

## New Data Reduction Tools and their Application to The Geysers Geothermal Field

M. Ali Khan<sup>1</sup>, Rich Estabrook<sup>2</sup>

<sup>1</sup>Division of Oil, Gas, and Geothermal Resources, 50 D Street # 300, Santa Rosa, CA 95404 (USA)

<sup>2</sup>Bureau of Land Management, 2550 North State Street, Ukiah, CA 95482 (USA)

<sup>1</sup>akhan@conservation.ca.gov, <sup>2</sup>Richard\_Estabrook@ca.blm.gov

DOGGR WEBSITE: <http://www.conservation.ca.gov>

**Keywords:** The Geysers, Production, Injection, Superheat, Decline Curve, Data Reduction, Visualization

### ABSTRACT

Microsoft Excel based (using Visual Basic for Applications) data-reduction and visualization tools have been developed that enable the user to numerically reduce large sets of geothermal data to any size. The data can be quickly sifted through and graphed to allow their study. The ability to analyze large data sets can yield responses to field management procedures that would otherwise be undetectable. Field-wide trends such as decline rates, response to injection, evolution of superheat, recording instrumentation problems and data inconsistencies can be quickly queried and graphed. Here we demonstrate the application of these tools to data from The Geysers Geothermal field. We believe these data-reduction tools will also be useful in other applications, such as oil and gas field data, and well log data. A copy of these tools may be requested by contacting the authors.

### 1. INTRODUCTION

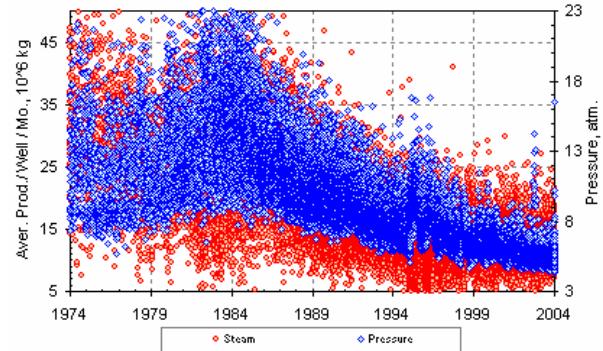
The California Department of Conservation, Division of Oil, Gas, and Geothermal Resources (DOGGR), and the US Bureau of Land Management (BLM) receive monthly production, injection, and related data from operators of oil and gas and geothermal wells in California. Most of these data are non-confidential and available through the DOGGR website (<http://www.conservation.ca.gov/>). For data visualization, very powerful Microsoft Excel based tools have been developed that may either be used directly or easily modified to fit individual needs. They may be used to: 1) easily organize and retrieve data-groups, 2) reduce large data sets to meaningful sizes, and 3) graphically present the data. With the help of these tools the user can quickly and easily review data in many different ways. Thus, data trends, as well as discrepancies, become more visible and easier to discern.

In Figure 1, field-wide average production and average wellhead pressure, reduced at a rate of 6:1, are plotted. There are just too many points to interpret, however, when the same data are “appropriately reduced,” in this case 1050:1, (Figure 2) a clear trend and useful information emerges (details are discussed in section 6.1).

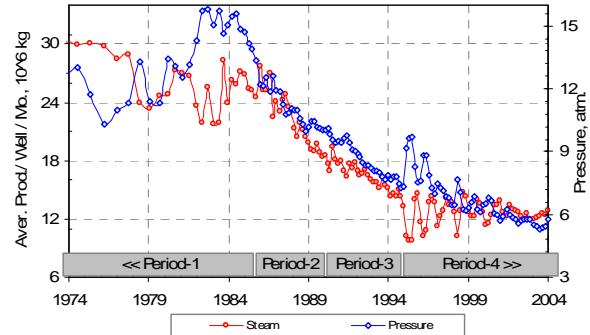
### 2. A BRIEF HISTORY OF THE GEYSERS

The Geysers Geothermal field, which is located about 70 miles north of San Francisco, California, USA, started production in 1960 with a 12 MW power plant. The field development picked up at a rapid pace from 1979 through 1989, although wellhead flowing pressure started showing a decline by 1984. Despite the drilling of new wells and an increase in installed capacity, the steam production peaked at 112 billion kg in 1987 (Figure 3). From 1976 through

1980 the mass replacement rate (i.e., the fluid re-injection rate) was about 24%, which is approximately the cooling tower recovery at The Geysers.



