KMS Technologies – KJT Enterprises, Inc.

Appendix 2
Data Format Standards

from

Strack, K.-M., 1992, *Exploration with Deep Transient Electromagnetics*: Elsevier, 373 pp.

Appendix 2
Data Format Standards

INTRODUCTION

One of the basic problems with electromagnetic methods is the lack of compatibility between different systems and with other geophysical techniques. Thus, it is very important to use already existing standards, rather than making new ones. This means initially more development work since all possible options and sidekick developments should be considered. In the long run, the return of increased application of the techniques will be larger than fighting forever incompatibilities.

From the beginning the choice was made to use a seismic standard. The only one suitable for original records is the SEGY standard (Barry et al, 1975). Using this standard also allows any seismic processing company to process electromagnetic data. This will bring a wealth of experience and new ideas into the processing and interpretation of EM data. The next step in standard would obviously be the adaption of the SEG 2 standard. At this stage this seems not feasible until sufficient users for the SEG 2 standard exist.

Transient electromagnetic data when recorded as original time series is in principle very similar to seismic data. The seismic source is substituted by the transmitter and the geophone traces by the transients. The need of a standard came about when larger scale LOTEM (a special transient EM configuration) measurements were conducted around the world.

This document describes the structure of the files as they are being used on the LOTEM Processing System (LOPS) software. At this stage, we have made the decision of either writing strictly SEGY tapes or a structure conforming internally with the SEGY standard adapted for the LOTEM Processing and Interpretation System (LOPIS). We strongly discourage you from using any other format, since the data will end up being non-interchangeable.

I would like to thank O. Engels, P.J. Bürger, J. Rossow for contributing all their thoughts and efforts to this standard. W. Schott spent countless hours checking the intermediate versions.

FILE FORMATS

On VAX machines the data is stored as binary random access file. The record size was selected to 256 bytes as a compromise between different field systems, number of real data points to be read into memory for selective stacking (1 record) and already existing computer standards. The selection for 256 bytes was made as being half of the VAX default record size. Even this moderate size gives you considerable storage problems when selectively stacking 1000 transient (i.e. reading 1000 records with 64 real words equals 256 kbytes of data). Effectively this can only be handled be virtual addressing operating systems.

There are three kinds of files used within VAX–based LOPIS: a raw or prestacked processed data file with 50 transients per file, a stacked data file with one transient per file, and an ASCII data file obtained after conversion to logarithmically spaced data. The latter is only of interest within LOPS and thus not discussed here. The prestack data file and the poststack data file have now the same format. This allows us to use any number of data points without changing the format.

On a PC the file format is identical to the VAX format except for the way the PC writes internally real numbers. Thus only binary files from the VAX must be converted to the PC format.

Format of a LOPS Data File

The example record numbers are calculated for 1024 data points. The respective values are given for 2048 data points in brackets in table A2.1. The record pointer is calculated from the number of data points in the file header. A graphic illustration is shown in figure A2.1

Table A2.1: Record structure for a LOPS data file containing several individual transients.

VAX RECORD No.	CONTENT:
1–15	File header information
16	Trace header of 1st transient
17–32(48)	Data of 1st transient
33(49)	Trace header of 2nd transient
34–49(50–81)	Data of 2nd transient
....	...
....	...
<EOF>	

This file format is presently used by LOPS. It is possible to have a variable file length. The standard maximum size (50 transients in a file) is:

- 865 records or 433 blocks for 1024 data points
- 1665 records or 833 blocks for 2048 data points.

Format of a LOPS Stacked Data File

The shown record numbers are calculated for 1024 data points. The respective values are given for 2048 data points in table A2.2 In brackets. The record pointer is calculated from the number of data points in the file header. A graphic illustration is shown in figure A2.1

Table A2.2: Record structure for a LOPS data file containing a stacked transient.

VAX RECORD No.	CONTENT:
1–15	File header information
16	Trace header of transient
17–32(48)	Data of transient
33–48(49–80)	Standard deviation of data
....	...
....	...
<EOF>	

Because there is only one transient in file, the file size can be calculated to

- 48 records or 24 blocks for 1024 data points.
- 80 records or 40 blocks for 2048 data points.

