SOLUTIONS IN SOIL SCIENCE -ANALYTICAL AND MEASUREMENT UNCERTAINTY

Robert A. Patterson
Lanfax Laboratories. Armidale
Corresponding author's email: rob@lanfaxlabs.com.au

ABSTRACT

The plethora of government and quazi-government agencies that require environmental assessment and on-going monitoring during operations of agricultural and industrial enterprises, large and small, seem confused between the possibility of measuring soil properties and the real need for those data. Modern field sampling techniques, whether deep drilling, electro-magnetic induction methods, excavation pits or simple spade examination allow a vast number of soil and geologic parameters to be gathered for laboratory assessment without any real input from the sample collector. Australian Standards for sample collection, packaging, and transport and agreed technical standards method have done little to quantify *in-situ* variations in paddock, and the skill of the field technician is fast fading.

Engineers traditionally use the physical properties of soils while agriculturalists favour chemical and biological properties to derive some sense from the complex inter-relationships that take place in the desired use of the soil. Soils for dam construction are assessed differently from soils for viticulture or housing developments. Different soil testing laboratories offer various testing regimes and even within the same elemental analysis a range of tests with different numerical results has become the norm, and often confusion follows.

This paper attempts to address some of the interpolation that can be derived by appropriate selection of physical and chemical tests without yielding to the regulatory authorities' desires to measure or monitor an extensive range of properties that either lead to the same intuitive answer or provide data that will never be used. In some cases, an appropriate site and soil assessment with moderate use of laboratory support can provide a clearer understanding of the likely outcomes of a proposed land use. Do we need tests to quantify the obvious?

Keywords: analytical methods, field assessment, soil monitoring, soil tests, soil variability,

1 INTRODUCTION

Many reasons can be advocated as a valid purpose for soil testing and the range of chemical, biological and physical tests available to meet those reasons are extensive. Whether the desired outcome from testing has anything to do with the practical use of the data often poses two important questions. Firstly, just because we can measure 'something', does it mean that we should? Secondly, are the monitoring requirements dictated by regulators in licences scientifically sound? Of course, the answer to the first question is simple – NO! But the reasoning can be extremely complex and one that requires serious thought and pre-planning. I will address that issue separately.

The answer to the second question is even more elusive because often the directive is in relation to a development application, an environmental study of some complexity or simply as a compliance requirement. Therein hides a real dilemma! Has the authority issuing the directive a clear understanding of the likely uncertainty in the results, the spatial and temporal variation of physical, chemical and biological properties, and the skill to assess the results towards a meaningful outcome? More importantly are all the tests specified required to make an informed opinion or could less tests with more samples provide a more accurate picture? Those aspects of soil monitoring and analytical results will be examined.

There are many environmental guidelines that form the basis for management plans of one description or another, for monitoring, for prescriptive obligations, for performance criteria or as a basis for

decision making. Not all guidelines (and Australian standards are included here) will generate "Best Practice" unless they are based upon science rather than policy or "wishful thinking". Some guidelines tend towards a shotgun approach to monitoring. The example of soil assessment and monitoring used in this paper is for soil tests conducted as part of a monitoring program for irrigation of reclaimed water (treated wastewater) onto land.

2 SOIL VARIABILITY

Most enthusiastic gardeners would know that chemical and biological properties of soil are not only variable across an urban block, but also vary throughout the year because of temperature, rainfall, evaporation, drainage, tillage, compaction or addition of compost (organic matter) because of the use of the garden (or lawn). Extrapolate this variability to broad land areas used for agriculture, horticulture, mining, forestry, subdivisions or natural reserves and the soil variability becomes exceedingly complex because of influences from landscape, regional climate as well as micro-climates. In the garden we may only require one or two soil samples, maybe surface only, whereas on the landscape the need may be for hundreds of samples to reflect the soil's variability. So where do we take representative samples, and how many do we need to be confident about what is happening in the soil environment when effluent is irrigated onto land?

