Chapter 3

Preliminary Results

In this chapter we explore the data by describing its elementary characteristics. Histograms and numerical summaries of the data including water depth and salinity from Boat 1 and Boat 2 are shown.

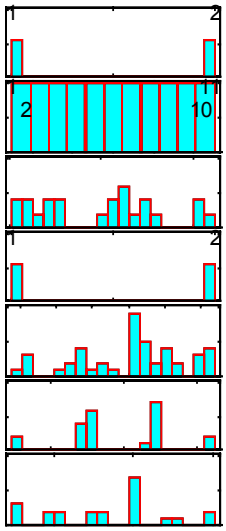
3.1 Depth

- **By Boat**

First we consider the depth of the water measured from each boat. Following the design described in Chapter 2, the depth was measured at each of 22 occasions from the two boats, one measurement from each side of each boat. In each boat there were two measurers. Each measurer had a fishing line and measured the depth of the water by slowly dropping the line until it reached the bottom. The water level was determined from a mark on the line and it was recorded in a logbook on each occasion. Thus we obtained 44 depth measurements. As described in Chapter 2, measurements were taken at the same position on several occasions according to a specified design. Depth was measured in centimeters. The determinants are boat (1 to 2), position (1 to 11), time, observer (1 to 2), latitude and longitude. These data are summarised in Figure 3.1.

Figure 3.1 shows that the mean of the depth measurements is 181.98 cm, and the standard deviation is 29.69 cm. The minimum and maximum values are 115.5 and 227.5 cm, respectively. The numerical summaries for the determinants are given primarily for data checking purposes.

Box plots of the depth measurements for Boat 1 and Boat 2 are shown in Figure 3.2. From this figure it is clear that the depth measurement in Boat 2 is higher than the depth in Boat 1, and the distribution in Boat 1 is more spread than that of Boat 2. The distributions are approximately symmetric, with no outliers. The medians in Boat 1 and Boat 2 are 162 cm and 208 cm, respectively.



col	variable	size	mean	st dev	min	max
1	Boat	44	1.5	0.506	1	2
2	Occasion	44	6	3.199	1	11
3	Time	44	10.002	0.67	9.03	11.23
4	Observer	44	1.5	0.506	1	2
5	Depth	44	180.977	29.693	115.5	227.5
6	Location(N)	44	54.948	0.355	54.24	55.62
7	Location(E)	44	16.335	0.37	15.74	17.01

Depth Measurements in Pattani Bay

Figure 3.1: Overall summary of the water depth data collected from Pattani Bay

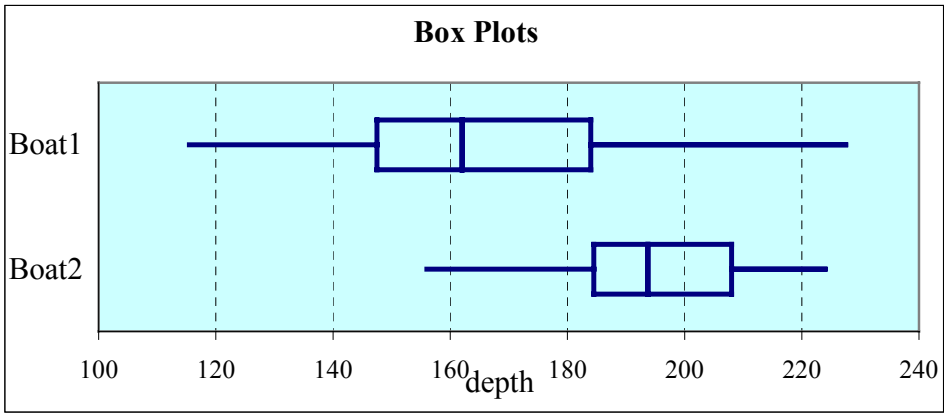


Figure 3.2: Box plots of the depth measured, by boat

Table 3.1 shows the result of the two-sample t-test for the depth measurements for Boat 1 and Boat 2. The difference is statistically significant (t = 4.121, p-value = 0.0002). The sum of squares of residuals is 94.69 and the r-squared goodness-of-fit is 90%. The residual standard deviation is 0.63, the difference between the sample means is 31.5 cm, and the 95% confidence interval for this difference ranges between 16.07 cm to 46.93 cm. The p-value for Levene's test that the populations have the same standard deviation is 0.033, indicating that the standard deviations are different.

Two-sample t-test: Resid SS: 26997.602 r-sq:0.288

Factor	*df*		*t*	*p-val*	*s*	*diff*	*CI/2*	*Lev p*
	v_1	v_2						
boat	1	42	4.121	0.0002	25.354	31.500	15.427	0.033

Table 3.1: Two-sample t-test of salinity measurements

Figure 3.3 shows the confidence intervals for the means of the depth measurements between Boat 1 and Boat 2. As stated above, the difference between the two boats is statistically significant. The mean depth from Boat 1 is approximately 165 cm while the mean depth from Boat 2 is approximately 198 cm. The confidence interval from Boat 1 is substantially wider than that from Boat 2.

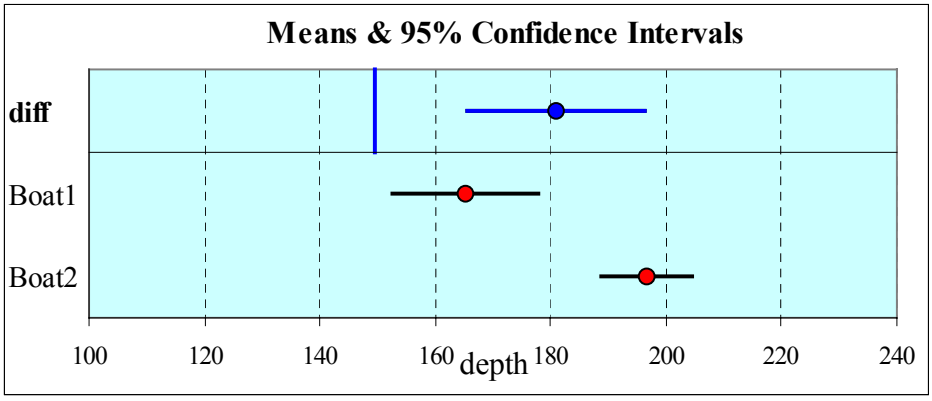


Figure 3.3: 95% Confidence interval of depth measured, by boat

The normal scores plot of the residual depth measurements is presented in Figure 3.4. This plot shows an approximately linear trend, suggesting that the normality assumption is reasonable for these data.

