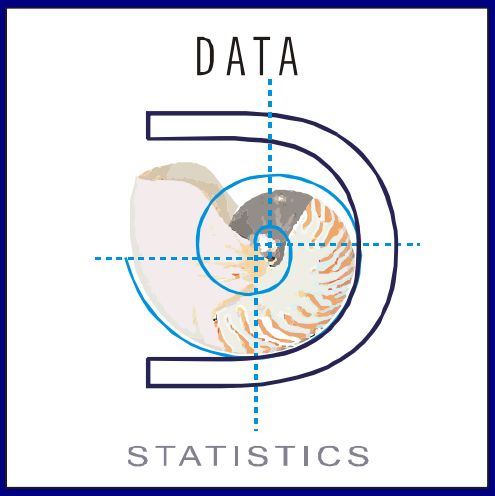

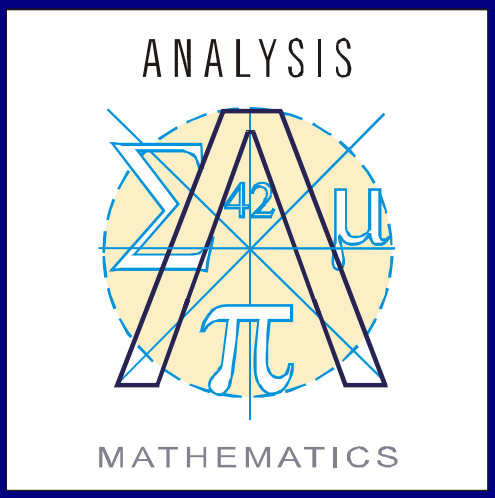

S T R A T E G I C
I N F O R M A T I O N
C O N S U L T A N T S

Analysis of Water Treatment Plant Workshop Data

April 2007

Project: WATER/55

DATA ANALYSIS AUSTRALIA PTY LTD

Analysis of Water Treatment Plant Workshop Data

April 2007

Client: Water Corporation

Project: WATER/55

Consultants: Dr John Henstridge
Natalie Smith

Data Analysis Australia Pty Ltd
97 Broadway
Nedlands, Western Australia 6009
(PO Box 3258 Broadway, Nedlands 6009)
Website: www.daa.com.au
Phone: (08) 9386 3304
Facsimile: (08) 9386 3202
Email: daa@daa.com.au

A.C.N. 009 304 956
A.B.N. 68 009 304 956

Table of Contents

1. Introduction

Mundaring Weir is the primary source of drinking water to about 100,000 people in the Hills suburbs of Perth, east of Hardey Road Glen Forrest, and those served by the Goldfields and Agricultural Water Supply Pipeline. Disinfection facilities have existed at Mundaring Weir for many years but as yet the water is not filtered. A water treatment plant is needed to introduce the filtration process for all water from Mundaring Weir to provide the best possible drinking water to the 100,000 people dependent on this source.

The Water Corporation conducted a series of public meetings and community forums in 2006, run by an independent facilitator, Colin Penter of Matrix Consulting. At the community forums site options for the water treatment plant were examined and participants in the forum were given the opportunity to rate the sites based on a number of criteria. Data Analysis Australia was commissioned by the Water Corporation to undertake an independent analysis of these ratings to identify the site(s) preferred by the community for the water treatment plant.

In carrying out the analysis Data Analysis Australia considered both the criteria being used and the rating processes. Consultants also attended a meeting of the Steering Group on 2 April 2007.

2. Background

During the forums, participants were asked to complete two forms:

- **Session One form** – Each participant selects the criteria by which to assess 11 sites to filter down to a smaller number of sites for further consideration. Participants may chose any or all of the 8 criteria, answering yes or no to each chosen one. They finally assess whether a site should be considered further by indicating yes or no.
- **Session Two form** – Of the sites they selected in the Session One form, each participant then assesses their selected sites in more detail. Participants were provided with a rating scale for each criterion, and chose to use all or some of the criteria to assess their chosen sites. The form also provides participants with the opportunity to comment on their selected sites, for example the reasons behind the ratings they gave.

The forms are given in Appendix A. The format of the forms was optimised for the decision making process in the forums, as it was essential to keep the burden on participants at a reasonable level. However the structure does not permit simple analysis and consequently more is required than simply adding up the scores.

A second complication was that not all participants followed the instructions perfectly. However no attempt was made to "reinterpret" intentions in these cases – data was used as provided wherever possible.

3. Analysis

3.1 Session One Forms

The first session aimed to focus participants upon sites that they might consider to be the most suitable. So that participants were not burdened with the task of considering all criteria and all sites together, they could use a subset of criteria and score all the sites by these using a simple yes or no response.

The analysis of the Session One data aimed only to give an understanding of how the criteria were being used. A simple statistical model was used to quantify how a yes or no response to a criterion (relative to not using that criterion) might influence the selection of a site for further consideration in Session Two.[1] The result of this analysis is shown in Figure 1.

