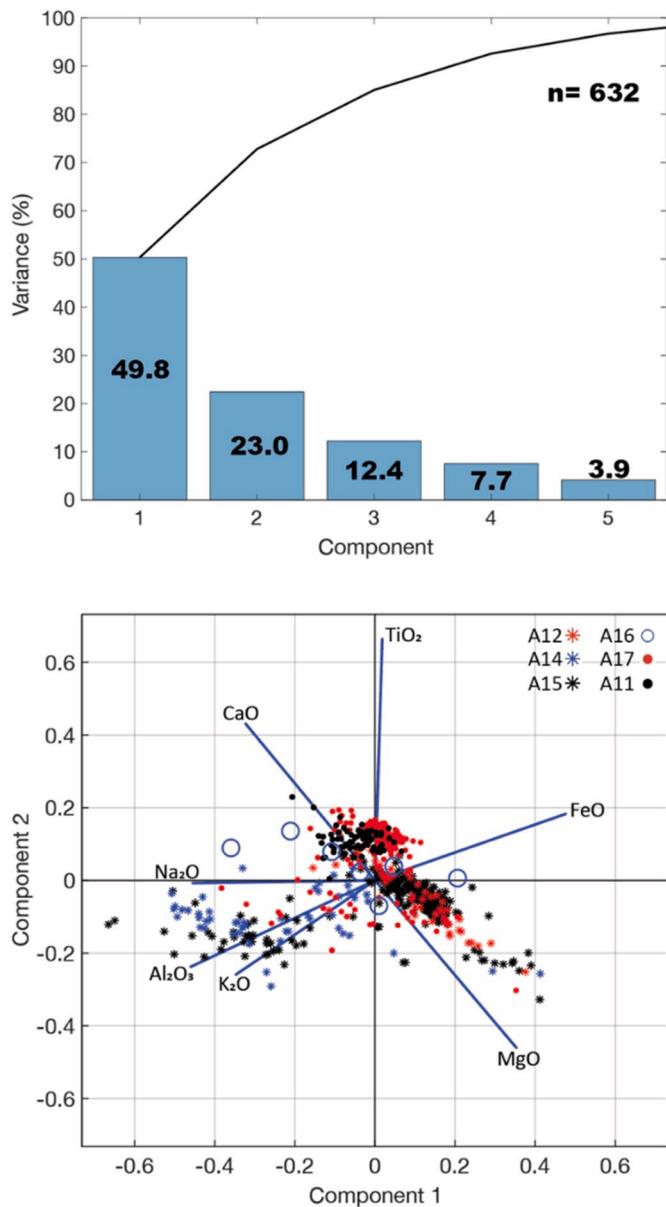
KCA iteratively assigns data to  $K$  discrete clusters based on randomly chosen  $K$  centroids, with an ideal end-goal of global minimization of the

average squared Euclidean distances between data in each cluster (MacQueen, 1967; Arthur and Vassilvitskii, 2006). In this form of centroid-based cluster analysis, some subset of the data is initially assigned to clusters. The centroid positions are then calculated, and each datum is assigned to the closest cluster based on those centroid positions. The process repeats until all of the data points are assigned to clusters and the centroids stabilize (i.e. convergence is reached). Here, the K-means++ algorithm is employed, after Arthur and Vassilvitskii (2006):

- (1) Randomly choose  $K$  initial centers  $C = \{c_1, \dots, c_K\}$ .
- (2) For each  $i$  that is an element of  $\{1, \dots, K\}$ , assign cluster  $C_i$  as the set of data points that are closer to  $c_i$  than to  $c_j$  for any  $j \neq i$ .
- (3) For each  $i$  that is an element of  $\{1, \dots, K\}$ , assign  $c_i$  as the center of mass for all points in  $C_i$  such that  $c_i = 1/(C_i)\exp(-1) \sum_{x \in C_i} x$ .
- (4) Steps 2 and 3 are repeated until  $C$  stabilizes.

As for PCA analysis, distribution shapes are not assumed (see Supplementary Fig. 6 for major element oxide distribution shapes).

In order to find the optimal configuration for any randomly-chosen  $K$  centroids, 1000 replicates per  $K$  were conducted to isolate the lowest local minimum, as KCA is sensitive to the initial choice of centroid locations, and individual runs often produce different local minima (Bradley and Fayyad, 1998). No previous assumptions were made regarding initial centroid placement, and so initial centroid locations were randomly selected for each replicate. Data preprocessing was conducted as for PCA described above, where the initial dataset was zero-meanned and normalized by unit variance.



**Fig. 4.** Pareto plot (top) and PCA biplot (bottom) reflecting trends in basalt characteristics for the lunar nearside as a whole. Data points represent major element oxide analyses for unique samples. Sample 14321, as part of the Apollo 14 sample set, reflects the five Dickenson groups (Dickenson et al., 1985).

K values from two to seven were evaluated with silhouette coefficients in combination with the K-means++ clustering algorithm. Silhouette coefficients range from -1 to +1 and represent how well a data point is matched to a cluster. Larger, more positive values indicate that data are well matched to their assigned cluster. A final number of clusters were determined such that a given K yielded the largest average silhouette coefficient. Selecting smaller K-values minimizes potential

overfitting of data to a greater number of clusters, and so only K ≤ 7 was considered. A silhouette value,  $S_i$ , was calculated for each data point,  $i$ , in the following manner:

$$S_i = (i_{min\_avg\_diff} - i_{min\_avg\_same})$$

where

$i_{min\_avg\_diff}$  represents the minimum average distance from each  $i$  to extra-cluster points, and  $i_{min\_avg\_same}$  represents the average distance from each  $i$  to other intra-cluster points.

Only the nearside group was considered here (the combination of all unique sample data), in a similar manner to PCA analysis as described above, for the purpose of investigating spatially large-scale trends in basalt characteristics.

#### 4. Results

Although geological processes responsible for the above trends in basalt characteristics cannot be directly discerned from principal component analysis or cluster algorithms, the results are discussed in the context of pattern evolution and cross-connectedness of characteristics through spatial-scale changes in Apollo site groupings (i.e. per-site to West-East to nearside). For initial exploratory interpretations, olivine, pyroxene, plagioclase, and ilmenite are considered to control major element distribution, with MgO and FeO largely present in olivine and pyroxene, Al<sub>2</sub>O<sub>3</sub> in plagioclase, and TiO<sub>2</sub> in ilmenite. Here, we probe the underlying data patterns in an attempt to reveal aspects of internal geologic conditions responsible for surface basalt patterns. Results are shown as 1) plots of variance per new component (Pareto plots) for a given set of basalt characteristics, 2) PCA biplots showing strength of the co-trends or anti-trends for basalt characteristics, 3) K-means cluster diagrams for nearside groups of major oxide elements and select mineral modes, and 4) summary tables for site-specific analyses.

##### 4.1. Principal component analysis: major element oxides

###### 4.1.1. Principal component analysis: site-specific major element oxides

Supplementary Fig. 1 displays variance distributions for up to six principal components for each of the Apollo site samples. Major element oxides from A16 samples show the greatest total intra-site variance based on PC1 and PC2 while the A11 samples display the smallest total variance for PC1 and PC2. A14 samples show the greatest difference between PC1 and PC2 variance (40.4%). A17 shows the overall flattest distribution in variance over PC1 and PC2. The strongest site-specific co-trends that produce loading values PC1 > ~0.4 are shown in Table 1 and Fig. 2.

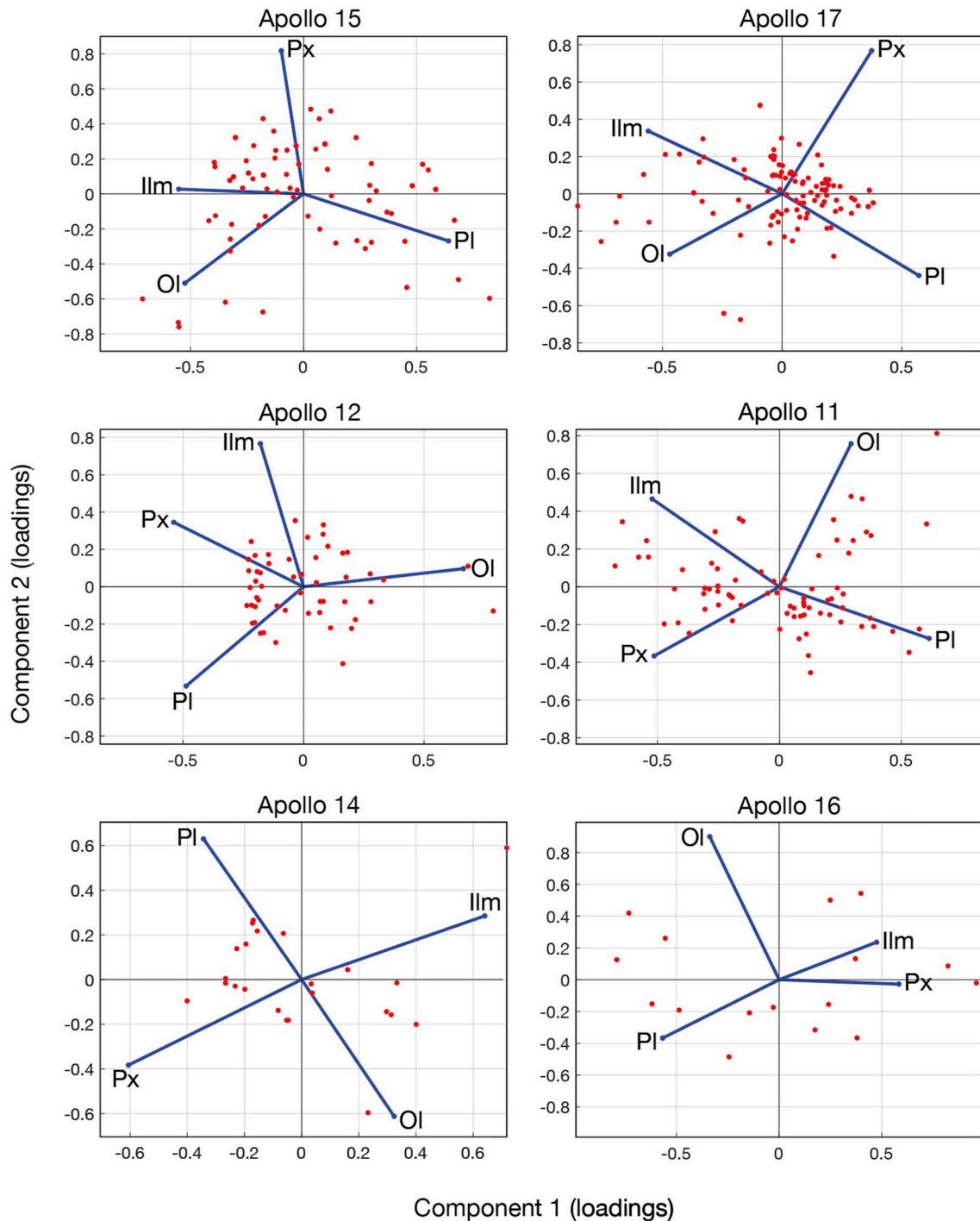
###### 4.1.2. Principal component analysis: west–east major element oxides

PC1 and PC2 from West and East groups account for 74.3% and 67.6% of total component variance, respectively (Fig. 3). Variance distribution is flatter for the East group, resulting in a standard deviation of 14.4 versus 21.1 for the West group, based on the component variances in Fig. 3. This increased flatness for the East group suggests decreased ranges of K<sub>2</sub>O and Na<sub>2</sub>O, particularly when compared to the West group. Co-trending major element oxides for the West group indicate Al<sub>2</sub>O<sub>3</sub>/Na<sub>2</sub>O pairing strongly correlates while FeO anti-correlates to the pair.

