

## Article

# A Study of Lunar Regolith Obtained during the Apollo and Luna Space Programs Based on Principal Component Analysis

Jacek Katzer <sup>1</sup> , Janusz Kobaka <sup>1,\*</sup> and Karol Seweryn <sup>2</sup>

<sup>1</sup> Faculty of Geoengineering, Institute of Geodesy and Civil Engineering, University of Warmia and Mazury in Olsztyn, 10-724 Olsztyn, Poland; jacek.katzer@uwm.edu.pl

<sup>2</sup> Space Research Centre PAS (CBK PAN), 00-716 Warszawa, Poland; kseweryn@cbk.waw.pl

\* Correspondence: janusz.kobaka@uwm.edu.pl

**Abstract:** In this study, a modern principal component analysis (PCA) of the chemical properties of lunar soils was conducted. American and Soviet results acquired during the Apollo and Luna missions, respectively, were analyzed and compared. The chemical composition of the lunar soil was the focus of our analysis, the main aim of which was to assess any possible differences between the results provided by the missions in question. The results were visualized in two- and three-dimensional spaces. The use of PCA virtual variables enabled the chemical composition of the lunar soil to be fully visualized—something impossible to achieve using traditional techniques—and key similarities and differences among the properties of the lunar soil samples were determined. The sources of any differences were then conceptualized. The work reported in this paper offers new directions for future studies, especially research into the design of new lunar soil simulants for lunar construction and civil engineering programs.

**Keywords:** lunar regolith; Apollo space program; Luna space program; PCA analysis



**Citation:** Katzer, J.; Kobaka, J.; Seweryn, K. A Study of Lunar Regolith Obtained during the Apollo and Luna Space Programs Based on Principal Component Analysis. *Aerospace* **2024**, *11*, 348. <https://doi.org/10.3390/aerospace11050348>

Academic Editor: Paolo Tortora

Received: 29 January 2024

Revised: 13 April 2024

Accepted: 23 April 2024

Published: 27 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, major space agencies such as NASA (Artemis program), ESA (Aragonaut mission), CNSA (Chang'e missions), Roskosmos (Luna missions), and ISRO (Chandrayaan missions) have focused on the Moon as a target for space exploration. In addition, many other space agencies are planning future activities associated with lunar missions, and these will likely result in a greatly improved understanding of the mechanical and geotechnical properties of the lunar regolith.

The onset of lunar exploration using space probes was closely related to the competition between the USA and the Soviet Union in the second half of the 20th century. During the American Apollo (1969–1972) and Soviet Luna (1970–1976) missions, samples of lunar soil were collected and transported back to Earth. In addition, some properties of the lunar regolith were studied in situ on the Moon [1].

Altogether, the Apollo and Luna missions resulted in the transportation of lunar rock to Earth in quantities of 383.4 kg and 0.3 kg, respectively (see Table 1).

In subsequent years, some of these rock samples were thoroughly tested, while others were kept sealed for possible future experiments. As a result of the numerous experiments that were conducted using the original samples, a great deal of knowledge was acquired about this lunar soil.

Using modern computers and software, it is possible to execute complex statistical analyses that were almost impossible when the above-listed samples were originally obtained. One of these analytical methods is principal component analysis (PCA). This was invented by Karl Pearson in 1901 [2] and later developed by Harold Hotelling in the 1930s [3,4].

**Table 1.** Lunar soil samples brought back to Earth by Apollo and Luna missions.

Mission (Start Date)	Samples Returned (kg)	Route Travelled on the Moon (km)	Type of Flight	Landing Coordinates	Landing Place Description
Apollo 11 (16 July 1969)	21.7	0.25	manned	0.67408° N 23.47297° E	mare
Apollo 12 (14 November 1969)	34.3	1.5	manned	3.01239° S 23.42157° W	mare
Apollo 14 (31 January 1971)	44.8	3.3	manned	3.64530° S 17.47136° W	highlands/mare
Apollo 15 (26 July 1971)	76.8	27.9	manned	26.1322° N 3.6339° E	mare
Apollo 16 (16 April 1972)	95.8	27	manned	8.97301° S 15.50019° E	highlands
Apollo 17 (07 December 1972)	110.0	30	manned	20.1908° N 30.7717° E	mare
Luna 16 (12 September 1970)	0.101	n/a	unmanned	0.5137° S 56.3638° E	mare
Luna 20 (14 February 1972)	0.030	n/a	unmanned	3.5333° N 56.5500° E	highlands
Luna 24 (9 August 1976)	0.170	n/a	unmanned	12.7145° N 62.2097° E	mare

The origins of PCA derive from the fact that the handling of large datasets comprising multiple objects with numerous variables and parameters poses a significant challenge for humans. Data presented in complex tables and multidimensional matrices are particularly difficult for human analysis. PCA computational algorithms are widely acknowledged as successful applications of linear algebra, enabling effective data reduction and the extraction of latent information from raw datasets. PCA has proven to be a very versatile statistical tool in fields such as biology [5], medicine [6], pharmacology, and climatology [7]. Today, PCA methodology is also successfully used in civil engineering (e.g., for the classification of steel fibers used as reinforcement for concrete [8] or for the design of concrete mixes [7]), as well as image processing and signal denoising [8]. The authors of the present paper proved in a previous study that PCA is highly effective in conducting quality assessments of lunar soil simulants (LSSs) [9]. Traditional approaches to selecting the appropriate parameters as benchmark criteria for the development of LSSs have often fallen short of scientific and engineering expectations. Consequently, the exploration and analysis of extensive datasets containing information on the physical, chemical, mechanical, and geometrical properties of lunar soil should perhaps rely instead on multivariate statistical methods such as PCA. Keeping all the above facts in mind, we decided to perform a PCA analysis of the lunar soil samples from the Apollo and Luna missions. The primary objective of our analysis was to evaluate potential disparities among the findings obtained by these missions. The first task of our research program, then, was to confirm the reliability of the Apollo and Luna datasets. It must be remembered that both these datasets were published during the Cold War by the two competing superpowers of the age. The activities of these powers were always susceptible to political bias and a vast array of national security issues. In this light, for the present study, we decided to focus on identifying the key chemical constituents that should be prioritized in the development of new simulants (especially those intended for civil engineering applications). A previous evaluation of existing LSSs conducted by us [7,8] proved that they were not suitable for use in the development of future lunar building materials and construction technologies.

