DEPARTMENT OF CHEMISTRY

UNIVERSITY OF SWAZILAND

C518 **EVALUATION OF ANALYTICAL DATA**

MAY 2009 FINAL EXAMINATION

Time Allowed: Three (2½) Hours

Instructions

1. **This examination has five (5) questions and one data sheet.**
2. **There are two sections in this examination.**

   Answer Any Three Questions.

   You are required to answer At Least One Question from Each Section;

   Diagrams should be clear, large and properly labeled. Marks will be deducted for improper units and lack of procedural steps in calculations.

   You can print answers for section B from the printer provided.
3. **Each question is worth 25 marks.**

Special Requirements

1. **Data sheet.**
2. **Graph Paper.**
3. **Statistical Tables.**
4. **Computer**
5. **Printer**

YOU ARE NOT SUPPOSED TO OPEN THIS PAPER UNTIL PERMISSION TO DO SO HAS BEEN GIVEN BY THE CHIEF INVIGILATOR.

SECTION A

Answer at least one question from this section.

QUESTION 1 [25]

a) The following absorbance data was obtained in triplicate during a standard additions determination of zinc in a soil sample using atomic absorption, AA, following classical dissolution of 500-mg portions:

Addition 0: **0.102**
Addition 1: **0.149**
Addition 2: **0.205**
Addition 3: **0.246**

where 0μL, 5μL, 10μL, and 15μL of a 100 ppm Zn standard solution was added to 10-mL aliquots of sample, respectively.

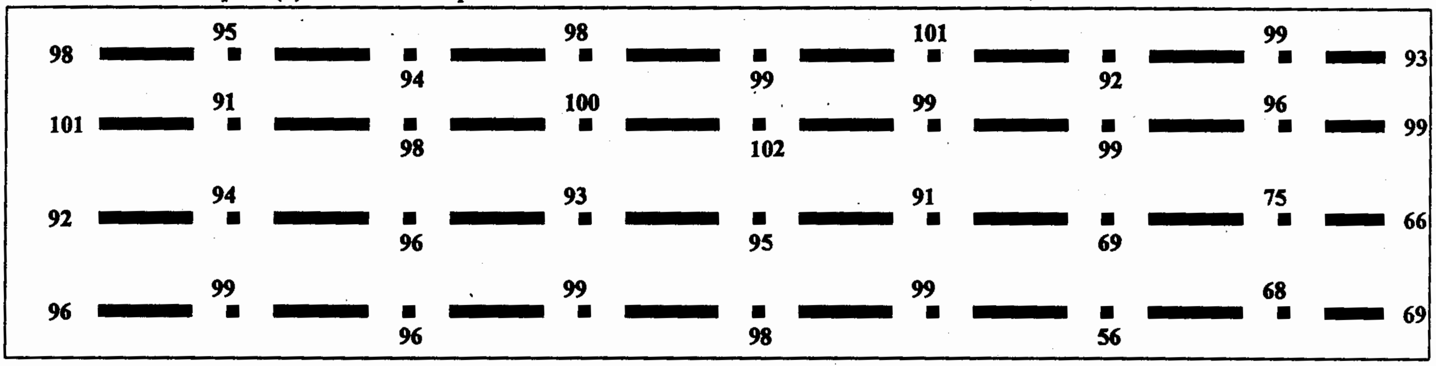
(i) Perform a linear regression on the calibration curve using the least squares method [4]

(ii) Calculate the relative error, in %, associated with the intercept, S_{vc} [2]

(iii) Calculate the absolute error, in ppm, associated with the analytical measurement, S_a, if five repetitive measurements of a sample solution aspirated into the AA gave the following results: 100 ppm, 99 ppm, 101ppm, 98 ppm, and 100 ppm [2]

(iv) Calculate the absolute subsampling uncertainty, S_{ss}, in ppm units if five 500-mg portions of the sample were found to contain 103ppm, 105ppm, 95 ppm, 101 ppm, and 108 ppm. [2]

b) Thirty six (36) samples of soil were taken from a field to map the spatial variability of zinc. 500-mg portions of each sample were digested and zinc measured by AA following the standard additions procedure on the same day and same instrument as in part (a) above. The spatial distribution of zinc was found to be as follows:



(i) Use the Kolmogorv-Smirnoff test to show that the distribution of zinc in the field is not Gaussian. [5]

(ii) What is meant by a "hot spot" or "coldspot" in analytical sampling? [2]

(iii) In this population, identify and map the points that have resulted in a "hot spot" or "coldspot" causing the non-Gaussian distribution of zinc in the population. [2]

(iii) Calculate the uncertainty due to the sampling operation in ppm units [2]

(iv) Use the Student's t-test equation to determine the minimum number of samples to be taken from the population if the average value of zinc is to be within the error due to sampling at the 95% confidence level. [4]

QUESTION 2 [25]

a) In data acquisition, noise is an important concept in instrumental analysis as it is a predominant factor in determining precision and detection limits.

