

# The Oz Machine: A Java applet for interactive instruction in geological log interpretation

Geoffrey C. Bohling

John H. Doveton

Kansas Geological Survey, 1930 Constant Avenue, Lawrence, Kansas 66047, USA

## ABSTRACT

Geophysical well logs represent measurements of a variety of properties of the rocks and fluids encountered by a well bore and are used by petroleum industry analysts to guide decisions regarding further well development and investigation. Nuclear logs of natural gamma rays, neutron moderation, electron density, and photoelectric absorption are extremely common and are sensitive measures of rock types and mineral compositions. The Oz Machine is a Java applet providing online, interactive instruction in geological interpretation of these nuclear well logs. It employs a simple Markov chain simulation to generate a synthetic sequence of lithologies (rock types) and then generates a suite of corresponding well logs based on a mineralogical recipe for each lithology and the typical log responses for each mineral. The resulting synthetic logs are displayed, and the student paints a geological interpretation of the logs into the depth track, selecting from a palette of lithologies presented next to the log display. The realism of the simulated log suite is enhanced by inclusion of random variation in the mineralogical composition of each lithology and application of a smoothing filter to emulate tool resolution effects. Despite the simplicity of the underlying simulation, the generated lithological and log sequences are surprisingly realistic, providing an essentially endless supply of "mirror-world" exercises in geological log interpretation.

**Keywords:** well logs, online instruction, Markov chains, stratigraphy, Java.

## INTRODUCTION

Petrophysicists have long used geophysical well logs to deduce the distribution of porosity and fluid saturations in petroleum reservoirs. Contemporary logging tools are also sensitive

to mineralogical variations in the surrounding rock matrix, providing a means for interpreting lithological variation in addition to variations in porosity and saturation. Along with providing information about the general geological framework of a region, the ability to infer lithological variation from well logs is critical in many reservoir characterization projects, since these variations often exert a first-order control on the distribution of porosity and permeability (Selley, 1998). The necessary interpretive skills are generally not acquired in a conventional university geological curriculum, but are of increasing importance to professional subsurface geologists. In this paper, we present a Java applet, the Oz Machine, which provides an interactive, online exercise in geological interpretation of wire-line logs. The Oz Machine was developed to accompany an introductory tutorial in geological interpretation of wire-line logs (*Reading the Rocks from Wireline Logs*, available at <http://www.kgs.ku.edu/PRS/ReadRocks/portal.html>), so that students could proceed directly from a narrative explanation of basic log interpretation principles to an interactive exercise employing those principles.

Since the recording of the first resistivity log in 1927, wire-line logs have been a principal source of information available to geologists for the study of rocks in the subsurface. Wire-line logs are physical properties of formations penetrated by a borehole and recorded by a sonde lifted on a cable ("wire line") from the bottom of the hole. The geological information extracted from these logs has generally been restricted to the determination of the tops and bottoms of key formations. By correlating tops between wells, subsurface geology can be mapped in terms of structure and lateral changes in thickness within a lithostratigraphic framework. The composition of correlative formations is given by observations on drill-cuttings and core data, when available. Resistivity logs provide limited information on rock compositions because the ability of a rock to conduct electrical current is

mostly controlled by the pore volume content of saline formation water. However, the marked difference in the resistivities of shales and rock types other than shales allows a simple log to be prepared that can be augmented by rock assignments made from ancillary geological information.

A significant increase in the geological information content of logs occurred when nuclear logging techniques were introduced to supplement electrical methods. Natural gamma rays were first recorded by a logging tool in the late 1930s, initially by a Geiger counter, which was soon replaced by a scintillation crystal device. Natural gamma rays emanating from formations in the borehole wall have sources in the potassium-40 isotope and isotopes of the uranium and thorium series. Since these isotopes tend to occur in greater abundance in shales, the major application of the gamma-ray log is in the discrimination of shales from other lithologies. While the gamma-ray log is a passive measurement of natural radiation, neutron and density logs are records of nuclear processes in the formations caused by radioactive sources on the logging tool. The neutron log is essentially a measure of the hydrogen concentration within a formation caused by the reduction in energy of fast neutrons in collisions with hydrogen nuclei. The density log is a measure of the electron density of the formation computed from the reduction of the gamma-ray flux emitted by a radioactive source on the tool, and it can be converted to a close approximation of the mass density. Finally, the photoelectric index records the interaction of low-energy gamma rays with formation nuclei and is a direct function of the aggregate atomic number.

