Salinity and Acidity in The Jimperding Brook in Western Australia

L. M. Bloom and J. M. Cross

School of Engineering and Mathematics, Edith Cowan University, Perth, Western Australia (l.bloom@ecu.edu.au)

Abstract: Water salinity and acidity are major problems in many streams and rivers in rural areas of Australia. The Jimperding Valley of Western Australia is an agricultural area located approximately 80km north east of Perth. Single land holdings range from 2 to 1600 hectares and land use ranges from small residential blocks to hobby farms to full scale farming operations. Jimperding Brook itself is an important tributary of the Avon River and has a catchment area of about 150 square kilometres (15 000 hectares). Some parts of this valley have for over ten years had an obvious salinity problem, as has Jimperding Brook itself. Water samples for pH and electrical conductivity (EC) were taken on a monthly basis from 1993 to 2000 and in this paper we present an analysis of these data. An extra factor here is the need to take into account some missing data due mainly to the fact that Jimperding Brook does, on occasion, dry up over the summer months. We combine time series and geostatistical techniques to obtain a model for use in predicting future acidity and salinity levels. We use ordinary kriging to estimate the missing pH data and then apply a time series model. The EC data exhibit strong seasonal variation and we first model this trend with reference to rainfall at one and two month lags. We then estimate the missing residuals and apply a further time series model. In addition, we investigate whether the awareness by landholders of this environmental problem over recent years has resulted in any practical improvement in the real situation.

Keywords: Time series, Salinity; Water Quality; Acidity, Missing Data

1. INTRODUCTION

The Jimperding Brook is located in Western Australia, approximately 80km northeast of Perth. This stream is an important tributary of the Avon River and has a catchment area of about 150 square kilometres (15000 hectares). The Jimperding Valley is an agricultural area with single land holdings ranging from 2 to 1600 hectares and land use ranging from small residential blocks to hobby farms to full scale farming operations. Some parts of this valley have an obvious salinity problem, as does Jimperding Brook itself.

In 1993 a local Landcare Group (known as the Jimperding Catchment Group) was formed in the area with the intention of improving farming practice to redress some of the problems of excessive salinity and acidity. In addition a water monitoring program was established (at a single site not far from the point where Jimperding Brook flows into the Avon River) with a view to modelling changes to water salinity and acidity over time. The aim was to predict future salinity and acidity levels and also to have a reference base from which to gauge possible deterioration in the Brook or measure any improvement. The intention was to incorporate this analysis could also be incorporated into a District Landcare Group GIS package and, in combination with the information on land and water use, provide a basis for future action to reduce the salinity and acidity problem in the area. In fact the Jimperding Catchment Group disbanded after four years. It was essentially replaced by the Deepdale Landcare Group, whose focus is on those sections of the Avon River and the Jimperding Brook close to where the two streams meet.

The statistical software package MINITAB and the geostatistical software GSLIB [Deutsch and Journel, 1998] have been used throughout for the analysis.

2. SAMPLING AND DATA SETS

Samples were taken on a monthly basis at a site not far from the point where Jimperding Brook flows into the Avon River. The site chosen was immediately downstream of the Lovers Lane floodway crossing. Initially, sampling was done also at a second site approximately 100m further downstream but there was no significant difference between the readings at these two sites and subsequent sampling was done only at the first site. Water salinity was measured by means of electrical conductivity (EC) measured in mS/cm on a hand-held conductivity meter and water acidity was measured by pH on a hand-held pH meter. The rainfall data used in this study has been obtained from the Perth Bureau of Meteorology and are monthly totals recorded at Gidgie Springs (Station 10302). The full data set, denoted by *Total*, is for the period April 1993 to June 1999. However, a subset, denoted by *Sample5*, comprising only the measurements for the first five years (from April 1993 to March 1998) has been used to determine the models to be used for pH and conductivity prediction purposes, with the remaining observations being reserved for the evaluation of these models.

Summary statistics for pH and EC from each of *Total* and *Sample5* are given in Table 1 and Table 2 respectively.