☞ NOTE

There are two coordinate systems in use. One is the absolute coordinate system used in the survey area. Please, use only rectangular coordinate systems. In FRG x (EAST) is the RECHTSWERT and y (NORTH) is the HOCHWERT. The second system is the reference coordinate system based on the transmitter with the origin (0,0) at the transmitter center. For this x, y coordinates are being used. When the transmitter information is available before acquisition the transformation can be done during the acquisition phase. Otherwise, it must be done during processing. E1 and E2 mark the respective electrode positions of the grounded wire dipole transmitter (source).

When the file contains records of an entire array the respective receiver position of the traces must be known. If the file contains single records at one site, the reference receiver position from the file header may be used.

☞ NOTE

Under a record we include all data recorded simultaneously during one source sweep (pulse, current switch etc). This means the i-th transient of a spread builds one record. Under a trace we include each individual transient recorded at each receiver site.

VAX Rec.	Byte pos.	Meaning	File structure
1	1	CARD IMAGE HEADER	FILE HEADER
.	to		
13	3200		
.	to	Identification and optional information	FILE HEADER (CONT.)
	3600		
15	– 3840	none (VAX spare)	
16	1 – 240	Trace header first transient	TRACE HEADER
17	256	Data of first transient	TRACE DATA
.	to		
32(48)	8448(16540)		
33(49)	cont.	Trace header of second transient	TRACE HEADER
34(50)		Data of second transient	TRACE DATA
.	.		
49(81)			
.		continued until last transient	.
End of file			

Fig.A2.1: Graphic illustration of the file records structure of LOTEM SEGY standard data files. Here the implementation for the VAX is shown (after Engels, pers. comm.).

FILE HEADER

The bytes positions 0001–3200 hold a 40x80–byte card image file header as summarized in table A2.3. In this card image file header the acquisition comments are written in the byte positions 0001–1520 and the processing comments in byte position 1521–3200. The above is contained as mask in block data for read and write. Internal read/write will have to fill the empty places. For the DEMS IV system, the acquisition comments (WDUM(70–100)) are written in the file header beginning with byte 894 and the processing comments beginning byte position 1537. None of the comments is presently written into the header. Future acquisition systems should do so.

Table A2.3: Card image header mask as defined by the SEGY standard modified for TEM data.

Use free spaces for your entry. The information is best coded in a block data module.

```
------- CARD IMAGE NUMBER
00000000011111111112222222222333333333344444444445555555555666666666677777777778
12345678901234567890123456789012345678901234567890123456789012345678901234567890
--------------------------------------------------------------------------------
C 1 CLIENT                 COMPANY                    CREW NO
C 2 LINE        AREA                  MAP ID
C 3 REEL NO        DAY-START OF REEL    YEAR     OBSERVER
C 4 INSTRUMENT: MFG          MODEL          SERIAL NO
C 5
C 6 SAMPLE INTERVAL       SAMPLES/TRACE      BITS/IN     BYTES/SAMPLE
C 7 RECORDING FORMAT       FORMAT THIS REEL        MEASUREMENT SYSTEM
C 8 SAMPLE CODE: FLOATING PT     FIXED PT     FIXED PT-GAIN     CORRELATED
C 9 GAIN TYPE: FIXED     BINARY    FLOATING POINT     OTHER
C10 FILTERS: ALIAS     HZ  NOTCH     HZ  BAND    -     HZ  SLOPE   -     DB/OCT
C11 SOURCE: TYPE          LENGTH          BEARING          AREA
C12 COMMENTS:
C13
C14
C15
C16
C17
C18
C19
C20 PROCESSING:
C21
C22
C23       !! THE CARD IMAGE IS AVAILABLE THROUGH COMMON!!!
C24
C25
C26
C27
C28
C29
C30
C31
C32
C33
C34
C35
C36
C37
C38
C39
C40
```

SEGY uses header bytes 3261 to 3600 as optional information. The file header should contain only the information which is not essential for the processing of a trace. If the header is needed, the information in these bytes should also be in the trace header. There must be an IBG (Interblock Gap) after byte 3200. Table A2.4 shows the definition of the file header sorted by bytes.

The coordinates given in meters cannot be represented as INTEGER*2, because their values are too large when using European GKK/UTM coordinates with 6 digits. This difference is presently under investigation.