Collectively, these nuclear logs are used mainly to produce the best estimates of pore volume in reservoir rocks free of the disruptive effects of rock compositional variation and changes in mineralogy. In an industry-wide standard convention, the gamma-ray, neutron, density, and photoelectric index logs are plotted

together so that petrophysicists can discern porosity and lithological variability immediately. The interpretation of rock types from log types is an inverse form of reasoning because it works from effect (the log responses) to cause (the earth model). The process is further complicated by potential ambiguities because several different earth models can cause similar log responses. By contrast, a forward model of logs generated from an earth model is an explicit result that is computed relatively easily by convolving a rock sequence with the physical properties of the mineral and fluid components. This observation is the basis for the Oz Machine, which was designed as an interactive teaching device when training neophyte petrophysicists to make lithological inferences from wire-line logs of common occurrence.

### Design Philosophy of the Oz Machine

The Oz Machine uses a simple one-dimensional Markov chain simulation to generate the vertical sequence of lithologies that serves as the basis for the log interpretation exercise. Vistelius (1949) first introduced the concept of applying Markov chain analysis to the study of sedimentary successions, and a wide variety of papers on the topic has appeared in the intervening years. However, the application of Markov chains within the Oz Machine is purely as a simulation device rather than a method for the analysis of outcrop or borehole sequences. The earliest publication that described the use of a transition probability matrix within a computer program to generate synthetic successions was written by Krumbein (1967). The book by Harbaugh and Bonham-Carter (1970) includes an extensive and useful review of Markovian sedimentary succession modeling and its implementation in simple computer programs. More recent publications have tackled the more difficult (but necessary) procedures that are involved in the simulation of lithologies in two or three spatial dimensions, such as a North Sea field application by Moss (1990) and simulation of fluvial fan deposits in the Loranca Basin of Spain by Elfeki and Dekking (2001). Carle and Fogg (1996, 1997) and Weissmann et al. (1999) presented applications of three-dimensional Markov chain simulation of facies distributions based on continuous-lag transition probability models, demonstrating that the transition probability approach provides a more geologically intuitive means of specifying the spatial structure than more traditional simulation techniques based on spatial covariance functions.

The name "Oz Machine" perhaps requires some explanation. The "Machine" component refers to the core engine of the applet that is the

operation of the transition probability matrix in the generation of a Markovian stratigraphic succession. "Oz" draws on the fictional associations with Kansas. Some of the characters that Dorothy met in the Land of Oz were caricatures of family members and friends back in Kansas. While the output of the Oz Machine is intended to be an acceptable simulation of the Kansas (and by extension, U.S. Midwest) subsurface, its limitations as a rigorous representation are obvious from its design features. First, it is a parametric model that is controlled by a set of transition probabilities. As an immediate consequence, simulated lithology thicknesses, which are dictated by the main diagonal elements of the transition probability matrix, follow a geometric distribution (Krumbein and Dacey, 1969), as contrasted with observed thicknesses, which are more closely matched by a log-normal distribution (Pettijohn, 1957). This limitation could be overcome by modifying the process to an embedded Markov chain in which the order of lithologies was simulated and linked with a distribution that more closely mimicked actual lithology thicknesses. However, the purpose of the Oz Machine is not to educate students in lithology thickness, but in the association between log responses and lithology. In the short (100 ft) sections generated by the Oz Machine, the lithology thicknesses are not unreasonable. At the same time, the student might notice that the succession of rock types might seem to be a little "accelerated," in the sense that there is generally more variety than would be seen in a typical Kansas subsurface section of equivalent length. Again, the spacing of the lithologies in the simulation is controlled by the transition probability matrix and can be computed explicitly as matrices of mean first passage times and their variances (Doveton and Duff, 1984). Although the enhanced vertical variation generated by the Oz Machine is slightly unrealistic, it is beneficial for the purposes of the exercise, in that the simulated sequences provide a more engaging challenge than they would if they displayed a more realistic (that is, more monotonous) level of variation.

In summary, the Oz Machine creates a "mirror world" (cf. Gelernter, 1993) of the Kansas subsurface in which the desired features of the associations between ideal rocks and log responses are honored and other aspects are emulated to varying degrees of credibility. A simple analogy can be made with the iconic London Underground Map, which was controversial when it was introduced as a circuit diagram in 1933 because the match between stations and their geographic locations was only approximate. But today, millions of travelers

instinctively use the mirror-world map of the Tube to find their way around London in the subsurface, but realize the limitations of the map as an accurate guide for street navigation.

At each step of the simulation, the log responses of the lithology are computed by convolving the proportions of the mineral content with the log properties of the mineral end members. The set of log response equations provides either exact (if ideal) solutions or close approximations to nonlinear relationships. The log responses of a wide variety of sedimentary minerals have been tabulated (including revisions and updates) for many years in chart books published by Schlumberger and are reported episodically on their Web site at [www.oilfield.slb.com](http://www.oilfield.slb.com). Because the output of the Oz Machine is a forward-modeled representation of log responses of ideal rocks based on their mineral properties, its education in the hypothetical context should be used as a precursor to the study of real logged successions with their lithological complexity, borehole environmental problems, and vagaries in tool performance.