(i) What is meant by "signal" in analytical data acquisition? [1]

(ii) What is meant by "noise" in analytical data acquisition? [1]

(iii) What is the significance of the concept "signal-to-noise ratio" in analytical data acquisition? [1]

(iv) Give the operational definition of "detection limit" in instrumental analysis. [2]

b) In regard to Johnson Noise,

(i) Explain it origins in analytical instrumentation [1]

(ii) Write down the equation relating the magnitude of this noise to the bandwidth, and explain all terms appearing in it [2]

(iii) Use diagrams to explain how difference amplifiers can be used to eliminate this type of noise. [3]

c) In regard to Shot Noise,

(i) Explain it origins in analytical instrumentation [1]

(ii) Write down the equation relating the magnitude of this noise to the bandwidth, and explain all terms appearing in it [2]

(iii) Use diagrams to explain how analog filters can be used to eliminate this type of noise. [3]

d) Although amplification does not eliminate noise, amplifiers are useful in analytical data acquisition. For the following operational amplifiers, draw the circuit and describe the output using the relevant equation.

(i) The current amplifier [2]

(ii) The summing amplifier [2]

(iii) The integrator [2]

(iv) The differentiator [2]

Question 3 [25]

a) Certified reference materials are useful in the evaluation of reliability and validity of analytical data, especially when the analyte is in a complex matrix. In the determination of copper in sugar cane leaves,

(i) What kind of certified reference materials would be suitable for this analysis? [1]

(ii) How would bulk sampling be carried out to source this material? [3]

(iii) Outline the processes that such a material would undergo during certification. [4]

(iv) Explain how this material would be used to evaluate validity and reliability of copper measurements in sugar cane leaves. [3]

b) Blind samples are useful in analytical quality control in a commercial water laboratory.

(i) What is meant by a blind sample?[1]

(ii) Explain how blind samples are used to evaluate validity and reliability of COD measurements in water. [3]

c) Quality control charts are useful in ensuring that repetitive day to day measurements are under statistical control. An in-house reference material was used to generate the following data over a period of 10 days of measurement of nickel in an ore:

Day #	1	2	3	4	5	6	7	8	9	10
Ni, ppm	103	101	104	99	150	101	110	89	102	100

(i) What is meant by an "in-house reference material"?[1]

(ii) Draw the quality control chart for the nickel determination, assuming that the in-house reference material is 101±4 ppm Ni.[3]

(iii) Which days were the measurements not under statistical control and why?[2]

d) Interlaboratory comparisons are useful in the evaluation of reliability and validity of analytical data. In the measurement of nitrates in a mine pit water sample by ion chromatography, "LAB A" ran ten replicate measurements on the sample, and requested "LAB B" to do the same with the remainder of the sample. The following results were obtained:

LAB A (ppm)	25	23	21	24	25	22	20	22	21	20
LAB B (ppm)	23	29	22	18	15	21	25	29	32	21

(i) Comment on the validity of the results at the 95% confidence level [2]

(ii) Comment on the relative precisions of the two laboratories at the 95% confidence level [2]

SECTION B

Answer at least one question from this section. You can print your answers from the printer provided.

