

## Chapter 4. Environmental Monitoring Database

### 4.1 Introduction

Regulatory agencies require that feedlot developments monitor their surrounds and the associated impacts. While this includes recording odour or dust nuisance complaints and other data applicable to social impacts (e.g. traffic flows), it is primarily focused on “environmental monitoring”. Environmental monitoring considers the measurement of chemical and physical attributes of surface and ground water and the soil within and surrounding the development. The inputs, changes of state and outputs are all measured and this information is used to develop an understanding of the nature of the system. The level of monitoring varies with:

- the size and nature of the development;
- site characteristics;
- real or perceived potential impact; and
- agency policy.

The regulatory framework that defines monitoring levels and parameters have in the past been based on regulations from overseas or the “best available information at the time”. In the latter instance, this has meant that tests considered to be useful based on past experience with, for example, sewage treatment plants, have been applied to intensive animal industries. Generally when this has occurred, the number and breadth of parameters requested as been widened so that maximum information is obtained. This has been done on the pretext that because environmental impact has not been previously defined through research, and the breadth of impact is unknown, many parameters should be tested to reduce the possibility that degradation of a specific part of the environment is not overlooked. While this regulatory approach is simple and perhaps easier for industry to comply with, it often incurs unnecessary tests and costs for industry.

While regulators for many industries now require environmental monitoring, there is little information available on the proper sample collection techniques and their subsequent handling, treatment and testing. There is almost no information available on environmental monitoring results and few outcomes that feed back information for the improvement of monitoring programs. The cattle feedlot operator has to make a range of decisions that affect the entire operation, from purchasing animals, feedlot rations, weather considerations, selling decisions and animal health and welfare issues along with the planting, tillage, fertiliser decisions that go along with an agronomic production system (Eigenberg *et al.* 1995). It is difficult for the feedlot operator to gain access to the information that is relevant and important to the decision making process (Eigenberg *et al.* 1995), especially environmental monitoring data. Therefore, the development of an Environmental Monitoring Database (EMD) will provide a tool to assist the decision making process required for the operation of a sustainable feedlot, or indeed any industry that is required to monitor its environment.

As part of the research for this study, an EMD was developed to store, manage and interrogate environmental data collected by the Cooperative Research Centre (CRC) for the Beef and Cattle Industry. Most data have been collected from the CRC's "Tullimba" cattle research facility and feedlot. They include information on soils, surface waters, ground waters, manure, effluent and crop matter associated with the development and operation of the feedlot. An important component of the research was the use of sample collection and handling techniques that did not compromise data integrity. These techniques are outlined in Appendix A.

## **4.2 Development of the Environmental Monitoring Database (EMD)**

Several key issues need to be addressed in database construction. Some of these were outlined by Gaunt *et al.* (1997), and include a recognition of the different spatial and temporal scales of the collected data, a mechanism for linking quantitative and qualitative data, understanding the variability and reliability of the collected data, and the ability to deal with missing data. To enable multidisciplinary use of data collected from an environmental system, it is important to coordinate the linkages of datasets collected from each of the sub-systems (Duijnhouwer & Dekkers 1997; Gaunt *et al.* 1997).

### **4.2.1 Purpose and Objectives of the Environmental Monitoring Database**

The EMD stores information collected for various projects conducted at the University of New England's cattle feedlot research facility ("Tullimba"). These projects have been undertaken to define the nutrient and salt cycles in feedlots and to develop better monitoring programs. Some data have also been collected from various commercial feedlots. While the EMD stores this information, its purpose is to allow the data from across the feedlot system to be linked and then investigated.

A secure and efficient data storage system is identified as the primary objective of the database. The next objective is to provide an interface that is capable of manipulating the datasets. The final objective is to develop "indices" and track their changes through time to provide a tool for the sustainable management of the dynamic system.

### **4.2.2 Structure of the Environmental Monitoring Database**

A systems approach was taken in the development of the EMD, with Microsoft Access used as the platform. Each sub-system was divided up into inputs, outputs, and external and internal system functions. Each sample type was identified as being one or more of these functions. For example, manure would be an output when considered as coming from the feedlot system, but would be an input when considered in the context of being incorporated into the soil for the purpose of adding nutrients for a crop.

The shell of the database was built based on the parameters that were identified as being important for each sample type. These parameters were recorded in separate analysis tables for each sample type. The analysis tables were:

- cattle faeces;

- clean water;
- crop;
- ground water;
- liquid effluent;
- manure;
- rain;
- runoff;
- soil; and
- spills.

This design proved to be extremely inefficient due to the large number of blank entries in the datasets when baseline data was input into the corresponding tables. This was due to previous analyses not necessarily corresponding to the analyses that would “ideally” be carried out in the future. To be efficient, Microsoft Access requires as few blank entries as possible.

The database developed for this research contains data from a number of different feedlots, in addition to the Tullimba feedlot and to accommodate this, the structure was changed by dividing the analyses tables up into the different geographic areas. For example, the soil analyses table was divided up into Tullimba soil analyses - archived and Tullimba soil analyses etc. The new tables were based on the analyses that were to be carried out each time a sample of each type was taken. The archived tables contained baseline data and became static in that no more data would be entered into them.

All the data was entered into this version of the database. Excel was used to query the database and an interface was built to display the time series of the selected parameter in the soil and/or ground water in the irrigation area. Ortho-P in ground water, sampled from the piezometers in and below the irrigation area, is used as an example of the interface in Figure 4.1. In Figure 4.1 the letter P followed by an identification number identifies a piezometer, for example, P5, P6 etc.

This database structure proved to be inadequate when creating relationships between the tables, as there were no primary keys used. A primary key is a field, or several fields, whose value or values uniquely identify each record in a table. In a relationship, a primary key is used to refer to specific records in one table from another table (Microsoft Access 1997).

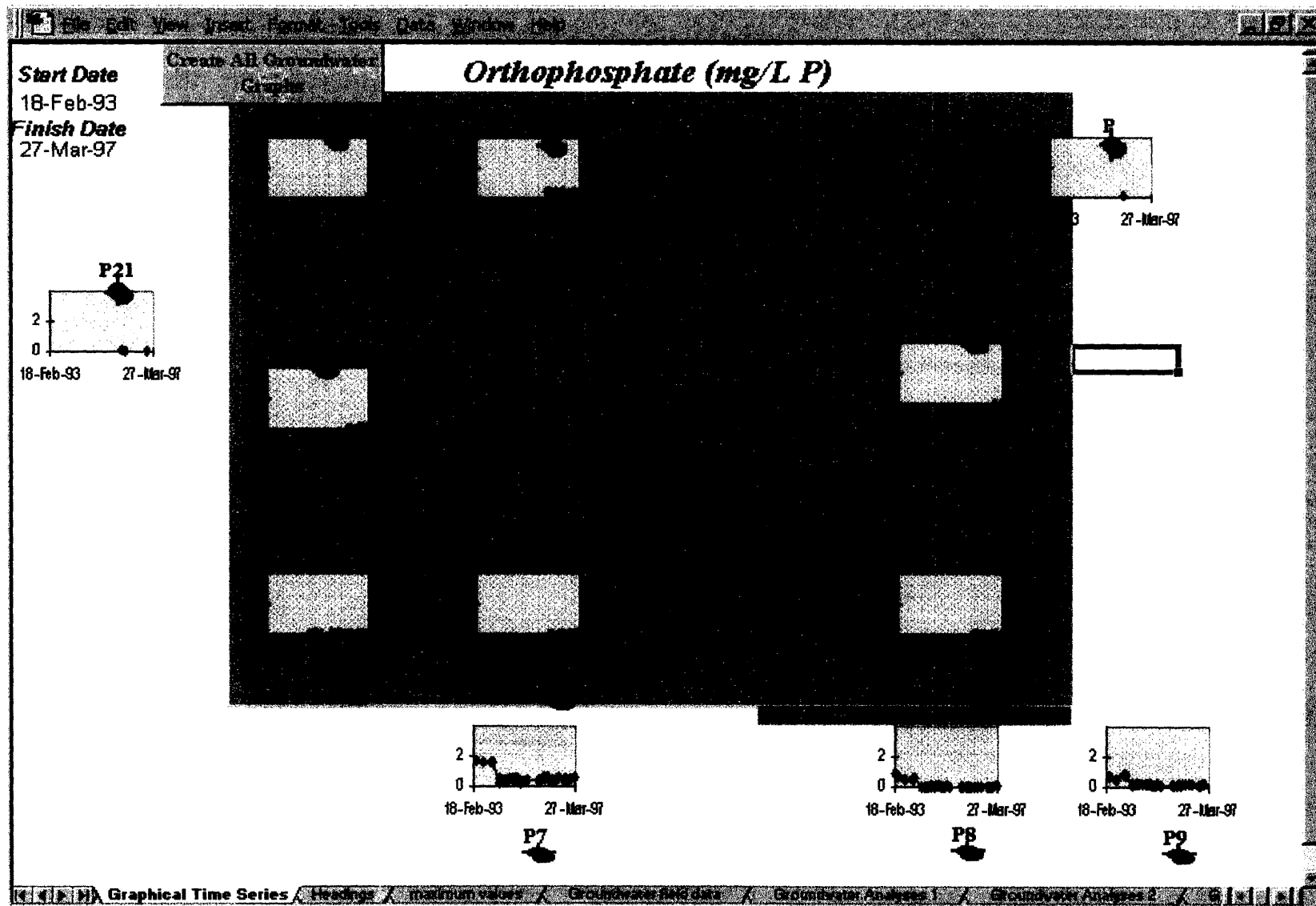


Figure 4.1. Excel Interface for Tullimba Ground Water

Using this database structure also proved difficult when incorporating data sets that measured different parameters to those of previous analyses. When the database tables were expanded to include these data, more blank entries were created. These problems were redressed through the complete restructuring of the database.

The database restructure included two central tables - a "Record Table" and an "Analyses Files Imported into Database" (AFID) table, with each of these tables having a primary key. The "Record Table" assigns each new sample a database record identification (ID) and records the database file number (which links it to the AFID table). It also records a sample number (assigned if the sample is sent to an external lab), the sample date, sample type, location, sampling point and laboratory ID. There are still some "holes" in this structure as not all the records have a sample number or laboratory ID. However, every other field should have an entry.

The analyses tables were then split into each parameter with the only two entries required for each record, being the database record ID (which links it to the "Record Table") and the value for the parameter. The heading for this value column in the table is exactly the same as the table name. The rationale behind this was to simplify the code when building the graphical interface.

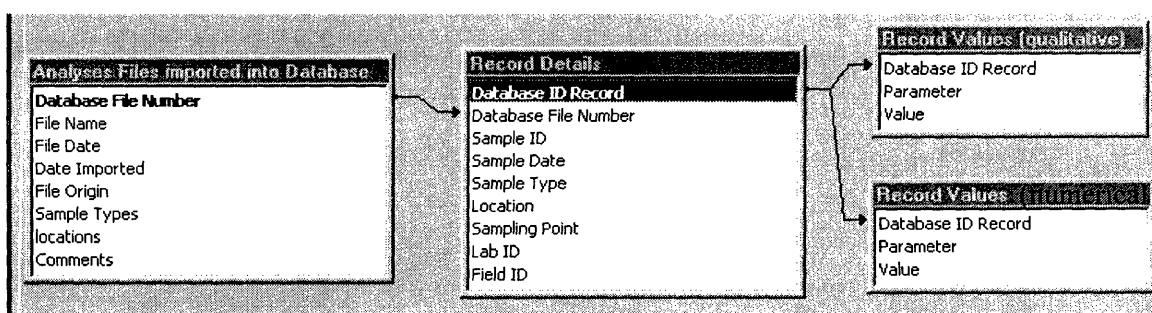
This structure was used for building queries and forms to provide graphical representation of the relationships that exist in the data by plotting data from different horizons in the soil. Data for a particular parameter from two horizons are plotted on a x-y scatter graph and a linear regression is applied to highlight any relationships that may exist. Even though regression implies dependency, its use here is simply as a visual aid to the graphing tool to indicate when further investigation may be warranted. Relationships in plant and soil data can also be investigated by plotting the concentration of a parameter in the soil against the concentration of the same parameter, or different parameter in the plant. A similar comparison of ground water data is also possible. Examples of these results will be presented in Section 4.5.5.