☞ NOTE

There must be an IBG between the card image header and the optional header. Otherwise IGBs are ONLY between traces. This means data and trace header are NOT separated (see SEG Digital Tape Standards, SEG 1980; Barry et al, 1973).

☞ NOTE

Variables are I*2 or I*4, except for the card image header. Bold face text gives important comments not directly part of the header explanation. RFU means that this variable is presently not used but already reserved for future use. Do not use these variables for your implementation. *Not used, dummy* marks presently unassigned bytes in the header. You may use these bytes for your specific implementation. Reference to LOPS variables is only given when these are different. generally for simplicity it is tried to keep the same variable names as in SEGY.

Table A.2.4: File header record and byte definition for TEM data adopting SEGY standard.

VAX REC.	byte position	SEGY variable	variable type	LOPS present variable name	LOPS meaning
1	0001-0256	C		CARDIMAGE(0001-0256)	--
2	0257-0512	A		CARDIMAGE(0257-0512)	--
3	0513-0768	R		CARDIMAGE(0513-0768)	--
4	0769-1024	D		CARDIMAGE(0769-1024)	WDUM(70-100)
			acquisition comments at byte 894-925		
5	1025-1280	I		CARDIMAGE(1025-1280)	--
6	1281-1536	M		CARDIMAGE(1281-1536)	PROCC(1-200)
			processing comments, 1537 up		
7	1537-1792	A		CARDIMAGE(1537-1792)	--
8	1793-2048	G		CARDIMAGE(1793-2048)	--
9	2049-2304	E		CARDIMAGE(2049-2304)	--
10	2305-2560	H		CARDIMAGE(2305-2560)	--
11	2561-2816	E		CARDIMAGE(2561-2816)	--
12	2817-3072	AD		CARDIMAGE(2817-3072)	--
13	3073-3200	ER		CARDIMAGE(3073-3200)	--
		HERE MUST BE AN IBG (interblock gap) on magnetic tapes			
	3201	SURNUM	I*4		survey id number. i.e. 8601. year,sequential.
	3205	LINENO	I*4		line number.
	3209	REELNO	I*4		tape reel number.
	3213	TRAREC	I*2		no. of traces per record.
	3215	AUXREC	I*2	IFILSOU	source code.
	3217	DT	I*2		srate in micro s. this reel.
	3219	IDELTO	I*2		srate in micro s. original.
	3221	NDAT	I*2		no. of samples per trace.
	3223	NDATO	I*2		no. of samples per trace. original.
	3225	SCODE	I*2		data sample format (i.e. 1=R*4. 2=I*4. 3=I*2. 4= other).
	3227	CDPFOLD	I*2		*not used. dummy*
	3229	TRACSOR	I*2		*trace sorting. RFU*
	3231	ISTACK	I*2		no. of sums per trace.
		3233-3248	**reserved for mixed frequency EM systems**		
	3233	SFSTART	I*2		*sweep frequency at start. RFU*

3235	SFEND	I*2		*sweep frequency at end, RFU*
3237	SWLENG	I*2		*sweep length in ms, RFU*
3239	SWCODE	I*2		*sweep type code, RFU*
3241	SWTRACE	I*2		*sweep trace number, RFU*
3243	TAPSTAR	I*2		*taper length at start, RFU*
3245	TAPEND	I*2		*taper length at end, RFU*
3247	TAPTYPE	I*2		*taper type code, RFU*
3249	ICORR	I*2		*correlated traces, RFU (0=yes; 1=no)*
3251	IGAIN	I*2		gain recovery (0=yes;1=no).
3253	AMPREC	I*2		*not used, dummy*
3255	IMEAS	I*2		meas. unit: 0=meters,1=feet.
3257	IMPSIG	I*2		*not used, dummy*
3259	VIBCOD	I*2		*not used, dummy*