**Figure 1: The Geysers. Field-wide average Production and average Injection vs. Time, reduced at 6:1.**

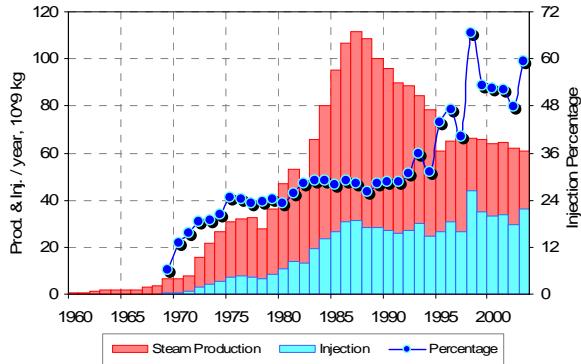


**Figure 2: The Geysers. Field-wide average monthly production rates per well month, and average wellhead flowing pressures.**

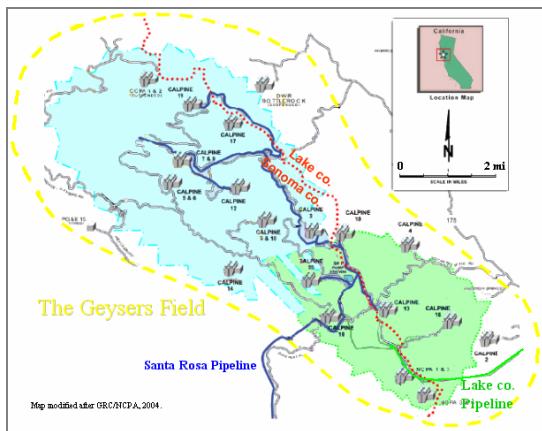
From 1980 through 1993, streams and creeks were tapped, thereby increasing the mass replacement rate to about 28%. From 1995 through 1997, due to major steam curtailments, and from 1997 onward due to additional Lake County pipeline injection (Figure 4), the mass replacement increased to about 55%. The Lake County 42-km pipeline transports about 1.05 million kg per month of secondary treated effluent to The Geysers for injection, which results in additional steam.

Encouraged by the success of the Lake County pipeline, the City of Santa Rosa, in collaboration with other municipalities, built a 65-km long pipeline to deliver about 1.25 million kg per month of tertiary treated effluent into The Geysers. The pipeline began operation in December 2003. The current mass replacement from both pipelines and other sources is about 80% of production (Stark et al.,

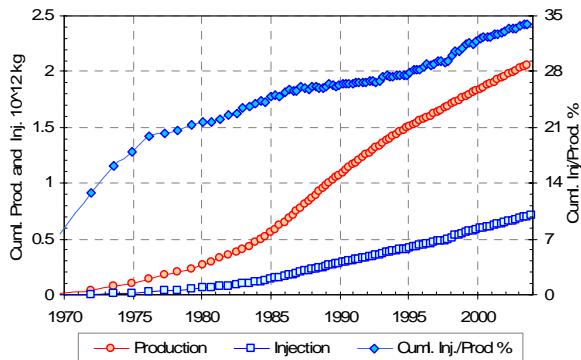
2004). This has resulted in a sustained increase in steam production, decrease in non-condensable gases, improved electric generation efficiency, and lower air emissions. The Geysers has become the largest heat mining operation in the world. By the end of 2003, The Geysers had produced 2,088 billion kg of steam (Figure 5), and injected 710 billion kg of fluids, resulting in a net mass replacement of 34%.