Continued use of the database indicated that this structure could be further improved as the database contained an excessive number of tables, which does not conform to "good" database design practices. In the new structure all the parameter tables have been "lumped" into one table, named "Record Values". This table has three columns – Database ID record, Parameter and value. The original "Record Table" is retained with the addition of only one column – a Field ID. This addition will be elaborated on in Section 4.4.1. The qualitative record values were entered into a separate table to the numerical record values, allowing the numerical record values to be manipulated mathematically via queries.

### 4.3 Querying the Database

A database “query” allows the linking of data in a variety of ways. For this to be successful, it is important that the relationships are defined properly from the outset. The latest database structure has this consideration incorporated within the design.

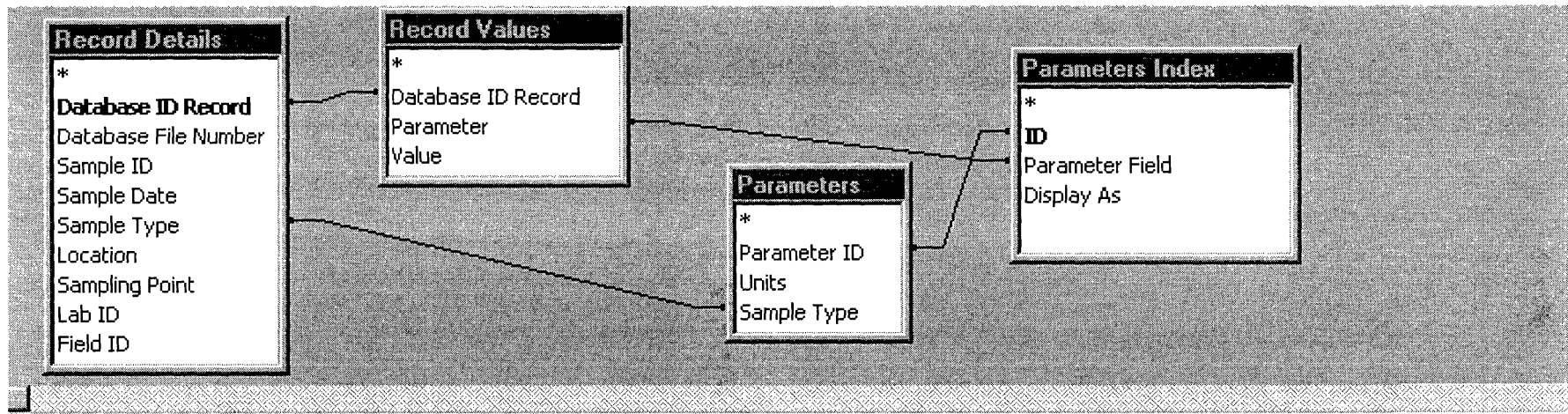
Figure 4.2 shows the relationships that are central to the database. The bold text indicates a primary key and the lines with an arrow at one end indicate a “one to many” relationship. For example, there is a one to many relationship from the “**Database File Number**” field in the “Analyses Files imported into Database” table to the “**Database File Number**” field in the “Record Detail” table. There are many other relationships defined in queries that have been built in the database and a proficient user of Access can build others quickly and easily.



**Figure 4.2. Relationships Central to the Database**

Crosstab queries are used to present the data in a spreadsheet format. “A crosstab query displays summarised values (sums, counts, and averages) from one field in a table and groups them by one set of facts listed down the left side of the datasheet and another set of facts listed across the top of the datasheet” (Microsoft Access 1997). These crosstab queries are constructed for each sample type or sub-system and are designed for use in a mass balance. For example, the animal system query includes analyses for ration, fresh faeces, pen manure and stock water and is designed for use in an animal system mass balance. To date, there are thirteen sample type/sub-system crosstab queries that describe the animal, water, ground water, effluent and crop systems.

The basic structure for all the above queries is the same with the criteria being changed to reflect the required sample type/s. Figure 4.3 is an example of this structure and shows the design view of the “Analyses - Fresh Faeces” query. The figure demonstrates the links between the “Database ID Record” fields in the “Record Details” and “Record Values” tables and the links between the “Sample Type” fields in the “Record Details” and the “Parameters” table. There is an additional link from the “Parameter ID” in the “Parameters” table and the “ID” field in the “Parameters Index” table. This is to enable the units to be returned in the “Parameter Field”.



Field:	Database ID Record	Sample Date	Parameter Field:	Sampling Point	Location	Sample Type	Value	Sample Type
Table:	Record Details	Record Details		Record Details	Record Details	Record Detail	Record Values	Record Details
Total:	Group By	Group By	Group By	Group By	Group By	Group By	Avg	Where
Crosstab:	Row Heading	Row Heading	Column Heading	Row Heading			Value	
Sort:	Ascending	Ascending						
Criteria:								"Fresh Faeces"
or	Parameter Field: [Parameters Index].[Parameter Field] & "(" & [Parameters].[Units] & ")"							

Figure 4.3. Design View of the "Analyses - Fresh Faeces" Query

The dimensional units for the record are not recorded when input into the database. This information is stored only once in the “Parameters” table. The units returned in the query are based on the sample type and parameter. The “Parameters Index” table stores a unique ID, parameter name and a “Display As” value for the 122 parameters that are currently in the database. These parameters and the units can vary for each sample type so the “Parameters” table stores the units for each parameter and sample type.

The above analyses queries are accessed through the “Sample Time Line” form. This form is a graphical display of all the samples collected from the “Tullimba” feedlot to date and is shown in Figure 4.4. Also included is the monthly rainfall data for the feedlot. This form is the point of entry for accessing data and the importance of the hydrological cycle on the dynamics of the system is highlighted by the inclusion of the monthly rainfall.

Field data are collected for most sample types and included in various tables such as “Crop Field Data”, “Surface Water Field Data” etc. Each of these tables has a “partner” table named “<sample type> Sample Dates”, for example, “Surface Water Sample Dates”. These are included in a union query that is utilised by the above “Sample Time Line” form. Union queries combine corresponding fields from two or more tables or queries into one field (Microsoft Access 1997). The number of tables that are required for particular field data is dependent on the sample type. For example, for the surface and ground water field data there are four tables: “Field Data”, “Field Values”, “Parameter ID”, and “Sample Dates”. These are used to display the data in a crosstab query, similar to the analyses crosstab queries described earlier.

Field data for the animal system mentioned previously are stored in the “Animal System Records” and “Animal System Sampling Dates” tables. Similarly, the samples collected at the same time as these from the holding ponds and terminal pond are stored in the “Water Cycle Records” and “Water Cycle Sampling Dates” tables. This structure makes it easier to collate the data for a mass balance. The queries that bring these tables together are named “Field Data – <sample type>”.

There are a few “static” tables that store data on piezometers, such as depth of the piezometer, coordinates etc., and data on the location and specifications of the ponds and dams. The other important tables are those that store the information concerning samples sent to external laboratories.



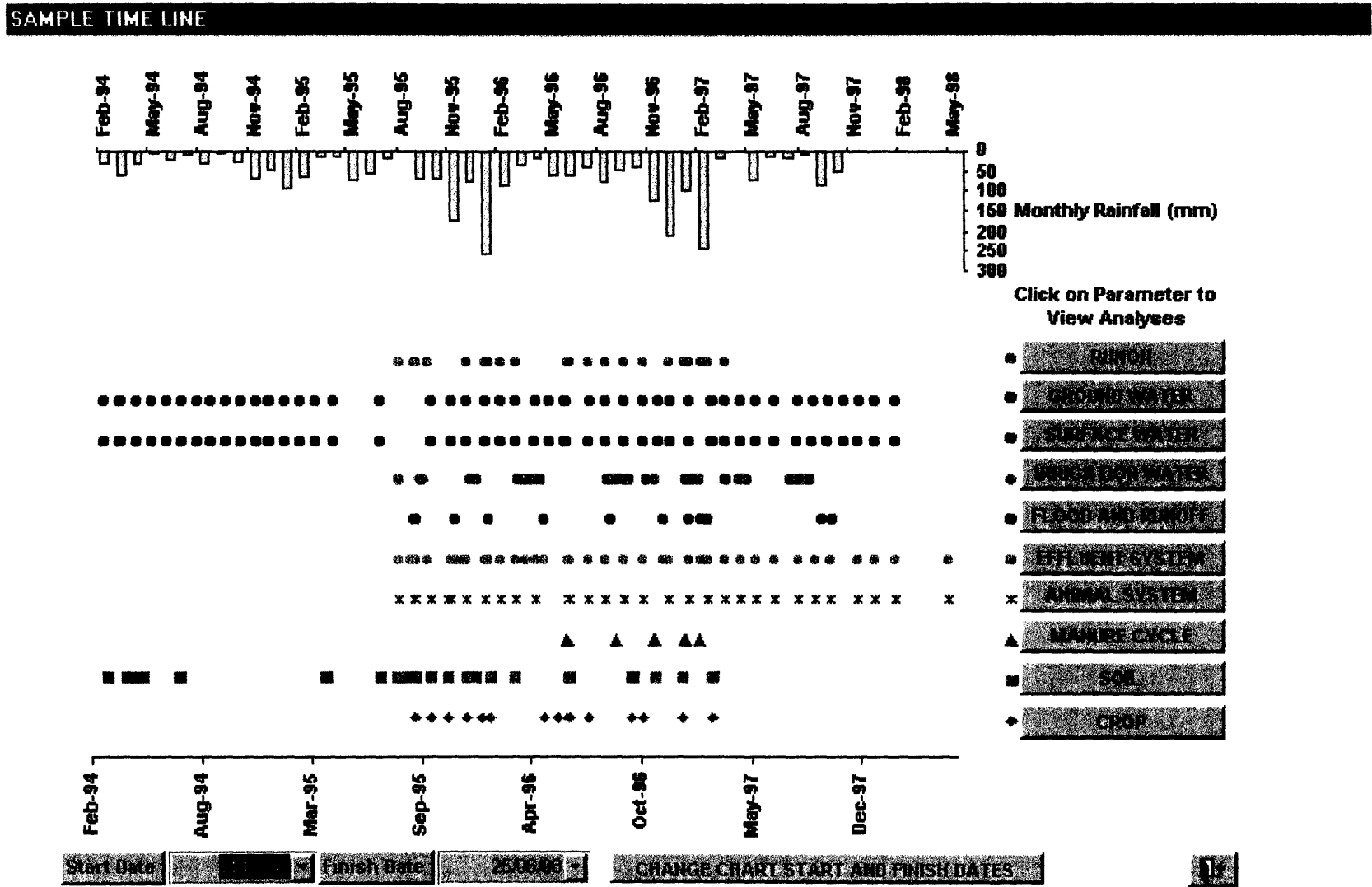


Figure 4.4. Sample Time Line Form

## **4.4 Data Management**

### **4.4.1 Keeping Track of the Sample**

Presently 2479 records (samples) are stored within the database. Each record is linked to a number of parameters, depending on the analysis carried out on the sample. It is important to have a system that can track samples because various people carry out sample collection and there have been a large number of samples collected over the past 5 years.

Several samples are taken on a regular basis – ground water, surface water, effluent, ration and manure (fresh faeces and pen manure). These samples are taken at the same time and put into freezers for storage. Each sample has data recorded, such as depth of pond or pen number and any other comments that are applicable. The date, sample point and location are stored in the “Water Cycle Records”, “Water Cycle Sample Dates”, “Animal System Records” and “Animal System Sample Dates” tables, described earlier.