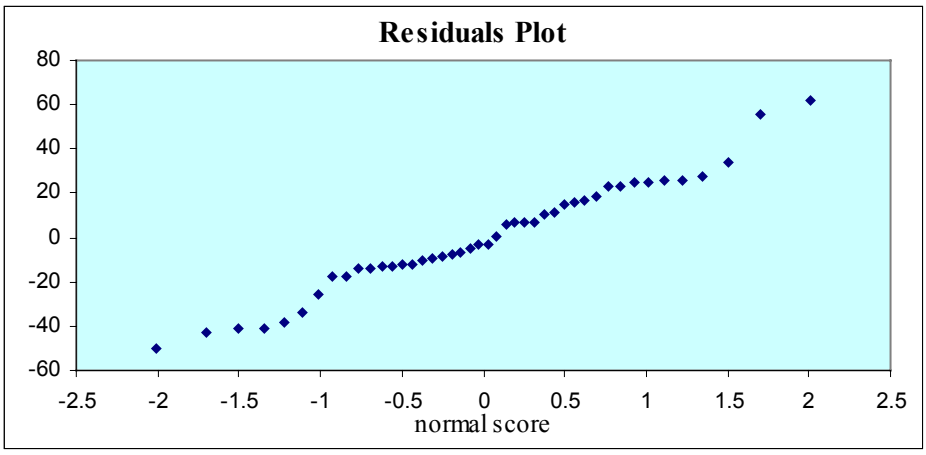


Figure 3.4: Normal score plot of depth measured, by boat

- **By Occasion**

Figure 3.5 shows box plots of the depth measurements for each of the 22 occasions. The figure shows some variation in the distributions. The median of the depth for these subsamples ranges from 120 cm (occasion 21) to 225 cm (occasion 16).

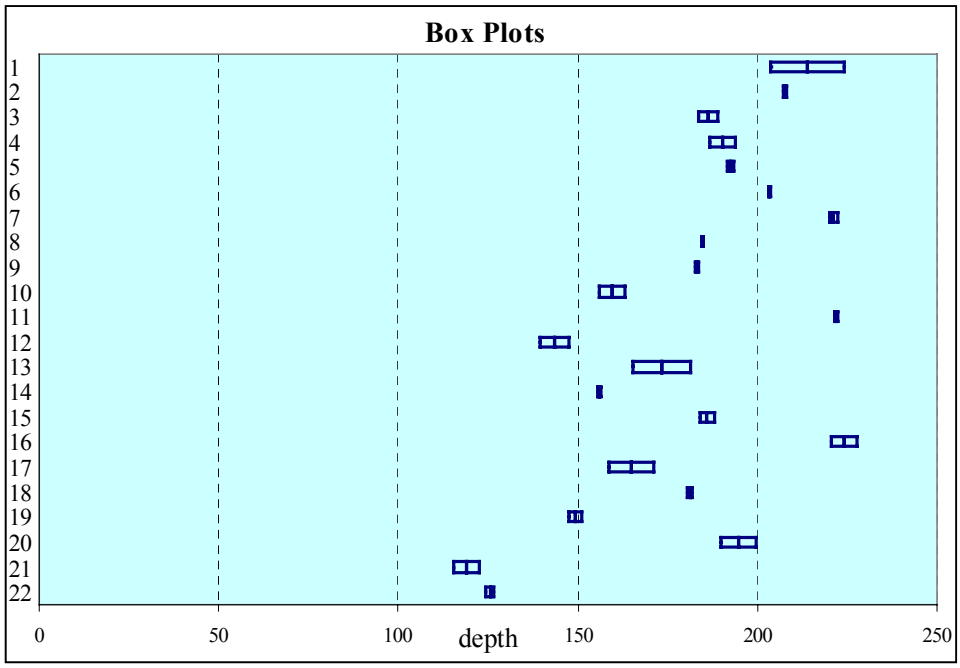


Figure 3.5: Box plots of the depth collected, by occasion

Table 3.2 shows the result of the one-way analysis of variance for the depth measurements on each occasion. The difference in occasions is statistically significant (F = 61.56, p-value = 0). The sum of squares of residuals is 634.44 and the r-squared

goodness-of-fit is 98%. The residual standard deviation is 5.37. The root-mean-squared difference between the sample means is 139.74. The 95% confidence interval for this difference ranges between 114.18 cm to 165.29 cm. The p-value for Levene's test that the populations have the same standard deviation is 0.00. Since this p-value is less than 0.05 the populations have different standard deviations.

Figure 3.6 shows confidence intervals for the means of the depth measurements by occasion. There is a significant difference between each occasion. The mean depth for these subsamples ranges from 120 cm (occasion 21) to 220 cm (occasion 17).

Figure 3.7 shows the normal scores plot of the residual depth measurements from this model. The normal scores plot is approximately linear, so the normality assumption is reasonable for this data.

