

Generating Missing Logs B Techniques and Pitfalls*

By

Michael Holmes¹, Dominic Holmes¹, and Antony Holmes¹

Search and Discovery Article #40107

*Adapted from “extended abstract” for presentation at the AAPG Annual Meeting, Salt Lake City, Utah, May 11-14, 2003.

¹Digital Formation, Inc., 6000 E. Evans Ave, Ste. 1-400, Denver, CO, 80222
(michael.holmes@digitalformation.com)

In most fields, log data are incomplete or unreliable for some intervals or entire wells. Neural networks are becoming a fashionable method to fill-in missing data, and they are powerful. The basic methodology is to train the system over intervals where the log of interest exists, and apply the training over missing log intervals. However, there are limitations and the approach can be easily abused. Inherent in the application is the assumption that reservoir characteristics remain similar over intervals where missing data are generated. For example, if training is established in hydrocarbon-bearing levels, and the application is in wet rocks, results might be unreliable.

A better approach is to use rigorous methodology to ensure data integrity and consistency:

- ❑ Despike porosity logs to eliminate bad hole data. Proprietary algorithms are applied, followed by hand editing as required.
- ❑ For extensive intervals of bad hole, pseudo logs are created using neural net training on intervals with reliable log traces and with similar petrophysical properties.

In wells with missing logs of crucial importance, pseudo logs are generated several ways:

- ❑ Using neural networks.
- ❑ Deterministic petrophysical modeling, using shale, matrix, and fluid properties from other existing curves.
- ❑ Stochastic modeling, where an approximate curve (perhaps from neural networks) is used as input, and the reconstructed curve is output.

The different pseudo logs can then be compared, and reasons for curve divergence (if any) can be examined. This approach can highlight where pseudo curves are reliable and where they are not.

Examples

The examples are from a well in the Wamsutter area of southwest Wyoming.

1) The Importance of Data Preparation

Two different neural networks are used to create a sonic log from density, neutron, and gamma ray logs:

- Using unedited (raw) data -- the system “learns” intervals of bad hole and faithfully reproduces sonic “spikes.”
- Using edited data -- corrected for bad hole -- the system reproduces the edited data, and is much more reliable.

In Figure 1, track #1 shows the comparison of the original sonic log with the despiked log, highlighting the spikes removed. The Synthetic Sonic #1 log was created using the original density, neutron, and gamma ray logs to predict the original sonic. Note that the spikes in the original sonic are faithfully reproduced by the neural network. The Synthetic Sonic #2 log was generated using the gamma ray and despiked density and neutron logs to model the despiked sonic. The result is a more appropriate model. The red bars highlight regions where spikes were reproduced in the first model, but are corrected in the second. The blue bar indicates a region where the logs used to model the sonic do not exhibit character sufficient to model the sonic in either case.

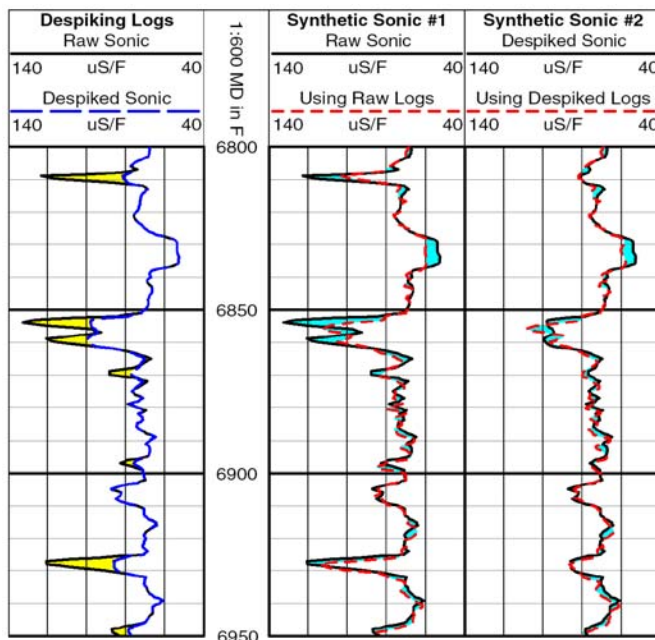


Figure 1: Comparing neural network models using the raw data with a model using despiked data.