For the present study, we chose to depict our results in both two- and three-dimensional spaces. By generating virtual variables (which convey the vast majority of the information carried by the original variables), we achieved a comprehensive visualization of the chemical composition of the lunar soil, a task that was previously unattainable using conventional techniques. The use of PCA analysis enabled the identification of a minimal set of essential parameters for a precise evaluation of the lunar soil using a simple experimental setup. Using this method, a detailed examination of the key similarities and differences among the properties of specific lunar soil samples is feasible. It should be remembered that statistical treatments should always be associated with deterministic behavioral frameworks, so that the theoretical consistency of the results may be confirmed. In the present study, we sought to assess the reliability of the results based on statistics.

The successful use of PCA for the analysis of the lunar soil would enable the creation, classification, standardization, and large-scale production of new LSSs, which would be reliable, available in large quantities, and reasonably affordable. Such LSSs are urgently needed by the research community (e.g., for testing possible future techniques and technologies for lunar construction). In short, a successful analysis of the lunar soil based on

multivariate statistics involving numerous parameters would allow a thorough and feasible assessment of its key properties.

## 2. Available Data

For our analysis, we used a number of sets of data on the chemical composition of the lunar soil. Table 2 presents data on the composition of the lunar soils obtained by the Apollo missions 11–17. These datasets were sourced from [9–14].

**Table 2.** Composition (wt. %) of lunar soil samples obtained by Apollo missions.

No.	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	P <sub>2</sub> O <sub>5</sub>	Mission
1	42.1	7.8	13.7	15.8	0.2	7.9	12	0.5	0.1	0.1	11
2	42.2	7.8	13.6	15.3	0.2	7.8	11.9	0.47	0.16	0.05	11
3	42.6	3.6	14.2	15.4	0.22	9.7	10.4	0.43	0.24	-	12
4	46	2.8	12.5	17.2	0.22	9.7	10.9	0.48	0.24	-	12
5	48.2	1.73	17.6	10.41	0.14	9.26	11.25	0.61	0.51	0.53	14
6	47.3	1.6	17.8	10.5	0.1	9.6	11.4	0.7	0.6	-	14
7	48.1	1.7	17.4	10.4	0.14	9.4	10.7	0.7	0.55	0.51	14
8	46.95	1.6	12.7	16.29	0.217	10.75	10.49	0.33	0.092	0.16	15
9	45.35	0.49	28.25	4.55	0.06	5.02	16.21	0.42	0.09	0.1	16
10	45.2	0.58	26.4	5.29	0.7	6.1	15.32	0.52	0.14	0.12	16
11	44.65	0.56	27	5.49	0.7	5.84	15.95	0.44	0.13	0.1	16
12	44.9	0.47	27.7	5.01	-	5.69	15.7	0.51	0.22	0.16	16
13	44.77	0.37	28.99	4.35	0.07	4.2	16.85	0.44	0.06	0.05	16
14	45	0.54	27.3	5.1	0.3	5.7	15.7	0.46	0.17	0.11	16
15	41.67	6.52	13.57	15.37	0.21	10.22	11.18	0.34	0.09	0.06	17
16	39.82	9.52	11.13	17.41	0.25	9.51	10.85	0.32	0.07	0.06	17
17	40.09	9.32	10.7	17.85	0.24	9.92	10.59	0.36	0.08	0.07	17
18	42.2	5.09	15.7	12.4	0.15	10.3	11.5	0.24	0.07	-	17

Table 3 presents data on the composition of the lunar soils obtained by Luna 16 (cases 19–52), Luna 20 (cases 53–59), and Luna 24 (cases 60–90). The datasets were sourced from [15–20] in the case of Luna 16, from [19,21–24] in the case of Luna 20, and from [25–31] in the case of Luna 24. The original datasets had only partial information regarding the contents of Cr<sub>2</sub>O<sub>3</sub>, Fe<sub>2</sub>O<sub>3</sub>, and S. Therefore, due to this lack of quantitative information, only 10 oxides were taken into account as variables in the PCA calculations (see Table 4). Cases 1–4, 8, 15–52, and 60–90 concern mare regions, cases 9–14 and 53–59 concern highlands, and cases 5–7 concern areas of transition between highlands and mare regions. In Tables 2 and 3, empty data slots can be observed. For the PCA calculations, these slots were filled with the average values for soil and basalt within each table.

**Table 3.** Composition (wt. %) of lunar soil samples obtained by Luna missions.

No.	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	P <sub>2</sub> O <sub>5</sub>	Mission
19	41.7	3.39	15.32	16.8	0.21	8.73	12.2	0.37	0.1	-	16
20	41.2	3.46	15.4	16.55	0.2	8.82	12.8	0.36	0.12	-	16
21	42.5	3.3	15.45	16.3	0.2	8.96	12.42	0.36	0.1	-	16
22	41.3	3.42	15.15	16.9	0.22	8.6	12.55	0.28	0.1	-	16
23	41.93	3.36	15.33	16.66	0.2	8.78	12.53	0.34	0.1	0.12	16
24	43.8	4.9	13.65	19.36	0.2	7.05	10.4	0.38	0.15	0.12	16
25	42.95	5.5	13.88	20.17	0.2	6.05	10.8	0.23	0.16	0.14	16
26	45.5	4.04	13.95	17.77	0.26	5.95	11.96	0.63	0.21	0.15	16
27	45.17	2.9	16.98	13.21	0.22	4.02	13.32	0.69	0.17	-	16
28	43.36	4.37	15.13	17.48	0.27	4.97	12.77	0.7	0.17	-	16

Table 3. Cont.

