Converting analog interpretive data to digital formats for use in database and GIS applications

By James G. Flocks

Report Series 2004-1070

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia 2004

For product and ordering information:
World Wide Web: *http://www.usgs.gov/pubprod*
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment:
World Wide Web: *http://www.usgs.gov*
Telephone: 1-888-ASK-USGS

Suggested citation:
Flocks, J.G., 2004, Converting analog interpretative data to digital formats for use in database and GIS applications: U.S. Geological Survey Open-File Report, 2004-1070.

Contents

Figures

Conversion Factors

Inch/Pound to SI

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
mile, nautical (nmi)	1.852	kilometer (km)
yard (yd)	0.9144	meter (m)
	Area	
acre	4,047	square meter (m^2)
acre	0.4047	hectare (ha)
acre	0.4047	square hectometer (hm^2)
acre	0.004047	square kilometer (km^2)
square foot (ft^2)	929.0	square centimeter (cm^2)
square foot (ft^2)	0.09290	square meter (m^2)
square inch (in^2)	6.452	square centimeter (cm^2)
section (640 acres or 1 square mile)	259.0	square hectometer (hm^2)
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Volume	
barrel (bbl), (petroleum, 1 barrel=42 gal)	0.1590	cubic meter (m^3)
ounce, fluid (fl. oz)	0.02957	liter (L)
pint (pt)	0.4732	liter (L)
quart (qt)	0.9464	liter (L)
gallon (gal)	3.785	liter (L)
gallon (gal)	0.003785	cubic meter (m^3)
gallon (gal)	3.785	cubic decimeter (dm^3)
million gallons (Mgal)	3,785	cubic meter (m^3)
cubic inch (in^3)	16.39	cubic centimeter (cm^3)
cubic inch (in^3)	0.01639	cubic decimeter (dm^3)
cubic inch (in^3)	0.01639	liter (L)
cubic foot (ft^3)	28.32	cubic decimeter (dm^3)
cubic foot (ft^3)	0.02832	cubic meter (m^3)

cubic yard (yd^3)	0.7646	cubic meter (m^3)
cubic mile (mi^3)	4.168	cubic kilometer (km^3)
acre-foot (acre-ft)	1,233	cubic meter (m^3)
acre-foot (acre-ft)	0.001233	cubic hectometer (hm^3)
	Flow rate	
acre-foot per day (acre-ft/d)	0.01427	cubic meter per second (m^3/s)
acre-foot per year (acre-ft/yr)	1,233	cubic meter per year (m^3/yr)
acre-foot per year (acre-ft/yr)	0.001233	cubic hectometer per year (hm^3/yr)
foot per second (ft/s)	0.3048	meter per second (m/s)
foot per minute (ft/min)	0.3048	meter per minute (m/min)
foot per hour (ft/hr)	0.3048	meter per hour (m/hr)
foot per day (ft/d)	0.3048	meter per day (m/d)
foot per year (ft/yr)	0.3048	meter per year (m/yr)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
cubic foot per second per square mile [(ft^3/s)/mi^2]	0.01093	cubic meter per second per square kilometer [(m^3/s)/km^2]
cubic foot per day (ft^3/d)	0.02832	cubic meter per day (m^3/d)
gallon per minute (gal/min)	0.06309	liter per second (L/s)
gallon per day (gal/d)	0.003785	cubic meter per day (m^3/d)
gallon per day per square mile [(gal/d)/mi^2]	0.001461	cubic meter per day per square kilometer [(m^3/d)/km^2]
million gallons per day (Mgal/d)	0.04381	cubic meter per second (m^3/s)
million gallons per day per square mile [(Mgal/d)/mi^2]	1,461	cubic meter per day per square kilometer [(m^3/d)/km^2]
inch per hour (in/h)	0 .0254	meter per hour (m/h)
inch per year (in/yr)	25.4	millimeter per year (mm/yr)
mile per hour (mi/h)	1.609	kilometer per hour (km/h)
	Mass	
ounce, avoirdupois (oz)	28.35	gram (g)
pound, avoirdupois (lb)	0.4536	kilogram (kg)
ton, short (2,000 lb)	0.9072	megagram (Mg)
ton, long (2,240 lb)	1.016	megagram (Mg)
ton per day (ton/d)	0.9072	metric ton per day
ton per day (ton/d)	0.9072	megagram per day (Mg/d)
ton per day per square mile [(ton/d)/mi^2]	0.3503	megagram per day per square kilometer [(Mg/d)/km^2]
ton per year (ton/yr)	0.9072	megagram per year (Mg/yr)
ton per year (ton/yr)	0.9072	metric ton per year

Pressure		
atmosphere, standard (atm)	101.3	kilopascal (kPa)
bar	100	kilopascal (kPa)
inch of mercury at 60ºF (in Hg)	3.377	kilopascal (kPa)
pound-force per square inch (lbf/in^2)	6.895	kilopascal (kPa)
pound per square foot (lb/ft^2)	0.04788	kilopascal (kPa)
pound per square inch (lb/in^2)	6.895	kilopascal (kPa)
Density		
pound per cubic foot (lb/ft^3)	16.02	kilogram per cubic meter (kg/m^3)
pound per cubic foot (lb/ft^3)	0.01602	gram per cubic centimeter (g/cm^3)
Energy		
kilowatthour (kWh)	3,600,000	joule (J)
Radioactivity		
picocurie per liter (pCi/L)	0.037	becquerel per liter (Bq/L)
Specific capacity		
gallon per minute per foot [(gal/min)/ft)]	0.2070	liter per second per meter [(L/s)/m]
Hydraulic conductivity		
foot per day (ft/d)	0.3048	meter per day (m/d)
Hydraulic gradient		
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)
Transmissivity*		
foot squared per day (ft^2/d)	0.09290	meter squared per day (m^2/d)
Application rate		
pounds per acre per year [(lb/acre)/yr]	1.121	kilograms per hectare per year [(kg/ha)/yr]
Leakance		
foot per day per foot [(ft/d)/ft]	1	meter per day per meter
inch per year per foot [(in/yr)/ft]	83.33	millimeter per year per meter [(mm/yr)/m]

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

°F=(1.8×°C)+32

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

°C=(°F-32)/1.8

Vertical coordinate information is referenced to the insert datum name (and abbreviation) here for instance, “North American Vertical Datum of 1988 (NAVD 88).”

Horizontal coordinate information is referenced to the insert datum name (and abbreviation) here for instance, “North American Datum of 1983 (NAD 83).”

Altitude, as used in this report, refers to distance above the vertical datum.

*Transmissivity: The standard unit for transmissivity is cubic foot per day per square foot times foot of aquifer thickness $[(ft^3/d)/ft^2]ft$. In this report, the mathematically reduced form, foot squared per day (ft^2/d), is used for convenience.

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25 °C).

Concentrations of chemical constituents in water are given either in milligrams per liter (mg/L) or micrograms per liter (µg/L).

NOTE TO USGS USERS: Use of hectare (ha) as an alternative name for square hectometer (hm^2) is restricted to the measurement of small land or water areas. Use of liter (L) as a special name for cubic decimeter (dm^3) is restricted to the measurement of liquids and gases. No prefix other than milli should be used with liter. Metric ton (t) as a name for megagram (Mg) should be restricted to commercial usage, and no prefixes should be used with it.

SI to Inch/Pound

Multiply	By	To obtain
	Length	
centimeter (cm)	0.3937	inch (in.)
millimeter (mm)	0.03937	inch (in.)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
kilometer (km)	0.5400	mile, nautical (nmi)
meter (m)	1.094	yard (yd)
	Area	
square meter (m^2)	0.0002471	acre
hectare (ha)	2.471	acre
square hectometer (hm^2)	2.471	acre
square kilometer (km^2)	247.1	acre
square centimeter (cm^2)	0.001076	square foot (ft^2)
square meter (m^2)	10.76	square foot (ft^2)
square centimeter (cm^2)	0.1550	square inch (ft^2)
square hectometer (hm^2)	0.003861	section (640 acres or 1 square mile)
hectare (ha)	0.003861	square mile (mi^2)
square kilometer (km^2)	0.3861	square mile (mi^2)
	Volume	
cubic meter (m^3)	6.290	barrel (petroleum, 1 barrel = 42 gal)
liter (L)	33.82	ounce, fluid (fl. oz)
liter (L)	2.113	pint (pt)
liter (L)	1.057	quart (qt)
liter (L)	0.2642	gallon (gal)
cubic meter (m^3)	264.2	gallon (gal)
cubic decimeter (dm^3)	0.2642	gallon (gal)
cubic meter (m^3)	0.0002642	million gallons (Mgal)
cubic centimeter (cm^3)	0.06102	cubic inch (in^3)
cubic decimeter (dm^3)	61.02	cubic inch (in^3)
liter (L)	61.02	cubic inch (in^3)
cubic decimeter (dm^3)	0.03531	cubic foot (ft^3)
cubic meter (m^3)	35.31	cubic foot (ft^3)
cubic meter (m^3)	1.308	cubic yard (yd^3)
cubic kilometer (km^3)	0.2399	cubic mile (mi^3)
cubic meter (m^3)	0.0008107	acre-foot (acre-ft)
cubic hectometer (hm^3)	810.7	acre-foot (acre-ft)

