From Paper to Binary: Problems and Solutions in Creating and Verifying Digital Log Databases

Do You Know What's Lurking in Your Data?

Deborah M.Olson, President, PayZone Inc.
Thomas M. Howard, Senior Petrophysicist, PayZone Inc.
Bakersfield, CA
Presented to SJWLS, January 20, 2010

Log Database: Foundation for All Reservoir Work

- In mature fields, most of what we know about the reservoir is tied to the logs
- Logs are used qualitatively (visual rock type assessments, fluid contents)
- Logs are used quantitatively (sand counts, correlation, depth references for wellbore work, petrophysical analysis)

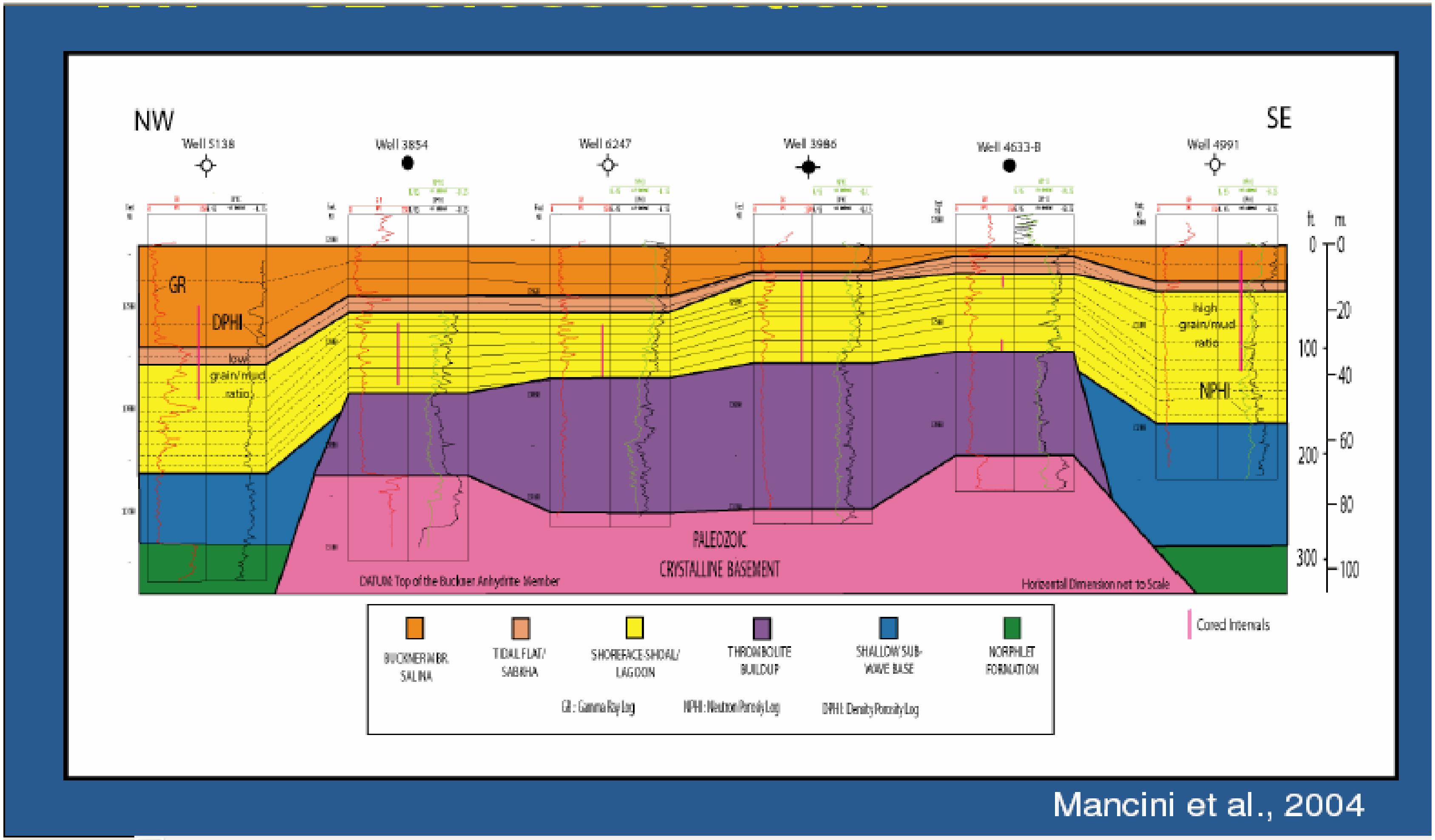
Log Database: Foundation for All Reservoir Work

- Qualitative Uses for Logs
  - Rock type (sand, shale, diatomite, etc.)
  - Rock quality (estimates of porosity, permeability, shaliness, diatomite phase)
  - Reservoir fluids (oil, gas, water)
  - Specialty logs (dipmeter, MRI, image logs, etc.)
  - Well conditions (cased hole logs)
  - Seismic tie-in (synthetics)