No.	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	P <sub>2</sub> O <sub>5</sub>	Mission
29	44.2	2.48	16.45	13.67	0.2	4.3	12.65	0.69	0.21	-	16
30	46.6	6.1	15.7	17.2	0.28	3.7	11.3	0.46	0.24	0.12	16
31	42.8	3.17	16.4	17.6	0.26	8.8	12.9	0.43	0.144	-	16
32	35.3	3.7	8.7	25.6	0.29	5.5	9.1	0.56	0.2	-	16
33	36.8	3.8	8.8	25.7	0.28	5.6	8.7	0.66	-	-	16
34	52	4.2	8.9	25.2	0.26	4.2	8.7	0.62	0.26	-	16
35	44.1	4.2	9.3	23.1	0.27	4.8	9.4	0.53	0.23	-	16
36	-	3.5	9.1	22.5	0.23	10.3	12.3	0.4	0.21	-	16
37	48.1	5.3	13.1	24.2	0.28	6.3	12	0.44	0.17	-	16
38	46.6	4.8	13	19.9	0.32	8.6	10.4	0.4	0.19	-	16
39	-	5.3	12.2	21.6	0.27	9.5	11.6	0.46	-	-	16
40	59.1	4.8	13.6	22.7	0.29	7.8	11.9	0.46	0.17	-	16
41	46	4.2	9.6	17.6	0.27	7	10.5	0.53	0.18	-	16
42	46.3	1.02	20.2	11.1	0.17	2.32	14.8	0.83	0.44	0.23	16
43	47.3	2.03	19	12.1	0.21	3.1	14.3	0.68	0.32	0.13	16
44	46.3	2.16	19.3	12.9	0.2	3.8	15	0.52	0.19	0.05	16
45	46.7	2.48	16	14.1	0.23	3.7	15.4	0.55	0.25	0.1	16
46	44.6	3.5	16.5	15.3	0.23	4.6	14.3	0.39	0.18	0.02	16
47	43.2	4.8	14.3	16.4	0.27	4.9	13.4	0.47	0.21	0.06	16
48	44.1	3.7	14.5	16.6	0.26	5.2	14.2	0.5	0.24	0.05	16
49	45.6	3.5	14.2	17.3	0.25	5.2	13.3	0.34	0.24	0.11	16
50	43.7	4.8	12.1	18.8	0.3	6.3	12.2	0.45	0.24	0.1	16
51	42.6	1.05	19.4	18.7	0.22	4.1	12.2	0.56	0.29	0.07	16
52	41.3	1.93	11.5	21.6	0.28	12	9.3	0.39	0.17	0.05	16
53	45.6	0.46	22.9	7.5	0.106	9.15	14.5	0.4	0.069	-	20
54	45.1	0.55	22.3	7	0.13	9.8	15.1	0.5	0.1	0.16	20
55	45.4	0.47	23.44	7.37	0.1	9.19	13.38	0.29	0.067	0.06	20
56	45.8	0.533	21.6	7.02	0.13	9.85	14.9	0.46	0.1	0.17	20
57	44.4	0.56	22.9	7.03	0.12	9.7	15.2	0.55	0.1	0.14	20
58	42.8	0.47	23.6	6.6	0.1	9.5	14.4	0.35	0.06	0.14	20
59	44.2	0.52	19.1	6.91	0.12	13.37	13.3	0.48	0.47	0.17	20
60	43.9	1.3	12.5	19.8	0.25	9.4	12.3	0.31	0.04	0.11	24
61	43.3	1.13	15.2	16.3	0.22	8.69	13.1	0.42	0.04	0.14	24
62	43.5	1.09	15.5	16.2	0.21	8.87	12.9	0.51	0.04	0.13	24
63	43.6	1.13	15.9	16.2	0.23	8.8	13.3	0.25	0.04	0.13	24
64	43.7	1.23	16	16.1	0.22	8.75	13.1	0.43	0.06	0.12	24
65	45.2	0.89	13.8	20.5	0.27	6.35	12.7	0.24	0.01	-	24
66	48	1	13.1	19.5	0.31	5.2	13.1	0.29	0.04	0.11	24
67	43.9	0.74	19	16.6	0.19	5.2	14	0.5	0.06	-	24
68	45.5	0.96	13.9	18.4	0.24	6.3	13.3	0.37	0.02	0.02	24
69	45.3	1.16	12.4	20.3	0.27	7.5	12.2	0.37	0.03	-	24
70	47.3	0.37	26.8	6.99	0.11	1.03	17.1	0.66	0.04	-	24
71	47.6	0.2	9.94	14.7	0.25	13	12.2	0.21	0.03	-	24
72	41.1	0.58	10.4	24.9	0.38	11.6	8.64	0.29	0.02	-	24
73	47.1	1.27	12.8	17.6	0.25	7.16	12.9	0.3	0.04	-	24
74	44.8	0.82	11.1	21.9	0.29	10.4	9.94	0.32	0.18	-	24
75	46.9	0.8	13	19.3	0.28	6.53	13.1	0.44	0.23	-	24
76	46.5	0.67	13.3	17.2	0.29	7.2	13.1	0.3	0.04	0.02	24
77	46.4	0.79	13.7	18.5	0.3	6.5	13.3	0.28	0.04	-	24
78	46.6	0.86	12.9	17.4	0.2	6.3	13.3	0.31	0.04	0.04	24
79	48.3	1.06	12	18.1	0.25	6.8	12.7	0.4	0.04	0.02	24
80	44.6	0.8	12.7	17.8	0.26	6.5	13.7	0.29	0.03	0	24
81	46.1	1.14	11.9	17.4	0.24	6.2	12.8	0.02	0.04	0.02	24
82	42.8	0.3	16.4	15.3	0.24	6.2	15.3	0.41	0.06	0.01	24
83	43.8	0.35	12.7	20.3	0.3	8	13.5	0.36	0.05	0.02	24
84	44.3	0.09	11.6	20.5	0.3	7.6	12.7	0.36	0.04	0.05	24

Table 3. Cont.

No.	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MnO	MgO	CaO	Na <sub>2</sub> O	K <sub>2</sub> O	P <sub>2</sub> O <sub>5</sub>	Mission
85	46	0.84	15.8	15.5	0.15	5.8	13.9	0.3	0.03	0.04	24
86	46.4	0.28	15.8	16.3	0.17	5.8	13.9	0.34	0.03	-	24
87	47.8	0.31	14.8	15.8	0.25	5.9	13.9	0.32	0.03	0.03	24
88	45.4	0.66	8.9	19.2	0.37	15.4	8.9	0.2	0.05	-	24
89	43.1	0.15	6.9	20.1	0.25	20.8	6.9	0.12	0.03	-	24
90	44.2	0.62	7.7	21.6	0.43	18.2	7	0.15	0.05	0.03	24

Table 4. Contributions of variables to PCA factors.

