TECHNICAL NOTE 2008/21

Department of Water, Land and Biodiversity Conservation

ATMOSPHERIC HYDROGEN SULPHIDE IN THE LOVEDAY BASIN: CAUSES AND MANAGEMENT OPTIONS

Mark Walter and Nicholas Souter

May 2008

ISBN 978-1-921218-99-6

Preferred way to cite this publication

Walter M, & Souter N, 2008, *Atmospheric hydrogen sulphide in the Loveday Basin: causes and management options,* DWLBC Technical Note 2008/21, Government of South Australia, through Department of Water, Land and Biodiversity Conservation, Adelaide.

Knowledge and Information Division

25 Grenfell Street, Adelaide
GPO Box 2834, Adelaide SA 5001

Telephone	National	(08) 8463 6946
	International	+61 8 8463 6946
Fax	National	(08) 8463 6999
	International	+61 8 8463 6999
Website	www.dwlbc.sa.gov.au	

Government of South Australia
Department of Water, Land and Biodiversity Conservation

CONTENTS

List of Figures

List of Tables

INTRODUCTION

Since 2002, reports of objectionable odour have been made by residents surrounding Loveday Basin (Burton, 2003). While the causes and mechanisms of this odour generation are not clearly defined, they are almost certainly associated with the sulphur (S) cycle in disposal basins (such as Loveday Basin) where water levels have been lowered (Hicks and Lamontagne, 2006). Wetlands may emit three main types of odorous sulphur gases: hydrogen sulphide (H_2S), volatile organic sulphur compounds (VOSC's) and sulphur dioxide (SO_2). These gases may be produced by a range of different inorganic and organic chemical pathways—pathways that in turn are influenced by the wetlands' salinity, wetting–drying regime, soil type and diurnal cycle (Hicks and Lamontagne, 2006). Thus the emission rates (amount produced per unit area of sediments per unit time) of these sulphur gases will vary according to the prevailing environmental conditions in the basin. Determining the emission rates of odorous gases is complicated at large sites such as Loveday, which exhibit considerable spatial and temporal variability, and requires the use of specialist equipment such as incubation chambers and micromet methods. However, instrumentation to measure the concentration of H_2S and SO_2 is readily available.

The main concern associated with the emission of sulphurous gases from Loveday has been the potential health and amenity effects at Cobdogla. The most frequently reported odour was rotten egg gas—the smell of H_2S—as opposed to the pungent/irritating smell produced by SO_2 (Hicks and Lamontagne, 2006). Towards the end of 2003 a monitoring program was implemented to measure the ambient H_2S concentration coming from Loveday Basin. It was expected that H_2S concentration would be highest at low water levels, in the summer and when the wind blows from the wetland. H_2S concentration was also expected to be greater in the morning, especially in summer, as temperature inversions at night would tend to trap H_2S closer to the wetland and elevate morning readings.

In this report we use exploratory statistical modelling to determine the relative importance of a range of factors that may effect H_2S concentration, and where required, use subsequent analysis to examine the effect of basin water level, season and wind direction on H_2S concentration. This information is then used to suggest how to best manage Loveday's odour problem.

METHODS

VARIABLES

Atmospheric hydrogen sulphide (H_2S) concentration was measured in parts per million (ppm) using a H_2S direct gold film collection analyser. The monitoring equipment was located at the Humphries pump compound part–way between the basin and the Cobdogla township (Fig. 1). Measurement of H_2S was recorded continuously in ten-minute averages from 25/11/2003 until 1/05/2007 using a Unidata 6004 Starlogger Datalogger. H_2S concentrations for 9am and 3pm on each day over the recording period were supplied to the Department of Water, Land and Biodiversity Conservation (DWLBC) and used for analysis.

As continuously–measured variables are commonly autocorrelated, we modelled for possible autocorrelation effects by adding an explanatory variable. Autocorrelation was modelled as the weighted moving average of H_2S over a period of 180 days.

Government of South Australia
Department of Water, Land and Biodiversity Conservation

Produced By: River Murray Assessments
Knowledge and Information Division
Department of Water, Land and Biodiversity Conservation

Datum: Geocentric Datum of Australia 1994 (GDA 94).