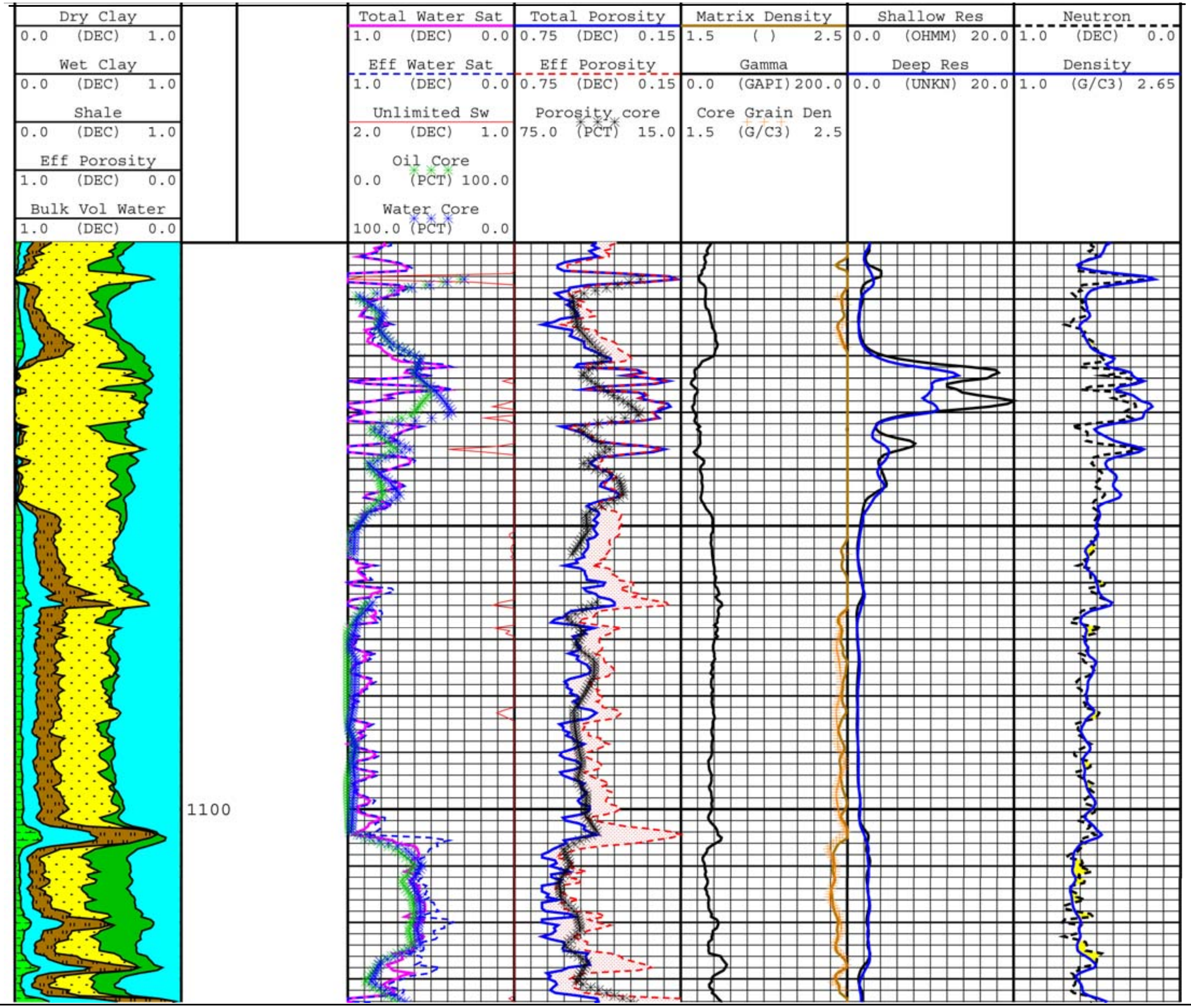
Log Database: Foundation for All Reservoir Work

- Quantitative Uses for Logs
  - Reservoir thickness (sand counts)
  - Geological correlations (depths) for structure and lithology
  - Depth-related wellbore activities, such as selecting perforations, setting casings/liners, correlating production log data to zone
  - Petrophysical analysis