SEGY optional file header 3261-3600

3261	—	I*2	ISTYPE	type of survey: 1=seismic (def.); 2=radar; 3=LOTEM.
3263	—	I*2	ITIMSC	time scale: 1=pico s; 2=nano s; 3=micro s (default); 4= milli s; 5=sec.
3265	—	I*2	ITYPREC	type of recording: 1=finite length (def.); 2= continuously.
3267	—		—	*not used, dummy*
3269	—		—	*not used, dummy*
3271	—		—	*not used, dummy*
3273	—	I*2	ITLEN	trans.length (see byte 3263).
3275	—	I*2	ITLEAD	lead time (see byte 3263).
3277	—	I*2	IVPERD	micro volts/division of ADC
3279	—	I*2	TRIGPOL	trigger reference polarity (1=positive, 2=negative, 0=undefined).
3281	—	I*2	ISRCELE	elev. of source center (see byte 3255).
3283	—	I*2	ISRCLEN	transmitter length (see byte 3255).
3285	—	I*2	ICURREN	source current in Amperes.
3287	—	I*4	JGKK(1)	transmitter coor. EAST E1.
3291	—	I*4	JGKK(2)	transmitter coor. NORTH E1.
3295	—	I*4	JGKK(3)	transmitter coor. EAST E2.
3299	—	I*4	JGKK(4)	transmitter coor. NORTH E2.
3303	—	I*4	JGKK(5)	receiver coor. EAST.
3307	—	I*4	JGKK(6)	receiver coor. NORTH.
3311	—	I*4	JXCOOR	x-coor. of receiver ref.
3315	—	I*4	JYCOOR	y-coor. of receiver ref.
3319	—	I*4	JZCOOR	elevation of rec. ref (see byte 3255).
3323	—	I*2	IRECREF	receiver reference (i.e. n= n-th receiver in spread).
3325	—	I*2	—	*not used, dummy*
3327	—	I*2	—	*not used, dummy*

Notch filter settings bytes 3329-3360 (0=out; 1=in)

14	3329	—	I*2	IA50S1	amplifier 50 Hz sett. 1.
	3331	—	I*2	IA50S2	amplifier 50 Hz sett. 2.
	3333	—	I*2	IA50S3	amplifier 50 Hz sett. 3.
	3335	—	I*2	IA50S4	amplifier 50 Hz sett. 4.
	3337	—	I*2	IA50S5	amplifier 50 Hz sett. 5.
	3339	—	I*2	IP50S1	preamplifier 50 Hz sett. 1.
	3341	—	I*2	IP50S2	preamplifier 50 Hz sett. 2.
	3343	—	I*2	IP50S3	preamplifier 50 Hz sett. 3.
	3345	—	I*2	IA16S1	amplifier 16 2/3 Hz sett. 1.
	3347	—	I*2	IA16S2	amplifier 16 2/3 Hz sett. 2.
	3349	—	I*2	IA16S3	amplifier 16 2/3 Hz sett. 3.
	3351	—	I*2	IA16S4	amplifier 16 2/3 Hz sett. 4.
	3353	—	I*2	IA16S5	amplifier 16 2/3 Hz sett. 5.
	3355	—	I*2	IP16S1	preamplifier 16 2/3 Hz.
	3357	—	I*2	IP16S2	preamplifier 16 2/3 Hz.

	3359	—	I*2	IP16S3	preamplifier 16 2/3 Hz.
	3361	—	I*2	ILAMP	lowpass frequency amplifier.
	3363	—	I*2	ILPAMP	lowpass frequency preamp.
	3365	—	I*2	IAGAIN	amplifier gain setting.
	3367	—	I*2	IPGAIN	preamplifier gain setting.
	3369	—	I*2	YEAR	year file created.
	3371	—	I*2	MONTH	month file created., 0= sequential days in 3373.
	3373	—	I*2	DAY	day file created.
	3375	—	I*2	HOUR	hour file created.
	3377	—	I*2	MINUTE	minute file created.
	3379	—	I*2	SECOND	second file created.
	3381	—	I*2	ITIMBA	time base: 1=local: 2=GMT: 3=other.
	3383	—	I*2	ISPEC	spectrum switch: 1=spectra. 0=raw data.
	3385	—	I*2	NSTACK	number of traces in this file (proc.)
	3387	—	I*2	IASTACK	average number of stacks.
	3389	—	I*2	MINSTK	min. no. of stacks (proc.)
	3391	—	I*2	MAXSTK	max. no. of stacks (proc.)
	3393-3584	—		—	*not used, dummy*
15	3585-3600	—		—	*not used, dummy*
	record filler				*VAX spare, non SEGY*

TRACE HEADER STRUCTURE

Suppose the number of data points is NDAT. The record pointer of the n–th trace (MOLD for LOPS) is calculated by:

$$REC=(MOLD-1)*(NDAT/64+1)+16$$

where 64 is the number of REAL data points per record. We have defined the record length on the VAX to be 256 bytes long (equals 64 real numbers). The value of 1 is added for the trace header of each trace and the value of 16 is added for the file header of each file. The definition of the trace header is given in table A2.5.

