. 12th New Zealand Geothermal Workshop 1990

CHEMOMETRICS APPLIED TO THE FLUID CHEMISTRY OF GEOTHERMAL FIELDS IN THE TAUPO VOLCANIC ZONE, NEW ZEALAND

N.V. Salvania **Keith Nicholson**

Geothermal Institute, University of Auckland, Private Bag, New Zealand

ABSTRACT

Weirbox water chemistries of geothermal fluids discharged from Kawerau, Rotorua, Waiotapu, Te Kopia, Orakeikorako, Ohaaki, Mokai, Rotokawa, Wairakei, Tauhara and Tokaanu geothermal fields were selected for trial application of multivariate statistical techniques. Correlation coefficients identified inter-element associations, which were further quantified by R-mode factor analysis. These methods identified sources and controls on the distribution of major and trace constituents in the water, and may form the basis of a future mixing model. Cluster analysis groups waters of similar chemistry, and shows promise as a method to recognise fields with a common reservoir, and to distinguish between wells within a field which discharge waters from shallow and deep reservoirs. The techniques require further development but, as this study shows, are a potential tool in the evaluation of geothermal fluid chemistry.

INTRODUCTION

Multivariate statistical analysis on geochemical data has been commonly employed in mineral exploration (eg. Howarth, 1983). More recently, it has also been successfully applied to lithogeochemical data as an aid to stratigraphic correlation (Nicholson and Anderton, 1989). The aim of this study was to apply similar statistical methods to hydrogeochemical data with the ultimate objective of correlating fluids with a common reservoir,

In this preliminary evaluation, we initially sought to determine whether meaningful results would be produced by the application of statistical methods *to* geothermal fluids. Accordingly, the elemental associations delineated by correlation coefficient and R-mode factor analysis were evaluated *to* ensure geochemical sensibility in the results, prior to attempting a reservoir-fluid correlation by cluster analysis.

METHODS: DATA SELECTION, TRANSFORMATION AND ANALYSIS

Chemical analyses of fluids from the following geothermal fields were included in this study: Kawerau, Rotorua, Waiotapu, Te Kopia, Orakeikorako, Ohaaki, Mokai, Rotokawa, Wairakei, Tauhara and Tokaanu. To reduce the variance in the data set introduced by poor sample selection and/or analysis, spring waters were excluded from the study and only chemical data for well discharge fluids were employed. To provide a common basis for the comparison of wells data, the discharge chemistry was recalculated to the weirbox composition. Determinations of the trace species are not routinely reported however, and it is not unusual for some major components to be excluded from some well analyses. To reduce the number of missing values in the final data matrix, only comprehensive analyses were selected, wherever possible. Finally, an ionic balance was performed on each of the remaining well data, and any analysis which showed a variation of more than 10% was rejected. This rigourous data selection left a total of 108 well analyses which were then used in the study. A complete list of all chemical analyses employed appears in Salvania (1989).

To yield a better fit to a Gaussian dismbution (a pre-requirement for multi-variate statistical analysis) the raw data were log-transformed prior to application of the correlation, factor and cluster analyses. The SAS statistical computer package (PC version) was employed for all analyses in this study (SAS, 1985; 1987).

RESULTS AND DISCUSSION

A statistical summary of the data employed in the study (Table A1), together with the full results of the correlation, factor and cluster analyses, are tabulated and presented in the appendix. These tables are referred to in the text, and are numbered with an "A" prefix.

Correlation analysis

Correlation coefficients can define associations between elements *or* species. Interpretation of these associations lies with the geochemist, but when applied to geothermal fluids the controls on associations are likely to be recognisable, specific chemical or physical processes.

Results. The full data set of the correlation coefficients is presented in Table **A2.** In evaluation this data, the following criteria were used:

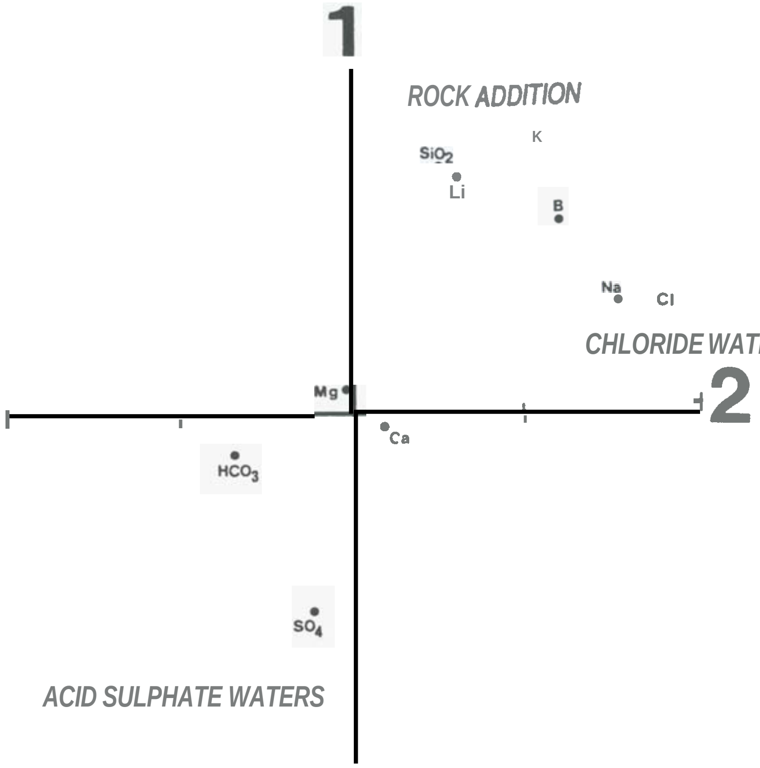

FIGURE 1. R-mode factor analysis on geothermal well chemistry; Factor 1 vs Factor 2.

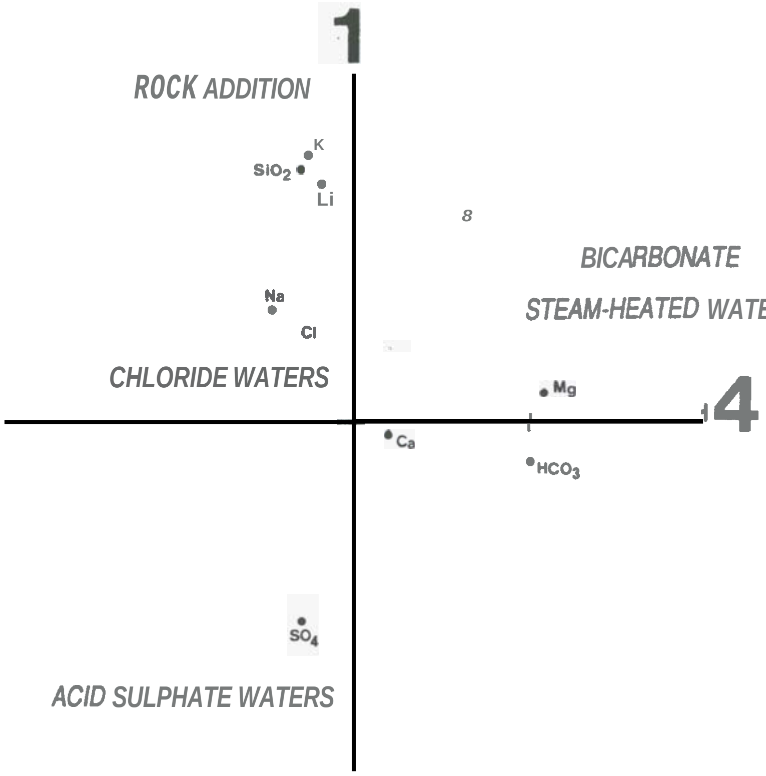

FIGURE 2. R-mode factor analysis on geothermal well chemistry; Factor 1 vs Factor 4.

coefficient value	*correlation significance*
<0.40	nil
0.41-0.49	weak
0.50–0.69	moderate
S.70	strong

On this basis, the following association was defined by the analysis: Li-Na-K-Cl-SiO_2-B. All species in this association consistently show a positive intercorrelation, while Ca, Mg, SO_4 and HCO_3 do not positively correlate with any species. Moreover, SO_4 shows negative. correlations with K, SiO_2 and B; while HCO_3-Cl and Ca-SiO_2 are also weakly negatively correlated.

Correlation analysis was also performed on the trace constituents (Table A3). As the number of samples (25) is approaching the number of variables (16) some caution has to be applied in interpreting the results. Nonetheless, it is clear from the consistent correlations that Rb, Cs, F and Br form part of the Li-Na-K-Cl-SiO_2-B association. Ammonia correlates with HCO_3, B and Mg, while iodide did not correlate with any species.