The example demonstrates the importance of preparing the data prior to using any neural network technique. The neural network will reproduce whatever the input data demonstrates. If the input includes bad data, the neural network will “learn” to predict bad data. Thus it is crucial to take measures to ensure the input data is valid and consistent.

2) Using Appropriate Amounts of Data for the Training Intervals

Different neural networks are used to predict a sonic log from density, neutron, and gamma ray logs. In each case, different training intervals are used. The differences can be quite large if the training points are not chosen extremely wisely (very subjective). A final case demonstrates using far more of the data to create a better model which more appropriately models the reservoir.

In Figure 2, the Synthetic Sonic #1 was created using only the training regions highlighted in yellow, whereas Synthetic Sonic #2 used the entire well as the training region. As expected, the first model is extremely accurate over the training regions. The problems arise over the other regions, for which it becomes obvious that the training intervals did not fully represent the data that was being modeled. As highlighted in red, there are now regions where the first model generated erroneous spikes in the sonic due to insufficient data being provided initially. This type of selective interval approach can lead to many such problems. Although selecting more training intervals can help resolve specific issues, it becomes a very subjective model. A better approach is to begin with as much data as possible and let the neural network incorporate the maximum amount of valid data.

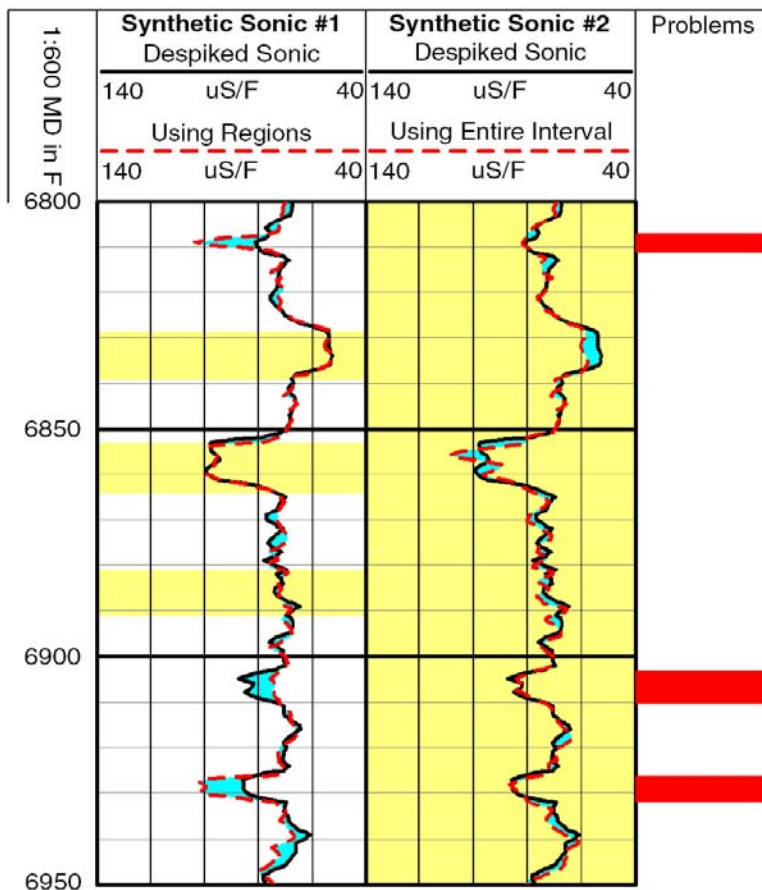


Figure 2: Comparing synthetic sonics created using specific regions as the training intervals for a neural network, with a model using the entire well as the training interval.

3) Differences using Fluid Substitution

Pseudo sonic logs, calculated deterministically and including the effects of gas substitution:

- a) Liquid-filled
- b) Residual gas
- c) Gas remote from the wellbore

It is clear in Figure 3 that in the gas-bearing sand, the sonic “sees” no gas. Density/Neutron and Resistivity responses clearly indicate gas, at depths of investigation beyond those measured by the sonic log.

