GEOTHERMAL DATA MINING -- VALUE IN ANTIQUES

Roland N. HORNE[1]

[1]Stanford Geothermal Program, Stanford University, USA.

Keywords: data analysis, nonparametric regression, wavelets

SUMMARY - Analysis of the production history in developed geothermal fields has been used extensively to forecast the future life of the reservoir. Mature reservoirs such as Wairakei offer a unique opportunity to take a retrospective look at how well reservoir engineering methods have been able to predict the future. In general, reservoir models have performed adequately at describing the bulk volume-pressure behaviour of reservoirs such as Wairakei, however there are other significant production behaviours such as thermal breakthrough from injection wells that have been almost impossible to predict. These fracture-dominated phenomena can still be modelled, however may need a different approach. The use of data mining procedures that allow the reservoir to reveal its own internal dependencies can provide very good indications of future behaviour, even for such events as thermal breakthrough. This paper will describe well-to-well correlation techniques based on nonparametric regression. The techniques have been applied successfully to real field histories.

1 INTRODUCTION

The task of reservoir engineering is essentially to forecast the future performance of a reservoir based on measurements of its past performance. Later in the life of a development, it is possible to take a retrospective look at how validly the forecasts predicted the reservoir behaviour. It may even be said that the reservoir models will achieve a perfect prediction, of the past, on the day the plant is closed.

Modelling is conducted at a range of levels, from "whole field" models during the feasibility studies when a go/no-go decision must be made, to well or regional models later in the field life when specific strategies or operational modifications need to be evaluated. In musing on the first 20 years of production at Wairakei, Bixley (1980) opined that reservoir modelling had not been of particularly great value in predicting the life of the reservoir, and that specific well-by-well phenomena had often been a surprise when they occurred. Similar surprises still occur today.

10 years into the life of Wairakei, Bolton (1970) presented a rather famous view of the reservoir performance, a representation of which is replicated here in Figure 1. This particular diagram was the focus of reservoir modellers throughout the 1970s, and the questions of whether Wairakei was in imminent danger of demise and/or what could be done about it were the cause of much wringing of hands in the geothermal community. As observed later by Bixley (1980), few of the models told the field operators what they wanted to know, at the time they wanted to know it. Fortunately, the demise of Wairakei was postponed at least 30 years further into the future, and has yet to occur (nor do we yet know when it will).