Discussion. The consistent positive association Li-Na-K-Cl-SiO_2-B clearly reflects species conmbuted to the discharge by the chloride fluid. The outlying species (Ca, Mg, SO_4 and HCO_3) however, are contributed by other, largely independent, sources; eg. retrograde mineral solubilities, groundwater, acid sulphate waters and steam heated waters, respectively. Considering the limited results for the trace species: Rb-Cs-F-Br is a further chloride fluid association; NH_3-HCO_3-B-Mg is interpreted as a steam heating association (with Mg probably produced from near-surface reactions with these fluids), while uncorrelated iodide emanates from another unidentified source which does not significantly contribute any other species, possibly organic matter. The meaningful correlations achieved in this analysis encourage the use of more advanced statistical methods. However, given the limited trace constituent data, these will not be considered further.

Factor analysis

To confirm and further quantify the associations defined by the correlation analysis, R-mode factor analysis was performed on the major-species data set. By reducing the number of variables under consideration to a few dominant controls (or "factors";' typically 3-6) factor analysis can greatly assist in the interpretation of element/species associations. Of the options available, principal factor; principal components, and maximum-likelihood factor analyses were undertaken, each with Varimax rotation. The results of each analysis were not significantly different, consequent only the principal factor analysis will be discussed.

Results. The eigenvalues implied at least a 3-factor model should be used; while the scree test (Malinosky and Howery, 1980) indicated a 4-factor model would be appropriate (Fig. A1) and this is presented in Table A4. This model accounts for almost all the variance in the data set (98.3%), the remainder probably being random error, giving confidence to the overall analysis. The communality of all the variables is S.65, except for SO_4, Mg and HCO_3 which have communalities of 0.32-0.39, indicating that this model does not adequately describe the variance of these species in the data set. The following associations are defined by the analysis:

	+ve loading	*-ve loading*
Factor 1:	Li-K-SiO_2-B;	SO_4
Factor 2:	Na-K-Cl-B	
Factor 3:	Ca	SiO_2
Factor 4:	Mg-HCO_3	

Discussion. The interpretation of these associations is readily illustrated in Figure 1. Factors 1 and 2 represent components added to the rock/retained in the fluid on alteration. Naturally, there is a spread between these end members as, for example, not all the K in the fluid is lost to the rock. The position of the constituents on the arc between these two extremes reflects the relative fluid-rock affinity of the species in question. The negative aspect of Factor 1 represents species in steam heated fluids.

Factor 4 probably represents dilution by bicarbonate steam-heated waters or groundwater. The antipathetic inclusion of Ca and SiO_2 in Factor 3 indicates that this is probably a mineral solubility control, with silica having an opposite sign to Ca because the concentration of the latter is controlled by minerals of retrograde solubility. The relative contribution of each solute by chloride fluid, steam heated waters and groundwaters is neatly illustrated in Figure 2.

Factor analysis can therefore be employed to help define controls on the distribution of the species in a geothermal fluid. An extension of this application, which awaits further study, would be the use of the factor loadings to determine the relative proportions of a given species contributed by different sources. These could be plotted to produce a mixing model complementary to those already in common use.

Cluster analysis

There are a variety of techniques, collectively termed cluster analysis, which aim to assign observations to distinct observations, and thereby form the basis of a classification. In terms of chemical data, this mean that the analysis is considered as a whole, and samples with similar chemistries are grouped together. Of the options available in the SAS package, "Fastclus" was selected for this study. This method was considered most appropriate as Fastclus is a non-hierarchical technique which clusters observations in such a way that each observation is placed into one, and only one, cluster. in an additional analysis, the "Aceclus" procedure was also applied prior to clustering with Fastclus. however, this did not produce any significant change in the outcome of the analysis and will not be discussed further.