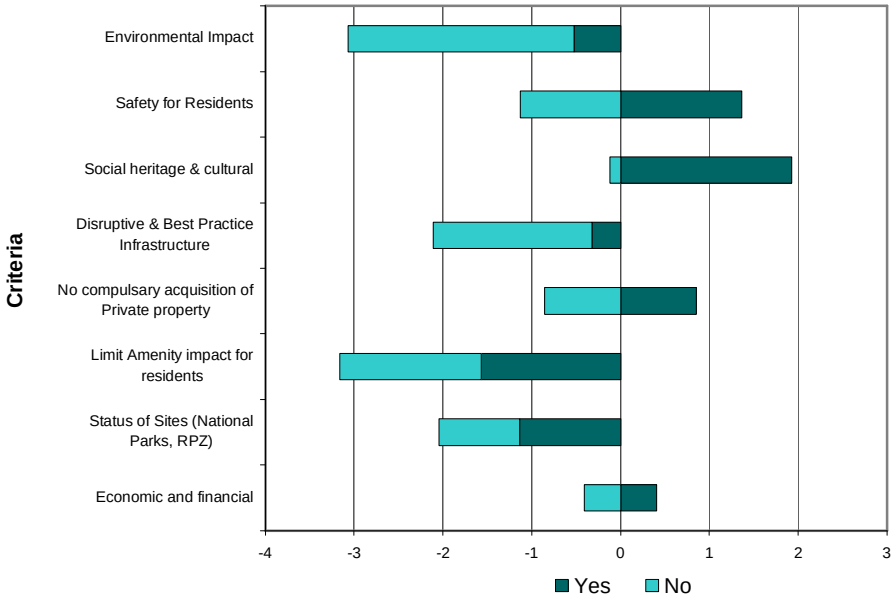

Figure 1. Effects of the various criteria on the likelihood of a participant selecting a site for further consideration in Session Two. The criteria are ranked according to the difference in impact of yes and no responses.

[1] The statistical method used was logistic regression and was implemented in the statistical package R. The analysis is approximate in that it assumed independence between the ratings given to each site by a participant.

Several observations can be made on this analysis:

- The yes responses were uniformly better linked to selecting a site than no responses. This suggests that the criteria were reasonably well understood.
- The criterion of most influence was "environmental impact" and the criterion of least influence was "economic and financial".
- For a number of criteria both yes and no responses would lead to a reduced likelihood of selecting a site. (Criteria where the bar is wholly to the left of the y-axis in Figure 1.) This at first paradoxical result is explained by the observation that participants who chose these criteria were less likely to chose as many sites for further consideration.

A key finding from the analysis is that the selection of sites for further investigation was clearly based upon people preferring those sites – participants implicitly stated that they would rank the sites they had not selected lower than those selected. This has important implications for the interpretation of Session Two results.

3.2 Session Two Forms

In Session Two participants assessed the sites they chose in Session One. This was done by rating each site against a number of criteria. These ratings were summed to give an overall rating or score of their chosen sites. A simple averaging of these scores is presented in Figure 2.

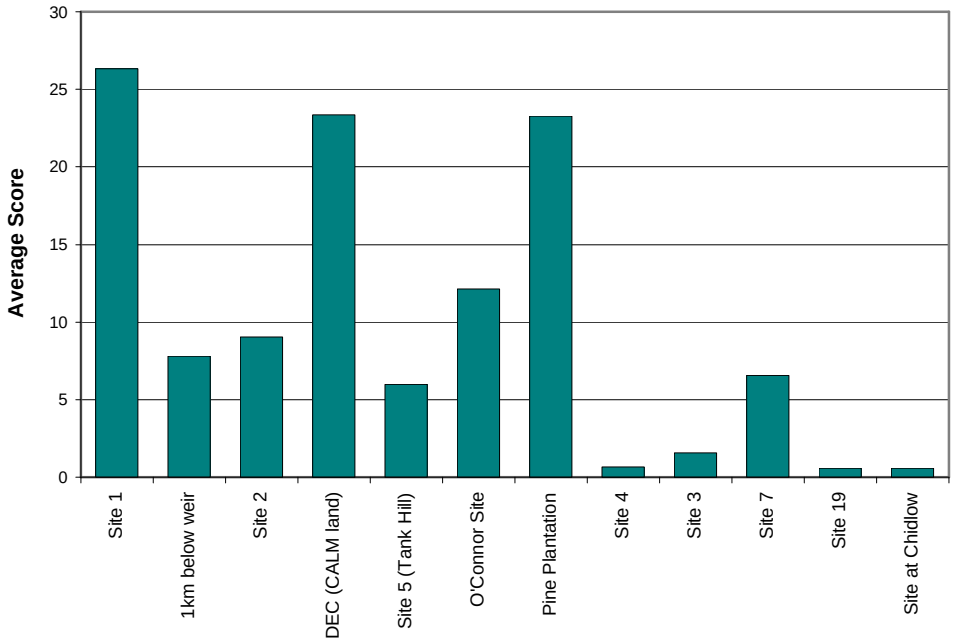

Figure 2. Simple averages of the scores given for each of the sites. (These averages implicitly give a zero score when a participant did not rate a site.)