**Table 2**

Variances for principal components 1 and 2 for site-specific mineral modes and strongest/weakest mineral mode co-trends. Diversity factor = df.

Mission	n	PC1 (%)	PC2 (%)	(PC1+PC2) %	(PC1-PC2) %	df	Maximum co-trending mineral/s (PC1)	Anti-trend along PC1
A11	76	53.3	27.2	80.5	26.1	3.1	Plagioclase	Pyroxene
A12	51	50.3	30.3	80.6	20.0	4.0	Olivine	Pyroxene
A14	23	43.6	30.9	74.5	12.7	5.9	Ilmenite	Pyroxene
A15	76	44.7	34.3	79	10.4	7.6	Plagioclase	Ilmenite
A16	16	64.5	21.6	86.1	42.9	2.0	Ilmenite/Pyroxene	Plagioclase
A17	107	52	22.9	74.9	29.1	2.6	Ilmenite/Olivine (weak)	no strong anti-trend



**Fig. 5.** PCA biplots for four mineral modes (olivine, pyroxene, plagioclase, and ilmenite), reflecting component variance from Supplementary Fig. 2, for each of the Apollo sites.

No pronounced major element oxide pairings occur in the East group, but  $\text{Al}_2\text{O}_3$  and  $\text{FeO}$  strongly anticorrelate along PC1.

#### 4.1.3. Principal component analysis: nearside major element oxides

Nearside major element oxide trends suggest that 72.8% of component variance is tied solely to PC1 and PC2 (Fig. 4), similar to the averaged value for ( $\text{PC1} + \text{PC2} = 73.9\%$ ) from Supplementary Fig. 1 and Table 1. PCA biplot patterns from the Nearside group (Fig. 4) resemble those of the West group (Fig. 3), suggesting that the West group controls more major element oxide diversity versus the East group.  $\text{TiO}_2$  accounts for almost all of the variance bound to PC2 (Fig. 4). The  $\text{Na}_2\text{O}/\text{Al}_2\text{O}_3$  co-trend along PC1 anti-correlates with  $\text{FeO}$ .

#### 4.2. Principal component analysis: mineral modes

##### 4.2.1. Principal component analysis: site-specific mineral modes

Intra-site mineral mode patterns are summarized in Table 2, incorporating results from Figs. 6 and 7. Based on loading values in Fig. 5, pyroxene and ilmenite modes increase together at A16 with all other trends being non-correlative or anti-correlative.

##### 4.2.2. Principal component analysis: west-east mineral modes

Both West and East groups display predominantly flat variance distributions across components. Mineral mode anti-correlation pairs (ilmenite/plagioclase, olivine/pyroxene) characterize the East group and an olivine-dominant olivine/ilmenite co-trend occurs for the West group with no strong anti-correlative pairings.

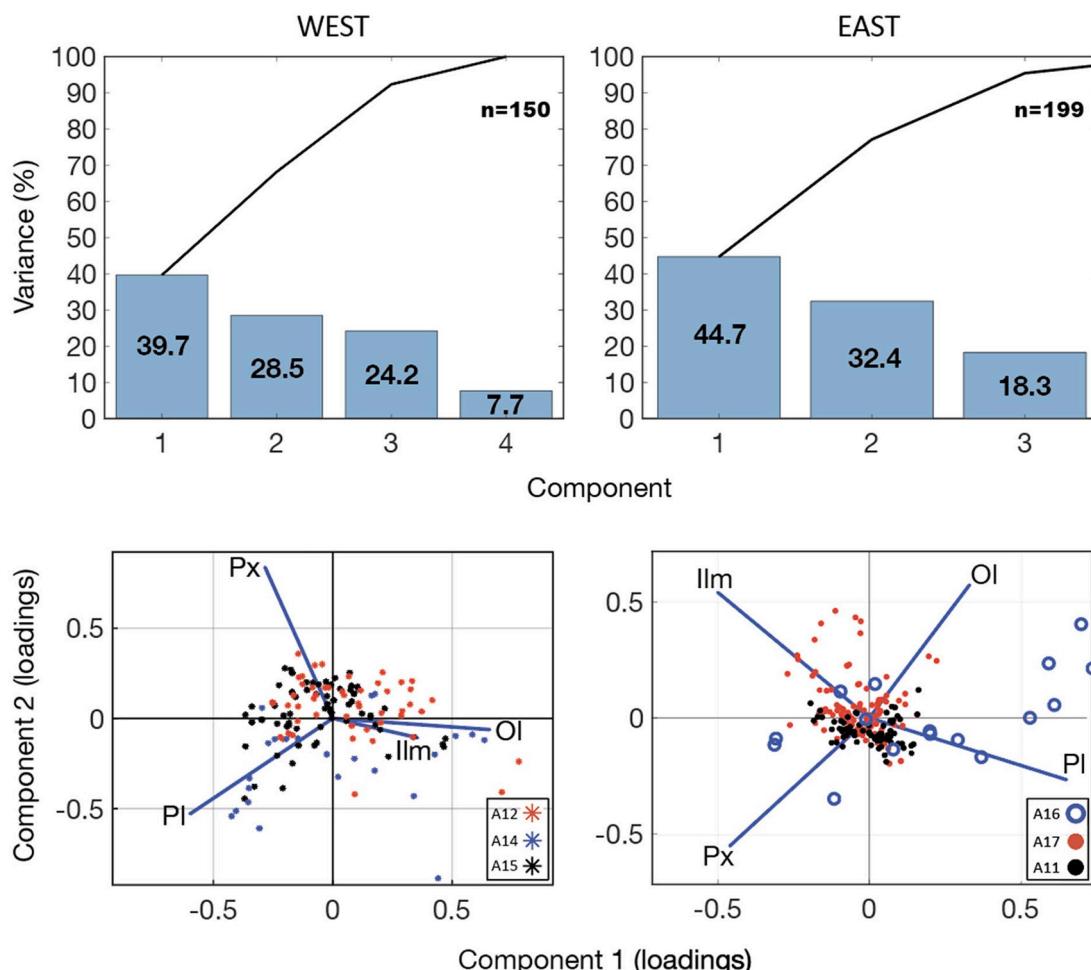
#### 4.2.3. Principal component analysis: nearside mineral modes

In general, nearside mineral mode trends for ilmenite (PC1 dominant) and olivine (PC2 dominant) are non-correlative but not strongly anti-correlative (Fig. 7). Plagioclase constitutes a “complex” loading value and contributes moderately more to PC1 than PC2. Pyroxene in general provides weak loading values for both PC1 and PC2.

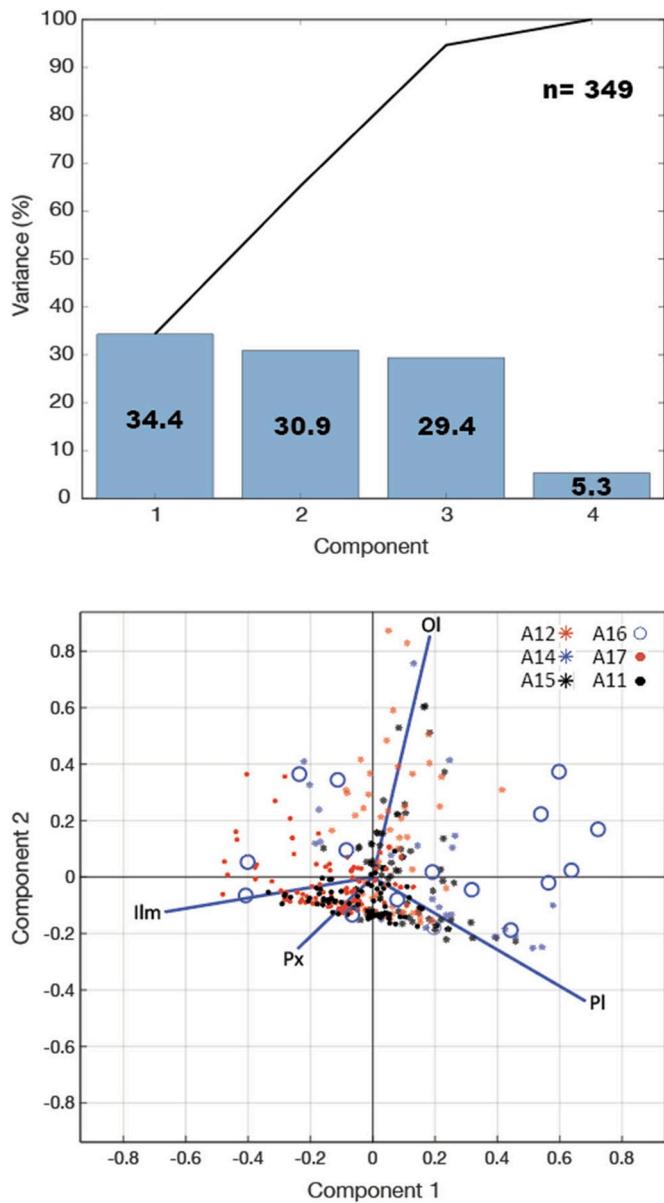
#### 4.3. Principal component analysis: ages, textures, and major element oxides

##### 4.3.1. Principal component analysis: site-specific ages, textures, and major element oxides

Variance accounted for by PC1 is the most pronounced for A15 and A12, and decreasingly so for A14, A11, and A17 (Table 3 and Supplementary Fig. 3). A11 basalt characteristics display co-trending behavior along PC1 for age/texture/ $\text{Al}_2\text{O}_3$  with no strong anti-correlative characteristics. A12 samples suggest an  $\text{Al}_2\text{O}_3/\text{CaO}$  pairing that anti-correlates with  $\text{MgO}$  along PC1. Characteristics for A14 basalts display the strongest anti-correlative pairings along PC1 as  $\text{CaO}/\text{Al}_2\text{O}_3$  and  $\text{MgO}/\text{FeO}$ . A16 basalt characteristic analyses were not performed due a lack of published characteristics for any single A16 basalt sample. A17 produced the flattest overall distribution in variance over seven components (Supplementary Fig. 3). PC2 for sites for A12, A15, and A14 is mostly controlled by  $\text{TiO}_2$  while PC2 for sites A11 and A17 is predominantly affected by  $\text{MgO}$  and  $\text{Al}_2\text{O}_3$ , respectively (Fig. 8).