**Figure 3: The Geysers yearly steam production, and injection rates and mass replacement percentages.**



**Figure 4: The Geysers Geothermal Field. Areas of Lake County injection project are indicated in green and Santa Rosa injection project in blue.**



**Figure 5: The Geysers. Field-wide cumulative production, cumulative injection and cumulative net mass replacement.**

### 3. THE DATA SET

DOGGR maintains and makes available through its website the monthly well reports that operating companies are required to file for oil, gas, and geothermal wells. This also includes other related wells such as disposal and observation wells. For geothermal wells, these data consist of monthly production, injection, wellhead pressure (mostly flowing pressure), temperature, instantaneous production and injection rates, well status and well type. Well name, well identification numbers (API), and well location are also available. A detailed description of the data set may be accessed at the DOGGR website. Data submitted by the operators are checked against a set of numerical constraints, and appended into a protected database.

### 4. THE TOOLS

There are over one million data-points in The Geysers field database. With such a large number of data-points, a graph showing even one field is too crowded to indicate any meaningful trend (Figure 1). Even when a graph is generated, changing data-points for repeated analysis is cumbersome and slow. Therefore, to improve the data processing, a set of tools was developed and is presented here with examples from The Geysers field. Microsoft Excel is the front-end working platform for all of these computer tools. The data may be imported seamlessly from Microsoft Access or a similar database. Some of these tools are available as manual operations within Access and Excel, but automation has made it easier and faster to analyze the data in many different ways. As we will demonstrate once the process is automated, the data takes on a whole new meaning.

Following are the three main tool components

#### 4.1 Data-Access

An Excel macro-based dialog box uses a Microsoft Access database to automatically retrieve data-groups based on pre-defined queries. These queries can easily be modified to suit changing needs, or link to different database environments. Another option is to copy and paste data into the working area for data-reduction tools.

#### 4.2 Data- Reduction

Generally, large data-sets are reduced by using existing criteria within the data set such as, year, month, a physical boundary, a certain well, etc. The data-reduction technique provided here needs no such criteria—it is simple, yet powerful. The user may choose any numerical data-reduction rate. For example, if the user chooses the data-reduction rate of 100:1, the program will process each consecutive group of 100 data points and reduce it to a single data-point. The data set for each reduction-group is selected sequentially from the top of the data-table to the last value at the bottom. Prior to applying data-reduction, the user may pre-sort the data-table as needed. For The Geysers, we pre-sorted the data-table by year, month, and a random number. This sorted the data-table into an in-time sequence without any other bias. In certain instances, introducing the random factor may have contributed to the unnecessary scatter of data, when sequential data-sets may have been more appropriate. The user can choose how the single data-points are generated. Some of the choices include: an average, an average ignoring some highest and lowest values, summation, cumulative, median, mode, and largest or smallest number.

When compared to traditional data-reduction by criteria, some results may appear to be unusual. For example, if data are reduced by selecting average or cumulative, the results will be similar to the normal data-reduction techniques, but if reducing by "summing," the resulting sums will yield higher values with a larger data-reduction rate.

Detailed instructions are provided in the Help menu of each tool. Special care has been taken to maximize the automation, so the user can run as many different combinations as possible in the shortest amount of time.

#### 4.3 Data-Graphing

The graphical representation of the above-mentioned reduced-data has also been automated. The users can easily and quickly change different combinations of data sets for graphing, comparison, offset realignment and curve fitting.

### 5. ASSUMPTIONS

When analyzing such a large data set, the effect of instrument uncertainties and random variations in the data measured tends to be minimized. In the absence of smaller variations, the larger variations become more conspicuous and relatable to the actual events. Authors also observed this benefit of "aggregate analysis" of data (Barker and Pingol, 1997, Khan, 1993). However, instrument and measurement bias will not be minimized when using a large data set, and unless corrected, can lead to erroneous conclusions. The old adage of "garbage in – garbage out" is just as true here as with any computational tool.

