

Viking GCMS Digital Data Set
Crowdsourcing Restoration Intro

Purpose of document: to present a summary for crowd-sourcing of the restoration project of the Viking Lander GCMS digital data set.

Introduction:

The fundamental data generated by the Viking GCMS instrument are mass scans from the mass spectrometer. It appears from the existing data set that the mass range (m/z) is 12-215. Mass scans were made every 10.24 seconds. Up to 500 scans were produced for each GCMS run (Biemann et al., 1977). The total data are 16 GCMS runs (two blanks, five at VL-1, and nine at VL-2) focused on the analysis of four soil samples, as listed in Table 1. The surviving form of the data set is explained below.

Table 1: A summary of the Viking GCMS data set.

Sample	Run ID Number	Date of Analysis	Oven Temp
VL-1 Blank	10008	-	-
VL-1 Sample 1	10015	sol 17	200
	10018	sol 23	500
VL-1 Sample 2	10023	sol 32	350
	10024	sol 37	500
	10025	sol 43	500
VL-2 Blank	10007	-	-
VL-2 Sample 1	10032	sol 24	200
	10032	sol 26	350
	10034	sol 35	500
	10035	sol 37	500
VL-2 Sample 2	10036	sol 41	50
	10037	sol 43	200
	10038	sol 45	350
	10039	sol 47	500
	10041	sol 61	500

The Viking GCMS digital data is currently archived in the National Space Science Data Center at NASA Goddard. The GCMS data currently exists in 3 forms:

- 1) Raw form on IBM-compatible tapes,
- 2) Presented as bar graphs on 16-mm microfilm, and
- 3) Full and reduced versions of the data stored in files with .phys extension.

Form 3, the digital version, is incoherent. The purpose of this document will be to present the format of Form 3 in order to restore this data set. NASA Ames currently has Form 2 and 3 and hope to use Form 2 (the microfilm) to assist in the restoration of Form 3 (the digital data).

Summary of data files: The data is stored in 32 .phys files. The data files are split between four folders: two folders include all sampling data while two folders include reduced versions of that same data. Each GCMS run is on a separate file, and there is one record for each spectral scan, including mass spectrum data and engineering data. There is no data set catalog with information to give details on the formats, units, etc. Table 2 shows the known breakdown and the size of each file.

Table 2: A summary of the Viking GCMS digital data as collected from the NSSDC.

Folder ID/NSSDC ID	Lander	File Name (Sample Run)	Description
5289/ PSPG-00248	VL-1 GCMS Full data set	DR005289_F00001.phys	6 files: VL-1 Blank Run Sample 1, Run 1 Sample 1, Run 2 Sample 2, Run 1 Sample 2, Run 2 Sample 2, Run 3
		DR005289_F00002.phys	
		DR005289_F00003.phys	
		DR005289_F00004.phys	
		DR005289_F00005.phys	
		DR005289_F00006.phys	
5631/ PSPG-00010	VL-1 GCMS Compressed	DR005631_F00001.phys	
		DR005631_F00002.phys	
		DR005631_F00003.phys	
		DR005631_F00004.phys	
		DR005631_F00005.phys	
		DR005631_F00006.phys	
5967/ PSPG-00246	VL-2 GCMS Full data set	DR005967_F00001.phys	10 files: VL-2 Blank Run Sample 1, Run 1 Sample 1, Run 2 Sample 1, Run 3 Sample 1, Run 4 Sample 2, Run 1 Sample 2, Run 2 Sample 2, Run 3 Sample 2, Run 4 Sample 2, Run 5
		DR005967_F00002.phys	
		DR005967_F00003.phys	
		DR005967_F00004.phys	
		DR005967_F00005.phys	
		DR005967_F00006.phys	
		DR005967_F00007.phys	
		DR005967_F00008.phys	
		DR005967_F00009.phys	
		DR005967_F00010.phys	
5388/ PSPG-00324	VL-2 GCMS Compressed	DR005388_F00001.phys	
		DR005388_F00002.phys	
		DR005388_F00003.phys	
		DR005388_F00004.phys	
		DR005388_F00005.phys	
		DR005388_F00006.phys	
		DR005388_F00007.phys	
		DR005388_F00008.phys	
		DR005388_F00009.phys	
		DR005388_F00010.phys	

Microfilm:

Form 2, the microfilm data, is potentially useful for decoding the digital data. Below is a description of the Viking GCMS microfilm data set. Each mass scan includes 2 graphs, an upper panel (f) which shows a complete spectrum normalized to most abundant ion and a lower panel which shows a spectrum with peaks below m/z 47 deleted and the plot is normalized to most abundant ion $> m/z$ 47, as shown in Figure 1.