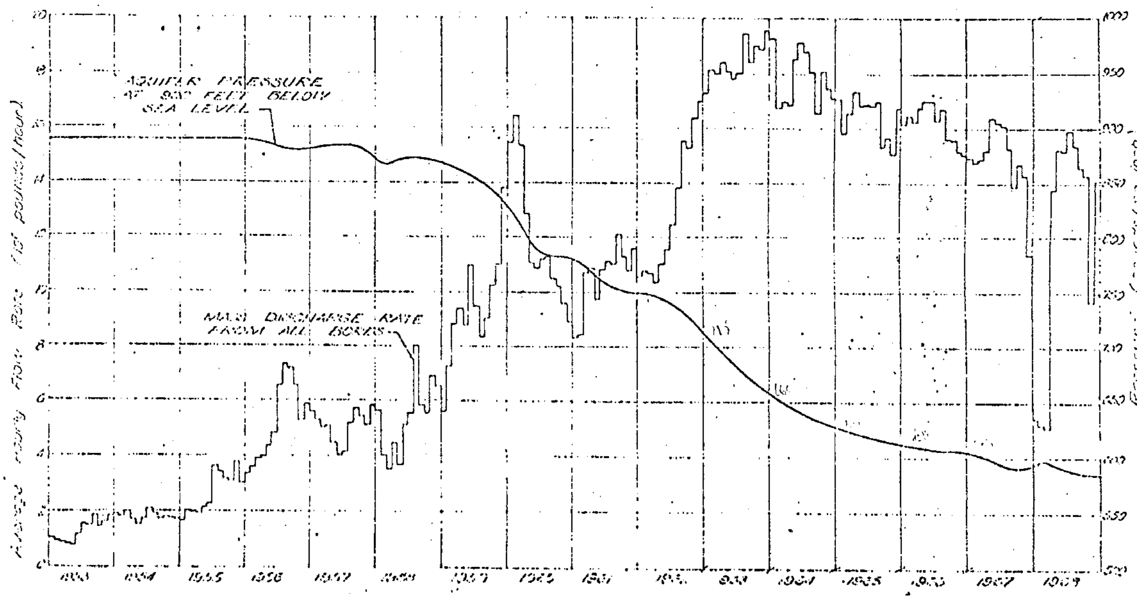

Figure 1: Early production history at Wairakei, from Bolton (1970).

One of the philosophical underpinnings of reservoir modelling is the mental step of the creation of the model. Herein lies a rich opportunity for misunderstanding, bias and wishful thinking. Sometimes the old adage “what’s the answer?” – “what would you like it to be?” creeps into the process. Many forms of reservoir modelling, such as simulation, decline curve analysis, tracer test analysis, well test analysis, material balance analysis, etc., have a central set of physical assumptions explicitly or implicitly included into their underlying models. The predictive effectiveness of the model will be constrained by the accuracy of its assumptions, even if the data are perfectly accurate and even if the model match is precise. One of the principal reasons geothermal reservoir modelling is so difficult is that the reservoir behaviour is usually governed in important ways by the location and properties of fractures, neither of which is ever known clearly.

A different philosophical approach is to let the data define the model. Individual wells and their hydraulic neighbours interact throughout their connecting fracture network, in ways that are characteristic of those fractures. Relating the interwell connectivity provides a useful modelling tool for the understanding of at least regional behaviour of the reservoir. Often such connectivity interpretations use models and are again constrained by model assumptions, however it is also possible to use nonparametric approaches, as will be described here, to avoid the definition of a model (or more correctly, to let the data define the model).

It has been common to analyse the movement of fluids in the reservoir by monitoring the production of chloride, which changes as a function of time because the reinjected water is elevated in chloride concentration due to the separation of steam. A classic paper that illuminated this approach was Harper and Jordan (1985), which quantified the rate of return of reinjection water at Palinpinon-I field by analysing chloride (among other variables).

In 1991, Urbino (Macario) and Horne used a correlation method to relate the chloride histories of production and injection wells, for example the well pair shown in Figure 2. This figure reveals a clear relationship between chloride injected at one location and the chloride produced at another. One of the variations of the approach was to subtract a linear-with-time trend (a model) from the data, which was an attempt to decipher short-term fluctuations from well histories that show a continuously increasing chloride concentration. The correlation approach was expanded further by Sullera and Horne (2001), who used wavelet decomposition to assist the illumination of the chloride fluctuations at different time resolutions, as seen in Figure 3.

A distinction between the approaches of Macario and Sullera is that Macario assumed an underlying correlation model (which was a simple algebraic sum) whereas Sullera's wavelet decomposition reveals well-to-well relationships directly from the data. For example the gray boxes in Figure 3 show clear relationships between OK-7D production and PN-6RD injection in the later part of the diagram, and between OK-7D production and PN-9RD injection in the earlier part of the diagram. A model was still needed to quantify the degree of well-to-well connectivity based on these observations.