Cross Section with well logs

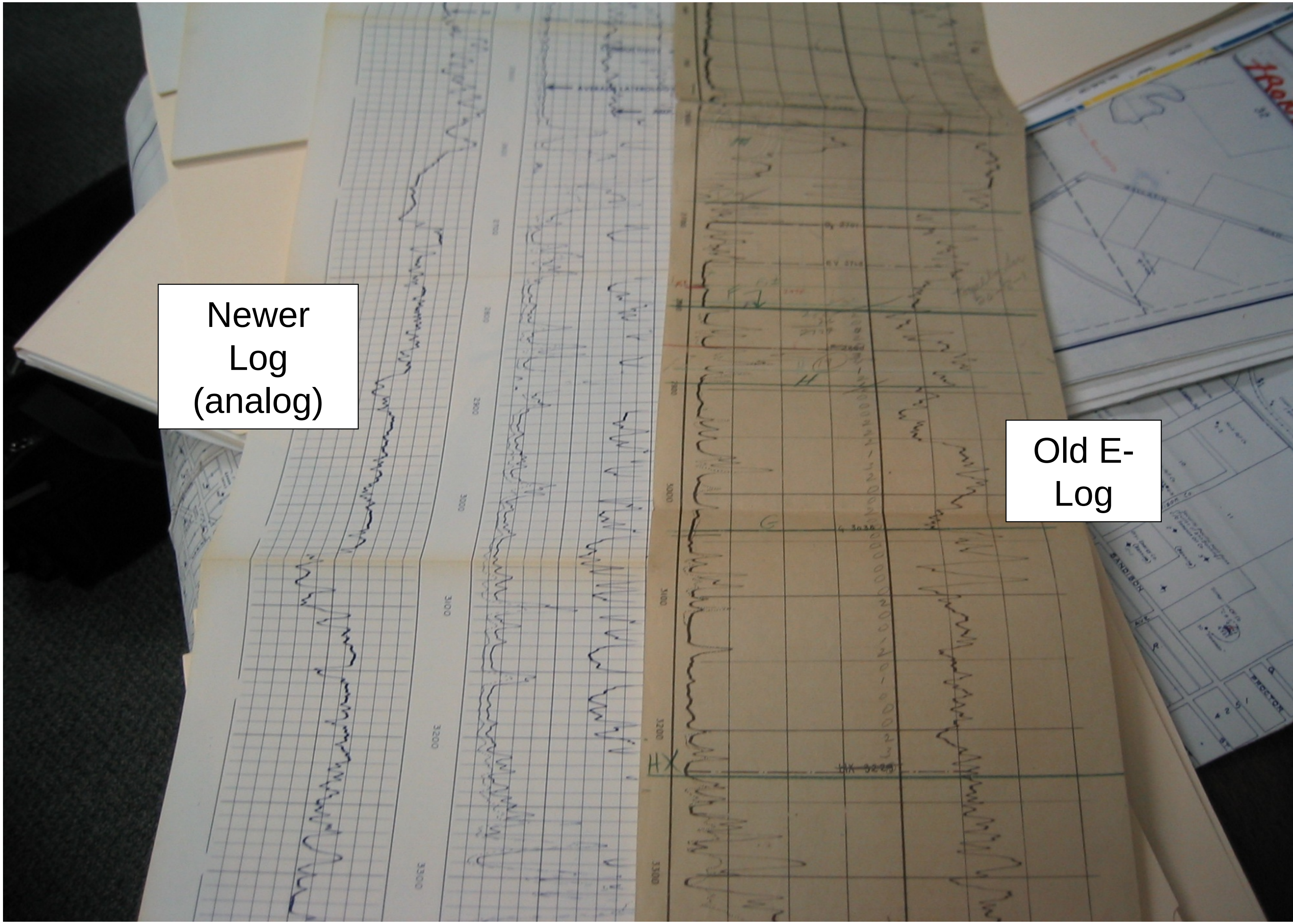


Results Log from Petrophysical Analysis

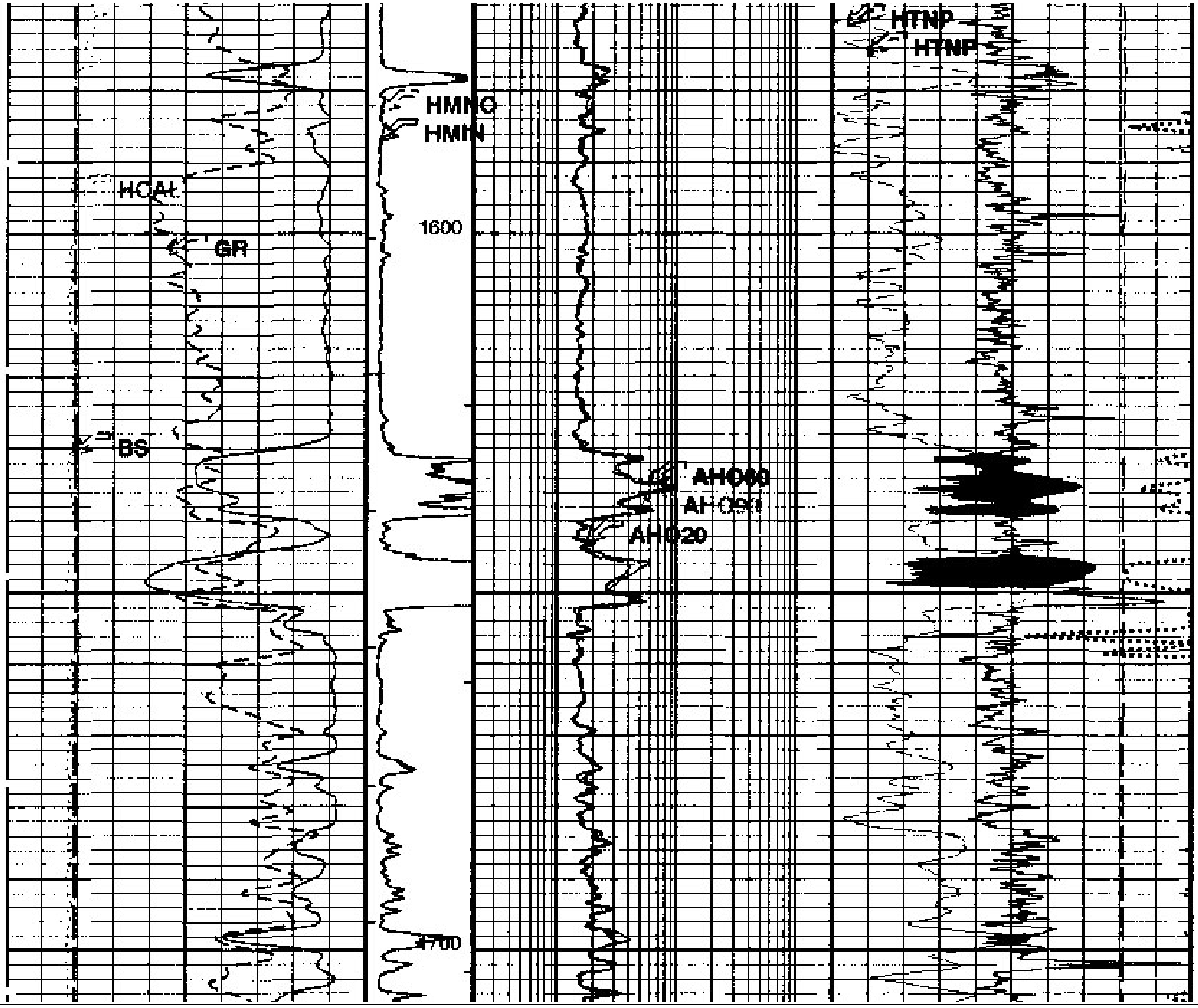


Types of Log Data

- Paper logs
  - No other record for wells pre-computer logging
  - Available for more recent wells in addition to digital data (often all you have if your company didn't own the well)
- Digital log data
  - First efforts in 1960's, for example dipmeters on mag tape using analog-to-digital conversion
  - Captured directly on modern digital logging units
    - First unit was Gearhart DDL in 1975
    - Starting in late 1970's in the SJV
  - Captured by digitizing paper logs


Newer Log (analog)
Old E-Log

Modern Composite Resistivity-Porosity Log From a Digital Logging Unit



Sources for Log Data

- Company files (paper logs)
- Company databases (digits, sometimes log scan images)
- State agencies
  - DOG in California is scanning their files and making image data available on-line for logs and histories
- Data vendors – log scans and/or digits

Digital vs Paper Log Data

- Most companies have a large amount of digital log data now
  - Loaded to a variety of computer systems
  - Used as-is without reference to the original logs
- Paper files are archived, sometimes poorly organized, hard to access, and seldom viewed
- Scanned log/history data may be available

Where Did My Digital Log Data Come From?

- Do you know?
- Most of the clients we have worked with don’t know the origin of the log digits they have.
- Log digits in local company databases likely came from a variety of sources over the last 25-35 years since computer systems that use digital log data first came into widespread distribution.

Source of Digits

- A significant fraction of digits in corporate databases came from hand digitizing
  - Large-scale digitizing projects were undertaken starting in the late 1970s to capture vast libraries of paper data
  - Just because the log was run on a digital truck does not mean YOUR digits came from that source! Many of these logs were hand digitized
    - The well belonged to another operator and digits were not available
    - Resources to load LIS tapes were not available or inconvenient

Sources of Digits

- Digital log databases have grown over the years
  - Transferred from one computer system to another
  - Been passed around as properties are sold
- LOTS and LOTS of people have had access to those databases
  - Adding and deleting curves
  - Renaming curves
  - Changing them in various ways
  - In most cases with *no documentation for what was done*

How Good is My Log Database

?