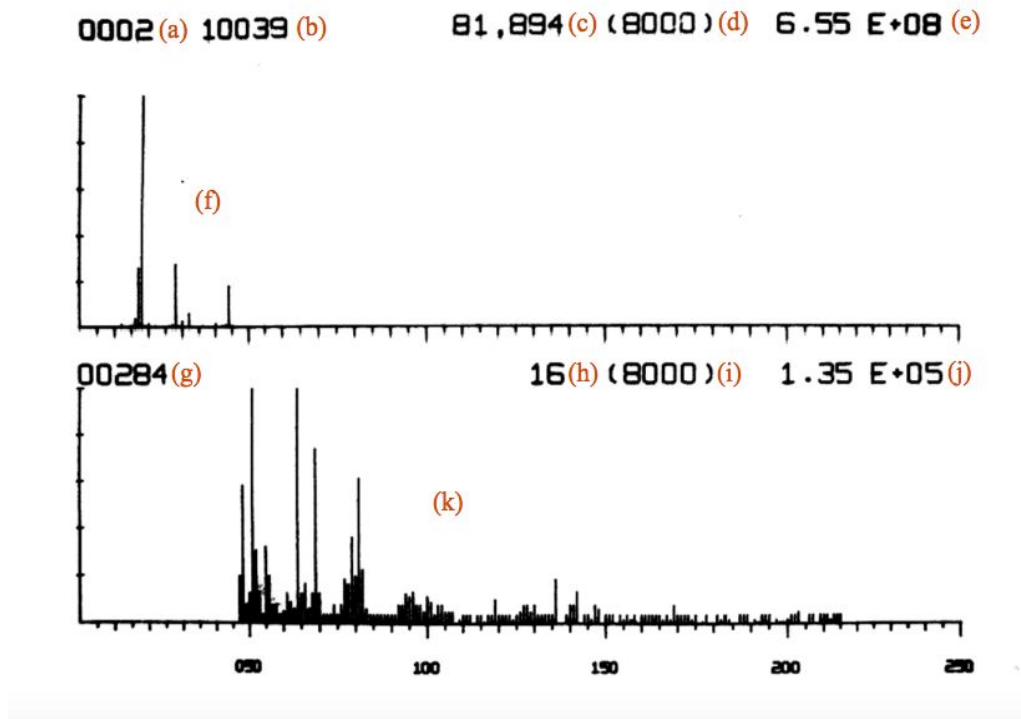


Figure 1: Representative Viking GCMS mass spectra as preserved on microfilm.

Microfilm Key:

Key	Description
a	MIT Scan Number
b	GCMS Run ID Number
c and h	Intensity of most abundant ion in arbitrary units and adjusted for effluent divider status
d and i	Effluent divider status
e and j	Total parts in arbitrary units and uncorrected for effluent divider status
f	Upper spectrum, complete spectrum normalized to most abundant ion
g	Mission scan number
k	Lower spectrum, peaks below m/z 47 are deleted and the plot is normalized to most abundant ion > m/z 47

Comparison of digital data and microfilm:

It may be possible to use the microfilm as a comparison in order to decipher the digital data. And comparison of the two data sets in their current form is explained below.

Single mass scan (digital data)-

Note: this is taken from the compressed form of the digital data.

The following bytes represent a single mass scan, Scan 148 from VL-2, Sample 2, GCMS Run 10039.