The time base is referred to the default onset (IONSET) which is in most instances 20% of the total trace length. Once the onset (IONSET in sample number) is known, this is set to be zero reference time. The pretrigger and delay time is referenced to this time (and corresponding sample number).

(All values are either INTEGER*4 or INTEGER*2 defined by its length.)

Table A.2.5: Trace header byte allocation table for TEM data adopting SEGY standard.

byte pos. length	SEGY variable	LOPS present value	LOPS meaning
001 4	TRACNO		trace number within line, numbers increase for add. reels.
005 4	TRAREEL		trace number in reel.
009 4	ORIREEL	JORREC	original record number.
013 4	TRAORI		trace number within original record.
017 4	ENEPT	ISRCNUM	source point number.
021 4	CDPENS		*gather number, RFU*
025 4	TRENSNM		*num. code for current trace within one gather, RFU.*
029 2	TRACID		trace id: 01-31 reserved for seismic data, 32=unknown, 33=LOTEM data raw,34=system response, 35=LOTEM data stacked, 36=LOTEM logarithmic (equidist.) data.
031 2	NSTACK		number of stacked traces yielding this one.
033 2	HORSTAC		*number of horizontal stacks yielding this trace, RFU*
035 2	USAGE		data code: 1=production, 2=test.
037 4	OFFSET		source to receiver (this trace) offset (ref. coor. syst.)
041 4	JZCOOR		receiver elevation (depending on byte 69-70).
045 4	SRCELE		source center elevation (depending on byte 69-70).
049 4	SRCDEP		source length (depending on byte 69-70).
053 4	DATELE		receiver datum elevation (depending on byte 69-70).
057 4	DATSRCE		source center datum elevation (depending on byte 69-70).
061 4	WATDEPS	JCURREN	source current in amperes.
065 4	WATDEPG		*not used, dummy*
069 2	SCALAR		scale factor for bytes 41-60 (negative = divisor).
071 2	SCALAR1		scale factor for bytes 73-88 (negative = divisor).
073 4	SRCCOX	JXREF	receiver x coordinate w.r.t. reference receiver (depending on byte 71-72).
077 4	SRCCOY	JYREF	receiver y coordinate w.r.t. reference receiver in m (depending on byte 71-72).
081 4	GRPCORX		receiver x-coordinate (depending on byte 71-72).
085 4	GRPCORY		receiver y-coordinate (depending on byte 71-72).
089 2	CORUNI		coordinate system used: 1=length, 2= seconds of arc.
091 2	WEAVEL		*not used, dummy*
093 2	SUBVEL		*not used, dummy*
095 2	UPTSRC		*not used, dummy*
097 2	UPTGRP		*not used, dummy*
099 2	SRCSTA		*not used, dummy*
101 2	GRPSTA		*not used, dummy*
103 2	TOTSTAT	IAMEXP	amplifier gain setting (if binary, exponent only).
105 2	LAGTIA	IONSET	number of samples before onset.
107 2	LAGTIB	IPRETRIG	pretrigger in base time scale.
109 2	DELAYT		delay time in base time scale of synchronization trigger.
111 2	MUTETS		*mute time start, RFU*
113 2	MUTETE		*mute time end, RFU*
115 2	NDAT		number of data points (samples) in this trace.
117 2	DT		sample interval (depending on file header byte 3263-3264).
119 2	GAININS		gain type: 1=fixed; 2=binary; 3=IFP.
121 2	GAINCON	IPAEXP	preamplifier gain setting (if binary, exponent only).
123 2	INIGAIN		initial gain in db.
125 2	ICORR		*correlated traces (0=yes: 1=no), RFU*
	byte 127-140 reserved for mixed frequency EM systems		
127 2	SFSTART		*sweep frequency at start, RFU*
129 2	SFEND		*sweep frequency at end, RFU*
131 2	SWLENG		*sweep length in ms, RFU*
133 2	SWCODE		*sweep type code, RFU*
135 2	TAPSTART		*taper length at start, RFU*
137 2	TAPEND		*taper length at end, RFU*
139 2	TAPTYPE		*taper type code, RFU*