Data-points generated by these tools are purely results of mathematical operations with no regard to the "relative location" or "weight" of each individual point. However, logical selection of individual wells to form data-groups will produce meaningful results. When analyzing data using this kind of mathematical "averaging," and using "uncorrected" data, conclusions should be general and relative rather than absolute, unless the user can tie these relative conclusions to some corrected data points.

Throughout this paper, wellhead data are used without any correction as to the downhole reservoir conditions. This is mainly because only publicly available data were used. Even if other information were used, it may not have significantly added to the data quality due to many uncertainties (Goyal, 1998). These include: influence of heat losses as steam travels through the borehole, production rate, cross-tie to other wells, and placement of recording gauges

At The Geysers, pressure data typically comes from transmitters that are part of the flow meter, most of which are downstream of the flow control valve. Therefore, recorded pressures are influenced by factors other than reservoir performance. Pipeline frictional losses, other wells (most wells are cross-tied), and power plant inlet pressures can all influence the recorded pressure.

The objective of this paper is not to provide an interpretation of The Geysers data or future forecasting, but to present some examples of using these data-reduction tools with real life data.

## 6. RESULTS

### 6.1 Field-wide Results

In Figure 2, average steam production per well per month and average wellhead pressure for all wells (about 700) are plotted for The Geysers field. At a reduction rate of 1050:1, we condensed 137,000 data-points per field-column to a mere 131 records per field and plotted them on this graph. The pressures used in the study are supposed to be flowing pressures, but when the wells are shut-in or throttled, the reported flowing pressure may approach shut-in pressures. Despite such drastic data reduction and varied conditions in different parts of the field, an inverse relationship of production to pressure and certain other field-wide conditions are clearly visible. One reason is that most of these 700 wells are cross-tied; therefore, the recorded wellhead flowing pressures are already "averaged" to some extent. Another possible reason is that the highly fractured Geysers geothermal reservoir facilitates more communication between wells than a typical oil and gas reservoir.

Following are some of the field-wide changes observed in Figure 2.

During Period-1, average production rates and average pressures are fluctuating immensely. This is the result of many new wells initially shut-in (i.e., pressure increases), but later brought into production (i.e., sudden pressure drop) when a new power plant comes on line.

During Period-2, most of the field development is complete and a relatively steady steam production and injection is maintained. Some fluctuation of pressures is visible as a result of well-throttling (Barker and Pingol, 1997).

During Period-3, production and pressure have the steadiest decline, because the majority of the wells are producing at open valve conditions. Throughout the 44-year history of the field, this may be the most stable condition, and hence the most suitable period for decline-curve fitting.

Period-4 is characterized by huge changes. During the winters of 1995-1997, major production curtailments resulted from sale agreement conditions. By September 1997, mass replacement increased to about 55% as the Lake County pipeline began operation (Figures 3 and 4). By December 2003, mass replacement increased to about 80% as additional water was being injected from the Santa Rosa pipeline. As our data end in December 2003, we do not see the effect of the Santa Rosa pipeline. However, beginning in 1998, the production rate per well remains almost constant, while the pressure decline remains unchanged. This "additional" production is attributed to the additional injection.

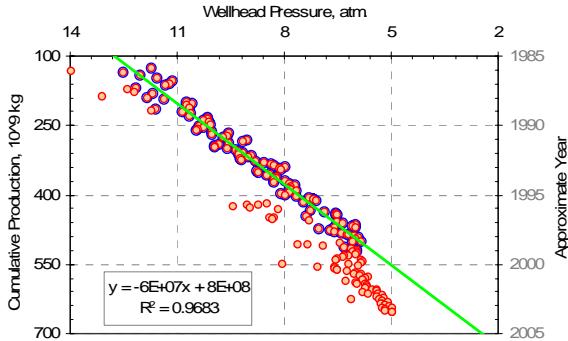
### 6.2 Southeast Geysers Results

The Southeast Geysers area is loosely defined as the one most affected by the injection of an additional 1.05 million kg of fluids per month brought in by the Lake County pipeline since September 1997. This increased the mass replacement from about 30% to 70%. There are about 152 production and 28 injection wells in this part of The Geysers.