Date: March 08

Figure 1 **Location of Loveday Basin, Loxton Research Centre meteorology station and hydrogen sulphide monitoring station**

Water level data was manually collected intermittently from 25/11/2003 until 11/3/2005. During this period 57 measurements were made. Logged water level was recorded twice daily at 9am and 3pm from 11/03/2005 until 11/04/2005 and from 10/6/2005 until 10/4/2007. In the intervening month between periods of logging and from 11/4/2007 until 1/5/2007, four manual measurements were taken.

A range of meteorological data were obtained from the closest Bureau of Meteorology (BoM) station, Loxton Research Centre (024024) at Loxton (Fig. 1), which was 28 km from the H_2S monitoring equipment. Data obtained for the period of 25/11/2003 until 1/05/2007 are listed in Table 1.

Table 1 Meteorological variables obtained from Loxton Research Centre from 25/11/2003 to 1/05/2007

Meteorological variable	Frequency collected	Unit
Minimum Air Temperature	Daily	°C
Maximum Air Temperature	Daily	°C
Precipitation to 9am	Daily	mm
Air Temperature	9am and 3pm	°C
Relative Humidity	9am and 3pm	%
Heating degree-days {base 20°C}	Daily	degree-days
Heat Accumulation {base 20°C}	Daily	degree-days
Evaporation to 9am	Daily	mm
Total Cloud Amount	9am and 3pm	oktas
Mean Sea Level Pressure	9am and 3pm	hPa
Bright Sunshine	Daily	hours
Speed of Maximum Wind Gust	Daily	km/h
Direction of Maximum Wind Gust	Daily	° true
Time of Maximum Wind Gust	Daily	local
Wind Run above 3 m	Daily	km
Wind Speed	9am and 3pm	km/h
Wind Direction	9am and 3pm	° true

ANALYSIS

Statistical modelling of the relationship between hydrogen sulphide concentration and the range of potential explanatory variables was undertaken using regression trees. Regression trees explain the variation of a single numerical response variable by one or more explanatory variables (De'ath and Fabricius, 2000). A regression tree is constructed by recursively partitioning the data and sample space (Loh, 2002). These partitions or splits separate the data into two mutually exclusive groups that are considered to be significantly different, with earlier splits representing stronger and more significant splits. At each terminal node a regression model is derived for the dataset applicable to the particular terminal node. Regression tree analysis was undertaken using Loh's (2002) GUIDE algorithm.

The complete timeseries dataset, including missing variables, was used to undertake the analysis. Initial exploration of regression trees was performed using a 0.5-Standard Error (SE) rule, whereby the best tree is taken as the smallest tree, such that its estimated error rate was within half a standard error of the minimum value for a tree of any size. This enabled the data to be explored in some detail without a tree achieving an unwieldy size. Once the most appropriate data had been chosen, this was then increased to a 1-SE rule as suggested by Breiman et al. (1984). After examination of the resulting tree this was further increased to 1.5-SE as it was considered the use of some climate variables from Loxton (28 kms distant) induced a degree of background noise that justified a stronger focus on the higher splits. A 1.5-SE rule ensured that the splits focussed on the strongest and most significant divisions.

RESULTS

H_2S CONCENTRATION AND WATER DEPTH

H_2S concentration over the period of 25/11/2003 until 1/05/2007 fluctuated considerably over the monitoring period (Fig. 2). High H_2S concentration appeared to be associated with low water level and conversely low H_2S concentration was associated with high water level (Fig. 2).