The Oz Machine has been beta-tested for two years at the University of Kansas as a component of a course in geological log analysis. Following introductory lectures on tool theory and log analysis methods, the students are provided with the Oz Machine to hone their skills in the basics of the interpretation of geology from logs. The student grade is based entirely on their completion of an individualized "millipede," a thousand feet of a Kansas subsurface succession logged by spectral gamma-ray, lithodensity, and neutron porosity tools. The students are encouraged to consider this as a petrophysical mapping project to be approached in a similar spirit to their observations and description of a lengthy outcrop in the field. Following this analogy, the Oz Machine then becomes an introductory lab module where students would be presented with selected examples of common lithologies as a preparation for fieldwork on successions with wider variabilities in rock properties and weathering aspects. For interested readers, the substance of this course is contained in Doveton (2004) as a CD reissue of the Society for Sedimentary Geology (SEPM) Short Course Notes #29 published in 1994. The most common request from beta-users at locations remote from Kansas has been for the inclusion of alternative transition probability matrices with lithologies keyed to sedimentary successions elsewhere. Such adaptations would be conceptually trivial (but labor-intensive) to implement and could be developed either from hypothetical transitions or based on statistics counted from outcrop or core.

Simulation Details

A first-order Markov chain represents a sequence of discrete states in which the transitions between states can be described according to a fixed set of one-step transition probabilities (Ross, 2000; Duda et al., 2001). That is, if  $\{X_n, n = 0, 1, 2, \dots\}$  represents the sequence, with the possible states represented by a set of integer values, then a first-order Markov chain is characterized by the set of one-step transition probabilities  $P_{ij} = \Pr[X_{n+1} = j | X_n = i]$ , the probability that state  $j$  will follow state  $i$ . Within the Oz Machine, the simulation process employs a matrix that specifies single-step upward transition probabilities between eighteen different lithologies. Each run generates a 100 ft vertical section, composed of fifty 2 ft intervals, that is built from the bottom up. The entire set of upward transition probabilities,  $P_{ij}$ , is stored in an  $18 \times 18$  transition probability matrix. This is schematically illustrated in Figure 1, and the grouping into six evaporite lithologies, nine lithologies typical of marine carbonate-shale successions, and four typical of clastic coal-bearing successions is shown. If interval number  $n$  consists of lithology number  $i$ , then the lithology for interval  $n + 1$  is determined by selecting the transition probabilities from row  $i$  of the transition probability matrix, converting these to a cumulative density function for the lithology of interval  $n + 1$ , and drawing a lithology at random from this cumulative density function. That is, the lithology for interval  $n + 1$  is taken as the first lithology,  $j$ , for which

$$\sum_{k=1}^j P_{ik} \geq u, \quad (1)$$

where  $u$  is a uniform random number. Figure 2 illustrates the process of selecting the lithology for an interval overlying a dolomite interval. The lithology for the first (deepest) interval in the section is determined by a random draw from a cumulative density function built from a vector of eighteen entry probabilities.

Following the creation of a sequence of lithologies, the program generates a corresponding suite of logs through a three-step process. First, each lithology is mapped to a set of component minerals according to a set of prescribed recipes with a small degree of random variation. So, for example, the quantitative composition of a “dolomitic limestone” is dictated by its syntax of calcite as the dominant mineral, dolomite as the subordinate mineral, and a pore volume as would be typical for a dolomitic limestone. Each two-foot interval of constant lithology is subdivided into four half-foot intervals, to match the digital sampling for