2.1 Soil Sampling

Where to take a soil sample may well be the most difficult starting point in any soil monitoring program because it assumes that you know where the impact (change or no change) will have occurred and can compare that with a control (not influenced by activity). The area irrigated with effluent will show different responses with changes in soil type, and unless you know how the soil properties change across the irrigation area you need to select an identified monitoring point within each major soil type. Different parts of the landscape are more susceptible to change than others. This environmental monitoring area (EMA) will be used for each subsequent monitoring event so that the variability (uncertainty) between readings will be minimised. Each EMA should cover an area of about 25 m^2 from within which replicate samples are taken and should be selected in an area that is more likely to reflect changes from the activity.

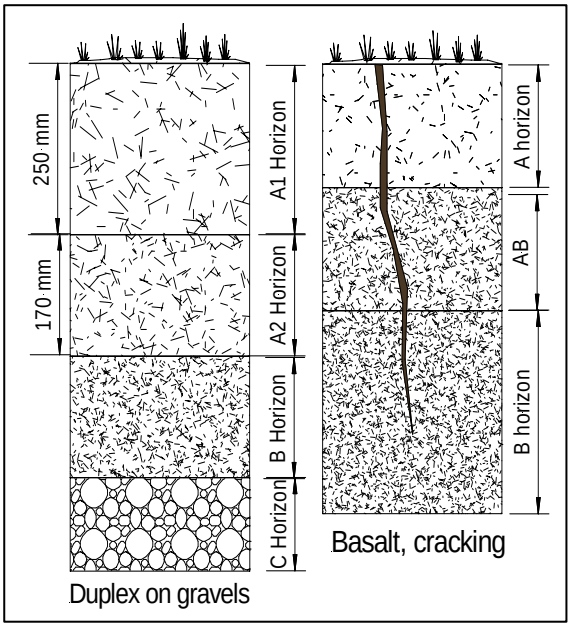

Figure 1 Soil profiles on different geology

Other methods of 'zig zag' across a paddock, a path from one corner to another or just a 'hit and miss' sample simply invites uncertainty into the results and may be difficult to replicate in later years when changes may be expected.

Select the EMA that best represents a particular soil type and under the influence of the irrigation while avoiding tracks, drainage lines, sheep camp or influences other than effluent irrigation. Where there are several soil types in the one paddock (not uncommon), you will require two or more EMA.

At each EMA (recorded with GPS coordinates), the first sample set will determine how many replicates are taken within the EMA, and down the profile. It is relatively easy to take many samples from the surface (usually 0 – 100 mm), bulk them into a bucket or clean plastic bag, remove stones and vegetation, mix well and sub-sample into a clean labelled sample bag.

Sampling down the profile is often dictated in management plans as every 100 mm depth to a maximum of 1200 mm in the profile, and replicated three times. The real problem with this procedure,

while it is simple on paper, is that each hole requires 12 samples, whereas the profile has three or maybe four horizons as shown in Figure 1. Since the influence of the effluent will be uniquely with the properties of each horizon, more detailed sampling increases cost (sampling time, analytical costs) and may complicate data assessment. Why have 12 samples when three will suffice? Sampling at each horizon, however, requires the person taking the sample to have some basic soil pedology.

Figures 2 and 3 show the surface soil variation in soil pH and exchangeable sodium percentage (ESP) derived from 40 soil cores taken at a regular grid across an 8 ha paddock (330 m x 240 m) that had been top-dressed with feedlot manure 12 months previously. The best method for spreading the dried manure was employed to ensure an even application and a crop of sorghum, grown in the ploughed paddock was harvested prior to soil sampling. The contours indicate the complexity of simple chemical analyses and how single site sampling could be non-representative. An analysis of composite sampling from zig-zag, diagonal and up to six random site collections was no more convincing than at an EMA, yet considerably more labour intensive. With repeated sampling over future seasons, the changes in EMA properties are likely to reflect changes to the whole paddock, at least that is the expectation.