Problems in Digital Databases

- We have cleaned up large fieldwide log databases for various clients over 20+ years
- Log database problems are abundant and severe in nearly every project
- If not recognized, many of the problems we have seen can lead to incorrect results when the data are used
- If recognized, much staff time will be wasted trying to resolve the problems on an ad-hoc basis instead of completing their work assignments

Some Statistics (% total project wells)

- Poor quality digitizing – 5% - 48%
- Curve did not belong to well – 9% - 23%
- Missing data – 25% - 35%
- Off depth – 3% - 27%
- Digitized at wrong scale – 3% - 6%
- Original hole/Redrill ID – 2% - 26%
- Well Name/API Problems – 1% - 5%
- Other corrections – 5% - 12%

Problems in Digital Databases

- Problems can be generally grouped into several types, each with subtypes
  - Well name/Database identifier (usually API numbers)
  - Curve names
  - Missing/incorrect curves
  - Digitizing errors
  - Unique individual problems
    - Many in this category are actually field errors that look like data errors until you check the log print; “Good log” but the curve(s) need repair

Well Name/API Numbers

- API numbers (or other well identifier) are used in many places, and they all must match or log digits can land in the wrong well when loaded (and try to find that later without a lot of detective work!)
  - API numbers are used in different formats, and even if the number is right computers will not match it if they are not formatted the same way in all places

Well Name/API Numbers

- API number must be matched to the correct well *(how many Well #1s are there in Section 12?)*
- API redrill number must be matched to the correct borehole
  - Often the redrill numbers have not been assigned and you have to create them as you identify the redrills and sort out the logs

Well Name/API Numbers

- Goals for completed database:
  - Select consistent set of well identifiers
    - API # is most frequently applied since state agencies and commercial data vendors use it
  - Choose format of API #:
    - Full number is 14 digits (recommended)
    - DOG uses 8 digits (no state code [2], no trailing 0's [4])
  - Identify correct well name
    - Well names sometimes are changed
    - DOG changes operators on non-abandoned wells when properties are transferred

Curve names

- Creative curve nomenclature: Alphabet Soup in 4-8 Characters
  - 60+ years of service company mnemonics
  - Digitizers can come up with unusual names
  - Every company has their own set of names
  - Individuals who worked on the well at some point often leave behind curves with no explanation of what they are (many look like petrophysical results curves)

Curve names

- Beware of overwritten curves!
  - Many log database systems, particularly older ones, are/were curve name driven
    - Only one instance of each name is allowed
    - If the same curve name is present in a newly-loaded dataset, the old curve will be overwritten, often without warning
  - Curves from multiple runs with the same name might be overwritten, deleting the other runs
  - Curves from different tools with the same name, such as GR or temperature, can be overwritten

Curve names

- Goals for completed database:
  – Assign all curves consistent names indicating what they are
    - Service company- and tool-specific names should be retained where they exist
    - Old e-log curves should receive names indicating log type and spacing (AM64, AO19, etc.)
  – Curves should then be copied to a standard set of names for uses such as cross sections

Curve Names in Petrophysical Processing

- Petrophysical field study processing is inefficient unless all input curves have the same names (alias tables are not the complete solution)
- Ideally, the database should:
  - Retain raw curves with original names so the analyst always knows what type of curve it is
  - Preserve curves through each stage of preparation (depth shifting, environmental correction, normalization) with common curve names

Missing/incorrect curves

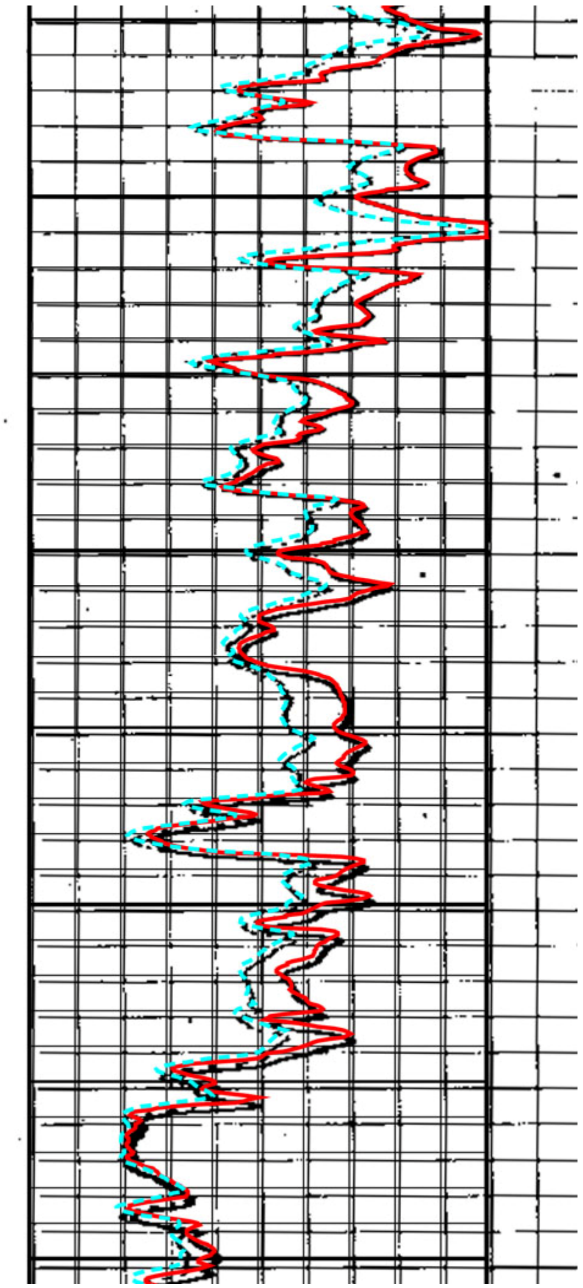
- Curves or parts of curves missing
- OH and RD curves mixed together
- Cased hole curves improperly labeled as such
- Curves present that belong to some other well

Missing/Incorrect Curves

- Goals for completed database:
  - Each well contains valid digital copies of all available (and useful) curves over the complete depth range
  - Missing curves or curve sections are digitized and merged with existing data, if needed
  - *Invalid curves are deleted!*

Good Digitizing

The digitized curves fit the original log extremely well, as shown by the overlay of digital data (color) on an image of the log. Note that the original log is slightly stretched but that the digitizing setup compensated for it. The overlay is fitted to the image at the top but if it is shifted down slightly both the grid and the digital data will fit the log image.