The problem with simple average scores is that it makes implicit assumptions on the sites that a participant had excluded in Session One, effectively giving them a score of zero. It also takes at face value the way in which participants have used the five point scale for each criterion, while experience has shown that individuals differ in how strongly they emphasise differences. It is hard to justify these assumptions.

An analysis that used averages based on ratings for each site only from those participants who rated that site in Session Two is also flawed. Even a rudimentary view of the data indicated that all the averages would be similar and not helpful in distinguishing between sites.

After investigating various alternatives Data Analysis Australia decided to use a ranking method. For each participant, the ratings were converted into a ranking of the sites from 1 to *n*, where *n* was the number of sites selected by that participant in Session One. Where a participant gave two sites the same rating, they were given equal ranks.

The ranking of a number of items by several persons is a common task. In many of these situations there is a need to combine the opinions or assessments of the participants to give a single ranking that provides the best possible match to the individual rankings. The formal method of combining these individual rankings as described in Appendix B is based upon maximising the concordance of the common ranking with each of the individual rankings.

In an ideal situation each participant considers every item (sites in this case) and gives a *complete ranking*. In this case the mathematics provides several approaches that give reasonably straightforward methods of finding a common ranking. However in our situation participants did not provide complete rankings. Participants selected a subset of sites in Session One and only these sites were ranked in Session Two, resulting in a relatively sparse set of rankings.

An import issue to consider in this analysis is the implied rankings of the sites that were *not* chosen in Session One. It is not possible to compare or rank the sites not selected, but it is reasonable to rank them below all the sites that were selected. Therefore these sites were given a ranking on $n+1$.

3.2.1 Results

The results of the analysis are given below in Table 1:

Ranking	Assuming a rank of n+1 for sites not selected *(preferred method)*	Not assuming a rank for sites not selected
1	Site 1	Site 1
2	DEC (CALM land)	DEC (CALM land)
3	Pine Plantation	Pine Plantation
4	O'Connor Site	1km below weir
5	1km below weir	O'Connor Site
6	Site 2	Site 5 (Tank Hill)
7	Site 7	Site 4
8	Site 5 (Tank Hill)	Site 3
9	Site 3	Site 2
10	Site 4	Site 19
11	Site 19	Site at Chidlow
12	Site at Chidlow	Site 7

Table 1. Common rankings, with and without the assumption that not selected sites should be ranked lower.

The first ranking given in Table 1 is the preferred solution, although it is interesting that the second ranking, that makes no assumptions about the meaning of Session One, gives a very similar result. The principal difference is in the fourth and fifth rankings.

4. Conclusions

Whatever analysis method is used, three sites stand out as being most widely acceptable – Site 1, DEC (CALM Land) and the Pine Plantation in that order, although the level of support for these is very similar. It would seem appropriate that any further discussion of sites should include consideration of these.

The ranking of sites after these three is slightly less clear, although the most logical ranking would place the O'Connor site fourth and the site 1 Km below the weir fifth. If only four sites were to be given further consideration then the fourth should logically be the O'Connor site, although it could be argued that if resources permit the site 1 Km below the weir should also be considered, giving five in all.

A. Appendix: Community Forum Forms

Session One

Filtering of Sites to a smaller number

Name: ____________________

Signature: ____________________

Are you a registered forum member ? Yes/No ____________________

Group: ____________________

Cost	Land Tenure	Site	Essential						Important		Should the site be considered for further assessment? Yes/No
			Safety for Residents	No compulsory acquisition of Private property	Disruptive & Best Practice Infrastructure	Environment Impact	Limit Amenity Impact for residents	Social, heritage & cultural	Economic and financial	Status of Sites (National Parks, RPZ)	
$125 M	WC's Reserve	**Site 1**									
$128 M	National Park	**1 km below Weir**									
$143 M	National Park, RPZ	**Site 2**									
$145 M	Part DEC & part state forest	**DEC (CALM land)**									
$147 M	National Park	**Site 5 (Tank Hill)**									
$150 M	State forest	**O'Connor Site**									
$156 M	State forest, RPZ	**Pine Plantation**									
$161 M	Private Property	**Site 4**									
$166 M	Private Property	**Site 3**									
$188 M	WC's Property	**Site 7**									
$203 M	Private property	**Site 19**									
$222 M	Private property	**Site at Chidlow**									

Step 1:
Decide which criteria should be used to filter the sites to a smaller number. Circle the criteria you have chosen.