One-way Results: Resid SS: 634.437 r-sq: 0.983

Factor	*df*		*F*	*p-val*	*s*	*diff*	*CI/2*	*Lev p*
	v_1	v_2						
occasion	21	22	61.555	0.000	5.370	139.737	25.555	0.00

Table 3.2: One-way analysis of variance of depth measurements

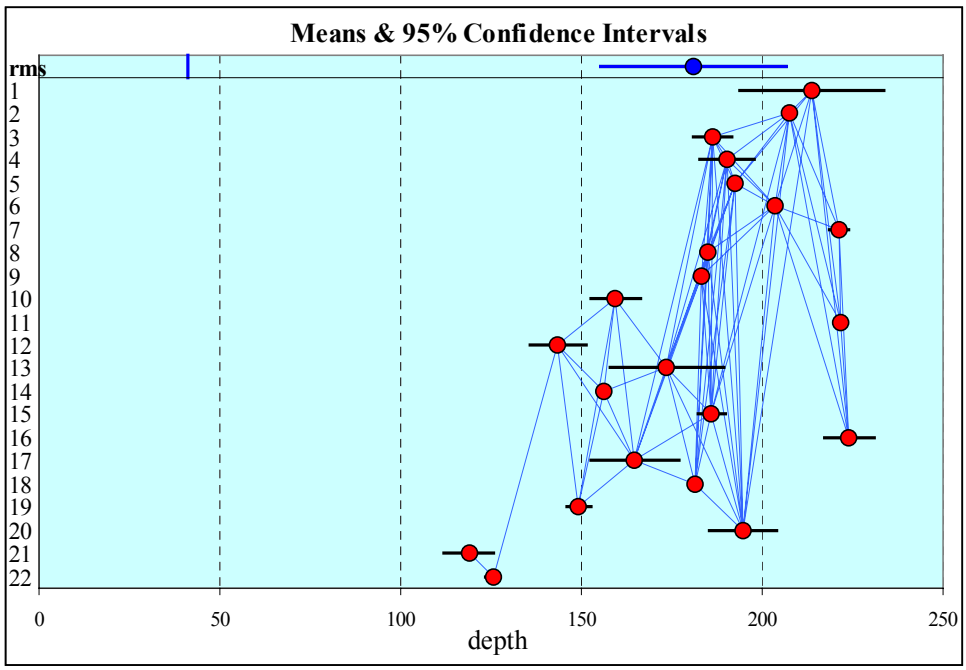


Figure 3.6: 95% Confidence intervals of the depth measured, by occasion

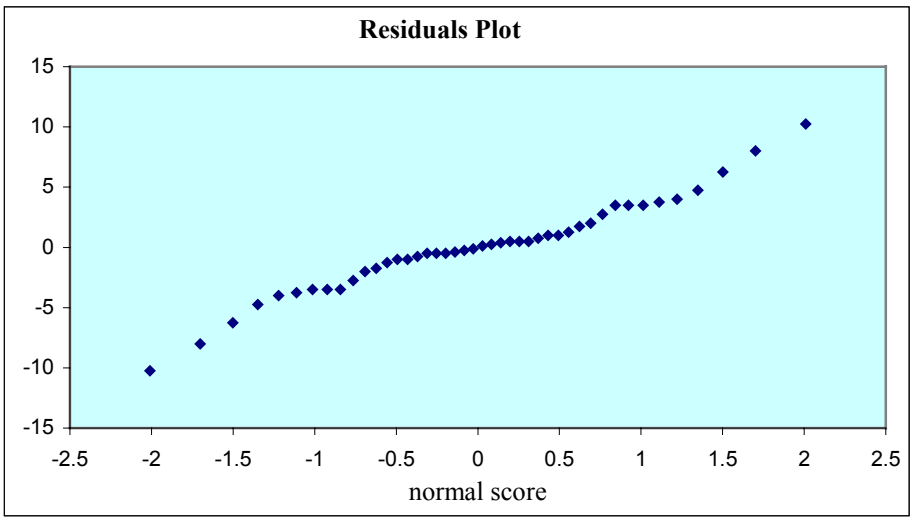


Figure 3.7: Normal score plot of residual depth measured

- **By Time**

We now consider how the depth at each location depends on the time of measurement. We obtained 9 locations in all 22 occasions as follows (occasion 1 to 11 from Boat 1 and occasion 12 to 22 from Boat 2):

Location	Occasion
A	12, 17, 22
B	1, 6, 11, 21
C	10
D	13, 5, 18
E	2, 7, 19, 20
F	9
G	15, 16
H	14, 4
I	3, 8

Table 3.3: Number of occasions for each location

Figure 3.8 shows the depth at these nine locations depending on the time of measurement. These depths vary because of tidal movements. Considering location B, at 9.05 am the estimated depth was 143.5 cm. By 10.00 am the estimated depth had increased to 165 cm. By 11.01 am depth at location B had decreased to 125.5 cm. However the depth increased again to 160 cm at 11.10 am. At location A at 9.03 am the depth was 213.75 cm. At 10.02 am the tide decreased and the depth became 203 cm. At 11 am the tide increased and the depth at location A became 222 cm. The depth did not change much in locations D, E, G and I. At locations C and F, the depth was measured only once. At location H the depth was measured twice at time 9.38 am and 9.44 am and the depth was 186.25 cm and 186 cm, respectively.