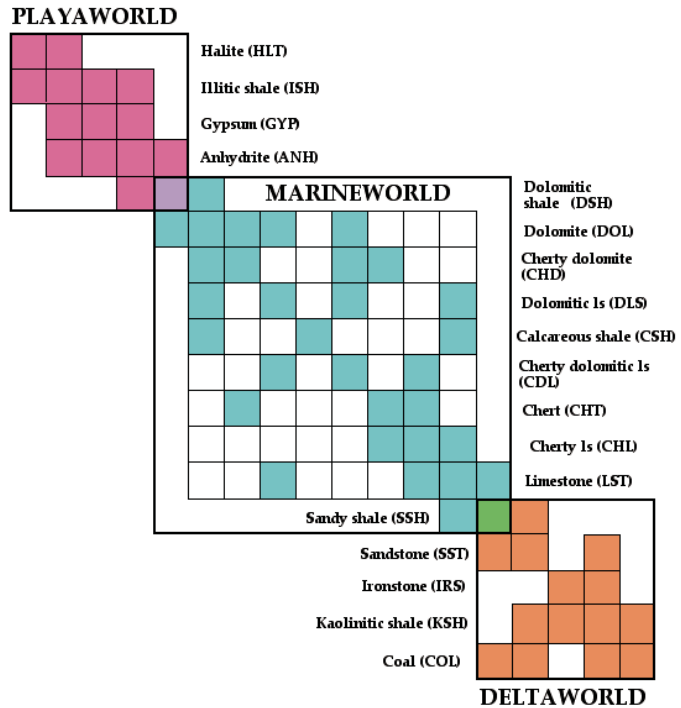


Figure 1. Schematic representation of transition probability matrix employed in the Oz Machine. The nonzero transition probabilities, indicating possible upward transitions lithology  $i$  (rows) to lithology  $j$  (columns, in same order as rows), are shaded (ls—limestone).

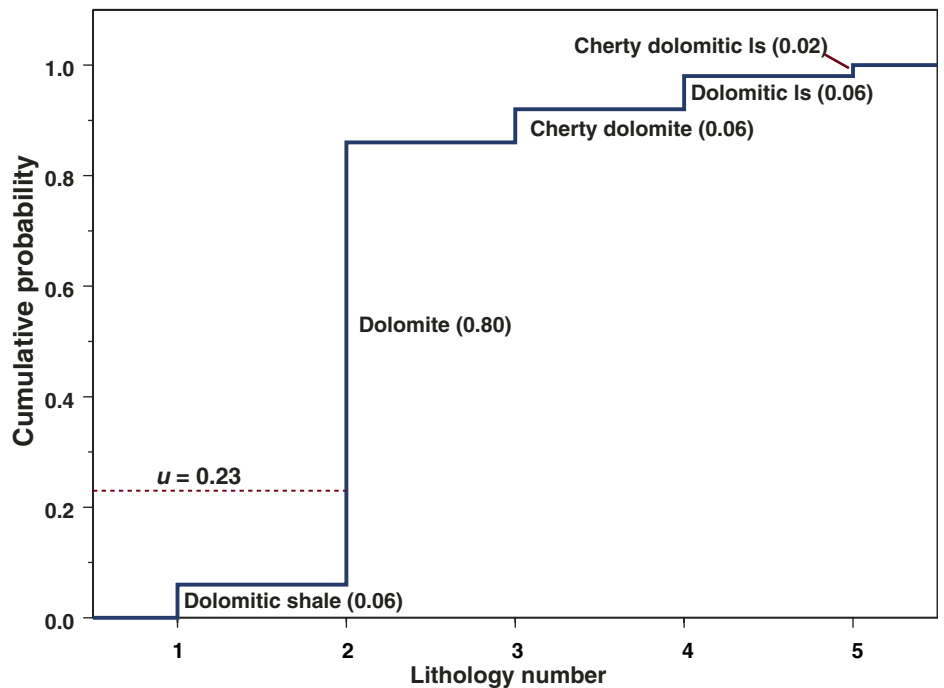


Figure 2. Illustration of lithology selection for an interval overlying an interval of dolomite, including only those lithologies with nonzero upward transition probabilities (in parentheses) from dolomite. Next lithology up is that for which cumulative probability exceeds a uniform random number,  $u$ .

a typical logging run, and the mineral composition for each half-foot interval is given by adding a small degree of noise (averaging out to a few percent variation) to each nonzero component in the ideal mineral composition and then rescaling the result to 100%. The minor random component added introduces some compositional variability but is limited in size so that lithological integrity is maintained without, for example, the transmutation of dolomitic limestones to limestones, dolomites, or calcitic dolomites. The component minerals employed are halite, gypsum, anhydrite, illite, dolomite, quartz, calcite, iron, kaolinite, coal, and finally a mineral named “porosity,” which is equated with pore fluids, represented by freshwater mud-filtrate.

Next, the volumetric mineral composition vector,  $\mathbf{V}$ , for each half-foot interval is mapped to the vector of log responses,  $\mathbf{L}$ , through the forward composition equation,

$$\mathbf{L} = \mathbf{C}\mathbf{V}, \quad (2)$$

where  $\mathbf{C}$  is a matrix specifying the log properties for each component (Doveton, 1994). So, for example, modeling of the density log response of a cherty dolomitic limestone by this algorithm would be given by:

$$\rho_b = L \cdot \rho_L + D \cdot \rho_D + Q \cdot \rho_Q + F \cdot \rho_F, \quad (3)$$

which is a mass-balance relationship that computes the bulk density as the sum of the products of the proportions of calcite, dolomite, chert, and pore fluid ( $L$ ,  $D$ ,  $Q$ ,  $F$ ) by their respective log densities. Application of the matrix algorithm allows four synthetic logs to be computed and displayed by the Oz Machine as gamma ray, neutron porosity, density porosity, and photoelectric index curves. The log property matrix,  $\mathbf{C}$ , is “hard-wired” in the code and is specified without noise, so that variations in the synthetic logs are due to the minor variations in the mineralogical composition for each lithology, discussed in the previous paragraph, except for the gamma-ray log. A small degree of additional random variation is added to the gamma-ray log to reflect the relatively higher degree of stochastic variation in this log.