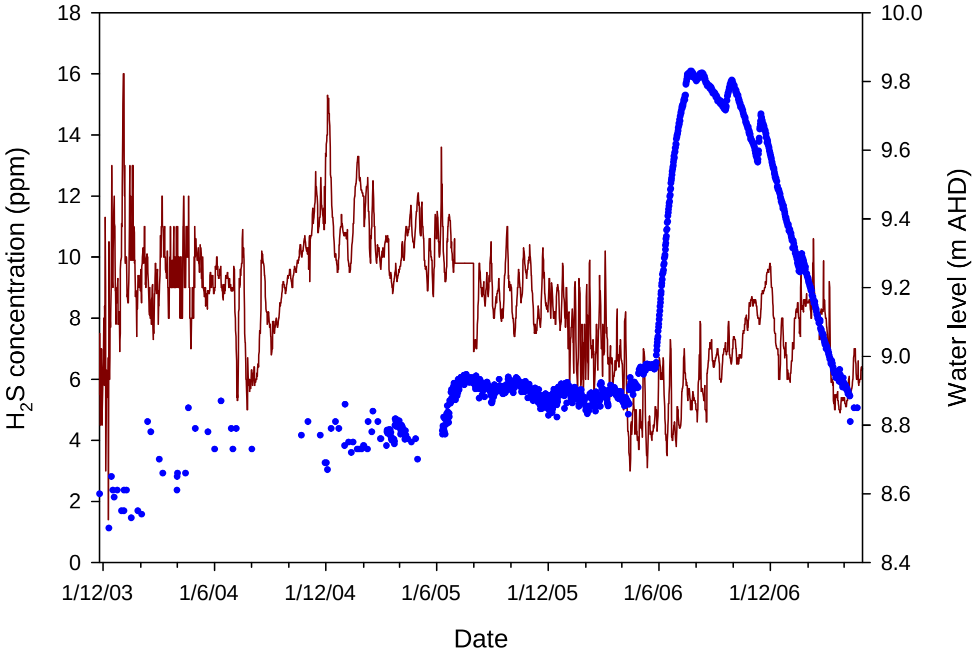

Figure 2 **H_2S concentration and basin water level for the period of 25/11/2003 to 1/05/2007 (brown line – H_2S concentration; blue line – north basin water level)**

DAILY DIFFERENCES IN H_2S CONCENTRATION BY SEASON AND WIND DIRECTION

An examination of the difference in H_2S concentration between the morning and afternoon (morning minus afternoon concentration) showed a possible increased frequency of higher afternoon H_2S concentrations in spring (Table 2) with no evidence of higher morning temperatures in any season, as there was little difference in either the frequency or mean values for positive (higher morning temperatures) or negative (higher afternoon temperatures) differences. In summer, the frequency of a change in H_2S concentration over the day was more frequent, as was the mean magnitude of this change (Table 2; Fig. 3).

Table 2 **Frequency and mean (±SE) magnitude of change in H_2S concentration (ppm) from each season over the period of 25/11/2003 to 1/5/2007**

	Positive difference frequency (%)	Negative difference frequency (%)	No difference frequency (%)	Mean (±SE) positive difference	Mean (±SE) negative difference
Spring	26	40	34	0.269±0.035	-0.208±0.024
Summer	42	40	18	0.521±0.064	-0.511±0.048
Autumn	35	36	29	0.396±0.035	-0.384±0.039
Winter	37	35	28	0.326±0.036	-0.335±0.037