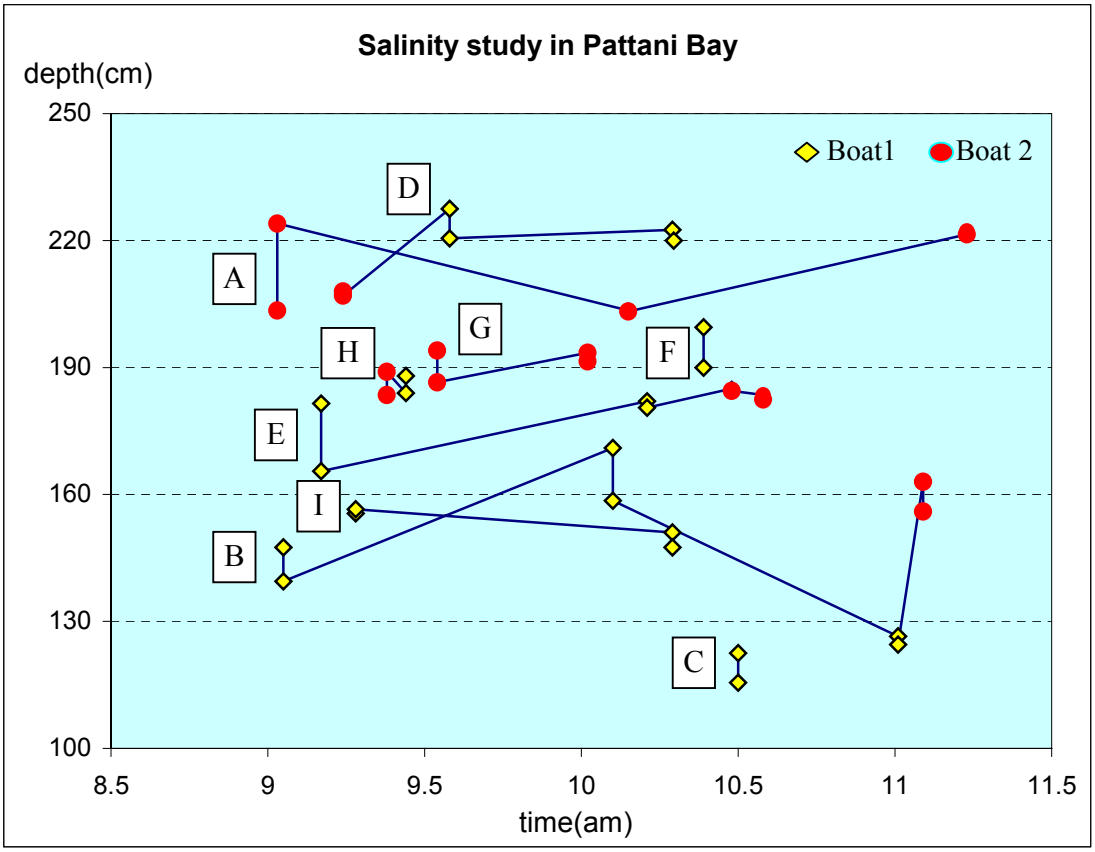


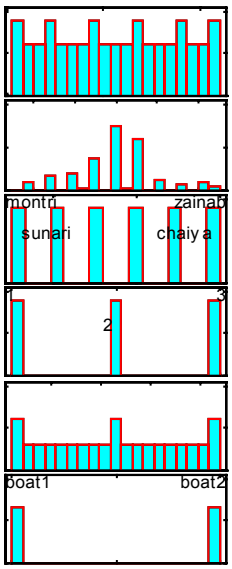
Figure 3.8: The depth at each location by time

3.2 Salinity

- **By Boat**

Second, we investigate the salinity of the water samples collected from the two boats. Following the design described in Chapter 2, two water samples were taken in the Bay on various occasions from the two boats, one from each side of each boat. These samples were stored in an icebox in a sealed plastic film container, and analysed later in the day for their salinity content. Each sample was measured independently by six readers sharing three salinometers. Thus we obtained 264 salinity samples. The response variable is salinity in parts per thousand. The determinants are reader (1-6), instrument (1-3) and occasion (1-22). These data are summarised in Figure 3.9.

Figure 3.9 shows that the histogram of salinity is unimodal with slight positive skewness (coefficient 0.16). However the kurtosis is close to zero, which is the value for a normal distribution. The mean of the salinity measurements is 23.98 ppt, and the standard deviation is 1.92 ppt. The minimum and maximum values are 19 and 29 ppt, respectively. The summaries for the determinants are given purely for data checking purposes.



col	variable	size	mean	st dev	skew	kurt	min	max
1	Sample	264	22.5	12.723	0	-1.201	1	44
2	Salinity(ppt)	264	23.975	1.923	0.163	0.497	19	29
3	Reader	264	3.5	1.711	0	-1.27	1	6
4	Instrument	264	2	0.818	0	-1.506	1	3
5	Occasion	264	11.5	6.356	0	-1.205	1	22
6	Boat	264	1.5	0.501	0	-2.015	1	2

Salinity Measurements in Pattani Bay

Figure 3.9: Overall summary of the salinity data collected from Pattani Bay

Figure 3.10 shows the box plots of the salinity measurements for Boat 1 and Boat 2. From this figure it seems clear that the salinity measurements from Boat 2 are higher than the measurements from Boat 1. The distribution from Boat 2 is approximately symmetric, while for Boat 1 it is skewed to the left. The medians in Boat 1 and Boat 2 are 23.8 and 25, respectively.

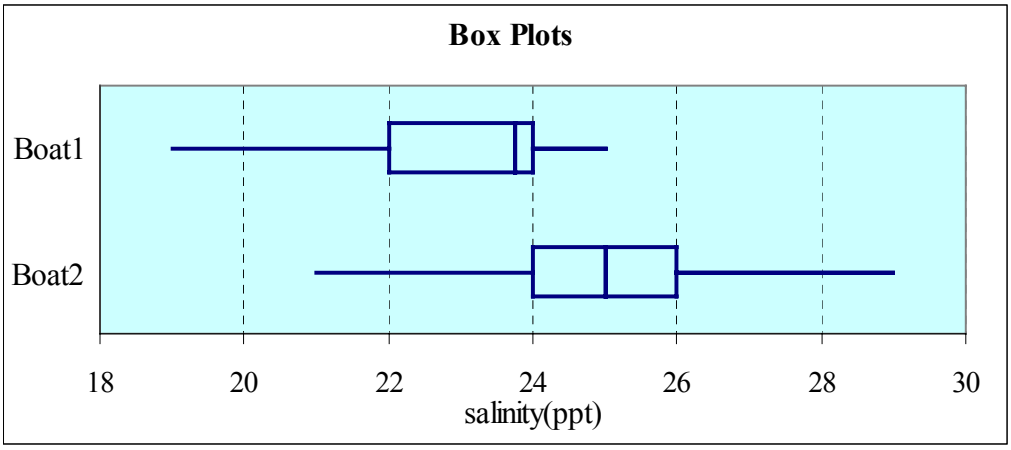


Figure 3.10: Box plots of salinity data collected, by boat

Table 3.4 shows the result of the two-sample t-test. The salinity measurements for Boat 1 and Boat 2 are statistically significantly different (t = 8.50, p-value = 0.00). The sum of squares of residuals is 762.51 and the r-squared goodness-of-fit is 22%. The residual standard deviation is 1.71, the difference between the sample means is 1.78, the 95% confidence interval for this difference are 1.37 and 2.20. The p-value for Levene's test that the populations have the same standard deviation is 0.65, indicating that the statistical assumption of equal variances is reasonable.