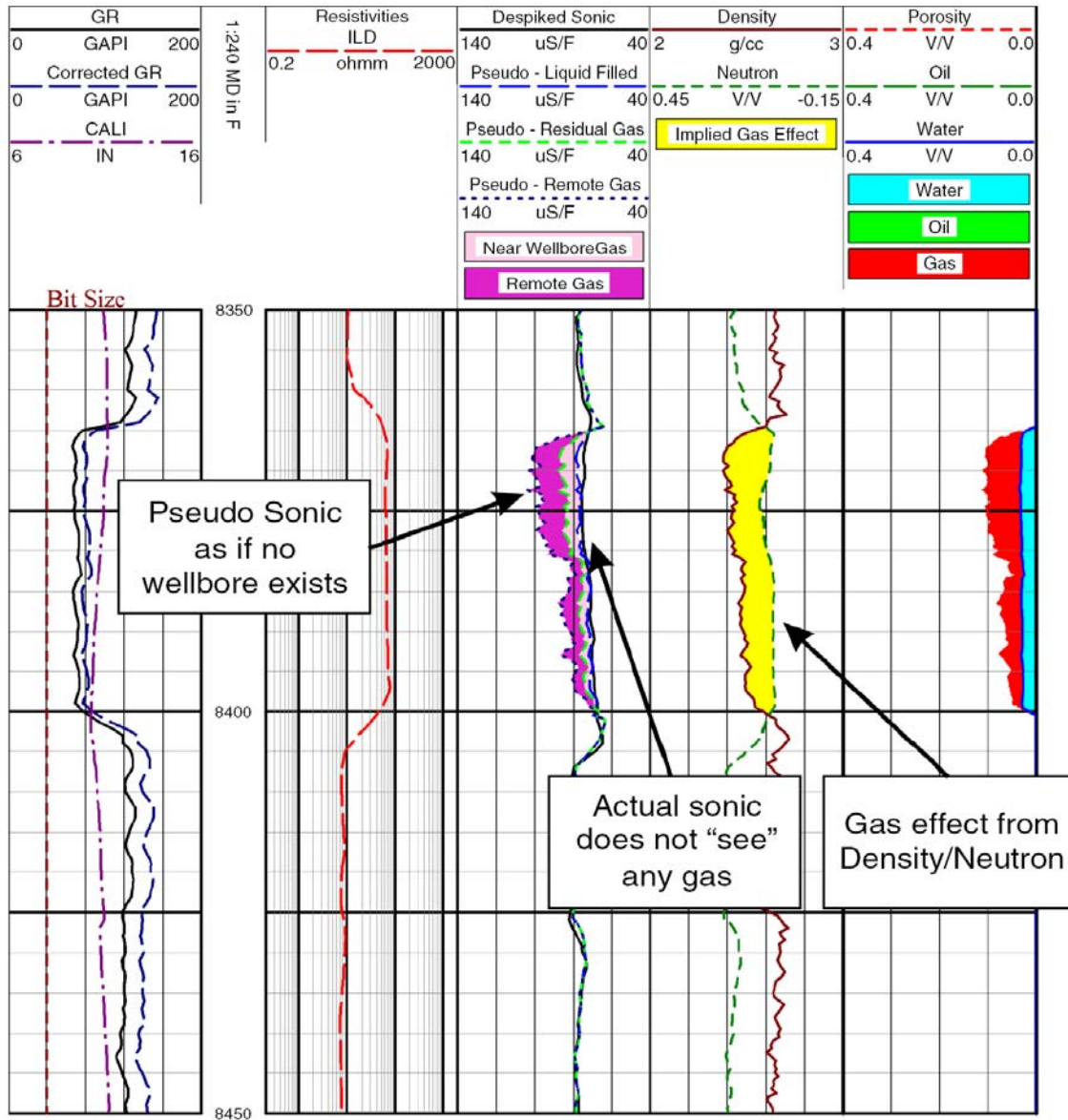


Figure 3: Comparisons of three pseudo sonic logs with the original log.