Flow rate		
cubic meter per second (m^3/s)	70.07	acre-foot per day (acre-ft/d)
cubic meter per year (m^3/yr)	0.000811	acre-foot per year (acre-ft/yr)
cubic hectometer per year (hm^3/yr)	811.03	acre-foot per year (acre-ft/yr)
meter per second (m/s)	3.281	foot per second (ft/s)
meter per minute (m/min)	3.281	foot per minute (ft/min)
meter per hour (m/hr)	3.281	foot per hour (ft/hr)
meter per day (m/d)	3.281	foot per day (ft/d)
meter per year (m/yr)	3.281	foot per year ft/yr)
cubic meter per second (m^3/s)	35.31	cubic foot per second (ft^3/s)
cubic meter per second per square kilometer [(m^3/s)/km^2]	91.49	cubic foot per second per square mile [(ft^3/s)/mi^2]
cubic meter per day (m^3/d)	35.31	cubic foot per day (ft^3/d)
liter per second (L/s)	15.85	gallon per minute (gal/min)
cubic meter per day (m^3/d)	264.2	gallon per day (gal/d)
cubic meter per day per square kilometer [(m^3/d)/km^2]	684.28	gallon per day per square mile [(gal/d)/mi^2]
cubic meter per second (m^3/s)	22.83	million gallons per day (Mgal/d)
cubic meter per day per square kilometer [(m^3/d)/km^2]	0.0006844	million gallons per day per square mile [(Mgal/d)/mi^2]
cubic meter per hour (m^3/h)	39.37	inch per hour (in/h)
millimeter per year (mm/yr)	0.03937	inch per year (in/yr)
kilometer per hour (km/h)	0.6214	mile per hour (mi/h)
Mass		
gram (g)	0.03527	ounce, avoirdupois (oz)
kilogram (kg)	2.205	pound avoirdupois (lb)
megagram (Mg)	1.102	ton, short (2,000 lb)
megagram (Mg)	0.9842	ton, long (2,240 lb)
metric ton per day	1.102	ton per day (ton/d)
megagram per day (Mg/d)	1.102	ton per day (ton/d)
megagram per day per square kilometer [(Mg/d)/km^2]	2.8547	ton per day per square mile [(ton/d)/mi^2]
megagram per year (Mg/yr)	1.102	ton per year (ton/yr)
metric ton per year	1.102	ton per year (ton/yr)
Pressure		
kilopascal (kPa)	0.009869	atmosphere, standard (atm)
kilopascal (kPa)	0.01	bar
kilopascal (kPa)	0.2961	inch of mercury at 60°F (in Hg)
kilopascal (kPa)	0.1450	pound-force per inch (lbf/in)

kilopascal (kPa)	20.88	pound per square foot (lb/ft^2)
kilopascal (kPa)	0.1450	pound per square inch (lb/ft^2)
Density		
kilogram per cubic meter (kg/m^3)	0.06242	pound per cubic foot (lb/ft^3)
gram per cubic centimeter (g/cm^3)	62.4220	pound per cubic foot (lb/ft^3)
Energy		
joule (J)	0.0000002	kilowatthour (kWh)
Radioactivity		
becquerel per liter (Bq/L)	27.027	picocurie per liter (pCi/L)
Specific capacity		
liter per second per meter [(L/s)/m]	4.831	gallon per minute per foot [(gal/min)/ft]
Hydraulic conductivity		
meter per day (m/d)	3.281	foot per day (ft/d)
Hydraulic gradient		
meter per kilometer (m/km)	5.27983	foot per mile (ft/mi)
Transmissivity*		
meter squared per day (m^2/d)	10.76	foot squared per day (ft^2/d)
Application rate		
kilograms per hectare per year [(kg/ha)/yr]	0.8921	pounds per acre per year [(lb/acre)/yr]
Leakance		
meter per day per meter [(m/d)/m]	1	foot per day per foot [(ft/d)/ft]
millimeter per year per meter [(mm/yr)/m]	0.012	inch per year per foot [(in/yr)/ft]

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

°F=(1.8×°C)+32

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

°C=(°F-32)/1.8

Vertical coordinate information is referenced to the insert datum name (and abbreviation) here, for instance, “North American Vertical Datum of 1988 (NAVD 88)”

Horizontal coordinate information is referenced to the insert datum name (and abbreviation) here, for instance, “North American Datum of 1983 (NAD 83)”

Altitude, as used in this report, refers to distance above the vertical datum.

*Transmissivity: The standard unit for transmissivity is cubic foot per day per square foot times foot of aquifer thickness $[(ft^3/d)/ft^2]ft$. In this report, the mathematically reduced form, foot squared per day (ft^2/d), is used for convenience.

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25°C).

Concentrations of chemical constituents in water are given either in milligrams per liter (mg/L) or micrograms per liter (µg/L).

NOTE TO USGS USERS: Use of hectare (ha) as an alternative name for square hectometer (hm^2) is restricted to the measurement of small land or water areas. Use of liter (L) as a special name for cubic decimeter (dm^3) is restricted to the measurement of liquids and gases. No prefix other than milli should be used with liter. Metric ton (t) as a name for megagram (Mg) should be restricted to commercial usage, and no prefixes should be used with it.

Converting analog interpretive data to digital formats for use in database and GIS applications

By James G. Flocks

Abstract

There is a growing need by researchers and managers for comprehensive and unified nationwide datasets of scientific data. These datasets must be in a digital format that is easily accessible using database and GIS applications, providing the user with access to a wide variety of current and historical information. Although most data currently being collected by scientists are already in a digital format, there is still a large repository of information in the literature and paper archive. Converting this information into a format accessible by computer applications is typically very difficult and can result in loss of data. However, since scientific data are commonly collected in a repetitious, concise matter (i.e., forms, tables, graphs, etc.), these data can be recovered digitally by using a conversion process that relates the position of an attribute in two-dimensional space to the information that the attribute signifies. For example, if a table contains a certain piece of information in a specific row and column, then the space that the row and column occupies becomes an index of that information. An index key is used to identify the relation between the physical location of the attribute and the information the attribute contains. The conversion process can be achieved rapidly, easily and inexpensively using widely available digitizing and spreadsheet software, and simple programming code.

In the geological sciences, sedimentary character is commonly interpreted from geophysical profiles and descriptions of sediment cores. In the field and laboratory, these interpretations were typically transcribed to paper. The information from these paper archives is still relevant and increasingly important to scientists, engineers and managers to understand geologic processes affecting our environment. Direct scanning of this information produces a raster facsimile of the data, which allows it to be linked to the electronic world. But true integration of the content with database and GIS software as point, vector or text information is commonly lost. Sediment core descriptions and interpretation of geophysical profiles are usually portrayed as lines, curves, symbols and text information. They have vertical and horizontal dimensions associated with depth, category, time, or geographic position. These dimensions are displayed in consistent positions, which can be digitized and converted to a digital format, such as a spreadsheet. Once this data is in a digital, tabulated form it can easily be made available to a wide variety of imaging and data manipulation software for compilation and world-wide dissemination.