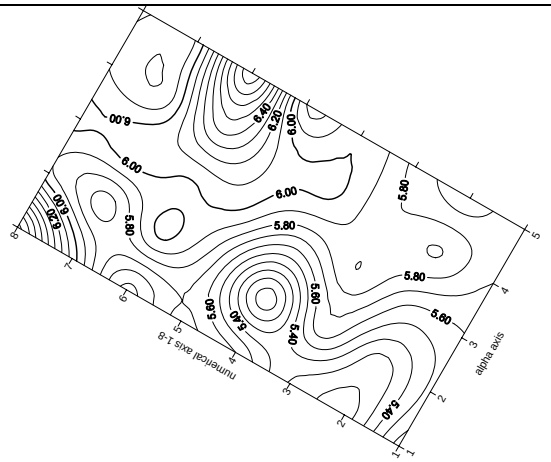

Figure 2 Variation in soil pH

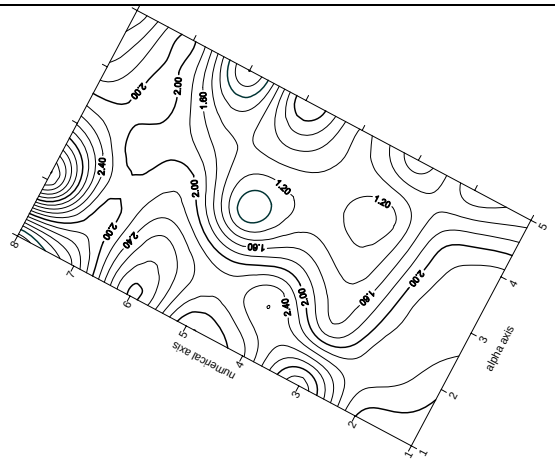

Figure 3 Variation in ESP

2.2 Number of soil samples

Many soil tests yield highly variable results simply because of the variability of the soil. A statistical analysis may be required to confirm the probable results, hence more than one test will be required. The measurement of hydraulic conductivity in the field requires more than one sample, but how many? That remains a serious consideration when planning a monitoring program -how many replicates at each site? Simple statistics are likely to indicate that for statistical acceptability, many more tests than has been set aside for the program will be required. Some compromise to meet financial constraints may need to be made, while maintaining confidence in the results.

2.3 Soil sample collection, storage and transport

While the legal requirements for "Chain of Custody Reports" may satisfy bureaucratic correctness, nothing can overcome the effects of poor packaging, poor transport and slow delivery unless protocols are in place to ensure all samples are handled appropriately and always handled the same way.

Protocols need to consider whether samples require cooling and chilling to reduce oxidation, as for acid sulphate soils, or they can be stored dry for long periods. The size of the sample (volume or weight), identification of sample (unique labelling), special packaging and storage are best derived from discussions with the laboratory tasked with the analyses. Often the most expensive part of the sampling routine is transport to the site and obtaining a sample. Excess soil sampled is unlikely to be a problem and may allow reserve samples to be stored if confirmation is required at a later time.

Water samples deteriorate more rapidly than soil samples and protocols for sampling need to take into account freight and receival in the laboratory. A water sample collected on Friday may have to wait until freight is available on Monday, for receipt by laboratory on Tuesday, by which time the sample is at least three days old. Many water samples have maximum holding periods of 24 hours. Soil samples that are to be held for extended periods may need air-drying, chilling or freezing.

2.4 Monitoring Schedule

Unlike licence requirements that may be annual, quarterly or some other specified period, monitoring as part of a management plan will be more flexible around the "need to know" or "incident based" rather than the "order to prove" particular soil properties or changes to characteristics at regular intervals. Soil properties are slow to change in response to small regular influences of farming, fertiliser application, biosolids spreading or irrigation with wastewater. Buffering capacity of the soil is naturally high, and in clay soils the extremely large expression of surface area, requires significant addition or removal of cations, or change to pH to show a response to that activity. For this reason, soil sampling on a seasonal or annual basis is preferred. Some licence conditions allow for annual sampling for the first three years, and, pending favourable review by soil scientist, further testing is conducted at extended frequencies.

2.5 Suite of soil analytes

The suite of soil analytical methods is extensive. Soils may be mixed with reagents, digested, fused with a chemical, heated, ignited, radiated by X-rays, ravaged by microbes to reveal particular chemical properties. Other suites of tests are used to measure physical properties such as water holding capacity, bulk density, particle size distribution, linear shrinkage. What tests are chosen will depend upon the sensitivity of the method and the interferences from other soil properties. Some laboratories have preference for one method over another for a variety of reasons, not the least is operator competence and Occupational Health and Safety (OH&S) issues. So the question arises "does it make any difference which test method is used?" to which the answer is "most definitely yes".