Question 4 [25]

a) Briefly define the term Principal Component Analysis. In your description include possible applications of principal component analysis in any example of your choice. You may use diagrams and equations to illustrate your answer. [6]

b) In Principal Component Analysis it is essential that data must be standardized (scaled or weighted and centered). Give reasons. [2]

Five samples were taken for analysis of Cadmium and Lead from a contaminated sample site, see data below:

Code	Cu	Pb
A1	10.3	1.2
A2	10.2	1.2
B1	9.01	1
C1	9.85	1.2
E1	9.5	1.3

Calculate:

i) the eigen values and eigen vectors [2]

ii) the loadings and the score factors [2]

Show your working. You may use Unscrambler software or Statistica on the PC provided to confirm your working.

c) Using the answers from (c) above give a plot of both the loadings and score factors. Briefly discuss the two plots highlighting any observed groupings and correlations. [2]

e) Using the Unscrambler software on the computer provided open the data " water " perform the following tasks:

i) Using basic univariate statistics determine the dominant variables of the data given. Give justifications for your answer. [1]

ii) Perform a Principal Component Analysis using the default values given below:

Weighting	:	1/Sdev.
Validation	:	Leverage
Centering	:	Mean Centering
PC's	:	10

Show the score, loadings, residual and explained variance plots. [2]

iii) Briefly discuss your findings in your Principal Component Analysis above (e(ii)). In your discussion include comments on samples groups, variable correlations, dominant pollutants and their likely sources, outliers and any observations of vital importance in your result output.

[8]

Question 5 [25]

a) Write short notes on the following chemometric terms [5]
i) Eigen value
ii) Multivariate techniques (MVA)
iii) Scores factors
iv) Loading factors
iv) Correlation

Use examples and graphs, where necessary, to illustrate your point.

b) Five samples were taken for analysis of sodium, Na, and Lead, Pb, from a contaminated sample site, see data below:

Code	Na	Mg	Ca
A1	0.3	2.8	13.9
D3	0.13	2.81	13.95
E1	0.18	3.4	13
E2	0.19	3.5	13.3
E3	0.17	3.3	12.9

Calculate:
i) the eigen values and eigen vectors [4]
ii) the loadings and the score factors [4]
Show your working. You may use any chemometrics software on the PC provided to confirm your working.

c) Using the Statistica software on the computer provided enter the data in the table 1 attached perform the following tasks:

i) Using appropriate method determine the dominant variables of the data given. Give justifications for your answer. [2]
ii) Perform a Principal Component Analysis.

Show the score and loadings. [2]

iii) Briefly discuss your findings in your Principal Component Analysis above (c(ii)). In your discussion include comments on samples groups, variable correlations, dominant pollutants and their likely sources, outliers and any observations of vital importance in your result output. [8]

SOIL SAMPLES

	pH	TOC	CHLORIDE	AROMATIC	AMINES	SUGARS	RCH_2COOH	SO^-_4	RCH_2SH	MOISTURE	FLOW	PSIZE
S1_WNT1	3.92	4.44	3.83	3.75	3.89	5.14	3.83	3.06	3.92	5.44	4.17	3.81
S2_WNT1	3.89	4.33	4.28	4.39	4.36	4.69	4.19	2.92	3.19	5.33	4.11	3.83
S3_WNT1	3.89	4.86	4.75	4.67	4.00	3.14	4.47	3.69	3.53	5.31	4.81	4.53
S4_WNT1	4.44	4.97	4.67	4.92	4.64	4.31	4.31	3.22	2.97	5.72	4.31	3.83
S1_SPN3	8.44	8.03	5.11	3.83	3.64	4.78	4.42	3.67	3.94	6.03	3.19	4.28
S2_SPN3	7.75	7.64	5.03	4.06	3.81	4.39	4.33	3.47	3.83	6.03	3.39	4.44
S3_SPN3	6.64	7.08	4.97	5.44	4.92	3.81	4.92	3.50	2.69	6.06	3.53	3.89
S4_SPN3	5.92	6.39	5.08	5.86	5.75	4.97	5.22	2.36	2.14	6.53	3.69	3.33
S1_AUT2	6.17	6.06	4.28	3.72	3.69	5.58	3.50	2.92	4.31	6.03	4.14	3.42
S2_AUT2	4.97	5.61	4.78	4.78	3.94	3.69	4.44	3.39	3.89	6.08	3.75	3.78
S3_AUT2	6.03	6.44	5.17	5.56	5.39	4.69	4.72	2.47	2.75	6.22	3.67	3.19
S4_AUT2	5.14	5.83	5.44	5.61	5.31	3.58	4.72	3.25	2.64	6.14	4.22	3.75

This data represents sediment sample results from river water. The analytes measured are indicated with the physical properties of the sediment samples represented by pH, TOC, moisture, flow and particle size.

Where:

TOC stands for Total Organic Carbon
PSIZE stands for soil particle size
Flow stands for suspension viscosity.