Finally, the sequence of log values generated from the mineral compositions is then convolved with a five-point smoothing filter as a simple emulation of tool resolution. If  $x_i$  represents the original value of a particular log at depth  $i$ , then the smoothing process replaces this value with

$$\tilde{x}_i = 0.05x_{i-2} + 0.20x_{i-1} + 0.50x_i + 0.20x_{i+1} + 0.05x_{i+2}. \quad (4)$$

Although this smoothing filter is not designed to reproduce the actual smoothing characteristics of any particular logging tool, it yields a degree of realism that is quite adequate for the purposes of this learning tool.

### Program Design

The program consists of a small set of Java classes with a fairly straightforward information flow. The Markov chain class generates a sequence of integer values, each representing the lithology for each two-foot interval. This sequence is generated from the lowest interval upward, based on the  $18 \times 18$  upward transition probability matrix specified directly in the Java code. The Lithology class simply specifies the properties for each lithology, including its type mineral composition, its displayed name, and the name of the image file containing the displayed symbol for that lithology. Thus, the Lithology class is required to turn the sequence of integers generated by the Markov chain class into a sequence of lithologies. The Log Suite class is responsible for converting this sequence of lithologies into a sequence of values for all four of the displayed logs. The code in Log Suite steps through the entire sequence at half-foot increments and, at each depth, calls a routine in the Lithology class to return a mineral composition based on the lithology at that depth (the type composition for that lithology plus a

random perturbation), and then computes the set of logs corresponding to that mineral composition. The Log Suite code then adds additional random noise to the gamma-ray log and applies the five-point smoothing filter to each log, as described above, to produce the final set of synthetic logs.

The main class, Oz Machine, implements the graphical user interface (GUI) and coordinates the activities of the other classes. As is typical with many programs of this nature, significantly more code is devoted to handling interactions with the user than to the underlying mathematical computations. However, Java’s Swing library of high-level GUI components considerably eased development of the user interface code.

### Program Operation

We have tried to make the program operation as simple as possible so that the student can expend his or her mental effort learning about geological log interpretation, rather than figuring out how to use the software. The introductory Web page (at <http://www.kgs.ku.edu/PRS/ReadRocks/OzIntro.html>) contains a brief introduction to the software along with links to further explanatory material and a link to the online tutorial in geologic interpretation of logs that the Oz Machine was originally designed to accompany. There is also a link to a site from which the student can download

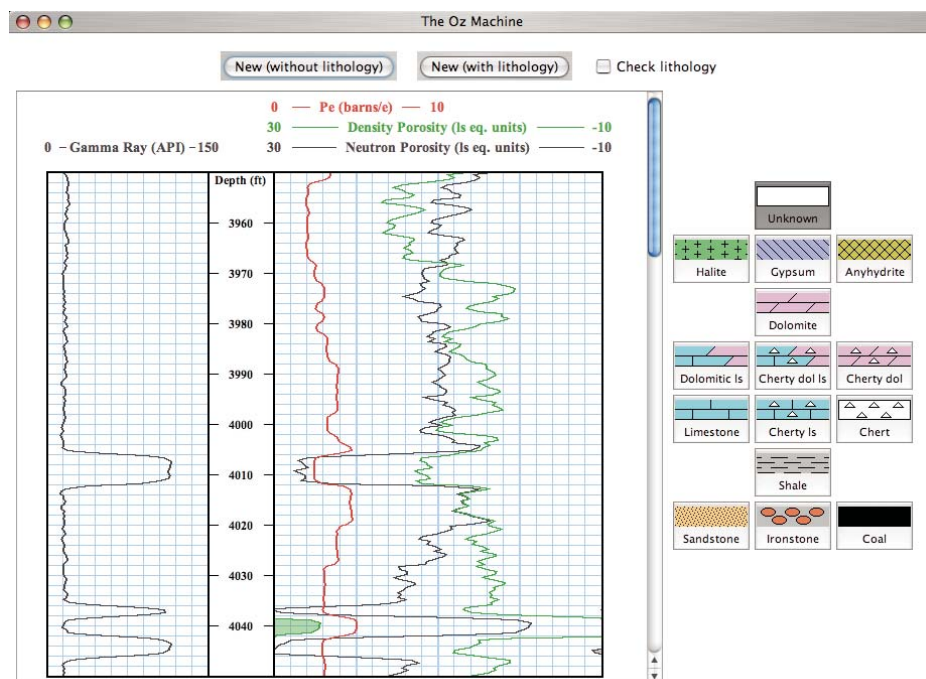


Figure 3. An example Oz Machine display in its initial state, before the student has started filling in the depth track with selected lithologies.



a recent version of the Java runtime environment, should that be required.