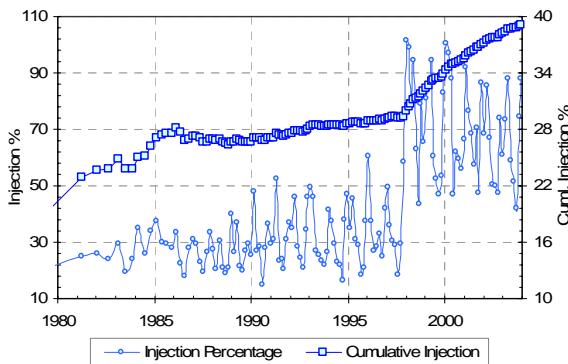
Figure 6 is a cross-plot of the average wellhead flowing pressures and cumulative steam production (reduced at 200:1) for the Southeast Geysers area. The relationship is linear from about 1985 through 1997—when mass replacement was in the range of 29% to 33% (Figure 7).

This linear relationship does not seem to be a manifestation of any averaging by the data reduction-tools, as many wells in The Geysers exhibit a similar linear relationship. This linear relationship essentially describes the steam production per pressure decline rate. Without an analytical explanation, such a relationship is analogous to empirical decline-curve equations (Khan 1998) that can give a viable estimate of extrapolation and forecasting as long as the reservoir and field parameters remain unchanged. Reyes et. al. (2004) reported a linear relationship between normalized steam production rates and reciprocal cumulative steam production, which according to the authors, ends at the start of the dry-out period. Starting in September 1997, the mass replacement increased to about 70%. As a result, from about 1998 through 2001, the pressure vs. cumulative production trend (Figure 6) takes on almost a vertical trend, indicating a lessening of the pressure decline rate. From about 2001 onward, as the reservoir re-saturates, the trend seems to be reverting to the original decline rate.

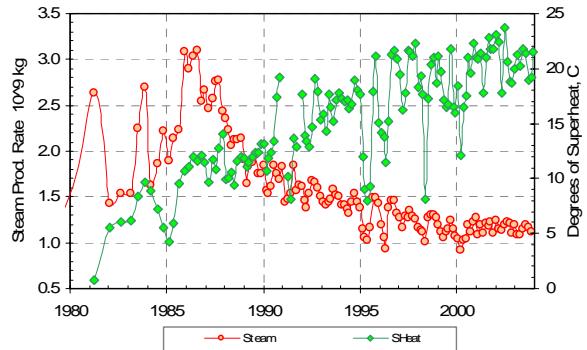
In the graph in Figure 6, we took the liberty of breaking the tradition of plotting  $p/z$  on the y-axis, and cumulative production on the x-axis; we reversed the axes. Thus, in this case, the before-mentioned trend is easy to distinguish by looking at the graph.



**Figure 6: Southeast Geysers area. Cross-plot of wellhead flowing pressures and cumulative steam production.**



**Figure 7: Southeast Geysers area. Injection history.**



**Figure 8: Southeast Geysers area. Steam production rates versus superheat.**

At The Geysers, after 44 years of production, with net mass extraction of  $1.2 \times 10^{12}$  kg, and no known significant natural fluid recharge (e.g., Beall et al., 2001), the reservoir pressure has declined from about 500 psi to about 100 psi. Most of the field has seen some degree of superheat. Reduced pressures and increased superheat (Enedy, 1989) have been used as criteria for successful injection strategies. Reyes and Horne (2003) described the dry-out state at The Geysers, where locally the mobile vapors (i.e. steam and non-condensable gases) have been produced, and immobile water has been boiled. Thus, the superheat is essentially a measure of how "dry" the reservoir is, which in turn is an indicator of how much excess heat is available to evaporate the water that is injected into the reservoir. This boiled water may be highly mobile and will rapidly flow toward the pressure sink (production well). Consequently, a reservoir with a low superheat (i.e., near saturation) would not have enough heat available to boil a significant fraction of the injected water.