Table 1 sets out a range of tests suitable for monitoring the soils in an effluent irrigation area with a reasonable expectation of providing valuable inputs to management and some confidence in predicting the changes to the soil environment. But which tests can be deleted and inferred from other tests?

Table 1. A list of soil tests that may be required of a sewage irrigation scheme

pH	Electrical conductivity (EC)	Total dissolved solids (TDS)
Soil pH in water and $CaCl_2$	Electrical conductivity (salinity) as above	Soil ammonium
Nitrite and nitrate	Total Kjeldahl nitrogen (TKN)	Total nitrogen (TN)
Extractable phosphorus	Total Phosphorus	Exchangeable cations (Na, Ca, K, Mg, Al)
Trace minerals (As, B, Cr, Cu, Co, Fe, Pb, Hg, Ni, Se, Dr, Zn)	Soil Chloride	Organochlorines and Organophosphates
Hydraulic conductivity	Phosphorus sorption index	Extractable sulphur
Cation exchange capacity	Soil organic carbon	Total organic carbon

2.6 Pollutant or nutrient?

One of the early decisions in the selection of appropriate tests is whether the particular component is a pollutant or a nutrient, since this may determine the frequency of analysis as well as the expected levels in the sample. Phosphorus in water may be either a pollutant or a nutrient, depending upon the fate of the water. Phosphorus in Coca-Cola™ is a food acid (338 – phosphoric acid) and an emulsifier in cheese but a vital nutrient for proliferation of Cyanobacteria ("blue-green algae") in water.

Phosphorus discharges to water need to be minimised. In the context of the effluent reuse scheme, firstly removing phosphorus from the effluent using lime or ferric salts creates a by-product with little benefits compared to removing the element with high nutrient value to crops and pastures.

2.7 Do we measure because we can?

Monitoring requirements, imposed by regulators, may appear to have no practical application or whose impact on the environment is not well understood. It is common for regulators to demand collection of vast quantities of data that are never accumulated into a knowledge base. Each EPA licence holder is required to collect data and report annually. Where are those data promulgated by DEC as knowledge for the community in which they are gathered? Where are the synergies from the collection of this wealth of data and the beneficial application to future activities?

After several monitoring periods, a skilled soil scientist will be able to make correlations between test results and observations to determine which tests are not required to make an accurate assessment of the impact the effluent irrigation is having on the soils. For example, the author (Patterson, 1997) showed that there was sufficiently robust correlation between soil TKN and soil organic carbon (OC) that is was cost effective to discontinue the more expensive TKN test and maintain the OC test.

2.8 Policy or Science?

Instances when regulators misinterpret data and prepare guidelines based upon that misunderstanding are not uncommon. The NSW Environment Protection Authority (2004) (now Department of Environment and Conservation states that "An EAT of 8 (Emerson Aggregate Test) means that the soil is so stable that it cannot be penetrated by roots". Class 8 soils are water stable, non-swelling such as the surface of a vertosol (black cracking clay) on which extensive crops are grown across the country. Indeed a water stable aggregate is a desirable characteristic and one that often reflects large amounts of soil organic carbon. The opposite to water stable aggregates are soils that rapidly disperse (usually a sodicity effect) or soils that slake when wet (usually low in organics). So is the first stage of a monitoring program educating the regulator on soil science?