When sampling ground water, the water level, temperature, pH and electrical conductivity are measured on site and are entered into the database under the “Ground Water Field Data”, “Ground Water Field Values” and “Ground Water Sample Dates” tables. The surface water field tables are similar, however water levels are not recorded and dissolved oxygen is an additional parameter measured in the field. When the samples are later analysed these data are entered into the record tables and matched to the field data by a field ID.

Other samples are taken at irregular intervals - eg. spills from the holding or terminal ponds, runoff above the feedlot catchment (runon), soil samples prior to sowing, crop samples during crop growth and at harvest. The information pertaining to these samples are recorded in the field and they are stored in various tables. When the analyses are returned the records are matched (via queries) to field data.

## **4.5 Exploring the Datasets**

A valid interpretation of the data is only possible through consideration and an understanding of the systems inputs and outputs and management history (USDA NRCS 1996). Feedback of the outputs to the inputs can affect management decisions and to enhance the likelihood of making the optimum decision, data should be used to provide the maximum feedback (Sadeh & Griffin 1997). This feedback can be provided by graphical representation of data.

Investigating the data from the perspective of considering the utilisation area as a system with outputs that supply feedback to enable the manipulation of the inputs will enhance the likelihood of a sustainable system. This systems approach will not only be beneficial from an environmental point of view but also for improving production efficiencies. Having an understanding of the history of nutrient content will also provide a basis for putting any individual sample in context (Shapiro 1996), and highlight the cyclic processes the nutrient is subject to. The following sections present examples of the output from the queries built in the EMD. Some of the graphs were created after transferring the data across to Excel.

#### 4.5.1 Soils

Figure 4.5 is a time series of Ortho-P in the soils receiving effluent and manure from the three feedlots. Feedlots B & D are situated in South East Queensland and their soil types are heavy montmorillonite clay. Feedlot B has had effluent and manure applied since 1975 and Feedlot D since 1982. Feedlot I is the “Tullimba” research feedlot. The irrigation area of this feedlot is on a duplex soil with a sandy loam A horizon and sodic B horizon. Note the difference in the scales on all the graphs. The levels of P in the A1 horizon of feedlots B and D have increased dramatically from the application of manure and effluent, however it is apparent that the levels have now decreased. The lower horizons do not follow the same pattern, which suggests that there has been minimal movement downwards of this nutrient. However, in Feedlot I, the trend of the increase in the B horizon suggests that there is some translocation of P downwards.

Figure 4.6 is a graph of Colwell P versus organic C in the Tullimba Soil. The first graph includes the data for all horizons and the scatter plot suggests that there is some relationship between the two elements. The next two graphs suggest that the relationship is stronger in the A horizon than it is in the B horizon. When all the Tullimba data was included with the clay soil data from a feedlot in South East Queensland, it was found that a third order polynomial fitted the relationship between Colwell P and organic C ( $r^2=0.90$ ) (see Lott *et al.* 1997b). The organic C content of the Tullimba soil ranges from 0 to 3 %, with a corresponding rise in Colwell P from 1 to 40 mg/kg. The linear portion of the third order polynomial fitted to all the data from both the feedlots ( $-3.97x^3+56.3x^2-85.3x+32.1$ ), starts at 3% organic carbon and around 100 mg/kg Colwell P. The slope of the linear portion of the third order polynomial is one order of magnitude steeper than the slope of the Tullimba A1 horizon data. Thus suggesting there are in fact two distinct linear relationships that are dependent on soil type.

#### 4.5.2 Ground Waters

Figure 4.7 shows two x-y scatter graphs for the ground water at Tullimba. The first is the relationship between total N and nitrate + nitrite ions. Even though the data set is not large, the graph suggests there is a relationship between Total N and mineral N in the ground water. If further results support this trend the less expensive chemical analysis for mineral N may provide equivalent data for total N, at least for the Tullimba feedlot. Total P versus ortho-P in ground water is the second graph and shows that at present there is insufficient data to define a relationship between these two elements.

#### 4.5.3 Surface Waters

Figure 4.8 shows the changes in EC through time at three sampling points on Headstation creek at Tullimba, with rainfall data also included. There is a drop in EC for all three points following substantial rainfall from Oct 95 to Jan 96. This indicates that the EC of the runoff entering the creek is less than that in the creek and therefore there is a dilution effect from the rainfall. These data are used in Section 4.6 to investigate any changes in Headstation creek and the ground water as a result of the feedlot operations.

#### 4.5.4 Effluent

Figure 4.9 presents time series graphs for the cations in the two holding ponds at Tullimba. These graphs show some of the variability in the irrigation water over time, with K showing the greatest variability. Rainfall and subsequent runoff are likely to be the biggest factor affecting the variability of the concentrations of cations. The greater variation in K indicates that this cation is either more variable in the feed ration or more volatile in nature than the other cations included in the graph.

#### 4.5.5 Interrelationships

The interface developed to extract data relationships is used to produce the graphs shown in Figure 4.10. The graphs show the relationship between the average Colwell P in the soil versus the average of the Total P extracted by the crop, for all horizons and for each separate horizon. The final graph is the time series for both the parameters. The interface automatically applies a linear regression to the data. As stated previously, this is merely to highlight areas for further investigation and is not purported to stand up to any rigorous statistical analyses.

Figure 4.11 shows the variability of nitrate in the ground water of the irrigation area compared with the relative static values in the piezometers beneath the irrigation area. The likely explanation for this occurrence is effluent in excess of plant requirements moving to the top of the B horizon, which is a relatively impermeable clay, and then possibly flowing laterally down the slope and being intercepted by the terminal pond. If this is occurring the B horizon is in effect acting as a tile drain and thus nutrients are retained within the system.

Visual representation of data is a valuable tool in providing an understanding of the system. Indeed, this understanding is necessary before the statistical analysis phase is entered. For example, the time series data of the average P in the various soil horizons at three feedlots (Figure 4.5) highlights the variation of the P contents in the different soil types. The figure also shows that although the levels at feedlot B reached very high concentrations, these levels have now been reduced to near the original levels. At feedlot I, some P has obviously moved through the sandy A horizon into the B horizon, a characteristic that is not evident in the soils of the other two feedlots, these soils being much heavier clays. Figure 4.6 indicates that the organic P build up in the soil is related to the increase of organic carbon.

The interface for comparing concentrations of a particular parameter in the soil with the same, or different parameter in the plant is useful for extracting relationships which may lead to the reduction of the number of parameters tested in the soil. At this point of time, data is limited but if more data confirms the relationship between Colwell P in the soil and total P in the plant this indicates less soil testing for Colwell P may be required. That is, the concentration of P in the plant can be used as an indicator for Colwell P in a duplex soil, with a sodic B horizon receiving feedlot manure and effluent. Olsen and Barber (1977) found that P solubility and plant availability increased with increasing organic matter. This could account for the observation of increasing P uptake with increased organic matter.