Step 2:
Then on each criteria selected, decide whether the site should be considered for further assessment.

Answer Yes, if you want this site to be considered further.

Answer No, if you do not want this site to be considered further.

Step 3:
Then assess whether the site should be considered for further assessment.

Session Two

Detailed assessment of a smaller number of Sites

Name: _____

Signature: _____

Are you a registered forum member ? Yes/No _____

Group: _____

Step 1
In this column, tick the sites you have selected in the Session One.

Step 2
Assess only the sites you have selected (ticked) using the rating scale below. Use all the criteria.

You do not have to derive the total. The computer will do that when Water Corporation analyses the results.

				Essential						Important			
Cost	Land Tenure	Site	Please tick the sites that you have selected in Session One	Safety for Residents	No compulsory acquisition of Private property	Disruptive & Best Practice Infrastructure	Environment Impact	Limit Amenity impact for residents	Social, heritage & cultural	Economic and financial	Status of Sites (National Parks, RPZ)	My Total	**Your thoughts.** Please give any thoughts you have about your selected sites. Include anything you think is relevant to the selection or that you think should be taken into consideration when considering this site. If you need more space write at the back of this sheet.
$125 M	WC's Reserve	**Site 1**											
$128 M	National Park	**1 km below Weir**											
$143 M	National Park, RPZ	**Site 2**											
$145 M	Part DEC & part state forest	**DEC (CALM land)**											
$147 M	National Park	**Site 5 (Tank Hill)**											
$150 M	State forest	**O'Connor Site**											
$156 M	State forest, RPZ	**Pine Plantation**											
$161 M	Private Property	**Site 4**											
$166 M	Private Property	**Site 3**											
$188 M	WC's Property	**Site 7**											
$203 M	Private property	**Site 19**											
$222 M	Private property	**Site at Chidlow**											

Rating Scale								
1	very unsafe	significant impact	substantial disruption	significant impact	significant impact	significant impact	most expensive	significant impact
2								
3								
4								
5	very safe	no impact	no disruption	no impact	no impact	no impact	cheapest	no impact

Note:
The higher the score the more you like the site.

DATA ANALYSIS AUSTRALIA PTY LTD

B. Appendix: Concordance Algorithm

It is necessary to introduce some limited notation to accurately describe the concepts that follow. We assume that there are R participants and n items being rated. The ith participant provides a ranking $r_i=(r_{i1},r_{i2},\ldots r_{in})$, giving each item a number indicating how it compares with the others; 1 for the best item possibly through to n for the worse item. To allow for the possibility that some cannot be ranked we may consider some rankings to be missing. These rankings must be compared against a common ranking $\boldsymbol{r}=(\boldsymbol{r}_1,\boldsymbol{r}_2,..,\boldsymbol{r}_n)$.

A basic principle that has been shown in statistical contexts to have effective and useful properties is the concept of concordance. This considers just two rankings and two items at a time. If the two rankings order the two items in the same way, they are said to concur. If they disagree it is called a discordance. The discordance between two rankings is the number of such pairwise discordances.

The aim is to find a common ranking that minimises the total discordance with the individual ratings. That is, the number of instances where $r_{ij}<r_{ik}$ and $\boldsymbol{r}_j>\boldsymbol{r}_k$. This concept of discordance has been used in many statistical constructs such as rank based non-parametric tests and measures of association such as Kendall's τ. (See for example, Kendall and Gibbons, 1990.) It also has the advantage of being mathematically manageable.

The concordance algorithm works by four stages:

- A preliminary ordering of the items is given using the average rankings. Where a participant has not ranked an item, that participant is ignored for the calculation of the average for that item.
- Since it is particularly easy to calculate the effect on the total discordance of transposing two adjacent items, the preliminary ordering is refined using a bubble sort type algorithm until the total discordance cannot be further reduced.
- An exhaustive *branch and bound search* is then applied starting from this ordering, limiting the search to permutations of the first twelve items. (This limitation is the only non-optimal step in the approach - obviously if only twelve or less items are being evaluated the approach is optimal. If it was critical to be certain of the global optimum, this restriction could be relaxed provided sufficient computing resources are made available.) Crucial to this stage is the order in which the orderings or permutations are generated. Standard algorithms (see for example Reingold *et al*, 1977) generate them in an order that does not allow a branch and bound technique. An alternative generation algorithm was used here.

The result from the third stage will in general not be unique - there may be several orderings that score equally in terms of total discordances. Since the usual application only requires choosing the best items - typically the top three - in many cases this is not a problem since the multiple solutions might only diverge after the first ones. However to provide a general solution the fourth stage scores each of the items in each of the solutions and ranks the items using the average score. The output is both the ranks and the scores so that if ambiguities remain they can be readily identified.