Introduction

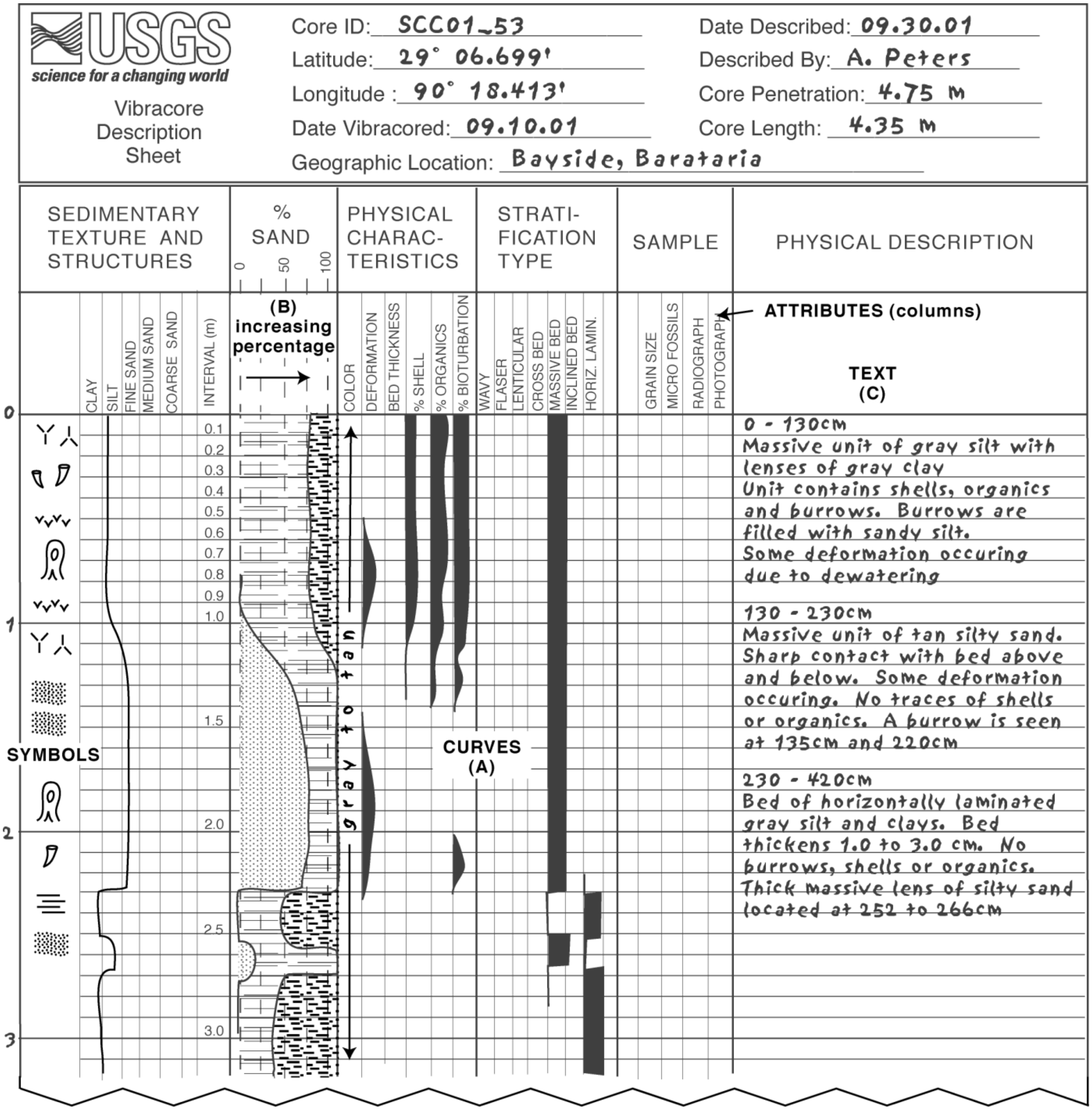
In the geological sciences subsurface features are typically characterized using sediment sampling and geophysical techniques. Data generated from these techniques, as well as interpretations of the data, are commonly documented on paper copies. Two examples of paper copies commonly found in research institutions are sediment core description sheets and seismic profiles.

Increasingly, such data is both generated and represented in digital format. The obvious advantage is that the data can then be directly utilized by computer applications to facilitate interpretation and accessibility. However, an abundant amount of data is still being generated or currently exists as archived paper copies. In many cases generating hard copies of interpretative data is still an important and practical part of the analytical process. Computer software has replaced the colored pencil during visual interpretations of seismic profiles, but transcribing visual descriptions of sediment cores directly into a computer is not a practical option to standardized paper templates when working in a core-archive laboratory. Furthermore, paper copies resolve potential difficulties with incompatible or out-dated computer file formats when archiving interpretative data. So the need still exists to convert this data to a digital format; often a difficult task that can result in loss of data. For example, direct digital scanning of the paper copies can produce raster facsimiles of the data, but this seldom provides any quantitative information.

Since a paper copy of any image is essentially a two-dimensional representation, quantitative representations of digitally scanned paper copies can be retained if the scans are subsequently referenced in the two-dimensional plane and this reference is retained in the digital format. A commonly used example of this method is geo-referenced raster images of topographic maps. The maps have been digitally scanned and converted to a format where the constraints of the map (i.e. latitude and longitude) are retained in the digital file (e.g. GEO-TIFF). This reference information can then be accessed by GIS applications and projected over the entire surface of the map. However, vector information that is visually evident on the original map; roads, rivers, etc., require further digitization if they are to be used quantitatively as database layers.

This report presents methods of digitizing paper copies of two specific, commonly used analog representations of geologic data: core description sheets and seismic profiles. The reader should be familiar with these two data formats. Examples of their use can be found widely in coastal stratigraphic studies (Suter and others, 1991; Kindinger and others, 2001). Digitizing converts the information they contain into two-dimensional (X,Y) space. Computer programs can then identify the x-y relationship and extract the associated information as well as incorporate data that otherwise would be lost during other data entry methods. The extracted information is then converted to a tabulated digital format for use in database and GIS applications. Batch mode processing allows for rapid analysis of numerous sheets or profiles, thereby increasing the speed at which the data can be converted.

Interpretations of data in geologic applications are commonly expressed in a consistent format that is familiar to the user and allows for correlation across various geologic investigations. For example, description of sediment cores are often rendered to a form sheet that has various fields necessary to adequately describe most sedimentary characteristics: texture, physical properties, depositional parameters, etc. (Fig. 1). In geophysical surveys, persistent amplitude peaks across a profile may represent stratigraphic boundaries and are commonly mapped using line drawings directly on the printed profile (Fig. 2). These paper copies make it easy to produce rapid and direct interpretations of the geologic data. The hard-copies also provide a stable archive of both the original data and the interpretations. However, the data is available only on-hand, and digitization of the data is required not only to disseminate the information but also to spatially reference the data so that subsurface interpretations can be mapped. Currently, digitizing paper copies is not only time-intensive, but is often selective of the type of data that is retained. For example, information from core description sheets are often hand-typed into a tabulated format, where the rows represent depth or thickness of the stratigraphic unit being described and the columns contain the text from the description part of the sheet (Fig. 1 (C)). Semi-quantitative information such as observed sand percentages are not included unless it is estimated during data entry from the sand-percent curve (Fig. 1 (B)) and included into a column. With seismic profiles, interpretations can be redrawn using computer applications, but because the profiles are commonly very long in one dimension, the extent that is redrawn is often limited. Also, the geographic reference of the profile is seldom retained.



Core ID: SCC01_53

Latitude: 29° 06.699'

Longitude: 90° 18.413'

Date Vibracored: 09.10.01

Geographic Location: Bayside, Barataria

Date Described: 09.30.01

Described By: A. Peters

Core Penetration: 4.75 m

Core Length: 4.35 m

0 - 130cm
Massive unit of gray silt with lenses of gray clay
Unit contains shells, organics and burrows. Burrows are filled with sandy silt.
Some deformation occuring due to dewatering

130 - 230cm
Massive unit of tan silty sand. Sharp contact with bed above and below. Some deformation occuring. No traces of shells or organics. A burrow is seen at 135cm and 220cm

230 - 420cm
Bed of horizontally laminated gray silt and clays. Bed thickens 1.0 to 3.0 cm. No burrows, shells or organics. Thick massive lens of silty sand located at 252 to 266cm

Figure 1. Example portion of core description sheet shows how attributes pertaining to sediment core are described. Data is entered as symbols, curves or text.