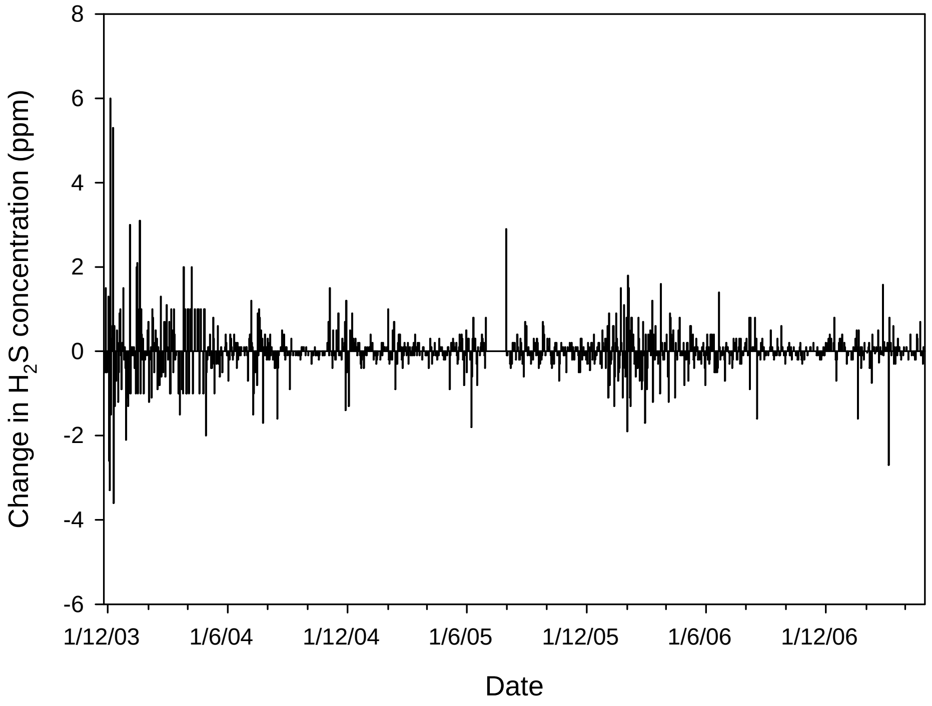

Figure 3 **Change in H_2S concentration (ppm) over the period of 25/11/2003 to 1/5/2007**

EXPLORATORY STATISTICAL MODELLING

With regard to variance explained, the best regression tree was Tree 3 (see Table 3). The most significant factor in the development of this tree was the use of daylight hours to represent seasonality and categorical data for both the wind direction and the direction of maximum wind gust. The first two regression trees used numerical datasets (direction measured in degrees). This however, produced spurious results due to the presence of the artificial break created by the crossover from 359^{o} (1^{o} west of North) to 0^{o} (directly north). As a result, these variables were transformed into 16 categories (e.g. NE, SSE, W, etc.). Tree 3 however, was ultimately discarded as the grouping of categories was found to be inconsistent on inspection (e.g. in one case direction of maximum wind gust from either the NE or SSE was considered to positively influence hydrogen sulphide concentration, with none of the intervening categories considered to have any affect). This led to both the wind direction and the direction of maximum wind gust variables being excluded from all further analyses, resulting in a small drop in the variance explained but more sensible results.

Table 3 **Atmospheric H_2S regression trees**

	Variables in dataset	Variables used by models	SE – rule used	Nodes	Terminal Nodes	Variance Explained	Comments
Tree 1	20	19	0.5	41	21	87.97	Meaningless splits (query seasonality)
Tree 2	21	16	0.5	13	7	83.71	Spurious results (query artificial breaks)
Tree 3	21	20	0.5	21	11	88.53	Meaningless results (query categorical data)
Tree 4	19	17	0.5	21	11	86.19	Sensible results, Overly large
Tree 5	19	16	1	19	10	85.69	Unnecessary splits
Tree 6	19	12	1.5	11	6	82.75	Unused variables
Tree 7	12	12	1.5	11	6	82.76	Small tree, generally appropriate results

Tree 5 represents the most appropriate tree using the 1-SE rule, however this was still considered to be overly large, especially as it made multiple splits along a number of branches based upon the same variable. Ultimately Tree 7 was considered to achieve the greatest balance of overall accuracy (variance explained = 82.75), small tree size and meaningful splits.

Tree 7 used three variables to create five splits (Fig. 4, Table 4), the first split being based upon day length less than or equal to 14.308 hours. In practice this equates to 8 January through to 5 December. While this is a small split (a greater component of the summer period was expected), it is unsurprising that the summer period was identified as being one of the most significant divisions likely to result in increased H_2S concentration. Both of the splits along either branch from the first node were then split based on water level in the North Basin with levels below approximately 9 m AHD associated with high H_2S concentration. The impact of these splits is clearly demonstrated in Figure 5, where the slope of the trendline if water level is below 8.899 m AHD (the split used in node 3) is negative, whereas the trendline for values above 8.899 m AHD is almost horizontal, indicating that exceeding this threshold would have little impact on the generation of odours.

Other splits were based upon evaporation and, once again, day length, with higher evaporation leading to higher H_2S concentration, and day length of less than 10.617 hours (approx. 6 May to 8 August) surprisingly increasing the likelihood of high H_2S concentration. This may however, be due to a lag effect in seasonality, as although the shortest day of the year occurs in mid to late June, the mid point for the season of winter is considered to be mid July. This is clear in Figure 6 where July, August and September exhibit significantly lower median and maximum H_2S values when compared with May and June.