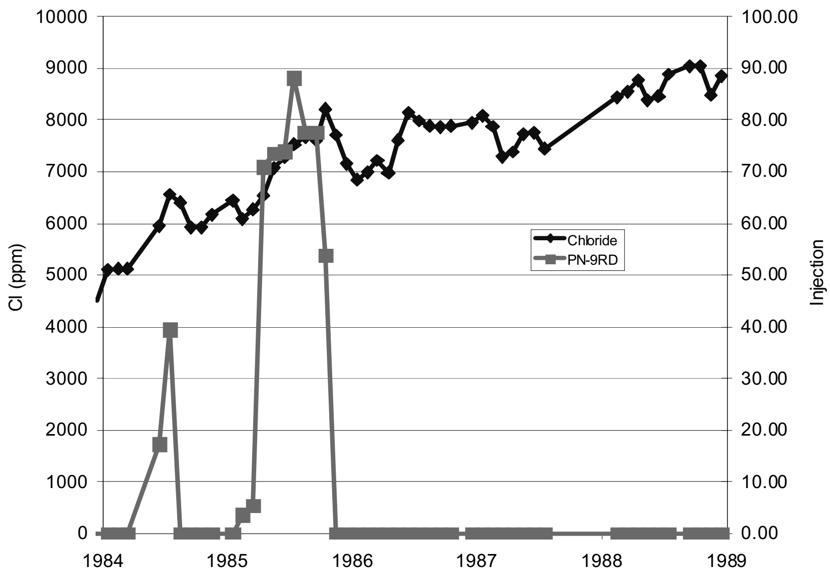

Figure 2: Example data from Palinpinon-I, showing chloride production in well OK-7 and injection rate of well PN-9RD, from Sullera and Horne (2001).

These approaches were reasonably successful, and were shown to be qualitatively consistent with the results of tracer tests. Nonetheless, the philosophical difficulty with this style of analysis is the requirement to make assumptions of the mathematical form of the model (for example that background chloride increased linearly with time). The reservoir physics may result in the relationships being something other than simple forms. This is a weakness of the approach.

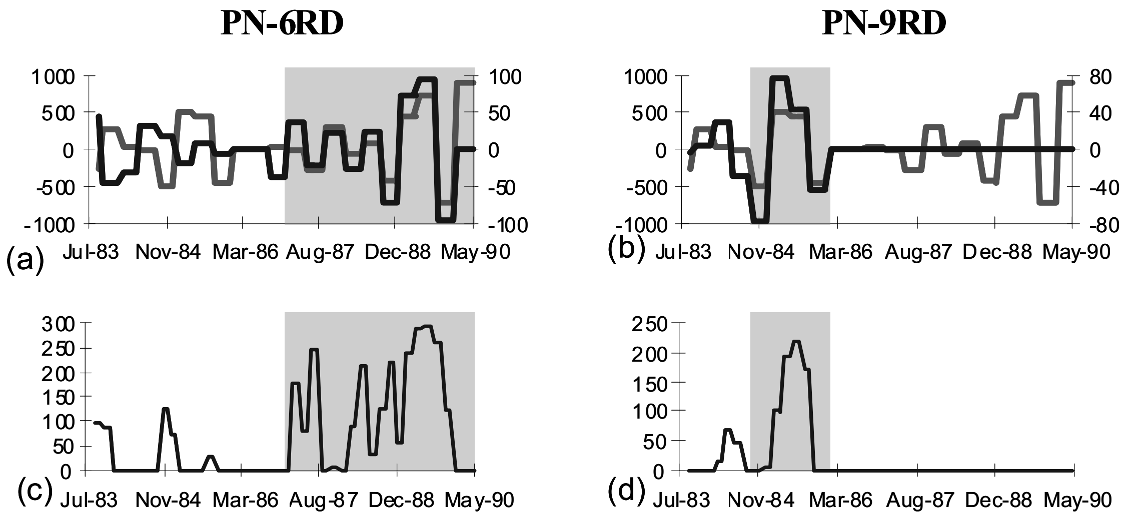

Figure 3: (a) Level 3 detail of OK-7D chloride - light line; level 3 detail of PN-6RD injection rate - dark line. (b) Level 3 detail of OK-7D chloride - light line; level 3 detail of PN-9RD injection rate - dark line. (c) PN-6RD injection rate. (d) PN-9RD injection rate. From Sullera and Horne (2001).

In an attempt to address this weakness, Horne and Szucs (2006) investigated the use of nonparametric regression. The fundamental idea of nonparametric regression is to match the data without making assumptions about the underlying form of the relationships. In fact, a major advantage of the approach is that the nature of the relationship is revealed in the process. The magnitude of the connectivities can also be estimated, and these values are then useful for reinjection analysis and design.

