Study of batch pre-processing techniques for logging curves of different periods

Na Lu

Geological Research Institute of the Fourth Oil Production Plant of Daqing Oilfield Co., Daqing 163511, China

Abstract: The sampling spacing of different series of old well data curves is varied and does not facilitate multi-well batch interpretation. This study achieves automatic interception of invalid well sections and reads valid sections into memory, increasing the speed of computing and the number of wells processed in batch; it also achieves automatic error detection and curve normalization, and outputs data files with useful curves and interpretation results, increasing the speed of computer operation and the number of wells processed in batch for old wells.

Keywords: Multiple logging series, data sampling, curve name standardization.

1. Analysis of the data characteristics of the TXT data format

At present, all the old well laterals and 581 series data are converted from the original digital scan data list format data to form the new txt text format data, while the DLS series is converted from the original LA file, which also generates the same txt file as the 581 series data, all files consist of 24 curves, and the data format has the following characteristics.

(1) Irregularity of sampling spacing

With multiple sampling formats such as 0.025m, 0.125m, 0.05m and 0.02m, the multiple sampling spacing is not conducive to automatic computer stratification, so we use linear interpolation to resample the LAS data so that the sampling spacing becomes a uniform 0.05m.

(2) Fixed curve names and number of curves

The curves are composed of 24 lines, the curve names are fixed and the values of useless curves are represented by 0, which increases the memory overhead of the computer to read them;

(3) Missing and abnormal data

The old series of well network data was converted from the original digitized scanned data, and the original curve digitization techniques were limited, with individual curves missing and abnormal.

2. Research on resampling techniques for different series of old well data

There are various types of LAS format data sampling spacing such as 0.025 m, 0.125 m, 0.05 m and 0.02 m. These sampling spacing are not conducive to automatic computer stratification, so we use linear interpolation to resample the data so that the sampling spacing becomes a uniform 0.05 m.

It is assumed that we know the coordinates $(x0, y0)$ and $(x1, y1)$, to get the y-value of a position x on a straight line in the interval $[x0, x1]$, where x0, x1 are the depths of two points adjacent to the original sampling, with a sampling spacing of 0.025 m, and y0 and y1 are the curve values corresponding to x0 and x1.
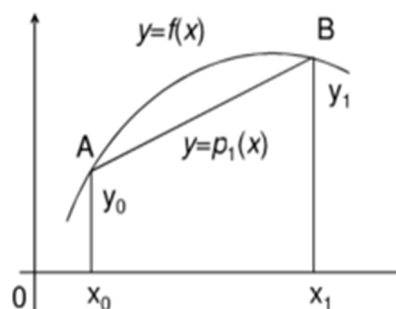

Figure 1. Schematic diagram of the linear interpolation method

According to Figure 3, assuming a point (x,y) on AB, two similar triangles can be made and we get

$$(y - y0)/(x - x0) = (y1 - y0)/(x1 - x0)..(1)$$

It is assumed that the value of α on both sides of the equation, then this value is the interpolation factor - the ratio of the distance from x0 to x to the distance from x0 to x1. Since the value of x is known, the value of α can be obtained from the equation

$$\alpha = (x - x0)/(x1 - x0) \qquad (2)$$

Likewise： $\alpha = (y - y0)/(y1 - y0)$ ....(3)

In this way, it can be expressed algebraically as

$$y = (1 - \alpha)y0 + \alpha y1 \ ......(4)$$

Or $\quad y = y0 + \alpha(y1 - y0)$ ....(5)

The y-value is the curve value calculated by interpolation, and the shape of the logging curve remains unchanged after resampling. After such processing, one is to facilitate automatic computer stratification interpretation, and the other is to harmonize with the current sampling spacing of LA data. As the lower limit of resolution of the logging curve on the logging lateral map is 0.1m, and the lower limit of thickness for various parameter interpretation is 0.2m, the 0.05m sampling spacing has no effect on the shape of the curve and does not affect the accuracy of parameter interpretation.