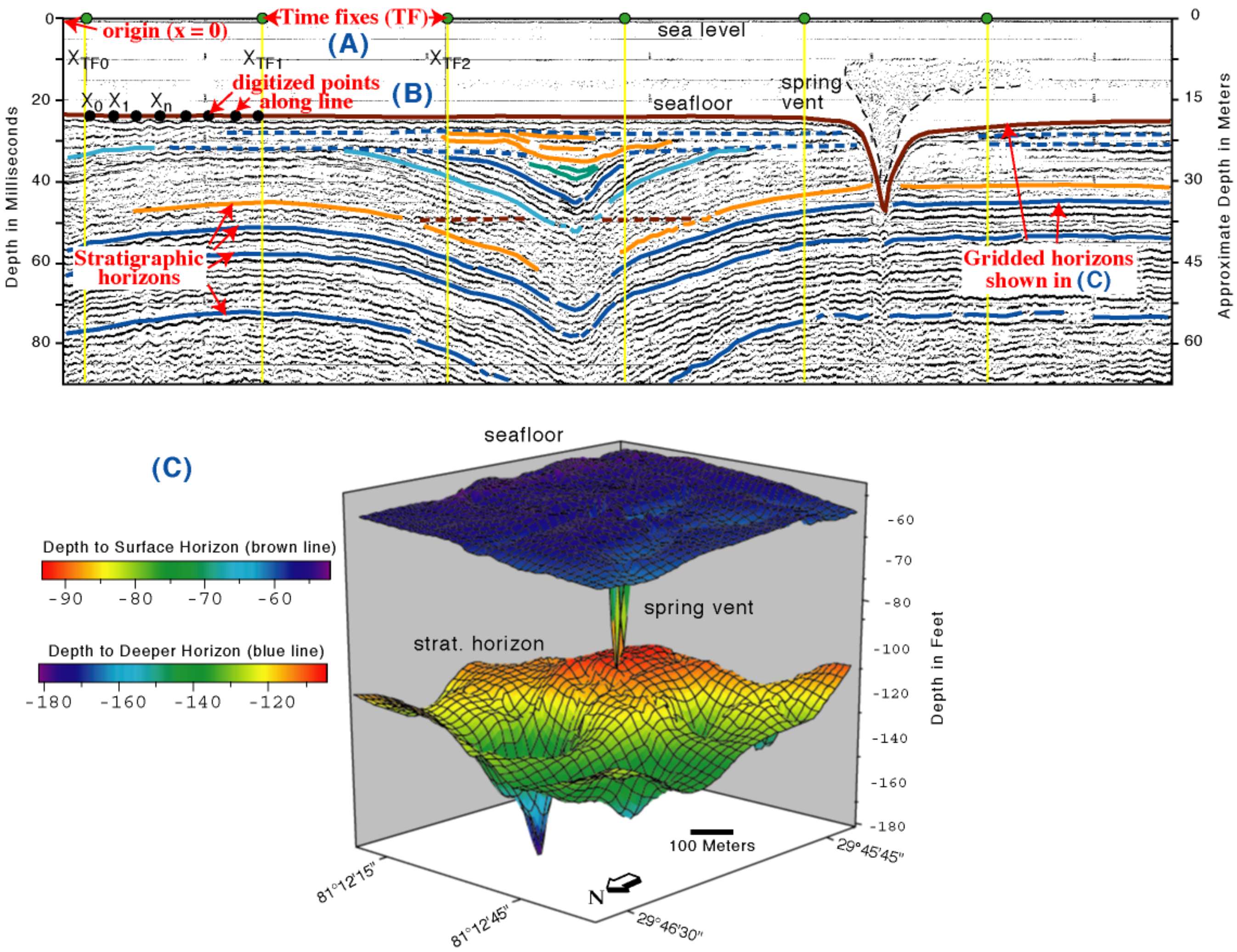


Figure 2. Example of a seismic profile with line-drawing interpretations, illustrating the digitizing process (A, B). Three-dimensional block diagram of digitized output from multiple line interpretations (C).

Although core description sheets and analog seismic profiles represent different kinds of data, they have a common fundamental relation: two-dimensional representations of data. This does not exclude three-dimensional information or other attributes such as text descriptions, it means simply that since they exist on a two dimensional surface, every attribute can be represented with an x and y distance from an origin, regardless of the attribute. The x, y position becomes a place-marker for the attribute, reducing it to a numerical array that can be reconstructed according to a key that outlines the relation. For example, a core description sheet may contain a column that represents the amount of shell material down-core (Fig. 1(A)). Variations in shell content are represented by filling-in the column semi-quantitatively, so that more fill in the column represents increased shell content. This diagram can be expressed as two x-y expressions: the shell percent (x) versus core depth (y); and the column position (x,y) on the paper. Thus, even though the shell content on a core description sheet is not represented numerically, the shell percent relative to core depth can still be recovered by knowing how far from an origin (i.e. the edge of the description area) the shell percent column exists.

On paper copies of seismic profile data, the amplitude traces are overlain by navigational fixes and shot numbers at set intervals (Fig. 2(A)). The distance to these navigational fixes on the paper can be measured, so that geographic relationships can be tagged to interpretations on the profile. Likewise, stratigraphic depth can be calculated relative to the vertical position (y) of the interpretation on the paper. The following sections describe the processes to convert these analog formats to digital files. Although mechanically similar, for clarity the methods for converting analog core description sheets and seismic profile data will be described separately.

Methods

Sediment core description sheets

Direct sampling of the subsurface is the only way to positively identify geologic characteristics such as stratigraphic horizons, textural variability or trace element distributions. Sediment sampling is also used in conjunction with geophysical measurements to ground truth or confirm the geophysical data, which provides a wider spatial coverage. A common method to portray the information acquired from sediment cores is to describe the core-section using specific physical and textural attributes: sand/silt/clay ratios, massive versus laminated stratification, etc. These attributes are estimated, or measured, by the geologist and transcribed onto a form sheet that contains spaces where the presence and magnitude of each attribute down-core can be displayed (Fig. 1). The forms also often contain fields where the visual descriptions of the sediments can be included using commonly accepted terminology, providing a fairly comprehensive description of the sediment core. These form sheets, called core description sheets, provide the basic sedimentologic parameters required to adequately characterize the geology of a study area, and are fairly consistent among research institutions. As a result, core descriptions can be readily utilized for correlations between different data sets and by different research projects.

A common method for converting core description sheets into a digital format involves typing the text descriptions into a spreadsheet, with stratigraphic intervals representing the rows of the spreadsheet. For example, a sediment core from the marine environment may contain 130 centimeters of gray silt containing shells, burrows and organic material (Fig. 1(C)). In a spreadsheet this information may be transcribed as one row containing the stratigraphic interval (0-130cm), a midpoint (65cm) and the subsequent text descriptions (gray silt with shells...). Although converting this information to a spreadsheet has tremendous potential over the hard copy, a lot of information recorded on the core description sheet has not been retained. Additionally, the constituents within the stratigraphic unit are either tied to the unit thickness or are averaged to a single midpoint, losing the higher resolution variations recorded in the hard copy. Spreadsheet data for use in GIS and database applications often contain incremented rows of a constant value, every centimeter for example. This requires that the data from the core description sheet be integrated in user specified intervals (i.e. 1 cm) across the stratigraphic interval, for better correlation with conventional database structure (Fig. 3).

A:

Depth_ start (cm)	Depth_ end (cm)	Depth_ median (cm)	structure	silt	fine_sand	sand%	silt%	clay%	color	deformation %	shell %	organic %	bioturb %	cross_ bed %	massive_ bed %	horiz_ lam %
0	1	0.5				0	73.28	26.72		0.86	45.86	47.59	53.86	0	100	3.61
1	2	1.5				0	73.28	26.72		0.86	45.86	47.59	53.86	0	100	3.61
2	3	2.5				0	73.28	26.72		0.86	45.86	47.59	53.86	0	100	3.61
3	4	3.5				0	73.28	26.72	tan	0.86	46.03	47.59	53.61	0	100	3.61
4	5	4.5	organics			0	73.28	26.72	tan	0.86	46.21	47.44	53.36	0	100	3.61
5	6	5.5	organics			0	73.28	26.72	tan	0.86	46.38	47.29	53.11	0	100	3.61
6	7	6.5	organics	silt		0	73.28	26.72	tan	0.86	46.56	47.15	52.86	0	100	3.61
7	8	7.5	organics	silt		0	73.31	26.69	tan	0.86	46.74	47	52.61	0	100	3.61
8	9	8.5	organics	silt		0	73.34	26.66	tan	0.86	46.91	46.86	52.37	0	100	3.61
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
104	105	104.5	organics		fine_sand	9.29	80.43	10.29	tan	0	45.03	37.28	56.61	0.24	100	2.59
105	106	105.5	organics		fine_sand	10.32	80.22	9.46	tan	0	44.16	35.73	53.74	0.24	100	2.55
106	107	106.5	organics		fine_sand	11.35	80.02	8.63	tan	0	43.29	34.18	50.88	0.24	100	2.52
107	108	107.5	organics		fine_sand	12.38	79.81	7.81	tan	0	42.42	32.63	48.02	0.25	100	2.49
108	109	108.5	organics		fine_sand	13.41	79.43	7.16	tan	0	41.55	31.08	45.15	0.25	100	2.46
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
230	231	230.5	planar_beds	silt		0	66.46	33.54	tan	20.34	6.43	0	6.46	0.68	20.58	100
231	232	231.5	planar_beds	silt		0	66.46	33.54	tan	17.46	6.43	0	5.11	0.69	10.29	100
232	233	232.5	planar_beds	silt		0	66.46	33.54	tan	14.58	6.43	0	3.77	0.69	0	100
233	234	233.5	planar_beds	silt		0	66.46	33.54	tan	11.71	6.43	0	3.89	0.69	0.1	100

B

B:

strat_type	color	sed_texture	phys_char.	Additional_Text_Attributes	Date	Time (GMT)	Proj_ id	Year	Core_ id
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
	gray	clay silt	%shell %organic %bioturb deformation	sandy-silt filled burrows	10-Sep	10:53	SCC	2001	53
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
horiz_lam	gray	clay silt		thick lens of silty-sand @ 252 & 266cm	10-Sep	10:53	SCC	2001	53
horiz_lam	gray	clay silt		thick lens of silty-sand @ 252 & 266cm	10-Sep	10:53	SCC	2001	53
horiz_lam	gray	clay silt		thick lens of silty-sand @ 252 & 266cm	10-Sep	10:53	SCC	2001	53
horiz_lam	gray	clay silt		thick lens of silty-sand @ 252 & 266cm	10-Sep	10:53	SCC	2001	53

C

C:

File name	Deg	Min	Deg	Min	W.D. (ft)	Pen_ Length (ft)	Pen_ Corr (ft)	Meas_ Length (in)	Location	Described by
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters
scc01_53b.txt	29	6.69941	90	18.413	8	15.6	15.6	171	bayside	A. Peters

Figure 3. Example spreadsheet output after core sheet from Figure 1 has been digitized. Columns on core description sheet that do not contain data have been omitted. Various sections of the core are shown, gaps are represented by black circles. Stratigraphic units are integrated across user-defined intervals (1 cm). Section (B) shows how the text description portion is displayed in the output file. Section (C) shows how the header information is displayed in the output file. Since the header information pertains to the whole core, output is repeated for each interval.