**Fig. 6.** Pareto diagrams (top row) and PCA biplots (bottom row) for the four mineral modes based on West (Apollo 15, 12, and 14) and East (Apollo 17, 11, and 16) divisions, reflecting site positioning with respect to the eastern PKT rift boundary (Fig. 1).



**Fig. 7.** Pareto (top) and PCA biplot (bottom) reflecting all unique sample modal analyses for the entire lunar nearside.

#### 4.3.2. Principal component analysis: west-east textures, ages, and major element oxides

Texture/age/ $\text{Al}_2\text{O}_3$  co-trends account for most of the variance in East/PC1 but virtually none for West/PC1 (Fig. 9).  $\text{Al}_2\text{O}_3/\text{Na}_2\text{O}$  define the bulk of variance for West/PC1 where FeO serves as the strongest anti-correlative basalt characteristic.

#### 4.3.3. Principal component analysis: nearside textures, ages, and major element oxides

Fig. 10 shows nearside treatment of nine lunar basalt characteristics produced the anti-correlative pairing  $\text{FeO}/\text{Al}_2\text{O}_3$  along PC1.  $\text{Na}_2\text{O}$  weakly co-trends with  $\text{Al}_2\text{O}_3$ , as does age to a lesser degree.  $\text{TiO}_2$  defines the bulk of the variance along PC2 and anti-correlative texture only weakly so.

#### 4.4. K-means cluster analysis

##### 4.4.1. K-means cluster analysis: nearside major element oxides

Silhouette evaluations of the standardized nearside basalt major element oxide dataset produced an optimal K-value of 2 (i.e. two clusters, hereafter referred to as C1 and C2) based on the K-means++ algorithm. The resulting cluster indices were then assigned to the original, non-standardized element oxide weight percents. Selected major element oxide K-means cluster plots are shown in Fig. 11.

##### 4.4.2. K-means cluster analysis: nearside mineral modes

Silhouette evaluations of the standardized nearside mineral mode dataset consistently favored K-values of 5. Cluster indices reflecting five clusters were then assigned to the original, non-standardized, mineral mode dataset (Fig. 12). Olivine proportions <1% either represent reported zero-values or reflect “no visible olivine” reports in published literature. In some cases, when olivine was not referred to in a petrological description, a value of zero was assumed.

#### 5. Discussion: combining PCA and KCA for lunar basalt characteristics

The results are discussed in a geological context suggested by the cross-connectedness and evolution of basalt characteristics through spatial-scale changes in Apollo site groupings (i.e. per-site to West-East to nearside). For initial exploratory interpretations, olivine, pyroxene, plagioclase, and ilmenite are considered to control major element distribution, with  $\text{MgO}$  and  $\text{FeO}$  largely present in olivine and pyroxene,  $\text{Al}_2\text{O}_3$  in plagioclase, and  $\text{TiO}_2$  in ilmenite. Here, we probe the underlying data patterns in an attempt to reveal aspects of geologic conditions potentially responsible for surface basalt patterns.

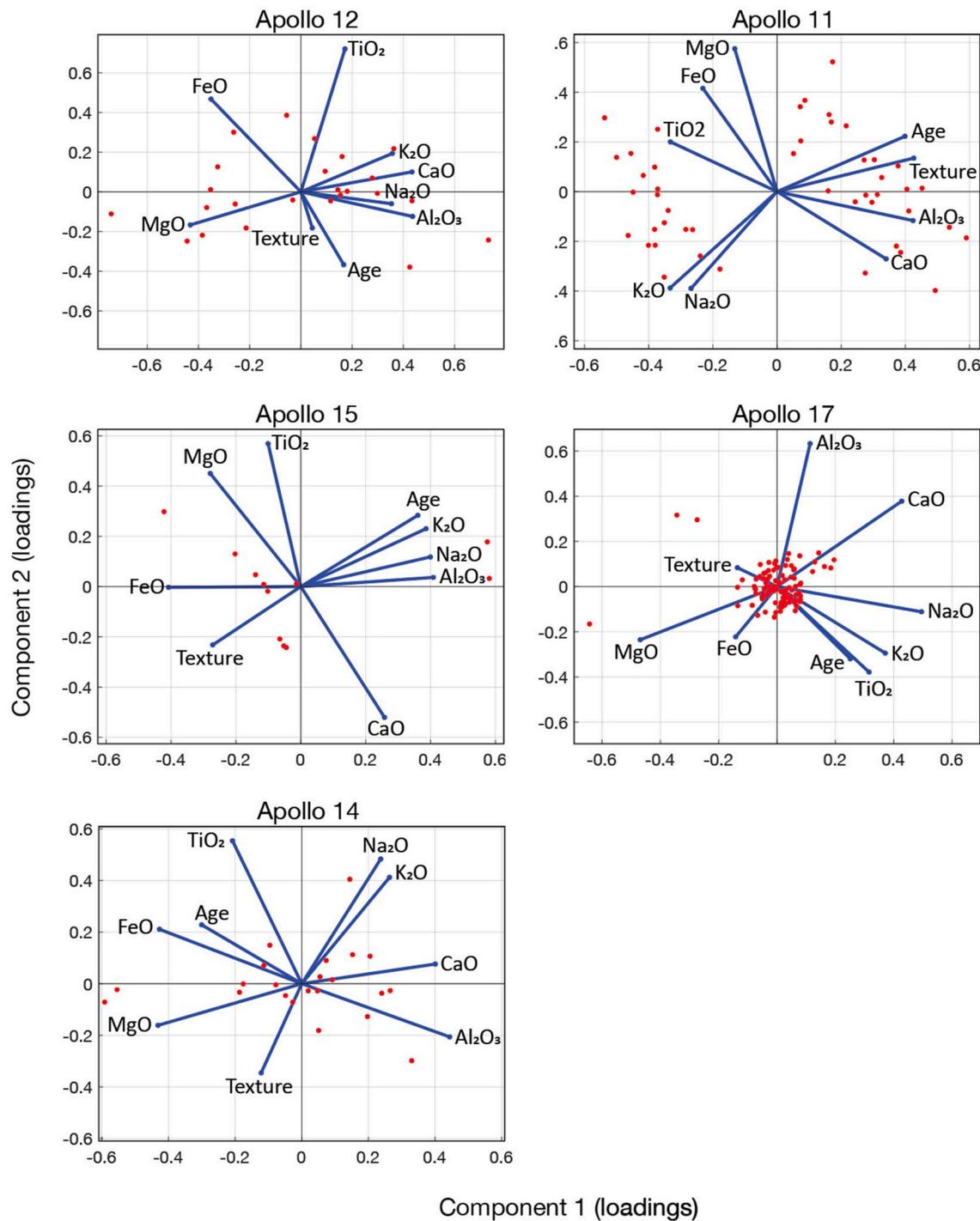
##### 5.1. Site-specific major element oxide and mineral mode patterns

Based on reported major element oxides on a per-site basis, sites A15, A12, and A14 produce the smallest diversity factors (2.3, 2.3, and 1.7, respectively; Table 1). These three sites also lie west of the eastern arm of the PKT Rift boundary and are more centrally located within the Procellarum KREEP Terrane (Fig. 1, red dashed line). The strongest generalized trends along PC1 are defined primarily by increases in  $\text{Al}_2\text{O}_3/\text{CaO}$  with anti-correlative  $\text{MgO}$  (A12) and/or anti-correlative  $\text{FeO}$  (Table 1). Increases in  $\text{Na}_2\text{O}$  are bound to  $\text{Al}_2\text{O}_3$  at A14 and A15 and to a lesser extent at A12.  $\text{TiO}_2$  variance is tied to PC2 with weak  $\text{K}_2\text{O}$  anti-correlations at all three sites. Variance among components for sites A17, A11, and A16 (the East group) are more flatly distributed ( $\text{df} = 4.7$ , 2.6, and 2.8, respectively). In general, these three sites suggest  $\text{Al}_2\text{O}_3$  diversity controls PC1 in addition to anti-correlative  $\text{FeO}$  and/or  $\text{MgO}$ .

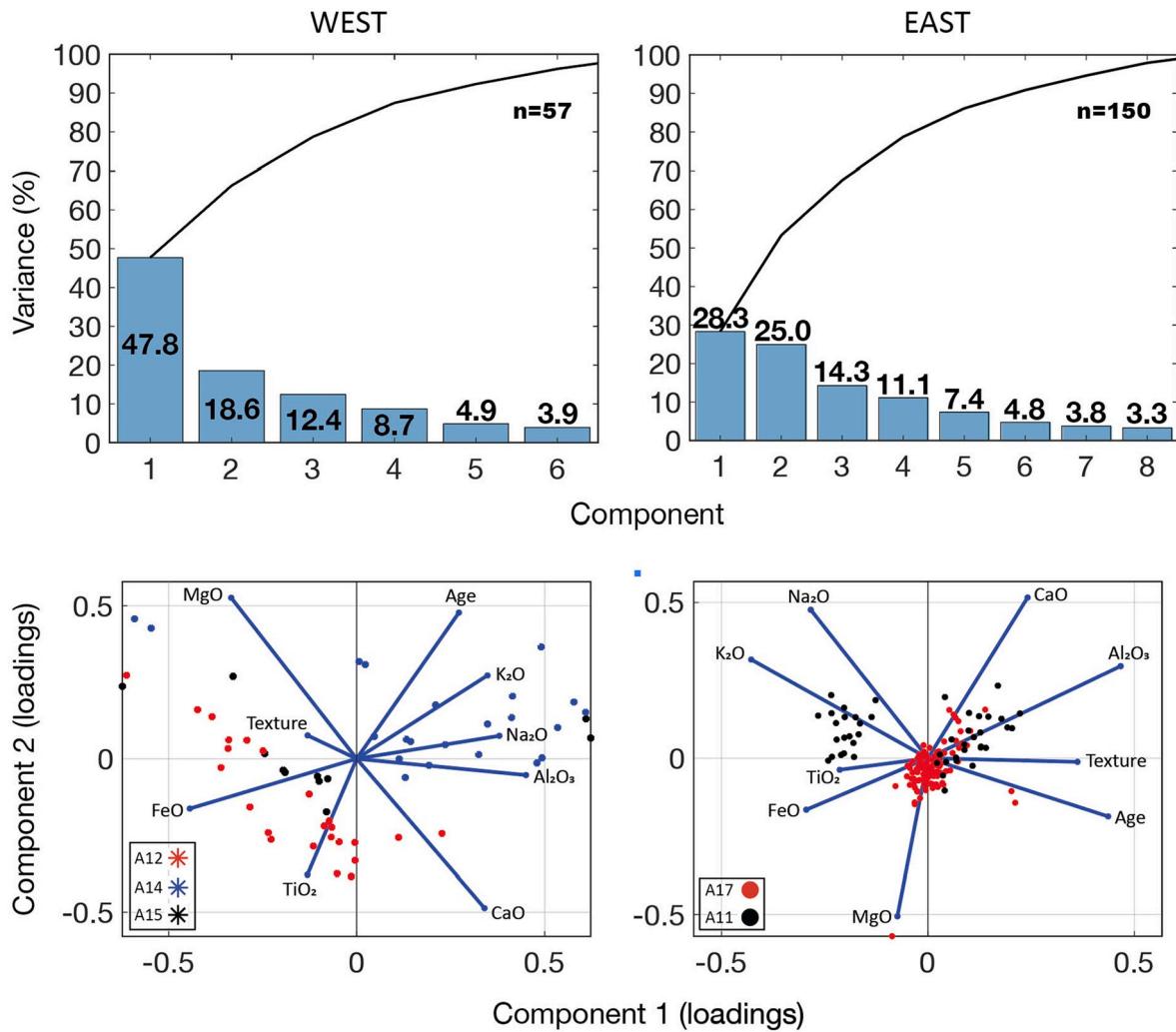
**Table 3**

Variances for principal components 1 and 2 for nearside textures, ages, and major element oxides, as well as strongest/weakest co-trends.