Table 1. Summary Statistics: *Total.*

	EC	pH
Number	67	67
Mean	8.96	7.72
Std. Dev.	3.72	0.25
Maximum	19.67	8.17
Upper Quartile	11.25	7.89
Median	8.34	7.81
Lower Quartile	5.89	7.53
Minimum	2.91	6.90

Table 2. Summary Statistics: *Sample5.*

	EC	pH
Number	52	52
Mean	8.51	7.73
Std. Dev.	3.70	0.24
Maximum	19.67	8.17
Upper Quartile	10.26	7.89
Median	7.88	7.82
Lower Quartile	5.81	7.56
Minimum	2.91	6.90

Due to the seasonal rainfall pattern, it is not unusual for little or no rain to fall during the summer months of December to February. When this happens Jimperding Brook becomes merely a series of non-flowing pools or even dries up completely. It is therefore not possible to obtain the pH and conductivity readings until the flow recommences. In the period under consideration, for *Sample5*, no observations were possible for this reason in March 1994, January and February 1995, February and March 1996 and January and February 1998. In addition, no pH reading was made in December 1994 due to pH meter problems and no EC reading was made in January 1996 due to conductivity meter problems.

3. pH ANALYSIS

The time series plot for the pH data from *Sample5* is given in Figure 1.

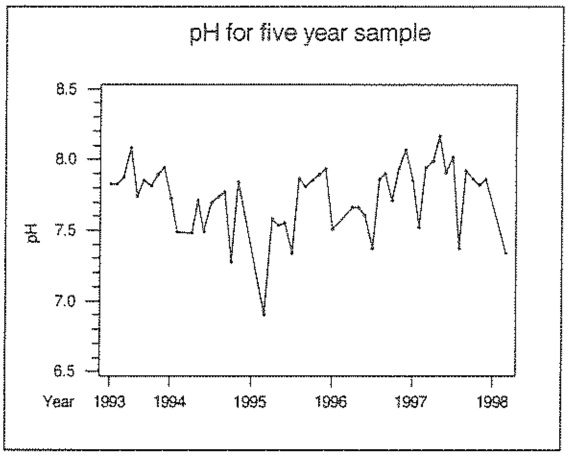

Figure 1. Plot for pH from *Sample5.*

As expected, the pH values exhibit no obvious seasonal trend. It can be seen from Figure 1 that the pH values recorded lie in the neutral to low basic range. The minimum value of 6.90, recorded in March 1995, in an outlier for these data. This was the first reading after the stream started flowing again after a dry period in early 1995 and this could account for this unusually low pH value. We note here that a study [Smith, 1996] was carried out in October 1995, involving one-off measurements of (among other things) pH and conductivity levels in the various pools and tributaries of the Avon River. One site chosen was (from the description given) in the same general section of Jimperding Brook as our own measurements. Smith recorded a neutral or basic pH value at all but two of the 21 sites considered, with Jimperding Brook being one of these exceptions. This anomaly is not reflected in our measurements, which, with the one exceptional measurement indicated, show Jimperding Brook over a six-year period as having stream pH values typical for the region.

There are various methods for determining missing values, including numerical interpolation and structural time series models [Chatfield 1996; Harvey 1989]. However for pH we decided to use the geostatistical method of Ordinary Kriging [see Goovaerts, 1997] to estimate the missing values. In order to do this, the experimental semivariogram for pH was calculated and this was modelled with a nugget of 0.02 together with a single exponential structure with sill 0.05 and effective range 12. This experimental semivariogram, together with the fitted model, is shown in Figure 2.

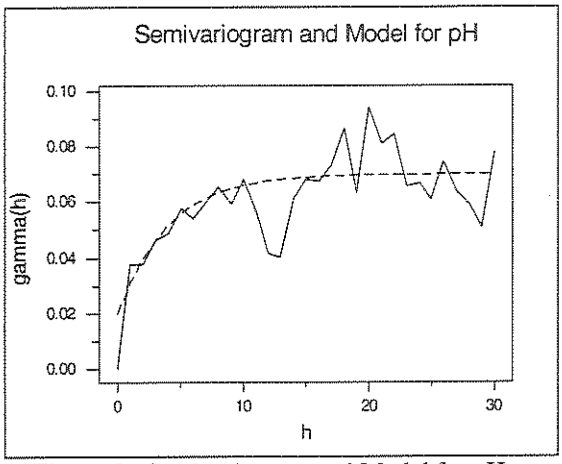

Figure 2. Semivariogram and Model for pH.