In all of the regression models (see App. 1), the autocorrelation factor was one of the most significant coefficients, indicating that periods of high H_2S concentration were related to preceding conditions as opposed to being a result of the immediate conditions. The importance of day length was further emphasised with its presence in four of the six regression models and as one of the more significant coefficients in three of these. This was generally as a positive coefficient although in node 16 this was negatively related to H_2S concentration. This was most likely as a result of the lag effect in seasonality, as discussed earlier.

The use of water level in the north basin varied across the regression models. It was also used in four of the six models—in two cases (nodes 9 and 6) this was as one of the stronger coefficients where it was negatively correlated. This is as expected and matches with the trends demonstrated in the splits at nodes 2 and 3 (e.g. an increase in water level leads to a decrease in H_2S concentration). In the other two cases (nodes 16 and 7) it is positively related which is surprising, although in both of these instances it is as one of the weaker coefficients.

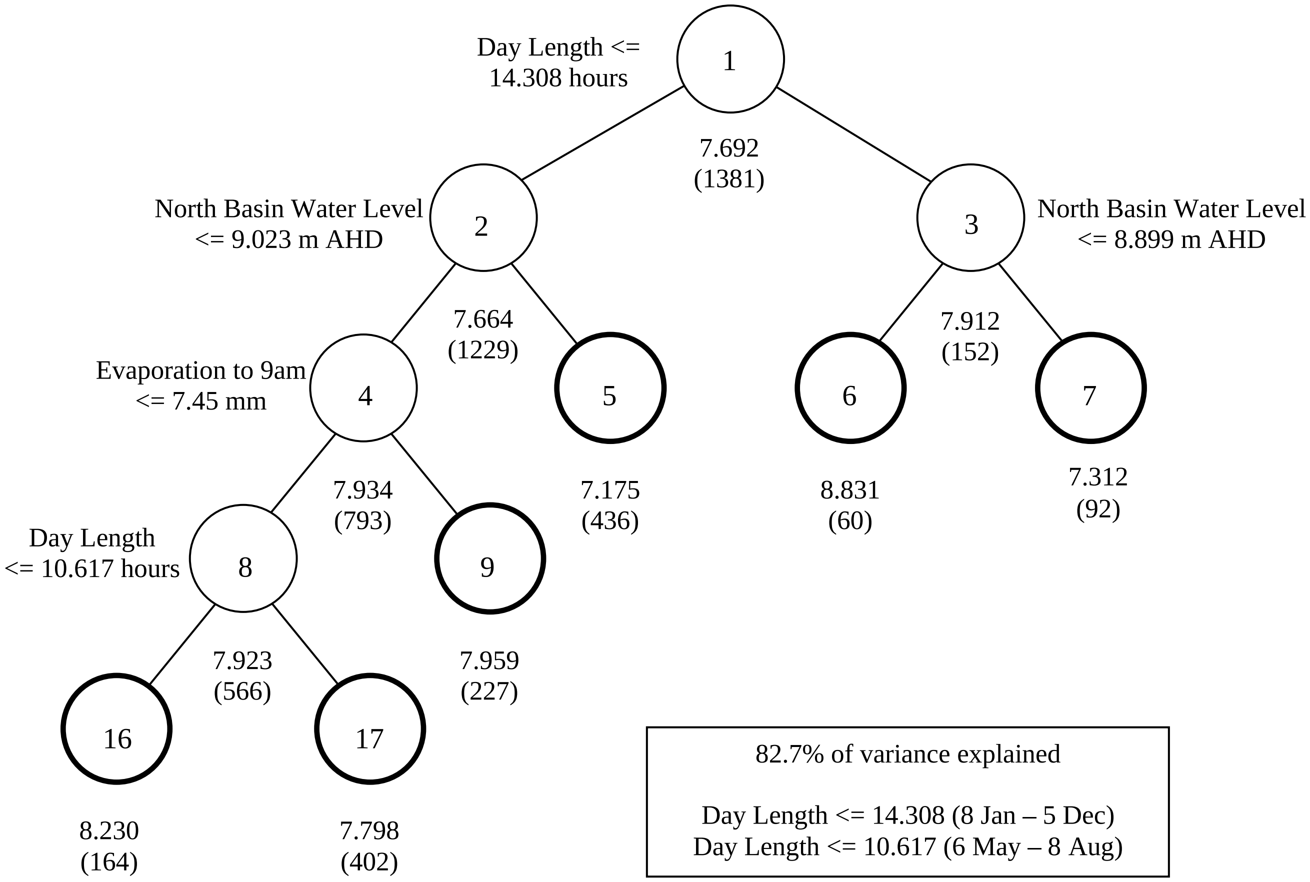

Figure 4 H$_2$S concentration regression tree (Tree 7) using data from 25/11/2003 to 1/5/2007. The splitting rule is presented next to each node and, if true, the node splits to the left. Below each node the mean H$_2$S concentration of all values applicable to the node and, in brackets, the number of applicable values.

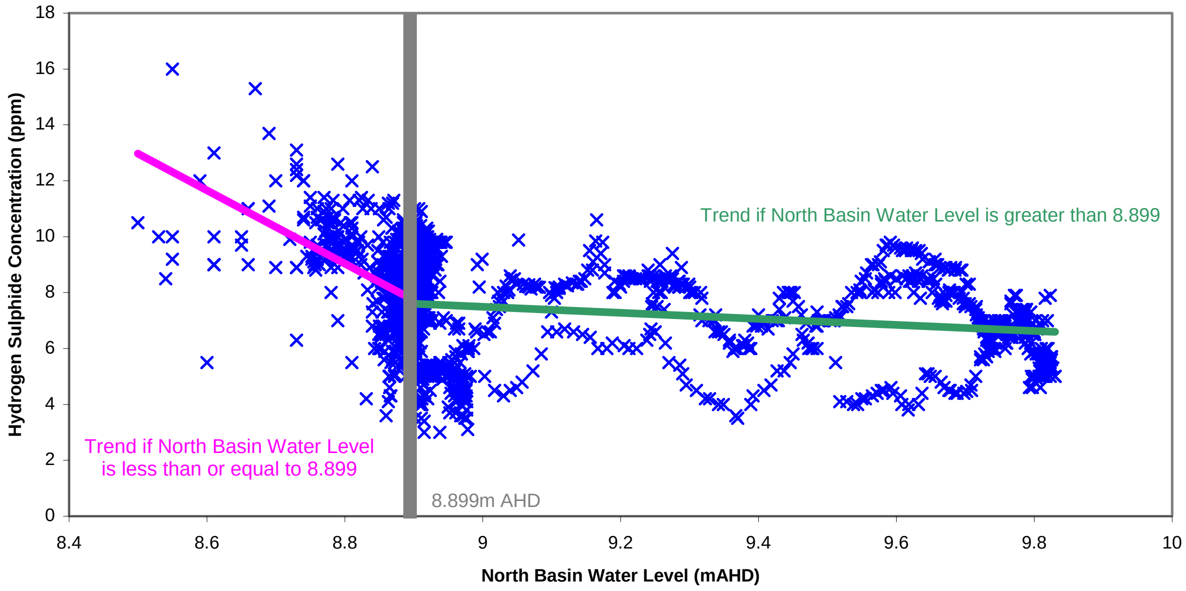

Figure 5 H$_2$S concentration against water level taken from the North Basin. Linear trend lines are shown for two datasets, which are separated based upon the node 3 split