Mission	n	PC1 (%)	PC2 (%)	(PC1 + PC2) %	(PC1-PC2) %	df	Maximum co-trending characteristics (PC1)	Anti-trend along PC1
A11	46	49	17.6	66.6	31.4	2.1	$\text{Al}_2\text{O}_3/\text{Age}/\text{Texture}$	$\text{TiO}_2$ (weak)
A12	24	53.1	15.2	68.3	37.9	1.8	$\text{CaO}/\text{Al}_2\text{O}_3$ ; $\text{K}_2\text{O}/\text{Na}_2\text{O}$ (weak)	$\text{TiO}_2$ (weak)
A14	22	49.4	16.5	65.9	32.9	2	$\text{CaO}/\text{Al}_2\text{O}_3$ ; $\text{MgO}/\text{FeO}$	pairs anticorrelate
A15	11	63.8	20.3	84.1	43.5	1.9	$\text{Al}_2\text{O}_3/\text{Na}_2\text{O}/\text{K}_2\text{O}$ (weak)/ Age(weak)	$\text{FeO}$
A16	0	NA	NA	NA	NA	NA	NA	NA
A17	104	30.1	21.4	51.5	8.7	5.9	$\text{Na}_2\text{O}/\text{CaO}$ (weak)	$\text{MgO}$



**Fig. 8.** PCA biplots for each of the Apollo site basalt sample characteristics. Apollo 16 was omitted due to a lack of published data such that age, texture and major element oxide information was not available for a given sample from the database.



**Fig. 9.** Pareto plots (top) and PCA biplot (bottom) for texture, age, and major element oxide data based on West and East groups as a reflection of location with respect to the eastern border of the PKT rift boundary (Fig. 1). Apollo 16 data were omitted.

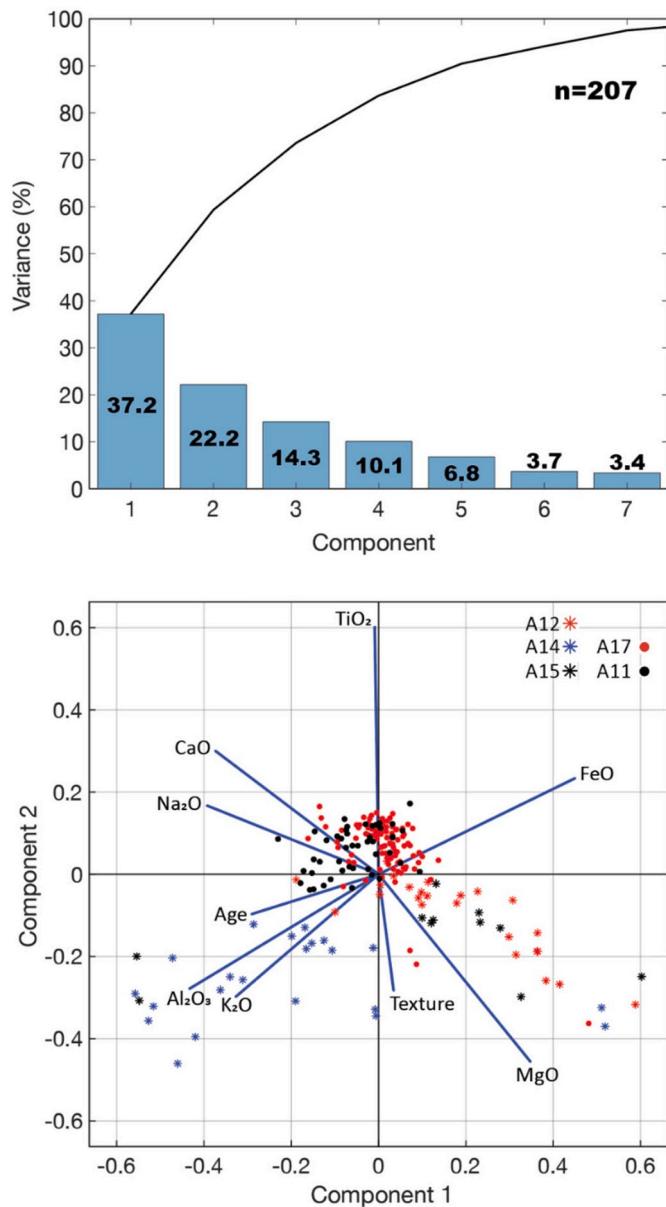
An early interpretation points to increases in plagioclase corresponding to decreases in pyroxenes at A15 and A14 while A12 suggests olivine diversity plays a more prominent role than pyroxene diversity. However, this conclusion cannot answer whether or not major element differences are directly linked to magma source characteristics and/or mantle mixing dynamics. Here, we expect that quantitative textural analysis (i.e. crystal size distributions) could directly help answer this question. At A11, TiO<sub>2</sub> also anti-correlates with Al<sub>2</sub>O<sub>3</sub>, suggesting lower ilmenite with increasing plagioclase volumes. This trend is also suggested in mineral mode biplots (Fig. 5).

Assuming plagioclase serves as the major host for CaO and Al<sub>2</sub>O<sub>3</sub> and that site-homogenization of samples are representative of their respective site population parameters, the general trend in Al<sub>2</sub>O<sub>3</sub> wt% diversity suggests relative per-site changes in plagioclase volumes for all sites. At A11, A12, A14, and A16 (southernmost sites), this Al<sub>2</sub>O<sub>3</sub> behavior strongly co-trends with increasing CaO wt%. The symmetry of loading lines about PC1 for the Al<sub>2</sub>O<sub>3</sub>/CaO pairing at A11, A12, A14, and A16 more strongly suggests plagioclase modal changes versus changes in CaO wt% strictly due to pyroxene speciation, as pyroxene is the other

major Ca-bearing phase. At sites A15 and A17 (northernmost sites), CaO wt% diversity only weakly correlates to that for Al<sub>2</sub>O<sub>3</sub>, and so CaO content is likely tied to either variable modal pyroxene or variety in pyroxene composition (e.g. augite versus pigeonite). As pigeonite is typically preserved in fast-cooled volcanic rocks and inverts to orthopyroxene during slow cooling, any potential pigeonite preservation would rest with a first-order approximation based on cooling-controlled texture. Based on our early analyses of CaO behavior, differences between the northernmost and southernmost sites suggest that the presence of pyroxenes may create reduced correlations between Al<sub>2</sub>O<sub>3</sub> and CaO.

Al<sub>2</sub>O<sub>3</sub> and anti-correlative FeO and MgO roughly suggests reduction in modal pyroxene and/or olivine (A11, A12, A14, and A16), while anti-correlative MgO defines A12 over anti-correlative FeO at A11. Given this anti-correlative behavior and the slight asymmetry of the CaO-Al<sub>2</sub>O<sub>3</sub> pairing at A11 compared to A12, A11 is estimated to contain relatively more pyroxene, and A12 likely contains a more variable (and perhaps primitive) olivine-pyroxene component.

Per-site mineral mode patterns (Figs. 6 and 7) suggest sites A11, A15,



**Fig. 10.** Pareto plot (top) and PCA biplot (bottom) for the lunar nearside for nine lunar basalt characteristics. No data for Apollo 16 (A16) samples were included due to a lack of relevant published data for single samples.

and A16 (and to a lesser extent A12 and A17) reflect greater ranges in modal plagioclase versus A12 and A14.  $\text{Al}_2\text{O}_3$  wt% biplot interpretations generally agree with modal results, but similar conclusions for modal plagioclase in A12 and A14 PCA biplots are obscured by complex loading values. The behavior of CaO as it is attributed to the interpreted presence of pyroxene at A15 and A17 agrees with the A15 modal biplot and to the A17 biplot to a lesser extent, in terms of PC2. Modal PCA biplot correlative mineral pairings along with major element oxide supporting behavior are summarized in Table 4.