TABLE 1
Wells in each cluster, listed by geothermal field

Field	Cluster 1	2	3	4	5	6	7	8	Total
Kawerau				7				2	9
Mokai	1						3		4
Ohaaki	1		1	24		1	1		28
Orakeikorako								3	3
Rotokawa				3	2				5
Rotorua			12					13	25
Tauhara			2	4			3		9
Te Kopia								1	1
Tokaanu	9	1			4				14
Waiotapu			1				1	2	4
Wairakei							6		6
Total	11	1	16	38	6	1	14	21	108

Results. Eight clusters were defined by the analysis, each of which grouped together a number of wells on the basis of chemical similarity. The results are shown listed by geothermal field in Table 1; individual wells and their respective cluster are presented in Table A5. This can be summarised as follows:

cluster	*dominant fields (% of wells in* cluster)
1	**Tokaanu *(64%)***
2	Tokaanu **(7%)**
3	Rotorua **(48%);** Tauhara (22%)
4	Ohaaki (86%); Kawerau ***(78%)***
5	Rotokawa (40%)
6	Ohaaki **(4%)**
7	Wairakei (100%); Mokai (75%); Tauhara (33%)
8	Te Kopia (100%); Orakeikorako (100%); Rotorua (52%)

Discussion. Some caution is needed in examining the above, since this form of representation is distorted as the numbers of wells each field included in the analysis are not equal, and clusters 2 and 6 contain only one well (from the Tokaanu and Ohaaki fields respectively). The objective of this analysis was to determine whether wells discharging from the same reservoir would fall into a single cluster. In grouping wells of the Wairakei-Tauhara field into cluster 7 (and perhaps also the grouping of the Te Kopia and Orakeikorako wells in cluster 8) it is clear that the analysis has been partially successful. However, although the wells from some fields do dominantly fall into one cluster (eg. Ohaaki, cluster **4**: Wairakei, cluster 7), the wells from other fields are included into two or more clusters (eg. Rotorua. clusters 3 and 8).. Furthermore, the clusters do not always contain wells from fields that could be realistically tapping the same reservoir (eg. Kawerau and Tauhara. cluster **4).**

Examining the raw data of the well discharges (Salvania, 1989) in conjunction with the cluster pattern, it is clear that the analysis has indeed clustered well discharges according to chemistry, but that the major-species composition of the fluids is too similar to permit a discrimination of the deep reservoirs. This can be shown **by** the Cl-HCO_3 plot of Figure 3. These components were chosen as geothermal fluids are commonly classified by anion chemistry, but other species could have been selected to illustrate the point. Wells from the individual fields are shown bounded by boxes of the different clusters. This plot indicates that the waters have been classified (or clustered) largely by relative salinity-dilution relations. The boxes **are** not of regular shape, and cluster **4** significantly overlaps cluster 3 because the analysis considers the overall chemistry of the discharge. Species other than chloride and bicarbonate therefore **also** play a role in the discrimination of clusters. but these are not taken into account in Figure 3. The analysis has therefore essentially yielded a classification of water types based not on selected species, but on the composition of the discharges considered as a whole.

CONCLUSIONS

This study demonstrates that meaningful results are achieved by the application of chemometrics to the chemistry of geothermal fluids. Inter-element correlation coefficients and particularly R-mode factor analysis can be used to identify different sources and controls on given species. Cluster analysis effectively groups waters of similar chemistry, and shows promise as a method to recognise fields with a common reservoir, and to distinguish between wells within **a** field which discharge waters from shallow and deep reservoirs.

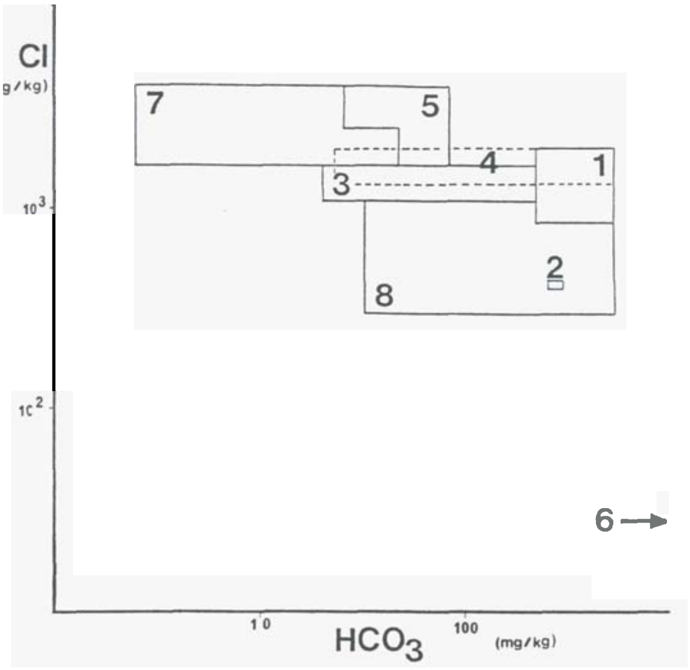

FIGURE 3. Cl-HCO_3 discharge chemistry. Wells enclosed within boxes fall in the same cluster. Cluster 4 box is shown in dashed lines.