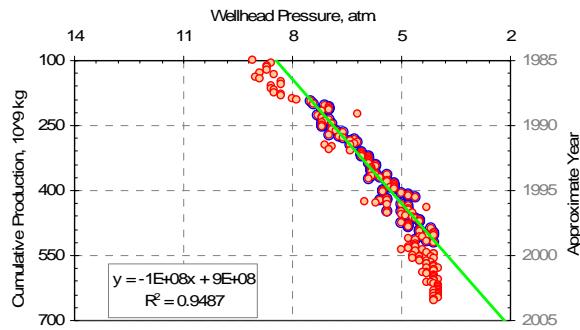
As noted earlier, for this study we did not correct surface data for the bottom hole reservoir conditions. Nevertheless, a reasonable indication of reservoir conditions can be obtained from surface data recordings of pressure and temperature. Figure 8 is a plot of steam production rate vs. degrees of superheat. As expected, prior to the dry-out period (about 1989) the rate of extraction and superheat is directly proportional. From about 1990 to about 1994, as the dry-out phase sets in, there is a steep increase in superheat. From 1995 onward, as a result of production curtailment (1995-1996) and then additional injection (1997 onward), there is some re-saturation of the reservoir; hence the superheat levels are about 20° C. Even though the superheat does indicate areas that might be suitable for injection, the degree of superheat is not necessarily indicative of the enthalpy. With continued injection, temperatures and enthalpies will eventually decline (Goyal, 1998).

## 7. PITFALLS

As is the case with any statistical tools, the user must be careful with what type of information is grouped and how the data are reduced. Unrelated groups, wrong selection of data-reduction criteria, grouping with statistical irrelevance, or statistical insignificance may give "viable looking" results, but they may not necessarily be valid. For example, Figure 8 shows a cross-plot of the average of three minimum wellhead flowing pressure values for a reduction-rate of 200:1 and cumulative production for the entire Southeast Geysers area. The graph and ensuing trend look viable. However, 3 data-points out of 200 data-points constitute only 1.5 percentile. Moreover, these 3 data-points

are at one end of the spectrum, and therefore not representative points. Hence, in this case, the results are statistically insignificant and irrelevant. On the other hand, if we used the same scheme and “excluded” those 3 data-points from our 200:1 data-reduction scheme, the results would be viable. Similarly, in the same scheme, if we had used 3 data-points out of 10 (instead of 200), that would have constituted 30% of the data-points and that may have been a viable solution.

Another factor that will influence the results is the sorting of the data before data reduction begins. The current data-reduction technique is biased toward the number of records. We are working to provide more options in that regard.



**Figure 9: Southeast Geysers area. Cross-plot of average of three minimum values per deduction-rate.**

## 8. CONCLUSIONS

The data-reduction tools presented here and tested with The Geysers data are simple, yet viable tools that can be used to rapidly sift through large amounts of data. The user can select any combination of single wells, or groups to form data sets and reduce them until just the right amount of information may be visualized. Of course, as with any data-reduction tool, the user must be careful with what type of information is grouped and how the data are reduced. Unrelated groups or wrong selection of data-reduction criteria may give “viable looking” results, but they may not necessarily be meaningful. Another factor that will influence the results is sorting of the data before data reduction begins. The current data-reduction technique is biased toward the number of records. Our future efforts will be to provide other options to offset this bias. Having clarified the pros and cons of using this data-reduction technique, our conclusion is, with due diligence, this technique, creates endless possibilities of making sense out of the data. In addition, what field (column) is sorted prior to data-reduction adds even more dimension to this process. We believe these data-reduction tools will also be useful in other applications, such as oil and gas field data, and well log data.