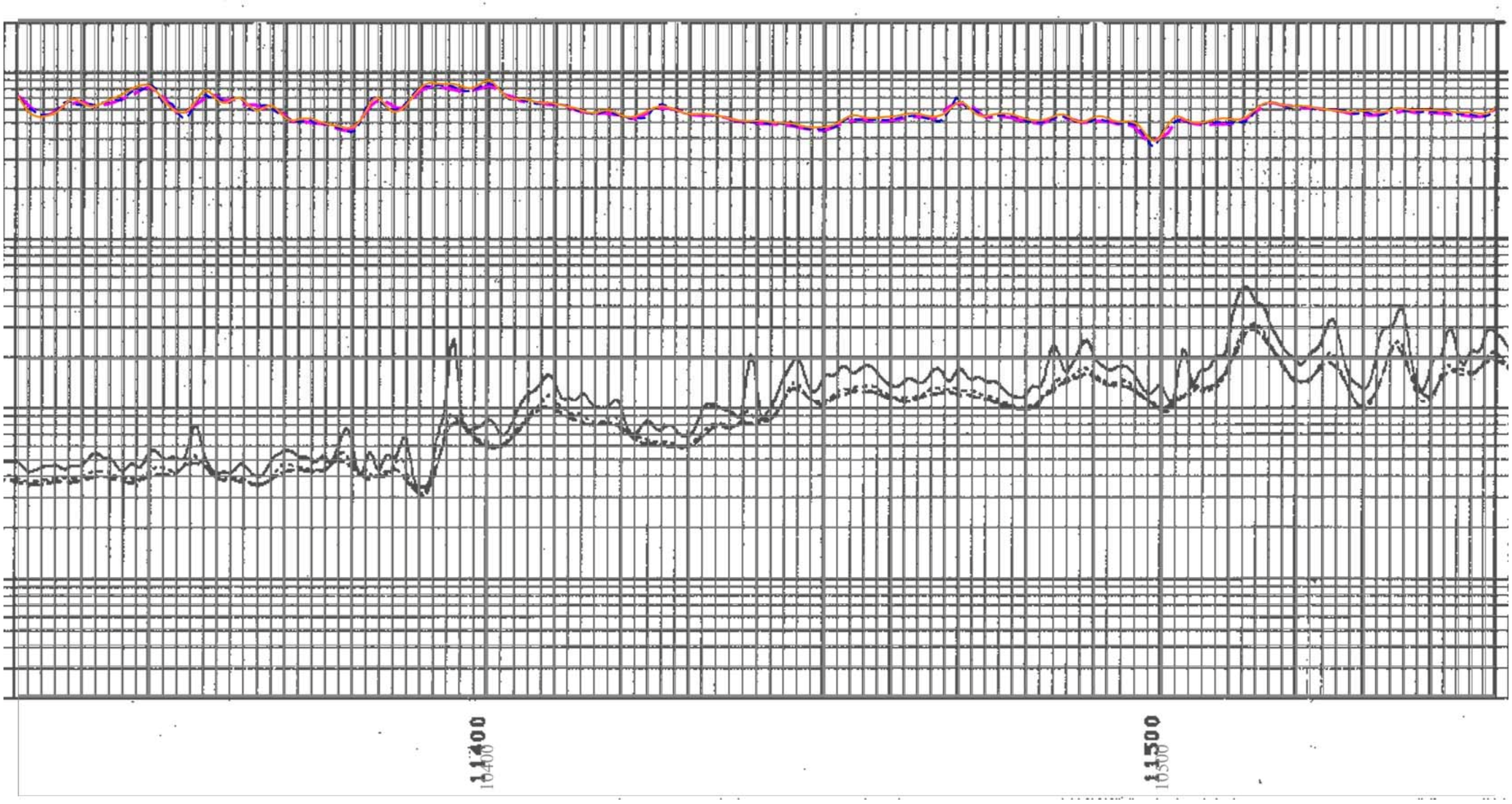


Digitizing Errors

- Wrong scales
  - Linear vs logarithmic
  - Backwards
  - Same as a different curve in the same track
  - Mis-scaled backup section
  - Just plain wrong
  - Missed scale changes
- Off depth
- Poor line quality
- Warped baseline not accounted for

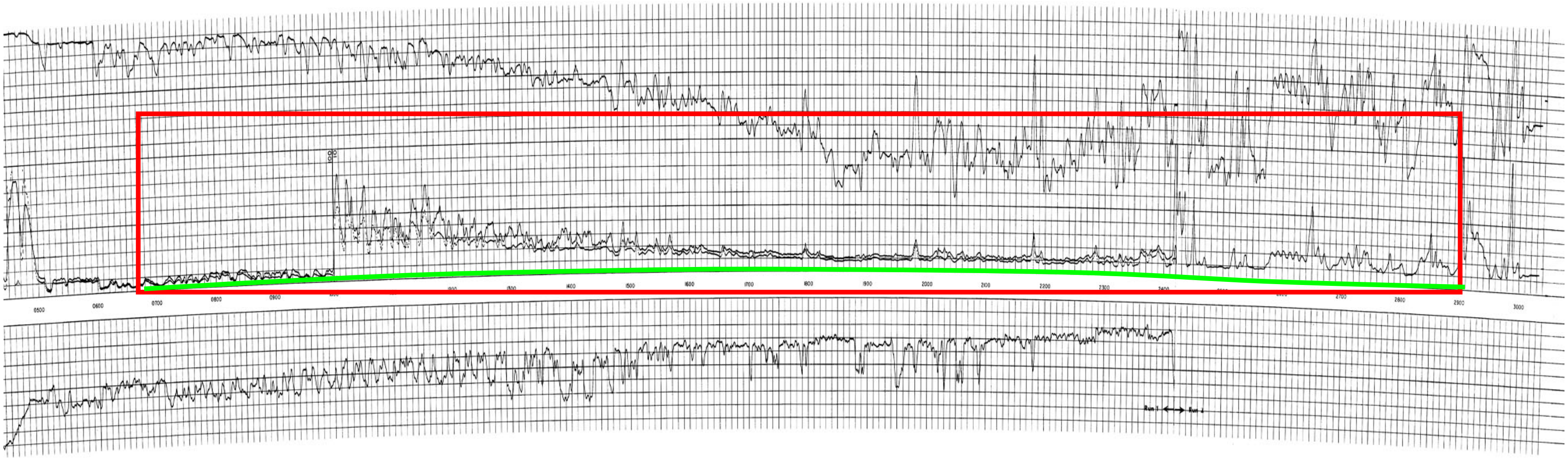
Linear to Logarithmic Scale

The colored curves are exactly the same data as the original log, and the scales were correct, but the digitizer forgot to check the “logarithmic scale” option. It can be easily fixed with a rescaling utility in your log editing system.