Variable Designation	Chemical Assignment	Contribution of the Variable [%]		
		F1	F2	F3
1	SiO <sub>2</sub>	2.42	0.02	0.24
2	TiO <sub>2</sub>	2.99	14.06	<b>8.69</b>
3	Al <sub>2</sub> O <sub>3</sub>	<b>23.51</b>	3.61	0.53
4	FeO	<b>20.96</b>	2.44	7.81
5	MnO	3.68	0.76	8.2
6	MgO	7.85	0.04	<b>48.41</b>
7	CaO	<b>18.61</b>	10.9	3.56
8	Na <sub>2</sub> O	10.31	<b>16.86</b>	7.36
9	K <sub>2</sub> O	4.19	<b>31.34</b>	0.33
10	P <sub>2</sub> O <sub>5</sub>	5.48	<b>19.97</b>	<b>14.87</b>

### 3. Methodology

The sequence of the five steps involved in the PCA methodology is summarized in Figure 1.

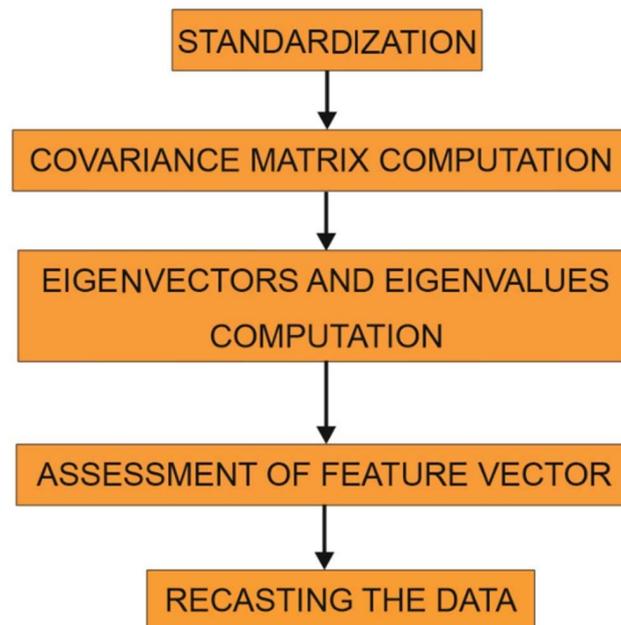


Figure 1. The five-step PCA process [32].

The PCA methodology is essentially a data transformation technique consisting of five consecutive steps [32]. In the first step, the range of the continuous original variables is standardized so that each contributes equally to the analysis. For the PCA to function, all

variables should have the same scale. The analysis of an exemplary vector  $v$ , onto which it is possible to project any point, is, therefore, conducted as follows:

$$V_1 = [2 \ 0 \ 1]; \quad ||v||^2 = 2^2 + 0^2 + 1^2 = 5 \quad (1)$$

The normalized vector is then obtained, as follows:

$$V_1 = \left[ \frac{1}{\sqrt{5}} \ 0 \ \frac{1}{\sqrt{5}} \right] \quad (2)$$

The second step involves covariance matrix computation. Essentially, this is carried out to assess how the original variables vary from the mean value with respect to each other. In the third step, eigenvalues are computed to establish the principal components of the dataset. In the fourth step, components of lesser significance (characterized by low eigenvalues) are discarded. Finally, in the fifth step, data are reoriented from the original axes to those represented by the principal components. In the present study, a PCA classification of the Apollo and Luna soil samples was conducted using the Statistica v. 13.3 computer program.

The acquisition of uncorrelated linear combinations of the original variables by PCA methodology is summarized in Figure 2. New virtual variables (called principal components) are created that cover the majority of the information provided by the original untreated variables. It should be remembered that the principal components are usually more difficult to interpret in comparison to physical variables, and also that they do not have any direct physical meaning, because they are artificially computed as linear combinations of the original variables.