Readers are encouraged to contact the authors for a copy of the tools and share their results.

## ACKNOWLEDGMENTS

Many colleagues helped with this project in one form or another. In particular we would like to thank, M. Lippmann of LBNL, K. Goyal, A. Pingol and M. Stark of Calpine, Steve Enedy of NCPA, G. Robin of EPA, M. Woods, E. Johnson, and L. Tabilio of DOGGR, and A. Truesdell and

P. Akhtar for their assistance. We would also like to thank DOGGR and BLM for letting us publish this.

## REFERENCES

Arps, J.J., 1945. Analysis of Decline Curves, AIME Transactions, pp. 228-247.

Barker, B.J., and Pingol, A.S.: “Geysers Reservoir Performance—an Update,” Proceedings of the Twenty-Second Workshop on Geothermal Reservoir Engineering, Stanford Univ., (1997) p. 55-62.

Beall, J.J., Adams M.C., and Smith, J.L.: “Geysers Reservoir Dry Out and Partial Restoration Evidenced by Twenty-Five Years of Tracer Tests,” proceedings of the GRC annual meeting, San Diego, USA; GRC Trans. V. 25 (2001) p. 725-729.

Dellinger, M and Allen, L.: “Lake County Success,” GRC Bulletin, V. 32 (2004) p. 119.

Chatas, A.T., and Yankie, Jr.: “1957, Application of Statistics to the Analysis of Production Decline Data,” JPT 2015, (1957) p. 52-54.

Enedy, K.L.: “Downhole Enthalpy and Superheat Evolution of Geysers Steam Wells,” proceedings of the GRC annual meeting, Santa Rosa, USA; GRC Trans., V.13, (1989) p. 377-382.

Enedy, K.L.: “Downhole Enthalpy and Superheat Evolution of Geysers Steam Wells,” proceedings of the GRC annual meeting, Santa Rosa, USA; GRC Trans., V.13, (1989) p. 377-382.

Goyal, K.P., and Box, T.W.: “Geysers Performance Update Through 2002,” Proceedings of the Twenty-Ninth Workshop on Geothermal Reservoir Engineering, Stanford Univ., (2004) p. 5-10.

Goyal, K.P.: “Injection Related Cooling in the Unit 13 Area of The Southeast Geysers,” Proceedings of the Twenty-Third Workshop on Geothermal Reservoir Engineering, Stanford Univ., (1998) p. 397-405.

Jericho, J.L.P., Li, K., and Horne, R.N.: “A New Decline Curve Analysis Method Applied to The Geysers,” Proceedings of the Twenty-Ninth Workshop on Geothermal Reservoir Engineering, Stanford Univ., (2004) p. 144-151.

Jericho, J.L.P., and Horne, R.N.: “Inferred Water Saturation in The Geysers Based on Well Performance Data,” Proceedings of the Twenty-Eighth Workshop on Geothermal Reservoir Engineering, Stanford Univ., (2003) p. 21-30.

Khan, M. A.: “Historical Data Aids Search for Oil,” American Assoc. of Petroleum Geologists-Explorer, V. 14 (1993) p. 38-39.

Sanyal, S.K.: “Forty Years of Production History at The Geysers Geothermal Field, California—the Lessons Learned,” Proceedings of the GRC annual meeting, San Francisco, USA; GRC Trans. V. 24 (2000) p. 317-323.

Stark, M. A., et al.: “Santa Rosa—Geysers Recharge Project, Geysers Geothermal Field, California, USA,” Proceedings of the World Geothermal Congress, Antalya, Turkey, 2005 (In press).