Figure 1-2 shows the comparison before and after interpolation of the curve of the lateral logging series of old well XXX, from the original sampling spacing of 0.025 to 0.05. As can be seen from the figure, the curve pattern has not changed much and will not affect the accuracy of the interpretation of the reservoir parameters of the old well.
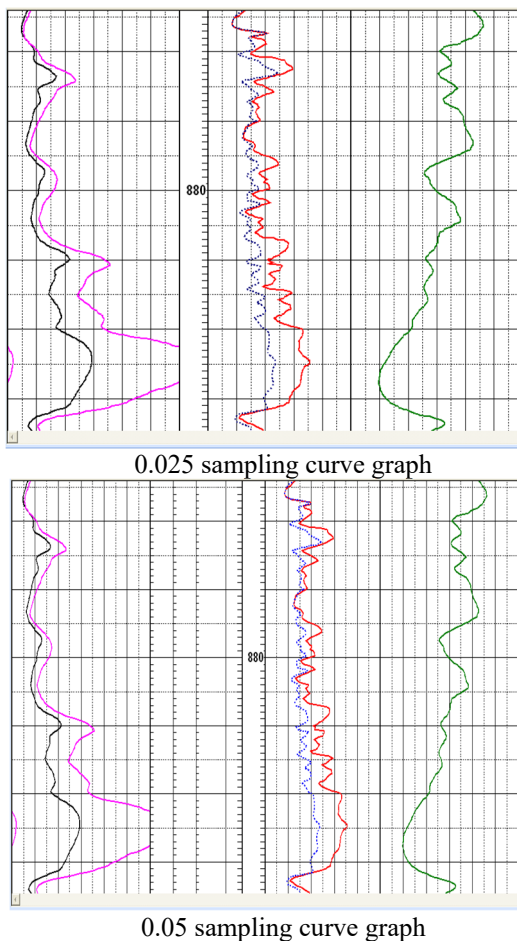

0.025 sampling curve graph

0.05 sampling curve graph

Figure 2. Comparison of curves before and after X-linear interpolation

3. Curve filtering and interception processing technology research

The old well batch interpretation is very strict on data processing. For previously processed intermediate result curves, no value curves and duplicate curves have to be filtered out, oil extraction X plant old well data files are saved 24 curves. For different series many curves are useless zero value curves, these curves need to be filtered out.

Reading of only the relevant landmark curves for different series.

Lateral logging series including: 0.25m, 0.45m, RMN, RMG, 2.5m, SP, etc.

581 logging series including: 0.25m, 0.45m, RL3S, RL3D, AC, CAL, RMN, RMG, SP, GR.

DLS logging series including: RLSLS, RLLD, RXO, HAC, DEN, GR, CAL, RMN, RMG, SP.

4. Curve format conversion and invalid data handling techniques

The TTX data format consists of two parts, the curve name information, and the data body, as shown in Figure 3.

Figure 3. Txt data format

The curve name information includes the curve names of all curves that the curve names contain English names. The same logging project has more than one curve name, such as deep lateral curve, there are three names HR3D, RL3D and RLLD. In the process of processing interpretation, it needs to be converted into standard English curve names, by reading the given standard curve name card file into the standard English curve names.

The data body information is the value corresponding to each curve. Generally, there are invalid values in the head and tail of the data and these illegal values should be filtered out.

Invalid data are indicated by -9999.0000, and invalid well sections can be identified by the invalid value of a particular curve.

5. Automatic curve error detection technology research

5.1 Curve illegal value detection

One or more rows of illegal data in different series of old wells during data conversion, where the illegal data changes the original data structure, causing problems with data decoding, which may lead to different problems such

as incorrectly read curves or deadlines, resulting in errors in multi-well batch processing.
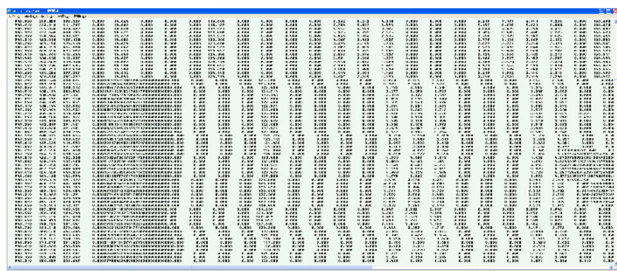

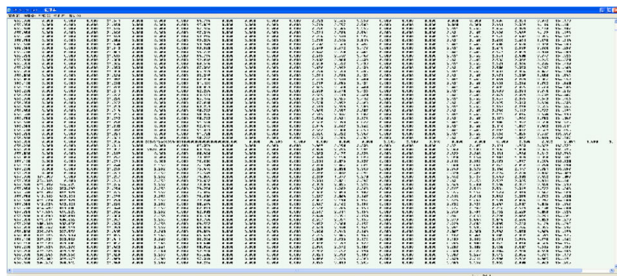**Figure 4.** XXX multiple curves with illegal values

Figure 5. Illegal values in the middle of XXX horizontal series data

5.2 Curve name reversal detection

The lateral and 581 series well networks have problems with inverted curve names, such as inverted deep lateral and shallow lateral curve names, inverted micro-electric and micro-gradient curve names, and inverted R025 and R045 curve names.