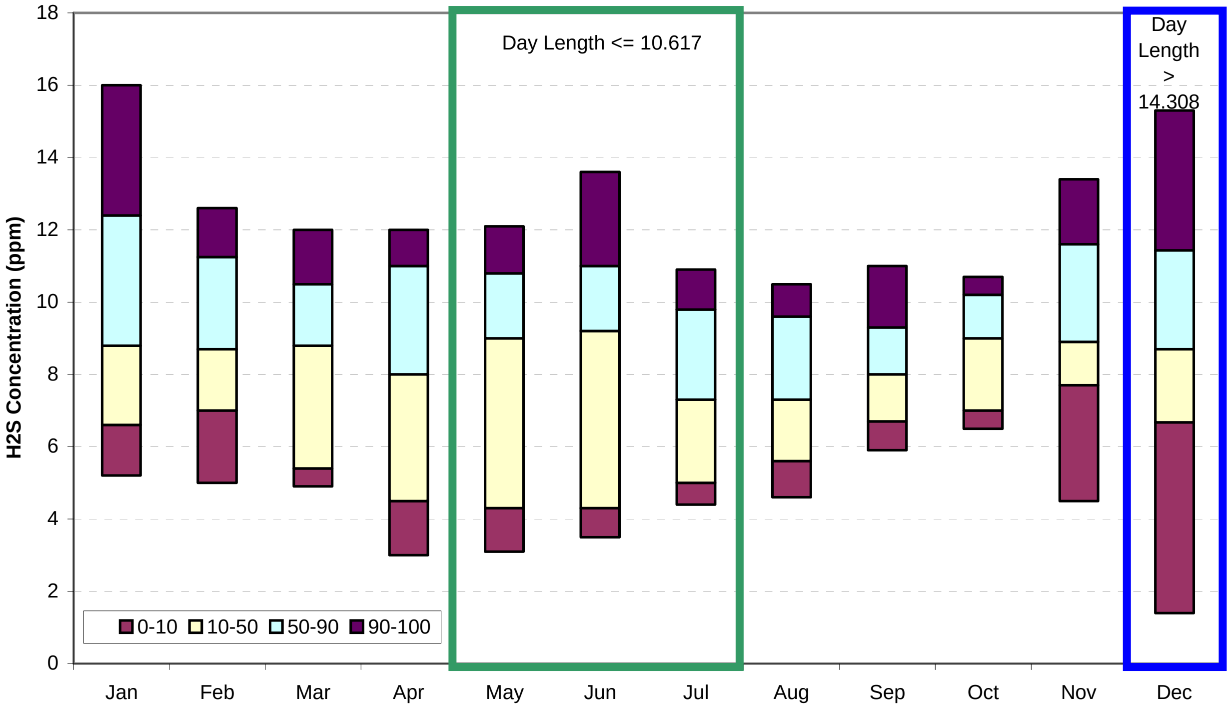

Figure 6 H_2S concentration percentiles by month

It is interesting to note those variables that were not used by the model. That the two variables dealing with wind speed were not used may be due to the lack of direction data, which would have provided them with further context. The use of the wind run variable appears to suggest that wind variables do play some part in atmospheric H_2S concentration, although this was only used in the regression model for node 17 where it had a relatively minor impact. Wind was not expected to be a driver in the production of H_2S, however it was hypothesised that it may regulate its concentration and dispersion.

It is also worth noting that no direct measure of temperature was used in any modelling. This also goes against prior expectations but may indicate that temperature is not a direct factor but rather a surrogate for seasonality and other factors such as evaporation and heat accumulation.

DISCUSSION

As expected, low water level in the north basin corresponded with higher levels of H_2S. The branching of the regression tree at nodes 2 and 3 indicates the presence of a strong threshold, whereby the risk of higher H_2S concentration increases noticeably if the water level falls below 9 m AHD. Away from this value, it appears that the impact of water level lessens to the point where it is at times positively or negatively correlated to hydrogen sulphide concentration within the regression models for various terminal nodes (Fig. 5).

It was clear from the branching of the regression tree that the summer period poses the greatest risk of high H_2S concentration. This however, does not suggest that other periods are of no concern. It is clear from Figure 3 that the maximum H_2S concentration measured in each month has been greater than 10 ppm—well above the target threshold this project has been working with of 8 ppm (Burton, 2003). Similar to the summer period, the majority of risk factors are

Table 4 Variables used by Tree 7

Dependent Variables	
	Hydrogen Sulphide Concentration (ppm)
Variables used in creating Splits	
	Day Length (hours)
	Evaporation to 9am (mm)
	Water Level – North Basin (m AHD)
Variables used to fit regression equations	
	Auto-Correlation Factor
	Bright Sunshine (hours)
	Day Length (hours)
	Evaporation to 9am (mm)
	Heat Accumulation (degree-days:base 20^{o}C)
	Mean Sea Level Pressure (hPa)
	Precipitation to 9am (mm)
	Relative Humidity (%)
	Total Cloud Amount (oktas)
	Water Level – North Basin (m AHD)
	Wind Run above 3m (km)
Unused Variables	
	Current Temperature
	Heating (degree-days:base 20^{o}C)
	Maximum Temperature (^{o}C)
	Minimum Temperature (^{o}C)
	Speed of Max. Wind Gust (km/h)
	Wind Speed (km/h)
Discarded Variables	
	Direction of Max Wind Gust (^{o} true)
	Wind Direction (^{o} true)

environmental variables that cannot be directly managed (such as evaporation). The focus here is on what variables can be managed and this is clearly the water level in the North Basin.