2 NONPARAMETRIC REGRESSION – ACE

The ACE (alternating conditional expectation) method was presented by Breiman and Friedman (1985) as a nonparametric approach to modelling data without knowing the model in advance.

The ACE method works by inferring a decomposition of the signal in the following form:

$$g^*(y) = \sum_{i=1}^{p} f_i^*(x_i) + e^* \qquad (1)$$

where e^* is the remaining error not captured by the functional form, and which is assumed to be normally distributed. It is important to note that $g(y)$ and $f(x)$ are not known in advance but are extracted as a result of the algorithm.

It should be noted that the ACE method can be applied with more than one independent x variable. Hence it is a very suitable way of investigating the relationships between outputs (for example, chloride production at a well) and inputs (for example, injection rates at many other wells). Horne and Szucs (2006) applied the ACE approach to analyse chloride production histories from Palinpinon field in the Philippines, and showed good success in predicting the independently measured tracer returns, as will be shown here.

3 APPLICATION TO PALINPINON-I CHLORIDE DATA

Having introduced the ACE approach, it is clear that this method offers advantages over the inherently "parametric" approaches used earlier by Urbino (Macario) and Horne (1991) and Sullera and Horne (2001). It is no longer necessary to make any explicit or implicit assumptions about how the input and out variables depend on each other. To compare the approaches, Horne and Szucs (2006) reexamined the same data set used in these two earlier studies, namely the production and injection histories of Palinpinon-I field over the period between 1983 and 1989.

Typical results are shown for well OK-7 in Figures 4 and 5. Figure 4 can be compared to the original data shown in Figure 2. Looking first at Figure 4, which has been simplified by including only the functions due to time (red line) and due to well PN-9RD (pink squares), it can be seen that the ACE procedure extracts a simpler picture of the relationships between input and output signals. Importantly, the time dependence is not linear, as was assumed in the earlier studies. The details of the PN-9RD function are somewhat deceptive, as the well was not injecting for much of the time. Hence it is the magnitude of the positive values of the transform function that indicate the degree of connection between this well and OK-7. Figure 5 adds the transform functions for all of the injection wells – the relative sizes of their positive components shows their connectivity to well OK-7.