The Hunter River Salinity Trading Scheme (EPA, 2003) suggests that in the 22 000 km^2 catchment of the Hunter River, salinity is calculated "so that the salt concentration does not go above 900EC (*sic*) in the middle and lower sectors of the river, or above 600EC (*sic*) in the upper sector" (p.4). There are two significant problems with this policy, enforceable through legislation. The units of measure as stated are some bureaucratic shorthand for what should be microSiemens per centimetre. This example of simplified policy leads to proliferation of incorrect units across the community. The second and most serious concern is that salinity is the criterion against which the trading scheme is measured, monitored and regulated. Salinity, as correlated with an electrical conductivity measurement (there is considerable variation in the correlation coefficient) is a measure of all the salts in the water. A salt is a compound that dissociates in water to form cations (positive ions) and anions (negative ions) which balance electrically. Environmental salts can be any combination of Na^+, Ca^{2+}, K^+, Mg^{2+}, Fe^{2+}, Al^{3+}, Zn^{2+}, Cu^{2+}, Mn^{2+}, other metals with anions of NO_3^-, PO_4^-, SO_4^{2-}, Cl^-, CO_3^{2-}, HCO_3^- and many others. To suggest that all the salts are detrimental to water quality is contrary to agricultural, horticultural, and engineering understanding of what constitutes a problem. Lime (as salt of calcium and carbonate) is added to soil to modify its pH and ameliorate sodic soils, gypsum (calcium sulphate) is used to ameliorate sodic soils and flocculate clays, and each of these additions works by increasing the salinity of the soil – but – the beneficial element is calcium. The problem with salinity units is exacerbated for soil measurements when there are two common methods for determining soil salinity, one based upon a soil water suspension and the second on a saturation extract. Often, the method is not reported and interpretation less than ideal.

Unfortunately salt in the environment is often considered synonymous with table salt (sodium chloride) and this is not necessarily the case. So good policy has to be based upon science, or policy will lead to bad outcome, irrespective of enforcement of policy. But who 'picks up the tab' when the outcomes of bad policy are bad outcomes? There's no such thing as bad science, just wrong inputs.

3 LABORATORIES AND METHODS

3.1 Which Laboratory?

The choice of a laboratory may be one of convenience (nearby) or one that provides reasonably priced analyses. However, the more important aspects are that the laboratory undertakes quality assurance (acceptable methods) and quality control (internal standards), both of which are complemented by inter-laboratory performance programs. The National Association of Testing Authorities (NATA) provides accreditation of laboratories to an international standard, but certification does not guarantee accuracy, simply that the protocols are in place to minimise uncertainty from likely sources of error (methods, operator competence, calibration, standardisation). The cost of NATA accreditation is high, both in monetary terms and in scheduled proficiency programs and many smaller laboratories can perform at equally high levels and not have accreditation. The Australasian Soil and Plant Analysis Council (ASPAC) also run inter-laboratory proficiency programs on Australian soil and plants. These tests satisfy the measure of competency required by NATA but may not result in uniform results as shown in Figure 4 where variation is significant.

3.2 Uncertainty

The uncertainty of a sample result is the combination of all errors (induced and natural) that impinge on the final results. Unfortunately, where laboratory analysis may aim at low levels of uncertainty (low coefficients of variability < 10%), in-field soil sampling brings with it a level of uncertainty that may be as high as 200% for no other reason than natural variability. Water samples tend not to vary significantly over short distances, but within a reservoir may show considerable variation from the limnetic (open water) zone to the benthic (bottom) region. Compounding this variability may be sampling techniques as discussed in Section 2.1 and packaging and transport induced changes, and delays in commencing analysis. That some regulatory authorities view soil sample changes of 10-20% as significant, many components may change by 100% and remain insignificant. Understanding which changes are relevant leads to selecting tests that better indicate those changes in a cost effective and timely manner.

Recent proficiency tests (Proficiency Testing Australia, June 2006) for water returned CVs of 8.8% for aluminium, 10.6% for boron, while two samples of manganese had CVs of 7.7% and 5.2%. When these CVs are translated into the range of acceptable values from NATA certified laboratories and a range of ± 3 times an estimate of the standard deviation, significant spread is acceptable. Translating these values to include other uncertainty from sample collection and transport shows that data must be critically assessed for their accuracy rather than a bland expectation of a particular scale of change.