00 00 00 00-00 00 00 00-00 05 00 94-27 37 00 01	0a 58 48 20-86 5c 71 d6-7a 58 4e 87-18 57 77 f1
01 ae 00 26-00 00 00 00-7f ff 01 1b-00 01 00 00	e5 56 41 09-0d 53 00 00-00 00 56 c9-f1 58 77 ea
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	39 58 43 3e-84 5a 62 ee-ee 5d 56 c9-f1 58 4d 74
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	99 58 60 3a-bf 55 40 39-c2 57 5f 5e-49 57 54 7e
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	94 59 66 5c-68 58 40 b4-5e 5a 00 00-00 00 4a 86
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	8c 56 77 f1-e5 56 59 74-5f 57 66 5c-68 58 77 4e
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	a7 5a 49 45-68 59 5f 5e-49 57 4e 87-18 57 54 7e
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	94 59 73 d4-8c 5d 5e d4-53 5a 43 d9-e1 5f 49 45
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	68 59 66 5c-68 58 77 f1-e5 56 77 f1-e5 56 77 f1
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	e5 56 00 00-00 00 42 61-67 59 40 b4-5e 5a 61 bd
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	ed 5a 5f 5e-49 57 49 45-68 59 49 17-0a 58 00 00
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	00 00 42 61-67 59 41 9d-a3 5c 5f fc-ca 5f 66 5c
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	68 58 61 bd-ed 5a 4d 74-99 58 58 79-07 59 7b 21
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	3f 5a 4d 74-99 58 61 bd-ed 5a 77 ea-39 58 60 c8
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	0b 59 49 45-68 59 43 3e-84 5a 7c 6d-7e 59 4a 89
00 04 00 00-00 00 00 00-00 00 00 00-00 00 00 00	05 5e 73 e8-09 5e 69 b8-29 62 73 d4-8c 5d 73 e8
00 60 00 62-00 72 00 7d-00 86 00 8e-00 91 00 a0	09 5e 54 aa-cb 60 70 61-bd 5f 6c c0-16 5d 4f 48
01 b5 00 81-00 77 00 55-01 49 00 b9-00 7f 00 4d	bd 5c 54 80-38 5d 66 2b-78 5e 54 7e-94 59 66 5c
00 26 01 ff-01 2a 00 09-00 0f 00 0e-00 94 01 ae	68 58 6e 1b-be 59 77 76-aa 5c 4b 23-7a 5a 4a 89
09 b5 00 00-00 00 00 00-00 00 00 00-00 00 00 00	05 5e 7b 5f-25 5d 7c 6d-7e 59 41 92-ae 5b 57 18
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	1c 5b 77 4e-a7 5a 4f 3b-69 5b 53 65-1a 5a 41 92
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	ae 5b 59 f1-80 5c 54 80-38 5d 66 09-3f 5c 73 d4
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	8c 5d 5f bc-87 5b 7b 35-d2 5b 70 4e-d9 5e 63 10
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	1c 5f 4f 70-99 5f 62 ee-ee 5d 5a 00-a2 5d 5a 2d
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	e1 60 54 9c-8f 5f 43 f0-a3 61 5d 6f-39 66 6c f6
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	c2 60 54 9c-8f 5f 7b 88-86 5f 60 2d-0b 62 70 61
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	bd 5f 7b 73-c2 5e 51 ee-68 5e 57 43-fe 5e 4f 48
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	bd 5c 48 20-86 5c 53 65-1a 5a 41 f5-d7 64 4c f9
00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00	3c 5f 74 35-b4 62 54 b8-f1 61 60 6d-8d 66 4f 98
00 00 00 00-00 00 59 74-5f 57 00 00-00 00 00 00	88 62 66 4d-bd 60 48 44-c9 5f 65 f8-35 5b 40 b4
00 00 00 00-00 00 5c c3-bf 5b 69 4e-2e 5c 4a 95	5e 5a 4d 06-2e 60 6e 1b-be 59 6d 09-16 61 46 29
79 5f 00 00-00 00 5f 5e-49 57 40 39-c2 57 6f d7	9e 62 71 1e-83 69 46 70-46 68 4a e0-9e 65 57 8d
b5 56 4a 86-8c 56 7a 20-8e 55 4a 86-8c 56 00 00	30 63 4f 8b-3e 61 69 94-de 60 4f a5-d5 63 00 00
00 00 4a 86-8c 56 00 00-00 00 59 74-5f 57 00 00	00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00
00 00 00 00-00 00 54 7e-94 59 00 00-00 00 41 92	00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00
ae 5b 77 f1-e5 56 4e 87-18 57 60 3a-bf 55 48 ea	00 00 00 00-00 00 00 00-00 00 00 00-00 00 00 00
7f 55 4a 86-8c 56 60 3a-bf 55 00 00-00 00 60 3a	00 00 00 00-00 00 00 00-00 00 00 05-00
bf 55 00 00-00 00 00 00-00 00 00 00-00 00 6c 26	
87 57 00 00-00 00 7c 6d-7e 59 59 74-5f 57 60 3a	
bf 55 48 ea-7f 55 4a 86-8c 56 00 00-00 00 00 00	
00 00 4e 87-18 57 59 74-5f 57 6f d7-b5 56 6f d7	
b5 56 6f d7-b5 56 4e 87-18 57 00 00-00 00 56 c9	
f1 58 6f d7-b5 56 40 39-c2 57 48 ea-7f 55 7a 20	
8e 55 00 00-00 00 00 00-00 00 66 5c-68 58 7b 21	
3f 5a 4a 86-8c 56 60 3a-bf 55 6b 26-f6 56 6f d7	
b5 56 59 74-5f 57 5f 5e-49 57 00 00-00 00 77 ea	
39 58 6f d7-b5 56 77 f1-e5 56 4a 86-8c 56 49 17	

Single mass scan (microfilm data)-

The following image is the microfilm mass scan of the same scan: scan 148 from VL-2, Sample 2, GCMS Run 10039.