The synthetic seismograms in Figure 4 show significant differences dependent on the pseudo sonic used. The lesson of the example is that if you do not know what fluid the sonic log is measuring (and it is not always residual gas), then any synthetic seismograms using the sonic will also have problematic meaning.

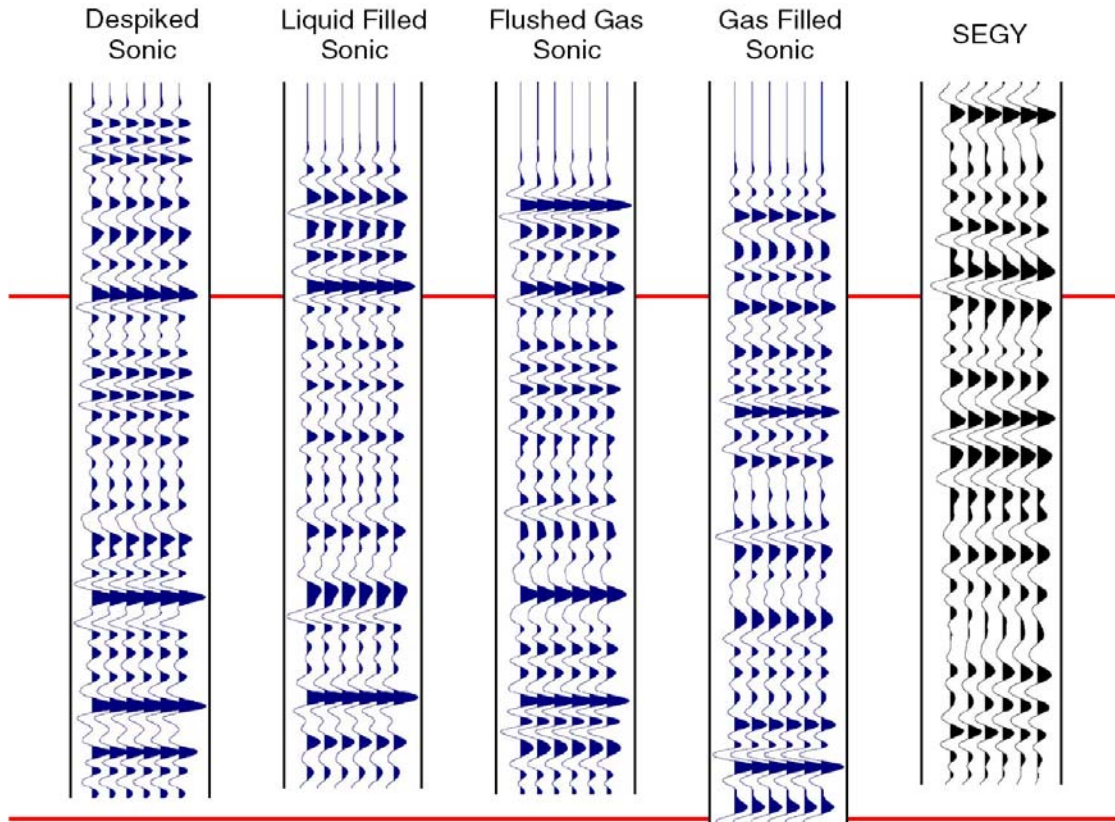


Figure 4: Comparing synthetic seismograms using four different synthetic sonic approaches.

Additionally, if a missing log is a shallow-reading device (such as a sonic log) but is generated from deeper reading devices (such as the deep resistivity, neutron, or density), then the resulting pseudo log has questionable value.

Summary

In each case, missing data can be generated using different methods. However, care must be used to ensure that the generated information has integrity and is appropriate for the reservoir.

- 1) Steps must be taken to clean-up and validate all input data to any synthetic generation method. This rather obvious step is often one of the easiest to overlook.
- 2) It is important for the interpreter to have an understanding of what the correct answer might be. When generating synthetic data, it is conceivable to generate nearly any answer. It is up to the interpreter to ensure the answer used makes geological and geophysical sense.