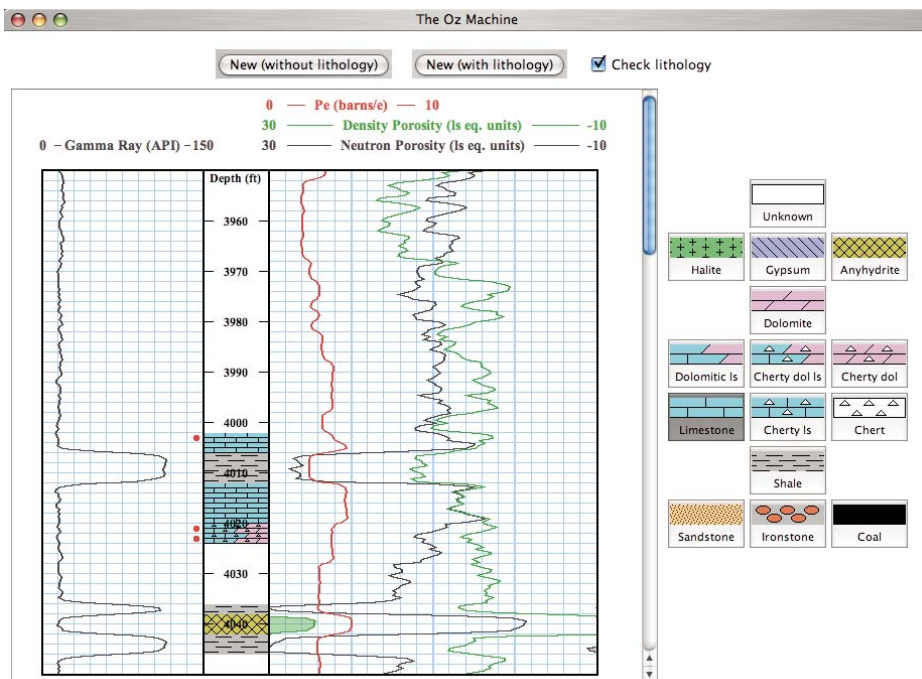
Upon launching the Oz Machine, the student will see a display like the one shown in Figure 3, although with a different depth range and a different log sequence. (The top depths for each sequence are randomly generated.) The gamma-ray log is displayed in the left track and the neutron porosity, density porosity, and Pe logs are all displayed in the right track, in a display that follows petroleum industry log-plotting conventions fairly closely. The student begins the exercise by selecting one of the lithology symbols in the palette on the right and then clicking in the depth track to paint in that lithology at selected depths. Each assignment (click) fills in a two-foot interval with the chosen lithology, where the intervals correspond to the depth increments in the blue-line grid. The “Unknown” lithology button at the top of the palette serves as an eraser, allowing the user to return any two-foot interval to its unassigned state.

Figure 4 shows the same example as Figure 3 after the student has worked partway through the assignment. The student has first picked the shale intervals, indicated by their high gamma-ray values, and then the anhydrite interval, indicated by its anomalous porosity log values. Then, the task of unraveling the more subtle log variations among the carbonate rocks (limestones and dolomites) that constitute most of the rest of the section has been started. The “Check lithology” check box has also been clicked, so that the program indicates when the user has made an incorrect assignment. Incorrect assignments are marked by a red circle to the left of the corresponding two-foot interval in the depth track. Figure 5 shows the display after the student has successfully completed the exercise.