The expectation that H_2S concentration would be higher in the morning, especially in the summer was not supported by the results. Generally the differences between morning and afternoon evened out over each season, with the exception of spring. The higher frequency of spring afternoon readings may have been caused by the southerly winds that are prevalent at this time of year. However, this was not supported by the modelling—perhaps due to the distance between the H_2S measuring equipment and the meteorological station. The lack of any detectable difference in summer may have been due to either the expected phenomenon not occurring or the phenomenon not being able to be detected by the monitoring equipment, which was placed away from the lake edge towards the town.

Recommendations

1. Maintain water level at or above 9 m AHD at all times.
2. During summer it would be prudent to ensure a small buffer is in place due to the increased rate of evaporation (e.g. 9.2 m AHD).

REFERENCES

Breiman L., Friedman JH, Olshen RA and Stone CJ, 1984, *Classification and Regression Trees*, Wadsworth, Belmont

Burton RS, 2003, Loveday Basin Odour Control Management Program for the Department of Water, Land & Biodiversity Conservation – Draft, RS Burton – Environment, Science & Technology

De'ath G and Fabricius K, 2000, 'Classification and regression trees: a powerful yet simple technique for ecological data analysis', in *Ecology* 81(11), pp 3178–3192

Hicks W and Lamontagne S, 2006, *A guide to sulphur gas emissions from wetlands and disposal basins: implications for salinity management*, CSIRO

Loh W-Y, 2002, 'Regression trees with unbiased variable selection and interaction detection', in *Statistica Sinica*, vol. 12, pp 361–386

APPENDIX 1 – TERMINAL NODE REGRESSION MODELS

Table A1 Regression model for Terminal node 16

Variable	Coefficient	t-stat
Constant	9.8215	0.86
Auto-Correlation	1.0182	28.3
Day Length	-1.6852	-5.49
Mean Sea Level Pressure	-0.034245	-4.23
Water Level – North Basin	4.5782	3.14
Precipitation to 9am	0.075211	2.99
Relative Humidity	0.0071189	2.34

Table A2 Regression model for Terminal node 17

Variable	Coefficient	t-stat
Constant	-6.3221	-9.98
Auto-Correlation	1.2752	39.38
Day Length	0.2995	5.39
Evaporation to 9am	-0.068602	-2.5
Bright Sunshine	-0.037761	-2.46
Wind Run above 3km	0.001017	2.15

Table A3 Regression model for Terminal node 9

Variable	Coefficient	t-stat
Constant	-6.4985	-0.38
Auto-Correlation	1.2097	15.17
Water Level – North Basin	-4.8552	-4.09
Mean Sea Level Pressure	0.043878	3.46
Relative Humidity	-0.011267	-3.14
Day Length	0.23609	2.76
Precipitation to 9am	0.12272	2.28

Table A4 Regression model for Terminal node 5

Variable	Coefficient	t-stat
Constant	11.196	2.2
Day Length	0.49283	9.79
Auto-Correlation	0.60561	6.96
Heat Accumulation	-0.059851	-3.67
Mean Sea Level Pressure	-0.013663	-2.8

Table A5 Regression model for Terminal node 6

Variable	Coefficient	t-stat
Constant	95.606	9.93
Water Level - North Basin	-10.707	-9.83
Auto-Correlation	0.81179	5.56
Total Cloud Amount	0.12718	3.05
Precipitation at 9am	0.34072	2.35

Table A6 Regression model for Terminal node 7

Variable	Coefficient	t-stat
Constant	33.936	2.98
Auto-Correlation	1.8283	8.52
Mean Sea Level Pressure	-0.050588	-4.18
Bright Sunshine	0.087296	3.42
Evaporation at 9am	0.052638	2.38
Water Level - North Basin	0.93874	2.07