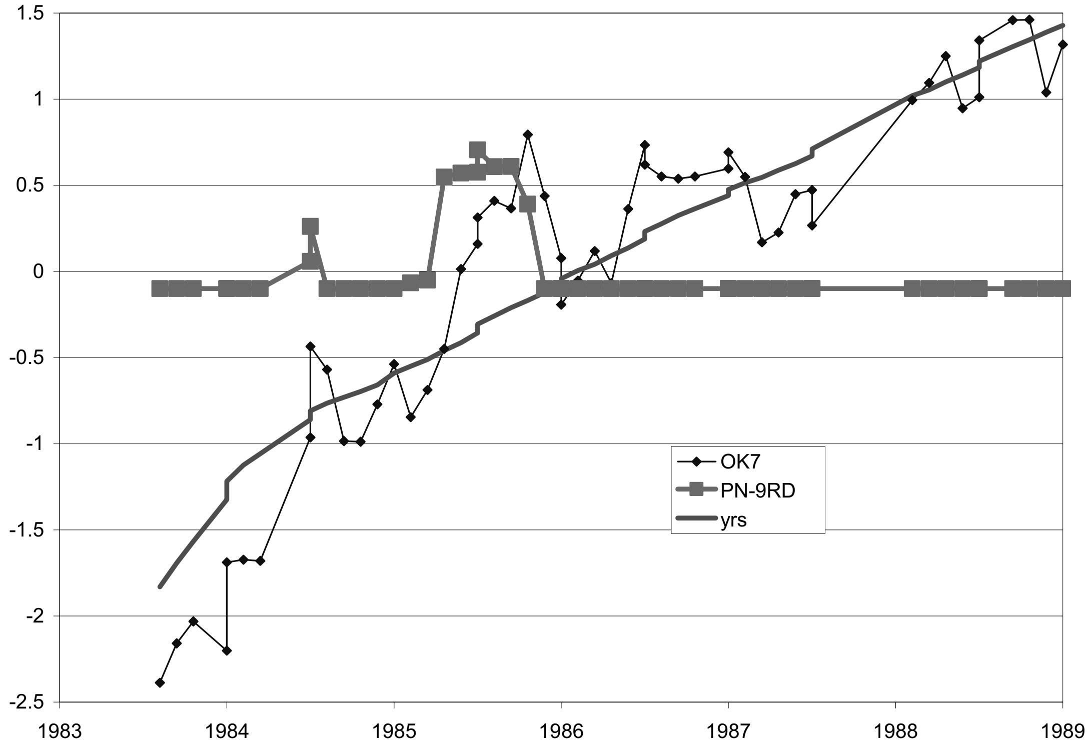

Figure 4: Extracted model functions from OK-7 data (thin line), showing dependence on time (red line), and dependence on injection into PN-9RD (pink squares). From Horne and Szucs (2006).

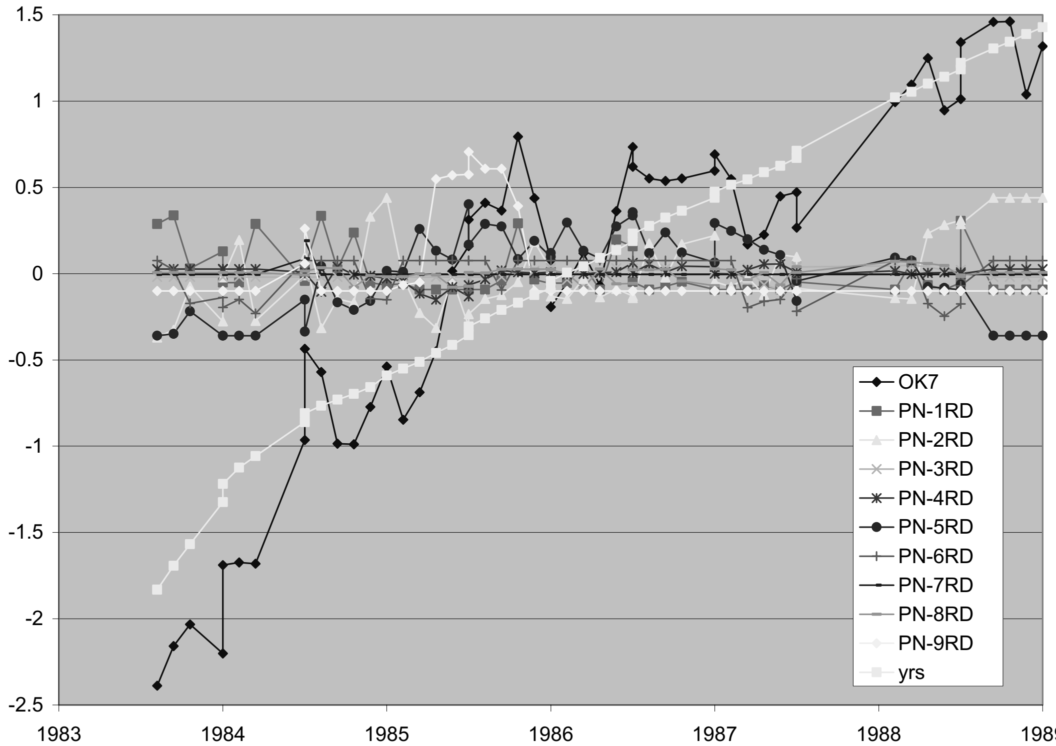

Figure 5: Extracted model functions from OK-7 data, showing dependence on time, and dependence on injection into all injection wells. From Horne and Szucs (2006).