## 5.2. West-east major element oxide and mineral mode patterns

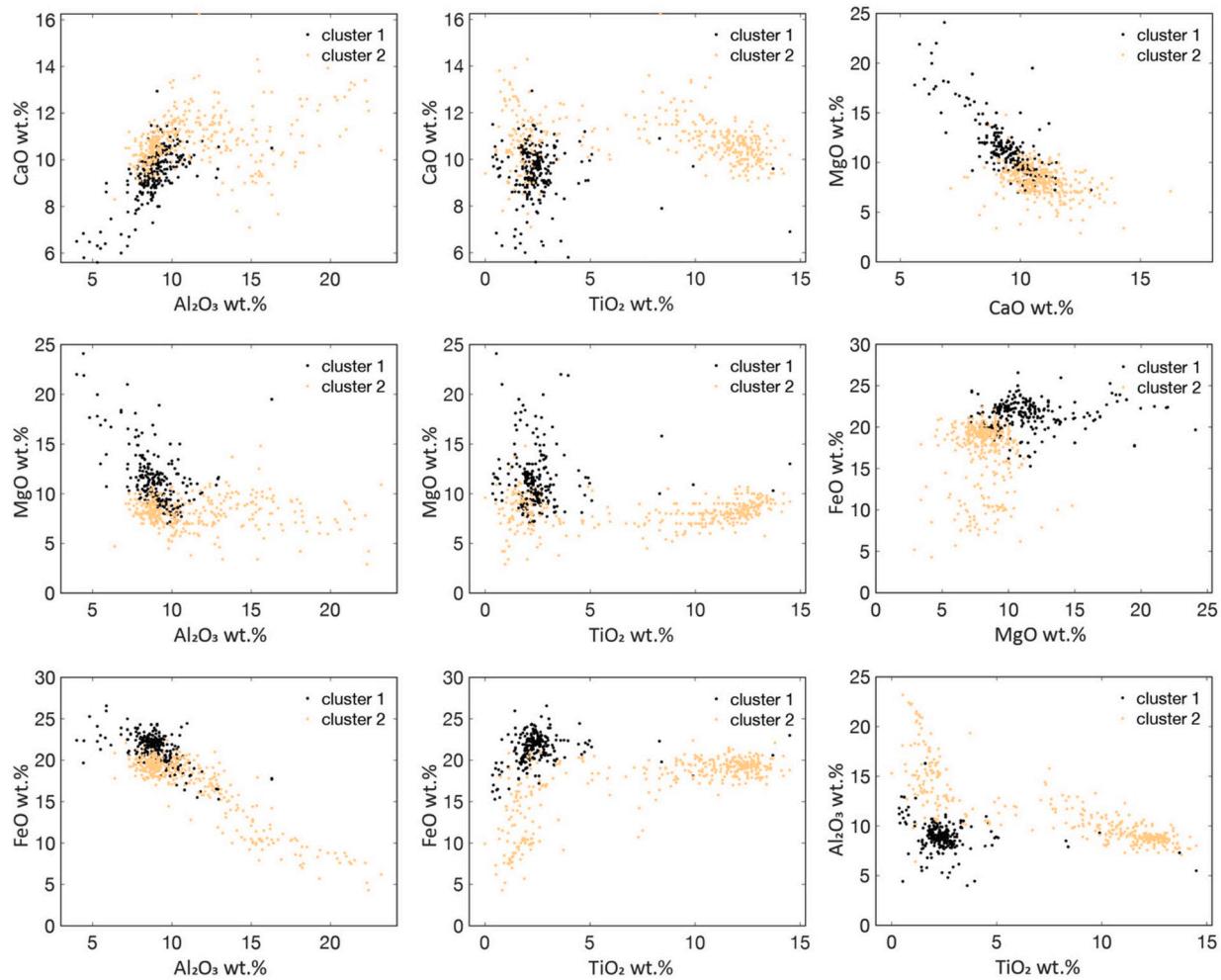
Considering only major element oxide patterns within West-East groups,  $\text{Al}_2\text{O}_3$  and CaO weakly co-trend (East) or weakly anti-correlate (West). FeO and MgO are non-correlative pairings for both groups but MgO is more anti-correlative with CaO for the West site than for the East site. Given CaO is more proximal in PCA biplot space to  $\text{Al}_2\text{O}_3$  for the East site, pyroxene content decreases with increasing plagioclase content as would be expected during progressive crystal fractionation of an ascending mafic magma, which is reinforced by PCA showing FeO/ $\text{Al}_2\text{O}_3$  anti-correlation. MgO/CaO anti-correlation in the West group implies potential opposing behavior of olivine and pyroxene abundances. An increase in plagioclase abundance (using  $\text{Al}_2\text{O}_3$  variance as a proxy) is also related to a decrease in ilmenite proportion (using  $\text{TiO}_2$  as the main proxy and FeO to a lesser degree). In the context of the West-East major element oxide interpretations above, East mineral mode PCA biplots support an increasing-plagioclase/decreasing-pyroxene pairing as well as decreasing-olivine/increasing-pyroxene pairing in the West. Ilmenite and plagioclase anti-correlate in the East and to a lesser extent in the West. The only correlative mineral pairing is suggested by olivine/ilmenite in the West. Diversity factors reflect increased variance flatness for the East group ( $df = 6.1$ ) versus the West group ( $df = 2.0$ ), as suggested by per-site analyses (Table 1). However,  $df_{EAST}$  reflects a greater value than not only individual values for sites A17, A11, and A16, but for A17 (the most diverse site). This pattern suggests that reduction in major element oxide diversity in the East group may be tied to the lack of strong co-trending major element pairings with  $\text{Al}_2\text{O}_3$ . In this sense,  $\text{Al}_2\text{O}_3$  diversity does not exhibit weight percent range increases that can be definitely tied to either plagioclase or pyroxene modal changes. Indirectly, this trend suggests that pyroxene may have a stronger presence in the East group; this is supported by higher pyroxene loading values for PC1 (Fig. 6).

## 5.3. Nearside major element oxide and mineral mode patterns

Comparisons of nearside groups with major element oxide non-correlative pairings of  $\text{Al}_2\text{O}_3/\text{CaO}/\text{MgO}/\text{FeO}$  (Fig. 4) suggest that olivine, pyroxene and plagioclase should have an equal presence in terms of modal range, and this general trend is indeed reflected in Fig. 7, despite a different (and reduced) basalt sample dataset. However, the range in FeO wt% is not reflected in the short PCA biplot line for pyroxene modal abundance. In this case, FeO distribution may instead reflect increased partitioning between ilmenite, pyroxene and olivine.  $\text{TiO}_2$  variance accounts for most of PC2 and ilmenite for most of PC1. Both nearside PCA biplots roughly suggest that increasing plagioclase ( $\text{Al}_2\text{O}_3$ ) corresponds with variably decreasing ilmenite ( $\text{TiO}_2$ ). The similarity of  $df$  values for East group major element oxides and mineral modes (6.1 and 6.3, respectively) may suggest that the major element data used for analysis here represent equivalent modal counterparts from the database.

## 5.4. Major element oxides, age, and texture patterns

Preliminary textural categorizations of lunar basalt samples resulted in indeterminate relationships with major element oxides and ages for sites A12 and A17. However, age and texture correlate at A11, non-correlate at A14, and weakly anti-correlate at A15. Texture in general anti-correlates with co-trending Na<sub>2</sub>O/K<sub>2</sub>O (i.e. increasing coarsening with decreasing alkalis). Pronounced trends for A11, A14, and A15 emphasizing age and/or texture may be summarized as follows:



**Fig. 11.** Two-dimensional cluster diagrams based on multidimensional (seven major element oxides) analyses through the K-means++ algorithm, considering  $K = 2$ . Major element oxide weight percent clustering reflects the lunar nearside as a single group, where  $n = 632$ .

A11: Co-trend groups (PC1): age/texture/ $\text{Al}_2\text{O}_3$ ; anti-correlations: none.

A11: Co-trend groups (PC2):  $\text{MgO}$ .

A14: Co-trend groups (PC1):  $\text{Al}_2\text{O}_3/\text{CaO}$  anticorrelates with  $\text{FeO}/\text{MgO}/\text{age}$ .

A14: Co-trend groups (PC2): texture anticorrelates with  $\text{Na}_2\text{O}/\text{K}_2\text{O}$ .

A15: Co-trend groups (PC1): age (weak)/ $\text{Al}_2\text{O}_3/\text{Na}_2\text{O}/\text{K}_2\text{O}$  anticorrelates  $\text{FeO}$ .

A15: Co-trend groups (PC2):  $\text{TiO}_2$  anticorrelates with  $\text{CaO}$ .

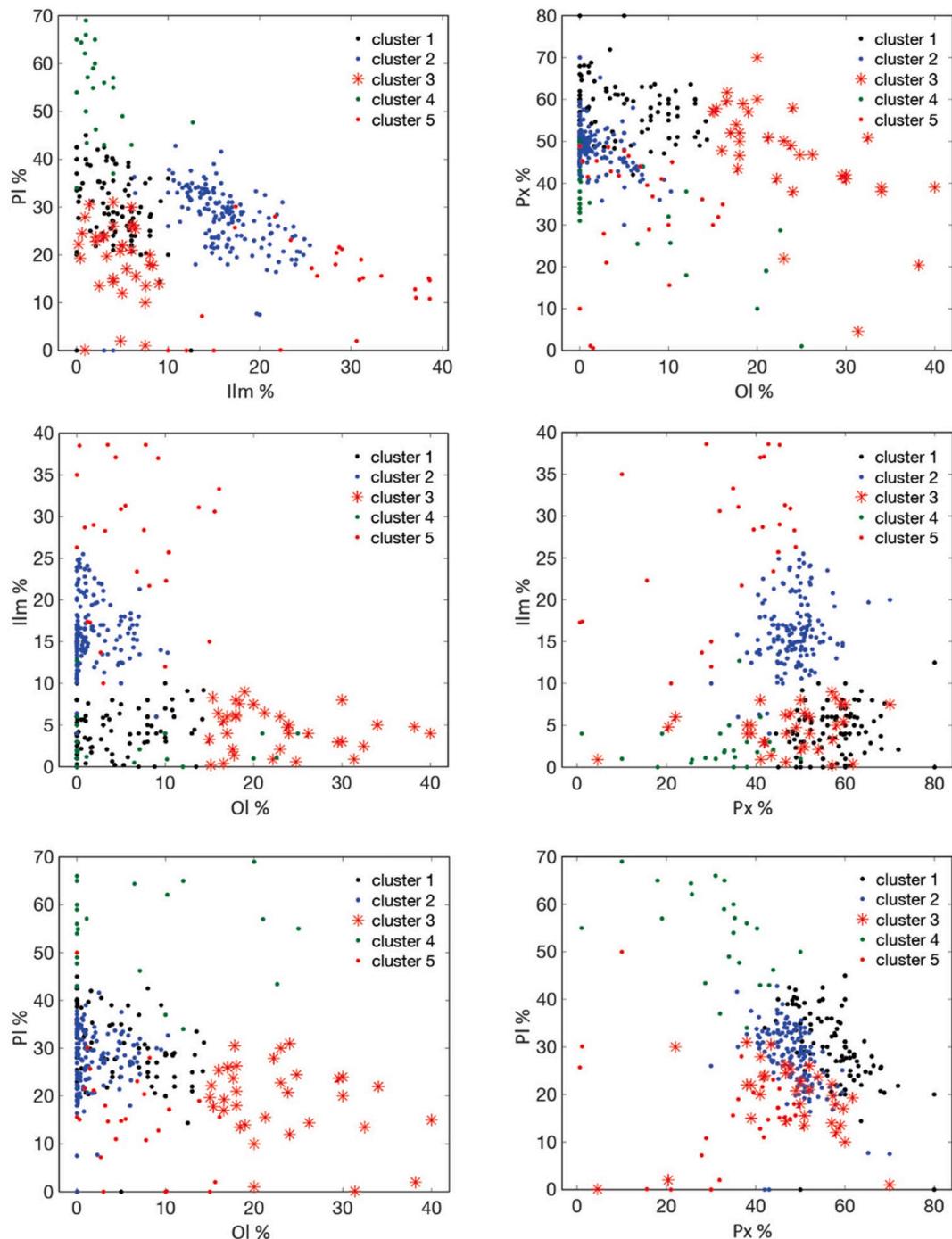
West-East groups emphasizing age and texture in combination with major element oxide data suggest that the strongest age and texture co-trends are only apparent for the East group with minor anti-correlation of  $\text{FeO}$  and  $\text{TiO}_2$ , possibly reflecting the degree of  $\text{TiO}_2/\text{ilmenite}$  availability in local upper mantle sources. The result reflects a coarsening-with-increasing-age relationship also defined by data from A11, a site that produces some of the oldest basalt ages ( $\sim 3.9$  Ga, along with A14). Coarsening-with-increasing-age trends may reflect reduced eruption speeds in early lunar history, and in combination with a minor anti-correlation in  $\text{TiO}_2$ , these trends suggest the possibility of a  $\text{TiO}_2$  conduit that may have brought up  $\text{TiO}_2$  from depth over time as reflected in finer-grained, younger basalts. Texture and age show no pronounced trends in the West group.

Nearside evaluations primarily show texture as anti-correlative with  $\text{TiO}_2$  along PC2,  $\text{FeO}$  as anti-correlative with  $\text{Al}_2\text{O}_3$  along PC1.