Sediment Cores - Data input

The information on core description sheets can be divided into four basic formats: symbols, curves, sections (intervals), and text. The first column in Figure 1 shows sedimentary structures and associated features represented by standard classification symbols. The "%sand" column shows examples of curve data, where a line represents percentage of constituent. The column from left to right indicates increasing percent range. The "color" column in the figure shows how a section of the core may be referred to without a quantitative attribute, in this case the whole core is included as "grey clay to tan sand". This format is often used in the "samples" section of core description sheet to show which intervals of core have been sampled for analysis. Finally, the text on the right side of the figure is a common way to include observations of stratigraphic units (Fig. 1(C)). The data from the core description sheets will be digitized and processed according to these four format types.

The process begins by using a standard digitizer tablet to calibrate the area being analyzed on the core descriptions sheet (Fig. 4(A)). Software that receives the signal from the tablet and outputs x,y positions as an ASCII file is necessary. Similarly, the core description sheets can be digitized on a computer by first generating raster images of the sheets and using a digitizing software that outputs x,y positions in ASCII format. Tablets have been found to be quicker and more comfortable for the user when doing large batch jobs, however, whether a sheet is digitized on a tablet or computer screen is not important as long as the output data is the same. The horizontal (x) direction of the sheet is arbitrary or could be dependant on the resolution of the tablet. The vertical (y) direction can be equal to the maximum penetration depth of the sediment core. This will facilitate processing of the downcore information without requiring an extra conversion formula. Assuming that a collection of core description sheets have the same layout, a index table can be developed by the user that describes the range from an origin (i.e. the left edge of the core description sheet) that each attribute column occupies (Fig. 5). Included on this index table is a keyword of what each column represents, this description is usually at the top of each column on the core description sheet. The index table also includes a keyword designating whether the attribute is a curve, symbol or text; this determines which processing routine will be used. The actual digitizing process involves reducing the information on the sheet to x,y values, as shown in Figure 6. Once presence and magnitude of an attribute has been established, the information can be recorded versus depth where they occur. For example, if a core description sheet depicts the presence of shell material from 0 to 130 cm core depth, the abundance profile of this parameter can be recorded as x and y values relative to an origin (Fig. 4(B)). Data that is not represented as a percentage is recorded only by the top and bottom of distribution (Fig. 4(C)). All of the non-text attributes on the sheet can be recorded in this fashion, converting the text and symbol descriptions requires extra referencing.

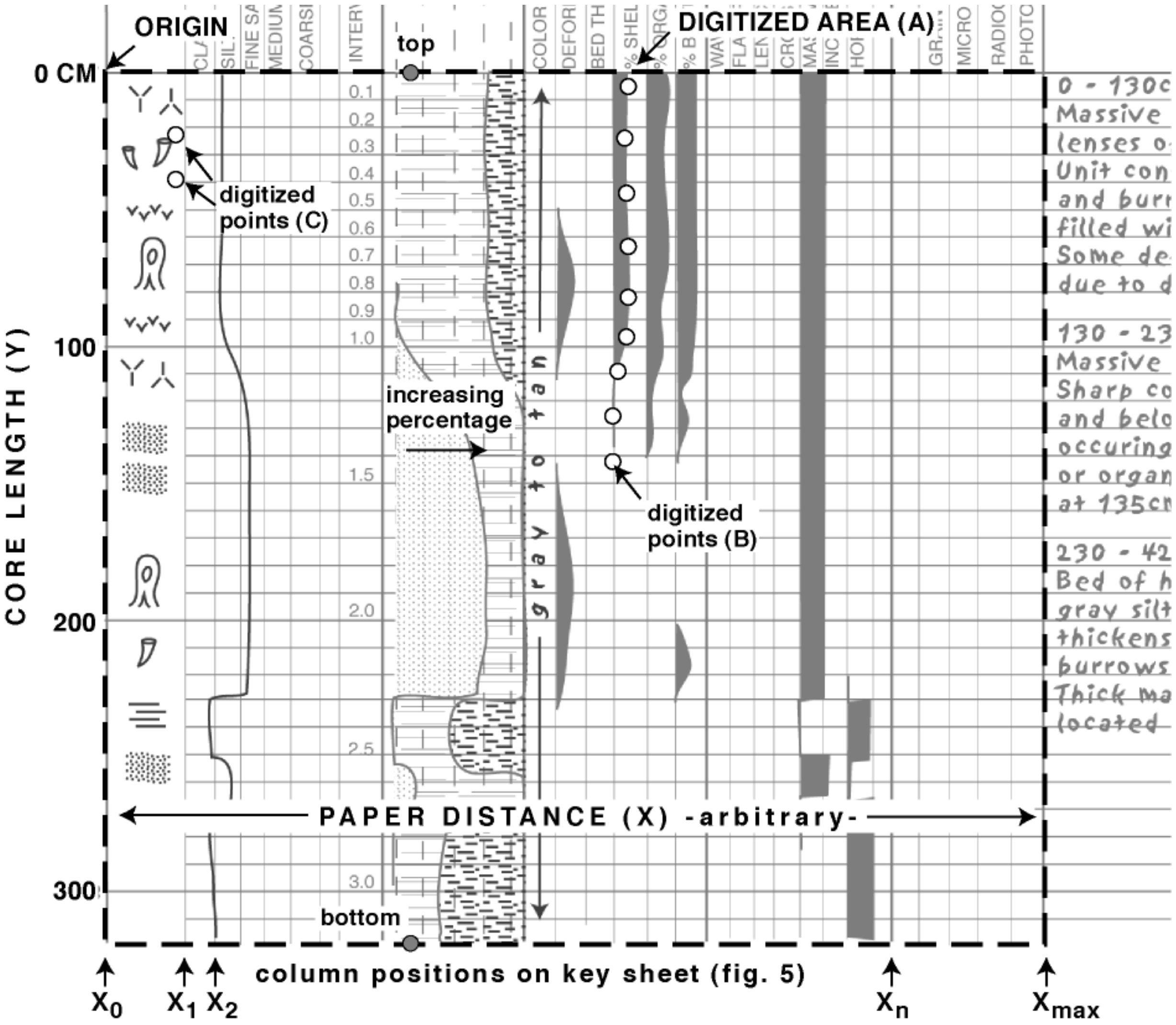


Figure 4. Expanded section of core description sheet from figure 1 showing area being digitized (A). Columns are recognized by their distance (x-value) from origin, vertical dimension (y) is either core length or depth of core penetration relative to a datum. Curve attributes are digitized along their outer edge (B). Symbols and text are digitized according to their distribution (C).

start	0	nodata	x-beds
clay	99.808	symbol	deformed_beds
silt	128.330	colcurve	planar_beds
fine_sand	156.521	colcurve	lenticular_beds
med_sand	185.071	colcurve	organics
coarse_sand	212.728	colcurve	sand
granule	242.172	colcurve	burrows
blank	270.846	interval	shell_fragments
interval	308.588	curve	gray
sand%	365.975	curve	tan
silt%	423.359	curve	brown
color	461.796	interval	black
deformation	488.817	curve	
bed_thickness	518.567	curve	
%shell	545.433	curve	
%organic	574.822	curve	
%bioturb	604.764	curve	
wavy	631.537	curve	
flaser	659.110	curve	
lenticular	688.072	curve	
cross_bed	714.806	curve	
massive_bed	744.359	curve	
inclined_bed	772.441	curve	
horiz_lam	801.307	curve	
blank2	829.705	curve	
grain_size	857.333	interval	
heavy_minerals	885.636	interval	
micro_fossils	914.433	interval	
radio	943.770	interval	
radio2	972.458	interval	
photo	1000.00	interval	
text		text	

Figure 5. Numbers in index table are distance from origin of columns on the core description sheet. Text attributes are indicated by x-distance > 1000. Rightmost column are the attributes listed on the symbol menu (see Fig. 6(C)).