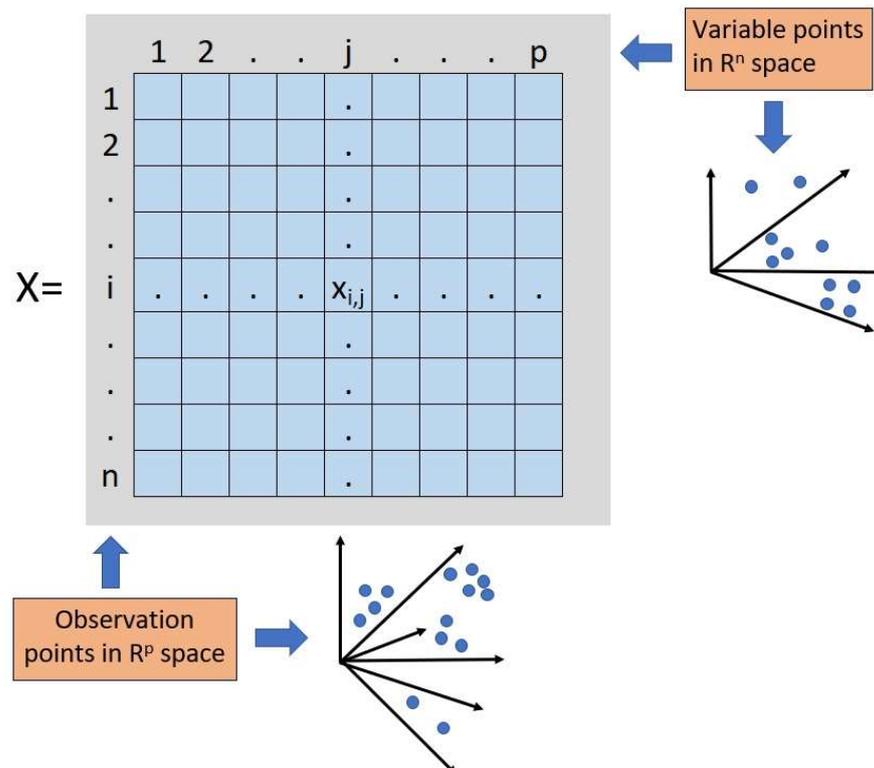


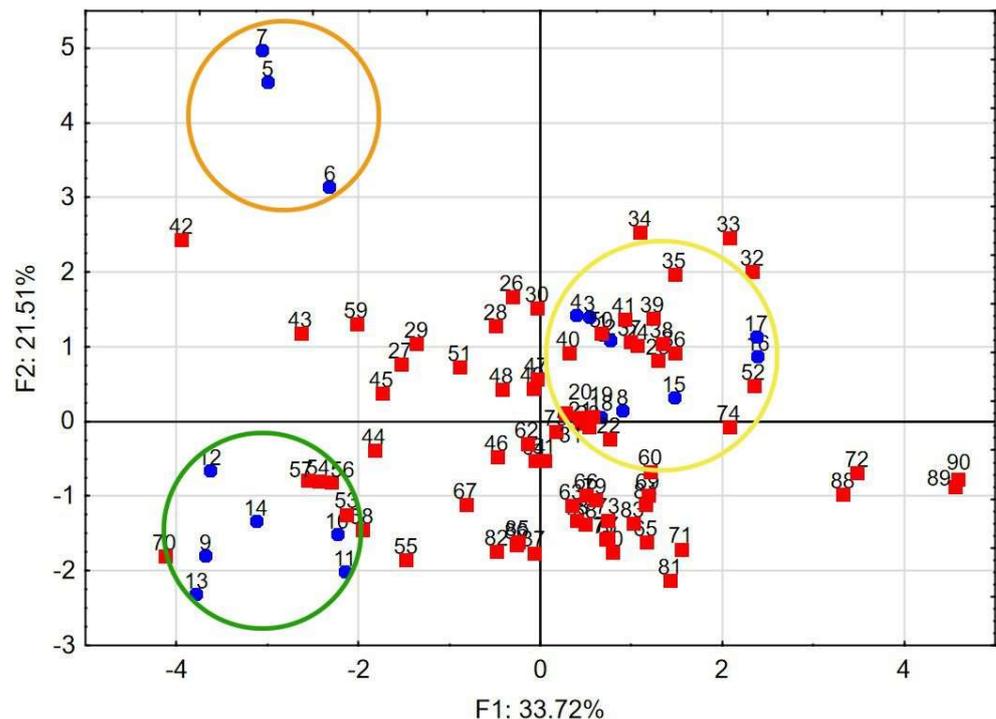
Figure 2. Raw data matrix  $X$  used for PCA.

#### 4. Results and Discussion

The results of the PCA processing of the datasets shown in Tables 2 and 3 are summarized in Table 4. The variables characterized by the highest contributions to three factors of the performed PCA analysis (marked with bold red) were as follows: factor 1 (F1)—

$\text{Al}_2\text{O}_3$ ,  $\text{FeO}$ , and  $\text{CaO}$ ; factor 2 (F2)— $\text{K}_2\text{O}$ ,  $\text{P}_2\text{O}_5$ , and  $\text{Na}_2\text{O}$ ; and factor 3 (F3)— $\text{MgO}$ ,  $\text{P}_2\text{O}_5$ , and  $\text{TiO}_2$ .

A PCA plot in two-dimensional space of factors F1 and F2 is presented in Figure 3. The PCA identifies the smallest number of factors or components necessary to explain all the variance, or as much as possible. In this context, a factor or component is a set of variables that, when combined in a linear fashion, explains some proportion of the observed variance [33]. In Figure 3, the two factors explain over 55% of the variability, and the lunar soil samples are clearly separated along both axes (factor F1 and F2).



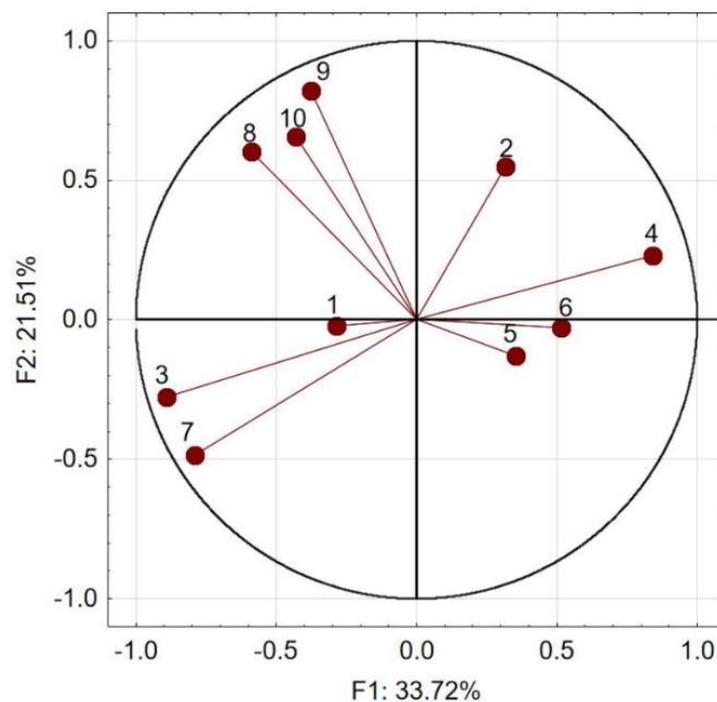
**Figure 3.** Factor scores F1 and F2 in two-dimensional space. Cases 1–18 (blue circles) are from Apollo missions and cases 19–90 (red squares) are from Luna missions.

It can be seen that the samples from Apollo 14 (orange circle) and Apollo 16 (green circle) are significantly differentiated from the rest of the population. However, the samples from the other Apollo missions (yellow circle) fit with the samples from the Luna missions.