The numerical value of the results is important in terms of the precision of the measurement and the relative importance of result. NATA, and other proficiency programs, set basic protocols for reporting significant figures. For example:

- Values less than 1 mg/kg are reported to three decimal places e.g. 0.256 mg/kg;
- Values between 1 and 10 mg/kg reported to two decimal places, e.g. 9.25 mg/kg;
- Values 10 to 100 mg/kg reported to one decimal place, e.g. 25.6 mg/kg; and
- Values over 100 mg/kg reported in whole numbers e.g. 244 mg/kg

3.3 Split samples

In the event of a sample result not being within the expected range and the laboratory being the first target to attract the blame, some regulators and program managers employ split sample analysis. That is, half of a sample is sent to one laboratory and the other half sent to another laboratory. Such an exercise is similar to Russian Roulette. As an example, Figure 4 shows the variability of a water sample tested for 5-day Biochemical oxygen demand. The spread of acceptable results was discussed above as from -3 to +3 on the graph, that is from 78.4 mg/L to 104 mg/L (using the significant figures required), as spread of 25.2 mg/L. So that if the results for the split samples were at each end of the graphs (there is no way of knowing where the laboratory will perform), results as far apart as 25 mg/L could be expected. Similar results are repeated for other water and soil analyses.

3.4 Method selection and sensitivity

Of the methods selected for the range of tests required, there is no one test that will give all the results simultaneously. When X-ray instruments are involved with determining numerous elements, not every element will be determined at its unique concentration in the soil, simply because of instrument sensitivity to low or high concentrations. Similarly, the same determination will not show the species differences between total and plant available elements. The Mehlich 3 test is used in the United States as a one-stop shop for most of the elemental analysis for agriculture. Australian soil science is less than convinced judging by comparison of proficiency samples against accepted methods.

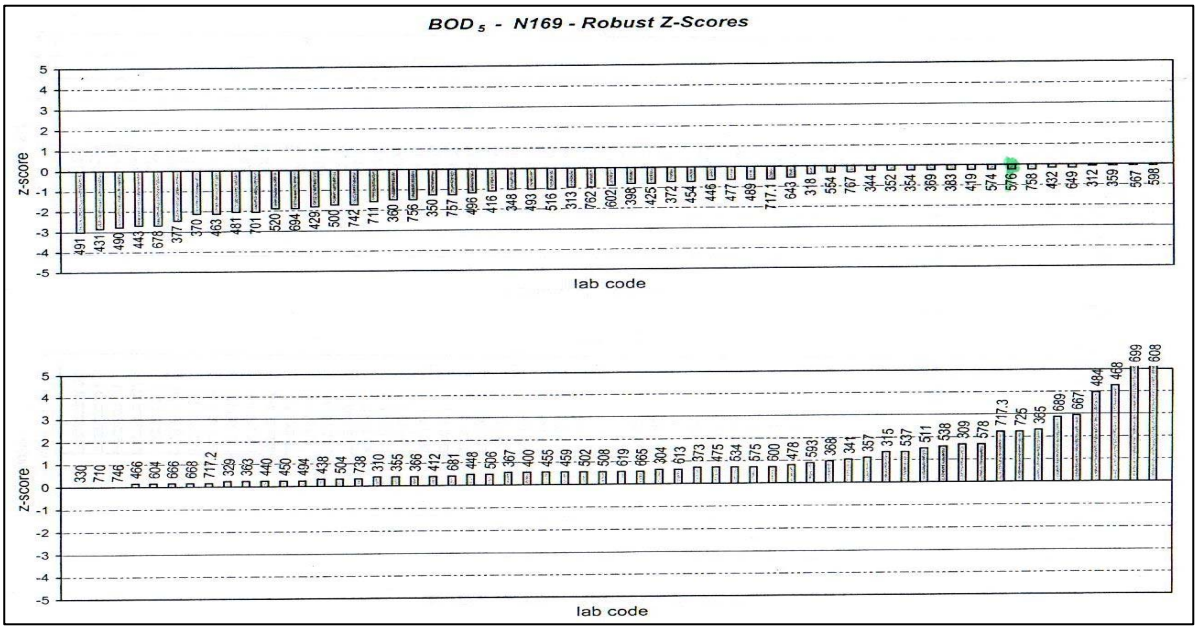

Figure 4 Comparison of performance of 5-day BOD test

Unfortunately analysts have developed a range of tests to account for interference from other soil components. Take for example the measurement of plant available phosphorus where the range of tests accounts for soil pH, phosphorus that is readily soluble in soil water to that which is released slowly from the soil matrix. These tests are shown in Table 2 together with the results from a recent ASPAC proficiency testing program. It would be unwise to try to make comparisons between different tests. Find a test that meets your requirements and stick with it!