141	2	ALIASF		alias filter frequency (0=not used).
143	2	ALIASSL		alias filter slope (0=passive).
145	2	NOTCHF		notch filter base frequency (0=not used).
147	2	NOTCHSL		notch filter slope (0=passive).
149	2	HIPASS		high pass filter frequency (0=not used).
151	2	LOPASS		low pass filter frequency (0=not used), amplifier.
153	2	HIGHPSL		high pass filter slope (0=passive).
155	2	LOWPASL		low pass filter slope (0=passive), amplifier.
157	2	YEAR		year of data recording.
159	2	DAY		day of data recording.
161	2	HOUR		hour of data recording.
163	2	MINUTE		minute of hour recording.
165	2	SECOND		second of minute recording.
167	2	TIMBASE		time basis: 1=local, 2=GMT, 3=other.
169	2	TRAWEI		trace weighting factor.
171	2	GRPROL	MONTH	month of data recording (0=sequential days in 159).
173	2	GRPONE		station no. of first trace in original record.
175	2	GRPLAST		station no. of last trace in original record.
177	2	NUMGRP		*number of receivers per group. RFU*
179	2	OVRTRAV	IRCVCD	receiver coil code.

SEGY optional trace header 181–240

181	2	—	ILOPAPA	low pass filter frequency (0=not used), preamplifier.
183	2	—	ILOPAPS	low pass filter slope (0=passive), preamplifier.
185	2	—	JSTATI	station number increment, used when more than one transient occurs at the same site (multiple transmitters, tests, etc.).
187	2	—	IRECSTAT	receiver station number.
189	2	—	FFID	field file identification
191	2	—	IA16S(1)	amplifier 16 2/3 Hz setting 1.
193	2	—	IA16S(2)	amplifier 16 2/3 Hz setting 2.
195	2	—	IA16S(3)	amplifier 16 2/3 Hz setting 3.
197	2	—	IA16S(4)	amplifier 16 2/3 Hz setting 4.
199	2	—	IA16S(5)	amplifier 16 2/3 Hz setting 5.
201	2	—	IREMTOT	total number of remote units.
203	2	—	NCHAN	nos of receivers per spread.
205	2	—	IA50S(1)	amplifier 50Hz setting 1.
207	2	—	NFIRST	first channel number.
209	2	—	IA50S(2)	amplifier 50Hz setting 2.
211	4	—	IEDL	length of dipole, if 215 =1 or 2 or receiver equivalent area in square units otherwise.
215	2	—	IFIELD	0-HZ, 1-EX, 2-EY, 3-HX, 4-HY.
217	2	—	ISYSTEM	receiver system code.
219	2	—	ICHAN	recording channel in use.
221	2	—	IPHYSADD	physical address of remote unit.
223	2	—	IA50S(3)	amplifier 50Hz setting 3.
225	2	—	IA50S(4)	amplifier 50Hz setting 4.
227	2	—	IA50S(5)	amplifier 50Hz setting 5.
229	2	—	IP50S(1)	preamplifier 50 Hz setting 1.
231	2	—	IP50S(2)	preamplifier 50 Hz setting 2.
233	2	—	IP50S(3)	preamplifier 50 Hz setting 3.
235	2	—	IP16S(1)	preamplifier 16 2/3 Hz setting 1.
237	2	—	IP16S(2)	preamplifier 16 2/3 Hz setting 2.
239	2	—	IP16S(3)	preamplifier 16 2/3 Hz setting 3.

KMS Technologies – KJT Enterprises, Inc.

6420 Richmond Ave., Suite 610
Houston, Texas 77057, USA
Tel: +1 713.532.8144
Fax: +1 832.204.8418

www.KMSTechnologies.com