These initial results are encouraging, and only introduce some of the possible applications of chemometrics to geothermal fluid chemistry. It is proposed to expand the study to include fields outside the TVZ in the data set, to include trace elements in the statistics and to perform similar analyses on the discharge chemistries of individual fields.

ACKNOWLEDGEMENTS

We wish to thank Pat Browne and Cedric Malate for helpful discussions during the course of this study, and to Anne Tan-Law for assistance with the SAS program. NVS also extends his thanks *to* DoDo Savella and Bing Vargas.

REFERENCES

HOWARTH, R.J., **(1983).** Statistics and data analysis in geochemical prospecting. Elsevier, Amsterdam.

MALINOSKY, E.R. AND HOWERY, D.G., **(1980).** Factor analysis in Chemistry. Wiley, New York.

NICHOLSON, **K.** and ANDERTON, R. **(1989).** Stratigraphy, structure, faulting, geochemistry and mineralisation of the Dalradian rocks around the Lecht, Banffshire, Scotland. Trans. Roy. Soc. Edin. Earth Sci., **80, 143-157.**

SALVANIA, N.V., **(1989).** Multi-variate statistical analysis of fluid chemistry from geothermal fields in Taupo Volcan Zone, N.Z. Dip. Geotherm. Energy Technology Report No. **89.24,** University of Auckland, 54pp. (unpubl.).

SAS (1985). SAS Introductory guide for personal computers. SAS Institute Inc, North Carolina.

SAS **(1987).** SAS application guide. **SAS** Institute Inc, North Carolina.

APPENDIX

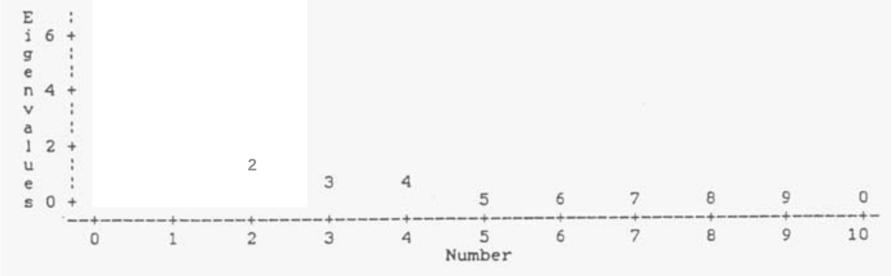

FIGURE A1. Scree plot indicating a 4-factor model is appropriate for the data set.

TABLE A1

Descriptive statistics of the data set

ELEMENT	NO.	MEAN	MIN.	MAX.	STANDARD DEVIATION	VARIANCE
Li	108	8.54	0.29	32	5.545	30.797
Na	108	907.39	220	1790	324.920	105573.1
K	108	135.53	13	500	84.084	7070.049
Ca	88	20.79	1	260	40.569	1645.822
Mg	68	1.03	0.005	37	9.687	21.972
Cl	108	1956.98	28	3879	676.299	957379.0
SO4	107	34.18	3	192	27.598	758.903
SiO2	106	537.96	157	1407	279.798	75486.28
B	108	22.71	0.42	105	29.699	610.058
HCO3	94	199.59	2.7	2284	329.766	108795.7