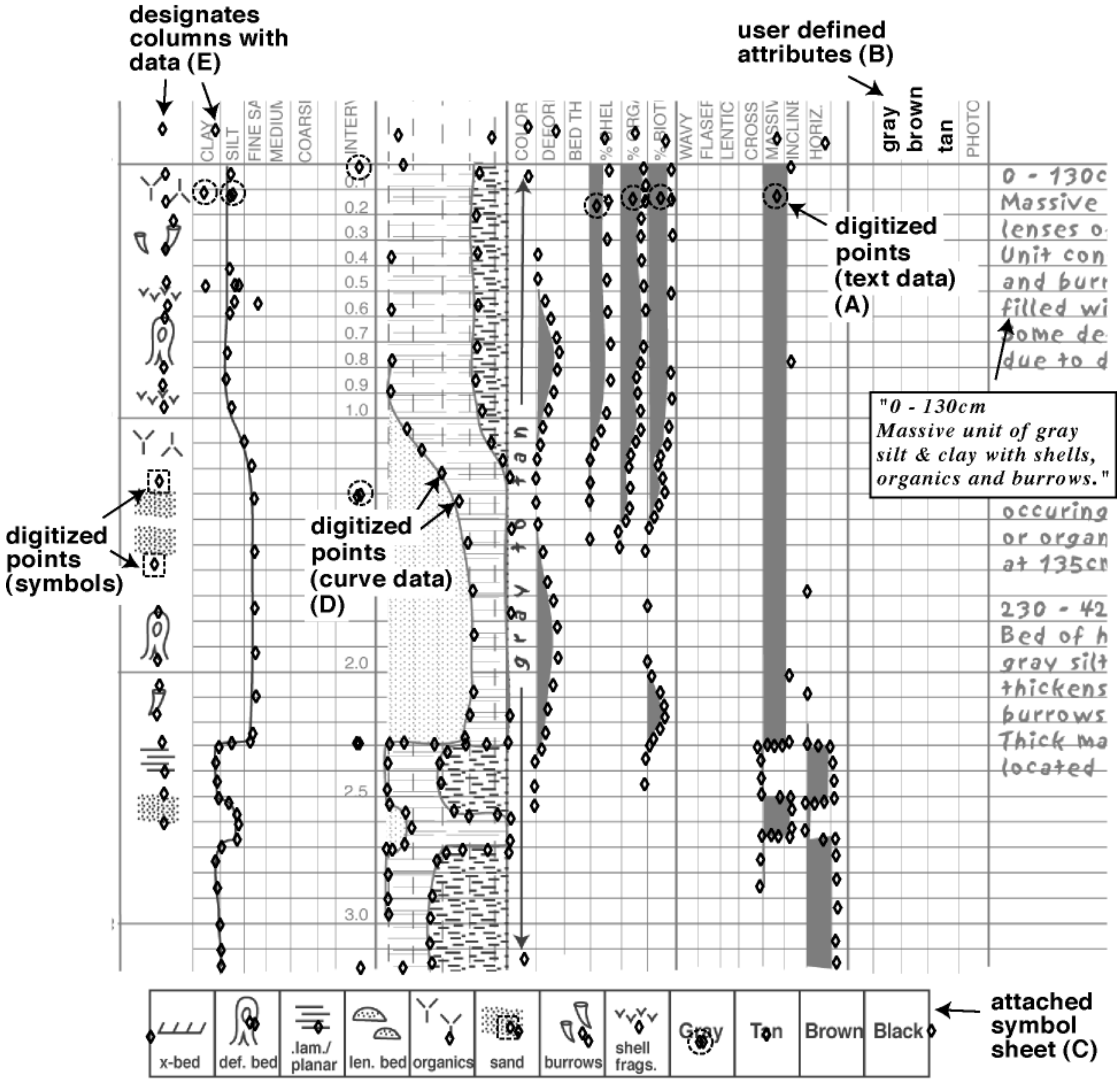


Figure 6. Overlay of digitized points (black diamonds) onto the core description sheet, showing how the information is reduced to points. The text information on the right-hand side of the sheet is accounted for using column descriptions and the symbol menu (A), using depth intervals of the stratigraphic unit (shown as points with dash circles). Digitized points with dash squares show how symbol data is rendered. Unused columns can be re-designated to include applicable attributes (B). A symbol key can be added to digitize commonly used symbols or

colors (C). Columns are delimited by digitizing a point within the column, but above the top of the sediment core (E).

Typing the text descriptions into a spreadsheet can be a tedious process. This step can be avoided since most of the relevant information in the text description can also be described using the other attributes on the core description sheet. For example, in figure 1(C), the first 130 cm of the core is described as “massive grey silt with abundant shell and organics”. This description can be represented by existing fields on the core description sheet that reiterate the text statement. Note that most of the descriptors in the sentence already have columns on the sheet that reflect downcore presence and magnitude. Thus, the above description would be digitized as: "0, 130, massive, grey, silt, shell, organics" (Fig. 6(A)). The extent and magnitudes of these constituents within the unit is already recorded using their respective profiles. Depth (y) ranges for each text description, commonly stratigraphic units, is recorded by designating a column on the core description sheet to be used as depth. When the processing program detects the x-values for the depth column, it knows that a core depth interval is being defined. Until a new core depth interval is defined it will assume that all the text attributes being digitized pertain to that stratigraphic unit. The program will then loop through the user-defined intervals and list text attributes repetitively until a new text description occurs. This will retain the text information with all the other attributes and force the text descriptions of the core description sheet to conform to spreadsheet requirements, which most database applications recognize. Additional fields can be added in specific locations on the sheet to further capture text information not represented by columns. For example, note that this core description sheet does not have fields corresponding to specific sediment colors. Blank columns or non-used fields can be adopted to include common descriptors such as “grey”, “tan”, or “contact” (Fig. 6(B)). Additionally, a symbol menu can be developed and attached to the sheets during the digitizing process (Fig. 6(C)). Attaching text descriptors to the symbol menu, however, should be minimized. The advantage to restricting text descriptions to existing (and a few extra user defined) fields on the core description sheets is that the terminology becomes standardized for all the cores described using the specific core description form. This way, interpretations from any sediment core survey that used the specific core sheet shown in figure 6 will have the same text descriptors, this can increase the correlation potential between cores even if they are from different projects or were described by different interpreters. If additional columns are to be used for text descriptors, their listing is required in the index table (Fig. 5). A set of index tables can be collected that describe different core description forms, or modifications to the same form. The user will be prompted to select the specific index table during the next step of this process.

Sediment cores - Recognizing the attributes

At this point the type of data on the form is irrelevant. It is the key sheet that establishes the link between the digitized values and the form attributes. Since the key sheet is customized for each style of core description form, it is not necessary to modify the processing program when different core description forms are encountered. The key sheet only picks one x-value for each attribute on the core description sheet (Fig. 4). However, since the attributes occupy space (i.e. columns), it is necessary to pick one edge or a midpoint of the attribute space. In this process, the right edge of the attribute column describes the x-value of that attribute. So that:

$$\textit{Attribute} = \textit{Attribute (x-value)} - \textit{Previous attribute (x-value)}$$

For curves, such as shell percent (Fig. 1(A)), the left edge of the column equals zero percent presence of attribute and the right edge equals 100 percent. Thus, the percentage of attribute per core depth interval can be calculated as:

$$\textit{Attribute percent (y)} = (\textit{Attribute (x-value)} - \textit{Xn-1}) / (\textit{Xn} - \textit{Xn-1})$$