Table 2 Variation in three extractable phosphorus tests (After ASPAC 2003)

Method	Units of Measure	Soil # 150	Soil # 154
Colwell Extractable P	mg/kg	41.2 – 85.2	7.1 – 20.5
Olsen Extractable P	mg/kg	14.2 – 45.6	1.1 – 8.1
Bray Extractable P	mg/kg	30.2 – 58.6	1.2 – 3.8

It is clear from Table 2 that swapping between tests exacerbates the variability and there is no correlation coefficient to standardise the results with previous but different tests. Interpretation of the test method based upon local knowledge (or assessor's experience) should not be discounted. Many agronomists are comfortable with understanding Colwell P Method and they should stick with it.

Since laboratories have developed particular methods based upon previous contracts for particular tests, equipment and operator competency, similar instrumentation within methods needs to be considered. Table 3 shows the variability from proficiency sampling results for measuring calcium in water based upon instrumentation. It has to be assumed that competent operators and quality controls were used to derive these results. Which result is accurate? The atomic absorption spectrometer (AAS)

with air-acetylene flame gave the widest results, but individual results may have been as accurate as that from the Inductively Coupled Plasma (ICP).

Table 3 Variation in sample analysis with equipment (after NATA subprogram 40, March 1999)

Method	N113 (mg/L)	N115 (mg/L)
AAS Air - acetylene	8.11 – 11.89	0.001 – 4.76
AAS nitrous oxide – acetylene	8.67 – 11.13	1.16 – 3.44
ICP	8.91 – 11.49	2.04 – 2.94

3.5 Methods reporting

Whatever method is used, the laboratory should indicate to the client the method used and the limits of detection for that method. Limits of detection (LD), or Practical Limits of Quantification (PLQ) are the lowest practical result that the laboratory can repeatedly expect to determine. Different instruments have different limits of detection. However, there is no such result as ZERO. An analysis that failed to detect sulphur in the soil will be reported as less than the limit of detection. For one method that may be < 3 mg/kg, while a more sensitive instrument may be able to determine that sample had 0.2 mg/kg, but could only determine down to 0.1 mg/kg. While precision may be a valuable goal, for most soils the relative changes to soil properties are more important for assessment of soil changes and minute changes are obscured by the natural variability.

3.6 Nomenclature and units of measure

Without a common vocabulary using System d'Internationale (SI) units, much of the scientific efforts will be lost through misinterpretation. The previous example of referring to units of electrical conductivity as EC rather than microSiemens per centimetre creates false impressions. The term 'non-filtrable residue" was displaced from the water testing vocabulary in 1986 and replaced with total suspended solids (that material remaining on a filter paper). Some state departments use the abbreviation SS in their meaning for "suspended solids" when the rest of the world uses TSS (total suspended solids) to avoid confusion with Settleable Solids (SS).

4 CONCLUSION

Practical skills, acquired by persons involved with field collection of samples, packaging, and transportation will include knowledge of what particular tests will be conducted on the soils so that variability at the collection end will not detract from technological. An understanding of the analytical variation, within and between laboratory analyses, uncertainty, and specific tests for specific purposes will lead to improved performance of both regulators and monitors. There is no valid reason why policy and regulations should not be based upon science, what needs to be common ground is agreement on what constitutes environmental harm and what is simple environmental readjustment to new inputs (effluent) and removals (leaching and nutrient extraction).

5 REFERENCES

Australasian Soil and Plant Analysis Council (2003) Soil Proficiency Testing Program Report.

Environment Protection Authority (2003) *Hunter River Salinity Trading Scheme. Working together to protect river quality and sustain economic development.* NSW EPA, Sydney

National Association of Testing Authorities (1999) Water Proficiency Testing Sub-program 40 (Cations), PTAC Report No. 283. National Association of Testing Authorities. Rhodes NSW

Proficiency Testing Australia (2006) Summary Results Waters Sub-program 86, June 2006.

Patterson, R.A. (1997) Environmental Indicators of Effective Sewage Effluent Re-use. Project Report, Master of Engineering. Deakin University, Geelong.