After estimating the missing values by use of ordinary kriging a time series model was fitted to the *Sample5* pH data. It was found that these data can be adequately represented by the Box Jenkins ARIMA model

$$pH(t) = 4.78501 + 0.3792\, pH(t-1) + Z(t)$$

where Z(t) is a purely random noise process with mean 0 and standard deviation 0.2183. The model fit is shown in Figure 3. The dashed line graph depicts the fitted values while the actual values are shown by the solid line graph.

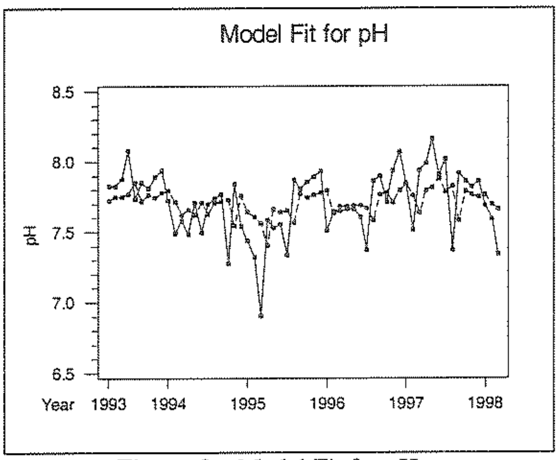

Figure 3. Model Fit for pH.

The model obtained is an Autoregressive process of order 1, which does not provide reliable forecasts beyond the month following the last data value.

4. CONDUCTIVITY ANALYSIS

The time series plot for the EC data from *Sample5* is given in Figure 4. It can be seen there that the electrical conductivity exhibits a strong seasonal trend and that this stretch of the Jimperding Brook has high (> 6 mS/cm) salinity levels for most of the time. From Figure 4 it can also be seen that the underlying EC levels appear to have fallen steadily from 1993 to 1997, but to have risen sharply early in 1998.

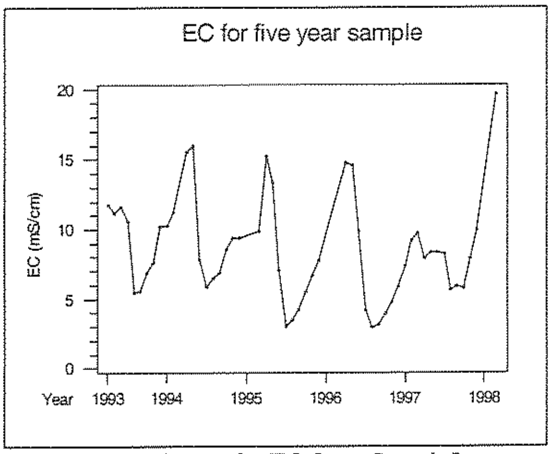

Figure 4. Plot for EC from *Sample5*.

In this area the majority of the rain falls in the winter months, flushing out streams such as the Jimperding Brook and reducing salinity levels. The time series plot for rainfall (recorded officially at Gidgie Springs) is given in Figure 5.

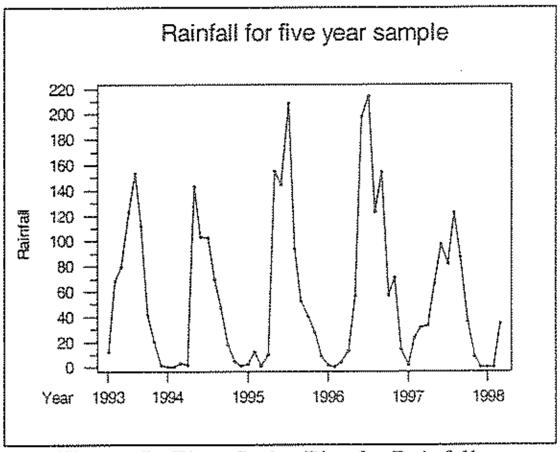

Figure 5. Time Series Plot for Rainfall.

A reasonable conjecture, and one seemingly borne out by the time series plots for EC (Figure 4) and rainfall (Figure 5), is that both are seasonal with EC varying inversely with rainfall but with a one to two month lag. Our aim here is to quantify such a predictive model and compare the predicted results with reality for the full data set.