Two-sample t-test: Resid SS: 762.513 r-sq: 0.216

Factor	***df***		***t***	***p-val***	***S***	***Diff***	***CI/2***	***Lev p***
	v_1	v_2						
boat	1	262	8.496	0.0000	1.706	1.784	0.413	0.652

Table 3.4: Two-sample t-test of salinity measurements

Figure 3.11 shows the confidence intervals for the means of the salinity measurements in Boat 1 and Boat 2. There is a significant difference in salinity measurements between the two boats. The mean salinity in Boat 1 is approximately 23 ppt and the

mean in Boat 2 is approximately 25 ppt. The confidence intervals are relatively narrow.

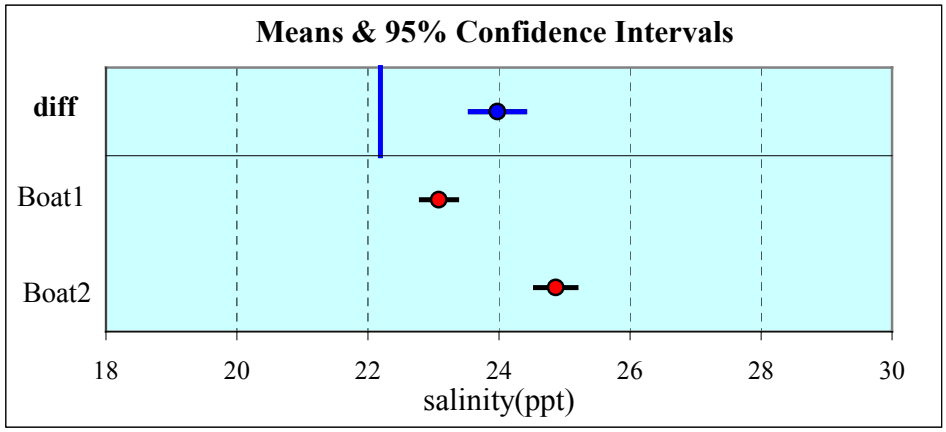


Figure 3.11: 95% Confidence interval of salinity data collected, by boat

- **By occasion**

Figure 3.12 shows box plots of the salinity measurements for each of the 22 occasions. In this figure there are four low outliers corresponding to the first, eighteenth and twentieth occasion, and there are two high outliers corresponding to in the fifth and twenty-second occasion. The median salinity for these sub samples ranges from 20 ppt (occasion 21) to 28 ppt (occasion11).

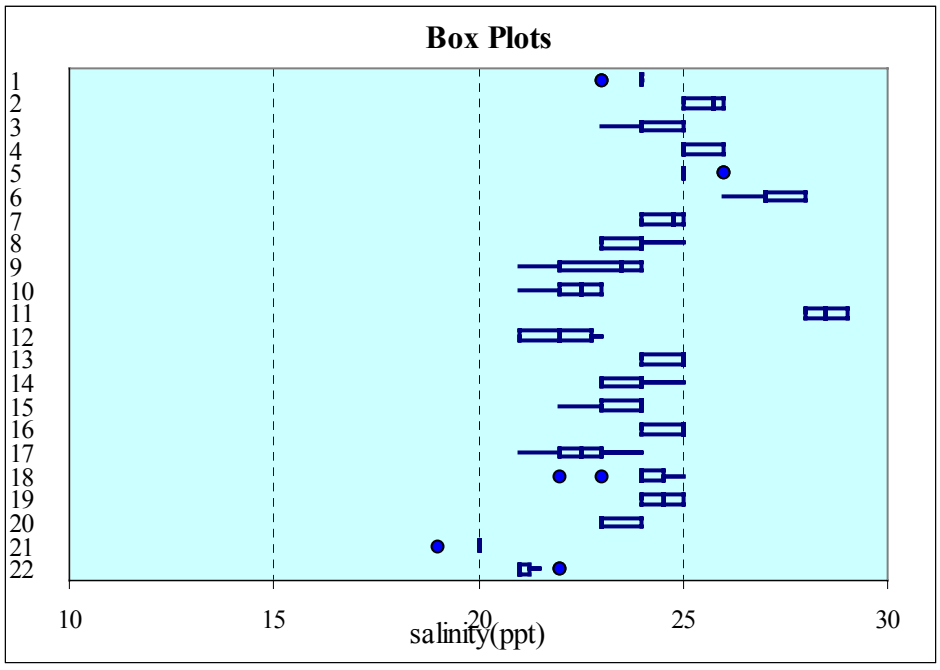


Figure 3.12: Box plots of salinity data collected, by occasion

Table 3.5 shows that the salinity measurements for each 22 occasions are statistically significantly different (F = 106.844, p-value = 0.0). The sum of squares of residuals is 94.69 and the r-squared goodness-of-fit is 90%. The residual standard deviation is 0.63, the root-mean-squared difference between the means is 8.76 and the 95% confidence interval for this difference ranges from 7.69 to 9.83. The p-value for Levene's test (0.004) indicates that the populations do not have the same standard deviation.

One-way Results: Resid SS: 94.687 r-sq: 0.903

Factor	*df*		*F*	*p-val*	*s*	*diff*	*CI/2*	*Lev p*
	v_1	v_2						
occasion	21	242	106.844	0.000	0.626	8.755	1.071	0.004

Table 3.5: One-way analysis of variance of salinity measurements

Figure 3.13 shows the differences in the means of the salinity measurements on the different occasions. There are three occasions that shows differences from the others. Occasion six and eleven have salinity levels which are higher than the others, while occasion 21 is lower.