### 5.5. K-means clustering of nearside major element oxides and mineral modes

Fig. 11 represents two clusters (C1 and C2) showing the original nearside major element oxide dataset values from ApolloBasaltDB after cluster assignment of the associated standardized dataset. C1 shows tighter clustering over C2, the latter displaying some strongly bimodal distributions. C1  $\text{TiO}_2$  clustering correlates strongly to low-Ti classification (Supplementary Fig. 4) but C2  $\text{TiO}_2$  represents a bimodally distributed, wider range of wt% values (Supplementary Fig. 5). C2 distributions retain their dominant distributional character compared to pre-cluster assignment (Supplementary Fig. 6). The two clusters establish the following major element oxide patterns and distributions, reported in Table 5 and Supplementary Figs. 4 and 5.

C2 major element oxides are defined by relatively greater mean abundances of  $\text{Ti}$ ,  $\text{Al}$ , and  $\text{Ca}$ , whereas C1 displays greater mean abundances of  $\text{Fe}$ - and  $\text{Mg}$ -oxides. On a first-order basis, this suggests mineral-wise discretization into corresponding clusters of ilmenite/plagioclase (C2) and olivine/pyroxene (C1). Partial overlaps of oxide wt % ranges for unique sample analyses (e.g. 0–5 wt%  $\text{TiO}_2$  between C1 and C2) occur in part due to the inherent multidimensional spatial structure of the data. The mean composition of both clusters (Table 5), assuming <1% total  $\text{MnO} + \text{Cr}_2\text{O}_3$  and that  $\text{SiO}_2$  is the sole remaining major element oxide present, forces terrestrial classification equivalents of low-alkali basalt and picro-basalt (LeMaitre et al., 2002). Comparisons to mean compositions of ocean ridge basalts modeled by Gale et al.



**Fig. 12.** Mineral mode K-means cluster plots for olivine (OI), pyroxene (Px), plagioclase (PI), and ilmenite (Ilm). Two-dimensional plots represent the lunar nearside as a single sample set,  $n = 207$ , based on multidimensional analyses of the four mineral modes.

**Table 4**

Mineral mode correlative and anticorrelative pairings as well as major element oxide supporting behavior.

Mission	Correlative modal pairings	Noncorrelative modal pairings	Anticorrelative modal pairings	Supporting MEO behavior
A11	None	Ol/Pl, Ol/Ilm, Ilm/Px	Pl/Ilm, Pl/Px, Ol/Px	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) TiO <sub>2</sub> (Ilm)
A12	None	Ol/Ilm, Pl/Px, Ilm/Pl	Ol/Pl, Ol/Px	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) TiO <sub>2</sub> (Ilm)
A14	None	Pl/Px, Pl/Ilm, Ilm/Ol, Ol/Px	Pl/Ol, Ilm/Px	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) MgO/FeO (Px)
A15	Ilm/Ol (weak)	Ilm/Px, Pl/Px	Pl/Ilm	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) CaO (Px)
A16	Ilm/Px (strong)	Ol/Ilm, Ol/Pl, Ol/Px	Pl/Ilm, Pl/Px	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) MgO/FeO (Px)
A17	Ilm/Ol (weak)	Ilm/Px, Pl/Px	Pl/Ilm, Ol/Px	Al <sub>2</sub> O <sub>3</sub> /CaO (Pl) TiO <sub>2</sub> (Ilm) CaO (Px)

**Table 5**

K-Means++ clustering for major element oxide wt% as a global nearside dataset, K = 2, n = number of samples per cluster.

	Major element oxide						K = 2	Apollo mission sample frequency						
	Ti	Al	Fe	Mg	Ca	Na	K	Cluster	A11	A12	A14	A15	A16	A17
Mean	2.4	9.0	21.4	11.3	9.5	0.3	0.0	1 (n = 268)	1	42	5	203	2	15
S.D.	1.4	1.4	1.8	2.8	1.1	0.1	0.1							
Mean	8.0	11.3	17.0	8.1	10.8	0.5	0.2	2 (n = 365)	77	13	47	52	5	171
S.D.	4.7	3.4	3.9	1.5	1.1	0.2	0.3							

**Table 6**

K-Means++ clustering results for mineral modes as a global nearside dataset, K = 5, n = number of samples per cluster.

	Mineral mode				K = 5	Apollo mission sample frequency					
	Ol	Px	Pl	IIm		Cluster	A11	A12	A14	A15	A17
Mean	6.2	27.6	54.1	4.8	1 (n = 22)	0	1	8	4	8	1
S.D.	8.6	12.5	10.0	7.9							
Mean	7.1	38.2	12.7	27.6	2 (n = 23)	1	0	3	0	0	19
S.D.	5.0	9.2	7.6	8.5							
Mean	1.8	48.4	27.6	16.1	3 (n = 159)	71	0	4	1	1	82
S.D.	2.4	5.2	6.7	4.1							
Mean	4.7	57.2	29.9	4.3	4 (n = 109)	4	29	6	61	5	4
S.D.	4.6	7.0	8.0	2.8							
Mean	23.1	47.3	19.2	4.6	5 (n = 35)	0	21	2	10	2	0
S.D.	7.0	12.7	7.8	2.5							

(2013) place both clusters as comparatively Mg-Fe-Ti-enriched and Al-Ca deficient. Although major element oxide distributions for C2 (Supplementary Fig. 5) are less linearly cohesive than those for C1 (Supplementary Fig. 4), C2 patterns suggest that previously established lunar basalt classifications based on TiO<sub>2</sub> wt% are incapable of capturing degrees of multicollinearity. For example, Fig. 4 suggests the majority of nearside major element controls on variance are based primarily on co-trending Na<sub>2</sub>O/Al<sub>2</sub>O<sub>3</sub> with anti-correlated FeO, but secondarily on TiO<sub>2</sub>. As a complement to these trends, cluster diagrams can further disentangle particular oxide behavior among the discretized groups in the following way: the spatial arrangement of TiO<sub>2</sub> for C2 displays a greater range than that for C1, and it is this extended range that imposes a unique aspect to the TiO<sub>2</sub> data patterns for the two clusters, which may also be viewed as two basic lunar basalt types. A similar extended-range distribution occurs for Al<sub>2</sub>O<sub>3</sub> in C2. These extended domains of major element oxides, despite regions of restricted overlaps, represent potential alternative classification schemes for lunar basalts based on multi-variable patterns. The overlaps delineate shared wt% ranges and so alternative lunar basalt classifications should focus on regions lying outside these overlaps.

Table 6 contains Apollo site mission characteristics based on K = 5 for nearside mineral mode cluster analyses. Distributions are shown in Supplementary Fig. 7 and selected pair-wise relationships in Fig. 12. In multidimensional space, and as previously demonstrated with major element oxide clustering, the five mineral clusters contain partially overlapping ranges but maintain some degree of restricted space. C1 contains orthogonal trends in olivine distribution that parallel the x- and y-axis orientations. C2 and C5 olivine distributions are globular, C3 is roughly linear, and C4 is sub-globular to sub-linear. C3 and C4 are more

sharply defined by relatively less olivine (low-Ol groups; also see Table 6) compared to other clusters. Orthogonal olivine distributions of C1 reinforce the idea that discrete ranges of mineral modes by themselves may be insufficient to define meaningful basalt groupings. A11 and A17 data are strongly (but not entirely) confined to C3, but their similarities display as roughly linear trends in C3 pair-wise mineral mode distributions.

As shown through cluster diagrams, classification schemes based on generic criteria, such as sharply defined major element oxide or mineral mode cutoff values, may require revision in order to reflect the inherent, overlapping ranges in basalt characteristics across multiple basalt groupings. Rather than terrestrial classifications that focus on a singular aspect (e.g. geochemical MORB classifications or textural classifications such as vitreous or doleritic), a single multi-characteristic basalt grouping may ultimately hold more meaning than multiple discrete classifications for lunar basalts. In this manner, data typically viewed as outliers (e.g. impact contaminants) would not be removed from analysis but included for the purpose of characterizing degree of contamination.

In terms of teasing out geologic processes from our analyses, particularly with respect to sample locations about the PKT rift boundary, we assume that surface basalts provide representative ages of their sample region although older basalts may be buried by more recent flows. Mare basalt model ages trend older from west to east (Hiesinger et al., 2010) based on crater size-frequency distributions, and we note a coarsening-with-increasing-age trend at A11 (the southern end of Mare Tranquillitatis) that is coupled with increasing plagioclase but decreasing Fe/Mg diversity. This trend might suggest that older basalts with limited mafic components were sourced from more shallow mantle regions than western basalts. Anti-correlations of plagioclase and

ilmenite at these eastern sites may hint at reduced KREEP contamination or limited access to  $\text{TiO}_2$  at depth.

Data-imposed limitations have been recognized during this study. For example, not all studies report the same petrological characteristics for a given sample, with rare earth element contents often being more common than major element oxides. In addition, the routine examination of crystal size distributions (Donohue and Neal, 2015; Neal et al., 2015; Fagan and Neal, 2016) would bolster quantitative textural interpretations and provide insight into residence times and eruptive styles. Spectral and hyperspectral elemental and mineral data would further complement the database and provide for an expanded and more rigorous spatial analysis. The authors recognize that trace elements provide detailed insight into magma reservoir characteristics and fractionation processes and these will be included in future versions of ApolloBasaltDB as well as in future machine learning analyses. Although small sample counts can still provide robust results for multidimensional analysis (Shaikat et al., 2016), a critical issue was that a given sample was not analyzed for multiple characteristics, thus preventing the exact same dataset from being used for major element oxide, mineral mode, age, and texture analyses. Non-Gaussian distributions as shown in distribution plots do not render either PCA or KCA ineffective, but predictive modeling (e.g. establishing confidence intervals) would require either synthetic data or resampling methods to compensate. These issues, in addition to our findings, underscore the need for a comprehensive, multi-characteristic lunar database.

## 6. Conclusions

Traditional terrestrial basalt classifications such as the total alkali-silica (TAS) compositional scheme and the quartz-alkali feldspar-plagioclase-feldspar-feldspathoid (QAPF) mineralogical scheme may not meaningfully represent the petrological variation of lunar basalts given the Moon's LMO evolution and lack of Earth-style plate tectonics. Diversity factors based on principal components show that major element oxide content from Apollo landing sites 15, 12, and 14 encompass a greater variance in  $\text{Al}_2\text{O}_3$ . These sites lie west of the eastern leg of the PKT rift boundary. Northern sites (A15 and A17) show further diversity relating to pyroxene abundance. These variations in mineral modal ranges suggest that geochemical upper-mantle structure may vary spatially based on PKT rift boundary proximity, revealing the depth of cumulate mantle overturn dynamics. Although our analyses do not definitively answer questions regarding the full dynamics of the cumulate mantle overturn event, our findings do suggest that further investigation into spatially and temporally dependent differences in lunar basalt characteristics is warranted.

Lunar basalt characteristic patterns may be preliminarily related to proximity to the PKT rift boundary. East sites (A17, A11, and A16) all show strong anti-correlations in modal ilmenite and plagioclase, correlating to major element oxide PCA biplots, with A16 showing the weakest trend. Site A11 shows anti-correlation between modal plagioclase and modal ilmenite through opposing trends in  $\text{FeO}/\text{TiO}_2$  and  $\text{Al}_2\text{O}_3$ , suggesting that increased ilmenite abundance is at the expense of plagioclase abundance. This trend is reinforced by ilmenite/plagioclase anti-correlations in mineral mode PCA biplots. This trend is also observed for A17 samples, which also lies along the PKT rift boundary in proximity to site A11.