Warped Baseline

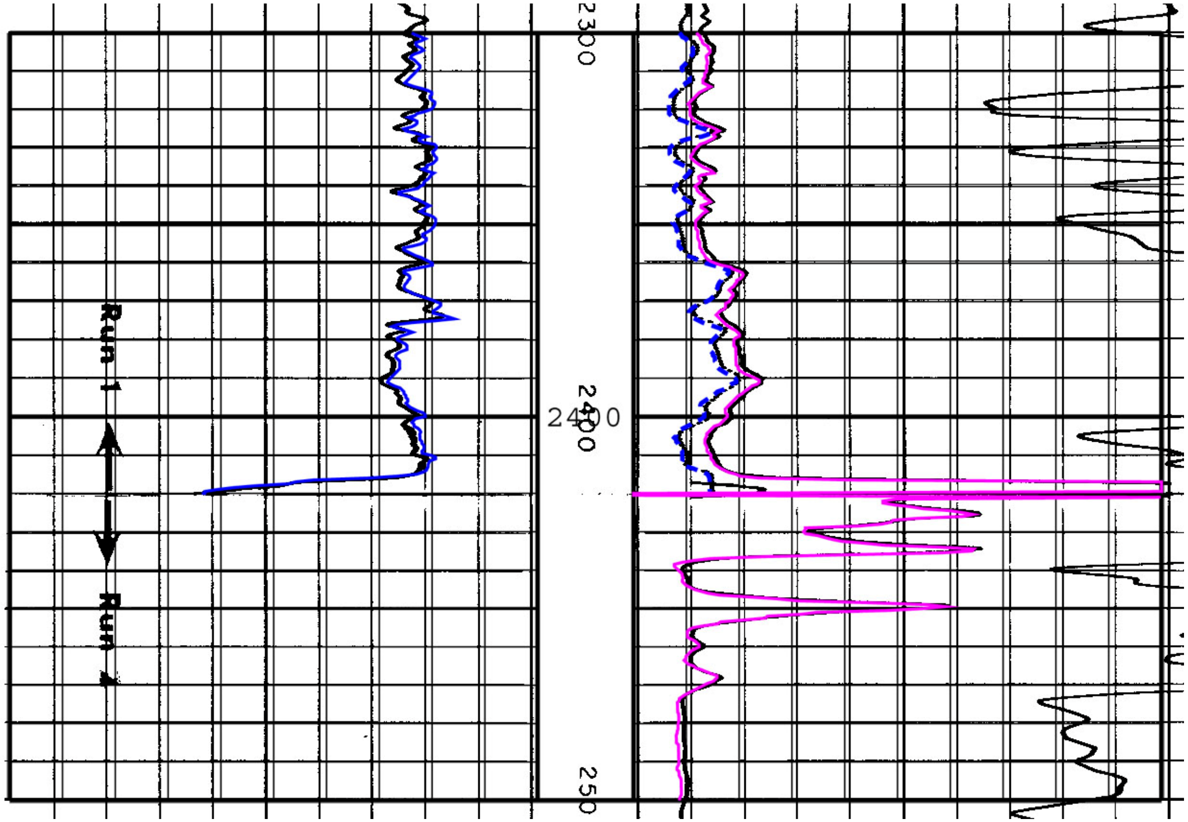
When log prints were made from film, they were sometimes skewed. Scans can also be skewed. A warped baseline (green line) will cause distortion in the digital log values when set up at the corners (red box). If this was not repaired at the time, it will cause the digital log data to drift away from the original over the length of the warped section, then drift back.



Digitizing Errors

- Digitizing parts of different curves as one
  - Sometimes difficult to distinguish one line from another
  - On spliced logs, similar line coding is different curve in each section
- Digital curve captured from 2” log compared to 5” log
  - Not really an error, but line quality appears poor on large-scale logs

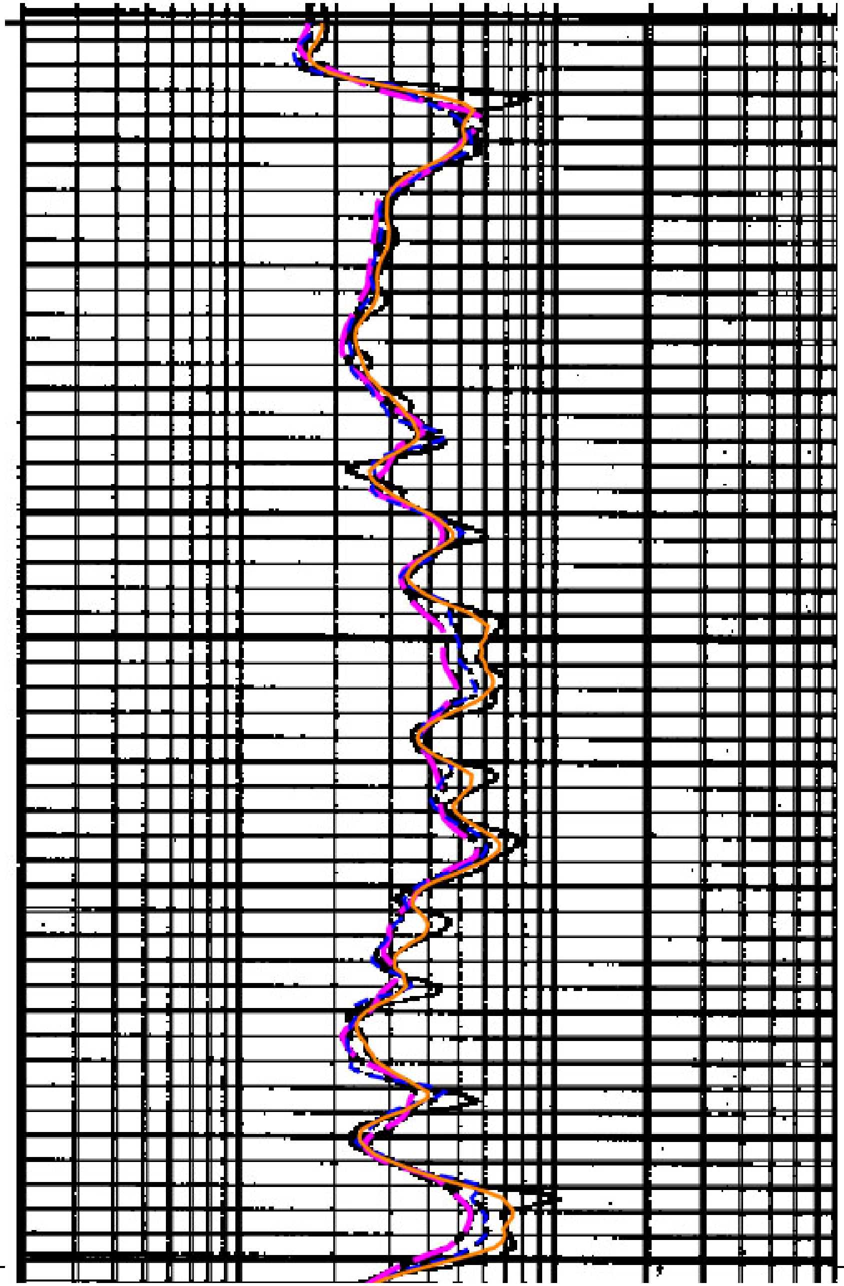
Digitizer Did Not Understand the Logs: IES Spliced to Induction Run in Oil-Base Mud