For EC the missing data occur right at the upper turning points. We therefore decided to use the available data in *Sample5* to identify the relationship between the conductivity and the rainfall by means of an appropriate multiple regression model using without first attempting to estimate the missing values.

As mentioned earlier we conjectured that electrical conductivity here is inversely proportional to rainfall, but with an appropriate lag. To test this conjecture the cross correlation coefficient for Rainfall and 1/Conductivity(1/EC) from *Sample5* was calculated, with the missing values dates removed. As there are unequal time intervals in the resulting data the cross correlation coefficients will be subject to error; however there are sufficiently many remaining observations to provide an indication of the correlation between the rainfall and 1/conductivity for different lags. These are given in Table 4. It can be seen there that there are significant correlation coefficients of 0.807 and 0.773 for the rainfall corresponding respectively to a lag of two months and one month prior to the measurement of conductivity.

Table 4. Cross Correlation: rainfall and 1/EC.

Lag	Correlation
-4	0.069
-3	0.452
-2	0.807
-1	0.773
0	0.453
1	-0.061
2	-0.381
3	-0.505
4	-0.465
5	-0.306
6	-0.023

The multiple regression model was then applied to obtain the relationship

$$1/EC(t) = 0.0682 + 0.000525\ Rain(t-1) + 0.000642\ Rain(t-2)$$

The values in the diagnostic Table 5 confirm that the rainfall accounts for approximately 82% of the variation in 1/EC and that the both *rainfall at lag one month* and *rainfall at lag two months* are significant variables. The model fit for the 1/EC values from *Sample5* is given in Figure 6. The dashed line graph depicts the fitted values while the actual values are shown by the solid line graph. In this figure the times where there are missing values have also been included.

Table 5. Diagnostic Table for the Multiple Regression Model.

Predictor	Coef.	StDev	T	P
Constant	0.0682	0.00676	10.08	0.0
Rain(t-1)	0.000525	0.00009	5.69	0.0
Rain(t-2)	0.000642	0.00009	7.02	0.0

$S = 0.0302$ $R^2 = 82.3\%$ $R^2\ (adj) = 81.5\%$

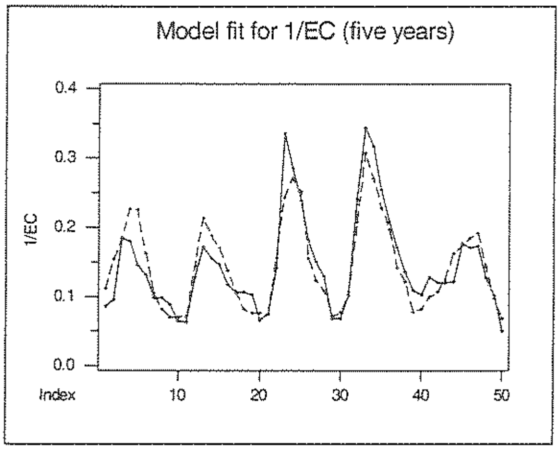

Figure 6. Model fit for 1/Conductivity (5 years).

It can be seen from Figure 6 that the model values for 1/EC correspond quite well to the actual data with the expected exception of the values at the turning points. Using the above regression model the predicted results for electrical conductivity itself from May 1998 until June 1999 were obtained.

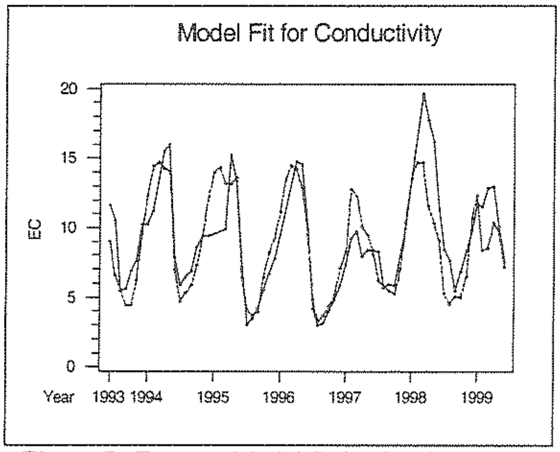

Figure 7. Forecast Model fit for Conductivity.