Figure 4 shows the PCA distribution of the variables set in 2D-factor-loading space. Samples characterized by high values of  $\text{Al}_2\text{O}_3$  (variable 3) and  $\text{CaO}$  (variable 7) are located in the lower-left quarter of the chart (see Table 4 and Figure 4), which includes samples from the Apollo 16 mission (see Table 2, samples 9–14; and Figure 3) and Luna 20 mission (see Table 3, samples 53–59). These results are characteristic of samples taken from highlands (see Table 1). Samples characterized by high values of  $\text{Na}_2\text{O}$  (variable 8),  $\text{K}_2\text{O}$  (variable 9), and  $\text{P}_2\text{O}_5$  (variable 10) are located in the upper-left quarter of the chart (see Table 4 and Figure 4), which mainly applies to samples from the Apollo 14 mission taken from areas of transition between highlands and mares (see Table 1). Samples characterized by high values of  $\text{TiO}_2$  (variable 2) and  $\text{FeO}$  (variable 4) are located in the upper-right quarter of the chart (see Table 4 and Figure 4), which applies to samples from the Apollo 11, 12, 15, and 17 missions (see Table 2 and Figure 3) and Luna 16 mission (see Table 3). These results are characteristic of samples taken from mare regions (see Table 1).

In the two-dimensional space presented in Figure 3, some reasonably clear clusters of results may be identified. The first covers results from Apollo 14 (surrounded by an orange circle), and the second covers results from Apollo 16 (surrounded by a green circle). The rest of the results are scattered in the central part of the two-dimensional space, forming

a loose cloud compatible with the Apollo 11, 12, 15, and 17 missions (marked with the yellow circle). When analyzing Figure 3, it should be remembered that, during the Apollo 14 mission, samples of lunar soil from the mare–highland border region near the Fra Mauro crater were collected (see Table 1 for landing coordinates). Samples 5 and 7 (see Figure 3) can be identified as outliers without any knowledge of landing sites; however, together with sample 6, their position in the chart is justified. This is because breccias are dominant in the samples taken by Apollo 14. This is in contrast to those missions that landed at mares, where the predominant rocks were basalts. Apollo 16 obtained highland lunar soil. The rest of the Apollo missions obtained mare/oceanus samples. With regard to the Soviet missions, Luna 16 landed in Mare Fecunditatis, Luna 20 landed in highlands [34,35], and Luna 24 landed in a mare area.

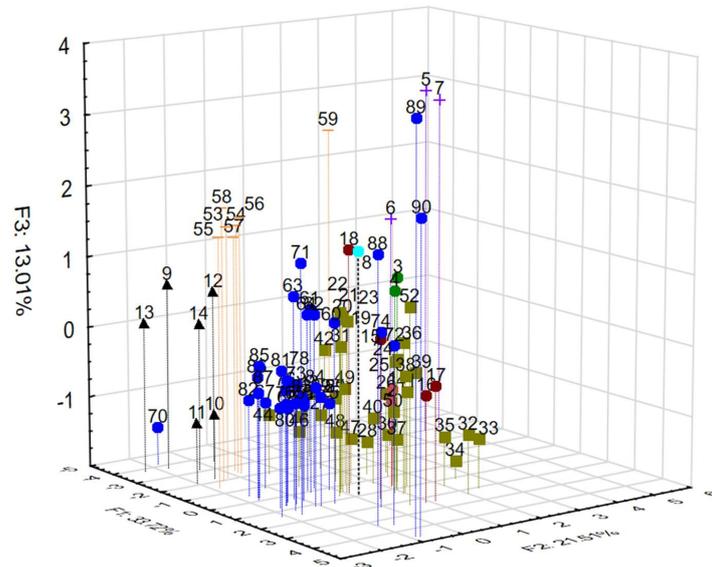


**Figure 4.** PCA projection of variables set in 2D-factor-loading space (variable designations and chemical assignments are shown in Table 4).

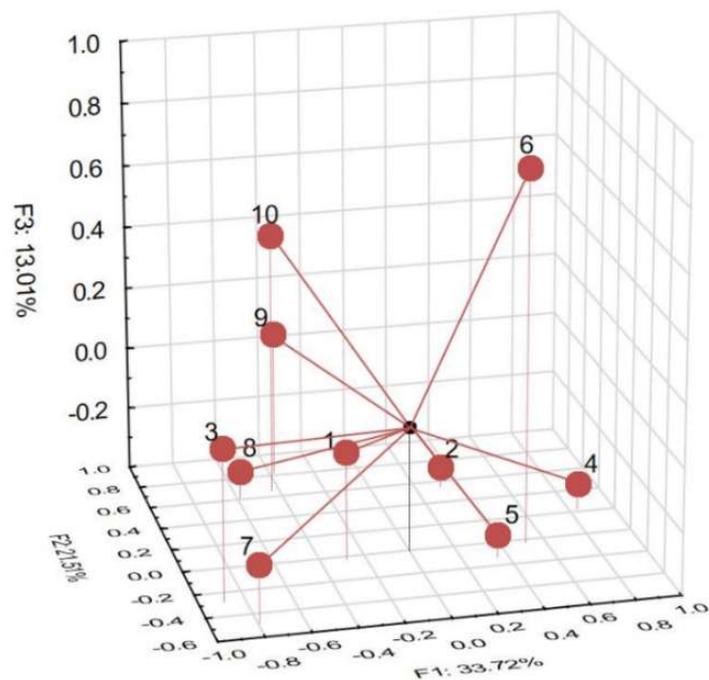
A two-factor analysis showed that the highland and border-region lunar samples differed significantly from the mare samples. The samples from highlands were characterized by relatively high amounts of  $\text{Al}_2\text{O}_3$  and  $\text{CaO}$  and low amounts of  $\text{TiO}_2$  and  $\text{FeO}$  (variables 3, 7, 2, and 4, respectively; see Table 4 and Figure 4). The specimens from mares varied, forming only a loose cluster of points with some significant outliers (e.g., samples 42, 70, 89, and 90). To obtain more definite outcomes, the authors decided to conduct a PCA in three-dimensional space. A PCA performed in 3D space (factor scores F1, F2, and F3) makes it possible to obtain more detailed information about differences between cases than is possible using a 2D analysis. The resulting PCA plot in 3D space for factors F1, F2, and F3 is presented in Figure 5.