Nearside analyses of both major element oxides and mineral modes suggest plagioclase (and  $\text{Al}_2\text{O}_3$ ) diversity is a result of ilmenite ( $\text{TiO}_2$ ) diversity. Site A11 suggests coarsening texture with increasing age and increasing  $\text{Al}_2\text{O}_3$ , but age and textural data provide inconclusive patterns in general, and more data is required to extract meaningful interpretations. Major element oxide PCA biplots strongly reflect the behavior of mineral mode PCA biplots, despite the use of partially overlapping datasets. This suggests that major element oxide data alone may suffice in approximating mineral occurrence, creating the opportunity for meaningful, real-time data manipulation during remote

sensing missions.

## Data availability

The ApolloDB database is freely available in MS Excel format from: [https://people.mines.edu/kcone/wp-content/uploads/sites/105/2019/08/ApolloBasaltDB\\_V1\\_2019.xlsx](https://people.mines.edu/kcone/wp-content/uploads/sites/105/2019/08/ApolloBasaltDB_V1_2019.xlsx)

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.icarus.2020.113787>.

## References

- Albarède, F., 1995. *Introduction to Geochemical Modeling*. University Press, Cambridge.
- Allègre, C.L., Schiano, P., Lewin, E., 1995. Differences between oceanic basalts by multi-trace element ratio topology. *Earth Planet. Sci. Lett.* 129, 1–12. [https://doi.org/10.1016/0012-821X\(94\)00235-Q](https://doi.org/10.1016/0012-821X(94)00235-Q).
- Andre, C.G., El-Baz, F., 1982. Regional chemical setting of the Apollo 16 landing site and the importance of the Kant plateau. In: *Lunar and Planetary Science Conference Proceedings*, 12, pp. 767–779.
- Andrews-Hanna, J.C., Besserer, J., Head III, J.W., Howett, C.J., Kiefer, W.S., Lucey, P.J., McGovern, P.J., Melosh, H.J., Neumann, G.A., Phillips, R.J., Schenk, P.M., Smith, D.E., Solomon, S.C., Zuber, M.T., 2014. Structure and evolution of the lunar Procellarum region as revealed by GRAIL gravity data. *Nature* 514 (7520), 68–72. <https://doi.org/10.1038/nature13697>.
- Arthur, D., Vassilvitskii, S., 2006. K-means++: the advantages of careful seeding. In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics Philadelphia, pp. 1027–1035.
- Beatty, D.W., Albee, A.L., 1978. Comparative petrology and possible genetic relations among the Apollo 11 basalts. In: *Lunar and Planetary Science Conference Proceedings*, 9, pp. 359–463.
- Bellucci, J.J., Nemchin, A.A., Grange, M., Robinson, K.L., Collins, G., Whitehouse, M.J., Snape, J.F., Norman, M.D., Kring, D.A., 2019. Terrestrial-like zircon in a clast from an Apollo 14 breccia. *Earth Planet. Sci. Lett.* 510, 173–185. <https://doi.org/10.1016/j.epsl.2019.01.010>.
- Boukaré, C.-E., Parmentier, E.M., Parman, S.W., 2018. Timing of mantle overturn during magma ocean solidification. *Earth Planet. Sci. Lett.* 491, 216–225. <https://doi.org/10.1016/j.epsl.2018.03.037>.
- Bradley, P.S., Fayyad, U.M., 1998. Refining initial points for K-means clustering. In: *International Conference on Machine Learning (ICML98)*, pp. 91–99.
- Brown, G.M., Peckett, A., Emeleus, C.H., Phillips, R., Pinsent, R.H., 1975. Petrology and mineralogy of Apollo 17 mare basalts. In: *Lunar and Planetary Science Conference Proceedings*, 6, pp. 1–13.
- Butler, J.C., 1976. Principal components analysis using the hypothetical closed array. *Math. Geol.* 8, 25–36.
- Condie, K.C., Aster, R.C., van Hunen, J., 2016. A great thermal divergence in the mantle beginning 2.5 Ga: geochemical constraints from greenstone basalts and komatiites. *Geosci. Front.* 7, 543–553. <https://doi.org/10.1016/j.gsf.2016.01.006>.
- Delano, J.W., 1979. Apollo 15 green glass-chemistry and possible origin. In: *Lunar and Planetary Science Conference Proceedings*, 10, pp. 275–300.
- Dickinson, T., Taylor, G.J., Keil, K., Schmitt, R.A., Hughes, S.S., Smith, M.R., 1985. Apollo 14 aluminous mare basalts and their possible relationship to KREEP. *Journal of Geophysical Research: Solid Earth* (S02), 90. <https://doi.org/10.1029/JB090iS02p0C365>.
- Donohue, P.H., Neal, C.R., 2015. Quantitative textural analysis of ilmenite in Apollo 17 high-titanium mare basalts. *Geochim. Cosmochim. Acta* 149, 115–130. <https://doi.org/10.1016/j.gca.2014.11.002>.
- Dudzinski, M.L., Norris, J.M., Chmura, J.T., Edwards, C.B.H., 1975. Repeatability of principal components in samples: normal and non-normal data sets compared. *Multivar. Behav. Res.* 10 (1), 109–117. [https://doi.org/10.1207/s15327906mbr1001\\_8](https://doi.org/10.1207/s15327906mbr1001_8).
- Dymek, R.F., 1986. Characterization of the Apollo 15 feldspathic basalt suite. In: *Geology and Petrology of the Apollo 15 Landing Site* (52 pp).
- Elardo, S.M., Draper, D.S., Shearer, C.K., 2011. Lunar Magma Ocean crystallization revisited: bulk composition, early cumulate mineralogy, and the source regions of the highlands Mg-suite. *Geochim. Cosmochim. Acta* 75, 3024–3045. <https://doi.org/10.1016/j.gca.2011.02.033>.
- Elkins-Tanton, L.T., Van Orman, J.A., Hager, B.H., Grove, T.L., 2002. Re-examination of the lunar magma ocean cumulate overturn hypothesis: melting or mixing is required. *Earth Planet. Sci. Lett.* 196, 239–249. [https://doi.org/10.1016/S0012-821X\(01\)00613-6](https://doi.org/10.1016/S0012-821X(01)00613-6).
- Elkins-Tanton, L.T., Burgess, S., Yin, Q.Z., 2011. The lunar magma ocean: reconciling the solidification process with lunar petrology and geochronology. *Earth Planet. Sci. Lett.* 304, 326–336. <https://doi.org/10.1016/j.epsl.2011.02.004>.
- Fagan, A.L., Neal, C.R., 2016. A new lunar high-Ti basalt type defined from clasts in Apollo 16 breccia 60639. *Geochim. Cosmochim. Acta* 173, 352–372.
- Furnes, H., Dilek, Y., De Wit, M., 2014. Precambrian greenstone sequences represent different ophiolite types. *Gondwana Res.* 27, 649–685. <https://doi.org/10.1016/j.gr.2013.06.004>.