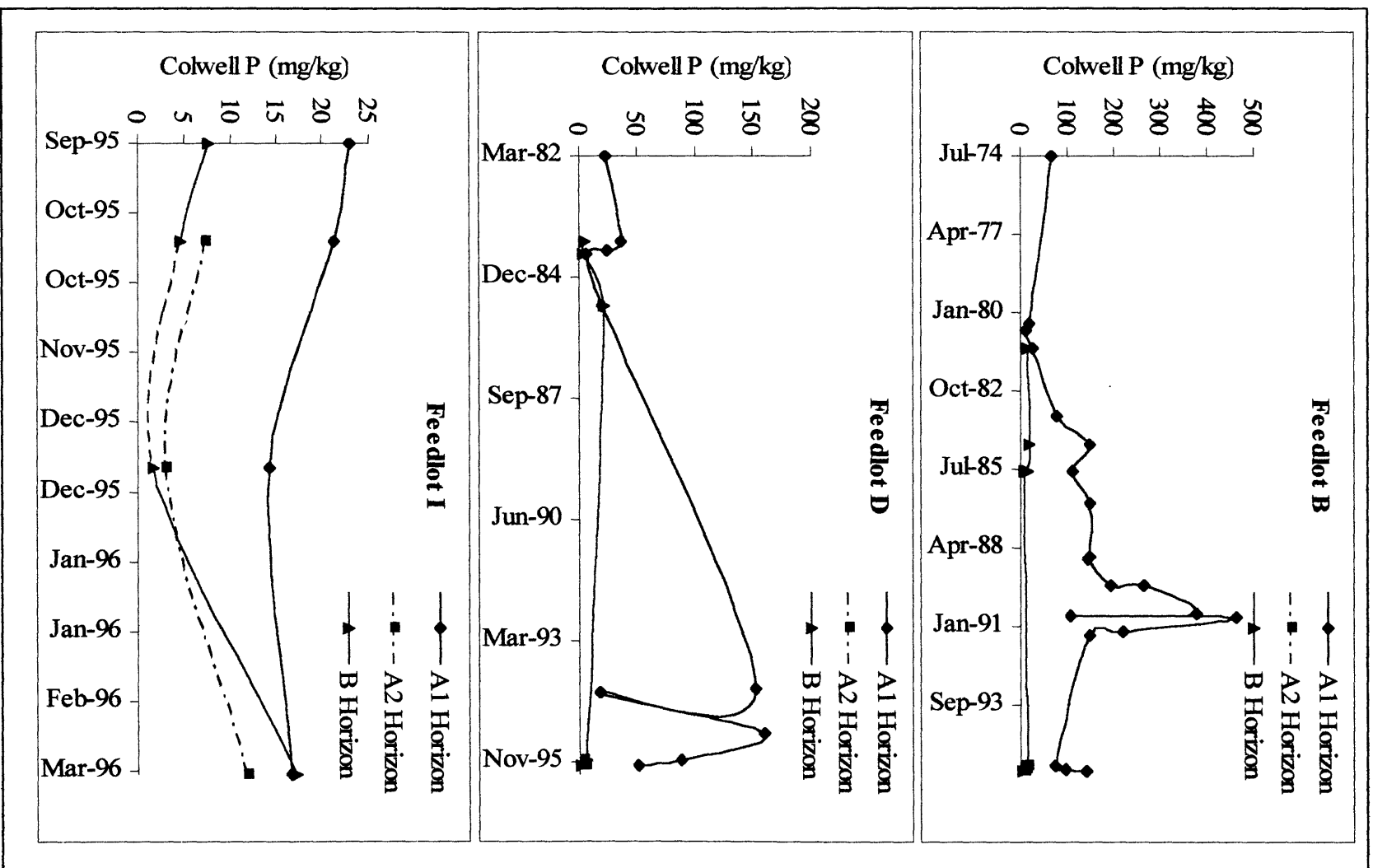
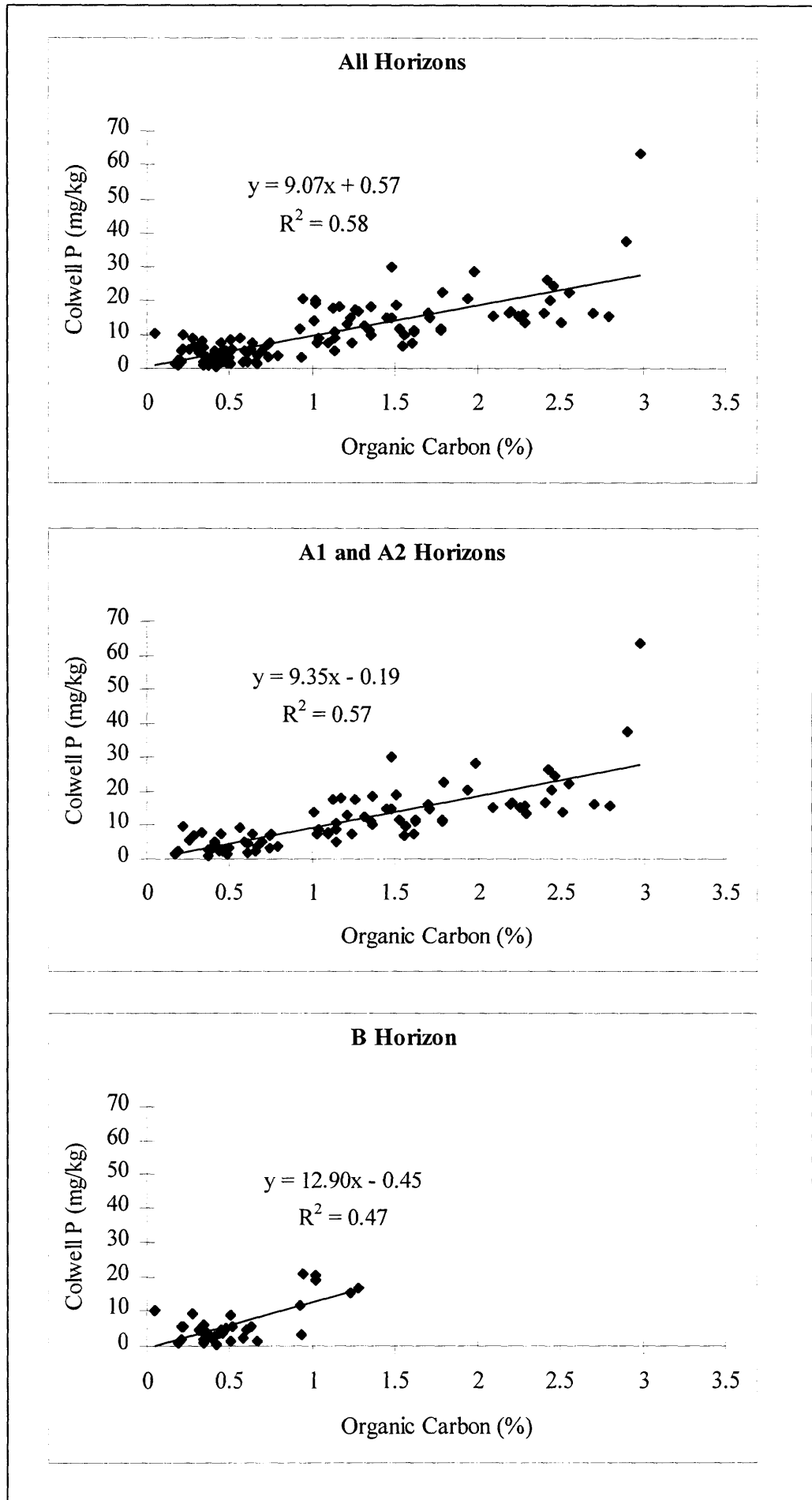
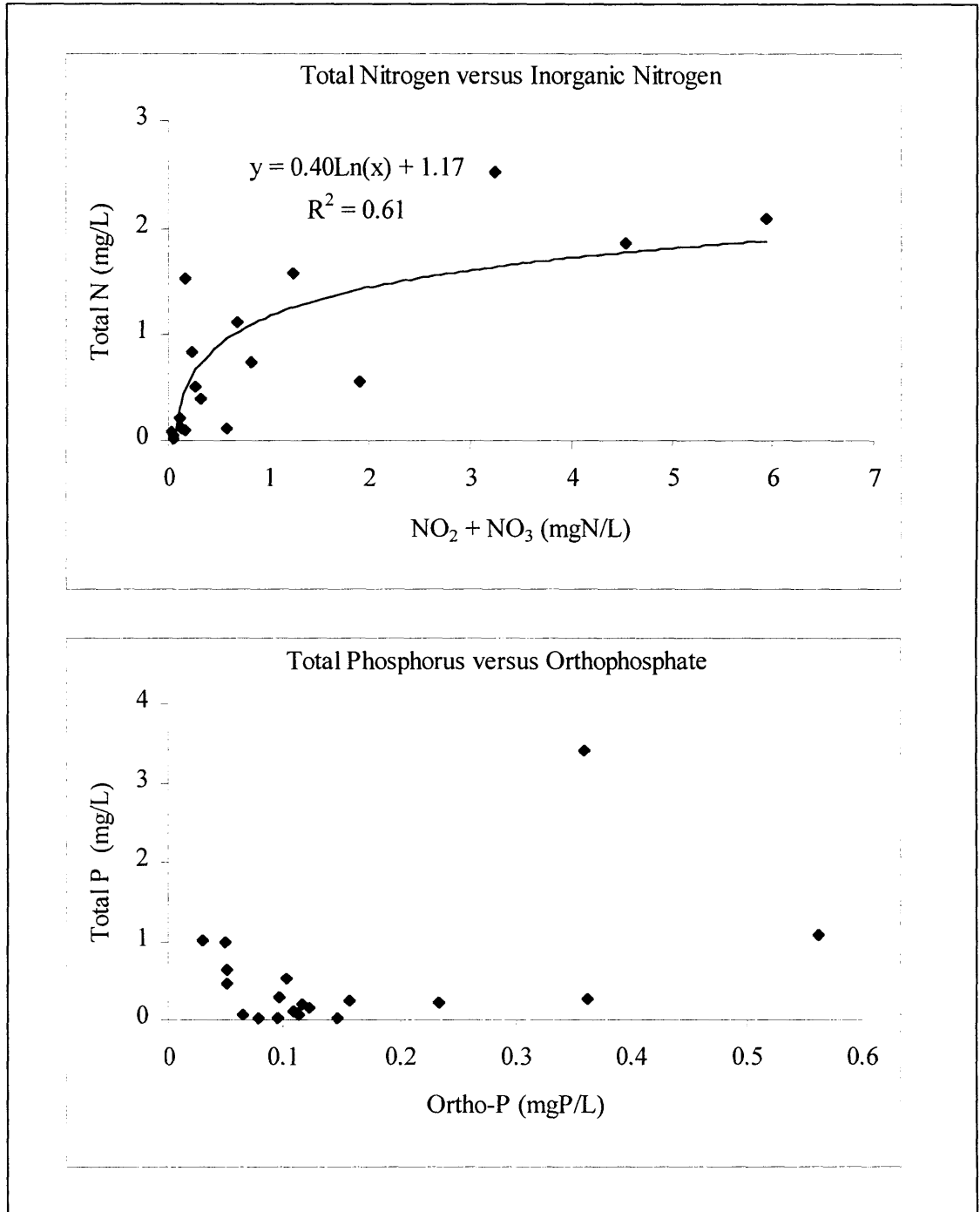


Figure 4.5. Soil Colwell P Across all Feedlots



**Figure 4.6. Colwell P versus Organic Carbon in Various Soil Horizons at Feedlot I (Tullimba)**



**Figure 4.7. Nitrogen and Phosphorus Relationships in Ground Water at Feedlot I (Tullimba)**

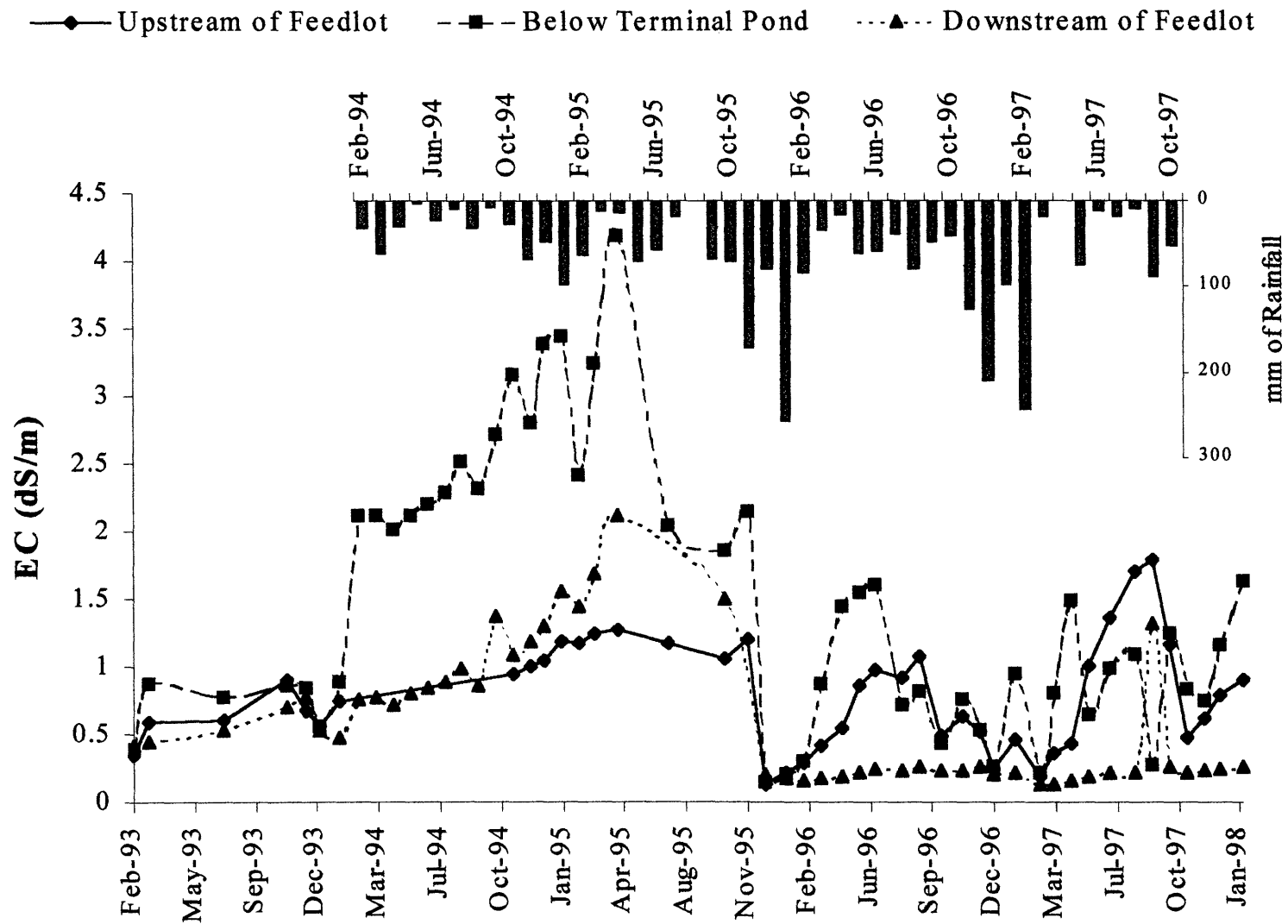


Figure 4.8. Changes in EC Through Time in Feedlot I (Tullimba) Surface Water at Headstation Creek with Rainfall Data