The three factors explain over 68% of the variability. The graph shows that the lunar soil samples are clearly separated along all three axes (factors F1, F2, and F3). The factor F3 samples collected by Apollo 14 (violet) and 16 (black) form clusters separate from the rest of the Apollo samples and most of the Luna samples; however, the Luna 20 cluster, which was close to the Apollo 16 cluster on the 2D chart, is placed significantly higher than the Apollo 16 cluster on the 3D chart due to the high amounts of  $\text{MgO}$  and  $\text{P}_2\text{O}_5$ .

In Figure 6, the PCA distribution of the variables set in 3D-factor-loading space is presented. The variables that most heavily influence the distribution of the samples along the F3 axis are 6 (MgO), 10 (P<sub>2</sub>O<sub>5</sub>), and 2 (TiO<sub>2</sub>).



**Figure 5.** PCA plot in three-dimensional space (factor scores F1, F2, and F3). Cases 1–18 are from Apollo missions (cases 1–2, Apollo 11; cases 3–4, Apollo 12; cases 5–7, Apollo 14; case 8, Apollo 15; cases 9–14, Apollo 16; and cases 15–18, Apollo 17). Cases 19–90 are from Luna missions (cases 19–52 (khaki green squares), Luna 16; cases 53–59 (orange dashes), Luna 20; cases 60–90 (blue circles), Luna 24).



**Figure 6.** PCA distribution of the variables set in 3D-factor-loading space (see Table 4 for variable designation and chemical assignment).

Before drawing conclusions, it is worth noting that the Apollo missions resulted in a total mass of samples that was a thousand times greater than that obtained by the Luna missions. This discrepancy may be readily explained. The Luna missions were stationary. The automatic lander and, consequently, any collection of samples, was confined to the

specific landing point. During the Apollo missions, human astronauts were able to walk around the landers. During missions 15, 16, and 17, they could also drive a Lunar Rover. Table 1 summarizes the routes travelled on the lunar surface during particular missions. Differences in lunar soil results may also be explained by variations in the procedures by which samples were collected and tested by different missions. However, the analysis carried out in the present study confirms the high consistency of results with respect to the chemical composition of the lunar regolith obtained during the Apollo and Luna missions.

## 5. Conclusions

The analysis conducted in the present study allows a number of conclusions to be drawn, as follows:

- The PCA technique enables the swift and reliable categorization of soil samples obtained from both mare and highland areas of the lunar surface.
- The calculation method used allows for the identification of chemical factors that may contribute to the grouping of objects within clusters depicted on PCA graphs.
- In terms of chemical composition, the samples obtained by the American Apollo missions appear to be very similar to those obtained by the Soviet Luna missions; the reliability of the Apollo and Luna datasets is, therefore, confirmed.
- The analysis reveals close similarities in the chemical compositions of samples originating from the same type of land, i.e., highlands or mares.
- The PCA method may be applied to distinguish the types of rocks contained in tested samples of lunar regolith.
- The creation of a new type of LSS (dedicated for civil engineering applications) is enabled.

**Author Contributions:** Conceptualization, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; methodology, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; software, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; writing—original draft preparation, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; writing—review and editing, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; visualization, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S.; supervision, J.K. (Jacek Katzer), J.K. (Janusz Kobaka) and K.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Science Centre (Poland), grant number DEC-2020/38/E/ST8/00527.

**Data Availability Statement:** All data generated or analyzed during this study are included in this published article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Seweryn, K.; Skocki, K.; Banaszekiewicz, M.; Grygorczuk, J.; Kolano, M.; Kuciński, T.; Mazurek, J.; Morawski, M.; Białek, A.; Rickman, H.; et al. Determining the Geotechnical Properties of Planetary Regolith Using Low Velocity Penetrometers. *Planet. Space Sci.* **2014**, *99*, 70–83. [[CrossRef](#)]
2. Pearson, K. LIII. On Lines and Planes of Closest Fit to Systems of Points in Space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [[CrossRef](#)]
3. Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24*, 417–441. [[CrossRef](#)]
4. Hotelling, H. Relations Between Two Sets of Variates. *Biometrika* **1936**, *28*, 321. [[CrossRef](#)]
5. Ghosh, A.; Barman, S. Application of Euclidean Distance Measurement and Principal Component Analysis for Gene Identification. *Gene* **2016**, *583*, 112–120. [[CrossRef](#)] [[PubMed](#)]
6. Argota-Perez, R.; Robbins, J.; Green, A.; Herk, M.V.; Korreman, S.; Vásquez-Osorio, E. Evaluating Principal Component Analysis Models for Representing Anatomical Changes in Head and Neck Radiotherapy. *Phys. Imaging Radiat. Oncol.* **2022**, *22*, 13–19. [[CrossRef](#)] [[PubMed](#)]
7. Kobaka, J. Principal Component Analysis as a Statistical Tool for Concrete Mix Design. *Materials* **2021**, *14*, 2668. [[CrossRef](#)] [[PubMed](#)]
8. Zarzycki, P.K.; Katzer, J.; Domski, J. Fast Classification of Fibres for Concretebased on Multivariate Statistics. *Comput. Concr.* **2017**, *20*, 23–29. [[CrossRef](#)]