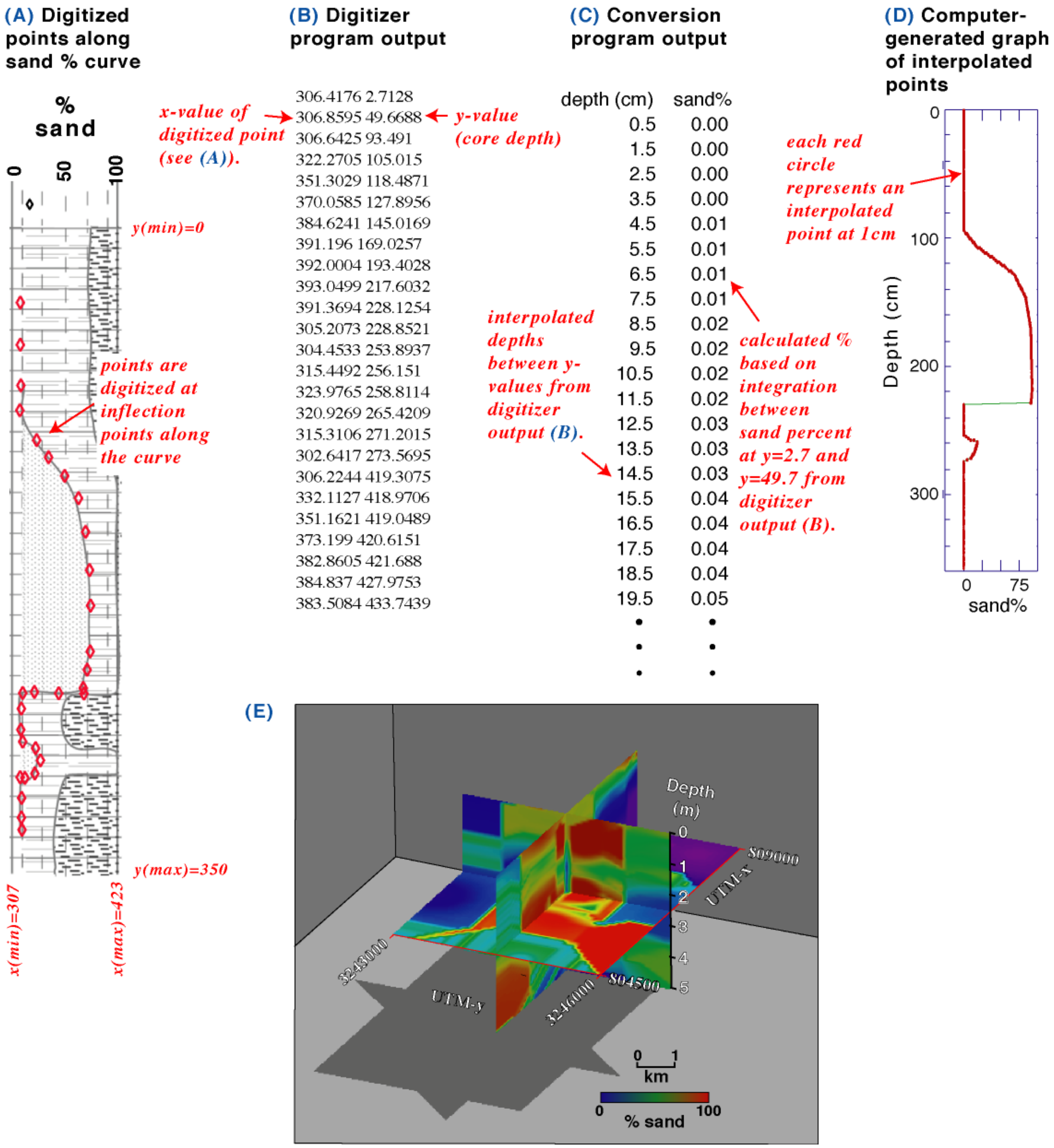
Where: *Attribute (x-value) = x-value of digitized point*

All of the non-text attributes are digitized as x-y values, the computer program uses the above formulas and the index table to identify the attributes by their x-value, calculate percent abundance and output the values relative to core depth (Fig. 3).

Since the attribute percentages are recorded per user-defined core interval (y), x-values for each core interval needs to be determined. It would be very difficult to digitize a point for each y-interval. Typically points are digitized only where there are inflection points in the data. For example, if a core contains sand only between 130 and 230 cm core depth, one might digitize the zero percentage above 130 cm, and then follow the sand percentage curve, digitizing where the curve deviates from a straight line (Fig. 7(D)). Since these digitized points do not correspond to any consistent core interval, it is necessary to interpolate x-values at the core intervals in between digitized values. This is accomplished by considering the vertical (y) distance between digitized points and integrating an x-value depending on the x-values at the digitized points:

*X value (n) = ((X value (dig1) – X value (dig0)) * (Y value (n) – Y value (dig0)) / (Y value (dig1) – Y value (dig0))*

Where dig0 and dig1 represent the previous and subsequent digitized points, respectively, around the core interval for which an x value is being interpreted. The processing program will perform a rolling loop to calculate an x value for every down core interval, throughout the length of the core. This process provides a greater detail of information than that obtained by constraining the core description to just stratigraphic units.



(A) Digitized points along sand % curve

(B) Digitizer program output

(C) Conversion program output

(D) Computer-generated graph of interpolated points

(E)

```
306.4176 2.7128
306.8595 49.6688
306.6425 93.491
322.2705 105.015
351.3029 118.4871
370.0585 127.8956
384.6241 145.0169
391.196 169.0257
392.0004 193.4028
393.0499 217.6032
391.3694 228.1254
305.2073 228.8521
304.4533 253.8937
315.4492 256.151
323.9765 258.8114
320.9269 265.4209
315.3106 271.2015
302.6417 273.5695
306.2244 419.3075
332.1127 418.9706
351.1621 419.0489
373.199 420.6151
382.8605 421.688
384.837 427.9753
383.5084 433.7439
```

```
depth (cm)   sand%
      0.5     0.00
      1.5     0.00
      2.5     0.00
      3.5     0.00
      4.5     0.01
      5.5     0.01
      6.5     0.01
      7.5     0.01
      8.5     0.02
      9.5     0.02
     10.5     0.02
     11.5     0.02
     12.5     0.03
     13.5     0.03
     14.5     0.03
     15.5     0.04
     16.5     0.04
     17.5     0.04
     18.5     0.04
     19.5     0.05
       •        •
       •        •
       •        •
```

x-value of digitized point (see (A)).

y-value (core depth)

points are digitized at inflection points along the curve

y(min)=0

y(max)=350

x(min)=307

x(max)=423

interpolated depths between y-values from digitizer output (B).

calculated % based on integration between sand percent at y=2.7 and y=49.7 from digitizer output (B).

each red circle represents an interpolated point at 1cm

Figure 7. Output-flow of sand percent curve from points digitized from profile (A), raw x,y digitizer-output values (B), conversion to depth versus sand percent (C) and resulting sand percent plot (D). Latitudinal and longitudinal slices through a 3-dimensional model of a 3 km^2 sand deposit (E), extending 5 meters below the sea floor. The

image was generated from the digitized and gridded output of the observed sand percent of 15 sediment cores, integrated at 1 cm intervals.

The attributes can be determined only by comparing the x-value from the index table to the digitized x-values. This may create a problem if there is any deviation from the defined columns, say if a point in the shell percent column is inadvertently digitized slightly left of the column (i.e. < 0%), or if there are any distortions between the facsimilies of the sheets being digitized. This problem can be avoided by incorporating variable margins to the attribute constraints. The index table defines where the percent-shell attribute begins and ends on the core description sheet, but the processing program can apply a certain amount of leniency to those x-values. As an example, if there is no shell material between 130cm and 230cm in the core (Fig. 4(B)), the user may digitize down the zero percent line (left most portion of shell% column). While doing this, it would be easy to “miss” the line and digitize a point in the previous column, which on this particular core description sheet describes bed thickness. The processing program would then interpret this digitized value to refer to a bed thickness attribute. However, if the computer program were to incorporate a ‘slop’ range about the column constraints, then any point within this range could be included with the appropriate attribute values. The discrepancy can then being converted to 0% or 100% depending on which side of the column it occurred. Alternatively, if each column were clearly identified at the onset of digitizing, then points beyond the column range can be set to the range. For example, if the first digitized point for each column was set squarely in the middle of the column, but out of the actual core depth range, then this point could be used as an indicator of which column the subsequent points pertain to (Fig. 6(E)). Thus, points digitized to the left of the column in this set would be interpreted as 0% percent and those to the right as 100%. Such refinements in the program allows for flexibility and improved efficiency in converting large amounts of data.

Cores - Further refinements

The processing program prompts the user for the sediment core identification number and retains it as the output file name. The rest of the header information (elevation, location, etc.) will need to be incorporated into the output dataset. This will either be done by hand, or if this data already exists in a digital file a further processing step occurs where this file is accessed and the data extracted using a “find” routine that locates the core identification number. Since this type of information commonly already exists in a digital format within metadata or navigation data sets, the information should be compiled for each sediment core survey into one large spreadsheet for easy access by the processing program. The end of the spreadsheet table shown in Figure 3 shows how the header data from the core description sheet was included for each user-defined interval in the output file.

Multiple core description sheets can be digitized at one session. If the sheets do not each occupy the same physical space on the paper, for example if the sheets are photocopies and the images move around on the pages, then a three point calibration will be necessary prior to digitizing each page. This calibration can be avoided if the area to be digitized is outlined on a clear plastic overlay. This template can then be attached to the digitizing tablet and each core description sheet can be placed under the overlay in the proper position. This template can be used only in the situation where all of the sheets are pre-printed original forms; photocopies of an original form may produce distortions and unreliable results. If this is the case, then three-point calibrations must be preformed prior to digitizing each core.

Once numerous core description sheets have been digitized, the digital files can be batch-processed by the processing program to greatly increase conversion speed. Once the conversions have finished, macro routines can be called up to graph the output data. This will allow the user to compare the computed values to the actual core description sheet (Fig. 7 (A & D)). Finally, the output of the data can be customized to fit database requirements, such as specific field positions of attributes within a spreadsheet.