IES to 2420’ has solid line coding for SN and dashed for induction. Below 2420, a single induction log was run in oil-base mud with solid line coding. The logs were then spliced together. The digitizer followed the solid line coding and captured the induction below 2420’ as part of the SN curve.

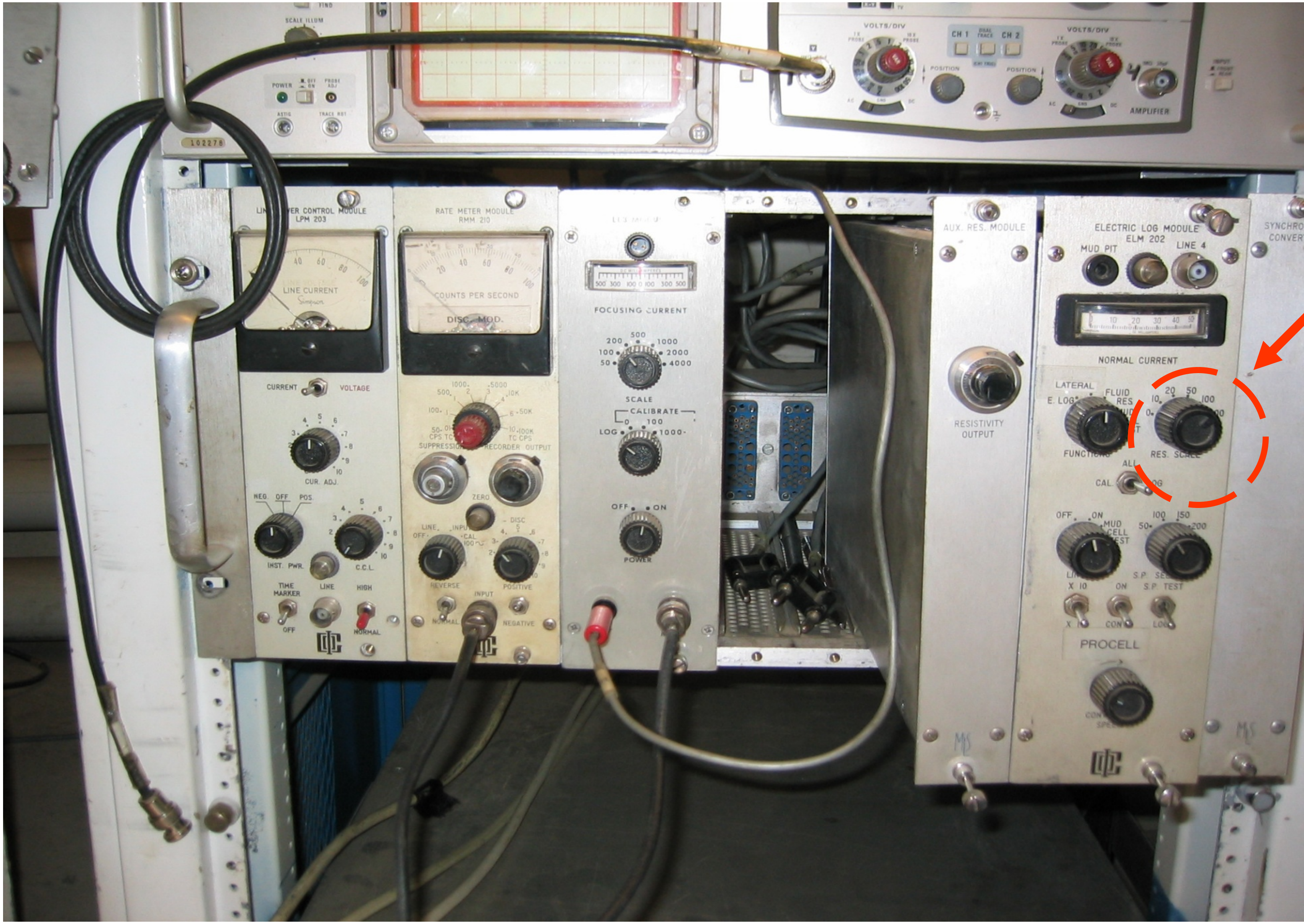
Digits from 2” Log, Compared to 5” Log

This is not really an error, but can be a data issue under some circumstances. Note that the shallow resistivity (orange curve) has missed or averaged many of the peaks on the log. The deep resistivity (pink) with its lower vertical resolution, is less affected. The medium resistivity (blue) was digitized from a 5” plot (it is not present on the typical 2”) and is a good fit. Because the medium curve came from a different paper log and digitizing shop, it could be off depth from the other logs (it is on depth in this plot)