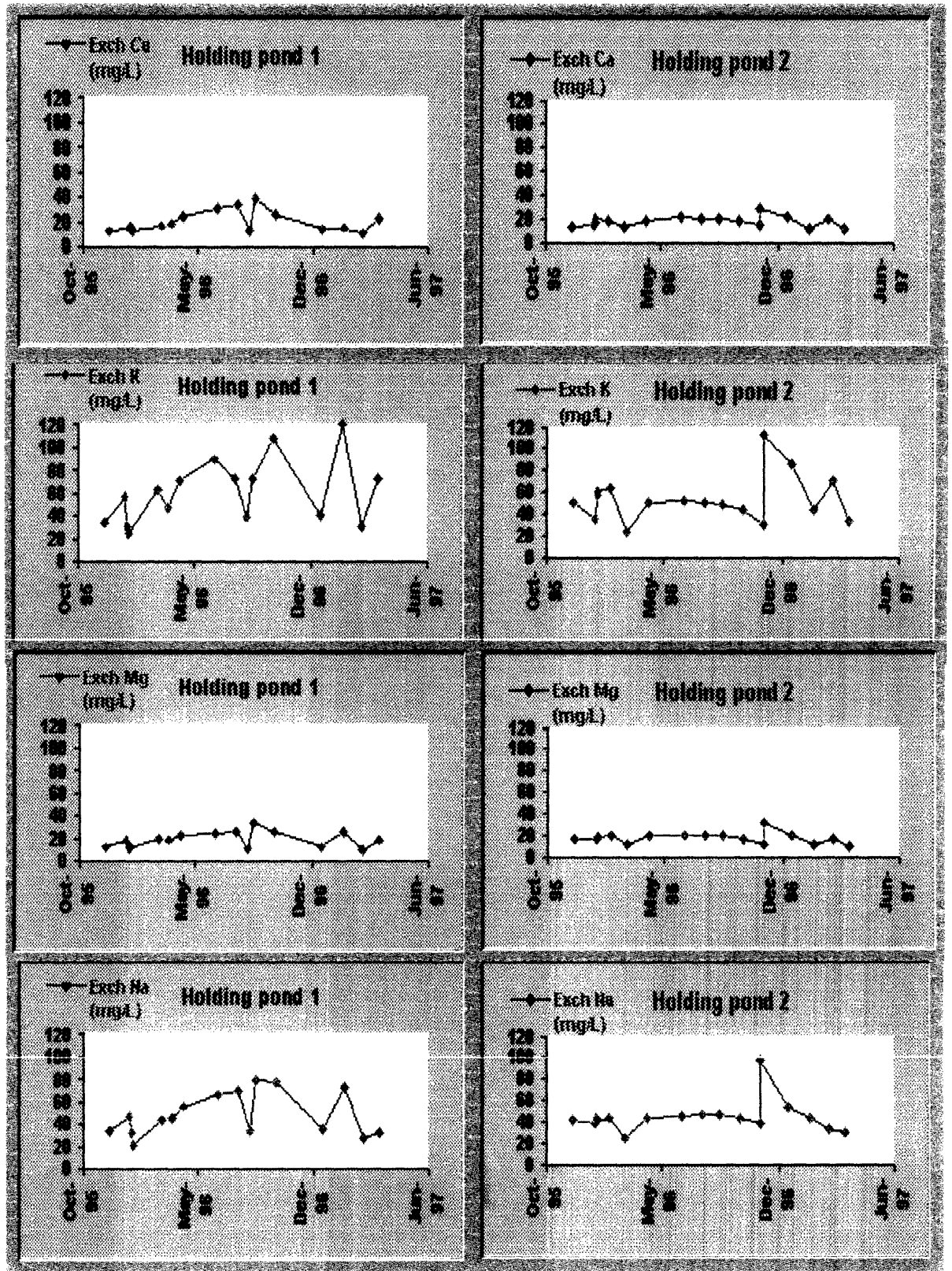


Figure 4.9. Changes in the Chemical Characteristics of Irrigation Water at Feedlot I (Tullimba)

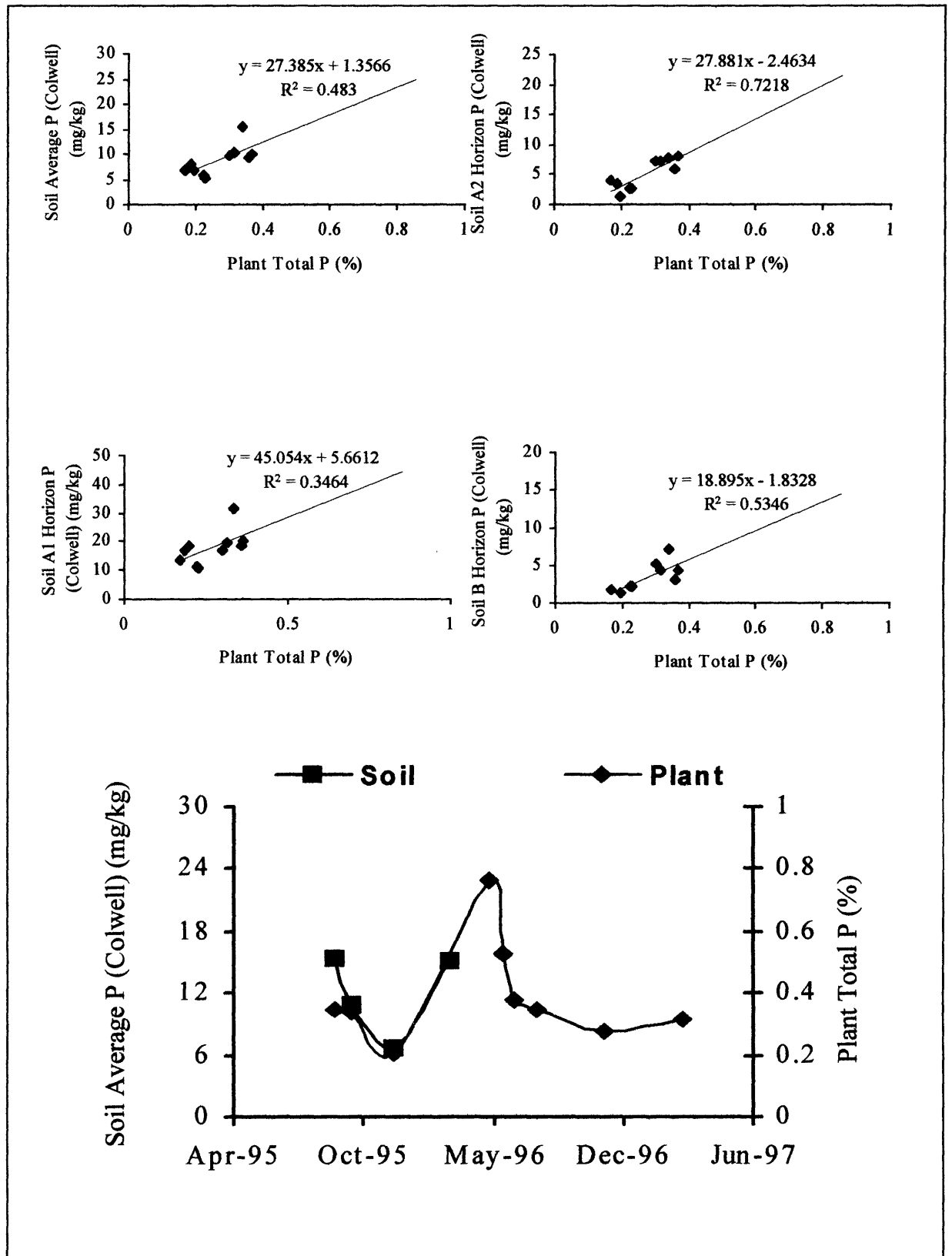


Figure 4.10. Average Colwell P in Soil versus Total P in the Plant at Feedlot I (Tullimba)

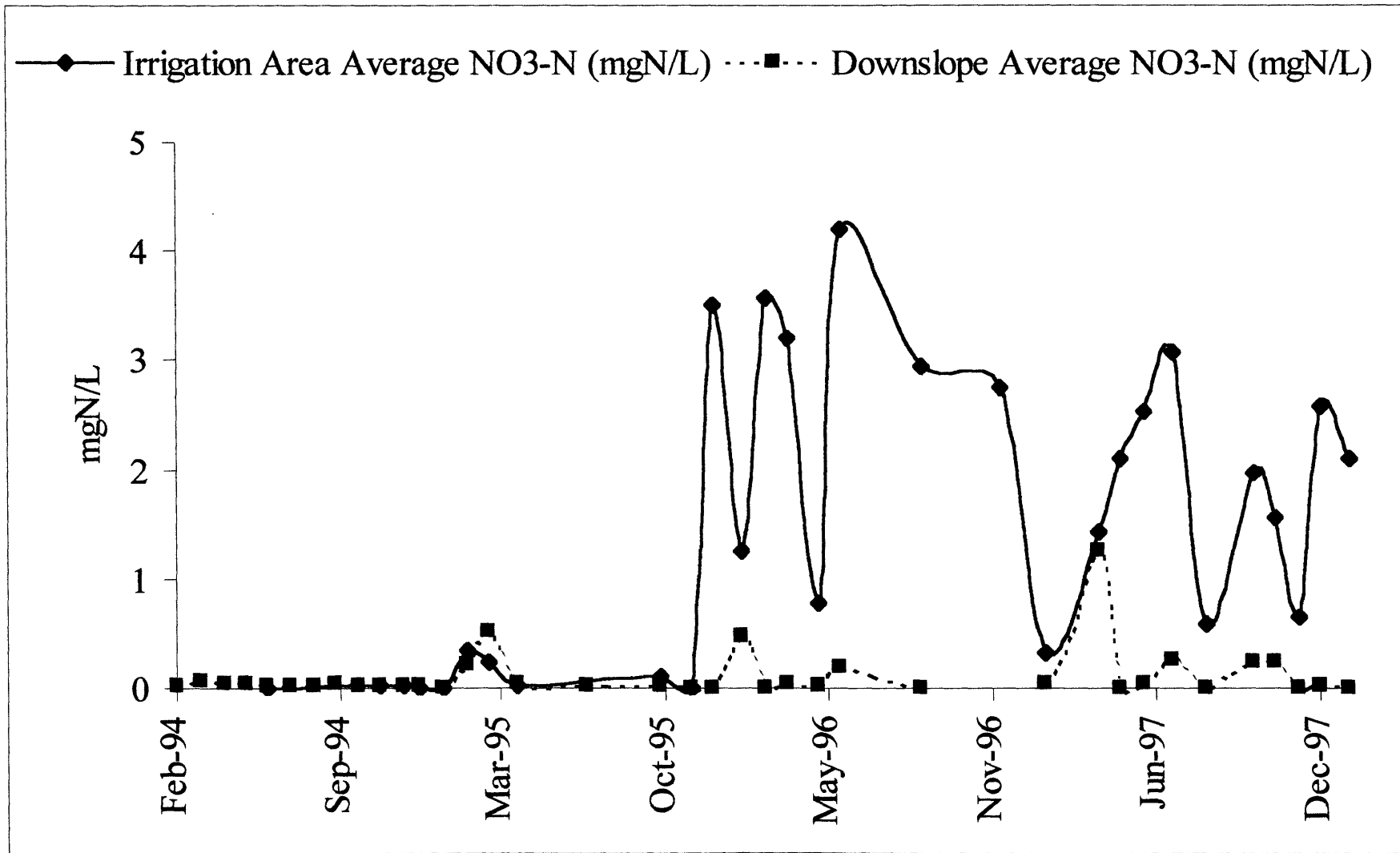


Figure 4.11. Variability of Nitrate in Irrigation Area Ground Water Compared with a Site Below the Irrigation Area at Feedlot I (Tullimba)