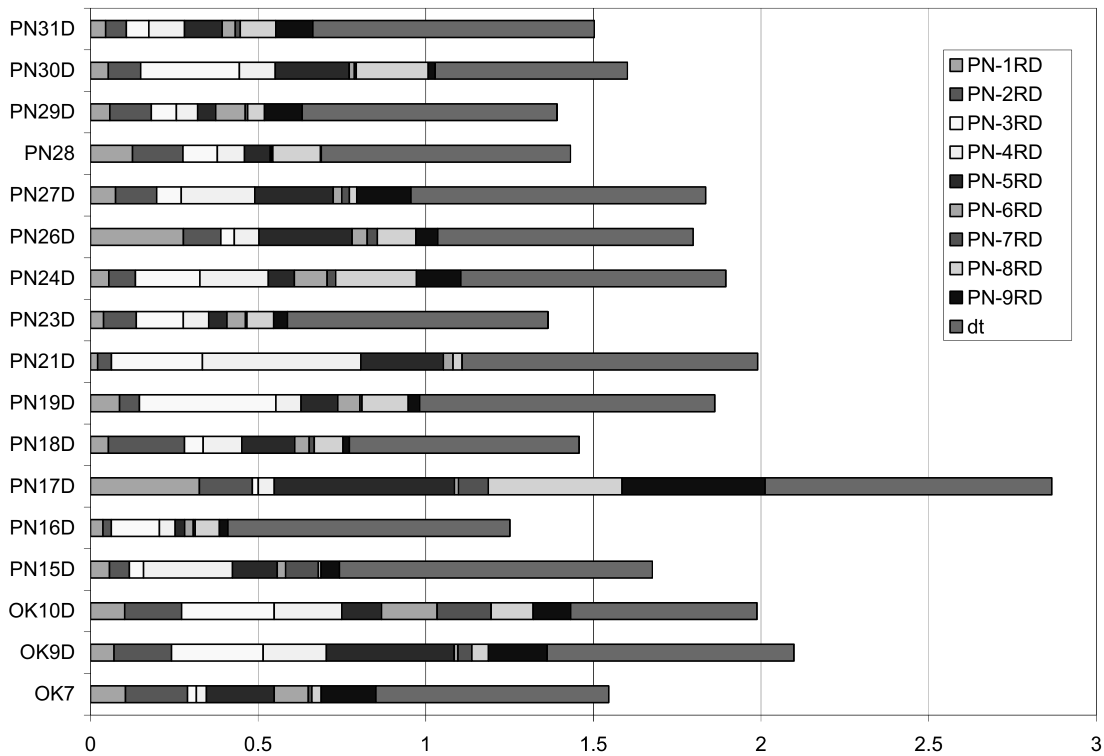

Figure 6: Summary of "connection indices" based on ACE function magnitudes. Total length of bars indicate impact on well of reinjection returns. Rightmost element represents time dependence. From Horne and Szucs (2006).

Based on the transform functions shown in Figure 5, we can compute a "connection index" for well-to-well connectivity. Horne and Szucs (2006) experimented with different ways to do this, and eventually decided on an index defined as:

$$I_i = \frac{1}{n}\sum_{j=1}^{n}\left|f_i(x_i(t_j)\right| \qquad (2)$$

These indices are illustrated graphically in Figure 6. Figure 6 includes the overall dependence on time, to reveal the size of the impact of reinjection at each production well (the total length of the bar). Individual bar segments in Figure 6 indicate the strength of the connections from specific injection wells. These same connection strengths are shown in Figure 7, allowing a quick visualization of the largest connections – the ones most likely to result in thermal breakthrough.