TABLE A2

Correlation coefficients between major constituents

n = 108

	Li	Na	K	Ca	Mg	Cl	SO4	SiO2	B	HCO3
Li	1.00000	0.50653	0.68833	0.24975	0,03346	0.46553	−0.38020	0.46199	0.57918	−0.27887
Na	0.50653	1.00000	0.73146	0.14367	−0.16784	0.84222	−4.24325	0.39978	0.56900	−0.33316
K	0.68933	0.73946	1.00000	−0.03739	−0.01097	0,72206	−0.46962	0.77287	0.70712	−0.33153
Ca	0.24975	0.14367	−0.03739	1.00000	0.12177	0.05771	0.17099	−0.48608	0.07064	0.02578
Mg	0.03346	−0.16784	−0.01097	0.12177	1.00000	−0.07902	−0.06474	−0.01845	0.26837	0.24204
Cl	0.46553	0.84222	0.72208	0.06771	−0.07902	1.00000	−0.26648	1.14920	0.69483	−0.47042
SO4	−0.38020	−0.24325	−0.48962	0.07099	−0.06474	−0.26648	1.00000	−0.49325	−0.43805	−0.04508
SiO2	0.16199	0.39978	0.77287	−0.48608	−0.06845	0.44920	−0.49325	1.00000	0.54300	−0.28199
B	0.57918	0.56900	0.70712	0.07064	0.26837	0.69483	−1.43805	0.54300	1.00000	−0.14358
HCO3	−0.27887	−0.33316	−6.33153	0.02578	0.24204	−0.47042	−0.04509	−0.28199	−0.14356	1.00000

TABLE A3

Correlation coefficients between trace and major constituents

n = 25

	Rb	Cs	F	Br	I	NH3
Li	0.81184	0.88116	0.71007	0.38068	-0.10977	-0.36223
Na	0.89474	0.82063	0.80891	0.72840	-0.24873	-0.36992
K	0.91762	0.84490	0.79557	0.79329	-0.13967	-0.21424
Ca	0.14193	0.33624	-0.03510	-0.15931	0.30533	-0.47676
Mg	-0.84360	-0.79180	-0.67854	-0.60444	0.14920	0.62732
Cl	0.91944	0.84452	0.74726	0.86766	-0.16417	-0.52080
SO4	-0.08253	0.03814	-0.15308	-0.33187	0.25219	-0.37332
SiO2	0.54578	0.33061	0.53881	0.77617	-0.29362	-0.14663
B	0.11762	-0.16630	0.35471	0.40444	-0.44163	0.42482
HCO3	-0.27355	-0.18279	-0.16499	-0.30435	0.09599	0.81870
Rb	1.00000	0.89393	0.86795	0.70827	-0.16928	-0.41964
Cs	0.89393	1.00000	0.69822	0.51725	0.05366	-0.43629
F	0.86795	0.69822	1.00000	0.55602	-0.37572	-0.23205
Br	0.70827	0.51725	0.55602	1.00000	-0.21912	-0.39487
I	-0.16928	0.05366	-0.37572	-0.21912	1.00000	0.07845
NH3	-0.41964	-0.43629	-0.23205	-0.39487	0.07845	1.00000

TABLE A4

R-mode factor analysis using Varimax rotation

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	COMMUNALITY
Li	0.68363	0.30798	0.27709	-0.09996	0.69760
Na	0.32639	0.77912	0.12952	-0.24312	0.78817
K	0.76056	0.55197	-0.06816	-0.13375	0.90511
Ca	-0.03270	0.03369	0.79453	0.03501	0.65019
Mg	0.08474	-0.01285	0.08819	0.55095	0.31812
Cl	0.29071	0.88203	0.01322	-0.16944	0.89147
SO4	-0.56112	-0.12855	0.11299	-0.15092	0.36680
SiO2	0.72638	0.26326	-0.59103	-0.14750	0.91141
B	0.56899	0.61093	0.00829	0.30535	0.78963
HC03	-0.11767	-0.34221	0.00804	0.50355	0.38459

TABLE A5

Individual well-cluster association for each field

Well	Cluster	Well	Cluster
OHAAKI		**ROTOKAWA**	
1	4	RK1	4
2	4	RK2	5
3	4	RK3	5
4	4	RK4	4
5	3	RK5	4
6	6		
7	7	**ORAKEIKORAKO**	
8	4	OK2	8
9	4	OK4	8
10	4	OK6	8
11	4		
12	4	**WAITAPOU**	
13	4	6	8
14	4	7	3
15	4	4	7
16	1	3	8
17	4		
18	4	**MOKAI**	
19	4	MK1	1
20	4	MK2	7
21	4	MK3	7
22	4	MK4	7
23	4		
24	4	**WAIRAKEI**	
25	4	4/1	7
27	4	20	7
28	4	44	7
29	4	27	7
		28	7
KAWERAU		41	7
3	4		
7A	4	**TE KOPIA**	
e	4	TK1	8
14	4		
16	4		
17	4		
4	8		
10	8		
12	4		