## 4.6 Applying Preliminary Statistical Models

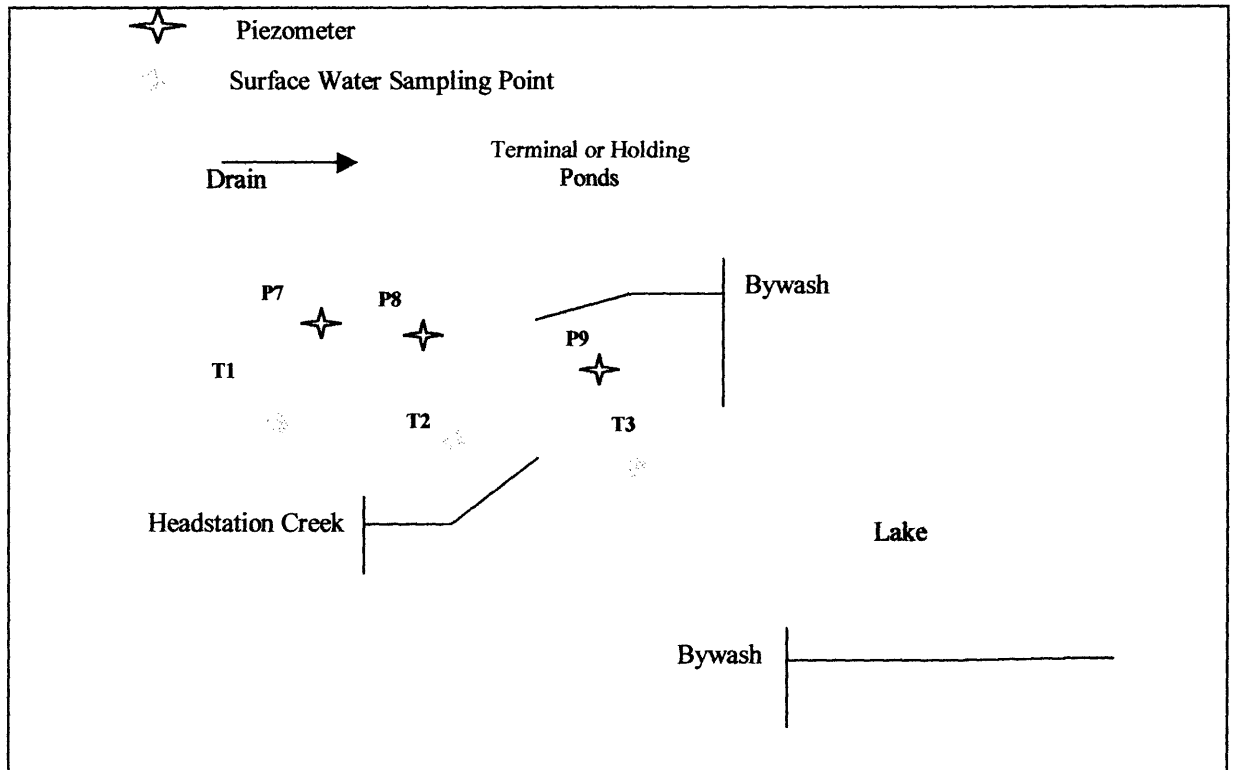
Results presented in the preceding sections highlight the complexities of dealing with environmental monitoring data. Time series graphs of data extracted in the queries give a broad view of the changes occurring in each parameter. For example, the time series of the electrical conductivity (EC) data in the creek running through the feedlot (Figure 4.8), indicates that there has been a shift in the moving average of this parameter upstream, below and downstream of the feedlot since the irrigation area has been operational. Superimposing the monthly rain data suggests a strong relationship between rainfall and EC in the creek and further investigation and time series statistical analyses should provide more evidence.

“S-plus” was used to investigate the characteristics of the EC datasets from three sampling points in Headstation Creek, which flows below the research feedlot. The three sampling points, T1, T2 and T3 are shown in Figure 4.12, along with the relative positions of three shallow piezometers (P7, P8 and P9). The relative position of the holding and terminal pond system is also indicated in the schematic.

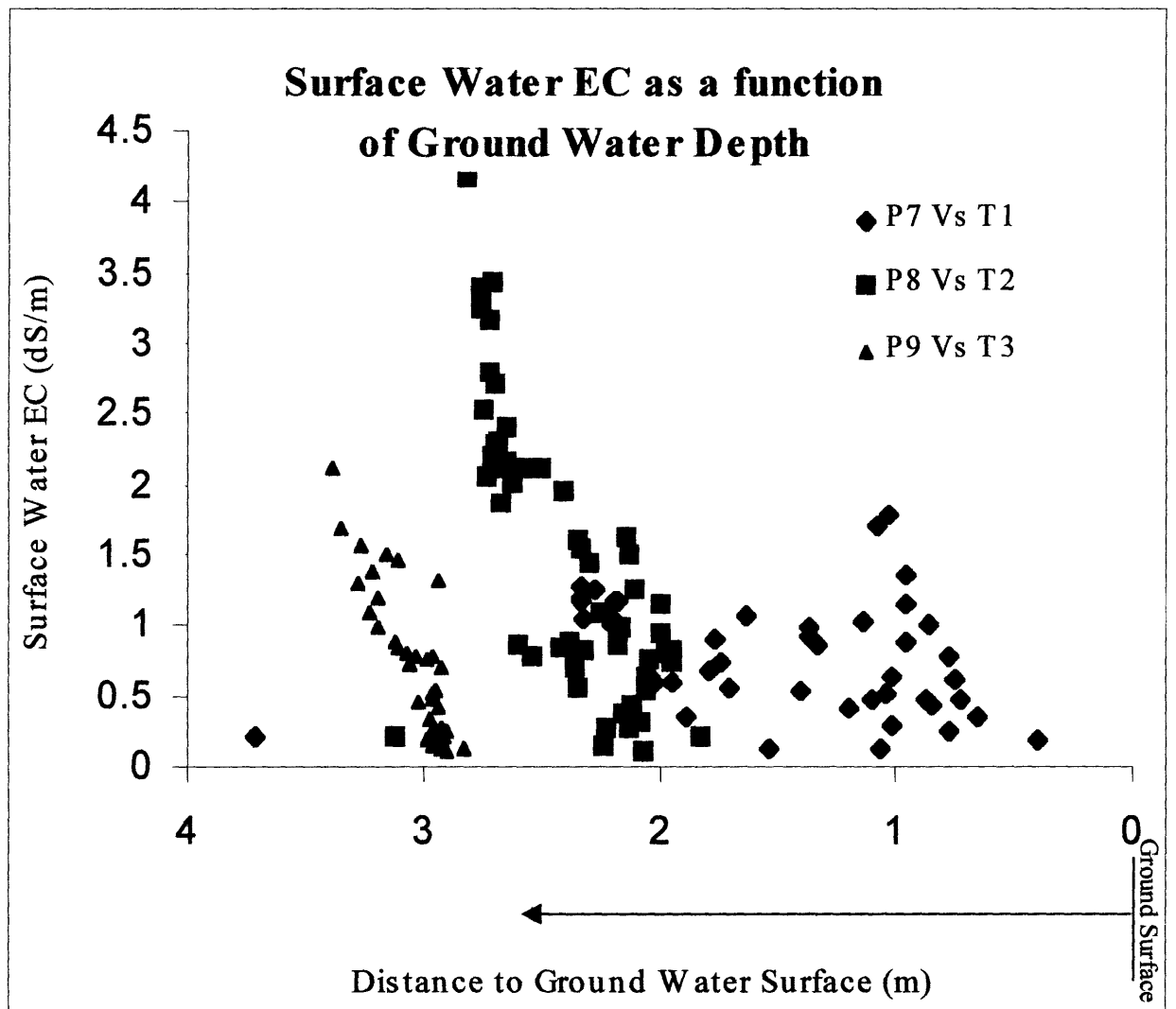
There is considerable difference in the time series characteristics of the three sampling points, as shown in Figure 4.8. T3 is adjacent to the lake, which has a maximum volume of 1200 ML, and the EC measurements taken over time have been more constant at this point. The other two sampling points are at locations where the creek is non-perennial and therefore these sampling points become isolated pools of water.

The database was used to investigate relationships between the EC of the water at the sampling points and the standing water level (SWL) in the adjacent piezometer (see Figure 4.12). The SWL in the adjacent piezometer is used, as there is no record for creek flow. The creek base flow is derived from shallow ground water and it is assumed there is a linear relationship between the level in the piezometer and creek flow. Figure 4.13 is a scatter plot for all three locations and shows an apparent linear relationship for P8 versus T2 and P9 versus T3.

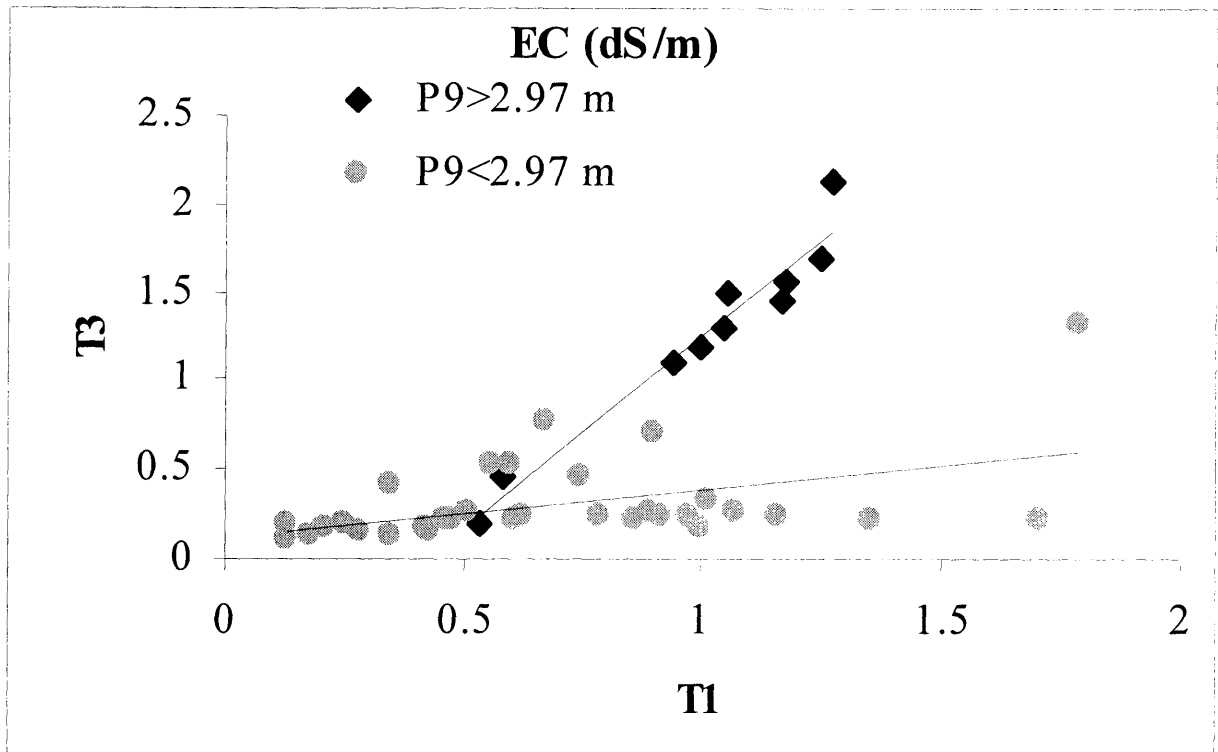
Figure 4.14 shows two distinct relationships in the EC at T3 versus the EC at T1. The relationship seems to be a function of the SWL in the adjacent piezometer (P9). This suggests that for high base flows there are corresponding decreases in the EC of the water (the depth shown (i.e. 2.97m) is the depth to the water table). A similar relationship was observed between T1 and T2, however, the relationship between T2 and T3 is not as clearly defined.



**Figure 4.12. Schematic of Surface Water Sampling Points**



**Figure 4.13. Scatter Plot of Surface Water EC versus Piezometer Standing Water Level**



**Figure 4.14. EC at T1 and T3 as a Function of the SWL in P9**

The EC of the water was statistically modelled in “S-plus”. A linear relationship was assumed with the EC being a function of:

- the total rainfall in the previous 10 to 60 days;
- the number days of rainfall in that period;
- the depth of effluent applied to the irrigation area in the same period, and the number of days it was applied over; and
- the water level in the adjacent piezometer in the current month and the preceding month (to account for the antecedent conditions of the catchment).

The holding pond depths were also included, though this dataset is considerably smaller than the rest. The inverse of the EC data was also used as a response variable as investigation of the database showed that the relationship between the EC and total rainfall was of this form.