Seismic profile interpretations

Seismic profiles are commonly used to map stratigraphic horizons across a study area. These horizons are identified as laterally continuous amplitude peaks on the profiles. Commonly, the stratigraphy is represented by delineating these consistent acoustic reflections on the profile and converting depth of the reflection from two-way travel time (seconds) to depth (i.e. meters) (Fig. 2). Geographic fixes, tied to the seismic profile by shot and/or time, are recorded with the data so that position of the profiles can be determined (Fig. 2(A)). Complex computer programs allow the geologist to perform this task completely in a digital format, automatically keeping the interpretations geo-referenced and computing spatial extents such as unit volumes. However, a large portion of this data still exists as hard copies. In this case, stratigraphic interpretation is typically drawn onto the profile, thus it becomes necessary to convert the interpretations to a digital format so that they can be geo-referenced, correlated to data from neighboring profiles, or perform other spatial computations. This can be accomplished by using a digitizing table to digitize the interpreted horizons, reducing the data to x-y values that correspond to distance along the profile and acoustic depth in milliseconds (Fig. 2(B)).

Seismic profiles – Data input

Seismic profiles are annotated with navigational fixes that correspond to geographic positions recorded concurrently by global positioning systems (GPS). Since these navigation fixes occupy a space on the profile (Fig. 2(A)) they can be digitized to return an x-value relative to an origin on the paper. Every point on an interpreted horizon can then be assigned a geographic position relative to its neighboring time fixes, as a ratio of the respective x-positions:

$$Gx(n) = GxTF1 - (((XTF1 - X(n)) / (XTF1 - XTF0)) * (GxTF1 - GxTF0))$$

Where:

Gx = *geographic position in the x direction (i.e. longitude)*

X = *x distance on paper*

$TF0, TF1$ = *neighboring (previous and next, respectively) navigational fixes about digitized point (n)*

The geographic-y direction (i.e. latitude) is calculated at the same time using the same equation. The x-values prior to conversion are arbitrary. The vertical (y) positions on the profiles correspond to depth in milliseconds, also annotated on the profiles. Since it is useful for the user to monitor their progress on a computer screen as they digitize, it is convenient to assign x-values the same physical dimension as the depth. This distance will be dropped during the conversion of x-distance to geographical positions, but at the time of digitizing will keep the x/y ratio close to one, so that the horizons can be seen on the computer screen rather than some highly skewed rendering.

Paper copies of seismic profiles are typically rolls or folded volumes that can be meters in length. Digitizing interpretations on these profiles are managed by digitizing the interpreted horizons in sections that occupy the width of the digitizing table. Before each section is digitized, three points on the paper are referenced and their x, y values are inserted into the digital file. This referencing later allows the conversion program to sort the sections based on their x-values and then align the profiles relative to depth (y), producing continuous output of values where the interpreted horizons occur.

Seismic profiles - Converting the interpretations

The process first involves developing a key that correlates the digitized points to geographic positions. This process is accomplished by digitizing the time fixes on the profile (Fig. 2(A) green circles). When the x-values for the time fixes are known, they can be compared to a navigation file by their time stamp to determine their corresponding geographic reference. The horizons on the seismic profile are then digitized, producing distance (x) and depth (y) values (Fig. 8). The processing program then loops through the navigational time fixes and gathers the digitized points from the horizons that occur between the current and previous navigational time fix, based on their respective distance (x) value (Fig. 2(B)). Using the formula listed above, geographic positions are then interpolated for every digitized point between the navigational time fixes. This allows for a very high-resolution representation of the interpreted horizon that can be imported into GIS programs for correlation between neighboring seismic profiles, as well as subsequent gridding of distribution and depth (Fig. 2(C)), or calculating spatial parameters such as isopachs.

```
->shot01
  0.1073861E-01 -0.8173988E+01
  0.1168612E+02 -0.8245041E+01
  0.2340040E+02 -0.8137177E+01
  0.3515490E+02 -0.8247704E+01
  0.4705719E+02 -0.8238556E+01
  0.5900904E+02 -0.8288834E+01
  0.7084188E+02 -0.8180580E+01
  0.8266519E+02 -0.8191559E+01
  0.9433040E+02 -0.8173248E+01
  0.1121494E+03 -0.8204330E+01
->seaf01
  0.8534532E+00 -0.2745066E+02
  0.1308091E+01 -0.2753857E+02
  0.2434021E+01 -0.2750509E+02
  0.4428402E+01 -0.2724030E+02
  0.5466539E+01 -0.2754482E+02
  0.6216959E+01 -0.2747284E+02
  0.8449114E+01 -0.2742580E+02
  0.9071310E+01 -0.2739397E+02
  0.1009863E+02 -0.2742041E+02
  0.1119548E+02 -0.2755587E+02
  0.1251919E+02 -0.2760121E+02
  0.1368459E+02 -0.2755766E+02
  0.1490925E+02 -0.2751392E+02
        •             •
        •             •
        •             •
```

Figure 8. Example of output from digitized seismic profile. First section (->shot01) contains x and y values, respectively, of time fixes from profile (see Fig. 2(A)). X-values are distance from origin in arbitrary units, y-values are depth in milliseconds. Next section (->seaf01) begins digitized output of first interpreted line (brown line in Fig. 2 (B)). Multiple lines can be included in each file and are indicated by a user-defined attribute name preceded by a "->".

Programming

The digitizing program is a commercial product that registers a tablet, table or computer monitor over a user defined range. To provide the proper output format, the digitizing software must be able to produces an ASCII File output of x,y locations. This is a very basic function that most cartographic or CAD digitizing software packages are capable of, the least complicated software available is recommended. Consult the software manual for directions on how to achieve the desired output; some additional editing may be necessary to produce the format needed by the processing program. During the development of this process, Didger ©* software by Golden Software ©* and Digitize ©* by Waccom ©* were used. If a digitizing tablet is not available, a raster facsimile of the core description sheet can be digitized directly from the computer monitor, using specific software that provides x,y coordinates in an ASCII format. The processing program is necessary to convert the x-y values generated by the digitizing program to interpreted values as defined in the index table. The routine is essentially nested loops that use incremented intervals to cycle through the digitized x-values from the input files and compare the values to neighboring x-values assigned to the core description attributes or navigation time stamps. When the x-values compare, the program assigns the attributes associated with the key sheet, or geographic fixes associated with the navigation time stamps to the digitized x value. The y-values correspond to either core depth or depth in seismic profile, this value is converted to appropriate distance values (i.e. milliseconds to meters) and carried through to the output dataset. While the programs are iterating through their routines, the output data is being generated and written to a tabulated ASCII file. The output files contain all of the input information and their corresponding interpolated x-relationships, such as constituent percentages for the core descriptions (Fig. 3) or geographic positions of the interpreted stratigraphic horizons (Fig. 8). These output files are then ready for the next step of data processing using GIS or database applications.

The processing program to convert the core description sheet data used for this report was written in Visual Basic as a Microsoft Excel © Macro. As such it is transferable across platforms and can run on any computer that contains Microsoft Excel ©. The macro and directions for it's use are available as a USGS Open-File report (Flocks, 2004). The dig_out files need not be Excel spreadsheets, as long as they are tab-delimited ASCII text files. The format requirements of the dig_out files can be modified in the conversion program. Output is an Excel spreadsheet.

The seismic profile processing program was written in FORTRAN, and requires a FORTRAN compiler to run. However, the equations outlined in this report are basic arithmetic equations that can be included into any computing language, the rest of the program is simply file reading and iteration routines. And as these processing programs are run after the digitizing procedure, they need not be associated with the digitizing software. During the use of the digitizing program for seismic interpretations, the horizon names should be relevant to the interpreted horizons to avoid confusion during batch processing of multiple files or lines.

Acknowledgements

This work is the result of an identified need to convert analog data into a digital format for use in database applications, specifically usSEABED, and for regional programs such as the National Marine Aggregate Resource Assessment. The author would like to acknowledge J. Kindinger and J. Williams for their persistence in resolving this issue, and their support and review of this publication. The comments of J. Reid, C. Jenkins and M. Kulp improved the processing output, and C. Dreher and N. Ferina were instrumental in the digitizing process. Core descriptions from the University of New Orleans Department of Geology and Geophysics were used as templates to generate the processing code for the down-core digitizing system, A. Peters provided the description for the sheet used in this paper.

References Cited

Flocks, J., 2004, Directions for converting analog sediment core description sheets and seismic profile interpretations to digital, spreadsheet format: USGS Open File Report, in review, 17p.

Kindinger, J., Flocks, J., Kulp, M., Penland, S., and Britsch, L, 2001, Sand resources, regional geology, and coastal processes for the restoration of the Barataria Barrier shoreline: USGS Open File Report, 01-384, 69 p.

Suter, J., Penland, S., and Ramsey, K., 1991, Nearshore sand resources of the Mississippi River delta plain: Marsh Island to Sandy Point, Louisiana Coastal Geology Technical Report, No. 8, 129 p.