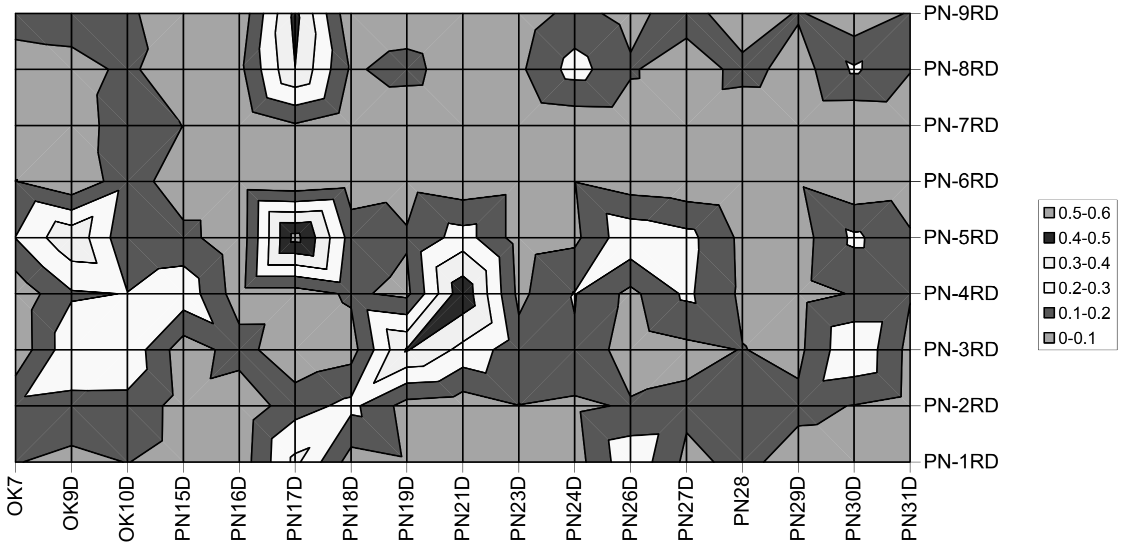

Figure 7: Summary of "connection indices" based on ACE functions, showing magnitude of well to well connections. From Horne and Szucs (2006).

4 COMPARISONS TO TRACER TESTS

We can compute these connection indices using ACE, but do they have physical meaning compared to actual fluid movements in the reservoir? One way this could be investigated was by comparing the indices to well-to-well connectivity measurements obtained by other approaches. Fortunately, a series of tracer tests conducted at Palinpinon-I allowed Horne and Szucs (2006) to make such a comparison.

During the early life of the field, PNOC-EDC conducted a number of tracer test campaigns at Palinpinon, as described by Urbino, Zaide, Malate, and Bueza (1986). These tracer test records showed the transit time of the tracers from one well to another, as well as the total fraction of the tracer recovered. Horne and Szucs (2006) used the reciprocal of the transit time as an indicator of the connectivity, based on the premise that a fast (short) transit time represents a strong connection.

Figure 8 shows the results of the tracer test with injection into PN-1RD, compared to the connectivity from well PN-1RD estimated in the ACE analysis. The results are very consistent.