It was found that the inverse transformed response variables (EC at each point) produced the linear models with the best fit. The predictor variables were the water levels for the current month in the piezometers P7 and P8, the total rainfall for the preceding 30 days and the number of days that rainfall occurred over. These predictor variables were found to account for 69% of variation of the transformed response variable at T1, and 62% of the variation of the transformed response variable at T2, while still retaining a level of significance for the coefficients. The variables used in the models are shown in Table 4.1. The output from “S-plus” for the models obtained for the inverse of the EC at T1 and T2 as the response variables are shown in Table 4.2. The linear equations obtained from this modelling exercise are given in Equation 4.1 and Equation 4.2, for T1 and T2 respectively.

**Table 4.1. Variables in Model**

Formula Variable	Variable in "S-plus" Output	Name
Y <sub>T1</sub>	1/T1	EC (at T1 or T2)
X1	P7	P7 SWL
X2	p8	P8 SWL
X3	Rain30	Rain in previous 30 days
X4	Dor30	Number of days of rain over that period

$1/Y_{T1} = 11.5 + 2.7X_1 - 6.1X_2 + 0.02X_3 - 0.19X_4 + e$  ..... **Equation 4.1**

$1/Y_{T2} = 9.6 + 2.3X_1 - 5.2X_2 + 0.02X_3 - 0.13X_4 + e$  ..... **Equation 4.2**

**Table 4.2. Output of Model Fitting**

```

Call: lm(formula = (1/T1) ~ p7 + p8 + rain30 + dor30, na.action = na.omit)
Residuals:
  Min    1Q  Median    3Q   Max
-1.613 -0.4913 -0.05131 0.4444 3.211

Coefficients:
              Value Std. Error  t value Pr(>|t|)
(Intercept) 11.4922  3.1115   3.6935 0.0007
          p7    2.7437  0.7954   3.4494 0.0014
          p8   -6.1071  1.7444  -3.5010 0.0012
        rain30  0.0230  0.0030   7.6410 0.0000
          dor30 -0.1940  0.0616  -3.1472 0.0032

Residual standard error: 0.8379 on 38 degrees of freedom
Multiple R-Squared: 0.6981
F-statistic: 21.96 on 4 and 38 degrees of freedom, the p-value is 1.874e-09

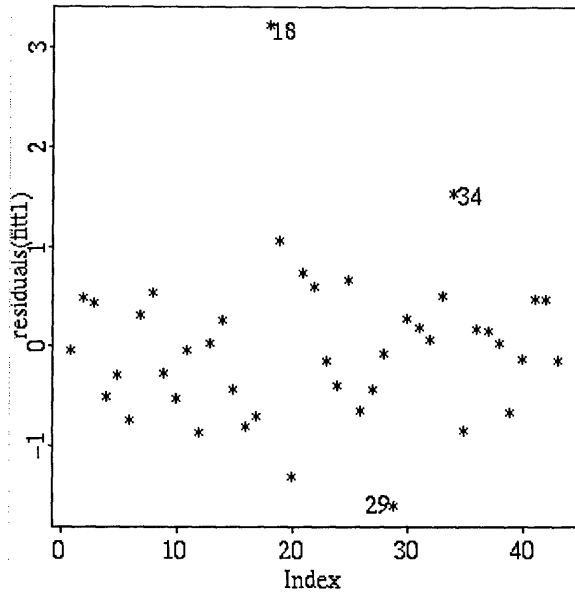
Call: lm(formula = (1/T2) ~ p7 + p8 + rain30 + dor30, na.action = na.omit)
Residuals:
  Min    1Q  Median    3Q   Max
-1.158 -0.4709 -0.09165 0.1705 3.421

Coefficients:
              Value Std. Error  t value Pr(>|t|)
(Intercept)  9.5990  3.2158   2.9850 0.0045
          p7    2.2908  0.8264   2.7720 0.0080
          p8   -5.2477  1.8008  -2.9141 0.0054
        rain30  0.0200  0.0031   6.3597 0.0000
          dor30 -0.1229  0.0610  -2.0148 0.0497

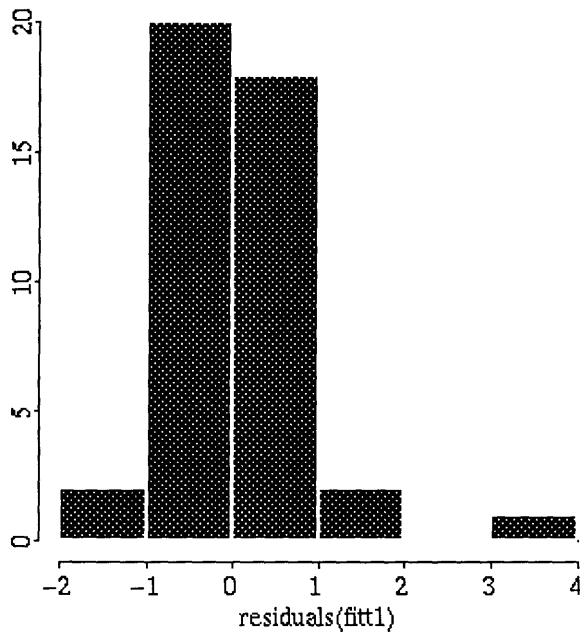
Residual standard error: 0.8753 on 47 degrees of freedom
Multiple R-Squared: 0.6246
F-statistic: 19.55 on 4 and 47 degrees of freedom, the p-value is 1.568e-09

```

The error terms (e) in Equation 4.1 and Equation 4.2 are an indication of the goodness of fit of the model. For the model to be successful the residuals are required to be independent and identically distributed with zero mean. Additionally, the residual standard error should be less than 0.3. These models have residual standard errors above 0.8, indicating that the error terms are relatively large for both models. The plot of the residuals for T1 are shown in Figure 4.15 and indicate that a few outliers should be evaluated in detail before using this model (Kufs 1992). The Kolmogrov-Smirnov test indicates that the residuals are not normally distributed at a 5% confidence interval. Figure 4.16 is a histogram of the residuals for T1.



**Figure 4.15. Plot of Residuals from T1 Model**



**Figure 4.16. Histogram of T1 Model Residuals**

Given the magnitude of the residuals in the previous two models, another model was applied to investigate the effects of the ground water only on the EC of the surface water. The results from this exercise are shown in Table 4.3 for EC at T2 as a function of the depth to the ground water at P7, P8 and P9 for the current month and the depth to the ground water at P7 and P8 for the previous month (to account for the antecedent conditions). Including the logarithm of P9 improved the fit with 81% of the variation in T2 explained by the predictor variables.



**Table 4.3. EC at T2 as a Function of Piezometer SWLs**

```
Call: lm(formula = T2 ~ p7 + p8 + log(p9) + p7lag + p8lag, na.action = na.omit)
Residuals:
    Min       1Q   Median       3Q      Max
-1.244 -0.1555  0.07807  0.2629  0.9198
Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept) -21.8612   2.0063  -10.8962  0.0000
           p7  -1.6547   0.4632   -3.5726  0.0009
           p8   4.2950   1.0822    3.9690  0.0003
    log(p9)  11.4812   1.9980    5.7464  0.0000
       p7lag  -0.8304   0.3669   -2.2636  0.0285
       p8lag   1.8040   0.8378    2.1533  0.0367

Residual standard error: 0.4482 on 45 degrees of freedom
Multiple R-Squared: 0.817
F-statistic: 40.18 on 5 and 45 degrees of freedom, the p-value is 1.665e-15
```

The residuals from this model are still slightly high but an improvement on the previous models, and the residuals do not show any autocorrelation. When the same model was applied to T1 data the Multiple R-Squared value was considerably less than the model for T2. These models are exploratory only and confirm that the ground water flow is an important factor in the surface water quality.

The ground water flow is a pathway for the nutrients applied in the manure and effluent and once they enter the ground water the nutrients can become a pollutant. Considerably more data and a detailed hydrological investigation of the ground water system are required to provide evidence to confirm that the feedlot is having an adverse effect on the adjacent water quality. This hydrological investigation is outside the scope of this study.

Exploratory analysis is important in the initial phase of extracting indices of environmental impact resulting from the utilisation of the manure and effluent. To expand on this exploratory analysis, the EMD was further developed to include a dynamic mass balance of the nutrients through their various pathways in the irrigation area.

## **4.7 Dynamic Mass Balance of Nutrients in the Utilisation Area**

Cycling of nutrients is a dynamic process that alters with time and changes of state in and around the system. A good understanding of nutrient cycling can be obtained by stepping through the time sequence of a particular elemental mass balance and a dynamic mass balance model with a graphical interface was developed to aid in understanding the nutrient dynamics. Time series data are used in this model to graphically represent the inputs, outputs and changes of state that have actually occurred in the effluent irrigation area, on a daily basis.

### **4.7.1 Gathering Data**

There is a considerable amount of data manipulation required before it is in a form that is useable for the graphical simulation. Collation of data occurs in a module that is separate from the actual simulation model and has two parts. The first part is the building of the tables to store collated data and the second part is the manipulation and transfer of these data.

The collation of data results in tables for each element that are common for the soil, crop, effluent and manure. At present there are thirteen such parameters (see Figure 4.17). The tables store data for each day there is an event that affects the soil store of the particular element, for example, an irrigation event or crop harvest. Also included are all soil analyses data, with the units of kg/ha. This information is stored in the "Irrigation Run Data" table. Other tables accessed during the graphical simulation are the "Irrigation - Volume and Depth" and "Rainfall and Evaporation" tables. The "Irrigation - Volume and Depth" table stores data for each day there is an irrigation event and the "Rainfall and Evaporation" table stores data for each day.

#### **4.7.2 Structure of the Dynamic Mass Balance**

In essence the above tables provide the basis for queries that are built at run time. They are select queries made up of two tables – the "Rainfall and Evaporation" and the "Irrigation Model <Parameter>" tables (built in the first module explained above), for each of the parameters selected (up to a maximum of five). The link in the query is a "one-to-many" join from the "Date" field in the "Rainfall and Evaporation" table to the "Sample Date" field in the "Irrigation Model <Parameter>" table. In a "one-to-many" relationship, a record in Table A can have many matching records in Table B, but a record in Table B has only one matching record in Table A (Microsoft Access 1997). This results in all the records being returned from the "Rainfall and Evaporation" table (i.e. a record for every day). Although there are a lot of gaps in the datasets this structure simplifies the code.

There are two blocks ("A" and "B") on either side of the irrigation area, each with five bays (see Figure 4.17). The interface consists of a number of rectangles representing each irrigation bay and above each block of irrigation bays are seven labels. When the form is first opened, five of these labels are blank and the last two have the heading "Soil Water Depth (mm)" and "Cum Irrig Vol (ML)". To the left of the irrigation bays are a list of available parameters to be simulated. Below this the user can input the initial and maximum soil water conditions. The default values are 70 and 300 mm respectively. Above the irrigation bays is a time-line that is incremented for each day of the data set. The soil option in the bottom left hand corner allows the user to select one of three options, which are outlined below.

1. Initial soil store - sets the initial soil values (baseline data and shows the changes in the soil store through the application of effluent and crop uptake).
2. Update Soil - this option updates the soil store of each element with values whenever they are available and reports the closing errors in the mass balance.
3. No soil – this option shows calculations using only the effluent input and the crop uptake. This allows the user to take a 'black box' approach to studying the soil store.

When the dynamic mass balance is run using existing data, the maximum values from the data sets are found. Based on these maximum values, scale factors are used to size coloured boxes that represent the quantity of a nutrient in the soil and the amount taken up by the crop for each parameter. Similar boxes are

used for the soil water depth and cumulative irrigation volume. For each daily time step, these boxes are re-sized to reflect the changes occurring. Calculations for each day are:

- soil water content (mm) based on the actual rainfall, evaporation and irrigation events;
- cumulative irrigation volumes for each bay; and
- inputs, outputs and changes of state within the soil (kg/ha) as a result of the addition of nutrients in the effluent and manure and extraction of nutrients by crop uptake.