9. Kobaka, J.; Katzer, J.; Zarzycki, P.K. Pilbara Craton Soil as a Possible Lunar Soil Simulant for Civil Engineering Applications. *Materials* **2019**, *12*, 3871. [[PubMed](#)]
10. Zarzycki, P.K.; Katzer, J. Assessment of Lunar Soil Simulants Based on Multivariate Statistics. In Proceedings of the Earth and Space 2018: Engineering for Extreme Environments, Cleveland, OH, USA, 9–12 April 2018. [[CrossRef](#)]
11. Hill, E.; Mellin, M.J.; Deane, B.; Liu, Y.; Taylor, L.A. Apollo Sample 70051 and High- and Low-Ti Lunar Soil Simulants MLS-1A and JSC-1A: Implications for Future Lunar Exploration. *J. Geophys. Res. Planets* **2007**, *112*. [[CrossRef](#)]
12. Wang, K.T.; Tang, Q.; Cui, X.M.; He, Y.; Liu, L.P. Development of Near-Zero Water Consumption Cement Materials via the Geopolymerization of Tektites and Its Implication for Lunar Construction. *Sci. Rep.* **2016**, *6*, 29659. [[CrossRef](#)] [[PubMed](#)]
13. Wallace, W.T.; Taylor, L.A.; Liu, Y.; Cooper, B.L.; McKay, D.S.; Chen, B.; Jeevarajan, A.S. Lunar Dust and Lunar Simulant Activation and Monitoring. *Meteorit. Planet. Sci.* **2009**, *44*, 961–970. [[CrossRef](#)]
14. Zheng, Y.; Wang, S.; Ouyang, Z.; Zou, Y.; Liu, J.; Li, C.; Li, X.; Feng, J. CAS-1 Lunar Soil Simulant. *Adv. Space Res.* **2009**, *43*, 448–454. [[CrossRef](#)]
15. Vinogradov, A.P. Preliminary Data on Lunar Ground Brought to Earth by Automatic Probe “Luna-16”. *Lunar Planet. Sci. Conf. Proc.* **1971**, *2*, 1.
16. Albee, A.L.; Chodos, A.A.; Gancarz, A.J.; Haines, E.L.; Papanastassiou, D.A.; Ray, L.; Tera, F.; Wasserburg, G.J.; Wen, T. Mineralogy, Petrology, and Chemistry of a Luna 16 Basaltic Fragment, Sample B-1. *Earth Planet. Sci. Lett.* **1972**, *13*, 353–367. [[CrossRef](#)]
17. Grieve, R.A.F.; McKay, G.A.; Weill, D.F. Microprobe Studies of Three Luna 16 Basalt Fragments. *Earth Planet. Sci. Lett.* **1972**, *13*, 233–242. [[CrossRef](#)]
18. Keil, K.; Kurat, G.; Prinz, M.; Green, J.A. Lithic Fragments, Glasses and Chondrules from Luna 16 Fines. *Earth Planet. Sci. Lett.* **1972**, *13*, 243–256. [[CrossRef](#)]
19. Cimbalknikova, A.; Palivova, M.; Frana, J.; Mastalka, A. Chemical Composition of Crystalline Rock Fragments from Luna 16 and Luna 20 Fines. In Proceedings of the The Soviet-American Conference on Cosmochemistry of the Moon and Planets, Moscow, Russia, 4–8 June 1977; pp. 263–275.
20. Kurat, G.; Kracher, A.; Keil, K.; Warner, R.; Prinz, M. Composition and Origin of Luna 16 Aluminous Mare Basalts. *Lunar Planet. Sci. Conf. Proc.* **1976**, *2*, 1301–1321.
21. Korotev, R.L.; Haskin, L.A.; Lindstrom, M.M. A Synthesis of Lunar Highlands Compositional Data. *Lunar Planet. Sci. Conf. Proc.* **1980**, *1*, 395–429.
22. McKay, D.S.; Heiken, G.; Basu, A.; Blanford, G.; Simon, S.; Reedy, R.; French, B.M.; Papike, J. The Lunar Regolith. In *Lunar Sourcebook: A Users Guide to the Moon*; Al, H., Ed.; Cambridge Univ. Press: Cambridge, UK, 1991.
23. Philpotts, J.A.; Schumann, S.; Bickel, A.L.; Lum, R.K.L. Luna 20 and Apollo 16 Core Fines: Large-Ion Lithophile Trace-Element Abundances. *Earth Planet. Sci. Lett.* **1972**, *17*, 13–18. [[CrossRef](#)]
24. Vinogradov, A.P. Preliminary Data on Lunar Soil Collected by the Luna 20 Unmanned Spacecraft. *Geochim. Cosmochim. Acta* **1973**, *37*, 721–729. [[CrossRef](#)]
25. McKay, D.S.; Basu, A.; Waits, G. Grain Size and the Evolution of Luna 24 Soils. In *Mare Crisium: The view from Luna 24*; Pergamon Press: Oxford, UK, 1978; pp. 125–136.
26. Barsukov, V.L. Preliminary Data for the Regolith Core Brought to Earth by the Automatic Lunar Station Luna 24. *Lunar Planet. Sci. Conf. Proc.* **1977**, *3*, 3303–3318.
27. Vaniman, D.T.; Papike, J.J. Ferrobasalts from Mare Crisium: Luna 24. *Geophys. Res. Lett.* **1977**, *4*, 497–500. [[CrossRef](#)]
28. Taylor, G.J.; Keil, K.; Warner, R.D. Very Low-Ti Mare Basalts. *Geophys. Res. Lett.* **1977**, *4*, 207–210. [[CrossRef](#)]
29. Barsukov, V.L.; Tarasov, L.S.; Dmitriev, L.V.; Kolesov, G.M.; Shevaleevskii, I.D.; Garanin, A.V. The Geochemical and Petrochemical Features of Regolith and Rocks from Mare Crisium (Preliminary Data). *Lunar Planet. Sci. Conf. Proc.* **1977**, *3*, 3319–3332.
30. Taylor, L.A.; Onorato, P.I.K.; Uhlman, D.R.; Coish, R.A. Subophitic Basalts from Mare Crisium: Cooling Rates. In *Mare Crisium: The view from Luna 24*; Pergamon Press: Oxford, UK, 1978; pp. 473–482.
31. Ryder, G.; Mcsween, H.Y.; Marvin, U.B. Basalts from Mare Crisium. *Moon* **1977**, *17*, 263–287. [[CrossRef](#)]
32. Jolliffe, I.T. Principal Component Analysis, Second Edition. *Encycl. Stat. Behav. Sci.* **2002**, *30*, 487. [[CrossRef](#)]
33. Mulaik, S.A. Blurring the Distinctions Between Component Analysis and Common Factor Analysis. *Multivar. Behav. Res.* **1990**, *25*, 53–59. [[CrossRef](#)] [[PubMed](#)]
34. Korotev, R.L. Lunar Geochemistry as Told by Lunar Meteorites. *Chem. Der Erde* **2005**, *65*, 297–346. [[CrossRef](#)]
35. Frank, M. The Lunar and Planetary Science XXXIII Conference. *Powder Diffr.* **2002**, *17*, 254. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.