- Gale, A., Dalton, C.A., Langmuir, C.H., Su, Y., Schilling, J., 2013. The mean composition of ocean ridge basalts. *Geochim. Geophys. Geosyst.* 14, 489–518.
- Gancarz, A.J., A'ee, A.L., Chodos, A.A., 1971. Petrologic and mineralogic investigation of some crystalline rocks returned by the Apollo 14 mission. *Earth Planet. Sci. Lett.* 12 (1), 1–18. [https://doi.org/10.1016/0012-821X\(71\)90050-1](https://doi.org/10.1016/0012-821X(71)90050-1).
- Ganne, J., Feng, X., 2017. Primary magmas and mantle temperatures through time. *Geochem. Geophys. Geosyst.* 18, 872–888. <https://doi.org/10.1002/2016GC006787>.
- Haase, I., Oberst, J., Scholten, F., Wählisch, M., Gläser, P., Karachevtseva, I., Robinson, M.S., 2012. Mapping the Apollo 17 landing site area based on lunar reconnaissance orbiter camera images and Apollo surface photography. *Journal of Geophysical Research: Planets* 117 (E12). <https://doi.org/10.1029/2011JE003908>.
- Hazen, R.M., Grew, E.S., Downs, R.T., Golden, J., Hystad, G., 2015. Mineral ecology: chance and necessity in the mineral diversity of terrestrial planets. *Can. Mineral.* 53, 295–323. <https://doi.org/10.3749/canmin.1400086>.
- Heiken, G., Vaniman, D., French, B.M., 1991. *Lunar Sourcebook: A User's Guide to the Moon*. University Press, Cambridge.
- Hiesinger, H., & Head III, J.W. (2006). New views of lunar geoscience: an introduction and overview. *Rev. Mineral. Geochem.*, 60(1), 1–81.
- Hiesinger, H., Head III, J.W., Wolf, U., Jaumann, R., Neukum, G., 2010. Ages and stratigraphy of lunar mare basalts in Mare Frigoris and other nearside maria based on crater size-frequency distribution measurements. *J. Geophys. Res.* 115, 22. <https://doi.org/10.1029/2009JE003380>.
- Hodges, C.A., Muehlberger, W.R., Ulrich, G.E., 1973. Geologic setting of Apollo 16. In: *Lunar and Planetary Science Conference Proceedings*, 4, p. 1.
- Iwamori, H., Yoshida, K., Nakamura, H., Kuwatani, T., Hamada, M., Haraguchi, S., Ueki, K., 2017. Classification of geochemical data based on multivariate statistical analyses: complementary roles of cluster, principal component, and independent component analyses. *Geochim. Geophys. Geosyst.* 18, 995–1012. <https://doi.org/10.1002/2016GC006663>.
- Jerde, E.A., Snyder, G.A., Taylor, L.A., Liu, Y.G., Schmitt, R.A., 1994. The origin and evolution of lunar high-Ti basalts: periodic melting of a single source at Mare Tranquillitatis. *Geochim. Cosmochim. Acta* 58 (1), 515–527. [https://doi.org/10.1016/0016-7037\(94\)90480-4](https://doi.org/10.1016/0016-7037(94)90480-4).
- Jolliff, B.L., Gillis, J.J., Haskin, L.A., Korotev, R.L., Wieczorek, M.A., 2000. Major lunar crustal terranes: surface expressions and crust-mantle origins. *Journal of Geophysical Research: Planets* 105 (E2), 4197–4216. <https://doi.org/10.1029/1999JE001103>.
- Jolliff, B.L., Wieczorek, M.A., Shearer, C.K., Neal, C.R. (Eds.), 2006. *New Views of the Moon*. Reviews in Mineralogy & Geochemistry, 60. Mineralogical Society of America, Chantilly, VA (721 pp).
- Keller, C.B., Schoene, B., 2012. Statistical geochemistry reveals disruption in secular lithospheric evolution about 2.5 Gyr ago. *Nature* 485, 490–493. <https://doi.org/10.1038/nature11024>.
- Khan, A., Mosegaard, K., Rasmussen, K.L., 2000. A new seismic velocity model for the moon from a Monte Carlo inversion of the Apollo lunar seismic data. *Geophys. Res. Lett.* 27, 1591–1598. <https://doi.org/10.1029/1999GL008452>.
- LeMaitre, R.W., 1982. Numerical petrology. Statistical interpretation of geochemical data. In: *Developments in Petrology*, 8. Elsevier, Amsterdam.
- LeMaitre, R.W., Streckeisen, A., Zanettin, B., Le Bas, M.J., Bonin, B., Bateman, P., 2002. *Igneous Rocks: A Classification and Glossary of Terms*. University Press, Cambridge.
- Li, H., Zhang, N., Wu, B., Dygert, N.J., Huang, J., Parmentier, E.M., 2019. Lunar cumulate mantle overturn: a model constrained by Ilmenite rheology. *Journal of Geophysical Research: Planets* 124, 1357–1378. <https://doi.org/10.1029/2018JE005905>.
- Lindstrom, M.M., Haskin, L.A., 1978. Causes of compositional variations within mare basalts suites. In: *Lunar and Planetary Science Conference*, 1, pp. 465–486.
- Lognonné, P., Gagnepain-Beyneix, J., Chenet, H., 2003. A new seismic model for the moon: implications for structure, thermal evolution and formation of the moon. *Earth Planet. Sci. Lett.* 211, 27–44. [https://doi.org/10.1016/S0012-821X\(03\)00172-9](https://doi.org/10.1016/S0012-821X(03)00172-9).
- Longhi, J., 1992. Experimental petrology and petrogenesis of mare volcanics. *Geochim. Cosmochim. Acta* 56, 2235–2251. [https://doi.org/10.1016/0016-7037\(92\)90186-M](https://doi.org/10.1016/0016-7037(92)90186-M).
- Longhi, J., Walker, D., Grove, T.L., Stolper, E., Hays, J.F., 1974. The petrology of the Apollo 17 mare basalts. In: *Lunar and Planetary Science Conference Proceedings*, 5, pp. 447–469.
- MacQueen, J.B., 1967. Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, pp. 281–297.
- McGetchin, T.R., Settle, M., Head, J.W., 1973. Radial thickness variation in impact crater ejecta: implications for lunar basin deposits. *Earth Planet. Sci. Lett.* 20 (2), 226–236. [https://doi.org/10.1016/0012-821X\(73\)90162-3](https://doi.org/10.1016/0012-821X(73)90162-3).
- Morrison, S.M., Liu, C., Eleish, A., Prabhu, A., Li, C., Ralph, J., Downs, R.T., Golden, J.J., Fox, P., Hummer, D.R., Meyer, M.B., Hazen, R.M., 2017. Network analysis of mineralogical systems. *Am. Mineral.* 102 (8), 1588–1596. <https://doi.org/10.2138/am-2017-6104CCBYNCND>.
- Neal, C.R., Hacker, M.D., Snyder, G.A., Taylor, L.A., Liu, Y.G., Schmitt, R.A., 1994. Basalt generation at the Apollo 12 site, part 1: new data, classification, and re-evaluation. *Meteoritics* 29 (3), 334–348. <https://doi.org/10.1111/j.1945-5100.1994.tb00597.x>.
- Neal, C.R., Donohue, P.H., Fagan, A.L., O'Sullivan, K.M., Oshrin, J., Hui, H., 2015. Distinguishing between pristine mare basalts and lunar impact melts: a non-destructive method using quantitative petrography. *Geochim. Cosmochim. Acta* 148, 62–80. <https://doi.org/10.1111/j.1945-5100.1994.tb00597.x>.
- Nimmo, F., McKenzie, D., 1998. Volcanism and tectonics on Venus. *Annu. Rev. Earth Planet. Sci.* 26 (1), 23–51. <https://doi.org/10.1146/annurev.earth.26.1.23>.
- Palin, R.M., Dyck, B., 2018. Metamorphic consequences of secular changes in oceanic crust composition and implications for uniformitarianism in the geological record. *Geosci. Front.* 9 (4), 1009–1019. <https://doi.org/10.1016/j.gsf.2018.04.004>.
- Papike, J.J., Hodges, F.N., Bence, A.E., Cameron, M., Rhodes, J.M., 1976. Mare basalts: crystal chemistry, mineralogy, and petrology. *Rev. Geophys.* 14, 475–540.
- Pieters, C.M., Wilhelms, D.E., 1985. Origin of olivine at Copernicus. *Journal of Geophysical Research: Solid Earth* 90 (S02), C415–C420. <https://doi.org/10.1029/JB090iS02p0C415>.
- Pieters, C.M., McCord, T.B., Zisk, S., Adams, J.B., 1973. Lunar black spots and nature of the Apollo 17 landing area. *J. Geophys. Res.* 78 (26), 5867–5875. <https://doi.org/10.1029/JB078i026p05867>.
- Rajmon, D., Reid, A.M., Spudis, P., 1999. *Geology and stratigraphy of Mare Tranquillitatis*. *Meteoritics Planet Sci.* 34 (4), A96.
- Rhodes, J.M., Hubbard, N.J., 1973. Chemistry, classification, and petrogenesis of Apollo 15 mare basalts. In: *Lunar and Planetary Science Conference Proceedings*, 4, p. 1127.
- Rhodes, J.M., Hubbard, N.J., Wiesmann, H., Rodgers, K.V., Brannon, J.C., Bansal, B.M., 1976. Chemistry, classification, and petrogenesis of Apollo 17 mare basalts. In: *Lunar and Planetary Science Conference Proceedings*, 7, pp. 1467–1489.
- Ridley, W.I., Reid, A.M., Warner, J.L., Brown, R.W., 1973. Apollo 15 green glasses. *Phys. Earth Planet. Inter.* 7 (2), 133–136. [https://doi.org/10.1016/0031-9201\(73\)90002-2](https://doi.org/10.1016/0031-9201(73)90002-2).
- Ryder, G., Schuraytz, B.C., 2001. Chemical variations of the large Apollo 15 olivine-normative mare basalt rock samples. *J. Geophys. Res. Planets* 106, 1435–1451. <https://doi.org/10.1029/2000JE001368>.
- Shaukat, S.S., Rao, T.A., Khan, M.A., 2016. Impact of sample size on principal component analysis ordination of an environmental data set: effects on eigenstructure. *Int. J. Ecol. Problems Biosphera* 35, 173–190. <https://doi.org/10.1515/eko-2016-0014>.
- Shervais, J.W., Taylor, L.A., Lindstrom, M.M., 1985. Apollo 14 mare basalts: petrology and geochemistry of clasts from consortium breccia 14321. *Journal of Geophysical Research: Solid Earth* 90 (S02), C375–C395. <https://doi.org/10.1029/JB090iS02p0C375>.
- Shih, C.Y., Nyquist, L.E., Bansal, B.M., Wiesmann, H., 1992. Rb-Sr and Sm-Nd chronology of an Apollo 17 KREEP basalt. *Earth Planet. Sci. Lett.* 108 (4), 203–215. [https://doi.org/10.1016/0012-821X\(92\)90023-O](https://doi.org/10.1016/0012-821X(92)90023-O).
- Spudis, P.D., 1984. Apollo 16 site geology and impact melts: implications for the geologic history of the lunar highlands. *Journal of Geophysical Research: Solid Earth* 89 (S01), C95–C107. <https://doi.org/10.1029/JB089iS01p00C95>.
- Spudis, P.D., Ryder, G., 1985. Geology and petrology of the Apollo 15 landing site: past, present, and future understanding. *Eos, Transac. Am. Geophys. Union* 66 (43), 721–726. <https://doi.org/10.1029/EO066i043p00721>.
- Steele, I.M., Smith, J.V., 1971. Mineralogy of Apollo 15415 “genesis rock”: source of anorthosites on moon. *Nature* 234 (5325), 138–142. <https://doi.org/10.1038/234138b0>.
- Sutton, R.L., Schaber, G.G., 1971. Lunar locations and orientations of rock samples from Apollo missions 11 and 12. In: *Lunar and Planetary Science Conference Proceedings*, 2, p. 17.
- Swann, G.A., Bailey, N.G., Batson, R.M., Eggleton, R.E., Hait, M.H., Holt, H.E., Schafer, J.P., 1971. Preliminary geologic investigations of the Apollo 14 landing site. In: *Apollo 14*, 39–85.
- Taylor, S.R., Jakes, P., 1974. The geochemical evolution of the Moon. In: *Lunar and Planetary Science Conference Proceedings*, 5, pp. 1287–1305.
- Thomas, R.J., Rothery, D.A., Conway, S.J., Anand, M., 2014. Long-lived explosive volcanism on mercury. *Geophys. Res. Lett.* 41 (17), 6084–6092. <https://doi.org/10.1002/2014GL061224>.
- Ueki, K., Iwamori, H., 2017. Geochemical differentiation processes for arc magma of the Sengen volcanic cluster, northeastern Japan, constrained from principal component analysis. *Lithos* 290–291, 60–75. <https://doi.org/10.1016/j.lithos.2017.08.001>.
- Wade, J., Dyck, B., Palin, R.M., Moore, J.D., Smye, A.J., 2017. The divergent fates of primitive hydroscopic water on Earth and Mars. *Nature* 552 (7685), 391–395. <https://doi.org/10.1038/nature25031>.
- Wänke, H., Baddenhausen, H., Dreibus, G., Jagoutz, E., Kruse, H., Palme, H., Teschke, F., 1973. Multi-element analyses of Apollo 15, 16, and 17 samples and the bulk composition of the moon. In: *Lunar and Planetary Science Conference Proceedings*, 4, p. 1461.
- Wieczorek, M.A., Neumann, G.A., Nimmo, F., Kiefer, W.S., Taylor, G.J., Melosh, J., Phillips, R.J., Solomon, S.C., Andrews-Hanna, J.C., Asmar, S.W., Konopliv, A.S., Lemoine, F.G., Smith, D.E., Watkins, M.M., Williams, J.G., Zuber, M.T., 2013. The crust of the moon as seen by GRAIL. *Science* 339, 671–675. <https://doi.org/10.1126/science.1231530>.
- Wilhelms, D.E., 1984. The moon. In: Carr, M.H. (Ed.), *The Geology of Terrestrial Planets*, 469. NASA Special Publication, Washington, D.C, pp. 107–206.
- Wilhelms, D.E., 1987. The geologic history of the moon. *U. S. Geol. Surv. Prof. Pap.* 1348. <https://doi.org/10.3133/pp1348>.
- Wilhelms, D.E., McCauley, J.F., 1971. Geologic Map of the Near Side of the Moon. US Geological Survey, Washington, DC. <https://doi.org/10.3133/i703>.
- Wood, J.A., Marvin, U.B., Reid Jr, J.B., Taylor, G.J., Bower, J.F., Powell, B.N., & Dickey Jr., J. S. (1971). Mineralogy and petrology of the Apollo 12 lunar sample. *SAO Special Rep.*, 333.
- Zussman, J., 1972. The mineralogy, petrology and geochemistry of lunar samples – a review. *The Moon* 5 (3–4), 422–435.