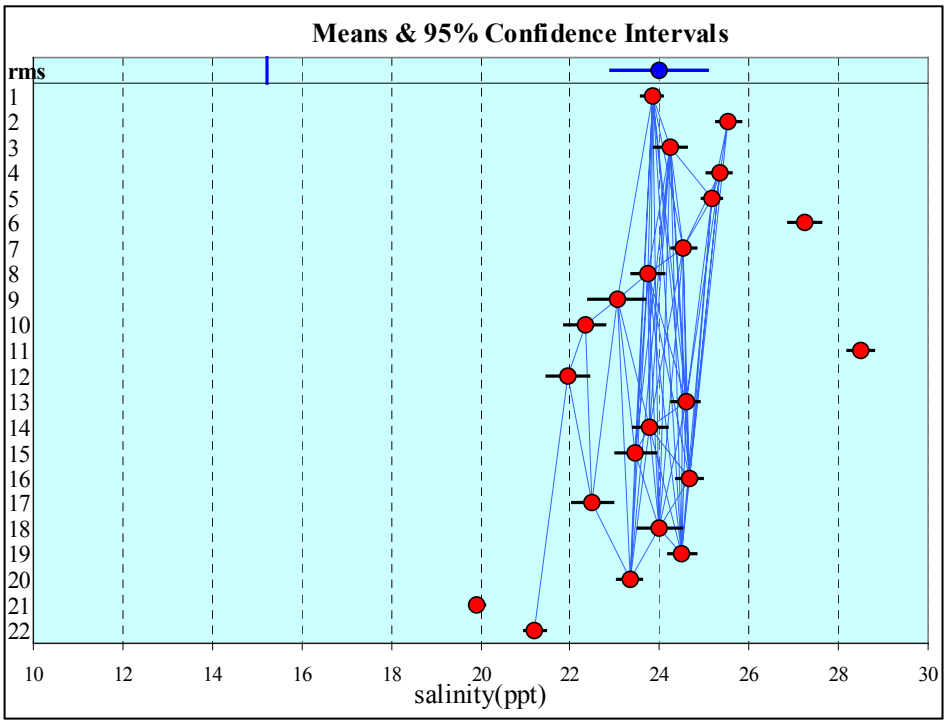


Figure 3.13: 95% Confidence interval of salinity data collected, by occasion

Figure 3.14 shows the normal scores plot of the residual salinity measurements. The normal scores plot exhibits an acceptably linear, suggesting that the normality assumption is reasonable for these data.

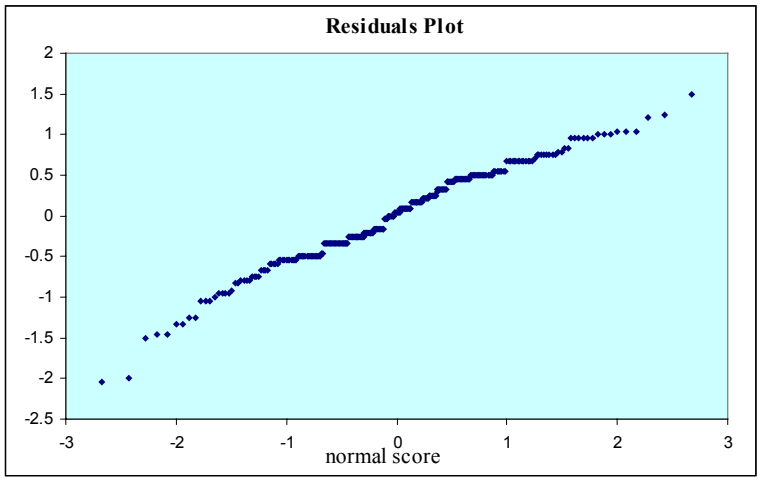


Figure 3.14: Normal score plot of salinity data collected, by occasion

Next, we investigate the salinity measurement by occasion adjusted for reader. As described in Chapter 2, each sample was measured independently by six readers sharing three salinometers.

Figure 3.15 shows box plots of salinity measurement for each occasion adjusted for the six readers. There are two outliers corresponding to occasions 17 and 18.

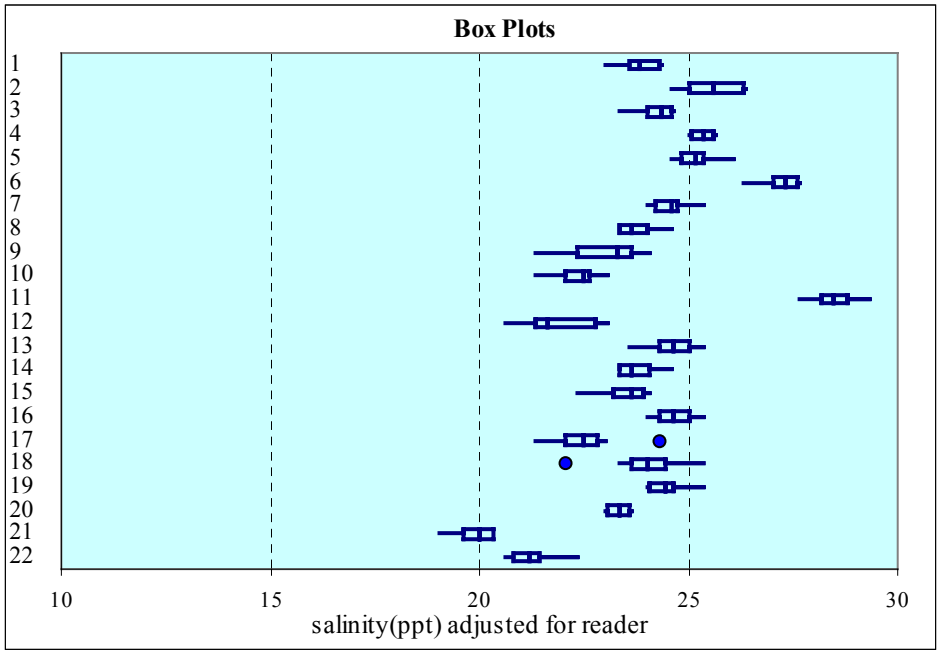


Figure 3.15: Box plots of salinity data collected, by occasion adjusted for reader

Table 3.6 shows that the salinity measurements for occasion adjusted by reader are statistically significantly different (F = 137.80, p-value = 0.00). The sum of squares of residuals is 71.90 and the r-squared goodness-of-fit is 93%. The residual standard deviation is 0.55, the root-mean-squared difference between the sample means is 8.76, and the 95% confidence interval for this difference ranges from 7.81 to 9.70. The p-value for Levene's test (0.02) indicates that the populations do not have the same standard deviation.