Such problems can be rectified automatically by the software, can be accurately identified by comparing the curve values of the effective thickness of the thick layer, and can be correctly matched by curve name and curve by interchanging curve data.
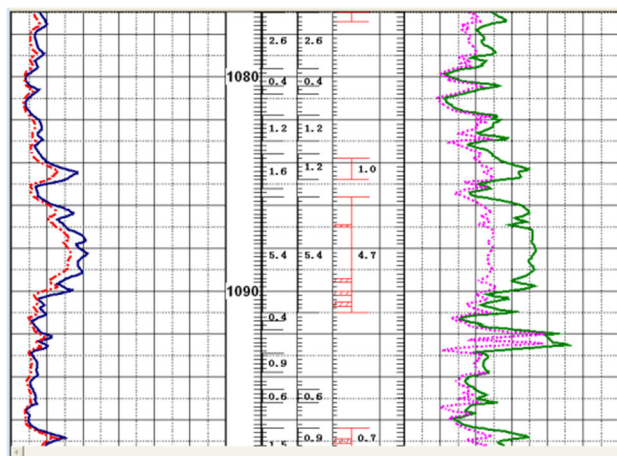

Figure 6. XXX curve reversal diagram

5.3 Automatic detection of missing curves

The number of curves in the different series is essentially fixed, and the main curves must not be missing; if they are missing, the interpretation method will not work, and the longevity of the data will be affected.

The program automatically monitors the names of the different series of curves and if a curve is missing, the correct curve name and the missing curve name are recorded by well number and saved in the "curve error message file" file to digitize the missing curve again.
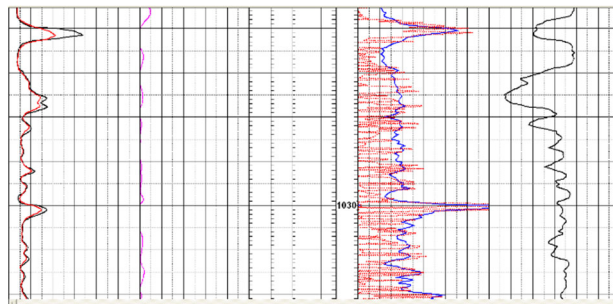

Figure 7. Schematic diagram of automatic curve name detection

6. Automatic detection of curve anomalies

The old series of well networks had outliers in the curves due to the digitization technology of the time. The outliers were generally of two types, either large flat sections or curve values that were not within the correct range.

Flat sections are identified using the effective thickness; if a well curve has more than two meters of flat sections within the effective thickness, the curve is considered to be abnormally flat.

(a) Different curves have different ranges of values, and a curve is classified as a curve straight anomaly if it exceeds the normal range in several places within the effective thickness range, e.g. if a flat section of more than 5 m occurs, the curve is designated as a curve straight anomaly. Curve anomaly identification method is that the curve generally has a normal range of values in the better range of reservoirs, but if more than 50% of the normal range, it is generally considered that there is an anomaly, for example, the density curve should be between 2 and 3, if there is a value less than 1 or greater than 4.5, it is determined that the curve is anomalous.

(a) A resistivity curve is considered abnormal if there are successive flat segments and values not less than zero within the effective thickness of the curve.

The program automatically monitors different series of curves for abnormalities and if a curve is abnormal, it is recorded by well number and saved in the "curve error message file" file to digitize the abnormal curve again.

7. Conclusion and Awareness

(1) Using linear interpolation to resample the data, the sampling spacing is made uniform at 0.05 m. The shape of the curve does not change much and does not affect the accuracy of interpretation of reservoir parameters in older wells.

(2) For the problem of inverted curve names in the transverse and 581 series well networks, we can realize automatic software rectification, firstly by comparing the

curve values of effective thickness layers, the problem can be accurately identified, and finally by interchanging the curve data, thus realizing the correct matching of curve names and curves.

(3) The program automatically monitors different series of curves for abnormalities and if a curve is abnormal, it is recorded by well number and saved in the "curve error message file" file to digitize the abnormal curve again.

References

1. Visual c++ Project Development Guide Tsinghua University Press October 2000

2. Visual c++ Programming Examples for Object-Oriented and Visual Programming Tsinghua University Press Nov. 2000

3. Estimation of oil-water relative permeability curves using production log data. Petroleum Exploration and Development, 2004, Issue 02

4. Vc++ Programming Techniques and Difficulties Analysis Tsinghua University Press April 2002

5. Visual c++ Supporting Reference Materials Beijing Hope Computer Company April 1999

6. Zhang Yunying, Wang Tieyan. Version 2.0 Workstation Daqing Adjustment Well Processing Interpretation System 2003

7. Yu Yalou, Du Guibin, et al. Report on the Development of Logging Data Processing and Interpretation Database and Tools Software 2003

8. Liu Jiang, Zhang Meiling et al. Summary Report on the Development of Application Module for Log Data Processing and Interpretation of Adjusted Wells 2003