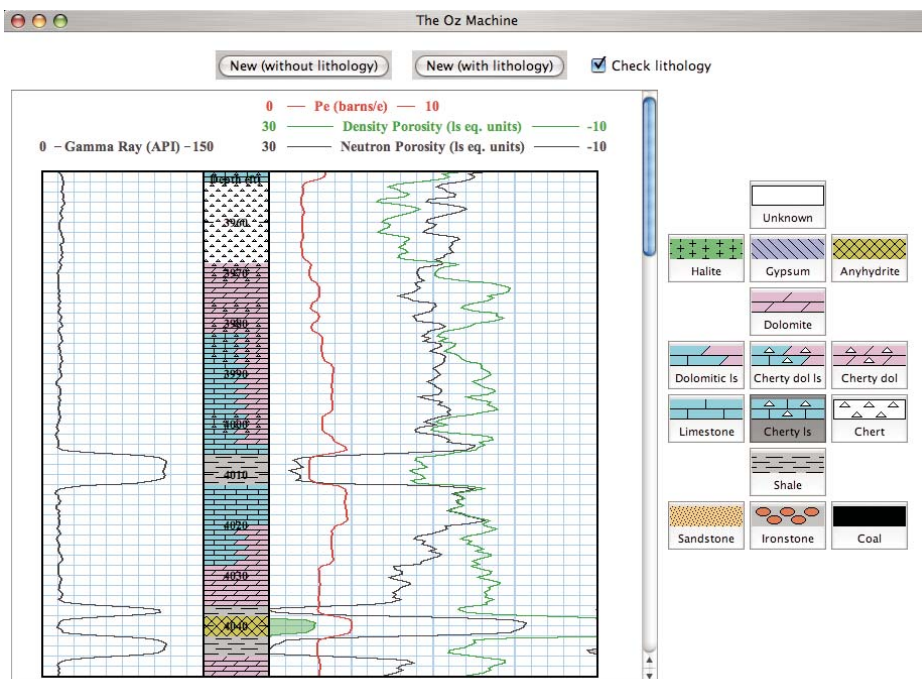
At any time, a click of the **New (without lithology)** button will generate a new display with a different log suite and a clear depth track waiting to be filled in. Alternatively, clicking **New (with lithology)** will generate a display with the depth track filled in with the true lithology, allowing students to generate examples for familiarizing themselves with typical log-lithology associations. Figure 6 shows a montage of such displays, which emphasizes that the Markov chain simulation process is capable of emulating a variety of depositional sequence styles.

**CONCLUDING REMARKS**

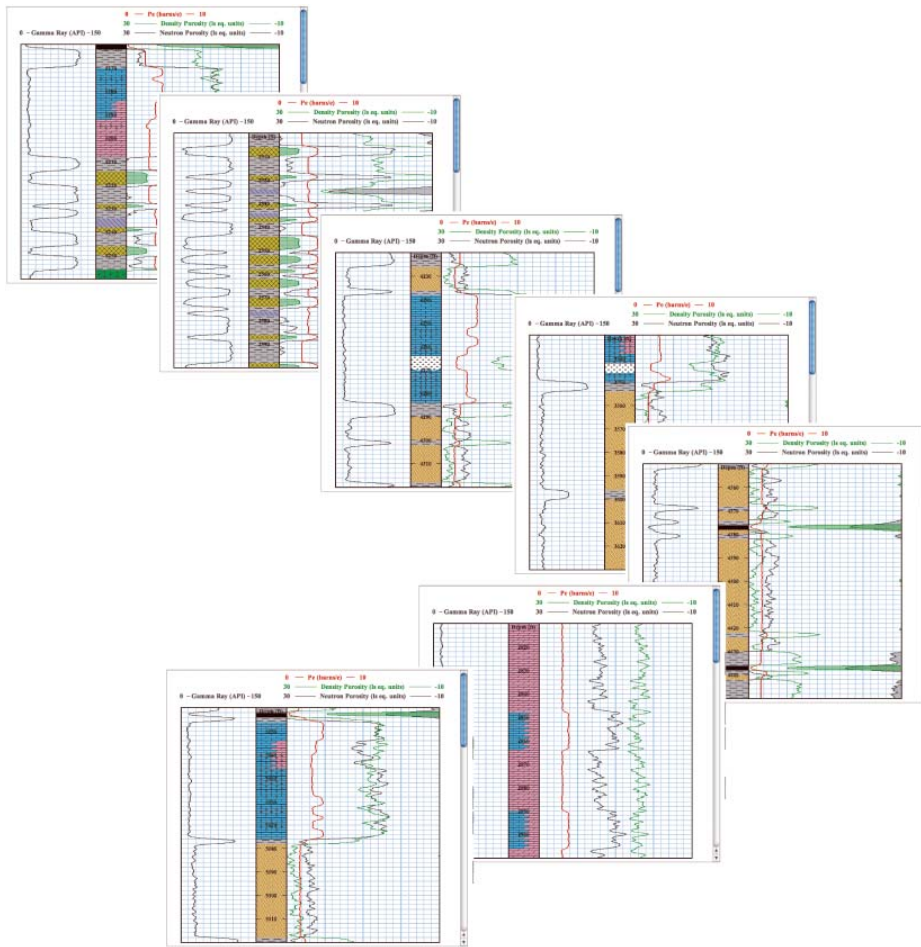
Despite the relatively limited lithological palette and the use of a transition probability matrix tuned to the North American Midcontinent, the Oz Machine renders a variety of compellingly realistic log suites, from wildly oscillating evaporite sequences through monotonous marine



**Figure 4.** An example Oz Machine display where the student has started filling in the depth track with lithology picks and has turned on “Check lithology,” which flags incorrect picks with red circles.