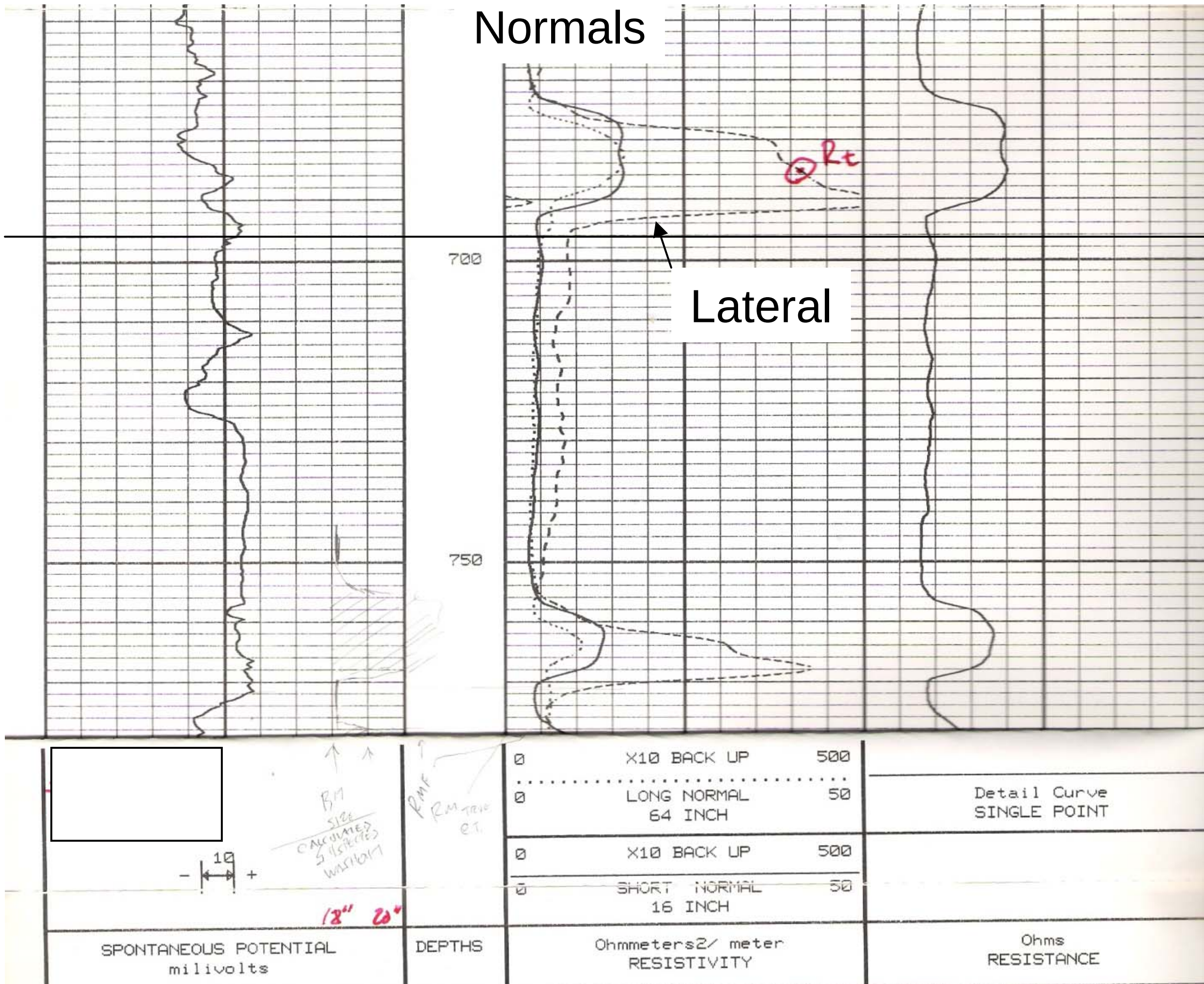


Errors from the Field

- Poor depth control, tools mismatched in depth
- Tool mis-calibrated
- Curves missing, all or part
- Tool run on wrong setting/configuration
- Tool calibrated, but ***not run correctly (e.g., too fast)***
- Poor data transmission
- Primary data errors (log spikes, cycle skips, magnetized drum on SP, etc.)


VOLTS/DIV
CH 1
CH 2
POSITION
AMPLIFIER
POWER
RATE METER MODULE
RMM 210
COUNTS PER SECOND
DISC. MOD.
LINE CURRENT
CURRENT
VOLTAGE
CUR. ADJ.
INST. PWR.
C.C.L.
TIME MARKER
LINE
HIGH
OFF
NORMAL
SUPPRESSION
RECORDER OUTPUT
ZERO
REVERSE
INPUT
POSITIVE
NEGATIVE
FOCUSING CURRENT
SCALE
CALIBRATE
LOG
OFF
ON
POWER
AUX. RES. MODULE
RESISTIVITY OUTPUT
ELECTRIC LOG MODULE
ELM 202
MUD PIT
LINE 4
NORMAL CURRENT
LATERAL
E. LOG
FLUID RES
RES. SCALE
CAL.
PROCELL

Example: wrong setting. Solution: resistivity from Lateral should be divided by two



Go Back to the Files

- There is *no substitute* for going back to the well files to resolve database problems
  - Digital curves must be plotted and compared to the paper logs; it is best if they are overlaid from top to bottom
  - Sometimes it is necessary to read the well history to resolve issues such as redrills
- Paper logs and histories should be scanned (or purchased/downloaded) and placed in a reference directory available to all

What Should a Complete Log Database Contain?

- All wells correctly named and assigned a unique database identifier (API #)
- Valid digital traces for all logs in the well, as determined from paper log copies
- All curves correctly named
- Multiple runs merged, if individual run logs are saved they have appropriate incremental numbers
- Set of curves copied to common curve names

Database Verification

- Verifying and correcting a digital log database is ***not fast and it is not cheap***
- The cost may be generally estimated ahead of time, but in our experience the magnitude of the problems can never be known in advance and either high estimates containing contingency factors or cost overruns are to be expected
- *But how much does it cost when everybody has to stop their work and look for or fix data?*
- ***How much does it cost to get the wrong answer?***

Database Verification: What it Takes

- Results will be best with a dedicated team devoted to the database task and not distracted with other work
- A common plan should be followed for each field area to ensure consistency throughout the database, with allowance for unique local issues
- Although well-trained technicians can do much of the work, personnel very experienced in logs *must* supervise the project, be available for questions, and troubleshoot the difficult problems

Database Preparation for Petrophysical Analysis

- Everything discussed so far is needed just for the general corporate log database
- More work is required to prepare the logs for petrophysical analysis
- Must be done by experienced petrophysical technicians supervised by log analysts
- 60% of the time and cost of a field study is spent on database verification and preparation

Database Preparation for Petrophysical Analysis

- If the log data are not valid, the petrophysical results will not be valid either
- Additional steps for an analysis database (after log verification) include:
  - Depth shifting
  - Environmental corrections
  - Normalization

Our Best Resource

- Our most valuable information resource for developed reservoirs is the log database
- Shouldn’t we get it right?