Two-way Results: Resid SS: 71.899 r-sq: 0.926

Factor	*df*		*F*	*p-val*	*s*	*rms*	*CI/2*	*Levp*
	v_1	v_2						
Occasion	21	242	106.844	0.0000	0.626	8.755	1.071	0.004
Reader+	5	237	15.024	0.0000	0.551			
Occasion+	21	237	137.802	0.0000	0.551	8.755	0.944	0.020

Table 3.6: Two-way analysis of variance of salinity measurements

Figure 3.16 shows how the means of the salinity measurement are different. The means from occasion six and 11 are higher than other occasions. The mean from occasion 21 is lower than the others.

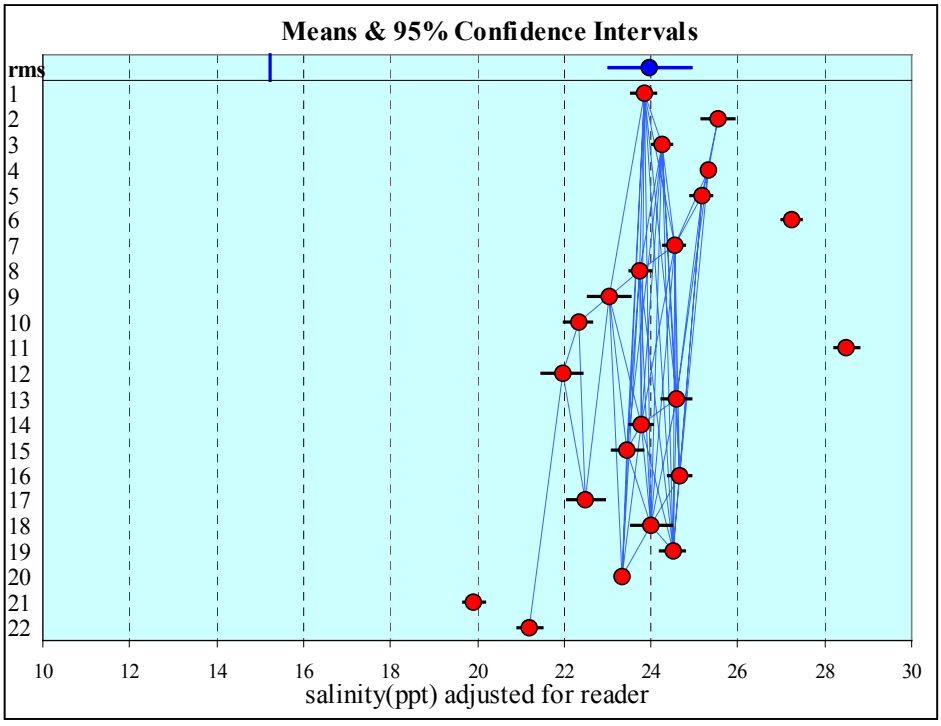


Figure 3.16: 95% Confidence intervals of salinity data, by occasion adjusted for reader

The normal scores plot of the salinity measurements adjusted for occasion and reader is depicted in Figure 3.17. The normal scores plot shows a smooth linear trend, suggesting that the normality assumption is reasonable for these data. However there are two more high and two more low outliers.

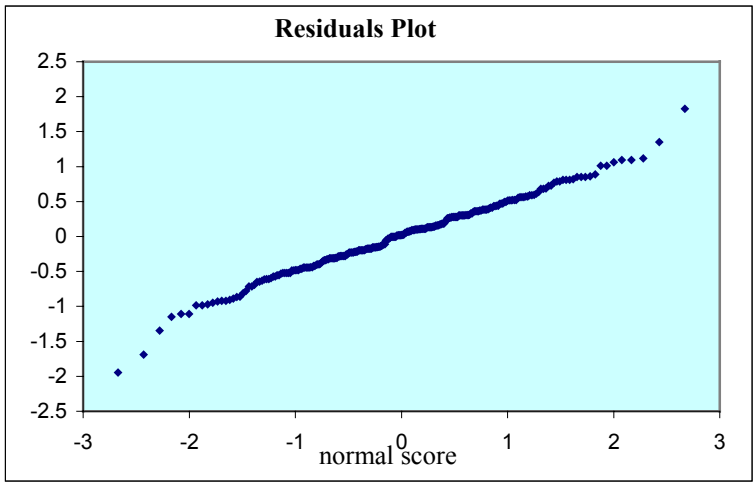


Figure 3.17: Normal score plot of salinity data collected, by occasion adjusted for reader

Finally, we investigate the salinity measurements by reader adjusting for occasion. Figure 3.18 shows box plots of salinity measurement for the six readers, adjusted for occasion. There are two low outliers, corresponding to reader one and five, and one high outlier, corresponding to reader three. Median salinity levels range from 23.53 (reader three) to 24.5 (reader five). The spread in reader four's measurement appears to be wider than the other readers.