**Figure 5.** An example Oz Machine display where the student has successfully completed the exercise.



**Figure 6. Montage demonstrating a variety of lithological and log sequences generated by the Oz Machine.**

carbonate sequences to blocky terrestrial sand-shale sequences and admixtures of the three. Currently, the Oz Machine provides an exercise for self-motivated students and does not include any code for scoring or evaluating the student's performance other than the flagging of incorrect lithology picks. However, we feel that the program is simple and compelling enough to hold the attention of users who are interested in learning about geological log interpretation, be they students in a graduate course or oil industry professionals looking for a crash course or refresher on this topic.

The transition probability matrix that is the core of the Oz Machine is a simple template for hypothetical rock successions in the subsurface of the U.S. Midcontinent. Future program enhancements will allow users to enter their own transition probability matrices and matrix component log properties, derived from real successions using measurements from outcrop or core or representing more generalized depositional models; this will provide the ability to mimic stratigraphic sequences at any location. Finally, the expansion of the Oz

Machine to incorporate oil and gas reservoirs could be achieved with the addition of a capillary pressure simulator to model fluid saturations within the pore space and to compute a matching resistivity log.

#### REFERENCES CITED

- Carle, S.F., and Fogg, G.E., 1996, Transition probability-based indicator geostatistics: *Mathematical Geology*, v. 28, no. 4, p. 453–476, doi: 10.1007/BF02083656.
- Carle, S.F., and Fogg, G.E., 1997, Modeling spatial variability with one and multidimensional continuous-lag Markov chains: *Mathematical Geology*, v. 29, no. 7, p. 891–918, doi: 10.1023/A:1022303706942.
- Doveton, J.H., 1994, Geologic log analysis using computer methods: American Association of Petroleum Geologists (AAPG) Computer Applications in Geology, No. 2: Tulsa, Oklahoma, The American Association of Petroleum Geologists, 169 p.
- Doveton, J.H., 2004, Geologic Log Interpretation: CD-ROM reprint of Society for Sedimentary Geology Short Course Notes 29, 169 p.
- Doveton, J.H., and Duff, P.McL.D., 1984, Passage-time characteristics of Pennsylvanian sequences in Illinois: Ninth International Carboniferous Congress: *Compte Rendu*, v. 3, p. 599–604.
- Duda, R.O., Hart, P.E., and Stork, D.G., 2001, Pattern Classification (second edition): New York, John Wiley & Sons, 654 p.
- Elfeki, A., and Dekking, M., 2001, A Markov chain model for subsurface characterization: Theory and applications: *Mathematical Geology*, v. 33, no. 5, p. 569–589, doi: 10.1023/A:1011044812133.
- Gelernter, D., 1993, *Mirror Worlds*: New York, Oxford University Press, 256 p.
- Harbaugh, J.W., and Bonham-Carter, G., 1970, *Computer Simulation in Geology*: New York, Wiley-Interscience, 575 p.
- Krumbein, W.C., 1967, FORTRAN IV computer programs for Markov chain experiments in geology: *Kansas Geological Survey Computer Contribution* 13, 38 p.
- Krumbein, W.C., and Dacey, M.F., 1969, Markov chains and embedded Markov chains: *Mathematical Geology*, v. 1, no. 1, p. 79–96, doi: 10.1007/BF02047072.
- Moss, B.P., 1990, Stochastic reservoir description: A methodology, in Hurst, A., Lovell, M.A., and Morton, A.C., eds., *Geological Applications of Wireline Logs*: Geological Society of London Special Publication 48, p. 57–75.
- Pettijohn, F.P., 1957, *Sedimentary Rocks*: New York, Harper, 718 p.
- Ross, S.M., 2000, *Introduction to Probability Models* (seventh edition): San Diego, Academic Press, 693 p.
- Selley, R.C., 1998, *Elements of Petroleum Geology* (second edition): San Diego, Academic Press, 470 p.
- Weissmann, G.S., Carle, S.F., and Fogg, G.E., 1999, Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics: *Water Resources Research*, v. 35, no. 6, p. 1761–1770, doi: 10.1029/1999WR900048.
- Vistelius, A.B., 1949, On the question of the mechanism of the formation of strata: *Academy of Sciences of the USSR, Earth Science Section*, v. 65, p. 191–194.

MANUSCRIPT RECEIVED 14 MARCH 2006  
 REVISED MANUSCRIPT RECEIVED 5 JUNE 2006  
 MANUSCRIPT ACCEPTED 23 JUNE 2006