Losses through leaching and runoff are not considered in the dynamic mass balance, but are included in the simulation model developed in Chapter 5.

Where possible actual values for input parameters are used for each of the irrigation bays. If these data are not available, average values are used. In the case of the effluent composition, if data are not available, the average is taken of the proceeding and following three values. If there are not three previous and three subsequent values, the first six or last six values are used to calculate the average.

The crop yields used are the average of the total weights recorded by the feedlot manager. The yield data set is incomplete for each irrigation bay and is therefore averaged over the whole block. Future crop harvest sampling should include obtaining yields from each irrigation bay.

There is soil data for some individual irrigation bays, however, most are averaged over the whole block. The total mass of a particular nutrient in the soil is calculated to the bottom of the B horizon. There were some records that were missing sampling depths and in these cases, the average sampling depth for that horizon completed the record.

## **4.8 Running the Dynamic Mass Balance**

This section demonstrates the database interface for a dataset spanning 3 years. This dataset is from Tullimba – the University of New England’s feedlot research facility. The parameters used in the run are Cl, K, Na, Mn and S and Figure 4.17 captures the initial conditions. The initial store of Na is relatively high, as expected in a sodic soil. Note the relative amounts of K, Na and Mn when viewing Figure 4.17 and Figure 4.18. Figure 4.18 is the screen capture at the end of the run. This figure indicates a considerable change in the balance of these three cations in the soil. The crop is mining the soil store of K, while the store of Na is increasing, with minimal amounts being removed by the crop. Figure 4.18 shows that the situation for Cl is similar. This suggests the predominant salt in the soil is NaCl. Therefore, agronomic management practices dealing with problems created by accumulation of Na need to be implemented if the Ca store in the soil begins to decrease. That is, addition of gypsum and lime and the inclusion of an irrigation leaching fraction. This is an example of the “data simulation” being used to extract information about the production process to update operational decisions (Sadeh & Griffin 1997).

**PARAMETER**

AI  
Ca

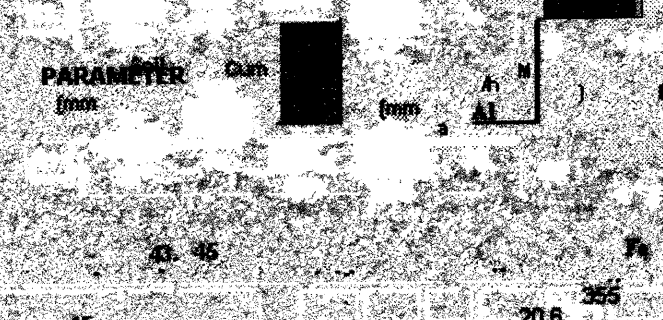
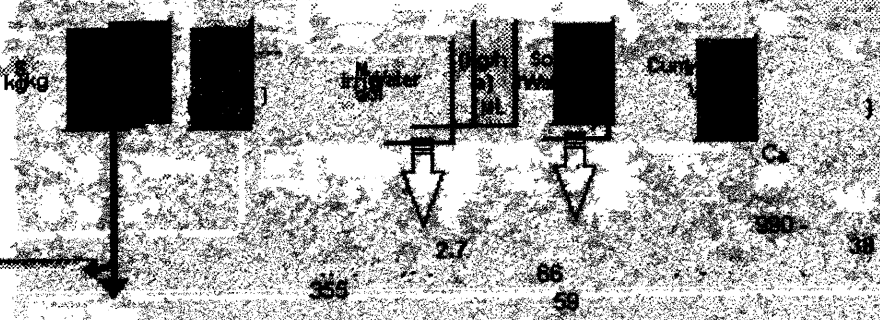
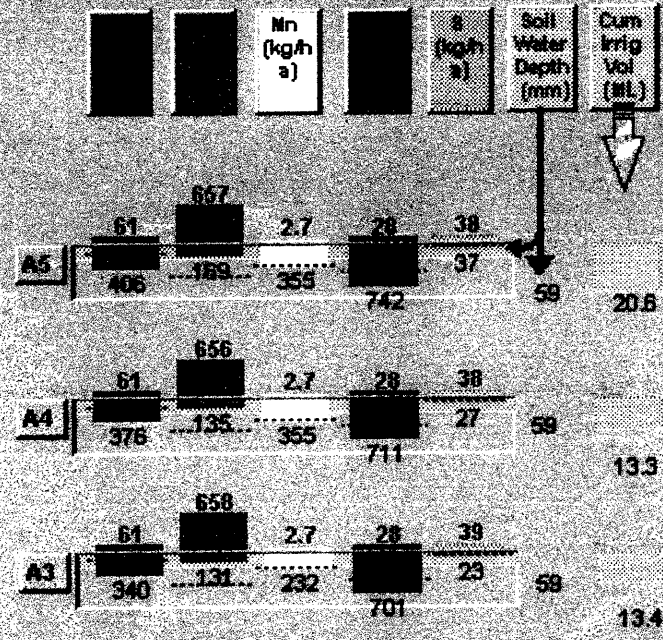
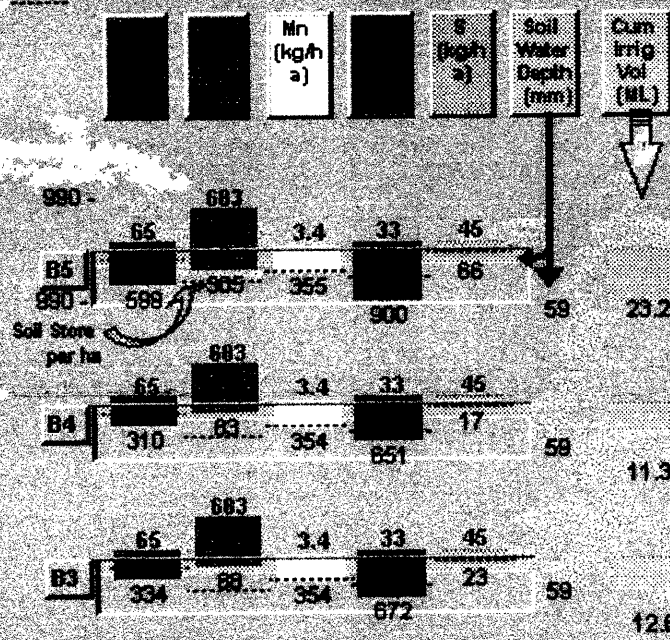
Fa

N

Initial Soil  
Water Depth

70

Maximum Soil



**PARAMETER**

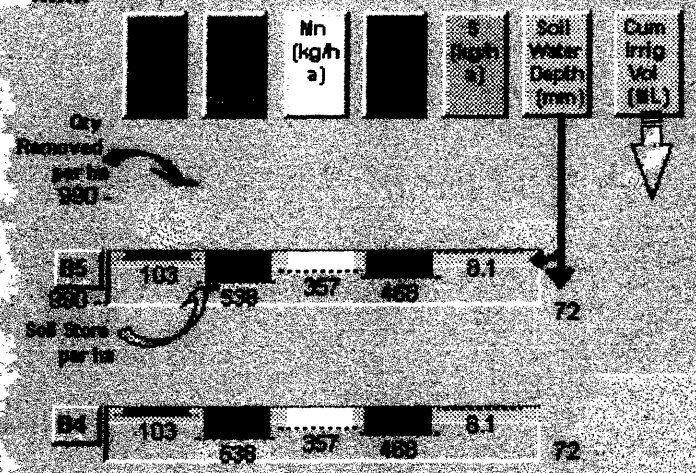
Depth (mm)

**PARAMETER**

Depth (mm)

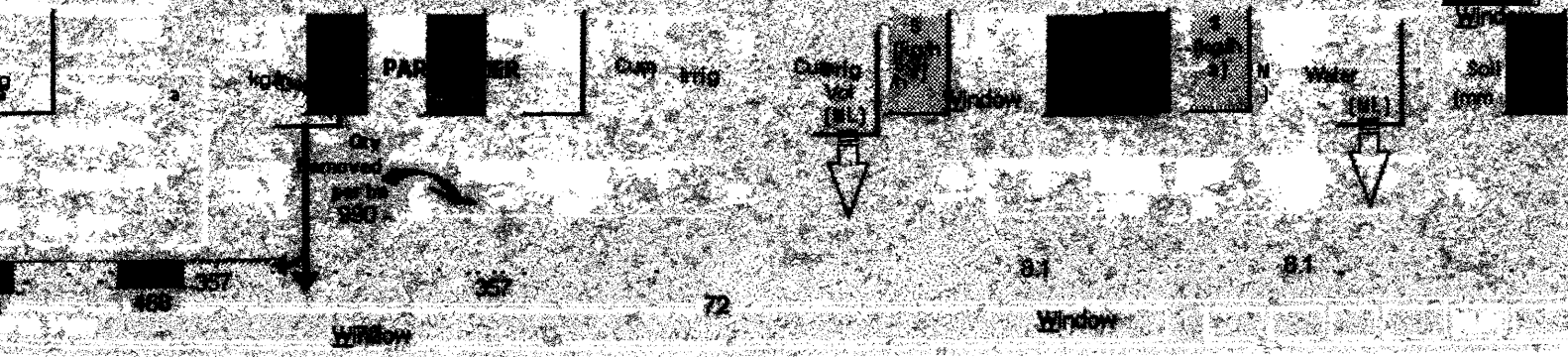
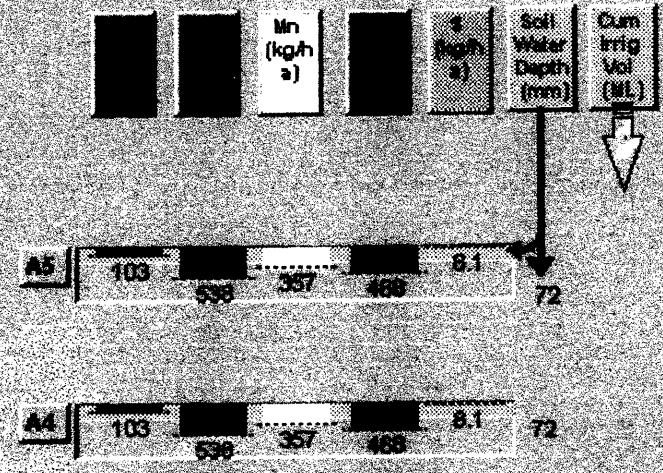
PARAMETER

A1



Dry Removed per ha 390

Soil Store per ha



Window

Window

The dynamic representation of the elements in the irrigation area is useful for visualising the relative quantities of the various elements and changes that are occurring on a daily basis. When combined with the use of the “static” queries in the EMD, the information obtained can indicate the direction further investigation should follow. For example, the cation exchange imbalance occurring in the soil is of concern and the results presented suggest some modelling is required in order to predict the effects different management practices will have on this trend.

The output from an integrated agricultural system utilising feedlot manure and effluent depends on stochastic processes. Climate, crop growth, manure and effluent composition and mineralisation of nutrients are all examples of stochastic variables. The next step for the database is to model these stochastic processes and use a simulation technique to quantify the sustainability of the system. Hamblin (1991) stated that some form of modelling or statistical analyses is required before point source measurements in time and space can be extrapolated and the use of a model to analyse and simulate the system can compensate for the limited observations that might be available (Hansen & Jones 1996). Simulation models can also provide an effective technique for predicting data (Hu 1995). Final justification for the modelling phase also comes from Hansen and Jones (1996) - sustainability deals with the future and therefore cannot be readily observed.

Chapter 5 further develops the dynamic mass balance as a predictive tool for the utilisation area using Monte Carlo simulation techniques. Monte Carlo is a modelling simulation technique that randomly selects a value for a particular input variable from a nominated statistical distribution. The model inputs are the statistical distributions obtained from the database and the model is run many times to produce a statistical distribution of the outputs.