A comparison of predicted and actual values is given in Figure 7. Again, the dashed line graph depicts the fitted values while the actual values are shown by the solid line graph. It can be seen that the model has been successful in predicting the turning points.. The model has not reproduced the unusually high peak in early 1998 but this data anomaly could be due to the fact that no rainfall at all was recorded from December 1997 to February 1998. The relatively high rainfall following this in March 1998 may well have picked up a lot of the salt that had settled, after evaporation of the water, into the sand of the stream bed. While there is some variation in 1999 also, this can be accounted for by the impact of a heavy cyclonic rainstorm in January 1999. A large quantity of rain fell over the period of only a few hours and this caused flash flooding and the depositing of vast amounts of debris over the sampling area. This has caused a blip in the accuracy of the predicted values over the three-month period February 1999 to April 1999. However, as the effects of this subsided, the predicted values came back on track with little difference between the measured and predicted values for May1999 and June 1999.

An analysis of the residuals between actual conductivity and model values was undertaken. The graph of the residuals is given in Figure 8. It can be seen that there appears to be a nonlinear trend in the residuals, which indicates that there are other factors apart from the rainfall that influence the conductivity level. This is confirmed following an examination of the autocorrelation and partial autocorrelation of the residuals for the multiple regression model which indicate that the residuals are highly correlated at lag 1 and are therefore not a purely random process.

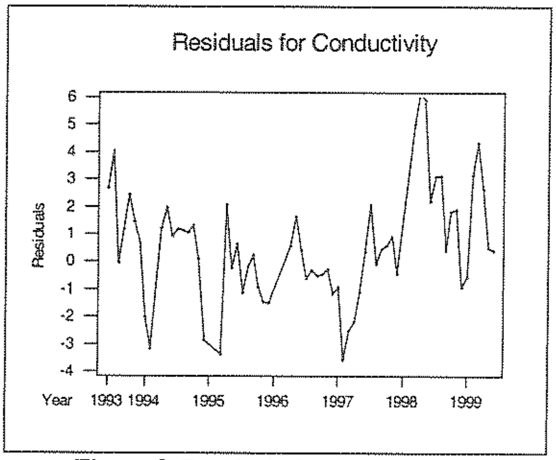

Figure 8. Residuals for Conductivity

After estimating the missing conductivity values it was possible to obtain the improved transfer function model

$$1/EC(t) = 0.0711 + 0.005\ Rain(t-1) + 0.0006\ Rain\ (t-2) + N(t)$$

where $N(t) = 0.4394\ N(t-1) + Z(t)$
and $Z(t)$ is a purely random process with mean 0 and standard deviation 0.074 .

It should be noted that the coefficients corresponding to rain at lag one and two are the same as the multiple regression model presented earlier. The transfer model accounted for 88.6% of the variation in the data and can be used to determine the forecast confidence intervals.

5. CONCLUDING DISCUSSION

The pH values fluctuate about a mean of 7.73 while remaining, with one anomalous reading, in the neutral to lower basic range.

The water salinity, as measured by the EC values, can be linked directly to the rainfall over the preceding two months. The EC values themselves, as shown in Figure 7, show a declining trend for the period from 1993 to 1997 but appear to be on the increase again after that. It may be coincidental that the timing of this increase reflects the demise of the Landcare Group involved.

Based on these results, a decision has subsequently been made by the Toodyay District Landcare Committee to carry out further monitoring at this and a number of other sites at selected parts to the Jimperding Brook. The objective of this is not only to obtain a continuing picture of the health of this steam but also to attempt to identify particular sources of its salinity problem.

6. REFERENCES

Chatfield C, *The Analysis of Time Series, Fifth Edition*, Chapman and Hall, 281 pp, London, 1996.

Deutsch, C. and A. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, 2nd. Ed.,Oxford University Press, 369 pp, New York, 1998.

Goovaerts, P. *Geostatistics for Natural Resources Evaluation*, Oxford University Press, 483 pp, New York, 1997.

Harvey, A.C., *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, Cambridge University Press, 554 pp, Cambridge, 1989.

Smith, T., *The water quality and invertebrate fauna of the Avon River: a highly saline but productive aquatic ecosystem.* Unpublished Honours thesis, Murdoch University, Perth, 1996.

1966