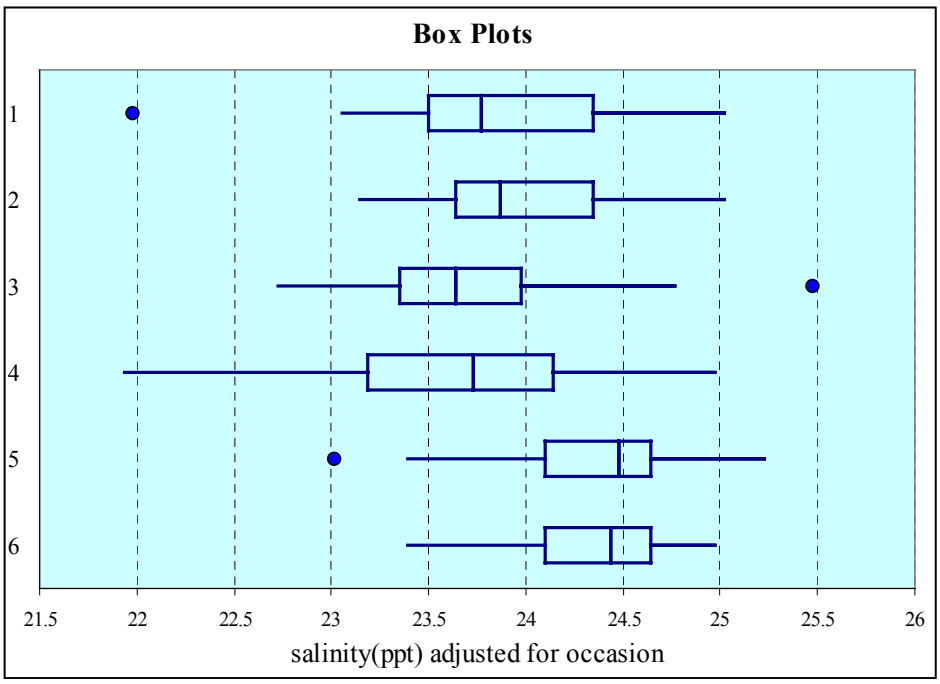


Figure 3.18: Box plots of salinity data, by reader adjusted for occasion

Table 3.7 shows how the salinity measurements for the reader are different (F = 15.02, p-value = 0). The sum of squares of residuals is 71.90 and the r-squared goodness-of-fit is 93%. The residual standard deviation is 0.55, the difference between the sample means is 0.788, the 95% confidence interval for this difference is (–0.48, 1.09). The p-value for Levene's test that the populations do not have the same standard deviation is 0.04. Reader four has higher standard deviation as shown in Table 3.8.

Two-way Results: Resid SS: 71.899 r-sq: 0.926

Factor	***df***		***F***	***p-val***	***s***	***rms***	***CI/2***	***Levp***
	v_1	v_2						
reader	5	258	1.238	0.2917	1.919	0.788	1.063	0.391
Occasion+	21	237	137.802	0.0000	0.551			
reader+	5	237	15.024	0.0000	0.551	0.788	0.305	0.040

Table 3.7: Two-way analysis of variance of salinity measurements

reader	*Size*	*Mean*	*SE*	*St.Dev*
1	44	23.920	0.286	1.898
2	44	23.955	0.264	1.751
3	44	23.648	0.270	1.790
4	44	23.625	0.324	2.152
5	44	24.364	0.298	1.978
6	44	24.341	0.289	1.916

Table 3.8: Comparison of the mean and standard deviation of the six readers

Figure 3.19 shows that the means of the salinity measurements are different between readers 1 to 4 and readers 5 and 6. The means for readers 5 and 6 are approximately 0.8 ppt high than the others. There is no difference in salinity measurements between readers one to four.

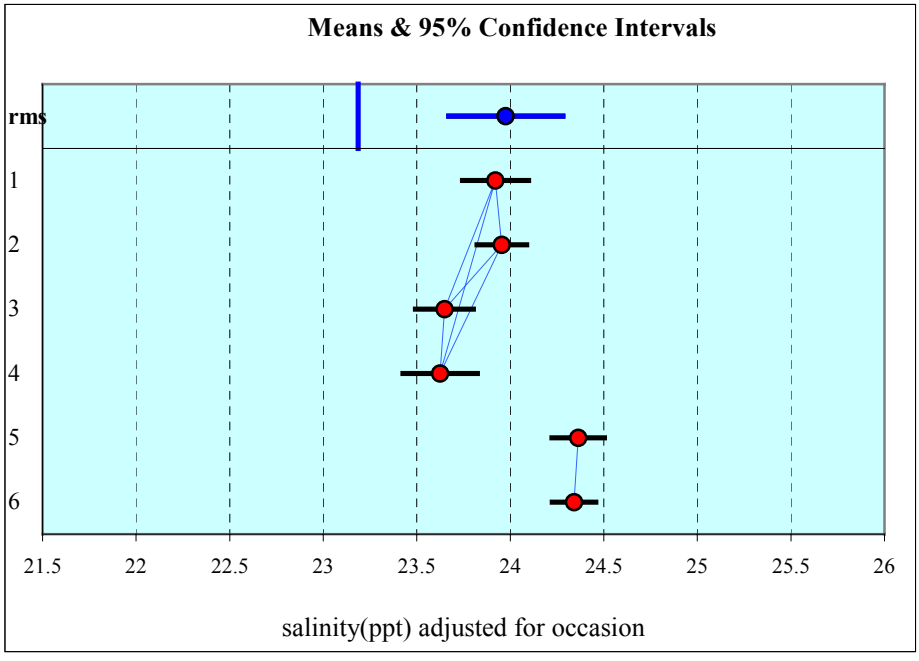


Figure 3.19: 95% Confidence intervals of salinity by reader adjusted for occasion

In fact the differences between the readers is accounted for by differences between the three salinometers. Readers 1 and 2 used salinometer 1, readers 3 and 4 used salinometer 2, and readers 5 and 6 used salinometer 3. Thus salinometer 3 gave higher readings than the other two instruments, possibly due to some residual effect because

of improper cleaning. If this explanation is true, the salinity measurements made with instrument 3 should be adjusted by subtracting 0.8 ppt.

3.3 Summary

The preliminary results may be summarized as follows.

First, the depth varies with both time and location. The variation with time is due to the effect of the tide, while the variation with location is due to the fact that the Bay does not have a flat bottom. The minimum depth was measured as 120 centimeters, and the maximum depth was found to be 225 centimeters. The maximum tidal variation at any location was seen to be 75 centimeters.

Second, salinity depends on occasion (time and location) and measuring instrument. The variation with occasion is due to the surface tidal current in estuary. The minimum salinity was measured as 19 ppt, and the maximum salinity was found to be 29 ppt. One of the three salinometers used gave readings that were consistently higher, by 0.8 ppt, than the other two.