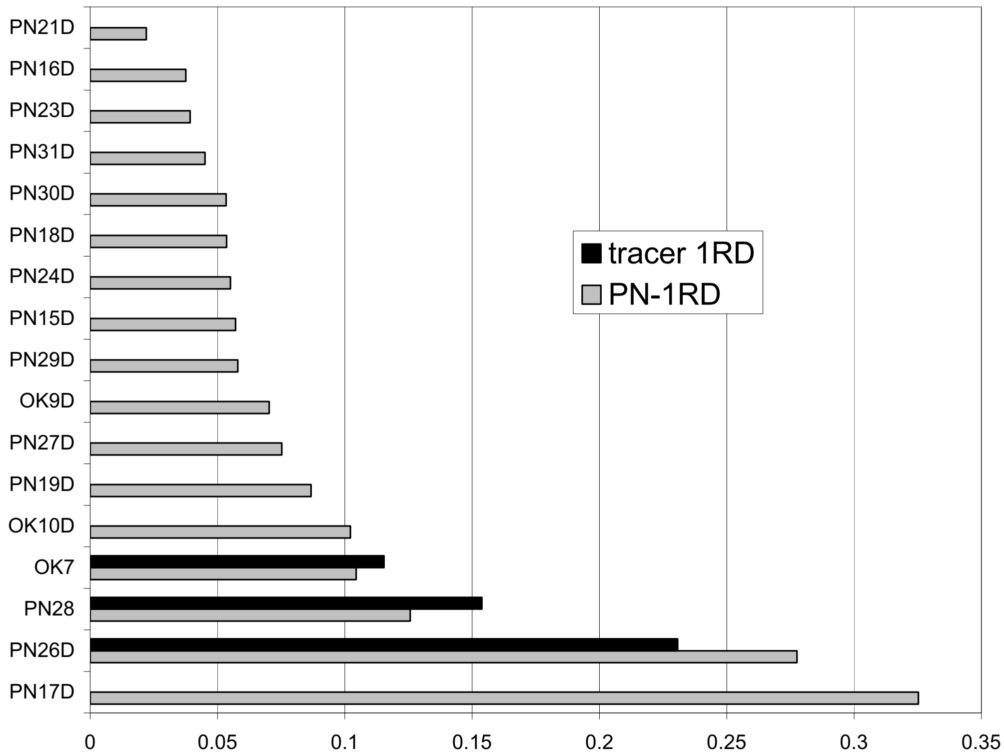

Figure 8: Comparison of connection indices from injector PN-1RD, compared to results of tracer test into PN-1RD (reciprocal of arrival time). From Horne and Szucs (2006).

5 CONCLUSION

The application of the nonparametric ACE method to Palinpinon-I data of production well chloride as a function of reinjection well injection rate showed that the well-to-well connectivity indices computed in this way are consistent with tracer test results. The advantage of inferring well connectivity by this approach is that it can be done with routinely measured production and geochemical data and does not require the expense and operational disruption that would be needed with a tracer test. Attaining an understanding of local connections between wells is very useful in designing a strategy for reinjection, and predicting where thermal breakthroughs are likely to occur. No modelling is involved.

The message from these studies is that sometimes we need to let the reservoir speak for itself. Reservoir engineers and geoscientists will always require deep insight into the physical and chemical mechanisms taking place underground, and sometimes it is a mistake to assume that this insight has been embedded invisibly inside an intricate model.

Geothermal energy production is effectively a mining exercise – we mine the heat energy of the earth, while simultaneously mining the data it provides to understand how to produce that energy as efficiently and effectively as possible.

6 REFERENCES

Bolton, R.S.: "The Behaviour of the Wairakei Geothermal Field During Exploitation." *Geothermics*, Special Issue 2, (1970) pp. 1426-1439.

Bixley. P.F.: "Modelling from a User's Point-of View," Proceedings of the 1980 Stanford Geothermal Workshop, Stanford California, December 1980.
http://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/1980/Bixley.pdf

Breiman L. and Friedman J.H.: "Estimating Optimal Transformations for Multiple Regression and Correlation," *Journal of American Statistical Association*, **80**, (September 1985), 580-619.

Harper, R T. and Jordan, O.T.: "Geochemical Changes in Response to Production and Reinjection for Palinpinon-I Geothermal Field, Negros Oriental, Philippines," Proceedings, New Zealand Geothermal Workshop, 1985, 39-44.
http://pangea.stanford.edu/ERE/pdf/IGAstandard/NZGW/1985/Harper.pdf

Horne, R.N., and Szucs, P.: "Inferring Well-to-Well Connectivity Using Nonparametric Regression on Well Histories," Proceedings of the 2006 Stanford Geothermal Workshop, Stanford, CA.
http://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2007/horne.pdf

Sullera, M.M., and Horne, R.N.: "Inferring Injection Returns from Chloride Monitoring Data," *Geothermics*, **30**(6), December (2001), 591-616.

Urbino, M.E.G., and R.N. Horne: "Optimizing Reinjection Strategy at Palinpinon, Philippines, Based on Chloride Data," Proceedings, 16th Stanford Geothermal Workshop, Jan. 1991, Stanford, CA.
http://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/1991/Urbino.pdf

Urbino, M.E.G., Zaide, M. C., Malate, R.C.M. and Bueza, E.L.: "Structural Flowpaths of Reinjected Fluids Based on Tracer Tests - Palinpinon I, Philippines," Proceedings, New Zealand Geothermal Workshop 1986, 53-58.
http://pangea.stanford.edu/ERE/pdf/IGAstandard/NZGW/1986/Urbino.pdf